Uvodno predavanje

Linearna Algebra IŠRM

Jaka Cimprič

FMF UL

oktober 2020

Predstavitev predavatelja

predavatelj: Jaka Cimprič

email: jaka.cimpric@fmf.uni-lj.si

kabinet: Jadranska 21, soba 4.21

govorilne ure: samo po emailu

Predstavitev predmeta

predmet: Linearna algebra (IŠRM)

obseg: Celoleten predmet, dve uri predavanj in dve uri vaj na teden

režim: Štirje kolokviji in trije izpitni roki (odvisno od razmer).

ocena: Oceno iz predmeta dobite na ustnem izpitu. Za pristop k ustnemu izpitu rabite vsaj 50% točk iz kolokvijev ali iz pisnega izpita. Veljajo samo kolokviji in pisni izpiti v tem šolskem letu.

obveščanje: preko spletne učilnice `https://ucilnica.fmf.uni-lj.si/` Čimprej se vpišite, sicer ne boste prejemali obvestil o tem predmetu.

izvajanje: Hibridno v predavalnici in na daljavo. Na prvem predavanju 5. 10. so v predavalnici tisti s lihimi vpisnimi številkami. Ostali spremljajo predavanje po Zoomu (link je na spletni učilnici). Drugi teden obratno.

Predstavitev snovi

Prvi semester:

- Vektorji v $\mathbb{R}^n$
- Sistemi linearnih enačb
- Matrike
- Determinante
- Algebraične strukture
- Vektorski prostori

Drugi semester:

- Linearne preslikave
- Lastne vrednosti in lastni vektorji
- Prostori s skalarnim produktom
- Adjungirana linearna preslikava
- Singularni razcep
- Kvadratne forme

Osnovna literatura:

- Koširjeva skripta (čez celo leto),
- moja skripta (samo 4 poglavja).

Dostopno preko spletne učilnice.

Vektorji v $\mathbb{R}^n$

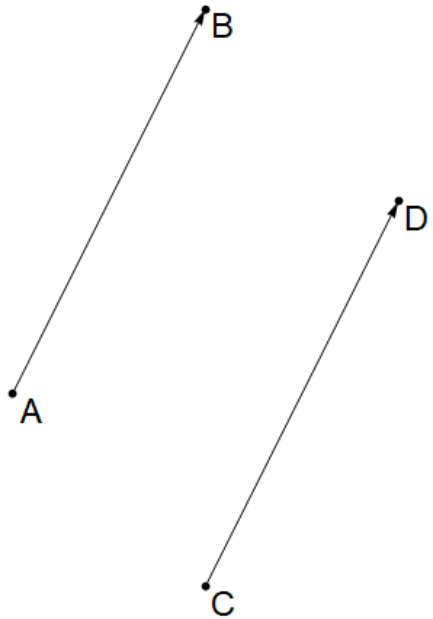
Geometrijski in krajevni vektorji

Za lažjo predstavo se omejimo na ravnino, čeprav bo vse veljalo tudi za prostor, pa tudi za višje dimenzije.

Definicija geometrijskega vektorja

Geometrijski vektor je usmerjena daljica med dvema točkama. Začetni točki pravimo **rep**, končni točki pa **glava** vektorja. Dva geometrijska vektorja sta **enaka**, če sta vzporedna, enako dolga in kažeta v isto smer.

Primer, ko je $\overrightarrow{AB} = \overrightarrow{CD}$



Izberimo sedaj v ravnini neko točko, ki ji pravimo **izhodišče**.

Definicija krajevnega vektorja

Krajevni vektor je geometrijski vektor, ki ima rep v izhodišču.

Vsak geometrijski vektor je enak natanko enemu krajevnemu vektorju. Lahko ga namreč vzporedno premaknemo tako, da rep pade v izhodišče.

Vsak krajevni vektor je natanko določen s svojo glavo. Če skozi izhodišče potegnemo koordinatni sistem, potem je glava natanko določena s svojimi koordinatami. Torej lahko identificiramo naslednje pojme:

- Urejeni pari realnih števil.
- Točke v ravnini.
- Krajevni vektorji v ravnini.
- Množice paroma enakih geometrijskih vektorjev v ravnini.

Podobna identifikacija velja tudi za prostor in višje dimenzije.

Primer: Koordinate krajevnega vektorja, ki je enak geometrijskemu vektorju iz točke $(x_1, \ldots, x_n)$ v točko $(y_1, \ldots, y_n)$, so $(y_1 - x_1, \ldots, y_n - x_n)$.

Operacije z vektorji

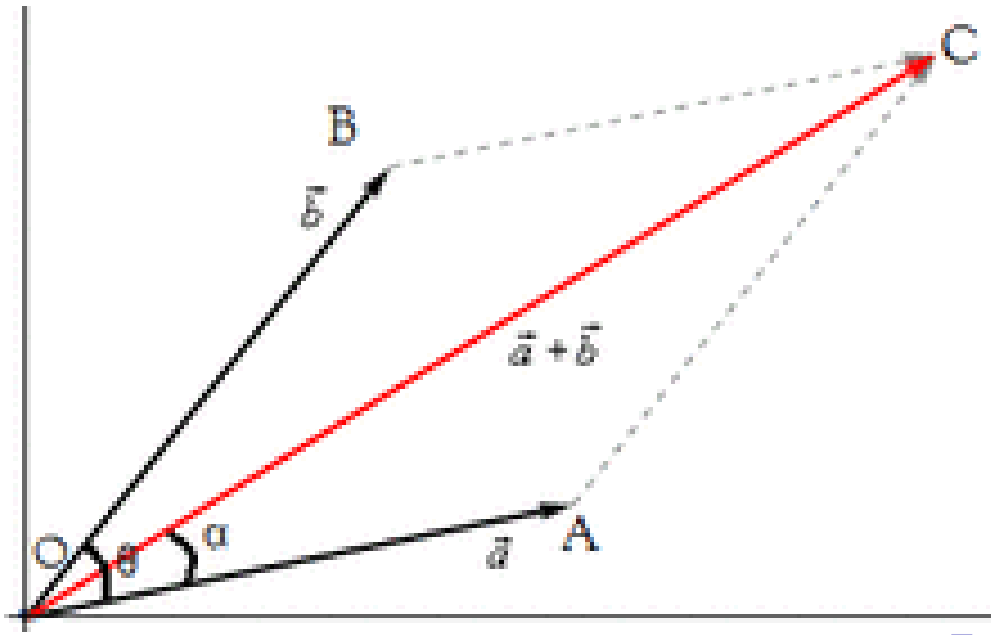
Krajevnim vektorjem bomo v nadaljevanju rekli kar **vektorji**. Njihova glavna prednost pred geometrijskimi vektorji je, da z njimi lahko računamo.

Definicija vsote

Vsota dveh vektorjev je algebraično definirana z

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) := (x_1 + y_1, \ldots, x_n + y_n).$$

Geometrijsko vsoto pojasnimo s paralelogramskim pravilom:



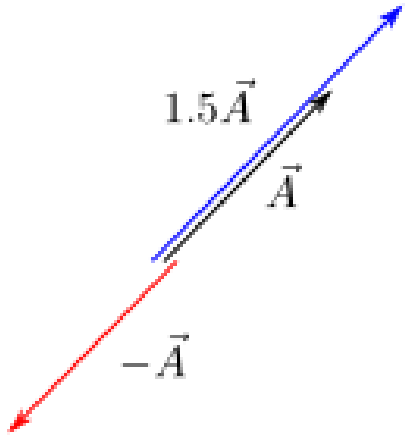
Realnim številom bomo v nadaljevanju pogosto rekli **skalarji**.

Definicija produkta s skalarjem

Produkt vektorja s skalarjem je algebraično definiran z

$$\alpha(x_1, \ldots, x_n) := (\alpha x_1, \ldots, \alpha x_n).$$

Geometrijski pomen produkta vektorja s skalarjem α je, da vektor raztegnemo za faktor α. Če je $\alpha < 0$, potem vektor raztegnemo za faktor $|\alpha|$ in ga prezrcalimo čez izhodišče.



Ker sta obe operaciji definirani po komponentah, lahko njune lastnosti izpeljemo iz lastnosti realnih števil. Označimo $\mathbf{0} := (0, \ldots, 0)$ in

$$-(x_1, \ldots, x_n) := (-x_1, \ldots, -x_n) = (-1)(x_1, \ldots, x_n).$$

Potem veljajo naslednje lastnosti:

Lastnosti vsote

- $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$,
- $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$,
- $\mathbf{x} + \mathbf{0} = \mathbf{x}$,
- $\mathbf{x} + (-\mathbf{x}) = \mathbf{0}$.

Lastnosti množenja s skalarjem:

- $\alpha(\mathbf{x} + \mathbf{y}) = \alpha\mathbf{x} + \alpha\mathbf{y}$,
- $(\alpha + \beta)\mathbf{x} = \alpha\mathbf{x} + \beta\mathbf{x}$,
- $(\alpha\beta)\mathbf{x} = \alpha(\beta\mathbf{x})$,
- $1\mathbf{x} = \mathbf{x}$, $0\mathbf{x} = \mathbf{0}$.

Te lastnosti bomo kasneje vzeli za aksiome abstraktnega vektorskega prostora.

Linearne kombinacije vektorjev

Definicija linearne kombinacije

Vektor $\mathbf{v} \in \mathbb{R}^n$ je **linearna kombinacija** vektorjev $\mathbf{v}_1, \ldots, \mathbf{v}_m \in \mathbb{R}^n$, če obstajajo taki skalarji $\alpha_1, \ldots, \alpha_m \in \mathbb{R}$, da velja

$$\mathbf{v} = \alpha_1 \mathbf{v}_1 + \ldots + \alpha_m \mathbf{v}_m \in \mathbb{R}^n.$$

Opomba: Če hočemo preveriti, ali je vektor $\mathbf{v}$ linearna kombinacija vektorjev $\mathbf{v}_1, \ldots, \mathbf{v}_m$, moramo rešiti sistem n linearnih enačb v m neznankah. S sistemi linearnih enačb se bomo ukvarjali kasneje.

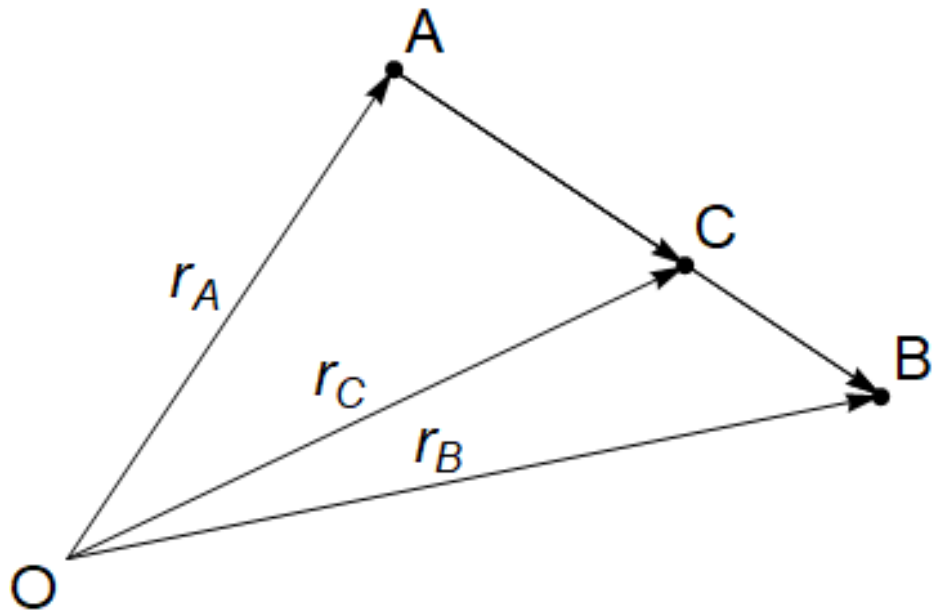
Primer linearne kombinacije

Preverimo, ali je vektor $(0, 2, -3)$ linearna kombinacija vektorjev $(1, 1, -1)$ in $(2, 0, 1)$. Če enačbo $(0, 2, -3) = \alpha(1, 1, -1) + \beta(2, 0, 1)$ razpišemo po komponentah, dobimo $0 = \alpha + 2\beta$, $2 = \alpha$, $-3 = -\alpha + \beta$. Uganemo rešitev $\alpha = 2, \beta = -1$, torej je odgovor na vprašanje pozitiven.

Primer: Delitev daljice v danem razmerju

Dana je daljica AB. Iščemo tako točko C na tej daljici, ki jo deli v razmerju $3:2$. To pomeni, da velja $\overrightarrow{AC} = \frac{3}{5}\overrightarrow{AB}$.
Označimo z $\mathbf{r}_A$, $\mathbf{r}_B$ in $\mathbf{r}_C$ krajevne vektorje, ki ustrezajo točkam A, B in C. Radi bi izrazili $\mathbf{r}_C$ z $\mathbf{r}_A$ in $\mathbf{r}_B$. Ker je $\overrightarrow{AB} = \mathbf{r}_B - \mathbf{r}_A$, dobimo

$$\mathbf{r}_C = \overrightarrow{0A} + \overrightarrow{AC} = \mathbf{r}_A + \frac{3}{5}(\mathbf{r}_B - \mathbf{r}_A) = \frac{2}{5}\mathbf{r}_A + \frac{3}{5}\mathbf{r}_B.$$



Primer: Središče točk

Središče točk $\mathbf{r}_1, \ldots, \mathbf{r}_m$ je točka

$$\frac{1}{m}(\mathbf{r}_1 + \ldots + \mathbf{r}_m).$$

Oglejmo si geometrijsko konstrukcijo središča: Naj bo $\mathbf{p}_1 = \mathbf{r}_1$. Za vsak $i = 2, \ldots, m$ naj bo $\mathbf{p}_i$ taka točka na daljici med $\mathbf{p}_{i-1}$ in $\mathbf{r}_i$, ki to daljico deli v razmerju $1 : (i-1)$. Trdimo, da je potem $\mathbf{p}_m$ iskana točka.

Dokaz: S popolno indukcijo bomo dokazali, da za vsak $i = 1, \ldots, m$ velja $\mathbf{p}_i = \frac{1}{i}(\mathbf{r}_1 + \ldots + \mathbf{r}_i)$. Ker je $\mathbf{p}_1 = \mathbf{r}_1$, trditev drži za $i = 1$. Po prejšnjem primeru za vsak $i = 2, \ldots, m$ velja $\mathbf{p}_i = \frac{i-1}{i}\mathbf{p}_{i-1} + \frac{1}{i}\mathbf{r}_i$. Ko vstavimo indukcijsko predpostavko $\mathbf{p}_{i-1} = \frac{1}{i-1}(\mathbf{r}_1 + \ldots + \mathbf{r}_{i-1})$ in uredimo, dobimo $\mathbf{p}_i = \frac{i-1}{i}(\frac{1}{i-1}(\mathbf{r}_1 + \ldots + \mathbf{r}_{i-1})) + \frac{1}{i}\mathbf{r}_i = \frac{1}{i}(\mathbf{r}_1 + \ldots + \mathbf{r}_i)$, kar smo trdili.

Linearna neodvisnost vektorjev

Definicija linearne neodvisnosti vektorjev

Vektorji $\mathbf{v}_1, \ldots, \mathbf{v}_m$ so **linearno odvisni**, če je eden od njih enak linearni kombinaciji preostalih. Sicer so **linearno neodvisni**.

En vektor je linearno neodvisen, če ni enak nič. Dva vektorja sta linearno neodvisna, če ne ležita na isti premici skozi izhodišče. Trije vektorji so linearno neodvisni, če ne ležijo na isti ravnini skozi izhodišče.

Oglejmo si še, kako linearno neodvisnost vektorjev računsko preverimo.

Trditev

Vektorji $\mathbf{v}_1, \ldots, \mathbf{v}_m$ so linearno neodvisni natanko tedaj, ko ima sistem linearnih enačb $\alpha_1\mathbf{v}_1 + \ldots + \alpha_m\mathbf{v}_m = \mathbf{0}$ samo eno rešitev, to je trivialno rešitev $\alpha_1 = \ldots = \alpha_m = 0$.

Dokaz: Očitno je $\alpha_1 = \ldots = \alpha_m = 0$ vedno rešitev sistema enačb $\alpha_1\mathbf{v}_1 + \ldots + \alpha_m\mathbf{v}_m = \mathbf{0}$. Če bi imel ta sistem še kakšno drugo rešitev, potem bi za to rešitev veljalo $\alpha_i \neq 0$ za nek i. Potem bi lahko izrazili

$$\mathbf{v}_i = (-\frac{\alpha_1}{\alpha_i})\mathbf{v}_1 + \ldots + (-\frac{\alpha_{i-1}}{\alpha_i})\mathbf{v}_{i-1} + (-\frac{\alpha_{i+1}}{\alpha_i})\mathbf{v}_{i+1} + \ldots + (-\frac{\alpha_m}{\alpha_i})\mathbf{v}_m.$$

Po definiciji bi to pomenilo, da je $\mathbf{v}_i$ enak linearni kombinaciji preostalih vektorjev, kar bi pomenilo, da so vektorji $\mathbf{v}_1, \ldots, \mathbf{v}_m$ linearno odvisni.

Velja tudi obratno. Če bi bili $\mathbf{v}_1, \ldots, \mathbf{v}_m$ linearno odvisni, potem bi lahko enega od njih izrazili kot linearno kombinacijo preostalih, recimo $\mathbf{v}_i = \beta_1\mathbf{v}_1 + \ldots + \beta_{i-1}\mathbf{v}_{i-1} + \beta_i\mathbf{v}_{i+1} + \ldots + \beta_{m-1}\mathbf{v}_m$. Potem bi sistem $\alpha_1\mathbf{v}_1 + \ldots + \alpha_m\mathbf{v}_m = \mathbf{0}$ imel, poleg trivialne rešitve, še rešitev $\alpha_1 = \beta_1, \ldots, \alpha_{i-1} = \beta_{i-1}, \alpha_i = -1, \alpha_{i+1} = \beta_i, \ldots, \alpha_m = \beta_{m-1}$.

Posledica

Vektorji $\mathbf{v}_1, \ldots, \mathbf{v}_m \in \mathbb{R}^n$ so linearno neodvisni natanko tedaj, ko se da vsak vektor iz $\mathbb{R}^n$ na **kvečjemu en** način izraziti kot linearna kombinacija vektorjev $\mathbf{v}_1, \ldots, \mathbf{v}_m$.

Dokaz: Če so $\mathbf{v}_1, \ldots, \mathbf{v}_m$ linearno neodvisni, in če se da nek vektor $\mathbf{v} \in \mathbb{R}^n$ izraziti kot

$$\mathbf{v} = \alpha_1\mathbf{v}_1 + \ldots + \alpha_m\mathbf{v}_m = \beta_1\mathbf{v}_1 + \ldots + \beta_m\mathbf{v}_m,$$

za neke skalarje $\alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_m$, potem je

$$\mathbf{0} = \mathbf{v} - \mathbf{v} = (\alpha_1 - \beta_1)\mathbf{v}_1 + \ldots + (\alpha_m - \beta_m)\mathbf{v}_m.$$

Po trditvi odtod sledi, da je $\alpha_i - \beta_i = 0$ za vsak $i = 1, \ldots, m$.

Obratno, če se da vektor $\mathbf{0}$ na kvečjemu en način izraziti kot linearna kombinacija vektorjev $\mathbf{v}_1, \ldots, \mathbf{v}_m$, potem iz

$$\alpha_1\mathbf{v}_1 + \ldots + \alpha_m\mathbf{v}_m = \mathbf{0} = 0\mathbf{v}_1 + \ldots + 0\mathbf{v}_m$$

sledi, da je $\alpha_i = 0$ za vsak $i = 1, \ldots, m$. Po trditvi odtod sledi, da so vektorji $\mathbf{v}_1, \ldots, \mathbf{v}_m$ linearno neodvisni.

Definicija ogrodja

Vektorji $\mathbf{v}_1, \ldots, \mathbf{v}_m \in \mathbb{R}^n$ so **ogrodje** natanko tedaj, ko se da vsak vektor iz $\mathbb{R}^n$ na **vsaj en** način izraziti kot linearna kombinacija vektorjev $\mathbf{v}_1, \ldots, \mathbf{v}_m$.

Definicija baze

Vektorji $\mathbf{v}_1, \ldots, \mathbf{v}_m \in \mathbb{R}^n$ so **baza** natanko tedaj, ko so ogrodje in ko so linearno neodvisni. Se pravi natanko takrat ko se da vsak vektor iz $\mathbb{R}^n$ na **natanko en** način izraziti kot linearna kombinacija vektorjev $\mathbf{v}_1, \ldots, \mathbf{v}_m$.

Opomba: Izkaže se, da ima vsaka baza v $\mathbb{R}^n$ natanko n elementov.

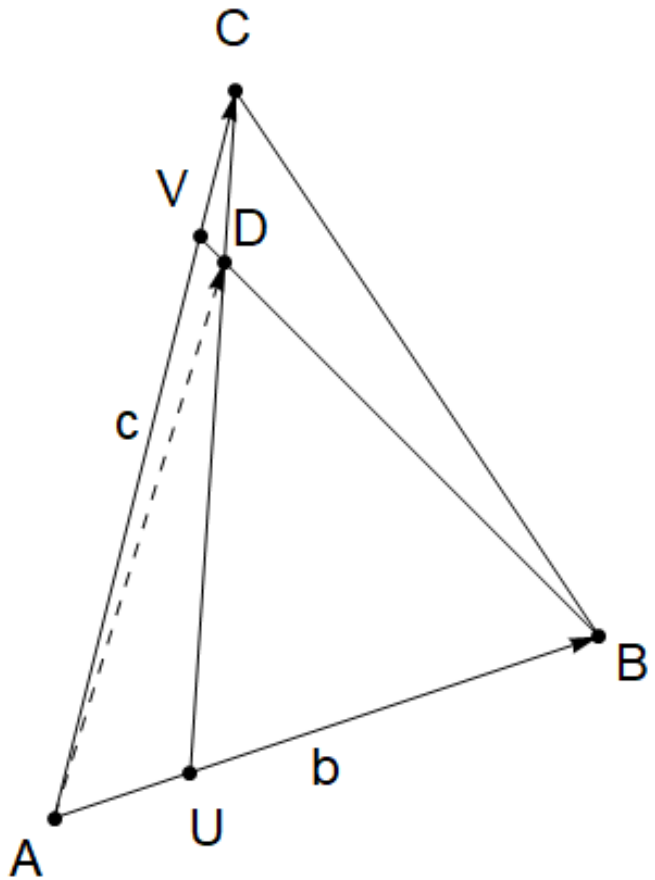
Primer baze

Standardna baza za $\mathbb{R}^n$ je

$$\mathbf{e}_1 = (1, 0, \ldots, 0),\ \mathbf{e}_2 = (0, 1, \ldots, 0),\ \ldots,\ \mathbf{e}_n = (0, 0, \ldots, 1),$$

Primer uporabe linearne neodvisnosti

Vzemimo trikotnik s oglišči A, B, C. Naj bo U taka točka na daljici AB, da je $\overline{AU} : \overline{UB} = 1 : 3$ in naj bo V taka točka na daljici AC, da velja $\overline{AV} : \overline{VC} = 4 : 1$. Naj bo točka D presečišče daljic CU in BV. (Glej sliko). Izrazimo vektor $\overrightarrow{AD}$ z vektorjema $\mathbf{b} = \overrightarrow{AB}$ in $\mathbf{c} = \overrightarrow{AC}$!



Ker je $\overrightarrow{AU} = \frac{1}{4}\mathbf{b}$, $\overrightarrow{AV} = \frac{4}{5}\mathbf{c}$, $\overrightarrow{UC} = \mathbf{c} - \frac{1}{4}\mathbf{b}$ in $\overrightarrow{VB} = \mathbf{b} - \frac{4}{5}\mathbf{c}$, velja

$$\overrightarrow{AD} = \overrightarrow{AU} + \lambda\overrightarrow{UC} = \frac{1}{4}\mathbf{b} + \lambda(\mathbf{c} - \frac{1}{4}\mathbf{b}) = \frac{1-\lambda}{4}\mathbf{b} + \lambda\mathbf{c}$$

in

$$\overrightarrow{AD} = \overrightarrow{AV} + \mu\overrightarrow{VB} = \frac{4}{5}\mathbf{c} + \mu(\mathbf{b} - \frac{4}{5}\mathbf{c}) = \mu\mathbf{b} + \frac{4(1-\mu)}{5}\mathbf{c},$$

kjer skalarjev λ in μ še ne poznamo. Ker sta $\mathbf{b}$ in $\mathbf{c}$ linearno neodvisna, odtod sledi, da je

$$\frac{1-\lambda}{4} = \mu \quad \text{in} \quad \lambda = \frac{4(1-\mu)}{5}.$$

Rešitev tega sistema je $\lambda = \frac{3}{4}$ in $\mu = \frac{1}{16}$. Torej je $\overrightarrow{AD} = \frac{1}{16}\mathbf{b} + \frac{3}{4}\mathbf{c}$.

Vektorji v $\mathbb{R}^n$, 2.del

Norma

Definicija norme

Norma vektorja $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ je skalar

$$\|\mathbf{x}\| = \sqrt{\sum_{i=1}^{n} x_i^2}.$$

Geometrijski pomen: Po Pitagorovem izreku je $\|\mathbf{x}\|$ ravno oddaljenost točke $\mathbf{x}$ od izhodišča, razdaljo med dvema točkama $\mathbf{x}$ in $\mathbf{y}$ pa lahko izrazimo kot $\|\mathbf{y} - \mathbf{x}\|$. Namesto $\|\overrightarrow{AB}\|$ pišemo raje $\overline{AB}$.

Osnovne lastnosti norme so:

- $\|\mathbf{x}\| \geq 0$ za vsak $\mathbf{x} \in \mathbb{R}^n$. Enačaj velja natanko tedaj, ko je $\mathbf{x} = \mathbf{0}$.
- $\|\alpha\mathbf{x}\| = |\alpha|\|\mathbf{x}\|$ za vsak $\alpha \in \mathbb{R}$ in $\mathbf{x} \in \mathbb{R}^n$.
- $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ za vse $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

Prvi dve lastnosti sledita direktno iz definicije norme. Tretjo lastnost bomo dokazali kasneje. Pravimo ji **trikotniška neenakost**.

Skalarni produkt

Definicija skalarnega produkta

Skalarni produkt vektorjev $\mathbf{x} = (x_1, \ldots, x_n)$ in $\mathbf{y} = (y_1, \ldots, y_n)$ je skalar

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^{n} x_i y_i.$$

Osnovne lastnosti skalarnega produkta. Za vse α, β in vse $\mathbf{x}, \mathbf{y}, \mathbf{z}$ velja:

- $\langle \mathbf{x}, \mathbf{x} \rangle = \|\mathbf{x}\|^2$ (**pozitivna definitnost**),
- $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$ (**simetričnost**),
- $\langle \alpha\mathbf{x} + \beta\mathbf{y}, \mathbf{z} \rangle = \alpha\langle \mathbf{x}, \mathbf{z} \rangle + \beta\langle \mathbf{y}, \mathbf{z} \rangle$ (**linearnost** v prvem faktorju),
  $\langle \mathbf{x}, \alpha\mathbf{y} + \beta\mathbf{z} \rangle = \alpha\langle \mathbf{x}, \mathbf{y} \rangle + \beta\langle \mathbf{x}, \mathbf{z} \rangle$ (**linearnost** v drugem faktorju).

Te lastnosti preverimo z direktnim računom.

Oglejmo si preprost primer uporabe skalarnega produkta.

Primer (Paralelogramsko pravilo)

Dokažimo, da za vsak paralelogram $ABCD$ velja

$$\overline{AC}^2 + \overline{BD}^2 = 2(\overline{AB}^2 + \overline{AD}^2).$$

Pišimo $\mathbf{a} = \overrightarrow{AB}$ in $\mathbf{b} = \overrightarrow{AD}$. Potem je

$$\overline{AB} = \|\mathbf{a}\|, \quad \overline{AD} = \|\mathbf{b}\|, \quad \overline{AC} = \|\mathbf{a}+\mathbf{b}\| \quad \text{in} \quad \overline{BD} = \|\mathbf{a}-\mathbf{b}\|.$$

Formula sedaj sledi iz računa

$$\|\mathbf{a}+\mathbf{b}\|^2 + \|\mathbf{a}-\mathbf{b}\|^2 = \langle \mathbf{a}+\mathbf{b}, \mathbf{a}+\mathbf{b}\rangle + \langle \mathbf{a}-\mathbf{b}, \mathbf{a}-\mathbf{b}\rangle =$$
$$= (\langle \mathbf{a},\mathbf{a}\rangle + \langle \mathbf{b},\mathbf{b}\rangle + 2\langle \mathbf{a},\mathbf{b}\rangle) + (\langle \mathbf{a},\mathbf{a}\rangle + \langle \mathbf{b},\mathbf{b}\rangle - 2\langle \mathbf{a},\mathbf{b}\rangle) =$$
$$= 2(\langle \mathbf{a},\mathbf{a}\rangle + \langle \mathbf{b},\mathbf{b}\rangle) = 2(\|\mathbf{a}\|^2 + \|\mathbf{b}\|^2).$$

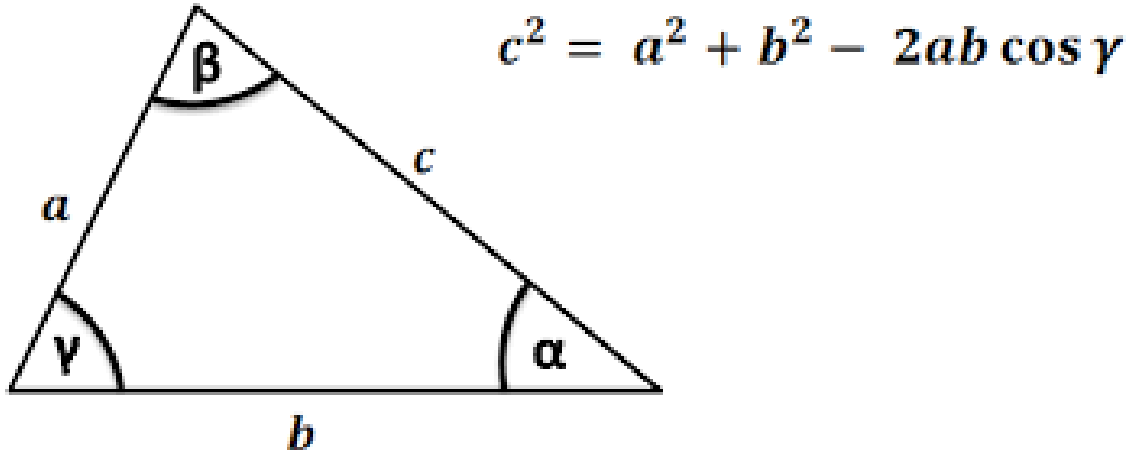
Oglejmo si sedaj, kako si skalarni produkt geometrijsko predstavljamo.

Trditev

Naj bo ϕ kot med vektorjema $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Potem velja

$$\langle \mathbf{x}, \mathbf{y} \rangle = \|\mathbf{x}\| \, \|\mathbf{y}\| \cos \phi.$$

Dokaz: V kosinusni izrek (slika spodaj) vstavimo $a = \|\mathbf{x}\|$, $b = \|\mathbf{y}\|$, $c = \|\mathbf{x} - \mathbf{y}\|$ in $\gamma = \phi$. Dobimo $\|\mathbf{x} - \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2\|\mathbf{x}\|\|\mathbf{y}\| \cos \phi$. Iz osnovnih lastnosti skalarnega produkta dobimo $\|\mathbf{x} - \mathbf{y}\|^2 = \langle \mathbf{x} - \mathbf{y}, \mathbf{x} - \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{x} \rangle + \langle \mathbf{y}, \mathbf{y} \rangle - \langle \mathbf{x}, \mathbf{y} \rangle - \langle \mathbf{y}, \mathbf{x} \rangle = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2\langle \mathbf{x}, \mathbf{y} \rangle$. Odštejemo obe formuli in krajšamo 2, pa dobimo trditev.



$$c^2 = a^2 + b^2 - 2ab \cos \gamma$$

Posledica (Kriterij za pravokotnost)

Vektorja $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ sta pravokotna natanko tedaj, ko je $\langle \mathbf{x}, \mathbf{y} \rangle = 0$.

Dokaz: Če sta oba vektorja neničelna, potem je $\langle \mathbf{x}, \mathbf{y} \rangle = 0$ natanko tedaj, ko je $\cos \phi = 0$. Slednje velja natanko tedaj, ko je $\phi = \frac{\pi}{2} + k\pi$, se pravi, ko je ϕ pravi kot. Če je vsaj eden od vektorjev ničeln, drži oboje.

Posledica (Cauchy-Schwartzova neenakost)

Za vsaka $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ velja $|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\| \, \|\mathbf{y}\|$.

Dokaz: Oceno $-1 \leq \cos \phi \leq 1$ pomnožimo z $\|\mathbf{x}\| \, \|\mathbf{y}\|$ in uporabimo trditev.

Posledica (Trikotniška neenakost)

Za vsaka $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ velja $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$.

Dokaz: Če Cauchy-Schwartzovo neenakost $-\|\mathbf{x}\| \, \|\mathbf{y}\| \leq \langle \mathbf{x}, \mathbf{y} \rangle \leq \|\mathbf{x}\| \, \|\mathbf{y}\|$ pomnožimo z 2, prištejemo $\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$ in korenimo, dobimo neenakost

$$|\|\mathbf{x}\| - \|\mathbf{y}\|| \leq \|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$$

Vektorski produkt

Za dva vektorja iz $\mathbb{R}^3$ lahko definiramo tudi njun **vektorski produkt**. Če je

$$\mathbf{x} = (x_1, x_2, x_3) \quad \text{in} \quad \mathbf{y} = (y_1, y_2, y_3),$$

potem definiramo

$$\mathbf{x} \times \mathbf{y} = (x_2y_3 - x_3y_2, x_3y_1 - x_1y_3, x_1y_2 - x_2y_1).$$

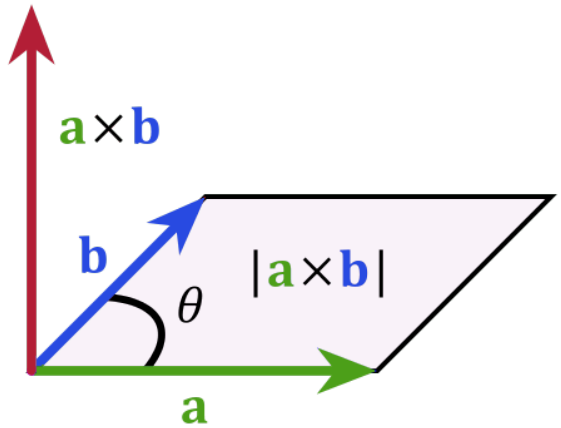
Osnovne algebrske lastnosti vektorskega produkta so:

- $\mathbf{x} \times \mathbf{x} = 0$ za vsak $\mathbf{x} \in \mathbb{R}^3$,
- $\mathbf{y} \times \mathbf{x} = -(\mathbf{x} \times \mathbf{y})$ za vsak $\mathbf{x} \in \mathbb{R}^3$ in $\mathbf{y} \in \mathbb{R}^3$,
- $(\alpha\mathbf{x} + \beta\mathbf{y}) \times \mathbf{z} = \alpha(\mathbf{x} \times \mathbf{z}) + \beta(\mathbf{y} \times \mathbf{z})$ za vsak $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^3$ in $\alpha, \beta \in \mathbb{R}$.

Te lastnosti preverimo z direktnim računom.

Opomba: Vektorskega produkta se ne da posplošiti na $n \neq 3$. Pri $n = 2$ mu je nekoliko podoben **vnanji produkt** $(x_1, x_2) \wedge (y_1, y_2) = x_1y_2 - x_2y_1$.

Kako si vektorski produkt geometrijsko predstavljamo?



- Dolžina $\mathbf{x} \times \mathbf{y}$ je enaka ploščini paralelograma, ki ga oklepata $\mathbf{x}$ in $\mathbf{y}$. Torej $\|\mathbf{x} \times \mathbf{y}\| = \|\mathbf{x}\|\|\mathbf{y}\| \sin \phi$, kjer je ϕ kot med vektorjema $\mathbf{x}$ in $\mathbf{y}$.
- Vektor $\mathbf{x} \times \mathbf{y}$ leži na premici skozi izhodišče, ki je pravokotna tako na vektor $\mathbf{x}$ kot na vektor $\mathbf{y}$.
- Smer vektorja $\mathbf{x} \times \mathbf{y}$ določimo s **pravilom desnega vijaka**. Desni mezinec položimo na vektor $\mathbf{x}$ tako, da prsti kažejo v smeri vektorja $\mathbf{y}$. Potem palec kaže v smeri vektorja $\mathbf{x} \times \mathbf{y}$.

Najprej izpeljemo $\|\mathbf{x} \times \mathbf{y}\|^2 + \langle \mathbf{x}, \mathbf{y} \rangle^2 = \|\mathbf{x}\|^2 \|\mathbf{y}\|^2$, kar nam da prvo lastnost, potem pa še $\langle \mathbf{x} \times \mathbf{y}, \mathbf{x} \rangle = \langle \mathbf{x} \times \mathbf{y}, \mathbf{y} \rangle = 0$, odkoder sledi druga lastnost.

Mešani produkt

Definicija mešanega produkta

Mešani produkt za tri vektorje iz $\mathbb{R}^3$ definiramo z

$$[\mathbf{x}, \mathbf{y}, \mathbf{z}] = \langle \mathbf{x} \times \mathbf{y}, \mathbf{z} \rangle.$$

Njegov geometrijski pomen je, da je (do predznaka) enak prostornini paralelepipeda, ki ga razpenjajo vektorji $\mathbf{x}, \mathbf{y}$ in $\mathbf{z}$. Velja namreč

$$[\mathbf{x}, \mathbf{y}, \mathbf{z}] = \|\mathbf{x} \times \mathbf{y}\| \, \|\mathbf{z}\| \cos\theta,$$

kjer je θ kot med $\mathbf{x} \times \mathbf{y}$ in $\mathbf{z}$. Osnovna ploskev paralelepipeda je paralelogram s stranicama $\mathbf{x}$ in $\mathbf{y}$. Njegova ploščina je $\|\mathbf{x} \times \mathbf{y}\|$. Višina paralelepipeda je do predznaka enaka $\|\mathbf{z}\| \cos\theta$. Če $\mathbf{z}$ leži na isti strani osnovne ploskve kot $\mathbf{x} \times \mathbf{y}$, potem se predznaka ujemata, sicer pa ne.

Njegov algebraičen pomen je, da je enak 3×3 determinanti z vrsticami $\mathbf{x}, \mathbf{y}, \mathbf{z}$. To bomo obravnavali pri geometrijskem pomenu determinante.

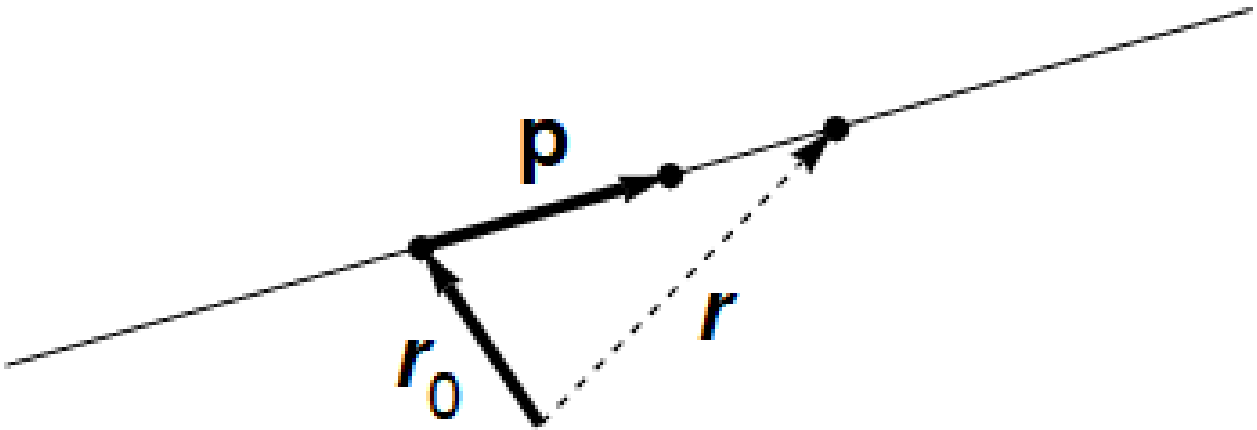
Osnovni lastnosti mešanega produkta sta linearnost v vsakem faktorju in poševna simetričnost (če zamenjaš dva faktorja, se spremeni predznak.)

Premice v $\mathbb{R}^n$

Premico v $\mathbb{R}^n$ najpogosteje podamo s točko na premici in vektorjem v smeri premice. Če je $\mathbf{r}_0$ dana točka na premici in $\mathbf{p}$ dani vektor v smeri premice, potem poljubno točko $\mathbf{r}$ na tej premici izrazimo kot

$$\mathbf{r} = \mathbf{r}_0 + t\mathbf{p},$$

kjer je t poljubno realno število. Spremenljivki t pravimo **parameter**, enačbi pa **parametrična enačba** premice.



Če je $\mathbf{r}_0 = (x_{01}, \ldots, x_{0n})$ in $\mathbf{p} = (p_1, \ldots, p_n)$, potem lahko parametrično enačbo zapišemo po komponentah kot

$$x_1 = x_{01} + tp_1, \ldots, x_n = x_{0n} + tp_n.$$

Če se požvižgamo na morebitno deljenje z nič, potem lahko zapišemo

$$t = \frac{x_1 - x_{01}}{p_1} = \ldots = \frac{x_n - x_{0n}}{p_n}.$$

Tej enačbi pravimo **normalna enačba** premice.

Če je premica podana z normalno enačbo, jo najprej pretvorimo v parametrično enačbo, saj je ta najprimernejša za računanje.

Primer

Poiščimo normalno enačbo premice v $\mathbb{R}^3$, ki gre skozi točki

$$\mathbf{r}_1 = (0, -2, 1) \quad \text{in} \quad \mathbf{r}_2 = (3, -2, -1).$$

Najprej poiščemo točko $\mathbf{r}_0$ na premici in vektor $\mathbf{p}$ v smeri premice. Vzemimo na primer

$$\mathbf{r}_0 = \mathbf{r}_1 = (0, -2, 1) \quad \text{in} \quad \mathbf{p} = \mathbf{r}_2 - \mathbf{r}_1 = (3, 0, -2).$$

Normalna enačba se potem glasi

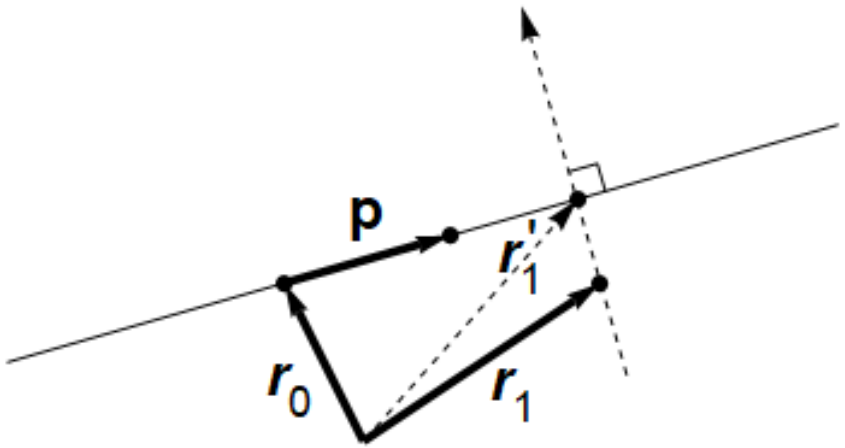
$$t = \frac{x_1}{3} = \frac{x_2 + 2}{0} = \frac{x_3 - 1}{-2}.$$

Osnovne naloge s premicami

Najpogostejše naloge s premicami so:

- pravokotna projekcija točke na premico,
- zrcaljenje točke čez premico in
- razdalja točke od premice.

Recimo torej, da bi radi pravokotno projecirali točko $\mathbf{r}_1$ na premico $\mathbf{r} = \mathbf{r}_0 + t\mathbf{p}$. Iskano točko označimo z $\mathbf{r}_1'$.



Ker točka $\mathbf{r}_1'$ leži na premici, obstaja tak skalar t, da velja

$$\mathbf{r}_1' = \mathbf{r}_0 + t\mathbf{p}.$$

Skalar t moramo določiti tako, da je vektor $\mathbf{r}_1 - \mathbf{r}_1'$ pravokoten na vektor $\mathbf{p}$. Potem je $0 = \langle \mathbf{r}_1 - \mathbf{r}_1', \mathbf{p} \rangle = \langle \mathbf{r}_1 - \mathbf{r}_0, \mathbf{p} \rangle - t\langle \mathbf{p}, \mathbf{p} \rangle$. Odtod sledi, da je

$$t = \frac{\langle \mathbf{r}_1 - \mathbf{r}_0, \mathbf{p} \rangle}{\langle \mathbf{p}, \mathbf{p} \rangle}.$$

Torej je

$$\mathbf{r}_1' = \mathbf{r}_0 + \frac{\langle \mathbf{r}_1 - \mathbf{r}_0, \mathbf{p} \rangle}{\langle \mathbf{p}, \mathbf{p} \rangle}\mathbf{p}.$$

Sedaj lahko izračunamo tudi oddaljenost točke $\mathbf{r}_1$ od premice $\mathbf{r} = \mathbf{r}_0 + t\mathbf{p}$ in njeno zrcalno sliko $\mathbf{r}_1''$ glede na to premico. Velja

$$d = \|\mathbf{r}_1 - \mathbf{r}_1'\| \quad \text{in} \quad \mathbf{r}_1'' = 2\mathbf{r}_1' - \mathbf{r}_1.$$

Zadnja formula sledi iz dejstva, da je $\mathbf{r}_1'$ ravno razpolovišče daljice med $\mathbf{r}_1$ in $\mathbf{r}_1''$, se pravi $\mathbf{r}_1' = \frac{1}{2}(\mathbf{r}_1 + \mathbf{r}_1'')$.

Kadar smo v $\mathbb{R}^3$, lahko pri računanju oddaljenosti točke $\mathbf{r}_1$ od premice $\mathbf{r} = \mathbf{r}_0 + t\mathbf{p}$ uporabimo vektorski produkt. Velja

$$d = \frac{\|(\mathbf{r}_1 - \mathbf{r}_0) \times \mathbf{p}\|}{\|\mathbf{p}\|}.$$

Tako leva kot desna stran sta namreč enaki $\|\mathbf{r}_1 - \mathbf{r}_0\| \sin\phi$, kjer je ϕ kot med vektorjema $\mathbf{r}_1 - \mathbf{r}_0$ in $\mathbf{p}$.

Ravnine in hiperravnine v $\mathbb{R}^n$

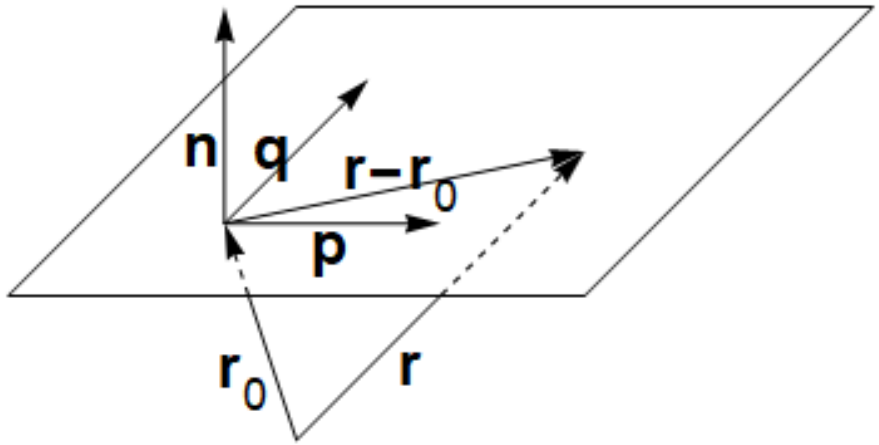
Ravnino v $\mathbb{R}^3$ lahko podamo na dva načina:

- S točko $\mathbf{r}_0$ na ravnini in dvema linearno neodvisnima smernima vektorjema $\mathbf{p}$ in $\mathbf{q}$. V tem primeru poljubna točka $\mathbf{r}$ na ravnini zadošča **parametrični enačbi**

$$\mathbf{r} = \mathbf{r}_0 + s\mathbf{p} + t\mathbf{q}.$$

- S točko $\mathbf{r}_0$ na ravnini in z neničelno normalo $\mathbf{n}$. V tem primeru poljubna točka $\mathbf{r}$ na ravnini zadošča **normalni enačbi**

$$\langle \mathbf{n}, \mathbf{r} - \mathbf{r}_0 \rangle = 0.$$



Primer s parametrično in normalno enačbo ravnine

Poiščimo normalno enačbo ravnine v $\mathbb{R}^3$, ki ima parametrično enačbo

$$\mathbf{r} = (1, 2, -1) + s(2, 1, 3) + t(-1, 1, 1).$$

Zapišimo enačbo po komponentah:

$$x = 1 + 2s - t, \quad y = 2 + s + t, \quad z = -1 + 3s + t.$$

Iz prve enačbe izrazimo s in ga vstavimo v drugo in tretjo enačbo. Dobimo

$$\frac{3}{2}t = -\frac{1}{2}x + y - \frac{3}{2} \quad \text{in} \quad \frac{5}{2}t = -\frac{3}{2}x + z + \frac{5}{2}.$$

Iz prve enačbe izrazimo t in ga vstavimo v drugo enačbo. Dobimo

$$0 = -\frac{2}{3}x - \frac{5}{3}y + z + 5.$$

To je iskana normalna enačba ravnine.

Prvi način podajanja ravnin deluje tudi v $\mathbb{R}^n$. Parametrična enačba

$$\mathbf{r} = \mathbf{r}_0 + s\mathbf{p} + t\mathbf{q}$$

določa ravnino v $\mathbb{R}^n$, ki gre skozi točke $\mathbf{r}_0$, $\mathbf{r}_0 + \mathbf{p}$, $\mathbf{r}_0 + \mathbf{q}$.

Drugi način podajanja ravnin se ne posploši na $\mathbb{R}^n$. Enačba

$$\langle \mathbf{n}, \mathbf{r} - \mathbf{r}_0 \rangle = 0$$

določa neko $n-1$ razsežno podmnožico v $\mathbb{R}^n$, ki ji pravimo **hiperravnina**.

Opomba 1: Hiperravnine v $\mathbb{R}$ so točke. Hiperravnine v $\mathbb{R}^2$ so premice. Hiperravnine v $\mathbb{R}^3$ so ravnine. Hiperravnine v $\mathbb{R}^4$ so trirazsežne množice.

Opomba 2: Hiperravnine v $\mathbb{R}^n$ so ravno množice rešitev netrivialnih linearnih enačb v n spremenljivkah. Če v enačbo hiperravnine vstavimo $\mathbf{n} = (a_1, \ldots, a_n)$, $\mathbf{r}_0 = (x_{1,0}, \ldots, x_{n,0})$ in $\mathbf{r} = (x_1, \ldots, x_n)$, dobimo linearno enačbo $a_1x_1 + \ldots + a_nx_n = b$, kjer je $b = a_1x_{1,0} + \ldots + a_nx_{n,0}$. Velja tudi obratno. Vsako linearno enačbo $a_1x_1 + \ldots + a_nx_n = b$, kjer $a_i \neq 0$ za nek i, lahko zapišemo v obliki $\langle \mathbf{n}, \mathbf{r} - \mathbf{r}_0 \rangle = 0$, kjer je $\mathbf{r}_0 = (0, \ldots, \frac{b}{a_i}, \ldots, 0)$.

Osnovne naloge z ravninami in hiperravninami

Najprej si poglejmo, kako poiščemo pravokotno projekcijo točke $\mathbf{r}_1$ na parametrično podano ravnino $\mathbf{r} = \mathbf{r}_0 + s\mathbf{p} + t\mathbf{q}$. Iskana točka naj bo $\mathbf{r}_1'$.

Ker točka $\mathbf{r}_1'$ leži na ravnini, obstajata taka parametra s in t, da velja

$$\mathbf{r}_1' = \mathbf{r}_0 + s\mathbf{p} + t\mathbf{q}. \tag{1}$$

Ker je vektor $\mathbf{r}_1' - \mathbf{r}_1$ pravokoten na ravnino, je pravokoten tudi na oba smerna vektorja, torej mora veljati:

$$\langle \mathbf{r}_1' - \mathbf{r}_1, \mathbf{p} \rangle = 0 \quad \text{in} \quad \langle \mathbf{r}_1' - \mathbf{r}_1, \mathbf{q} \rangle = 0 \tag{2}$$

Ko enačbo (1) vstavimo v enačbi (2) in uredimo, dobimo enačbi:

$$\begin{aligned} s\langle \mathbf{p}, \mathbf{p} \rangle + t\langle \mathbf{q}, \mathbf{p} \rangle &= \langle \mathbf{r}_1 - \mathbf{r}_0, \mathbf{p} \rangle \\ s\langle \mathbf{p}, \mathbf{q} \rangle + t\langle \mathbf{q}, \mathbf{q} \rangle &= \langle \mathbf{r}_1 - \mathbf{r}_0, \mathbf{q} \rangle \end{aligned} \tag{3}$$

Rešimo sistem (3) po s in t in vstavimo rezultat v (1). Dobimo $\mathbf{r}_1'$.

Poglejmo si še, kako poiščemo pravokotno projekcijo točke $\mathbf{r}_1$ na implicitno podano hiperravnino $\langle \mathbf{n}, \mathbf{r} - \mathbf{r}_0 \rangle = 0$. Iskana točka naj bo $\mathbf{r}_1'$.

Iskana točka $\mathbf{r}_1'$ leži na preseku hiperravnine in premice, ki gre skozi točko $\mathbf{r}_1$ in je pravokotna na hiperravnino. Velja torej

$$\langle \mathbf{n}, \mathbf{r}_1' - \mathbf{r}_0 \rangle = 0 \tag{4}$$

in obstaja tak parameter t, da velja:

$$\mathbf{r}_1' = \mathbf{r}_1 + t\mathbf{n} \tag{5}$$

Vstavimo enačbo (5) v enačbo (4) in izrazimo t. Dobimo

$$t = -\frac{\langle \mathbf{r}_1 - \mathbf{r}_0, \mathbf{n} \rangle}{\langle \mathbf{n}, \mathbf{n} \rangle}. \tag{6}$$

Iskano točko $\mathbf{r}_1'$ dobimo tako, da skalar (6) vstavimo v enačbo (5).

Primer projekcije točke na ravnino

Izračunaj pravokotno projekcijo točke $(1,1,1)$ na ravnino $x+2y+3z=4$.

V formulo

$$\mathbf{r}_1' = \mathbf{r}_1 - \frac{\langle \mathbf{r}_1 - \mathbf{r}_0, \mathbf{n} \rangle}{\langle \mathbf{n}, \mathbf{n} \rangle} \mathbf{n}$$

vstavimo $\mathbf{n} = (1,2,3)$, $\mathbf{r}_0 = (4,0,0)$, $\mathbf{r}_1 = (1,1,1)$ in dobimo

$$\mathbf{r}_1' = \frac{1}{7}(6,5,4).$$

Sedaj lahko določimo tudi oddajenost točke $\mathbf{r}_1$ od (hiper)ravnine

$$d = \|\mathbf{r}_1 - \mathbf{r}_1'\|$$

in pa zrcalno sliko točke $\mathbf{r}_1$ glede na (hiper)ravnino.

$$\mathbf{r}_1'' = 2\mathbf{r}_1' - \mathbf{r}_1.$$

Sistemi linearnih enačb

Linearna enačba

Definicija linearne enačbe

Linearna enačba v n spremenljivkah je enačba oblike

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n = b, \tag{1}$$

kjer so $a_1, a_2, \ldots, a_n, b$ dana realna števila, $x_1, x_2, \ldots, x_n$ pa so spremenljivke. Vsaki n-terici realnih števil, ki zadoščajo (1), pravimo **rešitev** enačbe (1).

Opomba: Enačba $2x = 1$ ni isto kot enačba $2x + 0y = 1$. Množica rešitev prve je točka v $\mathbb{R}$, množica rešitev druge pa je premica v $\mathbb{R}^2$.

Opomba: Če je $a_1 = \ldots = a_n = 0$, potem pravimo, da je enačba (1) **trivialna**. Ločimo dva primera trivialne enačbe. Če je $b = 0$, potem je njena množica rešitev enaka $\mathbb{R}^n$. Če je $b \neq 0$, potem je njena množica rešitev prazna.

Trditev

Če je linearna enačba (1) netrivialna, potem je njena množica rešitev hiperravnina v $\mathbb{R}^n$. Vsako hiperravnino v $\mathbb{R}^n$ lahko dobimo na ta način.

Dokaz: Če je enačba (1) netrivialna, obstaja tak $i \in \{1, \ldots, n\}$, da je $a_i \neq 0$. Torej jo lahko zapišemo v obliki

$$a_1 x_1 + \ldots + a_i(x_i - \frac{b}{a_i}) + \ldots + a_n x_n = 0. \tag{2}$$

Opazimo, da je vektor $\mathbf{n} := (a_1, \ldots, a_i, \ldots, a_n)$ neničeln, ker je $a_i \neq 0$. Če definiramo še $\mathbf{r} := (x_1, \ldots, x_i, \ldots, x_n)$ in $\mathbf{r}_0 := (0, \ldots, \frac{b}{a_i}, \ldots, 0)$, potem je enačba (2) ravno enačba hiperravnine $\langle \mathbf{n}, \mathbf{r} - \mathbf{r}_0 \rangle = 0$.

Vzemimo sedaj poljubno hiperravnino v $\mathbb{R}^n$. Potem obstajata tak neničeln $\mathbf{n} = (a_1, \ldots, a_n)$ in tak $\mathbf{r}_0 = (c_1, \ldots, c_n)$ v $\mathbb{R}^n$, da je hiperravnina množica rešitev vektorske enačbe $\langle \mathbf{n}, \mathbf{r} - \mathbf{r}_0 \rangle = 0$. Če razpišemo, dobimo netrivialno linearno enačbo $a_1 x_1 + \ldots + a_n x_n = b$, kjer je $b = a_1 c_1 + \ldots + a_n c_n$.

Uvod v sisteme linearnih enačb

Definicija sistema linearnih enačb

Sistem m **linearnih enačb v** n **spremenljivkah** je sistem enačb oblike

$$\begin{aligned} a_{1,1}x_1 + \ldots + a_{1,n}x_n &= b_1 \\ \vdots \qquad\qquad & \vdots \quad \vdots \\ a_{m,1}x_1 + \ldots + a_{m,n}x_n &= b_m \end{aligned} \tag{3}$$

kjer so $a_{i,j}$ in b_k realna števila, $x_1, \ldots, x_n$ pa spremenljivke.

Rešitev sistema (3) je taka n-terica realnih števil, ki zadošča vsaki od m linearnih enačb v sistemu.

Opomba: Sistemu m linearnih enačb v n spremenljivkah bomo na kratko rekli kar $m \times n$ **sistem.**

Opomba: Številom $a_{i,j}$ bomo rekli **koeficienti** sistema (3), številom b_k pa **desne strani** sistema (3).

Oglejmo si najprej geometrijski pomen množice rešitev sistema.

Trditev

Če so vsi koeficienti in vse desne strani $m \times n$ sistema enaki nič, potem je njegova množica rešitev enaka $\mathbb{R}^n$. V vseh ostalih primerih je množica rešitev presek hiperravnin v $\mathbb{R}^n$. To vključuje tudi primer, ko je množica rešitev prazna, saj je prazna množica presek dveh vzporednih hiperravnin.

Dokaz: Množica rešitev sistema je presek množic rešitev posameznih enačb. Vemo že, da je množica rešitev vsake netrivialne enačbe hiperravnina. Če so torej vse enačbe v sistemu netrivialne, je njegova množica rešitev presek hiperravnin.

Oglejmo si še kako trivialne enačbe vplivajo na množici rešitev. Če ima ena od trivialnih enačb sistema prazno množico rešitev, potem ima tudi sistem prazno množico rešitev, torej trditev drži.

Oglejmo si še nasprotni primer, ko imajo vse trivialne enačbe sistema za množico rešitev $\mathbb{R}^n$. Take trivialne enačbe lahko izpustimo, saj ne vplivajo na množico rešitev sistema. Torej trditev drži tudi v tem primeru.

Primeri 2×2 sistemov

Množica rešitev 2×2 sistema je lahko

- cela ravnina (samo pri sistemu $0x + 0y = 0, 0x + 0y = 0$)
- premica v ravnini (npr. pri sistemu $2x + y = 1, 4x + 2y = 2$ prva enačba določa isto premico kot druga enačba, torej je množica rešitev sistema presek dveh enakih premic.)
- točka v ravnini (npr. pri sistemu $x + y = 1, x - y = 3$ je množica rešitev presek dveh nevzporednih premic)
- prazna množica (npr. pri sistemu $2x + 1 = 1, 4x + 2y = 3$) je množica rešitev presek dveh različnih vzporednih premic).

Opomba: Tudi množica rešitev 3×2 sistema je bodisi cela ravnina bodisi premica v ravnini bodisi točka v ravnini bodisi prazna množica.

Opomba: Množica rešitev 2×3 sistema je bodisi cel prostor $\mathbb{R}^3$ bodisi ravnina v prostoru bodisi premica v prostoru bodisi prazna množica. Ne more pa biti množica rešitev 2×3 sistema točka v prostoru.

Oglejmo si še algebraičen pomen rešitev sistema

$$\begin{aligned} a_{1,1}x_1 + \ldots + a_{1,n}x_n &= b_1 \\ \vdots \qquad\qquad\quad & \quad \vdots \quad \vdots \\ a_{m,1}x_1 + \ldots + a_{m,n}x_n &= b_m \end{aligned}$$

Ta sistem lahko zapišemo kot

$$\begin{aligned} (b_1, \ldots, b_n) &= (a_{1,1}x_1 + \ldots + a_{1,n}x_n, \ldots, a_{m,1}x_1 + \ldots + a_{m,n}x_n) \\ &= x_1(a_{1,1}, \ldots, a_{m,1}) + \ldots + x_n(a_{1,n}, \ldots, a_{m,n}), \end{aligned}$$

se pravi kot razvoj v linearno kombinacijo

$$\mathbf{b} = x_1\mathbf{a}_1 + \ldots + x_n\mathbf{a}_n,$$

kjer $\mathbf{b} = (b_1, \ldots, b_n)$ in $\mathbf{a}_1 = (a_{1,1}, \ldots, a_{m,1}), \ldots, \mathbf{a}_m = (a_{1,n}, \ldots, a_{m,n})$. Vektorju $\mathbf{b}$ pravimo tudi **vektor desnih strani** sistema, vektorjem $\mathbf{a}_1, \ldots, \mathbf{a}_n$ pa tudi **stolpci** sistema. Množica rešitev sistema nam da torej vse možne razvoje vektorja desnih strani po stolpcih sistema.

Klasifikacija sistemov linearnih enačb

Sisteme linearnih enačb lahko klasificiramo glede na velikost, rešljivost in obliko desnih strani.

Glede na velikost delimo sisteme na

- kvadratne (število enačb je enako številu spremenljivk),
- predoločene (število enačb je večje od števila spremenljivk),
- poddoločene (število enačb je manjše od števila spremenljivk).

Glede na rešljivost delimo sisteme na

- nerešljive (če je množica rešitev prazna)
- enolično rešljive (če ima množica rešitev en element)
- neenolično rešljive (če ima množica rešitev več kot en element)

Glede na obliko desnih strani delimo sisteme na

- homogene (če so vse desne strani enake nič)
- nehomogene (če je vsaj ena desna stran različna od nič)

Opomba: Če sta $\mathbf{x}$ in $\mathbf{y}$ dve rešitvi istega sistema, potem je za vsak realen t tudi $(1-t)\mathbf{x}+t\mathbf{y}$ rešitev tega sistema. Vsak neenolično rešljiv sistem ima torej neskončno množico rešitev.

Opomba: Pogosto (nikakor pa ne vedno) je kvadraten sistem enolično rešljiv, predoločen sistem nerešljiv in poddoločen sistem neenolično rešljiv.

Opomba: Homogen sistem je vedno rešljiv. Ima namreč trivialno rešitev $(0,\ldots,0)$. Glavno vprašanje je, kdaj je neenolično rešljiv. Pokazali bomo, da to velja za vse poddoločene homogene sisteme.

Reševanje sistemov z izločanjem spremenljivk

Oglejmo si preprosto metodo za reševanje sistema linearnih enačb.

Prvi del: Najprej iz prve enačbe izrazimo eno spremenljivko in ta izraz vstavimo v vse nadaljnje enačbe. Dobimo nov sistem linearnih enačb z eno enačbo manj in eno spremenljivko manj. Metodo sedaj ponovimo na tem novem sistemu. Spet dobimo nov sistem linearnih enačb, ki ima dve enačbi manj in dve spremenljivki manj kot prvotni sistem. Postopek ponavljamo, dokler ne zmanjka bodisi enačb bodisi spremenljivk. Če nam zmanjka enačb, potem je prvotni sistem rešljiv. Če nam zmanjka spremenljivk, potem imamo na koncu eno ali več enačb oblike $0 = c$, kjer je $c \in R$. Če je katera od teh enačb protislovna (recimo $0 = 2$), potem končamo.

Drugi del: Izraz za zadnjo izraženo spremenljivko vstavimo v vse izraze za prejšnje izražene spremenljivke. Nato izraz za predzadnjo izraženo spremenljivko vstavimo v vse izraze za prejšnje izražene spremenljivke. Nadaljujemo, dokler ne pridemo do prve izražene spremenljivke. Rezultat je, da smo nekatere od spremenljivk izrazili s preostalimi spremenljivkami.

Primer reševanja sistema z izločanjem spremenljivk

Rešimo naslednji sistem linearnih enačb

$$\begin{aligned} 2x - y + 3u - 2v &= 1, \\ x + y + 2u - 2v &= 0, \\ x + 2u - 3v &= -2. \end{aligned}$$

Iz prve enačbe izrazimo

$$x = \frac{1 + y - 3u + 2v}{2}.$$

Ko formulo za x vstavimo v drugo in tretjo enačbo, dobimo novi enačbi

$$\begin{aligned} \frac{3}{2}y + \frac{1}{2}u - v &= -\frac{1}{2}, \\ \frac{1}{2}y + \frac{1}{2}u - 2v &= -\frac{5}{2}. \end{aligned}$$

Iz prve nove enačbe dobimo

$$y = \frac{-\frac{1}{2} - \frac{1}{2}u + v}{\frac{3}{2}} = \frac{-1 - u + 2v}{3}.$$

Ko izraz za y vstavimo v drugo novo enačbo in uredimo, dobimo

$$\frac{1}{3}u - \frac{5}{3}v = -\frac{7}{3}.$$

Odtod izrazimo

$$u = \frac{-\frac{7}{3} + \frac{5}{3}v}{\frac{1}{3}} = -7 + 5v.$$

Sedaj nam je zmanjkalo enačb, torej je sistem rešljiv. Ko formulo za u vstavimo v formuli za y in x, dobimo

$$y = \frac{6 - 3v}{3} = 2 - v \quad \text{in} \quad x = \frac{22 + y - 13v}{2}.$$

Sedaj še formulo za y vstavimo v formulo za x in dobimo

$$x = \frac{24 - 14v}{2} = 12 - 7v.$$

Končna rešitev je $x = 12 - 7v$, $y = 2 - v$ in $u = -7 + 5v$, se pravi

$$\mathbf{r} = (x, y, u, v) = (12-7v, 2-v, -7+5v, v) = (12, 2, -7, 0)+v(-7, -1, 5, 1).$$

Množica rešitev je torej premica v $\mathbb{R}^4$. Sistem je neenolično rešljiv.

Gaussova metoda za reševanje sistemov

Množica rešitev sistema ne spremeni, če na njem uporabimo eno od naslednjih **elementarnih vrstičnih transformacij**:

- Zamenjamo vrstni red dveh enačb.
- Eno od enačb pomnožimo z neničelno konstanto.
- K eni od enačb prištejemo večkratnik druge enačbe.

Gaussova metoda nam pove, kako s pomočjo elementarnih vrstičnih transformacij prevedemo sistem v obliko iz katere se da odčitati rešitev.

Ker se pri elementarnih vrstičnih transformacijah nič ne dogaja z spremenljivkami, ni potrebe, da jih pišemo. Zato najprej pretvorimo:

$$\begin{array}{ccccc} a_{1,1}x_1 + \ldots + a_{1,n}x_n & = & b_1 \\ \vdots & & \vdots \\ a_{m,1}x_1 + \ldots + a_{m,n}x_n & = & b_m \end{array} \longrightarrow \left[\begin{array}{ccc|c} a_{1,1} & \ldots & a_{1,n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m,1} & \ldots & a_{m,n} & b_m \end{array}\right]$$

Temu zapisu pravimo **razširjena matrika sistema**.

Na razširjeni matriki napravimo naslednje zaporedje korakov:

- Poišči prvi stolpec, ki vsebuje kak neničeln element. Označimo z j_1 njegovo zaporedno številko. Velja torej $a_{i,j} = 0$, za vse (i, j) z $j < j_1$.
- Če je $a_{1,j_1} = 0$ in $a_{r,j_1} \neq 0$, potem zamenjamo prvo in r-to vrstico. Torej lahko predpostavimo, da je $a_{1,j_1} \neq 0$.
- Prvo vrstico delimo z a_{1,j_1}. Torej lahko predpostavimo, da je $a_{1,j_1} = 1$.
- Za vsak $i = 2, \ldots, m$ odštejemo od i-te vrstice z a_{i,j_1} pomnoženo prvo vrstico. Torej lahko predpostavimo, da je $a_{2,j_1} = \ldots = a_{m,j_1} = 0$.

S temi koraki smo razširjeno matriko prevedli v obliko

$$\left[\begin{array}{ccccccc|c} 0 & \ldots & 0 & 1 & a'_{1,j_1+1} & \ldots & a'_{1,n} & b'_1 \\ 0 & \ldots & 0 & 0 & a'_{2,j_1+1} & \ldots & a'_{2,n} & b'_2 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & \ldots & 0 & 0 & a'_{m,j_1+1} & \ldots & a'_{m,n} & b'_m \end{array}\right]$$

Sedaj za hip pozabimo na prvo vrstico in ponovimo gornje zaporedje korakov na zadnjih $m - 1$ vrsticah. Dobimo:

$$\left[\begin{array}{cccccccccc|c}
0 & \dots & 0 & 1 & a'_{1,j_1+1} & \dots & a'_{1,j_2-1} & a'_{1,j_2} & a'_{1,j_2+1} & \dots & a'_{1,n} & b'_1 \\
0 & \dots & 0 & 0 & 0 & \dots & 0 & 1 & a''_{2,j_2+1} & \dots & a''_{2,n} & b''_2 \\
0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & a''_{3,j_2+1} & \dots & a''_{3,n} & b''_3 \\
\vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots \\
0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & a''_{m,j_2+1} & \dots & a''_{m,n} & b''_m
\end{array}\right]$$

Sedaj za hip pozabimo na prvi dve vrstci in ponovimo gornje zaporedje korakov na preostalih $m-2$ vrsticah. Postopek ponavljamo, dokler gre. (Se pravi, dokler ne zmanjka neničelnih vrstic.) Na koncu dobimo:

$$\left[\begin{array}{ccccccccccccccc|c}
0 & \dots & 0 & 1 & \times & \dots & \times & \times & \times & \dots & \times & \times & \times & \dots & \times & \dots & \times \\
0 & \dots & 0 & 0 & 0 & \dots & 0 & 1 & \times & \dots & \times & \times & \times & \dots & \times & \dots & \times \\
0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 1 & \times & \dots & \times & \dots & \times \\
0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & \dots & \times \\
\vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & & \vdots \\
0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & \dots & \times
\end{array}\right]$$

Temu pravimo **vrstična stopničasta forma** sistema. Simbol $\times$ označuje, da je tam število, ki je lahko neničelno. Pod "stopniščem" so same ničle.

Opomba: Če vrstična stopničasta forma vsebuje vrstico oblike $[0 \dots 0 \mid \times]$, kjer je element $\times$ neničeln, potem sistem ni rešljiv. Torej končamo.

Recimo, da ima vrstična stopničasta forma k neničelnih vrstic. Recimo, da se začetne enke teh vrstic (=**pivoti**) nahajajo na mestih $(1, j_1), \dots, (k, j_k)$. Najprej z zadnjim pivotom "uničimo" vse elemente nad njim. (To pomeni, da od prvih $k-1$ vrstic odštejemo ustrezne večkratnike k-te vrstice, da se na mestih $(1, j_k), \dots, (k-1, j_k)$ pojavijo ničle.) Nato s predzadnjim pivotom uničimo vse elemente nad njim. Postopek ponavljamo, dokler ne pridemo do prvega pivota (nad katerim ne moremo nič uničiti). Dobimo:

$$\left[\begin{array}{ccccccccccccccc|c}
0 & \dots & 0 & 1 & \times & \dots & \times & 0 & \times & \dots & \times & 0 & \times & \dots & \times & 0 & \dots & \times \\
0 & \dots & 0 & 0 & 0 & \dots & 0 & 1 & \times & \dots & \times & 0 & \times & \dots & \times & 0 & \dots & \times \\
0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 1 & \times & \dots & \times & 0 & \dots & \times \\
0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 1 & \dots & \times \\
\vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots \\
0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & \dots & \times
\end{array}\right]$$

Temu pravimo **reducirana vrstična stopničasta forma** sistema. Sedaj spet pišemo spremenljivke. Iz dobljenih enačb izrazimo spremenljivke $x_{j_1}, \dots, x_{j_k}$ s pomočjo preostalih spremenljivk. To je iskana rešitev.

Primer reševanja sistema z Gaussovo metodo

Rešimo naslednji sistem linearnih enačb

$$\begin{aligned} 2x - y + 3u - 2v &= 1, \\ x + y + 2u - 2v &= 0, \\ x + 2u - 3v &= -2. \end{aligned}$$

Najprej zapišemo razširjeno matriko sistema:

$$\left[\begin{array}{cccc|c} 2 & -1 & 3 & -2 & 1 \\ 1 & 1 & 2 & -2 & 0 \\ 1 & 0 & 2 & -3 & -2 \end{array}\right]$$

Na poziciji $(1,1)$ (se pravi levo zgoraj) bi radi dobili 1. Namesto da prvo vrstico delimo z 2, bomo raje zamenjali prvo in drugo vrstico:

$$\left[\begin{array}{cccc|c} 1 & 1 & 2 & -2 & 0 \\ 2 & -1 & 3 & -2 & 1 \\ 1 & 0 & 2 & -3 & -2 \end{array}\right]$$

Spodnja dva elementa v prvem stolpcu bi radi spravili na nič. To naredimo tako, da k drugi vrstici prištejemo z -2 pomnoženo prvo vrstico in k tretji vrstici prištejemo z -1 pomnoženo prvo vrstico. Dobimo:

$$\left[\begin{array}{cccc|c} 1 & 1 & 2 & -2 & 0 \\ 0 & -3 & -1 & 2 & 1 \\ 0 & -1 & 0 & -1 & -2 \end{array}\right]$$

Na poziciji $(2,2)$ bi radi dobili enko. Najprej tretjo vrstico pomnožimo z -1 in nato zamenjamo drugo in tretjo vrstico. Dobimo:

$$\left[\begin{array}{cccc|c} 1 & 1 & 2 & -2 & 0 \\ 0 & 1 & 0 & 1 & 2 \\ 0 & -3 & -1 & 2 & 1 \end{array}\right]$$

Spodnji element v drugem stolpcu spravimo na nič tako, da k tretji vrstici prištejemo s 3 pomnoženo drugo vrstico. Dobimo:

$$\left[\begin{array}{cccc|c} 1 & 1 & 2 & -2 & 0 \\ 0 & 1 & 0 & 1 & 2 \\ 0 & 0 & -1 & 5 & 7 \end{array}\right]$$

Tretjo vrstico delimo z -1 in dobimo vrstično stopničasto formo:

$$\left[\begin{array}{cccc|c} 1 & 1 & 2 & -2 & 0 \\ 0 & 1 & 0 & 1 & 2 \\ 0 & 0 & 1 & -5 & -7 \end{array}\right]$$

Sedaj moramo dobiti še ničle na pozicijah $(2,3)$, $(1,3)$ in $(1,2)$.
Če k prvi vrstici prištejemo z -2 pomnoženo tretjo vrstico, dobimo:

$$\left[\begin{array}{cccc|c} 1 & 1 & 0 & 8 & 14 \\ 0 & 1 & 0 & 1 & 2 \\ 0 & 0 & 1 & -5 & -7 \end{array}\right]$$

Na koncu k prvi vrstici prištejemo z -1 pomnoženo drugo vrstico.
Dobimo reducirano vrstično stopničasto formo:

$$\left[\begin{array}{cccc|c} 1 & 0 & 0 & 7 & 12 \\ 0 & 1 & 0 & 1 & 2 \\ 0 & 0 & 1 & -5 & -7 \end{array}\right]$$

Če reducirano vrstično formo zapišemo s spremenljivkami, dobimo

$$\begin{aligned} x + 7v &= 12, \\ y + v &= 2, \\ u - 5v &= -7. \end{aligned}$$

Odtod izrazimo spremenljivke x, y, u s spremenljivko v:

$$\begin{aligned} x &= 12 - 7v, \\ y &= 2 - v, \\ u &= -7 + 5v. \end{aligned}$$

Opomba: Pojasnimo zvezo med Gausovo metodo in metodo izločanja spremenljivk. Če npr. iz i-te enačbe sistema izrazimo spremenljivko x_1 in ta izraz vstavimo v j-to linearno enačbo, dobimo novo enačbo

$$(a_{j,2} - \frac{a_{i,2}}{a_{i,1}} a_{j,1})x_2 + \ldots + (a_{j,n} - \frac{a_{i,n}}{a_{i,1}} a_{j,1})x_n = b_j - \frac{b_i}{a_{i,1}} a_{j,1}.$$

Enak rezultat bi dobili, če bi k j-ti enačbi prišteli $-\frac{a_{j,1}}{a_{i,1}}$ kratnik i-te enačbe.

Homogeni sistemi

Spomnimo se, da je sistem linearnih enačb **homogen**, če je oblike:

$$\begin{array}{ccc} a_{1,1}x_1 + \ldots + a_{1,n}x_n & = & 0 \\ \vdots & \vdots & \vdots \\ a_{m,1}x_1 + \ldots + a_{m,n}x_n & = & 0 \end{array} \tag{4}$$

Opazimo, da je $(0, \ldots, 0)$ vedno rešitev tega sistema. Dokazali bomo:

- Poddoločen homogen sistem ima vsaj eno netrivialno rešitev.
- Linearna kombinacija dveh rešitev homogenega sistema je spet rešitev homogenega sistema.
- Splošna rešitev nehomogenega sistema je vsota posebne (=partikularne) rešitve nehomogenega sistema in splošne rešitve pripadajočega homogenega sistema.

Opomba: V tem razdelku bo izraz "nehomogen sistem" pomenil "ne nujno homogen sistem", se pravi splošen linearen sistem.

Trditev

Vsak poddoločen homogen sistem ima vsaj eno netrivialno rešitev.

Dokaz: Dokazujemo s popolno indukcijo po številu vrstic.

Dokažimo najprej za eno linearno enačbo $a_1x_1 + \ldots + a_nx_n = b$, kjer $n \geq 2$. Če je $a_n = 0$, potem je $(0, \ldots, 0, 1)$ netrivialna rešitev. Če je $a_n \neq 0$, potem je $(0, \ldots, 0, -a_n, a_{n-1})$ netrivialna rešitev.

Recimo, da trditev drži za vse homogene sisteme z $m-1$ vrsticami. Vzemimo poljuben homogen sistem z m vrsticami $n > m$ stolpci. Ločimo dva primera. Če so vsi koeficienti pri x_n enaki nič, potem je $(0, \ldots, 0, 1)$ netrivialna rešitev. V nasprotnem primeru lahko iz ene od enačb izrazimo x_n s preostalimi spremenljivkami. Ta izraz potem vstavimo v preostalih $m-1$ enačb in uredimo. Dobimo homogen sistem $m-1$ enačb v $n-1$ spremenljivkah $x_1, \ldots, x_{n-1}$. Po indukcijski predpostavki ima novi sistem netrivialno rešitev, recimo $(\alpha_1, \ldots, \alpha_{n-1})$. To rešitev vstavimo v izraz za x_n in dobimo α_n. Potem je $(\alpha_1, \ldots, \alpha_{n-1}, \alpha_n)$ netrivialna rešitev prvotnega sistema.

Trditev

Linerna kombinacija dveh rešitev homogenega sistema je spet rešitev tega sistema.

Dokaz: Če sta $(s_1, \ldots, s_n)$ in $(t_1, \ldots, t_n)$ dve rešitvi sistema (4), potem je

$$\begin{array}{rcl} a_{1,1}s_1 + \ldots + a_{1,n}s_n & = & 0 \\ \vdots & \vdots & \vdots \\ a_{m,1}s_1 + \ldots + a_{m,n}s_n & = & 0 \end{array} \qquad \begin{array}{rcl} a_{1,1}t_1 + \ldots + a_{1,n}t_n & = & 0 \\ \vdots & \vdots & \vdots \\ a_{m,1}t_1 + \ldots + a_{m,n}t_n & = & 0 \end{array}$$

Preverimo, da je $(\alpha s_1 + \beta t_1, \ldots, \alpha s_n + \beta t_n)$ rešitev (4) za vse $\alpha, \beta \in \mathbb{R}$

$$a_{1,1}(\alpha s_1 + \beta t_1) + \ldots + a_{1,n}(\alpha s_n + \beta t_n) =$$
$$= \alpha(a_{1,1}s_1 + \ldots + a_{1,n}s_n) + \beta(a_{1,1}t_1 + \ldots + a_{1,n}t_n) = 0$$
$$\vdots$$
$$a_{m,1}(\alpha s_1 + \beta t_1) + \ldots + a_{m,n}(\alpha s_n + \beta t_n) =$$
$$= \alpha(a_{m,1}s_1 + \ldots + a_{m,n}s_n) + \beta(a_{m,1}t_1 + \ldots + a_{m,n}t_n) = 0$$

Oglejmo si, kakšna je zveza med rešitvami nehomogenega sistema linearnih enačb in rešitvami **pripadajočega homogenega sistema** linearnih enačb, ki ga dobimo tako, da vse desne strani spremenimo v nič.

Trditev

Recimo, da je nehomogen sistem linearnih enačb rešljiv in recimo, da je n-terica $\mathbf{p}$ ena od njegovih rešitev. Vzemimo poljubno n-terico $\mathbf{x}$ in se vprašajmo, kdaj je $\mathbf{x}$ rešitev našega nehomogenega sistema.

To je res natanko tedaj, ko obstaja taka rešitev $\mathbf{h}$ pripadajočega homogenega sistema, da je $\mathbf{x} = \mathbf{p} + \mathbf{h}$.

Opomba: Množico rešitev nehomogenega sistema torej dobimo tako, da za $\mathbf{p}$ premaknemo množico rešitev pripadajočega homogenega sistema.

Dokaz: Če je $\mathbf{p}$ rešitev nehomogenega sistema (3) in če je $\mathbf{h}$ rešitev pripadajočega homogenega sistema (4), potem kot v dokazu prejšnje trditve vidimo, da je $\mathbf{p} + \mathbf{h}$ rešitev nehomogenega sistema (3).

Obratno, če sta $\mathbf{p}$ in $\mathbf{x}$ dve rešitvi nehomogenega sistema (3), potem kot v dokazu prejšnje trditve vidimo, da je $\mathbf{x} - \mathbf{p}$ rešitev pripadajočega homogenega sistema (4). Torej je $\mathbf{x} = \mathbf{p} + (\mathbf{x} - \mathbf{p})$ želene oblike.

Sistemi linearnih enačb, 2.del

Predoločeni sistemi

Spomnimo se, da je sistem linearnih enačb **predoločen**, če ima več enačb kot spremenljivk. Spomnimo se tudi, da je predoločen sistem običajno nerešljiv. Radi bi posplošili definicijo rešitve sistema tako, da bo vsak predoločen sistem posplošeno rešljiv.

Definicija posplošene rešitve sistema linearnih enačb

Posplošena rešitev sistema linearnih enačb

$$\begin{aligned} a_{1,1}x_1 + \ldots + a_{1,n}x_n &= b_1 \\ \vdots \qquad\qquad \vdots &\quad \vdots \\ a_{m,1}x_1 + \ldots + a_{m,n}x_n &= b_m \end{aligned} \tag{1}$$

je taka n-terica realnih števil $(x_1, \ldots, x_n)$ pri kateri je vektor levih strani

$$(a_{1,1}x_1 + \ldots + a_{1,n}x_n, \ldots, a_{m,1}x_1 + \ldots + a_{m,n}x_n)$$

najbližje vektorju desnih strani $(b_1, \ldots, b_m)$.

Opomba: Če je sistem (1) rešljiv v običajnem smislu, potem se posplošena rešitev ujema z običajno rešitvijo.

Opomba: Posplošeni rešitvi pravimo tudi (posplošena) **rešitev po metodi najmanjših kvadratov**, ker minimizira izraz

$$(a_{1,1}x_1 + \ldots + a_{1,n}x_n - b_1)^2 + \ldots + (a_{m,1}x_1 + \ldots + a_{m,n}x_n - b_m)^2 \quad (2)$$

Poglejmo si zdaj, kako poiščemo posplošeno rešitev. Izraz (2) najprej predelamo v vektorsko obliko

$$\|x_1\mathbf{a}_1 + \ldots + x_n\mathbf{a}_n - \mathbf{b}\|^2 \quad (3)$$

kjer je $\mathbf{a}_1 = (a_{1,1}, \ldots, a_{m,1}), \ldots, \mathbf{a}_n = (a_{1,n}, \ldots, a_{m,n}), \mathbf{b} = (b_1, \ldots, b_m)$. Geometrijsko gledano je minimizacija izraza (3) ekvivalentna pravokotni projekciji vektorja $\mathbf{b}$ na množico vseh linearnih kombinacij vektorjev $\mathbf{a}_1, \ldots, \mathbf{a}_n$. Poiskati moramo torej take skalarje $x_1, \ldots, x_n$, da je vektor $x_1\mathbf{a}_1 + \ldots + x_n\mathbf{a}_n - \mathbf{b}$ pravokoten na vektorje $\mathbf{a}_1, \ldots, \mathbf{a}_n$.

Če pravokotnostne pogoje

$$\langle x_1\mathbf{a}_1 + \ldots + x_n\mathbf{a}_n - \mathbf{b}, \mathbf{a}_1\rangle = \ldots = \langle x_1\mathbf{a}_1 + \ldots + x_n\mathbf{a}_n - \mathbf{b}, \mathbf{a}_n\rangle = 0$$

razpišemo, dobimo sistem linearnih enačb

$$\begin{array}{ccc} \langle \mathbf{a}_1, \mathbf{a}_1\rangle x_1 + \ldots + \langle \mathbf{a}_n, \mathbf{a}_1\rangle x_n & = & \langle \mathbf{b}, \mathbf{a}_1\rangle \\ \vdots \qquad\qquad\qquad\quad \vdots & & \vdots \\ \langle \mathbf{a}_1, \mathbf{a}_n\rangle x_1 + \ldots + \langle \mathbf{a}_n, \mathbf{a}_n\rangle x_n & = & \langle \mathbf{b}, \mathbf{a}_n\rangle \end{array} \tag{4}$$

Posplošene rešitve sistema (1) se torej ujemajo z običajnimi rešitvami sistema (4).

Opomba: Izkaže se, da je sistem (4) vedno rešljiv. Enolično je rešljiv natanko tedaj, ko so vektorji $\mathbf{a}_1, \ldots, \mathbf{a}_n$ linearno neodvisni. Kadar ni enolično rešljiv, poiščemo njegovo najkrajšo rešitev. S temi vprašanji se bomo ukvarjali na koncu drugega semestra v razdelku o psevdoinverzih.

Preprost primer predoločenega sistema

Poiščimo posplošeno rešitev 3×2 sistema

$$x = 0, \quad y = 0, \quad x + y = 1.$$

Sistem je očitno nerešljiv. Zapišimo ga v vektorski obliki

$$(0, 0, 1) = (x, y, x + y) = x(1, 0, 1) + y(0, 1, 1)$$

Razliko leve in desne strani skalarno množimo z $(1, 0, 1)$ in z $(0, 1, 1)$:

$$\langle x(1, 0, 1) + y(0, 1, 1) - (0, 0, 1), (1, 0, 1) \rangle = 0$$
$$\langle x(1, 0, 1) + y(0, 1, 1) - (0, 0, 1), (0, 1, 1) \rangle = 0.$$

Ko uredimo, dobimo sistem enačb

$$2x + y = 1, \quad x + 2y = 1.$$

Običajna rešitev tega sistema je $x = y = \frac{1}{3}$. To je potem tudi posplošena rešitev prvotnega sistema.

Regresijska premica in posplošitve

Iščemo tako premico $y = ax + b$, ki se najbolje prilega danim točkam

$$(x_1, y_1), \ldots, (x_m, y_m).$$

Če bi šla iskana premica skozi vse dane točke, bi koeficienta a in b zadoščala naslednjemu $m \times 2$ sistemu linearnih enačb

$$ax_1 + b = y_1 \quad \ldots \quad ax_m + b = y_m \tag{5}$$

Seveda se pri $m > 2$ skoraj nikoli ne zgodi, da bi dane točke ležale na isti premici. Zato se bomo zadovoljili s posplošeno rešitvijo sistema (5), ki jo dobimo tako, da minimiziramo izraz

$$(ax_1 + b - y_1)^2 + \ldots + (ax_m + b - y_m)^2 \tag{6}$$

Opomba: Izraz (6) je enak $d_1^2 + \ldots + d_m^2$, kjer je $d_i := |ax_i + b - y_i|$ razdalja med dano točko (x_i, y_i) in točko $(x_i, ax_i + b)$ na iskani premici.

Izraz (6) predelamo v vektorsko obliko

$$\|a(x_1, \ldots, x_m) + b(1, \ldots, 1) - (y_1, \ldots, y_m)\|^2 \tag{7}$$

Iščemo taka a in b, pri katerih je izraz (7) minimalen. To je poseben primer naloge o pravokotni projekciji točke na parametrično podano ravnino v $\mathbb{R}^m$. Vemo, da mora biti vektor $a(x_1, \ldots, x_m) + b(1, \ldots, 1) - (y_1, \ldots, y_m)$ pravokoten na vektorja $(x_1, \ldots, x_m)$ in $(1, \ldots, 1)$. Ko razpišemo izraza

$$\langle a(x_1, \ldots, x_m) + b(1, \ldots, 1) - (y_1, \ldots, y_m), (x_1, \ldots, x_m) \rangle = 0$$
$$\langle a(x_1, \ldots, x_m) + b(1, \ldots, 1) - (y_1, \ldots, y_m), (1, \ldots, 1) \rangle = 0$$

dobimo sistem dveh enačb za a in b:

$$\begin{array}{ccccc} a(\sum_{i=1}^{m} x_i^2) & + & b(\sum_{i=1}^{m} x_i) & = & \sum_{i=1}^{m} x_i y_i \\ a(\sum_{i=1}^{m} x_i) & + & bm & = & \sum_{i=1}^{m} y_i \end{array} \tag{8}$$

Rešitev tega sistema vstavimo v $y = ax + b$ in dobimo iskano premico.

