Lekcija 1

Sustavi Linearnih Jednadžbi

Studij linearne algebre počinjemo razmatranjem triju problema. Iako ti problemi na prvi pogled izgledaju vrlo jednostavno, oni će nas dovesti do iskustva da moramo biti vrlo oprezni kako ne bismo stvorili krive zaključke. Kad jednom usvojimo potrebnu dozu opreznosti, linearna algebra će se razvijati preko pitanja i odgovora na niz vrlo jednostavnih pitanja.

Tri problema glase: pomoću elementarnih aritmetičkih operacija nađi sve trojke realnih brojeva x_1, x_2, x_3 koji istovremeno zadovoljavaju sljedeće jednadžbe:

Problem 1

$$\begin{aligned} x_1 + x_3 &= 0 \\ x_2 - x_3 &= 1 \\ x_1 + 2x_2 + x_3 &= 1 \end{aligned}$$

Problem 2

$$\begin{aligned} x_1 + x_3 &= 0 \\ x_2 - x_3 &= 1 \\ x_1 + 2x_2 - x_3 &= 1 \end{aligned}$$

Problem 3

$$\begin{aligned} x_1 + x_3 &= 0 \\ x_2 - x_3 &= 1 \\ x_1 + 2x_2 - x_3 &= 2. \end{aligned}$$

Jednadžbe u tim problemima izgledaju vrlo slično; u stvari jednadžbe drugog problema razlikuju se od prvog samo na jednom mjestu u trećoj jednadžbi, a jednadžbe trećeg problema razlikuju se od drugog opet na samo jednom mjestu u trećoj jednadžbi. Sličnost vara, kao što ćemo uskoro vidjeti.

Jedan od zahtjeva u rješavanju tih problema jest korištenje osnovnih aritmetičkih operacija. Mogli bismo krenuti nasumce, kao što smo radili s dvije jednadžbe i dvije nepoznanice u osnovnoj školi. Ipak, u izvršavanju tih operacija poželjno je imati neki red. Zgodno odabran redoslijed operacija će nas dovesti do dobro definiranog algoritma odnosno programa ako će se algoritam implementirati na računalu.

Prije razmatranja gornjih triju problema, uvest ćemo tri pravila ili drukčije gledano tri transformacije, koje će olakšati razmatranje. Ta tri pravila će činiti osnovu za skoro sva računanja vezana uz rješavanje sustava.

Budući da se ista pravila mogu koristiti i za rješavanje sličnih, možda i kompliciranijih problema, korisno je izreći ih u što općenitijem obliku. Promotrimo zato općeniti sustav od m jednadžbi sa n nepoznanica oblika

$$\begin{array}{l} f_1(x_1, x_2, ..., x_n) = b_1 \\ f_2(x_1, x_2, ..., x_n) = b_2 \\ \dots\dots\dots\dots\dots\dots \\ f_m(x_1, x_2, ..., x_n) = b_m\,. \end{array} \tag{1.1}$$

Ovdje su f_i realne (ili kompleksne) funkcije od n realnih (ili kompleksnih) varijabli. **Riješiti** sustav znači odrediti sve $n-$torke brojeva $x_1,\, x_2,\, \ldots, x_n$ tako da svaka $n-$torka uvrštena u gornji sustav od m jednadžbi na lijevoj strani daje točno one vrijednosti (u danom redosljedu) koje se pojavljuju na desnoj strani jednadžbi. Svaku $n-$torku sa tim svojstvom zovemo **rješenje** sustava jednadžbi (1.1). Tri pravila koja ćemo definirati daju način na koji možemo dani sustav jednadžbi zamijeniti drugim sustavom, koji je ekvivalentan (istovjetan) polaznom sustavu. Za dva sustava jednadžbi kažemo da su **ekvivalentni**, ako je svako rješenje prvog sustava ujedno i rješenje drugog sustava i ako je svako rješenje drugog sustava također rješenje prvog sustava. Drugim rječima, dva sustava jednadžbi su ekvivalentna, ako zamjenom jednog sustava drugim niti gubimo rješenja niti dobivamo neko novo rješenje.

Tri pravila koja koristimo u dobivanju ekvivalentnog sustava jednadžbi danom sustavu jesu:

1. Zamjena bilo kojih dviju jednadžbi sustava,

2. Množenje neke jednadžbe sustava brojem koji nije nula,

3. Dodavanje jednoj jednadžbi druge jednadžbe koja je pomnožena nekim brojem.

Primjena bilo kojeg od tih pravila na dani sustav dovodi do sustava koji je ekvivalentan polaznom sustavu. Uvjerimo se u to.

Pravilo 1. dovodi do ekvivalentnog sustava, jer poredak u kojem se pojavljuju jednadžbe u sustavu (1.1) nema važnosti za rješenje. Za rješenje je bitno da nakon uvrštenja, lijeva strana svake jednadžbe bude jednaka desnoj srani.

Primijenimo sada pravilo 2. na sustav (1.1). Uzmimo npr. $i-$tu jednadžbu

$$f_i(x_1, x_2, ..., x_n) = b_i \,, \tag{1.2}$$

i pomnožimo ju sa α koji nije nula. Dobivamo

$$\alpha f_i(x_1, x_2, ..., x_n) = \alpha b_i \,, \quad \alpha \neq 0 \,. \tag{1.3}$$

Sve ostale jednadžbe novog sustava su identične odgovarajućim jednadžbama starog sustava. Ako n - torka brojeva $x_1,\ x_2,\ \ldots, x_n$ zadovoljava jednadžbu (1.2), onda će zadovoljavati i jednadžbu (1.3). Tvrdnja vrijedi bez obzira da li je α nula ili nije. Budući da su sve druge jednadžbe ostale iste, svako rješenje starog sustava bit će rješenje novog sustava. Pođimo sada od jednog rješenja $x_1,\ x_2,\ \ldots, x_n$ novog sustava. S obzirom da je $\alpha \neq 0$ (dakle, tek ovdje koristimo pretpostavku netrivijalnosti α), možemo pomnožiti $i-$tu jednadžbu novog sustava sa $1/\alpha$. Sada na isti način kao i prije zaključujemo da je $x_1,\ x_2,\ \ldots, x_n$ rješenje "još novijeg" sustava, koji je zapravo stari sustav. Prema tome, svako rješenje novog sustava je istovremeno rješenje starog sustava. Dakle novi sustav je ekvivalentan starom sustavu.

Pokažimo da 3. pravilo također vodi na ekvivalentan sustav. Izabiremo dvije jednadžbe polaznog sustava, recimo $i-$tu i $j-$tu jednadžbu:

$$\begin{aligned} f_i(x_1, x_2, ..., x_n) &= b_i \\ f_j(x_1, x_2, ..., x_n) &= b_j \,. \end{aligned} \tag{1.4}$$

Nakon primjene trećeg pravila $i-$ta i $j-$ta jednadžba novog sustava glase

$$\begin{aligned} f_i(x_1, x_2, ..., x_n) &= b_i \,, \\ f_j(x_1, x_2, ..., x_n) + \alpha f_i(x_1, x_2, ..., x_n) &= b_j + \alpha b_i \,, \end{aligned} \tag{1.5}$$

dok su sve ostale jednadžbe obaju sustava identične. Ako je $x_1, \ldots, x_n$ rješenje od (1.4), onda očito zadovoljava (1.5). S druge strane, ako je x_1, x_2, $\ldots, x_n$ rješenje sustava (1.5), onda uvrštavanjem b_i umjesto $f_i(x_1, x_2, ..., x_n)$ u drugoj jednadžbi od (1.5), jednadžbe (1.5) postaju (1.4). To znači da je x_1, x_2, $\ldots, x_n$ ujedno rješenje od (1.4). Prema tome, x_1, x_2, $\ldots, x_n$ zadovoljava (1.4) onda i samo onda ako zadovoljava (1.5). Budući da su ostale jednadžbe nepromijenjene, novi sustav je ekvivalentan staromu.

Sada ćemo upotrijebiti tri osnovna pravila u rješavanju problema postavljenih na početku ove lekcije. Kod problema 1., zamijenjujemo treću jednadžbu razlikom treće i prve jednadžbe (to je pravilo 3., jer prvu jednadžbu množimo s -1 i dodajemo ju trećoj jednadžbi). Dobivamo

$$\begin{aligned} x_1 + x_3 &= 0\,, \\ x_2 - x_3 &= 1\,, \\ 2x_2 &= 1\,. \end{aligned}$$

Riješivši treću jednadžbu, dobivamo vrijednost za x_2. Uvrstivši tu vrijednost u drugu jednadžbu dobivamo vrijednost za x_3. Uvrstivši vrijednost za x_3 u prvu jednadžbu dobivamo vrijednst za x_1. Tako smo dobili $x_1 = 1/2$, $x_2 = 1/2$, $x_3 = -1/2$. Dakle postoji jedna i samo jedna trojka brojeva, $1/2$, $1/2$, $-1/2$ koja zadovoljava sustav jednadžbi iz problema 1. Čitalac će odmah primijetiti da sve što smo do sada koristili jesu aritmetičke operacije, sva tri pravila zapravo se sastoje samo od aritmetičkih operacija.

Vratimo se problemu 2. Ako zamijenimo treću jednadžbu problema 2 s jednadžbom koja je razlika nje i prve jednadžbe, dobit ćemo ekvivalentan sustav

$$\begin{aligned} x_1 + x_3 &= 0\,, \\ x_2 - x_3 &= 1\,, \\ 2x_2 - 2x_3 &= 1\,. \end{aligned}$$

Zamijenimo treću jednadžbu novog sustava jednadžbom koja se dobije tako da od treće jednadžbe oduzmemo drugu pomnoženu s -2 (opet 3. pravilo), dobit ćemo ekvivalentan sustav

$$\begin{aligned} x_1 + x_3 &= 0\,, \\ x_2 - x_3 &= 1\,, \\ 0 &= -1\,. \end{aligned}$$

Zadnja jednadžba nam kaže: kad bi postojalo rješenje sustava ono bi bilo takvo da za njega vrijedi $0 = -1$. Budući da to očito nije istina, ne postoji rješenje sustava. Dakle, iako se sustavi u problemu 1 i problemu 2 razlikuju samo u jednom predznaku u trećoj jednadžbi, njihova rješenja su vrlo različita: problem 1 ima jedno i samo jedno rješenje a problem 2 nema niti jedno rješenje.

Promotrimo i treći problem. Primjenom trećeg pravila dva puta (prvo od treće jednadžbe oduzmemo prvu, a onda još oduzmemo drugu jednadžbu pomnoženu s 2), dobivamo ekvivalentan sustav:

$$\begin{aligned} x_1 + x_3 &= 0 \\ x_2 - x_3 &= 1 \\ 0 &= 0\,. \end{aligned}$$

Dakle, dobili smo sustav s dvije jednadžbe i tri nepoznanice

$$\begin{aligned} x_1 + x_3 &= 0 \\ x_2 - x_3 &= 1 \end{aligned}$$

Ako jednadžbe zapišemo u obliku $x_1 = -x_3$, $x_2 = 1 + x_3$, vidimo da za svaki realni (ili kompleksni) broj β, sustav ima rješenje

$$x_1 = -\beta\,, \quad x_2 = 1 + \beta\,, \quad x_3 = \beta\,.$$

Ovaj sustav jednadžbi nema samo jedno rješenje, ima ih beskonačno mnogo.

Sličnost problema 1., 2. i 3. pokazala se varljivim. Problem 1 ima jedno i samo jedno rješenje, problem 2 nema niti jedno, a problem 3 ima beskonačno mnogo rješenja. Način na koji smo dobili ta rješenja (primjenom triju pravila kojima možemo dani sustav zamijeniti ekvivalentnim sustavom) daje nam algoritam za rješavanje sustava triju linearnih jednadžbi s tri nepoznanice. U općem slučaju, primjenom spomenutih pravila, polazni sustav se svodi na sustav jednostavnijeg oblika koji možemo izravno riješiti. Npr. jedan takav sustav je onaj u kojem se u drugoj jednadžbi ne pojavljuje nepoznanica x_1, u trećoj nepoznanice x_1 i x_2 itd., u zadnjoj jednadžbi se pojavljuje samo jedna (ona s najvećim indeksom) nepoznanica. Metoda koja svodi polazni sustav na takav sustav zove se Gaussova metoda eliminacija.

U sljedećoj točki počinjemo studij linearne algebre. Kao što je gotovo svaki puta slučaj, novoj temi je potreban novi sustav notacije. Notacija treba olakšati čitljivost i razumljivost teksta.

Zadaci

1. Koristeći tri "osnovna" pravila svedite sustav jednadžbi na ekvivalentan sustav koji ima **gornju trokutastu formu:**

$$\begin{aligned} a_{11}x + a_{12}y + a_{13}z &= b_1 \\ a_{22}y + a_{23}z &= b_2 \\ a_{33}z &= b_3\,. \end{aligned}$$

Navedite korišteno pravilo u svakom koraku i nađite sva moguća rješenja za slijedeće sustave:

(a)

$$\begin{aligned} x_1 + x_2 + x_3 &= 4 \\ x_1 - x_2 + x_3 &= 2 \\ 2x_1 + x_2 - x_3 &= 1 \end{aligned}$$

(b)

$$\begin{aligned} x_1 + 2x_2 - x_3 &= 5 \\ 2x_1 - 3x_2 + x_3 &= 0 \\ 3x_1 + 16x_2 + x_3 &= 21 \end{aligned}$$

(c)

$$\begin{aligned} x_1 + 2x_2 + x_3 &= 2 \\ 2x_1 - 5x_2 + x_3 &= 11 \\ 3x_1 + 2x_2 - x_3 &= -10 \end{aligned}$$

(d)

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 1 \\ 2x_1 + x_2 - x_3 &= 2 \\ 5x_2 - 3x_3 &= 0 \end{aligned}$$

(e)

$$\begin{aligned} x_1 + x_2 + x_3 &= 1 + 4i \\ -x_1 + x_2 - x_3 &= -1 \\ x_1 - x_2 - x_3 &= -(1 + 2i) \end{aligned}$$

(f)

$$\begin{aligned} ix_1 + 3x_2 - 2ix_3 &= 8 + 3i \\ 2x_1 + (-1+i)x_2 + (2+3i)x_3 &= -11 + 8i \\ x_1 + x_2 + x_3 &= 1 + 5i \end{aligned}$$

2. Odredite α tako da svaki sustav ima najmanje jedno rješenje i nađite sva rješenja sustava.

(a)

$$\begin{aligned} x_1+3x_2-x_3 &= 1\\ 2x_1-x_2+x_3 &= \alpha+2\\ -x_1+11x_2-5x_3 &= 5 \end{aligned}$$

(b)

$$\begin{aligned} 2x_1-x_3 &= -1\\ x_1+3x_2-14x_3 &= \alpha-25\\ x_1-x_2+4x_3 &= 11 \end{aligned}$$

(c)

$$\begin{aligned} 2x_1+x_2+x_3 &= 6\\ x_1-x_2-x_3 &= -3\\ x_1 &= \alpha \end{aligned}$$

(d)

$$\begin{aligned} 3x_1+9x_2-15x_3 &= 6\\ 2x_1+6x_2-10x_3 &= 4\\ 4x_1+12x_2-20x_3 &= 1+\alpha \end{aligned}$$

(e)

$$\begin{aligned} x_1+x_2+2x_3 &= 2+3i\\ x_1-x_2+x_3 &= 1-\alpha i\\ 4x_1-3x_2+5x_3 &= 5 \end{aligned}$$

(f)

$$\begin{aligned} 2x_1+3x_2+x_3 &= -5+4i\\ ix_1+x_2-x_3 &= -(3+2i)\\ x_1+\tfrac{4}{2+i}x_3 &= \alpha\tfrac{4-i}{2+i} \end{aligned}$$

3. Dokažite da ovi sustavi nemaju rješenja:

(a)

$$\begin{aligned} x_1+x_2-3x_3 &= 1\\ 2x_1-x_2+x_3 &= 2\\ -3x_2+7x_3 &= 3 \end{aligned}$$

(b)

$$\begin{aligned} x_1+x_2-3x_3 &= 1\\ 2x_1-x_2+x_3 &= 2\\ 3x_1-2x_3 &= 1 \end{aligned}$$

(c)

$$\begin{aligned} 2x_1+3x_2-x_3 &= 0\\ x_1-3x_2+x_3 &= 1\\ -3x_2+x_3 &= 3 \end{aligned}$$

(d)

$$\begin{aligned} -x_1+2x_2-x_3 &= 0\\ x_1+8x_2+x_3 &= 9\\ x_1+3x_2+x_3 &= 5 \end{aligned}$$

(e)

$$\begin{aligned} x_1-x_2+3x_3 &= 1\\ x_1-5x_2-x_3 &= 4\\ 2x_1-6x_2+2x_3 &= 6 \end{aligned}$$

(f)

$$\begin{aligned} 2x_1-11x_2+11ix_3 &= 10\\ 2ix_1-3x_2+ix_3 &= 7\\ (1-i)x_1-4x_2+5ix_3 &= 1 \end{aligned}$$

4. Konstruirajte sustav od tri jednadžbe s tri nepoznanice tako da ima

(a) jedno i samo jedno rješenje,

(b) više od jednog rješenja,

(c) niti jedno rješenje.

U svakom od tri slučaja reducirajte sustav na ekvivalentan sustav u gornjoj trokutastoj formi.

Uvod u vektorske prostore

U ovom poglavlju prvo upoznajemo specijalne vektorske prostore uređenih n-torki realnih i kompleksnih brojeva, a zatim definiramo vektorski prostor kao algebarsku strukturu. Kao što smo napomenuli ranije, prvo uvodimo oznake. Sa **R** i **C** označavamo skupove realnih odnosno kompleksnih brojeva. Realne i kompleksne brojeve označavamo malim grčkim ili arapskim slovima, α, β, a, b, ... Realne i kompleksne brojeve još ćemo zvati (realni ili kompleksni) **skalari**. Skup svih cijelih brojeva je **Z**, a prirodnih brojeva (tu su cijeli pozitivni brojevi) je **N**.

Neka je $n \in \mathbf{N}$. **Uređena n-torka** skalara je svaki niz od n skalara. Uređene n-torke ćemo pisati tako da niz skalara omeđimo parom zagrada. Npr. ako su $x_1, x_2, \ldots, x_n$ zadani skalari, tada će $(x_1, x_2, \ldots, x_n)$ označavati jednu uređenu n-torku tih skalara. Pritom zagrade ukazuju da je x_1 na prvom mjestu u toj n-torki, x_2 na drugom mjestu, itd. Uočimo da je $(x_2, x_1, x_3, \ldots, x_n)$ jedna druga uređena n-torka istih skalara.

Sa V_n (ili $\mathbf{R}_n$) ćemo označiti skup svih uređenih n-torki realnih brojeva. Riječ "svih" je važna jer ukazuje da V_n uključuje baš sve n-torke, od kojih je svaka oblika $x = (x_1, x_2, \ldots, x_n)$, pri čemu je svaki x_i realni broj. Same n-torke označavamo malim arapskim slovima. Ako pišemo $x, y, z \in \mathrm{V}_n$ to će značiti da su x, y i z uređene n-torke realnih brojeva. Njihove komponente će biti x_i, y_i, z_i, respektivno.

Slično, n-torke kompleksnih brojeva ćemo isto tako označiti malim arapskim slovima, ali ćemo naznačiti da su im komponente kompleksni brojevi. Skup svih takovih n-torki ćemo označiti sa $\mathbf{C}_n$. Npr. $z = (z_1, z_2, \ldots, z_n)$ i $w = (w_1, w_2, \ldots, w_n)$ su iz $\mathbf{C}_n$, ako su $z_i, w_i \in \mathbf{C}$ za svako $1 \leq i \leq n$.

U daljem tekstu ćemo često uređene n-torke skalara kraće nazivati

n-torkama skalara.

Opće vektorske prostore ćemo također označavati velikim, a njihove elemente – vektore, malim arapskim slovima.

Skupove i podskupove ćemo označavati velikim arapskim ili kaligrafskim slovima.

2.1 Vektorski prostor V_n

Ako imamo dvije n-torke $x = (x_1, x_2, \ldots, x_n)$ i $y = (y_1, x_y, \ldots, y_n)$ iz V_n, postavlja se pitanje što bi moglo značiti $x = y$? Ako se n-torke razlikuju u nekoj komponenti na istom mjestu, sigurno ne možemo prihvatiti da su jednake. Zato definiramo da je $x = y$ ako je $x_i = y_i$ za sve indekse $1 \leq i \leq n$. To znači da jednakost $x = y$ ima isto značenje kao i n jednakosti

$$x_1 = y_1, \quad x_2 = y_2, \quad \ldots \quad x_n = y_n. \tag{2.6}$$

Sljedeće bismo željeli uvesti neke operacije nad n-torkama. Počnimo od zbrajanja. Ako su x i y kao gore, tada definiramo $x + y$ kao n-torku $z = (z_1, z_2, \ldots, z_n)$ sa svojstvom

$$z_1 = x_1 + y_1, \quad z_2 = x_2 + y_2, \quad \ldots \quad z_n = x_n + y_n. \tag{2.7}$$

Kraće, to pišemo

$$z = x + y.$$

Jednako kao što nema smisla pitati da li je $x = y$ kad x i y nisu iz istog skupa V_n (npr. $x \in \mathrm{V}_n$, $y \in \mathrm{V}_m$, $n \neq m$), tako nema smisla niti pitati što je $x + y$ za $x \in \mathrm{V}_n$, $y \in \mathrm{V}_m$, $n \neq m$. Operacija zbrajanja n-torki ima jedno važno svojstvo: za svako $x \in \mathrm{V}_n$ i svako $y \in \mathrm{V}_n$ vrijedi $x + y \in \mathrm{V}_n$, jer je svaka komponenta od $x + y$ kao zbroj realnih brojeva realni broj. Dakle polazeći od proizvoljnih elemenata iz V_n, njihova suma je ponovo u V_n. Zato se kaže da je skup V_n zatvoren s obzirom na operaciju zbrajanja (elemenata iz V_n). Da bi naglasili činjenicu da je na V_n definirana operacija zbrajanja i da je V_n zatvoren s obzirom na nju, često se piše $(\mathrm{V}_n, +)$.

Sljedeća operacija koju možemo jednako jednostavno definirati na V_n je operacija množenja n-torki realnim brojevima. Ako je $x \in \mathrm{V}_n$ i

$\alpha \in \mathbf{R}$, tada je αx n-torka $z = (z_1, z_2, \ldots, z_n)$ za koju vrijedi

$$z_1 = \alpha x_1, \quad z_2 = \alpha x_2, \quad \ldots \quad z_n = \alpha x_n. \tag{2.8}$$

Kraće relaciju (2.8) pišemo

$$z = \alpha x.$$

Iz definicije množenja sa skalarom vidimo da za svaki $\alpha \in \mathbf{R}$ i svaki $x \in \mathrm{V}_n$, αx ima kao komponente realne brojeve (jer je produkt realnih brojeva opet realni broj). Drugim riječima, V_n je zatvoren i u odnosu na množenje skalarom. Da bi to naglasili pišemo $(\mathrm{V}_n\,,\,\cdot)$. Ako promatramo V_n zajedno sa obje operacije $+$ i $\cdot$, tada notacija $(\mathrm{V}_n +\,,\,\cdot)$ naglašava činjenicu da je V_n zatvoren s obzirom na obje operacije. Treba uočiti da kod operacije množenja sa skalarom, n-torku množimo nečim što nije n-torka, dakle nečim što ne pripada skupu V_n. Ako želimo naglasiti skupove kojima pripadaju operandi koji sudjeluju u operacijama onda pišemo

$$+ \,:\, \mathrm{V}_n \times \mathrm{V}_n \mapsto \mathrm{V}_n\,, \qquad \cdot \,:\, \mathbf{R} \times \mathrm{V}_n \mapsto \mathrm{V}_n.$$

Zbrajanje n-torki i njihovo množenje skalarom konstruirane su na bazi operacija zbrajanja i množenja realnih brojeva. Svojstva realnih brojeva u odnosu na operacije zbrajanja i množenja omogućavaju izvesti neka svojstva skupa $(\mathrm{V}_n +\,,\,\cdot)$.

Teorem 2.1 *Vrijede sljedeće tvrdnje*

(i) $x + y = y + x$ *za sve* $x, y \in V_n$, *(komutativnost)*

(ii) $x+(y+z) = (x+y)+z$ *za sve* $x, y, z \in V_n$ *(asocijativnost).*

(iii) *Postoji (neutralan) element* $o \in V_n$ *takav da je* $x + o = x$ *za svaki* $x \in V_n$.

(iv) *Za svaki* $x \in V_n$ *postoji inverzni (ili suprotni) element* $-x \in V_n$ *za koji vrijedi* $x + (-x) = o$.

(v) $\alpha(x + y) = \alpha x + \alpha y$ *za sve* $x, y \in V_n$, $\alpha \in \mathbf{R}$ *(distributivnost množenja prema zbrajanju u* V_n*),*

(vi) $(\alpha+\beta)x = (\alpha+\beta)x$ *za sve* $x \in V_n$, $\alpha, \beta \in \mathbf{R}$
(distributivnost množenja prema zbrajanju u $\mathbf{R}$*),*

(vii) $\alpha(\beta x) = (\alpha\beta)x$ *za sve* $x \in V_n$, $\alpha, \beta \in \mathbf{R}$
(kompatibilnost množenja),

(viii) $1\,x = x$ *za svoko* x *(netrivijalnost množenja).*

Dokaz: Dokazi svih tvrdnji su vrlo jednostavni. Za primjer, dokazat ćemo tvrdnje (i) i (v). Za dokaz prve tvrdnje uzmimo proizvoljne $x, y \in \mathrm{V}_n$ i iskoristimo komutativnost zbrajanja realnih brojeva.

$$\begin{aligned} x+y &= (x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) \\ &= (x_1+y_1, x_2+y_2, \ldots, x_n+y_n) \\ &= (y_1+x_1, y_2+x_2, \ldots, y_n+x_n) \\ &= (y_1, y_2, \ldots, y_n) + (x_1, x_2, \ldots, x_n) = y+x. \end{aligned}$$

Za dokaz pete tvrdnje uzmimo proizvoljne $x, y \in \mathrm{V}_n$ i proizvoljni $\alpha \in \mathbf{R}$, te iskoristimo distributivnost množenja u odnosu na zbrajanje realnih brojeva,

$$\begin{aligned} \alpha(x+y) &= \alpha(x_1+y_1, x_2+y_2, \ldots, x_n+y_n) \\ &= (\alpha(x_1+y_1), \alpha(x_2+y_2), \ldots, \alpha(x_n+y_n)) \\ &= (\alpha x_1+\alpha y_1, \alpha x_2+\alpha y_2, \ldots, \alpha x_n+\alpha y_n)) \\ &= (\alpha x_1, \alpha x_2, \ldots, \alpha x_n) + (\alpha y_1, \alpha y_2, \ldots, \alpha y_n) \\ &= \alpha(x_1, x_2, \ldots, x_n) + \alpha(y_1, x_y, \ldots, y_n) = \alpha x + \alpha y. \qquad \blacksquare \end{aligned}$$

Uočimo da postoji samo jedan neutralni element $o = (0, 0, \ldots, 0)$ i da za $x \in \mathrm{V}_n$ postoji samo jedan inverzni element $-x = (-x_1, -x_2, \ldots, -x_n)$. Pritom se inverzni element dobije formulom

$$-x = (-1) \cdot x. \tag{2.9}$$

Na slični način možemo definirati dvije operacije i na drugim skupovimaza te pokazati da vrijede svojstva (i)—(viii). Neka je X takav skup. Operaciju koja pridjeljuje paru elemenata $x, y \in X$ element iz X zovimo **zbrajanje** i označimo ju sa $+$, a onu koja paru koji se sastoji od jednog skalara i jednog elementa iz X pridružuje element iz X, zovimo

množenje sa skalarom i označimo ju s $\cdot$. Ako je X zatvoren s obzirom na te dvije operacije i ako te operacije zadovoljavaju svih osam svojstava opisanih u teoremu 2.1, onda $(X, +, \cdot)$ zovemo **vektorski** ili **linearni prostor**. U kraćoj oznaci, kad se operacije podrazumijevaju, sam X će se zvati vektorski prostor. Elementi od X se onda zovu vektori, bez obzira na to što su oni kao matematički objekti.

Primjer 2.2 *Npr. neka $\mathbf{C}_{[a,b]}$ označava skup svih realnih neprekidnih funkcija na intervalu $[a, b]$. Ako za proizvoljne $f, g \in \mathbf{C}_{[a,b]}$ i realni broj $\alpha \in \mathbf{R}$ definiramo $h = f + g$ i $p = \alpha f$ formulama*

$$\begin{aligned} h(t) &= f(t) + g(t), \quad t \in [a, b] \quad i \\ p(t) &= \alpha f(t), \quad t \in [a, b], \end{aligned}$$

tada su h i p također realne neprekidne funkcije na intervalu $[a, b]$. Dakle je $\mathbf{C}_{[a,b]}$ zatvoren s obzirom na operacije zbrajanja i množenja skalarom. Lako se provjeri da operacije $+$ i $\cdot$ zadovoljavaju svih osam svojstava iz teorema 2.1. Dakle je $\mathbf{C}_{[a,b]}$ vektorski prostor, pa se f i g iz $\mathbf{C}_{[a,b]}$ mogu nazivati vektori iz $\mathbf{C}_{[a,b]}$. ■

Da bi razumjeli važnost zatvorenosti vektorskog prostora u odnosu na operacije zbrajanja i množenja skalarom, vratimo se (vektorskom) prostoru V_n. Neka je $x = (x_1, \ldots, x_n) \in \mathrm{V}_n$. Promotrimo sljedeće vektore:

$$\begin{aligned} e_1 &= (1, 0, 0, \ldots, 0) \\ e_2 &= (0, 1, 0, \ldots, 0) \\ e_2 &= (0, 0, 1, \ldots, 0) \\ \\ e_n &= (0, 0, 0, \ldots, 1). \end{aligned} \tag{2.10}$$

Jer je V_n zatvoren s obzirom na operacije zbrajanja i množenja skalarom, znamo da je

$$y = x_1 e_1 + x_2 e_2 + \cdots + x_n e_n \tag{2.11}$$

također element iz V_n. Naime, množenje svakog vektora e_i realnim brojem x_i ne izvodi nas iz skupa V_n, a isto tako bilo koji redoslijed sumiranja tih izmnoženih vektora ne izvodi nas iz V_n. Koristeći definiciju (2.8) množenja i (2.7) zbrajanja n-torki, dobivamo

$$\begin{aligned} y &= x_1e_1 + x_2e_2 + \cdots x_ne_n = (x_1, 0, \ldots, 0) + (0, x_2, \ldots, 0) + \\ &\quad \cdots + (0, 0, \ldots, x_n) = (x_1, \ldots, x_n) = x. \end{aligned}$$

Kako je x bio proizvoljni vektor, zaključujemo da za svaki $x \in \mathrm{V}_n$ vrijedi

$$x = x_1e_1 + x_2e_2 + \cdots x_ne_n. \tag{2.12}$$

Jer se svaki $x \in \mathrm{V}_n$ može prikazati kao suma multipala x_ie_i vektora e_i, kažemo da se svaki vektor iz V_n može prikazati kao **linearna kombinacija** skupa vektora $\{e_1, \ldots, e_n\}$ ili kraće, kao linearna kombinacija vektora $e_1, \ldots, e_n$. Zato kažemo da je V_n **razapet** skupom vektora $\{e_1, \ldots, e_n\}$ ili kraće, vektorima $e_1, \ldots, e_n$.

Vektorski potprostori od V_n

Označimo sa Z skup svih vektora iz V_n koji se mogu zapisati u obliku

$$z = (z_1, z_2, 0, 0, \ldots, 0), \tag{2.13}$$

gdje su z_1 i z_2 proizvoljni realni brojevi. Drugim riječima, Z se sastoji od svih onih n-torki koje imaju svojstvo da su im $n-2$ zadnje komponente nula. Tvrdimo da je skup Z zajedno sa operacijama zbrajanja n-torki i množenja n-torki skalarom, vektorski prostor. Da bi to provjerili, prvo trebamo dokazati da je Z zatvoren u odnosu na obje operacije. Uzmimo zato proizvoljne $u, v \in Z$ i $\alpha \in \mathbf{R}$. Tada je

$$u = (u_1, u_2, 0, \ldots, 0) \;\text{ i }\; v = (v_1, v_2, 0, \ldots, 0),$$

pa je

$$\begin{aligned} u + v &= (u_1 + v_1, u_2 + v_2, 0, \ldots, 0), \\ \alpha u &= (\alpha u_1, \alpha u_2, 0, \ldots, 0). \end{aligned}$$

Dakle i $u + v$ i αu imaju zadnje $n - 2$ komponente nula pa su elementi skupa Z. Time je dokazano da je Z zatvoren u odnosu na dane operacije. Da bi bio vektorski prostor, moraju na skupu Z operacije $+$ i $\cdot$

imati svih osam svojstava iz teorema 2.1. Za svojstva (i),(ii),(v)—(viii) to je jasno jer su elementi iz Z ujedno i elementi iz V_n. Neutralni element $o \in \mathrm{V}_n$ ima sve, pa zato i zadnjih $n-2$ komponenata nula, pa leži u Z. Konačno, suprotni element od $(z_1, z_2, 0, \ldots, 0)$ je $(-z_1, -z_2, 0, \ldots, 0)$ pa je on također u Z. Dakle, $(Z, +, \cdot)$ je vektorski prostor. Kako je on kao skup podskup od V_n, naziva se vektorski (ili linearni) **potprostor** od V_n.

Nije svaki podskup od V_n vektorski potprostor. Npr, ako uzmemo $\mathcal{S} = \{(1, u_2, 0, \ldots, 0); u_2 \in \mathbf{R}\} \subset \mathrm{V}_n$, tada će svaki zbroj dva vektora iz V_n biti oblika $(2, t, 0, \ldots, 0)$, $t \in \mathbf{R}$, pa skup $\mathcal{S}$ nije zatvoren s obzirom na zbrajanje. Na slični način se pokaže da $\mathcal{S}$ nije zatvoren niti u odnosu na množenje skalarom. Još lakše se vidi da $\mathcal{S}$ ne može biti vektorski prostor provjerom da li je nul-element iz V_n u $\mathcal{S}$. Naime, svaki potprostor od V_n mora sadržavati neutralni element iz V_n. Jasno da $\mathcal{S}$ ne sadrži o jer o nema prvu komponentu jedan već nula.

Opišimo sada način kako se iz danog skupa vektora $\{a_1, a_2, \ldots, a_p\} \subset \mathrm{V}_n$ može izgraditi najmanji vektorski potprostor od V_n koji sadrži sve te vektore. Neka je $L(a_1, a_2, \ldots, a_p)$ skup svih linearnih kombinacija vektora $a_1, a_2, \ldots, a_p$. To možemo matematički zapisati kao

$$L(a_1, a_2, \ldots, a_p) = \{\alpha_1 a_1 + \alpha_2 a_2 + \cdots + \alpha_p a_p;\ \alpha_1, \alpha_2, \ldots, \alpha_p \in \mathbf{R}\}.$$

Prvo uočimo da je $L(a_1, a_2, \ldots, a_p) \subseteq \mathrm{V}_n$. Zatim uočimo da zbroj dviju linearnih kombinacije iz $L(a_1, a_2, \ldots, a_p)$,

$$\begin{aligned}(\alpha_1 a_1 + \alpha_2 a_2 + \cdots + \alpha_p a_p)\ &+\ (\beta_1 a_1 + \beta_2 a_2 + \cdots + \beta_p a_p) = \\ &= (\alpha_1 + \beta_1) a_1 + (\alpha_2 + \beta_2) a_1 + \cdots + (\alpha_p + \beta_p) a_1\end{aligned}$$

daje opet linearnu kombinaciju iz $L(a_1, a_2, \ldots, a_p)$. Isto vrijedi i za produkt sa skalarom,

$$\alpha(\beta_1 a_1 + \beta_2 a_2 + \cdots + \beta_p a_p) = (\alpha\beta_1) a_1 + (\alpha\beta_2) a_1 + \cdots + (\alpha\beta_p) a_1.$$

Dakle je $L(a_1, a_2, \ldots, a_p)$ zatvoren u odnosu na operacije zbrajanja i množenja skalarom. Lako se provjeri da $(L(a_1, a_2, \ldots, a_p), +, \cdot)$ zadovoljava sve uvjete nabrojene u teoremu 2.1. Npr. neutralni element se dobije linearnom kombinacijon vektora $a_1, a_2, \ldots, a_p$ u kojoj su svi skalari α_i nule. Inverzni element od $\alpha_1 a_1 + \alpha_2 a_2 + \cdots + \alpha_p a_p$ je $-\alpha_1 a_1 -$

$\alpha_2 a_2 - \cdots - \alpha_p a_p$. Dakle $L(a_1, a_2, \ldots, a_p)$ je vektorski potprostor od V_n. On sadrži vektor a_i jer je $a_i = 0 \cdot a_1 + \cdots + 0 \cdot a_{i-1} + 1 \cdot a_i + 0 \cdot a_{i+1} + \cdots + 0 \cdot a_p$. Kako to vrijedi za svako i, skup vektora $\{a_1, a_2, \ldots, a_p\}$ je sadržan u vektorskom potprostoru $L(a_1, a_2, \ldots, a_p)$. To je ujedno najmanji potprostor od V_n sa tim svojstvom, jer svaki potprostor od V_n koji sadrži $\{a_1, a_2, \ldots, a_p\}$ nužno sadrži i sve linearne kombinacije skupa $\{a_1, a_2, \ldots, a_p\}$ pa sadrži i $L(a_1, a_2, \ldots, a_p)$. Iz definicije se vidi da $L(a_1, \ldots, a_p)$ ovisi o skupu vektora $\{a_1, \ldots, a_p\}$, pa je $L(a_1, \ldots, a_p) = L(a_{\pi(1)}, \ldots, a_{\pi(p)})$ za svaku permutaciju π skupa $\{1, \ldots, p\}$. Zbog te određenosti $L(a_1, a_2, \ldots, a_p)$ se naziva vektorski (linearni) potprostor od V_n razapet skupom vektora $\{a_1, a_2, \ldots, a_p\}$ (ili kraće: razapet vektorima $a_1, a_2, \ldots, a_p$). Ako je $L(a_1, \ldots, a_p) = X$, tada se $\{a_1, a_2, \ldots, a_p\}$ još zove **razapinjući skup** za X.

Zapravo, X ima beskonačno razapinjućih skupova, a zapis $X = L(a_1, \ldots, a_p)$ samo ukazuje da je $\{a_1, a_2, \ldots, a_p\}$ jedan takav skup. Isto vrijedi i za vektorske potprostore, jer su oni sami po sebi vektorski prostori.

Primjer 2.3 *Neka je Z prvi potprostor od V_n kojeg smo upoznali. Tada je $Z = L(e_1, e_2)$. Uočimo da je Z razapet npr. svakim od skupova $\{(\xi, \eta, 0, \ldots, 0),\ (\xi, -\eta, 0, \ldots, 0)\}$ gdje su ξ, η bilo kakvi od nule različiti realni brojevi.* ■

Neka je A podprostor od V_3 razapet vektorima (trojkama) $(1, 1, 0)$, $(-1, 1, 1)$. Tada možemo postaviti pitanje: je li npr. vektor $(3, 1, -1)$ u A ? Iz definicije slijedi da je $(3, 1, -1) \in L((1, 1, 0), (-1, 1, 1))$ ako postoje realni brojevi α_1, α_2 takvi da je

$$(3, 1, -1) = \alpha_1(1, 1, 0) + \alpha_2(-1, 1, 1) = (\alpha_1 - \alpha_2, \alpha_1 + \alpha_2, \alpha_2).$$

Jer lijeva trojka mora biti jednaka desnoj trojci, dobivamo tri jednadžbe sa dvije nepoznanice,

$$\begin{aligned} \alpha_1 - \alpha_2 &= 3 \\ \alpha_1 + \alpha_2 &= 1 \\ \alpha_2 &= -1. \end{aligned}$$

Iz zadnje jednadžbe dobivamo $\alpha_2 = -1$, a iz druge $\alpha_1 = 1 - \alpha_2 = 2$. Uvrštavajući dobivene vrijednosti za α_1, α_2 u prvu jednadžbu dobivamo

$\alpha_1 - \alpha_2 = 2 - (-1) = 3$ pa je i ona zadovoljena. Zaključujemo da je

$$(3, 1, -1) = 2 \cdot (1, 1, 0) + (-1) \cdot (-1, 1, 1),$$

pa je $(3, 1, -1) \in A$. Iz ovog primjera vidimo da za odgovor na upit pripada li neki vektor nekom potprostoru moramo riješiti jedan sustav linearnih jednadžbi gdje se čak broj nepoznanica ne poklapa s brojem jednadžbi. To nas vodi prema problemu rješavanja općeg sustava od m linearnih jednadžbi sa n nepoznanica. Tom problemu ćemo se vratiti kasnije.

Kartezijevi koordinatni sustavi

Geometrijska interpretacija algebarskih koncepata je uvijek korisna, ako je moguća, jer nam omogućava vizualizaciju i problema i rješenja problema. U tu svrhu posebno su važni slučajevi kad je $n = 2$ i $n = 3$. Označimo sa E_2 i E_3 dvodimenzionalne i trodimenzionalne brojevne sustave, poznate pod imenom Kartezijevi koordinatni sustavi.

Točka P u E_3 je označena sa tri realna broja (p_1, p_2, p_3) koji predstavljaju koordinate te točke s obzirom na Kartezijev koordinatni sustav (ako koordinate označimo kao u analitičkoj geometriji s x, y i z, tada su koordinate od P: $x = p_1$, $y = p_2$, $z = p_3$). Sa čime u E_3 povezati vektore iz V_3? Vektor $v = (v_1, v_2, v_3) \in \mathrm{V}_3$ povezujemo (pridružujemo ga s) usmjerenom dužinom u E_3 koja ima početak u ishodištu (točki s koordinatama $(0, 0, 0)$), a kraj u točki s koordinatama (v_1, v_2, v_3). Važno je zapamtiti da usmjerena dužina koja reprezentira neki element iz V_3 uvijek starta iz ishodišta. Zato takove usmjerene dužine još zovemo radij-vektori[1]. Jednakost dvaju elemenata $a, b \in \mathrm{V}_3$ odgovara u E_3 jednakosti pripadnih radij-vektora (oni završavaju u istoj točki). Sumi elemenata $a, b \in \mathrm{V}_3$ odgovara radij vektor koji se dobije kao suma pripadnih radij-vektora pomoću pravila paralelograma. Dakle, ako su $a = (a_1, a_2, a_3)$, $b = (b_1, b_2, b_3)$ onda je $a+b = (a_1+b_1, a_2+b_2, a_3+b_3)$, a pripadni radij-vektori završavaju u točkama s koordinatama (a_1, a_2, a_3), (b_1, b_2, b_3) i $(a_1 + b_1, a_2 + b_2, a_3 + b_3)$, respektivno. Jasno je da suma

[1] Kasnije ćemo uvesti pojam "paralelnih usmjerenih dužina" i "klasa ekvivalencije paralelnih usmjerenih dužina" koje će nam omogućiti da usmjerene dužine odvojimo od ishodišta

dva radij-vektora uvijek leži u ravnini razapetoj radij-vektorima koji ulaze u sumu. Množenje elementa $a \in \mathrm{V}_3$ s brojem α rezultira u produljenju (ako je $\mid \alpha \mid > 1$) ili skraćenju (ako je $\mid \alpha \mid < 1$) radij-vektora koji odgovara vektoru a. Naime, ako je $a = (a_1, a_2, a_3)$ tada je završna točka radij-vektora koji odgovara vektoru a odnosno αa određena koordinatama (a_1, a_2, a_3) odnosno $(\alpha a_1, \alpha a_2, \alpha a_3)$. Odmah vidimo da oba radij-vektora leže na istom pravcu dok im je smjer isti (suprotan) ako je $\alpha > 0$ ($\alpha < 0$). Nulvektor u V_3 odgovara ishodištu, a to se događa u ovom slučaju kad je $\alpha = 0$.

S obzirom da je svaki potprostor od V_3 zatvoren u odnosu na operacije zbrajanja i množenja skalarom, te na osnovu pravila za množenje radij-vektora skalarom i za zbrajanje radij-vektora, zaključujemo da

- 0-dimenzionalnom potprostoru u V_3 odgovara u E_3 ishodište;
- 1-dimenzionalnom potprostoru u V_3 odgovara u E_3 pravac koji prolazi ishodištem;
- 2-dimenzionalnom potprostoru u V_3 odgovara u E_3 ravnina koja prolazi ishodištem;
- 3-dimenzionalnom potprostoru u V_3 odgovara u E_3 cijeli prostor.

2.2 Vektorski prostor $\mathbf{C}_n$

Sa $\mathbf{C}_n$ označimo skup svih n-torki kompleksnih brojeva. Zbrajanje takovih n-torki definiramo na isti način kao i u slučaju realnih n-torki. Kod množenja skalarom dopuštamo da skalar bude kompleksan broj. Dakle, ako su $z = (z_1, z_2, \ldots, z_n), w = (w_1, w_2, \ldots, w_n) \in \mathbf{C}_n$ i $\zeta \in \mathbf{C}$ proizvoljni, tada je

$$z + w = (z_1 + w_1, z_2 + w_2, \ldots, z_n + w_n), \quad \zeta z = (\zeta z_1, \zeta z_2, \ldots, \zeta z_n).$$

Odmah se vidi da je $\mathbf{C}_n$ zatvoren u odnosu na operaciju zbrajanja n-torki i množenja n-torki skalarom. Također se lako provjeri da operacije $+$ i $\cdot$ na $\mathbf{C}_n$ zadovoljavaju svih osam uvjeta iz teorema 2.1. Stoga je $(\mathbf{C}_n, +, \cdot)$ vektorski prostor. Kako komponente n-torki i skalari koji ih množe pripadaju skupu **C**, $(\mathbf{C}_n, +, \cdot)$ zovemo vektorski prostor n-torki nad **C** ili vektorski prostor kompleksnih n-torki. Kao i prije, jedinstveni neutralni element je $o = (0, 0, \ldots, 0)$, a inverzni element od z je $-z = (-z_1, -z_2, \ldots, -z_n)$. Linearna kombinacija vektora $a_1, \ldots, a_k \in \mathbf{C}_n$ je vektor $z = \alpha_1 a_1 + \alpha_2 a_2 + \cdots + \alpha_k a_k$, pri čemu su skalari $\alpha_1, \ldots, \alpha_k$ općenito kompleksni. Skup svih linearnih kombinacija vektora $a_1, \ldots, a_k$ označavamo s $L(a_1, \ldots, a_k)$. Taj podskup od $\mathbf{C}_n$ opskrbljen operacijama $+$ i $\cdot$ iz $\mathbf{C}_n$ čini vektorski prostor koji je potprostor od $(\mathbf{C}_n, +, \cdot)$.

Osim standardnih binarnih operacija $+$ i $\cdot$, na $\mathbf{C}_n$ je prirodno definirana unarna operacija kompleksne konjugacije $^-$ koja svakoj n-torki $z = (z_1, z_2, \ldots, z_n) \in \mathbf{C}_n$ pridružuje n-torku $\bar{z} = (\bar{z}_1, \bar{z}_2, \ldots, \bar{z}_n)$. Vektor $\bar{z} \in \mathbf{C}_n$ se zove (kompleksno) konjugirani vektor z.

2.3 Opći vektorski prostor

Vektorski prostori V_n i $\mathbf{C}_n$ sagrađeni su nad skupovima **R** i **C** koji imaju posebno svojstvo: na njima su definirane operacije zbrajanja i množenja. Te operacije imaju karakterisike koje daju skupovima **R** i **C** posebnu matematičku strukturu koja se zove **polje**. Označimo sa Φ **R** ili **C**. Tada Φ ima sljedeća svojstva:

1. $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$ za sve $\alpha, \beta, \gamma \in \Phi$
(asocijativnost zbrajanja).

2. Postoji (neutralan za zbrajanje) element $0 \in \Phi$ takav da je $\alpha + 0 = 0 + \alpha = \alpha$ za svako $\alpha \in \Phi$.

3. Za svaki $\alpha \in \Phi$ postoji inverzni (u odnosu na zbrajanje) element $-\alpha$ za koji vrijedi: $\alpha + (-\alpha) = (-\alpha) + \alpha = 0$.

4. $\alpha + \beta = \beta + \alpha$ za sve $\alpha, \beta \in \Phi$, (komutativnost zbrajanja).

5. $(\alpha\beta)\gamma = \alpha(\beta\gamma)$ za bilo koje $\alpha, \beta, \gamma \in \Phi$,
(asocijativnost množenja u Φ),

6. Postoji (neutralan za množenje) element $1 \in \Phi$, takav da je $\alpha \cdot 1 = 1 \cdot \alpha = \alpha$ za svako $\alpha \in \Phi$.

7. Za svaki $\alpha \in \Phi$ koji nije 0, postoji inverzni (u odnosu na množenje) element α^{-1} za koji vrijedi: $\alpha\alpha^{-1} = \alpha^{-1}\alpha = 1$.

8. $\alpha(\beta+\gamma) = \alpha\beta + \alpha\gamma$ za sve $\alpha, \beta, \gamma \in \Phi$ (lijeva distributivnost)
$(\beta+\gamma)\alpha = \beta\alpha + \gamma\alpha$ za sve $\alpha, \beta, \gamma \in \Phi$ (desna distributivnost)

9. $\alpha\beta = \beta\alpha$ za sve $\alpha, \beta \in \Phi$ (komutativnost množenja).

2.3.1 Grupa

Zamislimo sada da je Φ neki skup na kojem je definirana samo jedna binarna operacija koju označimo sa $\circ$. Dakle $\circ$ pridružuje svakom paru elemenata iz Φ opet element iz Φ i pritom je Φ zatvoren s obzirom na operaciju $\circ$. Ako pritom vrijede svojstvo asocijativnosti, svojstvo postojanja neutralnog elementa i ako za svaki element od Φ postoji inverzni element, onda se $(\Phi, \circ)$ zove **grupa**. Ako je operacija $\circ$ slična operaciji zbrajanja (množenja) onda grupa dobiva pridjev aditivna (multiplikativna). Ako za $(\Phi, \circ)$ još vrijedi i svojstvo 4., onda se zove Abelova ili komutativna grupa.