Primer

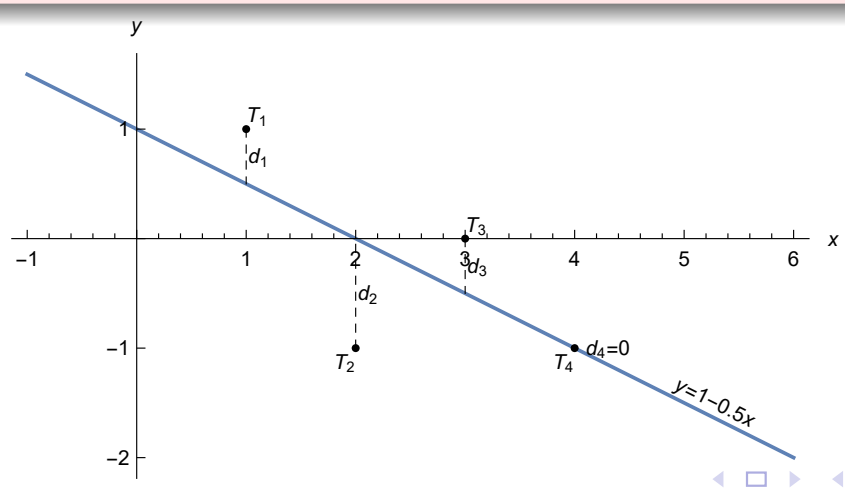
Poišči premico, ki se najbolje prilega točkam $(1,1),(2,-1),(3,0),(4,-1)$. Najprej izračunamo

$$\sum_{i=1}^{4} x_i^2 = 30, \quad \sum_{i=1}^{4} x_i = 10, \quad \sum_{i=1}^{4} x_i y_i = -5, \quad \sum_{i=1}^{4} y_i = -1$$

in rezultate vstavimo v sistem (8). Dobimo sistem enačb

$$30a + 10b = -5, \quad 10a + 4b = -1$$

katerega rešitev je $a = -\frac{1}{2}, b = 1$. Iskana premica je torej $y = 1 - 0.5x$.



Opomba: Premici, ki se najbolje prilega danim točkam pravimo tudi **regresijska premica**.

Opomba: Na podoben način poiščemo parabolo $y = ax^2 + bx + c$, ki se najbolje prilega danim točkam $(x_1, y_1), \ldots, (x_m, y_m)$. Rešiti moramo sistem

$$\begin{aligned} a\sum x_i^4 + b\sum x_i^3 + c\sum x_i^2 &= \sum x_i^2 y_i \\ a\sum x_i^3 + b\sum x_i^2 + c\sum x_i &= \sum x_i y_i \\ a\sum x_i^2 + b\sum x_i + c\sum 1 &= \sum y_i \end{aligned}$$

Opomba: Podobno poiščemo ravnino $z = ax + by + c$, ki se najbolje prilega danim točkam $(x_1, y_1, z_1), \ldots, (x_n, y_n, z_n)$. Rešiti moramo sistem

$$\begin{aligned} a\sum x_i^2 + b\sum x_i y_i + c\sum x_i &= \sum x_i z_i \\ a\sum x_i y_i + b\sum y_i^2 + c\sum y_i &= \sum y_i z_i \\ a\sum x_i + b\sum y_i + c\sum 1 &= \sum z_i \end{aligned}$$

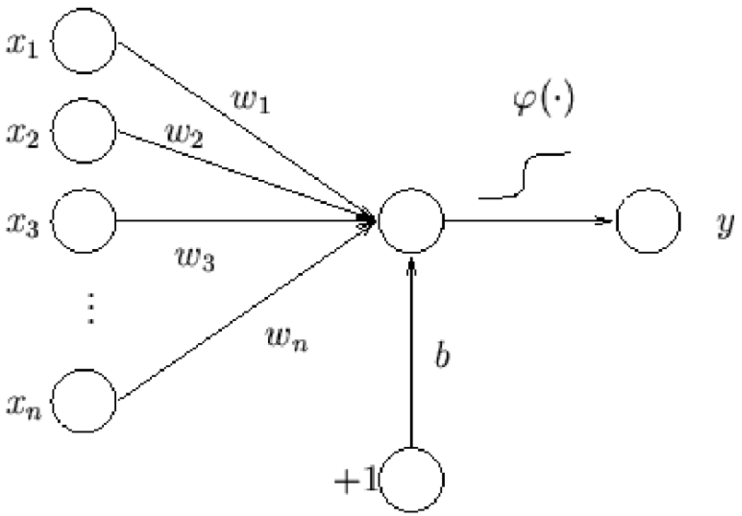
Tako nalogo dobimo, kadar treniramo perceptron.

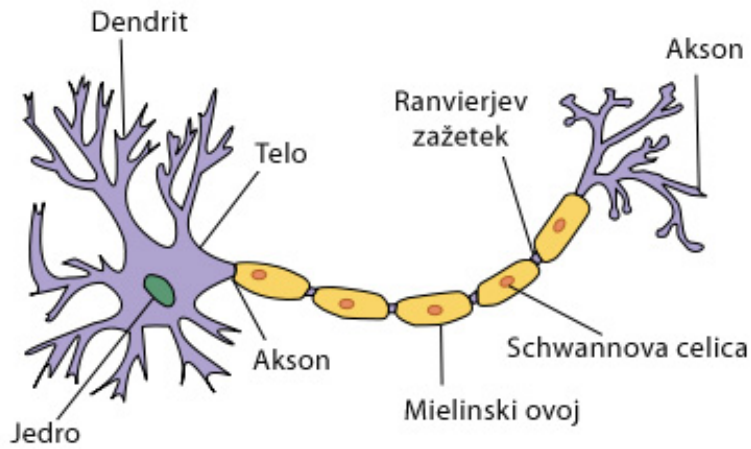
Perceptron je poenostavljen model biološkega nevrona.

- Dendriti prejmejo vhodne signale $x_1, \ldots, x_n$.
- Do telesa pride signal $z = w_1 x_1 + \ldots + w_n x_n + b$, kjer so w_i, b uteži.
- Akson pošlje izhodni signal $y = \varphi(z)$, kjer je φ aktivacijska funkcija. (Običajno je $\phi(z) = 1$, če $z > 0$ in 0 sicer. Pri nas je $\phi(z) = z$.)

Radi bi nevron natrenirali tako, da se bodo dobljeni izhodi čimbolj ujemali s pričakovanimi izhodi $(x_{1,1}, \ldots, x_{1,n}, y_1), \ldots, (x_{m,1}, \ldots, x_{m,n}, y_m)$, kjer je $m > n + 1$. Iščemo torej take uteži $w_1, \ldots, w_n, b$, ki "rešijo" sistem

$$w_1 x_{i,1} + \ldots + w_n x_{i,n} + b = y_i, \quad i = 1, \ldots, m.$$





Poddoločeni sistemi

Spomnimo se, da je sistem linearnih enačb **poddoločen**, če ima manj enačb kot neznank.

Trditev

Če je poddoločen sistem linearnih enačb rešljiv, potem ima neskončno rešitev.

Dokaz: To sledi iz naslednjih dveh rezultatov o homogenih sistemih.

- Vsak poddoločen homogen sistem ima neskončno rešitev.
- Če je **p** rešitev nehomogenega sistema in če je **h** rešitev pripadajočega homogenega sistema, potem je $\mathbf{p} + \mathbf{h}$ rešitev nehomogenega sistema.

Opomba: Seveda se lahko zgodi, da je nehomogen poddoločen sistem nerešljiv. Recimo 2×3 sistem $x + y + z = 1, x + y + z = 2$ je nerešljiv. Tudi 1×2 sistem $0x + 0y = 1$ je nerešljiv.

Kadar je poddoločen rešljiv, nas pogosto zanima njegova najkrajša rešitev.

Geometrijsko gledano iščemo pravokotno projekcijo izhodišča na presek hiperravnin. Ponovimo najprej, kako to naredimo za eno hiperravnino. Iščemo presek hiperravnine $\langle \mathbf{n}, \mathbf{r} \rangle = b$ s premico $\mathbf{r} = t\mathbf{n}$, ki gre skozi izhodišče in je pravokotna na hiperravnino. Vstavimo drugo enačbo v prvo in izrazimo $t = \frac{b}{\langle \mathbf{n},\mathbf{n} \rangle}$. Torej je iskana najkrajša rešitev $\mathbf{r} = \frac{b}{\langle \mathbf{n},\mathbf{n} \rangle}\mathbf{n}$.

Podobno poiščemo pravokotno projekcijo izhodišča na presek dveh hiperravnin $\langle \mathbf{n}_1, \mathbf{r} \rangle = b_1$ in $\langle \mathbf{n}_2, \mathbf{r} \rangle = b_2$. Naloge se lotimo v dveh korakih:

- Poiščemo ravnino, ki gre skozi izhodišče in je pravokotna na presek obeh hiperravnin. To je ravnina $\mathbf{r} = t_1\mathbf{n}_1 + t_2\mathbf{n}_2$.
- Izračunamo presek te ravnine s presekom obeh hiperravnin. Enačbo ravnine vstavimo v enačbi hiperravnin in uredimo. Dobimo sistem

$$t_1\langle \mathbf{n}_1, \mathbf{n}_1 \rangle + t_2\langle \mathbf{n}_1, \mathbf{n}_2 \rangle = b_1, \quad t_1\langle \mathbf{n}_2, \mathbf{n}_1 \rangle + t_2\langle \mathbf{n}_2, \mathbf{n}_2 \rangle = b_2.$$

  Rešimo po t_1 in t_2 in rešitev vstavimo v enačbo ravnine.

Enak postopek deluje tudi pri preseku več kot dveh hiperravnin.

Primer iskanja najkrajše rešitve poddoločenega sistema

Poišči najkrajšo rešitev sistema

$$\begin{aligned} 2x - y + 3u - 2v &= 1, \\ x + y + 2u - 2v &= 0, \\ x + 2u - 3v &= -2. \end{aligned}$$

Najprej sistem zapišemo v obliki

$$\langle \mathbf{n}_1, \mathbf{r} \rangle = 1, \quad \langle \mathbf{n}_2, \mathbf{r} \rangle = 0, \quad \langle \mathbf{n}_3, \mathbf{r} \rangle = -2,$$

kjer je $\mathbf{r} = (x, y, u, v)$ in

$$\mathbf{n}_1 = (2, -1, 3, -2), \quad \mathbf{n}_2 = (1, 1, 2, -2), \quad \mathbf{n}_3 = (1, 0, 2, -3).$$

Rešitev iščemo z nastavkom $\mathbf{r} = t_1\mathbf{n}_1 + t_2\mathbf{n}_2 + t_3\mathbf{n}_3$.

Dobimo sistem

$$\begin{aligned} 18t_1 + 11t_2 + 14t_3 &= 1, \\ 11t_1 + 10t_2 + 11t_3 &= 0, \\ 14t_1 + 11t_2 + 14t_3 &= -2, \end{aligned}$$

katerega rešitev je

$$t_1 = \frac{3}{4}, \quad t_2 = \frac{22}{19}, \quad t_3 = -\frac{137}{76}.$$

Ko te vrednosti vstavimo v nastavek, dobimo

$$\mathbf{r} = t_1\mathbf{n}_1 + t_2\mathbf{n}_2 + t_3\mathbf{n}_3 = (\frac{65}{76}, \frac{31}{76}, \frac{73}{76}, \frac{121}{76}).$$

To je iskana najkrajša rešitev prvotnega sistema.

Opomba: Drug način reševanja je, da sistem najprej rešimo in nato projeciramo izhodišče na množico rešitev.

Matrike

Definicija in primeri matrik

Definicija matrike

Matrika velikosti $m \times n$ je urejena m-terica urejenih n-teric realnih števil, torej element prostora $(\mathbb{R}^n)^m$. Namesto

$$A = ((a_{1,1}, \ldots, a_{1,n}), \ldots, (a_{m,1}, \ldots, a_{m,n}))$$

pišemo raje

$$A = \begin{bmatrix} a_{1,1} & \ldots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \ldots & a_{m,n} \end{bmatrix}. \tag{1}$$

Opomba: Matrike velikosti 1×1 lahko identificiramo s skalarji.

Opomba: Matrikam velikosti $m \times 1$ pravimo **stolpčni vektorji**. Matrikam velikosti $1 \times n$ pravimo **vrstični vektorji**. Tako stolpčne kot vrstične vektorje lahko identificiramo z običajnimi vektorji.

Opomba: Stolpčnim vektorjem

$$\begin{bmatrix} a_{1,1} \\ \vdots \\ a_{m,1} \end{bmatrix}, \begin{bmatrix} a_{1,2} \\ \vdots \\ a_{m,2} \end{bmatrix}, \ldots, \begin{bmatrix} a_{1,n} \\ \vdots \\ a_{m,n} \end{bmatrix}$$

pravimo **stolpci** matrike (1). Matriko (1) lahko zapišemo kot

$$A = \begin{bmatrix} \begin{bmatrix} a_{1,1} \\ \vdots \\ a_{m,1} \end{bmatrix}, \begin{bmatrix} a_{1,2} \\ \vdots \\ a_{m,2} \end{bmatrix}, \ldots, \begin{bmatrix} a_{1,n} \\ \vdots \\ a_{m,n} \end{bmatrix} \end{bmatrix}$$

Vrstičnim vektorjem

$$\begin{bmatrix} a_{1,1} & \ldots & a_{1,n} \end{bmatrix}, \begin{bmatrix} a_{2,1} & \ldots & a_{2,n} \end{bmatrix}, \ldots, \begin{bmatrix} a_{m,1} & \ldots & a_{m,n} \end{bmatrix}$$

pravimo **vrstice** matrike (1). Na preseku i-te vrstice in j-tega stolpca matrike (1) leži (i,j)-ti **element** (=(i,j)-ti **vhod**) matrike (1).

Operacije z matrikami

Produkt matrike s skalarjem je definiran po komponentah

$$\alpha \begin{bmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \dots & a_{m,n} \end{bmatrix} := \begin{bmatrix} \alpha a_{1,1} & \dots & \alpha a_{1,n} \\ \vdots & & \vdots \\ \alpha a_{m,1} & \dots & \alpha a_{m,n} \end{bmatrix}.$$

Tudi **vsota dveh matrik** enake velikosti je definirana po komponentah

$$\begin{bmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \dots & a_{m,n} \end{bmatrix} + \begin{bmatrix} b_{1,1} & \dots & b_{1,n} \\ \vdots & & \vdots \\ b_{m,1} & \dots & b_{m,n} \end{bmatrix} :=$$

$$:= \begin{bmatrix} a_{1,1} + b_{1,1} & \dots & a_{1,n} + b_{1,n} \\ \vdots & & \vdots \\ a_{m,1} + b_{m,1} & \dots & a_{m,n} + b_{m,n} \end{bmatrix}$$

Lastnosti teh dveh operacij so enake kot pri vektorjih.

Produkt dveh matrik velikosti $m \times n$ in $n \times p$

$$A = \begin{bmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \dots & a_{m,n} \end{bmatrix} \quad B = \begin{bmatrix} b_{1,1} & \dots & b_{1,p} \\ \vdots & & \vdots \\ b_{n,1} & \dots & b_{n,p} \end{bmatrix}$$

je taka matrika $C = AB$ velikosti $m \times p$, katere (i, j)-ti element je

$$c_{i,j} = \sum_{k=1}^{n} a_{i,k} b_{k,j}$$

Torej je $c_{i,j}$ ravno skalarni produkt i-te vrstice matrike A in j-tega stolpca matrike B. Definicijo $c_{i,j}$ lahko zapišemo tudi takole

$$\begin{bmatrix} c_{i,j} \end{bmatrix} = \begin{bmatrix} a_{i,1} & \dots & a_{i,n} \end{bmatrix} \begin{bmatrix} b_{1,j} \\ \vdots \\ b_{n,j} \end{bmatrix} = \begin{bmatrix} \sum_{k=1}^{n} a_{i,k} b_{k,j} \end{bmatrix}$$

Primer množenja dveh matrik

Zmnožimo matriki

$$A = \begin{bmatrix} 1 & 2 \\ -2 & 0 \end{bmatrix} \quad \text{in} \quad B = \begin{bmatrix} -1 & 1 \\ 3 & 2 \end{bmatrix}.$$

Velja

$$AB = \begin{bmatrix} 5 & 5 \\ 2 & -2 \end{bmatrix} \quad \text{in} \quad BA = \begin{bmatrix} -3 & -2 \\ -1 & 6 \end{bmatrix}.$$

Torej ni vseeno, v kakšnem vrstnem redu zmnožimo dve matriki.

Onovne lastnosti množenja matrik so

- $(AB)C = A(BC)$,
- $(A + B)C = AC + BC$,
- $A(B + C) = AB + AC$,
- $(\lambda A)B = A(\lambda B) = \lambda(AB)$,

Transponiranka matrike A je taka matrika A^T, da za vsak i in j velja: (i,j)-ti element A^T je enak (j,i)-temu elementu A. S formulo:

$$\begin{bmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \dots & a_{m,n} \end{bmatrix}^T := \begin{bmatrix} a_{1,1} & \dots & a_{m,1} \\ \vdots & & \vdots \\ a_{1,n} & \dots & a_{m,n} \end{bmatrix}$$

Primer transponiranja matrike

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}.$$

Osnovne lastnosti transponiranja matrik so

- $(A^T)^T = A$,
- $(AB)^T = B^T A^T$,
- $(A+B)^T = A^T + B^T$,
- $(\lambda A)^T = \lambda A^T$.

Ničelna matrika je taka matrika, ki ima vse elemente enake nič. Označimo z $0_{m,n}$ ničelno matriko velikosti $m \times n$. Kadar sta m in n znana iz konteksta pišemo kar 0 namesto $0_{m,n}$. Velja:

- Če k matriki A prištejemo ničelno matriko enake velikosti, dobimo A.
- Če matriko A z leve ali desne pomnožimo z ničelno matriko ustrezne velikosti, potem dobimo ničelno matriko.
- Če transponiramo ničelno matriko, dobimo ničelno matriko.

Identična matrika je taka kvadratna matrika, ki ima po glavni diagonali same enke, drugod pa same ničle. Identično matriko velikosti $n \times n$ označimo z I_n. Kadar je n znan iz konteksta, pišemo kar I namesto I_n.

- Če matriko A z leve ali desne pomnožimo z identično matriko ustrezne velikosti, potem spet dobimo matriko A.
- Če transponiramo identično matriko, spet dobimo identično matriko.

Matrični zapis sistema linearnih enačb

Sistem linearnih enačb

$$\begin{array}{ccccccc} a_{1,1}x_1 & + & \dots & + & a_{1,n}x_n & = & b_1 \\ \vdots & & & & \vdots & & \vdots \\ a_{m,1}x_1 & + & \dots & + & a_{m,n}x_n & = & b_m \end{array}$$

zapišemo s pomočjo matričnega množenja takole

$$\begin{bmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \dots & a_{m,n} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{1,1}x_1 + \dots + a_{1,n}x_n \\ \vdots \\ a_{m,1}x_1 + \dots + a_{m,n}x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

Na kratko to zapišemo kot $A\mathbf{x} = \mathbf{b}$, kjer je

$$A = \begin{bmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \dots & a_{m,n} \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

Matrični zapis Gaussove metode

Spomnimo se, da sistem rešujemo z uporabo naslednjih treh tipov **elementarnih transformacij**:

- k eni od enačb lahko prištejemo večkratnik druge enačbe;
- lahko zamenjamo vrstni red dveh enačb;
- enačbo lahko pomnožimo z neničelno konstanto.

Poskusimo te transformacije zapisati z matrikami.

Definirajmo **elementarne** $m \times m$ **matrike** $E_{i,j}(\alpha)$, $P_{i,j}$ in $E_i(\beta)$ takole:

- matriko $E_{i,j}(\alpha)$ dobimo tako, da v identični matriki I_m k i-ti vrstici prištejemo α-kratnik j-te vrstice;
- matriko $P_{i,j}$ dobimo tako, da v I_m zamenjamo i-to in j-to vrstico;
- matriko $E_i(\beta)$ dobimo tako, da v I_m množimo i-to vrstico z β.

Primeri elementarnih 2×2 matrik

$$E_{1,2}(7) = \begin{bmatrix} 1 & 7 \\ 0 & 1 \end{bmatrix}, \quad P_{1,2} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad E_2(5) = \begin{bmatrix} 1 & 0 \\ 0 & 5 \end{bmatrix}$$

Krajši račun pokaže, da lahko elementarne transformacije opišemo s pomočjo elementarnih matrik takole:

1. če v sistemu $A\mathbf{x} = \mathbf{b}$ k i-ti enačbi prištejemo α-kratnik j-te enačbe, potem dobimo sistem $E_{i,j}(\alpha)A\mathbf{x} = E_{i,j}(\alpha)\mathbf{b}$;
2. če v sistemu $A\mathbf{x} = \mathbf{b}$ zamenjamo i-to in j-to enačbo, potem dobimo sistem $P_{i,j}A\mathbf{x} = P_{i,j}\mathbf{b}$;
3. če v sistemu $A\mathbf{x} = \mathbf{b}$ pomnožimo i-to enačbo z β, potem dobimo sistem $E_i(\beta)A\mathbf{x} = E_i(\beta)\mathbf{b}$.

Gaussovo metodo lahko povemo takole: sistem $A\mathbf{x} = \mathbf{b}$ toliko časa množimo z elementarnimi matrikami z leve, dokler ne dobimo sistema $A'\mathbf{x} = \mathbf{b}'$, pri katerem je A' **reducirane stopničaste oblike**.

$$\left[\begin{array}{cccccccccccccccc}
0 & \dots & 0 & 1 & \times & \dots & \times & 0 & \times & \dots & \times & 0 & \times & \dots & \times & 0 & \dots \\
0 & \dots & 0 & 0 & 0 & \dots & 0 & 1 & \times & \dots & \times & 0 & \times & \dots & \times & 0 & \dots \\
0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 1 & \times & \dots & \times & 0 & \dots \\
0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 1 & \dots \\
\vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \\
0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & \dots
\end{array}\right]$$

Matrični zapis metode najmanjših kvadratov

Kadar je sistem

$$A\mathbf{x} = \mathbf{b} \tag{2}$$

nerešljiv, poiščemo tak $\mathbf{x}_0$, ki minimizira izraz

$$\|A\mathbf{x} - \mathbf{b}\|$$

Takemu $\mathbf{x}_0$ pravimo **posplošena rešitev** sistema (2).

Primerjajmo matrični zapis zgoraj z vektorskim zapisom od zadnjič. Če so $\mathbf{a}_1, \ldots, \mathbf{a}_n$ stolpci matrike A, potem je $A\mathbf{x} = x_1\mathbf{a}_1 + \ldots + x_n\mathbf{a}_n$. Izraz $\|x_1\mathbf{a}_1 + \ldots + x_n\mathbf{a}_n - \mathbf{b}\|$ je minimalen če je $x_1\mathbf{a}_1 + \ldots + x_n\mathbf{a}_n - \mathbf{b}$ pravokoten na $\mathbf{a}_1, \ldots, \mathbf{a}_n$. Če to razpišemo, dobimo

$$\begin{bmatrix} \langle \mathbf{a}_1, \mathbf{a}_1 \rangle & \ldots & \langle \mathbf{a}_1, \mathbf{a}_n \rangle \\ \vdots & \ddots & \vdots \\ \langle \mathbf{a}_n, \mathbf{a}_1 \rangle & \ldots & \langle \mathbf{a}_n, \mathbf{a}_n \rangle \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \langle \mathbf{a}_1, \mathbf{b} \rangle \\ \vdots \\ \langle \mathbf{a}_n, \mathbf{b} \rangle \end{bmatrix}$$

kar se v matričnem zapisu glasi $A^T A\mathbf{x} = A^T\mathbf{b}$.

Skicirali smo dokaz naslednje trditve. Bolj podroben dokaz je spodaj.

Trditev

Posplošene rešitve sistema $A\mathbf{x} = \mathbf{b}$ so enake običajnim rešitvam sistema

$$A^T A\mathbf{x} = A^T \mathbf{b} \tag{3}$$

Opomba: Sistem (3) je vedno rešljiv v običajnem smislu (v 2. semestru).
Opomba: V dokazu bomo večkrat uporabili formulo $\langle \mathbf{c}, \mathbf{d} \rangle = \mathbf{c}^T \mathbf{d}$.

Dokaz: Naj bo $\mathbf{x}_0$ tak, da $A^T A\mathbf{x}_0 = A^T \mathbf{b}$. Potem za vsak vektor $\mathbf{x}$ velja

$$\langle A(\mathbf{x} - \mathbf{x}_0), A\mathbf{x_0} - \mathbf{b} \rangle = (A(\mathbf{x} - \mathbf{x}_0))^T (A\mathbf{x_0} - \mathbf{b}) = (\mathbf{x} - \mathbf{x}_0)^T A^T (A\mathbf{x_0} - \mathbf{b})$$

kar je po predpostavki enako nič. Po Pitagorovem izreku odtod sledi

$$\|A\mathbf{x} - \mathbf{b}\|^2 = \|A(\mathbf{x} - \mathbf{x}_0) + A\mathbf{x}_0 - \mathbf{b}\|^2 = \|A(\mathbf{x} - \mathbf{x}_0)\|^2 + \|A\mathbf{x}_0 - \mathbf{b}\|^2$$

Odtod sledi $\|A\mathbf{x} - \mathbf{b}\| \geq \|A\mathbf{x}_0 - \mathbf{b}\|$, torej $\mathbf{x}_0$ res minimizira izraz $\|A\mathbf{x} - \mathbf{b}\|$.

Dokaz v nasprotno smer je nekoliko daljši. Recimo, da je $\mathbf{x}_0$ posplošena rešitev sistema $A\mathbf{x} = \mathbf{b}$. Torej za vsak $\mathbf{x}$ velja $\|A\mathbf{x} - \mathbf{b}\| \geq \|A\mathbf{x}_0 - \mathbf{b}\|$. Posebej za vsak $\mathbf{z}$ in t velja $\|A(\mathbf{x}_0 + t\mathbf{z}) - \mathbf{b}\| \geq \|A\mathbf{x}_0 - \mathbf{b}\|$. Pokažimo, da odtod sledi, da sta $\mathbf{c} := A\mathbf{x}_0 - \mathbf{b}$ in $\mathbf{d} := A\mathbf{z}$ pravokotna.

Po predpostavki je $\|\mathbf{c} + t\mathbf{d}\| \geq \|\mathbf{c}\|$ za vsak $t \in \mathbb{R}$. Če $\langle \mathbf{c}, \mathbf{d} \rangle \neq 0$, potem za $t_0 := -\frac{\langle \mathbf{c}, \mathbf{d} \rangle}{\langle \mathbf{d}, \mathbf{d} \rangle}$ velja $t_0\mathbf{d} \neq 0$. Po definiciji t_0 je $\langle \mathbf{c} + t_0\mathbf{d}, \mathbf{d} \rangle = 0$, torej je $\|\mathbf{c} + t_0\mathbf{d}\|^2 + \| - t_0\mathbf{d}\|^2 = \|\mathbf{c}\|^2$, kar nam da protislovje $\|\mathbf{c} + t_0\mathbf{d}\| < \|\mathbf{c}\|$.

Iz pravokotnosti $\mathbf{c}$ in $\mathbf{d}$ sledi

$$\langle A^T(A\mathbf{x}_0 - \mathbf{b}), \mathbf{z} \rangle = (A^T(A\mathbf{x}_0 - \mathbf{b}))^T\mathbf{z} = (A\mathbf{x}_0 - \mathbf{b})^T A\mathbf{z} = \mathbf{c}^T\mathbf{d} = 0.$$

Če vstavimo $\mathbf{z} = A^T(A\mathbf{x}_0 - \mathbf{b})$, dobimo $\mathbf{z} = \mathbf{0}$. Torej je $A^T A\mathbf{x}_0 = A^T\mathbf{b}$.

Primer posplošene rešitve sistema

Posplošena rešitev sistema

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

se ujema z običajno rešitvijo sistema

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Ko zmnožimo matrike, dobimo sistem

$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

ki ima rešitev $x = y = \frac{1}{3}$. To je posplošena rešitev prvotnega sistema.

Primer - Regresijska premica v matričnem zapisu

Premico $y = ax + b$, ki se najbolje prilega točkam $(x_i, y_i), i = 1, \ldots, m$, dobimo kot posplošeno rešitev sistema

$$\begin{bmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_m & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}$$

se pravi kot običajno rešitev sistema

$$\begin{bmatrix} x_1 & \ldots & x_m \\ 1 & \ldots & 1 \end{bmatrix} \begin{bmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_m & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} x_1 & \ldots & x_m \\ 1 & \ldots & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}$$

Ko zmnožimo matrike, dobimo sistem

$$\begin{bmatrix} \sum_{i=1}^m x_i^2 & \sum_{i=1}^m x_i \\ \sum_{i=1}^m x_i & m \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^m x_i y_i \\ \sum_{i=1}^m y_i \end{bmatrix}$$

Najkrajša rešitev sistema

Kadar je sistem

$$A\mathbf{x} = \mathbf{b}$$

neenolično rešljiv, nas zanima najkrajša rešitev, se pravi rešitev z najmanjšo normo. Pokazali bomo, da najkrajšo rešitev dobimo z nastavkom $\mathbf{x} = A^T\mathbf{y}$.

Trditev

Najkrajša rešitev sistema $A\mathbf{x} = \mathbf{b}$ je $A^T\mathbf{y}_0$, kjer je $\mathbf{y}_0$ poljubna rešitev

$$AA^T\mathbf{y} = \mathbf{b}. \tag{4}$$

Opomba: Da iz rešljivosti sistema $A\mathbf{x} = \mathbf{b}$ sledi rešljivost sistema $AA^T\mathbf{y} = \mathbf{b}$ bomo znali dokazati šele v drugem semestru.

Dokaz: Naj bo $\mathbf{x}_0$ poljubna rešitev sistema $A\mathbf{x} = \mathbf{b}$ in naj bo $\mathbf{y}_0$ poljubna rešitev sistema $AA^T\mathbf{y} = \mathbf{b}$. Radi bi pokazali, da je $\|\mathbf{x}_0\| \geq \|A^T\mathbf{y}_0\|$.

Pokažimo najprej, da sta $\mathbf{x}_0 - A^T\mathbf{y}_0$ in $A^T\mathbf{y}_0$ pravokotna. To sledi iz

$$\langle \mathbf{x}_0 - A^T\mathbf{y}_0, A^T\mathbf{y}_0 \rangle = (\mathbf{x}_0 - A^T\mathbf{y}_0)^T A^T\mathbf{y}_0 =$$
$$= (A(\mathbf{x}_0 - A^T\mathbf{y}_0))^T\mathbf{y}_0 = (\mathbf{b} - \mathbf{b})^T\mathbf{y}_0 = 0$$

Če uporabimo Pitagorov izrek, dobimo

$$\|\mathbf{x}_0\|^2 = \|\mathbf{x}_0 - A^T\mathbf{y}_0 + A^T\mathbf{y}_0\|^2 =$$
$$\|\mathbf{x}_0 - A^T\mathbf{y}_0\|^2 + \|A^T\mathbf{y}_0\|^2 \geq \|A^T\mathbf{y}_0\|^2$$

odkoder sledi želena neenakost

$$\|\mathbf{x}_0\| \geq \|A^T\mathbf{y}_0\|$$

Opomba: iz dokaza sledi, da za poljubni rešitvi $\mathbf{y}_1$ in $\mathbf{y}_2$ sistema $AA^T\mathbf{y} = \mathbf{b}$ velja $\|A^T\mathbf{y}_1\| = \|A^T\mathbf{y}_2\|$, torej je res vseeno, katero rešitev izberemo.

Primer najkrajše rešitve sistema

Poiščimo najkrajšo rešitev sistema

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -2 \\ 0 \end{bmatrix}$$

Pomagamo si z nastavkom

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}$$

Ko vstavimo nastavek v sistem in uredimo, dobimo nov sistem

$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} -2 \\ 0 \end{bmatrix}$$

katerega rešitev je $u = -\frac{4}{3}$, $v = \frac{2}{3}$.

Najkrajša rešitev prvotnega sistema je torej

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -\frac{4}{3} \\ \frac{2}{3} \end{bmatrix} = \begin{bmatrix} -\frac{4}{3} \\ -\frac{2}{3} \\ \frac{2}{3} \end{bmatrix}$$

Za konec naredimo primerjavo matričnega in vektorskega zapisa metode. Sistem $A\mathbf{x} = \mathbf{b}$ lahko zapišemo kot

$$\langle \mathbf{n}_1, \mathbf{x} \rangle = b_1 \quad \dots \quad \langle \mathbf{n}_m, \mathbf{x} \rangle = b_m \tag{5}$$

kjer so $\mathbf{n}_i$ transponirane vrstice matrike A. Rešitev iščemo z nastavkom

$$\mathbf{x} = y_1\mathbf{n}_1 + \ldots + y_m\mathbf{n}_m \tag{6}$$

Ko nastavek (6) vstavimo v sistem (5) in uredimo, dobimo sistem

$$\begin{bmatrix} \langle \mathbf{n}_1, \mathbf{n}_1 \rangle & \dots & \langle \mathbf{n}_1, \mathbf{n}_m \rangle \\ \vdots & \ddots & \vdots \\ \langle \mathbf{n}_m, \mathbf{n}_1 \rangle & \dots & \langle \mathbf{n}_m, \mathbf{n}_m \rangle \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

ki se v matričnem zapisu glasi $AA^T\mathbf{y} = \mathbf{b}$. Potem je $\mathbf{x} = A^T\mathbf{y}$.

Inverz matrike

Inverz kvadratne matrike A je taka kvadratna matrika B, da velja

$$AB = BA = I.$$

Matrika ima lahko kvečjemu en inverz. Če sta namreč B_1 in B_2 dva inverza matrike A, potem velja $AB_1 = I, B_1A = I, AB_2 = I$ in $B_2A = I$, torej je

$$B_1 = IB_1 = (B_2A)B_1 = B_2(AB_1) = B_2I = B_2.$$

Če matrika A ima inverz, ga označimo z A^{-1}. Ničelna matrika seveda nima inverza. Zanimivo pa je, da inverza nimajo tudi nekatere neničelne matrike.

Primeri matrik, ki nimajo inverza

Če ima matrika A ničelno vrstico, potem nima inverza. Za vsako matriko B ima namreč produkt AB tudi ničelno vrstico, medtem ko matrika I nima ničelne vrstice. Torej je $AB \neq I$ za vsak B.

Če ima matrika A ničeln stolpec, potem nima inverza. Za vsako matriko B ima namreč produkt BA tudi ničeln stolpec, medtem ko matrika I nima ničelnega stolpca. Torej je $BA \neq I$ za vsak B.

Primeri matrik, ki imajo inverz

Elementarne matrike imajo inverze. Velja namreč

$$E_{i,j}(\alpha)^{-1} = E_{i,j}(-\alpha), \quad P_{i,j}^{-1} = P_{i,j} \quad \text{in} \quad E_i(\beta)^{-1} = E_i(\frac{1}{\beta}).$$

Primer

Če imajo matrike $A_1, \ldots, A_n$ inverze, potem ima inverz tudi njihov produkt. Velja namreč

$$(A_1 \cdots A_n)^{-1} = A_n^{-1} \cdots A_1^{-1}.$$

Inverze matrik lahko uporabimo pri reševanju kvadratnih linearnih sistemov. Če $A\mathbf{x} = \mathbf{b}$ pomnožimo z leve z A^{-1}, dobimo

$$\mathbf{x} = I\mathbf{x} = (A^{-1}A)\mathbf{x} = A^{-1}(A\mathbf{x}) = A^{-1}\mathbf{b}.$$

Naredimo še preizkus

$$A(A^{-1}\mathbf{b}) = (AA^{-1})\mathbf{b} = I\mathbf{b} = \mathbf{b}.$$

Računanje inverza z Gaussovo metodo

Matriko A razširimo na desno z identično matriko iste velikosti. Dobimo matriko $[A|I]$. To matriko obdelujemo z elementarnimi transformacijami po vrsticah toliko časa, dokler levo od črte ne dobimo bodisi matrike z ničelno vrstico bodisi identične matrike. V prvem primeru matrika A nima inverza, v drugem primeru, pa je inverz tisto, kar stoji desno od črte.

Primer

Izračunajmo inverz matrike

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}.$$

Razširjena matrika je

$$[A|I] = \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \end{array}\right].$$

Če k drugi vrstici prištejemo z -3 pomnoženo prvo vrstico, dobimo

$$\left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & -2 & -3 & 1 \end{array}\right].$$

Če drugo vrstico delimo s -2, dobimo

$$\left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 1 & \frac{3}{2} & -\frac{1}{2} \end{array}\right].$$

Če k prvi vrstici prištejemo z -2 pomnožemo drugo vrstico, dobimo

$$\left[\begin{array}{cc|cc} 1 & 0 & -2 & 1 \\ 0 & 1 & \frac{3}{2} & -\frac{1}{2} \end{array}\right].$$

Torej je

$$A^{-1} = \left[\begin{array}{cc} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{array}\right].$$

Dokažimo sedaj, da ta metoda vedno deluje.

Naj bodo $E_1, \ldots, E_n$ take elementarne matrike, da ima matrika $S = E_n \cdots E_1 A$ reducirano stopničasto obliko. Ločimo dva primera.

Če ima matrika S ničelno vrstico, potem ni obrnljiva, saj ima potem tudi matrika ST ničelno vrstico za vsak T (glej prvi primer). Odtod sledi, da tudi matrika A ni obrnljiva (uporabi drugi in tretji primer).

Če pa matrika S nima ničelne vrstice, potem je enaka identični matriki. Odtod sledi $A^{-1} = E_n \cdots E_1$. To pa je ravno tisto, kar dobimo na desni v

$$[A|I] \xrightarrow{E_1} [E_1 A|E_1] \xrightarrow{E_2} \ldots \xrightarrow{E_n} [E_n \cdots E_1 A|E_n \cdots E_1] = [I|A^{-1}].$$

Spotoma smo dokazali, da ima matrika A inverz natanko tedaj, ko je enaka produktu elementarnih matrik.

Matrike, 2. del

Karakterizacije obrnljivih matrik

Pravimo, da je kvadratna matrika **obrnljiva**, če ima inverz. Dokazali bomo, da je obrnljivost ekvivalentna mnogim drugim lastnostim matrike.

Vsako $n \times n$ matriko A lahko zapišemo kot $A = [\mathbf{a}_1 \dots \mathbf{a}_n]$, kjer so $\mathbf{a}_1, \dots, \mathbf{a}_n$ stolpci matrike A. Potem za vsako $n \times n$ matriko C velja

$$CA = C[\mathbf{a}_1 \dots \mathbf{a}_n] = [C\mathbf{a}_1 \dots C\mathbf{a}_n] \tag{1}$$

in za vsak (stolpčni) vektor $\mathbf{x}$ velja

$$A\mathbf{x} = A \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1\mathbf{a}_1 + \dots + x_n\mathbf{a}_n. \tag{2}$$

Formuli (1) in (2) bomo večkrat potrebovali v nadaljevanju.

Trditev 1

- Stolpci matrike A so linearno neodvisni natanko tedaj, ko za vsak $\mathbf{x} \in \mathbb{R}^n$, ki zadošča $A\mathbf{x} = \mathbf{0}$, velja $\mathbf{x} = 0$.
- Stolpci matrike A so ogrodje natanko tedaj, ko za vsak $\mathbf{b} \in \mathbb{R}^n$ obstaja tak $\mathbf{x} \in \mathbb{R}^n$, da velja $A\mathbf{x} = \mathbf{b}$.

Dokaz: Spomnimo se, da so stolpci matrike A **linearno neodvisni**, če za vsake $x_1, \ldots, x_n \in \mathbb{R}$, ki zadoščajo $x_1\mathbf{a}_1 + \ldots + x_n\mathbf{a}_n = 0$, velja $x_1 = \ldots = x_n = 0$. S pomočjo formule (2) to zapišemo takole: Za vsak $\mathbf{x} \in \mathbb{R}^n$, ki zadošča $A\mathbf{x} = \mathbf{0}$, velja $\mathbf{x} = 0$.

Spomnimo se, da so stolpci matrike A **ogrodje**, če za vsak $\mathbf{b} \in \mathbb{R}^n$ obstajajo taki $x_1, \ldots, x_n \in \mathbb{R}$, da velja $\mathbf{b} = x_1\mathbf{a}_1 + \ldots + x_n\mathbf{a}_n$.
S pomočjo formule (2) to zapišemo takole: Za vsak $\mathbf{b} \in \mathbb{R}^n$ obstaja tak $\mathbf{x} \in \mathbb{R}^n$, da velja $A\mathbf{x} = \mathbf{b}$.

Trditev 2

Naj bo A poljubna $n \times n$ matrika in C poljubna obrnljiva $n \times n$ matrika.

- Če so stolpci matrike A linearno neodvisni, potem so tudi stolpci matrike CA linearno neodvisni.
- Če so stolpci matrike A ogrodje, potem so tudi stolpci matrike CA ogrodje.

Dokaz: Recimo, da so stolpci matrike A linearno neodvisni, se pravi, da iz $A\mathbf{x} = 0$ sledi $\mathbf{x} = 0$. Radi bi pokazali, da so potem tudi stolpci matrike CA linearno neodvisni, se pravi, da iz $CA\mathbf{x} = 0$ sledi $\mathbf{x} = 0$. Iz $CA\mathbf{x} = 0$ sledi $A\mathbf{x} = C^{-1}(CA\mathbf{x}) = C^{-1}\mathbf{0} = \mathbf{0}$, torej je res $\mathbf{x} = \mathbf{0}$.

Pokažimo še drugi del. Recimo, da so stolpci matrike A ogrodje. Torej za vsak $\mathbf{b} \in \mathbb{R}^n$ obstaja tak $\mathbf{x} \in \mathbb{R}^n$, da velja $A\mathbf{x} = C^{-1}\mathbf{b}$. Če pomnožimo s C, dobimo, da za vsak $\mathbf{b} \in \mathbb{R}^n$ obstaja tak $\mathbf{x} \in \mathbb{R}^n$, da velja $CA\mathbf{x} = \mathbf{b}$. Torej so stolpci matrike CA res ogrodje.

Trditev 3

Če je kvadratna matrika A reducirana vrstična stopničasta forma, potem so naslednje trditve ekvivalentne:

1. Stolpci A so linearno neodvisni.
2. Stolpci A so ogrodje.
3. A je identična matrika.

Dokaz: Očitno iz točke 3 sledita točki 1 in 2. Pokažimo še, da iz negacije točke 3 sledita negaciji točk 1 in 2.

Če kvadratna vrstična kanonična forma ni identiteta, potem ima stopnico, ki je daljša od 1. Ker je stolpec, ki je na drugem mestu v neki stopnici, linearna kombinacija prejšnjih stolpcev, stolpci take matrike niso linearno neodvisni.

Po drugi strani iz obstoja stopnice, ki je daljša od 1, sledi, da stopnišče ne pride do zadnje vrstice. Zadnja vrstica je torej ničelna. Linearna ogrinjača stolpcev torej ne vsebuje vektorjev, ki imajo zadnjo komponento neničelno. Stolpci matrike A zato niso ogrodje.

Posledica 1

Za vsako kvadratno matriko A so ekvivalentne trditve:

1. Stolpci A so linearno neodvisni.
2. Stolpci A so ogrodje.
3. Matrika A je produkt elementarnih matrik.

Dokaz: Ker ima identična matrika linearno neodvisne stolpce in ker so elementarne matrike obrnljive, ima po Trditvi 2 tudi produkt elementarnih matrik linearno neodvisne stolpce. Torej iz točke 3 sledi točke 1. Podobno dokažemo, da iz točke 3 sledi točka 2.

Dokažimo sedaj, da iz točke 1 sledi točka 3. Dokaz, da iz točke 2 sledi točka 3 je podoben. Po Gaussovem algoritmu obstajajo take elementarne matrike $E_1, \ldots, E_k$, da je $A' := E_k \cdots E_1 A$ reducirana vrstična stopničasta forma. Če so stolpci A linearno neodvisni, so po Trditvi 2 tudi stolpci A' linearno neodvisni. Po Trditvi 3 sledi, da je $A' = I$. Ker so E_i obrnljive matrike, sledi $A = E_1^{-1} \cdots E_k^{-1}$. Ker je inverz elementarne matrike elementarna matrika, je torej A produkt elementarnih matrik.

Dokažimo še naslednjo posledico:

Posledica 2

Za vsako kvadratno matriko A so ekvivalentne trditve:

1. A je obrnljiva.
2. Obstaja taka matrika B, da je $AB = I$.
3. Stolpci A so ogrodje.

Dokaz: Očitno iz točke 1 sledi točka 2. Po Posledici 1 iz točke 3 sledi točka 1, saj je produkt elementarnih matrik obrnljiva matrika.

Dokažimo še, da iz točke 2 sledi točka 3. Po Trditvi 1 je dovolj dokazati, da za vsak $\mathbf{b} \in \mathbb{R}^n$ obstaja tak $\mathbf{x} \in \mathbb{R}^n$, da velja $A\mathbf{x} = \mathbf{b}$. Lahko vzamemo kar $\mathbf{x} = B\mathbf{b}$.

Povzemimo vse rezultate v naslednji izrek:

Izrek o karakterizaciji obrnljivih matrik

Za vsako kvadratno matriko A so ekvivalentne trditve:

1. A je obrnljiva.
2. Obstaja taka matrika B, da je $AB = I$.
3. Stolpci A so ogrodje.
4. Za vsak $\mathbf{b} \in \mathbb{R}^n$ obstaja tak $\mathbf{x} \in \mathbb{R}^n$, da velja $A\mathbf{x} = \mathbf{b}$.
5. Stolpci A so linearno neodvisni.
6. Za vsak $\mathbf{x} \in \mathbb{R}^n$, ki zadošča $A\mathbf{x} = \mathbf{0}$, velja $\mathbf{x} = 0$.
7. A je produkt elementarnih matrik.
8. Vrstična kanonična forma za A je identiteta.

V Posledici 3 bomo dodali še točke 9 - 12. V poglavju o determinantah bomo pokazali še točko 13: A je obrnljiva natanko tedaj, ko je $\det A \neq 0$.

Trditev 4

Transponiranka obrnljive matrike je obrnljiva matrika.

Dokaz: Če je A obrnljiva, potem je $AA^{-1} = A^{-1}A = I$, odkoder sledi $(A^{-1})^T A^T = A^T (A^{-1})^T = I^T = I$. Torej je A^T obrnljiva in velja $(A^T)^{-1} = (A^{-1})^T$.

Če uporabimo izrek na matriki A^T namesto A in če upoštevamo Trditev 4, potem dobimo naslednjo posledico:

Posledica 3

Za vsako kvadratno matriko A so ekvivalentne trditve:

1. A je obrnljiva.
9. A^T je obrnljiva.
10. Obstaja taka matrika B, da je $BA = I$.
11. Vrstice A so ogrodje.
12. Vrstice A so linearno neodvisne.

Determinante

Definicija determinante

Vsaki kvadratni matriki A bomo priredili realno število $\det A$, ki mu pravimo **determinanta** matrike A. Povedali bomo, kako izračunamo determinante matrik velikosti 1×1 in kako se determinante matrik velikosti $n \times n$ izražajo z determinantami matrik velikosti $(n-1) \times (n-1)$. Potem bomo znali izračunati determinanto matrik poljubne velikosti.

Najprej definirajmo determinanto za matrike velikosti 1×1.

$$\det [a] = a.$$

Za vsako matriko A velikosti $n \times n$ in za vsak $i, j = 1, \ldots, n$ označimo z $A_{i,j}$ matriko velikosti $(n-1) \times (n-1)$, ki jo dobimo tako, da v matriki A zbrišemo i-to vrstico in j-ti stolpec. Definirajmo

$$\det A = a_{1,1} \det A_{1,1} - a_{1,2} \det A_{1,2} + a_{1,3} \det A_{1,3} - \ldots + (-1)^{n+1} a_{1,n} \det A_{1,n}.$$

Na kratko to zapišemo s formulo $\det A = \sum_{i=1}^{n} (-1)^{1+i} a_{1,i} \det A_{1,i}$.

Primer - determinata 2×2 matrike

Izpeljimo formulo za determinante matrik velikosti 2×2. Če je

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix},$$

potem je $A_{1,1} = [a_{2,2}]$ in $A_{1,2} = [a_{2,1}]$. Po definiciji determinant matrik velikosti 1×1 je $\det A_{1,1} = a_{2,2}$ in $\det A_{1,2} = a_{2,1}$. Zato velja

$$\det A = a_{1,1} \det A_{1,1} - a_{1,2} \det A_{1,2} = a_{1,1}a_{2,2} - a_{1,2}a_{2,1}.$$

Primer - determinata 3×3 matrike

Izpeljimo še formulo za determinanto 3×3 matrike. Naj bo

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{bmatrix}$$

Potem je

$$\det A = a_{1,1} \det A_{1,1} - a_{1,2} \det A_{1,2} + a_{1,3} \det A_{1,3}$$

kjer je

$$A_{1,1} = \begin{bmatrix} a_{2,2} & a_{2,3} \\ a_{3,2} & a_{3,3} \end{bmatrix}, \quad A_{1,2} = \begin{bmatrix} a_{2,1} & a_{2,3} \\ a_{3,1} & a_{3,3} \end{bmatrix} \quad \text{in} \quad A_{1,3} = \begin{bmatrix} a_{2,1} & a_{2,2} \\ a_{3,1} & a_{3,2} \end{bmatrix}.$$

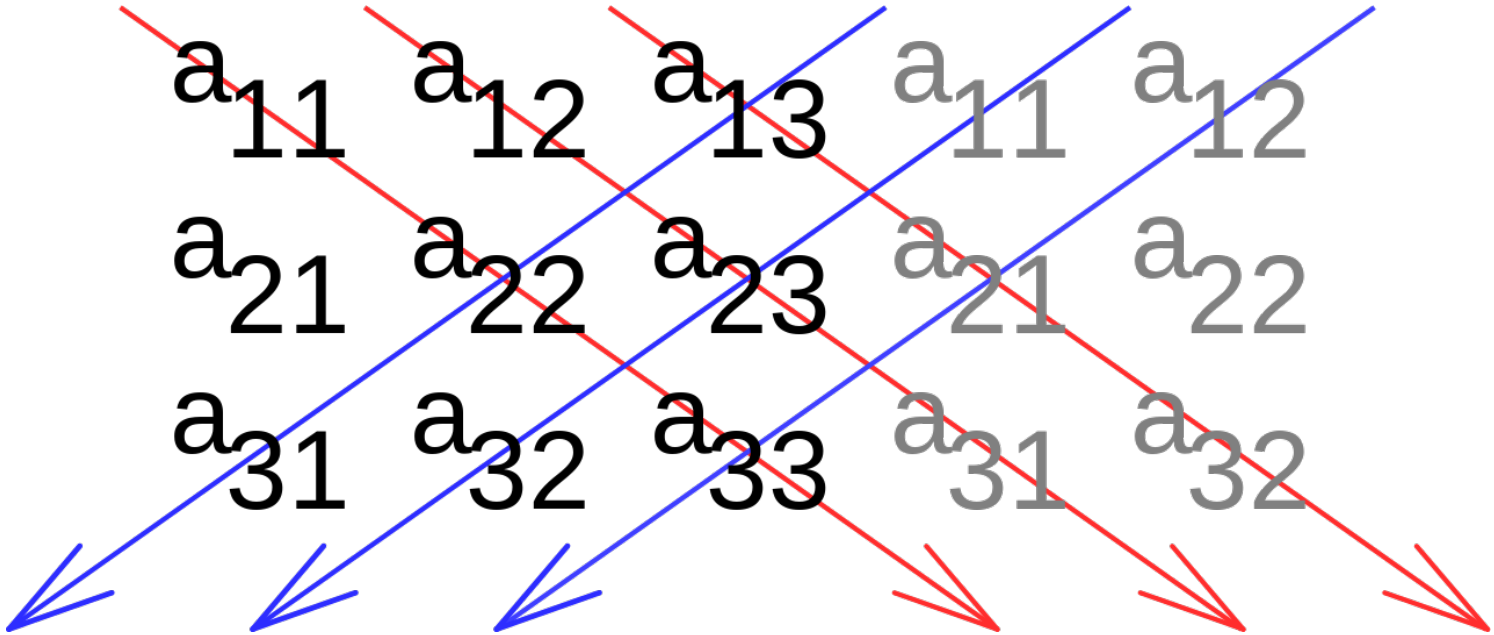
Po formuli iz prejšnjega primera je

$$\begin{aligned} \det A_{1,1} &= a_{2,2}a_{3,3} - a_{2,3}a_{3,2}, \\ \det A_{1,2} &= a_{2,1}a_{3,3} - a_{2,3}a_{3,1}, \\ \det A_{1,3} &= a_{2,1}a_{3,2} - a_{2,2}a_{3,1}. \end{aligned}$$

Zato velja

$$\begin{aligned} \det A &= a_{1,1}(a_{2,2}a_{3,3} - a_{2,3}a_{3,2}) - a_{1,2}(a_{2,1}a_{3,3} - a_{2,3}a_{3,1}) + a_{1,3}(a_{2,1}a_{3,2} - a_{2,2}a_{3,1}) \\ &= a_{1,1}a_{2,2}a_{3,3} + a_{1,2}a_{2,3}a_{3,1} + a_{1,3}a_{2,1}a_{3,2} - a_{1,3}a_{2,2}a_{3,1} - a_{1,2}a_{2,1}a_{3,3} - a_{1,1}a_{2,3}a_{3,2}. \end{aligned}$$

Formulo za determinanto 3×3 matrike si zapomnimo s Sarrusovim pravilom:



Najprej k matriki pripišemo njena prva dva stolpca. Vsaka od šestih puščic predstavlja produkt treh elementov, ki jih prečka. Rdeče puščice upoštevamo s pozitivnim predznakom, modre pa z negativnim predznakom.

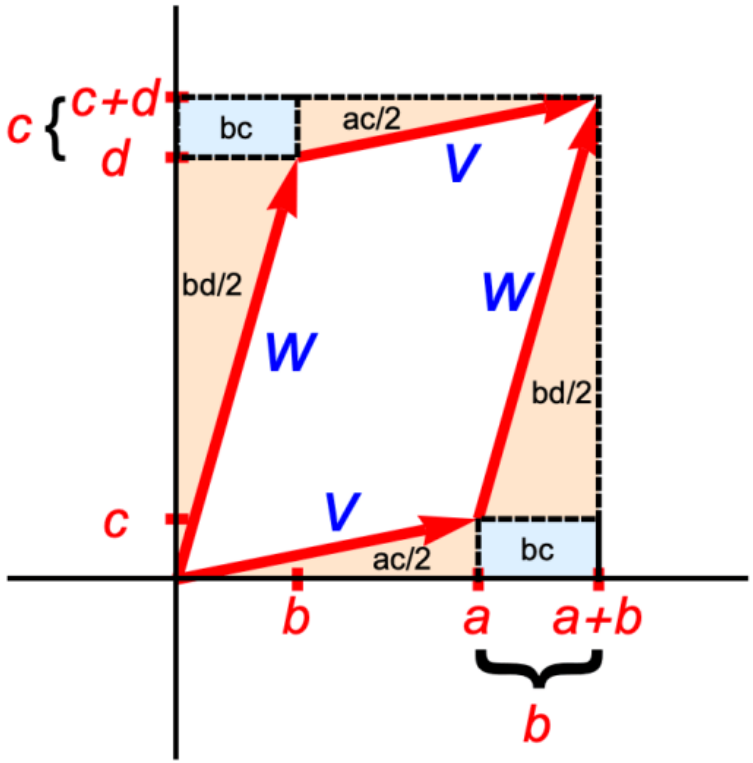
Opomba: Sarrusovo pravilo **ne** velja za 4×4 in večje determinante.

Geometrijski pomen determinante

Absolutna vrednost determinante

$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$$

je enaka ploščini paralelograma, ki ga razpenjata stolpca te determinante.



Paralelogram, ki ga razpenjata vektorja $\mathbf{v} = \begin{bmatrix} a \\ c \end{bmatrix}$ in $\mathbf{w} = \begin{bmatrix} b \\ d \end{bmatrix}$ ima ploščino

$$(a+b)(c+d) - 2bc - 2\frac{ac}{2} - 2\frac{bd}{2} =$$

$$= ad - bc = \det A.$$

Če bi $\mathbf{w}$ ležal na drugi strani $\mathbf{v}$ bi dobili $bc - ad = -\det A$. Predznak $\det A$ je torej povezan z orientacijo $\mathbf{v}$ in $\mathbf{w}$.

Determinanta 3×3 matrike s stolpci $\mathbf{u}, \mathbf{v}, \mathbf{w}$ je enaka

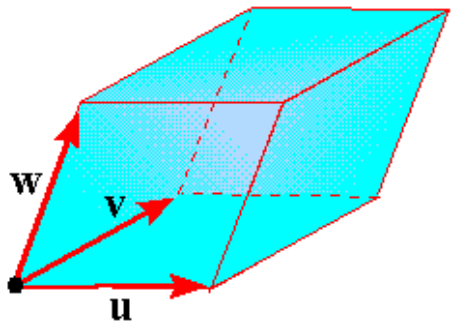
$$u_1(v_2w_3 - v_3w_2) - v_1(u_2w_3 - u_3w_2) + w_1(u_2v_3 - u_3v_2)$$

kar je enako mešanemu produktu

$$\langle \mathbf{u} \times \mathbf{v}, \mathbf{w} \rangle = (u_2v_3 - u_3v_2)w_1 + (u_3v_1 - u_1v_3)w_2 + (u_1v_2 - u_2v_1)w_3.$$

Vemo, da je absolutna vrednost mešanega produkta $\langle \mathbf{u} \times \mathbf{v}, \mathbf{w} \rangle$ enaka volumnu paralelepipeda, ki ga razpenjajo vektorji $\mathbf{u}, \mathbf{v}, \mathbf{w}$. Torej je

$$|\det \begin{bmatrix} \mathbf{u} & \mathbf{v} & \mathbf{w} \end{bmatrix}| = \text{volumen}$$



Opomba: Predznak $\det \begin{bmatrix} \mathbf{u} & \mathbf{v} & \mathbf{w} \end{bmatrix}$ je povezan z orientacijo vektorjev $\mathbf{u}, \mathbf{v}, \mathbf{w}$. Če zadoščajo pravilu desnega vijaka, je predznak pozitiven.

Razvoj determinante po vrstici ali stolpcu

Brez dokaza povejmo naslednji formuli.

Izrek o razvoju determinante

Če je A matrika velikosti $n \times n$, potem za poljubna $i, j \in \{1, \ldots, n\}$ veljata naslednji formuli:

- formula za razvoj $\det A$ po i-ti vrstici

$$\det A = \sum_{k=1}^{n} a_{i,k}(-1)^{i+k} \det A_{i,k};$$

- formula za razvoj $\det A$ po j-tem stolpcu

$$\det A = \sum_{k=1}^{n} a_{k,j}(-1)^{k+j} \det A_{k,j}.$$

Formula za razvoj po prvi vrstici se ujema z definicijo determinante.

Ponavadi razvijemo determinanto po tisti vrstici ali stolpcu, ki vsebuje največ ničel, saj to najbolj skrajša računanje.

Primer

Izračunajmo determinanto matrike

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 0 & 1 & -1 \\ 1 & -1 & 3 \end{bmatrix}$$

z razvojem po prvem stolpcu. Velja

$$\det A = 2 \cdot \det \begin{bmatrix} 1 & -1 \\ -1 & 3 \end{bmatrix} - 0 \cdot \det \begin{bmatrix} 1 & 1 \\ -1 & 3 \end{bmatrix} + 1 \cdot \det \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} =$$

$$= 2 \cdot 2 - 0 \cdot 4 + 1 \cdot (-2) = 2.$$

Determinante, 2. del

Linearnost determinante

Pokažimo najprej na primeru, da determinanta v splošnem ni linearna.

Primer

Naj bo

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

Potem je $\det(A+B) = \det I = 1$ in $\det A + \det B = 0 + 0 = 0$. Torej je

$$\det(A+B) \neq \det A + \det B$$

Naj bo sedaj $C = I$ in $\alpha = 2$. Potem je $\det(\alpha C) = 4$ in $\alpha \det C = 2$, torej

$$\det(\alpha C) \neq \alpha \det C$$

V nadaljevanju bomo pokazali, da je determinanta linearna v vsaki vrstici in stolpcu.

Trditev 1

Velja naslednja formula, ki ji pravimo linearnost determinante v i-ti vrstici:

$$\det \begin{bmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & & \vdots \\ a_{i-1,1} & \dots & a_{i-1,n} \\ \beta b_{i,1} + \gamma c_{i,1} & \dots & \beta b_{i,n} + \gamma c_{i,n} \\ a_{i+1,1} & \dots & a_{i+1,n} \\ \vdots & & \vdots \\ a_{n,1} & \dots & a_{n,n} \end{bmatrix} = \tag{1}$$

$$= \beta \det \begin{bmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & & \vdots \\ a_{i-1,1} & \dots & a_{i-1,n} \\ b_{i,1} & \dots & b_{i,n} \\ a_{i+1,1} & \dots & a_{i+1,n} \\ \vdots & & \vdots \\ a_{n,1} & \dots & a_{n,n} \end{bmatrix} + \gamma \det \begin{bmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & & \vdots \\ a_{i-1,1} & \dots & a_{i-1,n} \\ c_{i,1} & \dots & c_{i,n} \\ a_{i+1,1} & \dots & a_{i+1,n} \\ \vdots & & \vdots \\ a_{n,1} & \dots & a_{n,n} \end{bmatrix}$$

Dokaz formule (1): Razvijmo determinanto matrike

$$A(\beta,\gamma) = \begin{bmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & & \vdots \\ a_{i-1,1} & \dots & a_{i-1,n} \\ \beta b_{i,1} + \gamma c_{i,1} & \dots & \beta b_{i,n} + \gamma c_{i,n} \\ a_{i+1,1} & \dots & a_{i+1,n} \\ \vdots & & \vdots \\ a_{n,1} & \dots & a_{n,n} \end{bmatrix}$$

po i-ti vrstici. Dobimo

$$\begin{aligned} \det A(\beta,\gamma) &= \sum_{k=1}^{n} (-1)^{i+k} (\beta b_{i,k} + \gamma c_{i,k}) \det A_{i,k} \\ &= \beta \sum_{k=1}^{n} (-1)^{i+k} b_{i,k} \det A_{i,k} + \gamma \sum_{k=1}^{n} (-1)^{i+k} c_{i,k} \det A_{i,k} \end{aligned}$$

Odtod sledi želena formula

$$\det A(\beta,\gamma) = \beta \det A(1,0) + \gamma \det A(0,1)$$

Zamenjava dveh vrstic v determinanti

V nadaljavanju bomo potrebovali naslednjo trditev:

Trditev 2

Če v matriki zamenjamo dve vrstici, se njeni determinanti spremeni predznak. Se pravi $\det P_{i,j}A = -\det A$.

Dokaz: Trditev velja za 2×2 matrike, ker je

$$\det \begin{bmatrix} a_{2,1} & a_{2,2} \\ a_{1,1} & a_{1,2} \end{bmatrix} = a_{1,2}a_{2,1} - a_{1,1}a_{2,2} = -\det \begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix}.$$

Predpostavimo, da trditev velja za $(n-1) \times (n-1)$ matrike, kjer je $n \geq 3$.

Vzemimo poljubno $n \times n$ matriko A in poljubna i in j, ki sta različna. Potem za poljuben k, ki je različen od i in j, lahko $\det P_{i,j}A$ izračunamo tako, da jo razvijemo po k-ti vrstici, upoštevamo indukcijsko predpostavko, izpostavimo -1 in upoštevamo formulo za razvoj $\det A$ po k-ti vrstici. Dobimo ravno $-\det A$.

Iz trditve sledi naslednja:

Posledica

Če ima matrika dve vrstici enaki, potem je njena determinanta enaka nič.

Dokaz: Če sta i-ta in j-ta vrstica matrike A enaki, potem je $P_{i,j}A = A$. Odtod sledi $\det P_{i,j}A = \det A$. Toda zgoraj smo dokazali, da je $\det P_{i,j}A = -\det A$. Od tod sledi $\det A = -\det A$, torej je res $\det A = 0$.

Trditev 3

Če v matriki A k eni vrstici prištejemo večkratnik druge vrstice, potem se njena determinanta ne spremeni. Se pravi $\det(E_{i,j}(\alpha)A) = \det A$.

Dokaz: Po Trditvi 1 in po Posledici je

$$\det(E_{i,j}(\alpha)A) = \det \begin{bmatrix} \vdots & \vdots \\ a_{i,1} + \alpha a_{j,1} & a_{i,n} + \alpha a_{j,n} \\ \vdots & \vdots \\ a_{j,1} & a_{j,n} \\ \vdots & \vdots \end{bmatrix} =$$

$$= \det \begin{bmatrix} \vdots & \vdots \\ a_{i,1} & a_{i,n} \\ \vdots & \vdots \\ a_{j,1} & a_{j,n} \\ \vdots & \vdots \end{bmatrix} + \alpha \det \begin{bmatrix} \vdots & \vdots \\ a_{j,1} & a_{j,n} \\ \vdots & \vdots \\ a_{j,1} & a_{j,n} \\ \vdots & \vdots \end{bmatrix} =$$

$$= \det A + \alpha \cdot 0 = \det A$$

Gaussova metoda za računanje determinant

Gaussova metoda nam pove, kako dano matriko s pomočjo elementarnih vrstičnih transformacij prevedemo na reducirano vrstično stopničasto formo. Da bi to metodo uporabili za računanje determinant, potrebujemo odgovora na naslednji vprašanji:

- Kako elementarne vrstične transformacije vplivajo na determinanto?
- Kako izračunamo determinanto reducirane stopničaste vrstične forme?

Na prvo vprašanje smo že odgovorili v prejšnjih razdelkih.

Trditev

- Če v matriki A k eni vrstici prištejemo večkratnik druge vrstice, potem se njena determinanta ne spremeni.
- Če v matriki A zamenjamo dve vrstici, potem se njeni determinanti spremeni predznak.
- Če v matriki A eno vrstico pomnožimo z β, potem se tudi njena determinanta pomnoži z β.

Ker je reducirana vrstična stopničasta forma vedno zgornje trikotna, dobimo odgovor na drugo vprašanje iz naslednje trditve:

Trditev

Determinanta zgornje trikotne matrike je enaka produktu elementov na njeni glavni diagonali.

Dokaz: Če večkrat uporabimo formulo za razvoj po prvem stolpcu, dobimo

$$\det \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ 0 & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{n,n} \end{bmatrix} = a_{1,1} \det \begin{bmatrix} a_{2,2} & \dots & a_{2,n} \\ \vdots & \ddots & \vdots \\ 0 & \dots & a_{n,n} \end{bmatrix} =$$

$$= a_{1,1}a_{2,2} \det \begin{bmatrix} a_{3,3} & \dots & a_{3,n} \\ \vdots & \ddots & \vdots \\ 0 & \dots & a_{n,n} \end{bmatrix} = \dots = a_{1,1}a_{2,2} \cdots a_{n,n}.$$

Primer

S pomočjo Gaussove metode izračunaj determinanto matrike

$$A = \begin{bmatrix} 1 & 2 & 1 & 2 \\ 2 & 0 & 1 & -1 \\ -1 & 1 & 1 & -1 \\ 0 & 2 & 2 & 1 \end{bmatrix}.$$

S transformacijama $E_{2,1}(-2)$ in $E_{3,1}(1)$ dobimo

$$\det A = \det \begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & -4 & -1 & -5 \\ 0 & 3 & 2 & 1 \\ 0 & 2 & 2 & 1 \end{bmatrix}.$$

S transformacijama $E_{3,2}(\frac{3}{4})$ in $E_{4,2}(\frac{1}{2})$ dobimo

$$\det A = \det \begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & -4 & -1 & -5 \\ 0 & 0 & \frac{5}{4} & -\frac{11}{4} \\ 0 & 0 & \frac{3}{2} & -\frac{3}{2} \end{bmatrix}.$$

S transformacijo $E_{4,3}(-\frac{6}{5})$ dobimo

$$\det A = \det \begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & -4 & -1 & -5 \\ 0 & 0 & \frac{5}{4} & -\frac{11}{4} \\ 0 & 0 & 0 & \frac{9}{5} \end{bmatrix}.$$

Formula za determinanto zgornje trikotne matrike nam da

$$\det A = 1 \cdot (-4) \cdot \frac{5}{4} \cdot \frac{9}{5} = -9.$$

Determinanta produkta matrik

Dokažimo, da je determinanta produkta enaka produktu determinant.

Izrek o determinanti produkta matrik

Za poljubni kvadratni matriki A in B velja $\det AB = \det A \det B$.

Dokaz: Če v formule

$$\begin{aligned} \det(E_{i,j}(\alpha)B) &= \det B, \\ \det(P_{i,j}B) &= -\det B, \\ \det(E_i(\beta)B) &= \beta \det B \end{aligned} \tag{2}$$

vstavimo namesto B identično matriko, dobimo

$$\begin{aligned} \det(E_{i,j}(\alpha)) &= 1, \\ \det(P_{i,j}) &= -1, \\ \det(E_i(\beta)) &= \beta. \end{aligned} \tag{3}$$

Iz formul (2) in (3) sledi

$$\begin{aligned} \det(E_{i,j}(\alpha)B) &= \det E_{i,j}(\alpha)\det B, \\ \det(P_{i,j}B) &= \det P_{i,j}\det B, \\ \det(E_i(\beta)B) &= \det E_i(\beta)\det B. \end{aligned} \tag{4}$$

Torej formula $\det AB = \det A \det B$ velja, ko je A elementarna matrika.

Formula $\det AB = \det A \det B$ velja tudi v primeru, ko ima matrika A v eni vrstici same ničle, saj ima potem tudi matrika AB v isti vrstici same ničle, torej je $\det A = \det AB = 0$, kar vidimo iz razvoja po tej vrstici.

Lotimo se sedaj splošnega primera. Po Gaussu obstajajo take elementarne matrike $E_1, \ldots, E_n$, da ima matrika $S = E_n \cdots E_1 A$ reducirano vrstično stopničasto obliko. Ločimo dva primera.

Če ima matrika S ničelno vrstico, potem velja

$$\det E_n \cdots \det E_1 \det AB = \det(E_n \cdots E_1 A)B = \det SB = 0,$$

$$\det E_n \cdots \det E_1 \det A = \det E_n \cdots E_1 A = \det S = 0.$$

Po krajšanju dobimo $\det A = \det AB = 0$. Sledi $\det AB = \det A \det B$.

Če pa matrika S nima ničelne vrstice, potem je $S = I$. Odtod sledi, da je

$$\det E_n \cdots \det E_1 \det AB = \det(E_n \cdots E_1 A)B = \det IB = \det B,$$

$$\det E_n \cdots \det E_1 \det A = \det E_n \cdots E_1 A = \det I.$$

Drugo formulo pomnožimo z $\det B$ in primerjamo leve strani obeh formul. Po krajšanju spet dobimo $\det AB = \det A \det B$.

Determinanta obrnljive matrike

Kako z determinanto preverimo, ali je matrika obrnljiva?

Izrek o determinanti obrnljive matrike

Kvadratna matrika A je obrnljiva natanko tedaj, ko je $\det A \neq 0$.

Dokaz: Če je matrika A obrnljiva, potem obstaja taka matrika B, da velja $AB = I$. Če na obeh straneh uporabimo determinanto, dobimo

$$\det A \det B = \det AB = \det I = 1.$$

Odtod sledi, da je $\det A \neq 0$.

Recimo sedaj, da je $\det A \neq 0$. Izberimo take elementarne matrike $E_1, \ldots, E_n$, da je $S := E_n \cdots E_1 A$ reducirana vrstična stopničasta forma. Ker so elementarne matrike obrnljive, imajo neničelne determinante. Po predpostavki $\det A \neq 0$ odtod sledi, da je $\det S \neq 0$. Torej S ne more imeti ničelne vrstice. Odtod sledi, da je $S = I$. Torej je $A = E_1^{-1} \ldots E_n^{-1}$ obrnljiva matrika.

Determinanta transponirane matrike

Če matriko transponiramo, se njena determinanta ne spremeni.

Izrek o determinanti transponirane matrike

Za vsako kvadratno matriko A velja $\det A^T = \det A$.

Dokaz. Najprej opazimo, da izrek velja za elementarne matrike:

- Iz $P_{i,j}^T = P_{i,j}$ sledi $\det P_{i,j}^T = \det P_{i,j}$.
- Iz $E_i(\beta)^T = E_i(\beta)$ sledi $\det E_i(\beta)^T = \det E_i(\beta)$.
- Iz $E_{i,j}(\alpha)^T = E_{j,i}(\alpha)$ sledi $\det E_{i,j}(\alpha)^T = 1 = \det E_{i,j}(\alpha)$.

Izberimo take elementarne matrike $E_1, \ldots, E_n$, da je $S := E_n \cdots E_1 A$ reducirana vrstična stopničasta forma. Ločimo dva primera:

- Če ima S ničelno vrstico, potem ima S^T ničeln stolpec, torej je $\det S^T = 0 = \det S$.
- Če je $S = I$, potem $\det S^T = 1 = \det S$, ker je $I^T = I$.

Torej je v obeh primerih $\det S^T = \det S$.

Ker je determinanta produkta produkt determinant, velja

$$\begin{aligned} \det S &= \det(E_n \cdots E_1 A) \\ &= \det E_n \cdots \det E_1 \ \det A \end{aligned} \tag{5}$$

in

$$\begin{aligned} \det S^T &= \det(E_n \cdots E_1 A)^T \\ &= \det(A^T E_1^T \cdots E_n^T) = \\ &= \det A^T \ \det E_1^T \cdots \det E_n^T \end{aligned} \tag{6}$$

Iz (5) in (6) sledi

$$\det A^T = \det A \tag{7}$$

ker je $\det S^T = \det S$ in $\det E_i^T = \det E_i \neq 0$ za vsak $i = 1, \ldots, n$.

Determinanta bločno trikotne matrike

Kvadratna matrika je **bločno zgornje trikotna**, če je oblike

$$\begin{bmatrix} a_{1,1} & \dots & a_{1,m} & b_{1,1} & \dots & b_{1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{m,1} & \dots & a_{m,m} & b_{m,1} & \dots & b_{m,n} \\ 0 & \dots & 0 & c_{1,1} & \dots & c_{1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & c_{n,1} & \dots & c_{n,n} \end{bmatrix} = \begin{bmatrix} A & B \\ 0 & C \end{bmatrix}$$

Izrek o determinanti bločno zgornje trikotne matrike

Za vsako bločno zgornje trikotno matriko velja

$$\det \begin{bmatrix} A & B \\ 0 & C \end{bmatrix} = \det A \det C$$

Dokaz: Najprej opazimo, da velja

$$\begin{bmatrix} A & B \\ 0 & C \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & C \end{bmatrix} \begin{bmatrix} A & B \\ 0 & I \end{bmatrix}$$

Odtod sledi

$$\det \begin{bmatrix} A & B \\ 0 & C \end{bmatrix} = \det \begin{bmatrix} I & 0 \\ 0 & C \end{bmatrix} \det \begin{bmatrix} A & B \\ 0 & I \end{bmatrix}$$

Dokažimo sedaj, da je

$$\det \begin{bmatrix} A & B \\ 0 & I \end{bmatrix} = \det \begin{bmatrix} A & 0 \\ 0 & I \end{bmatrix}$$

S pomočjo elementarnih tranformacij tipa $E_{i,j}(\alpha)$ lahko namreč eno matriko prevedemo na drugo. Natančneje

$$\begin{bmatrix} A & B \\ 0 & I \end{bmatrix} = E \begin{bmatrix} A & 0 \\ 0 & I \end{bmatrix}$$

kjer je

$$E = E_{1,m+1}(b_{1,1}) \cdots E_{m,m+1}(b_{m,1}) E_{1,m+n}(b_{1,n}) \cdots E_{m,m+n}(b_{m,n})$$

Z večkratno uporabo razvoja po prvi vrsti dobimo.

$$\det \begin{bmatrix} I_m & 0 \\ 0 & C \end{bmatrix} = \det \begin{bmatrix} I_{m-1} & 0 \\ 0 & C \end{bmatrix} = \ldots = \det \begin{bmatrix} I_1 & 0 \\ 0 & C \end{bmatrix} = \det C$$

Z večkratno uporabo razvoja po zadnji vrsti dobimo

$$\det \begin{bmatrix} A & 0 \\ 0 & I_n \end{bmatrix} = \det \begin{bmatrix} A & 0 \\ 0 & I_{n-1} \end{bmatrix} = \ldots = \det \begin{bmatrix} A & 0 \\ 0 & I_1 \end{bmatrix} = \det A$$

Torej je

$$\det \begin{bmatrix} A & B \\ 0 & C \end{bmatrix} = \det \begin{bmatrix} I_m & 0 \\ 0 & C \end{bmatrix} \det \begin{bmatrix} A & 0 \\ 0 & I_n \end{bmatrix} = \det C \, \det A$$

Primer

$$\det \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 7 & 8 & 9 & 10 & 11 & 12 \\ 0 & 0 & 13 & 14 & 15 & 16 \\ 0 & 0 & 17 & 18 & 19 & 20 \\ 0 & 0 & 0 & 0 & 21 & 22 \\ 0 & 0 & 0 & 0 & 23 & 24 \end{bmatrix} = \det \begin{bmatrix} 1 & 2 \\ 7 & 8 \end{bmatrix} \det \begin{bmatrix} 13 & 14 & 15 & 16 \\ 17 & 18 & 19 & 20 \\ 0 & 0 & 21 & 22 \\ 0 & 0 & 23 & 24 \end{bmatrix}$$

$$= \det \begin{bmatrix} 1 & 2 \\ 7 & 8 \end{bmatrix} \det \begin{bmatrix} 13 & 14 \\ 17 & 18 \end{bmatrix} \det \begin{bmatrix} 21 & 22 \\ 23 & 24 \end{bmatrix} = (-6)(-4)(-2) = -48$$

Opomba: Podobna formula velja tudi za spodnjetrikotne matrike

$$\det \begin{bmatrix} A & 0 \\ B & C \end{bmatrix} = \det \begin{bmatrix} A & 0 \\ B & C \end{bmatrix}^T = \det \begin{bmatrix} A^T & B^T \\ 0 & C^T \end{bmatrix} =$$

$$= \det A^T \det C^T = \det A \det C$$

Schurov komplement

Zanima nas ali se da formulo

$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$$

posplošiti na primer, ko so a, b, c, d matrike. Odgovor ni preprost.

Trditev

Naj bodo A, B, C in D matrike velikosti $m \times m$, $m \times n$, $n \times m$ in $n \times n$.

- Če je matrika D obrnljiva, potem velja

$$\det \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \det D \, \det(A - BD^{-1}C)$$

- Če je matrika A obrnljiva, potem velja

$$\det \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \det A \, \det(D - CA^{-1}B)$$

Dokaz: Če je matrika D obrnljiva, potem velja

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} A - BD^{-1}C & B \\ 0 & D \end{bmatrix} \begin{bmatrix} I_m & 0 \\ D^{-1}C & I_n \end{bmatrix}$$

Odtod sledi, da je

$$\begin{aligned} \det \begin{bmatrix} A & B \\ C & D \end{bmatrix} &= \det \begin{bmatrix} A - BD^{-1}C & B \\ 0 & D \end{bmatrix} \det \begin{bmatrix} I_m & 0 \\ D^{-1}C & I_n \end{bmatrix} \\ &= \det(A - BD^{-1}C) \det D \det I_m \det I_n \end{aligned}$$

Če je matrika A obrnljiva, potem velja

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} I_m & 0 \\ CA^{-1} & I_n \end{bmatrix} \begin{bmatrix} A & B \\ 0 & D - CA^{-1}B \end{bmatrix}$$

Odtod sledi, da je

$$\begin{aligned} \det \begin{bmatrix} A & B \\ C & D \end{bmatrix} &= \det \begin{bmatrix} I_m & 0 \\ CA^{-1} & I_n \end{bmatrix} \det \begin{bmatrix} A & B \\ 0 & D - CA^{-1}B \end{bmatrix} \\ &= \det I_m \det I_n \det A \det(D - CA^{-1}B) \end{aligned}$$

Determinante, 3. del

Cramerovo pravilo

Cramerovo pravilo je eksplicitna formula za rešitev sistema linearnih enačb

$$A\mathbf{x} = \mathbf{b} \tag{1}$$

v primeru, ko je A kvadratna matrika z neničelno determinanto.

Naj bo n število stolpcev matrike A. Za vsak $i = 1, \ldots, n$ označimo z $A_i(\mathbf{b})$ matriko, ki jo dobimo iz matrike A tako, da njen i-ti stolpec zamenjamo z vektorjem $\mathbf{b}$.

Izrek - Cramerovo pravilo

Če je A $n \times n$ matrika z $\det A \neq 0$, potem je rešitev sistema (1) podana z

$$x_i = \frac{\det A_i(\mathbf{b})}{\det A}, \quad i = 1, \ldots, n. \tag{2}$$

Dokaz. Naj bodo $\mathbf{a}_1, \ldots, \mathbf{a}_n$ stolpci matrike A. Za vsak $i = 1, \ldots, n$ je

$$A_i(\mathbf{b}) = \begin{bmatrix} \mathbf{a}_1 & \ldots & \mathbf{a}_{i-1} & \mathbf{b} & \mathbf{a}_{i+1} & \ldots & \mathbf{a}_n \end{bmatrix}$$

Definirajmo še matriko

$$I_i(\mathbf{x}) = \begin{bmatrix} \mathbf{e}_1 & \ldots & \mathbf{e}_{i-1} & \mathbf{x} & \mathbf{e}_{i+1} & \ldots & \mathbf{e}_n \end{bmatrix}$$

kjer so $\mathbf{e}_1, \ldots, \mathbf{e}_n$ stolpci identične matrike I in je $\mathbf{x}$ rešitev sistema $A\,\mathbf{x} = \mathbf{b}$. Opazimo, da velja

$$\begin{aligned} A\,I_i(\mathbf{x}) &= A \begin{bmatrix} \mathbf{e}_1 & \ldots & \mathbf{e}_{i-1} & \mathbf{x} & \mathbf{e}_{i+1} & \ldots & \mathbf{e}_n \end{bmatrix} \\ &= \begin{bmatrix} A\mathbf{e}_1 & \ldots & A\mathbf{e}_{i-1} & A\mathbf{x} & A\mathbf{e}_{i+1} & \ldots & A\mathbf{e}_n \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{a}_1 & \ldots & \mathbf{a}_{i-1} & \mathbf{b} & \mathbf{a}_{i+1} & \ldots & \mathbf{a}_n \end{bmatrix} \\ &= A_i(\mathbf{b}) \end{aligned}$$

Odtod sledi

$$\det A \det I_i(\mathbf{x}) = \det A_i(\mathbf{b}).$$

Z razvojem $\det I_i(\mathbf{x})$ po i-ti vrstici dobimo še $\det I_i(\mathbf{x}) = x_i$.

Reševanje 2×2 sistema s Cramerovim pravilom

Rešimo naslednji sistem s Cramerovim pravilom:

$$\begin{aligned} 2x + y &= -1, \\ x + 3y &= 2. \end{aligned}$$

Rešitev se glasi

$$x = \frac{\det \begin{bmatrix} b_1 & a_{1,2} \\ b_2 & a_{2,2} \end{bmatrix}}{\det \begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix}} = \frac{\det \begin{bmatrix} -1 & 1 \\ 2 & 3 \end{bmatrix}}{\det \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}} = \frac{-5}{5} = -1,$$

$$y = \frac{\det \begin{bmatrix} a_{1,1} & b_1 \\ a_{2,1} & b_2 \end{bmatrix}}{\det \begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix}} = \frac{\det \begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix}}{\det \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}} = \frac{5}{5} = 1.$$

Reševanje 3×3 sistema s Cramerovim pravilom

S pomočjo Cramerovega pravila rešimo sistem

$$\begin{aligned} x + y + 2z &= 1 \\ 2x - y &= 1 \\ 3y + z &= -1. \end{aligned}$$

Ker je

$$\det A = \det \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{bmatrix} = \det \begin{bmatrix} 1 & 1 & 2 \\ 2 & -1 & 0 \\ 0 & 3 & 1 \end{bmatrix} = 9,$$

je rešitev sistema

$$x = \frac{\det \begin{bmatrix} b_1 & a_{1,2} & a_{1,3} \\ b_2 & a_{2,2} & a_{2,3} \\ b_3 & a_{3,2} & a_{3,3} \end{bmatrix}}{\det A} = \frac{\det \begin{bmatrix} 1 & 1 & 2 \\ 1 & -1 & 0 \\ -1 & 3 & 1 \end{bmatrix}}{\det A} = \frac{2}{9},$$

$$y = \frac{\det \begin{bmatrix} a_{1,1} & b_1 & a_{1,3} \\ a_{2,1} & b_2 & a_{2,3} \\ a_{3,1} & b_3 & a_{3,3} \end{bmatrix}}{\det A} = \frac{\det \begin{bmatrix} 1 & 1 & 2 \\ 2 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}}{\det A} = \frac{-5}{9},$$

$$z = \frac{\det \begin{bmatrix} a_{1,1} & a_{1,2} & b_1 \\ a_{2,1} & a_{2,2} & b_2 \\ a_{3,1} & a_{3,2} & b_3 \end{bmatrix}}{\det A} = \frac{\det \begin{bmatrix} 1 & 1 & 1 \\ 2 & -1 & 1 \\ 0 & 3 & -1 \end{bmatrix}}{\det A} = \frac{6}{9} = \frac{2}{3}.$$

Formula za inverz matrike

Vemo, da ima kvadratna matrika A inverz natanko tedaj, ko je $\det A \neq 0$. Vemo tudi, kako A^{-1} poiščemo z Gaussovo metodo. Nimamo pa še eksplicitne formule za A^{-1}. Izpeljali jo bomo s Cramerovim pravilom.

Definicija kofaktorske matrike

Naj bo A kvadratna matrika. Če v matriki A vsak element $a_{i,j}$ zamenjamo z njegovim **kofaktorjem** $(-1)^{i+j} \det A_{i,j}$, dobimo **kofaktorsko matriko** matrike A. Oznaka zanjo je $\tilde{A}$. (Spomnimo se, da matriko $A_{i,j}$ dobimo tako, da v matriki A pobrišemo i-to vrstico in j-ti stolpec.)

Primer kofaktorske matrike

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \Rightarrow \quad \tilde{A} = \begin{bmatrix} d & -c \\ -b & a \end{bmatrix}$$

$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \quad \Rightarrow \quad \tilde{A} = \begin{bmatrix} ei - fh & fg - di & dh - eg \\ ch - bi & ai - cg & bg - ah \\ bf - ce & cd - af & ae - bd \end{bmatrix}$$

Izrek - Formula za A^{-1}

Naj bo A $n \times n$ matrika, ki zadošča $\det A \neq 0$. Potem velja

$$A^{-1} = \frac{1}{\det A} \tilde{A}^T \tag{3}$$

kjer je $\tilde{A}$ kofaktorska matrika matrike A.

Dokaz: Matrika A^{-1} je rešitev matrične enačbe

$$AX = I.$$

Naj bodo $\mathbf{x}_1, \ldots, \mathbf{x}_n$ stolpci matrike X. Potem velja

$$A\mathbf{x}_1 = \mathbf{e}_1 \quad \ldots \quad A\mathbf{x}_n = \mathbf{e}_n$$

kjer so $\mathbf{e}_1, \ldots, \mathbf{e}_n$ stolpci matrike I.

Vzemimo poljubna i in j in izračunajmo (i, j)-ti element matrike X. Velja

$$x_{i,j} = i\text{-ti element vektorja } \mathbf{x}_j = \frac{\det A_i(\mathbf{e}_j)}{\det A}$$

Pri drugem enačaju smo uporabili Cramerovo pravilo za sistem $A\mathbf{x}_j = \mathbf{e}_j$.

Determinanto $B = A_i(\mathbf{e}_j)$ izračunamo z razvojem po i-tem stolpcu:

$$\det B = \sum_{k=1}^{n} (-1)^{k+i} b_{k,i} \det B_{k,i}$$

Upoštevamo, da je $B_{k,i} = A_{k,i}$ in da je

$$b_{k,i} = k\text{-ti element } \mathbf{e}_j = \begin{cases} 0 & k \neq j \\ 1 & k = j \end{cases}$$

pa dobimo

$$\det A_i(\mathbf{e}_j) = (-1)^{i+j} \det A_{j,i}$$

odkoder sledi

$$x_{i,j} = (-1)^{i+j} \frac{\det A_{j,i}}{\det A} = \frac{1}{\det A} (\text{kofaktor } a_{j,i})$$

Torej je res

$$A^{-1} = X = \frac{1}{\det A} \tilde{A}^T$$

Primer - Inverz 2×2 matrike

Če je

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

potem je

$$A^{-1} = \frac{1}{\det A} \tilde{A}^T = \frac{1}{ad - bc} \begin{bmatrix} d & -c \\ -b & a \end{bmatrix}^T = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Primer - Inverz 3×3 matrike

$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \Rightarrow A^{-1} = \frac{1}{\det A} \begin{bmatrix} ei - fh & ch - bi & bf - ce \\ fg - di & ai - cg & cd - af \\ dh - eg & bg - ah & ae - bd \end{bmatrix}$$

kjer $\det A = aei + bfg + cdh - ceg - afh - bdi$.

Determinante in permutacije

Za vsako naravno število n označimo $\mathbb{N}_n = \{1, 2, \ldots, n\}$. Za vsako preslikavo $\sigma : \mathbb{N}_n \to \mathbb{N}_n$ so ekvivalentne naslednje lastnosti:

- σ je injektivna (=slika različne elemente v različne elemente),
- σ je surjektivna (=vsak element je slika nekega elementa),
- σ je bijektivna (=injektivna in surjektivna).

Preslikavi, ki zadošča eni od teh treh ekvivalentnih lastnosti, pravimo **permutacija** množice $\mathbb{N}_n$. Množico vseh permutacij $\mathbb{N}_n$ označimo s S_n.

Primeri permutacij

Vseh preslikav iz $\mathbb{N}_3$ v $\mathbb{N}_3$ je 27. Od teh jih je 6 bijektivnih:

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} \quad \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \quad \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \quad \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$$

Naj bo $\mathbf{e}_i$ stolpčni vektor vellikosti n, ki ima na i-tem mestu enko, drugod pa same ničle. Vsaki permutaciji $\sigma \in S_n$ priredimo matriko

$$P_\sigma := \begin{bmatrix} \mathbf{e}_{\sigma(1)} & \mathbf{e}_{\sigma(2)} & \cdots & \mathbf{e}_{\sigma(n)} \end{bmatrix}$$

Determinanti te matrike pravimo **signatura** permutacije σ. Oznaka je

$$\operatorname{sgn}(\sigma) := \det P_\sigma.$$

Primeri

Če za σ vzamemo identično preslikavo, dobimo $P_\sigma = I$ in $\operatorname{sgn}(\sigma) = 1$.