Neka je $\mathcal{S} \subset \Phi$ neki podskup skupa Φ. Ako su $x, y \in \mathcal{S}$, tada je očito $x \circ y \in \Phi$. Ako međutim za svaka dva elementa $x, y \in \mathcal{S}$ vrijedi $x \circ y \in \mathcal{S}$, tada je $\mathcal{S}$ zatvoren u odnosu na operaciju $\circ$. Za svaki element x iz $\mathcal{S}$ postoji inverzni element x^{-1} u Φ i on nije nužno u $\mathcal{S}$. Ako je $\mathcal{S}$ zatvoren u odnosu na $\circ$, ako je neutralni element grupe Φ u $\mathcal{S}$ i ako je za svaki element x od $\mathcal{S}$, x^{-1} također u $\mathcal{S}$, tada se skup $\mathcal{S}$ zajedno sa (naslijeđenom) operacijom $\circ$ zove *podgrupa* grupe Φ. Dakle, da bi bio podgrupa od Φ, podskup $\mathcal{S}$ mora biti i sam grupa uz istu binarnu operaciju.

Pretpostavimo da imamo dvije grupe $(G_1, \circ)$, $(G_1, \diamond)$ i preslikavanje $f: G_1 \mapsto G_2$. Ako je f surjekcija (tj. $f(G_1) = G_2$) i vrijedi

$$f(g_1 \circ g_2) = f(g_1) \diamond f(g_2) \quad \text{za svako } g_1,\, g_2 \in G_1$$

tada se f zove izomorfizam grupa G_1 i G_2. Lako se pokaže da izomorfizam preslikava neutralni element grupe G_1 u neutralni element grupe G_2 i inverzni element g^{-1} od $g \in G_1$ prebacuje u inverzni element $[f(g)]^{-1}$ elementa $f(g)$. Stoga izomorfne grupe imaju istovjetnu strukturu.

Neka osnovna svojstva grupe, kao na primjer jedinstvenost neutralnog elementa te za svaki element grupe jedinstvenost njemu inverznog elementa, bit će dokazani u sklopu dokaza propozicije 2.8.

Grupa permutacija

Lijep primjer grupe je tzv. grupa permutacija.

Neka je N_n neki konačan skup od n elemenata. *Permutacija* p je svaka bijekcija skupa N_n na sebe. Skup svih permutacija skupa N_n označavamo sa Π_n.

Radi jednostavnosti razmatranja, umjesto sa općim skupom N_n radit ćemo sa sa skupom

$$S_n = \{1, 2, \dots, n\}\,.$$

Permutaciju skupa S_n označavamo sa

$$p = \begin{pmatrix} 1 & 2 & \dots & n \\ p(1) & p(2) & \dots & p(n) \end{pmatrix}.$$

Na primjer, ako je $n = 5$, imamo $S_n = \{1, 2, 3, 4, 5\}$ i

$$p = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 3 & 1 & 2 & 4 \end{pmatrix}. \tag{2.14}$$

je jedna permutacija. Na pitanje koliko ima takovih permutacija odgovara

Propozicija 2.4 *Skup Π_n ima $n! = 1 \cdot 2 \cdot \ldots \cdot n$ elemenata.*

Dokaz: Dokaz provodimo "sortiranjem" svih mogućih permutacija skupa S_n. Sve permutacije najprije rasporedimo u n grupa, tako da su u prvoj grupi sve permutacije p koje $1 \in S_n$ prevode u sebe: $p(1) = 1$, u drugoj grupi su sve permutacije za koje vrijedi $p(1) = 2$ itd. Svaku od tih "velikih" grupa dijelimo na $n-1$ manjih grupa koje su karakterizirane zahtjevima oblika $p(2) = 1$, $p(2) = 2$ itd., s tim da ako je u pripadnoj velikoj grupi bilo npr. $p(1) = 1$, onda $p(2) = 1$ otpada (jer su sve permutacije bijekcije). Ukupno sada imamo $n(n-1)$ grupa karakteriziranih svim mogućim preslikavanjima elemenata 1 i 2.

Postupak nastavljamo dijeljenjem grupa na podgrupe prema kriteriju kamo se preslikava element 3. Tako dobivamo $n(n-1)(n-2)$ manjih grupa karakteriziranih svim mogućim preslikavanjima elemenata 1, 2 i 3. Postupak nastavljamo sve dok na kraju ne ostane u svakoj grupi točno jedna permutacija, a takovih jednočlanih grupa ima $n \cdot (n-1) \cdot \cdots 2 \cdot 1 = n!$. ∎

Produkt permutacija definiramo kao kompoziciju funkcija. Npr. ako su $p, q \in \Pi_n$, tada je produkt $q \circ p$ definiran relacijom

$$(q \circ p)(i) = q(p(i)), \quad 1 \leq i \leq n.$$

Ako je npr. p kao u relaciji (2.14), a

$$q = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 4 & 3 & 2 & 1 \end{pmatrix},$$

tada je

$$\begin{aligned} q \circ p &= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 4 & 3 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 3 & 1 & 2 & 4 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 3 & 5 & 4 & 2 \end{pmatrix}. \end{aligned}$$

Teorem 2.5 $(\Pi_n, \circ)$ *je multiplikativna grupa.*

Dokaz: Kompozicija dviju bijekcija je opet bijekcija, pa je produkt permutacija opet permutacija. Svojstvo asocijativnosti slijedi iz činjenice da asocijativnost vrijedi općenito za kompoziciju triju funkcija. Neutralni element e je očito definiran s $e(i) = i$ za sve $1 \leq i \leq n$, dok za inverzni element vrijedi

$$\begin{pmatrix} 1 & 2 & \dots & n \\ p(1) & p(2) & \dots & p(n) \end{pmatrix}^{-1} = \begin{pmatrix} p(1) & p(2) & \dots & p(n) \\ 1 & 2 & \dots & n \end{pmatrix}, \quad (2.15)$$

pa preostaje samo gornji i donji red u oznaci za permutaciju na desnoj strani jednako ispresložiti, tako da prvi red postane $1\ 2 \cdots n$. ∎

Npr. za p iz relacije (2.14) imamo

$$p^{-1} = \begin{pmatrix} 5 & 3 & 1 & 2 & 4 \\ 1 & 2 & 3 & 4 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 4 & 2 & 5 & 1 \end{pmatrix}.$$

Ako za bilo koju permutaciju p definiramo

$$\begin{aligned} p \circ \Pi_n &= \{q;\ q = p \circ s,\ s \in \Pi_n\} \\ \Pi_n \circ p &= \{q;\ q = s \circ p,\ s \in \Pi_n\}, \end{aligned}$$

onda iz teorema 2.5 slijedi

$$p \circ \Pi_n = \Pi_n \circ p = \Pi_n \qquad \text{za svako } p \in \Pi_n. \tag{2.16}$$

Dakle, ako q prolazi cijelom grupom permutacija, isto čini i $p \circ q$ odnosno $q \circ p$.

Promotrimo specijalne permutacije

$$p_{ij} = \begin{pmatrix} 1 & \dots & i & \dots & j & \dots & n \\ 1 & \dots & j & \dots & i & \dots & n \end{pmatrix}, \quad 1 \leq i < j \leq n \tag{2.17}$$

koje zovemo transpozicije (ili zamjene). Za njih vrijedi $p_{ij} \circ p_{ij} = e$, pa je

$$p_{ij}^{-1} = p_{ij}\,. \tag{2.18}$$

Može se pokazati da se svaka permutacija može prikazati kao produkt transpozicija.

2.3.2 Prsten, tijelo, polje

Neka je sada Φ skup na kome su definirane dvije binarne operacije koje označimo sa $+$ i $\cdot$. Ako $(\Phi, +, \cdot)$ zadovoljava uvjete 1.—4., 5. i 8., tada se zove **prsten**. Dakle, $(\Phi, +, \cdot)$ je prsten ako je aditivna Abelova grupa na kojoj je definirano množenje elemenata koje zadovoljava svojstvo asocijativnosti i obostrane distributivnosti. Ako još vrijedi uvjet 6., tada govorimo o prstenu sa jedinicom. Ako vrijedi uvjet 9., tada je prsten komutativan.

Ako za $(\Phi, +, \cdot)$ vrijede svojstva 1.—8., tada se takova algebarska struktura zove **tijelo**. Dakle tijelo ima to svojstvo da je $(\Phi, +)$ aditivna Abelova grupa, $(\tilde{\Phi}, \cdot)$ je multiplikativna grupa, a vrijedi i svojstvo obostrane disrtibutivnosti. Ovdje je $\tilde{\Phi} = \Phi \setminus \{0\}$ skup Φ bez neutralnog elementa za zbrajanje. Ako je još $(\tilde{\Phi}, \cdot)$ komutativna multiplikativna grupa, onda se $(\Phi, +, \cdot)$ zove **polje**.

Primjer 2.6 Neka je $\mathcal{P}$ skup svih polinoma realne varijable s realnim koeficijentima. Uvedimo u taj skup operaciju zbrajanja. Za dva polinoma $p, q \in \mathcal{P}$,

$$\begin{aligned} p(x) &= a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0, \\ q(x) &= b_m x^m + b_{m-1}x^{m-1} + \cdots + b_1 x + b_0 \end{aligned}$$

zbroj $p+q$ je polinom $s \in \mathcal{P}$ koji ćemo definirati na sljedeći način.
Ako je $n > m$, napišimo q u obliku

$$q(x) = b_n x^n + b_{n-1}x^{n-1} + \cdots + b_m x^m + b_{m-1}x^{m-1} + \cdots + b_1 x + b_0,$$

gdje su $b_n = b_{n-1} = \cdots = b_{m+1} = 0$.
Ako je $m \geq n$, napišimo p u obliku

$$p(x) = a_m x^m + \cdots + a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0,$$

gdje su $a_m = \cdots = a_{n+1} = 0$ (ako je $m = n$, p ne mijenjamo).
Tada je s polinom stupnja $r = \max\{m, n\}$, za koji vrijedi

$$s(x) = (a_r + b_r)x^r + (a_{r-1} + b_{r-1})x^{r-1} + \cdots + (a_1 + b_1)x + (a_0 + b_0)$$

Pokažite da za tako definiranu operaciju zbrajanja polinoma vrijede svojstva asocijativnosti i komutativnosti. Neutralni element je nula-polinom koji ima sve koeficijente nula. Inverz od p je $-p$ ima sve koeficijente suprotnog predznaka od koeficijenata od p. Stoga je $(\mathcal{P}, +)$ aditivna Abelova grupa.

Definirajmo i množenje polinoma: $t = p \cdot q$, gdje su p i q kao prije,

$$\begin{aligned} t(x) &= (a_m b_n)x^{m+n} + (a_m b_{n-1} + a_{m-1}b_n)x^{m+n-1} + \cdots \\ &\quad + (a_1 b_0 + a_0 b_1)x + a_0 b_0. \end{aligned}$$

Pritom smo svaki član polinoma p pomnožili sa svakim članom polinoma q, tako dobijene članove zbrojili i posložili tako da potencije od x padaju.

Pokažite da je $(\mathcal{P}, +, \cdot)$ komutativni prsten s jedinicom. ■

Primjeri algebarskih struktura, grupa, prstenova, tijela i polja dani su kroz zadatke. Uočimo da su $\mathbf{R}$ i $\mathbf{C}$ polja. Zato govorimo da je V_n vektorski prostor nad poljem realnih brojeva, dok je $\mathbf{C}_n$ vektorski prostor nad poljem kompleksnih brojeva.

2.3.3 Apstraktni vektorski prostor

Općenito se vektorski prostor definira nad proizvoljnim tijelom. Dakle, za definiciju vektorskog prostora moramo imati dva skupa: skup elemenata koje zovemo vektori (u dosadašnjim primjerima n-torke) i skup elemenata koje zovemo skalari. Skup vektora označimo sa X, a skup skalara sa Φ. Skup Φ mora biti tijelo (ili polje), dakle moraju biti definirane operacije zbrajanja i množenja skalara sa posebnim svojstvima. Na skupu X moraju također biti definirane dvije operacije: zbrajanje vektora koje pretvara X u aditivnu Abelovu grupu i množenje vektora skalarom koje mora zadovoljavati uvjete 5.—8. iz teorema 2.1. Da u ovim uvodnim razmatranjima apstraktnog vektorskog prostora ne bi došlo do nejasnoća, operacije na Φ ćemo označiti sa $+$ i $\cdot$, a operacije na X sa $\oplus$ i $\otimes$. Kasnije ćemo $\oplus$ i $\otimes$ zamijeniti s $+$ i $\cdot$.

Definicija 2.7 *Vektorski ili linearni prostor je algebarska struktura koja se sastoji od dva skupa: X čije elemente zovemo vektori i Φ čije elemente zovemo skalari. Na skupu Φ su definirane dvije binarne operacije u oznaci $+$ i $\cdot$ koje čine $(\Phi, +, \cdot)$ tijelom. Na skupu X su definirane operacije zbrajanja vektora $\oplus$ i množenja $\otimes$ vektora skalarom iz Φ. Pritom je $(X, \oplus)$ aditivna Abelova grupa, a za operaciju $\otimes$ vrijede sljedeća svojstva kompatibilnosti:*

(a) $\alpha \otimes (x \oplus y) = \alpha \otimes x \oplus \alpha \otimes y$ *za bilo koje* $x, y \in X$, $\alpha \in \Phi$ *(distributivnost množenja prema zbrajanju u X),*

(b) $(\alpha + \beta) \otimes x = \alpha \otimes x \oplus \beta \otimes x$ *za bilo koje* $x \in X$, $\alpha, \beta \in \Phi$ *(distributivnost množenja prema zbrajanju u Φ),*

(c) $\alpha \otimes (\beta \otimes x) = (\alpha \cdot \beta) \otimes x$ *za bilo koje* $x \in X$, $\alpha, \beta \in \Phi$ *(kompatibilnost množenja),*

(d) *Ako je $1 \in \Phi$ neutralni element za množenje u Φ, tada je*
$1 \otimes x = x$ *za svako* $x \in X$ *(netrivijalnost množenja).*

Pokažimo neka važna svojstva algebarske strukture tijela i vektorskog prostora.

Propozicija 2.8 *Neka je $(X, \oplus, \otimes)$ vektorski prostor nad tijelom $(\Phi, +, \cdot)$.*

(i) *Neutralni elementi za zbrajanje i množenje u Φ su jedinstveni. Neutralni element za zbrajanje u X je jedinstven.*

(ii) *Za svaki $\alpha \in \Phi$ inverzni element od α s obzirom na $+$ je jedinstven. Ako $\beta \in \Phi$ nije neutralan za zbrajanje, njemu inverzni element s obzirom na $\cdot$ je jedinstven. Za svaki $x \in X$ inverzni element s obzirom na $\oplus$ je jedinstven.*

(iii) *Ako je $o \in X$ neutralni element s obzirom na $\oplus$, tada za svako $\alpha \in \Phi$ vrijedi $\alpha \otimes o = o$. Ako je $0 \in \Phi$ neutralni element s obzirom na $+$, tada za svako $\alpha \in \Phi$ vrijedi $\alpha \cdot 0 = 0 \cdot \alpha = 0$.*

(iv) *Ako je $0 \in \Phi$ neutralni element s obzirom na $+$, a $o \in X$ neutralni element s obzirom na $\oplus$, tada za svako $x \in X$ vrijedi $0 \otimes x = o$.*

(v) *Ako je $1 \in \Phi$ neutralni element s obzirom na $\cdot$, a -1 njegov inverzni element u Φ s obzirom na $+$, tada za svako $x \in X$ vrijedi $(-1) \otimes x = -x$, gdje je $-x$ inverzni element od x u X s obzirom na $\oplus$.*

Dokaz: (i) Dovoljno je pokazati da je u proizvoljnoj grupi $(G, \circ)$ neutralan element jedinstven. Pretpostavimo da postoje dva neutralna elementa e i e'. Tada mora vrijediti $e \circ e' = e'$ jer je e neutralan. Ali i e' je neutralan pa mora biti $e \circ e' = e$. Jer su u ovim jednakostima lijeve strane jednake moraju biti i desne, tj. mora biti $e' = e$.
(ii) Dovoljno je pokazati da u proizvoljnoj grupi $(G, \circ)$ svaki element ima jedinstven inverzni element. Pretpostavimo da element $g \in G$ ima dva inverza, g' i g''. Znači, pretpostavljamo da vrijedi $g \circ g' = g' \circ g = e$ i $g \circ g'' = g'' \circ g = e$, gdje je e neutralni element. Koristeći te relacije i svojstvo asocijativnosti odmah dobivamo

$$g'' = g'' \circ e = g'' \circ (g \circ g') = (g'' \circ g) \circ g' = e \circ g' = g'.$$

(iii) Jer je $o \in X$ neutralni element sigurno vrijedi $o = o \oplus o$. Prema svojstvu distributivnosti, za svako α vrijedi

$$\alpha \otimes o = \alpha \otimes (o \oplus o) = \alpha \otimes o \oplus \alpha \otimes o.$$

Dakle, vektor $y = \alpha \otimes o$ ima svojstvo $y = y \oplus y$. Dodajmo lijevoj i desnoj strani inverzni element $-y$ s obzirom na $\oplus$. Dobivamo $o = y$ što

je i trebalo dokazati.
Druga tvrdnja se dokazuje po istom obrascu. Za dokaz $\alpha \cdot 0 = 0$ $(0 \cdot \alpha = 0)$ se koristi lijeva (desna) distributivnost u Φ.
(iv) Ako pišemo $0 = 0 + 0$, tada zakon distributivnosti daje

$$0 \otimes x = (0 + 0) \otimes x = 0 \otimes x \oplus 0 \otimes x$$

pa $y = 0 \otimes x$ zadovoljava jednadžbu $y = y + y$. Kako je pokazano u dokazu tvrdnje (iii), $y = o$.
(v) Koristeći (iv) imamo

$$o = 0 \otimes x = (1 + (-1)) \otimes x = 1 \otimes x \oplus (-1) \otimes x,$$

odakle odmah, prema tvrdnji (ii), zaključujemo da je $(-1) \otimes x$ inverzni element od x s obzirom na $\oplus$. ■

Na vektorskom prostoru X možemo definirati razliku vektora formulom

$$x \ominus y = x \oplus (-1) \otimes y = x \oplus (-y) \tag{2.19}$$

Propozicija 2.9 *Neka je $(X, \oplus, \otimes)$ vektorski prostor nad tijelom $(\Phi, +, \cdot)$. Tada za operaciju odbijanja vrijede formule analogne formulama (a) i (b) iz definicije 2.7,*

$$\begin{aligned}
(\alpha - \beta) \otimes x &= \alpha \otimes x \ominus \beta \otimes x, \quad x \in X, \quad \alpha, \beta \in \Phi & (2.20) \\
\alpha \otimes (x \ominus y) &= \alpha \otimes x \ominus \beta \otimes x, \quad x, y \in X, \quad \alpha, \in \Phi & (2.21)
\end{aligned}$$

Dokaz: Koristeći redom svojstva (b) i (c) iz definicije 2.7 i definiciju razlike (2.19), dobivamo

$$\begin{aligned}
(\alpha - \beta) \otimes x &= (\alpha + (-\beta)) \otimes x = \alpha \otimes x \oplus (-\beta) \otimes x \\
&= \alpha \otimes x \oplus ((-1) \cdot \beta) \otimes x = \alpha \otimes x \oplus (-1) \otimes (\beta \otimes x) \\
&= \alpha \otimes x \ominus \beta \otimes x
\end{aligned}$$

Time je dokazana relacija (2.20). Za dokaz realcije (2.21), koristimo definiciju (2.19) te svojstva (a) i (c) iz definicije 2.7

$$\alpha \otimes (x \ominus y) = \alpha \otimes [x \oplus (-y)] = \alpha \otimes x \oplus \alpha \otimes (-y)$$

pa još treba pokazati da je $\alpha \otimes (-y) = -(\alpha \otimes y)$. Zbog tvrdnje (iii) propozicije 2.8 vrijedi

$$\alpha \otimes (-y) \oplus \alpha \otimes y = \alpha \otimes ((-y) \oplus y) = \alpha \otimes o = o$$

pa zaključujemo da je $\alpha \otimes (-y)$ inverzni element od $\alpha \otimes y$, što je i trebalo dokazati. ■

Pokažimo još da u vektorskom prostoru smijemo vektorske jednadžbe "kratiti" i sa skalarom i vektorom.

Propozicija 2.10 *Neka je $(X, \oplus, \otimes)$ vektorski prostor nad tijelom $(\Phi, +, \cdot)$.*

(i) *Ako je produkt $\alpha \otimes x$ skalara α i vektora x nulvektor i jedan od faktora nije nula, onda drugi faktor mora biti nula.*

(ii) *Ako je u jednadžbi $\alpha \otimes x = \beta \otimes x$, $\alpha, \beta \in \Phi$, $x \in X$, vektor $x \neq o$, tada vrijedi skalarna jednadžba $\alpha = \beta$.*

(iii) *Ako je u jednadžbi $\alpha \otimes x = \alpha \otimes y$, $\alpha \in \Phi$, $x, y \in X$, skalar $\alpha \neq o$, tada vrijedi vektorska jednadžba $x = y$.*

Dokaz: (i) Pokažimo prvo da $\alpha \otimes x = o$ i $\alpha \neq 0$ povlači $x = o$. Zaista, kako u tijelu Φ postoji inverz α^{-1}, dovoljno je pomnožiti jednadžbu $\alpha \otimes x = o$ s α^{-1} i iskoristiti uvjete (c) i (d) iz definicije 2.7 vektorskog prostora i tvrdnju (iii) propozicije 2.8. Pođimo sada od jednadžbe $\alpha \otimes x = o$ i pretpostavke $x \neq o$. Tvrdimo da je $\alpha = 0$. Doista, u protivnom bi $\alpha \neq 0$ i $\alpha \otimes x = o$ implicirale (kako je pokazano) $x = o$, a to se protivi pretpostavci $x \neq o$. Dakle, $\alpha \neq 0$ vodi u kontradikciju pa mora biti $\alpha = 0$.
(ii) Korištenjem propozicije 2.9(ii) imamo $\alpha \otimes x \ominus \beta \otimes x = (\alpha - \beta) \otimes x = o$, $x \neq o$. Sada iskaz slijedi iz tvrdnje (i) ove propozicije.
(iii) Korištenjem propozicije 2.9(i) imamo $\alpha \otimes x \ominus \alpha \otimes y = \alpha \otimes (x - y) = o$, $\alpha \neq 0$, pa iskaz opet slijedi iz tvrdnje (i) ove propozicije. ■

Potprostor vektorskog prostora $(X, \oplus, \otimes)$ je svaki podskup od X koji je uz operacije $\oplus$ i $\otimes$ i sam vektorski prostor nad istim poljem. Da bi to bio, treba a i dovoljno je da je zatvoren u odnosu na operacije $\oplus, \otimes$. Zaista, ako je $(S, \oplus, \otimes)$ potprostor od $(X, \oplus, \otimes)$ (kraće pišemo S

je potprostor od X) tada je nužno (po definiciji vektorskog prostora) zatvoren u odnosu na te operacije. Obratno, ako je podskup $S \subseteq X$ zatvoren u odnosu na operacije $\oplus, \otimes$, tada za njega automatski vrijede (jer je podskup od X) svi uvjeti iz definicije 2.7 pa je $(S, \oplus, \otimes)$ vektorski prostor. Dakle, vektorski potprostor je definiran upravo onim operacijama koje su naslijeđene od prostora. Trivijalni vektorski potprostori su jednočlan skup $\{o\}$ i cijeli X. Da bi neki skup $S \subseteq X$ bio potprostor nužno je i dovoljno da su ispunjeni sljedeći uvjeti:

(a) $\quad x \oplus y \in S$ za svaka dva vektora $x, y \in S$

(b) $\quad \alpha \otimes x \in S$ za sve $\alpha \in \Phi$ i $x \in S$

Naime, to su nužni i dovoljni uvjeti za zatvorenost skupa S u odnsu na operacije $\oplus$ i $\otimes$. Lako se pokaže (vidjeti zadatak ...) da su uvjeti (a) i (b) ekvivalentni uvjetu

(ab) $\quad \alpha \otimes x \oplus \beta \otimes y \in S$ za sve $x, y \in S$ i $\alpha, \beta \in \Phi$.

Neka su $a_1, a_2, \ldots, a_p$ vektori iz $(X, \oplus, \otimes)$. Linearna kombinacija vektora $a_1, a_2, \ldots, a_p$ (ili skupa vektora $\{a_1, a_2, \ldots, a_p\}$) je svaki vektor y oblika $y = \alpha_1 \otimes a_1 \oplus \alpha_2 \otimes a_2 \oplus \cdots \oplus \alpha_p \otimes a_p$, gdje su $\alpha_1, \alpha_2 \ldots, \alpha_p$ skalari iz Φ. Pokazat ćemo da je najmanji vektorski potprostor koji sadrži sve vektore $a_1, \ldots, a_p$, potprostor $(L(a_1, \ldots, a_p), \oplus, \otimes)$. Pritom su elementi od $L(a_1, \ldots, a_p)$ linearne kombinacije skupa $\{a_1, \ldots, a_p\}$,

$$L(a_1, \ldots, a_p) = \{\alpha_1 \otimes a_1 \oplus \cdots \oplus \alpha_p \otimes a_p \,;\, \alpha_1, \ldots, \alpha_p \in \Phi\}.$$

Zbog jednostavnosti pisanja označimo $L(a_1, \ldots, a_p)$ s S. Pokažimo da je S vektorski potprostor od X. Dovoljno je provjeriti uvjete (a) i (b). Ako su $x, y \in S$, tada za neke skalare α_i i β_i, $1 \leq i \leq p$ vrijedi $x = \alpha_1 \otimes a_1 \oplus \cdots \oplus \alpha_p \otimes a_p$ i $y = \beta_1 \otimes a_1 \oplus \cdots \oplus \beta_p \otimes a_p$. Stoga je

$$\begin{aligned}
x + y &= \alpha_1 \otimes a_1 \oplus \cdots \oplus \alpha_p \otimes a_p \quad \oplus \quad \beta_1 \otimes a_1 \oplus \cdots \oplus \beta_p \otimes a_p \\
&= (\alpha_1 + \beta_1) \otimes a_1 \oplus \cdots \oplus (\alpha_p + \beta_p) \otimes a_p \\
\alpha \otimes x &= \alpha \otimes (\alpha_1 \otimes a_1 \oplus \cdots \oplus \alpha_p \otimes a_p) \\
&= (\alpha \cdot \alpha_1) \otimes a_1 \oplus \cdots \oplus (\alpha \cdot \alpha_p) \otimes a_p,
\end{aligned}$$

pa su i $x \oplus y$ i $\alpha \otimes x$ linearne kombinacije skupa $\{a_1, \ldots, a_p\}$, a to znači da su u S. Dakle je S vektorski potprostor od X. I to najmanji od onih koji sadrže $\{a_1, \ldots, a_p\}$ jer svaki potprostor od X koji sadrži $\{a_1, \ldots, a_p\}$ nužno sadrži i sve linearne kombinacije od elemenata $a_1, \ldots, a_p$, pa zato sadrži i S.

U daljem ćemo umjesto oznaka $\oplus$, $\ominus$ i $\otimes$ koristiti jednostavnije oznake $+$, $-$, i $\cdot$, pri čemu ćemo $\cdot$ često izostavljati. Također, kad god su operacije $\oplus$ i $\otimes$ poznate, tj. jasne iz konteksta, umjesto $(X, \oplus, \otimes)$ ćemo pisati X. Ako je iz konteksta jasno da se radi o vektorskom prostoru ili potprostoru, atribut "vektorski" ili "linearni" ćemo kadkad izostavljati.

Lekcija 3

Skalarni Produkt i Norma

3.1 Skalarni produkt i norma u V_n

Operacije zbrajanja vektora i množenja vektora realnim brojem koje smo uveli u V_n imaju svojstvo da kao rezultat uvijek daju element iz skupa V_n. To svojstvo zatvorenosti je osnovna značajka vektorskog prostora $(\mathrm{V}_n, +, \cdot)$. Sljedeća korisna operacija koja dvama elementima skupa V_n pridružuje realni broj zove se skalarni produkt. Iako rezultat te operacije ne pripada skupu V_n, ta operacije je važna jer omogućuje uvođenje pojmova kao što su duljina (ili norma) vektora, ortogonalnost vektora i kut između vektora. Također daje vrlo prirodan način definiranja potprostora pomoću sustava linearnih jednadžbi.

Skalarni produkt u vektorskom prostoru $(\mathrm{V}_n, +, \cdot)$ definiramo na sljedeći način. Ako su $x = (x_1, x_2, \ldots, x_n)$ i $y = (y_1, y_2, \ldots, y_n)$ bilo koja dva elementa iz V_n, tada je **skalarni produkt** vektora x i y, u oznaci $(x \mid y)$ ili $x \circ y$, definiran relacijom

$$(x \mid y) = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n \,. \tag{3.1}$$

Rezultat te operacije je realni broj, jer je desna strana jednadžbe (3.1) suma produkata realnih brojeva. Opet ponavljamo da skalarni produkt ne daje rezultat koji pripada skupu V_n, već daje realni broj koristeći pravilo (3.1).

Primjer 3.1 *Ako su a i b elementi prostora V_4 i*

$$a = (1, 0, 2, -2), \quad b = (-4, 5, 8, 1),$$

tada pravilo (3.1) daje

$$\begin{aligned}(a \mid b) &= 1 \cdot (-4) + 0 \cdot 5 + 2 \cdot 8 + (-2) \cdot 1 \\ &= -4 + 0 + 16 - 2 = 10\,.\end{aligned}$$

■

Kao kod svake nove operacije, tako i kod skalarnog produkta, trebamo naći osnovna svojstva te operacije i njenu interakciju s postojećim operacijama $+$ i $\cdot$.

Teorem 3.2 *Za skalarni produkt vektora vrijede sljedeća svojstva:*

(i) $(x \mid x) \geq 0$ *za sve* $x \in V_n$

(ii) $(x \mid x) = 0$ *ako i samo ako* $x = 0$

(iii) $(x \mid y) = (y \mid x)$ *za sve* $x, y \in V_n$

(iv) $(\alpha x \mid y) = \alpha(x \mid y)$ *za sve* $x, y \in V_n$ *i svaki* $\alpha \in \mathbf{R}$

(v) $(x + y \mid z) = (x \mid z) + (y \mid z)$ *za sve* $x, y, z \in V_n$

Dokaz: (i) Ako koristeći definiciju (3.1) napišemo izraz za skalarni "kvadrat" $(x \mid x)$, dobijemo

$$(x \mid x) = x_1x_1 + x_2x_2 + \ldots + x_nx_n = \sum_{i=1}^{n} x_i^2\,. \tag{3.2}$$

Jer su svi x_i realni brojevi, tvrdnja (i) odmah slijedi.
(ii) Tvrdnja slijedi neposredno iz relacije (3.2).
(iii) Dokaz slijedi direktno iz definicije skalarnog produkta (3.1) zbog komutativnosti množenja realnih brojeva.
(iv) Ova tvrdnja slijedi zbog asocijativnosti množenja i distributivnosti množenja (u odnosu na zbrajanje) realnih brojeva.
(v) Tvrdnja slijedi zbog distributivnosti množenja prema zbrajanju realnih brojeva. ■

Uočimo da svojstva (iii) i (iv) povlače

$$(x \mid \alpha y) = (\alpha y \mid x) = \alpha(y \mid x) = \alpha(x \mid y) = (\alpha x \mid y). \tag{3.3}$$

Svojstva (i) i (ii) teorema 3.2 su posebno važna jer nam omogućuju definirati **duljinu**, odnosno **normu** elementa iz V_n. Normu od $x \in \mathrm{V}_n$, koju označavamo s $\|x\|$, definiramo relacijom

$$\|x\| = \sqrt{(x \mid x)}\,. \tag{3.4}$$

Ako iskoristimo relaciju (3.2), dobivamo

$$\|x\| = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}\,. \tag{3.5}$$

U V_2 odnosno u V_3 vrijedi

$$\|x\| = \sqrt{x_1^2 + x_2^2}\,, \qquad \|x\| = \sqrt{x_1^2 + x_2^2 + x_3^2}\,,$$

što zapravo nije ništa drugo do li iskaz Pitagorinog teorema u dvije, odnosno tri dimenzije. Definicija norme vektora iz V_n relacijom (3.5) nije ništa drugo nego poopćenje Pitagorinog teorema na n dimenzija. Na primjer, ako je $a = (1, 0, 2, -2) \in \mathrm{V}_4$, onda je

$$\|a\| = \sqrt{1^2 + 0^2 + 2^2 + (-2)^2} = \sqrt{9} = 3\,.$$

Primijetimo još, da iz svojstava (i), (ii) i (iv) skalarnog produkta slijedi

$$\|x\| \geq 0\,, \tag{3.6}$$

$$\|x\| = 0 \text{ ako i samo ako } x = 0\,, \tag{3.7}$$

$$\|\alpha x\| = |\alpha|\, \|x\|\,. \tag{3.8}$$

Zadnji rezultat je dobiven tako, da u definiciji (3.4) umjesto x pišemo αx, pa korištenjem svojstva (iv) i relacije (3.3) dobijemo faktor α^2 ispod korijena, što daje faktor $\sqrt{\alpha^2} = |\alpha|$.

Ako je $x \in \mathrm{V}_n$ takav da je $\|x\| = 1$, onda je x *jedinični vektor*. Jedinični vektor možemo dobiti iz svakog netrivijalnog vektora, dijeleći ga sa njegovom normom. Dakle, ako $y \in \mathrm{V}_n$ nije nulvektor, onda je

$$x = \left(\frac{1}{\|y\|}\right) y = \frac{1}{\|y\|} y$$

jedinični vektor u smjeru vektora y. On se još zove *normirani vektor* y. Jediničnu duljinu vektora x možemo provjeriti tako da izračunamo $\|x\|$ koristeći relaciju (3.8)

$$\|x\| = \left\| \left(\frac{1}{\|y\|} \right) y \right\| = \left| \frac{1}{\|y\|} \right| \|y\| = \frac{1}{\|y\|} \|y\| = 1 .$$

Jednadžba (3.8) nam kaže kako se norma ponaša u kombinaciji s operacijom množenja vektora sa skalarom. Sada ćemo pokazati nekoliko svojstava norme u kombinaciji sa operacijom zbrajanja u V_n. Korištenjem relacija (3.5) i (3.1) dobivamo

$$\begin{aligned} ||x \pm y||^2 &= (x \pm y \mid x \pm y) = (x \mid x) \pm (x \mid y) \pm (y \mid x) + (y \mid y) \\ &= \|x\|^2 \pm 2(x \mid y) + \|y\|^2 . \end{aligned} \tag{3.9}$$

Zbrajanjem i oduzimanjem dviju jednadžbi koje se kriju u relaciji (3.9), dobivamo

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2 \tag{3.10}$$

$$\|x + y\|^2 - \|x - y\|^2 = 4(x \mid y) . \tag{3.11}$$

Relacija (3.10) se zove relacija paralelograma jer je u svakom paralelogramu zbroj kvadrata duljina stranica jednak zbroju kvadrata duljina dijagonala. Druga relacija opet pokazuje da se skalarni produkt može izraziti preko kvadrata norme zbroja i kvadrata norme razlike ulaznih vektora.

Teorem 3.3 *Za sve $x, y \in V_n$ vrijedi*

$$|(x \mid y)| \leq \|x\| \, \|y\| \qquad \text{(Cauchy-Schwarzova nejednakost)}, \tag{3.12}$$

$$\|x + y\| \leq \|x\| + \|y\| \qquad \text{(nejednakost trokuta)}. \tag{3.13}$$

Dokaz: Ako je $x = 0$ ili $y = 0$, onda nejednakosti (3.12) i (3.13) postaju jednakosti pa (3.12) i (3.13) vrijede. Pretpostavimo da vrijedi $x \neq 0$ i $y \neq 0$ i izračunajmo

$$\begin{aligned} \|x - \alpha y\|^2 &= (x - \alpha y | x - \alpha y) \\ &= (x \mid x) - 2\alpha(x \mid y) + \alpha^2 (y \mid y) , \end{aligned}$$

gdje je α bilo koji realni broj. Korišteći činjenicu da je $\|x-\alpha y\|^2 \geq 0$ i relaciju (3.4), dobivamo

$$0 \leq \|x\|^2 - 2\alpha(x \mid y) + \alpha^2\|y\|^2.$$

Ova nejednakost vrijedi za svaki realni broj α. Ako uzmemo

$$\alpha = \frac{(x \mid y)}{\|y\|^2}\,,$$

dobivamo

$$\begin{aligned} 0 &\leq \|x\|^2 - 2\frac{(x \mid y)}{\|y\|^2}(x \mid y) + \left(\frac{(x \mid y)}{\|y\|^2}\right)^2 \|y\|^2 \\ &= \|x\|^2 - \frac{(x \mid y)^2}{\|y\|^2} \\ &= \|x\|^2 \left(1 - \frac{(x \mid y)^2}{\|x\|^2\,\|y\|^2}\right). \end{aligned}$$

Dijeljenjem sa pozitivnim brojem $\|x\|^2$ zaključujemo da je izraz u zagradi nenegativan. Množenjem sa pozitivnim brojem $\|x\|^2\|y\|^2$, prebacivanjem člana $-(x \mid y)^2$ na lijevu stranu nejednadžbe, te uzimanjem drugog korijena dobivamo nejednakost (3.12).

Koristeći nejednakost (3.12), dobivamo

$$\begin{aligned} \|x+y\|^2 &= \|x\|^2 + 2(x \mid y) + \|y\|^2 \leq \|x\|^2 + 2 \mid (x \mid y) \mid + \|y\|^2 \\ &\leq \|x\|^2 + 2\|x\|\,\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2\,. \end{aligned}$$

Uzimanjem drugog korijena lijeve i desne strane nejednakosti dobivamo nejednakost (3.13). ■

Nejednakost (3.12) kaže da vrijedi

$$-1 \leq \frac{(x \mid y)}{\|x\|\,\|y\|} \leq 1\,,$$

pa možemo relacijom $\cos\theta = (x \mid y)/|x\|\,\|y\|$ definirati kut θ. Uočimo, ako je $y = x$, dobivamo $\cos\theta = 1$, tj. $\theta = 0$, a ako je $y = -x$, dobivamo $\cos\theta = -1$, tj. $\theta = \pi$. Ako x i y interpretiramo pomoću radij-vektora, tada je jasno da kut između x i x mora biti 0, a između x i $-x$ mora biti π.

Sada relacija (3.9) prelazi u

$$\|x \pm y\|^2 = \|x\|^2 + \|y\|^2 \pm 2\|x\| \, \|y\| \, \cos\theta \,,$$

što nije ništa drugo nego kosinusov teorem. To nas motivira na sljedeću definiciju: *kosinus kuta između dva netrivijalna vektora* iz V_n je dan formulom

$$\cos\theta = \frac{(x \mid y)}{\|x\| \, \|y\|} \,, \quad x \neq 0 \,, y \neq 0 \,,\ 0 \leq \theta \leq \pi. \tag{3.14}$$

Ta se formula koristi i u analitičkoj geometriji kod definicije kuta između dva radij-vektora.

Definicija (3.14) je osnova za definiciju ortogonalnosti vektora. Kažemo da su vektori x i y iz V_n *ortogonalni* (ili okomiti) ako vrijedi

$$(x \mid y) = 0 \,. \tag{3.15}$$

Koristeći relaciju (3.15) odmah zaključujemo da je nulvektor ortogonalan na svaki vektor.

Neka je dan vektor $a \in \mathrm{V}_n$. Potražimo sve vektore $x \in \mathrm{V}_n$ koji su ortogonalni na a. Uvjet koji svaki vektor x mora zadovoljavati je jednostavan, $(a \mid x) = 0$. Skup svih vektora sa traženim svojstvom označimo sa $\mathcal{U}$,

$$\mathcal{U} = \{x \in \mathrm{V}_n ; (a \mid x) = 0\}.$$

Pokažimo da je $\mathcal{U}$ vektorski potprostor od V_n. Neka su $u_1, u_2 \in \mathcal{U}$ i $\alpha_1, \alpha_2 \in \mathbf{R}$. Korištenjem svojstava (iii)–(v) skalarnog produkta iz teorema 3.2 (specijalno svojstva (3.2)), dobivamo

$$\begin{aligned} (a \mid \alpha_1 u_1 + \alpha_2 u_2) &= (a \mid \alpha_1 u_1) + (a \mid \alpha_2 u_2) = \alpha_1 (a \mid u_1) + \alpha_2 (a \mid u_2) \\ &= \alpha_1 \cdot 0 + \alpha_2 \cdot 0 = 0 + 0 = 0, \end{aligned}$$

pa je $\alpha_1 u_1 + \alpha_2 u_2 \in \mathcal{U}$. Prema uvjetu (ab) iz prethodnog poglavlja, zaključujemo da je $\mathcal{U}$ vektorski potprostor od V_n. S obzirom da se uvjet $(a \mid x) = 0$ može zapisati u obliku homogene linearne algebarske jednadžbe

$$a_1 x_1 + a_2 x_2 + \ldots + a_n x_n = 0, \tag{3.16}$$

pokazali smo da je skup svih rješenja jednadžbe (3.16) vektorski potprostor od V_n. Sljedeća propozicija je jednostavna generalizacija te tvrdnje.

Propozicija 3.4 *Skup svih rješenja sustava od p homogenih linearnih jednadžbi*

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= 0 \\ \ldots \\ a_{pp}x_p + a_{p2}x_2 + \ldots + a_{pn}x_n &= 0 \end{aligned} \tag{3.17}$$

čini vektorski potprostor od V_n.

Dokaz: Ako označimo: $a_i = (a_{i1}, a_{i2}, \ldots, a_{in})$, $1 \leq i \leq p$, tada su $a_1, a_2, \ldots, a_p \in \mathrm{V}_n$, pa je skup svih rješenja homogenog sustava (3.17) opisan relacijom

$$\mathcal{V} = \{x \in \mathrm{V}_n; (a_i \mid x) = 0,\ 1 \leq i \leq p\}.$$

Uočimo da je $(0, 0, \ldots, 0) \in \mathcal{V}$, pa $\mathcal{V}$ nije prazan. Da bi pokazali da je $\mathcal{V}$ vektorski potprostor moramo provjeriti je li uvjet (ab) iz točke 2.3.3 ispunjen. Neka su $v_1, v_2 \in \mathcal{V}$ i $\alpha_1, \alpha_2 \in \mathbf{R}$. Tada za $1 \leq i \leq p$ vrijedi

$$(a_i \mid \alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 (a_i \mid v_1) + \alpha_2 (a_i \mid v_2) = \alpha_1 \cdot 0 + \alpha_1 \cdot 0 = 0,$$

pa je $\alpha_1 v_1 + \alpha_2 v_2 \in \mathcal{V}$. Time je propozicija dokazana. ■

U propoziciji je p proizvoljan. Ako je p dovoljno velik, možda će skup rješenja biti tek trivijalni potprostor $\{0\}$. Vidjet ćemo da je najmanji p sa tim svojstvom upravo broj nepoznanica n, ali pod dodatnim uvjetom da vektori a_i razapinju cijeli V_n.

Pitanje koje se sada nameće jest, kako naći sve vektore potprostora kojeg čine rješenja danog sustava linearnih homogenih jednadžbi. Jedan način za opisivanje potprostora je određivanje jednog njegovog razapinjućeg skupa. Na žalost, još nemamo sustavnu metodu nalaženja razapinjućeg skupa, počevši od sustava (3.17). Mogli bismo probati metodom pokušaja i promašaja naći vrlo mnogo rješenja sustava (3.17), ali tako nikad ne bismo bili sigurni da svi ti vektori razapinju cijeli potprostor. Sustavni postupak za to postoji i on je usko povezan sa nalaženjem minimalnog razapinjućeg skupa za $L(a_1, \ldots, a_p)$. Uočimo da $\{a_1, \ldots, a_p\}$ ne mora biti minimalni razapinjući skup. Kad jednom nađemo jedan minimalni razapinjući skup za $L(a_1, \ldots, a_p)$, tada je lakše odrediti skup svih vektora $\mathcal{Z}$ ortogonalnih na sve vektore spomenutog minimalnog skupa. Tada su svi vektori iz $\mathcal{Z}$ ortogonalni i na vektore $a_1, a_2, \ldots, a_p$, jer oni leže u $L(a_1, \ldots, a_p)$.

Traženje najmanjeg razapinjućeg skupa jednog potprostora se radi na najefikasniji način pomoću algebarske strukture koju zovemo matrična algebra, a koju ćemo razviti u sljedećem poglavlju.

3.2 Skalarni produkt i norma u $\mathbf{C}_n$

Skalarni produkt u $\mathbf{C}_n$ definiramo relacijom

$$(x \mid y) = x_1\bar{y}_1 + x_2\bar{y}_2 + \ldots + x_n\bar{y}_n\,. \tag{3.18}$$

Rezultat te operacije je kompleksni broj. Svojstva skalarnog produkta u $\mathbf{C}_n$ dana su sljedećim teoremom.

Teorem 3.5 *Za skalarni produkt vektora u $\mathbf{C}_n$ vrijede sljedeća svojstva:*

(i) $(x \mid x) \geq 0$ *za sve* $x \in \mathbf{C}_n$

(ii) $(x \mid x) = 0$ *ako i samo ako* $x = 0$

(iii) $(x \mid y) = \overline{(y \mid x)}$ *za sve* $x, y \in \mathbf{C}_n$

(iv) $(\alpha x \mid y) = \alpha(x \mid y)$ *za sve* $x, y \in \mathbf{C}_n$ *i svaki* $\alpha \in \mathbf{C}$

(v) $(x + y \mid z) = (x \mid z) + (y \mid z)$ *za sve* $x, y, z \in \mathbf{C}_n$

Dokaz: (i) Koristeći definiciju (3.18) za skalarni kvadrat $(x \mid x)$, dobivamo

$$(x \mid x) = x_1\bar{x}_1 + x_2\bar{x}_2 + \ldots + x_n\bar{x}_n = \sum_{i=1}^{n} \mid x_i \mid^2\ . \tag{3.19}$$

Jer su svi $\mid x_i \mid^2$ nenegativni realni brojevi, tvrdnja (i) odmah slijedi.
(ii) Tvrdnja slijedi neposredno iz relacije (3.19).
(iii), (iv) Dokazi slijede zbog istih razloga kao i za odgovarajuće tvrdnje

u teoremu 3.2. ■

Uočimo da svojstva (iii) i (iv) povlače

$$(x \mid \alpha y) = \overline{(\alpha y \mid x)} = \overline{\alpha (y \mid x)} = \bar{\alpha}\overline{(y \mid x)} = \bar{\alpha}(x \mid y). \tag{3.20}$$

Vidimo da kod skalarnog produkta u vektorskom prostoru $\mathbf{C}_n$ skalar koji množi drugi vektor izlazi ispred skalarnog produkta kao kompleksno konjugirani broj.

Kao i kod skalarnog produkta u V_n, preko svojstava (i) i (ii) možemo definirati duljinu, odnosno normu elementa $\mathbf{C}_n$. Norma od $x \in \mathbf{C}_n$ definirana je relacijom

$$\|x\| = \sqrt{(x \mid x)} = \sqrt{|x_1|^2 + |x_2|^2 + \ldots + |x_n|^2}\,. \tag{3.21}$$

Na slični način kao i u prostoru V_n, pokaže se da vrijedi

$$\|x\| \geq 0\,, \tag{3.22}$$

$$\|x\| = 0 \text{ ako i samo ako } x = 0\,, \tag{3.23}$$

$$\|\alpha x\| = |\alpha|\,\|x\|\,. \tag{3.24}$$

Vektor $x \in \mathbf{C}_n$ je jedinični ako je $\|x\| = 1$. Ako $y \in \mathbf{C}_n$ nije nulvektor, onda se na isti način kao prije, pokaže da je

$$x = \frac{y}{\|y\|}$$

jedinični vektor u smjeru vektora y. Kao i kod skalarnog produkta u V_n, direktnim računom možemo provjeriti osnovna svojstva skalarnog produkta u kombinaciji sa zbrajanjem vektora. Korištenjem osnovnih svojstava skalarnog produkta dobijemo,

$$\begin{aligned}\|\alpha x + \beta y\|^2 &= (\alpha x + \beta y \mid \alpha x + \beta y) \\ &= |\alpha|^2\|x\|^2 + \alpha\bar{\beta}(x \mid y) + \bar{\alpha}\beta\overline{(x \mid y)} + |\beta|^2\|y\|^2. \end{aligned} \tag{3.25}$$

Stavimo li u (3.25) $\alpha = 1,\ \beta = \pm 1$, dobivamo

$$\|x \pm y\|^2 = \|x\|^2 \pm 2\Re(x \mid y) + \|y\|^2, \tag{3.26}$$

gdje $\Re z$ označava realni dio kompleksnog broja z. Za $\alpha = 1,\ \beta = \pm i$, dobivamo

$$\|x \pm iy\|^2 = \|x\|^2 \pm 2\Im(x \mid y) + \|y\|^2, \tag{3.27}$$

gdje $\Im z$ označava imaginarni dio kompleksnog broja z. Koristeći relacije (3.26) i (3.27) dobivamo

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2\,, \tag{3.28}$$

$$(x \mid y) = \frac{1}{4}\left\{\|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2\right\}. \tag{3.29}$$

Dakle opet vrijedi relacija paralelograma (3.28). Relacija (3.29) izražava skalarni produkt u $\mathbf{C}_n$ pomoću normi odgovarajućih vektora.

Teorem 3.6 *Za sve $x, y \in \mathbf{C}_n$ vrijedi*

$$|(x \mid y)| \leq \|x\|\,\|y\| \quad \textit{(Cauchy-Schwarzova nejednakost)} \tag{3.30}$$

i

$$\|x+y\| \leq \|x\| + \|y\| \qquad \textit{(nejednakost trokuta)}. \tag{3.31}$$

Dokaz: Kao i u dokazu teorema 3.3, možemo pretpostaviti da su oba vektora netrivijalna. Stoga je $(y \mid y) \neq 0$, pa su dobro definirani brojevi

$$\alpha = \sqrt{(y \mid y)} = \|y\| \qquad \text{i} \qquad \beta = -\frac{(x \mid y)}{\|y\|}.$$

Uvrštenje u (3.25) daje

$$\begin{aligned}
0 \leq & \left\| \|y\|x - \frac{(x \mid y)}{\|y\|} y \right\|^2 = \\
& = \|y\|^2\|x\|^2 - \|y\|\frac{\overline{(x \mid y)}}{\|y\|}(x \mid y) - \|y\|\frac{(x \mid y)}{\|y\|}\overline{(x \mid y)} + \frac{|(x \mid y)|^2}{\|y\|^2}\|y\|^2 \\
& = \|y\|^2\|x\|^2 - \overline{(x \mid y)}(x \mid y) - (x \mid y)\overline{(x \mid y)} + |\,(x \mid y)\,|^2 \\
& = \|x\|^2\|y\|^2 - |\,(x \mid y)\,|^2\,.
\end{aligned} \tag{3.32}$$

Time je dokazana Cauchy-Schwarzova nejednakost.

Koristeći relaciju (3.26) i nejednakost (3.30), dobivamo

$$\begin{aligned}
\|x+y\|^2 &= \|x\|^2 + 2\Re(x \mid y) + \|y\|^2 \leq \|x\|^2 + 2\,|\,(x \mid y)\,| + \|y\|^2 \\
&\leq \|x\|^2 + 2\|x\|\,\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2\,,
\end{aligned}$$

pa uzimanjem drugog korijena lijeve i desne strane nejednakosti dobivamo nejednakost (3.31). ■

Za vektore $x, y \in \mathbf{C}_n$ kažemo da su ortogonalni (ili okomiti) ako vrijedi

$$(x \mid y) = 0\,.$$

Opet je nul-vektor okomit na sve vektore. Slično kao i prije, kut između vektora $x, y \in \mathbf{C}_n$ se može definirati preko formule

$$\cos\theta = \frac{|\,(x \mid y)\,|}{\|x\|\,\|y\|}\,, \quad x \neq 0\,, y \neq 0\,,\ 0 \leq \theta \leq \frac{\pi}{2}, \tag{3.33}$$

ali on ni u slučaju prostora $\mathbf{C}_2$, $\mathbf{C}_3$ nema jasno geometrijsko značenje.

3.3 Skalarni produkt i norma u općem vektorskom prostoru

Skalarni produkt i norma u općem vektorskom prostoru definiraju se samo za vektorske prostore nad poljima $\Phi = \mathbf{R}$ i $\Phi = \mathbf{C}$. Te dvije funkcije definiraju se nezavisno jedna od druge, pa se vektorski prostor u kojem postoji norma naziva normirani vektorski prostor, a onaj u kojem postoji skalarni produkt naziva se unitarni vektorski prostor. Vidjet ćemo da je svaki unitarni vektorski prostor također normirani vektorski prostor. Stoga razmatranje započinjemo s općenitijim, normiranim vektorskim prostorom. Do kraja ovog poglavlja Φ je ili $\mathbf{R}$ ili $\mathbf{C}$.

Definicija 3.7 *Neka je $(X, +, .)$ vektorski prostor nad Φ, pri čemu je $\Phi = \mathbf{R}$ ili $\Phi = \mathbf{C}$. Funkcija $\|\ \| : X \longrightarrow \mathbf{R}$ je* norma *ako vrijedi*

(i) $\|x\| \geq 0\,,$ *za svaki $x \in X$,* *(nenegativnost)*

(ii) $\|\alpha x\| = |\,\alpha\,|\ \|x\|\,,$ *za sve $\alpha \in \Phi$, $x \in X$,* *(homogenost)*

(iii) $\|x + y\| \leq \|x\| + \|y\|\,,$ *za sve $x, y \in X$, (nejednakost trokuta)*

(iv) $\|x\| = 0\,,$ *ako i samo ako $x = 0$* *(definitnost).*

Vektorski prostor na kojem je definirana norma zove se normirani vektorski prostor*. Funkcija $\|\ \| : X \longrightarrow \Phi$ koja zadovoljava samo svojstva (i), (ii) i (iii) se zove* polunorma *ili seminorma.*

U normiranom vektorskom prostoru, norma sume ili razlike vektora može se ocijeniti odozdo.

Propozicija 3.8 *Za svaku polunormu pa zato i normu vrijedi*

$$\|x \pm y\| \geq |\, \|x\| - \|y\|\, | \,. \tag{3.34}$$

Dokaz: Iz nejednakosti trokuta slijedi

$$\|x\| = \|x - y + y\| \leq \|x - y\| + \|y\| \,,$$

pa je

$$\|x - y\| \geq \|x\| - \|y\| \,.$$

Isto tako iz

$$\|y\| = \|y - x + x\| \leq \|y - x\| + \|x\| = \|x - y\| + \|x\|$$

slijedi

$$\|x - y\| \geq -\left(\|x\| - \|y\|\right) \,.$$

Zaključujemo da je

$$\|x - y\| \geq |\, \|x\| - \|y\|\, | \,.$$

Ako u posljednjoj nejednakosti zamijenimo y sa $-y$ i koristimo svojstvo (ii), dobijemo

$$\|x + y\| = \|x - (-y)\| \geq |\, \|x\| - \| - y\|\, | = |\, \|x\| - \|y\|\, | \,. \quad \blacksquare$$

Primjer 3.9 *U vektorskim prostorima V_n i $\mathbf{C}_n$ možemo definirati sljedeće funkcije*

$$\|x\|_1 = |x_1| + |x_2| + \ldots + |x_n| \,, \tag{3.35}$$

$$\|x\|_\infty = \max\{|x_1|,\, |x_2|,\, \ldots\, |x_n|\} \tag{3.36}$$

$$\|x\|_p = \sqrt[p]{|x_1|^p + |x_2|^p + \ldots + |x_n|^p}\,, \quad 1 \leq p \leq \infty \tag{3.37}$$

Lako je provjeriti da su $\|x\|_1$ i $\|x\|_\infty$ funkcije norme. Za opću, tzv. H"olderovu normu $\|x\|_p$ dokaz je složeniji. Uočimo da je H"olderova 2-norma $\|x\|_2$ upravo ona koju smo prije proučavali. $\blacksquare$

Dvije norme $\| \ \|_\alpha$ i $\| \ \|_\beta$ vektorskog prostora X su ekvivalentne, ako postoje realni pozitivni brojevi c_1 i c_2 takvi da vrijedi

$$c_1\|x\|_\alpha \leq \|x\|_\beta \leq c_2\|x\|_\alpha.$$

Npr. lako se pokaže da su norme $\| \ \|_2$, $\| \ \|_1$ i $\| \ \|_\infty$ u vektorskom prostoru V_n međusobno ekvivalentne. Pokušajte naći odgovarajuće konstante c_1 i c_2 za svaki par normi. Može se pokazati da su u vektorskim prostorima V_n i $\mathbf{C}_n$ sve norme međusobno ekvivalentne.

Za definiciju skalarnog produkta važna su svojstva (i)—(v) teorema 3.2 odnosno teorema 3.5

Definicija 3.10 *Neka je $(X, +, \cdot)$ vektorski prostor nad Φ, gdje je $\Phi = \mathbf{C}$. Funkcija $(\cdot \mid \cdot) : X \times X \longrightarrow \Phi$ se zove skalarni produkt ako vrijedi*

(i) $(x \mid x) \geq 0$ *za sve* $x \in X$

(ii) $(x \mid x) = 0$ *ako i samo ako* $x = 0$

(iii) $(x \mid y) = \overline{(y \mid x)}$ *za sve* $x, y \in X$

(iv) $(\alpha x \mid y) = \alpha(x \mid y)$ *za sve* $x, y \in X$ *i svaki* $\alpha \in \Phi$

(v) $(x + y \mid z) = (x \mid z) + (y \mid z)$ *za sve* $x, y, z \in X$.

Ako je $\Phi = \mathbf{R}$, tada u svojstvu (iii) treba ispustiti operaciju kompleksnog konjugiranja. Vektorski prostor na kojem je definiran skalarni produkt zove se unitarni vektorski prostor*.*

Primjenom svojstava (i), (ii) i (iii) iz definicije 3.10, dobivamo

$$\begin{aligned} (x + y \mid x + y) &= (x \mid x) + (x \mid y) + (y \mid x) + (y \mid y) \\ &= (x \mid x) + 2\Re(x \mid y) + (y \mid y)\,. \end{aligned} \tag{3.38}$$

Da bi pokazali kako je svaki unitarni vektorski prostor također normirani, trebat će nam

Propozicija 3.11 *U svakom unitarnom vektorskom prostoru sa skalarnim produktom $(\cdot \mid \cdot)$ vrijedi*

$$|\,(x \mid y)\,|^2 \leq (x \mid x)(y \mid y)\,. \tag{3.39}$$

Dokaz: Ako je $x = 0$ ili $y = 0$, onda relacija (3.39) vrijedi jer su obje strane nejednakosti nula. Pretpostavimo da vrijedi $x \neq 0$ i $y \neq 0$ i izračunajmo $(x - \alpha y \mid x - \alpha y)$ za realno α. Stavljajući $-\alpha y$ umjesto y u nejednakosti (3.38), dobivamo

$$0 \leq (x - \alpha y \mid x - \alpha y) = (x \mid x) - 2\alpha \Re(x \mid y) + \alpha^2 (y \mid y) \,.$$

Ova nejednakost vrijedi za svaki realni broj α. Ako uzmemo

$$\alpha = \frac{\Re(x \mid y)}{(y \mid y)} \,,$$

dobivamo

$$\begin{aligned} 0 &\leq (x \mid x) - 2\frac{\Re(x \mid y)}{(y \mid y)}\Re(x \mid y) + \left(\frac{\Re(x \mid y)}{(y \mid y)}\right)^2 (y \mid y) \\ &= (x \mid x) - \frac{|\Re(x \mid y)|^2}{(y \mid y)} \\ &= (x \mid x)\left(1 - \frac{|\Re(x \mid y)|^2}{(x \mid x)(y \mid y)}\right) \,. \end{aligned}$$

Ova nejednakost može biti zadovoljena onda i samo onda, ako vrijedi

$$\frac{|\Re(x \mid y)|^2}{(x \mid x)(y \mid y)} \leq 1 \,. \tag{3.40}$$

Zamijenimo li x sa ηx, gdje je η kompleksni broj modula 1, koji zadovoljava

$$(\eta x \mid y) = \eta(x \mid y) = \mid (x \mid y) \mid \,,$$

dobivamo

$$(\eta x \mid \eta x) = \eta\bar{\eta}(x \mid x) = \mid (x \mid x) \mid \,,$$

i

$$\frac{|(x \mid y)|^2}{(x \mid x)(y \mid y)} \leq 1 \,,$$

a to je isto što i (3.39). Uočimo da traženi η ovisi o x i postoji za svako x. Možemo ga definirati s

$$\eta = \begin{cases} \frac{\overline{(x|y)}}{|(x|y)|}, & \text{ako je } (x \mid y) \neq 0 \\ 1 & \text{ako je } (x \mid y) = 0 \end{cases} \,.$$

Kako je nejednakost (3.40) istinita za svako x i y isto vrijedi i za nejednakost (3.39). ■

Propozicija 3.12 *Ako je $(\cdot \mid \cdot)$ skalarni produkt unitarnog vektorskog prostora X, onda je pomoću $\|x\| = \sqrt{(x \mid x)}$ definirana norma na istom prostoru. Prema tome, svaki unitarni vektorski prostor je ujedno i normirani vektorski prostor.*

Dokaz: Treba provjeriti jesu li za funkciju $\|x\| = \sqrt{(x \mid x)}$ ispunjena svojstva norme (i)—(iv) iz definicije 3.7. Svojstva (i) i (iv) slijede direktno iz svojstava (i) i (ii) skalarnog produkta (vidi definiciju 3.10). Kombinacijom svojstava (ii) i (iii) iz definicije skalarnog produkta lako dobijemo

$$(\alpha x \mid \alpha x) = \alpha(x \mid \alpha x) = \alpha\bar{\alpha}(x \mid x) = \mid \alpha \mid^2 (x \mid x)\,.$$

Vađenjem drugog korijena lijeve i desne strane, dobivamo svojstvo (ii) norme. Konačno, koristeći relacije (3.38) i (3.39) dobivamo

$$\begin{aligned} \|x+y\|^2 &= \|x\|^2 + 2\Re(x \mid y) + \|y\|^2 \leq \|x\|^2 + 2|(x \mid y)| + \|y\|^2 \\ &\leq \|x\|^2 + 2\|x\|\,\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2\,, \end{aligned}$$

iz čega direktno slijedi preostalo svojstvo norme (nejednakost trokuta). ■

Jer je $\sqrt{(x \mid x)}$ norma (u oznaci $\|x\|$), propozicija 3.11 izriče da za nju vrijedi Cauchy-Schwarzova nejednakost. Normu iz propozicije 3.12 ćemo zvati norma pridružena skalarnom produktu ili norma izvedena (inducirana) iz skalarnog produkta.

Propozicija 3.13 *U svakom unitarnom vektorskom prostoru za pridruženu normu vrijedi relacija paralelograma,*

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2\,. \tag{3.41}$$

Dokaz: Relacija paralelograma se dobiva direktno zbrajanjem slijedećih dviju relacija

$$\begin{aligned} \|x+y\|^2 &= \|x\|^2 + \|y\|^2 + (x \mid y) + (y \mid x)\,, \\ \|x-y\|^2 &= \|x\|^2 + \|y\|^2 - (x \mid y) - (y \mid x)\,. \end{aligned}$$ ■

U našim dosadašnjim proučavanjima skalarnog produkta i norme došli smo do zaključka da je svaki unitarni vektorski prostor ujedno i normirani vektorski prostor. Postavlja se pitanje da li vrijedi i obratno, tj. da li svakoj normi $\| \ \|$ jednog normiranog vektorskog prostora možemo pridružiti neki skalarni produkt $(\cdot \mid \cdot)$, tako da vrijedi $\|x\| = \sqrt{(x \mid x)}$. Odgovor je ne, i to dokazuje sljedeća

Propozicija 3.14 *Za normu $\| \ \|_1$ vektorskog prostora V_n ne postoji niti jedan skalarni produkt $(\cdot \mid \cdot)$, za koji bi vrijedilo $\|x\|_1 = \sqrt{(x \mid x)}$.*

Dokaz: Kad bi postojao skalarni produkt sa tim svojstvom, onda bi za normu $\| \ \|_1$ morala vrijediti relacija paralelograma. Međutim, ako uzmemo npr. $x = (1, 2, 3, 0, \ldots, 0)$ i $y = (7, -1, -1, 0, \ldots, 0)$, imamo $x + y = (8, 1, 2, 0, \ldots, 0)$, $x - y = (-6, 3, 4, 0, \ldots, 0)$, pa dobivamo

$$\begin{aligned}
&\|x + y\|_1^2 + \|x - y\|_1^2 = (8 + 1 + 2)^2 + (6 + 3 + 4)^2 = 290\,, \\
&2\|x\|_1^2 + 2\|y\|_1^2 = 2(1 + 2 + 3)^2 + 2(7 + 1 + 1)^2 = 72 + 162 = 234\,.
\end{aligned}$$

Dakle relacija paralelograma ne vrijedi, pa ne postoji skalarni produkt kome bi bila pridružena norma $\| \ \|_1$. ■

Slično se može pokazati da ni za normu $\| \ \|_\infty$ u V_n relacija paralelograma ne vrijedi. Za koju normu u V_n relacija paralelograma vrijedi?

Prilikom dokazivanja da nekoj normi ne možemo pridružiti skalarni produkt koji ju inducira, poslužili smo se relacijom paralelograma. Pokazuje se, da je ta relacija odlučujuća i za odgovor na pitanje: kada možemo nekoj normi pridružiti skalarni produkt koji ju inducira? Taj rezultat navodimo bez dokaza.

Teorem 3.15 *Normi $\| \ \|$ nekog normiranog vektorskog prostora X možemo pridružiti skalarni produkt $(\cdot \mid \cdot)$ sa svojstvom $\|x\| = \sqrt{(x \mid x)}$ onda i samo onda ako norma zadovoljava relaciju paralelograma. U slučaju realnog normiranog prostora taj skalarni produkt određen je relacijom (3.11), tj.*

$$(x \mid y) = \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2)\,.$$

U slučaju kompleksnog normiranog prostora taj skalarani produkt je određen formulom (3.29), tj.

$$(x \mid y) = \frac{1}{4}\left\{\|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2\right\} .$$

■

Lekcija 4

Matrice

U ovom poglavlju ćemo uvesti novu algebarsku strukturu koja će značajno pojednostaviti i sistematizirati računanje rješenja sustava linearnih jednadžbi. Ta algebarska struktura se zove **matrična algebra**. Nova struktura je potrebna jer skalarni produkt $(\cdot \mid \cdot)$ koji je upotrebljen u konstrukciji sustava u obliku

$$(a_i \mid x) = b_i\,, \quad i = 1, \ldots, m \tag{4.1}$$

ne pripada $(\mathrm{V}_n, +, \cdot)$. Takove sustave ne možemo riješiti koristeći samo operacije $+$ i $\cdot$ koje pripadaju $(\mathrm{V}_n, +, \cdot)$. Za rješavanje sustava linearnih jednadžbi potrebna je veća struktura koja sadrži dodatne operacije.

Matrica je matematički objekt koji se sastoji od (realnih ili kompleksnih) brojeva koji su raspoređeni u retke i stupce. Ona se zapisuje u obliku pravokutne sheme. Brojeve od kojih se sastoji zovemo **elementima** matrice. Matrica A sa m redaka, n stupaca i s elementima a_{ij} zapisuje se kao

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \ldots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \ldots & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \ldots & a_{mn} \end{bmatrix} = (a_{ij})\,.$$

Takvu matricu zovemo $m \times n$ (čitaj: m-puta-n) matrica ili matrica tipa (reda, dimenzije) $m \times n$. Pritom je (a_{ij}) tek kraći zapis za pravoktnu shemu iz gornje relacije. Ako pišemo $A = (a_{ij})$ mislimo na cijelu shemu brojeva koji čine matricu A. Ako pišemo a_{ij} (ili $[A]_{ij}$ ili $(A)_{ij}$), mislimo na element koji se nalazi na presjeku $i-$tog retka i $j-$tog stupca matrice A. Pritom niz brojeva

$$a_{i1},\ a_{i2},\ a_{i3},\ \ldots,\ a_{in}$$

zovemo $i-$ti redak, a niz brojeva

$$\begin{matrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{matrix}$$

$j-$ti stupac matrice A. Ako vrijedi $m = n$, tj. ako je broj redaka matrice jednak broju stupaca, kažemo da je A **kvadratna** matrica reda n. Matricu sa samo jednim retkom zovemo **matrica redak** ili jednoretčana matrica, a matricu sa samo jednim stupcem zovemo **matrica stupac** ili jednostupčana matrica. Matricu, čiji su elementi realni brojevi, zovemo **realna matrica**, a matricu, čiji su elementi kako realni tako i kompleksni brojevi, zovemo **kompleksna matrica**.

Gore uvedena notacija podsjeća nas na notaciju koju smo koristili za vektore. To nije slučajno, jer ćemo kasnije vidjeti da vektore iz V_n možemo pistovjetiti sa jednostupčanim realnim matricama.

Praktično je imati oznaku za skup svih m-puta-n matrica, baš kao što je praktično imati simbol V_n za skup svih uređenih $n-$torki realnih brojeva. Sa $\mathbf{R}^{m\times n}$ ($\mathbf{C}^{m\times n}$) označavamo skup svih realnih (kompleksnih) $m \times n$ matrica. Ako nam nije važan podatak jesu li matrični elementi realni ili kompleksni brojevi, pisat ćemo M_{mn}. Ako pišemo $A \in M_{mn}$ (čitaj: "A pripada skupu M_{mn}"), to znači da je A m-puta-n matrica. U slučaju kad je $m = n$, oznaka M_{nn} se skraćuje na M_n.

Prva stvar, koju moramo definirati je jednakost matrica.

Definicija 4.1 *Matrica $A \in M_{mn}$ je jednaka matrici $B \in M_{pq}$ ako vrijedi $m = p$, $n = q$ i*

$$a_{ij} = b_{ij}, \quad \textit{za sve} \;\; 1 \le i \le m\,, \quad 1 \le j \le n\,. \tag{4.2}$$

U tom slučaju pišemo

$$A = B\,. \tag{4.3}$$

Prema tome, dvije matrice su jednake ako i samo ako imaju jednak broj redaka, jednak broj stupaca i svaki element jedne matrice jednak je odgovarajućem elementu druge matrice. Budući da matrica A ima $m \cdot n$ elementata, jednakost (4.3) zapravo znači $m \cdot n$ jednakosti realnih brojeva.

Sada ćemo uvesti osnovne operacije koje čine strukturu matrične algebre. Pritom ćemo elemente matrica A,B, C,... označiti sa odgovarajućim malim slovima a_{ij}, b_{ij}, c_{ij},...

Realne Matrice

U $\mathbf{R}^{m\times n}$ se zbrajanje matrica definira na sljedeći način.

Definicija 4.2 *Neka je $A \in \mathbf{R}^{m\times n}$ i $B \in \mathbf{R}^{p\times q}$. Ako je $m = p$ i $n = q$, onda matricu $C \in \mathbf{R}^{m\times n}$ s elementima*

$$c_{ij} = a_{ij} + b_{ij}, \qquad 1 \leq i \leq m\,, \quad 1 \leq j \leq n \tag{4.4}$$

zovemo **zbrojem** *ili sumom matrica A i B i pišemo*

$$C = A + B\,. \tag{4.5}$$

Zbrajanje matrica A i B znači $m \cdot n$ zbrajanja realnih brojeva, kao što je naznačeno u (4.4). Ono je definirano za svaki par matrica iz $\mathbf{R}^{m\times n}$ i rezulat je uvijek u $\mathbf{R}^{m\times n}$. Dakle je $\mathbf{R}^{m\times n}$ zatvoren u odnosu na operaciju zbrajanja. Zbrajanje u $\mathbf{R}^{m\times n}$ ima ova svojstva:

1. $A + (B + C) = (A + B) + C$ (asocijativnost)
2. $A + B = B + A$ (komutativnost)
3. Postoji matrica $O \in \mathbf{R}^{m\times n}$ (neutralni element) sa svojstvom da je $A + O = A$ za svaku matricu $A \in \mathbf{R}^{m\times n}$. Svi elementi matrice O jednaki su nuli.
4. Za svaki $A \in \mathbf{R}^{m\times n}$ postoji jedna i samo jedna matrica koju označavamo s $-A$ (inverzni element), takva da vrijedi $A+(-A) = O$. Ako je $A = (a_{ij})$, onda je $-A = (-a_{ij})$.

Iz navedenih svojstava možemo zaključiti, da skup matrica $\mathbf{R}^{m\times n}$ zajedno sa operacijom zbrajanja čini *Abelovu grupu* (vidi §2.3).

Sljedeća operacija koju možemo jednostavno definirati je množenje matrica sa skalarom.

Definicija 4.3 *Ako je $A \in \mathbf{R}^{m\times n}$ i $c \in \mathbf{R}$, matricu $B \in \mathbf{R}^{m\times n}$ s elementima*

$$b_{ij} = ca_{ij}\,, \quad 1 \le i \le m\,, \;\; 1 \le j \le n \tag{4.6}$$

zovemo **umnožak** *ili produkt matrice A sa skalarom c i označavamo*

$$B = cA\,. \tag{4.7}$$

Jednakost (4.7) zapravo označava da smo svaki od $m \cdot n$ elemenata matrice A pomnožili s c kako bismo dobili matricu B. Jasno je da za svaki realni skalar c i svaku matricu $A \in \mathbf{R}^{m\times n}$, mora $B = cA$ opet biti u $\mathbf{R}^{m\times n}$. Za proizvoljne $A, B \in \mathbf{R}^{m\times n}$ i $\alpha, \beta, c \in \mathbf{R}$ vrijedi

1. $c(A + B) = cA + cB$ (distributivnost množenja prema zbrajanju u $\mathbf{R}^{m\times n}$)
2. $(\alpha + \beta)A = \alpha A + \beta A$ (distributivnost množenja prema zbrajanju u $\mathbf{R}$)
3. $(\alpha\beta)A = \alpha(\beta A)$ (kompatibilnost množenja)
4. $1A = A$ (netrivijalnost množenja).

Za svaki par prirodnih brojeva m i n dobili smo strukturu $(\mathbf{R}^{m\times n}, +, \cdot)$. Iz svojstava zbrajanja matrica kao i množenja matrica skalarom zaključujemo da je $(\mathbf{R}^{m\times n}, +, \cdot)$ vektorski prostor nad $\mathbf{R}$.

Promotrimo sada vektorski prostor jednostupčanih matrica $\mathbf{R}^{n\times 1}$. On se još označava sa $\mathbf{R}^n$. Neka su $a_1, a_2, \ldots, a_n$ realni brojevi. Pomoću njih možemo načiniti n-torku $a = (a_1, a_2, \ldots, a_n) \in \mathrm{V}_n$ kao i jednostupčanu matricu

$$A = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \in \mathbf{R}^n.$$

Mijenjajući realne brojeve $a_1, \ldots, a_n$ na sve moguće načine, vektor a prolazi kroz cijeli skup V_n dok u isto vrijeme jednostupčana matrica prolazi kroz cijeli skup $\mathbf{R}^n$. Dakle, za svaki $a \in \mathrm{V}_n$ postoji $A \in \mathbf{R}^n$ i obrnuto za svako $A \in \mathbf{R}^n$ postoji $a \in \mathrm{V}_n$ sa istim elementima. Dakle,

$$a = (a_1, a_2, \ldots, a_n) \longleftrightarrow \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = A.$$

Primijetimo da se uz to pridruživanje, operacije $+$ i $\cdot$ na V_n i operacije $+$ i $\cdot$ na $\mathbf{R}^n$, jednako ponašaju, tj. ako je $a \longleftrightarrow A$ i $b \longleftrightarrow B$ onda je

$$a + b \longleftrightarrow A + B\,, \quad \alpha a \longleftrightarrow \alpha A\,, \quad \text{za svako} \;\; \alpha \in \mathbf{R}\,.$$

Stoga kažemo da su vektorski prostori $(\mathrm{V}_n, +, \cdot)$ i $(\mathbf{R}^n, +, \cdot)$ međusobno **izomorfni**. Slobodno govoreći, dvije strukture su izomorfne, ako postoji bijekcija između njih takva da se operacije definirane na tim strukturama identično ponašaju u odnosu na bijekciju.

Nema bitne razlike između vektorskih prostora $(\mathrm{V}_n, +, \cdot)$ i $(\mathbf{R}^n, +, \cdot)$. Stoga je uobičajeno da se jednostupčane matrice iz $(\mathbf{R}^n, +, \cdot)$ označavaju malim arapskim slovima, i nazivaju se vektorima. To je logično jer $\mathbf{R}^n$ je vektorski prostor. Ako skalarni produkt (ili normu) na vektorskom prostoru $\mathbf{R}^n$ definiramo na isti način kao i na vektorskom prostoru V_n , tada $(\mathbf{R}^n, +, \cdot)$ postaje unitarni (ili normiran) vektorski prostor, jednako kao i V_n. Kasnije ćemo vidjeti da je za rješavanje sustava linearnih jednadžbi i drugih problema linearne algebre, $\mathbf{R}^n$ mnogo značajniji od prostora V_n.

Sve što je rečeno za odnos između vektorskih prostora V_n i $\mathbf{R}^n$ može se reći i za odnos između vektorskih prostora V_n i $\mathbf{R}^{1\times n}$.

Ako su $a_1, \ldots, a_n \in \mathbf{R}^m$ neki vektori, tada notacija $A = [a_1, , \ldots, a_n]$ znači da je matrica A tako izgrađena da su $a_1, \ldots, a_n$ njeni stupci u redoslijedu u kojem su napisani. Slično, ako su $a'_1, \ldots, a'_m \in \mathbf{R}^{1\times n}$ jednoretčane matrice (još se zovu: vektori retci), tada oznaka

$$A = \begin{bmatrix} a'_1 \\ a'_2 \\ \vdots a'_m \end{bmatrix}$$

ukazuje da je $A \in \mathbf{R}^{m\times n}$ tako građena da su joj $a_1', \ldots, a_m'$ retci, u redosljedu koji je naznačen. Oznaka $A = [a_1, , \ldots, a_n]$ se još zove particija matrice po stupcima, a ona druga particija po retcima.

Matrična algebra sadrži osim operacija $+$ i $\cdot$ još dvije dodatne operacije, i te operacije su pravi razlog zbog kojeg uvodimo tu novu strukturu. Te dvije operacije čine osnovu s kojom ćemo kasnije moći riješiti mnoge probleme.

Definicija 4.4 *Neka je* $A \in \mathbf{R}^{m\times n}$ *i* $B \in \mathbf{R}^{n\times p}$. **Umnožak** *ili produkt matrica A i B je matrica*

$$C = A \cdot B \in \mathbf{R}^{m\times p} \tag{4.8}$$

čiji elementi su određeni formulom

$$\begin{aligned} c_{ij} &= a_{i1}b_{1j} + a_{i2}b_{2j} + a_{i3}b_{3j} + \ldots + a_{in}b_{nj} \\ &= \sum_{k=1}^{n} a_{ik}b_{kj} \qquad \text{za sve } \ 1 \le i \le m, \ 1 \le j \le p. \end{aligned} \tag{4.9}$$

Važno je primijetiti da je *produkt matrica A i B definiran samo onda kada je broj stupaca matrice A jednak broju redaka matrice B*. Operacija množenja matrica $\cdot$ može se sagledati kao preslikavanje koje uređenom paru matrica određenih dimenzija pridružuje matricu također određenih dimenzija, prema dijagramu

$$\cdot \ : \mathbf{R}^{m\times n} \times \mathbf{R}^{n\times p} \longrightarrow \mathbf{R}^{m\times p}.$$

Produkt matrica se obično piše (kao i produkt skalara) bez znaka množenja između faktora, dakle AB. Formula (4.9) izgleda na prvi pogled vrlo komplicirano. Pretpostavimo da želimo izračunati c_{ij},

$$C = \begin{bmatrix} \circ & \circ & \circ & \circ & \circ \\ \circ & \circ & \circ & \circ & \circ \\ \circ & \circ & \circ & \circ & \circ \\ \circ & \circ & c_{ij} & \circ & \circ \\ \circ & \circ & \circ & \circ & \circ \\ \circ & \circ & \circ & \circ & \circ \\ \circ & \circ & \circ & \circ & \circ \\ \circ & \circ & \circ & \circ & \circ \end{bmatrix},$$

pri čemu je $C = AB$. Tada iz matrice A trebamo imati na raspolaganju samo $i-$ti redak, a iz matrice B samo $j-$ti stupac:

$$AB = \begin{bmatrix} \circ & \circ & \circ & \circ & \circ & \circ & \circ \\ \circ & \circ & \circ & \circ & \circ & \circ & \circ \\ \circ & \circ & \circ & \circ & \circ & \circ & \circ \\ a_{i1} & a_{i2} & a_{i3} & \cdot & \cdot & \cdot & a_{in} \\ \circ & \circ & \circ & \circ & \circ & \circ & \circ \\ \circ & \circ & \circ & \circ & \circ & \circ & \circ \\ \circ & \circ & \circ & \circ & \circ & \circ & \circ \\ \circ & \circ & \circ & \circ & \circ & \circ & \circ \end{bmatrix} \begin{bmatrix} \circ & \circ & b_{1j} & \circ & \circ \\ \circ & \circ & b_{2j} & \circ & \circ \\ \circ & \circ & b_{3j} & \circ & \circ \\ \circ & \circ & \cdot & \circ & \circ \\ \circ & \circ & \cdot & \circ & \circ \\ \circ & \circ & \cdot & \circ & \circ \\ \circ & \circ & b_{nj} & \circ & \circ \end{bmatrix}.$$

Formula (4.9) kaže da jednostavno pomnožimo elemente $i-$tog retka matrice A sa odgovarajućim elementima $j-$tog stupca matrice B i produkte zbrojimo.

Primjer 4.5 *Ako su* $A \in \mathbf{R}^{3\times 2}$ *i* $B \in \mathbf{R}^{2\times 2}$

$$A = \begin{bmatrix} 1 & 3 \\ 0 & -1 \\ 5 & 1 \end{bmatrix}, \qquad B = \begin{bmatrix} 3 & -6 \\ 5 & 3 \end{bmatrix},$$

onda je $C = AB \in \mathbf{R}^{3\times 2}$ *i vrijedi*

$$C = \begin{bmatrix} 1\cdot 3 + 3\cdot 5 & 1\cdot(-6) + 3\cdot 3 \\ 0\cdot 3 + (-1)\cdot 5 & 0\cdot(-6) + (-1)\cdot 3 \\ 5\cdot 3 + 1\cdot 5 & 5\cdot(-6) + 1\cdot 3 \end{bmatrix} = \begin{bmatrix} 18 & 3 \\ -5 & -3 \\ 20 & -27 \end{bmatrix}.$$

Ako za B *uzmemo matricu stupac*

$$B = \begin{bmatrix} 2 \\ 4 \end{bmatrix},$$

onda je $C \in \mathbf{R}^3$ *i pritom je*

$$C = AB = \begin{bmatrix} 1 & 3 \\ 0 & -1 \\ 5 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 1\cdot 2 + 3\cdot 4 \\ 0\cdot 2 + (-1)\cdot 4 \\ 5\cdot 2 + 1\cdot 4 \end{bmatrix} = \begin{bmatrix} 14 \\ -4 \\ 14 \end{bmatrix}. \qquad \blacksquare$$

Neka je sada $A \in \mathbf{R}^{m\times n}$ i neka je $x \in \mathbf{R}^n$ jednostupčana matrica odnosno vektor. S obzirom da A ima onoliko stupaca koliko x ima redaka produkt Ax je definiran. Budući je

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & & \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix}, \qquad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix},$$

imamo

$$Ax = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \end{bmatrix} \in \mathbf{R}^m .$$

Ovo je zapravo lijeva strana sustava (4.1). Da bi sustav (4.1) u potpunosti opisali pomoću matrica, trebamo još napisati desnu stranu sustava pomoću vektora iz $\mathbf{R}^m$. Međutim, ako realne brojeve b_1, b_2,...,b_n smjestimo kao elemente vektora $b \in \mathbf{R}^m$, onda možemo sustav od m linearnih jednadžbi iz relacije (4.1) napisati jednom matričnom jednadžbom

$$Ax = b .$$

Ekonomičnost i efektivnost množenja matrica najbolje se ogleda u tome što nam omogućava sustav od m linearnih jednadžbi s n nepoznanica svesti na jednu matričnu jednadžbu koja uključuje samo operacije iz matrične algebre. To će nam omogućiti da taj sustav riješimo samo pomoću operacija iz matrične algebre.

Primjer 4.6 *Tri sustava s kojima smo započeli kurs linearne algebre, možemo u matričnoj notaciji ovako zapisati:*

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \qquad \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & 2 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix},$$

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & 2 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}.$$

■

Ako je $A \in \mathbf{R}^{m\times n}$ i $B \in \mathbf{R}^{n\times p}$, onda je produkt $AB \in \mathbf{R}^{m\times p}$ definiran. Postavlja se prirodno pitanje da li je, odnosno kada je BA definirano. Budući da je $B \in \mathbf{R}^{n\times p}$ i $A \in \mathbf{R}^{m\times n}$, produkt BA je definiran samo onda kada je broj stupaca matrice B jednak broju redaka matrice A, tj. ako vrijedi $p = m$. Prema tome, AB i BA su istovremeno definirani ako i samo ako je $A \in \mathbf{R}^{m\times n}$ i $B \in \mathbf{R}^{n\times m}$ za neke pozitivne cijele brojeve m i n. *To je pogotovo istina kad su A i B kvadratne matrice istog reda.*

Čak i ako obje matrice A i B pripadaju prostoru $\mathbf{R}^{n\times n}$, to još ne znači da vrijedi $AB = BA$. Na primjer, ako su

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix} \qquad \text{i} \qquad B = \begin{bmatrix} 1 & -1 \\ 2 & 0 \end{bmatrix},$$

tada je

$$AB = \begin{bmatrix} 5 & -1 \\ 5 & -3 \end{bmatrix}, \qquad BA = \begin{bmatrix} -2 & 1 \\ 2 & 4 \end{bmatrix}.$$

Prema tome, množenje matrica općenito **nije komutativno**. Ako za kvadratne matrice A i B vrijedi $AB = BA$, kažemo da **komutiraju**. Matrica $C(A,B) = AB - BA$ se naziva **komutator** matrica A i B. Jasno je da matrice komutiraju ako i samo ako je njihov komutator nul-matrica. Evo primjera matrica reda 3 koje komutiraju:

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 & 3 & 0 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{bmatrix}.$$

Svojstva množenja matrica su navedena u sljedećoj propoziciji.

Propozicija 4.7 *Za množenje matrica vrijede sljedeća svojstva:*

$$A(B+C) = AB + AC\,, \tag{4.10}$$
$$A(BC) = (AB)C\,, \tag{4.11}$$
$$\alpha(AB) = (\alpha A)B = A(\alpha B)\,. \tag{4.12}$$

Dokaz: Kako su na lijevoj i desnoj strani svake jednakosti - matrice, dovoljno je pokazati da je ij-ti element lijeve strane jednak ij-tom elementu desne strane.Prvo pokažimo svojstvo distributivnosti (4.10).

Koristeći distributivnost množenja prema zbrajanju u $\mathbf{R}$, dobivamo

$$\begin{aligned}
[A(B+C)]_{ij} &= \sum_k a_{ik}[B+C]_{kj} = \sum_k a_{ik}(b_{kj}+c_{kj}) \\
&= \sum_k (a_{ik}b_{kj}+a_{ik}c_{kj}) = \sum_k a_{ik}b_{kj} + \sum_k a_{ik}c_{kj} \\
&= [AB]_{ij} + [AC]_{ij} = [AB+AC]_{ij}.
\end{aligned}$$

Ovdje k prolazi vrijednosti od 1 do broja stupaca od A. Za drugu tvrdnju (asocijativnost matričnog produkta) koristimo asocijativnost množenja u $\mathbf{R}$,

$$\begin{aligned}
[A(BC)]_{ij} &= \sum_k a_{ik}[BC]_{kj} = \sum_k a_{ik} \sum_l b_{kl}c_{lj} \\
&= \sum_k \sum_l a_{ik}(b_{kl}c_{lj}) = \sum_k \sum_l a_{ik}b_{kl}c_{lj} \\
&= \sum_l \sum_k a_{ik}b_{kl}c_{lj} = \sum_l \left[\sum_k a_{ik}b_{kl}\right] c_{lj} \\
&= \sum_l [AB]_{il}c_{lj} = [(AB)C]_{ij}
\end{aligned}$$

Zadnja tvrdnja ide najlakše. Imamo

$$\begin{aligned}
[\alpha(AB)]_{ij} &= \alpha[AB]_{ij} = \alpha \sum_k a_{ik}b_{kj} \\
[(\alpha A)B]_{ij} &= \sum_k [\alpha A]_{ik}b_{kj} = \sum_k (\alpha a_{ik})b_{kj} \\
[A(\alpha B)]_{ij} &= \sum_k a_{ik}[\alpha B]_{kj} = \sum_k a_{ik}(\alpha b_{kj}).
\end{aligned}$$

Kako su zadnje sume u svim jednakostima jednake $\sum_k \alpha a_{ik}b_{kj}$ sve lijeve strane su međusobno jednake. ■

Jednu važnu klasu matrica čine **dijagonalne matrice**. Najpoznatiji primjer dijagonalne matrice je **jedinična** matrica

$$I = \begin{bmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & & & 1 \end{bmatrix} \in \mathbf{R}^{n\times n}\,. \tag{4.13}$$

Ona je jedinični element skupa $\mathbf{R}^{n\times n}$ s obzirom na množenje jer vrijedi

$$AI = IA = A \qquad \text{za svako } A \in \mathbf{R}^{n\times n}.$$

Dakle, ima istu ulogu kod množenja matrica, kao i broj 1 kod množenja realnih brojeva. Malo općenitije, vrijede relacije: $IA = A$ za svako $A \in \mathbf{R}^{n\times p}$ i $AI = A$ za svako $A \in \mathbf{R}^{p\times n}$. Jer je I reda n još se koristi oznaka I_n.

Općenitije, dijagonalna matrica je ona kod koje su svi izvandijagonalni (to su oni koji nisu dijagonalni) elementi jednaki nuli. Dijagonalnu matricu čiji su dijagonalni elementi $\alpha_1, \alpha_2, \ldots, \alpha_n$ označavamo s $\mathrm{diag}\,(\alpha_1, \alpha_2, \ldots, \alpha_n)$, dakle je

$$\mathrm{diag}\,(\alpha_1, \alpha_2, \ldots, \alpha_n) = \begin{bmatrix} \alpha_1 & 0 & 0 & \ldots & 0 \\ 0 & \alpha_2 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & & \vdots \\ 0 & 0 & & & \alpha_n \end{bmatrix}.$$

Odmah vidimo da produkt dviju netrivijalnih dijagonalnih matrica (npr. $\mathrm{diag}\,(1, 0, 2)$ i $\mathrm{diag}\,(0, -1, 0)$) može biti nul-matrica. Nul matrica također spada u dijagonalne. matrice. Dijegonalne matrice čine vektorski prostor koji je zatvoren i u odnosu na matrično množenje. Neka je $D = \mathrm{diag}\,(\alpha_1, \ldots, \alpha_n)$ i $A \in \mathbf{R}^{n\times p}$. Tada je i-ti redak od DA jednak i-tom retku od A pomnoženom sa α_i, pri čemu je $1 \leq i \leq n$. Slično, ako je $A \in \mathbf{R}^{p\times n}$, tada je i-ti stupac od AD jednak i-tom stupcu od A pomnoženom sa α_i. Ako su dijagonalni elementi od D pozitivni, operacija DA (AD) se naziva **skaliranje** redaka (stupaca). Dijagonalne matrice koje dobivamo množenjem jedinične matrice skalarom (tj. one oblika αI) nazivamo **skalarne** matrice. Ako je $D = \alpha I$, onda je $DA = \alpha A$, a također i $AD = \alpha A$, čim su dimenzije matrica takve da su produkti definirani.

Potencije kvadratne matrice A se ovako induktivno definiraju: $A^0 = I$, $A^{r+1} = A\,A^r$ za $r \geq 0$. Lako se pokaže (vidi zadatak 1) da vrijedi $A^pA^q = A^qA^p = A^{p+q}$ za sve nenegativne cijele brojeve p i q. Stoga je dobro definiran **matrični polinom** $p(A) = \alpha_k A^k + \alpha_{k-1}A^{k-1} + \cdots + \alpha_1 A + \alpha_0 I$, pri čemu su $\alpha_0, \ldots, \alpha_k$ realni brojevi.

Množenje matrica katkad daje iznenađujuće rezultate. Npr. ako je

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

onda je

$$A^2 = A \cdot A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

i

$$A^3 = A \cdot A^2 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = O\,.$$

Dakle, postoje 3×3 matrice za koje je $A \neq O$, $A^2 \neq O$ ali je $A^3 = O$. Na slični način bismo mogli pokazati da postoje $n \times n$ matrice za koje je prvih $n-1$ potencija različito od O, ali vrijedi $A^n = O$. Ako je $A \neq O$ i $A^k = O$ za neko $k \in \mathbf{N}$, tada se A naziva **nilpotentna** matrica.

Pokažimo još jedno iznenađujuće svojstvo matričnog množenja. Ako je

$$J = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

dobivamo

$$J^2 = JJ = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} = (-1) \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

tj.

$$J^2 = -I\,.$$

U aritmetici ne postoji realan broj čiji kvadrat je -1. Zato i jesu uvedeni kompleksni brojevi. S njima možemo riješiti npr. jednadžbu $x^2 = -1$. U matričnoj algebri postoji realna matrica, čiji kvadrat je jednak $-I$.

Postoji još jedna vrlo korisna operacija na matricama. Naziva se operacijom transponiranja.

Definicija 4.8 *Neka je* $A \in \mathbf{R}^{m\times n}$. *Matrica* $A^T \in \mathbf{R}^{n\times m}$ *se naziva* **transponirana** *matrica* A, *ako je svaki redak od* A^T *jednak odgovarajućem stupcu matrice* A.

Prema tome, transponiranu matricu dobivamo tako da stupce (retke) matrice zamijenimo njenim retcima (stupcima). Ako je

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

onda je

$$A^T = \begin{bmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \vdots & & & \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{bmatrix}.$$

Kraće to zapisujemo: ako je $A = (a_{ij})$, onda je $A^T = (a_{ji})$. Na primjer, ako je

$$A = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix}, \qquad \text{onda je} \qquad A^T = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}.$$

Operacija transponiranja je unarna operacija koja matrici pridružuje njoj transponiranu matricu. Da bi pokazali korisnost te operacije, vratimo se problemu određivanja svih $x \in \mathrm{V}_n$ za koje vrijedi

$$(a \mid x) = 0\,, \tag{4.14}$$

gdje je $a \in \mathrm{V}_n$ zadani vektor. Već smo vidjeli da svaki vektor iz V_n možemo identificirati sa vektorom (matricom stupcem) iz $\mathbf{R}^n$. Tu identifikaciju ćemo primijeniti na oba vektora a i x iz V_n. Označimo li i pripadne jednostupčane matrice također sa a i x, imamo

$$a = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}, \qquad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Jer

$$a^T = \begin{bmatrix} a_1 & a_2 & \dots & a_n \end{bmatrix},$$

ima stupaca koliko x ima redaka, produkt $a^T x$ je definiran i vrijedi

$$a^T x = a_1 x_1 + a_2 x_2 + \ldots + a_n x_n\,. \tag{4.15}$$

Iz definicije skalarnog produkta u V_n znamo da je desna strana jednadžbe (4.15) skalarni produkt u V_n. Kako je na isti način definiran i skalarni produkt u $\mathbf{R}^n$, $a^T x$ je skalarni produkt vektora a i x iz $\mathbf{R}^n$. To možemo ovako zapisati

$$(a \mid x) = a^T x \quad a\,, x\, \in \mathbf{R}^n\,, \tag{4.16}$$

pri čemu je $(\cdot \mid \cdot)$ oznaka za skalarni produkt u $\mathbf{R}^n$. Sada možemo sustav od p jednadžbi oblika

$$(a_k \mid x) = 0\,, \quad 1 \leq k \leq p, \tag{4.17}$$

zamijeniti odgovarajućim sustavom matričnih jednadžbi

$$a_k^T x = 0\,, \quad 1 \leq k \leq p\,. \tag{4.18}$$

Ako pišemo

$$a_k^T = \begin{bmatrix} a_{k1} & a_{k2} & \dots & a_{kn} \end{bmatrix},$$

i uvedemo matricu A pomoću redaka:

$$A = \begin{bmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_p^T \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & & \\ a_{p1} & a_{p2} & \dots & a_{pn} \end{bmatrix}, \tag{4.19}$$

onda sustav jednadžbi (4.18) prelazi u matričnu jednadžbu

$$Ax = 0\,, \tag{4.20}$$

gdje je $0 \in \mathbf{R}^p$ nul vektor.

Kako dvije matrice iz $\mathbf{R}^{n\times n}$ uvijek možemo množiti i dobivamo rezultat koji je opet u $\mathbf{R}^{n\times n}$, čitatelj bi mogao pomisliti, da se matrice možda mogu i dijeliti. Ponekad se matrice zaista mogu dijeliti (npr. matrice reda 1 koje su realni brojevi ili skalarne matrice), ali to nikako nije opći slučaj. Primijetimo da je npr.

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & \pi \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$

što znači, da ne postoji matrica koja podijeljena sa $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ daje $\begin{bmatrix} 0 & \pi \\ 0 & 0 \end{bmatrix}$. Promotrimo zato matričnu jednadžbu

$$AX = C$$

gdje su A i C eksplicitno zadane. Pitanje je možemo li "dijeliti" C sa A da bi dobili X? Da bi to postigli treba nam matrica W sa svojstvom

$$WA = I\,.$$

Naime, onda imamo

$$X = I\,X = (WA)X = W(AX) = WC\,.$$

Pritom smo iskoristili svojstvo asocijativnosti produkta matrica i svojstvo $WA = I$. Zapravo, načinili smo ono što radimo i sa skalarnom jednadžbom $\alpha x = \gamma$. Množimo ju s $\omega = 1/\alpha$. Na problem nalaženja matrice W ćemo se vratiti kasnije.

Nabrojimo sada glavna svojstva operacije transponiranja.

Propozicija 4.9 *Operacija transponiranja ima sljedeća svojstva,*

$$(A^T)^T = A\,, \tag{4.21}$$
$$(A+B)^T = A^T + B^T\,, \tag{4.22}$$
$$(AB)^T = B^T A^T\,. \tag{4.23}$$

Dokaz: Prvo treba provjeriti (za svaku jednakost) da je dimenzija matrice na lijevoj strani jednakosti jednaka dimenziji matrice na desnoj strani. Tu provjeru ostavljamo čitatelju. Zatim treba pakazati da je proizvoljni element matrice na lijevoj strani jednak odgovarajućem elementu matrice na desnoj strani.

Kako je ij-ti element matrice A^T jednak a_{ji}, mora biti ij-ti element matrice $(A^T)^T$ baš a_{ij}. No, to je ij-ti element od A, pa je prva jednakost (4.21) dokazana. Slično, ij-ti element od $(A+B)^T$ je $a_{ji}+b_{ji}$, a to je također ij-ti element desne strane u jednakosti (4.22). Konačno, ij-ti element od $(AB)^T$ je ji-ti element od AB, tj.

$$[(AB)^T]_{ij} = [(AB)]_{ji} = \sum_k a_{jk}b_{ki},$$

gdje k prolazi vrijednosti od 1 do broja stupaca od A (ili redaka od B). S druge strane, ij-ti element od B^TA^T je

$$[B^TA^T]_{ij} = \sum_k [B^T]_{ik}[A^T]_{kj} = \sum_k b_{ki}a_{jk} = \sum_k a_{jk}b_{ki},$$

čime je dokazana i zadnja jednakost. ∎

Primijetimo da su u zadnjem svojstvu matrice A i B zamijenile poredak.