Če za σ vzamemo **transpozicijo** elementov i in j, se pravi permutacijo

$$\sigma(k) = \begin{cases} j & \text{če } k = i \\ i & \text{če } k = j \\ k & \text{če } k \neq i, j \end{cases}$$

dobimo $P_\sigma = P_{i,j}$ in $\operatorname{sgn}(\sigma) = \det P_{i,j} = -1$.

Signature permutacij iz S_3

$$\operatorname{sgn}\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} = \det I = 1$$

$$\operatorname{sgn}\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} = \det P_{1,2} = -1$$

$$\operatorname{sgn}\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} = \det P_{1,3} = -1$$

$$\operatorname{sgn}\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = \det P_{2,3} = -1$$

$$\operatorname{sgn}\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = \det \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} = 1$$

$$\operatorname{sgn}\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} = \det \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} = 1$$

Dokažimo najprej, da za vsako permutacijo σ velja

$$\operatorname{sgn}(\sigma) \in \{-1, 1\}$$

Najprej opazimo, da velja

$$\begin{aligned} P_\sigma^T P_\sigma &= \begin{bmatrix} \mathbf{e}_{\sigma(1)}^T \\ \mathbf{e}_{\sigma(2)}^T \\ \vdots \\ \mathbf{e}_{\sigma(n)}^T \end{bmatrix} \begin{bmatrix} \mathbf{e}_{\sigma(1)} & \mathbf{e}_{\sigma(2)} & \cdots & \mathbf{e}_{\sigma(n)} \end{bmatrix} \\ &= \begin{bmatrix} \langle \mathbf{e}_{\sigma(1)}, \mathbf{e}_{\sigma(1)} \rangle & \langle \mathbf{e}_{\sigma(1)}, \mathbf{e}_{\sigma(2)} \rangle & \cdots & \langle \mathbf{e}_{\sigma(1)}, \mathbf{e}_{\sigma(n)} \rangle \\ \langle \mathbf{e}_{\sigma(2)}, \mathbf{e}_{\sigma(1)} \rangle & \langle \mathbf{e}_{\sigma(2)}, \mathbf{e}_{\sigma(2)} \rangle & \cdots & \langle \mathbf{e}_{\sigma(2)}, \mathbf{e}_{\sigma(n)} \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \mathbf{e}_{\sigma(n)}, \mathbf{e}_{\sigma(1)} \rangle & \langle \mathbf{e}_{\sigma(n)}, \mathbf{e}_{\sigma(2)} \rangle & \cdots & \langle \mathbf{e}_{\sigma(n)}, \mathbf{e}_{\sigma(n)} \rangle \end{bmatrix} \\ &= I \end{aligned}$$

Odtod sledi, da je

$$(\det P_\sigma)^2 = \det P_\sigma^T \det P_\sigma = \det P_\sigma^T P_\sigma = \det I = 1$$

Pokažimo še, da za vsaki permutaciji σ in τ velja

$$\operatorname{sgn}(\sigma \circ \tau) = \operatorname{sgn}(\sigma) \operatorname{sgn}(\tau)$$

Očitno za vsak i velja

$$P_\sigma \mathbf{e}_i = \mathbf{e}_{\sigma(i)}.$$

Odtod sledi

$$\begin{aligned} P_\sigma P_\tau &= P_\sigma \left[\mathbf{e}_{\tau(1)}, \mathbf{e}_{\tau(2)}, \ldots, \mathbf{e}_{\tau(n)} \right] \\ &= \left[P_\sigma \mathbf{e}_{\tau(1)}, P_\sigma \mathbf{e}_{\tau(2)}, \ldots, P_\sigma \mathbf{e}_{\tau(n)} \right] \\ &= \left[\mathbf{e}_{\sigma(\tau(1))}, \mathbf{e}_{\sigma(\tau(2))}, \ldots, \mathbf{e}_{\sigma(\tau(n))} \right] \\ &= P_{\sigma \circ \tau} \end{aligned}$$

Če uporabimo determinanto na obeh straneh, dobimo želeno formulo.

Explicitna izražava determinante

Radi bi dokazali naslednjo formulo

$$\det A = \sum_{\pi \in S_n} \operatorname{sgn}(\pi)\, a_{1,\pi(1)} a_{2,\pi(2)} \cdots a_{n,\pi(n)} \tag{4}$$

Primer

$$\det \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{bmatrix} =$$

$$= \operatorname{sgn} \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} a_{1,1} a_{2,2} a_{3,3} + \operatorname{sgn} \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} a_{1,2} a_{2,1} a_{3,3} +$$

$$+ \operatorname{sgn} \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} a_{1,3} a_{2,2} a_{3,1} + \operatorname{sgn} \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} a_{1,1} a_{2,3} a_{3,2} +$$

$$+ \operatorname{sgn} \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} a_{1,2} a_{2,3} a_{3,1} + \operatorname{sgn} \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} a_{1,3} a_{2,1} a_{3,2}$$

Ideja dokaza je, da vsako vrstico matrike A izrazimo kot

$$\begin{bmatrix} a_{i,1} & a_{i,2} & \dots & a_{i,n} \end{bmatrix} = \sum_{j=1}^{n} a_{i,j}\, \mathbf{e}_j^T$$

in nato upoštevamo linearnost determinante v tej vrstici. Dobimo

$$\det A = \det \begin{bmatrix} \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \end{bmatrix} \\ \vdots \\ \begin{bmatrix} a_{n,1} & a_{n,2} & \dots & a_{n,n} \end{bmatrix} \end{bmatrix} = \det \begin{bmatrix} \sum_{j_1=1}^{n} a_{i,j_1}\, \mathbf{e}_{j_1}^T \\ \vdots \\ \sum_{j_n=1}^{n} a_{i,j_n}\, \mathbf{e}_{j_n}^T \end{bmatrix} =$$

$$= \sum_{j_1=1}^{n} \dots \sum_{j_n=1}^{n} a_{i,j_1} \cdots a_{i,j_n} \det \begin{bmatrix} \mathbf{e}_{j_1}^T \\ \vdots \\ \mathbf{e}_{j_n}^T \end{bmatrix} = \sum_{(j_1,\dots,j_n)\in(\mathbb{N}_n)^n} a_{i,j_1} \cdots a_{i,j_n} \det \begin{bmatrix} \mathbf{e}_{j_1}^T \\ \vdots \\ \mathbf{e}_{j_n}^T \end{bmatrix}$$

Če sta dve izmed števil $j_1, \ldots, j_n$ enaki, potem je

$$\det \begin{bmatrix} \mathbf{e}_{j_1}^T \\ \vdots \\ \mathbf{e}_{j_n}^T \end{bmatrix} = 0,$$

ker ima determinanta dve vrstici enaki. Če pa so števila $j_1, \ldots, j_n$ paroma različna, potem je funkcija, ki pošlje vsak k v j_k, injektivna, se pravi

$$\begin{pmatrix} 1 & 2 & \ldots & n \\ j_1 & j_2 & \ldots & j_n \end{pmatrix} \in S_n$$

V tem primeru je

$$\det \begin{bmatrix} \mathbf{e}_{j_1}^T \\ \vdots \\ \mathbf{e}_{j_n}^T \end{bmatrix} = \det \begin{bmatrix} \mathbf{e}_{j_1} & & \mathbf{e}_{j_n} \end{bmatrix} = \operatorname{sgn} \begin{pmatrix} 1 & 2 & \ldots & n \\ j_1 & j_2 & \ldots & j_n \end{pmatrix}$$

Dokazali smo, da velja

$$\det A = \sum_{\left(\begin{smallmatrix} 1 & 2 & \dots & n \\ j_1 & j_2 & \dots & j_n \end{smallmatrix}\right) \in S_n} \operatorname{sgn} \begin{pmatrix} 1 & 2 & \dots & n \\ j_1 & j_2 & \dots & j_n \end{pmatrix} a_{i,j_1} \cdots a_{i,j_n}$$

kar na kratko zapišemo kot

$$\det A = \sum_{\pi \in S_n} \operatorname{sgn}(\pi)\, a_{1,\pi(1)} a_{2,\pi(2)} \cdots a_{n,\pi(n)}$$

Algebrske strukture, 1.del

Grupoidi in polgrupe

Operacija na množici nam pove, kako iz dveh elementov te množice zgradimo nov element te množice. Formalno jo definiramo takole:

Definicija operacije na množici

Operacija na množici M je funkcija iz $M \times M$ v M.

Opomba: Množica $M \times M$ se sestoji iz vseh urejenih parov (a, b), kjer sta a in b elementa M. Funkcija $\circ\colon M \times M \to M$ preslika urejeni par (a, b) v element $\circ(a, b)$. Namesto $\circ(a, b)$ bomo običajno pisali $a \circ b$.

Opomba: Na isti množici imamo lahko več različnih operacij. Da jih ločimo, uvedemo pojem grupoida. Grupoid je množica skupaj z izbrano operacijo na tej množici. Formalno ga definiramo takole:

Definicija grupoida

Grupoid je urejeni par $(M, \circ)$, kjer je M množica in je $\circ$ operacija na M.

Definicija polgrupe

Operacija $\circ$ na množici M je **asociativna**, če za vse a, b, c iz M velja

$$(a \circ b) \circ c = a \circ (b \circ c).$$

Polgrupa je tak grupoid $(M, \circ)$, kjer je $\circ$ asociativna operacija na M.
Operacija $\circ$ je **komutativna** če velja $a \circ b = b \circ a$ za vse $a, b \in M$.
Polgrupi s komutativno operacijo pravimo **komutativna polgrupa**.

Opomba: Iz asociativnosti operacije $\circ$ sledi, da je vrednost izraza $a_1 \circ a_2 \circ \ldots \circ a_n$ neodvisna od tega, kako postavimo oklepaje. Seveda pri postavljanju oklepajev ne smemo spreminjati vrstnega reda faktorjev.

Opomba: Če želimo, da je vrednost izraza $a_1 \circ a_2 \circ \ldots \circ a_n$ neodvisna tudi od vrstnega reda faktorjev, mora biti operacija $\circ$ tudi **komutativna**.

Primeri komutativnih polgrup

Naj bo $\mathbb{N} = \{1, 2, \ldots\}$ množica vseh naravnih števil. Množenje in seštevanje naravnih števil sta operaciji, ki sta tako komutativni kot asociativni. Torej sta $(\mathbb{N}, \cdot)$ in $(\mathbb{N}, +)$ komutativni polgrupi.

Polgrupa vseh $n \times n$ matrik

Naj bo $M_n(\mathbb{R})$ množica vseh $n \times n$ matrik in naj bo $\cdot$ množenje matrik. Potem je $(M_n(\mathbb{R}), \cdot)$ polgrupa. Če je $n \geq 2$, ta polgrupa ni komutativna.

Polgrupa vseh preslikav iz S v S

2) Naj bo S poljubna množica in naj bo F_S množica vseh funkcij iz S v S. Operacija $\circ$ naj bo kompozitum dveh funkcij. Potem je $(F_S, \circ)$ polgrupa. Za vse funkcije $f, g, h \in F_S$ in vse elemente $s \in S$ namreč velja

$$((f \circ g) \circ h)(s) = (f \circ g)(h(s)) = f(g(h(s))) = f((g \circ h)(s)) = (f \circ (g \circ h))(s)$$

Če ima S vsaj tri različne elemente potem polgrupa $(F_S, \circ)$ ni komutativna. Recimo, da so $a, b, c \in S$ paroma različni elementi. Naj bo f transpozicija elementov a in b, g pa transpozicija elementov a in c. Potem velja, da $f(g(a)) = f(c) = c$ ni enak $g(f(a)) = g(b) = b$. Torej $f \circ g \neq g \circ f$.

Polgrupa vseh besed s črkami iz A

Naj bo A neprazna množica črk. Naj bo M množica vseh besed, ki jih lahko sestavimo s črkami iz A. Operacija $\circ$ naj bo stikanje besed.

Recimo *aba* in *bab* sta dve besedi s črkami iz $A = \{a, b\}$. Če ju staknemo, dobimo besedo $aba \circ bab = ababab$. To ni isto kot $bab \circ aba = bababa$.

Očitno je stikanje besed asociativna operacija, torej je $(M, \circ)$ polgrupa. Če množica A vsebuje vsaj dve črki, potem je ta polgrupa nekomutativna.

Primer grupoida, ki ni polgrupa

Vektorski produkt je očitno operacija na $\mathbb{R}^3$. Pokažimo, da ta operacija ni niti asociativna niti komutativna. Velja namreč

$$\mathbf{e}_1 \times \mathbf{e}_2 = \mathbf{e}_3 \qquad (\mathbf{e}_1 \times \mathbf{e}_1) \times \mathbf{e}_2 = \mathbf{0}$$

$$\mathbf{e}_2 \times \mathbf{e}_1 = -\mathbf{e}_3 \qquad \mathbf{e}_1 \times (\mathbf{e}_1 \times \mathbf{e}_2) = \mathbf{e}_1 \times \mathbf{e}_3 = -\mathbf{e}_2$$

Grupoid $(\mathbb{R}^3, \times)$ torej ni komutativen in ni polgrupa.

Primer komutativnega grupoida, ki ni polgrupa

Označimo z $M_2(\mathbb{R})$ množico vseh 2×2 matrik. Jordanski produkt na $M_2(\mathbb{R})$ je definiran z

$$A \circ B = \frac{1}{2}(AB + BA)$$

Očitno je ta operacija komutativna. Pokažimo, da ni asociativna. Vzemimo

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

Očitno je $A \circ B = \frac{1}{2}I$ in $A \circ A = 0$. Torej je

$$(A \circ A) \circ B = 0, \qquad A \circ (A \circ B) = A \circ \frac{1}{2}I = \frac{1}{2}A$$

Torej je $(M_2(\mathbb{R}), \circ)$ komutativen grupoid, ki ni polgrupa.

Monoidi

Definicija enote

Naj bo $(M, \circ)$ grupoid. Element $e \in M$ je **enota** tega grupoida če velja $a \circ e = a$ in $e \circ a = a$ za vsak $a \in M$. Polgrupi z enoto pravimo **monoid**.

Opomba: Element $e \in M$ je **desna enota** grupoida $(M, \circ)$, če velja $a \circ e = a$ vsak $a \in M$. Element $e \in M$ je **leva enota** grupoida $(M, \circ)$, če velja $e \circ a = a$ vsak $a \in M$. Enota je torej tak element, ki je tako leva enota kot desna enota.

Trditev o enoličnosti enote

Vsak grupoid ima lahko največ eno enoto.

Dokaz: Naj bosta e in f dve enoti grupoida $(M, \circ)$. Potem za vsak $a \in M$ velja $e \circ a = a$ in za vsak $b \in M$ velja $b \circ f = b$. Če vstavimo $a = f$ in $b = e$, dobimo $e \circ f = f$ in $e \circ f = e$. Torej je $e = f$.

Opomba: Dokazali smo celo več, kot smo obljubili. Dokaz nam pove, da je vsaka leva enota grupoida enaka vsaki desni enoti grupoida. Odtod sledi, da grupoid z enoto nima drugih levih ali desnih enot.

Opomba: Če grupoid nima nobene leve enote, potem se lahko zgodi, da ima več desnih enot. Če grupoid nima nobene desne enote, potem se lahko zgodi, da ima več levih enot.

Primeri enot

- Polgrupa $(\mathbb{N}, \cdot)$ ima enoto 1. Polgrupa $(\mathbb{N}, +)$ nima enote, ker $0 \notin \mathbb{N}$.
- Polgrupa $(M_n(\mathbb{R}), \cdot)$ ima za enoto identično $n \times n$ matriko I_n.
- Polgrupa $(F_S, \circ)$ ima za enoto **identično preslikavo** iz S v S. To je preslikava definirana z $\mathrm{id}_S(x) = x$ za vsak $x \in S$.
- Polgrupa vseh besed s črkami iz A ima za enoto prazno besedo.
- Grupoid $(\mathbb{R}^3, \times)$ nima niti leve niti desne enote.

Primer polgrupe, ki ima več levih enot in nobene desne enote

Naj bo

$$M = \{ \begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix} \mid a, b \in \mathbb{R} \}$$

in naj bo $\circ$ običajno množenje matrik. Ker velja

$$\begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix} \begin{bmatrix} c & d \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} ac & ad \\ 0 & 0 \end{bmatrix}$$

je produkt dveh matrik iz M spet matrika iz M. Torej je $(M, \circ)$ grupoid. Ker je matrično množenje asociativno, je tudi operacija $\circ$ asociativna. Torej je $(M, \circ)$ polgrupa. Opazimo, da je vsak element oblike

$$\begin{bmatrix} 1 & x \\ 0 & 0 \end{bmatrix}$$

leva enota polgrupe $(M, \circ)$. Imamo torej neskončno levih enot. Odtod in iz prve opombe sledi, da nimamo nobene desne enote.

Grupe

Definicija inverza

Naj bo $(M, \circ)$ monoid z enoto e. Pravimo, da je element $a \in M$ **obrnljiv**, če obstaja tak element $b \in M$, ki zadošča $a \circ b = e$ in $b \circ a = e$. V tem primeru pravimo, da je element b **inverz** elementa a in pišemo $b = a^{-1}$. Monoidu, v katerem je vsak element obrnljiv, pravimo **grupa**.

Opomba: Element $b \in M$ je **desni inverz** elementa $a \in M$, če je $a \circ b = e$. Element $b \in M$ je **levi inverz** elementa $a \in M$, če je $b \circ a = e$. Inverz elementa $a \in M$ je torej tak element M, ki je tako levi kot desni inverz a.

Trditev o enoličnosti inverza

V vsakem monoidu ima vsak element največ en inverz.

Dokaz: Recimo, da sta elementa b in c inverza elementa a. Potem velja $a \circ b = e$, $b \circ a = e$, $a \circ c = e$ in $c \circ a = e$. Ker smo v monoidu, velja

$$b = b \circ e = b \circ (a \circ c) = (b \circ a) \circ c = e \circ c = c.$$

Opomba: V dokazu smo uporabili samo, da je b levi inverz a in da je c desni inverz a. Dokazali smo torej močnejšo trditev, da je vsak levi inverz elementa a enak vsakemu desnemu inverzu elementa a.

Opomba: Če element nima nobenega levega inverza, potem se lahko zgodi, da ima več desnih inverzov. Če element nima nobenega desnega inverza, potem se lahko zgodi, da ima več levih inverzov.

Opomba: Če monoid $(M, \circ)$ ni grupa, označimo z $M^\times$ množico vseh obrnljivih elementov v M. Opazimo, da $(M^\times, \circ)$ grupa, ker velja:

- Če $a, b \in M^\times$, potem tudi $a \circ b \in M^\times$. Velja namreč $(a \circ b)^{-1} = b^{-1} \circ a^{-1}$. Torej je $(M^\times, \circ)$ polgrupa.
- Velja $e \in M^\times$, ker je $e^{-1} = e$. Torej je $(M^\times, \circ)$ monoid.
- Če $a \in M^\times$, potem tudi $a^{-1} \in M^\times$. Velja namreč $(a^{-1})^{-1} = a$. Torej je $(M^\times, \circ)$ grupa.

Primer: Inverzi števil

- Množica vseh celih števil $\mathbb{Z}$ je grupa za operacijo seštevanja $+$. Očitno je namreč $(\mathbb{Z}, +)$ monoid z enoto 0 in vsak $x \in \mathbb{Z}$ ima inverz $-x \in \mathbb{Z}$.
- Tudi $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$, $(\mathbb{C}, +)$ so grupe.
- $(\mathbb{Q}, \cdot)$ je monoid z enoto 1, ampak ni grupa, ker element 0 ni obrnljiv. Vsak neničeln element $x \in \mathbb{Q}$ ima inverz $\frac{1}{x}$. Množica obrnljivih elementov v $\mathbb{Q}$ je torej $\mathbb{Q}^\times = \mathbb{Q} \setminus \{0\}$. Vemo, da je $(\mathbb{Q}^\times, \cdot)$ grupa.
- Tudi $(\mathbb{R}^\times, \cdot)$ in $(\mathbb{C}^\times, \cdot)$ sta grupi.

Opomba: Vse grupe v zgornjem primeru so komutativne. Komutativnim grupam pravimo tudi Abelove grupe.

Primer: Aditivna grupa matrik

Naj bo $M_{m,n}(\mathbb{R})$ množica vseh $m \times n$ matrik in naj bo operacija $+$ seštevanje matrik. Potem je $(M_{m,n}(\mathbb{R}), +)$ Abelova grupa z enoto $0_{m,n}$. Inverz matrike A v tej Abelovi grupi je matrika $-A$.

Primer: Inverzi matrik

Naj bo $M_n(\mathbb{R})$ množica vseh $n \times n$ matrik in naj bo $\cdot$ operacija množenja matrik. Pokazali smo, da je $(M_n(\mathbb{R}), \cdot)$ polgrupa z enoto I_n, torej monoid.

Vemo, da je matrika $A \in M_n(\mathbb{R})$ obrnljiva natanko tedaj, ko je $\det A \neq 0$. V tem primeru je $A^{-1} = \frac{1}{\det A}\tilde{A}^T$. Vemo tudi, da iz $AB = I$ sledi $BA = I$, torej pojmi levega inverza, desnega inverza in inverza sovpadejo.

Grupo vseh obrnljivih $n \times n$ matrik označimo z $\mathrm{GL}_n(\mathbb{R})$. Pravimo ji n-ta **glavna linearna grupa**.

Primer: Inverzi besed

Naj bo M množica vseh besed nad neko abecedo in naj bo $\circ$ stikanje besed. Potem je $(M, \circ)$ monoid, katerega enota je prazna beseda. Če staknemo neprazno besedo s poljubno besedo, dobimo neprazno besedo. Neprazne besede torej niso obrnljive, saj nimajo niti levega niti desnega inverza. Prazna beseda je seveda obrnljiva in njen inverz je spet prazna beseda.

Primer: Inverzi funkcij

Naj bo S neprazna množica in naj bo F_S množica vseh funkcij iz S v S. Vemo, da je F_S monoid za operacijo kompozitum funkcij. Pokazali bomo, da je funkcija $f : S \to S$ obrnljiva natanko tedaj, ko je bijektivna. Bijektivni funkciji iz S v S pravimo tudi **permutacija** množice S. Grupo vseh permutacij množice S bomo označili z $\mathcal{P}(S)$. Primer je $S_n := \mathcal{P}(\mathbb{N}_n)$.

Ekvivalentnost obrnljivosti in bijektivnosti sledi iz naslednjih dveh trditev:

- Funkcija $f : S \to S$ ima levi inverz natanko tedaj, ko je injektivna.
- Funkcija $f : S \to S$ ima desni inverz natanko tedaj, ko je surjektivna.

Dokaz: Če ima funkcija f levi inverz g, potem iz $f(x) = f(y)$ sledi $x = g(f(x)) = g(f(y)) = y$. Torej je f injektivna.

Če je funkcija f injektivna, potem definirajmo funkcijo $g : S \to S$ takole: Elemente, ki so v zalogi vrednosti funkcije f, preslikamo v njihove originale (ti so zaradi injektivnosti f enolično določeni). Elemente, ki niso v zalogi vrednosti funkcije f, preslikajmo v poljubne elemente. Po definiciji g velja $g(f(x)) = x$ za vsak $x \in S$, torej je g levi inverz f.

Če ima funkcija f desni inverz g, potem iz $y = f(g(y))$ sledi, da je f surjektivna, saj je vsak element $y \in S$ slika nekega elementa $g(y) \in S$.

Če je funkcija f surjektivna, potem definirajmo funkcijo $g\colon S \to S$ takole: $g(x) =$ poljuben tak $y \in S$, da velja $f(y) = x$. Obstoj y sledi iz definicije surjektivnosti. Iz definicije g sledi, da velja $f(g(x)) = x$ za vsak $x \in S$, torej je g desni inverz f. □

Za konec si oglejmo še konkreten primer. Naj bo $S = \mathbb{N}$ in naj bo funkcija $f : \mathbb{N} \to \mathbb{N}$ definirana z $f(x) = x + 1$. Očitno je f injektivna, ni pa surjektivna. Torej ima f levi inverz, nima pa desnega. Za vsak $a \in \mathbb{N}$ je

$$g(x) = \begin{cases} x - 1 & \text{če } x \geq 2 \\ a & \text{če } x = 1 \end{cases}$$

levi inverz funkcije f. Torej ima f neskončno levih inverzov.

Lahko poiščemo tudi funkcijo iz $\mathbb{N}$ v $\mathbb{N}$, ki ima neskončno desnih inverzov in nobenega levega. Primer je funkcija $f(2x) = x$ in $f(2x - 1) = 1$.

Podgrupoidi, podpolgrupe in podmonoidi

Naj bo $\circ$ operacija na množici M in naj bo N podmnožica v M. Pravimo, da je podmnožica N **zaprta** za operacijo $\circ$, če za vsaka elementa a in b iz N tudi element $a \circ b$ pripada N. V tem primeru lahko definiramo operacijo $\circ_N$ na N s predpisom $a \circ_N b := a \circ b$ za vsaka a in b iz N.

Operaciji $\circ_N$ pravimo **skrčitev** operacije $\circ$ na podmnožico N. Skrčitev $\circ_N$ podeduje številne lastnosti operacije $\circ$. Če je recimo $\circ$ asociativna, potem je tudi $\circ_N$ asociativna. Podobno velja tudi za komutativnost. Po drugi strani se obstoj enote ne podeduje vedno. Prav tako se ne obstoj inverza.

Definicije podstruktur

- Naj bo $(M, \circ)$ grupoid. Podmnožica $N \subseteq M$ je **podgrupoid** v $(M, \circ)$, če je zaprta za $\circ$. (Potem je $(N, \circ_N)$ grupoid.)
- Naj bo $(M, \circ)$ polgrupa. Podmnožica $N \subseteq M$ je **podpolgrupa** v $(M, \circ)$, če je podgrupoid v $(M, \circ)$. (Potem je $(N, \circ_N)$ polgrupa.)
- Naj bo $(M, \circ)$ monoid. Podmnožica $N \subseteq M$ je **podmonoid** v $(M, \circ)$, če je podpolgrupa in če vsebuje enoto. (Potem je $(N, \circ_N)$ monoid.)

Opomba: Če je $(M, \circ)$ polgrupa z enoto e in če je N podpolgrupa v $(M, \circ)$, potem je $N \cup \{e\}$ podmonoid v $(M, \circ)$.

Primeri

- Množica sodih naravnih števil je podpolgrupa v $(\mathbb{N}, +)$ in v $(\mathbb{N}, \cdot)$.
- Množica lihih naravnih števil je podmonoid v $(\mathbb{N}, \cdot)$.
- Množica vseh matrik oblike $\begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix}$ je podpolgrupa v $(M_2(\mathbb{R}), \cdot)$.
- Množica 2×2 matrik s elementi ≥ 0 je podmonoid v $(M_2(\mathbb{R}), +)$.
- Množica vseh besed, ki se začnejo z črko a je podpolgrupa v monoidu vseh besed. (Operacija je stikanje besed.)
- Množica vseh funkcij oblike $kx + n$ je podmonoid v monoidu vseh funkcij iz $\mathbb{R}$ v $\mathbb{R}$. (Operacija je kompozitum funkcij.)
- Množica vseh polinomov s pozitivnim vodilnim koeficientom je podpolgrupa v $(\mathbb{R}[x], +)$ ter podmonoid v $(\mathbb{R}[x], \cdot)$ in v $(\mathbb{R}[x], \circ)$.

Opomba: Če je $(M, \circ)$ polgrupa z enoto e in če je N podpolgrupa v $(M, \circ)$, potem se lahko zgodi, da ima $(N, \circ_N)$ enoto različno od e.

Primer podpolgrupe z različno enoto

Množica $M = \mathbb{N} \times \mathbb{N}$ je monoid za operacijo $(a, b) \circ (c, d) = (ac, bd)$. Monoid $(M, \circ)$ ima enoto $e = (1, 1)$. Podmnožica $N = \mathbb{N} \times \{0\}$ je zaprta za $\circ$ in ne vsebuje e. Torej je N podpolgrupa, ki ni podmonoid. Polgrupa $(N, \circ_N)$ ima enoto $f = (1, 0)$. Očitno f ni enaka enoti monoida $(M, \circ)$.

Oglejmo si še primer podgrupoida, ki ni podpolgrupa.

Primer podgrupoida

Vemo, da je $(\mathbb{R}^3, \times)$ grupoid, ki ni polgrupa. Naj bo $N \subseteq \mathbb{R}^3$ premica skozi izhodišče, se pravi linearna ogrinjača nekega vektorja. Za vsaka elementa $\mathbf{a}, \mathbf{b} \in N$ velja $\mathbf{a} \times \mathbf{b} = \mathbf{0}$, kar je element N.

Podgrupe

Spomnimo se, da je grupa tak monoid v katerem je vsak element obrnljiv.

Definicije podgrupe

Naj bo $(G, \circ)$ grupa. Podmnožica $H \subseteq G$ je **podgrupa** v $(G, \circ)$, če je podmonoid in če vsebuje inverz vsakega svojega elementa.

Opomba: Na dolgo povedano je podmnožica $H \subseteq G$ podgrupa, če:

- Za vsaka $a, b \in H$ velja $a \circ b \in H$.
- Velja $e \in H$.
- Za vsak $a \in H$ velja $a^{-1} \in H$.

Pokažimo, da lahko te tri lastnosti strnemo v eno samo.

Trditev

Naj bo množica G grupa za operacijo $\circ$. Podmnožica $H \subseteq G$ je **podgrupa** natanko tedaj, ko za vsaka $a, b \in H$ velja $a \circ b^{-1} \in H$.

Dokaz: Naj bo H taka podmnožica v G, ki zadošča lastnosti iz trditve. Preverimo, da H zadošča vsem trem lastnostim iz definicije podgrupe:

- Dokažimo, da $e \in H$. Če v lastnost iz trditve vstavimo $b = a$, dobimo $a \circ a^{-1} \in H$ za vsak $a \in H$. Torej je res $e \in H$.
- Dokažimo, da $a^{-1} \in H$ za vsak $a \in H$. Če v lastnost iz trditve vstavimo $a = e$ in $b = a$, dobimo $e \circ a^{-1} \in H$ za vsak $a \in H$.
- Dokažimo, da $a \circ b \in H$ za vsaka $a, b \in H$. Če $a, b \in H$, potem $a, b^{-1} \in H$ po prejšnji točki. Po lastnosti iz trditve je potem $a \circ (b^{-1})^{-1} \in H$. To pa je ravno $a \circ b \in H$.

Dokažimo še, da vsaka podgrupa H zadošča lastnosti iz trditve. Vzemimo poljubna $a, b \in H$. Po tretji lastnosti iz definicije podgrupe odtod sledi $a, b^{-1} \in H$. Po prvi lastnosti iz definicije podgrupe sledi $a \circ b^{-1} \in H$.

Primer: Ciklične podgrupe

Če je $(G, \circ)$ grupa, potem je za vsak element $a \in G$ množica $\langle a \rangle := \{a^m \mid m \in \mathbb{Z}\}$ podgrupa v $(G, \circ)$. Izraz a^m definiramo z $a^0 = e$, $a^n = \underbrace{a \circ \ldots \circ a}_{\text{n-krat}}$ in $a^{-n} = \underbrace{a^{-1} \circ \ldots \circ a^{-1}}_{\text{n-krat}}$ za vsak $n \in \mathbb{N}$.

Primeri podgrup v $\mathrm{GL}_n(\mathbb{R})$

Spomnimo se, da je $\mathrm{GL}_n(\mathbb{R}) = \{A \in M_n(\mathbb{R}) \mid \det A \neq 0\}$. Definirajmo:

- $\mathrm{SL}_n(\mathbb{R}) = \{A \in M_n(\mathbb{R}) \mid \det A = 1\}$.
- $\mathrm{UT}_n(\mathbb{R}) =$ vse zgornje trikotne $n \times n$ matrike z enkami po diagonali.
- $\mathrm{O}_n(\mathbb{R}) = \{A \in M_n(\mathbb{R}) \mid A^T A = I\}$.

Te množice so podgrupe v $\mathrm{GL}_n(\mathbb{R})$.

Primeri podgrup v S_n

Spomnimo se, da je S_n grupa vseh permutacij množice $\mathbb{N}_n = \{1, \ldots, n\}$.

- Podmnožica $A_n = \{\sigma \in S_n \mid \mathrm{sgn}(\sigma) = 1\}$ je podgrupa v S_n.
- Naj bo $\sigma = \begin{pmatrix} 1 & 2 & \ldots & n-1 & n \\ 2 & 3 & \ldots & n & 1 \end{pmatrix}$. Potem je $Z_n := \langle \sigma \rangle$ podgrupa v S_n.
- Naj bo R neka relacija na množici $\mathbb{N}_n$. Permutacija $\sigma \in S_n$ je **avtomorfizem** relacije R, če iz iRj vedno sledi $\sigma(i)R\sigma(j)$. Množica vseh avtomorfizmov relacije R je podgrupa v S_n.
- Označimo z R relacijo sosednosti oglišč v n-kotniku. Podgrupi vseh avtomorfizmov R pravimo n-ta **diedrska** grupa in jo označimo z D_n.

Lastnosti končnih grup

Zanimajo nas grupe s končno močjo (moč = število elementov).

Lagrangeov izrek

Moč podgrupe deli moč grupe.

Dokaz: Naj bo H podgrupa končne grupe $(G, \circ)$. Za vsak $g \in G$ naj bo

$$g \circ H := \{g \circ h \mid h \in H\}$$

Izrek sledi iz $G = \bigcup_{g \in G} g \circ H$ in iz naslednjih dveh pomožnih trditev:

- Za vsak g je moč množice $g \circ H$ enaka moči množice H.
- Če imata množici $g_1 \circ H$ in $g_2 \circ H$ neprazen presek, potem sta enaki.

Pokažimo najprej, da je $h \mapsto g \circ h$ bijektivna preslikava iz H v $g \circ H$. Surjektivnost sledi iz definicije $g \circ H$. Dokažimo še injektivnost. Če je $g \circ h_1 = g \circ h_2$, potem je $h_1 = g^{-1} \circ (g \circ h_1) = g^{-1} \circ (g \circ h_2) = h_2$. To nam da prvo pomožno trditev.

Če imata množici $g_1 \circ H$ in $g_2 \circ H$ neprazen presek, potem obstajata taka elementa $h_1, h_2 \in H$, da velja $g_1 \circ h_1 = g_2 \circ h_2$. Sledi $g_2^{-1} \circ g_1 = h_2 \circ h_1^{-1}$. Vzemimo sedaj poljuben element $x \in g_1 \circ H$. Potem obstaja tak element $h \in H$, da velja $x = g_1 \circ h$. Odtod sledi

$$x = e \circ g_1 \circ h = g_2 \circ g_2^{-1} \circ g_1 \circ h = g_2 \circ h_2 \circ h_1^{-1} \circ h \in g_2 \circ H.$$

Dokazali smo torej $g_1 \circ H \subseteq g_2 \circ H$. Če v dokazu zamenjamo g_1 in g_2, dobimo še obratno inkluzijo. To nam da drugo pomožno trditev. □

Primer

Grupa S_n ima $n!$ elementov. Oglejmo si njene podgrupe:

- A_n ima $\frac{n!}{2}$ elementov.
- D_n ima $2n$ elementov.
- Z_n ima n elementov.

V vseh primerih moč podgrupe deli moč grupe.

Definicija: Red elementa

Naj bo $(G, \circ)$ grupa. **Red elementa** $a \in G$ je najmanjše naravno število n, ki zadošča $a^n = e$. Če tak n ne obstaja, je red elementa a neskončen.

Trditev

Vsak element vsake končne grupe ima končen red, ki deli moč grupe.

Dokaz. Naj bo m moč grupe in a element grupe. Potem $e, a, a^2, \ldots, a^m$ ne morejo biti paroma različni, ker jih je več kot elementov grupe. Torej obstajata taki števili $k, l = 0, \ldots, m$, da je $k \neq l$ in $a^k = a^l$. Če je $k > l$, sledi $a^{k-l} = e$, sicer pa $a^{l-k} = e$. Torej ima a končen red. Če je red a enak n, potem je $\{e, a, a^2, \ldots, a^{n-1}\}$ podgrupa, torej n deli m. □

Posledica

Podmnožica končne grupe je podgrupa natanko tedaj, ko je zaprta za $\circ$.

Dokaz: Če je a element podmnožice in če je njegov red enak n, potem je $a^{-1} = a^{n-1}$ tudi element podmnožice, ker je ta zaprta za operacijo. Sledi, da podmnožica vsebuje enoto, torej izpolnjuje vse tri lastnosti podgrupe.

Primer: Red cikla

Oglejmo si najprej poseben primer permutacije, ki mu pravimo **cikel**. Naj bodo $a_1, \ldots, a_k$ paroma različni elementi množice $\mathbb{N}_n$. Označimo z

$$(a_1 \ a_2 \ \ldots \ a_k)$$

permutacijo, ki preslika $a_1 \mapsto a_2$, $a_2 \mapsto a_3$, $\ldots$, $a_{k-1} \mapsto a_k$, $a_k \mapsto a_1$, elemente iz $\mathbb{N}_n \setminus \{a_1, \ldots, a_k\}$ pa preslika same vase.

Pokažimo, da je red elementa $\sigma := (a_1 \ a_2 \ \ldots \ a_k)$ enak k. Velja

$$\sigma^l(a_i) = \begin{cases} a_{i+l} & \text{če } i + l \leq k \\ a_{i+l-k} & \text{če } i + l > k \end{cases}$$

za vsak $i, l = 1, \ldots, k$. Torej je $\sigma^k = \mathrm{id}$ in $\sigma \neq \mathrm{id}, \ldots, \sigma^{k-1} \neq \mathrm{id}$.

Primer: Red permutacije

Če cikla σ in τ nimata skupnih elementov, potem velja $\sigma \circ \tau = \tau \circ \sigma$. Potem za vsako naravno število k velja $(\sigma \circ \tau)^k = \sigma^k \circ \tau^k$. Ker σ^k in τ^k nimata skupnih elementov, je $\sigma^k \circ \tau^k = \mathrm{id}$ natanko tedaj, ko je $\sigma^k = \mathrm{id}$ in $\tau^k = \mathrm{id}$. Torej je $(\sigma \circ \tau)^k = \mathrm{id}$ natanko tedaj, ko red σ deli k in red τ deli k. Dokazali smo, da je red $\sigma \circ \tau$ enak najmanjšemu skupnemu večkratniku redov σ in τ. To trditev lahko posplošimo tudi na več ciklov.

Kratek premislek pokaže, da lahko vsako permutacijo zapišemo kot kompozitum disjunktnih ciklov. Torej je njen red enak najmanjšemu skupnemu večkratniku redov teh disjunktnih ciklov.

Izračunajmo red permutacije

$$\rho = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 5 & 1 & 3 & 2 \end{pmatrix}$$

Najprej izrazimo permutacijo kot kompozitum disjunktnih ciklov. Dobimo $\rho = (1\ 4\ 3) \circ (2\ 5)$. Red cikla $(1\ 4\ 3)$ je enak 3, red cikla $(2\ 5)$ pa je 2. Najmanjši skupni večkratnik 3 in 2 je 6. Torej je red permutacije ρ enak 6.

Algebrske strukture, 2.del

Homomorfizmi grupoidov, polgrup in monoidov

Homomorfizmi so preslikave, ki ohranjajo strukturo. Bolj natančno:

Definicija homomorfizma

Naj bosta $(M_1, \circ_1)$ in $(M_2, \circ_2)$ dva grupoida. Pravimo, da je preslikava $f : M_1 \to M_2$ **homomorfizem grupoidov**, če za vsaka $x, y \in M_1$ velja

$$f(x \circ_1 y) = f(x) \circ_2 f(y) \tag{1}$$

Homomorfizem polgrup je tak homomorfizem grupoidov, ki slika iz polgrupe v polgrupo. **Homomorfizem monoidov** je tak homomorfizem polgrup, ki slika iz monoida v monoid in preslika enoto v enoto.

Primeri homomorfizmov

Preslikava $f(x) = 2x$ je homomorfizem polgrup iz $(\mathbb{N}, +)$ v $(\mathbb{N}, +)$, ker velja $f(x + y) = 2(x + y) = 2x + 2y = f(x) + f(y)$ za vsaka $x, y \in \mathbb{N}$.

Preslikava $f(x) = x^2$ je homomorfizem monoidov iz $(\mathbb{N}, \cdot)$ v $(\mathbb{N}, \cdot)$, ker velja $f(xy) = (xy)^2 = x^2y^2 = f(x)f(y)$ za vsaka $x, y \in \mathbb{N}$ in $f(1) = 1$.

Primer: Homomorfizem polgrup, ki ne slika enote v enoto

Vemo, da je $(\mathbb{Z}, \cdot)$ polgrupa z enoto 1 in da je $(\mathbb{Z} \times \mathbb{Z}, \circ)$, kjer $(a, b) \circ (c, d) = (ac, bd)$, polgrupa z enoto $(1, 1)$. Preslikava

$$f : \mathbb{Z} \to \mathbb{Z} \times \mathbb{Z}, \quad f(x) = (x, 0)$$

je homomorfizem polgrup, ker $f(xy) = (xy, 0) = (x, 0) \circ (y, 0) =$
$= f(x) \circ f(y)$, ampak ne slika enote v enoto, ker $f(1) = (1, 0) \neq (1, 1)$.

Primer: Homomorfizem grupoidov

Naj bo $A = [\mathbf{a}_1\ \mathbf{a}_2\ \mathbf{a}_3]$ taka 3×3 matrika, katere stolpci zadoščajo $\mathbf{a}_1 \times \mathbf{a}_2 = \mathbf{a}_3$, $\mathbf{a}_2 \times \mathbf{a}_3 = \mathbf{a}_1$ in $\mathbf{a}_3 \times \mathbf{a}_1 = \mathbf{a}_2$. Potem je preslikava $\mathbf{x} \mapsto A\mathbf{x}$ homomorfizem grupoidov iz $(\mathbb{R}^3, \times)$ v $(\mathbb{R}^3, \times)$. Velja namreč

$$A\mathbf{x} \times A\mathbf{y} = (x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + x_3\mathbf{a}_3) \times (y_1\mathbf{a}_1 + y_2\mathbf{a}_2 + y_3\mathbf{a}_3)$$
$$= (x_1y_2 - x_2y_1)\mathbf{a}_1 \times \mathbf{a}_2 + (x_3y_1 - x_1y_3)\mathbf{a}_3 \times \mathbf{a}_1 + (x_2y_3 - x_3y_2)\mathbf{a}_2 \times \mathbf{a}_3$$
$$= (x_1y_2 - x_2y_1)\mathbf{a}_3 + (x_3y_1 - x_1y_3)\mathbf{a}_2 + (x_2y_3 - x_3y_2)\mathbf{a}_1 = A(\mathbf{x} \times \mathbf{y})$$

Pokažimo, da homomomorfizem monoidov slika inverze v inverze.

Trditev

Recimo, da je f homomorfizem monoidov iz monoida $(M_1, \circ_1)$ v monoid $(M_2, \circ_2)$. Če je a obrnljiv element v $(M_1, \circ_1)$, potem je $f(a)$ obrnljiv element v $(M_2, \circ_2)$ in velja $f(a)^{-1} = f(a^{-1})$.

Dokaz: Ker je element $a \in M_1$ obrnljiv, obstaja tak element $b \in M_1$, da velja $a \circ_1 b = e_1$ in $b \circ_1 a = e_1$. Ker je $f : M_1 \to M_2$ homomorfizem monoidov, odtod sledi $f(a) \circ_2 f(b) = f(a \circ_1 b) = f(e_1) = e_2$ in $f(b) \circ_2 f(a) = f(b \circ_1 a) = f(e_1) = e_2$. Torej je element $f(a)$ obrnljiv in velja $f(a)^{-1} = f(b)$. Ker je $b = a^{-1}$, sledi $f(a)^{-1} = f(a^{-1})$.

Primer

Determinanta $\det : M_n(\mathbb{R}) \to \mathbb{R}$ zadošča $\det AB = \det A \det B$ in $\det I_n = 1$. Torej je det homomorfizem monoidov iz $(M_n(\mathbb{R}), \cdot)$ v $(\mathbb{R}, \cdot)$. Po zgornji trditvi det slika obrnljive matrike v neničelna realna števila in velja $\det A^{-1} = \frac{1}{\det A}$.

Primer: Homomorfizmi monoidov

Naj bo M množica vseh funkcij iz $\mathbb{R}$ v $\mathbb{R}$ oblike $\phi_{k,l}(x) = kx + l$ in naj bo operacija $\circ$ kompozitum funkcij. Iz $k(k'x + l') + l = kk'x + kl' + l$ sledi $\phi_{k,l} \circ \phi_{k',l'} = \phi_{kk',kl'+l}$. Enota te polgrupe je $\phi_{1,0} = \mathrm{id}$. Preslikava

$$f(\phi_{k,l}) = \begin{bmatrix} k & l \\ 0 & 1 \end{bmatrix}$$

je homomorfizem monoidov iz $(M, \circ)$ v $(M_2(\mathbb{R}), \cdot)$, ker je

$$f(\mathrm{id}) = f(\phi_{1,0}) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2$$

$$f(\phi_{k,l} \circ \phi_{k',l'}) = f(\phi_{kk',kl'+l}) = \begin{bmatrix} kk' & kl'+l \\ 0 & 1 \end{bmatrix} =$$

$$= \begin{bmatrix} k & l \\ 0 & 1 \end{bmatrix} \begin{bmatrix} k' & l' \\ 0 & 1 \end{bmatrix} = f(\phi_{k,l}) f(\phi_{k',l'})$$

Homomorfizmi grup

Struktura grupe se sestoji iz treh delov: produktov, enote in inverzov. Homomorfizem grup slika produkte v produkte, enoto v enoto in inverze v izverze. Pokazali bomo, da druga in tretja lastnost sledita iz prve lastnosti. Za definicijo torej lahko vzamemo samo prvo lastnost.

Definicija homomorfizma grup

Homomorfizem grup iz grupe $(G_1, \circ_1)$ v grupo $(G_2, \circ_2)$ je taka preslikava $f : G_1 \to G_2$, ki zadošča $f(x \circ_1 y) = f(x) \circ_2 f(y)$ za vsaka $x, y \in M_1$.

Trditev

Homomorfizem grup slika enoto prve grupe v enoto druge grupe in inverz vsakega elementa iz prve grupe v inverz njegove slike.

Dokaz: Za vsak homomorfizem grup f velja

$$f(e_1) \circ_2 f(e_1) = f(e_1 \circ_1 e_1) = f(e_1) = e_2 \circ_2 f(e_1)$$

Če to pomnožimo z $f(e_1)^{-1}$ z desne, dobimo $f(e_1) = e_2$. Drugi del je posledica prvega dela in prejšnje trditve.

Primeri homomorfizmov grup

- Determinanta je homomorfizem grup iz $(\mathrm{GL}_n(\mathbb{R}), \cdot)$ v $(\mathbb{R}^\times, \cdot)$, ker velja $\det AB = \det A \det B$.
- Preslikava $\sigma \mapsto P_\sigma := [\mathbf{e}_{\sigma(1)} \dots \mathbf{e}_{\sigma(n)}]$ je homomorfizem grup iz $(S_n, \circ)$ v $(\mathrm{GL}_n(\mathbb{R}), \cdot)$, ker velja $P_{\sigma\circ\tau} = P_\sigma P_\tau$.
- Signatura permutacije je homomorfizem grup iz $(S_n, \circ)$ v $(\mathbb{R}^\times, \cdot)$, ker velja $\mathrm{sgn}(\sigma \circ \tau) = \mathrm{sgn}(\sigma)\mathrm{sgn}(\tau)$.

Opomba: Tretji homomorfizem je kompozitum prvega in drugega.

Pokažimo, da je kompozitum dveh homomorfizmov vedno homomorfizem.

Trditev

Če je f homomorfizem grup iz grupe $(G_1, \circ_1)$ v grupo $(G_2, \circ_2)$
in je g homomorfizem grup iz grupe $(G_2, \circ_2)$ v grupo $(G_3, \circ_3)$,
potem je $g \circ f$ homomorfizem grup iz grupe $(G_1, \circ_1)$ v grupo $(G_3, \circ_3)$.

Podobno velja tudi za homomorfizme grupoidov, polgrup in monoidov.

Dokaz: Vzemimo poljubna $x, y \in G_1$. Ker je f homomorfizem, velja

$$f(x \circ_1 y) = f(x) \circ_2 f(y).$$

Ker je g homomorfizem, velja

$$g(f(x) \circ_2 f(y)) = g(f(x)) \circ_3 g(f(y)).$$

Torej je

$$(g \circ f)(x \circ_1 y) = g(f(x \circ_1 y)) = g(f(x) \circ_2 f(y)) =$$
$$= g(f(x)) \circ_3 g(f(y)) = g(f(x)) \circ_3 g(f(y)) = (g \circ f)(x) \circ_3 (g \circ f)(y).$$

Izomorfizem grup

Definicija izomorfizma

Bijektivnemu homomorfizmu grup pravimo **izomorfizem** grup. Dve grupi sta **izomorfni**, če obstaja izomorfizem grup in ene v drugo.

Podobno definiramo tudi izomorfizme grupoidov, polgrup in monoidov.

Opomba: Izomorfizem grup je v resnici samo preimenovanje elementov. Pri tem se mora ustrezno preimenovati tudi tabela produktov.

Primer izomorfizma grup

Naj bo $G_1 = \{0, 1, 2\}$ in $G_2 = \{e, a, b\}$. Operaciji naj bosta definirani z

$\circ_1$	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

in

$\circ_2$	e	a	b
e	e	a	b
a	a	b	e
b	b	e	a

Potem sta $(G_1, \circ_1)$ in $(G_2, \circ_2)$ grupi. Izomorfizem je $0 \mapsto e, 1 \mapsto a, 2 \mapsto b$. S preimenovanjem elementov v tabeli za $\circ_1$ smo dobili ravno tabelo za $\circ_2$.

Pokažimo, da je inverz izomorfizma spet izomorfizem.

Trditev

Če je f homomorfizem grup iz grupe $(G_1, \circ_1)$ v grupo $(G_2, \circ_2)$ in če je f bijektivna preslikava, potem je preslikava f^{-1} tudi homomorfizem grup iz grupe $(G_2, \circ_2)$ v grupo $(G_1, \circ_1)$.

Podobno velja tudi za homomorfizme grupoidov, polgrup in monoidov.

Dokaz: Za poljubna elementa $x, y \in G_2$ velja

$$f(f^{-1}(x) \circ_1 f^{-1}(y)) = f(f^{-1}(x)) \circ_2 f(f^{-1}(y)) = x \circ_2 y = f(f^{-1}(x \circ_2 y))$$

Če upoštevamo, da je f injektivna, odtod sledi

$$f^{-1}(x) \circ_1 f^{-1}(y) = f^{-1}(x \circ_2 y).$$

Cayleyev izrek

Vsaka grupa je izomorfna kaki podgrupi v kaki grupi permutacij.

Dokaz: Naj bo $(G, *)$ grupa in naj bo $(\mathcal{P}(G), \circ)$ grupa vseh permutacij množice G. Ideja je, da konstruiramo injektiven homomorfizem grup ϕ iz $(G, *)$ v $(\mathcal{P}(G), \circ)$. Potem je $\phi(G)$ podgrupa v $(\mathcal{P}(G), \circ)$ in ϕ je bijektiven homomorfizem grup iz $(G, *)$ v $(\phi(G), \circ_{\phi(G)})$. Torej sta grupi $(G, *)$ in $(\phi(G), \circ_{\phi(G)})$ izomorfni.

Za vsak $g \in G$ lahko definiramo preslikavo $\phi_g \colon G \to G$, $\phi_g(x) = g * x$. Pokažimo, da je ϕ_g permutacija množice G. Če je $\phi_g(x) = \phi_g(y)$, potem je $x = g^{-1} * (g * x) = g^{-1} * (g * y) = y$, torej je ϕ_g injektivna. Iz $x = g * (g^{-1} * x) = \phi_g(g^{-1} * x)$ sledi, da je ϕ_g surjektivna.

Preslikavo $\phi \colon G \to \mathcal{P}(G)$ definirajmo z $\phi(g) := \phi_g$. Pokažimo najprej, da je ϕ injektivna. Iz $\phi_g = \phi_h$ sledi $g = \phi_g(e) = \phi_h(e) = h$. Pokažimo še, da je ϕ homomorfizem, se pravi, da je $\phi_{g*h} = \phi_g \circ \phi_h$ za vsaka $g, h \in G$. To sledi iz $\phi_{g*h}(x) = (g * h) * x = g * (h * x) = \phi_g(\phi_h(x)) = (\phi_g \circ \phi_h)(x)$.

Polkolobarji in kolobarji

Množici z dvema operacijama pravimo tudi **bigrupoid**. Operaciji običajno označimo s $+$ in $\cdot$, čeprav ni nujno, da imata enake lastnosti kot običajno seštevanje in množenje. Kadar med operacijama ni nobene zveze, je vseeno, če študiramo vsako zase. To pomeni, da je študij bigrupoida $(M, +, \cdot)$ enak ločenemu študiju grupoidov $(M, +)$ in $(M, \cdot)$.

Primer zanimive zveze med obema operacijama je **distributivnost**:

$$(x + y) \cdot z = (x \cdot z) + (y \cdot z) \quad \text{in} \quad z \cdot (x + y) = (z \cdot x) + (z \cdot y).$$

Definicija polkolobarja in kolobarja

Distributiven bigrupoid $(M, +, \cdot)$ je **polkolobar**, če je $(M, +)$ komutativna polgrupa. Polkolobar $(M, +, \cdot)$ je **kolobar**, če je $(M, +)$ Abelova grupa.

Opomba: Naj bo $(M, +, \cdot)$ kolobar. Enoto Abelove grupe $(M, +)$ označimo z 0. Iz distributivnosti sledi, da je $x \cdot 0 = 0$ in $0 \cdot x = 0$ za vsak $x \in M$.

Primeri polkolobarjev, ki niso kolobarji

- $(\mathbb{N}, +, \cdot)$ je polkolobar, ki ni kolobar. $(\mathbb{Z}, +, \cdot)$ je kolobar.
- Naj bo M_S množica vseh podmnožic dane množice S. Potem je M_S polkolobar za operaciji unija in presek množic.
- $(\mathbb{R}^{>0}, +, \cdot)$ je polkolobar, ki ni kolobar. $(\mathbb{R}, +, \cdot)$ je kolobar
- Na $\mathbb{R}$ vzemimo za $a + b$ maksimum a in b in za $a \cdot b$ običajno vsoto a in b. Potem dobimo polkolobar.

Definicija lastnosti kolobarjev

Kolobar $(M, +, \cdot)$ je

- **asociativen**, če je grupoid $(M, \cdot)$ asociativen.
- **komutativen**, če je grupoid $(M, \cdot)$ komutativen.
- **kolobar z enoto**, če ima grupoid $(M, \cdot)$ enoto.

Primeri kolobarjev

- $(\mathbb{R}^3, +, \times)$ je kolobar, ki ni asociativen, ni komutativen in nima enote.
- $(M_n(\mathbb{R}), +, \cdot)$ je asociativen kolobar z enoto, ki ni komutativen.
- $(\mathbb{Z}, +, \cdot)$ je komutativen in asociativen kolobar z enoto.
- $(\mathbb{R}[x], +, \cdot)$ je komutativen in asociativen kolobar z enoto.

Primer: Kolobar funkcij

Naj bo S neprazna množica. Označimo z $\mathbb{R}^S$ množico vseh funkcij iz S v $\mathbb{R}$. Za dve funkciji $f, g \in \mathbb{R}^S$ definiramo funkciji $f + g$ in $f \cdot g$ takole:

$$(f + g)(x) := f(x) + g(x) \quad \text{in} \quad (f \cdot g)(x) := f(x) \cdot g(x)$$

za vsak $x \in S$. $(\mathbb{R}^S, +, \cdot)$ je komutativen in asociativen kolobar z enoto.

Za vajo dokažimo asociativnost seštevanja funkcij. Ostale lastnosti kolobarja se dokaže podobno. Za vse $f, g, h \in \mathbb{R}^S$ in vse $x \in S$ velja

$$(f + (g + h))(x) = f(x) + (g + h)(x) = f(x) + (g(x) + h(x)) =$$
$$= (f(x) + g(x)) + h(x) = (f + g)(x) + h(x) = ((f + g) + h)(x)$$

Primer: Kolobar endomorfizmov Abelove grupe

Naj bo $(G,+)$ Abelova grupa. **Endomorfizem** $(G,+)$ je homomorfizem iz $(G,+)$ v $(G,+)$. Naj bo $\mathrm{End}(G,+)$ množica vseh endomorfizmov $(G,+)$. Vsota in produkt dveh endomorfizmov $\phi,\psi \in \mathrm{End}(G,+)$ definirajmo z

$$(\phi+\psi)(x) := \phi(x)+\psi(x) \quad \text{in} \quad (\phi\cdot\psi)(x) := \phi(\psi(x)).$$

Ker je kompozitum homomorfizmov homomorfizem, je $\phi\cdot\psi \in \mathrm{End}(G,+)$. Pokažimo, da je tudi $\phi+\psi \in \mathrm{End}(G,+)$. Ker je $(G,+)$ Abelova grupa, je

$$(\phi+\psi)(x+y) = \phi(x+y)+\psi(x+y) = \phi(x)+\phi(y)+\psi(x)+\psi(y) =$$
$$= \phi(x)+\psi(x)+\phi(y)+\psi(y) = (\phi+\psi)(x)+(\phi+\psi)(y).$$

Radi bi pokazali, da je $(\mathrm{End}(G,+),+,\cdot)$ asociativen kolobar z enoto.

Očitno je $\mathrm{End}(G,+)$ Abelova grupa za seštevanje endomorfizmov in očitno je množenje endomorfizmov asociativno. Dokažimo sedaj distributivnost.

Velja

$$((\phi + \psi) \cdot \rho)(x) = (\phi + \psi)(\rho(x)) = \phi(\rho(x)) + \psi(\rho(x)) =$$
$$= (\phi \cdot \rho)(x) + (\psi \cdot \rho)(x) = (\phi \cdot \rho + \psi \cdot \rho)(x)$$

in

$$(\rho \cdot (\phi + \psi))(x) = \rho((\phi + \psi)(x)) = \rho(\phi(x) + \psi(x)) =$$
$$= \rho(\phi(x)) + \rho(\psi(x)) = (\rho \cdot \phi)(x) + (\rho \cdot \psi)(x) = (\rho \cdot \phi + \rho \cdot \psi)(x)$$

Enota za množenje je identični endomorfizem.

Primer: Boolov kolobar

Naj bo M_S množica vseh podmnožic dane množice S. Potem je M_S kolobar za operaciji

$$A + B := (A \setminus B) \cup (B \setminus A)$$
$$A \cdot B := A \cap B$$

Podkolobarji

Definicija podkolobarja

Podkolobar kolobarja $(M, +, \cdot)$ je taka podmnožica $N \subseteq M$, da je N podgrupa Abelove grupe $(M, +)$ in podgrupoid grupoida $(M, \cdot)$.

Opomba: Preprosteje definicijo podkolobarja povemo takole. Podmnožica N v M je podkolobar, če za vsaka $x, y \in N$ velja $x - y \in N$ in $x \cdot y \in N$.

Opomba: Podkolobar spremenimo v kolobar tako, da ga opremimo s skrčitvami operacij $+$ in $\cdot$.

Opomba: Podobno definiramo tudi podpolkolobar in podbigrupoid. V tem primeru zahtevamo samo, da je podmnožica zaprta za operaciji $+$ in $\cdot$.

Primeri podkolobarjev in podpolkolobarjev

- $\mathbb{N}$ je podpolkolobar v $(\mathbb{Z}, +, \cdot)$.
- Števila deljiva s 3 so podkolobar v $(\mathbb{Z}, +, \cdot)$.
- $\mathbb{Z}$ je podkolobar v $(\mathbb{Q}, +, \cdot)$.

Primeri podkolobarjev v $M_n(\mathbb{R})$

- Zgornje trikotne $n \times n$ matrike so podkolobar v $(M_n(\mathbb{R}), +, \cdot)$.
- Matrike z ničelno zadnjo vrstico so podkolobar v $(M_n(\mathbb{R}), +, \cdot)$.
- Matrike z elementi iz $\mathbb{Z}$ so podkolobar v $(M_n(\mathbb{R}), +, \cdot)$.
- Matrike oblike $\begin{bmatrix} a & b \\ b & a \end{bmatrix}$, kjer $a, b \in \mathbb{R}$, so podkolobar v $(M_2(\mathbb{R}), +, \cdot)$.

Primeri podkolobarjev v $\mathbb{R}^S$

Množico vseh funkcij iz intervala $[a, b]$ v množico $\mathbb{R}$ označimo z $\mathbb{R}^{[a,b]}$.

- Podmnožica vseh zveznih funkcij je podkolobar v $(\mathbb{R}^{[a,b]}, +, \cdot)$.
- Podmnožica vseh odvedljivih funkcij je podkolobar v $(\mathbb{R}^{[a,b]}, +, \cdot)$.
- Podmnožica vseh omejenih funkcij je podkolobar v $(\mathbb{R}^{[a,b]}, +, \cdot)$.
- Množica vseh funkcij $f : [a, b] \to \mathbb{R}$, ki zadoščajo $f(b) = 0$ je podkolobar v $(\mathbb{R}^{[a,b]}, +, \cdot)$.

Homomorfizmi kolobarjev

Definicija homomorfizma

Homomorfizem kolobarjev iz kolobarja $(M_1, +_1, \cdot_1)$ v kolobar $(M_2, +_2, \cdot_2)$ je taka preslikava $f : M_1 \to M_2$, ki zadošča

$$f(x +_1 y) = f(x) +_2 f(y) \quad \text{in} \quad f(x \cdot_1 y) = f(x) \cdot_2 f(y)$$

za vsaka $x, y \in M_1$. Bijektivnemu homomorfizmu kolobarjev pravimo **izomorfizem kolobarjev**.

Opomba: Enako definiramo tudi homomorfizem/izomorfizem polkolobarjev/bigrupoidov. Pri homomorfizmih kolobarjev z enoto zahtevamo še, da slikajo multiplikativno enoto v multiplikativno enoto.

Opomba: Kompozitum dveh homomorfizmov/izomorfizmov kolobarjev je spet homomorfizem/izomorfizem kolobarjev. Inverz izomorfizma kolobarjev je spet izomorfizem kolobarjev.

Primer homomorfizma kolobarjev, ki ne slika enote v enoto

Preslikava

$$f\colon M_n(\mathbb{R}) \to M_{n+k}(\mathbb{R}), \quad f(A) := \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix}$$

je homomorfizem kolobarjev, ki ne slika enote v enoto.

Primer izomorfizma kolobarjev

Naj bo $B \in M_n(\mathbb{R})$ obrnljiva matrika. Preslikava

$$f\colon M_n(\mathbb{R}) \to M_n(\mathbb{R}), \quad f(A) := BAB^{-1}$$

je izomorfizem kolobarjev z enoto.

Primer

Preslikava

$$f\colon \mathbb{C} \to M_2(\mathbb{R}), \quad f(a+bi) := \begin{bmatrix} a & b \\ -b & a \end{bmatrix}$$

je homomorfizem kolobarjev z enoto.

Primer

Naj bo a neko realno število. Preslikava

$$f_a \colon \mathbb{R}[x] \to \mathbb{R}, \quad f_a(p) := p(a)$$

je homomorfizem kolobarjev z enoto. Pravimo ji **evalvacija** v točki a.

Primer

Množica $\mathbb{Z} \times \mathbb{Z}$ je kolobar za operaciji $(a, b) + (c, d) := (a + c, b + d)$ in $(a, b) \cdot (c, d) := (ac, bd)$. Preslikava

$$f \colon \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}, \quad f((a, b)) = a$$

je homomorfizem kolobarjev z enoto.

Algebrske strukture, 3.del

Obsegi in polja

Obsegi so zelo poseben tip kolobarjev. V kolobarjih lahko elemente seštevamo, odštevamo in množimo, v obsegih pa jih lahko tudi delimo (razen deljenja z nič seveda).

Definicija obsega

Obseg je tak kolobar, v katerem je množica neničelnih elementov grupa za množenje. Komutativnemu obsegu pravimo **polje**.

Opomba: Grupa je asociativen grupoid z enoto, v katerem je vsak element obrnljiv. Definicijo obsega lahko torej povemo tudi takole. **Obseg** je tak asociativen kolobar z enoto, v katerem je vsak neničeln element obrnljiv.

Primeri polj

Kolobarji $(\mathbb{Q}, +, \cdot)$, $(\mathbb{R}, +, \cdot)$ in $(\mathbb{C}, +, \cdot)$ so polja. Za vsako polje F naj bo $F(x)$ množica vseh racionalnih funkcij v spremenljivki x s koeficienti iz F. Ta množica je polje za običajno seštevanje in množenje racionalnih funkcij. Torej so $(\mathbb{Q}(x), +, \cdot)$, $(\mathbb{R}(x), +, \cdot)$ in $(\mathbb{C}(x), +, \cdot)$ polja.

Primer obsega, ki ni polje

Množica vseh matrik oblike

$$\begin{bmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{bmatrix}$$

kjer $\alpha, \beta \in \mathbb{C}$, je obseg za običajno seštevanje in množenje matrik. Pravimo mu **obseg kvaternionov**.

Če vstavimo $\alpha = a + bi$ in $\beta = c + di$, velja

$$\begin{bmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{bmatrix} = a\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + b\begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} + c\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} + d\begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}$$

Ta izraz lahko na kratko zapišemo kot $a\mathbf{1} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$. Matrika $\mathbf{1}$ je identična matrika, za matrike $\mathbf{i}, \mathbf{j}, \mathbf{k}$ pa velja

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -\mathbf{1}, \quad \mathbf{ij} = -\mathbf{ji} = \mathbf{k}, \quad \mathbf{jk} = -\mathbf{kj} = \mathbf{i}, \quad \mathbf{ki} = -\mathbf{ik} = \mathbf{j}.$$

Odtod med drugim sledi, da obseg kvaternionov ni komutativen.

Podobsegi in podpolja

Definicija podobsega in podpolja

Naj bo $(L, +, \cdot)$ obseg. Podmnožica $K \subseteq L$ je njegov **podobseg**, če je K podgrupa v $(L, +)$ in če je $K \setminus \{0\}$ podgrupa v $(L \setminus \{0\}, \cdot)$. Komutativen podobseg je **podpolje**.

Opomba: Na kratko povedano je $K \subseteq L$ podobseg v L, če je zaprta za odštevanje in deljenje.

Primeri podpolj

Očitno je $\mathbb{Q}$ podpolje polja $\mathbb{R}$ in $\mathbb{R}$ je podpolje polja $\mathbb{C}$.

Primer podpolja

Označimo s $\mathbb{Q}(\sqrt{3})$ množico vseh realnih števil oblike $a + b\sqrt{3}$, kjer $a, b \in \mathbb{Q}$. Pokažimo, da je $\mathbb{Q}(\sqrt{3})$ podpolje v $\mathbb{R}$.

Ker je

$$(a + b\sqrt{3}) - (c + d\sqrt{3}) = (a - c) + (b - d)\sqrt{3},$$

je množica $\mathbb{Q}(\sqrt{3})$ zaprta za odštevanje. Ker je

$$\frac{a + b\sqrt{3}}{c + d\sqrt{3}} = \frac{(a + b\sqrt{3})(c - d\sqrt{3})}{(c + d\sqrt{3})(c - d\sqrt{3})} = \frac{ac - 3bd}{c^2 - 3d^2} + \frac{bc - ad}{c^2 - 3d^2}\sqrt{3},$$

je množica $\mathbb{Q}(\sqrt{3}) \setminus \{0\}$ zaprta za deljenje. Pri tem smo upoštevali, da je $c + d\sqrt{3} \neq 0$ in $c^2 - 3d^2 \neq 0$, če $c \neq 0$ ali $d \neq 0$. V nasprotnem primeru bi namreč bilo $\sqrt{3}$ racionalno število.

Primer podpolja

Označimo s $\mathbb{Q}(\sqrt[3]{2})$ množico vseh realnih števil oblike $a + b\sqrt[3]{2} + c\sqrt[3]{4}$, kjer $a, b, c \in \mathbb{Q}$. Pokažimo, da je $\mathbb{Q}(\sqrt[3]{2})$ podpolje v $\mathbb{R}$.

Očitno je $\mathbb{Q}(\sqrt[3]{2})$ zaprta za odštevanje in množenje, torej je podkolobar. Pokazati je treba še, da za vsake $a, b, c \in \mathbb{Q}$, ki niso vsi nič, obstajajo taki $x, y, z \in \mathbb{Q}$, da je $(a + b\sqrt[3]{2} + c\sqrt[3]{4})^{-1} = x + y\sqrt[3]{2} + z\sqrt[3]{4}$. Treba je rešiti sistem $ax + 2cy + 2bz = 1, bx + ay + 2cz = 0, cx + by + az = 0$ z det $\neq 0$.

Homomorfizmi obsegov in polj

Definicija homomorfizma obsegov

Homomorfizem obsegov je tak homomorfizem kolobarjev z enoto, ki slika iz obsega v obseg. Enako definiramo **homomorfizem polj**.

Opomba: Na dolgo povedano je homomorfizem polj iz polja $(K, +_K, \cdot_K)$ v polje $(L, +_L, \cdot_L)$ taka preslikava $f : K \to L$, ki za vsaka $x, y \in K$ zadošča $f(x +_K y) = f(x) +_L f(y)$ in $f(x \cdot_K y) = f(x) \cdot_L f(y)$ in tudi $f(1_K) = 1_L$.

Opomba: Definicijo homomorfizma polj iz polja $(K, +_K, \cdot_K)$ v polje $(L, +_L, \cdot_L)$ lahko povemo tudi takole: To je taka preslikava iz K v L, ki je homomorfizem grup iz $(K, +_K)$ v $(L, +_L)$ in iz $(K \setminus \{0\}, \cdot_K)$ v $(L \setminus \{0\}, \cdot_L)$.

Primer homomorfizma obsegov

Preslikava iz realnih števil v kvaternione, ki je definirana z

$$f(a) := \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix}$$

je homomorfizem obsegov.

Trditev

Vsak homomorfizem obsegov je injektiven.

Dokaz: Recimo, da je $f : K \to L$ homomorfizem obsegov in da je $f(a_1) = f(a_2)$ za neka $a_1, a_2 \in K$. Radi bi pokazali, da je $a_1 = a_2$. Označimo $a := a_1 - a_2$. Potem je $f(a) = f(a_1) - f(a_2) = 0_L$. Če $a_1 \neq a_2$, potem je $a \neq 0$, torej obstaja tak b iz K, da je $ab = 1_K$. Odtod sledi, da je $0_L = 0_L f(b) = f(a)f(b) = f(ab) = f(1_K) = 1_L$, kar je protislovje. Torej je res $a_1 = a_2$.

Opomba: Bijektivnemu homomorfizmu obsegov pravimo **izomorfizem obsegov**. Inverz izomorfizma obsegov je spet izomorfizem obsegov.

Opomba: Če je $f : K \to L$ homomorfizem obsegov, potem je $f(K)$ podobseg v L. Poleg tega je f bijektiven homomorfizem obsegov iz obsega K v obseg $f(K)$. Obsega K in $f(K)$ sta zato izomorfna. Torej lahko smatramo K za podobseg v L.

Kolobarji $\mathbb{Z}_n$ in $F[x]/(p)$

V tem razdelku bomo konstruirali dva tipa komutativnih kolobarjev, v nadaljevanju pa se bomo ukvarjali s tem, kdaj so ti kolobarji polja.

Kolobar $\mathbb{Z}_n$

Vzemimo neko naravno število n in označimo $\mathbb{Z}_n = \{0, 1, \ldots, n-1\}$. Za vsaka $x, y \in \mathbb{Z}_n$ naj bo

$$x \oplus y := (x + y) \bmod n \quad \text{in} \quad x \odot y := (x \cdot y) \bmod n$$

kjer sta $+$ in $\cdot$ operaciji na $\mathbb{Z}$ in je $z \bmod n$ ostanek pri deljenju z z n. Trdimo, da je $(\mathbb{Z}_n, \oplus, \odot)$ komutativen in asociativen kolobar z enoto.

Komutativnost $\oplus$ in $\odot$ sledi direktno iz komutativnosti $+$ in $\cdot$.
Pokažimo asociativnost $\oplus$. Vzemimo $x, y, z \in \mathbb{Z}_n$ in označimo $u = x \oplus y$ in $v = y \oplus z$. Vzemimo take $i, j, k, l \in \mathbb{N}$, da je

$$x + y = in + u \qquad u + z = kn + (u \oplus z)$$
$$y + z = jn + v \qquad x + v = ln + (x \oplus v)$$

Odtod sledi

$$(x \oplus y) \oplus z = u \oplus z = u + z - kn = (x + y - in) + z - kn$$
$$x \oplus (y \oplus z) = x \oplus v = x + v - ln = x + (y + z - jn) - ln$$

torej je $(x \oplus y) \oplus z - x \oplus (y \oplus z) = (j + l - i - k)n$. Ker sta $(x \oplus y) \oplus z$ in $x \oplus (y \oplus z)$ med 0 in $n - 1$ in ker je njuna razlika deljiva z n, sta enaka.

Podobno dokažemo tudi asociativnost $\odot$ in distributivnost. Aditivna enota je 0, multiplikativna enota pa 1. Aditivni inverz elementa $x \neq 0$ je $n - x$.

Opomba: Preslikava

$$f : \mathbb{Z} \to \mathbb{Z}_n, \quad f(z) := z \bmod n$$

je homomorfizem kolobarjev z enoto iz $(\mathbb{Z}, +, \cdot)$ v $(\mathbb{Z}_n, \oplus, \odot)$.

Opomba: Če n ni praštevilo, potem kolobar $(\mathbb{Z}_n, \oplus, \odot)$ ni obseg. Iz razcepa $n = rs$, kjer $r, s < n$, namreč sledi, da je $r \odot s = 0$ in $r \neq 0$ in $s \neq 0$.

Konstrukcijo iz prejšnjega primera lahko razširimo tudi na polinome.

Kolobar $F[x]/(p)$

Naj bo F polje. Označimo s $F[x]$ množico vseh polinomov v spremenljivki x s koeficienti iz F. Običajno seštevanje in množenje polinomov označimo s $+$ in $\cdot$. Potem je $(F[x], +, \cdot)$ komutativen in asociativen kolobar z enoto.

Vzemimo nek nekonstanten polinom $p \in F[x]$ in označimo z $F[x]/(p)$ množico vseh polinomov iz $F[x]$, ki so nižje stopnje kot p. Za vsaka polinoma $r, s \in F[x]/(p)$ definirajmo polinoma

$$r \oplus s := r + s \quad \text{in} \quad r \odot s := (r \cdot s) \bmod p$$

kjer je $q \bmod p$ ostanek pri deljenju polinoma q s polinomom p.

Podobno kot v prejšnjem primeru pokažemo, da je $(F[x]/(p), \oplus, \odot)$ komutativen in asociativen kolobar z enoto.

Opomba: Preslikava

$$\phi\colon F[x] \to F[x]/(p), \quad \phi(q) := q \bmod p$$

je homomorfizem kolobarjev z enoto iz $(F[x], +, \cdot)$ v $(F[x]/(p), \oplus, \odot)$.

Definicija razcepnega in nerazcepnega polinoma

Polinom $p \in F[x]$ je **razcepen**, če obstajata taka polinoma $p_1, p_2 \in F[x]$ stopnje ≥ 1, da je $p = p_1 p_2$. Polinom, ki ni razcepen, je **nerazcepen**.

Opomba: Konstantni in linearni polinomi so nerazcepni.

Primer

Polinom $x^2 - 3$ leži tako v $\mathbb{Q}[x]$ kot v $\mathbb{R}[x]$. V $\mathbb{Q}[x]$ je nerazcepen, ker nima racionalne ničle. V $\mathbb{R}[x]$ je razcepen, ker velja $x^2 - 3 = (x + \sqrt{3})(x - \sqrt{3})$.

Opomba: Preslikavi $\phi\colon \mathbb{Q}[x]/(x^2 - 3) \to \mathbb{Q}(\sqrt{3})$, $\phi(q) = q(\sqrt{3})$ in $\psi\colon \mathbb{R}[x]/(x^2 - 3) \to \mathbb{R} \times \mathbb{R}$, $\psi(q) = (q(\sqrt{3}), q(-\sqrt{3}))$ sta izomorfizma kolobarjev. Ker je $\mathbb{Q}(\sqrt{3})$ polje, je tudi $\mathbb{Q}[x]/(x^2 - 3)$ polje. $\mathbb{R} \times \mathbb{R}$ ni polje.

Trditev

Če je polinom $p \in F[x]$ razcepen, potem kolobar $(F[x]/(p), \oplus, \odot)$ ni obseg.

Dokaz je podoben kot pri $\mathbb{Z}_n$. Če je p razcepen v $F[x]$, potem obstajata taka neničelna polinoma $p_1, p_2 \in F[x]/(p)$, da je $p_1 \odot p_2 = 0$.

Linearna diofantska enačba

V dokazu glavnega izreka bomo potrebovali naslednji tehnični rezultat.

Izrek o linearni diofantski enačbi

Če sta celi števili a_1 in a_2 tuji (= njun največji skupni delitelj je 1), potem obstajata taki celi števili x in y, da velja $a_1x + a_2y = 1$.

Dokaz: Brez škode lahko predpostavimo, da sta a_1 in a_2 naravni števili. Po izreku o deljenju z ostankom obstajajo taka naravna števila $k_1, \ldots, k_n$ in $a_3, \ldots, a_{n+1}$, da velja:

$$a_1 = k_1 a_2 + a_3 \quad \text{kjer } a_3 < a_2 \tag{1}$$

$$a_2 = k_2 a_3 + a_4 \quad \text{kjer } a_4 < a_3 \tag{2}$$

$$\vdots$$

$$a_{n-1} = k_{n-1} a_n + a_{n+1} \quad \text{kjer } a_{n+1} < a_n \tag{n-1}$$

$$a_n = k_n a_{n+1} \tag{n}$$

Postopek smo nadaljevali toliko časa, dokler ni ostanek padel na nič. Ker se ostanek v vsakem koraku zmanjša, je korakov samo končno mnogo.

Pokažimo najprej, da je $a_{n+1} = 1$. Iz enačbe (n) sledi, da a_{n+1} deli a_n Odtod in iz enačbe (n-1) sledi, da a_{n+1} deli a_{n-1}. Odtod in iz enačbe (n-2) sledi, da a_{n+1} deli a_{n-2}. Ta postopek nadaljujemo, dokler ne pridemo do prve enačbe. Torej a_{n+1} deli tako a_1 kot a_2. Ker sta a_1 in a_2 tuji, odtod sledi, da je $a_{n+1} = 1$.

Pokažimo sedaj, da za vsako naravno število $m = 1, \ldots, n$ obstajata taki celi števili x_m in y_m, da velja

$$a_1 x_m + a_2 y_m = a_{m+1}. \qquad (*)$$

Pri $m = 1$ lahko vzamemo kar $x_1 = 0$ in $y_1 = 1$. Iz enačbe (1) sledi $a_3 = a_1 - k_2 a_2$, torej lahko vzamemo $x_2 = 1$ in $y_2 = -k_2$. Izpeljimo sedaj še rekurzivni zvezi za x_m in y_m. Iz enačbe (m-1) sledi, da je $a_{m+1} = a_{m-1} - k_{m-1} a_m$, kar je po indukcijski predpostavki enako $(a_1 x_{m-2} + a_2 y_{m-2}) - k_{m-1}(a_1 x_{m-1} + a_2 y_{m-1})$. Če to primerjamo z želeno relacijo $(*)$, dobimo $x_m = x_{m-2} - k_{m-1} x_{m-1}$ in $y_m = y_{m-2} - k_{m-1} y_{m-1}$. Ko je m enak n, dobimo ravno izrek: $a_1 x_n + a_2 y_n = a_{n+1} = 1$. □

Naš glavni rezultat je:

Izrek o $\mathbb{Z}_p$

Če je p praštevilo, potem je kolobar $\mathbb{Z}_p$ polje.

Dokaz. Vemo že, da je $\mathbb{Z}_p$ komutativen in asociativen kolobar z enoto. Pokazati moramo še, da ima vsak neničelni element multiplikativen inverz. Vzemimo $a_2 = p$ in naj bo a_1 poljuben neničeln element v $\mathbb{Z}_p$. Vidimo, da sta a_1 in a_2 tuji števili. Po izreku o linearni diofantski enačbi obstajata taki celi števili x in y, da je $a_1x + a_2y = 1$. Odtod sledi, da je $a_1^{-1} = x \bmod p$.

Podobno dokažemo tudi naslednji rezultat:

Izrek o $F[x]/(p)$

Če je $p \in F[x]$ nerazcepen polinom stopnje ≥ 1, potem je $F[x]/(p)$ polje.

Dokaz: Dva polinoma sta tuja, če nimata skupnega faktorja stopnje ≥ 1. Če vzamemo $p_2 = p$ iz izreka in $p_1 \neq 0$ polinom, ki je nižje stopnje kot p, potem sta p_1 in p_2 tuja. Za vsaka tuja polinoma p_1 in p_2 konstruramo kot zgoraj taka polinoma q_1 in q_2, da je $p_1q_1 + p_2q_2 = 1$ in dobimo izrek.

Primer

Poiščimo inverz elementa 12 v polju $\mathbb{Z}_{41}$.

Iščemo tak $x \in \mathbb{Z}$, da je $12x \bmod 41 = 1$. To velja natanko tedaj, ko obstaja tak $y \in \mathbb{Z}$, da velja $12x + 41y = 1$. Evklidov algoritem nam da

$$\begin{aligned} 41 = 3 \cdot 12 + 5 \quad &\Rightarrow \quad 5 = 41 - 3 \cdot 12 \\ 12 = 2 \cdot 5 + 2 \quad &\Rightarrow \quad 2 = 12 - 2 \cdot 5 \\ 5 = 2 \cdot 2 + 1 \quad &\Rightarrow \quad 1 = 5 - 2 \cdot 2 \end{aligned}$$

Ko vstavimo prvo enačbo v drugo, dobimo

$$2 = 12 - 2 \cdot (41 - 3 \cdot 12) = -2 \cdot 41 + 7 \cdot 12$$

Ko to in prvo enačbo vstavimo v tretjo enačbo, dobimo

$$1 = (41 - 3 \cdot 12) - 2 \cdot (-2 \cdot 41 + 7 \cdot 12) = 5 \cdot 41 + (-17) \cdot 12$$

Torej je $x = -17$, kar pa ni v $\mathbb{Z}_{41}$. Sledi $12^{-1} = x \bmod 41 = 24$.

Primer

Izračunajmo inverz polinoma $x^3 - 2x + 2$ v polju $\mathbb{Q}[x]/(x^4 + 1)$.

Najprej uporabimo Evklidov algoritem

$$x^4 + 1 = x(x^3 - 2x + 2) + 2x^2 - 2x + 1$$

$$x^3 - 2x + 2 = \frac{x+1}{2}(2x^2 - 2x + 1) + \frac{3 - 3x}{2}$$

$$2x^2 - 2x + 1 = -\frac{4x}{3}\left(\frac{3 - 3x}{2}\right) + 1$$

Iz prve enačbe dobimo

$$2x^2 - 2x + 1 = (x^4 + 1) - x(x^3 - 2x + 2)$$

Iz druge enačbe potem dobimo

$$\begin{aligned}\frac{3 - 3x}{2} &= (x^3 - 2x + 2) - \frac{x+1}{2}(2x^2 - 2x + 1) \\ &= (x^3 - 2x + 2) - \frac{x+1}{2}((x^4 + 1) - x(x^3 - 2x + 2)) \\ &= -\frac{x+1}{2}(x^4 + 1) + \frac{x^2 + x + 2}{2}(x^3 - 2x + 2)\end{aligned}$$

Upoštevajmo sedaj oba izraza v tretji enačbi. Dobimo

$$\begin{aligned}
1 &= (2x^2 - 2x + 1) + \frac{4x}{3}\left(\frac{3-3x}{2}\right) \\
&= ((x^4+1) - x(x^3 - 2x + 2)) \\
&\quad + \frac{4x}{3}\left(-\frac{x+1}{2}(x^4+1) + \frac{x^2+x+2}{2}(x^3-2x+2)\right) \\
&= \frac{3-2x(x+1)}{3}(x^4+1) + \frac{-3x+2x(x^2+x+2)}{3}(x^3-2x+2)
\end{aligned}$$

Inverz polinoma $x^3 - 2x + 2$ v $\mathbb{Q}[x]/(x^4+1)$ je torej

$$\frac{-3x+2x(x^2+x+2)}{3} = \frac{2x^3+2x^2+x}{3}.$$

Produkt polinoma in njegovega inverza je res enak 1 v $\mathbb{Q}[x]/(x^4+1)$, ker

$$\frac{2x^3+2x^2+x}{3}(x^3-2x+2) = \frac{2x^2+2x-3}{3}(x^4+1) + 1.$$

Polja s p^n elementi

Radi bi opisali vsa končna polja. Ideja konstrukcije je naslednja:

- Vzemi praštevilo p in naravno število n. Vemo, da je $\mathbb{Z}_p$ polje.
- Dokaži, da v $\mathbb{Z}_p[x]$ obstaja nerazcepen polinom $q(x)$ stopnje n.
- Dokaži, da je $\mathbb{Z}_p[x]/(q(x))$ polje s p^n elementi.

Izrek o klasifikaciji končnih obsegov pravi:

- Vsako končno polje je izomorfno enemu od zgornjih polj.
- Dve končni polji z enakim številom elementov sta izomorfni.

Podrobnosti bomo izpustili. Raje si oglejmo primer.

Polje s štirimi elementi

Iščemo polje, ki ima štiri elemente. Kolobar $\mathbb{Z}_4$ sicer ima štiri elemente, ampak ni polje. Iskano polje je $\mathbb{Z}_2[x]/(x^2+x+1)$.

Množica $\mathbb{Z}_2[x]/(x^2+x+1)$ se sestoji iz vseh polinomov v $\mathbb{Z}_2[x]$, ki so nižje stopnje kot x^2+x+1. To so polinomi $0, 1, x, x+1$.

Operaciji na množici $\mathbb{Z}_2[x]/(x^2+x+1)$ sta seštevanje in množenje modulo x^2+x+1. Njuni tabeli sta:

$\oplus$	0	1	x	$x+1$
0	0	1	x	$x+1$
1	1	0	$x+1$	x
x	x	$x+1$	0	1
$x+1$	$x+1$	x	1	0

$\odot$	0	1	x	$x+1$
0	0	0	0	0
1	0	1	x	$x+1$
x	0	x	$x+1$	1
$x+1$	0	$x+1$	1	x

Pokazali smo že, da je $(\mathbb{Z}_2[x]/(x^2+x+1), \oplus, \odot)$ komutativen in asociativen kolobar z enoto. Iz tabele za $\odot$ se vidi, da ima vsak neničeln element inverz. Torej je ta kolobar polje.

Opomba: Tudi brez tabele za $\odot$ lahko dokažemo, da je ta kolobar polje. Zadošča dokazati, da je $x^2 + x + 1$ nerazcepen polinom v $\mathbb{Z}_2[x]$.

V $\mathbb{Z}_2[x]$ imamo dva polinoma stopnje 1 in štiri polinome stopnje 2. Polinoma stopnje 1 sta x in $x + 1$. Razcepni polinomi stopnje 2 so torej x^2, $x(x + 1) = x^2 + x$ in $(x + 1)^2 = x^2 + 1$. Polinom $x^2 + x + 1$ ni eden od teh, torej je nerazcepen.

Vektorski prostori 1.del

Definicija vektorskega prostora

Definicija vektorskega prostora

Vektorski prostor nad poljem $(F, \oplus, \odot)$ je podan s tremi podatki:

(1) neprazno množico V,

(2) tako operacijo $+$ na V, da je $(V, +)$ Abelova grupa,

(3) tako preslikavo $\cdot\colon F \times V \to V$, da velja:

- (i) $\alpha \cdot (u + v) = (\alpha \cdot u) + (\alpha \cdot v)$ za vse $\alpha \in F$ in $u, v \in V$,
- (ii) $(\alpha \oplus \beta) \cdot v = (\alpha \cdot v) + (\beta \cdot v)$ za vse $\alpha, \beta \in F$ in $v \in V$,
- (iii) $(\alpha \odot \beta) \cdot v = \alpha \cdot (\beta \cdot v)$ za vse $\alpha, \beta \in F$ in $v \in V$,
- (iv) $1 \cdot v = v$ za vse $v \in V$.

Opomba: Elementom množice V pravimo **vektorji**. Elementom množice F pravimo **skalarji**. Operaciji $+$ pravimo **vsota vektorjev**. Funkciji $\cdot$ pravimo **produkt vektorja s skalarjem**.

Opomba: Namesto $\alpha \oplus \beta$ bomo pisali kar $\alpha + \beta$. Namesto $\alpha \odot \beta$ bomo pisali kar $\alpha\beta$. Namesto $\alpha \cdot v$ bomo pisali kar αv. Produkt vektorja s skalarjem naj ima višjo prioriteto kot vsota vektorjev (manj oklepajev).

Opomba: Če v definiciji vektorskega prostora privzamemo samo, da je F kolobar, potem dobimo definicijo **modula**.

Vektorski prostor nad poljem F lahko definiramo tudi kot Abelovo grupo $(V, +)$ skupaj z homomorfizmom kolobarjev z enoto iz F v $\mathrm{End}(V, +)$, kjer je $\mathrm{End}(V, +)$ kolobar endomorfizmov Abelove grupe $(V, +)$.

Pokažimo zdaj, da sta obe definiciji ekvivalentni. Za vsak $\alpha \in F$ naj bo

$$\phi_\alpha \colon V \to V, \quad \phi_\alpha(v) := \alpha v.$$

Po lastnosti (i) je

$$\phi_\alpha(u + v) = \alpha(u + v) = \alpha u + \alpha v = \phi_\alpha(u) + \phi_\alpha(v),$$

torej je ϕ_α endomorfizem Abelove grupe $(V, +)$. Oglejmo si preslikavo

$$\Phi \colon F \to \mathrm{End}(V, +), \quad \Phi(\alpha) := \phi_\alpha.$$

Pokažimo, da je Φ homomorfizem kolobarjev z enoto.

Za vsaka $\alpha, \beta \in F$ in za vsak $v \in V$ po lastnostih (ii)-(iv) velja

$$\Phi(\alpha+\beta)(v) = (\alpha+\beta)v = \alpha v+\beta v = \Phi(\alpha)(v)+\Phi(\beta)(v) = (\Phi(\alpha)+\Phi(\beta))(v)$$
$$\Phi(\alpha\beta)(v) = (\alpha\beta)v = \alpha(\beta v) = \Phi(\alpha)(\Phi(\beta)(v)) = (\Phi(\alpha) \circ \Phi(\beta))(v)$$
$$\Phi(1)(v) = 1v = v = \mathrm{id}(v)$$

torej je $\Phi(\alpha + \beta) = \Phi(\alpha) + \Phi(\beta)$, $\Phi(\alpha\beta) = \Phi(\alpha) \circ \Phi(\beta)$ in $\Phi(1) = \mathrm{id}$.