Koristeći propoziciju 4.9 (4.23) lako se pokaže da osim "lijeve distributivnosti" vrijedi i "desna distributivnost". Najme, uz pomoć relacija (4.23), (4.21), (4.21) i (4.10) jednostavno dobivamo

$$\begin{aligned} (B+C)A &= \left[[(B+C)A]^T\right]^T = \left[A^T(B+C)^T\right]^T \\ &= \left[A^T(B^T+C^T)\right]^T = \left[A^TB^T + A^TC^T\right]^T \\ &= \left[(BA)^T + (CA)^T\right]^T = \left[(BA+CA)^T\right]^T \\ &= BA + CA\,. \end{aligned} \tag{4.24}$$

Relaciju (4.16) možemo iskoristiti za pisanje norme vektora iz $\mathbf{R}^n$ pomoću operacije množenja matrica. Budući da je $\|x\| = \sqrt{(x \mid x)}$, imamo

$$\|x\| = \sqrt{x^Tx}\,. \tag{4.25}$$

Formula (4.25) i jednakost (4.23) omogućavaju npr. napisati izraz za normu od Ax, gdje je $A \in \mathbf{R}^{m\times n}$ i $x \in \mathbf{R}^n$:

$$\|Ax\| = \sqrt{x^TA^TAx}\,. \tag{4.26}$$

Zadatak 4.10 *Neka su $A \in \mathbf{R}^{m\times n}$, $B \in \mathbf{R}^{n\times p}$. Dokažite sljedeće tvrdnje.*

(i) Ako je $B = [b_1, \ldots, b_p]$ *stupčana particija matrice* B*, tada je* $AB = [Ab_1, \ldots, Ab_p]$ *stupčana particija matrice* AB*.*

(ii) Za retčane particije matrica A *i* AB *vrijedi implikacija*

$$A = \begin{bmatrix} a_1^T \\ \vdots \\ a_m^T \end{bmatrix} \Longrightarrow AB = \begin{bmatrix} a_1^T B \\ \vdots \\ a_m^T B \end{bmatrix}.$$

(iii) Ako je $A = [a_1, \ldots, a_n]$ *stupčana particija matrice* A *i* $x \in \mathbf{R}^n$*, tada je* $Ax = x_1 a_1 + \ldots x_n a_n$*, gdje su* $x_1, \ldots, x_n$ *komponente vektora* x*.*

(iv) Ako je $y \in \mathbf{R}^n$*,* $y^T = [y_1, \ldots, y_n]$ *i ako je*

$$B = \begin{bmatrix} b_1^T \\ \vdots \\ b_n^T \end{bmatrix} \tag{4.27}$$

retčana particija od B*, tada je* $y^T B = y_1 b_1^T + \cdots + y_n b_n^T$*.*

(v) Ako je $A = [a_1, \ldots, a_n]$ *stupčana particija od* A *i ako je (4.27) retčana particija od* B*, tada je*

$$AB = a_1 b_1^T + \ldots + a_n b_n^T \,.$$

Pritom je svaka matrica $a_i b_i^T \in \mathbf{R}^{m \times p}$*.* ■

Za svaki par prirodnih brojeva m i n, struktura $(\mathbf{R}^{m\times n}, +, \cdot)$ je zatvorena u odnosu na operacije $+$ i $\cdot$. Množenje matrica iz $\mathbf{R}^{m\times n}$ je definirano tek ako je $m = n$. Također je važno da je skup $\mathbf{R}^{n\times n}$ zatvoren u odnosu na matrično množenje. Označimo li privremeno operaciju množenja matrica sa $\times$, možemo pisati $(\mathbf{R}^{n\times n}, +, \cdot, \times)$, pri čemu je $(\mathbf{R}^{n\times n}, +, \cdot)$ vektorski prostor, a množenje $\times$ zadovoljava svojstva (4.10), (4.24), (4.11) i (4.12). Takva struktura se zove **algebra**. Kako u $\mathbf{R}^{n\times n}$ postoji i neutralni element I za množenje $\times$, kaže se da je $(\mathbf{R}^{n\times n}, +, \cdot, \times)$ algebra s jedinicom. Dakle, **skup matrica reda n čini algebru s jedinicom**. Ona općenito nije komutativna. Na njoj je definirana i unarna operacija transponiranja koja zadovoljava svojstva (4.21)–(4.23).

Promotrimo još skup M_R svih realnih matrica,

$$M_R = \bigcup_{m,n\in\mathbf{N}} \mathbf{R}^{m\times n}.$$

Neka $\times$ označava operaciju množenja matrica, a T operaciju transponiranja matrica. Skup M_R je sigurno zatvoren u odnosu na operaciju množenja matrice sa skalarom. Isto tako je zatvoren u odnosu na operaciju transponiranja. Međutim, ostale dvije operacije nisu definirane za svaki par matrica iz M_R. Ipak, postoji lijepo svojstvo skupa M_R, da je zatvoren s obzirom na te dvije operacije uvijek kad je rezultat tih dviju matrica definiran.

Kompleksne matrice

Sa $\mathbf{C}^{m\times n}$ smo označili skup kompleksnih $m \times n$ matrica. U taj skup je uključen i skup realnih matrica $\mathbf{R}^{m\times n}$, pa je skup $\mathbf{C}^{m\times n}$ veći od $\mathbf{R}^{m\times n}$.

Zbroj dviju matrica $A, B \in \mathbf{C}^{m\times n}$ je definiran na isti način kao i u $\mathbf{R}^{m\times n}$, dakle: $C = A + B$ ako je $c_{ij} = a_{ij} + b_{ij}$ za sve $1 \leq i \leq m$, $1 \leq j \leq n$, gdje su $A = (a_{ij})$, $B = (b_{ij})$ i $C = (c_{ij})$. Kako su a_{ij} i b_{ij} općenito kompleksni brojevi, operacija $+$ na nivou matričnih elemenata je operacija zbrajanja kompleksnih brojeva. Na isti način kao i prije, lako se pokaže da je $(\mathbf{C}^{m\times n}, +)$ Abelova aditivna grupa. Neutralni element je matrica $O \in \mathbf{C}^{m\times n}$ čiji su elementi same nule. Inverzni (zapravo suprotni) element od $A \in \mathbf{C}^{m\times n}$ je matrica $-A = (-a_{ij})$. Uočimo da je O također u $\mathbf{R}^{m\times n}$. Što više, $(\mathbf{R}^{m\times n}, +)$ je netrivijalna podgrupa od $(\mathbf{C}^{m\times n}, +)$.

Kod množenja matrica iz $\mathbf{C}^{m\times n}$ skalarom, moramo dozvoliti množenje kompleksnim brojevima. Neka je $A \in \mathbf{C}^{m\times n}$ i $\omega \in \mathbf{C}$. Tada je $B = \omega \cdot A$ (ili kraće $B = \omega A$) ako je $b_{ij} = \omega a_{ij}$ za sve i, j. Ova operacija ima ista svojstva kao i operacija množenja realnim skalarom u $\mathbf{R}^{m\times n}$. Stoga $\mathbf{C}^{m\times n}$ uz operaciju zbrajanja i množenja skalarom iz $\mathbf{C}$ postaje vektorski prostor (nad poljem kompleksnih brojeva), koji označavamo s $(\mathbf{C}^{m\times n}, +, \cdot)$. Uočimo da $(\mathbf{R}^{m\times n}, +, \cdot)$ nije vektorski potprostor od $(\mathbf{C}^{m\times n}, +, \cdot)$ jer ti vektorski prostori nisu definirani nad istim poljem.

Elemente od $\mathbf{C}^{n\times 1}$ opet zovemo matrice stupci, jednostupčane matrice ili vektori. Oznaku $\mathbf{C}^{n\times 1}$ zamijenjujemo oznakom $\mathbf{C}^n$, a za sam vektorski prostor tih vektora koristimo oznaku $(\mathbf{C}^n, +, \cdot)$. Kao i prije, lako se pokaže da su vektorski prostori $(\mathbf{C}_n, +, \cdot)$ i $(\mathbf{C}^n, +, \cdot)$ izomorfni.

Operacija množenja kompleksnih matrica definirana je samo za parove matrica kod kojih lijevi faktor ima onoliko stupaca koliko desni faktor ima redaka. Množenje je definirano na isti način kao kod realnih matrica (vidi definiciju 4.4). Neutralni element s obzirom na matrično množenje je opet identiteta, tj. jedinična matrica I (koja je i u $\mathbf{R}^{n\times n}$ i u $\mathbf{C}^{n\times n}$, ako je reda n). Množenje matrica ima ista svojstva (4.10)–(4.12), uključujući i svojstvo (4.24) koje se dokazuje na isti način. Najme, operacija tranponiranja je definirana za svaku kompleksnu matricu na isti način kao i u slučaju realnih matrica. Time smo zapravo pokazali da je $(\mathbf{C}^{n\times n}, +, \cdot, \times)$ algebra s jedinicom, gdje smo načas matrično množenje označili s $\times$.

Glavna razlika u definicijama vezanim uz algebre $\mathbf{R}^{n\times n}$ i $\mathbf{C}^{n\times n}$ dolazi u momentu kad se želi u njima definirati skalarni produkt odnosno norma. Sjetimo se, u V_n odnosno u $\mathbf{R}^n$ skalarni produkt vektora a i b bio je definiran formulom $a_1b_1 + \cdots + a_nb_n$, gdje su $a_i\,, b_i$ komponente od $a\,, b$, respektivno. Zato se u $\mathbf{R}^n$ skalarni produkt mogao zapisati kao a^Tb. U $\mathbf{C}_n$ je skalarni produkt definiran formulom

$$(a \mid b) = a_1\bar{b}_1 + \cdots + a_n\bar{b}_n,$$

pa zato u $\mathbf{C}^n$ ne možemo pisati $(a \mid b) = a^Tb$, $a, b \in \mathbf{C}^n$. Da bi ipak $(a \mid b)$ napisali pomoću operacije matričog množenja moramo uvesti operaciju komleksnog konjugiranja i kompleksnog transponiranja.

Ako je $z = z_1 + iz_2$ kompleksni broj, tada je $\bar{z} = z_1 - iz_2$ kompleksno konjugirani broj. Na slični način definiramo i kompleksno konjugiranu matricu, a onda i kompleksno transponiranu matricu.

Definicija 4.11 *Neka je* $Z \in \mathbf{C}^{m\times n}$ *i* $Z = Z_1 + iZ_2$*, pri čemu su* $Z_1\,, Z_2 \in \mathbf{R}^{m\times n}$*. Tada se* $Z_1 - iZ_2$ *zove* **kompleksno konjugirana** *matrica i označava sa* $\overline{Z}$*.* **Kompleksno transponirana** *ili* **hermitski adjungirana** *matrica* Z *je matrica* $Z^* = Z_1^T - iZ_2^T$*.*

Uočimo da su obje operacije definirane za svaku kompleksnu matricu i pritom je $\overline{Z} \in \mathbf{C}^{m\times n}$ i $Z^* \in \mathbf{C}^{n\times m}$ ako je $Z \in \mathbf{C}^{m\times n}$. Operacije kompleksnog konjugiranja i kompleksnog transponiranja imaju sljedeća svojstva.

Propozicija 4.12 *Neka su $Z, W \in \mathbf{C}^{m\times n}$ i $\alpha \in \mathbf{C}$. Tada je*

(i) $$Z^* = \left(\overline{Z}\right)^T = \overline{Z^T}\,, \qquad \overline{\overline{Z}} = Z\,, \qquad Z^{**} = Z$$

(ii) $$\overline{Z+W} = \overline{Z} + \overline{W}\ , \qquad (Z+W)^* = Z^* + W^*$$

(iii) $$\overline{\alpha Z} = \bar{\alpha}\overline{Z}\,, \qquad (\alpha Z)^* = \bar{\alpha} Z^*$$

(iv) $$\overline{Z\,W} = \overline{Z}\,\overline{W}\ , \qquad (Z\,W)^* = W^*\,Z^*.$$

Dokaz: Sve tvrdnje izlaze iz osnovnih svojstava konjugiranja kompleksnih brojeva: $\bar{\bar{z}} = z$, $\overline{\alpha+\beta} = \bar{\alpha} + \bar{\beta}$, $\overline{\alpha\beta} = \bar{\alpha}\bar{\beta}$ i osnovnih svojstava operacije transponiranja. ■

Uočimo da vrijedi

$$\begin{aligned} (a \mid b) &= a_1\bar{b}_1 + \cdots + a_n\bar{b}_n \\ &= \bar{b}_1 a_1 + \cdots + \bar{b}_n a_n = \bar{b}^T a = b^* a\,, \end{aligned}$$

pa smo uspjeli skalarni produkt prikazati pomoću produkta matrica. Zato za normu vektora vrijedi

$$\|a\| = \sqrt{|\ a_1\ |^2 + \cdots + |\ a_n\ |^2} = \sqrt{a^* a}.$$

Uočite da tvrdnje (i) – (v) iz zadatka 4.10 vrijede i za kompleksne matrice.

Lekcija 5

Linearno nezavisni vektori

U ovom poglavlju pripremamo teren za rješavanje homogenog sustava od p linearnih jednadžbi s n nepoznanica, koji pišemo u obliku

$$Ax = 0\,. \tag{5.1}$$

Pritom su $A \in \mathbf{R}^{p\times n}$, $x \in \mathbf{R}^n$, a $0 \in \mathbf{R}^p$ je nul vektor. Prije nego sustav riješimo, trebamo uvesti pojmove minimalnog razapinjućeg skupa, dimenzije vektorskog prostora i ranga matrice. Oni su potrebni u opisu skupa svih rješenja sustava (5.1), a također i u uvjetima koji se postavljaju na matricu.

Naš polazni sustav se obično zapisuje u obliku

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= 0 \\ \cdots & \\ a_{p1}x_p + a_{p2}x_2 + \ldots + a_{pn}x_n &= 0\,. \end{aligned} \tag{5.2}$$

Ponovimo još jednom kako smo od sustava (5.2) došli do zapisa (5.1). Sa $a_i \in \mathbf{R}^n$ smo označili vektor čije su komponente koeficijenti $a_{i1}, \ldots, a_{in}$ i-te jednadžbe sustava. Zatim smo postavili zahtjev da x bude okomit na sve vektore a_i, $1 \le i \le p$. Kao skupna informacija o svim komponentama vektora a_i pojavila se matrica A. Uz oznake

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{bmatrix}, \quad a_i = \begin{bmatrix} a_{i1} \\ a_{i2} \\ \cdots \\ a_{in} \end{bmatrix}, \quad A = \begin{bmatrix} a_1^T \\ a_2^T \\ \cdots \\ a_n^T \end{bmatrix}, \tag{5.3}$$

okomitost vektora x na sve vektore a_i, može se zapisati kao

$$\begin{bmatrix} 0 \\ 0 \\ \cdots \\ 0 \end{bmatrix} = \begin{bmatrix} a_1^T x \\ a_2^T x \\ \cdots \\ a_p^T x \end{bmatrix} = \begin{bmatrix} a_1^T \\ a_2^T \\ \cdots \\ a_p^T \end{bmatrix} x = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \cdots & \cdots & & \cdots \\ a_{p1} & a_{p2} & \ldots & a_{pn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{bmatrix}$$

ili kraće $Ax = 0$.

5.1 Minimalni razapinjući skup

Na isti način kao i u propoziciji 3.4, lako se pokaže da skup svih rješenja sustava linearnih jednadžbi $Ax = 0$ čini vektorski potprostor. Taj potprostor je razapet vektorima koji su okomiti na sve vektore $a_1, a_2, \ldots, a_p$, pri čemu su a_i određeni relacijom (5.3). Kako razapinjućih vektora za neki potprostor može biti vrlo mnogo, postavlja se pitanje kako za dani potprostor pronaći razapinjući skup koji sadrži minimalni broj vektora.

Primjer 5.1 *Neka je potprostor $\mathcal{X} \subset \mathbf{R}^3$ razapet vektorima*

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 5 \\ -3 \\ 0 \end{bmatrix},$$

i neka je $\mathcal{Y} \subset \mathbf{R}^3$ razapet vektorima

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}.$$

U oba potprostora nalaze se svi vektori iz $\mathbf{R}^3$ čija je treća koordinata jednaka nuli. Zato je $\mathcal{Y} = \mathcal{X}$. Prema tome su vektori

$$\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 5 \\ -3 \\ 0 \end{bmatrix}$$

u definiciji potprostora $\mathcal{X}$ suvišni. ■

Svaki skup vektora vektorskog prostora razapinje neki potprostor, koji može biti i cijeli vektorski prostor. Mi želimo izdvojiti iz danog skupa vektora jedan njegov podskup, kao u primjeru 5.1, koji razapinje isti potprostor, ali koji je u nekom smislu minimalan.

Definicija 5.2 *Neka je $\mathcal{S}$ podskup vektorskog prostora X.* **Minimalni razapinjući podskup** *od $\mathcal{S}$ je svaki podskup $\mathcal{M}$ od $\mathcal{S}$, koji zadovoljava sljedeća dva svojstva:*

1. $L(\mathcal{M}) = L(\mathcal{S})$,
2. *Niti jedan pravi podskup skupa $\mathcal{M}$ ne razapinje $L(\mathcal{S})$.*

$\mathcal{M}$ se još zove **minimalni razapinjući skup** *od $L(\mathcal{S})$.* ■

Npr. ako je $\mathcal{S} = \{a_1, a_2, \ldots, a_p\}$, tada je $\mathcal{M} = \{u_1, u_2, \ldots, u_r\}$, $r \leq p$, gdje je svaki u_i neki od vektora iz $\mathcal{S}$. Sada se uvjet 1. iz definicije 5.2 zapisuje kao

$$L(u_1, \ldots, u_r) = L(a_1, \ldots, a_p).$$

Uočimo da u zadnjoj jednakosti poredak vektora $a_1, \ldots, a_p$ (odnosno $u_1, \ldots, u_r$) nije važan, jer je $L(a_1, \ldots, a_p)$ tek kraća oznaka za $L(\mathcal{S}) = L(\{a_1, \ldots, a_p\})$.

Ako imamo skup vektora koji razapinje određeni vektorski potprostor, pitanje je kako pronaći neki njegov minimalni razapinjući podskup. Trebat će nam algoritam koji polazeći od skupa $\{a_1, \ldots, a_p\}$ daje na izlazu minimalni razapinjući skup $\{u_1, \ldots, u_r\}$, $r \leq p$. U konstrukciji algoritma koristi se sljedeći rezultat.

Propozicija 5.3 *Neka je X vektorski prostor i $x_1, \ldots, x_k$ vektori iz X. Ako je y linearna kombinacija vektora x_i, $1 \leq i \leq k$, tada je*

$$L(y, x_1, \ldots, x_k) = L(x_1, \ldots, x_k) \tag{5.4}$$

Dokaz: Uvijek vrijedi $L(x_1, \ldots, x_k) \subseteq L(y, x_1, \ldots, x_k)$, bez obzira kakav je y. To je zato jer se svaka linearna kombinacija iz $L(x_1, \ldots, x_k)$ može prikazati kao neka linearna kombinacija iz $L(y, x_1, \ldots, x_k)$. Naime, svaku linearnu kombinaciju $\beta_1 x_1 + \cdots + \beta_k x_k$ možemo zapisati kao $0 \cdot y + \beta_1 x_1 + \cdots + \beta_k x_k$.

Dokaz će biti gotov ako pokažemo da za $y \in L(x_1, \ldots, x_k)$ vrijedi i obrnuta inkluzija $L(y, x_1, \ldots, x_k) \subseteq L(x_1, \ldots, x_k)$.

Neka je

$$y = \alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_k x_k. \tag{5.5}$$

Pođimo od proizvoljnog elementa $z \in L(y, x_1, \ldots, x_k)$ i pokažimo da je on u $L(x_1, \ldots, x_k)$. Za z postoje realni skalari β_i, $0 \leq i \leq k$, tako da vrijedi

$$z = \beta_0 y + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k.$$

Uvrštavajući izraz na desnoj strani jednakosti (5.5) umjesto y, dobivamo

$$\begin{aligned} z &= \beta_0(\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_k x_k) + \beta_1 x_1 + \cdots + \beta_k x_k \\ &= (\beta_1 + \beta_0 \alpha_1) x_1 + \cdots + (\beta_k + \beta_0 \alpha_k) x_k \,. \end{aligned}$$

Dakle je $z \in L(x_1, \ldots, x_k)$, pa zaključujemo da je $L(y, x_1, \ldots, x_k) \subseteq L(x_1, \ldots, x_k)$. ■

Jasno je da y u relaciji (5.4) može stajati ispred ili iza bilo kojeg vektora x_i. Opišimo sada algoritam za određivanje minimalnog razapinjućeg skupa.

Algoritam 5.4 *Polazimo od p vektora $x_1, x_2, \ldots, x_p$ vektorskog prostora X. Algoritam određuje jedan minimalni razapinjući podskup skupa $\{x_1, \ldots, x_p\}$. Algoritam koristi pomoćne vektore $u_1, \ldots, u_p$.*

0. *Inicijaliziraj brojač k i vektore u_i:*

$$k = p, \quad u_1 = x_1, \ldots, u_p = x_p;$$

1. *Ustanovi postoje li skalari $\alpha_1, \ldots, \alpha_k$, koji nisu svi jednaki nuli i za koje vrijedi*

$$\alpha_1 u_1 + \alpha_2 u_2 + \ldots + \alpha_k u_k = 0 \,. \tag{5.6}$$

Ako postoje, prijeđi na korak 2., inače prijeđi na korak 3.

2. *Neka je $\alpha_j \neq 0$.*
(a) ako je $j < k$, prenumeriraj vektore: $u_j = u_{j+1}, \ldots, u_{k-1} = u_k$ i prijeđi na (b); ako je $j = k$ odmah prijeđi na (b);
(b) Stavi $k = k - 1$ i vrati se na korak 1.

3. *Stavi* $r = k$, $\mathcal{M} = \{u_1, \ldots, u_r\}$. ■

U koraku 2 (a) algoritma važan je redoslijed: prvo u_{j+1} postaje u_j, zatim (ako je $j + 2 \leq k$) u_{j+2} postaje u_{j+1} itd.

Koraci 1. i 2. zahtijevaju obrazloženje. U koraku 1. zahtijeva se provjera egzistencije netrivijale linearne kombinacije koja je nul-vektor. To je za veće n netrivijalan problem za čije rješavanje postoje posebne metode (koje izlaze izvan ovog kursa linearne algebre). Međutim, za dobru definiranost algoritma bitno je da je pitanje egzistencije netrivijalne anihilirajuće linearne kombinacije uvijek rješivo: ili takve linearna kombinacija postoji ili ne postoji.

Obrazložimo kako algoritam funkcionira. Ako je u koraku 2.(a) $\alpha_j \neq 0$, tada se u_j može izraziti kao linearna kombinacija ostalih vektora:

$$u_j = -\frac{1}{\alpha_j}(\alpha_1 u_1 + \cdots + \alpha_{j-1} u_{j-1} + \alpha_{j+1} u_{j+1} + \cdots + \alpha_k u_k)$$

pa je po propoziciji 5.3,

$$L(u_1, \ldots u_k) = L(u_1, \ldots u_{j-1}, u_{j+1}, \ldots, u_k). \tag{5.7}$$

Dakle, iz skupa $\{u_1, \ldots, u_k\}$ smijemo izbaciti u_j, a da skup $\{u_1, \ldots u_{j-1}, u_{j+1}, \ldots, u_k\}$ i dalje razapinje isti potprostor. Da bi u novom skupu vektora,indekse brojali uzastopnim brojevima, smanjujuemo ih nakon indeksa $j - 1$ za jedan. Kako se broj vektora smanjio za jedan, ažurira se i broj vektora k.

Primjer 5.5 *Primijenite algoritam 5.6 na vektore iz primjera 5.1. U algoritmu imate mogućnost biranja koeficijenta* $\alpha_j \neq 0$. *Odabirajte* j *tako da postepeno izbacite sve vektore osim prva dva.* ■

Iz algoritma 5.4 vidimo da se postupak izbacivanja vektora ponavlja tako dugo dok se ne dobije skup vektora $\{u_1, \ldots, u_k\}$ za koji jednadžba

$$\alpha_1 u_1 + \cdots + \alpha_k u_k = 0 \tag{5.8}$$

u nepoznanicama $\alpha_1, \ldots, \alpha_k$ ima samo trivijalno rješenje: $\alpha_1 = \cdots = \alpha_k = 0$.

Propozicija 5.6 *Skup vektora* $\mathcal{M}$, *dobiven algoritmom 5.4 je minimalni razapinjući podskup skupa* $\{x_1, \ldots, x_p\}$.

Dokaz: Iz opisanog postupka je jasno da je $\mathcal{M} \subseteq \{x_1, \ldots, x_p\}$. Također, zbog jednakosti (5.7), nakon svakog prolaza kroz petlju u algoritmu, zaključujemo da vrijedi $L(u_1, \ldots, u_k) = L(x_1, \ldots, x_p)$. Stoga će vrijediti i $L(\mathcal{M}) = L(x_1, \ldots, x_p)$.

Jasno je i da niti jedan od vektora iz $\mathcal{M}$ nije nul-vektor, jer bi tada jednadžba (5.6) imala netrivijalno rješenje u skalarima α_i, pa bi se još jedan vektor u_j mogao izbaciti.

Preostaje pokazati da niti jedan pravi podskup od $\mathcal{M}$ ne razapinje $L(x_1, \ldots, x_p)$. Pretpostavimo suprotno. Tada bi iz skupa $\mathcal{M} = \{u_1, \ldots, u_r\}$ mogli maknuti neke vektore, označimo ih sa $\{u_{j_1}, \ldots, u_{j_k}\}$, pri čemu bi vrijedilo

$$L(\mathcal{M} \setminus \{u_{j_1}, \ldots, u_{j_k}\}) = L(x_1, x_2, \ldots, x_p). \tag{5.9}$$

Zbog relacije (5.9) vrijedi

$$\begin{aligned} L(x_1, x_2, \ldots, x_p) &= L(\mathcal{M} \setminus \{u_{j_1}, \ldots, u_{j_k}\}) \\ &\subseteq L(\mathcal{M} \setminus \{u_{j_1}\}) \\ &\subseteq L(\mathcal{M}) \\ &= L(x_1, x_2, \ldots, x_p), \end{aligned}$$

pa zključujemo

$$L(\mathcal{M} \setminus \{u_{j_1}\}) = L(x_1, x_2, \ldots, x_p).$$

Dakle, postoji neki vektor $u_j \in \mathcal{M}$ (npr. u_{j_1}), takav da je

$$L(x_1, x_2, \ldots, x_p) = L(u_1, \ldots, u_{j-1}, u_{j+1}, \ldots, u_r). \tag{5.10}$$

Budući da je $u_j \in L(x_1, x_2, \ldots, x_p)$, on je zbog relacije (5.10) i u $L(u_1, \ldots, u_{j-1}, u_{j+1}, \ldots, u_r)$, pa možemo pisati

$$u_j = \beta_1 u_1 + \ldots + \beta_{j-1} u_{j-1} + \beta_{j+1} u_{j+1} + \ldots + \beta_r u_r \,,$$

za neke skalare β_i, $1 \le i \le r$, $i \ne j$. Iz toga slijedi

$$\beta_1 u_1 + \ldots + \beta_{j-1} u_{j-1} + (-1) u_j + \beta_{j+1} u_{j+1} + \ldots + \beta_r u_r = 0,$$

a to je protivno pretpostavci da jednadžba (5.8) ima samo trivijalno rješenje $\alpha_1 = \ldots = \alpha_r = 0$. Dobivena kontradikcija pokazuje da se skup $\mathcal{M}$ ne može smanjiti, a da još uvijek razapinje $L(x_1, \ldots, x_p)$. $\blacksquare$

5.2 Linearno nezavisni vektori

Poželjno je vektorima, za koje jednadžba (5.8) ima samo trivijalno rješenje, dati neko ime.

Definicija 5.7 *Neka su x_1,...,x_r elementi vektorskog prostora X. Ako jednadžba*

$$\alpha_1 x_1 + \cdots + \alpha_r x_r = 0$$

u nepoznanicama $\alpha_1, \ldots, \alpha_r$ ima samo trivijalno rješenje $\alpha_1 = \cdots = \alpha_r = 0$, tada se vektori x_1, ...,x_r nazivaju **linearno nezavisnim**. *Vektori koji nisu linearno nezavisni nazivaju se* **linearno zavisni**. *Skup vektora $\{x_1, \ldots, x_r\}$ je linearno nezavisan (zavisan) ako su vektori $x_1, \ldots, x_r$ takvi.* ■

Najjednostavnija svojstva skupa linearno nezavisnih vektora dana su u sljedećoj propoziciji.

Propozicija 5.8 *Skup linearno nezavisnih vektora $\{u_1, \ldots, u_r\}$ ne sadrži nul-vektor. Ako je skup $\{u_1, \ldots, u_r\}$ linearno nezavisan, tada je i svaki njegov podskup linerano nezavisan. Ako je skup $\{u_1, \ldots, u_r\}$ linearno zavisan, tada je svaki njegov nadskup također linearno zavisan.*

Dokaz: Kad bi skup $\{u_1, \ldots, u_r\}$ sadržavao nul-vektor, tada bi postojala netrivijalna linearna kombinacija vektora koja bi iščezavala. Npr. kad bi bilo $u_j = 0$, tada bi $0u_1 + \cdot + 0u_{j-1} + 1u_j + 0u_{j+1} + \cdot + 0u_r$ bila jedna takova linearna kombinacija, pa bi $\{u_1, \ldots, u_r\}$ bio linearno zavisan skup vektora.

Kad bi linearna kombinacija nekih vektora iz skupa $\{u_1, \ldots, u_r\}$ iščezavala na netrivijalan način, lako bi ju proširili (pomoću skalara nula) do linearne kombinacije svih vektora koja bi također iščezavala na netrivijalan način. Dakle pretpostavka, da je neki podskup od $\{u_1, \ldots, u_r\}$ linearno zavisan, vodi u kontradikciju. Stoga je svaki podskup od $\{u_1, \ldots, u_r\}$ linearno nezavisan.

Ako neka linearna kombinacija vektora $u_1, \ldots, u_r$ iščezava na netrivijalan način, lako ju proširimo sa dodatnim vektorima koje množe nula skalari. Tako dobijemo linearnu kombinaciju vektora iz nadskupa od $\{u_1, \ldots, u_r\}$ koja iščezava na netrivijalan način. To znači da je taj nadskup linearno zavisan. ■

Iz definicija 5.7 i 5.2, propozicije 5.6 i relacije (5.8), možemo zaključiti da se minimalni razapinjući skup uvijek sastoji od linearno nezavisnih vektora. Uočimo da u koracima 1. i 2. algoritma 5.4 postoji sloboda u izboru anihilirajuće linearne kombinacije i u izboru netrivijalnog skalara α_j iz odabrane linearne kombinacije. To znači da minimalni razapinjući skup nije određen na jedinstveni način polaznim skupom vektora $\{x_1, \ldots, x_p\}$. Pogotovo je to istina ako pođemo ne od skupa, već od linearnog potprostora $\mathcal{X}$. Naime, tada možemo naći razne skupove vektora koji razapinju $\mathcal{X}$. Razmislite, da li ih ima beskonačno mnogo? Primjenom algoritma 5.4 na te skupove, dolazimo do još većeg broja minimalnih razapinjućih skupova od $\mathcal{X}$. Osnovno i vrlo važno svojstvo minimalnog razapinjućeg skupa je da *broj vektora u njemu ovisi samo o potprostoru kojeg razapinju.*

Propozicija 5.9 *Ako su $\{u_1, u_2, \ldots, u_r\}$ i $\{v_1, v_2, \ldots, v_s\}$ dva minimalna razapinjuća skupa istog potprostora, onda vrijedi $r = s$. Dakle broj vektora minimalnog razapinjućeg skupa nekog potprostora je invarijanta (konstanta) tog potprostora.*

Dokaz: Označimo, $\mathcal{X} = L(u_1, \ldots, u_r) = L(v_1, \ldots, v_s)$. Pretpostavimo da vrijedi npr. $r \leq s$.

Promotrimo niz vektora $v_1, u_1, \ldots, u_r$. Budući da je $v_1 \in \mathcal{X} = L(u_1, \ldots, u_r)$, vrijedi

$$v_1 = \alpha_1 u_1 + \ldots + \alpha_r u_r$$

za neke skalare $\alpha_1, \ldots, \alpha_r$, pa je zbog propozicije 5.3

$$L(u_1, \ldots, u_r) = L(v_1, u_1, \ldots, u_r).$$

Prema prethodnoj propoziciji, $v_1 \neq 0$, pa je barem jedan od brojeva α_i, $1 \leq i \leq r$ je različit od nule. Dakle, postoji u_j koji je linearna kombinacija ostalih vektora $v_1, u_1, \ldots, u_{j-1}, u_{j+1}, \ldots, u_r$. Prema propoziciji 5.3 možemo izbaciti u_j iz niza $v_1, u_1, u_2, \ldots, u_r$, a da još uvijek vrijedi

$$\mathcal{X} = L(v_1, u_1, \ldots, u_{j-1}, u_{j+1}, \ldots, u_r). \tag{5.11}$$

Sada iza v_1 dodajmo vektor v_2. Dakle promatramo niz vektora v_1, v_2, u_1, $\ldots$, $u_{j-1}, u_{j+1}, \ldots, u_r$. Jer je $v_2 \in \mathcal{X}$, zbog relacije (5.11) zaključujemo,

$$0 \neq v_2 = \beta_1 v_1 + \alpha_1 u_1 + \ldots + \alpha_{j-1} u_{j-1} + \alpha_{j+1} u_{j+1} + \ldots + \alpha_r u_r$$

sa nekim drugim skalarima α_i. Kad bi vrijedilo $\alpha_i = 0$ za sve i, onda bi vektori v_1 i v_2 bili linearno zavisni, a to je nemoguće jer su u minimalnom skupu $\{v_1, \ldots, v_s\}$ svi vektori linearno nezavisni. Dakle, postoji k takav da je $\alpha_k \neq 0$. Koristeći propoziciju 5.3, vektor u_k možemo izbaciti iz razapinjućeg skupa potprostora $\mathcal{X}$. Dakle je $\{v_1, v_2, u_1, \ldots, u_r\} \setminus \{u_j, u_k\}$ razapinjući skup za $\mathcal{X}$.

Označimo $u_j = u_{1'}$, $u_k = u_{2'}$. Koristeći iste argumente, postupak ubacivanja stalno novog vektora v_i i izbacivanja nekog vektora $u_{i'}$ možemo nastaviti sve dok ne izbacimo sve vektore u_i. Tako dolazimo do niza vektora $v_1, v_2, \ldots, v_r$ koji ima svojstvo $\mathcal{X} = L(v_1, v_2, \ldots, v_r)$. Dakle, dobili smo

$$L(v_1, v_2, \ldots, v_r) = \mathcal{X} = L(v_1, v_2, \ldots, v_s).$$

Kako je po pretpostavci $\{v_1, v_2, \ldots, v_s\}$ minimalni razapinjući skup za $\mathcal{X}$ ne može biti $r < s$, već je $r = s$.

Ako pođemo od pretpostavke da je na početku razmatranja bilo $s \leq r$, u dokazu treba samo zamijeniti uloge vektora v_i i u_i kao i brojeva r i s. Opet se dobije $r = s$. ■

Definicija 5.10 *Ako za vektorski prostor X postoji konačni minimalni razapinjući skup vektora $\{u_1, u_2, \ldots, u_r\}$ on se naziva* **konačno-dimenzionalnim vektorskim prostorom**. *Pritom se broj r naziva* **dimenzijom** *prostora X i označava sa $r = \dim X$. Dimenzija trivijalnog prostora $\{0\}$ je nula. Ako za vektorski prostor ne postoji konačni minimalni razapinjući skup vektora on se naziva* **beskonačno-dimenzionalni vektorski prostor**. ■

Beskonačno-dimenzionalnim vektorskim prostorima bavi se posebno područje matematike, koje se zove fukcionalna analiza. Mi ćemo se baviti samo konačno-dimenzionalnim prostorima. Stoga u daljem tekstu svako spominjanje vektorskog prostora X automatski znači da je on konačno-dimenzionalan.

Primjer 5.11 *Vektorski prostor V_n je konačno dimenzionalan, dimenzije n. To se vidi iz relacije*

$$V_n = L(e_1, e_2, \ldots, e_n),$$

gdje su $e_1, e_2, \ldots, e_n$ *dani sa (2.10). Na isti način se pokaže da vrijedi*

$$\mathbf{R}^n = L(e_1, e_2, \ldots, e_n), \qquad \textit{odnosno} \qquad \mathbf{R}^{1\times n} = L(e_1, e_2, \ldots, e_n)$$

gdje se e_i *definiraju na slični način, kao jedinični vektori-stupci odnosno vektori-retci, respektivno. Dakle,* $\mathbf{R}^n$ *i* $\mathbf{R}^{1\times n}$ *su također* n*-dimenzionalni vektorski prostori.* ■

Neka je X vektorski prostor dimenzije n i neka je $\mathcal{S} = \{x_1, \ldots, x_m\}$ podskup u X. Pretpostavimo prvo da su svi vektori x_i linearno nezavisni. Kako je svaki podskup skupa linearno nezavisnih vektora $\mathcal{S} = \{x_1, \ldots, x_m\}$ linearno nezavisan, kažemo da je r maksimalni broj linearno nezavisnih vektora u $\mathcal{S}$. Neka $\mathcal{S}$ nije linearno nezavisan skup vektora. Tada postoji minimalni razapinjući podskup $\mathcal{M}$ od $\mathcal{S}$ od r vektora iz $\mathcal{S}$ i pritom je $r < m$. Zbog propozicije 5.9 $r = \dim(\mathcal{S})$ je najveći broj linearno nezavisnih vektora u $\mathcal{S}$. Kraće kažemo da je r broj linearno nezavisnih vektora u $\mathcal{S}$.

Definicija 5.12 *Neka je* X *vektorski prostor. Uređeni skup vektora* $\mathcal{B}$ *iz* X *zove se* **baza** *vektorskog prostora* X *ako zadovoljava sljedeća dva uvjeta:*

1. $\mathcal{B}$ *je linearno nezavisan skup*

2. $L(\mathcal{B}) = X$. ■

Ako su $u_1, \ldots, u_n$ elementi baze $\mathcal{B}$, bazu ćemo poistovjetiti s uređenom n-torkom vektora $u_1, \ldots, u_n$ koja se koji put piše obrubljena zagradama: $(u_1, \ldots, u_r)$. Znači kod baze, za razliku od minimalnog razapinjućeg skupa, važan je uređaj vektora. Bazu čini niz, a ne skup, linearno nezavisnih vektora.

Iz prvog svojstva minimalnog razapinjućeg skupa $\mathcal{M}$ za X (tj. $L(\mathcal{M}) = X$) i svojstva linearne nezavisnosti od $\mathcal{M}$ (izvedenog primjenom algoritma 5.4), zaključujemo da je svaki uređeni minimalni razapinjući skup i baza od X. Vrijedi i obrat.

Propozicija 5.13 *Svaka baza vektorskog prostora* X *je neki uređeni minimalni razapinjući skup za* X.

Dokaz: Neka je $\mathcal{B} = (b_1, \ldots, b_n)$ baza od X. Tada je $L(b_1, \ldots, b_n) = X$. Primijenimo algoritam 5.4 na skup $\{b_1, \ldots, b_n\}$. Već u prvom prolazu, nakon 1. koraka, algoritam završava, ne mijenjajući skup $\{b_1, \ldots, b_n\}$. Dakle je po propoziciji 5.6, $\mathcal{B}$ (uređeni) minimalni razapinjući skup za X. ■

Na osnovu propozicije 5.9, definicije 5.10 i propozicije 5.13, zaključujemo da svaka baza ima točno $n = \dim X$ vektora.

Primjer 5.14 *Promotrimo u vektorskom prostoru* $\mathbf{R}^{2\times 2}$ *skup*

$$\mathcal{E} = \{E_{11}, E_{12}, E_{21}, E_{22}\},$$

gdje su

$$\begin{aligned} E_{11} &= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad E_{12} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \\ E_{21} &= \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad E_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}. \end{aligned}$$

Svaka matrica $C \in \mathbf{R}^{2\times 2}$ *se može napisati kao linearna kombinacija matrica* E_{ij}*,*

$$C = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} = c_{11}E_{11} + c_{12}E_{12} + c_{21}E_{21} + c_{22}E_{22},$$

pa je $L(E_{11}, E_{12}, E_{21}, E_{22}) = \mathbf{R}^{2\times 2}$*. Skup* $\mathcal{E}$ *je i linearno nezavisan jer jednadžba*

$$\alpha_1 E_{11} + \alpha_2 E_{12} + \alpha_3 E_{21} + \alpha_4 E_{22} = \begin{bmatrix} \alpha_1 & \alpha_2 \\ \alpha_3 & \alpha_4 \end{bmatrix} = O$$

povlači $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0$*. Dakle, svaki uređaj skupa* $\mathcal{E}$ *(npr.* $(E_{11}, E_{12}, E_{21}, E_{22})$*) predstavlja jednu bazu prostora* $\mathbf{R}^{2\times 2}$*. Vidimo da je dimenzija prostora* $\mathbf{R}^{2\times 2}$ *jednaka* 4*.*

Ovo razmatranje se lako generalizira na vektorski prostor $\mathbf{R}^{m\times n}$*. Kolika je dimenzija prostora* $\mathbf{R}^{m\times n}$ *?* ■

Sljedeći rezultat je posebno važan kad se promatra odnos vektorskog prostora i nekog njegovog potprostora.

Lema 5.15 *Ako su* $u_1, u_2, \dots, u_r$ *linearno nezavisni vektori u* n*-dimenzionalnom vektorskom prostoru* X*, tada je* $r \le n$*. Svaki niz linearno nezavisnih vektora* $u_1, u_2, \dots, u_r$ *u* X*, može se nadopuniti s* $n-r$ *vektora do baze od* X*.*

Dokaz: Jer je X konačno-dimenzionalni vektorski prostor, u X postoji minimalni razapinjući skup, a to znači da postoji i baza. Jer je X n-dimenzionalan, baza se sastoji od točno n baznih vektora. Označimo ih redom sa e_1, e_2, … ,e_n.

Jer je skup $\{u_1, u_2, \dots, u_r\}$ linearno nezavisan, prema propoziciji 5.8 među vektorima $u_1, u_2, \dots, u_r$ nema nul-vektora.

Pođimo od niza $u_1, e_1, \dots, e_n$. Jer je $u_1 \in X$, postoje skalari α_1,…, α_n, takvi da je $u_1 = \alpha_1 e_1 + \cdots + \alpha_n e_n$. Jer je $u_1 \neq 0$, zaključujemo da postoji j, $1 \le j \le n$, takav da je $\alpha_j \neq 0$. Stoga je e_j linearna kombinacija vektora u_1, e_1, …,e_{j-1}, e_{j+1}, …,e_n. Prema propoziciji 5.3 vrijedi

$$X = L(e_1, \dots, e_n) = L(u_1, e_1, \dots, e_n) = L(u_1, e_1^{(1)}, \dots, e_{n-1}^{(1)})$$

gdje smo s $e_1^{(1)}$, …,$e_{n-1}^{(1)}$ označili vektore e_1, …,e_{j-1}, e_{j+1}, …,e_n, respektivno. Kako je broj vektora u minimalnom razapinjućem skupu od X jednak n, zaključujemo da je niz $u_1, e_1^{(1)}, \dots, e_{n-1}^{(1)}$ baza od X.

Promotrimo sada niz vektora $u_1, u_2, e_1^{(1)}, \dots, e_{n-1}^{(1)}$. Iz činjenice da je $u_2 \in X = L(u_1, e_1^{(1)}, \dots, e_{n-1}^{(1)})$, zaključujemo da postoji iščezavajuća netrivijalna kombinacija

$$\beta_1 u_1 + \beta_2 u_2 + \alpha_1 e_1^{(1)} + \cdots + \alpha_{n-1} e_{n-1}^{(1)} = 0.$$

Ovdje su α_i neki drugi skalari nego prije. Uočimo da zahtjev $\alpha_1 = \cdots = \alpha_{n-1} = 0$, povlači $\beta_1 u_1 + \beta_2 u_2 = 0$, a to, zbog linearne nezavisnosti od u_1 i u_2, povlači $\beta_1 = \beta_2 = 0$. Dakle, pretpostavka $\alpha_1 = \cdots = \alpha_{n-1} = 0$ povlači trivijalnu kombinaciju, a to nije istina. Zaključujemo da postoji neki $\alpha_k \neq 0$, pa prema propoziciji 5.3 možemo izbaciti $e_k^{(1)}$, dobivajući novu bazu $u_1, u_2, e_1^{(2)}, \dots, e_{n-2}^{(2)}$ za X.

Nastavljajući ovaj postupak, nakon r ovakovih koraka izbacivanja, dolazimo do baze $u_1, u_2, \dots, u_r, e_1^{(r)}, \dots e_{n-r}^{(r)}$ za X. ■

Korolar 5.16 *Neka je* $\mathcal{X}$ *potprostor* n*-dimenzionalnog vektorskog prostora* X*. Ako je i* $\mathcal{X}$ *dimenzije* n *onda je* $\mathcal{X} = X$*.*

Dokaz: Neka je $e_1,\dots,e_n$ proizvoljna baza od $\mathcal{X}$. Nadopunimo ju vektorima $f_1,\dots,f_k$ do baze za X. Kad bi bilo $k \geq 1$, tada bi n-dimenzionalan vektorski prostor X imao bazu od barem $n+1$ vektora, što se protivi definiciji n dimenzionalnog vektorskog prostora. Zaključujemo da je $e_1,\dots,e_n$ baza i za X, pa je $\mathcal{X} = X$. ■

Propozicija 5.17 *U n-dimenzionalnom vektorskom prostoru X, sljedeće tvrdnje su ekvivalentne:*
1. $(u_1, \dots, u_n)$ je baza od X
2. $\{u_1, \dots, u_n\}$ je linearno nezavisan skup
3. $L(u_1, \dots, u_n) = X$.

Dokaz: Pokazat ćemo da vrijedi lanac implikacija: 1. $\Longrightarrow$ 2. $\Longrightarrow$ 3. $\Longrightarrow$ 1. Prva implikacija, 1. $\Longrightarrow$ 2. slijedi odmah iz definicije baze.

Za dokaz implikacije 2. $\Longrightarrow$ 3., označimo $\mathcal{X} = L(u_1, \dots, u_n)$. Jer je $\{u_1, \dots, u_n\}$ linearno nezavisan skup, vektorski potprostor $\mathcal{X}$ od X je n-dimenzionalan. Kako je i sam X n-dimenzionalan to je po korolaru 5.16 $\mathcal{X} = X$ (jedini potprostor od X iste dimenzije kao i X je on sam). Time je pokazano $L(u_1, \dots, u_n) = X$, tj. 3.

Da bi dokazali implikaciju 3. $\Longrightarrow$ 1., pretpostavimo da vrijedi tvrdnja 3. i primijenimo na vektore $u_1, \dots, u_n$ algoritam 5.4. Kad bi vektori $u_1, \dots, u_n$ bili linearno zavisni, algoritam bi izbacio barem jedan vektor, pa bi dobili minimalni razapinjući skup koji bi imao manje od n vektora. Kako je to po definiciji 5.10 nemoguće jer je X n-dimenzionalan, algoritam mora u prvom prolazu izbaciti minimalni razapinjući skup. To znači da vektori $u_1, \dots, u_n$ moraju biti linearno nezavisni. Zajedno s tvrdnjom 3. to daje tvrdnju 1. ■

Propozicija 5.18 *Za svaki niz vektora $a_1, a_2, \dots, a_p$, vrijedi*

$$\dim L(a_1, a_2, \dots, a_p) \leq p.$$

Dokaz: Dokaz slijedi iz činjenice, da minimalni razapinjući podskup skupa $\{a_1, a_2, \dots, a_p\}$ nema više elemenata od samog skupa. To garantira algoritam 5.4. ■

Važnost baze dolazi od jedinstvenosti prikaza vektora po baznim vektorima.

Propozicija 5.19 *Neka je X vektorski prostor i $\mathcal{B}$ baza u X. Tada se svaki vektor $x \in X$ na jednistven način može prikazati kao linearna kombinacija vektora baze $\mathcal{B}$.*

Dokaz: Neka je $\mathcal{B} = (u_1, \ldots, u_n)$. Jer je po svojstvu baze $L(\mathcal{B}) = X$, svaki vektor x dopušta prikaz

$$x = \alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_n u_n.$$

Pretpostavimo da postoji još jedan prikaz

$$x = \beta_1 u_1 + \beta_2 u_2 + \cdots + \beta_n u_n.$$

Izjednačavajući $\alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_n u_n = \beta_1 u_1 + \beta_2 u_2 + \cdots + \beta_n u_n$, dobivamo

$$(\alpha_1 - \beta_1)u_1 + (\alpha_2 - \beta_2)u_2 + \cdots + (\alpha_n - \beta_n)u_n = 0.$$

Po drugom svojstvu baze vektori u_i su linearno nezavisni. Stoga zaključujemo da vrijedi

$$\alpha_1 = \beta_1,\ \alpha_2 = \beta_2,\ \ldots, \alpha_n = \beta_n,$$

tj. prikaz svakog vektora x po baznim vektorima je jedinstven. ■

5.3 Rang matrice

Svaku matricu

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \ldots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \ldots & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \ldots & a_{mn} \end{bmatrix}$$

iz $\mathbf{R}^{m \times n}$ možemo promatrati kao niz njenih vektora-redaka

$$A = \begin{bmatrix} a_1^T \\ a_2^T \\ \ldots \\ a_m^T \end{bmatrix}, \quad a_i \in \mathbf{R}^n,$$

ili kao niz njenih vektora stupaca

$$A = [b_1, \ldots, b_n], \quad b_i \in \mathbf{R}^m\,.$$

Skup vektora-redaka (vektora-stupaca) matrice A razapinje vektorski potprostor od $\mathbf{R}^n$ ($\mathbf{R}^m$), koji zovemo *retčani* (*stupčani*) potprostor od A. Važno svojstvo tih potprostora je da imaju istu dimenziju.

Propozicija 5.20 *Retčani potprostor ima istu dimenziju kao i stupčani potprostor.*

Dokaz: Neka su $a_i^T \in \mathbf{R}^{1\times n}$, $1 \le i \le m$, retci i $b_j \in \mathbf{R}^m$, $1 \le j \le n$, stupci matrice A. Neka A ima najviše r linearno nezavisnih redaka i s linearno nezavisnih stupaca. To znači da između m redaka (n stupaca) matrice, postoji r (s) njih koji su linearno nezavisni, dok je svaki skup od $r+1$ redaka ($s+1$ stupaca) linearno zavisan.

Ako pokažemo da je $s \le r$ za svaku matricu, tada je propozicija zapravo dokazana. Naime, ako za svaku matricu, broj linearno nezavisnih stupaca nije veći od broja linearno nezavisnih redaka, onda to mora vrijediti kako za A, tako i za A^T. Međutim stupci (retci) od A^T su retci (stupci) od A, pa mora vrijediti

$$\begin{aligned} r &= \text{broj linearno nezavisnih redaka od } A \\ &= \text{broj linearno nezavisnih stupaca od } A^T \\ &\le \text{broj linearno nezavisnih redaka od } A^T \\ &= \text{broj linearno nezavisnih stupaca od } A \\ &= s\,, \end{aligned}$$

što zajedno sa $s \le r$ daje $r = s$.

Dokaz da je uvijek $s \le r$, provodimo razmatrajući dva slučaja.

1. Pretpostavimo prvo da su upravo a_1^T, ...,a_r^T linearno nezavisni. Tada oni prema propoziciji 5.17 čine bazu za $L(a_1^T, \ldots, a_r^T)$ pa su a_{r+1}^T, ...,a_m^T iz $L(a_1^T, \ldots, a_r^T)$. Stoga možemo pisati

$$\begin{aligned} a_1^T &= a_1^T \\ &\vdots \qquad\qquad\qquad \ddots \end{aligned}$$

$$\begin{array}{rcl} a_r^T & = & \qquad\qquad\qquad\qquad a_r^T \\ a_{r+1}^T & = & \beta_{r+1,1}a_1^T + \cdots + \beta_{r+1,r}a_r^T \\ & \vdots & \\ a_m^T & = & \beta_{m,1}a_1^T + \cdots + \beta_{m,r}a_r^T \,. \end{array}$$

Izdvojimo iz svake od ovih vektorskih jednadžbi, skalarnu jednadžbu koja se odnosi na j-tu komponentu lijeve i desne strane jednadžbe. To odgovara množenju svake jednadžbe sa jediničnim stupcem $e_j = [0, \ldots, 0, 1, 0, \ldots, 0]^T$ s desna. Za svako $1 \leq j \leq n$, dobijemo

$$\begin{array}{rcl} a_{1j} & = & a_{1j} \\ & \vdots & \qquad\qquad\quad \ddots \\ a_{rj} & = & \qquad\qquad\qquad\qquad a_{rj} \\ a_{r+1,j} & = & \beta_{r+1,1}a_{1j} + \cdots + \beta_{r+1,r}a_{rj} \\ & \vdots & \\ a_{mj} & = & \beta_{m,1}a_{1j} + \cdots + \beta_{m,r}a_{rj} \,. \end{array}$$

Ovaj niz skalarnih jednadžbi možemo zapisati kao jednu vektorsku jednadžbu

$$\begin{array}{rcl} b_j & = & a_{1j} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ \beta_{r+1,1} \\ \vdots \\ \beta_{m,1} \end{bmatrix} + \cdots + a_{rj} \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \beta_{r+1,r} \\ \vdots \\ \beta_{m,r} \end{bmatrix} \\ & = & a_{1j}w_1 + \cdots + a_{rj}w_r, \quad 1 \leq j \leq n. \end{array}$$

Dakle, $b_j \in L(w_1, \ldots, w_r)$ za sve $1 \leq j \leq n$, pa između $b_1, \ldots, b_n$ ne može biti više od r linearno nezavisnih. Dakle je $s \leq r$.

2. Pretpostavimo da su retci s indeksima $i_1, \ldots, i_r$ linearno nezavisni. Nadopunimo niz $i_1, i_2, \ldots, i_r$ preostalim brojevima iz skupa $\{1, \ldots, n\}$, tako da je $i_1, i_2, \ldots, i_r, i_{r+1}, \ldots, i_n$ neka permutacija niza $1, \ldots n$.

Načinimo matricu P čiji su stupci upravo $e_{i_1}, \ldots, e_{i_n}$, $P = [e_{i_1}, \ldots, e_{i_n}]$. Matrica P je jedna matrica permutacije koja ima svojstvo $P^T P = I_n = PP^T$. To slijedi zbog ortogonalnosti i normiranosti stupaca $e_1, \ldots, e_n$.

Retci matrice $P^T A$ su ispermutirani retci od A, pri čemu su prvih r redaka od $P^T A$ upravo $a_{i_1}^T, \ldots, a_{i_r}^T$. To slijedi iz činjenice da se k-ti redak (stupac) matrice dobije množenjem matrice s e_k^T s lijeva (s e_k s desna), pa je

$$e_k^T P^T A = (Pe_k)^T A = e_{i_k}^T A = a_{i_k}^T, \quad 1 \leq k \leq n\,.$$

Dakle su $e_k^T(P^T A)$, $1 \leq k \leq r$ linearno nezavisni.

Označimo broj linearno nezavisnih redaka i stupaca matrice $P^T A$ s $\tilde{r}$ i $\tilde{s}$, respektivno. Jer su u $P^T A$ ispermutirani retci od A, jasno je da vrijedi $\tilde{r} = r$. Primijenimo li razmatranje iz slučaja 1. na $P^T A$, a to možemo jer je prvih r redaka od $P^T A$ linearno nezavisno, dobivamo $\tilde{r} \leq \tilde{s}$. Ako pokažemo da vrijedi $s = \tilde{s}$, tada ćemo imati

$$r = \tilde{r} \leq \tilde{s} = s,$$

pa će propozicija biti dokazana.

Neka vrijedi tvrdnja

$$\alpha_{j_1} b_{j_1} + \cdots + \alpha_{j_k} b_{j_k} = 0 \ \text{ pri čemu je bar jedan } \alpha_{j_t} \neq 0. \tag{5.12}$$

Pomnožimo li jednadžbu u (5.12) s lijeva s P^T, dobivamo

$$\alpha_{j_1} P^T b_{j_1} + \cdots + \alpha_{j_k} P^T b_{j_k} = 0 \ \text{ pri čemu je bar jedan } \alpha_{j_t} \neq 0. \tag{5.13}$$

Obrnuto, ako pođemo od jednadžbe u (5.13) i pomnožimo ju s lijeva s P, zbog $PP^T = I_n$, dobivamo jednadžbu u (5.12). Dakle su tvrdnje (5.12) i (5.13) ekvivalentne jer jedna povlači drugu. Kako se tvrdnja (5.13) može zapisati kao

$$\alpha_{j_1}(P^T A)e_{j_1} + \cdots + \alpha_{j_k}(P^T A)e_{j_k} = 0 \ \text{ pri čemu je bar jedan } \alpha_{j_t} \neq 0\,,$$

pokazali smo da su stupci $b_{j_1}, \ldots, b_{j_k}$ matrice A linearno zavisni *onda i samo onda ako* su takvi i odgovarajući stupci matrice $P^T A$. Zadnja ekvivalencija se može (formalno logički koristeći tzv. obrat po kontrapoziciji) ovako iskazati: stupci $b_{j_1}, \ldots, b_{j_k}$ matrice A su linearno nezavisni onda i samo onda ako su takvi i odgovarajući stupci matrice $P^T A$. To povlači $\dim(b_1, \dots b_n) \subseteq \dim(Pb_1, \dots Pb_n)$ i također $\dim(Pb_1, \dots Pb_n) \subseteq \dim(b_1, \dots b_n)$, tj.

$$\dim(Pb_1, \dots Pb_n) = \dim(b_1, \dots b_n).$$

Dakle je $s = \tilde{s}$. ■

Broj linearno nezavisnih vektora redaka matrice A naziva se retčani rang matrice A. Broj linearno nezavisnih vektora stupaca matrice A naziva se stupčani rang matrice A. Kako se radi o istom broju izbacuje se pridjev retčani ili stupčani.

Definicija 5.21 *Broj linearno nezavisnih vektora redaka (stupaca) matrice A naziva se* **rang** *matrice A i označava s $r(A)$.* ■

Uz svaku matricu $A \in \mathbf{R}^{m\times n}$ vezana su dva važna potprostora *slika* i *jezgra* matrice. Evo kako su definirani.

Skup
$$\mathcal{R}(A) = \{Ax; x \in \mathbf{R}^n\}$$
se zove slika matrice $A \in \mathbf{R}^{m\times n}$ u $\mathbf{R}^m$.

Zadatak 5.22 *Pokažite da je skup $\mathcal{R}(A)$ potprostor od $\mathbf{R}^m$.*
Rješenje: *Dovoljno je uzeti proizvoljne elemente y_1, y_2 iz $\mathcal{R}(A)$, proizvoljne skalare α_1, α_2 iz $\mathbf{R}$ i pokazati da je $y_3 = \alpha_1 y_1 + \alpha_2 y_2$ također u $\mathcal{R}(A)$. No, za y_1 i y_2 postoje x_1 i x_2 iz $\mathbf{R}^n$ takvi da je $y_1 = Ax_1$, $y_2 = Ax_2$. Kako za vektor $x_3 = \alpha_1 x_1 + \alpha_2 x_2 \in \mathbf{R}^n$ vrijedi*

$$Ax_3 = A(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 Ax_1 + \alpha_2 Ax_2 = \alpha_1 y_1 + \alpha_2 y_2 = y_3\,,$$

mora y_3 biti u $\mathcal{R}(A)$. ■

Lema 5.23 *Za svako $A = [b_1, \dots, b_n] \in \mathbf{R}^{m\times n}$ vrijedi*

$$\mathcal{R}(A) = L(b_1, \dots, b_n)\,,$$

pa je $\mathcal{R}(A)$ stupčani potprostor od A. Stoga je dimenzija od $\mathcal{R}(A)$ jednaka $r(A)$.

Dokaz: Iz definicije ranga odmah slijedi da je $r(A)$ dimenzija potprostora $L(b_1,\ldots,b_n)$, pa druga tvrdnja slijedi iz prve tvrdnje.

Da bi dokazali prvu tvrdnju, dovoljno je pokazati da vrijedi

$$L(b_1,\ldots,b_n) \subseteq \mathcal{R}(A) \quad \text{i} \quad \mathcal{R}(A) \subseteq L(b_1,\ldots,b_n).$$

Prva inkluzija je jasna jer se svaki stupac b_i može zapisati kao $Ae_i \in \mathcal{R}(A)$, gdje je e_i i-ti stupac jedinične matrice I_n. Stoga je svaki b_i u $\mathcal{R}(A)$, a jer je $\mathcal{R}(A)$ vektorski prostor, mora biti $L(b_1,\ldots,b_n) \subseteq \mathcal{R}(A)$.

Da bi dokazali drugu inkluziju pokažimo da se svaki vektor Ax, $x \in \mathbf{R}^n$, može zapisati kao linearna kombinacija vektora b_i. Neka je $x \in \mathbf{R}^n$ prizvoljan. Razvijmo ga po bazi jediničnih vektora $e_i \in \mathbf{R}^n$. Dakle, postoje skalari x_1, $\ldots$,x_n, takvi da je $x = x_1e_1 + \cdots + x_ne_n$. Sada je

$$\begin{aligned} Ax &= A(x_1e_1 + \cdots + x_ne_n) = A(x_1e_1) + \cdots + A(x_ne_n) \\ &= x_1Ae_1 + \cdots + x_nAe_n = x_1b_1 + \cdots + x_nb_n \in L(b_1,\ldots,b_n), \end{aligned}$$

pa je tvrdnja dokazana. ■

Još prije smo pokazali da skup rješenja jednadžbe $Ax = 0$ čini vektorski potprostor koji se zove *jezgra* ili *nul-potprostor* od A. Označimo taj potprostor s

$$\mathcal{N}(A) = \{x \in \mathbf{R}^n\,;\ Ax = 0\}.$$

Postavlja se pitanje kolika je dimenzija jezgre. Odgovor daje

Teorem 5.24 *Neka je $A \in \mathbf{R}^{m\times n}$ ranga $r(A)$. Sva rješenja jednadžbe*

$$Ax = 0.$$

čine potprostor $\mathcal{N}(A) \subseteq \mathbf{R}^n$ dimenzije $n - r(A)$.

Dokaz: Neka je s dimenzija i $u_1, u_2, \ldots, u_s$ baza potprostora $\mathcal{N}(A)$. Treba pokazati da je $r(A) + s = n$.

Prema lemi 5.15 možemo niz u_1, u_2, $\ldots$,u_s nadopuniti vektorima $v_1, v_2, \ldots, v_r$ do baze u_1, u_2, $\ldots$,u_s, v_1, v_2, $\ldots$,v_r u $\mathbf{R}^n$. Jasno je da mora vrijediti $s + r = n$. Razvojem po toj bazi možemo svaki vektor x rastaviti na dva vektora, $x = x_0 + x_1$, gdje je x_0 onaj dio razvoja koji uključuje vektore u_1, u_2, ...,u_s, a x_1 preostali dio razvoja. Kako je svaki razvoj po bazi jedinstven, $x_0 \in \mathcal{N}(A)$ i $x_1 = x - x_0$ su jedinstveno određeni s x.

Za svaki $x \in \mathbf{R}^n$ vrijedi $Ax = A(x_0 + x_1) = Ax_0 + Ax_1 = Ax_1$, pa je svaki Ax neka linearna kombinacija vektora $Av_1, \ldots, Av_r$. Stoga je $\mathcal{R}(A) \subseteq L(Av_1, \ldots, Av_r)$. Kako je obrnuta inkluzija trivijalna, pokazali smo da je $L(Av_1, \ldots, Av_r) = \mathcal{R}(A)$. Kad bismo pokazali da su vektori $Av_1, \ldots, Av_r$ linearno nezavisni, oni bi činili bazu za $\mathcal{R}(A)$, a to bi značilo $r = r(A)$ i teorem bi bio dokazan.

Promotrimo sva rješenja $(\alpha_1, \ldots, \alpha_r)$ jednadžbe

$$\alpha_1 Av_1 + \cdots + \alpha_r Av_r = 0. \tag{5.14}$$

Jednadžba (5.14) se može zapisati kao $A(\alpha_1 v_1 + \cdots + \alpha_r v_r) = 0$. Ako je $(\alpha_1, \ldots, \alpha_r)$ neko rješenje jednadžbe (5.14), tada zadnja jednadžba pokazuje da je vektor $z = \alpha_1 v_1 + \cdots + \alpha_r v_r$ u $\mathcal{N}(A)$. Dakle, postoje skalari β_1, ...,β_s, takvi da je $z = \beta_1 u_1 + \cdots + \beta_s u_s$. Pišući $z = z$ pomoću dva razvoja,

$$\alpha_1 v_1 + \cdots + \alpha_r v_r = \beta_1 u_1 + \cdots + \beta_s u_s,$$

dobivamo iščezavajuću linearnu kombinaciju

$$\beta_1 u_1 + \cdots + \beta_s u_s + (-\alpha_1)v_1 + \cdots + (-\alpha_r)v_r = 0.$$

Kako je u_1, u_2, ...,u_s, v_1, v_2, ...,v_r baza u $\mathbf{R}^n$, a vektori baze su linearno nezavisni, dokazali smo da su svi α_i i β_j nule. Bitno je da smo dobili $\alpha_1 = \cdots = \alpha_r = 0$, pa jednadžba (5.14) dopušta samo trivijalno rješenje, čime je teorem dokazan. ■

Lekcija 6

Homogeni sustavi linearnih jednadžbi

U ovom poglavlju ćemo efektivno rješavati matričnu jednadžbu

$$Ax = 0,$$

gdje su $A \in \mathbf{R}^{m\times n}$, $x \in \mathbf{R}^n$ i 0 je nulvektor iz $\mathbf{R}^m$. Ta jednadžba je drugi zapis za sustav linearnih jednadžbi

$$\begin{array}{lcl} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n & = & 0 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n & = & 0 \\ \cdots & & \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n & = & 0. \end{array}$$

Tri pravila za dobivanje ekvivalentnog sustava, koja smo uveli u prvom poglavlju, svode se na specijalne operacije nad retcima matrice. Zvat ćemo ih **osnovne ili elementarne operacije nad (ili na) retcima** ili kraće **osnovne ili elementarne retčane operacije**. To su

1. zamjena dvaju redaka
2. množenje nekog retka brojem različitim od nule
3. dodavanje retka pomnoženog nekim skalarom drugom retku.

Ako je matrica B dobivena iz matrice A jednom osnovnom operacijom na retcima, tada se i polazna matrica A može dobiti iz matrice B

istovrsnom osnovnom retčanom operacijom. Npr. ako je operacija bila zamjena redaka, treba istu operaciju ponoviti na B. Ako je operacija bila množenje retka brojem različitim od nule, treba isti redak od B pomnožiti recipročnim brojem. Ako je operacija bila dodavanje i-tom retku j-ti redak pomnožen skalarom α, treba istom, i-tom retku matrice B, dodati j-ti redak od B pomnožen skalarom $-\alpha$. Dakle, za svaku elementarnu operaciju nad retcima postoji njoj inverzna operacija koja vraća matricu na polazno stanje, pri čemu je ta inverzna operacija također elementarna operacija nad retcima istog tipa.

6.1 Ekvivalentnost po retcima

Pomoću elementarnih retčanih operacija uvodimo pojam retčane ekvivalentnosti matrica.

Definicija 6.1 *Ako je matrica B dobivena iz matrice A pomoću jedne ili više osnovnih operacija nad retcima, za B se kaže da je ekvivalentna matrici A po retcima, ili da je* **retčano ekvivalentna** *matrici A. U tom slučaju pišemo $B\overset{r}{\sim}A$.* ■

Korištenjem inverznih elementarnih operacija, lako se pokaže da retčana ekvivalentnost ima svojstvo simetrije, tj. ako je $B\overset{r}{\sim}A$, tada je i $A\overset{r}{\sim}B$. Također, svaka matrica je retčano ekvivalentna samoj sebi (dovoljno je pomnožiti bilo koji redak matrice skalarom 1). Konačno, ako je $B\overset{r}{\sim}A$ i $C\overset{r}{\sim}B$, tada je $C\overset{r}{\sim}A$, jer je C dobivena iz A nizom elementarnih retčanih operacija. To pokazuje da retčana ekvivalentnost zadovoljava svojstva refleksivnosti, simetrije i tranzitivnosti, pa je relacija ekvivalencije na skupu matrica.

Teorem 6.2 *Ako je $B\overset{r}{\sim}A$, tada je $r(B) = r(A)$.*

Dokaz: Dovoljno je pokazati da osnovne retčane operacije ne mijenjaju rang matrice. Zato pretpostavimo da B nastaje iz A primjenom jedne elementarne retčane operacije.

- Neka B nastaje iz A zamjenom dvaju redaka. Ako je $a_1^T,\ldots,a_n^T$ poredak redaka u A, onda je poredak redaka od B neka permutacija tog niza, $a_{i_1}^T,\ldots,a_{i_n}^T$. Kako je

$$L(a_1^T, \ldots, a_n^T) = L(a_{i_1}^T, \ldots, a_{i_n}^T)\,,$$

maksimalni broj linearno nezavisnih redaka od A (tj. $\dim L(a_1^T,\ldots,a_n^T)$) mora biti jednak maksimalnom broju linearno nezavisnih redaka od B (tj. $\dim L(a_{i_1}^T,\ldots,a_{i_n}^T)$).

- Neka B nastaje iz A množenjem i-tog retka skalarom $\beta \neq 0$. Sada su retci od B redom a_1^T, …,βa_i^T, …,a_n^T. S obzirom da je

  $$L(a_1^T,\ldots,a_n^T) = L(a_1^T,\ldots,\beta a_i^T,\ldots,a_n^T)$$

  (a to je zato jer svaku linearnu kombinaciju $\beta_1 a_1^T + \cdots + \beta_i a_i^T + \cdots + \beta_n a_n^T$ možemo zapisati kao $\beta_1 a_1^T + \cdots + (\beta_i/\beta)\beta a_i^T + \cdots + \beta_n a_n^T$), zaključak je isti kao prije.

- Neka B nastaje iz A tako da se a_i^T zamijeni linearnom kombinacijom $a_i^T - \beta a_j^T$. Ako retke od B označimo s b_1^T,…,b_n^T, tada je $b_i^T = a_i^T - \beta a_j^T$, dok su ostali retci od B isti kao i odgovarajući retci od A.

  Pokažimo da je

  $$L(a_1^T,\ldots,a_n^T) \subseteq L(b_1^T,\ldots,b_n^T) \subseteq L(a_1^T,\ldots,a_n^T).$$

  Neka je $z \in L(a_1^T,\ldots,a_n^T)$ proizvoljan. Tada postoje skalari α_1,…,α_n, takvi da je $z = \alpha_1 a_1^T + \cdots \alpha_n a_n^T$. Ako iskoristimo

  $$\begin{aligned} \alpha_i a_i^T + \alpha_j a_j^T &= \alpha_i a_i^T - \alpha_i \beta a_j^T + \alpha_i \beta a_j^T + \alpha_j a_j^T \\ &= \alpha_i (a_i^T - \beta a_j^T) + (\alpha_i \beta + \alpha_j) a_j^T \\ &= \beta_i b_i^T + \beta_j b_j^T\,, \text{ gdje su } \beta_i = \alpha_i,\ \beta_j = \alpha_j + \alpha_i \beta\,, \end{aligned}$$

  vidimo da je z linearna kombinacija redaka $b_1^T,\ldots,b_n^T$, pa je $z \in L(b_1^T,\ldots,b_n^T)$. Time je dokazana prva inkluzija.

  Za dokaz druge inkluzije polazimo od proizvoljnog elementa $w \in L(b_1^T,\ldots,b_n^T)$. Za w postoje skalari β_1,…,β_n takvi da je $w = \beta_1 b_1^T + \cdots + \beta_n b_n^T$. Jer je

  $$\begin{aligned} \beta_i b_i^T + \beta_j b_j^T &= \beta_i (a_i^T - \beta a_j^T) + \beta_j a_j^T = \beta_i a_i^T - \beta_i \beta a_j^T + \beta_j a_j^T \\ &= \beta_i a_i^T + (\beta_j - \beta_i \beta) a_j^T \\ &= \alpha_i a_i^T + \alpha_j a_j^T\,, \text{ gdje su } \alpha_i = \beta_i\,,\ \alpha_j = \beta_j - \beta_i \beta\,, \end{aligned}$$

w je linearna kombinacija redaka od A, pa je $w \in L(a_1^T, \ldots, a_n^T)$. Zbog proizvoljnosti vektora w, dokazana je i druga inkluzija. Zaključujemo da je $L(a_1^T, \ldots, a_n^T) = L(b_1^T, \ldots, b_n^T)$, pa A i B imaju jednak broj linearno nezavisnih redaka. ■

Dakle retčana ekvivalentnost čuva rang matrice.

Znamo da svaka elementarna retčana operacija na matrici A prevodi polazni homogeni sustav linearnih jednadžbi u novi koji ima ista rješenja kao i polazni. Zato se elementarne transformacije nad retcima mogu iskoristiti kao alat za svađanje polaznog sustava na ekvivalentni sustav koji ima dovoljno jednostavnu matricu, tako da se rješenje jednostavno "pročita" iz matrice. Da bi se svađanje polazne matrice na jednostavniji oblik što više unificiralo, treba organizirati elementarne retčane operacije u algoritam.

6.2 Algoritam za redukciju matrice

Prikazat ćemo algoritam koji koristi elementarne transformacije nad retcima za svađanje polazne matrice $A \in \mathbf{R}^{m\times n}$ na matricu A_R koja je u specijalnom, tzv. reduciranom obliku, a retčano je ekvivalentna s A.

Algoritam 6.3 *Algoritam vrši elementarne retčane operacije na matrici X. Oznaka $X = (x_{ij})$ se koristi za iteriranu matricu. Sa x_k^T je označen k-ti redak od X. Na ulazu, X se poistovjeti s matricom $A \in \mathbf{R}^{m\times n}$. Na izlazu X sadrži matricu A_R u reduciranom obliku.*

1. *Stavi: $X = A, \;\; i = 1$.*

2. *Odredi j kao najmanji indeks stupca od X koji nema sve komponente x_{kj}, $i \le k \le m$, jednake nuli. Ako takav stupac ne postoji, idi na 7.*

3. *Ako je $x_{ij} \neq 0$, idi na 4. Ako je $x_{ij} = 0$, nađi najmanji indeks l za koji je $x_{lj} \neq 0$ i $l > i$. Zatim istovremeno zamijeni i-ti i l-ti redak matrice X.*

4. *Pomnoži i-ti redak matrice sa $1/x_{ij}$.*

5. *Za svako $k \neq i$, ako je $x_{kj} \neq 0$, zamijeni k-ti redak matrice X s linearnom kombinacijom k-tog i i-tog retka: $x_k^T - x_{kj} \cdot x_i^T$.*

6. *Stavi $i := i + 1$. Ako je $i = m + 1$ idi na 7., inače idi na 2.*

7. *Stavi $A_R = X$ i stani.* ■

Ovaj algoritma je praktičan za nekoga tko želi napisati kompjutorski program,[1] ali nije zgodan za razumijevanje same redukcije. Zato ćemo pokušati cijeli postupak slikovno pojasniti. Kao prvo, ako je A nul-matrica, onda smo odmah u koraku 2. gotovi, a matrica A ostaje nepromijenjena. Inače, polazna matrica A općenito izgleda ovako:

$$A = \begin{bmatrix} 0 & 0 & \dots & 0 & x & \dots & x \\ 0 & 0 & \dots & 0 & x & \dots & x \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 & x & \dots & x \end{bmatrix}.$$

Nakon 4. koraka matrica se transformira u

$$\begin{bmatrix} 0 & 0 & \dots & 0 & 1 & x & \dots & x \\ 0 & 0 & \dots & 0 & x & x & \dots & x \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 & x & x & \dots & x \end{bmatrix}.$$

a nakon 5. koraka u

$$\begin{bmatrix} 0 & 0 & \cdots & 0 & 1 & x & \cdots & x & \\ 0 & 0 & \cdots & 0 & 0 & x & \cdots & x & \\ \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \\ 0 & 0 & \cdots & 0 & 0 & x & & \cdots & x \end{bmatrix}.$$

U drugom prolazu kroz algoritam, promatramo podmatricu dobivenu na presjeku redaka $2,\dots,m$ i stupaca $j+1,\dots n$. Ako je ta podmatrica nul-matrica, onda povratkom na korak 2. algoritam završava (nakon skoka na 7. korak). Ako nije, onda ponavljamo korake 2., 3. i 4. na toj

[1] Ako želite napisati odgovarajući kompjutorski program, u 3. koraku zahtijevajte da je x_{lj} najveći po modulu između komponenti $x_{kj} \neq 0$, $i \leq k \leq m$. Ovo je potrebno jer računala rade s tzv. aritmetikom konačne preciznosti.

podmatrici. Korak 5. provodimo na retcima cijele matrice, ne samo podmatrice. Rezultat će općenito izgledati ovako:

$$\begin{bmatrix} 0 & 0 & \dots & 0 & 1 & x & \dots & 0 & x & \dots & x \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 1 & x & \dots & x \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & x & \dots & x \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & x & \dots & x \end{bmatrix}.$$

Na kraju trećeg prolaza kroz algoritam, dobijemo matricu ekvivalentnu matrici A, koja izgleda ovako:

$$\begin{bmatrix} 0 & \dots & 0 & 1 & x & \dots & x & 0 & x & \dots & x & 0 & x & \dots & x \\ & & & & & & & 1 & x & \dots & x & 0 & x & \dots & x \\ & & & & & & & & & & & 1 & x & \dots & x \\ & & & & & & & & & & & 0 & x & \dots & x \\ & & & & & & & & & & & \vdots & \vdots & \vdots & \vdots \\ & & & & & & & & & & & 0 & x & \dots & x \end{bmatrix}.$$

Postupak se ponavlja tako dugo dok preostali retci matrice ne postanu nul-retci, ili ih više nema.

6.3 Reducirani oblik matrice

Reducirani oblik matrice (koji se još zove retčana ešalonska forma matrice) određen je sljedećim uvjetima:

- Svi retci koji sadrže samo nule (ako takovih ima) nalaze se iza netrivijalnih redaka (ako takovih nema, cijela matrica je nul-matrica).

- Prvi od nule različit element svakog netrivijalnog retka je jedinica (tzv. pivotni ili stožerni element). Svi ostali elementi u stupcu stožernog elementa su nule.

- Ako su (i_1, j_1) i (i_2, j_2) pozicije dvaju stožernih elemenata, tada $i_1 < i_2$ povlači $j_1 < j_2$.

Treće svojstvo znači da, gledajući od prvog retka na niže, svaki sljedeći stožerni element leži desno od prethodnog (odatle i naziv retčana ešalonska forma). Može se pokazati da je svaka matrica retčano ekvivalentna samo jednoj matrici u reduciranom obliku.

Primjer 6.4 *Sljedeće matrice su u reduciranoj formi*

$$A = \begin{bmatrix} 1 & 0 & 4 & 2 & 3 \\ 0 & 1 & 3 & 5 & 4 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 & 0 & 1 & 5 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 & 6 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

$$C = \begin{bmatrix} 1 & 5 & 4 & 0 & 3 & 2 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 2 & 3 & 0 & 4 & 0 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 7 & 0 & 8 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \qquad D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Diskutirajte zašto ove matrice nisu u reduciranoj formi,

$$F = \begin{bmatrix} 1 & 5 & 4 & 2 & 3 \\ 0 & 1 & 3 & 5 & 4 \end{bmatrix}, \qquad G = \begin{bmatrix} 0 & 0 & 1 & 5 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 & 1 & 6 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

$$H = \begin{bmatrix} 1 & 5 & 4 & 0 & 3 & 2 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 2 & 3 & 0 & 4 & 0 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 7 & 0 & 8 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 6 \end{bmatrix}, \qquad K = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

■

Primjer 6.5 *Koristeći algoritam 6.3, svedimo matricu*

$$A = \begin{bmatrix} 0 & 0 & 3 & -1 \\ 0 & -1 & 4 & 7 \\ 0 & -1 & 7 & 6 \end{bmatrix}$$

na reduciranu formu.

U prvom koraku stavljamo $X = A$ *i* $i = 1$. *Kako je prvi stupac nul-vektor, a drugi nije, stavljamo* $j = 2$ *(korak 2). S obzirom da je* $x_{12} = 0$, *, prvi redak zamijenjujemo sa drugim retkom (korak 3),*

$$X = \begin{bmatrix} 0 & -1 & 4 & 7 \\ 0 & 0 & 3 & -1 \\ 0 & -1 & 7 & 6 \end{bmatrix}.$$

U 4. koraku prvi redak množimo sa $1/x_{12} = -1$,

$$X = \begin{bmatrix} 0 & 1 & -4 & -7 \\ 0 & 0 & 3 & -1 \\ 0 & -1 & 7 & 6 \end{bmatrix}.$$

U 5. koraku, od trećeg retka oduzimamo prvi redak pomnožen sa -1, *tj. trećem retku dodajemo prvi redak,*

$$X = \begin{bmatrix} 0 & 1 & -4 & -7 \\ 0 & 0 & 3 & -1 \\ 0 & 0 & 3 & -1 \end{bmatrix}.$$

U 6. koraku stavljamo $i = 2$ *i vraćamo se na korak 2. Promatramo podmatricu na presjeku drugog i trećeg retka i trećeg i četvrtog stupca. Budući da je* $x_{23} \neq 0$, *preskačemo korak 3., a u 4. koraku množimo drugi redak matrice sa* $1/x_{23} = 1/3$,

$$X = \begin{bmatrix} 0 & 1 & -4 & -7 \\ 0 & 0 & 1 & -1/3 \\ 0 & 0 & 3 & -1 \end{bmatrix}.$$

Zatim (u 5. koraku) od trećeg retka oduzimamo drugi redak pomnožen s 3, (tj. trećem retku dodajemo drugi redak pomnožen s -3*), i prvom retku dodajemo drugi redak pomnožen s* 4,

$$X = \begin{bmatrix} 0 & 1 & 0 & -25/3 \\ 0 & 0 & 1 & -1/3 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Sada stavljamo $i = 3$ *(korak 6) i vraćamo se na korak 2. Promatramo treći redak u kojem su same nule. Zato idemo na korak 7. Dobivena* *matrica*

$$X = A_R = \begin{bmatrix} 0 & 1 & 0 & -25/3 \\ 0 & 0 & 1 & -1/3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

je u reduciranoj formi. ■

Primjer 6.6 *Svedimo matricu*

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \\ 10 & 11 & 15 \end{bmatrix}$$

na reduciranu formu. Umjesto navođenja koraka iz algoritma 6.3 koje koristimo, postupak ćemo ubrzati tako da ispisujemo matricu tek onda kada se promijeni. Brojevi koraka koji su prijeđeni stavljeni su iznad strelica koje odvajaju matrice,

$$X \xrightarrow{1.-5.} \begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & -6 & -13 \\ 0 & -9 & -6 \end{bmatrix} \xrightarrow{6.,2.-4.} \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & -6 & -13 \\ 0 & -9 & -6 \end{bmatrix} \xrightarrow{5.} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & -1 \\ 0 & 0 & 12 \end{bmatrix}$$

$$\xrightarrow{6.,2.-4.} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \\ 0 & 0 & 12 \end{bmatrix} \xrightarrow{5.} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{6.,7.} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} = A_R\,.$$

■

Stupci matrice A_R koji sadrže pivotne elemente nazivaju se pivotni ili stožerni stupci. Njih stoga ima onoliko koliko ima pivotnih elemenata, odnosno onoliko koliko ima netrivijalnih (tj. od nule različitih) redaka.

Propozicija 6.7 *Neka je matrica A_R dobivena iz matrice $A \in \mathbf{R}^{m\times n}$ primjenom algoritma 6.3. Neka je r broj netrivijalnih redaka matrice A_R. Tada vrijede tvrdnje*

(i) *r je rang matrice A, $r \leq \min\{m\,,\,n\}$ i svi netrivijalni retci od A_R su linearno nezavisni.*

(ii) *Postoji r stupaca α_{j_1}, α_{j_2}, $\ldots$, α_{j_r} matrice A_R, koji su vektori kanonske baze iz $\mathbf{R}^m$, tj. $\alpha_{j_k} = e_k \in \mathbf{R}^m$ za $1 \leq k \leq r$.*

Dokaz: (i) Algoritam 6.3 koristi samo elementarne retčane transformacije. Zato su polazna matrica A i završna matrica A_R retčano ekvivalentne. Prema teoremu 6.2, $r(A_R) = r(A)$. Kako je A_R u reduciranom obliku, svaki njen netrivijalni redak započinje sa stožernom jedinicom koja je desno od prethodne stožerne jedinice kad matricu gledamo od prvog retka na niže. Stožerni stupci imaju osim stožernog elementa nule kao komponente. Zato svaka linearna kombinacija netrivijalnih redaka $\beta_1\alpha_1^T + \cdots + \beta_r\alpha_r^T$, gledana kao element iz $\mathbf{R}^{1\times n}$, ima kao komponente na pivotnim pozicijama $j_1, \ldots, j_r$ upravo multiplikatore $\beta_1, \ldots, \beta_r$ (vidi sljedeći primjer6.8). To znači da svaka iščezavajuća linearna kombinacija netrivijalnih redaka ima sve multiplikatore nule. Zato su svi netrivijalni retci od A_R linearno nezavisni, pa je njihov broj upravo $r(A_R)$, tj. $r = r(A_R)$. Po propoziciji 5.20 linearno nezavisnih stupaca ima točno onoliko koliko ima linearno nezavisnih redaka. Zato je $r \leq \min\{m\,,\, n\}$.

(ii) Pivotni stupac ima svojstvo da su mu sve komponente nule, osim jednog (pivotnog elemenata) koji je jedinica. Duljina svakog stupca u A_R je m, pa je svaki pivotni stupac neki element kanonske baze $(e_1, \ldots, e_m)$ u $\mathbf{R}^m$. Pivotnih stupaca u A_R ima onoliko koliko ima pivotnih elemenata, a to znači onoliko koliko ima netrivijalnih redaka, dakle r. Kako k-ti pivotni stupac ima jedinicu u k-tom retku, vrijedi $\alpha_{j_k} = e_k \in \mathbf{R}^m$ za $1 \leq k \leq r$. ■

Primjer 6.8 *Neka je*

$$A_R = \begin{bmatrix} 0 & 1 & 2 & 0 & -3 & 7 & 0 \\ 0 & 0 & 0 & 1 & 4 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

reducirani oblok matrice A. Pokažimo da su netrivijalni retci od A_R linearno nezavisni. Treba načiniti proizvoljnu linearnu kombinaciju tih redaka, izjednačiti ju s nulom (nul-retkom) i pokazati da moraju svi koeficijenti u linearnoj kombinaciji biti nula. Imamo

$$\alpha_1 \cdot [0\ 1\ 2\ 0\ \ -3\ 7\ 0] + \alpha_2 \cdot [0\ 0\ 0\ 1\ 4\ \ -2\ 0] + \alpha_3 \cdot [0\ 0\ 0\ 0\ 0\ 0\ 1]$$
$$= [0\ \alpha_1\ 2\alpha_1\ \alpha_2\ \ -3\alpha_1 + 4\alpha_2\ 7\alpha_1 - 2\alpha_2\ \alpha_3] = [0\ 0\ 0\ 0\ 0\ 0\ 0]$$

odakle slijedi gledajući drugu, četvrtu i sedmu komponentu, $\alpha_1 = 0$, $\alpha_2 = 0$ *i* $\alpha_3 = 0$. ■

6.4 Rješavanje homogenih sustava

Kao što znamo, skup rješenja jednadžbe

$$Ax = 0$$

jednak je skupu rješenja jednadžbe

$$A_R x = 0$$

jer smo matricu A_R dobili iz matrice A elementarnim operacijama nad retcima. Matrica A_R je u reduciranom obliku,

$$A_R = \begin{bmatrix} 0 & \ldots & 0 & 1 & x & \ldots & x & 0 & x & \ldots & x & 0 & x & \ldots & x \\ & & & & & & & 1 & x & \ldots & x & 0 & x & \ldots & x \\ & & & & & & & & & & & 1 & x & \ldots & x \\ & & & & & & & & & & & 0 & \ldots & \ldots & \vdots \\ & & & & & & & & & & & \vdots & \vdots & \vdots & \vdots \\ & & & & & & & & & & & 0 & 0 & \ldots & 0 \end{bmatrix},$$

sa r netrivijalnih redaka i isto toliko istaknutih (pivotnih) stupaca, $\alpha_{j_k} = e_k \in \mathbf{R}^m$, $1 \leq k \leq r$. Pritom je r jednak rangu matrice. Prema teoremu 5.24 znamo da je skup rješenja jednadžbe $A_R x = 0$ potprostor $\mathcal{N}(A)$ od $\mathbf{R}^n$ dimenzije $n - r(A_R) = n - r$.

Neka α_l nije pivotni stupac matrice $A_R = (\alpha_{ij})$. Dakle l nije u skupu $\{j_1, j_2, \ldots, j_r\}$. Za svaki takav stupac formirajmo vektor $x^{(l)}$ na sljedeći način. Prvo sve komponente od $x^{(l)}$ definirajmo nulama. Zatim stavimo $x^{(l)}_l = 1$, $x^{(l)}_{j_1} = -\alpha_{1l}$, $x^{(l)}_{j_2} = -\alpha_{2l}$, $\ldots$, $x^{(l)}_{j_r} = -\alpha_{rl}$. Npr. ako je $l > j_r$, tada je

$$x^{(l)} = [\underset{1}{0}, \ldots, 0, \underset{j_1}{-\alpha_{1l}}, 0, \ldots, 0, \underset{j_2}{-\alpha_{2l}}, 0, \ldots, 0, \underset{j_r}{-\alpha_{rl}}, 0, \ldots, 0, \underset{l}{1}, 0, \ldots, \underset{n}{0}]^T.$$

Množenjem matrice A_R vektorom $x^{(l)}$ dobivamo,

$$\begin{aligned}
A_R x^{(l)} &= 0 \cdot \alpha_1 + \cdots 0 \cdot \alpha_{j_1-1} + (-\alpha_{1l}) \cdot \alpha_{j_1} + (-\alpha_{2l}) \cdot \alpha_{j_2} + \cdots \\
&\quad +(-\alpha_{rl}) \cdot \alpha_{j_r} + 1 \cdot \alpha_l \\
&= (-\alpha_{1l}) \cdot e_1 + (-\alpha_{2l}) \cdot e_2 + \cdots + (-\alpha_{rl}) \cdot e_r + a_l \\
&= 0,
\end{aligned}$$

jer je po propoziciji 6.7(ii) $\alpha_{j_k} = e_k$, $1 \leq k \leq r$ i jer α_l ima razvoj po kanonskoj bazi u $\mathbf{R}^m$ $\alpha_l = \alpha_{1l}e_1 + \cdots + \alpha_{rl}e_r$. Naime, $\alpha_l = [\alpha_{1l}, \alpha_{2l}, \ldots, \alpha_{rl}, 0, \ldots, 0]^T$. Vidimo da je svaki vektor $x^{(l)}$ u $\mathcal{N}(A_R)$.

Vektora $x^{(l)}$ ima točno $n-r$ i međusobno su linearno nezavisni, jer u l-toj komponenti gdje je $x^{(l)}$ jednako 1, ostali vektori $x^{(i)}$, $i \neq l$, imaju vrijednost nula. Prema tome, u nul-potprostoru $\mathcal{N}(A_R)$ smo našli $n-r$ $= n - r(A_R)$ linearno nezavisnih vektora pa oni po propoziciji 5.17 čine bazu od $\mathcal{N}(A_R)$. Kako je $\mathcal{N}(A_R) = \mathcal{N}(A)$, ti vektori čine bazu za $\mathcal{N}(A)$. Time smo dokazali

Teorem 6.9 *Neka je $A \in \mathbf{R}^{m \times n}$ i neka je $A_R = (\alpha_{ij})$ dobivena iz matrice A algoritmom 6.3. Neka su $\alpha_{j_1}, \alpha_{j_2}, \ldots, \alpha_{j_r}$ stožerni stupci matrice A_R, za koje vrijedi $\alpha_{j_k} = e_k$, $1 \leq k \leq r$. Skup rješenja homogenog sustava $Ax = 0$ je potprostor od $\mathbf{R}^n$ razapet linearno nezavisnim vektorima $x^{(l)}$, $l \in \{1, \ldots, n\} \setminus \{j_1, \ldots, j_r\}$, za koje vrijedi*

$$x_l^{(l)} = 1, \qquad x_{j_k}^{(l)} = -\alpha_{kl}, \; 1 \leq k \leq r\,,$$

dok su ostale komponente od $x^{(l)}$ nule.

Primjer 6.10 *Matrica*

$$A = \begin{bmatrix} 0 & 0 & 3 & -1 \\ 0 & -1 & 4 & 7 \\ 0 & -1 & 7 & 6 \end{bmatrix}$$

iz primjera 6.5, ima reducirani oblik

$$A_R = \begin{bmatrix} 0 & 1 & 0 & -25/3 \\ 0 & 0 & 1 & -1/3 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

U toj matrici istaknuto mjesto imaju drugi i treći stupac, jer su oni vektori kanonske baze od $\mathbf{R}^3$. *Skup rješenja jednadžbe* $Ax = 0$ *je potprostor razapet vektorima iz* $\mathbf{R}^4$, *koje dobijemo iz preostalih stupaca (prvog i četvrtog) matrice* A_R:

$$x^{(1)} = [1, 0, 0, 0]^T, \quad x^{(4)} = [0, 25/3, 1/3, 1]^T.$$

Opće rješenje te jednadžbe možemo pisati u obliku

$$x = [\beta, \frac{25}{3}\gamma, \frac{1}{3}\gamma, \gamma]^T, \quad \beta\,, \gamma\ \in \mathbf{R}.$$

■

Primjer 6.11 *Za matricu iz primjera 6.6,*

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \\ 10 & 11 & 15 \end{bmatrix} \quad \textit{smo dobili} \quad A_R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Njena sva tri stupca su stožerna. Rang reducirane matrice je 3, *pa skup rješenja jednadžbe* $Ax = 0$ *ima dimenziju nula. Prema tome,* $\mathcal{N}(A)$ *ima samo trivijalno rješenje*

$$x = [0, 0, 0, 0]^T\,,$$

koje rješava homogeni sustav $Ax = 0$. ■

Lekcija 7

Nehomogeni sustav linearnih jednadžbi

U petom i šestom poglavlju proučavali smo riješenja jednadžbe $Ax = 0$, a u ovom poglavlju bavit ćemo se jednadžbom

$$Ax = b\,,$$

gdje je b bilo kakav vektor iz $\mathbf{R}^m$, dok su $A \in \mathbf{R}^{m\times n}$ i $x \in \mathbf{R}^n$. Ta matrična jednadžba je zapis nehomogenog sustava linearnih jednadžbi

$$\begin{array}{lcl} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n & = & b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n & = & b_2 \\ \cdots & & \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n & = & b_m \end{array}\,. \tag{7.1}$$

7.1 Rješavanje nehomogenog sustava

Tri osnovna pravila za rješavanje sustava jednadžbi koja smo uveli u prvom poglavlju vrijede kako za homogene tako i za nehomogene sustave jednadžbi, pa dakle vrijede i za rješavanje sustava (7.1). Matrična analogija tih pravila kod nehomogenih sustava linearnih jednadžbi su osnovne ili elementarne operacije nad retcima (koje smo uveli u šestom poglavlju), ali sada na proširenoj matrici

$$[A \mid b] = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array} \right] .$$

Primijenimo algoritam 6.3 na tu proširenu matricu. U dobivenoj reduciranoj matrici $[A \mid b]_R$, prvih n stupaca će izgledati isto kao i stupci matrice A_R koja nastaje iz A primjenom istog algoritma. Dakle, možemo pisati

$$[A \mid b]_R = [A_R \mid \tilde{b}] .$$

pri čemu je $\tilde{b}$ nastao iz b odgovarajućim elementarnim transformacijama komponenata od b. Neka je r broj netrivijalnih redaka matrice A_R. Razlikujemo dva slučaja.

1. Stupac $\tilde{b}$ je stožerni stupac matrice $[A \mid b]_R$. Jer matrica A_R ima r redaka različitih od nule, mora vektor $\tilde{b}$ imati $r+1$. komponentu jednaku 1. Stoga $r+1$. jednadžba sustava $A_R x = \tilde{b}$ glasi

$$0 = 1 .$$

   Dobivena konradikcija pokazuje da sustav $A_R x = \tilde{b}$ nema rješenje, jer niti za jedan $x \in \mathbf{R}^n$ neće biti zadovoljena $r+1$. jednadžba sustava. Kako je $A_R x = \tilde{b}$ ekvivalentan polaznom sustavu, niti polazni sustav $Ax = b$ nema rješenja.

   Uočimo da u ovom slučaju vrijedi $r(A) = r(A_R) = r$ i $r([A \mid b]) = r([A_R \mid \tilde{b}]) = r + 1$ pa je

$$r(A) < r([A \mid b]).$$

2. Stupac $\tilde{b}$ nije stožerni stupac matrice $[A \mid b]_R$. U tom slučaju su retci s indeksima $r+1$, $r+2, \dots, m$, matrice $[A \mid b]_R$ nul-retci. To znači da je prvih r redaka, ne samo matrice A_R već i matrice $[A_R \mid \tilde{b}]$, linearno nezavisno. Naime kad se "produže" linearno nezavisni retci (kod nas za jednu komponentu) onda produženi retci i dalje ostaju linearno nezavisni. Prema tome je

$$r(A) = r = r(A_R) = r([A_R \mid \tilde{b}]) = r([A \mid b]).$$

Neka su $j_1, j_2, \ldots, j_r$ indeksi stožernih stupaca matrice A_R. Jedno istaknuto rješenje predstavlja vektor $x^{(P)}$ definiran sa

$$\begin{aligned} x_{j_k}^{(P)} &= \tilde{b}_k, \quad k = 1, 2, \ldots, r, \\ x_j^{(P)} &= 0, \quad j \in \{1, 2, \ldots, n\} \backslash \{j_1, j_2, \ldots, j_r\}, \end{aligned} \tag{7.2}$$

dakle,

$$x^{(P)} = \begin{array}{l} [0, \ldots, 0, \tilde{b}_1, 0, \ldots, 0, \tilde{b}_2, 0, \ \ldots\ , 0, \tilde{b}_r, 0, \ldots, 0]^T . \\ \ 1 \qquad\quad j_1 \qquad\qquad\quad j_2 \qquad\qquad\qquad\ j_r \qquad\quad n \end{array}$$

Provjerimo da je $x^{(P)}$ rješenje sustava $A_R x = b$. Ako označimo stupce od A_R s α_1, α_2, $\ldots$,α_n i vodimo računa da su stožerni stupci $\alpha_{j_k} = e_k$ stupci jedinične matrice I_m, dobijemo

$$\begin{aligned} A_R x^{(P)} &= [\alpha_1, \ldots, \alpha_n] x^{(P)} = x_1^{(P)} \cdot \alpha_1 + \cdots + x_n^{(P)} \cdot \alpha_n \\ &= \tilde{b}_1 \cdot \alpha_{j_1} + \tilde{b}_2 \cdot \alpha_{j_2} + \cdots + \tilde{b}_r \cdot \alpha_{j_r} \\ &= \tilde{b}_1 \cdot e_1 + \cdots + \tilde{b}_r \cdot e_r \\ &= \tilde{b}. \end{aligned}$$

Dakle je $A_R x^{(P)} = \tilde{b}$, pa je i $Ax^{(P)} = b$. Rješenje $x^{(P)}$ nazivamo **partikularno rješenje** sustava $Ax = b$.

Neka je x bilo koje rješenje jednadžbe $Ax = b$. Tada za vektor $x - x^{(P)}$ vrijedi

$$A(x - x^{(P)}) = Ax - Ax^{(P)} = b - b = 0.$$

Dakle, kad x prolazi skupom rješenja sustava $Ax = b$, $x - x^{(P)}$ prolazi skupom rješenja homogenog sustava $Ax = 0$. Obrnuto, neka x_h prolazi skupom rješenja homogenog sustava $Ax = 0$. Tada za vektor $x^{(P)} + x_h$ vrijedi

$$A(x^{(P)} + x_h) = Ax^{(P)} + Ax_h = b + 0 = b.$$

Prema tome, za svako rješenje x sustava $Ax = b$ postoji rješenje $x^{(H)}$ homogenog sustava $Ax = 0$ tako da vrijedi

$$x = x^{(P)} + x^{(H)}, \tag{7.3}$$

i obrnuto, za svako rješenje $x^{(H)}$ homogenog sustava $Ax = 0$, postoji rješenje x sustava $Ax = b$, koje je dano relacijom (7.3). Skup svih rješenja $x^{(H)}$ čini nul-potprostor $\mathcal{N}(A)$ matrice A, a vektore iz $\mathcal{N}(A)$ određujemo koristeći teorem 6.9. U kontekstu rješavanja sustava $Ax = b$, potprostor $\mathcal{N}(A)$ se naziva skup **homogenih rješenja** jednadžbe $Ax = b$.