Pokažimo še obratno. Naj bo $\Phi\colon F \to \mathrm{End}(V, +)$ homomorfizem kolobarjev z enoto. Za vsaka $\alpha \in F$ in $v \in V$ definirajmo

$$\alpha v := \Phi(\alpha)(v)$$

Ker je $\Phi(\alpha)$ endomorfizem Abelove grupe $(V, +)$, velja

$$\alpha(u + v) = \Phi(\alpha)(u + v) = \Phi(\alpha)(u) + \Phi(\alpha)(v) = \alpha u + \alpha v.$$

Ker je Φ homomorfizem kolobarjev z enoto, velja

$$(\alpha+\beta)v = \Phi(\alpha+\beta)(v) = (\Phi(\alpha)+\Phi(\beta))(v) = \Phi(\alpha)(v)+\Phi(\beta)(v) = \alpha v+\beta v$$
$$(\alpha\beta)v = \Phi(\alpha\beta)(v) = (\Phi(\alpha) \circ \Phi(\beta))(v) = \Phi(\alpha)(\Phi(\beta)(v)) = \alpha(\beta v)$$
$$1v = \Phi(1)(v) = \mathrm{id}(v) = v$$

Primeri vektorskih prostorov

Oglejmo si najprej naš standardni primer.

Primer: Vektorski prostor F^n

Naj bo F polje in n naravno število. Označimo z F^n množico vseh n-teric elementov iz F. Vsota dveh vektorjev in produkt vektorja s skalarjem naj bosta definirana po komponentah.

$$(\alpha_1, , \ldots, \alpha_n) + (\beta_1, \ldots, \beta_n) := (\alpha_1 + \beta_1, \ldots, \alpha_n + \beta_n)$$
$$\gamma(\alpha_1, \ldots, \alpha_n) := (\gamma\alpha_1, \ldots, \gamma\alpha_n)$$

Potem je F^n vektorski prostor nad F.

Opomba: Pri $n = 1$ dobimo, da je F vektorski prostor nad F.

Primer: Vektorski prostor $M_{m,n}(F)$

Naj bo F polje in naj bosta m in n naravni števili. Označimo z $M_{m,n}(F)$ množico vseh $m \times n$ matrik z elementi iz F. Definirajmo vsoto dveh matrik in produkt matrike s skalarjem na običajen način, se pravi po komponentah. Potem je $M_{m,n}(F)$ vektorski prostor nad F.

Oglejmo si še posplošitev obeh prejšnjih primerov.

Primer: Vektorski prostor F^S

Naj bo F polje in naj bo S neprazna množica. Množico vseh funkcij iz S v F označimo z F^S. Vsota funkcij je definirana po elementih:

$$(f+g)(s) := f(s) + g(s)$$

Tudi produkt funkcije s skalarjem je definiran po elementih

$$(\gamma f)(s) := \gamma f(s)$$

Potem je F^S vektorski prostor.

Opomba: Če je $S = \{1, \ldots, n\}$ potem lahko F^S identificiramo s F^n. Če je $S = \{1, \ldots, m\} \times \{1, \ldots, n\}$ potem lahko F^S identificiramo s $M_{m,n}(F)$.

Primer: Trivialni vektorski prostor nad F

Množica $V = \{0\}$ je vektorski prostor nad poljem F. Vsota vektorjev je definirana z $0 + 0 = 0$. Produkt vektorja s skalarjem je definiran z $\alpha 0 = 0$.

Primer: Vektorski prostor polinomov

Naj bo F polje in naj bo $F[x]$ množica vseh polinomov v spremenljivki x s koeficienti iz F. Vsoto dveh polinomov in produkt polinoma s skalarjem definiramo na običajen način. Potem je $F[x]$ vektorski prostor nad F.

Primer: Razširitev polj

Če je F podpolje polja K, potem je K vektorski prostor nad F za običajno vsoto na K in običajen produkt na K. (Funkcija iz $F \times K$ v K je skrčitev operacije množenja, ki je funkcija iz $K \times K$ v K.)

Opomba: Ker je polje F podpolje polja racionalnih funkcij $F(x)$, je $F(x)$ vektorski prostor nad F.

Primer: Direktna vsota vektorskih prostorov

Naj bo F polje in naj bodo $V_1, \ldots, V_k$ vektorski prostori nad F. Označimo z $V_1 \oplus \ldots \oplus V_k$ množico vseh k-teric $(v_1, \ldots, v_k)$, kjer $v_i \in V_i$ za vse i. Vsota dveh k-teric in produkt k-terice s skalarjem sta definirana po komponentah. Potem je $V_1 \oplus \ldots \oplus V_k$ vektorski prostor nad F.

Vektorski podprostori

Definicija vektorskega podprostora

Naj bo V vektorski prostor nad poljem F. Neprazna podmnožica $U \subseteq V$ je **vektorski podprostor** v V, če velja:

- Za vsaka $u_1, u_2 \in U$ velja $u_1 + u_2 \in U$.
- Za vsak $u \in U$ in vsak $\alpha \in F$ velja $\alpha u \in U$.

Opomba: Vsak vektorski podprostor vsakega vektorskega prostora nad F je spet vektorski prostor nad F. (Vsoto dveh vektorjev v U izračunamo kot v V. Tudi produkt vektorja iz U s skalarjem izračunamo kot v V.)

Trditev

Vsak vektorski podprostor vsebuje ničelni vektor.

Dokaz: Naj bo V vektorski podprostor nad F. Enoto Abelove grupe V označimo z 0_V. Naj bo U vektorski podprostor v V in naj bo u poljuben element U. Po drugi lastnosti iz definicije vektorskega podprostora velja $0u \in U$. Iz $0_V + 0u = 0u = (0+0)u = 0u + 0u$ sledi $0_V = 0u \in U$.

Opomba: Če je W vektorski podprostor v V in če je U vektorski podprostor v W, potem je U vektorski podprostor v V.

Naslednja trditev nam da alternativno definicijo vektorskega podprostora.

Trditev

Naj bo V vektorski prostor nad poljem F. Neprazna podmnožica $U \subseteq V$ je **vektorski podprostor** v V natanko tedaj, ko za vsaka $u_1, u_2 \in U$ in vsaka $\alpha_1, \alpha_2 \in F$ velja $\alpha_1 u_1 + \alpha_2 u_2 \in U$.

Dokaz: Najprej dokažimo da ima vsak vektorski podprostor $U \subseteq V$ lastnost iz trditve. Vzemimo $u_1, u_2 \in U$ in $\alpha_1, \alpha_2 \in F$. Ker je U zaprta za množenje s skalarji, velja $\alpha_1 u_1 \in U$ in $\alpha_2 u_2 \in U$. Ker je U zaprta za seštevanje, odtod sledi $\alpha_1 u_1 + \alpha_2 u_2 \in U$.

Dokažimo še, da je vsaka podmnožica $U \subseteq V$, ki zadošča lastnosti iz trditve, vektorski podprostor v V. Za vsaka $u_1, u_2 \in U$ je po lastnosti iz trditve $1u_1 + 1u_2 \in U$. Torej je U zaprta za seštevanje. Za vsak $u \in U$ in vsak $\alpha \in F$ je po lastnosti iz trditve $\alpha u + 0u \in U$. Torej je U zaprta za množenje s skalarjem.

Trditev

Naj bodo $U_1, \ldots, U_k$ podprostori vektorskega prostora V. Potem velja:

- Njihov presek $U_1 \cap \ldots \cap U_k$ je vektorski podprostor v V.
- Njihova vsota $U_1 + \ldots + U_k$ je vektorski podprostor v V.
  (To je množica vseh vektorjev $u_1 + \ldots + u_k$, kjer $u_i \in U_i$ za vsak i.)

Dokaz za presek: Vzemimo $w, z \in U_1 \cap \ldots \cap U_k$ in $\alpha, \beta \in F$. Radi bi dokazali, da $\alpha w + \beta z \in U_1 \cap \ldots \cap U_k$. Za vsak $i = 1, \ldots, k$ je U_i tak podprostor v V, ki vsebuje tako w kot z, torej vsebuje tudi $\alpha w + \beta z$. Ker $\alpha w + \beta z$ leži v vseh U_i, leži tudi v preseku.

Dokaz za vsoto: Vzemimo $w, z \in U_1 + \ldots + U_k$ in $\alpha, \beta \in F$. Radi bi dokazali, da $\alpha w + \beta z \in U_1 + \ldots + U_k$. Po definiciji vsote podprostorov lahko w in z izrazimo kot $w = w_1 + \ldots + w_k$ in $z = z_1 + \ldots + z_k$, kjer $w_i, z_i \in U_i$ za vsak $i = 1, \ldots, k$. Ker so vsi U_i podprostori v V velja $\alpha w_i + \beta z_i \in U_i$ za vsak $i = 1, \ldots, k$. Odtod sledi

$$\alpha w + \beta z = (\alpha w_1 + \beta z_1) + \ldots + (\alpha w_k + \beta z_k) \in U_1 + \ldots + U_k.$$

Primeri vektorskih podprostorov

Primer: Trivialni in nepravi vektorski podprostor

Če je V vektorski prostor nad F, potem sta $\{0\}$ in V vedno vektorska podprostora v V. Prvemu pravimo **trivialni vektorski podprostor**, drugemu pa **nepravi vektorski podprostor**.

Če je $V = F$ sta to edina vektorska podprostora v V. Vsak netrivialen vektorski podprostor U v F namreč vsebuje neničeln element $\alpha \in F$. Potem za vsak $\beta \in F$ velja $\beta = (\beta\alpha^{-1})\alpha \in U$. Torej je $U = F$.

Primer: Vektorski podprostori v F^2

Naj bo F polje in $u \in F^2$. Množica $U := \{\alpha u \mid \alpha \in F\}$ je očitno zaprta za seštevanje in množenje s skalarjem, torej je vektorski podprostor v F^2.

Pokažimo, da je U pravi vektorski podprostor. Če $U = F^2$, bi obstajala taka $\alpha_1, \alpha_2 \in F$, da bi veljalo $(1,0) = \alpha_1 u$ in $(0,1) = \alpha_2 u$. Odtod bi sledilo, da je $\alpha_2(1,0) = \alpha_1(0,1)$, se pravi $\alpha_1 = \alpha_2 = 0$, kar da protislovje.

Naloga: Pokažite, da smo dobili vse prave vektorske podprostore v F^2.

Primer: Množica rešitev homogenega sistema linearnih enačb

Naj bo A $m \times n$ matrika z elementi iz F. Množico vseh n-teric $\mathbf{x} \in F^n$, ki rešijo homogeni sistem linearnih enačb $A\mathbf{x} = 0$ označimo z $\operatorname{Ker} A$ in ji pravimo **jedro** matrike A. Vemo, da je $\operatorname{Ker} A$ vektorski podprostor v F^n.

Primer: Matrični prostori

Naj bo $V = M_n(F)$ vektorski prostor $n \times n$ matrik. Naslednje podmnožice v V so vektorski podprostori v V.

- Vse simetrične matrike. (Matrika A je **simetrična**, če je $A^T = A$.)
- Vse antisimetrične matrike. (A je **antisimetrična**, če $A^T = -A$.)
- Vse matrike X, ki zadoščajo enačbi $A_1XB_1 + \ldots + A_kXB_k = 0$.
- Vse matrike, ki imajo sled enako nič.
- Vse matrike, ki imajo ničle na predpisanih mestih $(i_1, j_1), \ldots, (i_k, j_k)$. (Npr. vse diagonalne matrike, vse zgornje trikotne matrike, vse matrike z ničelnim prvim stolpcem, vse matrike z $a_{1,n} = 0$, itd.)

Primer: Prostori polinomov in racionalnih funkcij

Naj bo F polje. Množica vseh polinomov $F[x]$ je vektorski podprostor v vektorskem prostoru $F(x)$ vseh racionalnih funkcij. Množica $F[x]_{\leq n}$ vseh polinomov stopnje $\leq n$ je vektorski podprostor v vektorskem prostoru $F[x]$.

Primer: Funkcijski prostori

Naj bo $V = \mathbb{R}^{[a,b]}$ vektorski prostor vseh funkcij iz intervala $[a, b]$ v $\mathbb{R}$. Naslednje podmnožice v V so vektorski podprostori v V:

- Vse omejene funkcije.
- Vse integrabilne funkcije.
- Vse zvezne funkcije.
- Vse odvedljive funkcije.
- Vse dvakrat odvedljive funkcije $y(x)$, ki rešijo dano homogeno linearno diferencialno enačbo 2. reda $p(x)y''(x) + q(x)y'(x) + r(x)y(x) = 0$.
- Vse funkcije, ki zadoščajo $f(a) = 0$.
- Vse funkcije, ki zadoščajo $f(a) = f(b)$.

Linearna ogrinjača

Definicija linearne ogrinjače

Naj bo V vektorski podprostor nad F in naj bodo $v_1, \ldots, v_k$ elementi V. Pravimo, da je vektor $v \in V$ **linearna kombinacija** vektorjev $v_1, \ldots, v_k$, če obstajajo taki skalarji $\alpha_1, \ldots, \alpha_k \in F$, da velja $v = \alpha_1 v_1 + \ldots + \alpha_k v_k$.

Množico vseh linearnih kombinacij vektorjev $v_1, \ldots, v_k$ označimo z $\mathrm{Lin}\{v_1, \ldots, v_k\}$ in ji pravimo **linearna ogrinjača** množice $\{v_1, \ldots, v_k\}$.

S formulo: $\mathrm{Lin}\{v_1, \ldots, v_k\} := \{\alpha_1 v_1 + \ldots + \alpha_k v_k \mid \alpha_1, \ldots, \alpha_k \in F\}$.

Trditev

Za vsako končno podmnožico S v vektorskem prostoru V je njena linearna ogrinjača $\mathrm{Lin}\, S$ vektorski podprostor v V.

Dokaz: Če $u_1, u_2 \in \mathrm{Lin}\{v_1, \ldots, v_k\}$ in $\gamma_1, \gamma_2 \in F$, potem bi radi dokazali, da $\gamma_1 u_1 + \gamma_2 u_2 \in \mathrm{Lin}\{v_1, \ldots, v_k\}$. Izberimo take skalarje $\alpha_1, \ldots, \alpha_n$ in $\beta_1, \ldots, \beta_n$, da je $u_1 = \alpha_1 v_1 + \ldots + \alpha_k v_k$ in $u_2 = \beta_1 v_1 + \ldots + \beta_k v_k$. Sledi $\gamma_1 u_1 + \gamma_2 u_2 = (\gamma_1 \alpha_1 + \gamma_2 \beta_1) v_1 + \ldots + (\gamma_1 \alpha_k + \gamma_2 \beta_k) v_k \in \mathrm{Lin}\{v_1, \ldots, v_k\}$.

Vektorski prostori 2.del

Definicija baze

Naj bo V vektorski prostor. Vektor $v \in V$ se da izraziti kot **linearna kombinacija** vektorjev $v_1, \ldots, v_n \in V$, če obstajajo taki skalarji $\alpha_1, \ldots, \alpha_n$, da velja $v = \alpha_1 v_1 + \ldots + \alpha_n v_n$. Pogosto lahko te skalarje izberemo na več različnih načinov.

Definicija baze

Naj bo V vektorski prostor.

- Vektorji $v_1, \ldots, v_n \in V$ so **ogrodje**, če se da vsak vektor $v \in V$ na **vsaj en** način izraziti kot linearna kombinacija $v_1, \ldots, v_n$.
- Vektorji $v_1, \ldots, v_n \in V$ so **linearno neodvisni**, če se da vsak vektor $v \in V$ na **največ en** način izraziti kot linearna kombinacija $v_1, \ldots, v_n$.
- Vektorji $v_1, \ldots, v_n \in V$ so **baza**, če se da vsak vektor $v \in V$ na **natanko en** način izraziti kot linearna kombinacija $v_1, \ldots, v_n$.

Opomba: Očitno so vektorji $v_1, \ldots, v_n \in V$ baza natanko tedaj, ko so linearno neodvisni in ko so ogrodje.

Linearna ogrinjača vektorjev $v_1, \ldots, v_n \in V$ je definirana kot množica vseh vektorjev $v \in V$, ki se dajo izraziti kot linearna kombinacija vektorjev $v_1, \ldots, v_n$. Označimo jo z $\mathrm{Lin}\{v_1, \ldots, v_n\}$. Linearna ogrinjača prazne množice $\{\}$ naj bo $\{0\}$. Linearna ogrinjača je vektorski podprostor v V.

Opomba: Vektorji $v_1, \ldots, v_n \in V$ so **ogrodje** natanko tedaj, ko je njihova linearna ogrinjača enaka V.

Izrek 1 - Karakterizacije linearne neodvisnosti

Za vektorje $v_1, \ldots, v_n$ iz vektorskega prostora V so ekvivalentne trditve:

(1) Vektorji $v_1, \ldots, v_n$ so linearno neodvisni.

(2) Za vsake skalarje $\alpha_1, \ldots, \alpha_n$, ki zadoščajo $\alpha_1 v_1 + \ldots + \alpha_n v_n = 0$, velja $\alpha_1 = \ldots = \alpha_n = 0$.

(3) Nobeden od vektorjev $v_1, \ldots, v_n$ se ne da izraziti kot linearna kombinacija preostalih vektorjev.

Opomba: V nadaljevanju bomo uporabljali točko (2) kot (alternativno) definicijo linearne neodvisnosti vektorjev.

Dokaz: Dokazali bomo, da je (1) $\Leftrightarrow$ (2) in da je (2) $\Leftrightarrow$ (3).

(1) $\Rightarrow$ (2) Recimo, da so $v_1, \ldots, v_n$ linearno neodvisni. Potem se da vektor 0 na kvečjemu en način izraziti kot linearna kombinacija $v_1, \ldots, v_n$. Ena izražava je $0 = 0v_1 + \ldots + 0v_n$. Če imamo še eno izražavo, recimo $0 = \alpha_1 v_1 + \ldots + \alpha_n v_n$, potem je $\alpha_1 = \ldots = \alpha_n = 0$ zaradi enoličnosti.

(2) $\Rightarrow$ (1) Recimo, da iz $\alpha_1 v_1 + \ldots + \alpha_n v_n = 0$ sledi $\alpha_1 = \ldots = \alpha_n = 0$. Vzemimo nek vektor v, ki se da na dva načina izraziti kot linearna kombinacija vektorjev $v_1, \ldots, v_n$. Prvi način naj bo $v = \beta_1 v_1 + \ldots + \beta_n v_n$, drugi način pa $v = \gamma_1 v_1 + \ldots + \gamma_n v_n$. Ko oba načina odštejemo, dobimo $0 = v - v = (\beta_1 - \gamma_1)v_1 + \ldots + (\beta_n - \gamma_n)v_n$. Po predpostavki odtod sledi $\beta_i - \gamma_i = 0$ za vse i, se pravi $\beta_i = \gamma_i$ za vse i. Torej sta oba načina enaka.

(2) $\Rightarrow$ (3) Recimo, da velja (2) in da ne velja (3). Potem lahko pri nekem i izrazimo v_i kot linearno kombinacijo preostalih vektorjev, se pravi kot $v_i = \alpha_1 v_1 + \ldots + \alpha_{i-1} v_{i-1} + \alpha_{i+1} v_{i+1} + \ldots + \alpha_n v_n$. Ta izraz preoblikujemo v $\alpha_1 v_1 + \ldots + \alpha_{i-1} v_{i-1} + (-1)v_i + \alpha_{i+1} v_{i+1} + \ldots + \alpha_n v_n = 0$. Ker velja (2), odtod sledi $-1 = 0$ (in $\alpha_i = 0$ za vse i), kar je protislovje.

(3) $\Rightarrow$ (2) Recimo, da velja (3) in da ne velja (2). Torej obstajajo taki skalarji $\alpha_1, \ldots, \alpha_n$, ki niso vsi enaki nič, da je $\alpha_1 v_1 + \ldots + \alpha_n v_n = 0$. Vzemimo tak i, da je $\alpha_i \neq 0$. Potem lahko odtod izrazimo vektor v_i kot linearno kombinacijo ostalih vektorjev, kar je protislovje.

Primeri baz

V nadaljevanju bomo potrebovali, da se vse karakterizacije obrnljivih matrik nad poljem $\mathbb{R}$ posplošijo na vsako polje F.

Primer - Baze v F^n

Stolpčni vektorji $v_1, \ldots, v_n \in F^n$ so baza natanko tedaj, ko je

$$\det \begin{bmatrix} v_1 & \ldots & v_n \end{bmatrix} \neq 0.$$

To sledi iz karakterizacij obrnljivih matrik. Posebna primera sta:

- Standardna baza $e_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$, $e_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}$, $\ldots$, $e_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$.
- Vektorji $\begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$, $\begin{bmatrix} 1 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$, $\ldots$, $\begin{bmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$ so baza za F^n.

Opomba: Standardna baza v $M_{m,n}(F)$ so koordinatne matrike $E_{i,j}$. Matrika $E_{i,j}$ ima na (i,j)-tem mestu enko, drugod pa same ničle.

Primer - Lagrangeov interpolacijski polinom

Naj bo $n \geq 2$ naravno število, F polje z vsaj n elementi in $x_1, \ldots, x_n$ paroma različni elementi polja F. Naj bo $V = F[x]_{n-1}$ vektorski prostor vseh polinomov stopnje $\leq n-1$ s koeficienti iz F. Trdimo, da so polinomi

$$p_i(x) = \frac{(x-x_1)\cdots(x-x_n)}{x-x_i} = \prod_{j\in\mathbb{N}_n, j\neq i} (x-x_j), \quad \text{za } i = 1, \ldots, n$$

baza vektorskega prostora V. Začnimo z linearno neodvisnostjo. Če je

$$\sum_{i=1}^{n} \alpha_i p_i(x) = 0,$$

potem za vsak j vstavimo x_j. Ker je $p_i(x_j) = 0$ za vsak $i \neq j$, ostane $\alpha_j p_j(x_j) = 0$. Ker so x_k paroma različni, je $p_j(x_j) \neq 0$. Sledi $\alpha_j = 0$.

Dokažimo, da za vsak polinom $f(x) \in V$ velja formula

$$f(x) = \frac{f(x_1)}{p_1(x_1)}\, p_1(x) + \ldots + \frac{f(x_n)}{p_n(x_n)}\, p_n(x)$$

Odtod sledi, da so polinomi $p_i(x)$ ogrodje. Obe strani formule imata namreč stopnjo $\leq n-1$ in se ujemata v n različnih točkah $x_1, \ldots, x_n$.

Obstoj baze

Spomnimo se, da je ogrodje vektorskega prostora V taka podmnožica v V, da je njena linearna ogrinjača enaka V.

Definicija

Vektorski prostor je **končno-razsežen**, če ima končno ogrodje.

Spomnimo se, da je baza vektorskega prostora V taka podmnožica v V, ki je linearno neodvisna in je ogrodje za V.
Primer: Prazna množica je baza vektorskega prostora $\{0\}$.

Izrek 2 - Obstoj baze

Vsak končno-razsežen vektorski prostor ima končno bazo.

Dokaz: Naj bo V končno-razsežen vektorski prostor in naj bo $S = \{v_1, \ldots, v_n\}$ ogrodje z najmanj elementi. Trdimo, da je množica S linearno neodvisna. Dokazujemo s protislovjem.

Recimo, da S ni linearno neodvisna. Potem obstaja tak element $v_i \in S$, ki je linearna kombinacija elementov iz $S \setminus \{v_i\}$. Pokazali bomo, da je $S \setminus \{v_i\}$ ogrodje za V, kar je v protislovju z minimalnostjo ogrodja S.

Ker je v_i linearna kombinacija $v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n$, obstajajo taki $\gamma_j \in F$, da velja $v_i = \gamma_1 v_1 + \ldots + \gamma_{i-1} v_{i-1} + \gamma_{i+1} v_{i+1} + \ldots + \gamma_n v_n$. Vzemimo poljuben $v \in V$. Ker je S ogrodje za V, obstajajo taki $\alpha_j \in F$, da velja $v = \alpha_1 v_1 + \ldots + \alpha_i v_i + \ldots + \alpha_n v_n$. Ko vstavimo izraz za v_i in uredimo, dobimo $v = (\alpha_1 + \alpha_i\gamma_1)v_1 + \ldots + (\alpha_{i-1} + \alpha_i\gamma_{i-1})v_{i-1} + (\alpha_{i+1} + \alpha_i\gamma_{i+1})v_{i+1} + \ldots + (\alpha_n + \alpha_i\gamma_n)v_n$, kar je linearna kombinacija elementov iz $S \setminus \{v_i\}$. Torej je $S \setminus \{v_i\}$ res ogrodje za V. □

Primeri iz obstoja baze

Zanima nas, kako za dano množico $\{v_1, \ldots, v_m\} \subset F^n$ preverimo ali je ogrodje in kako poiščemo tako njeno podmnožico, ki je baza. Matriko $\begin{bmatrix} v_1 & \ldots & v_m \end{bmatrix}$ z Gaussovo metodo prevedemo v reducirano vrstično stopničasto formo R. Če ima R ničelno vrstico, potem $\{v_1, \ldots, v_m\}$ ni ogrodje. Če R nima ničelne vrstice in če so $i_1, \ldots, i_n$ zaporedne številke njenih pivotnih stolpcev potem je množica $\{v_{i_1}, \ldots, v_{i_n}\}$ baza za F^n.
Utemeljitev: Če matriko $\begin{bmatrix} v_1 & \ldots & v_m \end{bmatrix}$ z leve množimo z elementarno matriko, se ohranijo vse linearne relacije med stolpci.

Primer te metode

Pokažimo, da so stolpci matrike

$$A = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix}$$

ogrodje za F^4 in poiščimo take štiri stolpce, ki so baza za F^4.

Z Gaussovo metodo izračunamo njeno reducirano vrstično kanonično formo

$$R = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix}$$

Ker matrika R nima ničelne vrstice, so stolci matrike A ogrodje za F^4. Zaporedne številke pivotnih stolpcev v R so $1, 2, 3, 5$. Stolpci v A z istimi zaporednimi številkami so potem baza za F^4. Iz R lahko odčitamo tudi, kako se s temi stolpci izražata preostala stolpca matrike A.

Primer: $F[x]$ ni končno-razsežen

Spomnimo se, da je $F[x]$ vektorski prostor vseh polinomov v x s koeficienti iz F. Množica $\{1, x, x^2, \ldots\}$ je ogrodje $F[x]$. Če bi imeli končno ogrodje $\{p_1, \ldots, p_n\}$, potem bi vsak polinom p iz $F[x]$ lahko izrazili kot $p = \alpha_1 p_1 + \ldots \alpha_n p_n$, odkoder bi sledilo $\deg p \leq \max\{\deg p_1, \ldots, \deg p_n\}$, kar je protislovje, saj imajo lahko polinomi poljubno visoko stopnjo.

Enoličnost moči baze

Radi bi dokazali, da imajo vse baze istega vektorskega prostora enako elementov. Začnimo s pomožno trditvijo.

Trditev 2

Naj bo V vektorski prostor. Če so $u_1, \ldots, u_m$ linearno neodvisni vektorji v V in če so vektorji $v_1, \ldots, v_n$ ogrodje za V, potem je $m \leq n$.

Dokaz: Vsak vektor u_i razvijmo po vektorjih v_j:

$$\begin{aligned} u_1 &= \alpha_{1,1} v_1 + \ldots + \alpha_{1,n} v_n \\ &\vdots \\ u_m &= \alpha_{m,1} v_1 + \ldots + \alpha_{m,n} v_n \end{aligned} \tag{1}$$

Privzemimo sedaj, da je $m > n$ in poskušajmo dobiti protislovje.

Ker ima vsak podoločen homogen sistem netrivialno rešitev, obstajajo taki skalarji $x_1, \ldots, x_m$, ki niso vsi enaki nič, da velja:

$$\begin{aligned} \alpha_{1,1}x_1 + \ldots + \alpha_{m,1}x_m &= 0 \\ &\vdots \\ \alpha_{1,n}x_1 + \ldots + \alpha_{m,n}x_m &= 0 \end{aligned} \tag{2}$$

Iz (1) in (2) sledi

$$\begin{aligned} \sum_{i=1}^{m} x_i u_i &= \sum_{i=1}^{m} x_i(\alpha_{i,1}v_1 + \ldots + \alpha_{i,n}v_n) \\ &= (\sum_{i=1}^{m} \alpha_{i,1}x_i)v_1 + \ldots + (\sum_{i=1}^{m} \alpha_{i,n}x_i)v_n \qquad (3) \\ &= 0v_1 + \ldots + 0v_n = 0 \end{aligned}$$

Ker so vektorji $u_1, \ldots, u_m$ linearno neodvisni, sledi iz (3), da je $x_1 = \ldots = x_m = 0$. To je v protislovju z izbiro $x_1, \ldots, x_m$. Predpostavka $m > n$ je torej napačna. □

Oglejmo si nekaj posledic te trditve.

Izrek 3 - Enoličnost moči baze

Če ima vektorski prostor V bazo z n elementi, potem velja:

- Vsaka linearno neodvisna množica v V ima $\leq n$ elementov.
- Vsako ogrodje v V ima $\geq n$ elementov.
- Vsaka baza v V ima n elementov.

Dokaz: Recimo, da je $v_1, \ldots, v_n$ baza za V in so $u_1, \ldots, u_m \in V$ poljubni.

Prvo točko dokažemo takole. Predpostavimo, da so $u_1, \ldots, u_m$ linearno neodvisni. Ker so $v_1, \ldots, v_n$ ogrodje, je po Trditvi 2 $m \leq n$.

Drugo točko dokažemo takole. Predpostavimo, da so $u_1, \ldots, u_m$ ogrodje. Ker so $v_1, \ldots, v_n$ linearno neodvisni, je po Trditvi 2 $n \leq m$.

Očitno tretja točka sledi iz prvih dveh. □

Definicija dimenzije

Razsežnost (=dimenzija) vektorskega prostora je število elementov katerekoli njegove baze. Pravimo, da je vektorski prostor n-razsežen (=n-dimenzionalen), če ima razsežnost n.

Primer dimenzije

Vektorski prostor F^n je n-razsežen, ker ima standardna baza n elementov.

Vektorski prostor $M_{m,n}(F)$ $m \times n$ matrik nad F je mn-razsežen, ker ima bazo iz mn koordinatnih matrik $E_{i,j}$.

Iz dokaza Izreka 2 sledi:

Trditev 3

Naj bo V n-razsežen vektorski prostor. Potem je vsako ogrodje za V, ki ima n elementov, baza za V.

Primer iz enoličnosti moči baze

Naj bo F polje in S neskončna množica. Pokažimo, da vektorski prostor F^S vseh funkcij iz S v F ni končno-razsežen. Recimo, da ima F^S ogrodje iz n elementov. Konstruirali bomo linearno neodvisno množico v F^S, ki ima $n+1$ elementov, kar je v nasprotju z Izrekom 3.

Za vsak $s \in S$ označimo z δ_s funkcijo, ki pošlje element s v 1, elemente iz $S \setminus \{s\}$ pa v nič. Ker je S neskončna množica, lahko izberemo $n+1$ paroma različnih elementov $s_1, \ldots, s_{n+1} \in S$. Pokažimo, da so funkcije $\delta_{s_1}, \ldots, \delta_{s_{n+1}}$ linearno neodvisne. Recimo, da velja

$$\alpha_1 \delta_{s_1} + \ldots + \alpha_{n+1} \delta_{s_{n+1}} = 0$$

za neke skalarje $\alpha_1, \ldots, \alpha_{n+1} \in F$. Potem ze vsak $x \in S$ velja

$$\alpha_1 \delta_{s_1}(x) + \ldots + \alpha_{n+1} \delta_{s_{n+1}}(x) = 0$$

Ko vstavimo $x = s_i$ za $i = 1, \ldots, n+1$, in upoštevamo $\delta_j(s_i) = 0$ za vsak $j \neq i$, dobimo $\alpha_i \delta_{s_i}(s_i) = 0$. Ker je $\delta_{s_i}(s_i) = 1$, sledi $\alpha_i = 0$.

Dopolnitev linearno neodvisne množice do baze

V tem razdelku bomo dokazali, da vsako linearno množico dopolnimo do baze. Začnimo z pomožno trditvijo.

Trditev 4

Če so vektorji $v_1, \ldots, v_m$ linearno neodvisni, in če vektor v_{m+1} ne leži v njihovi linearni ogrinjači, potem so tudi vektorji $v_1, \ldots, v_m, v_{m+1}$ linearno neodvisni.

Dokaz: Recimo, da so vektorji $v_1, \ldots, v_m$ linearno neodvisni, in da vektor v_{m+1} ne leži v njihovi linearni ogrinjači. Vzemimo take $\alpha_1, \ldots, \alpha_m, \alpha_{m+1}$, da velja $\alpha_1 v_1 + \ldots + \alpha_m v_m + \alpha_{m+1} v_{m+1} = 0$. Radi bi dokazali, da velja $\alpha_1 = \ldots, \alpha_m = \alpha_{m+1} = 0$. Ločimo dva primera.

Če je $\alpha_{m+1} \neq 0$, potem velja $v_{m+1} = \frac{-1}{\alpha_{m+1}}(\alpha_1 v_1 + \ldots + \alpha_m v_m)$, kar je v nasprotju s predpostavko, da v_{m+1} ne leži v linearni ogrinjači $v_1, \ldots, v_m$.

Če je $\alpha_{m+1} = 0$, potem velja $\alpha_1 v_1 + \ldots + \alpha_m v_m = 0$. Ker so vektorji $v_1, \ldots, v_m$ linearno neodvisni, je $\alpha_1 = \ldots, \alpha_m = 0$. □

Trditev 5

Naj bo V n-razsežen vektorski prostor. Vsaka linearno neodvisna množica v V, ki ima n elementov, je baza.

Dokaz: Če so vektorji $v_1, \ldots, v_n \in V$ linearno neodvisni, niso pa baza, potem niso ogrodje. Torej obstaja tak vektor v_{n+1}, ki ne leži v linearni ogrinjači vektorjev $v_1, \ldots, v_n$. Po Trditvi 4 si vektorji $v_1, \ldots, v_n, v_{n+1}$ linearno neodvisni. Ker je V n-razsežen, ima ogrodje z n elementi. Po Trditvi 2 ima vsaka linearno neodvisna množica manj ali enako elementov kot vsako ogrodje. Torej je $n + 1 \leq n$, kar je protislovje. □

Izrek 4 - Dopolnitev do baze

Vsaka linearno neodvisna množica v končno-razsežnem vektorskem prostoru je vsebovana v neki bazi.

Dokaz: Če je V n-razsežen vektorski prostor in so vektorji $v_1, \ldots, v_m \in V$ linearno neodvisni, potem je $m \leq n$ po Izreku 3. Če je $m = n$, potem so vektorji $v_1, \ldots, v_m$ baza po Trditvi 5. Če je $m < n$, potem:

- Izberimo tak $v_{m+1} \in V$, da $v_{m+1} \notin \mathrm{Lin}\{v_1, \ldots, v_m\}$.
- Izberimo tak $v_{m+2} \in V$, da $v_{m+2} \notin \mathrm{Lin}\{v_1, \ldots, v_m, v_{m+1}\}$.
- Izberimo tak $v_{m+3} \in V$, da $v_{m+3} \notin \mathrm{Lin}\{v_1, \ldots, v_m, v_{m+1}, v_{m+2}\}$.
- ...
- Izberimo tak $v_n \in V$, da $v_n \notin \mathrm{Lin}\{v_1, \ldots, v_m, v_{m+1}, \ldots, v_{n-1}\}$.

Obstoj takih $v_{m+1}, \ldots, v_n$ sledi iz dejstva, da nobena podmnožica v n razsežnem prostoru, ki ima manj kot n elementov, ni ogrodje (drugi del Izreka 3). Če $n - m$ krat uporabimo Trditev 4, dobimo, da so vektorji $v_1, \ldots, v_m, v_{m+1}, \ldots, v_n$ linearno neodvisni. Po Trditvi 5 so torej baza. □

Primer iz dopolnjevanja linearno neodvisnih množic

Zanima nas, kako za dano množico $\{v_1, \ldots, v_m\}$ v F^n ugotovimo ali je linearno neodvisna in kako jo dopolnimo do baze. Tvorimo matriko

$$A = \left[\begin{array}{cccccc} v_1 & \ldots & v_m & e_1 & \ldots & e_n \end{array} \right]$$

kjer je $e_1, \ldots, e_n$ standardna baza za F^n. Z Gaussovo metodo prevedemo matriko A na reducirano vrstično stopničasto formo R.

Zaporedne številke pivotnih stolpcev matrike R naj bodo $i_1, \ldots, i_n$. Vektorji $v_1, \ldots, v_m$ so linearno neodvisni natanko tedaj, ko velja $i_1 = 1$ in $\ldots$ in $i_m = m$. Stolpci v A z zaporednimi številkami $i_{m+1}, \ldots, i_n$ so potem dopolnitev množice $\{v_1, \ldots, v_m\}$ do baze.

Primer

Dopolni vektorja $\begin{bmatrix} 1 \\ -1 \\ 0 \\ 2 \end{bmatrix}$ in $\begin{bmatrix} 2 \\ 2 \\ 0 \\ 1 \end{bmatrix}$ do baze za $\mathbb{R}^4$.

Najprej tvorimo matriko

$$A = \begin{bmatrix} 1 & 2 & 1 & 0 & 0 & 0 \\ -1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 2 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Njena reducirana vrstična stopničasta forma je

$$R = \begin{bmatrix} 1 & 0 & 0 & -\frac{1}{5} & 0 & \frac{2}{5} \\ 0 & 1 & 0 & \frac{2}{5} & 0 & \frac{1}{5} \\ 0 & 0 & 1 & -\frac{3}{5} & 0 & -\frac{4}{5} \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

Njeni pivotni stolpci imajo zaporedne številke $1, 2, 3$ in 5. Torej sta $\begin{bmatrix} 1 \\ -1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \\ 0 \\ 1 \end{bmatrix}$ linearno neodvisna in $\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$ ju dopolnita do baze.

Povzetek - Karakterizacije baz in razsežnosti

Povzemimo vse karakterizacije baz v en izrek.

Povzetek 1

Za vsako podmnožico S v n-razsežnem prostoru V so ekvivalentne trditve:

- S je baza.
- S je linearno neodvisna in je ogrodje.
- S je linearno neodvisna in ima n elementov.
- S je ogrodje in ima n elementov.
- S je linearno neodvisna in nima nobene prave linearno neodvisne nadmnožice.
- S je ogrodje in nima nobene prave podmnožice, ki bi bila ogrodje.

Povzemimo še vse karakterizacije razsežnosti v en izrek.

Povzetek 2

Za vsak vektorski prostor V so ekvivalentne naslednje trditve:

- V je n-razsežen.
- V ima ogrodje z n elementi, nima pa ogrodja z $n-1$ elementi.
- V ima bazo z n elementi.
- V ima linearno neodvisno množico z n elementi, nima pa linearno neodvisne množice z $n+1$ elementi.

Vektorski prostori 3.del

Razsežnosti podprostorov

Naj bo V končno-razsežen vektorski prostor. Njegovo razsežnost označimo z $\dim V$. Lahko jo definiramo na več ekvivalentnih načinov:

- Kot minimum moči ogrodij za V.
- Kot moč katerekoli baze za V. (Vse baze imajo enako moč.)
- Kot maksimum moči linearno neodvisnih množic v V.

Če V ni končno-razsežen, potem pišemo $\dim V = +\infty$.

Trditev

Če je V končno-razsežen vektorski prostor in če je U vektorski podprostor v V, potem je tudi U končno-razsežen vektorski prostor in $\dim U \leq \dim V$.

Dokaz: Naj bo $n = \dim V$. Če $\dim U = +\infty$ ali če $\dim U > n$, potem U nima ogrodja z $\leq n$ elementi. Torej obstajajo taki $u_1, u_2, \ldots, u_{n+1} \in U$ da $u_1 \notin \mathrm{Lin}\{\}$, $u_2 \notin \mathrm{Lin}\{u_1\}$, $u_3 \notin \mathrm{Lin}\{u_1, u_2\}$, $\ldots, u_{n+1} \notin \mathrm{Lin}\{u_1, \ldots, u_n\}$. Po Trditvi 3 od zadnjič odtod sledi, da so vektorji $u_1, u_2, \ldots, u_{n+1}$ linearno neodvisni v U (in zato tudi linearno neodvisni v V). Toda vsaka linearno neodvisna množica v V ima največ n elementov. To je protislovje.

Trditev (Dimenzijska formula za podprostore)

Če je V končno-razsežen vektorski prostor in če sta U_1 in U_2 vektorska podprostora v V, potem velja naslednja formula

$$\dim(U_1 + U_2) = \dim U_1 + \dim U_2 - \dim(U_1 \cap U_2).$$

Dokaz: Naj bo $f_1, \ldots, f_k$ baza za podprostor $U_1 \cap U_2$. Naj bo $g_1, \ldots, g_l$ njena dopolnitev do baze za U_1 in naj bo $h_1, \ldots, h_m$ njena dopolnitev do baze za U_2. Dokazali bomo, da je $f_1, \ldots, f_k, g_1, \ldots, g_l, h_1, \ldots, h_m$ baza za $U_1 + U_2$. Potem je $\dim(U_1 + U_2) = k + l + m$, $\dim(U_1 \cap U_2) = k$, $\dim U_1 = k + l$ in $\dim U_2 = k + m$. Torej formula v trditvi drži.

Pokažimo najprej, da so $f_1, \ldots, f_k, g_1, \ldots, g_l, h_1, \ldots, h_m$ ogrodje za $U_1 + U_2$. Po definiciji vsote dveh podprostorov lahko vsak vektor $u \in U_1 + U_2$ izrazimo kot $u = u_1 + u_2$, kjer $u_1 \in U_1$ in $u_2 \in U_2$. Ker je $f_1, \ldots, f_k, g_1, \ldots, g_l$ baza za U_1, lahko izrazimo u_1 kot linearno kombinacijo $f_1, \ldots, f_k, g_1, \ldots, g_l$. Ker je $f_1, \ldots, f_k, h_1, \ldots, h_m$ baza za U_2, lahko izrazimo u_2 kot linearno kombinacijo $f_1, \ldots, f_k, h_1, \ldots, h_m$. Torej lahko u izrazimo kot linearno kombinacijo $f_1, \ldots, f_k, g_1, \ldots, g_l, h_1, \ldots, h_m$.

Pokažimo še, da so $f_1, \ldots, f_k, g_1, \ldots, g_l, h_1, \ldots, h_m$ linearno neodvisni. Če $\alpha_1 f_1 + \ldots + \alpha_k f_k + \beta_1 g_1 + \ldots + \beta_l g_l + \gamma_1 h_1 + \ldots + \gamma_m h_m = 0$, potem $\alpha_1 f_1 + \ldots + \alpha_k f_k + \beta_1 g_1 + \ldots + \beta_l g_l = (-\gamma_1) h_1 + \ldots + (-\gamma_m) h_m$ leži v $U_1 \cap U_2$ (ker leva stran leži v U_1, desna pa v U_2.) Torej obstajajo taki skalarji $\delta_1, \ldots, \delta_k$, da je $(-\gamma_1) h_1 + \ldots + (-\gamma_m) h_m = \delta_1 f_1 + \ldots + \delta_k f_k$. Ker je $f_1, \ldots, f_k, h_1, \ldots, h_m$ baza za U_2, odtod sledi $\gamma_1 = \ldots = \gamma_m = 0$ (in $\delta_1 = \ldots = \delta_k = 0$). Sledi $\alpha_1 f_1 + \ldots + \alpha_k f_k + \beta_1 g_1 + \ldots + \beta_l g_l = 0$. Ker je $f_1, \ldots, f_k, g_1, \ldots, g_l$ baza za U_1, odtod sledi $\alpha_1 = \ldots = \alpha_k = 0$ in $\beta_1 = \ldots = \beta_l = 0$. Dokaz linearne neodvisnosti je s tem končan. □

Posledica dimenzijske formule

Za vsaka podprostora U_1 in U_2 v V velja $\dim(U_1 + U_2) \leq \dim U_1 + \dim U_2$. Enačaj velja natanko tedaj, ko je $U_1 \cap U_2 = \{0\}$.

Dokaz. V dimenzijski formuli upoštevamo, da je $\dim(U_1 \cap U_2) \geq 0$ in da velja $\dim(U_1 \cap U_2) = 0$ natanko tedaj, ko je $U_1 \cap U_2 = \{0\}$. □

Opomba: Dimenzijsko formulo lahko posplošimo tudi na tri podprostore:

$$\dim(U_1 + U_2 + U_3) = \dim U_1 + \dim U_2 + \dim U_3 + \dim(U_1 \cap U_2 \cap U_3) - \dim(U_1 \cap U_2) - \dim(U_1 \cap U_3) - \dim(U_2 \cap U_3)$$

Direktne vsote podprostorov

Pojme linearne ogrinjače, ogrodja, linearne neodvisnosti in baze lahko posplošimo iz vektorjev na vektorske podprostore. Terminologija ni ista.

Posplošitev pojma linearne ogrinjače vektorjev je pojem vsote vektorskih podprostorov.

Definicija vsote vektorskih podprostorov

Vsota vektorskih podprostorov $U_1, \ldots, U_n$ (nekega vektorskega prostora) je vektorski podprostor

$$U_1 + \ldots + U_n := \{u_1 + \ldots + u_n \mid u_1 \in U_1, \ldots, u_n \in U_n\}$$

Opomba: Vektorske podprostore $U_1, \ldots, U_n$ lahko smatramo za "ogrodje" vektorskega prostora V, če lahko vsak vektor $v \in V$ na **vsaj en** način izrazimo kot $v = u_1 + \ldots + u_n$, kjer $u_1 \in U_1, \ldots u_n \in U_n$. To velja natanko tedaj, ko je $V = U_1 + \ldots + U_n$.

Definicija (interne) direktne vsote podprostorov

Vsota vektorskih podprostor $U_1, \ldots, U_n$ je **direktna**, če za vsake vektorje $v_1, w_1 \in U_1, \ldots, v_n, w_n \in U_n$, ki zadoščajo $v_1 + \ldots + v_n = w_1 + \ldots + w_n$ velja $v_1 = w_1, \ldots, v_n = w_n$.

Opomba: Vektorske podprostore $U_1, \ldots, U_n$ lahko smatramo za "linearno neodvisne", če lahko vsak vektor $v \in V$ na **največ en** način izrazimo kot $v = u_1 + \ldots + u_n$, kjer $u_1 \in U_1, \ldots, u_n \in U_n$. To velja natanko tedaj, ko je vsota $U_1 + \ldots + U_n$ direktna.

Opomba: **Eksterna direktna vsota** podprostorov $U_1, \ldots, U_n$ je definirana kot množica vseh n-teric $(u_1, \ldots, u_n)$, kjer $u_i \in U_i$ za vsak $i = 1, \ldots, n$. Označimo jo z $U_1 \oplus \ldots \oplus U_n$. Naj bo Φ preslikava iz $U_1 \oplus \ldots \oplus U_n$ v V, ki pošlje $(u_1, \ldots, u_n)$ v $u_1 + \ldots + u_n$. Njena zaloga vrednosti je ravno vsota $U_1 + \ldots + U_n$. Preslikava Φ je injektivna natanko tedaj, ko je vsota $U_1 + \ldots + U_n$ direktna. Torej lahko preko Φ identificiramo interno direktno vsoto $U_1 + \ldots + U_n$ z eksterno direktno vsoto $U_1 \oplus \ldots \oplus U_n$. Zato tudi interno direktno vsoto označimo z $U_1 \oplus \ldots \oplus U_n$.

Osnovni primer

Za vse vektorje $v_1, \ldots, v_n$ v V velja

$$\mathrm{Lin}\{v_1\} + \ldots + \mathrm{Lin}\{v_n\} = \mathrm{Lin}\{v_1, \ldots, v_n\}$$

Torej so vektorji $v_1, \ldots, v_n$ ogrodje za V natanko tedaj, ko velja $\mathrm{Lin}\{v_1\} + \ldots + \mathrm{Lin}\{v_n\} = V$ ($\Leftrightarrow \mathrm{Lin}\{v_1\}, \ldots, \mathrm{Lin}\{v_n\}$ so "ogrodje").

Neničelni vektorji $v_1, \ldots, v_n$ v V so LN natanko tedaj, ko je vsota $\mathrm{Lin}\{v_1\} + \ldots + \mathrm{Lin}\{v_n\}$ direktna ($\Leftrightarrow \mathrm{Lin}\{v_1\}, \ldots, \mathrm{Lin}\{v_n\}$ so "LN").
D: Vsota $\mathrm{Lin}\{v_1\} + \ldots + \mathrm{Lin}\{v_n\}$ je direktna natanko tedaj, ko za vsake skalarje α_i, β_j velja, da iz $\alpha_1 v_1 + \ldots + \alpha_n v_n = \beta_1 v_1 + \ldots + \beta_n v_n$ sledi $\alpha_1 v_1 = \beta_1 v_1$, $\ldots$, $\alpha_n v_n = \beta_n v_n$. Ker so vektorji $v_1, \ldots, v_n$ neničelni je to ekvivalentno lastnosti, da iz $\alpha_1 v_1 + \ldots + \alpha_n v_n = \beta_1 v_1 + \ldots + \beta_n v_n$ sledi $\alpha_1 = \beta_1$, $\ldots$, $\alpha_n = \beta_n$. To pa je ravno linearna neodvisnost $v_1, \ldots, v_n$.

Opomba: Vektorske podprostore $U_1, \ldots, U_n$ lahko smatramo za "bazo" vektorskega prostora V, če lahko vsak vektor $v \in V$ na **natanko en** način izrazimo kot $v = u_1 + \ldots + u_n$, kjer $u_1 \in U_1$, $\ldots$, $u_n \in U_n$. To velja natanko tedaj, ko je $V = U_1 \oplus \ldots \oplus U_n$.

Karakterizacije direktnih vsot

Direktno vsoto vektorskih podprostorov bi radi opisali na čimveč različnih načinov. Mnoge od teh opisov bomo potrebovali v naslednjih poglavjih.

Izrek (Karakterizacije direktnih vsot)

Za vsako naravno število n in za vsake podprostore $U_1, \ldots, U_n$ v vsakem končno-razsežnem vektorskem prostoru V so ekvivalentne lastnosti:

(1) Vsota $U_1 + \ldots + U_n$ je direktna.

(2) Če vektorji $u_1 \in U_1, \ldots, u_n \in U_n$ zadoščajo $u_1 + \ldots + u_n = 0$ (v V), potem velja $u_1 = \ldots = u_n = 0$.

(3) Če so $u_{1,1}, \ldots, u_{1,r_1}$ baza za U_1 in $\ldots$ in so $u_{n,1}, \ldots, u_{n,r_n}$ baza za U_n, potem so $u_{1,1}, \ldots, u_{1,r_1}, \ldots, u_{n,1}, \ldots, u_{n,r_n}$ baza za $U_1 + \ldots + U_n$.

(4) Velja $\dim(U_1 + \ldots + U_n) = \dim U_1 + \ldots + \dim U_n$.

(5) Za vsak $i = 1, \ldots, n-1$ velja $(U_1 + \ldots + U_i) \cap U_{i+1} = \{0\}$.

(1) $\Rightarrow$ (2) Recimo da velja (1). Vzemimo take $u_i \in U_i$, da velja $0 = u_1 + \ldots + u_n$. Ker velja tudi $0 \in U_i$ za vse i in $0 = 0 + \ldots + 0$, sledi po (1), da je $u_1 = 0, \ldots, u_n = 0$. Torej velja (2).

(2) $\Rightarrow$ (1) Recimo, da velja (2). Vzemimo take $u_i, v_i \in U_i$, da velja $u_1 + \ldots + u_n = v_1 + \ldots + v_n$. Odtod sledi $(u_1 - v_1) + \ldots + (u_n - v_n) = 0$ in $u_i - v_i \in U_i$ za vse i. Iz (2) sledi $u_i - v_i = 0$ za vse i. Torej velja (1).

(2) $\Rightarrow$ (3) Recimo, da velja (2) in da je za vsak i $u_{i,1}, \ldots, u_{i,r_1}$ baza za U_i. Vsak vektor $u \in U_1 + \ldots + U_n$ lahko izrazimo kot $u = u_1 + \ldots + u_n$, kjer $u_i \in U_i$ za vse i. Vsak u_i lahko izrazimo kot linearno kombinacijo vektorjev $u_{i,1}, \ldots, u_{i,r_1}$, ker so ti ogrodje za U_i. Torej lahko u izrazimo kot linearno kombinacijo vektorjev $u_{1,1}, \ldots, u_{1,r_1}, \ldots, u_{n,1}, \ldots, u_{n,r_n}$. Torej so ti vektorji ogrodje za $U_1 + \ldots + U_n$. Pokažimo še, da so linearno neodvisni. Recimo, da velja $\alpha_{1,1}u_{1,1} + \ldots + \alpha_{1,r_1}u_{1,r_1} + \ldots + \alpha_{n,1}u_{n,1} + \ldots + \alpha_{n,r_n}u_{n,r_n} = 0$. Za vsak i označimo $u_i := \alpha_{i,1}u_{i,1} + \ldots + \alpha_{i,r_i}u_{i,r_i}$. Očitno $u_i \in U_i$ za vsak i in $u_1 + \ldots + u_n = 0$. Po točki (2) odtod sledi, da je $u_i = 0$ za vse i. Ker so $u_{i,1}, \ldots, u_{i,r_1}$ linearno neodvisni, res velja $\alpha_{i,1} = \ldots = \alpha_{i,r_i} = 0$ za vse i.

(3) $\Rightarrow$ (4) Upoštevamo, da je razsežnost enaka moči baze.

(4) $\Rightarrow$ (5) Če n-krat uporabimo posledico dimenzijske formule, dobimo

$$\dim(U_1 + \ldots + U_n) \leq \dim(U_1 + \ldots + U_{n-1}) + \dim U_n \leq$$
$$\leq (\dim(U_1 + \ldots + U_{n-2}) + \dim U_{n-1}) + \dim U_n \leq \ldots$$
$$\ldots \leq \dim U_1 + \ldots + \dim U_n$$

Če velja (4), potem se začetek in konec ujemata, torej imamo povsod enačaj. Odtod sledi, da za vsak $i = 1, \ldots, n-1$ velja

$$\dim(U_1 + \ldots + U_i + U_{i+1}) = \dim(U_1 + \ldots + U_i) + \dim U_{i+1}$$

Če za vsak i uporabimo posledico dimenzijske formule, dobimo (5).

(5) $\Rightarrow$ (2) Recimo, da velja (5). Vzemimo take vektorje $u_1 \in U_1$, $\ldots$, $u_n \in U_n$, da velja $u_1 + \ldots + u_n = 0$. Odtod sledi, da vektor $u_1 + \ldots + u_{n-1} = -u_n$ pripada podprostoru $(U_1 + \ldots + U_{n-1}) \cap U_n$, ki je po predpostavki enak $\{0\}$. Torej je $u_1 + \ldots + u_{n-1} = 0$ in $u_n = 0$. Odtod sledi, da vektor $u_1 + \ldots + u_{n-2} = -u_{n-1}$ pripada podprostoru $(U_1 + \ldots + U_{n-2}) \cap U_{n-1}$, ki je po predpostavki enak $\{0\}$. Torej je $u_1 + \ldots + u_{n-2} = 0$ in $u_{n-1} = 0$. Po nekaj korakih dobimo, da so vsi u_i enaki nič. torej velja (2). $\square$

Prehod na novo bazo

Naj bo V n-razsežen vektorski prostor in naj bosta $\mathcal{B} = \{u_1, \ldots, u_n\}$ ter $\mathcal{C} = \{w_1, \ldots, w_n\}$ dve bazi za V. ($\mathcal{B}$ je stara baza, $\mathcal{C}$ pa nova baza.)

Vsak vektor $v \in V$ lahko enolično razvijemo tako po bazi $\mathcal{B}$ kot po bazi $\mathcal{C}$:

$$v = \beta_1 u_1 + \ldots + \beta_n u_n \tag{1}$$

$$v = \gamma_1 w_1 + \ldots + \gamma_n w_n \tag{2}$$

V tem primeru pišemo

$$[v]_{\mathcal{B}} = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_n \end{bmatrix} \quad \text{in} \quad [v]_{\mathcal{C}} = \begin{bmatrix} \gamma_1 \\ \vdots \\ \gamma_n \end{bmatrix} \tag{3}$$

Konstruirali bomo tako matriko $P_{\mathcal{C}\leftarrow\mathcal{B}}$, da bo za vsak $v \in V$ veljalo

$$[v]_{\mathcal{C}} = P_{\mathcal{C}\leftarrow\mathcal{B}}[v]_{\mathcal{B}} \tag{4}$$

Taki matriki bomo rekli **prehodna matrika** iz baze $\mathcal{B}$ na bazo $\mathcal{C}$.

Vsak element baze $\mathcal{B}$ lahko razvijemo po bazi $\mathcal{C}$. Dobimo

$$\begin{aligned} u_1 &= \alpha_{1,1} w_1 + \ldots + \alpha_{1,n} w_n \\ &\vdots \\ u_n &= \alpha_{n,1} w_1 + \ldots + \alpha_{n,n} w_n \end{aligned} \tag{5}$$

Skalarje $\alpha_{i,j}$ potem zložimo v matriko

$$P_{\mathcal{C}\leftarrow\mathcal{B}} = \begin{bmatrix} \alpha_{1,1} & \ldots & \alpha_{n,1} \\ \vdots & & \vdots \\ \alpha_{1,n} & \ldots & \alpha_{n,n} \end{bmatrix} \tag{6}$$

Pozor, skalarje iz i-tega razvoja smo zložili v i-ti stolpec (in ne v i-to vrstico kot bi pričakovali). Na kratko definicijo $P_{\mathcal{C}\leftarrow\mathcal{B}}$ zapišemo kot

$$P_{\mathcal{C}\leftarrow\mathcal{B}} = \begin{bmatrix} [u_1]_{\mathcal{C}} & \ldots & [u_n]_{\mathcal{C}} \end{bmatrix} \tag{7}$$

Dokažimo sedaj formulo (4). Če formule (5) vstavimo v formulo (1) in uredimo, dobimo razvoj

$$\begin{aligned} v &= \beta_1 u_1 + \ldots + \beta_n u_n \\ &= \beta_1(\alpha_{1,1} w_1 + \ldots + \alpha_{1,n} w_n) + \ldots + \beta_n(\alpha_{n,1} w_1 + \ldots + \alpha_{n,n} w_n) \\ &= (\beta_1 \alpha_{1,1} + \ldots + \beta_n \alpha_{n,1}) w_1 + \ldots + (\beta_1 \alpha_{1,n} + \ldots + \beta_n \alpha_{n,n}) w_n \end{aligned} \tag{8}$$

Če primerjamo koeficiente v razvojih (2) in (8), dobimo

$$\begin{aligned} \gamma_1 &= \beta_1 \alpha_{1,1} + \ldots + \beta_n \alpha_{n,1} \\ &\vdots \\ \gamma_n &= \beta_1 \alpha_{1,n} + \ldots + \beta_n \alpha_{n,n} \end{aligned} \tag{9}$$

V matričnem zapisu se (9) glasi

$$\begin{bmatrix} \gamma_1 \\ \vdots \\ \gamma_n \end{bmatrix} = \begin{bmatrix} \alpha_{1,1} & \ldots & \alpha_{n,1} \\ \vdots & & \vdots \\ \alpha_{1,n} & \ldots & \alpha_{n,n} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_n \end{bmatrix} \tag{10}$$

Če v formuli (10) upoštevamo definiciji (3) in (6) dobimo formulo (4).

Primeri prehodnih matrik

Vzemimo dve bazi v $\mathbb{R}^3$

$$\mathcal{B} = \{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \} \quad \text{in} \quad \mathcal{C} = \{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \}$$

Izračunajmo prehodni matriki $P_{\mathcal{C}\leftarrow\mathcal{B}}$ in $P_{\mathcal{B}\leftarrow\mathcal{C}}$. Ker je

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = 1 \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 0 \cdot \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + 0 \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = (-1) \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 1 \cdot \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + 0 \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = 0 \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + (-1) \cdot \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + 1 \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

velja $P_{\mathcal{C}\leftarrow\mathcal{B}} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}$. Izračun $P_{\mathcal{B}\leftarrow\mathcal{C}}$ je preprostejši. Iz

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = 1 \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 0 \cdot \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + 0 \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = 1 \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 1 \cdot \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + 0 \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = 1 \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 1 \cdot \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + 1 \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

sledi, da je $P_{\mathcal{B}\leftarrow\mathcal{C}} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$. Opazimo, da je produkt obeh prehodnih matrik identična matrika. To velja tudi v splošnem.

Lastnosti prehodnih matrik

Doslej smo dokazali naslednjo lastnost prehodnih matrik.

Izrek 1 (Osnovna formula za prehodne matrike)

Naj bosta $\mathcal{B}$ in $\mathcal{C}$ dve bazi končno-razsežnega vektorskega prostora V. Potem za vsak vektor $v \in V$ velja $[v]_{\mathcal{C}} = P_{\mathcal{C}\leftarrow\mathcal{B}}[v]_{\mathcal{B}}$

Oglejmo si nekaj posledic izreka 1.

Izrek 2 (Produkt prehodnih matrik)

Naj bodo $\mathcal{B}$, $\mathcal{C}$ in $\mathcal{D}$ tri baze končno-razsežnega vektorskega prostora V. Potem velja $P_{\mathcal{D}\leftarrow\mathcal{B}} = P_{\mathcal{D}\leftarrow\mathcal{C}} P_{\mathcal{C}\leftarrow\mathcal{B}}$

Dokaz: Naj bo $\mathcal{B} = \{u_1, \ldots, u_n\}$. Upoštevajmo formulo (7) in Izrek 1.

$$\begin{aligned} P_{\mathcal{D}\leftarrow\mathcal{C}} P_{\mathcal{C}\leftarrow\mathcal{B}} &= P_{\mathcal{D}\leftarrow\mathcal{C}} \begin{bmatrix} [u_1]_{\mathcal{C}} & \ldots & [u_n]_{\mathcal{C}} \end{bmatrix} \\ &= \begin{bmatrix} P_{\mathcal{D}\leftarrow\mathcal{C}}[u_1]_{\mathcal{C}} & \ldots & P_{\mathcal{D}\leftarrow\mathcal{C}}[u_n]_{\mathcal{C}} \end{bmatrix} \\ &= \begin{bmatrix} [u_1]_{\mathcal{D}} & \ldots & [u_n]_{\mathcal{D}} \end{bmatrix} \\ &= P_{\mathcal{D}\leftarrow\mathcal{B}} \end{aligned}$$

Izrek 3 (Inverz prehodne matrike)

Naj bosta $\mathcal{B}$ in $\mathcal{C}$ dve bazi končno-razsežnega vektorskega prostora. Potem velja $(P_{\mathcal{C}\leftarrow\mathcal{B}})^{-1} = P_{\mathcal{B}\leftarrow\mathcal{C}}$

Dokaz: Če uporabimo Izrek 2 z $\mathcal{D} = \mathcal{B}$, dobimo $P_{\mathcal{B}\leftarrow\mathcal{C}} P_{\mathcal{C}\leftarrow\mathcal{B}} = P_{\mathcal{B}\leftarrow\mathcal{B}}$. Po definiciji prehodne matrike je $P_{\mathcal{B}\leftarrow\mathcal{B}}$ identična matrika. Odtod sledi, da je $P_{\mathcal{C}\leftarrow\mathcal{B}}$ obrnljiva matrika in da velja

$$(P_{\mathcal{C}\leftarrow\mathcal{B}})^{-1} = P_{\mathcal{B}\leftarrow\mathcal{C}}$$

Posledica

Naj bosta $\mathcal{B} = \{u_1, \ldots, u_n\}$ in $\mathcal{C} = \{w_1, \ldots, w_n\}$ bazi vektorskega prostora F^n. Potem velja $P_{\mathcal{C}\leftarrow\mathcal{B}} = \begin{bmatrix} w_1 & \ldots & w_n \end{bmatrix}^{-1} \begin{bmatrix} u_1 & \ldots & u_n \end{bmatrix}$.

Dokaz: Naj bo $\mathcal{S}$ standardna baza F^n. Potem je $[v]_{\mathcal{S}} = v$ za vsak $v \in F^n$. Iz formule (7) sledi $P_{\mathcal{S}\leftarrow\mathcal{B}} = \begin{bmatrix} [u_1]_{\mathcal{S}} & \ldots & [u_n]_{\mathcal{S}} \end{bmatrix} = \begin{bmatrix} u_1 & \ldots & u_n \end{bmatrix}$. Podobno dokažemo $P_{\mathcal{S}\leftarrow\mathcal{C}} = \begin{bmatrix} w_1 & \ldots & w_n \end{bmatrix}$. Po izrekih 2 in 3 je

$$P_{\mathcal{C}\leftarrow\mathcal{B}} = P_{\mathcal{C}\leftarrow\mathcal{S}} P_{\mathcal{S}\leftarrow\mathcal{B}} = \begin{bmatrix} w_1 & \ldots & w_n \end{bmatrix}^{-1} \begin{bmatrix} u_1 & \ldots & u_n \end{bmatrix}$$

Linearne preslikave 1. del

Definicija linearne preslikave

Zanimajo nas preslikave med vektorskimi prostori, ki ohranjajo strukturo (se pravi ohranjajo vsoto in produkt s skalarjem). Običajno jim pravimo linearne preslikave, včasih pa tudi homomorfizmi vektorskih prostorov.

Definicija linearne preslikave

Naj bosta U in V vektorska prostora nad istim poljem F.

- Preslikava $L\colon U \to V$ je **aditivna**, če velja $L(u_1 + u_2) = L(u_1) + L(u_2)$ za vsaka $u_1, u_2 \in U$.
- Preslikava $L\colon U \to V$ je **homogena**, če velja $L(\alpha u) = \alpha L(u)$ za vsak $u \in U$ in vsak $\alpha \in F$.
- Preslikava $L\colon U \to V$ je **linearna**, če je aditivna in homogena.

Opomba: Namesto $L(u)$ bomo pogosto pisali kar Lu. Za linearno preslikavo torej velja $L(u_1 + u_2) = Lu_1 + Lu_2$ in $L(\alpha u) = \alpha Lu$.

Opomba: Aditivna preslikava iz U v V je isto kot homomorfizem Abelovih grup iz $(U, +)$ v $(V, +)$.

Aditivnost in homogenost lahko združimo v eno lastnost:

Trditev

Naj bosta U in V vektorska prostora nad istim poljem F. Preslikava $L\colon U \to V$ je linearna natanko tedaj, ko velja

$$L(\alpha_1 u_1 + \alpha_2 u_2) = \alpha_1 L u_1 + \alpha_2 L u_2 \tag{1}$$

za vsaka $u_1, u_2 \in U$ in vsaka $\alpha_1, \alpha_2 \in F$.

Dokaz: Pokažimo najprej, da iz linearnosti sledi lastnost (1). Če je L aditivna, velja $L(\alpha_1 u_1 + \alpha_2 u_2) = L(\alpha_1 u_1) + L(\alpha_2 u_2)$. Če je L homogena, velja $L(\alpha_1 u_1) = \alpha_1 L u_1$ in $L(\alpha_2 u_2) = \alpha_2 L u_2$. Če je L tako aditivna kot homogena, potem torej velja (1).

Pokažimo še, da iz lastnosti (1) sledi linearnost. Če L zadošča (1), potem velja $L(u_1 + u_2) = L(1 \cdot u_1 + 1 \cdot u_2) = 1 \cdot L u_1 + 1 \cdot L u_2 = L u_1 + L u_2$ in $L(\alpha u) = L(\alpha \cdot u + 0 \cdot u) = \alpha \cdot L u + 0 \cdot L u = \alpha L u$. Torej je L tako aditivna kot homogena.

Primer - Projekcija na x os

Če točko (x, y) v ravnini projeciramo na x os, dobimo točko $(x, 0)$. Dokažimo, da je $L(x, y) = (x, 0)$ linearna preslikava. To sledi iz

$$L(\alpha_1(x_1, y_1) + \alpha_2(x_2, y_2)) = L(\alpha_1 x_1 + \alpha_2 x_2, \alpha_1 y_1 + \alpha_2 y_2) =$$
$$= (\alpha_1 x_1 + \alpha_2 x_2, 0) = \alpha_1(x_1, 0) + \alpha_2(x_2, 0) = \alpha_1 L(x_1, y_1) + \alpha_2 L(x_2, y_2).$$

Primer - Vrtež okrog izhodišča

Če točko $(x, y) = (r\cos\phi, r\sin\phi)$ v ravnini zavrtimo okrog izhodišča za kot τ v nasprotni smeri od urinega kazalca, potem dobimo točko $(r\cos(\phi + \tau), r\sin(\phi + \tau)) = (x\cos\tau - y\sin\tau, x\sin\tau + y\cos\tau)$, kjer smo uporabili adicijska izreka za cos in sin. Dokažimo, da je $L(x, y) = (x\cos\tau - y\sin\tau, x\sin\tau + y\cos\tau)$ linearna preslikava:

$$L(\alpha_1(x_1, y_1) + \alpha_2(x_2, y_2)) = L(\alpha_1 x_1 + \alpha_2 x_2, \alpha_1 y_1 + \alpha_2 y_2) =$$
$$= ((\alpha_1 x_1 + \alpha_2 x_2)\cos\tau - (\alpha_1 y_1 + \alpha_2 y_2)\sin\tau, (\alpha_1 x_1 + \alpha_2 x_2)\sin\tau + (\alpha_1 y_1 + \alpha_2 y_2)\cos\tau) =$$
$$= \alpha_1(x_1\cos\tau - y_1\sin\tau, x_1\sin\tau + y_1\cos\tau) + \alpha_2(x_2\cos\tau - y_2\sin\tau, x_2\sin\tau + y_2\cos\tau) =$$
$$= \alpha_1 L(x_1, y_1) + \alpha_2 L(x_2, y_2).$$

Primer - Odvajanje

Vemo, da za odvod velja

$$(f+g)' = f' + g' \quad \text{in} \quad (cf)' = cf'$$

kjer je c konstanta. Odvod je torej linearna preslikava iz vektorskega podprostora vseh odvedljivih funkcij v $\mathbb{R}^{[a,b]}$ v vektorski prostor $\mathbb{R}^{[a,b]}$.

Primer - Integriranje

Vemo, da je določeni integral definiran za zvezne funkcije in da velja

$$\int_a^b (f(x)+g(x))\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx$$

$$\text{in} \quad \int_a^b cf(x)\,dx = c\int_a^b f(x)\,dx$$

Torej je $\int_a^b$ linearna preslikava iz vektorskega podprostora vseh zveznih funkcij v $\mathbb{R}^{[a,b]}$ v vektorski prostor $\mathbb{R}$.

Posplošitve linearnih preslikav

Opomba: Če sta U in V vektorska prostora nad različnima poljema, potem ne moremo definirati linearne preslikave iz U v V. Lahko pa definiramo neko posplošitev linearne preslikave.

Recimo, da je U vektorski prostor nad poljem F in V vektorski prostor nad poljem K. Naj bo $\phi\colon F \to K$ homomorfizem polj. Preslikava $L\colon U \to V$ ϕ-**homogena**, če velja $L(\alpha u) = \phi(\alpha)L(u)$ za vsak $u \in U$ in vsak $\alpha \in F$. Preslikava $L\colon U \to V$ je ϕ-**linearna**, če je aditivna in ϕ-homogena.

Primer: Če je $F = K = \mathbb{C}$ in $\phi(z) = \overline{z}$ (se pravi $\phi(a+bi) = a-bi$), potem ϕ-linearni preslikavi iz U v V pravimo **konjugirano linearna** preslikava.

Opomba: Naj bosta U in V vektorska prostora nad poljem F in naj bo K podpolje v F. Preslikava $L\colon U \to V$ je K-**homogena**, če velja $L(\alpha u) = \alpha Lu$ za vsak $\alpha \in K$ in vsak $u \in U$. Preslikava $L\colon U \to V$ je K-**linearna**, če je aditivna in K-homogena.

Primer: Konjugiranje je $\mathbb{R}$-linearna preslikava iz $\mathbb{C}$ v $\mathbb{C}$, ni pa $\mathbb{C}$-linearna.

Opomba: Če sta U in V modula nad kolobarjem F, potem linearni preslikavi iz U v V rečemo **homomorfizem modulov**.

Linearni izomorfizmi

Definicija linearnega izomorfizma

Bijektivni linearni preslikavi pravimo **linearni izomorfizem**.

Trditev 1

Inverzna preslikava od linearnega izomorfizma je spet linearni izomorfizem.

Dokaz: Naj bosta U in V vektorska prostora nad poljem F in naj bo $L\colon U \to V$ bijektivna linearna preslikava. Radi bi pokazali, da je tudi njena inverzna preslikava $L^{-1}\colon V \to U$ linearna.

Za vsaka vektorja $v_1, v_2 \in V$ in vsaka skalarja $\alpha_1, \alpha_2 \in F$ velja

$$L(\alpha_1 L^{-1} v_1 + \alpha_2 L^{-1} v_2) = L(\alpha_1 u_1 + \alpha_2 u_2) = \alpha_1 L u_1 + \alpha_2 L u_2 =$$
$$= \alpha_1 L(L^{-1} v_1) + \alpha_2 L(L^{-1} v_2) = \alpha_1 v_1 + \alpha_2 v_2 = L(L^{-1}(\alpha_1 v_1 + \alpha_2 v_2))$$

kjer je $u_1 = L^{-1} v_1$ in $u_2 = L^{-1} v_2$. Ker je L injektivna preslikava, sledi

$$\alpha_1 L^{-1} v_1 + \alpha_2 L^{-1} v_2 = L^{-1}(\alpha_1 v_1 + \alpha_2 v_2)$$

Torej je L^{-1} res linearna preslikava.

Trditev 2

Kompozitum dveh linearnih izomorfizmov je spet linearni izomorfizem.

Dokaz: Zadošča dokazati, da je kompozitum dveh linearnih preslikav spet linearna preslikava. Naj bosta $L\colon U \to V$ in $K\colon V \to W$ dve linearni preslikavi. Potem velja

$$(K \circ L)(\alpha_1 u_1 + \alpha_2 u_2) = K(L(\alpha_1 u_1 + \alpha_2 u_2)) = K(\alpha_1 L u_1 + \alpha_2 L u_2) =$$
$$= \alpha_1 K(L u_1) + \alpha_2 K(L u_2) = \alpha_1 (K \circ L) u_1 + \alpha_2 (K \circ L) u_2$$

Opomba: Pravimo, da sta dva vektorska prostora (nad istim poljem) **linearno izomorfna**, če obstaja linearni izomorfizem iz prvega v drugega. Iz Trditev 1 in 2 sledi, da je linearna izomorfnost ekvivalenčna relacija. (Trditev 1 da simetričnost, Trditev 2 pa tranzitivnost. Refleksivnost sledi iz dejstva, da je identična preslikava linearni izomorfizem.)

Opomba: Iz naslednjega primera sledi, da sta katerakoli dva enako-razsežna vektorska prostora nad istim poljem linearno izomorfna.

Primer - Linearni izomorfizem iz F^n v n-razsežen vektorski prostor

Radi bi dokazali, da je vsak n-razsežen vektorski prostor nad F linearno izomorfen F^n.

Naj bo V n-razsežen vektorski prostor nad F in naj bo $\mathcal{B} = \{v_1, \ldots, v_n\}$ baza za V. Definirajmo preslikavo $\phi_{\mathcal{B}} \colon F^n \to V$ s predpisom

$$\phi_{\mathcal{B}}(x_1, \ldots, x_n) = x_1 v_1 + \ldots + x_n v_n$$

Ker je $\mathcal{B}$ ogrodje, je $\phi_{\mathcal{B}}$ surjektivna. Ker je $\mathcal{B}$ linearno neodvisna, je $\phi_{\mathcal{B}}$ injektivna. Pokažimo še, da je $\phi_{\mathcal{B}}$ linearna preslikava. To sledi iz

$$\begin{aligned}\phi_{\mathcal{B}}(\alpha(x_1, \ldots, x_n) + \beta(y_1, \ldots, y_n)) &= \phi_{\mathcal{B}}(\alpha x_1 + \beta y_1, \ldots, \alpha x_n + \beta y_n) = \\ &= (\alpha x_1 + \beta y_1)v_1 + \ldots + (\alpha x_n + \beta y_n)v_n = \alpha(x_1 v_1 + \ldots + x_n v_n) + \\ &+ \beta(y_1 v_1 + \ldots + y_n v_n) = \alpha\phi_{\mathcal{B}}(x_1, \ldots, x_n) + \beta\phi_{\mathcal{B}}(y_1, \ldots, y_n)\end{aligned}$$

Torej je $\phi_{\mathcal{B}}$ linearni izomorfizem. Njegova inverzna preslikava je $v \mapsto [v]_{\mathcal{B}}$. Po trditvi je tudi inverzna preslikava linearni izomorfizem.

Primer - Linearni izomorfizem iz $M_{m,n}(F)$ v $\mathcal{L}(F^n, F^m)$

Naj bo F polje in naj bosta m in n naravni števili. Naj bo $\mathcal{L}(F^n, F^m)$ vektorski prostor vseh linearnih preslikav iz F^n v F^m. Seštevanje je definirano z $(L_1 + L_2)u := L_1u + L_2u$, množenje s skalarjem pa z $(\alpha L)u := \alpha Lu$. Naj bo $M_{m,n}(F)$ vektorski prostor vseh $m \times n$ matrik nad F. Konstruirali bomo linearni izomorfizem iz $M_{m,n}(F)$ v $\mathcal{L}(F^n, F^m)$.

Za vsako $m \times n$ matriko $A = [a_{i,j}]$ definirajmo preslikavo L_A iz F^n v F^m

$$L_A(x_1, \ldots, x_n) = (a_{1,1}x_1 + \ldots + a_{1,n}x_n, \ldots, a_{m,1}x_1 + \ldots + a_{m,n}x_n)$$

To lahko zapišemo kot $L_Ax = Ax$, kjer x smatramo za stolpčni vektor. Iz lastnosti matričnega množenja sledi, da je L_A linearna preslikava.

Radi bi pokazali, da je preslikava $A \to L_A$ iz $M_{m,n}(F)$ v $\mathcal{L}(F^n, F^m)$ linearni izomorfizem. Linearnost sledi iz

$$L_{\alpha A+\beta B}x = (\alpha A + \beta B)x = \alpha Ax + \beta Bx = \alpha L_Ax + \beta L_Bx = (\alpha L_A + \beta L_B)x$$

Bijektivnost dokažemo tako, da konstruiramo inverzno preslikavo.

Vsaki linearni preslikavi $L\colon F^n \to F^m$ lahko priredimo $m \times n$ matriko $\begin{bmatrix} Le_1 & \dots & Le_n \end{bmatrix}$, kjer je $e_1, \dots, e_n$ standardna baza za F^n. Pokažimo, da je $L \mapsto \begin{bmatrix} Le_1 & \dots & Le_n \end{bmatrix}$ inverzna preslikava od preslikave $A \mapsto L_A$.

Najprej preverimo, da je kompozitum $A \mapsto L_A \mapsto \begin{bmatrix} L_A e_1 & \dots & L_A e_n \end{bmatrix}$ identična preslikava. Če upoštevamo, da je $\begin{bmatrix} e_1 & \dots & e_n \end{bmatrix} = I$, dobimo

$$\begin{bmatrix} L_A e_1 & \dots & L_A e_n \end{bmatrix} = \begin{bmatrix} Ae_1 & \dots & Ae_n \end{bmatrix} = A \begin{bmatrix} e_1 & \dots & e_n \end{bmatrix} = A$$

Preverimo še, da je kompozitum $L \mapsto \begin{bmatrix} Le_1 & \dots & Le_n \end{bmatrix} \mapsto L_{[Le_1 \dots Le_n]}$ identična preslikava. Za vsak $x \in F^n$ velja

$$L_{[Le_1 \dots Le_n]} x = \begin{bmatrix} Le_1 & \dots & Le_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} =$$

$$= x_1 Le_1 + \dots + x_n Le_n = L(x_1 e_1 + \dots + x_n e_n) = Lx$$

kjer smo pri prvem enačaju upoštevali definicijo L_A, pri drugem bločno množenje matrik in pri tretjem linearnost preslikave L. □

Matrika linearne preslikave

Linearni preslikavi med dvema vektorskima prostoroma bi radi priredili matriko. To prirejanje bo odvisno od izbire baz v obeh vektorskih prostorih.

Naj bosta U in V vektorska prostora nad istim poljem F in naj bo $L\colon U \to V$ linearna preslikava. Izberimo bazo $\mathcal{B} = \{u_1, \ldots, u_n\}$ za U in bazo $\mathcal{C} = \{v_1, \ldots, v_m\}$ za V. Razvijmo vektorje $Lu_1, \ldots, Lu_n$ po bazi $\mathcal{C}$:

$$\begin{aligned} Lu_1 &= \alpha_{1,1}v_1 + \ldots + \alpha_{1,m}v_m \\ &\vdots \\ Lu_n &= \alpha_{n,1}v_1 + \ldots + \alpha_{n,m}v_m \end{aligned} \tag{2}$$

Skalarje $\alpha_{i,j}$ iz (2) zložimo v matriko

$$[L]_{\mathcal{C}\leftarrow\mathcal{B}} := \begin{bmatrix} \alpha_{1,1} & \ldots & \alpha_{n,1} \\ \vdots & & \vdots \\ \alpha_{1,m} & \ldots & \alpha_{n,m} \end{bmatrix} \tag{3}$$

ki ji pravimo **matrika linearne preslikave L glede na bazi $\mathcal{B}$ in $\mathcal{C}$**.

Opomba: Bodite pozorni na to, da smo skalarje iz i-te vrstice v (2) zložili v i-ti stolpec v (3). To si najlažje zapomnimo tako, da (3) zapišemo kot

$$[L]_{\mathcal{C}\leftarrow\mathcal{B}} := \begin{bmatrix} [Lu_1]_{\mathcal{C}} & \dots & [Lu_n]_{\mathcal{C}} \end{bmatrix} \tag{4}$$

kjer je $[v]_{\mathcal{C}}$ stolpčni vektor iz koeficientov razvoja vektorja v po bazi $\mathcal{C}$.

Opomba: Do definicije matrike linearne preslikave lahko pridemo tudi preko linearnih izomorfizmov. Kompozitum linearnih preslikav

$$F^n \xrightarrow{\phi_{\mathcal{B}}} U \xrightarrow{L} V \xrightarrow{\phi_{\mathcal{C}}^{-1}} F^m$$

je linearna preslikava $\phi_{\mathcal{C}}^{-1} \circ L \circ \phi_{\mathcal{B}}$ iz F^n v F^m. Linearni izomorfizem iz $\mathcal{L}(F^n, F^m)$ v $M_{m,n}(F)$ ji priredi matriko

$$\begin{bmatrix} (\phi_{\mathcal{C}}^{-1} \circ L \circ \phi_{\mathcal{B}})(e_1) & \dots & (\phi_{\mathcal{C}}^{-1} \circ L \circ \phi_{\mathcal{B}})(e_n) \end{bmatrix} \tag{5}$$

Ker je $(\phi_{\mathcal{C}}^{-1} \circ L \circ \phi_{\mathcal{B}})(e_i) = \phi_{\mathcal{C}}^{-1}(L(\phi_{\mathcal{B}}(e_i))) = \phi_{\mathcal{C}}^{-1}(Lu_i) = [Lu_i]_{\mathcal{C}}$ za vsak $i = 1, \dots, n$, se matrika (5) ujema z matriko (4).

Primeri matrik linearnih preslikav

Primer

Naj bo U vektorski prostor vseh realnih polinomov stopnje ≤ 3 in V vektorski prostor vseh realnih polinomov stopnje ≤ 2. Za bazo prostora U vzemimo $\mathcal{B} = \{1, x, x^2, x^3\}$, za bazo prostora V pa $\mathcal{C} = \{1, x, x^2\}$. Naj bo $D\colon U \to V$ odvajanje polinomov. Iščemo matriko $[D]_{\mathcal{C}\leftarrow\mathcal{B}}$.

Razvijmo vektorje $D1, Dx, Dx^2, Dx^3$ po bazi $1, x, x^2$. Velja

$$\begin{aligned} D1 &= 0 &&= 0\cdot 1 + 0\cdot x + 0\cdot x^2 \\ Dx &= 1 &&= 1\cdot 1 + 0\cdot x + 0\cdot x^2 \\ Dx^2 &= 2x &&= 0\cdot 1 + 2\cdot x + 0\cdot x^2 \\ Dx^3 &= 3x^2 &&= 0\cdot 1 + 0\cdot x + 3\cdot x^2 \end{aligned}$$

Torej je

$$[D]_{\mathcal{C}\leftarrow\mathcal{B}} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

Primer

Če smatramo $\mathbb{C}$ za vektorski prostor nad $\mathbb{R}$, potem je konjugiranje linearna preslikava iz $\mathbb{C}$ v $\mathbb{C}$. Za bazi vzemimo $\mathcal{B} = \mathcal{C} = \{1, i\}$. Velja

$$\bar{1} = 1 = 1 \cdot 1 + 0 \cdot i$$
$$\bar{i} = -i = 0 \cdot 1 + (-1) \cdot i$$

Potem je $[L]_{\mathcal{C} \leftarrow \mathcal{B}} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$

Primer

Naj bo A $m \times n$ matrika nad F in naj bo L_A linearna preslikava iz F^n v F^m, ki je definirana z $L_A x = Ax$. Naj bo $\mathcal{B} = \{e_1, \ldots, e_n\}$ standardna baza za F^n in $\mathcal{C} = \{f_1, \ldots, f_m\}$ standardna baza za F^m. Velja

$$L_A e_1 = A e_1 = a_1 = a_{1,1} f_1 + \ldots + a_{m,1} f_m$$
$$\vdots$$
$$L_A e_n = A e_n = a_m = a_{1,n} f_1 + \ldots + a_{m,n} f_m$$

Torej je

$$[L_A]_{\mathcal{C}\leftarrow\mathcal{B}} = \begin{bmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \dots & a_{m,n} \end{bmatrix} = A$$

Primer

Naj bo V vektorski prostor in naj bo id_V identična preslikava iz V v V. Naj bosta $\mathcal{B} = \{u_1, \dots, u_n\}$ in $\mathcal{C} = \{v_1, \dots, v_n\}$ bazi za V. Razvijmo

$$\mathrm{id}_V(u_1) = u_1 = \alpha_{1,1}v_1 + \dots + \alpha_{1,n}v_n$$

$$\vdots$$

$$\mathrm{id}_V(u_n) = u_n = \alpha_{n,1}v_1 + \dots + \alpha_{n,n}v_n$$

Dobimo enake razvoje kot pri definiciji prehodne matrike. Torej je

$$[id_V]_{\mathcal{C}\leftarrow\mathcal{B}} = \begin{bmatrix} \alpha_{1,1} & \dots & \alpha_{n,1} \\ \vdots & & \vdots \\ \alpha_{1,n} & \dots & \alpha_{n,n} \end{bmatrix} = P_{\mathcal{C}\leftarrow\mathcal{B}}$$

Lastnosti matrik linearnih preslikav

Matrike linearnih preslikav imajo podobne lastnosti kot prehodne matrike.

Izrek 1 - Osnovna formula

Naj bo $L\colon U \to V$ linearna preslikava in naj bo u element U. Naj bo $\mathcal{B}$ baza za U in $\mathcal{C}$ baza za V. Potem velja $[Lu]_{\mathcal{C}} = [L]_{\mathcal{C}\leftarrow\mathcal{B}}[u]_{\mathcal{B}}$.

Dokaz: Razvijmo vektor u po bazi $\mathcal{B} = \{u_1, \ldots, u_n\}$. Dobimo

$$u = \beta_1 u_1 + \ldots + \beta_n u_n \tag{6}$$

Ker je L linearna preslikava, iz (6) sledi

$$Lu = \beta_1 Lu_1 + \ldots + \beta_n Lu_n \tag{7}$$

Za vsak $i = 1, \ldots, n$ razvijmo Lu_i po bazi $\mathcal{C} = \{v_1, \ldots, v_m\}$. Dobimo

$$Lu_i = \alpha_{i,1} v_1 + \ldots + \alpha_{i,m} v_m \tag{8}$$

Če razvoje (8) vstavimo v razvoj (7), dobimo

$$\begin{aligned} Lu &= \beta_1(\alpha_{1,1} v_1 + \ldots + \alpha_{1,m} v_m) + \ldots + \beta_n(\alpha_{n,1} v_1 + \ldots + \alpha_{n,m} v_m) \\ &= (\beta_1\alpha_{1,1} + \ldots + \beta_n\alpha_{n,1})v_1 + \ldots + (\beta_1\alpha_{1,m} + \ldots + \beta_n\alpha_{n,m})v_m \end{aligned} \tag{9}$$

Iz (9) sledi

$$[Lu]_{\mathcal{C}} = \begin{bmatrix} \beta_1\alpha_{1,1} + \ldots + \beta_n\alpha_{n,1} \\ \vdots \\ \beta_1\alpha_{1,m} + \ldots + \beta_n\alpha_{n,m} \end{bmatrix} =$$

$$= \begin{bmatrix} \alpha_{1,1} & \ldots & \alpha_{n,1} \\ \vdots & & \vdots \\ \alpha_{1,m} & \ldots & \alpha_{n,m} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_n \end{bmatrix} = [L]_{\mathcal{C}\leftarrow\mathcal{B}}[u]_{\mathcal{B}}$$

Izrek 2 - Matrika kompozituma linearnih preslikav

Naj bodo U, V in W vektorski prostori nad istem poljem in naj bosta $L\colon U \to V$ in $K\colon V \to W$ linearni preslikavi. Naj bo $\mathcal{B}$ baza za U, $\mathcal{C}$ baza za V in $\mathcal{D}$ baza za W. Potem velja $[K \circ L]_{\mathcal{D}\leftarrow\mathcal{B}} = [K]_{\mathcal{D}\leftarrow\mathcal{C}}[L]_{\mathcal{C}\leftarrow\mathcal{B}}$

Dokaz: Naj bo $\mathcal{B} = \{u_1, \ldots, u_n\}$. Po formuli (4) velja

$$[K \circ L]_{\mathcal{D}\leftarrow\mathcal{B}} = \begin{bmatrix} [(K \circ L)u_1]_{\mathcal{D}} & \ldots & [(K \circ L)u_n]_{\mathcal{D}} \end{bmatrix}$$
$$= \begin{bmatrix} [K(Lu_1)]_{\mathcal{D}} & \ldots & [K(Lu_n)]_{\mathcal{D}} \end{bmatrix} \qquad (10)$$

Po izreku 1 za vsak $i = 1, \ldots, n$ velja

$$[K(Lu_i)]_{\mathcal{D}} = [K]_{\mathcal{D}\leftarrow\mathcal{C}}[Lu_i]_{\mathcal{C}} \tag{11}$$

Če (11) vstavimo v (10), dobimo

$$[K \circ L]_{\mathcal{D}\leftarrow\mathcal{B}} = \begin{bmatrix} [K]_{\mathcal{D}\leftarrow\mathcal{C}}[Lu_1]_{\mathcal{C}} & \ldots & [K]_{\mathcal{D}\leftarrow\mathcal{C}}[Lu_n]_{\mathcal{C}} \end{bmatrix} =$$
$$= [K]_{\mathcal{D}\leftarrow\mathcal{C}} \begin{bmatrix} [Lu_1]_{\mathcal{C}} & \ldots & [Lu_n]_{\mathcal{C}} \end{bmatrix} = [K]_{\mathcal{D}\leftarrow\mathcal{C}}[L]_{\mathcal{C}\leftarrow\mathcal{B}}$$

Pri tretjem enačaju smo uporabili formulo (4). □

Izrek 3 - Prehod na novo bazo

Naj bo $L\colon U \to V$ linearna preslikava. Naj bosta $\mathcal{B}$ in $\mathcal{B}'$ bazi za U ter $\mathcal{C}$ in $\mathcal{C}'$ bazi za V. Potem velja $[L]_{\mathcal{C}'\leftarrow\mathcal{B}'} = P_{\mathcal{C}'\leftarrow\mathcal{C}}[L]_{\mathcal{C}\leftarrow\mathcal{B}}\left(P_{\mathcal{B}'\leftarrow\mathcal{B}}\right)^{-1}$

Dokaz: Najprej zapišemo L kot $\mathrm{id}_V \circ L \circ \mathrm{id}_U$. Po izreku 2 velja

$$[\mathrm{id}_V \circ L \circ \mathrm{id}_U]_{\mathcal{C}'\leftarrow\mathcal{B}'} = [id_V]_{\mathcal{C}'\leftarrow\mathcal{C}}[L]_{\mathcal{C}\leftarrow\mathcal{B}}[id_U]_{\mathcal{B}\leftarrow\mathcal{B}'}$$

Upoštevajmo še $[id_V]_{\mathcal{C}'\leftarrow\mathcal{C}} = P_{\mathcal{C}'\leftarrow\mathcal{C}}$ in $[id_U]_{\mathcal{B}\leftarrow\mathcal{B}'} = P_{\mathcal{B}\leftarrow\mathcal{B}'} = \left(P_{\mathcal{B}'\leftarrow\mathcal{B}}\right)^{-1}$.