Prethodnim razmatranjima dokazali smo

Teorem 7.1 *Neka je $A \in \mathbf{R}^{m\times n}$ i $b \in \mathbf{R}^m$. Ako je*

$r(A) < r([A \mid b])$ *onda sustav $Ax = b$ nema rješenja;*

$r(A) = r([A \mid b])$ *onda je opće rješenje sustava $Ax = b$ određeno relacijom (7.3), pri čemu je $x^{(H)}$ opće rješenje homogenog sustava $Ax = 0$, a $x^{(P)}$ je partikularno rješenje sustava $Ax = b$, određeno relacijom (7.2).* ■

Skup rješenja nehomogene linearne jednadžbe očito ima nekakvu strukturu. Takva struktura se zove **linearna mnogostrukost**.

Definicija 7.2 *Neka je X vektorski prostor, $\mathcal{Y} \subseteq X$ neki njegov potprostor i $x_0 \in X$ neki vektor. Skup*

$$\mathcal{M} = \{x \,:\, x = x_0 + y, \quad y \in \mathcal{Y}\}$$

naziva se linearna mnogostrukost u vektorskom prostoru X. **Dimenzija** *linearne mnogostrukosti $\mathcal{M}$ je dimenzija potprostora $\mathcal{Y}$. Skup $\mathcal{M}$ se obično označava $x_0 + \mathcal{Y}$.*

Linearna mnogostrukost $\mathcal{M}$ se kadkad naziva translatirani potprostor $\mathcal{Y}$ za vektor x_0. Kaže se i da je $\mathcal{M}$ paralelna s $\mathcal{Y}$. Ako je $x_0 \in \mathcal{Y}$, tada je $\mathcal{M} = \mathcal{Y}$, tj. linearna mnogostrukost je (netranslatirani) vektorski potprostor $\mathcal{Y}$. Ako je $\mathcal{Y} = \{0\}$, tada se $x_0 + \mathcal{Y}$ svodi na x_0. Dakle, svaki vektor je linearna mnogostrukost.

Sljedeći poznati rezultat je posljedica teorema 7.1. Zbog važnosti, zapisujemo ga u obliku teorema.

Teorem 7.3 (Kronecker-Cappeli) *Neka je $A \in \mathbf{R}^{m\times n}$ i $b \in \mathbf{R}^m$. Sustav $Ax = b$ ima rješenje ako i samo ako vrijedi*

$$r(A) = r([A \mid b]). \tag{7.4}$$

Ako jednakost (7.4) vrijedi, skup svih rješenja sustava $Ax = b$ čini linearnu mnogostrukost u $\mathbf{R}^n$ dimenzije $n - r(A)$.

Dokaz: Pretpostavimo da vrijedi jednakost (7.4). Prema teoremu 7.1 rješenje postoji, a skup svih rješenja čini linearnu mnogostrukost $x^{(P)} + \mathcal{N}(A)$. Prema teoremu 5.24, dimenzija nul-potprostora $\mathcal{N}(A)$ je $n - r(A)$.

Obrnuto, neka postoji rješenje $x = [x_1, \ldots, x_n]^T$ sustava $Ax = b$. Ako produkt Ax zapišemo pomoću stupaca a_j, $1 \leq j \leq n$, matrice A, imat ćemo

$$x_1 a_1 + \cdots + x_n a_n = b.$$

Dakle je b linearna kombinacija vektora a_j, pa je $b \in L(a_1, \ldots, a_n)$. To znači da je $L(a_1, \ldots, a_n) = L(a_1, \ldots, a_n, b)$. Zadnja jednakost se može zapisati kao $\mathcal{R}(A) = \mathcal{R}([A, \mid b\,])$. Uzimanjem dimenzija potprostora na obje strane, odmah se dobiva relacija (7.4). ∎

Primjer 7.4 *Neka je*

$$A = \begin{bmatrix} 0 & -2 & -4 & 0 & 6 & 0 \\ 0 & 1 & 2 & 3 & 15 & 0 \\ 0 & -3 & -6 & 1 & 15 & -5 \\ 0 & 3 & 6 & -2 & -21 & 6 \end{bmatrix}, \quad b = \begin{bmatrix} -8 \\ 10 \\ 0 \\ 5 \end{bmatrix}.$$

Riješite sustav $Ax = b$.

Rješenje. *Imamo*

$$\begin{aligned} [A \mid b\,] \;&=\; \left[\begin{array}{cccccc|c} 0 & -2 & -4 & 0 & 6 & 0 & -8 \\ 0 & 1 & 2 & 3 & 15 & 0 & 10 \\ 0 & -3 & -6 & 1 & 15 & -5 & 0 \\ 0 & 3 & 6 & -2 & -21 & 6 & 5 \end{array}\right] \\ &\overset{r}{\sim} \left[\begin{array}{cccccc|c} 0 & 1 & 2 & 0 & -3 & 0 & 4 \\ 0 & 0 & 0 & 3 & 18 & 0 & 6 \\ 0 & 0 & 0 & 1 & 6 & -5 & 12 \\ 0 & 0 & 0 & -2 & -12 & 6 & -7 \end{array}\right] \end{aligned}$$

$$\overset{r}{\sim}\left[\begin{array}{cccccc|c} 0 & 1 & 2 & 0 & -3 & 0 & 4 \\ 0 & 0 & 0 & 1 & 6 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & -5 & 10 \\ 0 & 0 & 0 & 0 & 0 & 6 & -3 \end{array}\right]$$

$$\overset{r}{\sim}\left[\begin{array}{cccccc|c} 0 & 1 & 2 & 0 & -3 & 0 & 4 \\ 0 & 0 & 0 & 1 & 6 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 9 \end{array}\right]$$

$$\overset{r}{\sim}\left[\begin{array}{cccccc|c} 0 & 1 & 2 & 0 & -3 & 0 & 0 \\ 0 & 0 & 0 & 1 & 6 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right] = [A \mid b]_R.$$

Vidimo da je $r(A) = 3 < 4 = r([A \mid b])$, pa sustav $Ax = b$ nema rješenja. ■

Primjer 7.5 *Riješite sustav $Ax = b$, ako je marica A ista kao u prethodnom primjeru, a vektor $b = \begin{bmatrix} -8 & 10 & 0 & -4 \end{bmatrix}^T$.*

Rješenje. *Imamo*

$$[A \mid b] = \left[\begin{array}{cccccc|c} 0 & -2 & -4 & 0 & 6 & 0 & -8 \\ 0 & 1 & 2 & 3 & 15 & 0 & 10 \\ 0 & -3 & -6 & 1 & 15 & -5 & 0 \\ 0 & 3 & 6 & -2 & -21 & 6 & -4 \end{array}\right]$$

$$\overset{r}{\sim}\left[\begin{array}{cccccc|c} 0 & 1 & 2 & 0 & -3 & 0 & 4 \\ 0 & 0 & 0 & 3 & 18 & 0 & 6 \\ 0 & 0 & 0 & 1 & 6 & -5 & 12 \\ 0 & 0 & 0 & -2 & -12 & 6 & -16 \end{array}\right]$$

$$\overset{r}{\sim}\left[\begin{array}{cccccc|c} 0 & 1 & 2 & 0 & -3 & 0 & 4 \\ 0 & 0 & 0 & 1 & 6 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & -5 & 10 \\ 0 & 0 & 0 & 0 & 0 & 6 & -12 \end{array}\right]$$

$$\overset{r}{\sim}\left[\begin{array}{cccccc|c} 0 & 1 & 2 & 0 & -3 & 0 & 4 \\ 0 & 0 & 0 & 1 & 6 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right] = [A \mid b]_R.$$

Vidimo da je $r(A) = r([A \mid b\,]) = 3$*, pa sustav* $Ax = b$ *ima rješenje. Rješenje je linearna mnogostrukost dimenzije* $6 - 3 = 3$*. Partikularno rješenje je prema relaciji (7.2) dano sa*

$$x^{(P)} = \begin{bmatrix} 0 & 4 & 0 & 2 & 0 & -2 \end{bmatrix}^T .$$

Preostala rješenja dobivamo koristeći homogena rješenja. Opće rješenje homogene jednadžbe $Ax = 0$ *je opisano u teoremu 6.9. Imamo*

$$x^{(H)} = \alpha_1 \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \alpha_2 \begin{bmatrix} 0 \\ -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \alpha_3 \begin{bmatrix} 0 \\ 3 \\ 0 \\ -6 \\ 1 \\ 0 \end{bmatrix} .$$

Stoga je opće rješenje nehomogenog sustava $Ax = b$*,*

$$x = x^{(P)} + x^{(H)} = \begin{bmatrix} 0 \\ 4 \\ 0 \\ 2 \\ 0 \\ -2 \end{bmatrix} + \alpha_1 \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \alpha_2 \begin{bmatrix} 0 \\ -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \alpha_3 \begin{bmatrix} 0 \\ 3 \\ 0 \\ -6 \\ 1 \\ 0 \end{bmatrix}$$

ili

$$x = \begin{bmatrix} \alpha_1 \\ 4 - 2\alpha_2 + 3\alpha_3 \\ \alpha_2 \\ 2 - 6\alpha_3 \\ \alpha_3 \\ -2 \end{bmatrix} .$$

■

Pretpostavimo da znamo neko rješenje z sustava $Ax = b$, koje nije nužno $x^{(P)}$ iz relacije (7.2). Dakle, za z vrijedi $Az = b$. Možemo li kazati da je skup svih rješenja sustava $Ax = b$ linearna mnogostrukost $z + \mathcal{N}(A)$? Potvrdan odgovor slijedi iz

Propozicija 7.6 *Neka je $\mathcal{M} = x_0 + \mathcal{Y}$ linearna mnogostrukost. Ako je $z \in \mathcal{M}$, tada je $\mathcal{M} = z + \mathcal{Y}$.*

Dokaz: Moramo pokazati da je $x_0+\mathcal{Y} \subseteq z+\mathcal{Y}$ i obrnuto $z+\mathcal{Y} \subseteq x_0+\mathcal{Y}$. Koristit ćemo činjenicu da je $z = x_0 + z_1$, za neko $z_1 \in \mathcal{Y}$.

Neka je $u \in x_0 + \mathcal{Y}$. Tada je on oblika $u = x_0 + y$ za neko $y \in \mathcal{Y}$. To možemo zapisati kao $u = x_0 + z_1 + (y - z_1) = z + w$. Jer je $\mathcal{Y}$ potprostor i $y, z_1 \in \mathcal{Y}$ mora i $w = y - z_1$ biti u $\mathcal{Y}$. Dakle je $u \in z + \mathcal{Y}$. Zbog proizvoljnosti od u zaključujemo da je $x_0 + \mathcal{Y} \subseteq z + \mathcal{Y}$.

Obratno, neka je $u \in z + \mathcal{Y}$. Tada je on oblika $u = z + y'$ za neko $y' \in \mathcal{Y}$. To možemo zapisati kao $u = x_0 + z_1 + y' = x_0 + w'$. Jer je $\mathcal{Y}$ potprostor i $z_1, y' \in \mathcal{Y}$ mora i $w' = z_1 + y'$ biti u $\mathcal{Y}$. Dakle je $u \in x_0 + \mathcal{Y}$. Zbog proizvoljnosti od u zaključujemo da je $z + \mathcal{Y} \subseteq x_0 + \mathcal{Y}$. ■

Posebno je zanimljiv slučaj kad rješenje sustava $Ax = b$ postoji, a rang matrice jednak je broju stupaca matrice. Za takvu matricu se kaže da ima puni stupčani rang[1].

Korolar 7.7 *Neka je $A \in \mathbf{R}^{m\times n}$, $m \geq n$ i $b \in \mathbf{R}^m$. Ako vrijedi*

$$r(A) = r([A \mid b]) = n,$$

onda sustav $Ax = b$ ima jedinstveno rješenje x. Ako se algoritam 6.3 primijeni na proširenu matricu $[A \mid b]$, tada se komponente rješenja dobiju iz zadnjeg stupca reducirane matrice $[A \mid b]_R = [A_R \mid \tilde{b}]$,

$$x_i = \tilde{b}_i, \quad i = 1, 2, \ldots, n; \quad \tilde{b} = (\tilde{b}_i).$$

Ako još vrijedi $m = n$, onda je

$$x = \tilde{b}.$$

Dokaz: Primijenimo algoritam 6.3 na proširenu matricu $[A, \mid b]$. Kako je rang matrice A (a to je i broj netrivijalnih redaka od A_R) jednak broju stupaca od A (i A_R), reducirana proširena matrica ima oblik

[1]Matrica $A \in \mathbf{R}^{m\times n}$ ima puni rang ako vrijedi $r(A) = \min\{m, n\}$. Kako $r(A) = n$ povlači $n \leq m$ možemo iskaz: "A ima puni stupčani rang", zamijeniti iskazom "A ima puni rang i vrijedi $n \leq m$".

$$[A_R \mid \tilde{b}\,] = \left[\begin{array}{cccccccc|c} 1 & 0 & 0 & 0 & 0 & \dots & & 0 & \tilde{b}_1 \\ & 1 & 0 & 0 & 0 & \dots & & 0 & \tilde{b}_2 \\ & & 1 & 0 & 0 & \dots & & 0 & \tilde{b}_3 \\ & & & 1 & 0 & \dots & & 0 & \tilde{b}_4 \\ & & & & 1 & \dots & & 0 & \tilde{b}_5 \\ 0 & & & & & \ddots & & & \vdots \\ & & & & & & & 1 & \tilde{b}_n \\ \hline 0 & & 0 & & 0 & & & 0 & 0 \\ 0 & & 0 & & 0 & & & 0 & 0 \\ \dots & & & & & & & & \vdots \\ 0 & & 0 & & 0 & & & 0 & 0 \end{array}\right] = \left[\begin{array}{c|c} I_n & \tilde{b}_0 \\ \hline 0 & 0 \end{array}\right].$$

Iz definicije (7.2) vektora $x^{(P)}$ direktno slijedi $x^{(P)} = \tilde{b}_0$. Budući da je rang matrice A jednak n, po teoremu 5.24 je dimenzija od $\mathcal{N}(A)$ jednaka $n - r(A) = n - n = 0$, pa homogeni sustav $Ax = 0$ ima samo trivijalno rješenje $x^{(H)} = 0$. Prema tome opće rješenje nehomogenog sustava je dano sa

$$x = x^{(P)} + x^{(H)} = x^{(P)}.$$

Ako je $m = n$ onda je i $x = x^{(P)} = \tilde{b}_0 = \tilde{b}$. ■

Primjer 7.8 *Riješite sustav $Ax = b$, ako je*

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

Rješenje.

$$[A \mid b] \;=\; \left[\begin{array}{cccc|c} 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & -1 & 1 \\ 1 & -1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & -2 \end{array}\right] \overset{r}{\sim} \left[\begin{array}{cccc|c} 1 & -1 & 0 & 0 & 1 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 1 & 1 & 2 \\ 1 & 1 & 0 & 0 & -2 \end{array}\right]$$

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

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

$$
\overset{r}{\sim}\left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & -1/2 \\ 0 & 1 & 0 & 0 & -3/2 \\ 0 & 0 & 1 & 0 & 3/2 \\ 0 & 0 & 0 & 1 & 1/2 \end{array}\right] = [A \mid b]_R \,.
$$

Rang matrice A kao i proširene matrice $[A \mid b]$ jednak je broju stupaca matrice A, tj. 4. Stoga postoji samo jedno rješenje sustava $Ax = b$, koje je dano sa

$$
x = \tilde{b} = \begin{bmatrix} -\frac{1}{2} & -\frac{3}{2} & \frac{3}{2} & \frac{1}{2} \end{bmatrix}^T .
$$

7.2 Kvadratni sustavi i Gaussove eliminacije

U praksi se najčešće javljaju sustavi kod kojih je broj linearnih jednadžbi jednak ili je veći od broja nepoznanica. U prvom slučaju je matrica sustava A kvadratna, pa sustave zovemo kvadratni. U drugom slučaju matrica ima više redaka nego stupaca. Kod takovih sustava općenito ne postoji vektor x koji bi zadovoljavao jednadžbu $Ax = b$, pa se traži onaj vektor x za koji je kvadrat Euklidske norme $\|Ax - b\|_2^2$ najmanji. Takovih vektora može biti više pa se obično traži vektor najmanje norme koji minimizira $\|Ax - b\|_2^2$. S obzirom da je Euklidska norma suma kvadrata (modula) komponenata, taj problem je poznat pod imenom: problem najmanjih kvadrata. Za njegovo rješavanje koriste se ortogonalne transformacije, posebno tzv. QR faktorizacija i singularna dekompozicija, pa rješavanje tog sustava nećemo razmatrati. Stoga ćemo se posvetiti za praksu još važnijem slučaju kvadratnih sustava. Do kraja ove lekcije, pretpostavljamo da je u sustavu $Ax = b$

matrica A reda n i ranga n. U tom slučaju postoji jedinstveno rješenje x. Takva matrica ima reduciranu matricu jediničnu, a postojanje i jedinstenost rješenja sustava $Ax = b$ slijede iz korolara 7.7.

Sustave s malo (npr. do 10) nepoznanica možemo uz malo spretnosti i strpljivosti, na osnovi algoritma 6.3 riješiti olovkom i papirom. Problem nastaje onda kad treba riješiti velike sustave. U primjenama nerijetko susrećemo sustave s matricama reda 1000 i više. Vrlo drastičan primjer je numeričko rješavanje parcijalnih diferencijalnih jednadžbi u trodimenzionalnom prostoru. Kod aproksimiranja diferencijalnih operatora pomoću diferencijskih shema ili korištenjem drugih tehnika (npr. konačnih elemenata), mogu se dobiti sustavi s ogromnim brojem nepoznanica. Kod tog aproksimativnog rješavanja diferencijalnih jednadžbi, svađanje na sustav linearnih jednadžbi uvijek znači da će točno rješenje sustava biti tek određena (možda loša) aproksimacija rješenja polaznog problema diferencijalnih jednadžbi. Koji puta za jednu decimalu točnosti rezultata polaznog problema, treba rješavati sustav s 1000 nepoznanica, za dvije decimale točnosti sustav s 10^6 nepoznanica, za tri decimale sustav s 10^9 nepoznanica, itd. Kod takovih matrica jedino što nam preostaje je napisati kompjutorski program koji će sustav riješiti u dogledno vrijeme.

Kod numeričkog rješavanja sustava uz pomoć računala, javlja se novi problem: računala koriste tzv. aritmetiku konačne preciznonosti. Tipičan zapis jednog broja u računalu ima oblik: $\pm\, m \times 2^e$, gdje je m mantisa, a e eksponent. Mantisa dopušta točno određen broj znamenaka, a za eksponent postoji točno određen interval u kojem se on može nalaziti. U računalu su i mantisa i eksponent smješteni kao binarni brojevi. Tipično, u tzv. aritmetici jednostruke (dvostruke) preciznosti mantisa ima prema IEEE standardu koji je prisutan kod gotovo svih računala, 24 (53) binarne znamenke, dok je je eksponent u zatvorenom inervalu $[-125\,,\,128]$ ($[-1021\,,\,1024]$). To znači da aritmetika radi s brojevima u rasponu od približno -3.4×10^{38} do $+3.4 \times 10^{38}$ (od -1.8×10^{308} do 1.8×10^{308}) pri čemu mantisa ima, preračunato oko 7 (16) dekadskih znamenki. Ako želimo u računalo učitati broj 3/7 ili π, ono će ga zapisati kao binarni broj sa 24 (ili 53) binarnih znamenaka, a ostale znamenke će odbaciti. Čak i brojevi kao 1/10, 1/5, 2/5, 3/5, ... ne mogu se egzaktno zapisati, jer njihove binarne mantise imaju beskonačno znamenki. To znači da već kod smještavanja matrice sustava A u računalo, većina smještenih elemenata će nositi grešku u zadnjoj binarnoj znamenci. To su tzv. relativne greške u brojevima koje dolaze od smještavanja brojeva u računalo. One su omeđene s $6.8 \cdot 10^{-8}$ za slučaj jednostruke odnosno $1.11 \cdot 10^{-16}$ za slučaj dvostruke preciznosti.

Kada se pokrene neki algoritam, gotovo svaka aritmetička operacija će unijeti malu relativnu grešku u rezultat. Npr. kad množimo brojeve čije mantise imaju t znamenaka, mantisa rezultata će općenito imati $2t$ znamenaka, pa će se morati zaokružiti na t mjesta. Slično zaokruživanje će postojati i kod drugih računskih operacija. Pa koliko ima operacija kod rješavanja sustava reda n? Pokazuje se da ih ima oko $(2/3)n^3$. To znači da za rješavanje sustava reda 1000 treba oko $6.7 \cdot 10^8$, a za rješavanje sustava reda 10^9 oko $6.7 \cdot 10^{26}$ računskih operacija. Dakle, možemo strahovati da toliki broj grešaka zaokruživanja može dosta utjecati na točnost izračunatog rezultata. Vidimo da je općenito teško doći do iole točnog rješenja gore spomenutih diferencijalnih jednadžbi, jer bolja aproksimacija diferencijalnih jednadžbi sustavom linearnih jednadžbi znači veliki red matrice, a onda nastupaju problemi sa računanjem zbog iskrslih grešaka zaokruživanja.

Što se može poduzeti? Prva mogućnost koja pada na pamet jest računati uz mnogo veći broj decimala, dakle uz aritmetiku veće preciznosti. Za to su potrebni posebni softwareski paketi, jer kupljena računala dolaze uz standardni software koji omogućuje rad samo s aritmetikama jednostruke i dvostruke preciznosti. Spomenuti softwareski paketi za računanje u povećanoj točnosti toliko će usporiti računanje, da će već rješavanje sustava reda 100 biti dugotrajan posao. Uočivši taj problem, neiskusan rješavač sustava će reći: “Kupimo onda jače računalo”. To će osim izazivanja velikih troškova problem samo malo ublažiti, ali ne i riješiti. Ne smijemo zaboraviti da i najveća svjetska računala nisu svemoćna. Tako se npr. uz pomoć najsnažnijih računala 24-satna vremenska prognoza za cijelu državu SAD računa tek do na točnost od ± 10 kilometara. Zato je uvijek dobar pristup maksimalno iskoristiti analitičko znanje o problemu i o rješenju, te umjesto standardnih algoritama izmisliti posebni algoritam, ili modificirati neki postojeći algoritam za dani problem.

Kako bi se utjecaj grešaka zaokruživanja na točnost rješenja sustava smanjio, potrebno je modificirati algoritam 6.3. Općenito je u algoritmu potrebno više zahvata:

1. umjesto svađanja na jediničnu matricu A_R, bolje je koristiti elementarne operacije za svađanje matrice na gornje-trokutasti oblik;

2. potrebno je izostaviti dijeljenje pivotnog retka sa pivotnim elementom jer kad je pivotni element mali to može dovesti do velikog povećanja elemenata (pa i prisutnih grešaka) u elementima pivotnog retka;

3. potrebno je izabirati pivotni element tako da je po mogućnosti što veći.

Ove modifikacije nas dovode do **Gaussove metode eliminacija**, za razliku od opisanog algoritma 6.3 koji se spominje (posebno u kontekstu rješavanja sustava) kao **Gauss-Jordanov** algoritam. Kod Gauss-Jordanovog postupka (svaka) nepoznanica x_k se eliminira iz svih jednadžbi osim iz k-te. Kod Gaussovih eliminacija, prvo se nepoznanica x_1 eliminira iz jednadžbi 2—n, zatim nepoznanica x_2 iz jednadžbi 3—n, onda x_3 iz jednadžbi 4—n, itd. Algoritam se koristi uz dvije strategije izbora pivotnih elemenata: parcijalnog i kompletnog pivotiranja. S obzirom da se u praksi najčešće koristi parcijalno pivotiranje, kompletno pivotiranje ćemo razmatrati tek na nivou ideje.

Parcijalno pivotiranje

Kod parcijalnog pivotiranja, pivotni element u r-tom koraku izabire se kao najveći po modulu element u pivotnom (tj. r-tom) stupcu, od r-tog retka na niže.

Algoritam 7.9 *Algoritam vrši elementarne retčane operacije na matrici X. Oznaka $X = (x_{ij})$ se koristi za iteriranu matricu. Sa x_k^T je označen k-ti redak od X. Na ulazu, X se poistovjeti s proširenom matricom $[A \mid b] \in \mathbf{R}^{n\times(n+1)}$. Na izlazu X je oblika $[R \mid \tilde{b}]$, gdje je R gornje-trokutasta matrica.*

1. *Stavi:* $X = [A \mid b], \;\; r = 1.$

2. *Pronađi najmanji indeks* r' *takav da je*

$$| \, x_{r'r} \, |= \max\{| \, x_{kr} \, |; \;\; r \leq k \leq n\}$$

3. *Ako je* $x_{r'r} = 0$*, idi na 6, inače idi na 4.*

4. *Zamijeni retke s indeksima* r *i* r'.

5. *Za* $i = r+1, \ldots, n$*, pomnoži* r*-ti redak matrice brojem* $m_{ir} = x_{ir}/x_{rr}$ *i tako dobiveni redak oduzmi od* i*-tog retka:* $x_i^T \leftarrow x_i^T - m_{ir} \cdot x_r^T$.

6. *Stavi* $r := r + 1$. *Ako je* $r = n$ *idi na 7, inače idi na 2.*

7. *Stavi* $[R \mid \tilde{b}\,] = X$ *i stani.*

U slučaju izostavljanja permutacije redaka (izostave se koraci 2., 3. i 4. algoritma) govori se metodi Gaussove eliminacije bez pivotiranja. Ona će stati ako je pivotni element x_{rr} nula odnosno generirati veće greške u matrične elemente ako je $|x_{rr}|$ mali.

Promotrimo algoritam 7.9. Neka je r prvi indeks za koji će se dogoditi skok s 3. na 6. korak. U tom slučaju je r-ti stupac tekuće matrice X zavisan od prethodnih stupaca (za to je dovoljno da su svi dijagonalni elementi a_{kk}, $1 \leq k \leq r-1$ različiti od nule) pa X ima rang manji od n. Kako elementarne operacije nad retcima ne mijenjaju rang, mora i matrica $[A \mid b\,]$ imati rang manji od n. To znači da je i $r(A) < n$. U tom slučaju rješenje ili ne postoji ili postoji, ali nije jedinstveno. Ta alternativa se razrješava u toku rješavanja sustava $Rx = \tilde{b}$. Ako se x_r dobije iz kvocijenta oblika nešto/0 rješenje ne postoji, a ako se dobije iz oblika 0/0 postojat će beskonačno mnogo rješenja. Primijetimo također da algoritam neće stati sve do prirodnog završetka kad r postane n.

Kako smo mi pretpostavili da je $r(A) = n$, algoritam će uvijek naići na netrivijalne pivotne elemente, koji će se na kraju procesa svi nalaziti na dijagonali matrice R.

Glavni dobitak kod pivotiranja je da su multiplikatori m_{ir} koji množe pivotne retke mali, tj. omeđeni su s 1 po modulu. To omogućuje istraživanje točnosti (još se kaže stabilnosti) algoritma.

Na izlazu algoritam 7.9 daje gornje-trokutastu matricu R i vektor $\tilde{b}$ koji je nastao iz vektora b primjenom istih retčanih operacija koje su svele A na R.

Preostaje još riješiti trokutasti sustav

$$Rx = \tilde{b}.$$

Ako taj sustav napišemo u obliku

$$\begin{aligned}
r_{11}x_1 + r_{12}x_2 + \cdots r_{1,n-1}x_{n-1} + r_{1n}x_n &= \tilde{b}_1 \\
r_{22}x_2 + \cdots r_{2,n-1}x_{n-1} + r_{2n}x_n &= \tilde{b}_2 \\
\vdots \quad &\quad \vdots \\
r_{n-1,n-1}x_{n-1} + r_{n-1,n}x_n &= \tilde{b}_{n-1} \\
r_{nn}x_n &= \tilde{b}_n
\end{aligned}$$

vidimo da možemo odmah iz zadnje jednadžbe dobiti x_n. Uvrštavajući dobivenu vrijednost za x_n u predzadnju jednadžbu, možemo dobiti x_{n-1}. Nastavljajući tako, dobijemo sve komponente vektora x. Kako ih dobivamo u redoslijedu: x_n, x_{n-1},. . . ,x_1, ovaj algoritam se zove **obratni postupak** ili **povratna supstitucija**. Povratni postupak ima daleko manje operacija od eliminacija. Dok eliminacije zahtijevaju oko $n^2/2$ dijeljenja, $n^3/3$ množenja i $n^3/3$ zbrajanja (odnosno oduzimanja) dotle povratni postupak traži samo n dijeljenja te oko $n^2/2$ množenja i $n^2/2$ zbrajanja (odnosno oduzimanja).

Primjer 7.10 *Pomoću Gaussove metode eliminacija s parcijalnim pivotiranjem riješite sustav $Ax = b$, gdje su:*

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 0 & 1 \\ 1 & -1 & 0 & -1 \\ 2 & 3 & -2 & -2 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 7 \\ 0 \\ 10 \end{bmatrix}.$$

Rješenje. *Prikazat ćemo matrice dobivene nakon zamjene redaka i nakon eliminacije nepoznanice (tj. nakon stvaranja nula u stupcu ispod dijagonalnog elementa))*

$$[A \mid b] = \left[\begin{array}{cccc|c} 1 & 2 & 3 & 4 & 1 \\ 4 & 2 & 0 & 1 & 7 \\ 1 & -1 & 0 & -1 & 0 \\ 2 & 3 & -2 & -2 & 10 \end{array}\right]$$

$$
\begin{aligned}
&\overset{r}{\sim}\left[\begin{array}{cccc|c} 4 & 2 & 0 & 1 & 7 \\ 1 & 2 & 3 & 4 & 1 \\ 1 & -1 & 0 & -1 & 0 \\ 2 & 3 & -2 & -2 & 10 \end{array}\right] \overset{r}{\sim}\left[\begin{array}{cccc|c} 4 & 2 & 0 & 1 & 7 \\ 0 & \frac{3}{2} & 3 & \frac{14}{4} & -\frac{3}{4} \\ 0 & -\frac{3}{2} & 0 & -\frac{5}{4}1 & -\frac{7}{4} \\ 0 & 2 & -2 & -\frac{5}{2}2 & \frac{13}{2} \end{array}\right] \\
&\overset{r}{\sim}\left[\begin{array}{cccc|c} 4 & 2 & 0 & 1 & 7 \\ 0 & 2 & -2 & -\frac{5}{2}2 & \frac{13}{2} \\ 0 & -\frac{3}{2} & 0 & -\frac{5}{4}1 & -\frac{7}{4} \\ 0 & \frac{3}{2} & 3 & \frac{14}{4} & -\frac{3}{4} \end{array}\right] \overset{r}{\sim}\left[\begin{array}{cccc|c} 4 & 2 & 0 & 1 & 7 \\ 0 & 2 & -2 & -\frac{5}{2} & \frac{13}{2} \\ 0 & 0 & -\frac{3}{2} & -\frac{25}{8} & \frac{25}{8} \\ 0 & 0 & \frac{9}{2} & \frac{45}{8} & -\frac{45}{8} \end{array}\right] \\
&\overset{r}{\sim}\left[\begin{array}{cccc|c} 4 & 2 & 0 & 1 & 7 \\ 0 & 2 & -2 & -\frac{5}{2} & \frac{13}{2} \\ 0 & 0 & \frac{9}{2} & \frac{45}{8} & -\frac{45}{8} \\ 0 & 0 & -\frac{3}{2} & -\frac{25}{8} & \frac{25}{8} \end{array}\right] \overset{r}{\sim}\left[\begin{array}{cccc|c} 4 & 2 & 0 & 1 & 7 \\ 0 & 2 & -2 & -\frac{5}{2} & \frac{13}{2} \\ 0 & 0 & \frac{9}{2} & \frac{45}{8} & -\frac{45}{8} \\ 0 & 0 & 0 & -\frac{10}{8} & \frac{10}{8} \end{array}\right] \\
= \quad & [R \mid \tilde{b}].
\end{aligned}
$$

Sustav $Rx = \tilde{b}$ rješavamo obratnim postupkom. Prvo se izračuna x_4,

$$x_4 = \frac{10}{8}/(-\frac{10}{8}) = -1.$$

Zatim se -1 uvrsti umjesto x_4 u treću jednadžbu, koja se onda riješi

$$x_3 = \frac{2}{9} \cdot \left[-\frac{45}{8} - \frac{45}{8} \cdot (-1)\right] = 0.$$

Nepoznanica x_2 se dobije iz druge jednadžbe. Prethodno se u nju uvrste vrijednosti za x_3 i x_4,

$$x_2 = \frac{1}{2} \cdot \left[\frac{13}{2} + 2 \cdot 0 + \frac{5}{2} \cdot (-1)\right] = 2.$$

Konačno, x_1 se dobije iz prve jednadžbe

$$x_2 = \frac{1}{4} \cdot [7 - 2 \cdot 2 - 0 \cdot 0 - 1 \cdot (-1)] = 1.$$

■

Napišimo algoritam Gaussovih eliminacija s parcijalnim pivotiranjem u obliku pogodnom za pisanje programa na računalu. Pritom je zgodno spremiti multiplikatore m_{ij} na poziciju (i, j), $i > j$ matrice A, dakle u element a_{ij} (nakon što on, prema algoritmu 7.9 postane nula).

Algoritam 7.11 *Neka je* $A \in \mathbf{R}^{n\times(n+1)}$*, pri čemu prvih* n *stupaca matrice čine stupci od matrice sustava, a* $n+1$*. stupac je vektor* b*. Na izlazu, prvih* n *stupaca matrice čine stupci od* R*, a zadnji stupac je rješenje polaznog sustava.*

1. *za* $r = 1, 2, \ldots, n-1$

1.1. *pronađi najmanji indeks* r' *takav da je*

$$| a_{r'r} | \geq | a_{kr} |, \; r \leq k \leq n$$

Ako je $a_{r'r} = 0$ *idi na 2. inače idi na 1.2*

1.2 *zamijeni retke s indeksima* r *i* r'

1.3 *za* $i = r+1, r+2, \ldots, n+1$

1.3.1 $a_{ir} = a_{ir}/a_{rr}$

1.3.2 *za* $j = r+1, r+2, \ldots, n+1$

1.3.3 $a_{ij} = a_{ij} - a_{ir}a_{rj}$

1.3.4 *kraj petlje po* j

1.4 *kraj petlje po* i

1.5 *kraj petlje po* r

Rješavanje sustava $Rx = \tilde{b}$*. Nakon provjere da su svi dijagonalni elementi različiti od* 0*, algoritam glasi.*

2. $a(n, n+1) = a(n, n+1)/a(n, n)$

2.1 *za* $i = n-1, n-2, \ldots, 1$

2.1.1 *za* $j = i+1, \ldots, n$

2.1.2 $a(i, n+1) = a(i, n+1) - a(i, j) * a(j, n+1)$

2.1.3 *kraj petlje po* j

2.2 $a(i, n+1) = a(i, n+1)/a(i, i)$

2.3. *kraj petlje po* i

Uočimo da Gaussov algoritam u slučaju beskonačno rješenja sustava ne daje bazu nul-potprostora, pa je u tom smislu slabiji od Gauss-Jordanovog algoritma. Ipak, iz dobivene matrice $[R, \tilde{b}]$ moguće je dobiti dodatnim elementarnim retčanim operacijama matricu $[A_R, \tilde{b}']$ pri čemu se iz A_R i $\tilde{b}'$ odredi opće rješenje sustava (partikularno rješenje i baza nul-potprostora). Cijena koja se za to plaća su osim dodatnih računskih operacija i potencijalno veliki multiplikatori koji prave nule u gornjem trokutu matrice (oni vrše eliminaciju nepoznanice iz prethodnih jednadžbi). To može dovesti na računalima do velikih grešaka u elementima A_R i u rješenju x. Kako se najčešće, unaprijed zna da je matrica sustava ranga n, na računalima se gotovo isključivo koristi Gaussova metoda eliminacija sa parcijalnim pivotiranjem.

7.3 Kompletno pivotiranje

Kompletno pivotiranje je još točniji algoritam od parcijalnog pivotiranja, ali zahtijeva mnogo više računskog vremena oko pronalaženja pivotnog elementa u svakom koraku, a osim zamjene redaka treba činiti i zamjenu stupaca matrice. Zato se u toku postupka osim proširene matrice pamti i vektor čija k-ta komponenta predstavlja indeks stupca koji se zamijenio sa k-tim (pivotnim) stupcem u k-tom koraku. U prvom koraku algoritma treba pronaći najveći po modulu element u cijeloj matrici i onda zamjenom redaka i stupaca dovesti ga na $(1, 1)$– poziciju. Zatim se izvrši eliminacija nepoznanice x_1 iz 2.—n. jednadžbe. Time je gotov prvi korak metode. Zatim se prvi korak ponavlja na podmatrici reda $n-1$ na presjeku zadnjih $n-1$ redaka i stupaca. Najbolje je raditi s proširenom matricom jer se sve retčane operacije ionako rade i na vektoru b. Na kraju se dobije gornje-trokutasta matrica R, vektor $\tilde{b}$ i vektor p ineksa pri zamjenama stupaca. Zatim se riješi trokutasti sustav $Rx = \tilde{b}$ i na kraju ispermutiraju komponente od x prema indeksima zapamćenim u vektoru p. Metoda se rijetko koristi jer je dosta sporija od metode s parcijalnim pivotiranjem, a ne daje bitno točnije rezultate.

Lekcija 8

Linearne Matrične Jednadžbe

Linearne matrične jednadžbe se najčešće pojavljuju u jednom od sljedećih oblika

$$
\begin{aligned}
A+X &= B && (8.1)\\
AX &= B && (8.2)\\
XA &= B && (8.3)\\
AX+B &= C && (8.4)\\
XA+B &= C && (8.5)\\
AX+XB &= C. && (8.6)
\end{aligned}
$$

U ovim jednadžbama A, B, C su poznate (zadane) matrice, a matrica X je ona koju tražimo. Da bi te jednadžbe imale smisla, moramo paziti na dimenzije matrica. Tako npr. u jednadžbi $A+X=B$ sve matrice moraju imati isti broj redaka i isti broj stupaca, a u jednadžbi $AX=B$, ako je $A \in \mathbf{R}^{m\times n}$, onda Xmora imati n redaka, B mora imati m redaka, a obje matrice X i B moraju imati isti broj stupaca. Zato se uzima da su $B \in \mathbf{R}^{m\times p}$ i $X \in \mathbf{R}^{n\times p}$ za neki prirodan broj p.

Rješenje jednadžbe $A+X=B$, kao i jednadžbe $X+A=B$ je

$$X = B - A\,,$$

i time je problem (8.1) riješen. Jednadžbu (8.3) možemo transponiranjem lijeve i desne strane svesti na oblik

$$A^T X^T = B^T$$

pa smo time problem (8.3) sveli na problem (8.2). Jednadžbu (8.4) oduzimanjem matrice B od lijeve i desne strane jednadžbe svodimo na oblik $AX = C - B$, pa je problem opet reduciran na (8.2). Isto tako, jednadžbu $XA + B = C$ oduzimanjem i transponiranjem svodimo na oblik (8.2). Preostaje jednadžba

$$AX + XB = C$$

koja se zove Ljapunovljeva jednadžba. Njeno rješavanje je dosta kompliciranije od rješavanja jednadžbe $AX = B$ i za sada ju nećemo razmatrati. Prema tome, preostaje riješiti jednadžbu (8.2). Preciznije, zanimat će nas pitanje egzistencije, jedinstvenosti i algoritma za taj problem.

Od posebnog interesa su i jednadžbe

$$AX = I$$

odnosno

$$XA = I$$

jer nas vode do pojmova desnog i lijevog inverza matrice. Te ćemo slučajeve razmatrati na kraju poglavlja.

8.1 Jednadžba $AX = B$

Promotrimo jednadžbu

$$AX = B, \quad A \in \mathbf{R}^{m\times n},\ X \in \mathbf{R}^{n\times p},\ B \in \mathbf{R}^{m\times p}.$$

Uvedimo stupčane particije

$$X = \left[\ x_1, x_2, \ldots, x_p\ \right] \quad \text{i} \quad B = \left[\ b_1, b_2, \ldots, b_p\ \right].$$

Po pravilu matričnog množenja vrijedi

$$AX = A\left[\ x_1, x_2, \ldots, x_p\ \right] = \left[\ Ax_1\,,\ Ax_2\,,\ \ldots\,,\ Ax_p\ \right] \tag{8.7}$$

Stoga se jednadžba $AX = B$ može zapisati kao p jednadžbi

$$Ax_1 = b_1, \quad Ax_2 = b_2, \quad \dots \quad Ax_p = b_p. \tag{8.8}$$

Jednadžba $AX = B$ je ekvivalentna sustavu od p jednadžbi (8.8). To znači: ako $AX = B$ ima rješenje, onda i svaka jednadžba $Ax = b_i$ ima rješenje i obrnuto, ako svaka jednadžba $Ax = b_i$ ima rješenje, tada ga ima i $AX = B$.

To se jednostavno vidi. Neka npr. $AX = B$ ima rješenje X. Tada zbog relacije (8.7), odmah slijedi (8.8). Pritom su rješenja sustava upravo stupci od X. Obrnuto, neka svaka jednadžba sustava (8.8) ima rješenje. Recimo, neka je x_i bilo koje rješenje od $Ax = b_i$ i tako za sve $1 \leq i \leq p$. Sada iz tih vektora x_i sagradimo matricu X, zahtijevajući da je x_i upravo i-ti stupac od X. Gledajući relaciju (8.7) i stupčanu particiju od X, odmah vidimo da vrijedi $AX = B$, pa smo dakle, sagradili rješenje X.

Sustav oblika $Ax = b_i$ smo razmatrali u 7. poglavlju i znamo da se može riješiti npr. algoritmom 6.3 koji svađa proširenu matricu na reducirani oblik ili korištenjem Gaussove metode eliminacija s parcijalnim ili kompletnim pivotiranjem. Budući da algoritam 6.3 (isto vrijedi i za Gaussovu metodu) ide "s lijeva u desno", prvo prevodi A na reducirani oblik A_R (odnosno na trokutastu matricu R), možemo izbjeći ponavljanje redukcije matrice A na A_R (ili R) $p-1$ puta, ako primijenimo algoritam samo jedamput na proširenu matricu $[A \mid B]$.

Kakva rješenja X možemo očekivati? Imamo li jedno, nijedno ili beskonačno rješenja? Iz sljedećeg teorema ćemo vidjeti da odgovor ovisi o odnosu ranga matrice A i ranga proširene matrice $[A \mid B]$.

Teorem 8.1 *Jednadžba*

$$AX = B \tag{8.9}$$

ima rješenje ako i samo ako vrijedi

$$r(A) = r([A \mid B]). \tag{8.10}$$

Ako je uvjet (8.10) ispunjen i ako je $X^{(P)}$ neko rješenje jednadžbe (8.9), onda je opće rješenje jednadžbe (8.9) oblika

$$X = X^{(P)} + X^{(H)},$$

gdje je $X^{(H)}$ opće rješenje homogene matrične jednadžbe

$$AX = O\,. \tag{8.11}$$

Specijalno, jednadžba (8.9) ima jedinstveno rješenje ako i samo ako homogena jednadžba (8.11) ima nul-matricu kao jedino rješenje.

Dokaz: Neka postoji rješenje X jednadžbe (8.9). Tada postoji i p rješenja x_i sustava (8.8). Po Kronecker-Capellijevom teoremu (teorem 7.3) za svako $1 \leq i \leq p$ vrijedi $r(A) = r([A \mid b_i)$. To znači $b_i \in \mathcal{R}(A) = L(a_1, \ldots, a_n)$ za sve i, gdje su a_1, $\ldots$,a_n stupci od A. Dakle je

$$\begin{aligned}
\mathcal{R}(A) &= L(a_1, \ldots, a_n) = L(a_1, \ldots, a_n, b_1) \\
&= L(a_1, \ldots, a_n, b_1, b_2) = \cdots \\
&= L(a_1, \ldots, a_n, b_1, \ldots, b_p) \\
&= \mathcal{R}([A \mid B]).
\end{aligned}$$

Kako su potprostori $\mathcal{R}(A)$ i $\mathcal{R}([A \mid B])$ isti, njihove dimenzije su jednake, a to su upravo $r(A)$ i $r([A \mid B])$.

Obratno, neka vrijedi $r(A) = r([A \mid B)$. Moramo pokazati da tada postoji rješenje jednadžbe $AX = B$. Pođimo od suprotnog, kad ne bi postojalo rješenje X, tada ne bi sve jednadžbe sustava (8.8) imale rješenje. Postojao bi neki i, takav da $Ax = b_i$ nema rješenje. To bi po Kronecker-Capellijevom teoremu značilo da je $r(A) < r([A \mid b_i])$. Kako uvijek vrijedi $r([A \mid b_i]) \leq r([A \mid B])$, pokazali smo da je $r(A) < r([A \mid B])$. Dobivena je kontradikcija s pretpostavkom $r(A) = r([A \mid B])$, pa zaključujemo da je tome bila kriva pretpostavka da ne postoji rješenje jednadžbe $AX = B$. Time je dokazan i drugi smjer prve tvrdnje.

Pretpostavimo sada, da vrijedi uvjet (8.10), tako da postoji rješenje jednadžbe $AX = B$ i pretpostavimo da imamo jedno rješenje $X^{(P)}$. Označimo sa X opće rješenje sustava $AX = B$, i neka je $Y = X - X^{(P)}$. Uvrštavanjem, dobivamo

$$AY = AX - AX^{(P)} = B - B = O,$$

pa je Y rješenje homogene jednadžbe (8.11). Dakle, kad X prolazi skupom svih rješenja sustava $AX = B$, Y prolazi skupom rješenja homogenog sustava $AX = O$.

Obrnuto, ako je Y opće rješenje homogenog sustava $AX = O$, tada $X = X^{(P)} + Y$ zadovoljava

$$AX = AX^{(P)} + AY = B + O = B,$$

pa je rješenje jednadžbe (8.9). Time je dokazana i druga tvrdnja.

Treća tvrdnja je neposredna posljedica druge tvrdnje. Ako je skup rješenja homogene jednadžbe (8.11) samo nul-matrica iz $\mathbf{R}^{n\times p}$, tada jednadžba (8.9) ne može imati više od jednog rješenja, jer bi u protivnom njihova razlika dala netrivijalno rješenje homogene jednadžbe.

S druge strane, ako $AX = B$ ima samo jedno rješenje, tada homogena jednadžba mora imati samo nul-matricu kao rješenje. U protivnom bi imali $O \neq Z \in \mathbf{R}^{n\times p}$, $AZ = O$, pa bi uz pretpostavljeno (jedinstveno) rješenje, nazovimo ga sa $X^{(P)}$, imali i dodatno rješenje $X^{(P)} + Z$ jednadžbe (8.9). ∎

Jer je $X \in \mathbf{R}^{n\times p}$ rješenja jednadžbe $AX = B$ leže u np-dimenzionalnom vektorskom prostoru $\mathbf{R}^{n\times p}$. Može se pokazati da rješenja homogene jednadžbe čine $(n - r(A))p$ dimenzionalni potprostor tog prostora, pa možemo reći da skup svih rješenja jednadžbe $AX = B$ čini linearnu mnogostrukost dimenzije $(n - r(A))p$.

8.2 Lijevi i desni inverz

Promotrimo sada specijalne jednadžbe:

$$AX = I, \quad A \in \mathbf{R}^{m\times n}, \quad I \in \mathbf{R}^{m\times m}, \quad X \in \mathbf{R}^{n\times m} \tag{8.12}$$

i

$$YA = I, \quad A \in \mathbf{R}^{m\times n}, \quad I \in \mathbf{R}^{n\times n}, \quad Y \in \mathbf{R}^{n\times m}. \tag{8.13}$$

U dokazivanju postojanja i jedinstvenosti rješenja tih jednadžbi ključnu ulogu će imati teorem 8.1.

Definicija 8.2 *Svako rješenje jednadžbe $AX = I$ naziva se* **desni inverz** *matrice A. Svako rješenje jednadžbe $YA = I$ naziva se* **lijevi inverz** *matrice A.*

Kadkada se lijevi (desni) inverz naziva lijevi (desni) generalizirani inverz od A. Iz relacija (8.12) i (8.13) se vidi da su obje matrice X i Y istog tipa. Na osnovu teorema 8.1 zaključujemo da A može imati beskonačno lijevih (ili desnih) inverza, da može imati točno jedan lijevi (ili desni) inverz, a također da možda A nema niti jedan lijevi (ili desni) inverz. Zato nam trebaju kriteriji za postojanje lijevih odnosno desnih inverza.

Propozicija 8.3 *Matrica $A \in \mathbf{R}^{m\times n}$ ima barem jedan*

(i) desni inverz ako i samo vrijedi $r(A) = m \leq n$ *;*

(ii) lijevi inverz ako i samo vrijedi $r(A) = n \leq m$ *.*

Dokaz: (i) Ako matrica ima linearno nezavisne retke, pa ju proširimo s dodatnim stupcima (koje umetnemo bilo ispred matrice, iza matrice ili između samih stupaca matrice), rang tako proširene matrice neće biti ni veći niti manji od ranga polazne matrice. Ako to zaključivanje primijenimo na I_m, zaključujemo da uvijek vrijedi

$$m = r(I_m) = r([A \mid I_m])\,. \tag{8.14}$$

Prema teoremu 8.1, $AX = I_m$ ima barem jedno rješenje ako i samo ako vrijedi

$$r(A) = r([A \mid I_m])\,.$$

Zbog relacije (8.14), zadnja jednakost vrijedi ako i samo ako vrijedi $r(A) = m$. Uočimo još da $r(A) = m$ povlači $m \leq n$ (jer bi inače bilo $r(A) \leq n < m$).