Linearne preslikave 2. del

Ponovitev definicij

Ponovimo definiciji vektorskega podprostora in linearne preslikave.

Definicija vektorskega podprostora

Naj bo V vektorski prostor nad poljem F. Podmnožica W v V je **vektorski podprostor**, če za vsaka $w_1, w_2 \in W$ in vsaka $\alpha_1, \alpha_2 \in F$ velja $\alpha_1 w_1 + \alpha_2 w_2 \in W$

Definicija linearne preslikave

Naj bosta U in V vektorska prostora nad poljem F. Preslikava $L\colon U \to V$ je **linearna**, če za vsaka $u_1, u_2 \in U$ in vsaka $\alpha_1, \alpha_2 \in F$ velja $L(\alpha_1 u_1 + \alpha_2 u_2) = \alpha_1 L u_1 + \alpha_2 L u_2$

Vsaki linearni preslikavi bomo priredili dva podprostora: jedro in sliko.

V nadaljevanju predpostavljamo, da so vsi vektorski prostori nad istim poljem F in da so končno-razsežni.

Jedro linearne preslikave

Definicija jedra

Jedro linearne preslikave $L\colon U \to V$ je množica

$$\operatorname{Ker} L = \{u \in U \mid Lu = 0\}$$

Opomba: Angleški izraz za jedro je "kernel". Nekateri uporabljajo tudi izraz "null-space" in potem jedro označijo z $\mathcal{N}(L)$.

Trditev 1

Jedro linearne preslikave $L\colon U \to V$ je vektorski podprostor v U

Dokaz: Radi bi pokazali, da za vsaka $u_1, u_2 \in \operatorname{Ker} L$ in $\alpha_1, \alpha_2 \in F$ velja $\alpha_1 u_1 + \alpha_2 u_2 \in \operatorname{Ker} L$. Po definiciji jedra iz $u_1, u_2 \in \operatorname{Ker} L$ sledi $Lu_1 = 0$ in $Lu_2 = 0$. Po definiciji linearne preslikave odtod sledi

$$L(\alpha_1 u_1 + \alpha_2 u_2) = \alpha_1 Lu_1 + \alpha_2 Lu_2 = \alpha_1 0 + \alpha_2 0 = 0.$$

Po definiciji jedra odtod sledi, da velja $\alpha_1 u_1 + \alpha_2 u_2 \in \operatorname{Ker} L$.

Definicija ničnosti

Ničnost linearne preslikave L je število

$$n(L) = \dim \operatorname{Ker} L$$

Opomba: Angleški izraz za ničnost je "nullity".

Trditev 2

Naslednje lastnosti linearne preslikave L so ekvivalentne:

(1) $\operatorname{Ker} L = \{0\}$,

(2) $\mathrm{n}(L) = 0$,

(3) L je injektivna.

Dokaz: Ekvivalenca med (1) in (2) je očitna.
Če velja (3), potem iz $Lu = L0$ sledi $u = 0$, torej velja (1).
Recimo, da velja (1). Vzemimo taka $u_1, u_2 \in U$, da velja $Lu_1 = Lu_2$. Ker je L linearna, je $L(u_1 - u_2) = Lu_1 - Lu_2 = 0$, torej je $u_1 - u_2 \in \operatorname{Ker} L$. Iz (1) sledi, da je $u_1 - u_2 = 0$, torej je $u_1 = u_2$. Torej iz (1) sledi (3).

Slika linearne preslikave

Definicija slike

Slika linearne preslikave $L\colon U \to V$ je množica

$$\operatorname{Im} L = \{Lu \mid u \in U\}$$

Opomba: Angleški izraz za sliko je "image". Nekateri uporabljajo tudi izraz "range" in potem sliko označijo z $\mathcal{R}(L)$.

Trditev 3

Slika linearne preslikave $L\colon U \to V$ je vektorski podprostor v V

Dokaz: Radi bi pokazali, da za vsaka $v_1, v_2 \in \operatorname{Im} L$ in vsaka $\beta_1, \beta_2 \in F$ velja $\beta_1 v_1 + \beta_2 v_2 \in \operatorname{Im} L$. Po definiciji slike iz $v_1, v_2 \in \operatorname{Im} L$ sledi, da obstajata taka $u_1, u_2 \in U$, da velja $v_1 = Lu_1$ in $v_2 = Lu_2$. Odtod sledi

$$\beta_1 v_1 + \beta_2 v_2 = \beta_1 Lu_1 + \beta_2 Lu_2 = L(\beta_1 u_1 + \beta_2 u_2)$$

ker je L linearna preslikava. Po definiciji slike odtod sledi, da element $\beta_1 v_1 + \beta_2 v_2$ pripada $\operatorname{Im} L$, saj je slika elementa $\beta_1 u_1 + \beta_2 u_2$ v U.

Definicija ranga

Rang linearne preslikave L je število

$$\mathrm{r}(L) = \dim \operatorname{Im} L.$$

Direktno iz definicij dobimo:

Trditev 4

Naslednje lastnosti linearne preslikave $L\colon U \to V$ so ekvivalentne:

(1) $\operatorname{Im} L = V$,

(2) $\mathrm{r}(L) = \dim V$,

(3) L je surjektivna.

Opomba: Linearna preslikava $L\colon U \to V$ je **ničelna**, če velja $Lu = 0$ za vsak $u \in U$. Če sta U in V končno-razsežna vektorska prostora, potem so ekvivalentne trditve: (1) L je ničelna, (2) $\operatorname{Ker} L = U$, (3) $\mathrm{n}(L) = \dim U$, (4) $\operatorname{Im} L = \{0\}$, (5) $\mathrm{r}(L) = 0$.

Osnovna formula

Osnovna formula povezuje ničnost in rang.

Izrek 1: Osnovna formula

Za vsako linearno preslikavo $L\colon U \to V$ velja

$$\mathrm{n}(L) + \mathrm{r}(L) = \dim U.$$

Dokaz: Naj bo $w_1, \ldots, w_k$ baza za $\operatorname{Ker} L$ in naj bo $u_1, \ldots, u_l$ njena dopolnitev do baze U. Velja torej $\dim \operatorname{Ker} L = k$ in $\dim U = k + l$. Dokazati moramo še, da velja $\dim \operatorname{Im} L = l$.

Zadošča dokazati, da je $Lu_1, \ldots, Lu_l$ baza za $\operatorname{Im} L$. Vzemimo poljuben $v \in \operatorname{Im} L$ in izberimo tak u, da je $v = Lu$. Razvijo u po bazi za U:

$$u = \alpha_1 w_1 + \ldots + \alpha_k w_k + \beta_1 u_1 + \ldots + \beta_l u_l$$

Ker je L linearna in ker je $Lw_1 = \ldots = Lw_k = 0$, odtod sledi

$$Lu = \alpha_1 Lw_1 + \ldots + \alpha_k Lw_k + \beta_1 Lu_1 + \ldots + \beta_l Lu_l = \beta_1 Lu_1 + \ldots + \beta_l Lu_l$$

Torej so $Lu_1, \ldots, Lu_l$ ogrodje za $\operatorname{Im} L$.

Pokažimo sedaj, da so vektorji $Lu_1, \ldots, Lu_l$ linearno neodvisni. Če je $\beta_1 Lu_1 + \ldots + \beta_l Lu_l = 0$, potem $\beta_1 u_1 + \ldots + \beta_l u_l \in \operatorname{Ker} L$. Ker je $w_1, \ldots, w_k$ baza za $\operatorname{Ker} L$, obstajajo taki $\gamma_1, \ldots, \gamma_l \in F$, da je

$$\beta_1 u_1 + \ldots + \beta_l u_l = \gamma_1 w_k + \ldots + \gamma_l w_k$$

Ker je $w_1, \ldots, w_k, u_1, \ldots, u_l$ baza za U, odtod sledi $\beta_1 = \ldots = \beta_l = 0$. □

Posledica 1

Za vsako linearno preslikavo $L\colon U \to V$ obstaja taka baza $\mathcal{B}$ za U in taka baza $\mathcal{C}$ za V, da velja

$$[L]_{\mathcal{C}\leftarrow\mathcal{B}} = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}$$

kjer je I identična matrika velikosti $\mathrm{r}(L)$.

Dokaz: Naj bodo $w_1, \ldots, w_k$ in $u_1, \ldots, u_l$ kot v dokazu osnovne formule. Naj bo $z_1, \ldots, z_s$ dopolnitev $Lu_1, \ldots, Lu_l$ do baze za V. Vzemimo

$$\mathcal{B} = \{u_1, \ldots, u_l, w_1, \ldots, w_k\} \quad \text{in} \quad \mathcal{C} = \{Lu_1, \ldots, Lu_l, z_1, \ldots, z_s\}$$

Ocena za ničnost

V nadaljevanju bomo potrebovali oceno za ničnost kompozituma.

Trditev 5

Naj bosta $L\colon U \to V$ in $K\colon V \to W$ linearni preslikavi. Potem velja

$$\mathrm{n}(K \circ L) \leq \mathrm{n}(K) + \mathrm{n}(L)$$

Dokaz: Če uporabimo osnovni izrek za linearno preslikavo

$$\tilde{L}\colon \operatorname{Ker}(K \circ L) \to \operatorname{Ker} K, \quad \tilde{L}(u) = L(u)$$

dobimo $\mathrm{n}(\tilde{L}) + \mathrm{r}(\tilde{L}) = \dim \operatorname{Ker}(K \circ L)$. Ker je $\operatorname{Ker} \tilde{L} \subseteq \operatorname{Ker} L$, velja

$$\mathrm{n}(\tilde{L}) = \dim \operatorname{Ker} \tilde{L} \leq \dim \operatorname{Ker} L = \mathrm{n}(L)$$

Ker je $\operatorname{Im} \tilde{L} \subseteq \operatorname{Ker} K$, velja

$$\mathrm{r}(\tilde{L}) = \dim \operatorname{Im} \tilde{L} \leq \dim \operatorname{Ker} K = \mathrm{n}(K)$$

Jedro in slika matrike

Vsaki $m \times n$ matriki A z elementi iz F lahko priredimo linearno preslikavo

$$L_A \colon F^n \to F^m \quad L_A v = Av$$

Jedro in sliko matrike A definiramo takole

$$\operatorname{Ker} A := \operatorname{Ker} L_A \quad \text{in} \quad \operatorname{Im} A := \operatorname{Im} L_A$$

Se pravi

$$\operatorname{Ker} A = \{v \in F^n \mid Av = 0\}$$

in

$$\operatorname{Im} A = \{Av \mid v \in F^n\}.$$

Elemente F^n si obakrat predstavljamo kot stolpčne vektorje.

Trditev 6

Slika matrike je enaka linearni ogrinjači njenih stolpcev. Rang matrike je enak maksimalnemu številu linearno neodvisnih stolpcev te matrike.

Dokaz: Če so $\mathbf{a}_1, \ldots, \mathbf{a}_n$ stolpci matrike A in $\alpha_1, \ldots, \alpha_n$ komponente vektorja v, potem je $Av = \alpha_1\mathbf{a}_1 + \ldots + \alpha_n\mathbf{a}_n$. Torej je

$$\operatorname{Im} A = \{\alpha_1\mathbf{a}_1 + \ldots + \alpha_n\mathbf{a}_n \mid \alpha_1, \ldots, \alpha_n \in F\} = \operatorname{Lin}\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$$

Izberimo take stolpce $\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_r}$, ki so linearno neodvisni in ki zadoščajo $\operatorname{Lin}\{\mathbf{a}_1, \ldots, \mathbf{a}_n\} = \operatorname{Lin}\{\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_r}\}$. Potem so stolpci $\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_r}$ baza podprostora $\operatorname{Im} A$. Odtod sledi, da je rang matrike A enak r, se pravi maksimalnemu številu linearno neodvisnih stolpcev matrike A. □

Opomba: Linearni ogrinjači stolpcev matrike A pravimo **stolpčni prostor** matrike A in jo označimo z $\operatorname{Col} A$. Dokazali smo, da je $\operatorname{Im} A = \operatorname{Col} A$.

Opomba: Linearni ogrinjači vrstic matrike A pravimo **vrstični prostor** matrike A in jo označimo z $\operatorname{Row} A$. Očitno elemente $\operatorname{Row} A$ dobimo tako, da transponiramo elemente $\operatorname{Col} A^T = \operatorname{Im} A^T$.

Primer

Izračunaj sliko in rang matrike

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{bmatrix}$$

Najprej izračunamo reducirano vrstično stopničasto formo

$$R = \begin{bmatrix} 1 & 0 & -1 & -2 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Ker elementarne vrstične transformacije ohranjajo linearne relacije med stolpci, odtod sledi, da je $\mathrm{r}(A) = 2$ in

$$\operatorname{Im} A = \operatorname{Lin}\{ \begin{bmatrix} 1 \\ 5 \\ 9 \end{bmatrix}, \begin{bmatrix} 2 \\ 6 \\ 10 \end{bmatrix} \}$$

Stolpčna in vrstična ekvivalentnost matrik

Pravimo, da sta matriki A in B **stolpčno ekvivalentni**, če lahko dobimo eno iz druge s pomočjo elementarnih stolpčnih transformacij, se pravi, če obstaja taka obrnljiva matrika Q, da je $B = AQ$.

Pravimo, da sta matriki A in B **vrstično ekvivalentni**, če lahko dobimo eno iz druge s pomočjo elementarnih vrstičnih transformacij, se pravi, če obstaja taka obrnljiva matrika P, da je $B = PA$.

Trditev 7

Naj bosta A in B $m \times n$ matriki nad poljem F.

(1) Če sta A in B stolpčno ekvivalentni, potem imata enak stolpčni prostor in enako sliko. Imata tudi enak rang in enako ničnost.

(2) Če sta A in B vrstično ekvivalentni, potem imata enak vrstični prostor in enako jedro. Imata tudi enako ničnost in enak rang.

Dokaz: Če sta A in B stolpčno ekvivalentni, potem obstaja taka obrnljiva matrika Q, da je $B = AQ$. Pokažimo najprej, da velja $\operatorname{Im} AQ = \operatorname{Im} A$.

Za vsak $v \in F^n$ so ekvivalentne trditve: (a) $v \in \operatorname{Im} AQ$, (b) $v = AQu$ za nek $u \in F^n$, (c) $v = Au'$ za nek $u' \in F^n$, (d) $v \in \operatorname{Im} A$. Ekvivalentnost med (b) in (c) sledi iz obrnljivosti matrike Q.

Iz $\operatorname{Im} A = \operatorname{Im} B$ sledi $\mathrm{r}(A) = \dim \operatorname{Im} A = \dim \operatorname{Im} B = \mathrm{r}(B)$. Po osnovni formuli odtod sledi $\mathrm{n}(A) = n - \mathrm{r}(A) = n - \mathrm{r}(B) = \mathrm{n}(B)$.

Če sta matriki A in B vrstično ekvivalentni, potem obstaja taka obrnljiva matrika P, da velja $B = PA$. Pokažimo, da je $\operatorname{Ker} PA = \operatorname{Ker} A$. Za vsak $v \in F^n$ je $Av = 0$ ekvivalentno s $PAv = 0$ zaradi obrnljivosti matrike P.

Iz $\operatorname{Ker} A = \operatorname{Ker} B$ sledi $\mathrm{n}(A) = \dim \operatorname{Ker} A = \dim \operatorname{Ker} B = \mathrm{n}(B)$. Po osnovni formuli je torej $\mathrm{r}(A) = n - \mathrm{n}(A) = n - \mathrm{n}(B) = \mathrm{r}(B)$. □

Opomba: Poglejmo si, kakšna je zveza med $\operatorname{Ker} A$ in $\operatorname{Row} A$. Očitno je vektor iz F^n v $\operatorname{Ker} A$ natanko tedaj, ko je "pravokoten" na vse vrstice matrike A, se pravi, natanko tedaj, ko je "pravokoten" na $\operatorname{Row} A$. Učeno se temu reče, da je $\operatorname{Ker} A$ "ortogonalni komplement" od $\operatorname{Im} A^T = (\operatorname{Row} A)^T$.

Ekvivalentnost matrik

Definicija

Matriki A in B sta **ekvivalentni** (oznaka $A \sim B$) natanko tedaj, ko obstajata taki obrnljivi matriki P in Q, da velja $B = PAQ$.

Opomba: Dve matriki sta ekvivalentni natanko tedaj ko lahko dobimo eno iz druge s pomočjo elementarnih vrstičnih in stolpčnih transformacij.

Trditev 8

Ekvivalentnost matrik je ekvivalenčna relacija.

Dokaz: Ekvivalentnost matrik A in B označimo z $A \sim B$. Dokažimo najprej $A \sim A$. To sledi iz $A = I_m A I_n$ in obrnljivosti matrik I_m in I_n.

Če je $A \sim B$, potem je $B = PAQ$ za obrnljivi matriki P, Q. Odtod sledi $A = P^{-1}BQ^{-1}$, kjer sta tudi P^{-1}, Q^{-1} obrnljivi. Torej je $B \sim A$.

Če je $A \sim B$ in $B \sim C$, potem je $B = P_1 A Q_1$ in $C = P_2 B Q_2$ za neke obrnljive matrike P_1, Q_1, P_2, Q_2. Odtod sledi, da je $C = P_2 P_1 A Q_1 Q_2$ in da sta matriki $P_2 P_1$ in $Q_1 Q_2$ obrnljivi. Torej je $A \sim C$.

Primer

Naj bo $L\colon U \to V$ linearna preslikava, $\mathcal{B}, \mathcal{B}'$ bazi za U in $\mathcal{C}, \mathcal{C}'$ bazi za V. Potem sta matriki $[L]_{\mathcal{C}\leftarrow\mathcal{B}}$ in $[L]_{\mathcal{C}'\leftarrow\mathcal{B}'}$ ekvivalentni. Velja namreč

$$[L]_{\mathcal{C}'\leftarrow\mathcal{B}'} = P_{\mathcal{C}'\leftarrow\mathcal{C}}[L]_{\mathcal{C}\leftarrow\mathcal{B}}P_{\mathcal{B}\leftarrow\mathcal{B}'}$$

in matriki $P_{\mathcal{C}'\leftarrow\mathcal{C}}$ ter $P_{\mathcal{B}\leftarrow\mathcal{B}'}$ sta obrnljivi.

Opomba: Velja tudi obratno. Če sta matriki A in A' ekvivalentni, potem obstaja taka linearna preslikava $L\colon U \to V$ in take baze $\mathcal{B}, \mathcal{B}'$ za U in $\mathcal{C}, \mathcal{C}'$ za V, da je $[L]_{\mathcal{C}\leftarrow\mathcal{B}} = A$ in $[L]_{\mathcal{C}'\leftarrow\mathcal{B}'} = A'$.

Dokaz: Recimo, da je $A' = PAQ$, kjer sta P in Q obrnljivi matriki. Naj bo $U = F^n$ in $V = F^m$ in naj bo $L = L_A$, se pravi linearna preslikava, ki vektor množi z matriko A. Naj bo $\mathcal{B}$ standardna baza za F^n in $\mathcal{C}$ standardna baza za F^m. Potem velja $[L]_{\mathcal{C}\leftarrow\mathcal{B}} = A$. Naj bodo $\mathcal{B}'$ stolpci matrike P^{-1} in naj bodo $\mathcal{C}'$ stolpci matrike Q. Potem velja $P_{\mathcal{C}\leftarrow\mathcal{C}'} = P^{-1}$ in $P_{\mathcal{B}\leftarrow\mathcal{B}'} = Q$. Torej je $[L]_{\mathcal{C}'\leftarrow\mathcal{B}'} = P_{\mathcal{C}'\leftarrow\mathcal{C}}[L]_{\mathcal{C}\leftarrow\mathcal{B}}P_{\mathcal{B}\leftarrow\mathcal{B}'} = PAQ = A'$.

Trditev 9

Vsaka matrika ranga r je ekvivalentna matriki $\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$.

Dokaz: Naj bo A $m \times n$ matrika nad F. Po Posledici 1 obstajata taka baza $\mathcal{B}$ za F^n in taka baza $\mathcal{C}$ za F^m, da velja

$$[L_A]_{\mathcal{C} \leftarrow \mathcal{B}} = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$$

Naj bosta $\mathcal{S}_n$ in $\mathcal{S}_m$ standardni bazi za F^n in F^m. Potem je

$$A = [L_A]_{\mathcal{S}_m \leftarrow \mathcal{S}_m} = P_{\mathcal{S}_m \leftarrow \mathcal{C}} [L_A]_{\mathcal{C} \leftarrow \mathcal{B}} P_{\mathcal{B} \leftarrow \mathcal{S}_n} = P_{\mathcal{S}_m \leftarrow \mathcal{C}} \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} P_{\mathcal{B} \leftarrow \mathcal{S}_n}$$

kjer sta $P_{\mathcal{S}_m \leftarrow \mathcal{C}}$ in $P_{\mathcal{B} \leftarrow \mathcal{S}_n}$ obrnljivi matriki. Torej je $A \sim \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$. □

Opomba: Bolj računski dokaz poteka takole.
Najprej matriko A s pomočjo elementarnih vrstičnih transformacij prevedemo v reducirano vrstično stopničasto formo R. Potem matriko R s pomočjo elementarnih stolpčnih transformacij prevedemo v želeno obliko.

Izrek 2 - Karakterizacija ekvivalentnosti matrik

Dve matriki sta ekvivalentni natanko tedaj, ko imata enako velikost in enak rang.

Dokaz: Recimo, da sta A in B ekvivalentni matriki. Potem obstajata taki obrnljivi matriki P in Q, da velja $B = PAQ$.

Ker so vse obrnljive matrike kvadratne, je matrika PAQ enake velikosti kot matrika A. Torej sta A in B enake velikosti.

Po Trditvi 7 velja $\mathrm{r}(PAQ) = \mathrm{r}(PA) = \mathrm{r}(A)$. Torej imata A in B enak rang.

Recimo sedaj, da imata matriki A in B enako velikost $m \times n$ in enak rang r. Po Trditvi 9 sta tako A kot B ekvivalentni $m \times n$ matriki $\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$.

Odtod po Trditvi 8 sledi, da sta A in B ekvivalentni. □

Enakost stolpčnega in vrstičnega ranga

Rangu matrike A pravimo tudi **stolpčni rang** matrike A, ker je enak maksimalnemu številu linearno neodvisnih stolpcev matrike A.

Rangu matrike A^T pravimo tudi **vrstični rang** matrike A, ker je enak maksimalnemu številu linearno neodvisnih vrstic matrike A.

Radi bi pokazali, da sta stolčni in vrstični rang matrike A enaka.

Izrek 3

Za vsako matriko A je $\mathrm{r}(A) = \mathrm{r}(A^T)$.

Dokaz: Če je A $m \times n$ matrika ranga r, potem po Trditvi 9 obstajata taki obrnljivi matriki P in Q, da velja

$$A = P \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} Q$$

Odtod sledi

$$A^T = Q^T \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}^T P^T.$$

Ker sta P in Q obrnljivi matriki, sta tudi P^T in Q^T obrnljivi.
Po Trditvi 7 odtod sledi

$$\mathrm{r}(A^T) = \mathrm{r}(Q^T \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}^T) = \mathrm{r}(\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}^T) = r$$

Torej je

$$\mathrm{r}(A) = r = \mathrm{r}(A^T)$$

Opomba: Matrika $\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$ je velikosti $m \times n$, matrika $\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}^T$ pa je velikosti $n \times m$. Imata pa obe enak rang r.

Linearne preslikave 3. del

Podobnost kvadratnih matrik

Definicija podobnosti matrik

Kvadratni matriki A in B sta **podobni**, če obstaja taka obrnljiva matrika P, da velja $B = PAP^{-1}$.

Trditev

Podobnost matrik je ekvivalenčna relacija.

Dokaz: Ker je $A = IAI^{-1}$, kjer je I identična matrika, je A podobna A.

Če je matrika A podobna matriki B, potem je $B = PAP^{-1}$ za neko obrnljivo matriko P. Odtod sledi, da je $A = P^{-1}BP = P^{-1}B(P^{-1})^{-1}$ in da je matrika P^{-1} obrnljiva. Torej je tudi matrika B podobna matriki A.

Če je matrika A podobna matriki B in če je matrika B podobna matriki C, potem velja $B = PAP^{-1}$ in $C = QBQ^{-1}$ za obrnljivi matriki P in Q. Odtod sledi $C = QPAP^{-1}Q^{-1} = (QP)A(QP)^{-1}$ in da je matrika QP obrnljiva. Torej je matrika A podobna matriki C. □

Opomba: Iz podobnosti sledi ekvivalentnost matrik, obratno pa ni res.

Primer: Podobnost ni ekvivalentnost

Matriki $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ in $B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ sta ekvivalentni, ker imata enako velikost in enak rang. Če bi bili matriki A in B podobni, potem bi obstajala taka obrnljiva matrika P, da bi veljalo $B = PAP^{-1}$. Odtod bi sledilo

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x & y \\ u & v \end{bmatrix} = \begin{bmatrix} x & y \\ u & v \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{kjer} \quad P = \begin{bmatrix} x & y \\ u & v \end{bmatrix}$$

Z množenjem matrik bi dobili

$$\begin{bmatrix} u & v \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} x & 0 \\ u & 0 \end{bmatrix}$$

Odtod pa bi sledilo $u = v = 0$. To bi bilo protislovje s predpostavko, da je matrika P obrnljiva.

Primer podobnih matrik

Naj bo $L\colon V \to V$ linearna preslikava in naj bosta $\mathcal{B}$ in $\mathcal{B}'$ bazi vektorskega prostora V. Potem sta matriki $[L]_{\mathcal{B}\leftarrow\mathcal{B}}$ in $[L]_{\mathcal{B}'\leftarrow\mathcal{B}'}$ podobni, ker velja

$$[L]_{\mathcal{B}'\leftarrow\mathcal{B}'} = P_{\mathcal{B}'\leftarrow\mathcal{B}}[L]_{\mathcal{B}\leftarrow\mathcal{B}}P_{\mathcal{B}\leftarrow\mathcal{B}'} = P_{\mathcal{B}'\leftarrow\mathcal{B}}[L]_{\mathcal{B}\leftarrow\mathcal{B}}\left(P_{\mathcal{B}'\leftarrow\mathcal{B}}\right)^{-1}$$

kjer je $P_{\mathcal{B}'\leftarrow\mathcal{B}}$ obrnljiva matrika.

Opomba: Velja tudi obratno. Če sta kvadratni matriki A in A' podobni, potem obstaja taka linearna preslikava $L\colon V \to V$ in taki bazi $\mathcal{B}$ in $\mathcal{B}'$ za V, da velja $[L]_{\mathcal{B}\leftarrow\mathcal{B}} = A$ in $[L]_{\mathcal{B}'\leftarrow\mathcal{B}'} = A'$.

Trditev

Podobni matriki imata enako determinanto.

Dokaz: $\det PAP^{-1} = \det P \det A \det P^{-1} = \det A \det PP^{-1} = \det A$.

Opomba: Odtod sledi, da je za vsako linearno preslikavo $L\colon V \to V$ vrednost izraza $\det[L]_{\mathcal{B}\leftarrow\mathcal{B}}$ neodvisna od izbire baze $\mathcal{B}$. To vrednost vzamemo za definicijo $\det L$.

Primer

Naj bo A kvadratna $n \times n$ matrika nad F. Recimo, da obstaja tak vektor $v \in F^n$, da so vektorji $v, Av, \ldots, A^{n-1}v$ linearno neodvisni (=baza). Razvijmo vektor $A^n v$ po tej bazi: $A^n v = \alpha_0 v + \alpha_1 Av + \ldots + \alpha_{n-1} A^{n-1} v$.

Trdimo, da je $A = PBP^{-1}$, kjer je

$$B = \begin{bmatrix} 0 & 0 & \ldots & 0 & \alpha_0 \\ 1 & 0 & \ldots & 0 & \alpha_1 \\ 0 & 1 & \ldots & 0 & \alpha_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & \alpha_{n-1} \end{bmatrix}$$

in je $P = \begin{bmatrix} v & Av & \ldots & A^{n-2}v & A^{n-1}v \end{bmatrix}$ obrnljiva matrika. To sledi iz

$$AP = \begin{bmatrix} Av & A^2 v & \ldots & A^{n-1}v & A^n v \end{bmatrix} =$$

$$= \begin{bmatrix} Av & A^2 v & \ldots & A^{n-1}v & \alpha_0 v + \alpha_1 Av + \ldots + \alpha_{n-1} A^{n-1} v \end{bmatrix} = PB$$

Definicija: B je **pridružena matrika** polinoma $p(x) = x^n - \sum_{i=0}^{n-1} \alpha_i x^i$.

Kaj želimo?

Glavni cilj

Za dano kvadratno matriko A iščemo čimpreprostejšo matriko (=matriko z veliko ničlami), ki ji je podobna.

Ekvivalentno: Za dano linearno preslikavo $L\colon V \to V$ iščemo tako bazo $\mathcal{B}$ za V, da bo matrika $[L]_{\mathcal{B}\leftarrow\mathcal{B}}$ čimpreprostejša.

Opomba: Morali se bomo omejiti na kompleksne matrike.

Preproste matrike so recimo:

- diagonalne matrike (Ni vsaka k. matrika podobna diagonalni.)
- zgornje trikotne matrike (Vsaka k. matrika je podobna zg. trikotni.)
- Jordanske kanonične forme (Zelo posebni primeri zgornje trikotnih. Vsaka kompleksna matrika je podobna Jordanski kanonični formi.)
- Frobeniusove kanonične forme (Bločno diagonalne matrike iz pridruženih matrik polinomov. Vsaka matrika je podobna taki.)

Invariantni podprostori

Definicija

Naj bo V vektorski prostor nad F in naj bo $L\colon V \to V$ linearna preslikava. Vektorski podprostor W v V je **invarianten** za L, če za vsak $w \in W$ velja $Lw \in W$.

Opomba: Kadar govorimo o invariantnih podprostorih matrike $A \in M_n(F)$ imamo v mislih invariantne podprostore pripadajoče linearne preslikave $L_A\colon F^n \to F^n$, $v \mapsto Av$.

Primeri invariantnih podprostorov

Naj bo $A \in M_n(F)$ in naj bo $\mathbf{e}_1, \ldots, \mathbf{e}_n$ standardna baza za F^n.

Če je A zgornje trikotna, potem so naslednji podprostori invariantni za A:

$$\mathrm{Lin}\{\mathbf{e}_1\}, \mathrm{Lin}\{\mathbf{e}_1, \mathbf{e}_2\}, \ldots, \mathrm{Lin}\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$$

Če je A diagonalna, potem so naslednji podprostori invariantni za A:

$$\mathrm{Lin}\{\mathbf{e}_1\}, \mathrm{Lin}\{\mathbf{e}_2\}, \ldots, \mathrm{Lin}\{\mathbf{e}_n\}$$

Primer invariantnega podprostora

Naj bo $V = F[x]$ vektorski prostor vseh polinomov s koeficienti v F in naj bo $D\colon V \to V$ odvajanje polinomov. Za vsako naravno število n je $F[x]_{\leq n}$ (se pravi vektorski podprostor vseh polinomov v $F[x]$, ki so stopnje $\leq n$) invarianten podprostor za D.

$\operatorname{Ker} L$ in $\operatorname{Im} L$ sta invariantna podprostora za L

Naj bo $L\colon V \to V$ linearna preslikava. Očitno L pošlje elemente $\operatorname{Ker} L$ v 0. Ker je $0 \in \operatorname{Ker} L$, torej L pošlje vse elemente $\operatorname{Ker} L$ v elemente $\operatorname{Ker} L$. Torej je $\operatorname{Ker} L$ invarianten podprostor za L.

Očitno L pošlje vse elemente V v elemente $\operatorname{Im} L$. Torej pošlje tudi vse elemente $\operatorname{Im} L$ v elemente $\operatorname{Im} L$. To pomeni, da je $\operatorname{Im} L$ invarianten podprostor za L.

Oglejmo si zdaj posplošitev zadnjega primera na polinome v L.

Naj bo V vektorski prostor nad F. Za vsak polinom

$$p(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_k x^k \in F[x].$$

in vsako linearno preslikavo $L\colon V \to V$ definirajmo

$$p(L) = a_0 \mathrm{id}_V + a_1 L + a_2 L^2 + \ldots + a_k L^k$$

kjer je $L^2 = L \circ L$, $L^3 = L \circ L \circ L$, itd. Očitno je $p(L)$ linearna preslikava iz V v V, ki pošlje vektor $v \in V$ v $p(L)v = a_0 v + a_1 Lv + \ldots + a_k L^k v$.

Opomba: Podobno za vsako matriko $A \in M_n(F)$ definiramo

$$p(A) = a_0 I_n + a_1 A + a_2 A^2 + \ldots + a_k A^k$$

Trditev

Naj bo V vektorski prostor nad F in naj bo $L\colon V \to V$ linearna preslikava. Za vsaka polinoma $p(x), q(x) \in F[x]$ sta podprostora $\operatorname{Ker} q(L)$ in $\operatorname{Im} q(L)$ invariantna za $p(L)$.

Dokaz: V dokazu bomo uporabili, da je

$$p(L) \circ q(L) = q(L) \circ p(L)$$

kar sledi iz $L^r \circ L^s = L^{r+s} = L^s \circ L^r$ z direktnim računom.

Če je $v \in \operatorname{Ker} q(L)$, potem je $q(L)v = 0$. Odtod sledi $p(L)(q(L)v) = 0$. Po prvem odstavku odtod sledi $q(L)(p(L)v) = 0$. Torej je $p(L)v \in \operatorname{Ker} q(L)$.

Če je $v \in \operatorname{Im} q(L)$, potem obstaja tak $u \in V$, da je $v = q(L)u$. Odtod sledi $p(L)v = p(L)(q(L)v) = q(L)(p(L)v) = q(L)u'$, kjer je $u' = p(L)v$ element V. Odtod sledi, da je $p(L)v \in \operatorname{Im} q(L)$. □

Invariantni podprostori in glavni cilj

Če ima linearna preslikava invarianten podprostor, potem ji lahko priredimo matriko z veliko ničlami.

Trditev

Naj bo $L\colon V \to V$ linearna preslikava in naj bo W invarianten podprostor za L. Izberimo bazo $w_1, \ldots, w_k$ za W in jo dopolnimo do baze $\mathcal{B}$ za V. Potem je $[L]_{\mathcal{B}\leftarrow\mathcal{B}}$ bločno zgornje trikotna matrike, se pravi

$$[L]_{\mathcal{B}\leftarrow\mathcal{B}} = \begin{bmatrix} A & B \\ 0 & C \end{bmatrix}$$

za neko $k \times k$ matriko A in neki matriki B in C.

Dokaz: Naj bodo $v_1, \ldots, v_l \in V$ taki vektorji, da je $w_1, \ldots, w_k, v_1, \ldots, v_l$ baza za V. Označimo to bazo z $\mathcal{B}$. Ker je W invarianten podprostor za L, so $Lw_1, \ldots, Lw_k$ v W. Torej lahko $Lw_1, \ldots, Lw_k$ razvijemo po $w_1, \ldots, w_k$:

$$Lw_1 = \alpha_{1,1}w_1 + \ldots + \alpha_{1,k}w_k$$

$$\vdots$$

$$Lw_k = \alpha_{k,1}w_1 + \ldots + \alpha_{k,k}w_k$$

Vektorje $Lv_1, \ldots, Lv_l$ razvijmo po bazi $w_1, \ldots, w_k, v_1, \ldots, v_l$

$$Lv_1 = \beta_{1,1}w_1 + \ldots + \beta_{1,k}w_k + \gamma_{1,1}v_1 + \ldots + \gamma_{1,l}v_l$$

$$\vdots$$

$$Lv_l = \beta_{l,1}w_1 + \ldots + \beta_{l,k}w_k + \gamma_{l,1}v_1 + \ldots + \gamma_{l,l}v_l$$

Označimo $\mathcal{B} = \{w_1, \ldots, w_k, v_1, \ldots, v_l\}$. Potem je

$$[L]_{\mathcal{B}\leftarrow\mathcal{B}} = \begin{bmatrix} \alpha_{1,1} & \ldots & \alpha_{k,1} & \beta_{1,1} & \ldots & \beta_{l,1} \\ \vdots & & \vdots & \vdots & & \vdots \\ \alpha_{1,k} & \ldots & \alpha_{k,k} & \beta_{1,k} & \ldots & \beta_{l,k} \\ 0 & \ldots & 0 & \gamma_{1,1} & \ldots & \gamma_{l,1} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \ldots & 0 & \gamma_{1,l} & \ldots & \gamma_{l,l} \end{bmatrix}$$

Direktna vsota invariantnih podprostorov

V nadaljevanju bomo naleteli na naslednjo situacijo:

Trditev

Naj bo vektorski prostor V direktna vsota podprostorov $W_1, \ldots, W_k$. Naj bo $L\colon V \to V$ taka linearna preslikava, da so $W_1, \ldots, W_k$ invariantni za L. Naj bo $\mathcal{B}_i$ baza za W_i za $i = 1, \ldots, k$ in naj bo $\mathcal{B} = \mathcal{B}_1 \cup \ldots \cup \mathcal{B}_k$. Potem obstajo take matrike $A_1, \ldots, A_k$ velikosti $\dim W_1, \ldots, \dim W_k$, da velja

$$[L]_{\mathcal{B}\leftarrow\mathcal{B}} = \begin{bmatrix} A_1 & 0 & \ldots & 0 \\ 0 & A_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & A_k \end{bmatrix}$$

Dokaz je podoben kot pri prejšnji trditvi.

Enorazsežni invariantni podprostori

Pokažimo, da ima vsaka linearna preslikava $L\colon V \to V$, kjer je V netrivialen vektorski prostor nad $\mathbb{C}$, enorazsežen invarianten podprostor.

Naslednje trditve so očitno ekvivalentne:

(1) L ima enorazsežen invarianten podprostor.

(2) Obstaja tak neničeln $v \in V$ in tak $\lambda \in \mathbb{C}$, da je $Lv = \lambda v$.

(3) Obstaja tak neničeln $v \in V$ in tak $\lambda \in \mathbb{C}$, da je $(L - \lambda \,\mathrm{id})v = 0$.

(4) Obstaja tak $\lambda \in \mathbb{C}$, da je $\mathrm{Ker}\,(L - \lambda \,\mathrm{id}) \neq \{0\}$.

(5) Obstaja tak $\lambda \in \mathbb{C}$, da linearna preslikava $L - \lambda \,\mathrm{id}$ ni injektivna.

(6) Obstaja tak $\lambda \in \mathbb{C}$, da linearna preslikava $L - \lambda \,\mathrm{id}$ ni obrnljiva.

(7) Obstaja tak $\lambda \in \mathbb{C}$, da velja $\det(L - \lambda \,\mathrm{id}) = 0$.

Ker smo v kompleksnih številih, ima polinomska enačba

$$\det(L - \lambda \,\mathrm{id}) = 0$$

vsaj eno netrivialno rešitev λ.

Kongruentnost matrik

Poleg ekvivalentnosti matrik in podobnosti matrik obstajajo še druge ekvivalenčne relacije na matrikah. Čisto na koncu semestra bomo pri teoriji kvadratnih form srečali še kongruentnost matrik.

Definicija kongruentnosti matrik

Pravimo, da sta kvadratni matriki A in B **kongruentni**, če obstaja taka obrnljiva matrika P, da velja $B = PAP^T$.

Opomba: Če sta dve matriki kongruentni, potem sta očitno tudi ekvivalentni. Obratno ni res. Recimo matriki $[1]$ in $[-1]$ sta ekvivalentni, nista pa kongruentni. Imamo torej naslednjo situacijo:

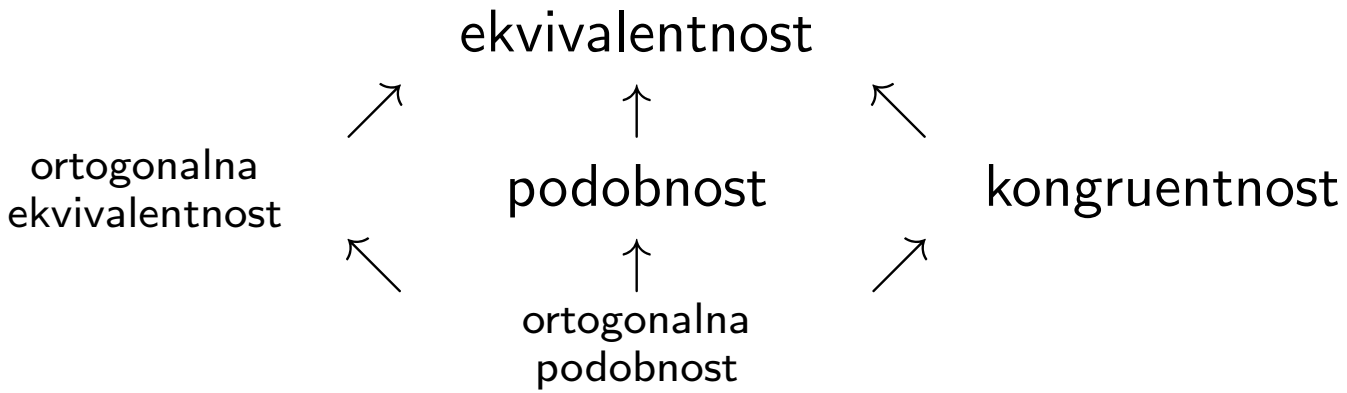
Če v definiciji ekvivalentnosti matrik in podobnosti matrik zamenjamo obrnljive matrike z ortogonalnimi matrikami, potem dobimo definiciji ortogonalne ekvivalentnosti matrik in ortogonalne podobnosti matrik.

Definicije

Kvadratna matrika P je **ortogonalna**, če velja $PP^T = I$.

Matriki A in B sta **ortogonalno ekvivalentni**, če obstajata taki ortogonalni matriki P in Q, da velja $B = PAQ^{-1}$ ($\Leftrightarrow B = PAQ^T$).

Kvadratni matriki A in B sta **ortogonalno podobni**, če obstaja taka ortogonalna matrika P, da velja $B = PAP^{-1}$ ($\Leftrightarrow B = PAP^T$).



Lastne vrednosti - 1. del

1. Lastni problem

Za dano kvadratno matriko iščemo čimpreprostejšo matriko (npr. diagonalno ali zgornje trikotno matriko), ki ji je podobna.

Vemo, da pri tem zelo pomaga, če ima matrika invarianten podprostor. Vprašajmo se, kdaj ima matrika enorazsežen invarianten podprostor.

Trditev 1

Za vsako matriko $A \in M_n(F)$ sta ekvivalentni trditvi:

(1) Obstaja enorazsežen podprostor v F^n, ki je invarianten za A.

(2) Obstaja tak skalar $\lambda \in F$ in tak neničeln vektor $v \in F^n$, da velja $Av = \lambda v$.

Dokaz: Če velja (2), potem je očitno $\mathrm{Lin}\{v\}$ enorazsežen invarianten podprostor za A. Torej velja (1).

Če velja (1), potem obstaja tak neničeln vektor $v \in F^n$, da je podprostor $\mathrm{Lin}\{v\}$ invarianten za A. Ker je $Av \in \mathrm{Lin}\{v\}$, obstaja tak skalar $\lambda \in F$, da velja $Av = \lambda v$. Torej velja (2). □

Trditev 1 služi kot motivacija za študij enačbe

$$Av = \lambda v$$

kjer je $A \in M_n(F)$ dana matrika, iščemo pa skalar $\lambda \in F$ in vektor $v \in F^n$.

Enačbi $Av = \lambda v$ rečemo **lastni problem** za matriko A. Trivialna rešitev lastnega problema je $v = 0$ in $\lambda =$ karkoli. Ta rešitev za nas ni zanimiva. Če je (λ, v) netrivialna rešitev lastnega problema, potem pravimo, da je λ **lastna vrednost** matrike A, v pa **lastni vektor** matrike A. Bolj natančno:

Definicija lastne vrednosti in lastnega vektorja

Skalar $\lambda \in F$ je **lastna vrednost** matrike $A \in M_n(F)$, če obstaja tak neničeln vektor $v \in F^n$, da velja $Av = \lambda v$. Vsakemu takemu vektorju v pravimo **lastni vektor** matrike A, ki pripada lastni vrednosti λ.

Pozor: Vsak lastni vektor matrike A je po definiciji neničeln vektor.

Posledica: Iz Trditve 1 sledi, da ima matrika A enorazsežen invarianten podprostor natanko tedaj, ko ima kako lastno vrednost.

Pri iskanju lastnih vrednosti in lastnih vektorjev matrike A si pomagamo z naslednjo trditvijo:

Trditev 2

Za vsako matriko $A \in M_n(F)$ in vsak skalar $\lambda \in F$ so ekvivalentne naslednje trditve:

(1) λ je lastna vrednost matrike A.

(2) $\mathrm{Ker}\,(A - \lambda I) \neq \{0\}$.

(3) Matrika $A - \lambda I$ ni obrnljiva.

(4) $\det(A - \lambda I) = 0$.

Dokaz: Enačbo $Av = \lambda v$ lahko zapišemo v obliki $(A - \lambda I)v = 0$. Velja namreč, da je $(A - \lambda I)v = Av - \lambda Iv = Av - \lambda v$. Odtod sledi, da je točka (1) ekvivalentna s točko (2).

Ekvivalentnost točk (2), (3) in (4) sledi iz karakterizacij obrnljivih matrik. Matrika $A - \lambda I$ je namreč obrnljiva natanko tedaj, ko je $\det(A - \lambda I) \neq 0$, in natanko tedaj, ko so stolpci matrike $A - \lambda I$ linearno neodvisni. □

Karakteristični polinom

Prvi korak pri reševanju lastnega problema za matriko A je izračun karakterističnega polinoma za matriko A.

Definicija karakterističnega polinoma

Karakteristični polinom matrike $A \in M_n(F)$ je polinom $p_A(x) = \det(A - xI)$. (I je identična matrika velikosti n.)

S pomočjo karakterističnega polinoma potem izračunamo vse lastne vrednosti matrike. Velja namreč:

Trditev 3

Skalar $\lambda \in F$ je lastna vrednost matrike $A \in M_n(F)$ natanko tedaj, ko je λ ničla karakterističnega polinoma matrike A.

Dokaz: To je ekvivalenca med (1) in (4) v Trditvi 2.

Dogovor: V nadaljevanju se bomo omejili na primer $F = \mathbb{C}$.

Razlog: Osnovni izrek algebre pravi, da ima vsak nekonstanten polinom s kompleksnimi koeficienti vsaj eno kompleksno ničlo. Odtod sledi, da ima vsaka kompleksna kvadratna matrika vsaj eno lastno vrednost.

Karakteristični polinom lahko razcepimo na linearne faktorje:

$$p_A(x) = (-1)^n(x - \lambda_1)^{n_1} \cdots (x - \lambda_k)^{n_k}$$

kjer so $\lambda_1, \ldots, \lambda_k$ vse paroma različne lastne vrednosti matrike A. Naravna števila $n_1, \ldots, n_k$ očitno zadoščajo $n_1 + \ldots + n_k = n$. Pravimo jim **algebraične večkratnosti** lastnih vrednosti $\lambda_1, \ldots, \lambda_k$. Bolj natančno:

Definicija algebraične večkratnosti

Če je lastna vrednost λ m-kratna ničla karakterističnega polinoma, potem pravimo, da je njena **algebraična večkratnost** enaka m.

Primer: 1 je lastna vrednost matrike I_2. Njena algebraična večkratnost je 2.

Karakteristični polinom 2×2 matrike

Karakteristični polinom matrike $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ je

$$p_A(x) = \det \begin{bmatrix} a-x & b \\ c & d-x \end{bmatrix} = (a-x)(d-x) - bc =$$

$$= x^2 - (a+d)x + ad - bc = x^2 - (\operatorname{sled} A)x + \det A$$

Opomba: Podobno izpeljemo, da je karakteristični polinom $n \times n$ matrike $A = [a_{i,j}]$ oblike $p_A(x) = c_n x^n + c_{n-1}x^{n-1} + \ldots + c_1 x + c_0$ kjer je

$$c_n = (-1)^n, \quad c_{n-1} = (-1)^{n-1} \sum_i a_{i,i} \quad \text{in} \quad c_0 = \det A$$

Ostali koeficienti so bolj kompliciran, npr.

$$c_{n-2} = (-1)^{n-2} \sum_{i<j} \det \begin{bmatrix} a_{i,i} & a_{i,j} \\ a_{j,i} & a_{j,j} \end{bmatrix}$$

Karakteristični polinom zgornje trikotne matrike

Če je

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ 0 & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{n,n} \end{bmatrix}$$

potem je $\det(A - xI) = (-1)^n(x - a_{1,1})(x - a_{2,2}) \dots (x - a_{n,n})$

Dokaz: Determinanta gornjetrikotne matrike je produkt njenih diagonalnih elementov, torej je $\det(A - xI) = (a_{1,1} - x)(a_{2,2} - x) \dots (a_{n,n} - x)$.

Trditev 4

Če sta matriki A in B podobni, potem je $\det(A - xI) = \det(B - xI)$. Torej imata A in B enake lastne vrednosti z enakimi algebraičnimi večkratnostmi.

Dokaz: Če je $A = PBP^{-1}$, potem je $A - xI = P(B - xI)P^{-1}$. Torej je $\det(A - xI) = \det P \det(B - xI) \det P^{-1} = \det P \det P^{-1} \det(B - xI) = \det(PP^{-1}) \det(B - xI) = \det I \det(B - xI) = \det(B - xI)$.

Lastni podprostori

Množica vseh lastnih vektorjev matrike A, ki pripadajo lastni vrednosti λ, je enaka $\mathrm{Ker}(A - \lambda I) \setminus \{0\}$. Ta množica je vedno neskončna, ker je vsak neničeln večkratnik vsakega lastnega vektorja spet lastni vektor.

Vektor 0 ni nikoli lastni vektor. Če ga dodamo k množici lastnih vektorjev, dobimo množico $\mathrm{Ker}(A - \lambda I)$, ki je vektorski podprostor v $\mathbb{C}^n$.

Definicija lastnega podprostora in geometrijske večkratnosti

Če je λ lastna vrednost matrike A, potem vektorskemu podprostoru $\mathrm{Ker}(A - \lambda I)$ pravimo **lastni podprostor** matrike A za lastno vrednost λ, njegovi dimenziji pa **geometrijska večkratnost** lastne vrednosti λ.

Opomba: Lastne podprostore matrike A poiščemo tako, da za vsako njeno lastno vrednost λ rešimo homogen sistem linearnih enačb $(A - \lambda I)v = 0$, kjer smatramo, da so komponente vektorja v spremenljivke.

Primer

Določi vse lastne vrednosti in lastne podprostore matrike $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$.

Rešitev: Karakteristični polinom je $p_A(x) = \det(A - xI) = x^2 + 1$.
Ima dve kompleksni ničli $\lambda_1 = i$ in $\lambda_2 = -i$. Lastna podprostora sta

$$\mathrm{Ker}(A - \lambda_1 I) = \mathrm{Ker} \begin{bmatrix} -i & 1 \\ -1 & -i \end{bmatrix} = \mathrm{Lin}\{\begin{bmatrix} 1 \\ i \end{bmatrix}\}$$

$$\mathrm{Ker}(A - \lambda_2 I) = \mathrm{Ker} \begin{bmatrix} i & 1 \\ -1 & i \end{bmatrix} = \mathrm{Lin}\{\begin{bmatrix} 1 \\ -i \end{bmatrix}\}$$

Opomba: Ker sta oba lastna podprostora enorazsežna, imata tako λ_1 kot λ_2 geometrijsko večkratnost 1. Poleg tega imata tako λ_1 kot λ_2 algebraično večkratnost enako 1 saj sta obe enostavni ničli $p_A(x)$.

Primer

Določi vse lastne vrednosti in lastne podprostore matrik

$$A_1 = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$$

Rešitev: Vse tri matrike imajo enak karakteristični polinom, namreč

$$(2-x)^3 = -(x-2)^3.$$

Torej je pri vseh 2 edina lastna vrednost in ima algebraično večkratnost 3. V prvem primeru je lastni podprostor lastne vrednosti 2 enak

$$\mathrm{Ker}\,(A_1 - 2I) = \mathrm{Lin}\{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\} = \mathbb{C}^3$$

torej je geometrijska večkratnost lastne vrednosti 2 enaka 3.

V drugem primeru je lastni podprostor lastne vrednosti 2 enak

$$\mathrm{Ker}\,(A_2 - 2I) = \mathrm{Lin}\{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\}$$

torej je geometrijska večkratnost lastne vrednosti 2 enaka 2.

V tretjem primeru je lastni podprostor lastne vrednosti 2 enak

$$\mathrm{Ker}\,(A_3 - 2I) = \mathrm{Lin}\{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\}$$

torej je geometrijska večkratnost lastne vrednosti 2 enaka 1. □

Naj bosta A in B podobni matriki in naj bo λ lastna vrednost za A. Po Trditvi 4 je λ tudi lastna vrednost za B. Poglejmo kakšna je zveza med lastnima podprostoroma $\mathrm{Ker}\,(A-\lambda I)$ in $\mathrm{Ker}\,(B-\lambda I)$.

Naj bo P taka obrnljiva matrika, da velja $B = PAP^{-1}$. Potem velja

$$\begin{aligned}
\mathrm{Ker}\,(B-\lambda I) &= \mathrm{Ker}\,P(A-\lambda I)P^{-1} \\
&= \{v \in F^n \mid P(A-\lambda I)P^{-1}v = 0\} \\
&= \{v \in F^n \mid (A-\lambda I)P^{-1}v = 0\} \\
&= \{Pw \mid w \in F^n, (A-\lambda I)w = 0\} \\
&= \{Pw \mid w \in \mathrm{Ker}\,(A-\lambda I)\} \\
&= P\,\mathrm{Ker}\,(A-\lambda I)
\end{aligned}$$

Odtod sledi, da podprostora $\mathrm{Ker}\,(A-\lambda I)$ in $\mathrm{Ker}\,(B-\lambda I)$ nista nujno enaka. Sta pa enaki njuni dimenziji.

Velja namreč

$$\dim \operatorname{Ker}(B-\lambda I) = \mathrm{n}(B-\lambda I) = \mathrm{n}(P(A-\lambda I)P^{-1}) =$$

$$= \mathrm{n}(P(A-\lambda I)) = \mathrm{n}(A-\lambda I) = \dim \operatorname{Ker}(A-\lambda I)$$

ker se ničnost ohranja pri vrstični in stolpčni ekvivalenci matrik.

Dokazali smo torej, da imata podobni matriki enake geometrijske večkratnosti lastnih vrednosti. Natančneje:

Trditev 5

Naj bosta A in B podobni matriki in naj bo λ lastna vrednost za A z geometrijsko večratnost m. Potem je λ tudi lastna vrednost za B z geometrijsko večkratnostjo m.

Opomba: Ker so pri diagonalni matriki geometrijske večkratnosti lastnih vrednosti enake algebraičnim večkratnostim, je to res tudi za vse matrike, ki so podobne diagonalnim. Odtod sledi, da matriki A_2 in A_3 iz zadnjega primera nista podobni diagonalnim matrikam.

Diagonalizacija matrik

S pomočjo lastnih vrednosti in lastnih vektorjev matrike A lahko včasih poiščemo diagonalno matriko, ki je podobna matriki A.

Recimo, da ima matrika A n linearno neodvisnih vektorjev $v_1, \ldots, v_n$ in naj bodo $\lambda_1, \ldots, \lambda_n$ pripadajoče lastne vrednosti, se pravi

$$Av_1 = \lambda_1 v_1, \ldots, Av_n = \lambda_n v_n$$

Odtod sledi, da je matrika

$$P = \begin{bmatrix} v_1 & \ldots & v_n \end{bmatrix}$$

obrnljiva in velja

$$AP = \begin{bmatrix} Av_1 & \ldots & Av_n \end{bmatrix} = \begin{bmatrix} \lambda_1 v_1 & \ldots & \lambda_n v_n \end{bmatrix} = PD$$

kjer je

$$D = \begin{bmatrix} \lambda_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \lambda_n \end{bmatrix}$$

Primer

Poišči diagonalno matriko, ki je podobna matriki $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$.

Rešitev: Ker sta lastna vektorja

$$v_1 = \begin{bmatrix} 1 \\ i \end{bmatrix} \quad \text{in} \quad v_2 = \begin{bmatrix} 1 \\ -i \end{bmatrix}$$

matrike A linearno neodvisna, vzamemo

$$P = \begin{bmatrix} v_1 & v_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix}$$

Ker sta njuni lastni vrednosti $\lambda_1 = i$ in $\lambda_2 = -i$, vzamemo

$$D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}.$$

Po gornjem računu je $A = PDP^{-1}$.

2. Schurov izrek

Vemo, da ni vsaka kvadratna matrika nad $\mathbb{C}$ podobna kaki diagonalni matriki. Pokažimo pa, da je vedno podobna kaki zgornje trikotni matriki. Kasneje se bo izkazalo, da je podobna zelo posebni zgornje trikotni matriki (Jordanski kanonični formi).

Schurov izrek

Vsaka kvadratna matrika nad $\mathbb{C}$ je podobna kaki zgornje trikotni matriki.

Dokaz bo z popolno indukcijo po velikosti matrike. Očitno trditev velja za matrike velikosti 1, ker so te že same zgornje trikotne. Recimo sedaj, da trditev velja za vse matrike velikosti $n-1$ in vzemimo poljubno matriko A velikosti n. Radi bi dokazali, da trditev velja tudi za matriko A.

Naj bo λ lastna vrednost matrike A in naj bo v_1 pripadajoči lastni vektor. Naj bodo $v_2, \ldots, v_n$ dopolnitev v_1 do baze za F^n. Potem je matrika

$$P = \begin{bmatrix} v_1 & \ldots & v_n \end{bmatrix}$$

obrnljiva. Razvijmo vektorje $Av_2, \ldots, Av_n$ po bazi $v_1, v_2, \ldots, v_n$.

$$Av_2 = \alpha_{2,1}v_1 + \alpha_{2,2}v_2 + \ldots + \alpha_{2,n}v_n$$

$$\vdots$$

$$Av_n = \alpha_{n,1}v_1 + \alpha_{n,2}v_2 + \ldots + \alpha_{n,n}v_n$$

Potem velja

$$AP = A\begin{bmatrix} v_1 & v_2 & \ldots & v_n \end{bmatrix} = \begin{bmatrix} Av_1 & Av_2 & \ldots & Av_n \end{bmatrix}$$

$$= \begin{bmatrix} \lambda v_1 & \alpha_{2,1}v_1 + \alpha_{2,2}v_2 + \ldots + \alpha_{2,n}v_n & \ldots & \alpha_{n,1}v_1 + \alpha_{n,2}v_2 + \ldots + \alpha_{n,n}v_n \end{bmatrix}$$

$$= \begin{bmatrix} v_1 & v_2 & \ldots & v_n \end{bmatrix} \begin{bmatrix} \lambda & \alpha_{2,1} & \ldots & \alpha_{n,1} \\ 0 & \alpha_{2,2} & \ldots & \alpha_{n,2} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \alpha_{2,n} & \ldots & \alpha_{n,n} \end{bmatrix} = P \begin{bmatrix} \lambda & c \\ 0 & B \end{bmatrix}$$

Po indukcijski predpostavki obstaja taka obrnljiva matrika Q velikosti $n-1$ in taka zgornje trikotna matrika T velikosti $n-1$, da velja

$$B = QTQ^{-1}$$

Odtod sledi

$$\begin{bmatrix} 1 & 0 \\ 0 & Q \end{bmatrix}^{-1} P^{-1}AP \begin{bmatrix} 1 & 0 \\ 0 & Q \end{bmatrix} =$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & Q^{-1} \end{bmatrix} \begin{bmatrix} \lambda & c \\ 0 & B \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & Q \end{bmatrix} =$$

$$= \begin{bmatrix} \lambda & c \\ 0 & Q^{-1}B \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & Q \end{bmatrix} =$$

$$= \begin{bmatrix} \lambda & cQ \\ 0 & Q^{-1}BQ \end{bmatrix} = \begin{bmatrix} \lambda & cQ \\ 0 & T \end{bmatrix}$$

kar je zgornje trikotna matrika. S tem je indukcijski korak dokazan. □

Primer

Poišči kako zgornje trikotno matriko, ki je podobna matriki

$$A = \begin{bmatrix} 3 & 4 \\ -1 & -1 \end{bmatrix}$$

Karakteristični polinom matrike A je

$$\det(A - xI) = (3 - x)(-1 - x) + 4 = 1 - 2x + x^2 = (1 - x)^2$$

torej je 1 lastna vrednost matrike A. Pripadajoči lastni vektor je v jedru matrike $A - I$. Vzemimo recimo

$$v_1 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

Dopolnimo ta vektor do baze za $\mathbb{C}^2$ z vektorjem

$$v_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Označimo

$$P = \begin{bmatrix} v_1 & v_2 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ -1 & 1 \end{bmatrix}$$

Potem velja

$$P^{-1}AP = \frac{1}{2}\begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 3 & 4 \\ -1 & -1 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ -1 & 1 \end{bmatrix}$$

$$= \frac{1}{2}\begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 2 & 4 \\ -1 & -1 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 2 & 4 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$$

Torej je matrika A podobna zgornje trikotni matriki

$$\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$$

Lastne vrednosti - 2. del

3. Obstoj diagonalizacije

Še vedno delamo s kvadratnimi matrikami, ki imajo kompleksne elemente.

Definicija diagonalizacije

Diagonalizacija matrike A je razcep $A = PDP^{-1}$, kjer je P obrnljiva matrika, D pa diagonalna matrika.

Naj bodo $v_1, \ldots, v_n$ stolpci matrike P, $d_1, \ldots, d_n$ pa diagonalni elementi matrike D. Očitno velja $AP = [Av_1 \ldots Av_n]$ in $PD = [d_1 v_1 \ldots d_n v_n]$. Torej velja $AP = PD$ natanko tedaj, ko je $Av_i = d_i v_i$ za vsak $i = 1, \ldots, n$. Matrika P je obrnljiva natanko tedaj, ko so $v_1, \ldots, v_n$ linearno neodvisni.

Povzetek:

- Diagonalni elementi matrike D so ravno lastne vrednosti matrike A.
- Stolpci matrike P so linearno neodvisni lastni vektorji matrike A.
- Matrika A ima diagonalizacijo natanko tedaj, ko ima n linearno neodvisnih lastnih vektorjev.

Primer matrike, ki ima diagonalizacijo

Poiščimo diagonalizacijo matrike

$$A = \begin{bmatrix} 0 & 2 & 2 \\ 1 & -1 & -2 \\ -2 & 4 & 5 \end{bmatrix}$$

Najprej izračunamo karakteristični polinom

$$p_A(x) = \det(A - xI) = 2 - 5x + 4x^2 - x^3 = -(x-1)^2(x-2)$$

Lastni podprostor, ki pripada lastni vrednosti 1 je

$$\operatorname{Ker}(A - I) = \operatorname{Ker} \begin{bmatrix} -1 & 2 & 2 \\ 1 & -2 & -2 \\ -2 & 4 & 4 \end{bmatrix} = \operatorname{Lin}\{ \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} \}$$

Lastni podprostor, ki pripada lastni vrednosti 2 je

$$\operatorname{Ker}(A-2I)=\operatorname{Ker}\begin{bmatrix}-2 & 2 & 2\\ 1 & -3 & -2\\ -2 & 4 & 3\end{bmatrix}=\operatorname{Lin}\{\begin{bmatrix}1\\ -1\\ 2\end{bmatrix}\}$$

Za matriki

$$P=\begin{bmatrix}2 & 2 & 1\\ 1 & 0 & -1\\ 0 & 1 & 2\end{bmatrix},\qquad D=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 2\end{bmatrix}$$

torej velja

$$AP=A\begin{bmatrix} v_1 & v_2 & v_3\end{bmatrix}=\begin{bmatrix} Av_1 & Av_2 & Av_3\end{bmatrix}=\begin{bmatrix} v_1 & v_2 & 2v_3\end{bmatrix}=PD$$

odkoder sledi

$$A=PDP^{-1}$$

Primer matrike, ki nima diagonalizacije

Pokažimo, da matrika $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ nima diagonalizacije.

Prvi način: Poskusimo najti tako obrnljivo matriko $P = \begin{bmatrix} x & y \\ u & v \end{bmatrix}$ in tako diagonalno matriko $D = \begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix}$, da velja $AP = PD$. Primerjava istoležnih elementov nam da $u = d_1 x$, $v = d_2 y$, $0 = d_1 u$, $0 = d_2 u$. Če prvo enačbo pomnožimo z u in upoštevamo tretjo, dobimo $u^2 = 0$. Če drugo enačbo pomnožimo z v in upoštevamo četrto, dobimo $v^2 = 0$. Matrika P torej ima ničelno vrstico, kar je v nasprotju z obrnljivostjo.

Drugi način: Izračunajmo lastne vrednosti in lastne vektorje. Karakteristični polinom je x^2, torej je 0 edina lastna vrednost. Pripadajoči lastni podprostor $\mathrm{Ker} A = \mathrm{Lin}\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}\}$ je enodimenzionalen, torej matrika A nima dveh linearno neodvisnih lastnih vektorjev.

Primer uporabe: S pomočjo diagonalizacije je preprosto izračunati potenco matrike, kar pride prav pri reševanju sistemov diferenčnih enačb. Velja

$$A^n = AA\cdots A = PDP^{-1}PDP^{-1}\cdots PDP^{-1} = PDD\cdots DP^{-1} = PD^nP^{-1}$$

kjer D^n izračunamo tako, da potenciramo vse diagonalne elemente.

Odtod sledi, da za vsak polinom (ali konvergentno potenčno vrsto) $f(x) = \sum_i c_i x^i$ velja

$$f(A) = \sum_i c_i A^i = \sum_i c_i PD^i P^{-1} = P(\sum_i c_i D^i)P^{-1} = Pf(D)P^{-1}$$

kjer $f(D)$ izračunamo tako, da uporabimo f na vseh diagonalnih elementih. To nam pride prav reševanju sistemov linearnih diferencialnih enačb s konstantnimi koeficienti, kjer moramo izračunati $e^A = \sum_{n=0}^{\infty} \frac{1}{n!} A^n$.

V nadaljevanju se bomo večkrat sklicevali na naslednjo trditev.

Trditev 1

Naslednje lastnosti matrike A so ekvivalentne

- Matrika A ima diagonalizacijo.
- Matrika A je podobna diagonalni matriki.
- Matrika A ima n linearno neodvisnih lastnih vektorjev.
- Vsota lastnih podprostorov matrike A je enaka $\mathbb{C}^n$.

Dokaz: Iz definicij sledi, da sta prva in druga lastnost ekvivalentni. Dokazali smo že, da sta druga in tretja lastnost ekvivalentni. Tretja in četrta trditev obe pravita, da so lastni vektorji A ogrodje za $\mathbb{C}^n$.

Naslednja trditev je bolj tehnične narave, ampak ima zanimive posledice.

Trditev 2

Lastni vektorji, ki pripadajo paroma različnim lastnim vrednostim, so linearno neodvisni.

Dokaz. Naj bodo $\lambda_1, \ldots, \lambda_k$ paroma različne lastne vrednosti matrike A in naj bodo $v_1, \ldots, v_k$ pripadajoči lastni vektorji. Se pravi

$$Av_1 = \lambda_1 v_1 \quad \ldots \quad Av_k = \lambda_k v_k \tag{1}$$

in $v_1, \ldots, v_k$ so neničelni vektorji. Z indukcijo po j bomo pokazali, da je za vsak $j \leq k$ množica $\{v_1, \ldots, v_j\}$ linearno neodvisna.

Baza indukcije: Ker je $v_1 \neq 0$, je množica $\{v_1\}$ linearno neodvisna.

Indukcijski korak: Recimo, da je množica $\{v_1, \ldots, v_j\}$ linearno neodvisna, kjer je $j < k$. Radi bi pokazali, da je potem tudi množica $\{v_1, \ldots, v_{j+1}\}$ linearno neodvisna. Recimo, da je

$$\alpha_1 v_1 + \ldots + \alpha_j v_j + \alpha_{j+1} v_{j+1} = 0. \tag{2}$$

Če (2) pomnožimo z leve z matriko A in upoštevamo (1), dobimo

$$\alpha_1 \lambda_1 v_1 + \ldots + \alpha_j \lambda_j v_j + \alpha_{j+1} \lambda_{j+1} v_{j+1} = 0. \tag{3}$$

Če od enačbe (3) odštejemo z λ_{j+1} pomnoženo enačbo (2), dobimo

$$\alpha_1(\lambda_1 - \lambda_{j+1})v_1 + \ldots + \alpha_j(\lambda_j - \lambda_{j+1})v_j = 0. \tag{4}$$

Po indukcijski predpostavki odtod sledi, da je

$$\alpha_1(\lambda_1 - \lambda_{j+1}) = 0, \ldots, \alpha_j(\lambda_j - \lambda_{j+1}) = 0. \tag{5}$$

Ker so $\lambda_1, \ldots, \lambda_k$ paroma različne, odtod sledi $\alpha_1 = 0, \ldots, \alpha_j = 0$. Iz enačbe (2) sedaj sledi, da je $\alpha_{j+1} v_{j+1} = 0$. Odtod sledi $\alpha_{j+1} = 0$, saj je $v_{j+1} \neq 0$. Dokazali smo, da iz (2) sledi $\alpha_1 = \ldots = \alpha_{j+1} = 0$, torej so $v_1, \ldots, v_{j+1}$ linearno neodvisni. □

Posledica 1

Če je A $n \times n$ matrika, ki ima n paroma različnih lastnih vrednosti, potem ima A diagonalizacijo.

Dokaz: Ker ima A n paroma različnih lastnih vrednosti, ima po Trditvi 2 tudi n linearno neodvisnih lastnih vektorjev. Odtod po Trditvi 1 sledi, da ima A diagonalizacijo.

Posledica 2

Vsota vseh lastnih podprostorov matrike je direktna.

Dokaz: Naj bodo $\lambda_1, \ldots, \lambda_k$ vse paroma različne lastne vrednosti matrike $A \in M_n(\mathbb{C})$. Pripadajoči lastni podprostori so potem $V_i = \mathrm{Ker}\,(A - \lambda_i I)$ za $i = 1, \ldots, k$. Trdimo, da je vsota podprostorov $V_1, \ldots, V_k$ direktna. Treba je dokazati, da so poljubni vektorji $v_1 \in V_1, \ldots, v_k \in V_k$, ki zadoščajo $v_1 + \ldots + v_k = 0$, enaki nič. To sledi iz Trditve 2.

Primer

Poiščimo diagonalizacijo $n \times n$ matrike

$$A = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ 1 & 0 & 0 & \dots & 0 \end{bmatrix}$$

Rešitev: Naj bo $\varepsilon = e^{2\pi i/n}$ in $v_k = \sum_{j=1}^{n} \varepsilon^{jk} \mathbf{e}_j$. Potem je

$$Av_k = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ 1 & 0 & 0 & \dots & 0 \end{bmatrix} \begin{bmatrix} \varepsilon^k \\ \varepsilon^{2k} \\ \varepsilon^{3k} \\ \vdots \\ \varepsilon^{nk} \end{bmatrix} = \begin{bmatrix} \varepsilon^{2k} \\ \varepsilon^{3k} \\ \varepsilon^{4k} \\ \vdots \\ \varepsilon^{k} \end{bmatrix} = \varepsilon^k \begin{bmatrix} \varepsilon^k \\ \varepsilon^{2k} \\ \varepsilon^{3k} \\ \vdots \\ \varepsilon^{nk} \end{bmatrix} = \varepsilon^k v_k$$

Ker so lastne vrednosti ε^k, $k = 1, \dots, n$, paroma različne, so lastni vektorji v_k, $k = 1, \dots, n$, linearno neodvisni. Torej ima A diagonalizacijo.

Naj bo

$$P = \begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix} \qquad D = \begin{bmatrix} \varepsilon & 0 & \dots & 0 \\ 0 & \varepsilon^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \varepsilon^n \end{bmatrix}$$

Potem je $AP = PD$, torej je iskana diagonalizacija

$$A = PDP^{-1}$$

Primer

Poiščimo diagonalizacijo $n \times n$ matrike

$$C = \begin{bmatrix} c_0 & c_1 & c_2 & \dots & c_{n-2} & c_{n-1} \\ c_{n-1} & c_0 & c_1 & \dots & c_{n-3} & c_{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ c_3 & c_4 & c_5 & \dots & c_1 & c_2 \\ c_2 & c_3 & c_4 & \dots & c_0 & c_1 \\ c_1 & c_2 & c_3 & \dots & c_{n-1} & c_0 \end{bmatrix}$$

Rešitev: Naj bo A matrika iz prejšnjega primera. Opazimo, da velja

$$A^2 = \begin{bmatrix} 0 & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 \\ 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & 0 & \dots & 0 \end{bmatrix}$$

Podobno dobimo $A^3, A^4, \dots$. (Diagonale iz enk se premikajo desno navzgor.). Odtod sledi

$$C = c_0 I + c_1 A + \dots + c_{n-1} A^{n-1}$$

Označimo $p(x) = c_0 + c_1 x + \dots + c_{n-1} x^{n-1}$. Potem je

$$C = p(A) = p(PDP^{-1}) = P p(D) P^{-1}$$

kjer sta P in D kot v prejšnjem primeru.

Opomba: Matrike z večkratnimi lastnimi vrednostmi včasih imajo diagonalizacijo (npr. $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$) včasih pa ne (npr. $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$).
Izkaže se, da je to povezano z algebraično in geometrijsko večkratnostjo lastnih vrednosti. V prvem primeru sta obe večkratnosti enaki ($a(1) = g(1) = 2$), v drugem pa sta različni ($a(0) = 2$ in $g(0) = 1$).

Lema

Za vsako lastno vrednost λ je njena geometrijska večkratnost $g(\lambda)$ manjša ali enaka njeni algebraični večkratnosti $a(\lambda)$.

Dokaz: Naj bo $v_1, \ldots, v_m$ baza za $\mathrm{Ker}(A - \lambda I)$ in naj bo $v_{m+1}, \ldots, v_n$ njena dopolnitev do baze $\mathcal{B}$ za $\mathbb{C}^n$. Običajen račun potem pokaže, da je

$$P^{-1}AP = \begin{bmatrix} \lambda I_m & B \\ 0 & C \end{bmatrix}$$

kjer je $P = \begin{bmatrix} v_1 & \ldots & v_m & v_{m+1} & \ldots & v_n \end{bmatrix}$ obrnljiva matrika. Sledi

$$\det(A - xI_n) = \det(\lambda I_m - xI_m)\det(C - xI_{n-m}) = (\lambda - x)^m \det(C - xI_{n-m})$$

Ker $(x - \lambda)^m$ deli karakteristični polinom, je $a(\lambda) \geq m = g(\lambda)$.

Trditev 3

Matrika A ima diagonalizacijo natanko tedaj, ko se za vsako lastno vrednost A njena geometrijska in algebraična večkratnost ujemata.

Dokaz: Naj bodo $\lambda_1, \ldots, \lambda_k$ vse paroma različne lastne vrednosti matrike A. Njen karakteristični polinom je $p_A(x) = (-1)^n(x-\lambda_1)^{n_1}\cdots(x-\lambda_k)^{n_k}$, kjer je n_i algebraična večkratnost za λ_i. Očitno je $n_1 + \ldots + n_k = n$. Geometrijska večkratnost za λ_i je $m_i := \dim \mathrm{Ker}(A - \lambda_i I)$.

Po Trditvi 1 ima matrika A diagonalizacijo natanko tedaj ko je $U = \mathbb{C}^n$, kjer je U vsota vseh lastnih podprostorov matrike A. Po Posledici 2 je $\dim U = m_1 + \ldots + m_k$. Torej ima matrika A diagonalizacijo natanko tedaj, ko je $m_1 + \ldots + m_k = n_1 + \ldots + n_k$. Ker je $m_i \leq n_i$ za vsak i (Lema), to velja natanko tedaj, ko je $m_i = n_i$ za vsak i. □

Trditev 4

Naj bodo $\lambda_1, \ldots, \lambda_k$ vse paroma različne lastne vrednosti matrike A. Potem ima matrika A diagonalizacijo natako tedaj, ko je produkt matrik $A - \lambda_1 I, \ldots, A - \lambda_k I$ enak ničelni matriki.

Dokaz. Označimo $L_i = A - \lambda_i I$ za $i = 1, \ldots, k$. Potrebovali bomo formulo

$$\dim \operatorname{Ker}(L_1 L_2 \cdots L_k) \leq \dim \operatorname{Ker}(L_1) + \dim \operatorname{Ker}(L_2) + \ldots + \dim \operatorname{Ker}(L_k). \tag{1}$$

Primer $k = 2$ formule (1) smo dokazali v prejšnjem poglavju. Za splošen k formulo (1) dokažemo z indukcijo.

Če velja $L_1 \cdots L_k = 0$, potem je $\operatorname{Ker}(L_1 \cdots L_k) = \mathbb{C}^n$, kjer je n velikost matrike A, torej je $\dim \operatorname{Ker}(L_1 \cdots L_k) = n$. Označimo z $m_i = \dim \operatorname{Ker}(L_i)$ geometrijsko večkratnost λ_i. Iz formule (1) torej sledi $n \leq m_1 + \cdots + m_k$. Naj bo n_i algebraična večkratnost λ_i. Iz $n = n_1 + \ldots + n_k$ in iz $m_i \leq n_i$ sledi, da je $m_i = n_i$ za vsak i. Po Trditvi 3 ima torej matrika A diagonalizacijo.

Dokažimo še drugo smer. Recimo, da ima matrika A diagonalizacijo $A = PDP^{-1}$. S permutacijo diagonalnih elementov matrike D dobimo matriko, ki ji je podobna. Zato lahko predpostavimo, da enake lastne vrednosti ležijo skupaj na diagonali D, se pravi, da je

$$D = \begin{bmatrix} \lambda_1 I_{m_1} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_k I_{m_k} \end{bmatrix} \tag{2}$$

Iz (2) sledi, da je i-ti diagonalni blok matrike $D - \lambda_i I$ enak 0, torej je

$$(D - \lambda_1 I)(D - \lambda_2 I) \cdots (D - \lambda_k I) = 0. \tag{3}$$

Iz (3) izpeljemo

$$P(D - \lambda_1 I)P^{-1} P(D - \lambda_2 I)P^{-1} \cdots P(D - \lambda_k I)P^{-1} = 0 \tag{4}$$

kjer je $P(D - \lambda_i I)P^{-1} = PDP^{-1} - \lambda_i I = A - \lambda_i I$. □

Lastne vrednosti - 3. del

4. Minimalni polinom matrike

Označimo s $\mathbb{C}[x]$ množico vseh polinomov v x s kompleksnimi koeficienti. Z $M_n(\mathbb{C})$ označimo množico vseh kompleksnih $n \times n$ matrik.

V polinom $p(x) = c_n x^n + c_{n-1} x^{n-1} + \ldots + c_1 x + c_0 \in \mathbb{C}[x]$, bi radi namesto x vstavili matriko $A \in M_n(\mathbb{C})$. Problem je v tem, ker so členi $c_n A^n, c_{n-1} A^{n-1}, \ldots, c_1 A$ matrike, člen c_0 pa je skalar. Zato v $p(x)$ člen c_0 pomnožimo z $x^0 = 1$, v $p(A)$ pa ga pomnožimo z $A^0 := I$. Torej je

$$p(A) := c_n A^n + c_{n-1} A^{n-1} + \ldots + c_1 A + c_0 I$$

Opomba: V nadaljevanju bomo večkrat potrebovali,
da iz $p(x) = p_1(x)p_2(x)$ sledi $p(A) = p_1(A)p_2(A)$.
To ni povsem očitno, ampak je treba napraviti kratek račun.

Sedaj smo nared za glavno definicijo tega razdelka:

Definicija minimalnega polinoma

Polinom $m \in \mathbb{C}[x]$ je **minimalni polinom** matrike $A \in M_n(\mathbb{C})$, če velja:
(1) $m(A) = 0$,
(2) m ima vodilni koeficient 1.
(3) med vsemi polinomi, ki zadoščajo (1) in (2), ima m najnižjo stopnjo.

Prepričajmo se najprej, da je definicija smiselna.

Trditev 1: (Obstoj in enoličnost minimalnega polinoma)

Vsaka matrika A ima natanko en minimalni polinom. Označimo ga z m_A.

Dokaz bomo razbili v več korakov.

Vzemimo poljubno matriko $A \in M_n(\mathbb{C})$.

1. korak Konstruirajmo najprej tak polinom $p \in \mathbb{C}[x]$, ki zadošča $p(A) = 0$ in ki ima vodilni koeficient enak ena.

Očitno je $M_n(\mathbb{C})$ vektorski prostor nad $\mathbb{C}$ dimenzije n^2. Ker ima množica $\{I, A, A^2, \ldots, A^{n^2}\}$ en element več od dimenzije, je linearno odvisna. Obstajajo torej taki $c_0, c_1, c_2, \ldots, c_{n^2} \in \mathbb{C}$, od katerih je vsaj en neničeln, da velja $\sum_{i=0}^{n^2} c_i A^i = 0$. Polinom $p(x)$ dobimo tako, da polinom $\sum_{i=0}^{n^2} c_i x^i$ delimo z njegovim vodilnim koeficientom.

2. korak Matrika A ima vsaj en minimalni polinom.

Naj bo $\mathcal{M}$ množica vseh polinomov, ki zadoščajo točkam (1) in (2) iz definicije minimalnega polinoma. Po 1. koraku je množica $\mathcal{M}$ neprazna. Vzemimo v množici $\mathcal{M}$ polinom z najnižjo stopnjo. Tak polinom potem zadošča vsem trem točkam iz definicije minimalnega polinoma.

3. korak Naj bo m minimalni polinom matrike A in naj bo p tak polinom, ki zadošča $p(A) = 0$. Potem m deli p.

Po izreku o deljenju z ostankom obstajata taka polinoma k in r, da velja $p = km + r$ in $\deg r < \deg m$. Ker je $m(A) = 0$ in $p(A) = 0$, je tudi $r(A) = 0$. Če bi bil r neničeln polinom, bi ga delili z vodilnim koeficientom in bi dobili tak polinom, ki zadošča točkam (1) in (2) iz definicije minimalnega polinoma in je nižje stopnje od m. To je v nasprotju s predpostavko, da m zadošča točki (3) iz definicije minimalnega polinoma. Torej je r ničeln polinom.

4. korak Če sta m_1 in m_2 minimalna polinoma matrike A, potem je $m_1 = m_2$.

Ker je $m_2(A) = 0$, po 3. koraku m_1 deli m_2. Ker je $m_1(A) = 0$, po 3. koraku m_2 deli m_1. Torej obstaja taka konstanta $c \neq 0$, da je $m_2 = cm_1$. Ker imata oba polinoma vodilni koeficient enak 1, je $c = 1$. □

Radi bi dokazali, da minimalni polinom matrike A deli karakteristični polinom matrike A. Po 3. koraku dokaza Trditve 2, je dovolj dokazati:

Trditev 2: (Cayley-Hamiltonov izrek)

Naj bo $p_A(x)$ karakteristični polinom matrike A. Potem velja $p_A(A) = 0$.

Opomba: Kaj je narobe z naslednjim "dokazom"? V $p_A(x) = \det(A - xI)$ vstavimo $x = A$ in dobimo $p_A(A) = \det(A - AI) = \det 0 = 0$.

Dokaz: Spomnimo se formule $B^{-1} = \frac{1}{\det B}\tilde{B}^T$ za inverz matrike, kjer matriko $\tilde{B}$ dobimo tako, da v matriki B vsak element $b_{i,j}$ zamenjamo z $(-1)^{i+j} \det B_{i,j}$. (B s pobrisano i-to vrstico in j-tim stolpcem je $B_{i,j}$).
Če to formulo pomnožimo z $\det(B)B$ z leve, dobimo $\det(B)I = B\tilde{B}^T$.
Sedaj vstavimo $B = A - xI$ in dobimo $p_A(x)I = (A - xI)\widetilde{(A - xI)}^T$.
Elementi matrike $\widetilde{(A - xI)}^T$ so polinomi v x stopnje $\leq n - 1$. (So namreč determinante velikosti $n - 1$, katerih elementi so polinomi stopnje ≤ 1.)

V enačbo $p_A(x)I = (A - xI)\widetilde{(A - xI)}^T$ vstavimo
$p_A(x) = c_0 + c_1x + \ldots + c_nx^n$, kjer $c_i \in \mathbb{C}$, in
$\widetilde{(A - xI)}^T = B_0 + xB_1 + \ldots + x^{n-1}B_{n-1}$, kjer $B_i \in M_n(\mathbb{C})$.
Primerjajmo koeficiente pri potencah x^i, kjer $i = 0, 1, 2, \ldots, n-1, n$:

$$\begin{aligned} c_0 I &= AB_0 \\ c_1 I &= AB_1 - B_0 \\ c_2 I &= AB_2 - B_1 \\ &\vdots \\ c_{n-1} I &= AB_{n-1} - B_{n-2} \\ c_n I &= -B_{n-1} \end{aligned}$$

Pomnožimo i-to enačbo z A^i z leve in vse enačbe seštejmo.
Dobimo $c_0 I + c_1 A + c_2 A^2 + \ldots + c_{n-1}A^{n-1} + c_n A^n =$
$AB_0 + A(AB_1 - B_0) + A^2(AB_2 - B_1) + \ldots + A^{n-1}(AB_{n-1} - B_{n-2}) + A^n(-B_{n-1})$.
Na desni strani odpravimo oklepaje in opazimo, da se vse pokrajša. □

Ker minimalni polinom $m_A(x)$ deli karakterističnega polinom $p_A(x)$ in ker je $p_A(x)$ oblike $p_A(x) = (-1)^n(x-\lambda_1)^{n_1}\cdots(x-\lambda_k)^{n_k}$, je $m_A(x)$ oblike

$$m_A(x) = (x-\lambda_1)^{r_1}\cdots(x-\lambda_k)^{r_k}$$

kjer je $r_1 \leq n_1, \ldots, r_k \leq n_k$. Radi bi pokazali še, da je $r_1 \geq 1, \ldots, r_k \geq 1$:

Trditev 3

Vsaka lastna vrednost matrike A je ničla minimalnega polinoma m_A.

Dokaz: Naj bo λ lastna vrednost matrike A in naj bo v pripadajoči lastni vektor. Iz $Av = \lambda v$ s popolno indukcijo izpeljemo, da velja $A^k v = \lambda^k v$ za vsak k. Če namreč velja $A^{k-1}v = \lambda^{k-1}v$ za nek k, potem je

$$A^k v = A(A^{k-1}v) = A(\lambda^{k-1}v) = \lambda^{k-1}Av = \lambda^{k-1}(\lambda v) = \lambda^k v.$$

Recimo, da je minimalni polinom oblike $m_A(x) = \sum_{i=0}^{r} c_i x^i$. Potem velja

$$0 = 0v = m_A(A)v = \sum_{i=0}^{r} c_i A^i v = \sum_{i=0}^{r} c_i \lambda^i v = m_A(\lambda)v.$$

Ker je v neničeln, odtod sledi $m_A(\lambda) = 0$.

Kakšna je zveza med minimalnim polinomom in obstojem diagonalizacije?

Trditev 4

Matrika ima diagonalizacijo natanko tedaj, ko njen minimalni polinom nima večkratnih ničel.

Dokaz: Naj bodo $\lambda_1, \ldots, \lambda_k$ vse paroma različne lastne vrednosti matrike A. Iz 3. razdelka vemo, da ima A diagonalizacijo natanko tedaj, ko je produkt matrik $A - \lambda_i I$ enak nič. Označimo z $m(x)$ produkt polinomov $x - \lambda_i$, se pravi $m(x) = (x - \lambda_1) \cdots (x - \lambda_k)$. Torej ima A diagonalizacijo natanko tedaj, ko je $m(A) = 0$. Po 3. koraku v dokazu Trditve 1 velja $m(A) = 0$ natanko tedaj, ko minimalni polinom $m_A(x)$ deli polinom $m(x)$. Po Trditvi 3 velja to natanko tedaj, ko je $m_A(x) = m(x)$. □

Primer

Matriki $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ in $B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ imata enak karakteristični polinom toda različna minimalna polinoma $m_A(x) = x - 1$ in $m_B(x) = (x - 1)^2$.

5. Korenski podprostori

Naj bodo $\lambda_1, \ldots, \lambda_k$ vse paroma različne lastne vrednosti matrike $A \in M_n(\mathbb{C})$. Vemo, da je karakteristični polinom matrike A oblike

$$p_A(x) = (-1)^n (x - \lambda_1)^{n_1} \cdots (x - \lambda_k)^{n_k},$$

minimalni polinom matrike A pa je oblike

$$m_A(x) = (x - \lambda_1)^{r_1} \cdots (x - \lambda_k)^{r_k}$$

kjer števila $r_1, \ldots, r_k$ zadoščajo $1 \leq r_1 \leq n_1, \ldots, 1 \leq r_k \leq n_k$.

Števila r_i potrebujemo v definiciji korenskega podprostora:

Definicija korenskih vektorjev in korenskega podprostora

Množica $\mathrm{Ker}(A - \lambda_i I)^{r_i}$ je vektorski podprostor v $\mathbb{C}^n$. Pravimo ji **korenski podprostor** matrike A za lastno vrednost λ_i. Njenim neničelnim elementom pravimo **korenski vektorji** matrike A za lastno vrednost λ_i.

Korenske podprostore matrike A bomo potrebovali pri konstrukciji jordanske kanonične forme matrike A v naslednjem razdelku. Oglejmo si najprej nekaj njihovih osnovnih lastnosti.

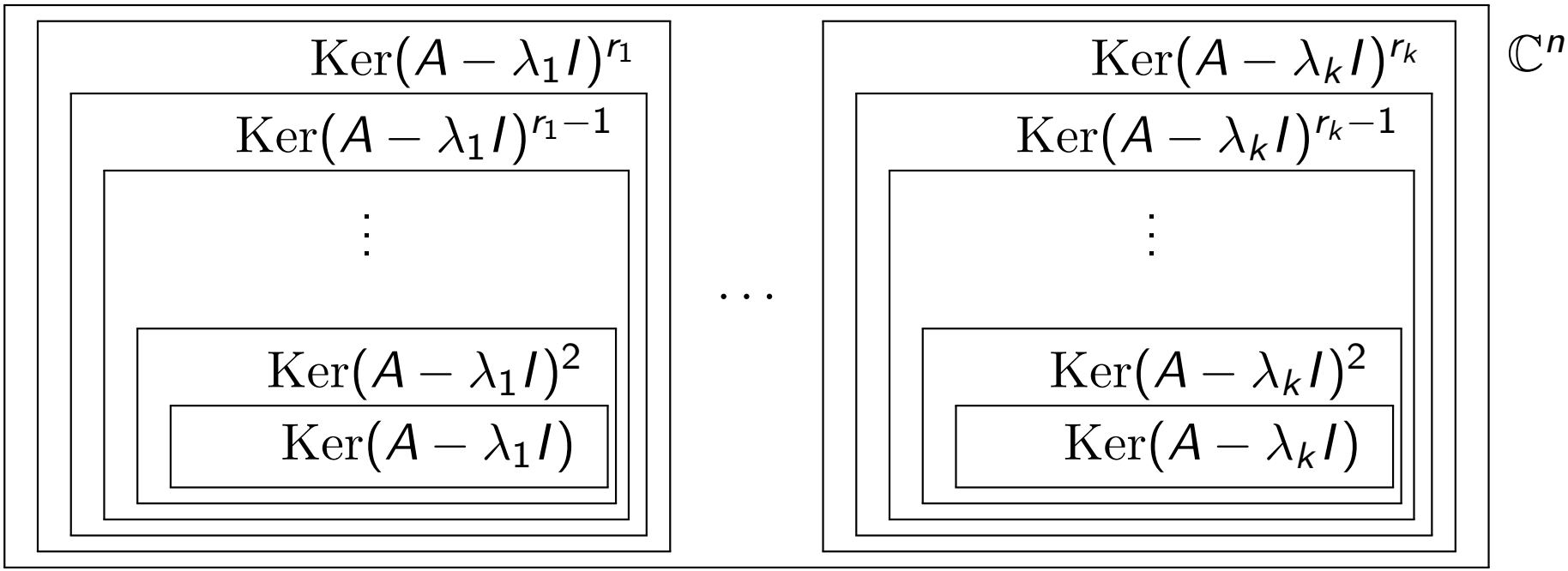
Izrek o korenskem razcepu

Naj bodo A, λ_i, n_i in r_i kot zgoraj.

1. Velja $\mathrm{Ker}(A-\lambda_i I) \subset \mathrm{Ker}(A-\lambda_i I)^2 \subset \ldots \subset \mathrm{Ker}(A-\lambda_i I)^{r_i} = \mathrm{Ker}(A-\lambda_i I)^{r_i+1} = \mathrm{Ker}(A-\lambda_i I)^{r_i+2} = \ldots$, kjer je inkluzija stroga.
2. $\dim \mathrm{Ker}(A-\lambda_i I)^{r_i} = n_i$, se pravi, da je dimenzija korenskega podprostora A za λ_i enaka algebraični večkratnosti λ_i.
3. $\mathbb{C}^n = \bigoplus_{i=1}^{k} \mathrm{Ker}(A-\lambda_i I)^{r_i}$, kar pomeni, da za vsak $w \in \mathbb{C}^n$ obstajajo natanko določeni $w_i \in \mathrm{Ker}(A-\lambda_i I)^{r_i}$, ki zadoščajo $w = \sum_{i=1}^{k} w_i$.

Označimo z $W_i = \mathrm{Ker}(A-\lambda_i I)^{r_i}$ korenski podprostor A za λ_i, z $V_i = \mathrm{Ker}(A-\lambda_i I)$ pa lastni podprostor A za λ_i. Iz prve točke sledi, da je V_i vsebovan v W_i in da sta enaka natanko tedaj, ko je $r_i = 1$.

Izrek o korenskem razcepu nam pove, da imamo naslednjo situacijo:



Velike škatle $\mathrm{Ker}(A-\lambda_i I)^{r_i}$ so korenski podprostori matrike A. Njihova vsota je $\mathbb{C}^n$. Male škatle $\mathrm{Ker}(A-\lambda_i I)$ so lastni podprostori matrike A.

Priprave na dokaz izreka

V dokazu izreka bomo potrebovali pojem invariantnega podprostora.

Definicija invariantnega podprostora

Vektorski podprostor U v $\mathbb{C}^n$ je **invarianten** za matriko $A \in M_n(\mathbb{C})$, če za vsak $u \in U$ velja $Au \in U$.

Primeri invariantnih podprostorov matrike A

- Trivialni podprostor $\{0\}$.
- Lastni podprostori matrike A.
- Korenski podprostori matrike A.
- Presek invariantnih podprostorov za A je invarianten podprostor za A.

Dokažimo invariantnost korenskega podprostora $W_i = \mathrm{Ker}(A - \lambda_i I)^{r_i}$. Za vsak $w \in W_i$ velja $(A - \lambda_i I)^{r_i} w = 0$. Če to pomnožimo z A, dobimo $A(A - \lambda_i I)^{r_i} w = 0$. Ker je $A(A - \lambda_i I)^{r_i} = (A - \lambda_i I)^{r_i} A$, odtod sledi $(A - \lambda_i I)^{r_i} Aw = 0$. Torej $Aw \in W_i$. Dokaz ostalih točk je še lažji.

Trditev 1

Vsak netrivialen invarianten podprostor za matriko A vsebuje vsaj en lastni vektor matrike A.

Dokaz: Naj bo U netrivialen invarianten podprostor za A. Ker je $U \neq \{0\}$, ima U bazo, recimo $\mathcal{B} = \{u_1, \ldots, u_m\}$. Ker je U invarianten za A, lahko definiramo linearno preslikavo $L\colon U \to U$ z $L(u) = Au$.

Naj bo $\begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix}$ lastni vektor matrike $[L]_{\mathcal{B}}$ za lastno vrednost λ. Očitno $v := x_1 u_1 + \ldots + x_m u_m$ pripada U. Pokažimo, da je v lastni vektor za A. Velja namreč $[Lv]_{\mathcal{B}} = [L]_{\mathcal{B}}[v]_{\mathcal{B}} = \lambda[v]_{\mathcal{B}} = [\lambda v]_{\mathcal{B}}$. Torej je $Lv = \lambda v$. □

Dokaz točke (3)

Dokaz točke (3) v izreku bomo razdelili na več trditev.

Trditev 2

Presek dveh različnih korenskih podprostorov matrike A je trivialen.

Dokaz. Ker sta korenska podprostora W_i in W_j invariantna za A, je invarianten za A tudi njun presek. Če je njun presek netrivialen, potem po Trditvi 1 vsebuje lastni vektor za A. Torej je $Av = \lambda v$ za nek neničeln $v \in W_i \cap W_j$. Odtod sledi, da za vsak polinom $p(x)$ velja $p(A)v = p(\lambda)v$. Če vzamemo $p(x) = (x - \lambda_i)^{r_i}$, dobimo $0 = (A - \lambda_i I)^{r_i} v = (\lambda - \lambda_i)^{r_i} v$, odkoder sledi $\lambda = \lambda_i$. Če pa vzamemo $p(x) = (x - \lambda_j)^{r_j}$, potem dobimo $0 = (A - \lambda_j I)^{r_j} v = (\lambda - \lambda_j)^{r_j} v$, odkoder sledi $\lambda = \lambda_j$. Torej je $\lambda_i = \lambda_j$. Odtod sledi $W_i = W_j$, kar je v nasprotju s predpostavko. □

Trditev 3

Če vektorji $w_1 \in W_1, \ldots, w_k \in W_k$ zadoščajo $w_1 + \ldots + w_k = 0$, potem velja $w_1 = \ldots = w_k = 0$.

Dokaz: Dokazali bomo, da za vsak $i = 1, \ldots, k$ velja naslednja trditev:

(T_i) Če vektorji $w_1 \in W_1, \ldots, w_i \in W_i$ zadoščajo $w_1 + \ldots + w_i = 0$, potem velja $w_1 = \ldots = w_i = 0$.

Baza indukcije: Trditev (T_1) je očitna.

Indukcijski korak: Recimo, da velja trditev (T_i), kjer $i < k$.
Radi bi dokazali trditev (T_{i+1}). Vzemimo take vektorje $w_1 \in W_1, \ldots, w_i \in W_i, w_{i+1} \in W_{i+1}$, da velja $w_1 + \ldots + w_i + w_{i+1} = 0$. Če to pomnožimo z $(A - \lambda_{i+1} I)^{r_{i+1}}$ z leve, dobimo $w_1' + \ldots + w_i' + 0 = 0$. Ker je vsak korenski podprostor W_j invarianten in ker vsebuje w_j, vsebuje tudi $w_j' := (A - \lambda_{i+1} I)^{r_{i+1}} w_j$. Iz $w_1' + \ldots + w_i' = 0$ torej po indukcijski predpostavki sledi, da velja $w_1' = \ldots = w_i' = 0$. Zato $w_1, \ldots, w_i \in W_{i+1}$. Po Trditvi 2 je $W_1 \cap W_{i+1} = \{0\}, \ldots, W_i \cap W_{i+1} = \{0\}$. Odtod sledi, da je $w_1 = \ldots = w_i = 0$. Odtod sledi še $w_{i+1} = 0$. □

Opomba: Recimo, da je $\mathcal{B}_i$ baza za korenski podprostor W_i. Iz Trditve 3 sledi, da je $\mathcal{B}_1 \cup \ldots \cup \mathcal{B}_k$ baza za vsoto $W_1 + \ldots + W_k$ vseh korenskih podprostorov. Torej je $\dim(W_1 + \ldots + W_k) = \dim W_1 + \ldots + \dim W_k$.

Trditev 4

Vsota vseh korenskih podprostorov matrike A je enaka $\mathbb{C}^n$.

Dokaz: Če v oceno $n(L_1 \cdots L_k) \leq n(L_1) + \ldots + n(L_k)$ iz prejšnjega poglavja, vstavimo $L_i = (A - \lambda_i I)^{r_i}$ in upoštevamo $L_1 \cdots L_k = m_A(A) = 0$ ter $n(L_i) = \dim \mathrm{Ker}(A - \lambda_i I)^{r_i} = \dim W_i$, dobimo, da velja ocena $n \leq \dim W_1 + \ldots + \dim W_k$. Po opombi zgoraj (ki ste jo delali na vajah) iz Trditve 3 sledi, da je $\dim(W_1 + \ldots + W_k) = \dim W_1 + \ldots + \dim W_k$. Ker je torej $\dim(W_1 + \ldots + W_k) \geq n$, velja $W_1 + \ldots + W_k = \mathbb{C}^n$. □

Opomba: Trditvi 3 in 4 nam dasta ravno točko (3) v izreku.

Dokaz točke (2)

Označimo $t_i := \dim W_i$ za vsak $i = 1, \ldots, k$. Radi bi pokazali, da je $t_i = n_i$. Naj bo $\mathcal{B}_i = \{v_{i,1}, \ldots, v_{i,t_i}\}$ baza za W_i. Ker je podprostor W_i invarianten za A, obstajajo taka števila $\alpha_{i,j,j'}$, da velja

$$\begin{aligned} Av_{i,1} &= \alpha_{i,1,1}v_{i,1} + \ldots + \alpha_{i,1,t_i}v_{i,t_i} \\ &\vdots \\ Av_{i,t_i} &= \alpha_{i,t_i,1}v_{i,1} + \ldots + \alpha_{i,t_i,t_i}v_{t_i,t_i} \end{aligned}$$

V matrični obliki to zapišemo kot $AP_i = P_iA_i$, kjer je

$$A_i = \begin{bmatrix} \alpha_{i,1,1} & \ldots & \alpha_{i,t_i,1} \\ \vdots & \ddots & \vdots \\ \alpha_{i,1,t_i} & \ldots & \alpha_{i,t_i,t_i} \end{bmatrix}, \quad P_i := [v_{i,1} \ \ldots \ v_{i,t_i}]$$

Odtod sledi

$$A\,[P_1 \ \ldots \ P_k] = [P_1 \ \ldots \ P_k] \begin{bmatrix} A_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & A_k \end{bmatrix} \tag{1}$$

Matrike $P := [P_1 \ \dots \ P_k]$ je obrnljiva, ker je (po že dokazani točki (3) v izreku) unija baz $\mathcal{B}_i$ baza za $\mathbb{C}^n$. Označimo z A' bločno diagonalno matriko iz A_i-jev. Po formuli (1) je $A' = P^{-1}AP$, torej imata A in A' enak karakteristični polinom

$$\det(A - xI) = \det(A_1 - xI) \cdots \det(A_k - xI). \tag{2}$$

Pokažimo, da je λ_i edina lastna vrednost matrike A_i. Potem je $\det(A_i - xI) = (\lambda_i - x)^{t_i}$. To vstavimo v formulo (2) in dobimo $t_i = n_i$.

Naj bo μ_i poljubna lastna vrednost matrike A_i in naj bo $u_i \in \mathbb{C}^{t_i}$ pripadajoči lastni vektor. Potem je

$$A(P_i u_i) = (AP_i)u_i = (P_i A_i)u_i = P_i(A_i u_i) = P_i(\mu_i u_i) = \mu_i P_i u_i \tag{3}$$

Ker je $P_i u_i$ linearna kombinacija stolpcev P_i in ker so stolpci P_i v W_i je $P_i u_i \in W_i$. Poleg tega je $P_i u_i \neq 0$, ker je $u_i \neq 0$ in ker so stolpci P_i linearno neodvisni. Po definiciji W_i je $(A - \lambda_i I)^{r_i} P_i u_i = 0$. Iz formule (3) sledi, da je $(A - \lambda_i I)^{r_i} P_i u_i = (\mu_i - \lambda_i)^{r_i} P_i u_i$. Torej je $\mu_i = \lambda_i$.

Dokaz točke (1)

Dokaz točke (1) v izreku bomo razdelili v več korakov.

1. korak Če je $m \leq m'$, potem velja $\mathrm{Ker}(A - \lambda_i I)^m \subseteq \mathrm{Ker}(A - \lambda_i I)^{m'}$.

Za vsak $v \in \mathrm{Ker}(A - \lambda_i I)^m$ velja $(A - \lambda_i I)^m v = 0$. Pomnožimo to z leve z $(A - \lambda_i I)^{m'-m}$ in dobimo $(A - \lambda_i I)^{m'} v = 0$. Torej je $v \in \mathrm{Ker}(A - \lambda_i I)^{m'}$.

2. korak Za vsak $r_i' \geq r_i$ je $\mathrm{Ker}(A - \lambda_i I)^{r_i'} = \mathrm{Ker}(A - \lambda_i I)^{r_i}$.

Označimo $W_i' := \mathrm{Ker}(A - \lambda_i I)^{r_i'}$ in $n_i' = \dim W_i'$. Po 1. koraku je $W_i \subseteq W_i'$, zato je $n_i \leq n_i'$. Kot v dokazu Trditve 3 vidimo, da za poljubne $w_i' \in W_i'$, ki zadoščajo $w_1' + \ldots + w_k' = 0$, velja $w_1' = \ldots = w_k' = 0$. Kot v opombi za Trditvijo 3 odtod sledi, da je $n_1' + \ldots + n_k' = n$. Odtod in iz $n_i \leq n_i'$ sledi, da je $n_i' = n_i$ za vsak i. Torej je $W_i' = W_i$ za vsak i.

3. korak Za vsak i je $\mathrm{Ker}(A-\lambda_i I)^{r_i-1} \neq \mathrm{Ker}(A-\lambda_i I)^{r_i}$.

Če bi za nek i veljal enačaj, potem bomo pokazali, da bi polinom $m_i(x) := m_A(x)/(x-\lambda_i)$ zadoščal $m_i(A)=0$, kar bi bilo v nasprotju z definicijo minimalnega polinoma. Vzemimo poljuben $w \in \mathbb{C}^n$.
Po Trditvi 4 obstajajo taki $w_j \in W_j$, da velja $w = \sum_{j=1}^{k} w_j$.
Za vsak $j \neq i$ velja $(A-\lambda_j)^{r_j} w_j = 0$, odkoder sledi $m_i(A)w_j = 0$.
Za $j = i$ pa velja $(A-\lambda_j)^{r_j-1} w_j = 0$, torej je tudi $m_i(A)w_j = 0$.
Odtod sledi $m_i(A)w = 0$. Ker je to res za vsak $w \in \mathbb{C}^n$, je $m_i(A) = 0$.

4. korak Če velja $\mathrm{Ker}(A-\lambda_i I)^m \neq \mathrm{Ker}(A-\lambda_i I)^{m+1}$ za nek $m \geq 1$ in nek $i = 1, \ldots, k$, potem velja tudi $\mathrm{Ker}(A-\lambda_i I)^{m-1} \neq \mathrm{Ker}(A-\lambda_i I)^m$.

Predpostavimo, da velja $\mathrm{Ker}(A-\lambda_i I)^{m-1} = \mathrm{Ker}(A-\lambda_i I)^m$.
Dokazujemo, da velja $\mathrm{Ker}(A-\lambda_i I)^m = \mathrm{Ker}(A-\lambda_i I)^{m+1}$. Vzemimo poljuben $v \in \mathrm{Ker}(A-\lambda_i I)^{m+1}$. Potem $(A-\lambda_i)v \in \mathrm{Ker}(A-\lambda_i I)^m$.
Torej $(A-\lambda_i)v \in \mathrm{Ker}(A-\lambda_i I)^{m-1}$ po predpostavki. Odtod sledi, da $v \in \mathrm{Ker}(A-\lambda_i I)^m$. Torej je $\mathrm{Ker}(A-\lambda_i I)^m \supseteq \mathrm{Ker}(A-\lambda_i I)^{m+1}$.
Obratna inkluzija sledi iz 1. koraka.

Lastne vrednosti - 4. del

6. Jordanska kanonična forma

Definicija jordanske kletke in jordanske matrike

Jordanska kletka je matrika oblike

$$\begin{bmatrix} \lambda \end{bmatrix}, \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}, \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}, \begin{bmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{bmatrix}, \ldots$$

kjer je λ kompleksno število. **Jordanska matrika** je matrika oblike

$$\begin{bmatrix} J_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & J_m \end{bmatrix},$$

kjer so $J_1, \ldots, J_m$ jordanske kletke (lahko različnih velikosti in različnih λ).

Spomnimo se, da ni vsaka matrika podobna diagonalni matriki. Velja pa:

Izrek o jordanski kanonični formi

Vsaka kompleksna kvadratna matrika A je podobna kaki jordanski matriki J. Pravimo, da je J **jordanska kanonična forma** za A.

Opomba: Jordanska kanonična forma matrike A ni enolična. Če namreč permutiramo njene jordanske kletke, potem spet dobimo jordansko kanonično formo matrike A. S primerno permutacijo jordanskih kletk lahko dosežemo, da jordanske kletke z istim λ sedijo skupaj in so urejene po velikosti od največje do najmanjše.

Opomba: Jordanska kanonična forma matrike A je uporabna za računanje funkcij matrike A. To bomo spoznali v naslednjem razdelku.

V dokazu izreka bomo potrebovali pojem jordanske verige.

Definicija jordanske verige

Naj bo λ lastna vrednost matrike A. **Jordanska veriga** dolžine k je tako zaporedje neničelnih vektorjev $v_1, \ldots, v_k$ iz $\mathbb{C}^n$, da velja $(A - \lambda I)v_1 = 0$, $(A - \lambda I)v_2 = v_1, \ldots, (A - \lambda I)v_k = v_{k-1}$.

Opomba: Iz definicije sledi, da je $Av_1 = \lambda v_1, Av_2 = v_1 + \lambda v_2, \ldots,$ $Av_k = v_{k-1} + \lambda v_k$. V matrični obliki to zapišemo kot

$$A \begin{bmatrix} v_1 & v_2 & \ldots & v_k \end{bmatrix} = \begin{bmatrix} v_1 & v_2 & \ldots & v_k \end{bmatrix} \begin{bmatrix} \lambda & 1 & \ldots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ & & & \\ 0 & 0 & \ddots & 1 \\ 0 & 0 & \ldots & \lambda \end{bmatrix} \qquad (*)$$

Matrika $\begin{bmatrix} v_1 & v_2 & \ldots & v_k \end{bmatrix}$ ni kvadratna. Stolpci so linearno neodvisni.

Definicija jordanske baze

Bazi, ki je unija jordanskih verig pravimo **jordanska baza**.

Če uspemo najti jordansko bazo za $\mathbb{C}^n$, je dokaz izreka o jordanski kanonični formi končan. Elemente te baze namreč zložimo v matriko

$$P = \begin{bmatrix} P_1 & \dots & P_m \end{bmatrix},$$

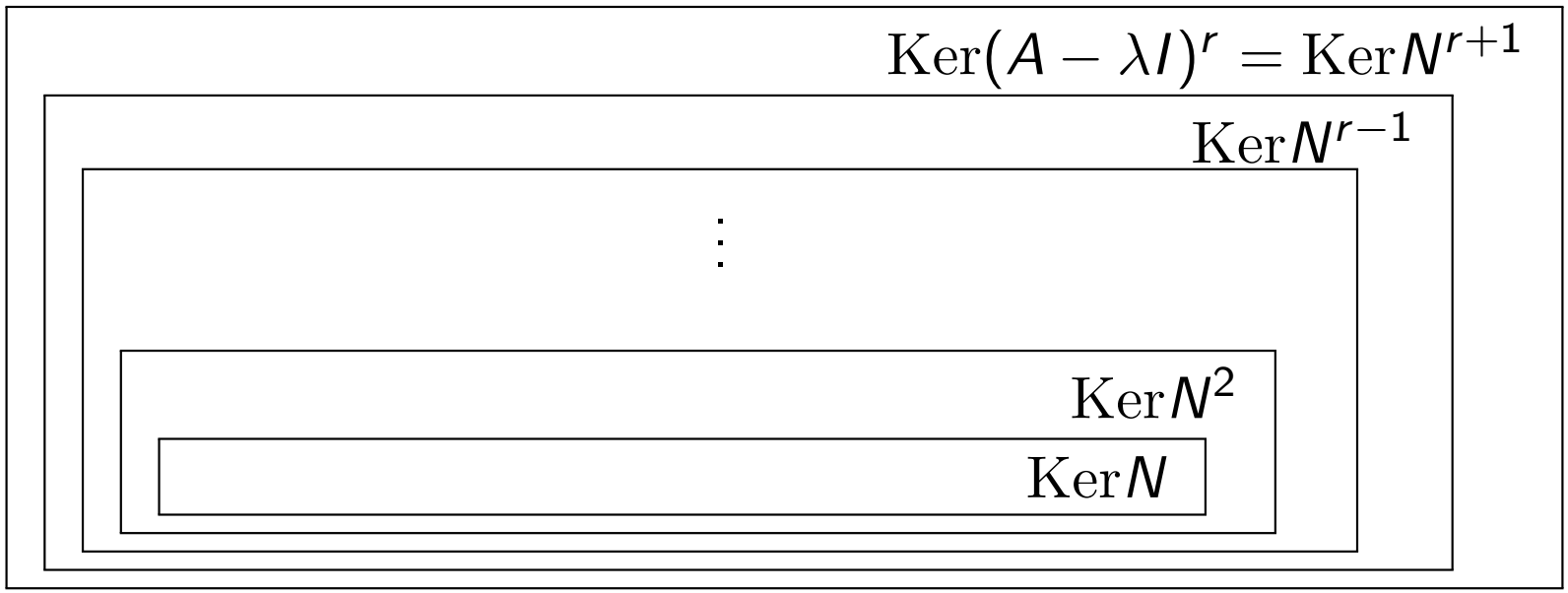
kjer so stolpci podmatrike P_i ravno elementi i-te Jordanske verige v tej bazi. Po formuli $(*)$ obstajajo take Jordanske kletke $J_1, \dots, J_m$, da velja $AP_i = P_i J_i$ za vsak i. Odtod sledi, da velja

$$AP = P \begin{bmatrix} J_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & J_m \end{bmatrix},$$

Torej je matrika A res podobna Jordanski matriki.

Da bi našli jordansko bazo za $\mathbb{C}^n$, je dovolj dovolj poiskati jordansko bazo za vsak korenski podprostor posebej. Po izreku o korenskem razcepu je namreč $\mathbb{C}^n$ direktna vsota vseh korenskih podprostorov matrike A, torej je unija baz vseh korenskih podprostorov baza za cel prostor $\mathbb{C}^n$.

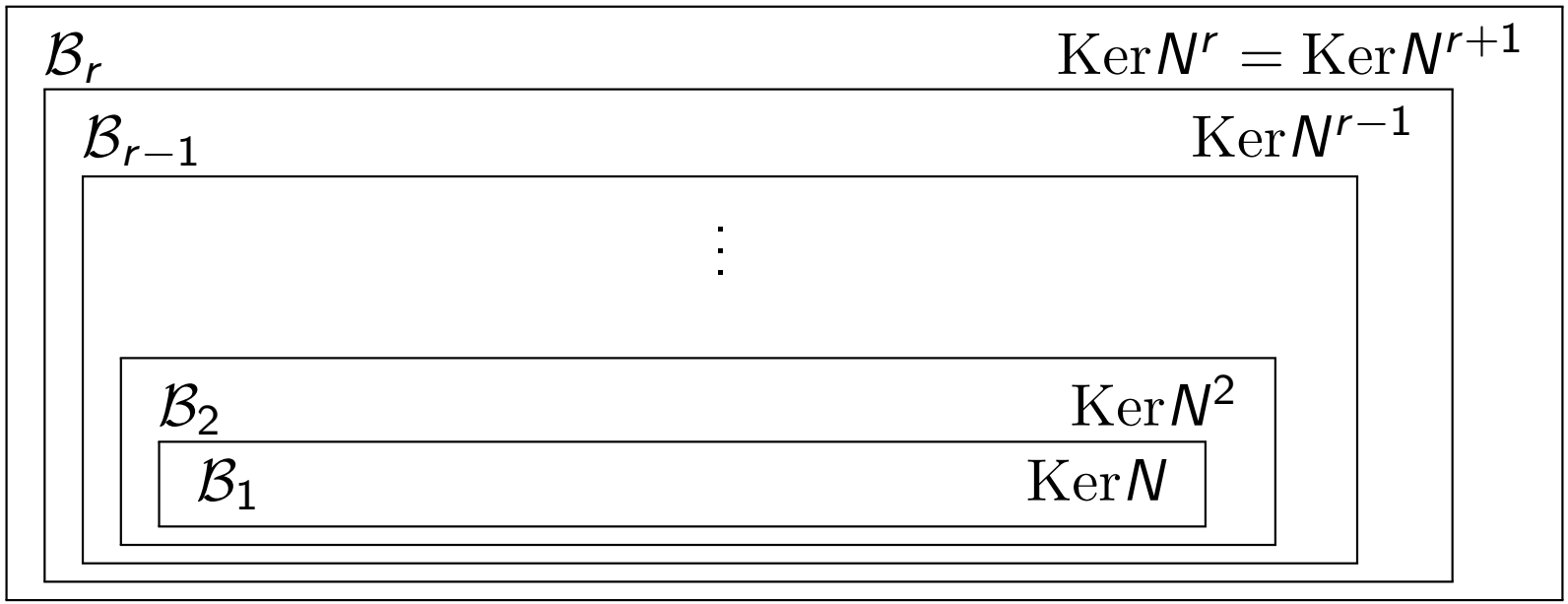
Korenski podprostor $\mathrm{Ker}(A-\lambda I)^r$, ki ustreza lastni vrednosti λ si predstavljamo kot veliko škatlo, v kateri gnezdijo majhne škatle. Najmanjša od teh škatel je lastni podprostor $\mathrm{Ker}(A-\lambda I)$ za λ.



Naj bo λ lastna vrednost matrike A. Označimo $N = A - \lambda I$. Opišimo konstrukcijo jordanske baze za korenski podprostor $\mathrm{Ker} N^r$.

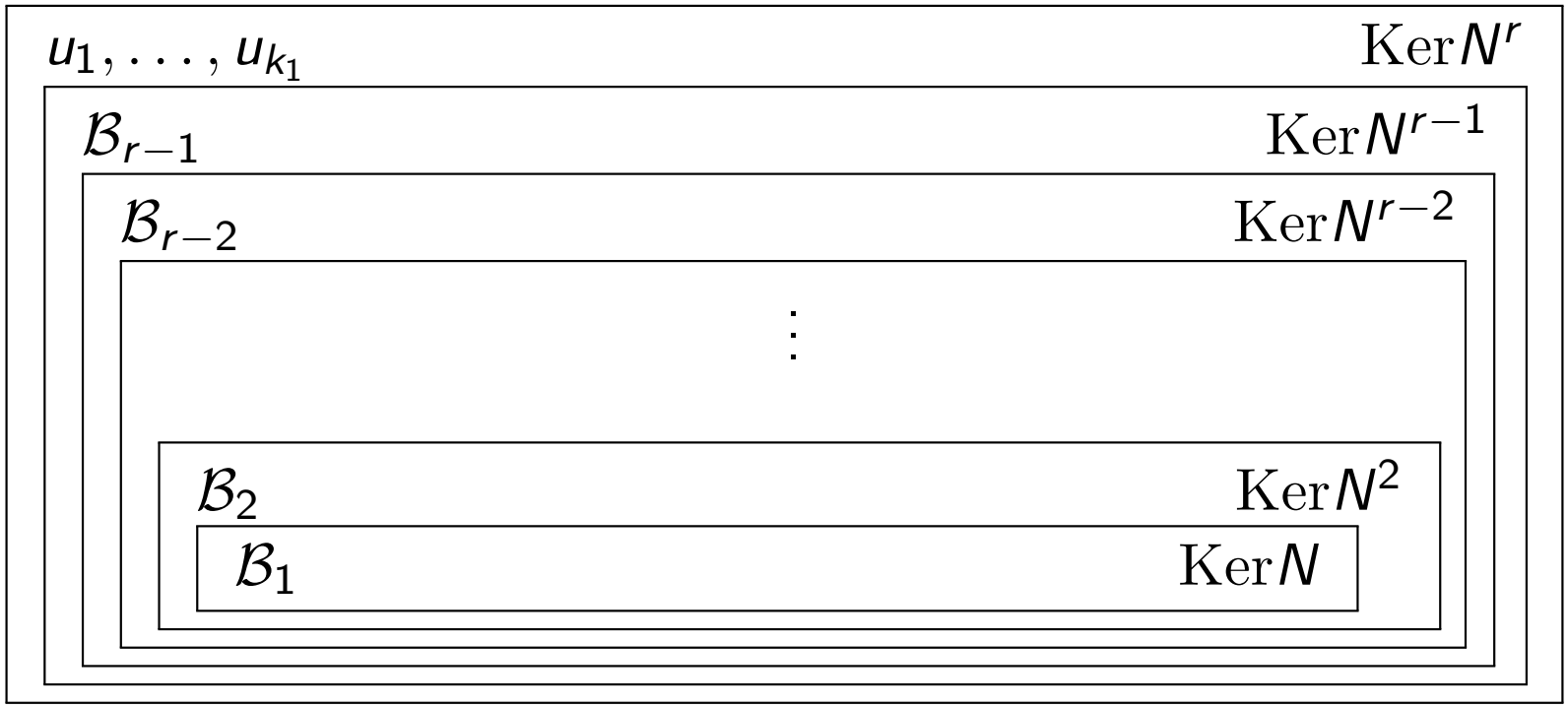
1. korak Računanje pomožnih baz.

Najprej za vsak $i = 1, \ldots, r$ izberemo poljubno bazo $\mathcal{B}_i$ za $\mathrm{Ker} N^i$.



2. korak Popravljanje pomožne baze $\mathcal{B}_r$.

Najprej izberimo take elemente $u_1, \ldots, u_{k_1} \in \mathcal{B}_r$, ki dopolnijo $\mathcal{B}_{r-1}$ do baze za $\mathrm{Ker} N^r$. Potem je $\mathcal{B}_{r-1} \cup \{u_1, \ldots, u_{k_1}\}$ popravek pomožne baze $\mathcal{B}_r$.

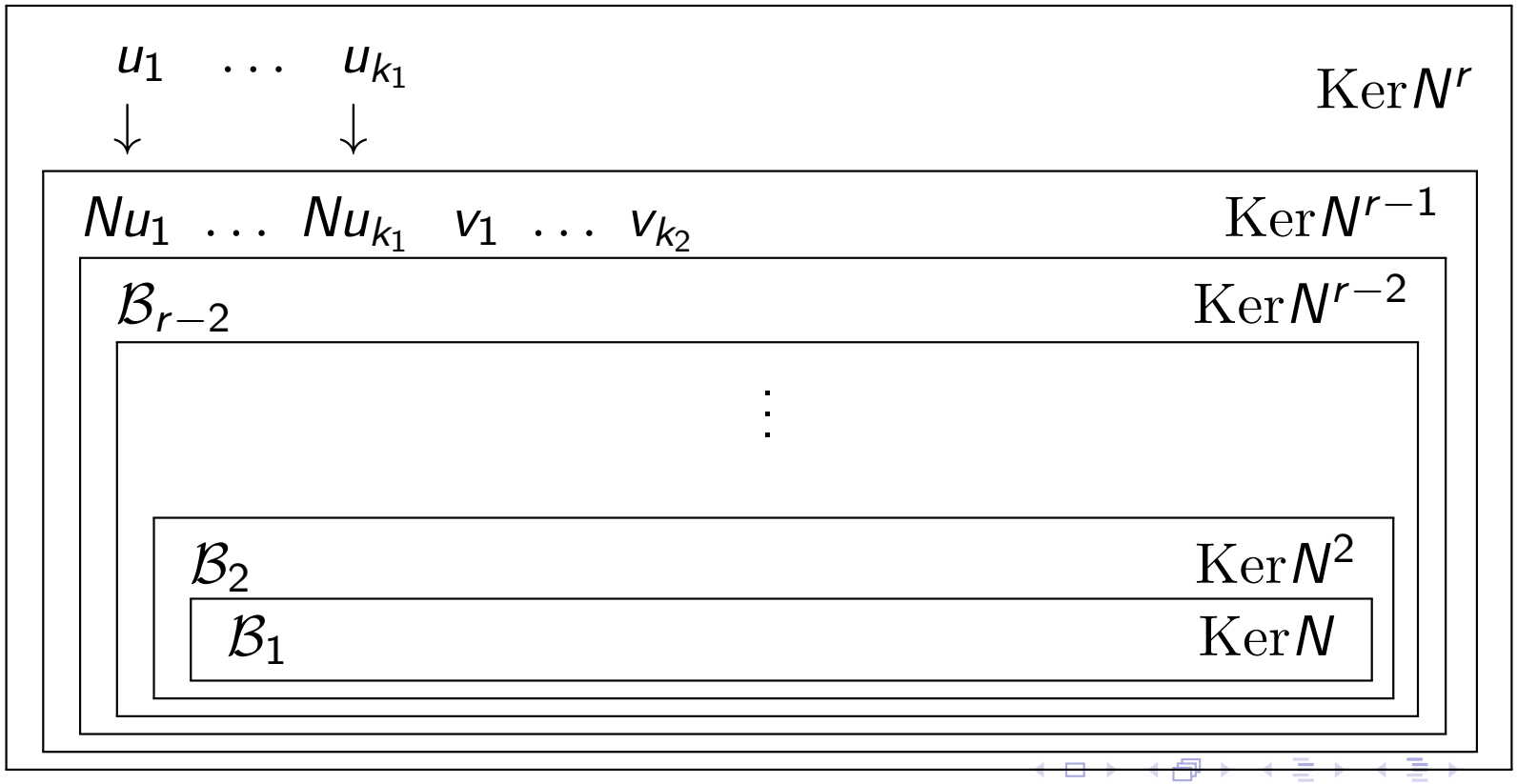


3. korak Popravljanje pomožne baze $\mathcal{B}_{r-1}$.

Najprej vektorje $u_1, \ldots, u_{k_1} \in \mathrm{Ker} N^r$ pomnožimo z matriko N.

- Dobljeni vektorji $Nu_1, \ldots, Nu_{k_1}$ ležijo v $\mathrm{Ker} N^{r-1}$.
- Množica $\mathcal{B}_{r-2} \cup \{Nu_1, \ldots, Nu_{k_1}\}$ je linearno neodvisna.

Izberimo take elemente $v_1, \ldots, v_{k_2} \in \mathcal{B}_{r-1}$, ki dopolnijo linearno neodvisno množico $\mathcal{B}_{r-2} \cup \{Nu_1, \ldots, Nu_{k_1}\}$ do baze za $\mathrm{Ker} N^{r-1}$. Potem je $\mathcal{B}_{r-2} \cup \{Nu_1, \ldots, Nu_{k_1}\} \cup \{v_1, \ldots, v_{k_2}\}$ popravek pomožne baze $\mathcal{B}_{r-1}$.

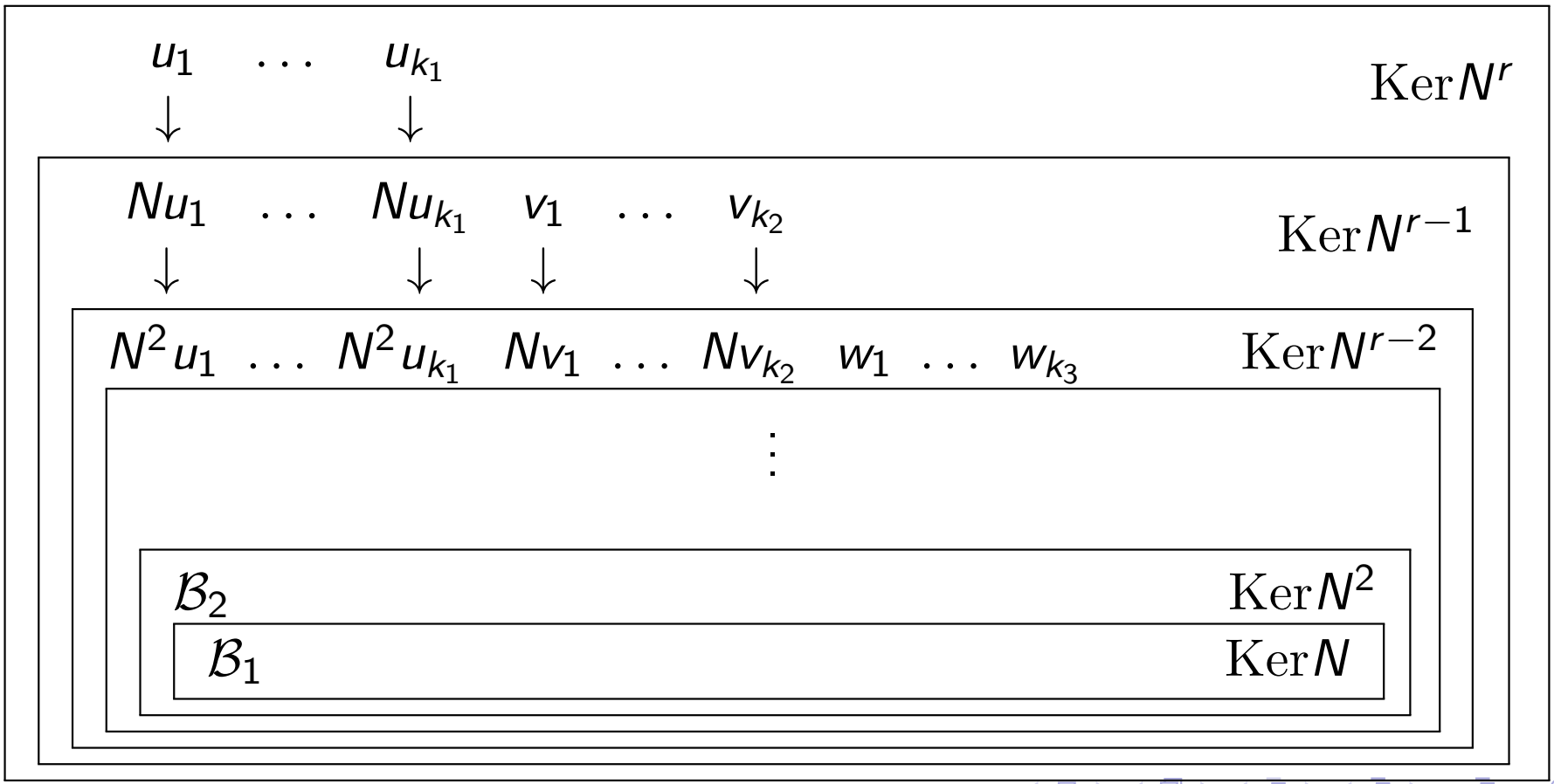


4. korak Popravljanje pomožne baze $\mathcal{B}_{r-2}$.

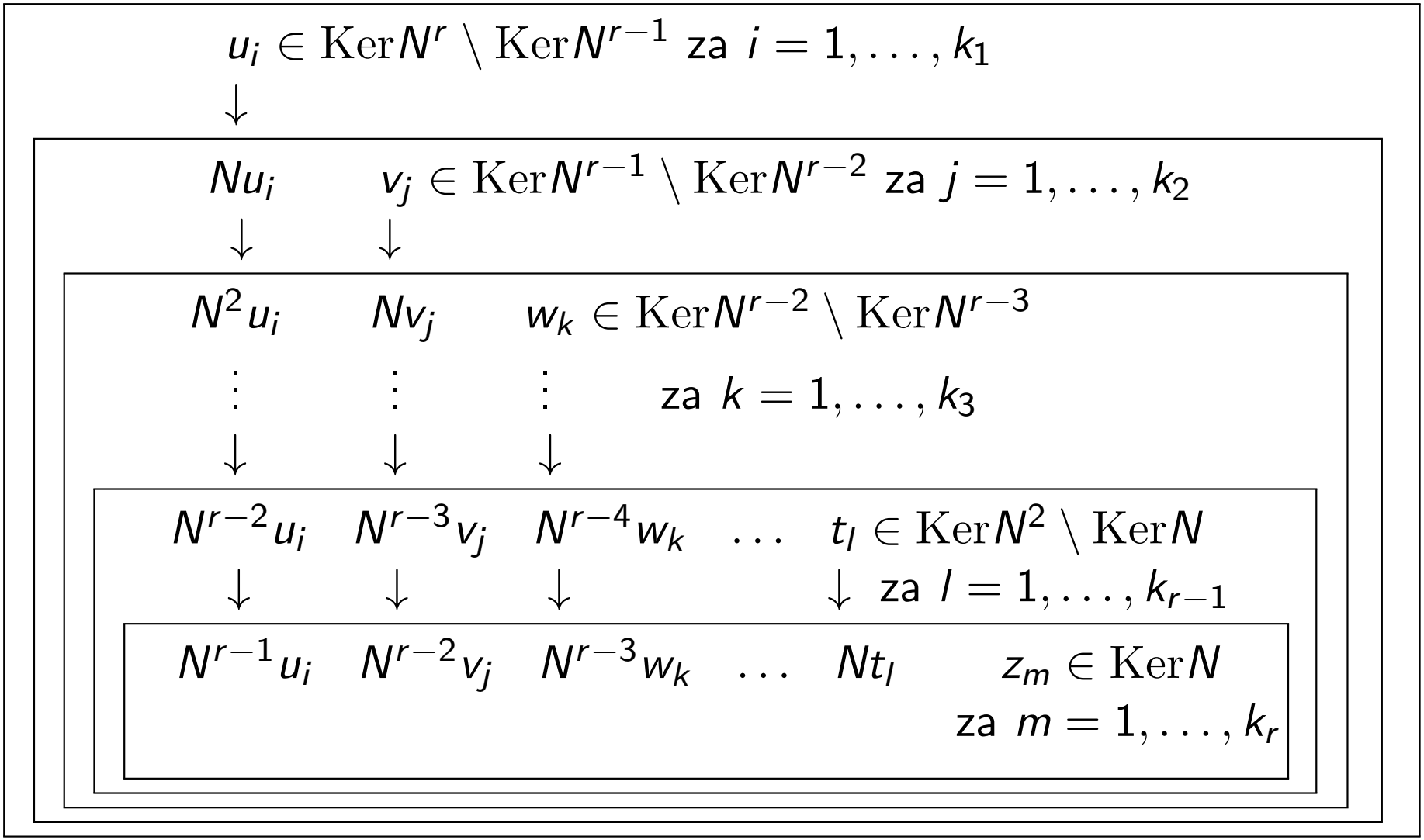
Najprej vektorje $Nu_1, \ldots, Nu_{k_1}, v_1, \ldots, v_{k_2} \in \mathrm{Ker}N^{r-1}$ pomnožimo z N.

- Dobljeni vektorji $N^2u_1, \ldots, N^2u_{k_1}, Nv_1, \ldots, Nv_{k_2}$ ležijo v $\mathrm{Ker}N^{r-2}$.
- Množica $\mathcal{B}_{r-3} \cup \{N^2u_1, \ldots, N^2u_{k_1}, Nv_1, \ldots, Nv_{k_2}\}$ je LN.

Izberimo take elemente $w_1, \ldots, w_{k_3} \in \mathcal{B}_{r-2}$, ki dopolnijo LN množico $\mathcal{B}_{r-3} \cup \{N^2u_1, \ldots, N^2u_{k_1}, Nv_1, \ldots, Nv_{k_2}\}$ do baze za $\mathrm{Ker}N^{r-2}$.



Postopek nadaljujemo, dokler ne popravimo vseh pomožnih baz. Na koncu dobimo naslednjo skico:



Vsak stolpec v zgornji skici nam da eno jordansko verigo. Imamo torej k_1 jordanskih verig dolžine r, k_2 jordanskih verig dolžine $r-1$, k_3 jordanskih verig dolžine $r-2$, $\ldots$, k_r jordanskih verig dolžine 1. Skupaj je to $k_1 + \ldots + k_r = \dim \mathrm{Ker} N$ jordanskih verig. Jordanskih verig za lastno vrednost λ je torej toliko kot je njena geometrijska večkratnost.

Treba je še utemeljiti, zakaj ta konstrukcija deluje. V vsakem od r korakov konstrukcije smo enkrat uporabili naslednjo trditev.

Trditev

Naj bo k naravno število in N matrika. Naj bo $\mathcal{B}_{k-1}$ baza za $\mathrm{Ker} N^{k-1}$, $\mathcal{B}_k$ baza za $\mathrm{Ker} N^k$ in naj bo $\mathcal{C}_{k+1}$ dopolnitev $\mathcal{B}_k$ do baze za $\mathrm{Ker} N^{k+1}$. Potem je množica $\mathcal{B}_{k-1} \cup N(\mathcal{C}_{k+1})$ linearno neodvisna, torej jo lahko dopolnimo do baze za $\mathrm{Ker} N^k$ s podmnožico od $\mathcal{B}_k$.

Dokaz: Naj bo $\mathcal{B}_{k-1} = \{e_1, \ldots, e_r\}$, $\mathcal{B}_k = \{f_1, \ldots, f_s\}$ in $\mathcal{C}_{k+1} = \{g_1, \ldots, g_t\}$. Ker je $\mathcal{C}_{k+1}$ podmnožica $\mathrm{Ker}N^{k+1}$, je $N(\mathcal{C}_{k+1}) := \{Ng_1, \ldots, Ng_t\}$ podmnožica $\mathrm{Ker}N^k$. Preverimo, da je množica $\mathcal{B}_{k-1} \cup N(\mathcal{C}_{k+1})$ linearno neodvisna. Recimo, da je

$$\alpha_1 e_1 + \ldots + \alpha_k e_k + \beta_1 Ng_1 + \ldots + \beta_t Ng_t = 0 \qquad (1)$$

Pomnožimo z N^{k-1} in upoštevajmo, da $e_1, \ldots, e_r \in \mathrm{Ker}N^{k-1}$. Dobimo $\beta_1 N^k g_1 + \ldots + \beta_t N^k g_t = 0$, torej $\beta_1 g_1 + \ldots + \beta_t g_t \in \mathrm{Ker}N^k$. Ker je $f_1, \ldots, f_s$ baza za $\mathrm{Ker}N^k$, lahko razvijemo

$$\beta_1 g_1 + \ldots + \beta_t g_t = \gamma_1 f_1 + \ldots + \gamma_s f_s.$$

Ker je $f_1, \ldots, f_s, g_1, \ldots, g_t$ baza za $\mathrm{Ker}N^{k+1}$, velja $\beta_1 = \ldots = \beta_t = 0$ (in $\gamma_1 = \ldots = \gamma_s = 0$), torej iz (1) dobimo $\alpha_1 e_1 + \ldots + \alpha_k e_k = 0$. Ker je $e_1, \ldots, e_r$ baza za $\mathrm{Ker}N^{k-1}$, odtod sledi $\alpha_1 = \ldots = \alpha_r = 0$. Torej je množica $\mathcal{B} := \{e_1, \ldots, e_r, Ng_1, \ldots, Ng_t\}$ res linearno neodvisna. Če je $\mathrm{Lin}\mathcal{B} = \mathrm{Ker}N^k$, potem končamo. Če $\mathrm{Lin}\mathcal{B} \neq \mathrm{Ker}N^k$ potem $\mathrm{Lin}\mathcal{B}$ ne more vsebovati vseh f_j, saj so ti baza za $\mathrm{Ker}N^k$. Torej obstaja tak indeks i_1, da $f_{i_1} \notin \mathrm{Lin}\mathcal{B}$. Podobno konstruiramo tak i_2, da $f_{i_2} \notin \mathrm{Lin}(\mathcal{B} \cup \{f_{i_1}\})$. S postopkom nadaljujemo dokler $\mathcal{B} \cup \{f_{i_1}, \ldots, f_{i_{s-t}}\}$ ni baza za $\mathrm{Ker}N^k$. □

Primer

Poišči jordansko kanonično formo matrike

$$A = \begin{bmatrix} 0 & 1 & -1 & 2 \\ 0 & 2 & 2 & 2 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

Rešitev: Najprej izračunamo karakteristični polinom

$$\det(A - xI) = x(x-2)^3.$$

Potem izračunamo lastne in korenske podprostore za lastno vrednost 0

$$\mathrm{Ker}A = \mathrm{Lin}\{\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}\} = \mathrm{Ker}A^2$$

in nato še za lastno vrednost 2

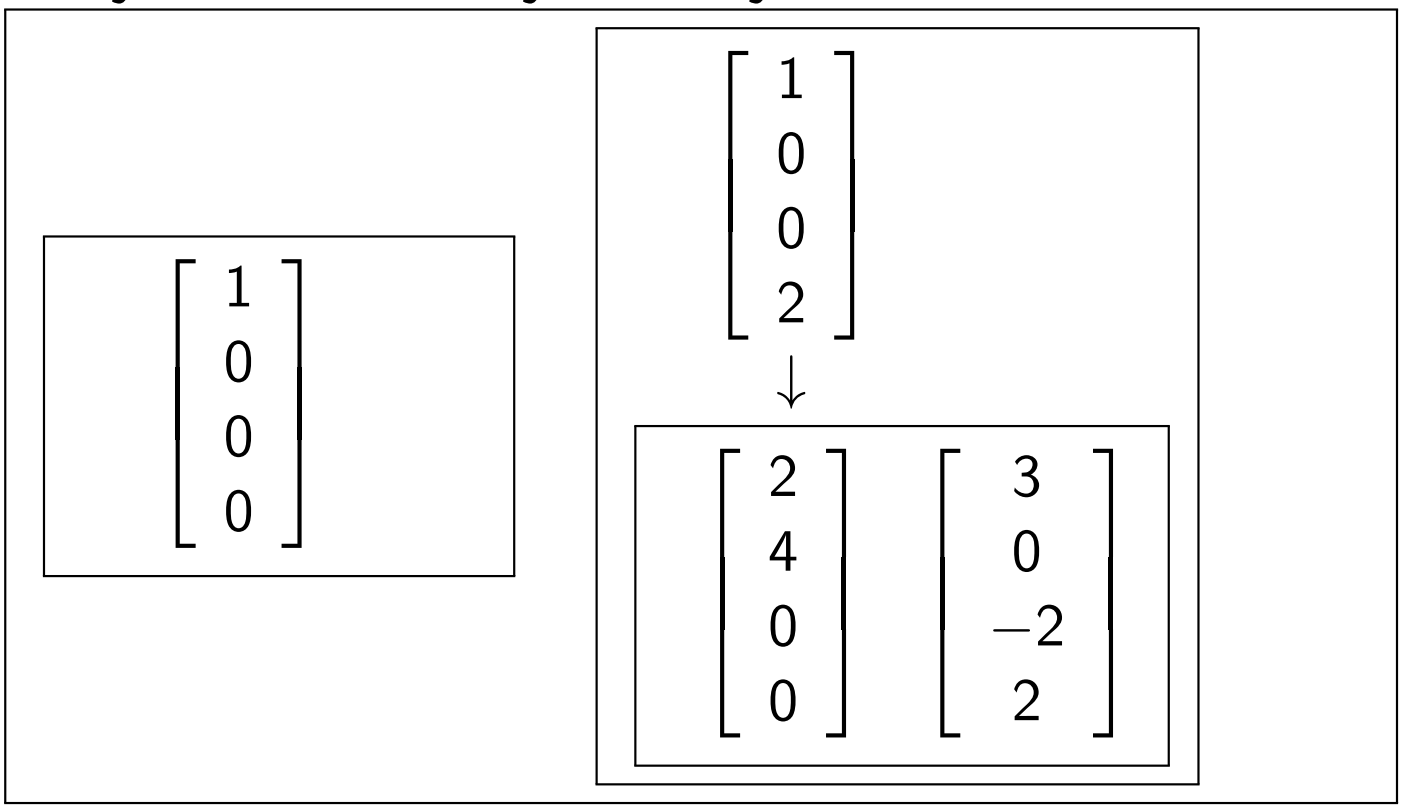
$$\mathrm{Ker}(A-2I) = \mathrm{Lin}\{\begin{bmatrix} 3 \\ 0 \\ -2 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 0 \\ 0 \end{bmatrix}\}$$

$$\mathrm{Ker}(A-2I)^2 = \mathrm{Lin}\{\begin{bmatrix} 1 \\ 0 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 0 \\ 0 \end{bmatrix}\} = \mathrm{Ker}(A-2I)^3$$

Opazimo, da je $\begin{bmatrix} 1 \\ 0 \\ 0 \\ 2 \end{bmatrix}$ dopolnitev baze $\mathrm{Ker}(A-2I)$ do baze $\mathrm{Ker}(A-2I)^2$

in da je $(A-2I)\begin{bmatrix} 1 \\ 0 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ 0 \\ 0 \end{bmatrix}$ linearno neodvisen od $\begin{bmatrix} 3 \\ 0 \\ -2 \\ 2 \end{bmatrix}$.

Torej imamo naslednjo situacijo.



Jordanske verige zložimo v matriko

$$P = \begin{bmatrix} 1 & 2 & 1 & 3 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & -2 \\ 0 & 0 & 2 & 2 \end{bmatrix}$$

Vsaki jordanski verigi dolžine k pripada ena jordanska kletka velikosti $k \times k$. Torej je jordanska kanonična forma matrike A enaka

$$J = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}.$$

Matriko A torej lahko izrazimo kot

$$A = PJP^{-1}.$$

7. Funkcije matrik

Če poznamo razcep $A = PJP^{-1}$ matrike A, potem se računanje potenc matrike A prevede na računanje potenc matrike J, saj velja

$$A^n = (PJP^{-1})(PJP^{-1}) \cdots (PJP^{-1}) = PJ^nP^{-1}.$$

Ker je J bločno diagonalna matrika sestavljena iz Jordanskih kletk, se potenciranje matrike J prevede na potenciranje Jordanskih kletk.

Potenciranje matrik je uporabno pri reševanju linearnih rekurzivnih enačb.

Primer sistema linearnih rekurzivnih enačb

$$x_{n+1} = -4x_n + 4y_n, \quad x_0 = 0$$

$$y_{n+1} = -x_n, \qquad y_0 = 1.$$

Najprej sistem zapišemo v matrični obliki

$$\begin{bmatrix} x_{n+1} \\ y_{n+1} \end{bmatrix} = A \begin{bmatrix} x_n \\ y_n \end{bmatrix}, \quad A = \begin{bmatrix} -4 & 4 \\ -1 & 0 \end{bmatrix}$$

Odtod sledi

$$\begin{bmatrix} x_n \\ y_n \end{bmatrix} = A^n \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} = \begin{bmatrix} -4 & 4 \\ -1 & 0 \end{bmatrix}^n \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Kot v prejšnjem razdelku izračunamo jordansko kanonično formo za A

$$A = PJP^{-1} = \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} -2 & 1 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 1 & -2 \end{bmatrix}^{-1}$$

S popolno indukcijo pokažemo, da je

$$J^n = \begin{bmatrix} -2 & 1 \\ 0 & -2 \end{bmatrix}^n = \begin{bmatrix} (-2)^n & n(-2)^{n-1} \\ 0 & (-2)^n \end{bmatrix}.$$

Odtod sledi, da je

$$A^n = PJ^nP^{-1} = \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} (-2)^n & n(-2)^{n-1} \\ 0 & (-2)^n \end{bmatrix} \begin{bmatrix} 2 & -3 \\ -1 & 2 \end{bmatrix}$$

$$= \begin{bmatrix} (-2)^n - 2n(-2)^{n-1} & 4n(-2)^{n-1} \\ -n(-2)^{n-1} & (-2)^n + 2n(-2)^{n-1} \end{bmatrix}$$

kar lahko še naprej poenostavimo. Če A^n pomnožimo z $\begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$, dobimo

$$\begin{bmatrix} x_n \\ y_n \end{bmatrix} = \begin{bmatrix} 4n(-2)^{n-1} \\ (-2)^n + 2n(-2)^{n-1} \end{bmatrix} = (-2)^n \begin{bmatrix} -2n \\ 1-n \end{bmatrix}$$

Oglejmo si, kako izračunamo potenco jordanske kletke.

Formula za potenco $k \times k$ jordanske kletke

$$\begin{bmatrix} \lambda & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & 1 \\ 0 & \dots & \dots & \dots & \lambda \end{bmatrix}^n = \begin{bmatrix} \lambda^n & n\lambda^{n-1} & \binom{n}{2}\lambda^{n-2} & \dots & \binom{n}{k-1}\lambda^{n-k+1} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \binom{n}{2}\lambda^{n-2} \\ \vdots & & & \ddots & n\lambda^{n-1} \\ 0 & \dots & \dots & \dots & \lambda^n \end{bmatrix}$$

Dokaz: Jordansko kletko lahko zapišemo v obliki $\lambda I + N$, kjer je N matrika, ki ima na prvi naddiagonali same enke, drugod pa same ničle. Potence matrike N je preprosto izračunati. Opazimo, da je N^i matrika, ki ima na i-ti naddiagonali same enke, drugod pa same ničle. Ker je N $k \times k$ matrika, nima k-te naddiagonale, zato je $N^k = 0$. Po binomski formuli

$$(\lambda I + N)^n = \sum_{i=1}^{n} \binom{n}{i} \lambda^{n-i} N^i = \sum_{i=1}^{k-1} \binom{n}{i} \lambda^{n-i} N^i \quad \square$$

Iz Analize vemo, da se dajo številne funkcije izraziti s potenčno vrsto, recimo

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

Ali lahko našo metodo za računanje potenc matrike posplošimo na potenčne vrste? Izkaže se, da lahko. Vzemimo recimo funkcijo

$$f(x) = \sum_{n=0}^{\infty} c_n x^n$$

in vstavimo $A = PJP^{-1}$ namesto x. Dobimo

$$f(A) = \sum_{n=0}^{\infty} c_n A^n = \sum_{n=0}^{\infty} c_n P J^n P^{-1} = P(\sum_{n=0}^{\infty} c_n J^n) P^{-1} = P f(J) P^{-1}$$

Računanje $f(J)$, kjer je J jordanska matrika, lahko prevedemo na primer, ko je J jordanska kletka. Velja namreč

$$f\left(\begin{bmatrix} J_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & J_m \end{bmatrix}\right) = \begin{bmatrix} f(J_1) & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & f(J_m) \end{bmatrix}.$$

Pri računanju funkcije jordanske kletke si pomagamo z naslednjo formulo.

Formula za funkcijo $k \times k$ jordanske kletke

$$f\left(\begin{bmatrix} \lambda & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & 1 \\ 0 & \dots & \dots & \dots & \lambda \end{bmatrix}\right) = \begin{bmatrix} f(\lambda) & f'(\lambda) & \frac{f''(\lambda)}{2} & \dots & \frac{f^{k-1}(\lambda)}{(k-1)!} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \frac{f''(\lambda)}{2} \\ \vdots & & & \ddots & f'(\lambda) \\ 0 & \dots & \dots & \dots & f(\lambda) \end{bmatrix}$$

Za $f(x) = x^n$ je to ravno formula za potenco jordanske kletke. Ker je formula linearna v funkciji f, odtod sledi, da velja za vse potenčne vrste.

Oglejmo si primer uporabe funkcij matrik. Iz Analize vemo, da je funkcija $y(t) = ce^{at}$ rešitev diferencialne enačbe $y'(t) = ay(t)$. To nam da idejo za reševanje naslednjega sistema diferencialnih enačb:

$$\begin{aligned} y_1'(t) &= a_{1,1}y_1(t) + \ldots + a_{1,d}y_d(t) \\ &\vdots \\ y_d'(t) &= a_{d,1}y_1(t) + \ldots + a_{d,d}y_d(t) \end{aligned}$$

Ta sistem najprej zapišemo v matrični obliki kot

$$\mathbf{y}'(t) = A\,\mathbf{y}(t).$$

Izkaže se, da je njegova rešitev vektorska funkcija

$$\mathbf{y}(t) = e^{At}\mathbf{c},$$

kjer je $\mathbf{c}$ konstanten vektor.

Matrično funkcijo e^{At} izračunamo tako, da v funkcijo $f(x) = e^{tx}$ vstavimo A namesto x. Kot zgoraj to prevedemo na primer e^{Jt}, kjer je J jordanska kletka. V tem primeru iz formule za funkcijo jordanske kletke dobimo

$$e^{Jt} = \begin{bmatrix} e^{\lambda t} & te^{\lambda t} & \frac{t^2}{2}e^{\lambda t} & \dots & \frac{t^{k-1}}{(k-1)!}e^{\lambda t} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \frac{t^2}{2}e^{\lambda t} \\ \vdots & & & \ddots & te^{\lambda t} \\ 0 & \dots & \dots & \dots & e^{\lambda t} \end{bmatrix}$$

Vektorski prostori s skalarnim produktom - 1. del

1. Skalarni produkt

V tem razdelku bomo definirali skalarni produkt na realnih in kompleksnih vektorskih prostorih ter si ogledali nekaj primerov.

Začnimo z realnimi vektorskimi prostori.

Definicija skalarnega produkta nad $\mathbb{R}$

Naj bo V vektorski prostor nad obsegom realnih števil $\mathbb{R}$. Preslikava, ki vsakemu paru vektorjev $u, v \in V$ priredi realno število $\langle u, v \rangle$ je **skalarni produkt**, če zadošča naslednjim lastnostim:

1. Za vsak neničeln $v \in V$ velja $\langle v, v \rangle > 0$.
2. Za vsaka $u, v \in V$ velja $\langle v, u \rangle = \langle u, v \rangle$.
3. Za vsake $u_1, u_2, v \in V$ ter $\alpha_1, \alpha_2 \in \mathbb{R}$ velja $\langle \alpha_1 u_1 + \alpha_2 u_2, v \rangle = \alpha_1 \langle u_1, v \rangle + \alpha_2 \langle u_2, v \rangle$.

Opomba: Prvi lastnosti pravimo pozitivna definitnost, drugi simetričnost, tretji pa linearnost v prvem faktorju.

Oglejmo si nekaj preprostih posledic aksiomov:

- Za vse $u, v_1, v_2 \in V$ ter vse $\beta_1, \beta_2 \in \mathbb{R}$ velja

$$\langle u, \beta_1 v_1 + \beta_2 v_2 \rangle = \beta_1 \langle u, v_1 \rangle + \beta_2 \langle u, v_2 \rangle$$

  Tej lastnosti pravimo linearnost v drugem faktorju. Dokažemo jo takole: $\langle u, \beta_1 v_1 + \beta_2 v_2 \rangle = \langle \beta_1 v_1 + \beta_2 v_2, u \rangle = \beta_1 \langle v_1, u \rangle + \beta_2 \langle v_2, u \rangle = \beta_1 \langle u, v_1 \rangle + \beta_2 \langle u, v_2 \rangle$ kjer smo pri prvem in tretjem enačaju uporabili simetričnost, pri drugem enačaju pa linearnost v prvem faktorju.

- Za vsak $v \in V$ velja

$$\langle v, 0 \rangle = \langle 0, v \rangle = 0.$$

  Če v linearnosti v prvem faktorju vstavimo $\alpha_1 = \alpha_2 = 0$, dobimo $\langle 0, v \rangle = \langle 0 \cdot u_1 + 0 \cdot u_2, v \rangle = 0 \langle u_1, v \rangle + 0 \langle u_2, v \rangle = 0$. Iz simetričnosti sledi $\langle v, 0 \rangle = 0$.

- Za vsak $v \in V$ velja $\langle v, v \rangle \geq 0$. Enačaj velja natanko tedaj, ko je $v = 0$. To sledi iz pozitivne definitnosti in iz $\langle 0, 0 \rangle = 0$.

Oglejmo si primere skalarnih produktov na realnih vektorskih prostorih.

Primer: Standardni skalarni produkt na $\mathbb{R}^n$

Običajnemu skalarnemu produktu iz prvega semestra bomo tu pravili **standardni skalarni produkt** na $\mathbb{R}^n$. Definiran je z

$$\langle (\alpha_1, \ldots, \alpha_n), (\beta_1, \ldots, \beta_n) \rangle = \alpha_1\beta_1 + \ldots + \alpha_n\beta_n$$

Pokazali smo že, da zadošča lastnostim (1),(2),(3).

Seveda standardni skalarni produkt ni edini skalarni produkt na $\mathbb{R}^n$. Klasifikacijo vseh skalarnih produktov na $\mathbb{R}^n$ bomo obdelali v naslednjem poglavju. Tu pokažimo samo, da jih je neskončno.

Dodatni primeri skalarnih produktov na $\mathbb{R}^n$

Za vsako n-terico $\delta = (\delta_1, \ldots, \delta_n)$ strogo pozitivnih realnih števil je s

$$\langle (\alpha_1, \ldots, \alpha_n), (\beta_1, \ldots, \beta_n) \rangle = \delta_1\alpha_1\beta_1 + \ldots + \delta_n\alpha_n\beta_n$$

definiran skalarni produkt na $\mathbb{R}^n$. (Preveri lastnosti (1),(2),(3)!)

Na funkcijskih prostorih je skalarni integral običajno definiran z integralom.

Primer: Standardni skalarni produkt na $\mathcal{C}[a, b]$

Za dve zvezni funkciji f, g iz intervala $[a, b]$ v $\mathbb{R}$ definirajmo.

$$\langle f, g \rangle = \int_a^b f(x)g(x)\,dx.$$

Preverimo lastnosti (1),(2),(3), torej je to skalarni produkt.

Konstruirajmo še neskončno mnogo drugih skalarnih produktov na $\mathcal{C}[a, b]$.

Dodatni primeri skalarnih produktov na $\mathcal{C}[a, b]$

Naj bo $w \in \mathcal{C}[a, b]$ taka funkcija, ki zadošča $w(x) > 0$ za vsak $x \in [a, b]$. S

$$\langle f, g \rangle = \int_a^b w(x)f(x)g(x)\,dx.$$

je definiran skalarni produkt na $\mathbb{R}^n$. (Preveri lastnosti (1),(2),(3)!)

Razširimo zdaj definicijo na kompleksne vektorske prostore. Spomnimo se, da za $z = a + bi$ (kjer $a, b \in \mathbb{R}$ in $i^2 = -1$) definiramo $\bar{z} = a - bi$.

Definicija skalarnega produkta nad $\mathbb{C}$

Naj bo V vektorski prostor nad obsegom kompleksnih števil $\mathbb{C}$. Preslikava, ki vsakemu paru vektorjev $u, v \in V$ priredi kompleksno število $\langle u, v \rangle$ je **skalarni produkt**, če zadošča naslednjim lastnostim:

1. Za vsak neničeln $v \in V$ velja $\langle v, v \rangle \in \mathbb{R}$ in $\langle v, v \rangle > 0$.
2. Za vsaka $u, v \in V$ velja $\langle v, u \rangle = \overline{\langle u, v \rangle}$.
3. Za vsake $u_1, u_2, v \in V$ ter $\alpha_1, \alpha_2 \in \mathbb{C}$ velja $\langle \alpha_1 u_1 + \alpha_2 u_2, v \rangle = \alpha_1 \langle u_1, v \rangle + \alpha_2 \langle u_2, v \rangle$.

Opomba: Lastnosti (2) pravimo konjugirana simetričnost. Lastnosti (1) in (3) sta enaki kot v realnem primeru.

Opomba: Pomembna razlika z realnim primerom je, da tu nimamo linearnosti v drugem faktorju ampak konjugirano linearnost, se pravi

$$\langle u, \beta_1 v_1 + \beta_2 v_2 \rangle = \bar{\beta}_1 \langle u, v_1 \rangle + \bar{\beta}_2 \langle u, v_2 \rangle$$

za vse $u, v_1, v_2 \in V$ ter vse $\beta_1, \beta_2 \in \mathbb{C}$. To sledi iz $\langle u, \beta_1 v_1 + \beta_2 v_2 \rangle = \overline{\langle \beta_1 v_1 + \beta_2 v_2, u \rangle} = \overline{\beta_1 \langle v_1, u \rangle + \beta_2 \langle v_2, u \rangle} = \bar{\beta}_1 \overline{\langle v_1, u \rangle} + \bar{\beta}_2 \overline{\langle v_2, u \rangle} = \\ = \bar{\beta}_1 \langle u, v_1 \rangle + \bar{\beta}_2 \langle u, v_2 \rangle$. Pri prvem in četrtem enačaju smo upoštevali konjugirano simetričnost, pri drugem enačaju linearnost v prvem faktorju. Pri tretjem enačaju smo upoštevali $\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2$ in $\overline{z_1 z_2} = \bar{z}_1 \bar{z}_2$.

Opomba: Tudi tu velja $\langle v, 0 \rangle = \langle 0, v \rangle = 0$ za vsak $v \in V$.

Opomba: Tudi tu velja $\langle v, v \rangle \geq 0$ za vsak $v \in V$. Enačaj velja natanko tedaj, ko je $v = 0$.

Primer: Standardni skalarni produkt na $\mathbb{C}^n$.

Za dve kompleksni n-terici definirajmo njun **standardni skalarni produkt**

$$\langle(\alpha_1, \ldots, \alpha_n), (\beta_1, \ldots, \beta_n)\rangle = \alpha_1\bar{\beta}_1 + \ldots + \alpha_n\bar{\beta}_n.$$

Pozor, za razliko od realnega primera imamo tu $\bar{\beta}_i$ namesto β_i.

Opomba: Če $\sum_i \alpha_i\bar{\beta}_i$ zamenjamo z $\sum_i \delta_i\alpha_i\bar{\beta}_i$, kjer so $\delta_i > 0$ fiksni, dobimo drug skalarni produkt na $\mathbb{C}^n$.

Standardni skalarni produkt na $\mathcal{C}([a, b], \mathbb{C})$

Za dve zvezni funkciji f, g iz intervala $[a, b]$ v $\mathbb{C}$ definirajmo.

$$\langle f, g\rangle = \int_a^b f(x)\overline{g(x)}\, dx.$$

Pozor, za razliko od realnega primera imamo tu $\overline{g(x)}$ namesto $g(x)$.

Opomba: Če vzamemo $\int_a^b w(x)f(x)\overline{g(x)}\, dx$, kjer je $w(x)$ zvezna in za vsak x realna in strogo pozitivna, potem dobimo drug skalarni produkt.

2. Norma porojena iz skalarnega produkta

Definicija norme je podobna kot v prvem semestru.

Definicija norme

Naj bo V realen ali kompleksen vektorski prostor s skalarnim produktom. Za vsak element $v \in V$ definirajmo njegovo **normo** z

$$\|v\| = \sqrt{\langle v, v \rangle}.$$

Dokažimo najprej pomožno trditev.

Trditev (Cauchy-Schwartzova neenakost)

Naj bo V vektorski prostor s skalarnim produktom. Za vsaka $u, v \in V$ je

$$|\langle u, v \rangle| \leq \|u\| \|v\|$$

V dokazu bomo večkrat potrebovali, da je $z\bar{z} = |z|^2$ za vsak $z \in \mathbb{C}$.

Dokaz: Naj bo $\alpha = \langle v, v \rangle$, $\beta = \langle u, v \rangle$ in $w = \alpha u - \beta v$. Potem velja

$$0 \leq \langle w, w \rangle = \alpha\bar{\alpha}\langle u, u \rangle - \alpha\bar{\beta}\langle u, v \rangle - \beta\bar{\alpha}\langle v, u \rangle + \beta\bar{\beta}\langle v, v \rangle.$$

Opazimo, da je

$$\alpha\bar{\beta}\langle u, v \rangle = \beta\bar{\alpha}\langle v, u \rangle = \beta\bar{\beta}\langle v, v \rangle = |\langle u, v \rangle|^2 \langle v, v \rangle = |\langle u, v \rangle|^2 \|v\|^2$$

in

$$\alpha\bar{\alpha}\langle u, u \rangle = |\langle v, v \rangle|^2 \langle u, u \rangle = \|v\|^4 \|u\|^2.$$

Torej je

$$0 \leq \langle w, w \rangle = \|v\|^4 \|u\|^2 - |\langle u, v \rangle|^2 \|v\|^2.$$

Če je $v \neq 0$, lahko krajšamo $\|v\|^2$ in dobimo

$$|\langle u, v \rangle|^2 \leq \|v\|^2 \|u\|^2.$$

S korenjenjem potem dobimo Cauchy-Schwartzovo neenakost. Če je $v = 0$, potem Cauchy-Schwartzova neenakost očitno drži. □

Trditev (Osnovne lastnosti norme)

Naj bo V vektorski prostor s skalarnim produktom. Pripadajoča norma zadošča naslednjim lastnostim:

1. Za vsak neničeln $v \in V$ je $\|v\| > 0$.
2. Za vsak $v \in V$ in vsak skalar α je $\|\alpha v\| = |\alpha|\|v\|$.
3. Za vsaka $u, v \in V$ je $\|u + v\| \leq \|u\| + \|v\|$.

Dokaz: Lastnost (1) sledi iz pozitivne definitnosti skalarnega produkta.

Lastnost (2) sledi iz $\|\alpha v\|^2 = \langle \alpha v, \alpha v \rangle = \alpha\bar{\alpha}\langle v, v \rangle = |\alpha|^2\|v\|^2$, kjer smo upoštevali linearnost v prvem in konjugirano linearnost v drugem faktorju.

Lastnost (3) sledi iz $\|u + v\|^2 = \langle u + v, u + v \rangle =$
$\langle u, u \rangle + \langle u, v \rangle + \langle v, u \rangle + \langle v, v \rangle = \langle u, u \rangle + 2\mathrm{Re}(\langle u, v \rangle) + \langle v, v \rangle \leq$
$\|u\|^2 + 2|\langle u, v \rangle| + \|v\|^2 \leq \|u\|^2 + 2\|u\|\|v\| + \|v\|^2 = (\|u\| + \|v\|)^2$.
Upoštevali smo, da za $z = a + ib$ velja $z + \bar{z} = 2a \leq 2\sqrt{a^2 + b^2} = 2|z|$.
Uporabili smo tudi Cauchy-Schwartzovo neenakost. □

Vemo, kako se norma izraža s skalarnim produktom. Ali znamo tudi skalarni produkt izraziti z normo? Odgovor nam daje:

Trditev (Polarizacijske identitete)

1. Če je V vektorski prostor nad $\mathbb{R}$, potem za vsaka $u, v \in V$ velja

$$\langle u, v \rangle = \frac{1}{4} \left(\|u + v\|^2 - \|u - v\|^2 \right)$$

2. Če je V vektorski prostor nad $\mathbb{C}$, potem za vsaka $u, v \in V$ velja

$$\langle u, v \rangle = \frac{1}{4} \sum_{k=0}^{3} i^k \|u + i^k v\|^2$$

Dokaz: V realnem primeru velja

$$\sum_{k=0}^{1}(-1)^k\|u+(-1)^k v\|^2=\sum_{k=0}^{1}(-1)^k\langle u+(-1)^k v,u+(-1)^k v\rangle=$$

$$=\sum_{k=0}^{1}\Big((-1)^k\langle u,u\rangle+(-1)^{2k}\langle v,u\rangle+(-1)^{2k}\langle u,v\rangle+(-1)^{3k}\langle v,v\rangle\Big)=$$

$$=\Big(\sum_{k=0}^{1}(-1)^k\Big)\langle u,u\rangle+\Big(\sum_{k=0}^{1}(-1)^{2k}\Big)\langle v,u\rangle+\Big(\sum_{k=0}^{1}(-1)^{2k}\Big)\langle u,v\rangle+$$

$$+\Big(\sum_{k=0}^{1}(-1)^{3k}\Big)\langle v,v\rangle=2\langle v,u\rangle+2\langle u,v\rangle=4\langle u,v\rangle,$$

pri čemer smo upoštevali, da je

$$\sum_{k=0}^{1}(-1)^k=\sum_{k=0}^{1}(-1)^{3k}=0\quad\text{in}\quad\sum_{k=0}^{1}(-1)^{2k}=2.$$

V kompleksnem primeru velja

$$\sum_{k=0}^{3} i^k \|u + i^k v\|^2 = \sum_{k=0}^{3} i^k \langle u + i^k v, u + i^k v \rangle =$$

$$= \sum_{k=0}^{3} \left(i^k \langle u, u \rangle + i^{2k} \langle v, u \rangle + i^k \bar{i}^k \langle u, v \rangle + i^{2k} \bar{i}^k \langle v, v \rangle \right) =$$

$$= \left(\sum_{k=0}^{3} i^k \right) \langle u, u \rangle + \left(\sum_{k=0}^{3} i^{2k} \right) \langle v, u \rangle + \left(\sum_{k=0}^{3} i^k \bar{i}^k \right) \langle u, v \rangle +$$

$$+ \left(\sum_{k=0}^{3} \left(i^{2k} \bar{i}^k \right) \langle v, v \rangle = 4 \langle u, v \rangle, \right.$$

pri čemer smo upoštevali, da je

$$\sum_{k=0}^{3} i^k = \sum_{k=0}^{3} i^{2k} = \sum_{k=0}^{3} i^{2k} \bar{i}^k = 0 \quad \text{in} \quad \sum_{k=0}^{3} i^k \bar{i}^k = 4.$$

3. Ortogonalne baze

Kot med dvema vektorjema iz $\mathbb{R}^n$ je pravi natanko tedaj, ko je njun standardni skalarni produkt enak nič. Za splošne skalarne produkte nimamo več geometrijske intuicije, zato pravokotnost definiramo s pomočjo skalarnega produkta.

Definicija ortogonalnosti vektorjev

Naj bo V vektorski prostor s skalarnim produktom. Vektorja $u, v \in V$ sta **pravokotna** (s tujko ortogonalna), če velja $\langle u, v \rangle = 0$.

Opomba: Ničelni vektor je pravokoten na vse vektorje, zato ni zanimiv.

Opomba: Neničeln vektor ne more biti pravokoten sam nase.

Primer

Vektorja $(1, 2)$ in $(2, -3)$ iz $\mathbb{R}^2$ sta pravokotna glede na skalarni produkt

$$\langle (\alpha_1, \alpha_2), (\beta_1, \beta_2) \rangle = 3\alpha_1\beta_1 + \alpha_2\beta_2,$$

nista pa pravokotna glede na standardni skalarni produkt.

Definicija ortogonalne množice

Naj bo V vektorski prostor s skalarnim produktom. Množica vektorjev iz V je **ortogonalna**, če ne vsebuje ničelnega vektorja in če sta vsaka dva različna vektorja iz te množice pravokotna.

Primer

Vektorji $(1,1,1,1,1,1,1,1)$, $(-1,-1,-1,-1,1,1,1,1)$, $(-1,-1,1,1,-1,-1,1,1)$ in $(-1,1,-1,1,-1,1,-1,1)$ tvorijo ortogonalno množico v $\mathbb{R}^8$ za standardni skalarni produkt.

Primer

Funkcije $f_k(x) = \sin \frac{k\pi x}{a}$, kjer $k = 1, 2, 3, \ldots$, tvorijo neskončno ortogonalno množico v $\mathcal{C}[0, a]$ za standardni skalarni produkt.

Dokaz: Velja $\langle f_k(x), f_l(x) \rangle = \int_0^a \sin \frac{k\pi x}{a} \sin \frac{l\pi x}{a}\, dx =$
$\frac{1}{2} \int_0^a \cos \frac{(k-l)\pi x}{a}\, dx - \frac{1}{2} \int_0^a \cos \frac{(k+l)\pi x}{a}\, dx = 0$, če $k \neq l$.

Trditev

Vsaka ortogonalna množica je linearno neodvisna.

Dokaz: Recimo, da so $v_1, \ldots, v_k$ neničelni paroma pravokotni vektorji iz V. Če je $\alpha_1 v_1 + \ldots + \alpha_k v_k = 0$ za neke skalarje $\alpha_1, \ldots, \alpha_k$, moramo pokazati, da velja $\alpha_1 = \ldots = \alpha_k = 0$. Opazimo, da za vsak i velja $0 = \langle 0, v_i \rangle =$
$= \langle \alpha_1 v_1 + \ldots + \alpha_k v_k, v_i \rangle = \alpha_1 \langle v_1, v_i \rangle + \ldots + \alpha_i \langle v_i, v_i \rangle + \ldots + \alpha_k \langle v_k, v_i \rangle$.
Ker je $\langle v_1, v_i \rangle = \ldots = \langle v_{i-1}, v_i \rangle = \langle v_{i+1}, v_i \rangle = \ldots = \langle v_k, v_i \rangle = 0$, je
$0 = \alpha_i \langle v_i, v_i \rangle$. Ker je $v_i \neq 0$ za vsak i, odtod sledi $\alpha_i = 0$ za vsak i. □

Definicija ortogonalne baze

Ortogonalna množica v V, ki je ogrodje za V, je **ortogonalna baza** za V.

Primer

Standardna baza v $\mathbb{R}^n$ je ortogonalna baza glede na standardni skalarni produkt.

Definicija ortonormirane baze

Vektorjem z normo 1 pravimo **normirani** vektorji. Ortogonalni množici, v kateri so vsi elementi normirani, pravimo **ortonormirana množica**. Ortonormirani množici, ki je baza, pravimo **ortonormirana baza**.

Opomba: Definicijo ortonormiranosti množice $\{v_1, \ldots, v_n\}$ se običajno na kratko pove z naslednjo formulo

$$\langle v_i, v_j \rangle = \begin{cases} 0 & \text{če } i \neq j \\ 1 & \text{če } i = j \end{cases}$$

Opomba: Vsak neničeln vektor lahko spremenimo v normiran vektor tako, da ga delimo z njegovo normo. Pravimo, da smo element normirali. Vsako ortogonalno množico lahko spremenimo v ortonormirano množico tako, da vse njene elemente normiramo.

Primer normiranja ortogonalne baze

Vektorji $(-\frac{1}{2}, 1, 1), (1, -\frac{1}{2}, 1), (1, 1, -\frac{1}{2})$ tvorijo ortogonalno bazo za $\mathbb{R}^3$, ki ni normirana. Če te vektorje normiramo (delimo s $\frac{3}{2}$), dobimo vektorje $(-\frac{1}{3}, \frac{2}{3}, \frac{2}{3}), (\frac{2}{3}, -\frac{1}{3}, \frac{2}{3}), (\frac{2}{3}, \frac{2}{3}, -\frac{1}{3})$, ki tvorijo ortonormirano bazo za $\mathbb{R}^3$.

Vektorski prostori s skalarnim produktom - 2. del

4. Obstoj ortogonalne baze

Najprej ponovimo osnovne definicije. Naj bo V vektorski prostor s skalarnim produktom. Množica $\{v_1, \ldots, v_n\} \subset V$ je

- **ortogonalna množica**, če velja $\langle v_i, v_j \rangle = 0$ za vsaka $i, j = 1, \ldots, n$, kjer $i \neq j$, in če velja $v_i \neq 0$ za vsak $i = 1, \ldots, n$.
- **ortonormirana množica**, če je ortogonalna množica in če velja $\|v_i\| = 1$ za vsak $i = 1, \ldots, n$.
- **ortogonalna baza**, če je ortogonalna množica in če je ogrodje.
- **ortonormirana baza**, če je ortonormirana množica in če je ogrodje.

Spomnimo se, da je vsaka ortogonalna množica linearno neodvisna. Obratno seveda ni res. Lahko pa vsako linearno neodvisno množico s preprostim postopkom predelamo v ortogonalno množico, ne da bi pri tem spremenili njeno linearno ogrinjačo. Temu postopku pravimo **Gram-Schmidtova ortogonalizacija**. Oglejmo si, kako poteka.

Trditev (Gram-Schmidtova ortogonalizacija)

Naj bo $\{u_1, \ldots, u_n\}$ baza. Definirajmo

$$
\begin{aligned}
v_1 &= u_1 \\
v_2 &= u_2 - \frac{\langle u_2, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 \\
v_3 &= u_3 - \frac{\langle u_3, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 - \frac{\langle u_3, v_2 \rangle}{\langle v_2, v_2 \rangle} v_2 \\
&\vdots \\
v_n &= u_n - \frac{\langle u_n, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 - \frac{\langle u_n, v_2 \rangle}{\langle v_2, v_2 \rangle} v_2 - \ldots - \frac{\langle u_n, v_{n-1} \rangle}{\langle v_{n-1}, v_{n-1} \rangle} v_{n-1}
\end{aligned}
$$

Potem je $\{v_1, \ldots, v_n\}$ ortogonalna baza.

Dokaz. S popolno indukcijo bomo dokazali, da je za vsak $k = 1, \ldots, n$ množica $\{v_1, \ldots, v_k\}$ ortogonalna.

Ker je $v_1 = u_1 \neq 0$, je $\{v_1\}$ ortogonalna.

Recimo, da trditev drži za nek $k < n$. S pomočjo formule

$$v_{k+1} = u_{k+1} - \sum_{i=1}^{k} \frac{\langle u_{k+1}, v_i \rangle}{\langle v_i, v_i \rangle} v_i \tag{1}$$

bomo dokazali, da trditev velja tudi za $k+1$. Če (1) skalarno pomnožimo z v_j za vsak $j = 1, \ldots, k$ in upoštevamo ortogonalnost $\{v_1, \ldots, v_k\}$, dobimo

$$\begin{aligned} \langle v_{k+1}, v_j \rangle &= \langle u_{k+1}, v_j \rangle - \sum_{i=1}^{k} \frac{\langle u_{k+1}, v_i \rangle}{\langle v_i, v_i \rangle} \langle v_i, v_j \rangle \\ &= \langle u_{k+1}, v_j \rangle - \frac{\langle u_{k+1}, v_j \rangle}{\langle v_j, v_j \rangle} \langle v_j, v_j \rangle = 0 \end{aligned}$$

Torej so elementi množice $\{v_1, \ldots, v_k, v_{k+1}\}$ paroma ortogonalni.

Da bi dokazali ortogonalnost množice $\{v_1, \ldots, v_k, v_{k+1}\}$, moramo še preveriti, da je $v_{k+1} \neq 0$. Če bi veljalo $v_{k+1} = 0$, potem bi dobili

$$u_{k+1} = \sum_{i=1}^{k} \frac{\langle u_{k+1}, v_i \rangle}{\langle v_i, v_i \rangle} v_i \in \mathrm{Lin}\{v_1, \ldots, v_k\}$$

Iz prvih k formul v Gram-Schmidtovi ortogonalizaciji sledi tudi

$$u_1, \ldots, u_k \in \mathrm{Lin}\{v_1, \ldots, v_k\}$$

Torej k-razsežen prostor $\mathrm{Lin}\{v_1, \ldots, v_k\}$ vsebuje linearno neodvisno množico $\{u_1, \ldots, u_k, u_{k+1}\}$, ki ima $k+1$ elementov. To je protislovje.

Izrek o obstoju ortogonalne baze

Vsak končnorazsežen vektorski prostor s skalarnim produktom ima ortogonalno bazo. Poleg tega lahko vsako ortogonalno množico dopolnimo do ortogonalne baze.

Dokaz. Prvi del dokažemo tako, da vzamemo poljubno bazo in jo s pomočjo Gram-Schmidtove ortogonalizacije predelamo v ortogonalno bazo. Drugi del dokažemo tako, da ortogonalno množico $\{u_1, \ldots, u_m\}$ najprej dopolnimo do običajne baze $\{u_1, \ldots, u_m, \ldots, u_n\}$ in potem uporabimo Gram-Schmidtovo ortogonalizacijo. Dobimo $v_1 = u_1, \ldots, v_m = u_m, \ldots$ torej dobljena ortogonalna baza res vsebuje množico $\{u_1, \ldots, u_m\}$. □

Odtod z normiranjem vektorjev dobimo:

Posledica

Vsak končnorazsežen vektorski prostor s skalarnim produktom ima ortonormirano bazo. Poleg tega lahko vsako ortonormirano množico dopolnimo do ortonormirane baze.

Primer

Naj bo V vektorski prostor vseh realnih polinomov stopnje ≤ 3 in naj bo skalarni produkt na V definiran z $\langle p, q \rangle = \int_{-1}^{1} p(x)q(x)\,dx$. Konstruirajmo ortonormirano bazo za V.