(ii) Matrična jednadžba $XA = I_n$ je ekvivalentna s $A^TX^T = I_n$, a ta jednadžba prema (i) ima barem jedno rješenje ako i samo ako vrijedi

$$r(A^T) = \text{ broj redaka od } A^T \ = \text{ broj stupaca od } A = n.$$

Kako je $r(A^T) = r(A)$, dokazali smo prvu tvrdnju od (ii). Druga tvrdnja, $n \leq m$, slijedi istim zaključivanjem kao u dokazu odgovarajuće tvrdnje od (i). ■

Definicija 8.4 *Matrica* $A \in \mathbf{R}^{m\times n}$ *ima*

puni retčani rang *ako vrijedi* $r(A) = m$*;*

puni stupčani rang *ako vrijedi* $r(A) = n$*;*

puni rang *ako vrijedi* $r(A) = \min\{m, n\}$*, tj. ako ima puni retčani ili stupčani rang.*

Propozicija 8.5 *Neka je* $A \in \mathbf{R}^{m\times n}$*. Tada su sljedeće tri tvrdnje ekvivalentne:*

(i) $m = n$ *i* $r(A) = n$ *.*

(ii) *A ima barem jedan lijevi i barem jedan desni inverz.*

(iii) *A ima jedinstveni lijevi inverz* Y*, jedinstveni desni inverz* X *i vrijedi* $X = Y$*.*

Dokaz: Dokazat ćemo krug implikacija: (i) → (ii) → (iii) → (i).

(i) → (ii) Prema propoziciji 8.3, A ima barem jedan lijevi i barem jedan desni inverz.

(ii) → (iii) Ako za X i Y vrijedi $AX = I_m$ i $YA = I_n$ onda prema propoziciji 8.3 vrijedi $m = n = r(A)$. Također vrijedi

$$Y = YI_m = Y(AX) = (YA)X = I_nX = X,$$

a to znači da je skup lijevih inverza jednak skupu desnih inverza. Pokažimo da je skup desnih inverza jednočlan. Kad bi imali dva desna međusobno različita inverza X_1 i X_2, tada bi za $Z = X_1 - X_2$ vrijedilo

$$AZ = A(X_1 - X_2) = AX_1 - AX_2 = I_n - I_n = O\,.$$

Specijalno, za svaki stupac z_i od Z bi vrijedilo $Az_i = 0$. Zbog činjenice da je rang od A jednak broju stupaca od A, korolar 7.7 implicira $z_i = 0$. Dakle je $z_i = 0$ za sve i, pa je $Z = O$, a to se protivi pretpostavci da je $X_1 \neq X_2$. Vidimo da je skup desnih (pa zato i lijevih) inverza najviše jednočlan, a kako po (ii) nije prazan, zaključujemo da je baš jednočlan.

(iii) → (i) Jer A ima i lijevi i desni inverz, po propoziciji 8.3 mora vrijediti i $r(A) = n$ i $r(A) = m$, pa je (i) dokazano. ■

Specijalno, ako je $A \in \mathbf{R}^{n\times n}$ i $r(A) = n$, iz propozicije 8.5 zaključujemo da postoji samo jedan lijevi inverz i samo jedan desni inverz i da su oni jednaki. Prirodno je zato zvati tu matricu jednostavno **inverz** matrice A i označiti ju s A^{-1}. Matrice za koje postoji inverz obično se nazivaju **invertibilne**.

8.3 Regularne matrice

Invertibilne matrice su vrlo važne pa nose još nekoliko naziva.

Definicija 8.6 *Matrica $A \in \mathbf{R}^{n\times n}$ je* **nesingularna** *(***singularna***) ako vrijedi $r(A) = n$ ($r(A) < n$). Nesingularna matrica se još zove* **regularna** *matrica.* ■

Evo jedne jednostavne karakterizacije singularnih i nesingularnih matrica.

Propozicija 8.7 *Matrica $A \in \mathbf{R}^{n\times n}$ je singularna ako i samo ako postoji netrivijalni vektor $x \in \mathbf{R}^n$ takav da je $Ax = 0$. Matrica A je nesingularna ako za svaki $x \neq 0$ vrijedi $Ax \neq 0$.*

Dokaz: Ako je $A \in \mathbf{R}^{n\times n}$ singularna, tada je $r(A) < n$ pa je defekt(A) $= n - r(A) \geq 1$. Dakle je dimenzija nul-potprostora $\mathcal{N}(A)$ barem jedan, pa postoji $x \in \mathcal{N}(A)$, $x \neq 0$.

Ako je $r(A) = n$, tada je $\mathcal{N}(A) = \{0\}$, pa $x \neq 0$ povlači $Ax \neq 0$. ■

Dakle, za regularne matrice je preslikavanje $x \mapsto Ax$ injekcija, a jer je $\dim(\mathcal{R}(A)) = n - \text{defekt}(A) = n$, ono je i bijekcija skupa $\mathbf{R}^n$ na sebe. Regularne matrice imaju puni rang pa im je reducirana forma jedinična matrica. One su i invertibilne, baš kao što su invertibilne matrice regularne.

Skup regularnih matrica ima posebnu važnost u primjenama jer uz operaciju matričnog množenja on postaje grupa. Pokažimo to.

Prvo moramo pokazati da je produkt dviju regularnih matrica opet regularna matrica. Neka su A i B regularne matrice. Tada su i invertibilne pa za njih postoje inverzi, A^{-1} i B^{-1}, respektivno. Pokažimo da je $B^{-1}A^{-1}$ inverz od AB. Imamo,

$$\begin{aligned}(B^{-1}A^{-1})(AB) &= B^{-1}(A^{-1}A)B = B^{-1}I_nB = B^{-1}B = I_n,\\ (AB)(B^{-1}A^{-1}) &= A(BB^{-1})A^{-1} = AI_nA^{-1} = AA^{-1} = I_n\,.\end{aligned}$$

Kako je $B^{-1}A^{-1}$ i lijevi i desni inverz od AB, po propoziciji 8.5 (implikacija (iii) $\to$ (i)) zaključujemo da je AB punog ranga, a to po definiciji 8.6 znači da je AB regularna. Dakle, skup regularnih matrica je zatvoren u odnosu na množenje.

Asocijativnost kao svojstvo grupe odmah slijedi iz asocijativnosti matričnog množenja. Neutralni element je jedinična matrica (ona je sama sebi inverz). Konačno, za svaku regularnu matricu A, inverzni element u grupi je A^{-1}.

Propozicija 8.8 *Skup regularnih matrica čini grupu u odnosu na matrično množenje. Grupa je zatvorena u odnosu na operaciju transponiranja. Ako su A_1, A_2, ...,A_k regularne matrice, tada je*

$$(A_1A_2\cdots A_k)^{-1} = A_k^{-1}\cdots A_2^{-1}A_1^{-1}.$$

Dokaz: Da je grupa regularnih matrica zatvorena u odnosu na transponiranje slijedi iz činjenice da je regularna matrica A karakterizirana uvjetima $r(A) = m = n$, a to onda vrijedi i za A^T jer je $r(A) = r(A^T)$. Zadnja tvrdnja slijedi provjerom da je $A_k^{-1}\cdots A_1^{-1}$ lijevi i desni inverz od $A_1A_2\cdots A_k$. ■

8.4 Elementarne matrice

Svaku od tri elementarne operacije nad recima matrice možemo zapisati pomoću množenja (s lijeva) tzv. elementarnim matricama. Neka je $A \in \mathbf{R}^{m\times n}$ matrica na koju se primijenjuje retčana operacija i $\tilde{A} \in \mathbf{R}^{m\times n}$ matrica A nakon operacije. Istaknuti reci neka su oni s indeksima i i j. Koristeći retčane particije tih matrica, možemo pisati

$$A = \begin{bmatrix} a_1^T \\ \vdots \\ a_i^T \\ \vdots \\ a_j^T \\ \vdots \\ a_m^T \end{bmatrix} \longmapsto \begin{bmatrix} \tilde{a}_1^T \\ \vdots \\ \tilde{a}_i^T \\ \vdots \\ \tilde{a}_j^T \\ \vdots \\ \tilde{a}_m^T \end{bmatrix} = \tilde{A}\,.$$

Promotrimo svaku retčanu operaciju ponaosob.

1. **Zamjena (transpozicija) redaka matrice**. U ovom slučaju je

   $$\tilde{a}_i^T = a_j^T\,,\ \ \tilde{a}_j^T = a_i^T \qquad \text{i} \qquad \tilde{a}_k^T = a_k^T\,,\ \text{za } k \in \{1,\ldots,m\}\backslash\{i,j\}\,.$$

   Isti efekt se postiže množenjem A s lijeva matricom

   $$I_{ij} = \begin{array}{cc} \begin{bmatrix} \ddots & & & & & & & & \\ & 1 & & & & & & & \\ & & 0 & & \cdots & & 1 & & \\ & & & 1 & & & & & \\ & & \vdots & & \ddots & & \vdots & & \\ & & & & & 1 & & & \\ & & 1 & & \cdots & & 0 & & \\ & & & & & & & 1 & \\ & & & & & & & & \ddots \end{bmatrix} & \begin{matrix} \\ \\ i \\ \\ \\ \\ j \\ \\ \\ \end{matrix} \end{array} \in \mathbf{R}^{m\times m}.$$

   Uočimo da je $I_{ij} = [e_1,\ldots,e_{i-1},e_j,e_{i+1},\ldots,e_{j-1},e_i,e_{j+1},\ldots,e_n]$, pa I_{ij} nastaje iz I_n permutacijom (točnije: transpozicijom) stupaca. Stoga je $r(I) = r(I_{ij}) = m$, tj. I_{ij} je invertibilna. Kako množenje s lijeva matricom I_{ij} zamjenjuje i-ti i j-ti redak, mora vrijediti $I_{ij}(I_{ij}I) = I$, što daje $I_{ij}I_{ij} = I$. Dakle je $I_{ij}^{-1} = I_{ij}$. Zbog $A = I_{ij}^{-1}\tilde{A} = I_{ij}\tilde{A}$, imamo

   $$A \quad \xmapsto{I_{ij}} \quad \tilde{A} \quad \xmapsto{I_{ij}} \quad A.$$

Vidimo da se inverzna operacije prve elementarne transformacije, također svodi na množenje s lijeva matricom I_{ij}. U slučaju $i = j$, transformacija zamjenjuje i-ti redak sa samim sobom, pa se definira $I_{ii} = I$. U slučaju $i > j$ definiramo $I_{ij} = I_{ji}$ jer kod zamjene redaka nije bitan uređaj odgovarajućih indeksa po veličini.

Uočimo još dvije stvari. Prvo, I_{ij} je simetrična, pa je I_{ij}^T opet elementarna transformacija. Drugo, množenje matrice $A \in \mathbf{R}^{m\times n}$ s $I_{ij} \in \mathbf{R}^{n\times n}$ s desna zamjenjuje i-ti i j-ti stupac matrice.

2. **Množenje i-tog retka matrice brojem** $\alpha \neq 0$, realizira se množenjem matrice A s lijeva, matricom

$$D_i(\alpha) = \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \alpha & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix} \in \mathbf{R}^{m\times m}$$

pri čemu je α i-ti dijagonalni element. Odmah se vidi da je $D_i(\alpha)$ simetrična i invertibilna, pri čemu je $D_i(\alpha)^{-1} = D_i(1/\alpha)$. Dakle je $D_i(\alpha)^{-1}$ opet elementarna matrica istog tipa. Jer je $A = D_i(\alpha)^{-1}\tilde{A}$, vrijedi

$$A \quad \overset{D_i(\alpha)}{\longmapsto} \quad \tilde{A} \quad \overset{D_i(1/\alpha)}{\longmapsto} \quad A\,.$$

Ovo odgovara promjeni i-tog retka: $a_i^T \longmapsto \tilde{a}_i^T = \alpha a_i^T \longmapsto a_i^T$. Uočimo da se množenje matrice A matricom $D_i(\alpha) \in \mathbf{R}^{n\times n}$ s desna, svodi na množenje i-tog stupca matrice A s α.

3. **Dodavanje j-tom retku i-tog retka pomnoženog brojem** α. U ovom slučaju je

$$\tilde{a}_i^T = a_i^T\,,\ \tilde{a}_j^T = a_j^T + \alpha a_i^T \ \text{ i } \tilde{a}_k^T = a_k^T\,,\ \text{za } k \in \{1,\ldots,m\} \setminus \{i,j\}\,,$$

pa je u slučaju $i < j$, odgovarajuća matrica

$$M_{ij}(\alpha) = \left[\begin{array}{ccccccc} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & \vdots & \ddots & & & \\ & & \alpha & \cdots & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{array}\right] \begin{array}{c} \\ \\ i \\ \\ j \\ \\ \\ \end{array} \in \mathbf{R}^{m\times m}$$

U slučaju $\alpha = 0$ imamo $M_{ij}(0) = I$. Matričnim množenjem, lako se provjeri jednakost

$$M_{ij}(\alpha)M_{ij}(\beta) = M_{ij}(\alpha+\beta), \tag{8.15}$$

pa je u slučaju $\beta = -\alpha$,

$$M_{ij}(\alpha)M_{ij}(-\alpha) = M_{ij}(-\alpha)M_{ij}(\alpha) = I.$$

Dakle je $M_{ij}(-\alpha)$ lijevi i desni inverz od $M_{ij}(\alpha)$, pa je po propoziciji 8.5

$$M_{ij}(\alpha)^{-1} = M_{ij}(-\alpha).$$

Množenje matrice $\tilde{A}$ matricom $M_{ij}(-\alpha)$ s lijeva, dodaje j-tom retku od $\tilde{A}$, i-ti redak pomnožen s $-\alpha$. Stoga vrijedi

$$A \quad \overset{M_{ij}(\alpha)}{\longmapsto} \quad \tilde{A} \quad \overset{M_{ij}(-\alpha)}{\longmapsto} \quad A\,,$$

što odgovara promjeni j-tog retka: $a_j^T \longmapsto \tilde{a}_j^T = a_j^T + \alpha a_i^T \longmapsto \tilde{a}_j^T - \alpha a_i^T = a_j^T$. Vidimo i u ovom slučaju da inverznoj retčanoj operaciji odgovara elementarna transformacija istog tipa.

Množenje matrice A s $M_{ij}(\alpha)$ s desna, dodaje j-tom stupcu od A i-ti stupac pomnožen s α.

Promotrimo još operaciju dodavanja j-tog retka i-tom retku. U ovom slučaju mijenja se samo i-ti redak, pa toj transformaciji odgovara množenje matrice A s lijeva matricom

$$M_{ij}(\alpha)^T = \begin{array}{@{}c@{\quad}c@{}} \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & \cdots & \alpha & & \\ & & & \ddots & \vdots & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix} & \begin{matrix} \\ \\ i \\ \\ j \\ \\ \\ \end{matrix} \end{array} \in \mathbf{R}^{m\times m}$$

Transponiranjem jednakosti (8.15), vidimo da za $M_{ij}(\alpha)^T$ vrijedi $M_{ij}(\beta)^T M_{ij}(\alpha)^T = M_{ij}(\alpha+\beta)^T$, pa je $M_{ij}(\alpha)^{-T} = M_{ij}(-\alpha)^T$. Slično kao i prije, imamo

$$A \quad \overset{M_{ij}(\alpha)^T}{\longmapsto} \quad \tilde{A} \quad \overset{M_{ij}(-\alpha)^T}{\longmapsto} \quad A\,,$$

što odgovara promjeni i-tog retka: $a_i^T \longmapsto \tilde{a}_i^T = a_i^T + \alpha a_j^T \longmapsto \tilde{a}_i^T - \alpha a_j^T = a_i^T$. Vidimo i u ovom slučaju da inverznoj retčanoj operaciji odgovara elementarna transformacija istog tipa.

Množenje matrice A s $M_{ij}(\alpha)^T$ s desna, dodaje i-tom stupcu od A j-ti stupac pomnožen s α.

Dakle, obje elementarne matrice $M_{ij}(\alpha)$ i $M_{ij}(\alpha)^T$ odgovaraju trećoj elementarnoj retčanoj transformaciji. To znači da je i u ovom slučaju transponirana elementarna matrica opet elementarna matrica istog tipa.

Matricu $M_{ij}(\alpha)^T$ ćemo također označavati s $M_{ij}^T(\alpha)$ odnosno sa $M_{ji}(\alpha)$, jer istovremena zamjena redaka i stupaca od $M_{ij}(\alpha)$ dovodi do $M_{ij}^T(\alpha)$.

Kao zaključak prikaza elementarnih matrica imamo

Lema 8.9 *Elementarne matrice su regularne. Inverzna i transponirana matrica elementarne matrice su elementarne matrice istog tipa. Specijalno vrijedi*

(i) $\quad I_{ij}^T = I_{ij}^{-1} = I_{ij}$,

(ii) $D_i(\alpha)^T = D_i(\alpha), \quad D_i(\alpha)^{-1} = D_i(\alpha^{-1}), \ \ \alpha \neq 0,$

(iii) $M_{ij}^T(\alpha) = M_{ji}(\alpha), \quad M_{ij}^{-1}(\alpha) = M_{ij}(-\alpha).$ ■

Primijenimo na proizvoljnu matricu $A \in \mathbf{R}^{m\times n}$ algoritam za svađanje na reducirani oblik A_R. Algoritam primjenjuje konačni niz elementarnih retčanih operacija na A. Kako svakoj elementarnoj retčanoj operaciji pripada množenje s lijeva nekom elementarnom matricom, algoritam možemo opisati pomoću elementarnih matrica na sljedeći način

$$\begin{aligned} A_R &= E_k(E_{k-1}(\cdots(E_2(E_1A))\cdots)) \\ &= E_kE_{k-1}\cdots E_2E_1A. \end{aligned} \tag{8.16}$$

Pritom je svaka E_r elementarna matrica jednog od tri tipa: I_{ij}, $D_i(\alpha)$ ili $M_{ij}(\alpha)$. Kako su sve matrice E_r, $1 \leq r \leq k$ regularne, takav je i njihov produkt (propozicija 8.8). Time smo dokazali

Lema 8.10 *Za svaku matricu $A \in \mathbf{R}^{m\times n}$ postoji regularna matrica $S \in \mathbf{R}^{m\times m}$, takva da je*

$$A_R = SA.$$

■

Pri opisu elementarnih matrica, vidjeli smo da svako množenje elementarnom matricom s desna čini jednu elementarnu operaciju nad stupcima matrice. Stoga se možemo zapitati na koji jednostavni oblik se može $A \in \mathbf{R}^{m\times n}$ svesti uz pomoć elementarnih operacija nad stupcima.

Pođimo od proizvoljne $A \in \mathbf{R}^{m\times n}$ i označimo sa $B = A^T \in \mathbf{R}^{n\times m}$. Svedimo B elementarnim retčanim operacijama ra reducirani oblik,

$$B_R = E_k\cdots E_2E_1B.$$

Transponiranjem dobivamo

$$B_R^T = B^TE_1^TE_2^T\cdots E_k^T.$$

Označimo li $A_R^c = B_R$, $F_r = E_r^T$, $1 \leq r \leq k$, imamo

$$A_R^c = AF_1\cdots F_k,$$

gdje su $A, A_R^c \in \mathbf{R}^{m\times n}$, a $F_1, \ldots, F_k$ su elementarne matrice. Pritom je A_R^c reducirani oblik na koji se A može svesti elementarnim operacijama nad stupcima. On je "transponirani reducirani oblik".

U slučaju regularne A, matrice A_R i A_R^c su jednake I. To je posljedica sljedećeg teorema.

Teorem 8.11 *Kvadratna matrica A je regularna ako i samo ako je produkt elementarnih matrica. Matrica A^T je regularna ako i samo ako je takva A.*

Dokaz: Matrica A je regularna ako i samo ako je kvadratna i punog ranga, a to je ako i samo ako ima reducirani oblik $A_R = I$. Prema relaciji (8.16) vrijedi $A_R = E_k \cdots E_1 A$ za neke elementarne matrice E_r, $1 \le r \le k$. Uvrštavanjem I na mjesto od A_R, te množenjem s lijeva matricama: E_k^{-1}, E_{k-1}^{-1}, …,E_1^{-1} u tom redoslijedu, dobijemo

$$A = E_1^{-1} E_2^{-1} \cdots E_k^{-1}.$$

Dakle A je regularna ako i samo ako ima takav prikaz. Kako je inverz svake elementarne matrice E_r^{-1} opet elementarna matrica (istog tipa) prva tvrdnja je dokazana.

Druga tvrnja je već dokazana u propoziciji 8.8, a ovdje dajemo drugi dokaz koji slijedi iz prve tvrdnje teorema. Neka je A regularna. Tada je ona produkt $F_1 \cdots F_k$ elementarnih matrica. Transponiranjem dobivamo $A^T = F_k^T \cdots F_1^T$. Kako je po lemi 8.9 svaka F_r^T također elementarna matrica, ona je i regularna. Dakle je A^T kao produkt regularnih matrica regularna. Obratno, ako je A^T regularna, onda po upravo dokazanom mora i $(A^T)^T = A$ biti regularna, čime je i drugi smjer dokazan. ■

Lekcija 9

Osnovne klase matrica

Od ovog poglavlja nadalje bavit ćemo se uglavnom samo kvadratnim matricama, dakle matricama iz skupa $\mathbf{R}^{n\times n}$, jer se one u primjeni najčešće pojavljuju. Od sada pa nadalje, ako nije izričito drugačije naznačeno, podrazumjevamo da se radi o matrici iz skupa $\mathbf{R}^{n\times n}$.

Definicija 9.1 *Matrica A je* **simetrična** *ako vrijedi $A^T = A$. Matrica A je* **antisimetrična** *ako vrijedi $A^T = -A$. Za proizvoljnu kvadratnu matricu X, matrica*

$$X_s = \frac{1}{2}(X + X^T)$$

se naziva **simetrični dio** *matrice X, a matrica*

$$X_a = \frac{1}{2}(X - X^T)$$

antisimetrični dio *matrice X.* ■

Vrlo jednostavno je pokazati, da je za svaku matricu X, X_s uvijek simetrična, a X_a je uvijek antisimetrična matrica i vrijedi

$$X = X_s + X_a.$$

Kod simetričnih (antisimetričnih) matrica je antisimetrični (simetrični) dio jednak nul-matrici.

Propozicija 9.2 *Ako vrijedi*

$$X = U + V, \quad U^T = U, \quad V^T = -V,$$

onda je

$$U = \frac{1}{2}(X + X^T), \quad V = \frac{1}{2}(X - X^T),$$

tj. $U = X_s$, $V = X_a$, *pa svaka matrica ima jedinstven rastav na simetrični i antisimetrični dio.*

Dokaz: Iz definicije matrica X_s i X_a znamo da je $X = X_s + X_a$, pri čemu je

$$X_s = \frac{1}{2}(X + X^T), \quad X_a = \frac{1}{2}(X - X^T).$$

Zbog pretpostavke propozicije, imamo

$$0 = X - X = (X_s + X_a) - (U + V) = (X_s - U) + (X_a - V)\,. \tag{9.1}$$

Budući je nul-matrica simetrična, transponiranjem prethodne jednakosti dobivamo

$$\begin{aligned} 0 &= 0^T = (X_s - U)^T + (X_a - V)^T = (X_s^T - U^T) + (X_a^T - V^T) \\ &= (X_s - U) + (-X_a - (-V)) = (X_s - U) - (X_a - V)\,. \end{aligned} \tag{9.2}$$

Zbrajanjem jednakosti (9.1) i (9.2) dobivamo $X_s - U = 0$, tj. $X_s = U$. Oduzimanjem istih jednakosti dobivamo $X_a - V = 0$, tj. $X_a = V$. ■

Sada ćemo uvesti jednu jednostavnu matričnu funkciju koja na vektorskom prostoru kvadratnih matrica ima korisna svojstva. Zovemo ju **trag** matrice. Ona matrici pridružuje broj, trag : $\mathbf{R}^{n\times n} \to \mathbf{R}$, pa se kaže da je funkcional na skupu matrica.

Definicija 9.3 *Trag matrice A je suma njezinih dijagonalnih elemenata:*

$$\mathrm{trag}\,(A) = \sum_{i=1}^{n} a_{ii}\,.$$

■

Propozicija 9.4 *Funkcija* trag *ima sljedeća svojstva:*

$$\begin{aligned}
\operatorname{trag}(A+B) &= \operatorname{trag}(A)+\operatorname{trag}(B),\\
\operatorname{trag}(cA) &= c\operatorname{trag}(A),\\
\operatorname{trag}(A^T) &= \operatorname{trag}(A),\\
\operatorname{trag}(AB) &= \operatorname{trag}(BA),\\
\operatorname{trag}(ABC) &= \operatorname{trag}(BCA)=\operatorname{trag}(CAB).
\end{aligned}$$

Dokaz: Ova propozicija se dokazuje jednostavnim uvrštavanjem elemenata matrice u pojedine jednadžbe, pa dokaz ostavljamao čitatelju. ■

Prva dva svojstva pokazuju da je trag linearna funkcija (linearni funkcional). Neosjetljivost traga na operaciju transponiranja je jasna jer se dijagonala matrice ne mijenja nakon transponiranja matrice. Invarijantnost traga u odnosu na komutaciju u produktu matrica je važno svojstvo koje trag čini privlačnim kako u teoretskim razmatranjima tako i u praktičnim numeričkim primjenama.

Već smo spomenuli da je trag matrice linearni funkcional na svakom vektorskom prostoru matrica $\mathbf{R}^{n\times n}$. Na vektorskom prostoru $\mathbf{R}^n$ možemo definirati kvadratični funkcional $q_A : \mathbf{R}^n \to \mathbf{R}$.

Definicija 9.5 *Ako je $A \in \mathbf{R}^{n\times n}$ simetrična matrica, onda funkciju q_A,*

$$q_A(x) = x^T A x\,, \quad x \in \mathbf{R}^n\,,$$

nazivamo **kvadratna forma**. ■

Iako je $q_A(x)$ dobro definiran za proizvoljnu kvadratnu matricu A, on ima lijepa svojstva ako je A simetrična matrica.

Definicija 9.6 *Za simetričnu matricu A kažemo da je* **pozitivno semidefinitna**, *ako vrijedi*

$$x^T A x \geq 0 \quad \text{za svaki } x \in \mathbf{R}^n\,.$$

Ako vrijedi

$$x^T A x > 0 \quad \text{za svaki } x \in \mathbf{R}^n\,,\ x \neq 0\,,$$

onda kažemo da je A **pozitivno definitna** *matrica.* ■

Dakle za simetričnu pozitivno semidefinitnu (pozitivno definitnu) matricu A vrijedi $q_A(x) \geq 0$ za svako x ($q_A(x) > 0$ za svako $x \neq 0$).

Propozicija 9.7 *Svaka pozitivno definitna simetrična matrica je punog ranga, tj. regularna matrica.*

Dokaz: Neka je A pozitivno definitna matrica. Kad bi bilo $r(A) < n$, onda bi jednadžba $Ax = 0$ imala barem jedno rješenje $x_0 \neq 0$, pa bi bilo

$$x_0^T A x_0 = x_0^T 0 = 0\,.$$

To je u suprotnosti sa definicijom pozitivno definitne matrice, pa zaključujemo da rang matrice ne može biti manji od n. ■

Propozicija 9.8 *Neka je $B \in \mathbf{R}^{m\times n}$.*

(i) Matrice $B^T B \in \mathbf{R}^{n\times n}$ i $BB^T \in \mathbf{R}^{m\times m}$ su pozitivno semidefinitne.

(ii) Ako je B punog stupčanog ranga, onda je $B^T B \in \mathbf{R}^{n\times n}$ pozitivno definitna matrica.

(iii) Ako je B punog retčanog ranga, onda je $BB^T \in \mathbf{R}^{m\times m}$ pozitivno definitna matrica.

Dokaz: (i) Iz svojstava transponiranja matrica dobivamo

$$\begin{aligned}(B^T B)^T &= B^T (B^T)^T = B^T B,\\ (BB^T)^T &= (B^T)^T B^T = BB^T\end{aligned}$$

pa su $B^T B$ i BB^T simetrične matrice. Također vrijedi

$$\begin{aligned}x^T B^T B x = (Bx)^T (Bx) = \|Bx\|^2 \geq 0 \quad \text{za svaki } x \in \mathbf{R}^n\,,\\ x^T BB^T x = (B^T x)^T (B^T x) = \|B^T x\|^2 \geq 0 \quad \text{za svaki } x \in \mathbf{R}^m\,.\end{aligned}$$

(ii) Ako je matrica B punog stupčanog ranga, onda jednadžba $Bx = 0$ ima samo trivijalno rješenje $x = 0$. Dakle, $x \neq 0$ povlači $Bx \neq 0$, pa je

$$x^T B^T B x = \|Bx\|^2 > 0 \quad \text{za svaki } x \in \mathbf{R}^n\,,\ x \neq 0\,.$$

(iii) Ako je B punog retčanog ranga, onda je B^T punog stupčanog ranga, pa jednadžba $B^T x = 0$ ima samo trivijalno rješenje $x = 0$. Dakle, $x \neq 0$ povlači $B^T x \neq 0$, pa je

$$x^T BB^T x = \|B^T x\|^2 > 0 \quad \text{za svaki } x \in \mathbf{R}^m\,,\ x \neq 0\,.$$

■

Primjedba 9.9 *Može se pokazati, da za svaku pozitivno semidefinitnu matricu A postoji kvadratna matrica B, takva da je $A = B^T B$. Pritom je B regularna ako i samo ako je takva A. Matrica B općenito nije jedinstvena, ali uz neke uvjete jest. Matrica B se može odabrati gornjom ili donjom trokutastom i u tom slučaju je B jedinstveno određena uvjetom regularnosti matrice A i pozitivnošću dijagonalnih elemenata od B.*

Ortogonalne matrice

Promotrimo sada matrice Q koje istovremeno zadovoljavaju jednadžbe

$$Q^T Q = I \tag{9.3}$$

$$Q Q^T = I\,. \tag{9.4}$$

Iz jednadžbe (9.3) slijedi da je Q^T lijevi inverz od Q, a iz jednadžbe (9.4) slijedi da je Q^T desni inverz od Q. Pomoću propozicije 8.5 zaključujemo da je Q nužno kvadratna i regularna. Pritom je lijevi inverz isto što i desni inverz i to je baš inverz matrice. Dakle je Q^T inverz od Q.

Matrice koje zadovoljavaju (9.3) imaju vrlo važna svojstva, pa nose i poseban naziv.

Definicija 9.10 *Kvadratna matrica Q koja ima svojstvo $Q^T Q = I$ zove se* **ortogonalna matrica**. ■

Ako je Q ortogonalna prema definiciji 9.10, tada zbog propozicije 8.3(ii) zaključujemo da $r(Q) = n$. Kako je $m = n = r(Q)$, Q je regularna, pa je lijevi inverz Q^T ujedno inverz, $Q^T = Q^{-1}$. Stoga također vrijedi i relacija (9.4). Time je pokazano da Q iz definicije 9.10 zadovoljava oba uvjeta (9.3) i (9.4).

Koristeći iste argumente, u zadnjoj definiciji smo umjesto uvjeta (9.3) mogli koristiti uvjet (9.4).

Propozicija 9.11 *Za ortogoinalne matrice Q vrijedi*

$$\|Qx\| = \|x\|,$$

tj. **ortogonalne matrice čuvaju duljinu vektora**. *Ortogonalne matrice su jedine kvadratne matrice koje posjeduju to svojstvo.*

Dokaz: Neka je Q ortogonalna matrica. Tada je

$$\|Qx\|^2 = (Qx)^T(Qx) = x^T Q^T Q x = x^T I x = x^T x = \|x\|^2.$$

Time je dokazano da ortogonalne matrice posjeduju dano svojstvo. Neka je sada $Q \in \mathbf{R}^{n\times n}$ matrica koja ima svojstvo

$$\|Qx\| = \|x\|.$$

Tada za svaki $x \in \mathbf{R}^n$ vrijedi

$$0 = x^T x - x^T x = x^T Q^T Q x - x^T I x = x^T (Q^T Q - I) x.$$

Matrica $Q^T Q - I$ je simetrična, a iz zadnje jednadžbe slijedi da je i pozitivno semidefinitna, pa postoji matrica B, takva da je $B^T B = Q^T Q - I$ (primjedba 9.9). Sada je

$$\|Bx\|^2 = x^T B^T B x = x^T (Q^T Q - I) x = 0,$$

pa je $Bx = 0$ za svaki $x \in \mathbf{R}^n$. Iz toga slijedi $B = 0$, pa je i $Q^T Q - I = 0$, tj. $Q^T Q = I$. ∎

Prisjetimo se definicije kuta između dvaju vektora iz trećeg poglavlja (vidi 3.14):

$$\cos\theta = \frac{(x \mid y)}{\|x\|\,\|y\|}, \quad x \neq 0\,, y \neq 0\,,\ 0 \leq \theta \leq \pi.$$

Propozicija 9.12 *Ako je Q ortogonalna matrica, onda je kut između vektora Qx i Qy jednak kutu između vektora x i y, tj.* **ortogonalne matrice čuvaju kut među vektorima**.

Dokaz: Imamo

$$\frac{(Qx \mid Qy)}{\|Qx\|\,\|Qy\|} = \frac{x^T Q^T Q y}{\|x\|\,\|y\|} = \frac{x^T y}{\|x\|\,\|y\|} = \frac{(x \mid y)}{\|x\|\,\|y\|}\,.$$

Budući da se kut između dvaju vektora uvijek nalazi u intervalu $[0, \pi]$, tvrdnja je dokazana. ∎

Obrat propozicije općenito ne vrijedi.

Posebno su značajne ortogonalne matrice tipa $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ čiji stupci su zarotirani kanonski vektori e_1 i e_2 u ravnini određenoj vektorima e_1 i e_2 oko ishodišta, za kut θ u pozitivnom smjeru. Zato se matrice tog oblika zovu matrice rotacije u $\mathbf{R}^2$ za kut θ.

Slično,

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}, \quad \begin{bmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{bmatrix}, \quad \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

su ortogonalne matrice ravninskih rotacije u $\mathbf{R}^3$ koje rotiraju vektore u ravninama određenim kanonskim vektorima e_2, e_3, odnosno e_1, e_3, odnosno e_1, e_2. Ortogonalne matrice imaju još jedno važno svojstvo. Njihovi stupci i retci čine ortonormirane baze prostora $\mathbf{R}^n$ i $\mathbf{R}^{1\times n}$.

Propozicija 9.13 *Neka su*

$$Q = \begin{bmatrix} q_1 & q_2 & \ldots & q_n \end{bmatrix} \ i \ Q = \begin{bmatrix} \tilde{q}_1^T \\ \tilde{q}_2^T \\ \vdots \\ \tilde{q}_n^T \end{bmatrix}$$

stupčana i retčana particija kvadratne matrice Q. Tada je Q ortogonalna ako i samo vrijedi

$$(q_i \mid q_j) = q_i^T q_j = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \quad i \quad (\tilde{q}_i \mid \tilde{q}_j) = \tilde{q}_i^T \tilde{q}_j = \begin{cases} 1 & i = j, \\ 0 & i \neq j, \end{cases}$$

tj. ako i samo ako ima ortonormirane retke i stupce.

Dokaz: Uočimo da je $Q^TQ = I$ ekvivalentno s $(Q^TQ)_{ij} = \delta_{ij}$ gdje je δ_{ij} Kroneckerova delta. Jer je

$$(Q^TQ)_{ij} = e_i^T Q^T Q e_j = (Qe_i)^T (Qe_j) = q_i^T q_j\,, \quad i, j \in \{1, \ldots, n\},$$

$Q^TQ = I$ je ekvivalentno s $q_i^T q_j = \delta_{ij}$. Drugim riječima $Q^TQ = I$ je ekvivalentno s tvrdnjom da Q ima ortonormirane stupce.

Na slični način se pokaže da je $QQ^T = I$ ekvivalentno s $\tilde{q}_i^T \tilde{q}_j = \delta_{ij}$ odnosno s tvrdnjom da Q ima ortonormirane retke.

Zaključujemo da Q zadovoljava obje relacije $Q^TQ = I$ i $QQ^T = I$ ako i samo ako ima ortonormirane stupce i retke. ∎

Sljedeća klasa ortogonalnih matrica je u bliskoj vezi sa permutacijama skupa $S_n = \{1, 2, \ldots, n\}$.

Definicija 9.14 *Matrica* **permutacije** *je matrica čiji elementi su nule ili jedinice i koja u svakom retku i u svakom stupcu ima točno jednu jedinicu.* ■

Iz definicije odmah slijedi da je svaka matrica permutacije kvadratna. Skup svih matrica permutacija reda n označit ćemo s $\mathbf{P}_n$. Kako svaka matrica permutacije ima ortonormirane stupce i retke, $\mathbf{P}_n$ je podskup skupa ortogonalnih matrica. Specijalno, matrice iz $\mathbf{P}_n$ su regularne.

Svakoj permutaciji $p \in \Pi_n$ možemo pridružiti matricu permutacije P formulom

$$Pe_i = e_{p(i)}, \quad 1 \leq i \leq n. \tag{9.5}$$

Relacijom (9.5) definirano je preslikavanje $\mathcal{J} : \Pi_n \mapsto \mathbf{P}_n$, takvo da je $\mathcal{J}(p) = P$.

Primjer 9.15 *Neka je $p \in \Pi_5$ kao u relaciji (2.14), tako da je*

$$p = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 3 & 1 & 2 & 4 \end{pmatrix}, \; p^{-1} = \begin{pmatrix} 5 & 3 & 1 & 2 & 4 \\ 1 & 2 & 3 & 4 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 4 & 2 & 5 & 1 \end{pmatrix}.$$

Tada je

$$\mathcal{J}(p) = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \mathcal{J}(p^{-1}) = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

■

Ako su $p, q \in \Pi_n$ i $P = \mathcal{J}(p)$, $Q = \mathcal{J}(q)$, tada $p \neq q$ povlači $P \neq Q$. Doista, $p \neq q$ znači postojanje bar jednog $i \in S_n$ za koji je $p(i) \neq q(i)$, pa je po relaciji (9.5) $Pe_i \neq Qe_i$, a to znači da se P i Q razlikuju barem u i-tom stupcu. Time je pokazano da je preslikavanje $\mathcal{J}$ injekcija (ili 1 – 1 preslikavanje). Pokažimo da je $\mathcal{J}$ također surjekcija, tj. da za svaku matricu permutacije P postoji $p \in \Pi_n$ tako da vrijedi (9.5). Neka je $P \in \mathbf{P}_n$. Po definiciji 9.14, Pe_i je za svako i koordinatni vektor, pa postoji funkcija $\tilde{p}: \; S_n \to S_n$, takva da je $Pe_i = e_{\tilde{p}(i)}$ za sve $1 \leq i \leq n$. Pokažimo da je $\tilde{p} \in \Pi_n$. Funkcija koja preslikava konačni skup u sebe je bijekcija ako i samo ako je injekcija. Kad $\tilde{p}$ ne bi bila bijekcija, morali bi postojati $i, j \in S_n$, $i \neq j$ za koje bi bilo $\tilde{p}(i) = \tilde{p}(j)$. Označimo $k = \tilde{p}(i)$. Sada bi zbog $Pe_i = Pe_j = e_k$ u k-tom retku matrice P imali barem dvije jedinice (zasigurno u stupcima i i j). Dakle, P ne bi bila matrica permutacije. Vidimo da pretpostavka da $\tilde{p}$ nije bijekcija vodi u kontradikciju. Stoga $\tilde{p}$ mora biti bijekcija pa je u Π_n. ■

Teorem 9.16 *Skup $\mathbf{P}_n$ zajedno s operacijom množenja matrica čini grupu. Preslikavanje $\mathcal{J}$ je izomorfizam grupa $(\Pi_n, \circ)$ i $(\mathbf{P}_n, \cdot)$.*

Dokaz: Već smo pokazali da je $\mathcal{J}$ bijekcija. Pokažimo da zadovoljava i uvjet

$$\mathcal{J}(q \circ p) = \mathcal{J}(q)\mathcal{J}(p). \tag{9.6}$$

Neka je $P = \mathcal{J}(p)$, $Q = \mathcal{J}(q)$ i $S = \mathcal{J}(q \circ p)$. Gledajući i-ti stupac od S i od QP,

$$\begin{aligned} Se_i &= e_{(q\circ p)(i)} = e_{q(p(i))}, \\ QPe_i &= Qe_{p(i)} = e_{q(p(i))}, \end{aligned}$$

zaključujemo da su oni jednaki. Kako to vrijedi za svako $1 \leq i \leq n$, mora biti $S = QP$, pa je (9.6) dokazano. Iz relacije (9.6) odmah slijedi da je $(\mathbf{P}_n, \cdot)$ multiplikativna grupa. Doista, kako je $\mathcal{J}$ surjekcija, relacija (9.6) pokazuje da je skup $\mathbf{P}_n$ zatvoren u odnosu na matrično množenje. Asocijativnost je posljedica asocijativnosti matričnog množenja, a neutralni element je jedinična matrica koja je u $\mathbf{P}_n$. Kako je matrica permutacije regularna, za svako $P \in \mathbf{P}_n$ postoji inverzna matrica P^{-1}, pa još treba pokazati da je $P^{-1} \in \mathbf{P}_n$. Za P postoji $p \in \Pi_n$, tako da je $P = \mathcal{J}(p)$. Zbog (9.6) i zbog činjenice da je $(\Pi_n, \circ)$ grupa, imamo

$$\mathcal{J}(p)\mathcal{J}(p^{-1}) = \mathcal{J}(p \circ p^{-1}) = \mathcal{J}(e) = I_n$$

pa je

$$\mathcal{J}(p^{-1}) = \mathcal{J}(p)^{-1} = P^{-1},$$

odakle zaključujemo da je $P^{-1} \in \mathbf{P}_n$. Time smo pokazali da je $(\mathbf{P}_n, \cdot)$ grupa pa je $\mathcal{J}$ izomorfizam grupa. ■

Koju ulogu u matričnom računu imaju matrice permutacije? Neka za matrice $P, Q \in \mathbf{P}_n$ vrijedi $P = \mathcal{J}(p)$ i $Q = \mathcal{J}(q)$. Promotrimo elemente matrice $A' = P^T AQ$. Za element a'_{ij} matrice A' vrijedi

$$\begin{aligned} a'_{ij} &= e_i^T A' e_j = e_i^T (P^T AQ) e_j = (e_i^T P^T) A (Qe_j) \\ &= (Pe_i)^T A (Qe_j) = e_{p(i)}^T A e_{q(j)} = a_{p(i)q(j)}. \end{aligned} \tag{9.7}$$

Specijalno, ako je $Q = I$, i-ti redak od $P^T A$ je $p(i)$-ti redak od A. Dakle, retci od $P^T A$ su (permutacijom p) ispermutirani retci od A. Ako je $P = I$, j-ti stupac od AQ je $q(j)$-ti stupac od A. Dakle, stupci od AQ su (permutacijom q) ispermutirani stupci od A. Ako je $P = Q$, tada vrijedi $a'_{ij} = a_{p(i)p(j)}$, pa je dijagonala od A', gledana kao niz brojeva, ispermutirana (permutacijom p) dijagonala od A.

Napomenimo još da simetrične, antisimetrične i ortogonalne matrice pripadaju široj klasi tzv. **normalnih** matrica koje su karakterizirane uvjetom $AA^T = A^T A$.

Lekcija 10

Determinante

Determinanta kvadratne matrice je važan pojam u linearnoj algebri. Determinanta ima više važnih svojstava i njih ćemo dokazati u ovoj lekciji. Posebno zanimljivo svojstvo je da se komponente rješenja sustava linearnih jednadžbi s regularnom matricom mogu elegantno izraziti pomoću kvocijenata određenih determinanti. Upravo zato su determinante kronološki prethodile matricama. Danas je njihova upotreba isključivo vezana uz matričnu teoriju. Naime, ne postoje jednostavni i relativno točni algoritmi za numeričko računanje determinante.

U definiciji determinante se pojavljuje broj inverzija u permutaciji skupa S_n. Kako će se pojam inverzije permutacije pojavljivati i u dokazima osnovnih svojstava determinante, prvo ćemo se upoznati s njim.

Definicija 10.1 **Inverzija** *u permutaciji p je svaki par $(p(i), p(j))$ za koji vrijedi $p(i) > p(j)$ kad je $i < j$. $I(p)$ je ukupni* **broj inverzija** *u permutaciji p. Permutacija p je* **parna** *(***neparna***) ako je $I(p)$ paran (neparan) broj.* **Parnost** *je funkcija $\pi : \Pi_n \longrightarrow \{-1, 1\}$ definirana s $\pi(p) = (-1)^{I(p)}$.* ■

Parnost se može definirati i pomoću algebarskog izraza.

Propozicija 10.2 *Za funkciju parnosti vrijedi*

$$\pi(p) = \prod_{1 \le i < j \le n} \frac{p(j) - p(i)}{j - i}, \quad n \geq 2. \tag{10.1}$$

Dokaz: Primijetimo da za svaki par brojeva (i,j), $1 \leq i < j \leq n$, izraz $j-i$ dolazi točno jednom u nazivniku izraza (10.1), a u brojniku dolazi točno jedan od izraza $j-i$ ili $i-j$, pa je rezultat zaista u skupu $\{-1,1\}$. Nazivnik je kao produkt pozitivnih brojeva uvijek pozitivan (jednak je $(n-1)!\cdot(n-2)!\cdots 2!$), dok u brojniku ima onoliko negativnih cijelih brojeva koliki je broj inverzija u permutaciji p. ■

Primjer 10.3 *Neka je p kao u relaciji (2.14),*

$$p = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 3 & 1 & 2 & 4 \end{pmatrix}.$$

Tada je $I(p) = 6$, jer postoji šest inverzija: $(5,3)$, $(5,1)$, $(5,2)$, $(5,4)$, $(3,1)$ i $(3,2)$. Dakle je p parna permutacija i $\pi(p) = (-1)^6 = 1$. Do istog rezultata se dolazi korištenjem formule iz propozicije 10.2. Ako npr. faktore u produktu uredimo tako da prvo j ima vrijednost 2, zatim 3, pa 4 i na kraju 5, dok i uvijek raste od 1 do $j-1$, dobijemo

$$\begin{aligned} \pi(p) &= \frac{3-5}{2-1}\cdot\frac{(1-5)(1-3)}{(3-1)(3-2)}\cdot\frac{(2-5)(2-3)(2-1)}{(4-1)(4-2)(4-3)} \\ &\quad \cdot\frac{(4-5)(4-3)(4-1)(4-2)}{(5-1)(5-2)(5-3)(5-4)} \\ &= \frac{-2}{1}\cdot\frac{(-4)\cdot(-2)}{2\cdot 1}\cdot\frac{(-3)\cdot(-1)\cdot 1}{3\cdot 2\cdot 1}\cdot\frac{(-1)\cdot 1\cdot 3\cdot 2}{4\cdot 3\cdot 2\cdot 1} \\ &= (-1)^6 = 1\,. \end{aligned}$$

■

Sjetimo se specijalnih permutacija p_{ij} (vidi relaciju (2.17)) koje zovemo transpozicije ili zamjene. Za njih vrijedi $p_{ij} \circ p_{ij} = e$, pa zato vrijedi relacija (2.18) tj. $p_{ij}^{-1} = p_{ij}$ za sve $i < j$. Postavlja se pitanje kako se funkcija parnost ponaša u odnosu na operacije uzimanja inverzne permutacije i na množenje permutacijom transpozicije.

Propozicija 10.4 *Neka je $p \in \Pi_n$. Tada vrijedi*

(i) $\pi(p^{-1}) = \pi(p)$

(ii) $\pi(p \circ p_{ij}) = -\pi(p)$

(iii) $\pi(p_{ij} \circ p) = -\pi(p)$

Dokaz: (i) Koristeći relaciju (2.15) i vodeći računa da vrijednost produkta ne ovisi o redosljedu faktora, imamo

$$\pi(p^{-1}) = \prod_{1\le i<j\le n} \frac{j-i}{p(j)-p(i)} = \frac{1}{\pi(p)} = \pi(p)\,.$$

Zadnja jednakost vrijedi jer je $\pi(p) \in \{1, -1\}$.

(ii) Pokažimo tvrdnju prvo za slučaj $j = i+1$, dakle za slučaj specijalne transpozicije $p_{i,i+1}$. Uvedimo oznaku $\tilde{p} = p \circ p_{i,i+1}$. Imamo

$$\begin{aligned} p &= \begin{pmatrix} 1 & \dots & i & i+1 & \dots & n \\ p(1) & \dots & p(i) & p(i+1) & \dots & p(n) \end{pmatrix}, \\ p \circ p_{i,i+1} = \tilde{p} &= \begin{pmatrix} 1 & \dots & i & i+1 & \dots & n \\ p(1) & \dots & p(i+1) & p(i) & \dots & p(n) \end{pmatrix}. \end{aligned}$$

Uočimo da je:

- $(p(r), p(s))$, $r < i$, $i+1 < s$, inverzija u p, ako i samo ako je $(\tilde{p}(r), \tilde{p}(s))$ inverzija u $\tilde{p}$,
- $(p(r), p(i))$, $r < i$, inverzija u p, ako i samo ako je $(\tilde{p}(r), \tilde{p}(i+1))$ inverzija u $\tilde{p}$,
- $(p(r), p(i+1))$, $r < i$, inverzija u p, ako i samo ako je $(\tilde{p}(r), \tilde{p}(i))$ inverzija u $\tilde{p}$,
- $(p(i), p(s))$, $i+1 < s$, inverzija u p, ako i samo ako je $(\tilde{p}(i+1), \tilde{p}(s))$ inverzija u $\tilde{p}$,
- $(p(i+1), p(s))$, $i+1 < s$, inverzija u p, ako i samo ako je $(\tilde{p}(i), \tilde{p}(s))$ inverzija u $\tilde{p}$,
- $(p(i), p(i+1))$ inverzija u p, ako i samo ako $(\tilde{p}(i), \tilde{p}(i+1))$ nije inverzija u $\tilde{p}$.

Stoga $\tilde{p}$ ima točno jednu više ili jednu manje inverziju od p, pa mora biti $\pi(\tilde{p}) = -\pi(p)$.

Neka je sada $j > i + 1$. Tada primijenjujući na p niz od $2(j - i) - 1$ transpozicije $p_{s,s+1}$ dobivamo $p \circ p_{ij}$. Pritom s prvo prolazi vrijednosti i,$i + 1$, ...,$j - 2$, a nakon toga vrijednosti $j - 1$, $j - 2$, ...,i,

$$\begin{aligned} p \circ p_{ij} &= (\cdots (((\cdots (p \circ p_{i,i+1}) \circ \cdots \circ p_{j-2,j-1}) \circ p_{j-1,j}) \\ &\quad \circ p_{j-2,j-1}) \circ \cdots \circ p_{i+1,i+2}) \circ p_{i,i+1} \,. \end{aligned}$$

Prema gore pokazanom, svaka primjena transpozicije $p_{s,s+1}$ mijenja parnost permutacije, a kako takovih transpozicija ima neparni broj, moraju parnosti od $p \circ p_{ij}$ i od p imati suprotne predznake.