Vzemimo standardno bazo

$$u_1 = 1, \quad u_2 = x, \quad u_3 = x^2, \quad u_4 = x^3.$$

Z Gram-Schmidtovo ortogonalizacijo dobimo

$$\begin{aligned}
v_1 &= 1 \\
v_2 &= x - \frac{\langle x, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 = x - 0v_1 = x \\
v_3 &= x^2 - \frac{\langle x^2, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 - \frac{\langle x^2, v_2 \rangle}{\langle v_2, v_2 \rangle} v_2 = x^2 - \frac{1}{3} \\
v_4 &= x^3 - \frac{\langle x^3, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 - \frac{\langle x^3, v_2 \rangle}{\langle v_2, v_2 \rangle} v_2 - \frac{\langle x^3, v_3 \rangle}{\langle v_3, v_3 \rangle} v_3 = x^3 - \frac{3}{5}x
\end{aligned}$$

Vektorji v_1, v_2, v_3, v_4 so (po dokazu izreka) ortogonalna baza za V. Izračunajmo sedaj njihove norme in njihove normalizacije:

$$\|v_1\|^2 = \int_{-1}^{1} 1\,dx = 2, \qquad \frac{v_1}{\|v_1\|} = \frac{1}{\sqrt{2}},$$
$$\|v_2\|^2 = \int_{-1}^{1} x^2\,dx = \frac{2}{3}, \qquad \frac{v_2}{\|v_2\|} = \sqrt{\frac{3}{2}}\,x,$$
$$\|v_3\|^2 = \int_{-1}^{1} (x^2 - \tfrac{1}{3})^2\,dx = \frac{8}{45}, \qquad \frac{v_3}{\|v_3\|} = \sqrt{\frac{45}{8}}(x^2 - \tfrac{1}{3}),$$
$$\|v_4\|^2 = \int_{-1}^{1} (x^3 - \tfrac{3}{5}x)^2\,dx = \frac{8}{175}, \qquad \frac{v_4}{\|v_4\|} = \sqrt{\frac{175}{8}}(x^3 - \tfrac{3}{5}x).$$

Normalizirani vektorji tvorijo ortonormirano bazo za V.

Opomba: Iz tega primera vidimo, da je ortonormirana baza običajno vključuje kompliciranе korene. Za računske namene so zato pogosto boljše ortogonalne baze. Za teoretične namene pa so boljše ortonormirane baze.

5. Prednosti ortogonalnih baz

Glavna prednost je, da je vektor preprosteje razviti po ortogonalni bazi.

Izrek (Fourierov razvoj)

Naj bo V vektorski prostor s skalarnim produktom in $\{v_1, \ldots, v_n\}$ ortogonalna baza za V. Potem za vsak element $v \in V$ velja

$$v = \sum_{i=1}^{n} \frac{\langle v, v_i \rangle}{\langle v_i, v_i \rangle} v_i.$$

Če je baza ortonormirana, se formula poenostavi v $v = \sum_{i=1}^{n} \langle v, v_i \rangle v_i$.

Dokaz: Ker je $\{v_1, \ldots, v_n\}$ baza za V, obstajajo taki skalarji $\alpha_1, \ldots, \alpha_n$, da velja $v = \sum_{i=1}^{n} \alpha_i v_i$. Če to enakost skalarno pomnožimo z v_j za nek j, dobimo $\langle v, v_j \rangle = \langle \sum_{i=1}^{n} \alpha_i v_i, v_j \rangle = \sum_{i=1}^{n} \alpha_i \langle v_i, v_j \rangle = \alpha_j \langle v_j, v_j \rangle$. Pri zadnjem enačaju smo upoštevali, da je $\langle v_i, v_j \rangle = 0$ za vse i, ki so različni od j. Odtod izrazimo $\alpha_j = \frac{\langle v, v_j \rangle}{\langle v_j, v_j \rangle}$. Sledi $v = \sum_{i=1}^{n} \alpha_i v_i = \sum_{i=1}^{n} \frac{\langle v, v_i \rangle}{\langle v_i, v_i \rangle} v_i$. □

Druga prednost ortogonalnih baz je, da je iz razvoja elementa po taki bazi zelo preprosto izračunati njegovo normo. Iz $v = \sum_{i=1}^{n} \alpha_i v_i$ namreč sledi $\|v\|^2 = \langle v, v \rangle = \langle \sum_{i=1}^{n} \alpha_i v_i, \sum_{j=1}^{n} \alpha_j v_j \rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \bar{\alpha}_j \langle v_i, v_j \rangle = \sum_{i=1}^{n} \alpha_i \bar{\alpha}_i \langle v_i, v_i \rangle = \sum_{i=1}^{n} |\alpha_i|^2 \langle v_i, v_i \rangle$. Odtod za $\alpha_i = \frac{\langle v, v_i \rangle}{\langle v_i, v_i \rangle}$ dobimo:

Izrek (Parsevalova identiteta)

Naj bo V vektorski prostor s skalarnim produktom in $\{v_1, \ldots, v_n\}$ ortogonalna baza za V. Potem za vsak element $v \in V$ velja

$$\|v\|^2 = \sum_{i=1}^{n} \frac{|\langle v, v_i \rangle|^2}{\langle v_i, v_i \rangle}.$$

Če je baza ortonormirana, se formula poenostavi v $\|v\|^2 = \sum_{i=1}^{n} |\langle v, v_i \rangle|^2$.

Tretja prednost ortogonalnih baz je, da je zelo preprosto izračunati ortogonalno projekcijo vektorja na podprostor.

Definicija ortogonalne projekcije

Naj bo V vektorski prostor s skalarnim produktom, v vektor iz V in W vektorski podprostor v V. Vektorju iz W, ki je najbližje vektorju v, pravimo **ortogonalna projekcija** vektorja v na podprostor W.

Opomba: Razdalja med dvema vektorjema v V je definirana kot norma njune razlike. Ortogonalna projekcija vektorja $v \in V$ na podprostor $W \subseteq V$ je torej tak vektor $v' \in W$, da velja

$$\|v - w\| > \|v - v'\|$$

za vsak vektor $w \in W$, ki je različen od vektorja v'.

Izrek o ortogonalni projekciji

Naj bo V vektorski prostor s skalarnim produktom in W končnorazsežen vektorski podprostor v V z ortogonalno bazo $\{w_1, \ldots, w_k\}$. Potem je ortogonalna projekcija vektorja $v \in V$ na podprostor W dana s formulo

$$v' = \sum_{i=1}^{k} \frac{\langle v, w_i \rangle}{\langle w_i, w_i \rangle} w_i.$$

Če je baza ortonormirana, se formula poenostavi v $v' = \sum_{i=1}^{k} \langle v, w_i \rangle w_i$.

Opomba: V dokazu bomo uporabili Pitagorov izrek, ki pravi, da za pravokotna vektorja a in b velja

$$\|a + b\|^2 = \|a\|^2 + \|b\|^2.$$

To dokažemo z direktnim računom
$\langle a + b, a + b \rangle = \langle a, a \rangle + \langle a, b \rangle + \langle b, a \rangle + \langle b, b \rangle = \langle a, a \rangle + \langle b, b \rangle$.

Dokaz izreka: Ker je v' linearna kombinacija vektorjev w_i iz W, je v' element W. Radi bi pokazali, da za vsak $w \in W$, ki je različen od v', velja $\|v - w\| > \|v - v'\|$. Vzemimo poljuben w iz W in ga razvijmo po bazi za W. Dobimo $w = \sum_{i=1}^{k} \beta_i w_i$, kjer je $\beta_i = \frac{\langle w, w_i \rangle}{\langle w_i, w_i \rangle}$ za vse i.

Pišimo $v' = \sum_{i=1}^{k} \alpha_i w_i$, kjer je $\alpha_i = \frac{\langle v, w_i \rangle}{\langle w_i, w_i \rangle}$ za vse i.

1.korak Dokažimo, da je vektor $v - v'$ pravokoten na vse w_i,

Velja $\langle v - v', w_j \rangle = \langle v, w_j \rangle - \langle v', w_j \rangle = \langle v, w_j \rangle - \sum_{i=1}^{k} \alpha_i \langle w_i, w_j \rangle = \langle v, w_j \rangle - \alpha_j \langle w_j, w_j \rangle = 0$.

2.korak Dokažimo, da je vektor $v - v'$ pravokoten na vektor $v' - w$,

Velja $\langle v - v', v' - w \rangle = \langle v - v', \sum_{i=1}^{k} (\alpha_i - \beta_i) w_i \rangle =$
$= \sum_{i=1}^{k} (\overline{\alpha_i - \beta_i}) \langle v - v', w_i \rangle = 0$ po 1. koraku.

3.korak Dokažimo, da je $\|v - w\|^2 = \|v - v'\|^2 + \|v' - w\|^2$.

To je ravno Pitagorov izrek za vektorja $a = v - v'$ in $b = v' - w$, ki sta pravokotna po 2. koraku.

Iz 3. koraka sledi, da je $\|v - w\|^2 > \|v - v'\|^2$, če je $v' \neq w$. □

6. Ortogonalni komplement

Definicija ortogonalnega komplementa

Naj bo V vektorski prostor s skalarnim produktom in naj bo S pomnožica v V. Pravimo, da je vektor $v \in V$ **ortogonalen** na množico S, če velja $\langle v, s \rangle = 0$ za vsak $s \in S$. Množico vseh vektorjev iz V, ki so ortogonalni na množico S, označimo z $S^\perp$ in ji pravimo **ortogonalni komplement** S.

Trditev

Za vsako podmnožico S v V je $S^\perp$ vektorski podprostor v V.

Dokaz: Če v_1 in v_2 pripadata $S^\perp$, je (po definiciji $S^\perp$) $\langle v_1, s \rangle = 0$ in $\langle v_2, s \rangle = 0$ za vsak $s \in S$. Potem je za vsaka skalarja α_1 in α_2 tudi $\langle \alpha_1 v_1 + \alpha_2 v_2, s \rangle = \alpha_1 \langle v_1, s \rangle + \alpha_2 \langle v_2, s \rangle = 0$ za vsak $s \in S$. Odtod sledi (spet po definiciji $S^\perp$), da je $\alpha_1 v_1 + \alpha_2 v_2 \in S^\perp$ za vsaka α_1 in α_2. □

Trditev

Za vsako podmnožico S v V je $S^{\perp} = (\mathrm{Lin}S)^{\perp}$.

Dokaz: Če je vektor v ortogonalen na elemente $s_1, \ldots, s_k$, potem je ortogonalen tudi na vsako njihovo linearno kombinacijo, saj velja

$$\langle v, \alpha_1 s_1 + \ldots + \alpha_k s_k \rangle = \bar{\alpha}_1 \langle v, s_1 \rangle + \ldots + \bar{\alpha}_k \langle v, s_k \rangle = 0 + \ldots + 0 = 0$$

Naslednji primer bo pomemben v nadaljevanju:

Primer ortogonalnega komplementa

Naj bo V končnorazsežen vektorski prostor s skalarnim produktom, $\{u_1, \ldots, u_k\}$ ortogonalna množica v V in $\{v_1, \ldots, v_l\}$ njena dopolnitev do ortogonalne baze za V. Potem je $(\mathrm{Lin}\{u_1, \ldots, u_k\})^{\perp} = \mathrm{Lin}\{v_1, \ldots, v_l\}$.

Dokaz: Vzemimo poljuben element $v \in V$ in ga razvijmo po bazi za V. Dobimo $v = \sum_{i=1}^{k} \alpha_i u_i + \sum_{j=1}^{l} \beta_j v_j$. Ker je $\langle v, u_i \rangle = \alpha_i \langle u_i, u_i \rangle$, je v ortogonalen na vse u_i natanko tedaj, ko so vsi α_i enaki nič. To pa velja natanko tedaj, ko je $v \in \mathrm{Lin}\{v_1, \ldots, v_l\}$.

Izrek o ortogonalnem razcepu

Naj bo V končnorazsežen prostor s skalarnim produktom in naj bo U podprostor v V. Potem velja naslednje:

1. $\dim U^\perp = \dim V - \dim U$.
2. $(U^\perp)^\perp = U$.
3. $V = U \oplus U^\perp$.

Dokaz: Naj bo $\{u_1, \ldots, u_k\}$ ortogonalna baza za U in naj bo $\{v_1, \ldots, v_l\}$ njena dopolnitev do ortogonalne baze za V.

Po primeru je $U^\perp = \mathrm{Lin}\{v_1, \ldots, v_l\}$. Torej je $\dim U + \dim U^\perp = k + l$, kar je enako $\dim V$. To nam da prvo trditev.

Drugi del dokažemo tako, da v primeru zamenjamo bazo iz u_i z bazo iz v_j. Ker je $\{v_1, \ldots, v_l\}$ ortogonalna baza za $U^\perp$, in ker je $\{u_1, \ldots, u_k\}$ njena dopolnitev do ortogonalne baze za V, je $(U^\perp)^\perp = \mathrm{Lin}\{u_1, \ldots, u_k\}$.

Tretji del sledi iz $V = \mathrm{Lin}\{u_1, \ldots, u_k\} \oplus \mathrm{Lin}\{v_1, \ldots, v_l\}$ in primera. □

Opomba: Formuli $V = U \oplus U^\perp$ pravimo **ortogonalni razcep** prostora V glede na podprostor U.

To formulo lahko dokažemo tudi s pomočjo izreka o ortogonalni projekciji. Za vsak element $v \in V$ lahko izračunamo njegovo projekcijo v' na podprostor U. Iz dokaza izreka o ortogonalni projekciji vemo, da je vektor $v - v'$ pravokoten na vse vektorje podprostora U. Velja torej $v - v' \in U^\perp$. Iz $v = v' + (v - v')$ torej sledi, da je vsak element iz V vsota elementa iz U in elementa iz $U^\perp$, kar dokaže formulo $V = U + U^\perp$. Treba je preveriti še $U \cap U^\perp = \{0\}$. Vsak element $u \in U$, ki je pravokoten na vse elemente iz u, je pravokoten tudi sam nase, torej je enak nič. □

5. Linearni funkcionali

Linearni funkcionali so zelo posebni primeri linearnih preslikav.

Definicija

Naj bo V vektorski prostor nad obsegom F. Linearni preslikavi iz V v F pravimo **linearen funkcional** na V.

Opomba: V definiciji smo upoštevali, da je tudi F vektorski prostor nad F. To je poseben primer vektorskega prostora F^n za $n = 1$.

Primera linearnih funkcionalov

Naj bo V vektorski prostor vseh realnih polinomov stopnje $\leq n$. Potem sta

$$\phi(p) = \int_0^1 p(x)\,dx \quad \text{in} \quad \psi(p) = p(1)$$

dva primera linearnih funkcionalov na V.

Podobno kot vsaki drugi linearni preslikavi, lahko tudi linearnemu funkcionalu priredimo matriko. Naslednje trditev je očitna.

Trditev (Matrika linearnega funkcionala)

Naj bo L linearen funkcional na vektorskem prostoru V, naj bo $\mathcal{B} = \{v_1, \ldots, v_n\}$ baza za V in naj bo $\mathcal{S} = \{1\}$ standardna baza za F. Matrika L glede na ti dve bazi ima eno vrstico in n stolpcev. Velja

$$[L]_{\mathcal{S}\leftarrow\mathcal{B}} = \begin{bmatrix} L(v_1) & \ldots & L(v_n) \end{bmatrix}.$$

Izračunajmo matriki linearnih funkcionalov iz primera:

Nadaljevanje primera

Naj bo $\mathcal{B} = \{1, x, \ldots, x^n\}$ in naj bosta ϕ ter ψ kot zgoraj. Potem je

$$[\phi]_{\mathcal{S}\leftarrow\mathcal{B}} = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \ldots & \frac{1}{n} \end{bmatrix} \qquad [\psi]_{\mathcal{S}\leftarrow\mathcal{B}} = \begin{bmatrix} 1 & 1 & 1 & \ldots & 1 \end{bmatrix}.$$

O linearnih funkcionalih na vektorskih prostorih s skalarnim produktom lahko povemo precej več. Začnimo s preprostim primerom.

Primer

Naj bo V vektorski prostor s skalarnim produktom in $w \in V$. Potem je

$$\phi_w(v) := \langle v, w \rangle$$

linearen funkcional na V.

Dokaz: To sledi iz linearnosti skalarnega produkta v prvem faktorju.

Izkaže se, da ta primer ni tako preprost, saj nam da vse linearne funkcionale na končnorazsežnih vektorskih prostorih s skalarnim produktom.

Rieszov izrek o reprezentaciji linearnih funkcionalov

Naj bo V končnorazsežen vektorski prostor s skalarnim produktom in naj bo ϕ linearen funkcional na V. Potem obstaja tak vektor $w \in V$, da je

$$\phi(v) = \langle v, w \rangle$$

za vsak $v \in V$. Vektor w je enolično določen.

Izrek lahko povemo tudi takole. Vsak linearen funkcional na V je oblike ϕ_w (glej primer) za natanko določen $w \in V$.

Dokaz: Naj bo $v_1, \ldots, v_n$ ortonormirana baza za V. Vzemimo poljuben $v \in V$ in ga razvijmo po tej bazi. Velja $v = \sum_{i=1}^{n} \langle v, v_i \rangle v_i$. Odtod sledi $\phi(v) = \sum_{i=1}^{n} \langle v, v_i \rangle \phi(v_i)$ zaradi linearnosti ϕ. Če upoštevamo še konjugirano linearnost skalarnega produkta v drugem faktorju, dobimo $\sum_{i=1}^{n} \langle v, v_i \rangle \phi(v_i) = \sum_{i=1}^{n} \langle v, \overline{\phi(v_i)} v_i \rangle = \langle v, \sum_{i=1}^{n} \overline{\phi(v_i)} v_i \rangle$. Označimo $w = \sum_{i=1}^{n} \overline{\phi(v_i)} v_i$ in opazimo, da ta vektor ni odvisen od v, ampak samo od ϕ in od baze. Dokazali smo, da velja $\phi(v) = \langle v, w \rangle$ za vsak $v \in V$.

Pokažimo še, da je vektor w enolično določen s funkcionalom ϕ.
Če velja $\phi(v) = \langle v, w_1 \rangle$ za vsak $v \in V$ in $\phi(v) = \langle v, w_2 \rangle$ za vsak $v \in V$, potem $0 = \phi(v) - \phi(v) = \langle v, w_1 \rangle - \langle v, w_2 \rangle = \langle v, w_1 - w_2 \rangle$ za vsak $v \in V$. Vstavimo $v = w_1 - w_2$ in dobimo $\langle w_1 - w_2, w_1 - w_2 \rangle = 0$. Vemo, da odtod sledi $w_1 - w_2 = 0$, se pravi $w_1 = w_2$. □

Adjungirana preslikava - 1. del

1. Rieszov izrek o reprezentaciji linearnih funkcionalov

Ponovimo od zadnjič:

Rieszov izrek o reprezentaciji linearnih funkcionalov

Naj bo V končnorazsežen vektorski prostor s skalarnim produktom in naj bo ϕ linearen funkcional na V. Potem obstaja tak vektor $w \in V$, da je

$$\phi(v) = \langle v, w \rangle$$

za vsak $v \in V$. Vektor w je enolično določen.

Naj bo $v_1, \ldots, v_n$ ortonormirana baza za V. Potem za vsak $v \in V$ velja

$$\phi(v) = \phi(\sum_{i=1}^{n} \langle v, v_i \rangle v_i) = \sum_{i=1}^{n} \langle v, v_i \rangle \phi(v_i) = \langle v, \sum_{i=1}^{n} \overline{\phi(v_i)} v_i \rangle = \langle v, w \rangle$$

Primer

Naj bo V vektorski prostor vseh realnih polinomov stopnje ≤ 3 in naj bo $\langle p, q\rangle = \int_{-1}^{1} p(x)q(x)\,dx$ skalarni produkt na V. Za linearni funkcional $\phi(p) = p(1)$ na V določi tak $r \in V$, da bo $\phi(p) = \langle p, r\rangle$ za vsak $p \in V$.

Prva rešitev: Najprej poiščemo ortonormirano bazo za V. Recimo

$$v_1 = \frac{1}{\sqrt{2}},\ v_2 = \sqrt{\frac{3}{2}}\,x,\ v_3 = \sqrt{\frac{45}{8}}(x^2 - \frac{1}{3}),\ v_4 = \sqrt{\frac{175}{8}}(x^3 - \frac{3}{5}x)$$

Vstavimo to bazo v formulo $w = \sum_{i=1}^{n} \overline{\phi(v_i)}v_i$ in dobimo

$$w = v_1(1)v_1 + v_2(1)v_2 + v_3(1)v_3 + v_4(1)v_4$$

$$= \left(\frac{1}{\sqrt{2}}\right)^2 + \left(\sqrt{\frac{3}{2}}\right)^2 x + \frac{2}{3}\left(\sqrt{\frac{45}{8}}\right)^2 (x^2 - \frac{1}{3}) + \frac{2}{5}\left(\sqrt{\frac{175}{8}}\right)^2 (x^3 - \frac{3}{5}x)$$

$$= -\frac{3}{4} - \frac{15}{4}x + \frac{15}{4}x^2 + \frac{35}{4}x^3.$$

Druga rešitev: Recimo, da je $r(x) = e + fx + gx^2 + hx^3$. Potem je

$$1 = \phi(1) = \langle 1, r \rangle = \int_{-1}^{1} (e + fx + gx^2 + hx^3)\, dx = 2e + \frac{2}{3}g$$

$$1 = \phi(x) = \langle x, r \rangle = \int_{-1}^{1} x(e + fx + gx^2 + hx^3)\, dx = \frac{2}{3}f + \frac{2}{5}h$$

$$1 = \phi(x^2) = \langle x^2, r \rangle = \int_{-1}^{1} x^2(e + fx + gx^2 + hx^3)\, dx = \frac{2}{3}e + \frac{2}{5}g$$

$$1 = \phi(x^3) = \langle x^3, r \rangle = \int_{-1}^{1} x^3(e + fx + gx^2 + hx^3)\, dx = \frac{2}{5}f + \frac{2}{7}h$$

Odtod dobimo $e = -\frac{3}{4}, g = \frac{15}{4}$ in $f = -\frac{15}{4}, h = \frac{35}{4}$.

Opomba: Prednost druge metode je, da nam ni treba računati ortonormirane baze, slabost pa je, da moramo rešiti sistem linearnih enačb. To metodo lahko spremenimo v alternativen dokaz Rieszovega izreka.

2. Definicija adjungirane linearne preslikave

V tem razdelku bomo delali z več vektorskimi prostori s skalarnim produktom, zato bomo pri vsakem skalarnem produktu označili iz katerega vektorskega prostora prihaja. Izbrani skalarni produkt na U bomo označili z $\langle u_1, u_2 \rangle_U$, izbrani skalarni produkt na V bomo označili z $\langle v_1, v_2 \rangle_V$, itd.

Definicija adjungirane linearne preslikave

Naj bo $L\colon U \to V$ linearna preslikava med dvema vektorskima prostoroma s skalarnim produktom. Pravimo, da je linearna preslikava $L^*\colon V \to U$ **adjungirana linearna preslikava** linearne preslikave L, če velja

$$\langle Lu, v \rangle_V = \langle u, L^* v \rangle_U$$

za vsak $u \in U$ in vsak $v \in V$.

Iz definicije še ne sledi, da L^* obstaja. Dokažimo obstoj in enoličnost v končnorazsežnem primeru.

Izrek o obstoju in enoličnosti adjungirane linearne preslikave

Vsaka linearna preslikava med končnorazsežnima vektorskima prostoroma s skalarnim produktom ima natanko eno adjungirano linearno preslikavo.

Dokaz: Recimo, da sta U in V končnorazsežna vektorska prostora s skalarnim produktom in da je $L\colon U \to V$ linearna preslikava.

Dokažimo najprej enoličnost: Naj bosta L^* in L' dve adjungirani linearni preslikavi linearne presklikave L. Potem za vsak $u \in U$ ter $v \in V$ velja

$$\langle Lu, v\rangle = \langle u, L^*v\rangle \quad \text{in} \quad \langle Lu, v\rangle = \langle u, L'v\rangle.$$

Če enačbi odštejemo, dobimo, da za vsak $u \in U$ in vsak $v \in V$ velja

$$\langle u, L^*v - L'v\rangle = \langle u, L^*v\rangle - \langle u, L'v\rangle = \langle Lu, v\rangle - \langle Lu, v\rangle = 0.$$

Če vstavimo $u = L^*v - L'v$, dobimo $\langle L^*v - L'v, L^*v - L'v\rangle = 0$, se pravi

$$L^*v - L'v = 0$$

za vsak $v \in V$. Torej je $L^* = L'$.

Dokažimo še obstoj adjungirane linearne preslikave.

Vzemimo poljuben vektor $v \in V$ in si oglejmo preslikavo

$$\phi(u) = \langle Lu, v \rangle_V,$$

ki slika iz U v skalarje. Preverimo najprej, da je ϕ linearen funkcional na U. To sledi iz $\phi(\alpha_1 u_1 + \alpha_2 u_2) = \langle L(\alpha_1 u_1 + \alpha_2 u_2), v \rangle_V =$
$= \langle \alpha_1 Lu_1 + \alpha_2 Lu_2, v \rangle_V = \alpha_1 \langle Lu_1, v \rangle_V + \alpha_2 \langle Lu_2, v \rangle_V = \alpha_1 \phi(u_1) + \alpha_2 \phi(u_2)$.

Po Rieszovem izreku o reprezentaciji linearnih funkcionalov obstaja natanko en vektor $w \in U$, ki zadošča $\phi(u) = \langle u, w \rangle_U$ za vsak $u \in U$. Definirajmo

$$L^* v := w.$$

Na dolgo to povemo takole: Za vsak v iz V je $L^* v$ tak vektor iz U, da velja $\langle Lu, v \rangle_V = \langle u, L^* v \rangle_U$ za vsak $u \in U$.

S tem smo definirali preslikavo L^* iz V v U, ki zadošča

$$\langle Lu, v \rangle_V = \langle u, L^* v \rangle_U$$

za vsak $u \in U$ in vsak $v \in V$.

Preverimo moramo še, da je L^* res linearna preslikava.

Vzemimo poljubna vektorja v_1 in v_2 iz V in poljubna skalarja α_1 in α_2. Po definiciji so L^*v_1, L^*v_2 in $L^*(\alpha_1 v_1 + \alpha_2 v_2)$ taki vektorji iz U, da velja

$$\langle u, L^*v_1\rangle_U = \langle Lu, v_1\rangle_V, \tag{1}$$
$$\langle u, L^*v_2\rangle_U = \langle Lu, v_2\rangle_V, \tag{2}$$
$$\langle u, L^*(\alpha_1 v_1 + \alpha_2 v_2)\rangle_U = \langle Lu, \alpha_1 v_1 + \alpha_2 v_2\rangle_V \tag{3}$$

za vsak $u \in U$. Iz (1),(2) in (3) sledi, da za vsak $u \in U$ velja

$$\begin{aligned}
&\langle u, L^*(\alpha_1 v_1 + \alpha_2 v_2) - \alpha_1 L^*v_1 - \alpha_2 L^*v_2\rangle_U \\
&= \langle u, L^*(\alpha_1 v_1 + \alpha_2 v_2)\rangle_U - \bar{\alpha}_1\langle u, L^*v_1\rangle_U - \bar{\alpha}_2\langle u, L^*v_2\rangle_U \\
&= \langle Lu, \alpha_1 v_1 + \alpha_2 v_2\rangle_U - \bar{\alpha}_1\langle Lu, v_1\rangle_U - \bar{\alpha}_2\langle Lu, v_2\rangle_U \\
&= \langle Lu, \alpha_1 v_1 + \alpha_2 v_2\rangle_U - \langle Lu, \alpha_1 v_1\rangle_U - \langle Lu, \alpha v_2\rangle_U = 0
\end{aligned}$$

Če v to enakost vstavimo $u = L^*(\alpha_1 v_1 + \alpha_2 v_2) - \alpha_1 L^*v_1 - \alpha_2 L^*v_2$, dobimo

$$L^*(\alpha_1 v_1 + \alpha_2 v_2) - \alpha_1 L^*v_1 - \alpha_2 L^*v_2 = 0$$

torej je preslikava L^* res linearna.

Primer računanja adjungirane preslikave

Naj bo $U = V$ vektorski prostor $\mathbb{C}^2$ z nestandardnim skalarnim produktom $\langle(\alpha_1, \alpha_2), (\beta_1, \beta_2)\rangle = 2\alpha_1\bar{\beta}_1 + 3\alpha_2\bar{\beta}_2$. Izračunaj adjungirano linearno preslikavo linearne preslikave $L(x, y) = (ax + by, cx + dy)$.

Ker je L^* linearna preslikava iz $\mathbb{C}^2$ v $\mathbb{C}^2$, je oblike

$$L^*(x, y) = (ex + fy, gx + hy).$$

Poleg tega mora veljati

$$\langle L(x_1, y_1), (x_2, y_2)\rangle = \langle(x_1, y_1), L^*(x_2, y_2)\rangle$$

za vse $x_1, x_2, y_1, y_2 \in \mathbb{C}$. Ko upoštevamo definiciji L in L^* dobimo

$$\langle(ax_1 + by_1, cx_1 + dy_1), (x_2, y_2)\rangle = \langle(x_1, y_1), (ex_2 + fy_2, gx_2 + hy_2)\rangle$$

za vse $x_1, x_2, y_1, y_2 \in \mathbb{C}$.

Ko upoštevamo še definicijo skalarnega produkta dobimo

$$2(ax_1 + by_1)\bar{x}_2 + 3(cx_1 + dy_1)\bar{y}_2 = 2x_1\overline{(ex_2 + fy_2)} + 3y_1\overline{(gx_2 + hy_2)}$$

za vse $x_1, x_2, y_1, y_2 \in \mathbb{C}$. Primerjajmo istoležne koeficiente:

$$\begin{aligned}
\text{koeficienti pri } x_1\bar{x}_2 : &\quad 2a = 2\bar{e} \\
\text{koeficienti pri } y_1\bar{x}_2 : &\quad 2b = 3\bar{g} \\
\text{koeficienti pri } x_1\bar{y}_2 : &\quad 3c = 2\bar{f} \\
\text{koeficienti pri } y_1\bar{y}_2 : &\quad 3d = 3\bar{h}
\end{aligned}$$

Odtod izrazimo e, f, g, h z a, b, c, d in vstavimo v definicijo L^*. Dobimo

$$L^*(x, y) = (\bar{a}x + \frac{3}{2}\bar{c}y, \frac{2}{3}\bar{b}x + \bar{d}y). \quad \square$$

3. Matrika adjungirane linearne preslikave

Naj bosta U in V končnorazsežna vektorska prostora s skalarnim produktom. Zanima nas, kakšna je zveza med matriko linearne preslikave $L\colon U \to V$ in matriko njene adjunjgirane linearne preslikave $L^*\colon V \to U$. Odgovor je odvisen od izbire baz za U in V. V splošnem ni nobene zveze, če pa za U in V izberemo ortonormirani bazi, potem velja naslednje:

Trditev

Naj bo $\mathcal{B}$ ortonormirana baza za U in $\mathcal{C}$ ortonormirana baza za V. Naj bo $L\colon U \to V$ linearna preslikava in $L^*\colon V \to U$ njena adjungirana linearna preslikava. Potem matriko $[L^*]_{\mathcal{B}\leftarrow\mathcal{C}}$ dobimo tako, da v matriki $[L]_{\mathcal{C}\leftarrow\mathcal{B}}$ vse elemente konjugiramo in dobljeno matriko še transponiramo.

Dokaz: Naj bo $\mathcal{B} = \{u_1, \ldots, u_m\}$ in $\mathcal{C} = \{v_1, \ldots, v_n\}$. Izračunajmo matriko $[L]_{\mathcal{C}\leftarrow\mathcal{B}}$. S pomočjo Fourierovega razvoja dobimo:

$$\begin{aligned} Lu_1 &= \langle Lu_1, v_1\rangle_V v_1 + \ldots + \langle Lu_1, v_n\rangle_V v_n \\ &\vdots \\ Lu_m &= \langle Lu_m, v_1\rangle_V v_1 + \ldots + \langle Lu_m, v_n\rangle_V v_n \end{aligned}$$

Torej je

$$[L]_{\mathcal{C}\leftarrow\mathcal{B}} = \begin{bmatrix} \langle Lu_1, v_1\rangle_V & \ldots & \langle Lu_m, v_1\rangle_V \\ \vdots & \ddots & \vdots \\ \langle Lu_1, v_n\rangle_V & \ldots & \langle Lu_m, v_n\rangle_V \end{bmatrix}$$

Podobno s Fourierovim razvojem dobimo

$$\begin{aligned} L^* v_1 &= \langle L^* v_1, u_1 \rangle_U u_1 + \ldots + \langle L^* v_1, u_m \rangle_U u_m \\ &\vdots \\ L^* v_n &= \langle L^* v_n, u_1 \rangle_U u_1 + \ldots + \langle L^* v_n, u_m \rangle_U u_m \end{aligned}$$

Torej je

$$[L^*]_{\mathcal{B} \leftarrow \mathcal{C}} = \begin{bmatrix} \langle L^* v_1, u_1 \rangle_U & \ldots & \langle L^* v_n, u_1 \rangle_U \\ \vdots & \ddots & \vdots \\ \langle L^* v_1, u_m \rangle_U & \ldots & \langle L^* v_n, u_m \rangle_U \end{bmatrix}$$

Opazimo, da je (i, j)-ti element matrike $[L^*]_{\mathcal{B} \leftarrow \mathcal{C}}$ enak

$$\langle L^* v_j, u_i \rangle_U = \overline{\langle u_i, L^* v_j \rangle_U} = \overline{\langle L u_i, v_j \rangle_V}$$

kar je enako konjugiranemu (j, i)-temu elementu matrike $[L]_{\mathcal{C} \leftarrow \mathcal{B}}$.

Če torej vse elemente matrike $[L]_{\mathcal{C} \leftarrow \mathcal{B}}$ konjugiramo in dobljeno matriko transponiramo, dobimo ravno matriko $[L^*]_{\mathcal{B} \leftarrow \mathcal{C}}$.

Zadnji rezultat motivira naslednjo definicijo.

Definicija adjungirane matrike

Naj bo A kompleksna $m \times n$ matrika in naj bo $\overline{A}$ matrika, ki jo dobimo tako, da v matriki A konjugiramo vse elemente. Potem matriki

$$A^* := (\overline{A})^T$$

pravimo **adjungirana matrika** matrike A.

Opomba: Torej je $[L^*]_{\mathcal{B}\leftarrow\mathcal{C}} = \left([L]_{\mathcal{C}\leftarrow\mathcal{B}}\right)^*$, če sta $\mathcal{B}$ in $\mathcal{C}$ ortonormirani bazi.

Primer

Naj bo $U = V$ vektorski prostor $\mathbb{C}^2$ z nestandardnim skalarnim produktom $\langle(\alpha_1, \alpha_2), (\beta_1, \beta_2)\rangle = 2\alpha_1\bar{\beta}_1 + 3\alpha_2\bar{\beta}_2$. Izračunajmo še enkrat adjungirano linearno preslikavo linearne preslikave $L(x, y) = (ax + by, cx + dy)$.

Najprej opazimo, da je

$$u_1 = \frac{1}{\sqrt{2}}(1,0), \quad u_2 = \frac{1}{\sqrt{3}}(0,1)$$

ortonormirana baza glede na dani skalarni produkt. Označimo jo z $\mathcal{B}$. Določimo sedaj matriko preslikave L glede na to bazo. Velja

$$Lu_1 = \frac{1}{\sqrt{2}}(a,c) = au_1 + c\sqrt{\frac{3}{2}}u_2$$

$$Lu_2 = \frac{1}{\sqrt{3}}(b,d) = b\sqrt{\frac{2}{3}}u_1 + du_2$$

Torej je

$$[L]_{\mathcal{B}\leftarrow\mathcal{B}} = \begin{bmatrix} a & b\sqrt{\frac{2}{3}} \\ c\sqrt{\frac{3}{2}} & d \end{bmatrix}$$

Odtod sledi, da je

$$[L^*]_{\mathcal{B}\leftarrow\mathcal{B}} = \left([L]_{\mathcal{B}\leftarrow\mathcal{B}}\right)^* = \begin{bmatrix} a & b\sqrt{\frac{2}{3}} \\ c\sqrt{\frac{3}{2}} & d \end{bmatrix}^* = \begin{bmatrix} \bar{a} & \bar{c}\sqrt{\frac{3}{2}} \\ \bar{b}\sqrt{\frac{2}{3}} & \bar{d} \end{bmatrix}$$

Torej je

$$\begin{aligned} L^*(x,y) &= L^*(\sqrt{2}xu_1 + \sqrt{3}yu_2) = \sqrt{2}xL^*(u_1) + \sqrt{3}yL^*(u_2) = \\ &= \sqrt{2}x(\bar{a}u_1 + \bar{b}\sqrt{\frac{2}{3}}u_2) + \sqrt{3}y(\bar{c}\sqrt{\frac{3}{2}}u_1 + \bar{d}u_2) = \\ &= (\bar{a}x + \frac{3}{2}\bar{c}y, \frac{2}{3}\bar{b}x + \bar{d}y) \end{aligned}$$

4. Lastnosti adjungiranja

Adjungiranje matrik je preslikava iz matrik v matrike, ki vsaki matriki A priredi adjungirano matriko $A^* = (\overline{A})^T$. Z direktnim računom lahko preverimo naslednje lastnosti te preslikave.

1. $(\alpha A + \beta B)^* = \bar{\alpha} A^* + \bar{\beta} B^*$,
2. $(A^*)^* = A$,
3. $(AB)^* = B^* A^*$,
4. $0^* = 0$, $I^* = I$.

Iz lastnosti (1) sledi, da je na realnih matrikah fiksne velikosti adjungiranje linearna preslikava. Na kompleksnih matrikah fiksne velikosti adjungiranje ni linearna ampak konjugirano linearna preslikava.

Dokaz: Vemo, da za konjugiranje kompleksnih števil velja

$$\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2, \quad \overline{z_1 z_2} = \bar{z}_1 \bar{z}_2, \quad \bar{\bar{z}} = z.$$

Odtod sledi, da za kompleksne matrike velja

$$\overline{\alpha A + \beta B} = \bar{\alpha}\bar{A} + \bar{\beta}\bar{B}, \quad \overline{AB} = \bar{A}\bar{B}, \quad \bar{\bar{A}} = A.$$

Upoštevajmo še lastnosti transponiranja

$$(\alpha A + \beta B)^T = \alpha A^T + \beta B^T, \quad (AB)^T = B^T A^T, \quad (A^T)^T = A.$$

Odtod sledi

$$\left(\overline{\alpha A + \beta B}\right)^T = (\bar{\alpha}\bar{A} + \bar{\beta}\bar{B})^T = \bar{\alpha}\bar{A}^T + \bar{\beta}\bar{B}^T,$$

$$\left(\overline{AB}\right)^T = (\bar{A}\bar{B})^T = \bar{B}^T \bar{A}^T.$$

Če upoštevamo, da je $\bar{A}^T = \overline{A^T}$, potem dobimo še $\overline{\bar{A}^T}^T = A$. □

Podoben razmislek lahko naredimo tudi za adjungiranje linearnih preslikav.

Naj bosta U in V vektorska prostora s skalarnim produktom. Označimo z $\mathcal{L}(U, V)$ množico vseh linearnih preslikav iz U v V. To množico lahko spremenimo v vektorski prostor, če za linearni preslikavi L_1, L_2 in skalarja α_1, α_2 definiramo linearno preslikavo $\alpha_1 L_1 + \alpha_2 L_2$ s predpisom $(\alpha_1 L_1 + \alpha_2 L_2)(u) = \alpha_1 L_1(u) + \alpha_2 L_2(u)$.

Adjungiranje je preslikava $\mathcal{L}(U, V)$ v $\mathcal{L}(V, U)$. Preverimo, da zadošča

$$(\alpha_1 L_1 + \alpha_2 L_2)^* = \bar{\alpha}_1 L_1^* + \bar{\alpha}_2 L_2^*. \tag{1}$$

Za vsaka $u \in U$ in $v \in V$ velja $\langle u, (\alpha_1 L_1 + \alpha_2 L_2)^*(v)\rangle =$
$= \langle (\alpha_1 L_1 + \alpha_2 L_2)(u), v\rangle = \langle \alpha_1 L_1 u + \alpha_2 L_2 u, v\rangle =$
$= \alpha_1 \langle L_1 u, v\rangle + \alpha_2 \langle L_2 u, v\rangle = \alpha_1 \langle u, L_1^* v\rangle + \alpha_2 \langle u, L_2^* v\rangle =$
$= \langle u, \bar{\alpha}_1 L_1^* v + \bar{\alpha}_2 L_2^* v\rangle = \langle u, (\bar{\alpha}_1 L_1^* + \bar{\alpha}_2 L_2^*)(v)\rangle.$
Če upoštevamo enoličnost adjungirane preslikave, odtod sledi (1).

Lastnost (1) bi lahko dokazali tudi s prehodom na matrike.

Lastnost

$$(L^*)^* = L \tag{2}$$

dokažemo tako, da preverimo, da sta L in $(L^*)^*$ adjungirani linearni preslikavi linearne preslikave L^*. Za vsaka $u \in U$ in $v \in V$ velja $\langle v, (L^*)^* u\rangle = \langle L^* v, u\rangle = \overline{\langle u, L^* v\rangle} = \overline{\langle Lu, v\rangle} = \langle v, Lu\rangle$. Lastnost

$$(L_2 L_1)^* = L_1^* L_2^* \tag{3}$$

dokažemo tako, da preverimo, da sta $(L_2 L_1)^*$ in $L_1^* L_2^*$ adjungirani linearni preslikavi linearne preslikave $L_2 L_1$. Recimo, da L_1 slika iz U v V, L_2 pa iz V v W. Za vsak $u \in U$ in vsak $w \in W$ velja naslednji račun $\langle u, (L_2 L_1)^* w\rangle_U = \langle L_2 L_1 u, w\rangle_W = \langle L_1 u, L_2^* w\rangle_V = \langle u, L_1^* L_2^* w\rangle_U$.

Seveda bi tudi lastnosti (2) in (3) lahko dokazali s prehodom na matrike.

Adjungirana preslikava - 2. del

5. Jedro in slika adjungirane preslikave

Recimo, da poznamo jedro in sliko linearne preslikave $L\colon U \to V$. Radi bi odtod dobili jedro in sliko adjungirane linearne preslikave $L^*\colon V \to U$. Spomnimo se, da velja $\langle Lu, v\rangle = \langle u, L^*v\rangle$ za vsak $u \in U$ in vsak $v \in V$.

Izrek

Za vsako linearno preslikavo $L\colon U \to V$, kjer sta U in V končnorazsežna vektorska prostora s skalarnim produktom, velja

$$\operatorname{Ker} L^* = (\operatorname{Im} L)^{\perp}. \tag{1}$$

Dokaz: Vzemimo poljuben vektor $v \in V$. Opazimo, da velja $L^*v = 0$ natanko tedaj, ko je $\langle u, L^*v\rangle = 0$ za vsak $u \in U$. Po definiciji L^* velja to natanko tedaj, ko je $\langle Lu, v\rangle = 0$ za vsak $u \in U$. Po definiciji $\operatorname{Im} L$ velja to natanko tedaj, ko je $\langle w, v\rangle = 0$ za vsak $w \in \operatorname{Im} L$. Po definiciji ortogonalnega komplementa velja to natanko tedaj, ko $v \in (\operatorname{Im} L)^{\perp}$. □

Oglejmo si nekaj preprostih posledic formule (1).

Če na obeh straneh (1) uporabimo ortogonalni komplement in upoštevamo, da je $((\operatorname{Im} L)^{\perp})^{\perp} = \operatorname{Im} L$, dobimo

$$\operatorname{Im} L = (\operatorname{Ker} L^*)^{\perp}. \tag{2}$$

Če v formuli (1) zamenjamo L z L^*, dobimo

$$\operatorname{Ker} L = (\operatorname{Im} L^*)^{\perp}. \tag{3}$$

Če v formuli (2) zamenjamo L z L^*, dobimo

$$\operatorname{Im} L^* = (\operatorname{Ker} L)^{\perp}. \tag{4}$$

Isto formulo dobimo, če na obeh straneh v (3) uporabimo ortogonalni komplement.

Pokažimo sedaj, da formule (1)-(4) veljajo tudi za matrike.

Najprej standardni skalarni produkt zapišemo v matrični obliki

$$\langle v, w \rangle = v_1 \bar{w}_1 + \ldots + v_n \bar{w}_n = \begin{bmatrix} \bar{w}_1 & \ldots & \bar{w}_n \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = w^* v. \qquad (5)$$

Naj bo A kompleksna $m \times n$ matrika. Opazimo, da iz (5) sledi

$$\langle Au, v \rangle = v^* Au = v^* A^{**} u = (A^* v)^* u = \langle u, A^* v \rangle \qquad (6)$$

za vsak $u \in \mathbb{C}^n$ in vsak $v \in \mathbb{C}^m$. V $\langle Au, v \rangle$ nastopa standardni skalarni produkt na $\mathbb{C}^m$, v $\langle u, A^* v \rangle$ pa standardni skalarni produkt na $\mathbb{C}^n$. Naj bo L_A linearna preslikava it $\mathbb{C}^n$ v $\mathbb{C}^m$ definirana z $L_A u = Au$. Potem velja

$$(L_A)^* = L_{A^*} \qquad (7)$$

saj po (6) velja $\langle u, (L_A)^* v \rangle = \langle L_A u, v \rangle = \langle Au, v \rangle = \langle u, A^* v \rangle = \langle u, L_{A^*} v \rangle$ za vsak $u \in \mathbb{C}^n$ in vsak $v \in \mathbb{C}^m$. Iz (7) sledi recimo verzija (3) za matrike

$$\operatorname{Ker} A = \operatorname{Ker} L_A = (\operatorname{Im} (L_A)^*)^\perp = (\operatorname{Im} L_{A^*})^\perp = (\operatorname{Im} A^*)^\perp.$$

Če je L slika iz U v V, potem L^*L slika iz U v U, LL^* pa iz V v V. Poleg tega iz lastnosti adjungiranja sledi

$$(L^*L)^* = L^*L \quad \text{in} \quad (LL^*)^* = LL^*. \tag{5}$$

Oglejmo si, kako izračunamo jedro in sliko preslikav L^*L in LL^*.

Trditev

Za vsako linearno preslikavo $L\colon U \to V$, kjer sta U in V končnorazsežna vektorska prostora s skalarnim produktom, velja

$$\operatorname{Ker} L^*L = \operatorname{Ker} L. \tag{6}$$

Dokaz: Vzemimo poljuben $u \in U$. Če $u \in \operatorname{Ker} L$, potem je $Lu = 0$. Odtod sledi $L^*Lu = L^*0 = 0$, torej $u \in \operatorname{Ker} L^*L$.
Dokažimo še obratno. Če $u \in \operatorname{Ker} L^*L$, potem $L^*Lu = 0$. Po definiciji L^* je $\langle Lu, Lu\rangle = \langle u, L^*Lu\rangle = \langle u, 0\rangle = 0$. Odtod sledi $Lu = 0$, torej je $u \in \operatorname{Ker} L$.

Oglejmo si nekaj posledic formule (6).

Če na obeh straneh formule (6) vzamemo ortogonalni komplement in upoštevamo, da po formuli (4) velja $\operatorname{Im} L^* = (\operatorname{Ker} L)^\perp$ in $\operatorname{Im} L^*L = \operatorname{Im} (L^*L)^* = (\operatorname{Ker} L^*L)^\perp$, potem dobimo

$$\operatorname{Im} L^*L = \operatorname{Im} L^*. \tag{7}$$

Če v formulah (6) in (7) zamenjamo L z L^*, potem dobimo

$$\operatorname{Ker} LL^* = \operatorname{Ker} L^*, \tag{8}$$

$$\operatorname{Im} LL^* = \operatorname{Im} L. \tag{9}$$

Opomba: Pravimo, da je $L\colon V \to V$ **normalna** linearna preslikava, če velja $LL^* = L^*L$. Za normalne linearne preslikave iz (6) in (8) sledi, da je $\operatorname{Ker} L^* = \operatorname{Ker} L$, iz (7) in (9) pa sledi, da je $\operatorname{Im} L^* = \operatorname{Im} L$.

6. Lastne vrednosti adjungirane matrike

Recimo, da je A kompleksna $n \times n$ matrika. Zanima nas, kakšna je zveza med lastnimi vrednostmi matrike A in lastnimi vrednostmi matrike A^*.

Izrek o lastnih vrednostih adjungirane matrike

Naj bo A kvadratna matrika nad $\mathbb{C}$ in naj bo $\lambda \in \mathbb{C}$. Potem je λ lastna vrednost za A natanko tedaj, ko je $\bar{\lambda}$ lastna vrednost za A^*.

Dokaz: Radi bi dokazali, da velja $\operatorname{Ker}(A - \lambda I) \neq \{0\}$ natanko tedaj, ko velja $\operatorname{Ker}(A^* - \bar{\lambda} I) \neq \{0\}$. Zadošča dokazati, da velja

$$\dim \operatorname{Ker}(A - \lambda I) = \dim \operatorname{Ker}(A^* - \bar{\lambda} I). \tag{1}$$

Označimo $B = A - \lambda I$. Iz lastnosti operacije adjungiranja sledi, da je $B^* = A^* - \bar{\lambda} I$. Torej zadošča dokazati, da je

$$\dim \operatorname{Ker} B = \dim \operatorname{Ker} B^*. \tag{2}$$

Iz izreka $\operatorname{Ker} B^* = (\operatorname{Im} B)^\perp$ sledi

$$\dim \operatorname{Ker} B^* = \dim (\operatorname{Im} B)^\perp \tag{3}$$

Po izreku o ortogonalnem razcepu, je

$$\dim (\operatorname{Im} B)^\perp = n - \dim \operatorname{Im} B \tag{4}$$

kjer je n velikost matrike A. Po osnovnem izreku je

$$n - \dim \operatorname{Im} B = \dim \operatorname{Ker} B. \tag{5}$$

Iz formul (3), (4) in (5) sledi formula (2). □

Oglejmo si še en dokaz, ki je bolj računski.

Najprej opazimo, da za vsako kompleksno kvadratno matriko B velja

$$\det B^* = \det(\overline{B})^T = \det \overline{B} = \overline{\det B}.$$

Če vstavimo $B = A - x\,I$, dobimo

$$\det(A^* - \bar{x}\,I) = \overline{\det(A - x\,I)},$$

torej za karakteristična polinoma A in A^* velja

$$p_{A^*}(\bar{x}) = \overline{p_A(x)}.$$

Torej je $p_A(x) = 0$ natanko tedaj, ko je $p_{A^*}(\bar{x}) = 0$, kar dokaže izrek. □

Opomba: Iz tega dokaza je razvidno, kako se $p_{A^*}(x)$ izraža z $p_A(x)$.
Iz $p_A(x) = c_0 + c_1 x + \ldots + c_n x^n$ namreč sledi, da velja
$p_{A^*}(x) = \overline{p_A(\bar{x})} = \overline{c_0 + c_1 \bar{x} + \ldots + c_n \bar{x}^n} = \bar{c}_0 + \bar{c}_1 x + \ldots + \bar{c}_n x^n$.

Primer

Določi lastne vrednosti in lastne vektorje matrike

$$A = \begin{bmatrix} 1 & 2 \\ i & 1 \end{bmatrix}$$

in njene adjungirane matrike A^*.

Rešitev: Karakteristični polinom matrike A je

$$\det(A - \lambda I) = (1-\lambda)^2 - 2i = (1-\lambda)^2 - (1+i)^2 = (\lambda + i)(\lambda - 2 - i)$$

torej sta $-i$ in $2 + i$ lastni vrednosti matrike A. Podobno je

$$\det(A^* - \lambda I) = (1-\lambda)^2 + 2i = (1-\lambda)^2 - (1-i)^2 = (\lambda - i)(\lambda - 2 + i)$$

karakteristični polinom matrike A^*. Torej sta i in $2 - i$ lastni vrednosti matrike A^*. Torej so lastne vrednosti za A^* res konjugirane lastnim vrednostim za A.

Lastna podprostora za A sta

$$\operatorname{Ker}(A + iI) = \operatorname{Lin}\begin{bmatrix} 1-i \\ -1 \end{bmatrix}, \quad \operatorname{Ker}(A - (2+i)I) = \operatorname{Lin}\begin{bmatrix} 1-i \\ 1 \end{bmatrix}.$$

Lastna podprostora za A^* pa sta

$$\operatorname{Ker}(A^* - iI) = \operatorname{Lin}\begin{bmatrix} 1-i \\ -2 \end{bmatrix}, \quad \operatorname{Ker}(A^* - (2-i)I) = \operatorname{Lin}\begin{bmatrix} 1-i \\ 2 \end{bmatrix}.$$

Odtod je razvidno, da $\operatorname{Ker}(A - \lambda I)$ in $\operatorname{Ker}(A^* - \bar{\lambda} I)$ nista nujno enaka.

Opomba: Pokazali bomo, da za normalne matrike vedno velja $\operatorname{Ker}(A - \lambda I) = \operatorname{Ker}(A^* - \bar{\lambda} I)$. Gornji primer kaže, da to v splošnem ni res. Vemo pa, da vedno velja $\dim \operatorname{Ker}(A - \lambda I) = \dim \operatorname{Ker}(A^* - \bar{\lambda} I)$.

7. Simetrične in hermitske matrike

Oglejmo si pomemben razred linearnih preslikav.

Definicija sebiadjungirane preslikave

Naj bo V končnorazsežen vektorski prostor s skalarnim produktom. Linearna preslikava $L\colon V \to V$ je **sebiadjungirana**, če velja $L = L^*$.

Opomba: Če je A matrika sebiadjungirane preslikave $L\colon V \to V$ glede na ortonormirano bazo $\mathcal{B}$ za V, potem je $A^* = \left([L]_{\mathcal{B}}\right)^* = [L^*]_{\mathcal{B}} = [L]_{\mathcal{B}} = A$. Taki matriki včasih pravimo **sebiadjungirana** matrika, vendar se pogosteje uporablja naslednja terminologija:

Definicija simetrične in hermitske matrike

Kompleksna matrika A je **hermitska**, če zadošča $A = A^*$. Realnim hermitskim matrikam pravimo tudi **simetrične** matrike.

Opomba: Za vsako matriko A sta A^*A in AA^* hermitski matriki. Vsaka potenca hermitske matrike je spet hermitska matrika.

Primer: Ortogonalni projektorji

Naj bo V vektorski prostor s skalarnim produktom, W podprostor v V in $L\colon V \to V$ taka linearna preslikava, ki vsakemu vektorju iz V priredi njegovo ortogonalno projekcijo na W. Pokažimo, da je $L = L^*$ in $L = L^2$.

Dokaz: Naj bo $w_1, \ldots, w_k$ ortonormirana baza za W. Potem za vsak $v \in V$ velja $Lv = \sum_{i=1}^{k} \langle v, w_i \rangle w_i$. Za vsak $v' \in V$ velja

$$\langle Lv, v' \rangle = \langle \sum_{i=1}^{k} \langle v, w_i \rangle w_i, v' \rangle = \sum_{i=1}^{k} \langle v, w_i \rangle \langle w_i, v' \rangle,$$

$$\langle v, Lv' \rangle = \langle v, \sum_{i=1}^{k} \langle v', w_i \rangle w_i \rangle = \sum_{i=1}^{k} \overline{\langle v', w_i \rangle} \langle v, w_i \rangle.$$

Torej je za vsaka $v, v' \in V$ velja $\langle Lv, v' \rangle = \langle v, Lv' \rangle$ kar pomeni $L = L^*$. Ker je $Lw_i = w_i$ za vsak i, je

$$L^2 v = \sum_{i=1}^{k} \langle v, w_i \rangle Lw_i = \sum_{i=1}^{k} \langle v, w_i \rangle w_i = Lv.$$

Kaj lahko povemo o lastnih vrednostih hermitskih matrik?

Trditev 1

Vse lastne vrednosti hermitska matrike so realne.

Dokaz: Naj bo λ lastna vrednost hermitske matrike in naj bo v pripadajoči lastni vektor. Potem za standardni skalarni produkt velja

$$\lambda\langle v, v\rangle = \langle \lambda v, v\rangle = \langle Av, v\rangle = \langle v, A^* v\rangle = \langle v, Av\rangle = \langle v, \lambda v\rangle = \bar{\lambda}\langle v, v\rangle$$

Ker je $v \neq 0$, lahko pokrajšamo $\langle v, v\rangle$ in dobimo $\lambda = \bar{\lambda}$. □

Kaj lahko povemo o lastnih vektorjih hermitskih matrik?

Trditev 2

Lastna vektorja hermitske matrike, ki pripadata različnim lastnim vrednostim, sta ortogonalna glede na standardni skalarni produkt.

Dokaz: Naj bo A hermitska matrika in naj bosta u in v njena lastna vektorja. Naj bosta λ in μ pripadajoči lastni vrednosti, se pravi

$$Au = \lambda u \quad \text{in} \quad Av = \mu v.$$

Po prejšnji trditvi sta λ in μ realni števili. Odtod sledi

$$\langle Au, v\rangle = \langle \lambda u, v\rangle = \lambda\langle u, v\rangle \tag{1}$$

$$\langle u, Av\rangle = \langle u, \mu v\rangle = \bar{\mu}\langle u, v\rangle = \mu\langle u, v\rangle \tag{2}$$

Ker je A hermitska matrika, velja

$$\langle Au, v\rangle = \langle u, A^* v\rangle = \langle u, Av\rangle \tag{3}$$

Iz (1), (2) in (3) sledi

$$\lambda\langle u, v\rangle = \mu\langle u, v\rangle \tag{4}$$

Če je $\lambda \neq \mu$, iz (4) sledi, da je

$$\langle u, v\rangle = 0.$$

Kaj lahko povemo o Jordanski formi hermitske matrike?

Trditev 3

Vsaka hermitska matrika je podobna diagonalni matriki.

Dokaz. Zadošča dokazati, da se korenski podprostori ujemajo z lastnimi podprostori. Odtod namreč sledi, da so vse Jordanske kletke velikosti 1.

Naj bo λ lastna vrednost hermitske matrike A. Radi bi dokazali, da za vsako naravno število m velja $\operatorname{Ker}(A-\lambda I)^m = \operatorname{Ker}(A-\lambda I)$.

Označimo $B = A - \lambda I$ in opazimo, da je tudi matrika B hermitska.
Iz formule $\operatorname{Ker} B^*B = \operatorname{Ker} B$ torej sledi, da je $\operatorname{Ker} B^2 = \operatorname{Ker} B$.
Če namesto B vstavimo B^{2^k} za $k = 1, \ldots, m-1$, dobimo

$$\operatorname{Ker} B^{2^m} = \operatorname{Ker} B^{2^{m-1}} = \operatorname{Ker} B^{2^{m-2}} = \ldots = \operatorname{Ker} B.$$

Ker je $2^m \geq m$ za vsak m, odtod sledi

$$\operatorname{Ker} B \subseteq \operatorname{Ker} B^m \subseteq \operatorname{Ker} B^{2^m} = \operatorname{Ker} B$$

Torej res za vsak m velja

$$\operatorname{Ker} B^m = \operatorname{Ker} B.$$

Definicija unitarne in ortogonalne matrike

Kompleksna $n \times n$ matrika je **unitarna**, če njeni stolpci tvorijo ortonomirano bazo za $\mathbb{C}^n$ glede na standardni skalarni produkt.

Realni unitarni matriki pravimo tudi **ortogonalna** matrika.

Opomba: Za unitarno matriko P velja $PP^* = I$. Za ortogonalno matriko P velja $PP^T = I$. V obeh primerih je torej $P^{-1} = P^*$.

Izrek

Za vsako hermitsko matriko A obstaja taka unitarna matrika P in taka realna diagonalna matrika D, da velja $A = PDP^{-1}$.

Za vsako simetrično matriko A obstaja taka ortogonalna matrika P in taka realna diagonalna matrika D, da velja $A = PDP^{-1}$.

Dokaz prvega dela: Naj bo A hermitska $n \times n$ matrika. Za vsak lastni podprostor matrike A izberimo ortonormirano bazo. Naj bo $\mathcal{B}$ unija izbranih ortonormiranih baz po vseh lastnih podprostorih matrike A. Ker so lastni podprostori matrike A paroma ortogonalni (po Trditvi 2), je $\mathcal{B}$ ortogonalna množica. Ker je $\mathbb{C}^n$ direktna vsota lastnih podprostorov matrike A (po Trditvi 3), je $\mathcal{B}$ ogrodje za $\mathbb{C}^n$. Torej je $\mathcal{B}$ ortonormirana baza za $\mathbb{C}^n$. Sestavljena je iz lastnih vektorjev matrike A.

Naj bo P matrika, katere stolpci so elementi baze $\mathcal{B}$. Potem je P unitarna matrika, katere stolpci so lastni vektorji matrike A, torej je

$$AP = A \begin{bmatrix} v_1 & \dots & v_n \end{bmatrix} = \begin{bmatrix} Av_1 & \dots & Av_n \end{bmatrix} =$$

$$= \begin{bmatrix} \lambda_1 v_1 & \dots & \lambda_n v_n \end{bmatrix} = \begin{bmatrix} v_1 & \dots & v_n \end{bmatrix} \begin{bmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{bmatrix} = PD \quad \square$$

Dokaz drugega dela. Naj bo A simetrična matrika. Naj bodo $\lambda_1, \ldots, \lambda_k$ vse paroma različne lastne vrednosti matrike A. Ker so matrike $A - \lambda_i I$ realne (po Trditvi 1), lahko za lastni podprostor $\mathrm{Ker}(A - \lambda_i I)$ izberemo ortonormirano bazo $\mathcal{B}_i$ iz realnih vektorjev. (Gram-Schmidtova ortogonalizacija nam realno bazo spremeni v realno ortogonalno bazo.) Ker so lastni podprostori matrike A paroma ortogonalni in ker je njihova vsota enaka $\mathbb{C}^n$ je $\bigcup_{i=1}^{k} \mathcal{B}_i$ ortonormirana baza za $\mathbb{C}^n$. Elemente te baze vzamemo za stolpce matrike P. □

Še en dokaz drugega dela: Dokažimo s popolno indukcijo, da ima $\mathbb{R}^n$ ortonormirano bazo iz lastnih vektorjev matrike A. Recimo, da je $\mathcal{B} = \{v_1, \ldots, v_r\}$ ortonormirana množica v $\mathbb{R}^n$ iz lastnih vektorjev A. Trdimo, da za vsak $w \in \mathcal{B}^\perp$ velja $Aw \in \mathcal{B}^\perp$. Za vsak $v \in \mathcal{B}$ je namreč $\langle Aw, v\rangle = \langle w, Av\rangle = \langle w, \lambda v\rangle = \bar{\lambda}\langle w, v\rangle = 0$. Ker je $\mathcal{B}^\perp$ invarianten za A, vsebuje vsaj en realen lastni vektor za A. Če ta lastni vektor normiramo, dobimo tak vektor v_{r+1}, da je $\{v_1, \ldots, v_r, v_{r+1}\}$ ortonormirana množica v $\mathbb{R}^n$ iz lastnih vektorjev matrike A. □

Primer

Naj bo

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 3 & -1 \\ 1 & -1 & 3 \end{bmatrix}$$

Poišči tako ortogonalno matriko P in tako diagonalno matriko D, da je

$$A = PDP^{-1}.$$

Karakteristični polinom matrike A je

$$\det(A - xI) = -x^3 + 7x^2 - 12x = -x(x-3)(x-4)$$

Lastni vektorji, ki pripadajo lastnim vrednostim 4, 3, 0 so

$$\begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$$

Očitno so to vektorji paroma ortogonalni. Če jih normiramo, dobimo

$$\begin{bmatrix} 0 \\ -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}, \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}, \begin{bmatrix} -2/\sqrt{6} \\ 1/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}$$

Iskana matrika P je torej

$$P = \begin{bmatrix} 0 & 1/\sqrt{3} & -2/\sqrt{6} \\ -1/\sqrt{2} & 1/\sqrt{3} & 1/\sqrt{6} \\ 1/\sqrt{2} & 1/\sqrt{3} & 1/\sqrt{6} \end{bmatrix}$$

iskana matrika D pa je

$$D = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Adjungirana preslikava - 3. del

8. Normalne matrike

Običajno matrika ne komutira s svojo adjungirano matriko. Kadar pa komutira, ima zelo lepe lastnosti.

Definicija normalne matrike

Kompleksna matrika A je **normalna**, če zadošča $A^*A = AA^*$.

Najprej zabeležimo preprosto ugotovitev.

Trditev

Vsaka normalna matrika je kvadratna.

Dokaz: Če je A $m \times n$ matrika, potem je $A^* = (\overline{A})^T$ $n \times m$ matrika. Matrika A^*A je potem $n \times n$ matrika, matrika AA^* pa $m \times m$ matrika. Če je A normalna, potem velja $n = m$, torej je A kvadratna matrika. □

Oglejmo si nekaj primerov normalnih matrik.

Primer 1: Hermitske matrike so normalne

Če je A hermitska, potem velja $A^* = A$. Odtod sledi $A^*A = A^2 = AA^*$. Torej je A normalna matrika.

Primer 2: Unitarne matrike so normalne

Če je A unitarna, potem je $A^* = A^{-1}$. Odtod sledi $A^*A = I = AA^*$. Torej je A normalna matrika.

Primer 3: Diagonalne matrike so normalne

Če je A diagonalna, potem je tudi A^* diagonalna. Dve diagonalni matriki vedno komutirata, zato je $A^*A = AA^*$. Torej je A normalna matrika.

Primer 4: Iz $A = A^T$ ne sledi nujno, da je A normalna

Matrika $A = \begin{bmatrix} 1 & i \\ i & 0 \end{bmatrix}$ zadošča $A = A^T$, vendar ni normalna.

Primer 5: Normalne 2×2 matrike

Matrika $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ je normalna $\iff |b| = |c|$ in $\bar{c}(a-d) = b(\bar{a}-\bar{d})$.

Dokaz: 2×2 matrika je normalna natanko tedaj, ko velja

$$\begin{bmatrix} \bar{a} & \bar{c} \\ \bar{b} & \bar{d} \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} \bar{a} & \bar{c} \\ \bar{b} & \bar{d} \end{bmatrix}$$

To pa velja natanko tedaj, ko je

$$\begin{aligned} \bar{a}a + \bar{c}c &= a\bar{a} + b\bar{b} & \bar{a}b + \bar{c}d &= a\bar{c} + b\bar{d} \\ \bar{b}a + \bar{d}c &= c\bar{a} + d\bar{b} & \bar{b}b + \bar{d}d &= c\bar{c} + d\bar{d} \end{aligned}$$

Prva in četrta enakost sta ekvivalentni z $b\bar{b} = c\bar{c}$, se pravi z

$$|b| = |c|$$

Tretja enakost je samo konjugirana druga enakost, ki je ekvivalentna z

$$\bar{c}(a-d) = b(\bar{a}-\bar{d})$$

Če je λ lastna vrednost matrike A, potem vemo, da je $\bar{\lambda}$ lastna vrednost matrike A^*. Vemo tudi, da v splošnem iz $Av = \lambda v$ ne sledi $A^*v = \bar{\lambda}v$. Velja pa to za normalne matrike, kar bomo dokazali v dveh korakih.

Trditev

Če je matrika A normalna, potem je za vsak $\lambda \in \mathbb{C}$ tudi $A - \lambda I$ normalna.

Dokaz. Označimo $B = A - \lambda I$. Iz lastnosti adjungiranja sledi, da je $B^* = A^* - \bar{\lambda} I$. Ker je A normalna, velja $A^*A = AA^*$. Odtod sledi $BB^* = AA^* - \lambda A^* - \bar{\lambda} A + \lambda\bar{\lambda} I = A^*A - \lambda A^* - \bar{\lambda} A + \lambda\bar{\lambda} I = B^*B$. □

Trditev

Če je matrika A normalna, potem imata A in A^* enake lastne vektorje.

Dokaz: Naj bo λ lastna vrednost normalne matrike A. Po prejšnji trditvi je tudi $B = A - \lambda I$ normalna matrika. Vemo že, da odtod sledi

$$\operatorname{Ker} B = \operatorname{Ker} B^*B = \operatorname{Ker} BB^* = \operatorname{Ker} B^*.$$

Torej za normalno matriko A velja

$$\operatorname{Ker} (A - \lambda I) = \operatorname{Ker} (A^* - \bar{\lambda} I).$$

Izrek

Vsaka normalna matrika je podobna diagonalni matriki.

Dokaz: Radi bi dokazali, da je vsak korenski podprostor normalne matrike že kar lastni podprostor. Iz korenskega razcepa potem sledi, da ima matrika n linearno neodvisnih lastnih vektorjev, torej se da diagonalizirati.

Naj bo A normalna matrika in naj bo λ njena lastna vrednost. Vemo, da je potem tudi matrika $B = A - \lambda I$ normalna. Radi bi dokazali, da velja

$$\operatorname{Ker} B^2 = \operatorname{Ker} B. \tag{1}$$

Odtod s popolno indukcijo dokažemo, da velja $\operatorname{Ker} B^{2^k} = \operatorname{Ker} B$ za vsak k. Ker je $2^k \geq k$ za vsak k, odtod sledi $\operatorname{Ker} B \subseteq \operatorname{Ker} B^k \subseteq \operatorname{Ker} B^{2^k} = \operatorname{Ker} B$, torej je $\operatorname{Ker} B^k = \operatorname{Ker} B$ za vsak k. To pomeni, da je korenski podprostor za λ res enak lastnemu podprostoru za λ.

Vemo že, da velja

$$\operatorname{Ker} B^*B = \operatorname{Ker} B. \tag{2}$$

Če v formulo (2) namesto B vstavimo B^*B, dobimo

$$\operatorname{Ker} (B^*B)^2 = \operatorname{Ker} B^*B. \tag{3}$$

Če v formulo (2) namesto B vstavimo B^2, dobimo

$$\operatorname{Ker} (B^2)^*B^2 = \operatorname{Ker} B^2. \tag{4}$$

Ker je B normalna preslikava, velja

$$(B^2)^*B^2 = B^*B^*BB = B^*BB^*B = (B^*B)^2. \tag{5}$$

Lastnost (1) sledi iz formul (2)-(5). □

Izrek

Lastni podprostori normalne matrike so paroma ortogonalni.

Dokaz: Recimo, da je A normalna matrika in da sta λ in μ različni lastni vrednosti za A. Radi bi pokazali, da je vsak vektor $u \in \mathrm{Ker}\,(A - \lambda I)$ pravokoten na vsak vektor $v \in \mathrm{Ker}\,(A - \mu I)$ glede na standardni skalarni produkt na $\mathbb{C}^n$.

Dokazali smo že, da velja $\mathrm{Ker}\,(A - \mu I) = \mathrm{Ker}\,(A^* - \bar{\mu} I)$, torej je

$$\mu\langle u, v\rangle = \langle u, \bar{\mu} v\rangle = \langle u, A^* v\rangle = \langle Au, v\rangle = \langle \lambda u, v\rangle = \lambda\langle u, v\rangle.$$

Ker je $\mu \neq \lambda$, odtod sledi $\langle u, v\rangle = 0$, kar smo želeli dokazati. □

Sedaj lahko povemo še več o diagonalizaciji normalnih matrik.

Izrek

Matrika A je normalna natanko tedaj, ko obstajata taka diagonalna matrika D in taka unitarna matrika P, da velja $A = PDP^{-1}$.

Dokaz: Če je $A = PDP^{-1}$, kjer je D diagonalna in P zadošča $P^{-1} = P^*$, potem velja

$$A^*A = (PDP^*)^*PDP^* = PD^*P^*PDP^* = PD^*DP^*$$

in

$$AA^* = PDP^*(PDP^*)^* = PDP^*PD^*P^* = PDD^*P^*.$$

Ker je $D^*D = DD^*$, sledi $A^*A = AA^*$.

Privzemimo sedaj, da je A normalna $n \times n$ matrika. Vemo, da je vsota vseh lastnih podprostorov matrike A enaka $\mathbb{C}^n$ in da so lastni podprostori paroma pravokotni glede na standardni skalarni produkt. Če za vsak lastni podprostor izberemo ortonormirano bazo, potem je unija teh baz ortonormirana baza za $\mathbb{C}^n$. Označimo to ortonormirano bazo z $v_1, \ldots, v_n$. Naj bodo $\lambda_1, \ldots, \lambda_n$ pripadajoče lastne vrednosti. Označimo

$$P = \begin{bmatrix} v_1 & \ldots & v_n \end{bmatrix} \quad \text{in} \quad D = \begin{bmatrix} \lambda_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \lambda_n \end{bmatrix}.$$

Ker je $v_1, \ldots, v_n$ ortonormirana baza glede na standardni skalarni produkt, je

$$P^* P = I.$$

Ker so $v_1, \ldots, v_n$ lastni vektorji z lastnimi vrednostmi $\lambda_1, \ldots, \lambda_n$, je

$$AP = \begin{bmatrix} \lambda_1 v_1 & \ldots & \lambda_n v_n \end{bmatrix} = PD.$$

Oglejmo si še normalne preslikave in njihovo zvezo z normalnimi matrikami.

Definicija

Naj bo V končnorazsežen vektorski prostor s skalarnim produktom. Linearna preslikava $L\colon V \to V$ je normalna, če velja $L^*L = LL^*$.

Trditev

Naj bo V končnorazsežen vektorski prostor s skalarnim produktom in $\mathcal{B}$ ortonormirana baza za V. Potem je linearna preslikava $L\colon V \to V$ normalna natanko tedaj, ko je njena matrika $[L]_{\mathcal{B}}$ normalna.

Dokaz: Vemo, da velja $[L^*]_{\mathcal{B}} = \big([L]_{\mathcal{B}}\big)^*$. Odtod sledi, da je $[L^*L]_{\mathcal{B}} = [L^*]_{\mathcal{B}}[L]_{\mathcal{B}} = \big([L]_{\mathcal{B}}\big)^*[L]_{\mathcal{B}}$ in $[LL^*]_{\mathcal{B}} = [L]_{\mathcal{B}}[L^*]_{\mathcal{B}} = [L]_{\mathcal{B}}\big([L]_{\mathcal{B}}\big)^*$. Torej velja $\big([L]_{\mathcal{B}}\big)^*[L]_{\mathcal{B}} = [L]_{\mathcal{B}}\big([L]_{\mathcal{B}}\big)^*$ natanko tedaj, ko je $[L^*L]_{\mathcal{B}} = [LL^*]_{\mathcal{B}}$. Slednje velja natanko tedaj, ko je $L^*L = LL^*$. □

Opomba: Naš glavni izrek za normalne matrike lahko povemo tudi takole: Linearna preslikava $L\colon V \to V$ je normalna natanko tedaj ko obstaja taka ortonormirana baza $\mathcal{B}$ za V, da je matrika $[L]_{\mathcal{B}}$ diagonalna. (Velja nad $\mathbb{C}$.)

8. Izometrije

Izometrije so preslikave, ki ohranjajo razdaljo. Ker razdaljo definiramo z normo, jih definiramo kot linearne preslikave, ki ohranjajo normo.

Definicija izometrije

Naj bosta U in V končnorazsežna vektorska prostora s skalarnim produktom. Pravimo, da je linearna preslikava $L\colon U \to V$ **izometrija**, če za vsak $u \in U$ velja

$$\|Lu\|_V = \|u\|_U.$$

Opomba: Vsaka izometrija je injektivna, saj iz $Lu = 0$ in iz $\|Lu\| = \|u\|$ sledi $u = 0$. Izometrija je surjektivna natanko tedaj, ko je $\dim U = \dim V$.

Primer izometrije, ki ni surjektivna

Preslikava $L\colon \mathbb{C}^2 \to \mathbb{C}^5$, kjer $L(x, y) = (x, y, 0, 0, 0)$ je izometrija, če sta $\mathbb{C}^2$ in $\mathbb{C}^5$ opremljena s standardnim skalarnim produktom.

Opomba: Kadar je $V = U$ pravimo izometriji tudi unitarna linearna preslikava. Vsaka unitarna linearna preslikava je bijektivna.

Oglejmo si najprej nekaj karakterizacij izometrij.

Trditev

Za linearno preslikavo $L\colon U \to V$, kjer sta U in V končnorazsežna vektorska prostora s skalarnim produktom, so ekvivalentne trditve:

1. L je izometrija
2. $\langle Lu, Lu' \rangle_V = \langle u, u' \rangle_U$ za vsaka $u, u' \in U$
3. $L^*L = \mathrm{id}_U$ (identična preslikava na U)
4. Za vsako ortonormirano bazo $u_1, \ldots, u_n$ v U je $Lu_1, \ldots, Lu_n$ ortonormirana množica v V.
5. Za neko ortonormirano bazo $u_1, \ldots, u_n$ v U je $Lu_1, \ldots, Lu_n$ ortonormirana množica v V.

Opomba: Te karakterizacije seveda veljajo tudi za unitarne linearne preslikave, saj so te posebni primeri izometrij.

Dokaz: Dokažimo najprej ekvivalenco med (1) in (2). Če velja (2), potem je $\|Lu\|^2 = \langle Lu, Lu\rangle = \langle u, u\rangle = \|u\|^2$ za vsak $u \in U$, torej je L izometrija. Recimo, da je L izometrija. V kompleksnem primeru velja

$$\begin{aligned}
\langle Lu, Lu'\rangle &= \frac{1}{4}\sum_{k=0}^{3} i^k \|Lu + i^k Lu'\|^2 \\
&= \frac{1}{4}\sum_{k=0}^{3} i^k \|L(u + i^k u')\|^2 \\
&= \frac{1}{4}\sum_{k=0}^{3} i^k \|u + i^k u'\|^2 \\
&= \langle u, u'\rangle.
\end{aligned}$$

Torej L zadošča (2). Podobno velja tudi v realnem primeru, samo drugo polarizacijsko identiteto uporabimo.

Dokažimo sedaj ekvivalenco med (2) in (3). Če velja (3), potem je

$$\langle Lu, Lu' \rangle = \langle u, L^* Lu' \rangle = \langle u, \mathrm{id}_U u' \rangle = \langle u, u' \rangle$$

za vsaka $u, u' \in U$. Torej velja (2). Obratno, če velja (2), potem je

$$\langle u, L^* Lu' - u' \rangle = \langle u, L^* Lu' \rangle - \langle u, u' \rangle = \langle Lu, Lu' \rangle - \langle u, u' \rangle = 0$$

za vsaka $u, u' \in U$. Če vstavimo $u = L^* Lu' - u'$, dobimo, da je $L^* Lu' = u'$ za vsak $u' \in U$. Torej velja (3).

Očitno iz (2) sledi (4) in iz (4) sledi (5). Dokažimo še, da iz (5) sledi (1). Naj bo $u_1, \ldots, u_n$ taka ortonormirana baza za U, da je $Lu_1, \ldots, Lu_n$ ortonormirana množica v V. Vzemimo poljuben element $u \in U$ in ga razvijmo po ortonormirani bazi; recimo $u = \alpha_1 u_1 + \ldots + \alpha_n u_n$. Po Parsevalovi identiteti velja $\|u\|^2 = |\alpha_1|^2 + \ldots + |\alpha_n|^2$. Podobno iz $Lu = \alpha_1 Lu_1 + \ldots + \alpha_n Lu_n$ in ortonormiranosti $Lu_1, \ldots, Lu_n$ sledi, da je $\|Lu\|^2 = |\alpha_1|^2 + \ldots + |\alpha_n|^2$. Torej je $\|Lu\| = \|u\|$ za vsak $u \in U$. □

Kaj lahko povemo o matriki izometrije?

Trditev

Naj bosta U, V končnorazsežna vektorska prostora s skalarnim produktom. Naj bo $\mathcal{B}$ ortonormirana baza za U in naj bo $\mathcal{C}$ ortonormirana baza za V. Potem je linearna preslikava $L\colon U \to V$ izometrija natanko tedaj, ko za njeno matriko $A = [L]_{\mathcal{C}\leftarrow\mathcal{B}}$ velja $A^*A = I$.

Dokaz: Vemo, da je linearna preslikava $L\colon U \to V$ izometrija natanko tedaj, ko je $L^*L = \mathrm{id}_U$. Očitno to velja natanko tedaj, ko je $[L^*L]_{\mathcal{B}} = I$. Upoštevajmo še, da je $[L^*L]_{\mathcal{B}} = [L^*]_{\mathcal{B}\leftarrow\mathcal{C}}[L]_{\mathcal{C}\leftarrow\mathcal{B}}$ in $[L^*]_{\mathcal{B}\leftarrow\mathcal{C}} = \left([L]_{\mathcal{C}\leftarrow\mathcal{B}}\right)^*$. Torej je L izometrija natanko tedaj, ko je $\left([L]_{\mathcal{C}\leftarrow\mathcal{B}}\right)^*\left([L]_{\mathcal{C}\leftarrow\mathcal{B}}\right) = I$. □

Opomba: Če ortonormirani bazi $\mathcal{B}$ in $\mathcal{C}$ spretno izberemo, se matrika izometrije zelo poenostavi. Izberimo tako ortonormirano bazo $\mathcal{B} = \{u_1, \ldots, u_n\}$ v U, da je $v_1 = Lu_1, \ldots, v_n = Lu_n$ ortonormirana množica v V. Naj bo $v_{n+1}, \ldots, v_m$ njena dopolnitev do ortonormirane baze $\mathcal{C}$ za V. Potem je matrika $[L]_{\mathcal{C}\leftarrow\mathcal{B}}$ oblike $\begin{bmatrix} I \\ 0 \end{bmatrix}$.

Trditev

Za pravokotno kompleksno matriko A velja $A^*A = I$ natanko tedaj, ko so stolpci A ortonormirana množica glede na standardni skalarni produkt.

Dokaz: Naj bodo $v_1, \ldots, v_n$ stolpci matrike A. Potem velja

$$A^*A = \begin{bmatrix} v_1^* \\ \vdots \\ v_n^* \end{bmatrix} \begin{bmatrix} v_1 & \ldots & v_n \end{bmatrix} = \begin{bmatrix} v_1^*v_1 & \ldots & v_1^*v_n \\ \vdots & \ddots & \vdots \\ v_n^*v_1 & \ldots & v_n^*v_n \end{bmatrix}$$

kjer je $v_j^*v_i = \langle v_i, v_j \rangle$ standardni skalarni produkt. Odtod sledi, da velja $A^*A = I$ natanko tedaj, ko je

$$\langle v_i, v_j \rangle = \begin{cases} 1 & \text{če } i = j \\ 0 & \text{če } i \neq j \end{cases}$$

torej natanko tedaj, ko je $\{v_1, \ldots, v_n\}$ ortonormirana množica glede na standardni skalarni produkt. □

Naše ugotovitve o matrikah izometrij povzemimo v naslednjo trditev:

Trditev

Za vsako kompleksno $m \times n$ matriko A, kjer je $m \geq n$, so ekvivalentne naslednje trditve:

1. $A^*A = I$.
2. Stolpci matrike A so ortonormirana množica v $\mathbb{C}^m$ glede na standardni skalarni produkt.
3. Linearna preslikava $L_A \colon \mathbb{C}^n \to \mathbb{C}^m, v \mapsto Av$ je izometrija.
4. Obstaja taka izometrija $L \colon U \to V$, kjer dim $U = n$ in dim $V = m$, in taki ortonormirani bazi $\mathcal{B}$ za U in $\mathcal{C}$ za V, da velja $A = [L]_{\mathcal{C} \leftarrow \mathcal{B}}$.
5. Obstaja taka unitarna $m \times m$ matrika P in taka unitarna $n \times n$ matrika Q, da velja $A = P \begin{bmatrix} I_n \\ 0 \end{bmatrix} Q$.

9. Ortogonalne in unitarne matrike

Začnimo kar z definicijo.

Definicija ortogonalne matrike in unitarne matrike

Naj bo A kompleksna kvadratna matrika, ki zadošča $A^*A = I$. Potem pravimo, da je A **unitarna** matrika. Realni unitarni matriki pravimo tudi **ortogonalna** matrika.

Opomba: Vemo, da za kvadratne matrike iz lastnosti $A^*A = I$ sledi $AA^* = I$. Odtod sledijo naslednje preproste lastnosti unitarnih matrik:

- Vsaka unitarna matrika je normalna.
- Vsaka unitarna matrika A je obrnljiva in zadošča $A^{-1} = A^*$.
- Če je A unitarna matrika, potem je tudi A^* unitarna matrika.

Opomba: Dokazali smo že, da je kvadratna matrika unitarna natanko tedaj, ko njeni stolpci tvorijo ortonormirano bazo glede na standardni skalarni produkt. To nam da veliko primerov unitarnih matrik.

Primer: Klasifikacija ortogonalnih 2×2 matrik

Za vsak realen φ sta matriki

$$\begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix} \quad \text{in} \quad \begin{bmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{bmatrix}$$

ortogonalni. Vsaka ortogonalna 2×2 matrike je ene od teh dveh oblik.

Dokaz: Obe matriki sta realni, kvadratni in zadoščata $A^*A = I$, torej sta ortogonalni. Če je matrika

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

ortogonalna, potem velja $A^TA = I$, se pravi

$$\begin{bmatrix} a^2+c^2 & ab+cd \\ ab+cd & b^2+d^2 \end{bmatrix} = \begin{bmatrix} a & c \\ b & d \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Iz $a^2 + c^2 = 1$ sledi, da obstaja tak realen φ, da velja $a = \cos\varphi$ in $c = \sin\varphi$. Podobno iz $b^2 + d^2 = 1$ sledi, da obstaja tak realen ψ, da je $b = \cos\psi$ in $d = \sin\psi$. Iz $ab + cd = 0$ potem sledi, da je

$$\cos(\psi - \varphi) = \cos\varphi\cos\psi + \sin\varphi\sin\psi = ab + cd = 0.$$

Odtod sledi, da je $\psi - \varphi = \frac{\pi}{2} + k\pi$. Torej je

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} \cos\varphi & \cos(\frac{\pi}{2} + k\pi + \varphi) \\ \sin\varphi & \sin(\frac{\pi}{2} + k\pi + \varphi) \end{bmatrix} = \begin{bmatrix} \cos\varphi & -(-1)^k\sin\varphi \\ \sin\varphi & (-1)^k\cos\varphi \end{bmatrix}.$$

Za sode k dobimo prvo matriko, za lihe k pa drugo matriko. □

Unitarne matrike imajo zanimivo algebraično strukturo.

Trditev

Množica vseh unitarnih $n \times n$ matrik je grupa za matrično množenje.

Dokaz: Če sta A in B unitarni matriki, potem velja

$$(AB)^*(AB) = (B^*A^*)(AB) = B^*(A^*A)B = B^*IB = B^*B = I,$$

torej je tudi AB unitarna matrika. Če je A unitarna matrika, potem velja $A^{-1} = A^*$. Odtod sledi

$$(A^{-1})^*A^{-1} = (A^{-1})^*A^* = (AA^{-1})^* = I^* = I,$$

torej je tudi A^{-1} unitarna matrika. □

Za vsak obseg F označimo z $\mathrm{GL}(n, F)$ grupo vseh obrnljivih $n \times n$ matrik z elementi iz F. Tej grupi pravimo **splošna linearna grupa**. Grupo vseh $n \times n$ matrik nad F z determinanto 1 označimo z $\mathrm{SL}(n, F)$ in ji pravimo **specialna linearna grupa**.

Grupo vseh unitarnih $n \times n$ matrik označimo z $U(n)$ in ji pravimo **splošna unitarna grupa**. Grupo vseh unitarnih $n \times n$ matrik z determinanto 1 označimo z $\mathrm{SU}(n)$ in ji pravimo **specialna unitarna grupa**.

Grupo vseh ortogonalnih $n \times n$ matrik označimo z $O(n)$ in ji pravimo **splošna ortogonalna grupa**. Grupo vseh ortogonalnih $n \times n$ matrik z determinanto 1 označimo z $\mathrm{SO}(n)$ in ji pravimo **specialna ortogonalna grupa**. Grupi $\mathrm{O}(2)$ in $\mathrm{SO}(2)$ smo eksplicitno izračunali.

$$\begin{array}{ccc} \mathrm{GL}(n,\mathbb{C}) & \supset & \mathrm{U}(n) \\ \cup & & \cup \\ \mathrm{SL}(n,\mathbb{C}) & \supset & \mathrm{SU}(n) \end{array} \qquad \begin{array}{ccc} \mathrm{GL}(n,\mathbb{R}) & \supset & \mathrm{O}(n) \\ \cup & & \cup \\ \mathrm{SL}(n,\mathbb{R}) & \supset & \mathrm{SO}(n) \end{array}$$

Kaj vemo o lastnih vrednostih in lastnih vektorjih unitarnih matrik?

Trditev

Lastne vrednosti unitarne matrike imajo absolutno vrednost 1.

Prvi dokaz: Naj bo λ lastna vrednost za unitarno matriko A in naj bo v pripadajoči lastni vektor. Ker je $A^*A = I$, velja

$$\langle v, v\rangle = \langle v, I\, v\rangle = \langle v, A^*Av\rangle = \langle Av, Av\rangle = \langle \lambda v, \lambda v\rangle = \lambda\bar{\lambda}\langle v, v\rangle.$$

Ker je $v \neq 0$, lahko krajšamo $\langle v, v\rangle$ in dobimo $1 = \lambda\bar{\lambda}$.

Drugi dokaz: Ker je vsaka unitarna matrika normalna, ima faktorizacijo $A = PDP^{-1}$, kjer je D diagonalna in P unitarna matrika. Odtod sledi

$$D^*D = (P^*AP)^*(P^*AP) = P^*A^*PP^*AP = P^*A^*AP = P^*P = I.$$

Torej imajo diagonalni elementi matrike D absolutno vrednost 1. Toda diagonalni elementi matrike D so ravno lastne vrednosti matrike A.

Primer

Vemo, da sta matriki

$$A = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix} \quad \text{in} \quad B = \begin{bmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{bmatrix}$$

ortogonalni za vsak φ. Lastni vrednosti matrike A sta $\cos\varphi \pm i\sin\varphi$. Lastni vrednosti matrike B sta ± 1.

O lastnih vektorjih ne moremo povedati nič novega.

Trditev

Različnim lastnim vrednostim unitarne matrike pripadata ortogonalna lastna vektorja.

Prvi dokaz: Vsaka unitarna matrika je normalna in za normalne matrike smo to že dokazali.

Drugi dokaz: Naj bo $Au = \lambda u$ in $Av = \mu v$ in $\lambda \neq \mu$. Ker je A unitarna, je $\lambda\bar{\mu}\langle u, v\rangle = \langle \lambda u, \mu v\rangle = \langle Au, Av\rangle = \langle u, A^*Av\rangle = \langle u, v\rangle$. Poleg tega iz $\lambda \neq \mu$ in $\lambda\bar{\lambda} = \mu\bar{\mu} = 1$ sledi, da je $\lambda\bar{\mu} \neq 1$. Odtod sledi $\langle u, v\rangle = 0$.

Singularni razcep - 1. del

1. Pozitivno semidefinitne matrike

Te matrike bomo potrebovali pri konstrukciji singularnega razcepa.

Definicija pozitivno semidefinitne in pozitivno definitne matrike

Matrika A je **pozitivno semidefinitna**, če je sebiadjungirana in če velja $\langle Av, v\rangle \geq 0$ za vsak vektor v. Matrika A je **pozitivno definitna**, če je sebiadjungirana in če velja $\langle Av, v\rangle > 0$ za vsak neničeln vektor v.

Oznake: Če je matrika A pozitivno semidefinitna pišemo $A \geq 0$.
Če je matrika A pozitivno definitna, pišemo $A > 0$.

Opomba: V obeh definicijah uporabljamo standardni skalarni produkt, ki je definiran z $\langle u, v\rangle = v^* u$. Torej je $\langle Av, v\rangle = v^* Av$.

Opomba: Pozitivno definitne matrike so poseben primer pozitivno semidefinitnih matrik. Te pa so poseben primer sebiadjungiranih matrik.

Opomba: Lahko bi definirali tudi pojem pozitivno (semi)definitne linearne preslikave iz V v V, kjer je V vektorski prostor s skalarnim produktom.

Oglejmo si nekaj lastnosti pozitivno semidefinitnih matrik.

Izrek (Karakterizacije pozitivno semidefinitnih matrik)

Za kompleksno matriko A so ekvivalentne naslednje trditve:

1. A je pozitivno semidefinitna.
2. A je sebiadjungirana in vse njene lastne vrednosti so ≥ 0.
3. Obstaja taka unitarna matrika P in taka diagonalna matrika D, ki ima vse elemente realne in ≥ 0, da velja $A = PDP^{-1}$.
4. Obstaja taka sebiadjungirana matrika B, da velja $A = B^2$.
5. Obstaja taka matrika B (ne nujno kvadratna), da velja $A = B^*B$.

Dokaz. Če velja (1), potem za vsako lastno vrednost λ matrike A in za vsak pripadajoči lastni vektor v velja $\lambda\langle v, v\rangle = \langle \lambda v, v\rangle = \langle Av, v\rangle \geq 0$. Ker je $v \neq 0$, odtod sledi $\lambda \geq 0$. Torej velja (2).

Vemo, da se da vsako sebiadjungirano matriko A izraziti kot $A = PDP^{-1}$, kjer je P unitarna matrika in D realna diagonalna matrika, ki ima po diagonali lastne vrednosti matrike A. Torej iz (2) sledi (3).

Če velja (3), lahko diagonalne elemente matrike D korenimo. Torej obstaja taka realna diagonalna matrika E, da je $D = E^2$. Definirajmo matriko $B := PEP^{-1}$. Ker je matrika P unitarna, je matrika B sebiadjungirana in velja $B^2 = (PEP^{-1})^2 = PE^2P^{-1} = PDP^{-1} = A$. Torej velja (4).

Očitno iz (4) sledi (5). Če velja (5), potem je $A^* = (B^*B)^* = B^*B = A$. Poleg tega je $\langle B^*Bv, v\rangle = \langle Bv, Bv\rangle \geq 0$ za vsak v. Torej velja (1). □

Sedaj lahko definiramo osnovni pojem, ki nastopa v singularnem razcepu.

Definicija singularne vrednosti matrike

Naj bo B kompleksna $m \times n$ matrika in σ nenegativno realno število. Pravimo, da je σ **singularna vrednost** matrike B, če je σ^2 lastna vrednost $n \times n$ matrike B^*B.

Opomba. Po izreku so vse lastne vrednosti matrike $A = B^*B$ realne in ≥ 0. Njihovi kvadratni koreni so ravno singularne vrednosti matrike B. Če so stolpci B linearno neodvisni, potem 0 ni lastna vrednost matrike B^*B (sledi iz $\operatorname{Ker} B^*B = \operatorname{Ker} B = \{0\}$) torej so vse singularne vrednosti > 0.

Pozitivno definitna verzija izreka

Za kompleksno matriko A so ekvivalentne trditve:

1. A je pozitivno definitna.
2. A je sebiadjungirana in vse njene lastne vrednosti so > 0.
3. $A = PDP^{-1}$ za unitarno matriko P in diagonalno matriko $D > 0$.
4. $A = B^2$ za obrnljivo sebiadjungirano matriko B.
5. $A = B^*B$ za matriko B z linearno neodvisnimi stolpci.

Realna verzija izreka

Za vsako realno matriko A so ekvivalentne trditve:

1. A je pozitivno semidefinitna.
2. A je simetrična in vse njene lastne vrednosti so ≥ 0.
3. $A = PDP^{-1}$ za ortogonalno matriko P in diagonalno matriko $D \geq 0$.
4. $A = B^2$ za simetrično matriko B.
5. $A = B^TB$ za realno matriko B, ki ni nujno kvadratna.

Opomba: Pogosto se zastavi vprašanje ali že iz pogoja $\langle Av, v\rangle \geq 0$ za vsak v sledi, da je A pozitivno semidefinitna. V realnem primeru je odgovor ne, v kompleksnem primeru pa je odgovor da.

Prvi del dokaza: Če je A neničelna realna matrika, ki zadošča $A^T = -A$, potem A ni sebiadjungirana, vendar velja $\langle Av, v\rangle = 0$ za vsak realen v. Velja namreč $v^T Av = (v^T Av)^T = v^T A^T v = -v^T Av$, torej je $v^T Av = 0$.

Drugi del dokaza. Pokažimo, da je kompleksna $n \times n$ matrika A sebiadjungirana natanko tedaj, ko velja $\langle Av, v\rangle \in \mathbb{R}$ za vsak $v \in \mathbb{C}^n$. Za $B := A - A^*$ velja $\langle Bv, v\rangle = \langle Av, v\rangle - \langle A^*v, v\rangle = \langle Av, v\rangle - \overline{\langle Av, v\rangle}$. Torej je $\langle Av, v\rangle \in \mathbb{R}$ natanko tedaj, ko je $\langle Bv, v\rangle = 0$. Če je $\langle Bv, v\rangle = 0$ za vsak v, potem vstavimo $v = u + u'$ in dobimo $\langle Bu, u'\rangle + \langle Bu', u\rangle = 0$ za vsaka u, u'. (Ker je $\langle Bu, u\rangle = \langle Bu', u'\rangle = 0$.) Če namesto u vstavimo iu, dobimo $i\langle Bu, u'\rangle - i\langle Bu', u\rangle = 0$ za vsaka u, u'. Če to enačbo delimo z i in jo prištejemo k gornji enačbi, dobimo $\langle Bu, u'\rangle = 0$ za vsaka u, u'. Če vstavimo $u' = Bu$, dobimo $Bu = 0$ za vsak u. Torej je $B = 0$.

2. Klasifikacija skalarnih produktov

V tem razdelku bomo opisali vse skalarne produkte na $\mathbb{R}^n$ in $\mathbb{C}^n$.

Standardni skalarni produkt je definiran z $\langle u, v \rangle = v^* u$. Na dolgo

$$\langle \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}, \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_n \end{bmatrix} \rangle = \alpha_1 \bar{\beta}_1 + \ldots + \alpha_n \bar{\beta}_n = \begin{bmatrix} \bar{\beta}_1 & \ldots & \bar{\beta}_n \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}$$

Vsak drug skalarni produkt bomo označili z $[u, v]$.

Trditev

Če je A kompleksna pozitivno definitna $n \times n$ matrika, potem je z

$$[u, v] := \langle Au, v \rangle$$

definiran skalarni produkt na $\mathbb{C}^n$. Če je A realna pozitivno definitna $n \times n$ matrika, potem nam ista formula da skalarni produkt na $\mathbb{R}^n$.

Dokaz: Preverimo, da $[u, v]$ zadošča vsem trem lastnostim iz definicije skalarnega produkta.

Ker je A pozitivno definitna matrika, velja $\langle Au, u\rangle > 0$ za vsak neničeln vektor u. Torej velja $[u, u] > 0$ za vsak neničeln vektor u.

Ker je $A = A^*$, velja za vsak u in vsak v

$$[v, u] = \langle Av, u\rangle = \langle v, A^*u\rangle = \langle v, Au\rangle = \overline{\langle Au, v\rangle} = \overline{[u, v]}$$

To je ravno konjugirana simetričnost.

Preverimo še linearnost v prvem faktorju. Velja

$$[\alpha_1 u_1 + \alpha_2 u_2, v] = \langle A(\alpha_1 u_1 + \alpha_2 u_2), v\rangle = \langle \alpha_1 Au_1 + \alpha_2 Au_2, v\rangle =$$

$$= \alpha_1\langle Au_1, v\rangle + \alpha_2\langle Au_2, v\rangle = \alpha_1[u_1, v] + \alpha_2[u_2, v]$$

za vektorje u_1, u_2, v in skalarja α_1, α_2. □

Pokažimo še, da smo na ta način dobili vse skalarne produkte na $\mathbb{R}^n$ in $\mathbb{C}^n$.

Trditev

Za vsak skalarni produkt $[u, v]$ na $\mathbb{C}^n$ obstaja taka kompleksna pozitivno definitna $n \times n$ matrika A, da za vsak $u \in \mathbb{C}^n$ in vsak $v \in \mathbb{C}^n$ velja

$$[u, v] = \langle Au, v \rangle.$$

Za vsak skalarni produkt $[u, v]$ na $\mathbb{R}^n$ obstaja taka realna pozitivno definitna $n \times n$ matrika A, da za vsaka $u, v \in \mathbb{R}^n$ velja $[u, v] = \langle Au, v \rangle$.

Dokaz: Naj bo $e_1, \ldots, e_n$ standardna baza za $\mathbb{C}^n$. Pokažimo, da je

$$A = \begin{bmatrix} [e_1, e_1] & \ldots & [e_n, e_1] \\ \vdots & \ddots & \vdots \\ [e_1, e_n] & \ldots & [e_n, e_n] \end{bmatrix}$$

iskana matrika.

Iz konjugirane simetričnosti skalarnega produkta $[u, v]$ sledi, da je

$$A^* = \begin{bmatrix} \overline{[e_1, e_1]} & \dots & \overline{[e_1, e_n]} \\ \vdots & \ddots & \vdots \\ \overline{[e_n, e_1]} & \dots & \overline{[e_n, e_n]} \end{bmatrix} = \begin{bmatrix} [e_1, e_1] & \dots & [e_n, e_1] \\ \vdots & \ddots & \vdots \\ [e_1, e_n] & \dots & [e_n, e_n] \end{bmatrix} = A.$$

Po drugi strani za $u = \sum_{i=1}^{n} \alpha_i e_i$ in $v = \sum_{j=1}^{n} \beta_j e_j$ velja

$$[u, v] = \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \bar{\beta}_j [e_i, e_j] =$$

$$= \begin{bmatrix} \bar{\beta}_1 & \dots & \bar{\beta}_n \end{bmatrix} \begin{bmatrix} [e_1, e_1] & \dots & [e_n, e_1] \\ \vdots & \ddots & \vdots \\ [e_1, e_n] & \dots & [e_n, e_n] \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} = v^* A u = \langle Au, v \rangle$$

Odtod sledi, da je A pozitivno definitna, saj za vsak neničeln u velja

$$\langle Au, u \rangle = [u, u] > 0.$$

3. QR razcep in razcep Choleskega

Oglejmo si dva znana razcepa matrik.

Izrek (QR razcep)

Če je matrika A obrnljiva kompleksna matrika, potem obstaja unitarna matrika Q in taka zgornje trikotna matrika R, da je $A = QR$. Če je A realna, lahko Q in R izberemo tako, da sta realni.

Dokaz: Naj bodo $v_1, \ldots, v_n$ stolpci matrike A, ker je A obrnljiva, so ti vektorji baza za $\mathbb{C}^n$. Napravimo Gram-Schmidtovo ortogonalizacijo

$$\begin{aligned} w_1 &= v_1, & e_1 &= \frac{w_1}{\|w_1\|} \\ w_2 &= v_2 - \langle v_2, e_1 \rangle e_1, & e_2 &= \frac{w_2}{\|w_2\|} \\ &\vdots & &\vdots \\ w_n &= v_n - \langle v_n, e_1 \rangle e_1 - \ldots - \langle v_n, e_{n-1} \rangle, & e_n &= \frac{w_n}{\|w_n\|} \end{aligned}$$

Pomnožimo i-to enačbo skalarno z e_i in dobimo $\langle w_i, e_i \rangle = \langle v_i, e_i \rangle$.

Torej je $w_i = \frac{\langle w_i, w_i \rangle}{\|w_i\|^2} w_i = \langle w_i, e_i \rangle e_i = \langle v_i, e_i \rangle e_i$ za vsak i. Odtod sledi

$$\begin{aligned} v_1 &= \langle v_1, e_1 \rangle e_1, \\ v_2 &= \langle v_2, e_1 \rangle e_1 + \langle v_2, e_2 \rangle e_2, \\ &\vdots \\ v_n &= \langle v_n, e_1 \rangle e_1 + \langle v_n, e_2 \rangle v_2 + \ldots + \langle v_n, e_n \rangle e_n. \end{aligned}$$

kar lahko po matrično zapišemo z

$$\begin{bmatrix} v_1 & v_2 & \ldots & v_n \end{bmatrix} =$$

$$= \begin{bmatrix} e_1 & e_2 & \ldots & e_n \end{bmatrix} \begin{bmatrix} \langle v_1, e_1 \rangle & \langle v_2, e_1 \rangle & \ldots & \langle v_n, e_1 \rangle \\ 0 & \langle v_2, e_2 \rangle & \ldots & \langle v_n, e_2 \rangle \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \langle v_n, e_n \rangle \end{bmatrix}.$$

Torej je res $A = QR$, kjer je Q unitarna in R zgornje trikotna. □

Izrek (Razcep Choleskega)

Če je matrika A pozitivno definitna, potem obstaja taka zgornje trikotna matrika R, da je $A = R^*R$.

Dokaz: Če je matrika A pozitivno definitna, potem obstaja taka obrnljiva matrika B, da velja $A = B^*B$. Naj bo $B = QR$, kjer je Q unitarna in R zgornje trikotna. Potem je $A = (QR)^*(QR) = R^*Q^*QR = R^*R$. □

Opomba. Matriko R v razcepu Choleskega lahko izračunamo tudi z naslednjim algoritmom. Začnimo z

$$A_1 := A$$

Recimo, da je A_i pozitivno definitna matrika oblike

$$A_i = \begin{bmatrix} I_{i-1} & 0 & 0 \\ 0 & a_{i,i} & \mathbf{b}_i^* \\ 0 & \mathbf{b}_i & B_i \end{bmatrix}.$$

Potem je $a_{i,i} > 0$, torej lahko definiramo obrnljivo zgornje trikotno matriko

$$R_i = \begin{bmatrix} I_{i-1} & 0 & 0 \\ 0 & \sqrt{a_{i,i}} & \frac{1}{\sqrt{a_{i,i}}}\mathbf{b}_i^* \\ 0 & 0 & I_{n-i} \end{bmatrix}.$$

Potem velja

$$A_i = R_i^* A_{i+1} R_i,$$

kjer je

$$A_{i+1} = \begin{bmatrix} I_{i-1} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & B_i - \frac{1}{a_{i,i}}\mathbf{b}_i\mathbf{b}_i^* \end{bmatrix}.$$

Ta postopek ponovimo n-krat in upoštevamo, da je $A_{n+1} = I$. Dobimo

$$A = R^* R, \qquad \text{kjer je} \qquad R = R_n \dots R_1.$$

Opomba: Oglejmo si še, kako QR razcep izpeljemo iz razcepa Choleskega.

Drugi dokaz QR razcepa. Naj bo A obrnljiva matrika. Potem je matrika A^*A pozitivno definitna, torej ima razcep Choleskega

$$A^*A = R^*R,$$

kjer je R zgornje trikotna matrika. Pokažimo, da je matrika

$$Q := AR^{-1}$$

unitarna. To sledi iz $Q^*Q = (R^{-1})^*A^*AR^{-1} = (R^{-1})^*R^*RR^{-1} = (R^{-1})^*R^* = (RR^{-1})^* = I^* = I$. Iz $Q = AR^{-1}$ sledi, da je $A = QR$. □

4. Singularni razcep (=SVD)

V tem in naslednjih razdelkih bomo imeli opravka s pravokotnimi (= ne nujno kvadratnimi) diagonalnimi matrikami.

Definicija pravokotne diagonalne matrike

Pravokotna matrika $A = [a_{i,j}]$ je **diagonalna**, če je $a_{i,j} = 0$ za vsaka $i \neq j$.

Primera pravokotnih diagonalnih matrik

Pravokotni matriki

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \end{bmatrix} \quad \text{in} \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{bmatrix}$$

sta diagonalni.

Izrek o singularnem razcepu

Naj bo A kompleksna $m \times n$ matrika. Potem obstajajo taka unitarna $m \times m$ matrika Q_1, taka unitarna $n \times n$ matrika Q_2 in taka diagonalna $m \times n$ matrika D, kjer $d_{i,i} \geq 0$ za vsak $i = 1, \ldots, \min\{m, n\}$, da velja

$$A = Q_1 D Q_2^{-1}.$$

Če je A realna matrika, potem lahko matriki Q_1 in Q_2 izberemo tako, da sta obe realni (se pravi ortogonalni). Matrika D je že itak realna.

Opomba: Singularnemu razcepu se v angleščini reče "singular value decomposition", kar se okrajša v "SVD".

Opomba: Iz $A = Q_1 D Q_2^{-1} = Q_1 D Q_2^*$ sledi, da je

$$A^* A = Q_2 D^* Q_1^* Q_1 D Q_2^* = Q_2 D^* D Q_2^*, \tag{1}$$

$$A A^* = Q_1 D Q_2^* Q_2 D^* Q_1^* = Q_1 D D^* Q_1^*. \tag{2}$$

Iz (1) sledi, da so stolpci Q_2 ortonormirana baza iz lastnih vektorjev matrike A^*A. Podobno iz (2) sledi, da so stolpci Q_1 ortonormirana baza iz lastnih vektorjev matrike AA^*. Če je $m > n$, je

$$D^* D = \begin{bmatrix} d_{1,1}^2 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & d_{n,n}^2 \end{bmatrix}, DD^* = \begin{bmatrix} d_{1,1}^2 & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & & \vdots \\ 0 & \dots & d_{n,n}^2 & \dots & 0 \\ \vdots & & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & 0 \end{bmatrix}, \tag{3}$$

kjer so $d_{1,1}^2, \dots, d_{n.n}^2$ lastne vrednosti matrike A^*A, torej so $d_{1,1}, \dots, d_{n,n}$ singularne vrednosti matrike A. Podobno je v primerih $m = n$ in $m < n$.

Dokaz izreka: Naj bo A kompleksna $m \times n$ matrika. (Realen primer ima skoraj enak dokaz.) Konstrukcijo singularnega razcepa za A razdelimo v štiri dele: konstrukcije Q_2, D in Q_1 ter dokaz formule $AQ_2 = Q_1 D$.

Konstrukcija matrike Q_2. Najprej uredimo lastne vrednosti matrike A^*A po velikosti. Nato izračunamo pripadajoče lastne podprostore. Za vsakega od njih določimo ortonormirano bazo. Elemente v uniji teh baz označimo z $v_1, \ldots, v_n$, pripadajoče lastne vrednosti pa z $\lambda_1, \ldots, \lambda_n$. Velja torej

$$A^*Av_i = \lambda_i v_i$$

za vsak $i = 1, \ldots, n$ in $\lambda_1 \geq \ldots \geq \lambda_n$. Definirajmo

$$Q_2 = \begin{bmatrix} v_1 & \ldots & v_n \end{bmatrix}$$

Ker je $v_1, \ldots, v_n$ ortonormirana baza za $\mathbb{C}^n$, je matrika Q_2 unitarna.

Konstrukcija matrike D. Naj bo r rang matrike A. Vemo, da je r enak rangu matrike A^*A, zato velja

$$\lambda_1 \geq \ldots \geq \lambda_r > 0, \quad \lambda_{r+1} = \ldots = \lambda_n = 0$$

Naj bo $\sigma_i = \sqrt{\lambda_i}$ za $i = 1, \ldots, r$ in naj bo D diagonalna $m \times n$ matrika, ki ima po diagonali $\sigma_1, \ldots, \sigma_r$ in morda ničle. Eksplicitno

$$D = \begin{bmatrix} \sigma_1 & \ldots & 0 & \ldots & 0 \\ \vdots & \ddots & \vdots & & \vdots \\ 0 & \ldots & \sigma_r & \ldots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \ldots & 0 & \ldots & 0 \end{bmatrix}$$

Konstrukcija matrike Q_1. Za $i = 1, \ldots, r$ definirajmo

$$u_i = \frac{1}{\sigma_i} A v_i.$$

Pokažimo, da ti vektorji tvorijo ortonormirano množico v $\mathbb{C}^m$. Velja

$$\langle u_i, u_j \rangle = \frac{1}{\sigma_i \sigma_j} \langle A v_i, A v_j \rangle = \frac{1}{\sigma_i \sigma_j} \langle A^* A v_i, v_j \rangle = \frac{1}{\sigma_i \sigma_j} \langle \lambda_i v_i, v_j \rangle =$$

$$= \frac{\lambda_i}{\sigma_i \sigma_j} \langle v_i, v_j \rangle = \frac{\sigma_i^2}{\sigma_i \sigma_j} \langle v_i, v_j \rangle = \frac{\sigma_i}{\sigma_j} \langle v_i, v_j \rangle = \begin{cases} 1 & \text{če } j = i \\ 0 & \text{če } j \neq i \end{cases}$$

za vsaka $i, j = 1, \ldots, r$. Naj bo $u_{r+1}, \ldots, u_m$ dopolnitev ortonormirane množice $u_1, \ldots, u_r$ do ortonormirane baze za $\mathbb{C}^m$ in naj bo

$$Q_1 = \begin{bmatrix} u_1 & \ldots & u_m \end{bmatrix}.$$

Za konec dokažimo še, da je res $AQ_2 = Q_1 D$. Iz

$$Q_2 = \begin{bmatrix} v_1 & \dots & v_r & v_{r+1} & \dots & v_n \end{bmatrix}$$

sledi, da je

$$\begin{aligned} AQ_2 &= \begin{bmatrix} Av_1 & \dots & Av_r & Av_{r+1} & \dots & Av_n \end{bmatrix} \\ &= \begin{bmatrix} Av_1 & \dots & Av_r & 0 & \dots & 0 \end{bmatrix} \end{aligned}$$

Pri drugem enačaju smo upoštevali, da za $i = r+1, \dots, n$ velja $A^*Av_i = \lambda_i v_i = 0$, odkoder zaradi $\operatorname{Ker} A^*A = \operatorname{Ker} A$ sledi $Av_i = 0$. Iz

$$Q_1 = \begin{bmatrix} u_1 & \dots & u_r & u_{r+1} & \dots & u_m \end{bmatrix}$$

in definicije D sledi, da je

$$\begin{aligned} Q_1 D &= \begin{bmatrix} \sigma_1 u_1 & \dots & \sigma_r u_r & 0 & \dots & 0 \end{bmatrix} \\ &= \begin{bmatrix} Av_1 & \dots & Av_r & 0 & \dots & 0 \end{bmatrix}. \end{aligned}$$

Pri drugem enačaju smo upoštevali, da je $u_i = \frac{1}{\sigma_i} Av_i$ za $i = 1, \dots, r$.

Oglejmo si še geometrijski pomen singularnega razcepa.

Posledica

Naj bo A realna $m \times n$ matrika in naj bo K enotska krogla v $\mathbb{R}^n$. Potem je množica $\{Av \mid v \in K\}$ elipsoid v podprostoru $\operatorname{Im} A \subseteq \mathbb{R}^m$. Polosi tega elipsoida so ravno neničelne singularne vrednosti matrike A.

Dokaz: Naj bo $A = Q_1 D Q_2^{-1}$ singularni razcep matrike A, ki smo ga konstruirali v dokazu prejšnjega izreka. Vzemimo poljuben vektor v iz enotske krogle v $\mathbb{R}^n$ in ga razvijmo po stolpcih $v_1, \ldots, v_n$ matrike Q_2. Velja $v = x_1 v_1 + \ldots + x_n v_n$, kjer $x_1^2 + \ldots + x_n^2 \leq 1$.

Naj bodo $\sigma_1, \ldots, \sigma_r$ neničelni elementi matrike D in naj bodo $u_1 = \frac{1}{\sigma_1} A v_1, \ldots, u_r = \frac{1}{\sigma_r} A v_r$ začetni stolpci matrike Q_1. Potem velja $Av = x_1 Av_1 + \ldots + x_n Av_n = x_1 \sigma_1 u_1 + \ldots + x_r \sigma_r u_r = y_1 u_1 + \ldots + y_r u_r$, kjer je $\frac{y_1^2}{\sigma_1^2} + \ldots + \frac{y_r^2}{\sigma_r^2} = x_1^2 + \ldots + x_r^2 \leq x_1^2 + \ldots + x_r^2 + \ldots + x_n^2 \leq 1$. Torej Av res leži v r-razsežnem elipsoidu v $\operatorname{Im} A$ s polosmi $\sigma_1, \ldots, \sigma_r$.

Velja tudi obratno. Vsak element $y_1 u_1 + \ldots + y_r u_r$ tega elipsoida je slika nekega elementa iz enotske krogle, namreč slika od $\frac{y_1}{\sigma_1} v_1 + \ldots + \frac{y_r}{\sigma_r} v_r$.

Singularni razcep - 2. del

5. Primeri singularnega razcepa

Primer

Izračunajmo singularni razcep matrike

$$A = \begin{bmatrix} 1 & 1 & -1 & -1 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix}$$

Rešitev: Najprej izračunamo lastne vrednosti in lastne vektorje matrike

$$A^T A = \begin{bmatrix} 2 & 1 & -2 & -1 \\ 1 & 2 & -1 & -2 \\ -2 & -1 & 2 & 1 \\ -1 & -2 & 1 & 2 \end{bmatrix}$$

Lastni polinom matrike A je $x^2(x-2)(x-6)$. Neničelni singularni vrednosti matrike A sta $\sigma_1 = \sqrt{6}$ in $\sigma_2 = \sqrt{2}$.

Lastni vektorji, ki pripadajo lastnim vrednostim $6, 2$ in 0 so

$$w_1 = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}, w_2 = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}, w_3 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, w_4 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

Vektorja w_3 in w_4 tvorita bazo za lastni podprostor lastne vrednosti 0. Na srečo sta že ortogonalna, zato ne rabimo Gram-Schmidtove ortogonalizacije. Odtod dobimo ortonormirano bazo za $\mathbb{R}^4$

$$v_1 = \frac{w_1}{\|w_1\|}, \quad v_2 = \frac{w_2}{\|w_2\|}, \quad v_3 = \frac{w_3}{\|w_3\|}, \quad v_4 = \frac{w_4}{\|w_4\|}.$$

Tvorimo matriki $Q = \begin{bmatrix} v_1 & v_2 & v_3 & v_4 \end{bmatrix}$ in D

$$Q_2 = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{2} & -\frac{1}{2} & 0 & \frac{1}{\sqrt{2}} \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{\sqrt{2}} & 0 \end{bmatrix}, \quad D = \begin{bmatrix} \sqrt{6} & 0 & 0 & 0 \\ 0 & \sqrt{2} & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Ostane nam še konstrukcija matrike Q_1. Najprej izračunajmo vektorja

$$u_1 = \frac{1}{\sigma_1} A v_1 = \frac{1}{\sqrt{6}} \begin{bmatrix} 2 \\ -1 \\ -1 \end{bmatrix}, \quad u_2 = \frac{1}{\sigma_2} A v_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$$

Poiščimo sedaj tak vektor u_3, ki bo u_1 in u_2 dopolnil do ortonormirane baze za $\mathbb{R}^3$. Običajno se to naredi z Gram-Schmidtom, ampak poglejmo si še eno metodo. Rešimo enačbo $AA^T u_3 = 0$ in rešitev normirajmo. Dobimo $u_3 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}^T$. Ker so u_1, u_2 in u_3 lastni vektorji matrike AA^T, ki pripadajo različnim lastnim vrednostim 6, 2 in 0, so paroma ortogonalni. Sedaj lahko tvorimo

$$Q_1 = \begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix} = \begin{bmatrix} \frac{2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{bmatrix}$$

Konstruirali smo razcep $A = Q_1 D Q_2^T$, kjer sta Q_1 in Q_2 ortogonalni matriki, D pa diagonalna matrika.

Primer.

Pojasni kako preprosto izračunamo singularni razcep normalne matrike.

Rešitev. Naj bo A $n \times n$ matrika, ki zadošča $A^*A = AA^*$.

Najprej izračunamo singularne vrednosti matrike A. Vemo, da ima A n linearno neodvisnih lastnih vektorjev $v_1, \ldots, v_n$. Pripadajoče lastne vrednosti označimo z $\lambda_1, \ldots, \lambda_n$. Ker za normalne matrike iz $Av_i = \lambda_i v_i$ sledi $A^*v_i = \bar{\lambda}_i v_i$, imamo $A^*Av_i = \bar{\lambda}_i \lambda_i v_i$ za vsak $i = 1, \ldots, n$. Torej so $\sigma_i := \sqrt{\bar{\lambda}_i \lambda_i} = |\lambda_i|$ vse singularne vrednosti A.

Permutirajmo vektorje $v_1, \ldots, v_n$ tako, da velja

$$|\lambda_1| \geq \ldots \geq |\lambda_r| > 0, \quad \lambda_{r+1} = \ldots = \lambda_n = 0$$

in označimo

$$Q_2 = \begin{bmatrix} v_1 & \ldots & v_n \end{bmatrix}$$

Upoštevamo, da za $i = 1, \ldots, r$ velja $\sigma_i = |\lambda_i|$ in dobimo

$$u_i = \frac{1}{\sigma_i} A v_i = \frac{\lambda_i}{|\lambda_i|} v_i$$

Za dopolnitev do baze $\mathbb{C}^n$ lahko vzamemo kar $v_{r+1}, \ldots, v_n$. Torej je

$$Q_1 = \begin{bmatrix} \frac{\lambda_1}{|\lambda_1|} v_1 & \cdots & \frac{\lambda_r}{|\lambda_r|} v_r & v_{r+1} & \cdots & v_n \end{bmatrix}$$

Naj bo $|D|$ matrika, ki ima po diagonali $|\lambda_1|, \ldots, |\lambda_n|$. Singularni razcep je

$$A = Q_1 |D| Q_2^*.$$

Dodatek: naredimo primerjavo singularnega razcepa z lastnim razcepom

$$A = Q_2 D Q_2^*$$

Označimo s $\operatorname{sgn}(D)$ matriko, ki ima po diagonali $\frac{\lambda_1}{|\lambda_1|}, \ldots, \frac{\lambda_r}{|\lambda_r|}, 1, \ldots, 1$.
Opazimo, da je $D = \operatorname{sgn}(D)|D|$ in $Q_1 = Q_2 \operatorname{sgn}(D)$, torej je

$$A = Q_2 D Q_2^* = Q_2 \operatorname{sgn}(D) \operatorname{sgn}(D)^* D Q_2^* = Q_1 |D| Q_2^*$$

Za pozitivno semidefinitne matrike se oba razcepa ujemata.

Izrek o singularnem razcepu lahko formuliramo tudi za linearne preslikave.

Posledica

Za vsako linearno preslikavo $L\colon U \to V$ med končnorazsežnima vektorskima prostoroma s skalarnim produktom obstaja taka ortonormirana baza $\mathcal{B}$ za U in taka ortonormirana baza $\mathcal{C}$ za V, da je matrika $[L]_{\mathcal{C}\leftarrow\mathcal{B}}$ diagonalna z elementi ≥ 0.

Dokaz: Najprej izberemo poljubni ortonormirani bazi $\mathcal{B}' = \{b'_1, \ldots, b'_m\}$ za U in $\mathcal{C}'$ za V. Potem naredimo singularni razcep $[L]_{\mathcal{C}'\leftarrow\mathcal{B}'} = Q_1 D Q_2^{-1}$. Če je $Q_2 = [r_{i,j}]$, potem z $b_i = \sum_{k=1}^{n} r_{k,i} b'_i$ definiramo novo bazo $\mathcal{B}$ za U. Ker so stolpci matrike Q_2 ortonormirani, je tudi $\mathcal{B}$ ortonormirana baza. Podobno definiramo tudi novo bazo $\mathcal{C}$ za V in dokažemo ortonormiranost. Velja $[L]_{\mathcal{C}\leftarrow\mathcal{B}} = P_{\mathcal{C}\leftarrow\mathcal{C}'}([L]_{\mathcal{C}'\leftarrow\mathcal{B}'})P_{\mathcal{B}'\leftarrow\mathcal{B}} = Q_1^{-1}(Q_1 D Q_2^{-1})Q_2 = D$. □

Opomba: Če ne zahtevamo ortonormiranosti baz $\mathcal{B}$ in $\mathcal{C}$, potem ju lahko izberemo tako, da ima $[L]_{\mathcal{C}\leftarrow\mathcal{B}}$ po diagonali enke in ničle, drugod pa ničle.

Še ena reformulacija singularnega razcepa.

Posledica

Za vsako matriko $A \in \mathbb{C}^{m\times n}$ ranga r obstaja taka ortonormirana množica $u_1, \ldots, u_r$ v $\mathbb{C}^m$ in taka ortonormirana množica $v_1, \ldots, v_r$ v $\mathbb{C}^n$, da velja

$$A = \sigma_1 u_1 v_1^* + \ldots + \sigma_r u_r v_r^*,$$

kjer so $\sigma_1, \ldots, \sigma_r$ neničelne singularne vrednosti matrike A.

Dokaz. Sledi iz singularnega razcepa $A = Q_1 D Q_2^*$ in formule

$$\begin{bmatrix} u_1 & \ldots & u_r & \ldots & u_m \end{bmatrix} \begin{bmatrix} \sigma_1 & \ldots & 0 & \ldots & 0 \\ \vdots & \ddots & \vdots & & \vdots \\ 0 & \ldots & \sigma_r & \ldots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \ldots & 0 & \ldots & 0 \end{bmatrix} \begin{bmatrix} v_1^* \\ \vdots \\ v_r^* \\ \vdots \\ v_n^* \end{bmatrix} = \sum_{i=1}^{r} \sigma_i u_i v_i^*$$

6. Aproksimacija z matrikami nizkega ranga

Naj bo $\mathbb{C}^{m\times n}$ vektorski prostor vseh kompleksnih $m \times n$ matrik. Na tem vektorskem prostoru lahko definiramo skalarni produkt z

$$\langle B, C\rangle := \operatorname{Sled}(BC^*) = \sum_{i=1}^{m}\sum_{j=1}^{n} b_{i,j}\overline{c_{i,j}}.$$

Pripadajoča norma je $\|B\| := \sqrt{\langle B, B\rangle} = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}|b_{i,j}|^2}$. Tej normi pravimo **Frobeniusova norma**. Pogosto se označuje z $\|B\|_F$.

Trditev

Če je Q_1 unitarna $m \times m$ matrika in Q_2 unitarna $n \times n$ matrika, potem za vsako $m \times n$ matriko A velja $\|Q_1^* A Q_2\| = \|A\|$.

Dokaz: Upoštevamo definicijo norme in lastnost $\operatorname{Sled}(XY) = \operatorname{Sled}(YX)$. Dobimo $\|Q_1^* A Q_2\|^2 = \operatorname{Sled}(Q_1^* A Q_2 Q_2^* A^* Q_1) = \operatorname{Sled}(Q_1^* A A^* Q_1) =$
$= \operatorname{Sled}(AA^* Q_1 Q_1^*) = \operatorname{Sled}(AA^*) = \|A\|^2$. □

Problem: Naj bo $A \in \mathbb{C}^{m\times n}$ dana matrika in $k \in \mathbb{N}$ dano število. Iščemo tako matriko A_k ranga $\leq k$, da za vsako drugo matriko B ranga $\leq k$ velja

$$\|A - A_k\| \leq \|A - B\|.$$

To pomeni, da se med vsemi matrikami v $\mathbb{C}^{m\times n}$, ki so ranga $\leq k$, matrika A_k najbolje prilega matriki A. Odgovor nam daje naslednji izrek.

Izrek (Eckart & Young 1936)

Naj bo $A \in \mathbb{C}^{m\times n}$ matrika ranga r in naj bo $k < r$. Naj bo $A = Q_1 D Q_2^*$ singularni razcep matrike A, kjer $\sigma_1 \geq \ldots \geq \sigma_r > 0$. Definirajmo matriko D_k tako, da v D zamenjamo $\sigma_{k+1}, \ldots, \sigma_r$ z nič. Med vsemi matrikami v $\mathbb{C}^{m\times n}$, ki so ranga $\leq k$, se $A_k := Q_1 D_k Q_2^*$ najbolje prilega A.

Dokaz: Najprej izračunajmo razdaljo med A in A_k:

$$\|A - A_k\|^2 = \|Q_1(D - D_k)Q_2^*\|^2 = \|D - D_k\|^2 = \sigma_{k+1}^2 + \ldots + \sigma_r^2.$$

Naj bo B $m \times n$ matrika ranga $\leq k$, ki se najbolje prilega matriki A. Radi bi dokazali, da je $\|A - B\|^2 = \sigma_{k+1}^2 + \ldots + \sigma_r^2$.

Začnimo tokrat s singularnim razcepom matrike B:

$$B = U \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} V^*,$$

kjer je D diagonalna $k \times k$ matrika in U, V unitarni matriki. Pišimo

$$A = U(U^*AV)V^* = U \begin{bmatrix} E & F \\ G & H \end{bmatrix} V^*,$$

kjer je E $k \times k$ matrika. Dokažimo, da je $F = G = 0$ in $E = D$. Naj bo

$$C = U \begin{bmatrix} E & F \\ 0 & 0 \end{bmatrix} V^* \quad \text{in} \quad C' = U \begin{bmatrix} E & 0 \\ G & 0 \end{bmatrix} V^*.$$

Ker sta $\begin{bmatrix} E & F \\ 0 & 0 \end{bmatrix}$ in $\begin{bmatrix} E & 0 \\ G & 0 \end{bmatrix}$ ranga $\leq k$, sta tudi C in C' ranga $\leq k$.

Iz definicije norme in lastnosti $\|UXV^*\| = \|X\|$ sledi

$$\|A-B\|^2 = \| \begin{bmatrix} E-D & F \\ G & H \end{bmatrix} \|^2 = \|E-D\|^2 + \|F\|^2 + \|G\|^2 + \|H\|^2$$

$$\|A-C\|^2 = \| \begin{bmatrix} 0 & 0 \\ G & H \end{bmatrix} \|^2 = \|G\|^2 + \|H\|^2$$

$$\|A-C'\|^2 = \| \begin{bmatrix} 0 & F \\ 0 & H \end{bmatrix} \|^2 = \|F\|^2 + \|H\|^2$$

Odtod sledi, da je

$$\begin{aligned} \|A-B\|^2 &= \|A-C\|^2 + \|E-D\|^2 + \|F\|^2 \\ &= \|A-C'\|^2 + \|E-D\|^2 + \|G\|^2 \end{aligned}$$

Ker se matrika B med vsemi matrikami ranga $\leq k$ najbolje prilega k, velja

$$\|A-B\|^2 \leq \|A-C\|^2 \quad \text{in} \quad \|A-B\|^2 \leq \|A-C'\|^2.$$

Odtod sledi, da je

$$\|E - D\|^2 + \|F\|^2 = 0 \quad \text{in} \quad \|E - D\|^2 + \|G\|^2 = 0.$$

Torej je $\|E - D\| = \|F\| = \|G\| = 0$; se pravi $E = D$ in $F = G = 0$.

Naj bo $H = WJZ^*$ singularni razcep matrike H. Odtod sledi, da je

$$A = U \begin{bmatrix} D & 0 \\ 0 & H \end{bmatrix} V^* = (U \begin{bmatrix} I & 0 \\ 0 & W \end{bmatrix}) \begin{bmatrix} D & 0 \\ 0 & J \end{bmatrix} (V \begin{bmatrix} I & 0 \\ 0 & Z \end{bmatrix})^* \quad (1)$$

Singularni razcep matrike B popravimo v

$$B = U \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} V^* = (U \begin{bmatrix} I & 0 \\ 0 & W \end{bmatrix}) \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} (V \begin{bmatrix} I & 0 \\ 0 & Z \end{bmatrix})^*. \quad (2)$$

Ker se B najbolje prilega A, so vsi elementi J pod vsemi elementi D.
Torej je (1) singularni razcep matrike A. Iz (1) in (2) sedaj sledi

$$\|A - B\|^2 = \sigma_{k+1}^2 + \ldots + \sigma_r^2. \quad \square$$

To se uporablja pri zgoščevanju slik. Recimo, da imamo sliko velikosti $m \times n$. Za vsako točko si moramo zapomniti njeno barvo, se pravi neko realno število. (Odtenek sive ali kombinacija odtenkov modre, zelene in rdeče.) Dobimo torej realno $m \times n$ matriko A.

Naredimo singularni razcep in si zapomnimo prvih k singularnih vrednosti, ter vektorje $u_1, \ldots, u_k$ in $v_1, \ldots, v_k$ za katere velja

$$A_k = \sigma_1 u_1 v_1^* + \ldots + \sigma_k u_k v_k^*$$

Vektorji u_i so dolžine m, vektorji v_i pa dolžine n, torej je A_k podana z $k(m+n+1)$ podatki. Razmerje z začetno količino podatkov je

$$\frac{k(m+n+1)}{mn} = k(\frac{1}{m} + \frac{1}{n} + \frac{1}{mn}),$$

kjer lahko zadnji člen zanemarimo. Če je recimo $m = n = 1000$ in $k = 50$, si moramo zapomniti samo 10% od začetne količine podatkov.

7. Psevdoinverz

Kot ponavadi, se omejimo na realna in kompleksna števila.

Za diagonalno $m \times n$ matriko D z neničelnimi elementi $\sigma_1, \ldots, \sigma_r$ definiramo njen psevdoinverz D^+ kot diagonalno $n \times m$ matriko z neničelnimi elementi $\frac{1}{\sigma_1}, \ldots, \frac{1}{\sigma_r}$.

Primer

$$\begin{bmatrix} \sigma_1 & 0 & 0 & 0 \\ 0 & \sigma_2 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}^+ = \begin{bmatrix} \frac{1}{\sigma_1} & 0 & 0 \\ 0 & \frac{1}{\sigma_2} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Opomba: Očitno velja $(D^+)^+ = D$ za vsako diagonalno matriko D.

Opomba: Če je D obrnljiva, potem velja $D^+ = D^{-1}$.

Za splošno matriko A najprej izračunamo singularni razcep

$$A = Q_1 D Q_2^*$$

in potem definiramo

$$A^+ = Q_2 D^+ Q_1^*.$$

Opomba: Za vsako matriko A je $(A^+)^+ = Q_1 (D^+)^+ Q_2^* = Q_1 D Q_2^* = A$.

Opomba: Če je matrika A obrnljiva, potem velja $A^+ = A^{-1}$.

Primer

Izračunajmo psevdoinverz matrike

$$A = \begin{bmatrix} 1 & 1 & -1 & -1 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix}$$

Rešitev: Najprej izračunamo singularni razcep matrike A.

$$A = \begin{bmatrix} \frac{2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{bmatrix} \begin{bmatrix} \sqrt{6} & 0 & 0 & 0 \\ 0 & \sqrt{2} & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{2} & -\frac{1}{2} & 0 & \frac{1}{\sqrt{2}} \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{\sqrt{2}} & 0 \end{bmatrix}^T$$

Potem je

$$A^+ = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{2} & -\frac{1}{2} & 0 & \frac{1}{\sqrt{2}} \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{\sqrt{2}} & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{6}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{bmatrix}^T$$

$$= \frac{1}{6} \begin{bmatrix} 1 & -2 & 1 \\ 1 & 1 & -2 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}$$

Oglejmo si, kako izračunamo psevdoinverz brez singularnega razcepa. Za pozitivno semidefinitno matriko B vemo, da se njen singularni razcep ujema z njenim lastnim razcepom $B = QDQ^*$, kjer je Q unitarna in D diagonalna z nenegativnimi elementi. Torej je $B^+ = QD^+Q^*$. Splošni primer lahko prevedemo na pozitivno semidefinitni primer takole:

Trditev

Za vsako kompleksno matriko A velja

$$A^+ = (A^*A)^+A^* = A^*(AA^*)^+$$

Dokaz: Če je $A = Q_1DQ_2^*$ singularni razcep, potem sta tudi $A^*A = Q_2D^*DQ_2^*$ in $AA^* = Q_1DD^*Q_1^*$ singularna razcepa, torej je $A^* = Q_2D^*Q_1^*$, $(A^*A)^+ = Q_2(D^*D)^+Q_2^*$ in $(AA^*)^+ = Q_1(DD^*)^+Q_1^*$. Trditev sedaj sledi iz $D^+ = (D^*D)^+D^* = D^*(DD^*)^+$. □

Posledica

- Če so stolpci A linearno neodvisni, je $A^+ = (A^*A)^{-1}A^*$.
- Če so vrstice A linearno neodvisne je $A^+ = A^*(AA^*)^{-1}$.

Ker singularni razcep ni enolično določen z matriko A, se zastavi vprašanje ali morda tudi psevdoinverz A^+ ni enolično določen z matriko A.

Izrek o enoličnosti psevdoinverza

Matrika A^+ je enolično določena z matriko A.

Dokaz: Dokažimo najprej nekaj identitet, ki jim zadošča matrika A^+

$$AA^+A = Q_1DQ_2^*Q_2D^+Q_1^*Q_1DQ_2^* = Q_1DD^+DQ_2^* = Q_1DQ_2^* = A,$$
$$A^+AA^+ = Q_2D^+Q_1^*Q_1DQ_2^*Q_2D^+Q_1^* = Q_2D^+DD^+Q_1^* = Q_2D^+Q_1^* = A^+,$$
$$(AA^+)^* = (Q_1DQ_2^*Q_2D^+Q_1^*)^* = Q_1(DD^+)^*Q_1^* = Q_1DD^+Q_1^* = AA^+,$$
$$(A^+A)^* = (Q_2D^+Q_1^*Q_1DQ_2^*)^* = Q_2(D^+D)^*Q_2^* = Q_2D^+DQ_2^* = A^+A$$

Če bi z dvema singularnima razcepoma dobili različna psevdoinverza B in C, potem bi oba zadoščala tem štirim identitetam. Odtod sledi

$$AB \overset{(1)}{=} ACAB \overset{(3)}{=} (AC)^*(AB)^* = C^*(ABA)^* \overset{(1)}{=} C^*A^* = (AC)^* \overset{(3)}{=} AC.$$

Podobno dokažemo, da je tudi $BA = CA$. Oboje uporabimo v računu

$$B \overset{(2)}{=} BAB = BAC = CAC \overset{(2)}{=} C. \quad \square$$

8. Najkrajša posplošena rešitev sistema linearnih enačb

Psevdoinverz se rabi pri iskanju rešitev sistemov linearnih enačb.

Naj bo A pravokotna matrika. Spomnimo se definicije posplošene rešitve sistema $Ax = b$. To je tak vektor x_0, da je Ax_0 najbližje vektorju b.

Posplošena rešitev ni enolično določena. Med vsemi posplošenimi rešitvami sistema $Ax = b$ poiščimo najkrajšo. Izkaže se, da je to ravno A^+b.

Izrek

Naj bo A kompleksna $m \times n$ matrika in b vektor iz $\mathbb{C}^m$. Najkrajša posplošena rešitev sistema $Ax = b$ je $x_0 = A^+b$.

Dokaz: Naj bo $A = Q_1 D Q_2^*$ singularni razcep. Označimo $b' = Q_1^* b$. Primerjajmo posplošene rešitve sistema $Ax = b$ s posplošenimi rešitvami sistema $Dx' = b'$. Za vsak vektor $x \in \mathbb{C}^m$ velja

$$\|Ax - b\| = \|Q_1 D Q_2^* x - b\| = \|Q_1(DQ_2^* x - b')\| = \|DQ_2^* x - b'\|$$

in $\|Q_2^* x\| = \|x\|$. Odtod očitno sledi dvoje:

- Če je x_0 posplošena rešitev sistema $Ax = b$, potem je $x_0' := Q_2^* x_0$ posplošena rešitev sistema $Dx' = b'$. Velja tudi $\|x_0'\| = \|x_0\|$.
- Če je x_0' posplošena rešitev sistema $Dx' = b'$, potem je $x_0 := Q_2 x_0'$ posplošena rešitev sistema $Ax = b$. Velja tudi $\|x_0\| = \|x_0'\|$.

Torej je x_0 najkrajša posplošena rešitev sistema $Ax = b$ natanko tedaj, ko je $x_0' = Q_2^* x_0$ najkrajša posplošena rešitev sistema $Dx' = b'$.

Poiščimo sedaj najkrajšo posplošeno rešitev sistema $Dx' = b'$, se pravi najkrajšega izmed vseh vektorjev x', ki minimizirajo izraz $\|Dx' - b'\|$.
Če vstavimo $x' = (x'_1, \ldots, x'_n)$ in $b' = (b'_1, \ldots, b'_m)$ dobimo

$$\|Dx' - b'\|^2 = (\sigma_1 x'_1 - b'_1)^2 + \ldots + (\sigma_r x'_r - b'_r)^2 + (b'_{r+1})^2 + \ldots + (b'_m)^2.$$

Minimum tega izraza dobimo pri

$$x' = (\frac{b'_1}{\sigma_1}, \ldots, \frac{b'_r}{\sigma_r}, x'_{r+1}, \ldots, x'_n),$$

kjer so $x_{r+1}, \ldots, x_n$ poljubni. Najkrajša od teh posplošenih rešitev je

$$x'_0 = (\frac{b'_1}{\sigma_1}, \ldots, \frac{b'_r}{\sigma_r}, 0, \ldots, 0).$$

Opazimo, da je izraz na desni enak $D^+ b'$. Odtod sledi, da je

$$x_0 := Q_2 x'_0 = Q_2 D^+ b' = Q_2 D^+ Q_1^* b = A^+ b$$

najkrajša posplošena rešitev sistema $Ax = b$.

Primer

Sistem linearnih enačb

$$\begin{aligned} x_1 + x_2 - x_3 - x_4 &= 1 \\ -x_1 + x_3 &= 0 \\ -x_2 + x_4 &= 1 \end{aligned}$$

ni rešljiv običajnem smislu. Poiščimo najkrajšo posplošeno rešitev.

Rešitev: Sistem najprej zapišemo v matrični obliki

$$\begin{bmatrix} 1 & 1 & -1 & -1 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

Najkrajša posplošena rešitev je

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 & 1 & -1 & -1 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix}^{+} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} =$$

$$= \frac{1}{6} \begin{bmatrix} 1 & -2 & 1 \\ 1 & 1 & -2 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} \\ -\frac{1}{6} \\ -\frac{1}{3} \\ \frac{1}{6} \end{bmatrix}$$

Singularni razcep - 3. del

9. Kovariančna in korelacijska matrika

Najprej osvežimo osnovne pojme iz opisne statistike. Ponavljamo poskus in beležimo podatke na merilnih napravah. Dobimo **tabelo podatkov**.

Če imamo eno merilno napravo, ima tabela podatkov en stolpec podatkov.

ponovitev	podatek
1.	x_1
2.	x_2
$\vdots$	$\vdots$
n-ta	x_n

Za ta stolpec izračunamo povprečje

$$\bar{x} = \frac{x_1 + \ldots + x_n}{n}.$$

Če nas zanima razpršenost podatkov okrog povprečja, potem najprej izračunamo odmike podatkov od povprečja, se pravi števila

$$x_i' = x_i - \bar{x}.$$

Dobimo **centralizirano tabelo podatkov**

ponovitev	centralizirani podatek
1.	$x_1' = x_1 - \bar{x}$
2.	$x_2' = x_2 - \bar{x}$
$\vdots$	$\vdots$
n-ta	$x_n' = x_n - \bar{x}$

Povprečen odmik od povprečja ni dobra mera za razpršenost, saj velja

$$\frac{x_1' + \ldots + x_n'}{n} = \frac{x_1 + \ldots + x_n - n\bar{x}}{n} = \bar{x} - \bar{x} = 0.$$

Boljša mera za razpršenost podatkov je kvadratična sredina odmikov

$$s = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i')^2} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}.$$

Temu številu pravimo **standardna deviacija** stolpca podatkov.

Če vse elemente v centralizirani tabeli podatkov delimo z s dobimo **standardizirano tabelo podatkov**

ponovitev	standardizirani podatek
1.	$x_1'' = x_1'/s = (x_1 - \bar{x})/s$
2.	$x_2'' = x_2'/s = (x_2 - \bar{x})/s$
$\vdots$	$\vdots$
n-ta	$x_n'' = x_n'/s = (x_n - \bar{x})/s$

Povprečje standardizirane tabele podatkov je 0, njena standardna deviacija pa je 1, saj je $\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i'')^2} = \sqrt{\frac{1}{ns^2}\sum_{i=1}^{n}(x_i')^2} = \sqrt{\frac{s^2}{s^2}} = 1.$

Če imamo dve merilni napravi, ima tabela podatkov dva stolpca podatkov.

ponovitev poskusa	podatek prve merilne naprave	podatek druge merilne naprave
1.	x_1	y_1
2.	x_2	y_2
$\vdots$	$\vdots$	$\vdots$
n-ta	x_n	y_n

Seveda lahko obdelamo vsako stolpec posebej kot pri eni merilni napravi.
Če stolpca centraliziramo, dobimo **centralizirano tabelo podatkov**.
Če ju standardiziramo, dobimo **standardizirano tabelo podatkov**.

Še bolj nas zanima koliko so podatki v enem stolpcu odvisni od podatkov v drugem stolpcu. Stopnjo odvisnosti merimo s **kovarianco**

$$K = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = \frac{1}{n}\sum_{i=1}^{n} x_i' y_i'$$

Če so podatki prve merilne naprave neodvisni od podatkov druge merilne naprave, je kovarianca enaka nič. Obratno ne velja.

Opomba: Pojem kovariance je tesno povezan s pojmom regresijske premice. Spomnimo se, da je to premica, ki se najbolje prilega točkam $(x_1, y_1), \ldots, (x_n, y_n)$. Izpeljali smo, da je njen smerni koeficient enak

$$k = \frac{\overline{xy} - \bar{x}\bar{y}}{\overline{x^2} - \bar{x}^2}.$$

Pokažimo, da je števec tega izraza ravno kovarianca. Velja namreč

$$\begin{aligned} K &= \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) \\ &= \frac{1}{n}\sum_{i=1}^{n} x_i y_i - \frac{1}{n}\sum_{i=1}^{n}\bar{x}y_i - \frac{1}{n}\sum_{i=1}^{n} x_i\bar{y} + \frac{1}{n}\sum_{i=1}^{n}\bar{x}\bar{y} \\ &= \overline{xy} - \bar{x}\bar{y} - \bar{x}\bar{y} + \bar{x}\bar{y} \\ &= \overline{xy} - \bar{x}\bar{y} \end{aligned}$$

Skoraj enak račun nam pokaže, da je imenovalec enak s_1^2, kjer je s_1 standardna deviacija prvega stolpca. Velja torej $k = \frac{K}{s_1^2}$.

S pomočjo kovariance in standardnih deviacij obeh stolpcev lahko definiramo **kovariančno matriko**

$$C = \begin{bmatrix} s_1^2 & K \\ K & s_2^2 \end{bmatrix}.$$

Če vstavimo definicije s_1, s_2 in K, dobimo

$$\begin{aligned} C &= \begin{bmatrix} \frac{1}{n}\sum_{i=1}^n (x_i')^2 & \frac{1}{n}\sum_{i=1}^n x_i' y_i' \\ \frac{1}{n}\sum_{i=1}^n x_i' y_i' & \frac{1}{n}\sum_{i=1}^n (y_i')^2 \end{bmatrix} \\ &= \frac{1}{n}\begin{bmatrix} x_1' & \dots & x_n' \\ y_1' & \dots & y_n' \end{bmatrix} \begin{bmatrix} x_1' & y_1' \\ \vdots & \vdots \\ x_n' & y_n' \end{bmatrix} \\ &= \frac{1}{n}(X')^T X' \end{aligned}$$

kjer je X' matrika centraliziranih podatkov.

Opomba: Matriko X' dobimo tako, da v matriki rezultatov X, od vsakega elementa odštejemo povprečje stolpca v katerem se nahaja.

Opazimo, da se kovarianca vedno nahaja med $-s_1 s_2$ in $s_1 s_2$. To sledi iz Cauchy-Schwartzove neenakosti

$$|\sum_{i=1}^{n} x_i' y_i'| \leq \sqrt{\sum_{i=1}^{n} (x_i')^2} \sqrt{\sum_{i=1}^{n} (y_i')^2}.$$

če obe strani delimo z $n = (\sqrt{n})^2$. Zato pridemo na idejo, da bi definirali

$$r = \frac{K}{s_1 s_2}$$

Temu izrazu pravimo **korelacija** (ali korelacijski koeficient) obeh stolpcev v matriki podatkov. S standardiziranimi podatki se izraža takole

$$r = \frac{1}{n s_1 s_2} \sum_{i=1}^{n} x_i' y_i' = \frac{1}{n} \sum_{i=1}^{n} x_i'' y_i''$$

Torej je korelacija enaka nič natanko tedaj, ko sta stolpca v standardizirani tabeli podatkov ortogonalna.

Definirajmo še **korelacijsko matriko**

$$R = \begin{bmatrix} 1 & r \\ r & 1 \end{bmatrix}$$

Opazimo, da velja

$$\begin{aligned} R &= \begin{bmatrix} \frac{1}{n}\sum_{i=1}^{n}(x_i'')^2 & \frac{1}{n}\sum_{i=1}^{n} x_i'' y_i'' \\ \frac{1}{n}\sum_{i=1}^{n} x_i'' y_i'' & \frac{1}{n}\sum_{i=1}^{n}(y_i'')^2 \end{bmatrix} \\ &= \frac{1}{n}\begin{bmatrix} x_1'' & \dots & x_n'' \\ y_1'' & \dots & y_n'' \end{bmatrix} \begin{bmatrix} x_1'' & y_1'' \\ \vdots & \vdots \\ x_n'' & y_n'' \end{bmatrix} \\ &= \frac{1}{n}(X'')^T X'' \end{aligned}$$

kjer je X'' matrika standardiziranih podatkov.

Opomba: Matriko X'' dobimo tako, da v matriki centraliziranih podatkov X' vsak stolpec delimo z njegovo standardno deviacijo.

Posplošimo sedaj definicije iz dveh na p merilnih naprav. Naj bo $x_{i,j}$ podatek j-te merilne naprave pri i-ti ponovitvi poskusa.

	1. merilnik	2. merilnik	...	p-ti merilnik
1. ponovitev	$x_{1,1}$	$x_{1,2}$	...	$x_{1,p}$
2. ponovitev	$x_{2,1}$	$x_{2,2}$	...	$x_{2,p}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$
n-ta ponovitev	$x_{n,1}$	$x_{n,2}$	...	$x_{n,p}$
povprečje	$\bar{x}_1$	$\bar{x}_2$	...	$\bar{x}_p$
stand. deviacija	s_1	s_2	...	s_p

Za vsak stolpec v tabeli smo izračunali povprečje in standardno deviacijo

$$\bar{x}_j = \frac{1}{n}\sum_{i=1}^{n} x_{i,j}, \qquad s_j = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_{i,j} - \bar{x}_j)^2}$$

Matriki $X = [x_{i,j}]$ pravimo **matrika podatkov**. Njena velikost je $n \times p$, kjer je n število ponovitev poskusa, p pa število merilnih naprav.

Matriki $X' = [x'_{i,j}]$, kjer je

$$x'_{i,j} = x_{i,j} - \bar{x}_j$$

odmik od povprečja, bomo rekli **centralizirana matrika podatkov**.

Matriki $X'' = [x''_{i,j}]$, kjer je

$$x''_{i,j} = x'_{i,j}/s_j = (x_{i,j} - \bar{x}_j)/s_j$$

standardiziran odmik, bomo rekli **standardizirana matrika podatkov**.

Kovarianca j-tega in k-tega stolpca matrike X je enaka

$$C_{j,k} = \frac{1}{n}\sum_{i=1}^{n}(x_{i,j} - \bar{x}_j)(x_{i,k} - \bar{x}_k) = \frac{1}{n}\sum_{i=1}^{n} x'_{i,j}x'_{i,k}$$

Opazimo, da je $C_{j,j} = s_j^2$, kjer je s_j standardna deviacija j-tega stolpca matrike X. **Kovariančna matrika** $C = [C_{j,k}]$ zadošča

$$C = \frac{1}{n}(X')^T X'$$

Korelacija med j-tim in k-tim stolpcem matrike X je enaka

$$r_{j,k} = C_{j,k}/(s_j s_k)$$

Velja $-1 \leq r_{j,k} \leq 1$ in $r_{j,j} = 1$. **Korelacijska matrika** $R = [r_{j,k}]$ zadošča

$$R = \frac{1}{n}(X'')^T X''$$

10. Metoda glavnih komponent (PCA)

Naj bo X' centralizirana matrika podatkov za nek poskus.
(Pozor, nekateri raje delajo s standardizirano matriko podatkov X''.)

Radi bi zmanjšali število stolpcev matrike X' (iz p na k), ne da bi pri tem izgubili veliko informacij. Ideja je, da si zapomnimo samo take linearne kombinacije stolpcev matrike X', ki imajo največjo standarno deviacijo.

Spomnimo se, da vsako linearno kombinacijo stolpcev X' lahko zapišemo kot $X'w$ za nek vektor w. Ponavljamo naslednji postopek:

- Poišči tak normiran vektor w_1, da ima vektor $y_1 := X'w_1$ največjo možno standardno deviacijo.
- Med vsemi normiranimi vektorji, ki so ortogonalni na w_1, poišči tak vektor w_2, da ima $y_2 := X'w_2$ največjo možno standardno deviacijo.
- Med vsemi normiranimi vektorji, ki so ortogonalni na w_1 in w_2, poišči tak vektor w_3, da ima $y_3 := X'w_3$ največjo možno stand. deviacijo.
- ...

Končamo, ko se X' ujema z $y_1 w_1^T + \ldots + y_k w_k^T$ do predpisane natančnosti.

Opomba: Vektorjem $y_1, \ldots, y_k$ bomo rekli **glavne komponente** matrike X', vektorjem $w_1, \ldots, w_k$ pa **glavne osi** matrike X'.
Pozor, terminologija se zelo razlikuje od avtorja do avtorja.

Trditev 1

Za vektorje $w_1, w_2, w_3, \ldots$ lahko vzamemo kar lastne vektorje kovariančne matrike C, ki ustrezajo prvi, drugi, tretji, ... največji lastni vrednosti.

Opomba: Lastni vektorji kovariančne matrike

$$C = \frac{1}{n}(X')^T X'$$

se ujemajo z lastnimi vektorji matrike

$$(X')^T X'$$

ti pa se ujemajo s stolpci matrike Q_2 v singularnem razcepu

$$X' = Q_1 D Q_2^T$$

Dokaz. Naj bodo $\lambda_1 \geq \ldots \geq \lambda_p$ lastne vrednosti matrike C in naj bodo $v_1, \ldots, v_p$ pripadajoči lastni vektorji. Poskrbimo, da so ortonormirani.

Vsak normiran vektor $w \in \mathbb{R}^p$ lahko zapišemo kot $w = \beta_1 v_1 + \ldots + \beta_p v_p$ kjer je $\beta_1^2 + \ldots + \beta_p^2 = 1$. Standardna deviacija vektorja $X'w$ je enaka

$$\begin{aligned}
\sigma_{X'w} &= \sqrt{\frac{1}{n}(X'w)^T(X'w)} \\
&= \sqrt{w^T C w} \\
&= \sqrt{(\beta_1 v_1 + \ldots + \beta_p v_p)^T(\beta_1 \lambda_1 v_1 + \ldots + \beta_p \lambda_p v_p)} \\
&= \sqrt{\beta_1^2 \lambda_1 + \ldots + \beta_p^2 \lambda_p} \\
&= \sqrt{(1 - \beta_2^2 - \ldots - \beta_p^2)\lambda_1 + \beta_2^2 \lambda_2 + \ldots + \beta_p^2 \lambda_p} \\
&= \sqrt{\lambda_1 - (\beta_2^2(\lambda_1 - \lambda_2) + \ldots + \beta_p^2(\lambda_1 - \lambda_p))} \\
&\leq \sqrt{\lambda_1}
\end{aligned}$$

Enačaj je očitno dosežen pri $w = v_1$. Torej lahko vzamemo $w_1 = v_1$.

Vsak normiran vektor $w \in \mathbb{R}^p$, ki je ortogonalen na v_1 lahko zapišemo kot

$$w = \gamma_2 v_2 + \ldots + \gamma_p v_p,$$

kjer je $\gamma_2^2 + \ldots + \gamma_p^2 = 1$. Standardna deviacija vektorja $X'w$ je enaka

$$\begin{aligned}
\sigma_{X'w} &= \sqrt{w^T C w} \\
&= \sqrt{\gamma_2^2 \lambda_2 + \ldots + \gamma_p^2 \lambda_p} \\
&= \sqrt{(1 - \gamma_3^2 - \ldots - \gamma_p^2)\lambda_2 + \gamma_3^2 \lambda_3 + \ldots + \gamma_p^2 \lambda_p} \\
&= \sqrt{\lambda_2 - (\gamma_3^2(\lambda_2 - \lambda_3) + \ldots + \gamma_p^2(\lambda_2 - \lambda_p))} \\
&\leq \sqrt{\lambda_2}
\end{aligned}$$

Enačaj je dosežen pri $w = v_2$, zato lahko vzamemo $w_2 = v_2$.

Normiran vektor $w \in \mathbb{R}^p$, ki je ortogonalen na v_1 in v_2 lahko zapišemo kot $w = \delta_3 v_3 + \ldots + \delta_p v_p$, kjer je $\delta_3^2 + \ldots + \delta_p^2 = 1$. Podobno kot zgoraj je $\sigma_{X'w} \leq \sqrt{\lambda_3}$ in enačaj je dosežen pri $w = v_3$. To lahko nadaljujemo. □

Trditev 3

Če je $i \neq j$, potem je kovarianca med stolpcema y_i in y_j enaka nič.

Dokaz: Ker sta stolpca y_i in y_j linearni kombinaciji stolpcev matrike X', sta njuni povprečji enaki nič. Torej je njuna kovarianca enaka

$$\begin{aligned} K &= \frac{1}{n}\langle y_i, y_j \rangle \\ &= \frac{1}{n} y_j^T y_i \\ &= \frac{1}{n}(X' w_j)^T (X' w_i) \\ &= w_j^T C w_i \\ &= w_j^T (\lambda_i w_i) \\ &= \lambda_i \langle w_i, w_j \rangle \\ &= 0 \end{aligned}$$

Opomba: Različne glavne komponente X' so torej nekorelirane.

V nadaljevanje nas bo zanimalo, kolikšno napako naredimo, če matriko X' zamenjamo z matriko $y_1 w_1^T + \ldots + y_k w_k^T$. Napako bomo merili s Frobeniusovo matrično normo.

Uvedimo nekaj oznak. Za vsak $k = 1, \ldots, p$ naj bo

$$W_k = \begin{bmatrix} w_1 & \ldots & w_k \end{bmatrix} \quad \text{in} \quad Y_k = \begin{bmatrix} y_1 & \ldots & y_k \end{bmatrix} = X' W_k$$

Očitno je

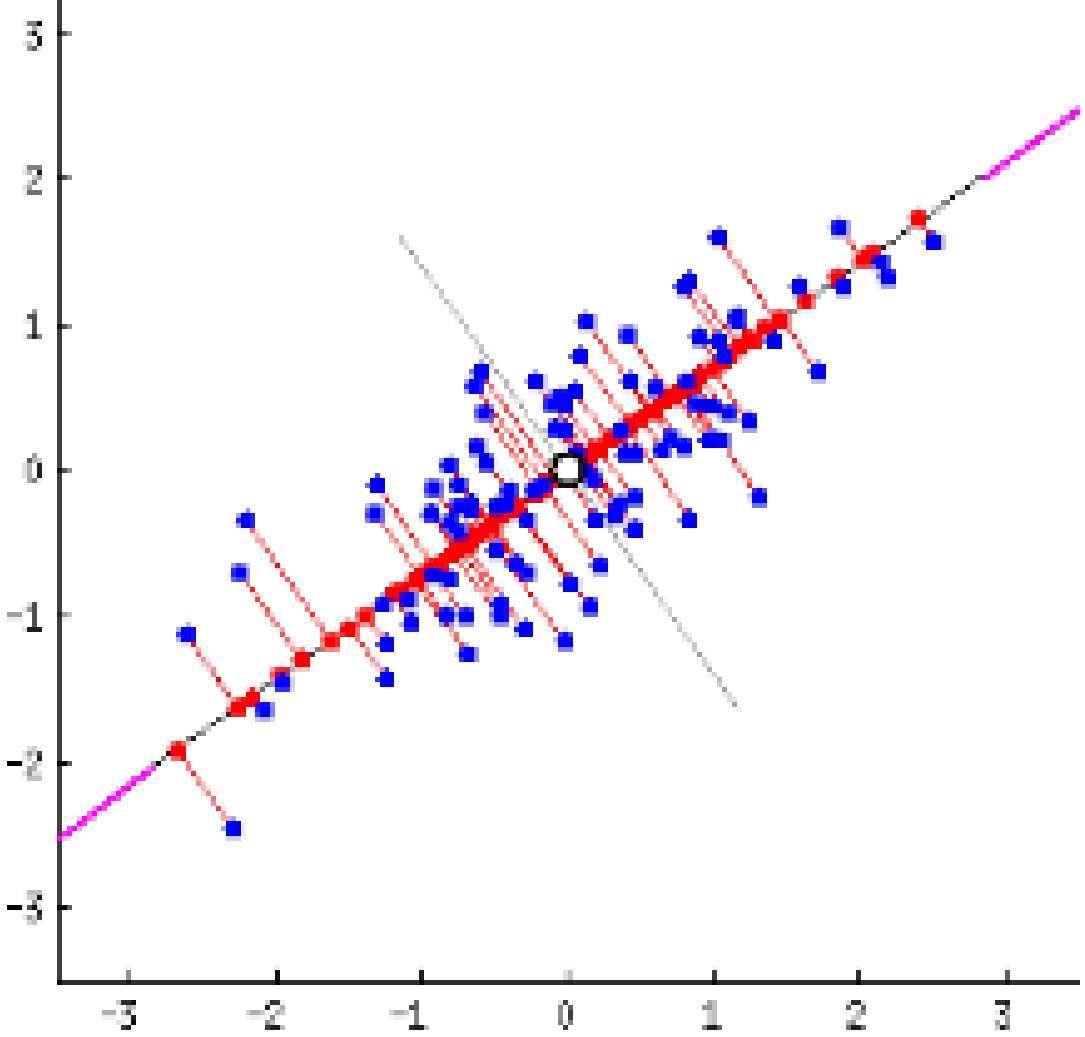
$$Y_k W_k^T = y_1 w_1^T + \ldots + y_k w_k^T$$

Iskana napaka je torej

$$\|X' - Y_k W_k^T\|$$

Trditev 4 (Geometrijski pomen $Y_k W_k^T$ in $\|X' - Y_k W_k^T\|$)

Če i-to vrstico matrike X' ortogonalno projeciramo na podprostor $\mathrm{Lin}\{w_1^T, \ldots, w_k^T\}$ v $\mathbb{R}^p$, potem dobimo ravno i-to vrstico matrike $Y_k W_k^T$. Vsota kvadratov oddaljenosti vrstic matrike X' od njihovih ortogonalnih projekcij na podprostor $\mathrm{Lin}\{w_1^T, \ldots, w_k^T\}$ je torej enaka $\|X' - Y_k W_k^T\|^2$.



Vsaka koordinatna os ustreza eni merilni napravi. Na skici je $p = 2$.
Vsaka modra točka ustreza eni vrstici matrike X', torej eni ponovitvi poskusa.
Modra premica ustreza podprostoru $\mathrm{Lin}\{w_1^T, \ldots, w_k^T\}$. Na skici je $k = 1$.
Vsaka rdeča točka ustreza eni vrstici matrike $Y_k W_k^T$.

Dokaz. Vrstice matrik X' in $Y_k W_k^T$ ter vektorji $w_1^T, \ldots, w_k^T$ so vrstični vektorji dolžine p. Ker so $w_1, \ldots, w_k$ ortonormirani, so tudi $w_1^T, \ldots, w_k^T$ ortonormirani. Naj bo $x^{(i)}$ i-ta vrstica matrike X'. Njena ortogonalna projekcija na podprostor $\mathrm{Lin}\{w_1^T, \ldots, w_k^T\}$ je enaka $\sum_{j=1}^{k} \langle x^{(i)}, w_j^T \rangle w_j^T$.

Najprej opazimo, da je skalarni produkt $\langle x^{(i)}, w_j^T \rangle$ enak matričnemu produktu $x^{(i)} w_j$. Poleg tega je

$$\sum_{j=1}^{k} \langle x^{(i)}, w_j^T \rangle w_j^T = \begin{bmatrix} x^{(i)} w_1 & \ldots & x^{(i)} w_k \end{bmatrix} \begin{bmatrix} w_1^T \\ \vdots \\ w_k^T \end{bmatrix} = x^{(i)} W_k W_k^T.$$

Tudi i-ta vrstica produkta $X'(W_k W_k^T)$ je enaka $x^{(i)} W_k W_k^T$.

Drugi del trditve sledi iz prvega dela in naslednjega računa. Za poljubni matriki A in B iste velikosti $n \times p$ velja

$$\|A - B\|^2 = \sum_{i=1}^{n} \sum_{j=1}^{p} (a_{i,j} - b_{i,j})^2 = \sum_{i=1}^{n} \|a^{(i)} - b^{(i)}\|^2. \quad \square$$

Trditev 5

Če so $\lambda_1 \geq \ldots \geq \lambda_p$ lastne vrednosti kovariančne matrike $C = \frac{1}{n}(X')^T X'$, potem je $\|X' - Y_k W_k^T\|^2 = n(\lambda_{k+1} + \ldots + \lambda_p)$.

Dokaz. Izračunajmo razdaljo med X' in $Y_k W_k^T$ v Frobeniusovi normi. Naj bo $X' = UDW^T$ singularni razcep matrike X' in naj bo $W = \begin{bmatrix} W_k & Z \end{bmatrix}$ za nek Z. Iz $I_p = W^T W = W^T \begin{bmatrix} W_k & Z \end{bmatrix} = \begin{bmatrix} W^T W_k & W^T Z \end{bmatrix}$ sledi $W^T W_k = \begin{bmatrix} I_k \\ 0 \end{bmatrix}$. Naj bo D_k matrika, ki jo dobimo iz D tako, da vse diagonalne elemente razen prvih k zamenjamo z nič. Opazimo, da velja

$$Y_k W_k^T = X' W_k W_k^T = UDW^T W_k W_k^T = UD \begin{bmatrix} I_k \\ 0 \end{bmatrix} W_k^T =$$

$$= UD_k \begin{bmatrix} I_k \\ 0 \end{bmatrix} W_k^T = UD_k \begin{bmatrix} W_k^T \\ 0 \end{bmatrix} = UD_k \begin{bmatrix} W_k^T \\ Z^T \end{bmatrix} = UD_k W^T$$

Odtod sledi, da je v Frobeniusovi normi

$$\|X' - Y_k W_k^T\| = \|U(D - D_k)W^T\| = \|D - D_k\|.$$

Če so $\lambda_1 \geq \ldots \geq \lambda_p$ lastne vrednosti matrike $C = \frac{1}{n}(X')^T X'$, potem so $n\lambda_1 \geq \ldots \geq n\lambda_p$ lastne vrednosti matrike $(X')^T X'$. Odtod sledi, da so $\sqrt{n\lambda_1} \geq \ldots \geq \sqrt{n\lambda_p}$ diagonalni elementi matrike D. Torej je

$$\|D - D_k\|^2 = (\sqrt{n\lambda_{k+1}})^2 + \ldots + (\sqrt{n\lambda_p})^2 = n(\lambda_{k+1} + \ldots + \lambda_p). \quad \square$$

Naslednja posledica Trditve 5 nam pove kako izbrati k, da dosežemo predpisano natančnost.

Posledica

Če za nek k velja $\lambda_{k+1} + \ldots + \lambda_p \leq \frac{\varepsilon^2}{n}$, potem je $\|X' - Y_k W_k^T\| \leq \varepsilon$.
Če pa za nek k velja $\frac{\lambda_{k+1}+\ldots+\lambda_p}{\lambda_1+\ldots+\lambda_k+\lambda_{k+1}+\ldots+\lambda_p} \leq \varepsilon$, potem je $\frac{\|X'-Y_k W_k^T\|^2}{\|X'\|^2} \leq \varepsilon$.

Za konec si oglejmo povzetek algoritma za računanje glavnih komponent.

1. Preberi matriko podatkov X in predpisano relativno natančnost ε.
2. Zapomni si povprečja stolpcev matrike X; $\bar{x} := \begin{bmatrix} \bar{x}_1 & \dots & \bar{x}_p \end{bmatrix}$.
3. Centraliziraj matriko podatkov; $X' := X - e^T \bar{x}$, kjer $e = [1, \dots, 1]$.
4. Poišči singularni razcep $X' = UDW^T$, kjer za diagonalne elemente matrike D (=singularne vrednosti matrike X') velja $\sigma_1 \geq \dots \geq \sigma_p$.
5. Poišči tak k, da velja $\frac{\sigma_1^2 + \dots + \sigma_k^2}{\sigma_1^2 + \dots + \sigma_k^2 + \dots + \sigma_p^2} \geq 1 - \varepsilon^2$.
6. Zapomni si prvih k stolpcev matrike W ($= W_k =$ glavne osi X') in prvih k stolpcev matrike UD ($= Y_k =$ glavne komponente X')
7. Velja $\frac{\|X' - Y_k W_k^T\|}{\|X'\|} \leq \varepsilon$. Torej je $Y_k W_k^T$ dober približek za X'.
   Odtod sledi, da je $Y_k W_k^T + e^T \bar{x}$ dober približek za matriko X.

Kvadratne forme

1. Uvod

Monom v spremenljivkah $x_1, \ldots, x_n$ je izraz oblike

$$c\, x_1^{d_1} \cdots x_n^{d_n}$$

kjer so $d_1, \ldots, d_n \in \mathbb{N}_0$ in $c \in \mathbb{R}$ (lahko tudi kako drugo polje). Število c je **koeficient** monoma. Število $d_1 + \ldots + d_n$ je **stopnja** monoma (če $c \neq 0$).

Polinom v spremenljivkah $x_1, \ldots, x_n$ je vsota monomov v $x_1, \ldots, x_n$.
S krajšanjem dosežemo, da imajo vsi ti monomi različne $(d_1, \ldots, d_n)$.
Stopnja polinoma je maksimum stopenj njegovih neničelnih monomov.
Polinom je **konstanten/linearen/kvadraten/kubičen**, če je stopnje 0/1/2/3. Polinom je **forma**, če so vsi njegovi monomi iste stopnje.

Primer

Vsak kvadraten polinom v spremenljivkah x, y je oblike

$$ax^2 + bxy + cy^2 + dx + ey + f$$

Opazimo, da je ta polinom vsota kvadratne forme $ax^2 + bxy + cy^2$, linearne forme $dx + ey$ in konstantne forme f

Primer

Vsaka kvadratna forma v spremenljivkah x, y, z je oblike

$$ax^2 + by^2 + cz^2 + dxy + eyz + fzx$$

Če vstavimo $z = 1$ dobimo splošen kvadraten polinom

$$ax^2 + by^2 + c + dxy + ey + fx = ax^2 + dxy + by^2 + fx + ey + c$$

Linearne forme v spremenljivkah $x_1, \ldots, x_n$ so oblike

$$\sum_{i=1}^{n} a_i x_i = a^T x,$$

kjer je $a = [a_i]$ vektor koeficientov in $x = [x_i]$ vektor spremenljivk.

Kvadratne forme v spremenljivkah $x_1, \ldots, x_n$ so oblike

$$\sum_{i=1}^{n} \sum_{j=1}^{n} a_{i,j} x_i x_j = x^T A x,$$

kjer je $A = [a_{i,j}]$ matrika koeficientov in je $x = [x_i]$ vektor spremenljivk.
Ker je $a_{i,j} x_i x_j + a_{j,i} x_j x_i = (a_{i,j} + a_{j,i}) x_j x_i = \frac{a_{i,j}+a_{j,i}}{2} x_i x_j + \frac{a_{i,j}+a_{j,i}}{2} x_j x_i$,
lahko predpostavimo, da je $a_{i,j} = a_{j,i}$, se pravi, da je A simetrična.

Primer

$$ax^2 + bxy + cy^2 = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

$$dx + ey = \begin{bmatrix} d & e \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Primer

$$ax^2 + by^2 + cz^2 + dxy + eyz + fzx =$$

$$= \begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} a & d/2 & f/2 \\ d/2 & b & e/2 \\ f/2 & e/2 & c \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Opomba: Vidimo, da je pri mešanih členih bolje namesto d, e, f pisati $2d, 2e, 2f$.

2. Zamenjava spremenljivk

Če v kvadratno formo

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_{i,j}x_i x_j = x^T A x$$

vstavimo

$$\begin{aligned} x_1 &= p_{1,1}x_1' + \ldots + p_{1,n}x_n' \\ &\vdots \qquad\quad \vdots \\ x_n &= p_{n,1}x_1' + \ldots + p_{n,n}x_n' \end{aligned}$$

se pravi $x = Px'$, kjer $P = [p_{i,j}]$, dobimo novo kvadratno formo

$$x^T A x = (x')^T P^T A P x' = (x')^T A' x'$$

Definicija

Matrika A je **kongruentna** matriki B, če obstaja taka obrnljiva matrika P, da velja $B = P^T AP$.

Opomba: Kongruentnost matrik je ekvivalenčna relacija.

Definicija

Signatura simetrične matrike A je urejeni par (k, l), kjer je k število strogo pozitivnih lastnih vrednosti A, l pa število strogo negativnih lastnih vrednosti A (oboje je šteto z večkratnostmi).

Sylvestrov izrek o inerciji

Simetrična matrika A je kongruentna matriki bločne oblike

$$\begin{bmatrix} I_k & 0 & 0 \\ 0 & -I_l & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

natanko tedaj, ko je (k, l) signatura matrike A.

Dokaz lažje smeri: Vemo, da za vsako simetrično matriko A obstaja taka ortogonalna matrika Q in taka diagonalna matrika $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, da velja $A = QDQ^{-1}$. S permutacijo diagonalnih elementov D in z ustrezno permutacijo stolpcev P lahko dosežemo, da je

$$\lambda_1 > 0, \ldots, \lambda_k > 0, \lambda_{k+1} < 0, \ldots, \lambda_{k+l} < 0, \lambda_{k+l+1} = 0, \ldots, \lambda_n = 0$$

Definirajmo matriko

$$E = \mathrm{diag}(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_k}, \sqrt{-\lambda_{k+1}}, \ldots, \sqrt{\lambda_{k+l}}, 1, \ldots, 1)$$

Ker je $D = E\,\mathrm{diag}(I_k, -I_l, 0)E^T$, je

$$A = QDQ^{-1} = QDQ^T = QE\,\mathrm{diag}(I_k, -I_l, 0)(QE)^T$$

Ker sta Q in E obrnljivi, je torej A kongruentna $\mathrm{diag}(I_k, -I_l, 0)$.

Dokaž težje smeri izpustimo.

Posledica

Dve simetrični matriki iste velikosti sta kongruentni natanko tedaj, ko imata enako signaturo.

Posledica

Vsako kvadratno formo

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_{i,j}x_i x_j$$

kjer je $[a_{i,j}]$ simetrična matrika s signaturo (k, l), lahko z linearno zamenjavo spremenljivk prevedemo na obliko

$$(x'_1)^2 + \ldots + (x'_k)^2 - (x'_{k+1})^2 - \ldots - (x'_{k+l})^2$$

Primer

Vsako kvadratno formo $ax^2 + bxy + cy^2$ lahko z linearno zamenjavo spremenljivk prevedemo na natanko eno od naslednjih kvadratnih form

$$0, \quad (x')^2, \quad -(x')^2, \quad (x')^2 + (y')^2, \quad (x')^2 - (y')^2, \quad -(x')^2 - (y')^2$$

3. Krivulje 2. reda

Krivulja drugega reda je množica rešitev enačbe

$$p(x, y) = 0,$$

kjer je $p(x, y)$ realen polinom druge stopnje, se pravi

$$p(x, y) = ax^2 + 2bxy + cy^2 + 2dx + 2ey + f,$$

kjer $a, b, c, d, e, f \in \mathbb{R}$. Primer $b = 0$ ste obravnavali že v gimnaziji, kjer ste se naučili, da se taka krivulja s primernim premikom prevede na elipso ali hiperbolo ali parabolo ali par premic. Oglejmo si še primer $b \neq 0$.

Množico rešitev enačbe $p(x, y) = 0$ označimo z

$$V(p) := \{ \begin{bmatrix} x \\ y \end{bmatrix} \mid p(x, y) = 0 \}.$$

Če definicijo polinoma $p(x, y)$ zapišemo v matrični obliki, dobimo

$$p(x,y) = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + 2 \begin{bmatrix} d & e \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + f \tag{1}$$

Iz poglavja o simetričnih matrikah vemo, da obstaja taka ortogonalna matrika P, da velja

$$\begin{bmatrix} a & b \\ b & c \end{bmatrix} = P \begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix} P^T. \tag{2}$$

Z morebitno spremembo predznaka prvega stolpca matrike P lahko dosežemo, da je $\det P = 1$. Pokazali smo, da je potem P vrtež

$$P = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix} \tag{4}$$

Definirajmo novi spremenljivki x' in y' z

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = P^T \begin{bmatrix} x \\ y \end{bmatrix} \tag{5}$$

Označimo še $\begin{bmatrix} d' & e' \end{bmatrix} = \begin{bmatrix} d & e \end{bmatrix} P$. Potem velja

$$p(x,y) = \begin{bmatrix} x' & y' \end{bmatrix} \begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} + 2 \begin{bmatrix} d' & e' \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} + f \tag{6}$$

$$= d_1(x')^2 + d_2(y')^2 + 2d'x' + 2e'y' + f.$$

Označimo

$$q(x,y) = d_1x^2 + d_2y^2 + 2d'x + 2e'y + f. \tag{7}$$

Množico rešitev enačbe $q(x,y) = 0$, se pravi

$$V(q) = \{ \begin{bmatrix} x \\ y \end{bmatrix} \mid q(x,y) = 0 \}$$

znamo narisati, ker $q(x,y)$ ne vsebuje člena xy.

Kako iz množice $V(q)$ dobimo množico $V(p)$? Iz (6) in (7) sledi

$$q(x', y') = p(x, y). \tag{8}$$

Če upoštevamo še zvezo (5), dobimo

$$\begin{aligned} V(p) &= \{\begin{bmatrix} x \\ y \end{bmatrix} \mid p(x, y) = 0\} \\ &= \{P \begin{bmatrix} x' \\ y' \end{bmatrix} \mid p(x, y) = 0\} \\ &= \{P \begin{bmatrix} x' \\ y' \end{bmatrix} \mid q(x', y') = 0\} \\ &= \{P \begin{bmatrix} x' \\ y' \end{bmatrix} \mid \begin{bmatrix} x' \\ y' \end{bmatrix} \in V(q)\} \end{aligned}$$

Torej množico $V(p)$ dobimo tako, da množico $V(q)$ zavrtimo za kot ϕ v pozitivno smer.

Primer

Nariši krivuljo $4x^2 + 4xy + 7y^2 = 1$

Rešitev: Matrika polinoma je

$$A = \begin{bmatrix} 4 & 2 \\ 2 & 7 \end{bmatrix}.$$

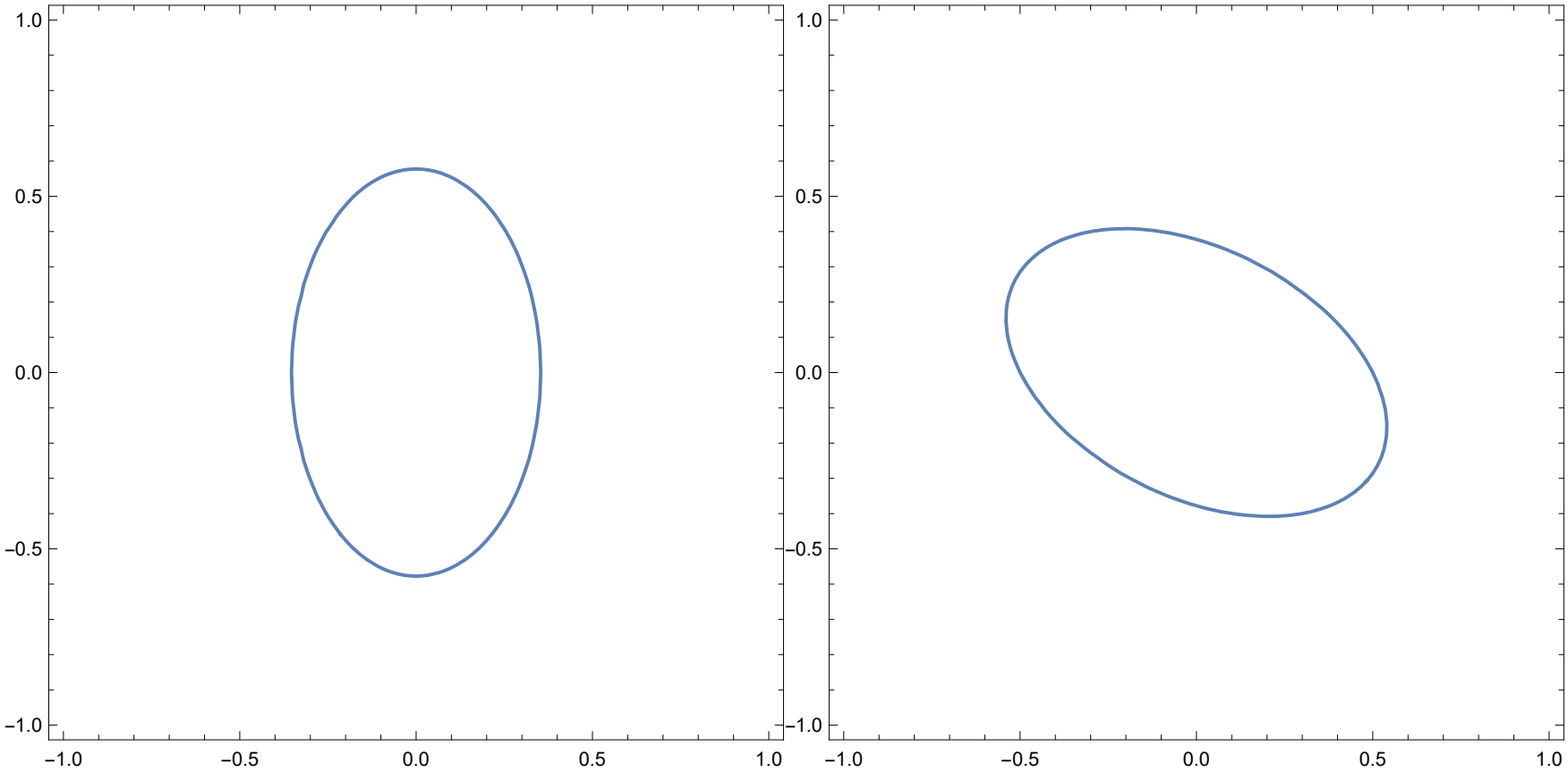
Lastni vrednosti sta 8 in 3. Ortogonalna matrika iz pripadajočih lastnih vektorjev je

$$P = \begin{bmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix}$$

To je vrtež za kot $\operatorname{arctg} 2$ kar je približno 63.4°. Elipso

$$8x^2 + 3y^2 = 1$$

znamo narisati. Če jo zavrtimo za 63.4° dobimo iskano krivuljo.



Leva krivulja je $8x^2 + 3y^2 = 1$, desna pa $4x^2 + 4xy + 7y^2 = 1$.

Primer

Če imamo dve merilni napravi in veliko ponovitev, potem elipsa

$$\begin{bmatrix} x & y \end{bmatrix} C^{-1} \begin{bmatrix} x \\ y \end{bmatrix} = 5.9912,$$

vsebuje približno 95% centraliziranih podatkov (=vrstic matrike X').
Njeni glavni osi se ujemata z lastnima podprostoroma matrike C.