(iii) Koristeći tvrdnju (i), relaciju $(p_{ij} \circ p)^{-1} = p^{-1} \circ p_{ij}^{-1}$, relaciju (2.18), tvrdnju (ii) i na kraju opet tvrdnju (i), dobivamo

$$\begin{aligned} \pi(p_{ij} \circ p) &= \pi\left((p_{ij} \circ p)^{-1}\right) = \pi(p^{-1} \circ p_{ij}^{-1}) = \pi(p^{-1} \circ p_{ij}) \\ &= -\pi(p^{-1}) = -\pi(p). \end{aligned}$$ ■

Zadatak 10.5 *Dokažite tvrdnju (ii) propozicije 10.4 uz pomoć formule (10.1).*

Rješenje: Pokazat ćemo da vrijedi

$$\frac{\pi(p \circ p_{ij})}{\pi(p)} = -1 \,. \tag{10.2}$$

Uvedimo opet oznaku $\tilde{p} = p \circ p_{ij}$. Prema propoziciji 10.2, $\pi(\tilde{p})$ i $\pi(p)$ su produkti od po $n(n-1)$ faktora od kojih su mnogi u oba produkta isti. Npr. za $n = 12$, $i = 4$, $j = 9$, dijagram na sljedećoj stranici pokazuje za koje parove indeksa r, s, faktori $(\tilde{p}(s) - \tilde{p}(r))/(s - r)$ u $\pi(\tilde{p})$ moraju biti jednaki faktorima $(p(s) - p(r))/(s - r)$ u $\pi(p)$ (označeni su sa $\cdot$), i za koje parove indeksa mogu biti različiti (označeni su sa x).

Općenito će vrijediti

$$\begin{aligned} \frac{\pi(\tilde{p})}{\pi(p)} &= \frac{\tilde{p}(j) - \tilde{p}(i)}{p(j) - p(i)} \cdot \prod_{k=1}^{i-1} \frac{\tilde{p}(i) - \tilde{p}(k)}{p(i) - p(k)} \cdot \frac{\tilde{p}(j) - \tilde{p}(k)}{p(j) - p(k)} \\ &\cdot \prod_{k=j+1}^{n} \frac{\tilde{p}(k) - \tilde{p}(i)}{p(k) - p(i)} \cdot \frac{\tilde{p}(k) - \tilde{p}(j)}{p(k) - p(j)} \cdot \prod_{k=i+1}^{j-1} \frac{\tilde{p}(k) - \tilde{p}(i)}{p(k) - p(i)} \cdot \frac{\tilde{p}(j) - \tilde{p}(k)}{p(j) - p(k)} \end{aligned}$$

$$
\begin{aligned}
&= \frac{p(i)-p(j)}{p(j)-p(i)} \cdot \prod_{k=1}^{i-1} \frac{p(j)-p(k)}{p(i)-p(k)} \cdot \frac{p(i)-p(k)}{p(j)-p(k)} \\
&\prod_{k=j+1}^{n} \frac{p(k)-p(j)}{p(k)-p(i)} \cdot \frac{p(k)-p(i)}{p(k)-p(j)} \cdot \prod_{k=i+1}^{j-1} \frac{p(k)-p(j)}{p(k)-p(i)} \cdot \frac{p(i)-p(k)}{p(j)-p(k)} \\
&= (-1)\cdot 1\cdot 1\cdot 1 = -1\,.
\end{aligned}
$$

	1	2	3	i	5	6	7	8	j	10	11	12	s
1		·	·	x	·	·	·	·	x	·	·	·	
2			·	x	·	·	·	·	x	·	·	·	
3				x	·	·	·	·	x	·	·	·	
$i=4$					x	x	x	x	x	x	x	x	
5						·	·	·	x	·	·	·	
6							·	·	x	·	·	·	
7								·	x	·	·	·	
8									x	·	·	·	
$j=9$										x	x	x	
10											·	·	
11												·	
r													

10.1 Determinante: osnovna svojstva

U ovom dijelu definiramo determinantu matrice i dokazujemo njena osnovna svojstva.

Definicija 10.6 **Determinanta** *matrice* $A=(a_{ij})\in M_n$ *je skalar*

$$\det(A) = \sum_{p\in\Pi_n} \pi(p) a_{1p(1)} a_{2p(2)} a_{3p(3)} \cdots a_{np(n)}\,.$$

Determinanta $\det(A)$ *je reda* n *ako je* A *reda* n. ■

Umjesto $\det(A)$ još se koristi oznaka $\det A$, $|A|$ ili

$$\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}\,.$$

Slično kao kod matrice, govorit ćemo o elementima, recima, stupcima i dijagonali determinante. Uočimo da se u svakom članu sume nalazi po jedan element iz svakog stupca i po jedan element iz svakog retka matrice. Stoga, u slučaju jedinične matrice, od $n!$ članova sume, samo jedan je različit od nule,

$$\det(I_n) = (I_n)_{11}(I_n)_{22}\cdots(I_n)_{nn} = 1\cdots 1 = 1\,. \tag{10.3}$$

Definicija 10.6 je matematički jasna, ali je za računanje determinante reda većeg od tri praktički neupotrebljiva. Naime, broj članova u sumi, pa zato i broj računskih operacija, strahovito brzo raste s n.

Primjer 10.7 *Izračunajte broj operacija potrebnih za računanje determinante reda n, na osnovu definicije 10.6. Pretpostavite da* $\det(A)$ *računate na brzom računalu koje izvršava jednu operaciju množenja za jednu miljarditinku sekunde. Koliko vremena je potrebno za računanje determinante reda* 10, 20, 30, 50 100 *?*

Rješenje Na računalu je operacija množenja nekoliko puta sporija od operacije zbrajanja ili oduzimanja. U formuli za determinantu ima n puta više operacija množenje od operacija zbrajanja. Zato možemo zanemariti aditivne operacije. Prema formuli iz definicije 10.6, postoji ukupno $n!$ pribrojnika, a svaki je produkt od n faktora. Za produkt od n faktora treba $n-1$ množenje, pa ukupno ima $op(n) = (n-1)\cdot n!$ operacija množenja. U sljedećoj tabeli izračunati su za dane vrijednosti od n, brojevi $op(n) = (n-1)\cdot n!$ i pripadna vremena: $t(n)$ u sekundama i $T(n)$ u godinama.

n	$op(n)$	$t(n)$	$T(n)$
10	$3.26592 \cdot 10^{007}$	$3.26592 \cdot 10^{-02}$	$1.0349 \cdot 10^{-09}$
20	$4.62251 \cdot 10^{019}$	$4.62251 \cdot 10^{010}$	$1.4648 \cdot 10^{003}$
30	$7.69233 \cdot 10^{033}$	$7.69233 \cdot 10^{024}$	$2.4376 \cdot 10^{017}$
50	$1.49029 \cdot 10^{066}$	$1.49029 \cdot 10^{057}$	$4.7224 \cdot 10^{049}$
100	$9.23929 \cdot 10^{159}$	$9.23930 \cdot 10^{150}$	$2.9277 \cdot 10^{143}$

Vidimo da bi već za $n = 30$ računali duže (desetak milijuna puta duže) od procijenjene starosti svemira, od velikog praska do danas! ■

Za veće n je najbolje računati determinantu koristeći elementarne transformacije nad retcima ili stupcima, ali prije toga moramo otkriti osnovna svojstva determinante da bi vidjeli kako se ona mijenja kad se na matricu primijenjuju elementarne operacije. Na kraju ćemo vidjeti da korištenjem Gaussove metode eliminacija s parcijalnim pivotiranjem možemo determinantu izračunati pomoću približno $n^3/3$ operacija množenja. Za $n = 1000$ to znači oko $10^9/3$ množenja ili na brzom računalu iz primjera 10.7, za oko trećinu sekunde.

Prvo ćemo odrediti izraze za determinante reda 2 i 3.

Primjer 10.8 *Neka je*

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

Skup Π_2 *svih permutacija skupa* S_2*, sastoji se samo iz dviju permutacija:*

$$p_1 = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} \quad i \quad p_2 = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}.$$

Pritom je očito $\pi(p_1) = 1$, $\pi(p_2) = -1$. *Prema definiciji 10.6, imamo*

$$\det A = a_{11}a_{22} - a_{12}a_{21} = ad - bc. \qquad \blacksquare$$

Primjer 10.9 *Neka je*

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}.$$

Skup Π_3 *se sastoji od šest permutacija:*

$$\begin{aligned} p_1 &= \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \quad p_2 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \quad p_3 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \\ p_4 &= \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \quad p_5 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \quad p_6 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}. \end{aligned}$$

Pritom su p_1, p_2, p_3 *parne, a* p_4, p_5, p_6 *neparne permutacije, pa je*

$$\begin{aligned} \det(A) = & \, a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ & -a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{31}a_{22}a_{13}. \end{aligned}$$

Ovaj izraz ćemo najlakše upamtiti ako determinanti (ili matrici) nadopišemo još jednom prvi i drugi stupac s desne strane

$$\left|\begin{matrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{matrix}\right|\begin{matrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{matrix}$$

i onda zbrojimo produkte elemenata od tri dijagonale u smjeru $\searrow$ i od te vrijednosti oduzmemo sumu produkata elemenata od tri dijagonale u smjeru $\nearrow$. pa Taj postupak se naziva Sarrusovo pravilo. ■

Propozicija 10.10 *Vrijedi*

$$\det(A^T) = \det(A)\,.$$

Dokaz: U matrici A^T, na mjestu (i,j) nalazi se a_{ji}. Stoga po definiciji determinante imamo

$$\det(A^T) = \sum_{p\in\Pi_n} \pi(p) a_{p(1)1} a_{p(2)2} a_{p(3)3} \cdots a_{p(n)n}\,.$$

Zato jer vrijedi

$$\begin{pmatrix} p(1) & p(2) & \dots & p(n) \\ 1 & 2 & \dots & n \end{pmatrix} = \begin{pmatrix} 1 & 2 & \cdots & n \\ p^{-1}(1) & p^{-1}(2) & \cdots & p^{-1}(n) \end{pmatrix} = p^{-1}\,,$$

$\pi(p) = \pi(p^{-1})$, kao i komutativnost množenja skalara, imamo

$$\pi(p) a_{p(1)1} a_{p(2)2} a_{p(3)3} \cdots a_{p(n)} = \pi(p^{-1}) a_{1p^{-1}(1)} a_{2p^{-1}(2)} a_{3p^{-1}(3)} \cdots a_{np^{-1}(n)}\,.$$

Dakle je

$$\det(A^T) = \sum_{p\in\Pi_n} \pi(p^{-1}) a_{1p^{-1}(1)} a_{2p^{-1}(2)} a_{3p^{-1}(3)} \cdots a_{np^{-1}(n)}\,.$$

Jer je $(\Pi_n, \circ)$ grupa, kad p prolazi skupom Π_n, p^{-1} također prolazi skupom Π_n (svaki element p grupe je inverz od točno jednog elementa grupe: od p^{-1}). Stoga, pišući $q = p^{-1}$, dobivamo

$$\det(A^T) = \sum_{q\in\Pi_n} \pi(q) a_{1q(1)} a_{2q(2)} a_{3(3)} \cdots a_{nq(n)} = \det(A)\,.$$

■

Propozicija 10.11 *Ako u determinanti zamijenimo dva retka (stupca), ona mijenja predznak.*

Dokaz: Dokaz provodimo za stupce, a za retke onda dokazujemo pomoću prethodne propozicije.

Pretpostavimo da smo u matrici A zamijenili stupce a_i i a_j i tako dobili matricu $\tilde{A} = (\tilde{a}_{rt})$. Tada je $\tilde{A} = AI_{ij}$, gdje je I_{ij} matrica transpozicije za koju vrijedi $I_{ij}e_t = e_{p_{ij}(t)}$. $1 \leq t \leq n$. Stoga po relaciji (9.7), vrijedi $\tilde{a}_{rt} = a_{rp_{ij}(t)}$, a kako u definiciju od $\det(\tilde{A})$ ulaze elementi $\tilde{a}_{rp(r)}$, imamo

$$\tilde{a}_{rp(r)} = a_{r,p_{ij}(p(r))} = a_{r,(p_{ij}\circ p)(r))} = a_{rq(r)} \quad \text{za svako } 1 \leq r \leq n,$$

pri čemu je $q = p_{ij} \circ p$ nova permutacija. Pritom zbog propozicije 10.4 (iii) vrijedi

$$\pi(q) = \pi(p_{ij} \circ p) = -\pi(p).$$

Relacija (2.16) garantira da će q prolaziti kroz cijeli skup Π_n, ako to čini p. Stoga vrijedi

$$\begin{aligned}
\det(\tilde{A}) &= \sum_{p\in\Pi_n} \pi(p)\tilde{a}_{1p(1)}\tilde{a}_{2p(2)} \cdots \tilde{a}_{np(n)} \\
&= \sum_{q\in\Pi_n} (-\pi(q))a_{1q(1)}a_{2q(2)} \cdots a_{nq(n)} \\
&= -\det(A)\,.
\end{aligned}$$

Dokaz za retke koristi simetričnost matrice transpozicije, tvrdnju za stupce i propoziciju 10.10 i

$$\det(I_{ij}A) = \det((I_{ij}A)^T) = \det(A^T I_{ij}) = -\det(A^T) = -\det(A). \qquad \blacksquare$$

Korolar 10.12 *Determinanta s dva jednaka retka ili stupca, ima vrijednost nula.*

Dokaz: Zamjenom dva jednaka retka ili stupca, matrica se ne mijenja, pa zato i njena determinanta ostaje ista. S druge strane, prema propoziciji 10.11, zamjenom dva retka ili stupca, determinanta mijenja predznak. Stoga mora vrijediti $\det(A) = -\det(A)$, pa je $\det(A) = 0$.

$\blacksquare$

Propozicija 10.13 *Ako se jedan redak (stupac) determinante pomnoži skalarom α, cijela determinanta se množi s α.*

Dokaz: Ako $\tilde{A}$ nastaje iz A množenjem i-tog retka sa α,

$$\tilde{A} = D_i(\alpha)A\,,$$

tada vrijedi

$$\begin{aligned}
\det(\tilde{A}) &= \sum_{p\in\Pi_n} \pi(p)\tilde{a}_{1p(1)}\cdots\tilde{a}_{np(n)} \\
&= \sum_{p\in\Pi_n} \pi(p)a_{1p(1)}\cdots(\alpha a_{ip(i)})\cdots a_{np(n)} \\
&= \alpha\sum_{p\in\Pi_n} \pi(p)a_{1p(1)}\cdots a_{ip(i)}\cdots a_{np(n)} \\
&= \alpha\det(A)\,.
\end{aligned}$$

Tvrdnja za stupce slijedi uz pomoć propozicije 10.10, tvrdnje za retke i simetričnosti matrice $D_i(\alpha)$,

$$\begin{aligned}
\det(AD_i(\alpha)) &= \det((AD_i(\alpha))^T) = \det(D_i(\alpha)A^T) \\
&= \alpha\det(A^T) = \alpha\det(A).
\end{aligned}$$

■

Korolar 10.14 *Ako determinanta ima jedan nul-redak ili nul-stupac, njena vrijednost je nula.*

Dokaz: U prethodnoj propoziciji treba uzeti $\alpha = 0$, a indeks i od $D_i(\alpha)$ je onaj od null-retka ili nul-stupca. ■

Specijalno, propozicija 10.13 implicira

$$\det(\alpha A) = \alpha^n\det(A), \quad A \in \mathbf{R}^{n\times n}\,. \tag{10.4}$$

Propozicija 10.15 *Neka su*

$$\begin{aligned}
A &= [a_1,\ldots a_{i-1},\ a_i\ ,a_{i+1},\ldots,a_n], \\
B &= [a_1,\ldots a_{i-1},\ b_i\ ,a_{i+1},\ldots,a_n] \quad i \\
C &= [a_1,\ldots a_{i-1}, a_i+b_i, a_{i+1},\ldots,a_n]
\end{aligned}$$

stupčane particije kvadratnih matrica A, B i C, respektivno. Tada vrijedi

$$\det(C) = \det(A) + \det(B).$$

Dokaz: Označimo elemente matrica: $A = (a_{rt})$, $B = (b_{rt})$, $C = (c_{rt})$. Izuzev u i-tom stupcu, elementi matrica A, B i C na istim pozicijama, su jednaki. Stoga za svaku permutaciju $p \in \Pi_n$ vrijedi

$$\begin{aligned}
\pi(p)c_{1p(1)} \cdots c_{np(n)} &= \pi(p)a_{p^{-1}(1),1} \cdots c_{p^{-1}(i),i} \cdots a_{p^{-1}(n),n} \\
&= \pi(p)a_{p^{-1}(1),1} \cdots (a_{p^{-1}(i),i} + b_{p^{-1}(i),i}) \cdots a_{p^{-1}(n),n} \\
&= \pi(p)a_{p^{-1}(1),1} \cdots a_{p^{-1}(i),i} \cdots a_{p^{-1}(n),n} \\
&\quad +\pi(p)a_{p^{-1}(1),1} \cdots b_{p^{-1}(i),i} \cdots a_{p^{-1}(n),n} \\
&= \pi(p)a_{1p(1)} \cdots a_{ip(i)} \cdots a_{np(n)} \\
&\quad +\pi(p)a_{1p(1)} \cdots b_{ip(i)} \cdots a_{np(n)},
\end{aligned}$$

odakle, uzimanjem sume po p, odmah slijedi

$$\det(C) = \det(A) + \det(B)\,. \qquad \blacksquare$$

Korolar 10.16 *Ako su*

$$A = \begin{bmatrix} a_1^T \\ \vdots \\ a_{i-1}^T \\ a_i^T \\ a_{i+1}^T \\ \vdots \\ a_n^T \end{bmatrix}, \quad B = \begin{bmatrix} a_1^T \\ \vdots \\ a_{i-1}^T \\ b_i^T \\ a_{i+1}^T \\ \vdots \\ a_n^T \end{bmatrix}, \quad C = \begin{bmatrix} a_1^T \\ \vdots \\ a_{i-1}^T \\ a_i^T + b_i^T \\ a_{i+1}^T \\ \vdots \\ a_n^T \end{bmatrix},$$

retčane particije kvadratnih matrica A, B i C, respektivno, tada vrijedi

$$\det(C) = \det(A) + \det(B).$$

Dokaz: Koristeći propozicije 10.10 i 10.15 dobijemo

$$\det(C) = \det(C^T) = \det(A^T) + \det(B^T) = \det(A) + \det(B). \qquad \blacksquare$$

Propozicija 10.17 *Ako jednom retku (stupcu) determinante dodamo neki drugi redak (stupac) pomnožen proizvoljnim skalarom, determinanta matrice se ne mijenja.*

Dokaz: Ako matrica $\tilde{A}$ nastaje iz A tako da i-tom stupcu dodamo j-ti stupac pomnožen s α, onda je

$$\tilde{A} = [a_1, \ldots, a_i + \alpha a_j, \ldots, a_j, \ldots a_n] .$$

Prema propoziciji 10.15, $\det(\tilde{A})$ se može napisati kao zbroj dviju determinanti, a onda se još iskoristi propozicija 10.13,

$$\begin{aligned} \det(\tilde{A}) &= \det(A) + \alpha \det([a_1, \ldots, a_j, \ldots, a_j, \ldots a_n]) \\ &= \det(A) + \alpha \cdot 0 \\ &= \det(A) . \end{aligned}$$

Ako se i-ti stupac dodaje j-tom stupcu, tada je

$$\tilde{A} = [a_1, \ldots, a_i, \ldots, \alpha a_i + a_j, \ldots, a_n] ,$$

a dokaz koristi isti obrazac.

Dokaz za retke se sprovodi na isti način, koristeći korolar 10.16 umjesto propozicije 10.15, a lako se dobije i iz tvrdnje za stupce uz pomoć propozicije 10.10. ■

Sljedeći teorem sažima gotovo sve dosadašnje rezultate.

Teorem 10.18 *Neka je A proizvoljna kvadratna matrica i neka je E matrica elementarne transformacije. Tada je*

$$\det(EA) = \det(E)\det(A) = \det(AE) . \tag{10.5}$$

Dokaz: Za E postoje tri mogućnosti.

1. $E = I_{ij}$. Nakon zamjene i-tog i j-tog retka, odnosno stupca, matrice A, prema propoziciji 10.11, vrijedi

$$\det(I_{ij}A) = -\det(A), \qquad \det(AI_{ij}) = -\det(A) .$$

Uzimanjem jedinične matrice umjesto A i korištenjem relacije (10.3), dobivamo $\det(I_{ij}) = -1$, pa je (10.5) pokazano.

2. $E = D_i(\alpha)$. U ovom slučaju je prema propoziciji 10.13,

$$\det(D_i(\alpha)A) = \alpha \det(A), \qquad \det(AD_i(\alpha)) = \alpha \det(A) .$$

Zamjenjujući A sa I_n i korištenjem relacije (10.3), odmah dobivamo $\det(D_i(\alpha)) = \alpha$, pa (10.5) vrijedi i u ovom slučaju.

3. $E = M_{ij}(\alpha)$ ili $E = M_{ij}^T(\alpha)$. Sada je prema propoziciji 10.17,

$$\begin{aligned}\det(M_{ij}(\alpha)A) &= \det(A) = \det(AM_{ij}(\alpha)) \\ \det(M_{ij}^T(\alpha)A) &= \det(A) = \det(AM_{ij}^T(\alpha))\,.\end{aligned}$$

Zamjenjujući opet A sa I_n, dobivamo $\det(M_{ij}(\alpha)) = 1 = \det(M_{ij}^T(\alpha))$, pa i u ovom slučaju vrijedi (10.5). ■

U dokazu teorema 10.18 smo pokazali da za elementarnu matricu $E \in \mathbf{R}^{n\times n}$ vrijedi

$$\det(E) = \begin{cases} -1, & \text{ako je } E = I_{ij} \\ \alpha, & \text{ako je } E = D_i(\alpha) \\ 1, & \text{ako je } E = M_{ij}(\alpha) \text{ ili } E = M_{ij}^T(\alpha) \end{cases} \tag{10.6}$$

pa zaključujemo da su determinante elementarnih matrica različite od nule. Kako su elementarne matrice nesingularne, postavlja se pitanje imaju li sve nesingularne matrice determinantu različitu od nule.

Korolar 10.19 *Kvadratna matrica A je regularna ako i samo ako je* $\det(A) \neq 0$.

Dokaz: Za svaku matricu A postoji konačni niz elementarnih matrica $E_1, \cdots, E_k$ (svađanje na reduciranu formu), takvih da je

$$A_R = E_k \cdots E_1 A\,.$$

Pritom je A_R (retčana) reducirana forma od A. Po teoremu 10.18 vrijedi

$$\begin{aligned}\det(A_R) &= \det(E_k(E_{k-1}\cdots E_1A)) \\ &= \det(E_k)\det(E_{k-1}(E_{k-2}\cdots E_1A)) \\ &= \det(E_k)\det(E_{k-1})\det(E_{k-2}\cdots E_1A) = \cdots \\ &= \det(E_1)\det(E_2)\cdots\det(E_k)\det(A).\end{aligned}$$

Kako su sve $\det(E_i) \neq 0$, zaključujemo da je

$$\det(A_R) \neq 0 \quad \text{ako i samo ako je} \quad \det(A) \neq 0. \tag{10.7}$$

Kako je A_R kvadratna, postoje dvije mogućnosti: ili je $A_R = I_n$ ili A_R ima samo nule u zadnjem retku. Po definiciji, A je nesingularna (singularna) ako A_R ima (nema) puni rang. Dakle, A je nesingularna (singularna) ako i samo ako je $A_R = I_n$ (zadnji redak od A_R je nul-redak) tj. ako i samo ako je $\det(A_R) = 1$ ($\det(A_R) = 0$), a to je zbog (10.7), ako i samo ako je $\det(A) \neq 0$ ($\det(A) = 0$). ■

Zbog $\det(A^T) = \det(A)$, odmah slijedi da je A singularna (regularna) ako i samo ako je ako A^T takva (to već znamo zbog teorema 8.11). Sljedeći važan teorem odnosi se na determinantu produkta matrica.

Teorem 10.20 (Binet-Cauchy) *Za kvadratne matrice A i B vrijedi*

$$\det(AB) = \det(A)\det(B). \tag{10.8}$$

Dokaz: Ako je barem jedna od matrica singularna, tada je i produkt AB singularna matrica (za dokaz vidi primjer 10.21) pa (10.8) vrijedi s nulama na svakoj strani jednakosti. Ako su A i B obje regularne, onda postoje elementarne matrice $E_1, \ldots, E_k$ i $F_1, \ldots, F_m$, takve da je $A = E_1 \cdots E_k$ i $B = F_1 \cdots F_m$. Na isti način kao u korolaru 10.19 pokaže se da je

$$\det(A) = \det(E_1) \cdots \det(E_k) \quad \text{i} \quad \det(B) = \det(F_1) \cdots \det(F_m).$$

Koristeći teorem 10.18 i zadnju relaciju dobije se

$$\begin{aligned}
\det(AB) &= \det((E_1E_2 \cdots E_k)(F_1 \cdots F_m)) \\
&= \det(E_1(E_2 \cdots E_kF_1 \cdots F_m)) \\
&= \det(E_1)\det(E_2(E_3 \cdots E_kF_1 \cdots F_m)) = \cdots \\
&= \det(E_1)\det(E_2) \cdots \det(E_k)\det(F_1) \cdots \det(F_m) \\
&= [\det(E_1) \cdots \det(E_k)]\,[\det(F_1) \cdots \det(F_m)] \\
&= \det(A)\det(B).
\end{aligned}$$

■

Primjer 10.21 *Dokažite da je produkt AB sinularna matrica ako i samo ako je bar jedna od matrica A ili B singularna.*

Dokaz: Neka je barem jedna od matrica A ili B singularna. Pokažimo da je tada i AB tada singularna matrica. Doista, ako je B singularna, tada postoji $x \neq 0$ takav da je $Bx = 0$. Stoga je $(AB)x = A(Bx) = A0 = 0$, pa je AB singularna. Ako je A singularna, tada je i A^T singularna, pa je zbog $(AB)^T x = (B^T A^T)x = B^T(A^T x) = B^T 0 = 0$ i $(AB)^T$ singularna, a to znači da je i AB singularna. Obrnuto, ako je AB singularna, tada postoji $x \neq 0$ takav da je $A(Bx) = 0$, pa je ili $Bx = 0$ ili $Bx \neq 0$ i $A(Bx) = 0$, tj. ili je B singularna ili je A singularna. ■

10.2 Razvoj determinante

Kao što smo vidjeli, računanje determinante većeg reda pomoću formule iz definicije je praktički nemoguće. U ovom dijelu ćemo dokazati rezultat koji to računanje znatno olakšava.

Neka je $A \in \mathbf{R}^{n\times n}$. Sa $A_{ij}^c \in \mathbf{R}^{(n-1)\times(n-1)}$ ćemo označiti podmatricu od A koja nastaje izbacivanjem njenog i-tog retka i j-tog stupca. Npr. ako je

$$A = \begin{bmatrix} a_{11} & \ldots & a_{1,j-1} & a_{1j} & a_{1,j+1} & \ldots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \ldots & a_{i-1,j-1} & a_{i-1,j} & a_{i-1,j+1} & \ldots & a_{i-1,n} \\ a_{i,1} & \ldots & a_{i,j-1} & a_{ij} & a_{i,j+1} & \ldots & a_{i,n} \\ a_{i+1,1} & \ldots & a_{i+1,j-1} & a_{i+1,j} & a_{i+1,j+1} & \ldots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \ldots & a_{n,j-1} & a_{nj} & a_{n,j+1} & \ldots & a_{nn} \end{bmatrix},$$

onda je

$$A_{ij}^c = \begin{bmatrix} a_{11} & \ldots & a_{1,j-1} & a_{1,j+1} & \ldots & a_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \ldots & a_{i-1,j-1} & a_{i-1,j+1} & \ldots & a_{i-1,n} \\ a_{i+1,1} & \ldots & a_{i+1,j-1} & a_{i+1,j+1} & \ldots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & \ldots & a_{n,j-1} & a_{n,j+1} & \ldots & a_{nn} \end{bmatrix}.$$

Definicija 10.22 **Minora** *elementa* a_{ij} *matrice* $A \in \mathbf{R}^{n\times n}$ *je determinanta matrice* A_{ij}^c*, koja nastaje iz* A *izbacivanjem* i*-tog retka i* j*-tog stupca.* **Algebarski komplement** *ili kofaktor elementa* a_{ij} *je skalar* $(-1)^{i+j}\det(A_{ij}^c)$*.* ■

Dakle algebarski komplement se razlikuje od odgovarajuće minore tek u faktoru $(-1)^{i+j}$, tj. eventualno do na predznak.

Lema 10.23 *Ako za elemente prvog retka matrice* $A \in \mathbf{R}^{n\times n}$ *vrijedi*

$$a_{1j} = 0, \quad 2 \le j \le n$$

onda je

$$\det(A) = a_{11}\det(A_{11}^c).$$

Dokaz: U definiciji determinante matrice A,

$$\det(A) = \sum_{p\in\Pi_n} \pi(p) a_{1p(1)} a_{2p(2)} a_{3p(3)} \cdots a_{np(n)}\,.$$

svaki član sume ima kao faktor jedan element iz prvog retka. Zbog uvjeta na matricu, preostaju samo članovi sume koji imaju faktor a_{11}, a to su točno oni za koje je $p(1) = 1$. Prema tome je

$$\det(A) = a_{11}\left(\sum_{\substack{p\in\Pi_n\\ p(1)=1}} \pi(p) a_{2p(2)} a_{3p(3)} \cdots a_{np(n)}\right).$$

Svakoj permutaciji $p \in \Pi_n$ koja zadovoljava uvjet $p(1) = 1$ pridružena je permutacija $\tilde{p}$ na skupu $\{2, 3, \ldots, n\}$,

$$\tilde{p} = \begin{pmatrix} 2 & 3 & \ldots & n \\ p(2) & p(3) & \ldots & p(n) \end{pmatrix}.$$

Obrnuto, svakoj permutaciju $\tilde{p}$ na skupu $\{2, 3, \ldots, n\}$, jednoznačno je pridružena permutaciju $p \in \Pi_n$ formulom

$$p = \begin{pmatrix} 1 & 2 & \ldots & n \\ 1 & \tilde{p}(2) & \ldots & \tilde{p}(n) \end{pmatrix}.$$

Parnost od p jednaka je parnosti od $\tilde{p}$, jer svaka inverzija u $\tilde{p}$ znači i odgovarajuću inverziju u p, a s druge strane $(p(1), p(j)) = (1, p(j))$ nije inverzija ni za jedno $2 \leq j \leq n$. Prema tome, kad p u sumi prolazi sve permutacije iz Π_n za koje vrijedi $p(1) = 1$, pridružena $\tilde{p}$ prolazi svim permutacijama na skupu $\{2, \ldots, n\}$ i pritom $\tilde{p}$ ima istu parnost kao i p. Stoga možemo pisati

$$\begin{aligned}
\det(A) &= a_{11} \cdot \sum_{\substack{p\in\Pi_n \\ p(1)=1}} \pi(p) a_{2p(2)} a_{3p(3)} \cdots a_{np(n)} \\
&= a_{11} \cdot \sum_{\tilde{p}} \pi(\tilde{p}) a_{2\tilde{p}(2)} a_{3\tilde{p}(3)} \cdots a_{n\tilde{p}(n)} \\
&= a_{11} \det(A^c_{ij}) .
\end{aligned}$$

■

Zadatak 10.24 *Pokažite pomoću formule (10.1) jednakost parnosti permutacija p i $\tilde{p}$ iz dokaza leme 10.23.*

Rješenje:

$$\begin{aligned}
\pi(p) &= \prod_{1\leq i<j\leq n} \frac{p(j)-p(i)}{j-i} = \prod_{j=2}^{n} \frac{p(j)-p(1)}{j-1} \cdot \prod_{2\leq i<j\leq n} \frac{p(j)-p(i)}{j-i} \\
&= \prod_{j=2}^{n} \frac{p(j)-1}{j-1} \cdot \prod_{2\leq i<j\leq n} \frac{\tilde{p}(j)-\tilde{p}(i)}{j-i} \\
&= 1 \cdot \prod_{2\leq i<j\leq n} \frac{\tilde{p}(j)-\tilde{p}(i)}{j-i} = 1 \cdot \pi(\tilde{p}) \\
&= \pi(\tilde{p}) .
\end{aligned}$$

■

Teorem 10.25 *Za $A \in \mathbf{R}^{n\times n}$, $A = (a_{ij})$ vrijedi*

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A^c_{ij}), , \quad 1 \leq i \leq n.$$

Ovu formulu nazivamo **razvoj determinante po i-tom retku**. *Isto tako vrijedi formula*

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(A^c_{ij}), , \quad 1 \leq j \leq n$$

koju nazivamo **razvoj determinante po j-tom stupcu**.

Dokaz: Prvo dokazujemo razvoj po retcima. Neka je $1 \leq i \leq n$ proizvoljan. Polazeći od matrice A i retka i, formiramo n novih matrica $A_i^{(1)}$, $A_i^{(2)}$, $\ldots$,$A_i^{(n)}$, tako da za njihove retke vrijedi

$$e_r^T A_i^{(j)} = \begin{cases} e_r^T A, & \text{ako je } r \neq i \\ a_{ij} e_i^T, & \text{ako je } r = i \end{cases} .$$

Dakle, za $1 \leq j \leq n$ vrijedi

$$A_i^{(j)} = \begin{bmatrix} a_{11} & \ldots & a_{1,j-1} & a_{1j} & a_{1,j+1} & \ldots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \ldots & a_{i-1,j-1} & a_{i-1,j} & a_{i-1,j+1} & \ldots & a_{i-1,n} \\ 0 & \ldots & 0 & a_{ij} & 0 & \ldots & 0 \\ a_{i+1,1} & \ldots & a_{i+1,j-1} & a_{i+1,j} & a_{i+1,j+1} & \ldots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \ldots & a_{n,j-1} & a_{nj} & a_{n,j+1} & \ldots & a_{nn} \end{bmatrix} .$$

Prema propoziciji 10.15 vrijedi

$$\det(A) = \det(A_i^{(1)}) + \det(A_i^{(2)}) + \ldots + \det(A_i^{(n)}).$$

Da bi odredili vrijednost determinante $\det(A_i^{(j)})$, načinimo $i-1$ transpoziciju susjednih redaka (prvo $i \leftrightarrow i-1$, zatim $i-1 \leftrightarrow i-2$, pa $i-2 \leftrightarrow i-3$, itd. na kraju $2 \leftrightarrow 1$) kako bi doveli element a_{ij} na poziciju $(1,j)$, a zatim načinimo $j-1$ transpoziciju susjednih stupaca (prvo $j \leftrightarrow j-1$, pa $j-1 \leftrightarrow j-2$, itd., na kraju $2 \leftrightarrow 1$) da bi doveli polazni a_{ij} na poziciju $(1,1)$. Taj postupak možemo matrično opisati

$$\tilde{A}_i^{(j)} = I_{12} I_{23} \cdots I_{i-1,i} A_i^{(j)} I_{j-1,j} \cdots I_{23} I_{12}. \tag{10.9}$$

Matrica $\tilde{A}_i^{(j)}$ ima oblik

$$\tilde{A}_i^{(j)} = \begin{bmatrix} a_{ij} & 0^T \\ a_1' & A_{ij}^c \end{bmatrix} ,$$

gdje je 0^T nul-redak dimenzije $n-1$, a a_1' vektor stupac duljine $n-1$. Vidimo da $\tilde{A}_i^{(j)}$ zadovoljava pretpostavku leme 10.23. Primijenimo determinantu na matricu s lijeve i desne strane jednakosti (10.9). Zatim iskoristimo lemu 10.23 za determinantu na lijevoj strani jednakosti i teorem 10.18 i relaciju (10.6) za determinantu na desnoj strani jednakosti. Dobivamo

$$\begin{aligned} a_{ij}\det(A_{ij}^c) &= \det(I_{12})\cdots\det(I_{i-1,i})\det(A_i^{(j)})\det(I_{j-1,j})\cdots\det(I_{12}) \\ &= (-1)^{i-1}\det(A_i^{(j)})(-1)^{j-1} = (-1)^{i+j-2}\det(A_i^{(j)}) \\ &= (-1)^{i+j}\det(A_i^{(j)}). \end{aligned}$$

Odatle odmah slijedi $\det(A_i^{(j)}) = (-1)^{i+j}a_{ij}\det(A_{ij}^c)$. Kako isti zaključak vrijedi za svako j, razvoj po retku i je dokazan. Jer je i proizvoljan prva tvrdnja teorema je dokazana.

Razvoj po stupcima se dokazuje na analogni način, korištenjem korolara 10.16 umjesto propozicije 10.15 ili direktno korištenjem razvoja po retcima na A^T uz korištenje propozicije 10.10. ■

Korolar 10.26 *Determinanta trokutaste matrice jednaka je produktu njenih dijagonalnih elemenata*

Dokaz: Neka je $A \in \mathbf{R}^{n\times n}$ donje-trokutasta matrica. Razvojem determinante po prvom retku (ili direktnom primjenom leme 10.23) dobivamo $\det(A) = a_{11}\det(B_1)$, gdje je $B_1 = A_{11}^c$. Kako je B_1 opet donje-trokutasta, razvoj po prvom retku determinante $\det(B_1)$ daje $\det(A) = a_{11}a_{22}\det(B_2)$, gdje je $B_2 = (B_1)_{11}^c$. Nastavljajući postupak lako dobivamo $\det(A) = a_{11}\cdots a_{nn}$.

Neka je sada A gornje-trokutasta. Tada je A^T donje-trokutasta sa istim dijagonalnim elementima kao i A, pa uz pomoć propozicije 10.10 dobivamo

$$\det(A) = \det(A^T) = a_{11}\cdots a_{nn}\,.$$ ■

Računanje determinante

Vratimo se problemu računanja determinante. Koristit ćemo metodu Gaussovih eliminacija s parcijalnim pivotiranjem. Ova metoda koristi samo dvije vrste elementarnih transformacija: zamjenu redaka i dodavanje retka pomnoženog brojem, drugom retku. Uz pomoć elementarnih matrica I_{ij} i M_{ij}, postupak na matrici $A \in \mathbf{R}^{n\times n}$ možemo opisati relacijom

$$\begin{aligned} R &= \left(M_{n-1,n} I_{n-1,(n-1)'}\right) \left(M_{n-2,n} M_{n-2,n-1} I_{n-2,(n-2)'}\right) \cdots \\ &\quad \cdots (M_{1n} M_{1,n-1} \cdots M_{12} I_{11'})\, A\,. \end{aligned}$$

Ovdje je $R = A^{(n)}$, $M_{ri} = M_{ri}(m_{ir})$, a $m_{ir} = a_{ir}^{(r)}/a_{rr}^{(r)}$, $r+1 \leq i \leq n$ su multiplikatori koji se koriste u r-tom koraku eliminacija kad se poništavaju elementi r-tog stupca od $A^{(r)}$, koji leže ispod dijagonalnog elementa $a_{rr}^{(r)}$. Uzimanjem determinante u zadnjoj jednadžbi i korištenjem teorema 10.18, dobivamo

$$\det(R) = \det(I_{n-1,(n-1)'}) \det(I_{n-2,(n-2)'}) \cdots \det(I_{11'}) \det(A).$$

Općenito je u $I_{rr'}$, $r' \geq r$, pa je

$$\det(I_{rr'}) = \begin{cases} -1, & \text{ako je } r' > r \\ 1, & \text{ako je } r' = r \end{cases} .$$

Stoga je

$$\det(I_{n-1,(n-1)'}) \det(I_{n-2,(n-2)'}) \cdots \det(I_{11'}) = (-1)^{\sigma}\,,$$

gdje je σ broj pravih zamjena redaka u tijeku Gaussovih eliminacija. Kako po korolaru 10.26, za gornje-trokutastu matricu $R = (r_{ij})$ vrijedi $\det(R) = r_{11} \cdots r_{nn}$, dobili smo

$$\det(A) = (-1)^{\sigma} r_{11} \cdots r_{nn}\,.$$

Pritom je ukupno potrebno svega oko $n^3/3$ operacija množenja.

Primjedba 10.27 *Kad se metodom Gaussovih eliminacija s parcijalnim pivotiranjem A svede na trokutastu matricu R, kod računanja produkta $r_{11} \cdots r_{nn}$ na računalu, treba biti oprezan. Posebno kod većeg reda matrice, prijeti dobivanje međurezultata koji prekoračuje najveći ili potkoraćuje najmanji broj prikaziv u računalu. Neka je npr. $n = 80$ i neka je prvih 40 dijagonalnih elemenata od R jednako 0.1, dok su preostali jednaki 10. Tada je $r_{11} \cdots r_{80,80} = 1$. Ako se produkt računa od prvog dijagonalnog elementa prema zadnjem, tada se nakon 39 množenja dobije $r_{11} \cdots r_{40,40} = 10^{-40}$ broj koji je manji od najmanjeg* *prikazivog pozitivnog broja (u tzv. "jednostrukoj preciznosti") pa zato taj produkt, a onda i produkt svih dijagonalnih elemenata postaje* 0*. Ako bi produkt računali od zadnjeg prema prvom dijagonalnom elementu, računalo bi javilo grešku zbog prekoračenja najvećeg broja koje računalo prikazuje (jer je* $r_{40,40}r_{39,39}\cdots r_{21,21} = 10^{40}$*) i prekinulo bi računanje produkta. Tom problemu se može doskočiti na razne načine, od kojih je najjednostavniji (iako ne najbolji) način, koristiti formulu*

$$| \det(A) |= \exp\left(\log(|r_{11}|) + \cdots + \log(|r_{nn}|)\right). \qquad \blacksquare$$

Kad želimo bez računala izračunati determinantu, onda kombiniramo više tehnika od kojih su najčešće: razvoj determinante po pogodnom retku ili stupcu, dodavanje retka (ili stupca) pomnoženog pogodnim brojem drugom retku (stupcu) i Binet-Cauchyjev teorem.

Primjer 10.28 *Neka je*

$$V(x_1,x_2,x_3,x_4) = \begin{vmatrix} 1 & 1 & 1 & 1 \\ x_1 & x_2 & x_3 & x_4 \\ x_1^2 & x_2^2 & x_3^2 & x_4^2 \\ x_1^3 & x_2^3 & x_3^3 & x_4^3 \end{vmatrix}$$

tzv. Vandermondeova determinanta reda 4*. Da bi odredili formulu za njeno računanje, pomnožimo ju s lijeva s pogodnom matricom poznate determinante i iskoristimo Binet-Cauchyjev teorem:*

$$\begin{aligned} V(x_1,x_2,x_3,x_4) &= \begin{vmatrix} 1 & & & \\ -x_1 & 1 & & \\ & -x_1 & 1 & \\ & & -x_1 & 1 \end{vmatrix} \begin{vmatrix} 1 & 1 & 1 & 1 \\ x_1 & x_2 & x_3 & x_4 \\ x_1^2 & x_2^2 & x_3^2 & x_4^2 \\ x_1^3 & x_2^3 & x_3^3 & x_4^3 \end{vmatrix} \\ &= \begin{vmatrix} 1 & 1 & 1 & 1 \\ 0 & x_2-x_1 & x_3-x_1 & x_4-x_1 \\ 0 & x_2(x_2-x_1) & x_3(x_3-x_1) & x_4(x_4-x_1) \\ 0 & x_2^2(x_2-x_1) & x_3^2(x_3-x_1) & x_4^2(x_4-x_1) \end{vmatrix}. \end{aligned}$$

Razvojem determinante po prvom stupcu dobivamo determinantu reda tri. U toj determinanti možemo izlučiti ispred determinante faktore $(x_2 - x_1)$*,* $(x_3 - x_1)$ *i* $(x_4 - x_1)$ *koji množe 1., 2. i 3. stupac. Tako* *dobijemo*

$$\begin{aligned} V(x_1,x_2,x_3,x_4) &= (x_2-x_1)(x_3-x_1)(x_4-x_1)\begin{vmatrix} 1 & 1 & 1 \\ x_2 & x_3 & x_4 \\ x_2^2 & x_3^2 & x_4^2 \end{vmatrix} \\ &= (x_2-x_1)(x_3-x_1)(x_4-x_1)V(x_2,x_3,x_4) \end{aligned}$$

Na isti način se pokaže da općenito vrijedi

$$V(x_1,x_2,\ldots,x_n) = (x_n-x_1)(x_{n-1}-x_1)\cdots(x_2-x_1)V(x_2,\ldots,x_n)\,.$$

Specijalno, vrijedi

$$\begin{aligned} V(x_2,x_3,x_4) &= (x_3-x_2)(x_4-x_2)V(x_3,x_4) \qquad i \\ V(x_3,x_4) &= (x_4-x_3)V(x_4) = (x_4-x_3)\cdot 1 = x_4-x_3\,, \end{aligned}$$

pa je

$$V(x_1,x_2,x_3,x_4) = (x_2-x_1)(x_3-x_1)(x_4-x_1)\cdot(x_3-x_2)(x_4-x_2)\cdot(x_4-x_3)\,.$$

Općenitije, lako se pokaže da vrijedi

$$V(x_1,x_2,\ldots,x_n) = \prod_{i<j}(x_j-x_i)\,.$$

Ovaj rezultat kaže da su vektori čije su komponente (iste) potencije međusobno različitih brojeva x_1, …,x_n, linearno nezavisni. ■

10.3 Adjunkta i Cramerovo pravilo

U ovom dijelu ćemo pokazati kako se komponente rješenja sustava linearnih jednadžbi sa regularnom matricom sustava, mogu jednostavno prikazati kao kvocijenti određenih determinanti.

Definicija 10.29 **Adjunkta** *kvadratne matrice A reda n, je matrica*

$$\mathrm{adj}\,(A) = \begin{bmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & \ddots & \vdots \\ \alpha_{n1} & \cdots & \alpha_{nn} \end{bmatrix}$$

pri čemu su $\alpha_{ij} = (-1)^{j+i}\det(A^c_{ji})$ algebarski komplementi elemenata a_{ji} matrice A. ■

Adjunkta ima vrlo specijalna svojstva.

Teorem 10.30 *Ako je $A \in \mathbf{R}^{n\times n}$, tada vrijedi*

$$A \,\mathrm{adj}\,(A) = \mathrm{adj}\,(A)\, A = \det(A)\, I_n\,. \tag{10.10}$$

Dokaz: Iz razvoja determinante po i-tom retku i odnosno i-tom stupcu, dobivamo

$$\begin{aligned} [A\,\mathrm{adj}\,(A)]_{ii} &= \sum_{k=1}^{n} a_{ik}\alpha_{ki} \\ &= \sum_{k=1}^{n} a_{ik}(-1)^{i+k}\det(A_{ik}^c) = \det(A)\,, \qquad (10.11) \\ [\mathrm{adj}\,(A)\,A]_{ii} &= \sum_{k=1}^{n} \alpha_{ik}a_{ki} \\ &= \sum_{k=1}^{n} (-1)^{k+i}\det(A_{ki}^c)a_{ki} = \det(A)\,. \qquad (10.12) \end{aligned}$$

Promotrimo sada nedijagonalne elemente matrica $A\,\mathrm{adj}\,(A)$ i $\mathrm{adj}\,(A)\,A$:

$$\begin{aligned} [A\,\mathrm{adj}\,(A)]_{ij} &= \sum_{k=1}^{n} a_{ik}\alpha_{kj} = \sum_{k=1}^{n} a_{ik}(-1)^{j+k}\det(A_{jk}^c)\,, \qquad (10.13) \\ [\mathrm{adj}\,(A)\,A]_{ij} &= \sum_{k=1}^{n} \alpha_{ik}a_{kj} = \sum_{k=1}^{n} (-1)^{k+i}\det(A_{ki}^c)a_{kj}\,. \qquad (10.14) \end{aligned}$$

Uočimo da je (10.13) razvoj determinante po j-tom retku od matrice s dva jednaka retka (matrica se razlikuje od A po tome što joj je j-ti redak zamijenjen s i-tim retkom). Kako je po korolaru 10.12 takva determinanta jednaka nuli, imamo $[A\,\mathrm{adj}\,(A)]_{ij} = 0$. Ovaj zaključak vrijedi za sve $i \neq j$.

Slično, (10.14) je razvoj determinante po i-tom stupcu od matrice s dva jednaka stupca (matrica se razlikuje od A po tome što joj je i-ti stupac zamijenjen s j-tim stupcem). Kako je po korolaru 10.12 takva determinanta jednaka nuli, imamo $[\mathrm{adj}\,(A)\,A]_{ij} = 0$. Zaključak vrijedi za sve $i \neq j$.

Dokazali smo

$$[A\,\mathrm{adj}\,(A)]_{ij} = [\mathrm{adj}\,(A)\,A]_{ij} = \left\{ \begin{array}{ll} \det(A), & i = j \\ 0, & i \neq j \end{array} \right. ,$$

a to je samo drugi zapis relacije (10.10). ∎

Korolar 10.31 *Ako je A nesingularna, tada je*

$$A^{-1} = \frac{\mathrm{adj}\,(A)}{\det(A)}\,. \tag{10.15}$$

Dokaz: Ako je A nesingularna, onda je $\det(A) \neq 0$, pa jednadžbu (10.10) možemo podijeliti sa $\det(A)$ i dobiti definicijsku jednadžbu za inverznu matricu. ∎

Promotrimo sad sustav

$$Ax = b$$

gdje je $A \in \mathbf{R}^{n\times n}$ nesingularna matrica. Definirajmo matrice

$$A_i = [a_1, \ldots, a_{i-1}, b, a_{i+1}, \ldots, a_n]\,, \quad 1 \leq i \leq n\,, \tag{10.16}$$

pri čemu je $A = [a_1, \ldots, a_n]$ stupčana particija od A.

Teorem 10.32 (Cramerovo pravilo) *Neka je $A \in \mathbf{R}^{n\times n}$ nesingularna, $b \in \mathbf{R}^n$ i neka je $x = (x_i)$ rješenje sustava $Ax = b$. Tada vrijedi*

$$x_i = \frac{\det(A_i)}{\det(A)}\,, \quad 1 \leq i \leq n\,. \tag{10.17}$$

Dokaz: Razvojem $\det(A_i)$ po i-tom stupcu, u kojem se nalazi vektor b, dobivamo

$$\begin{aligned} \det(A_i) &= \sum_{k=1}^{n} b_k\,(-1)^{k+i}\det(A_{ki}) \\ &= \sum_{k=1}^{n} b_k\alpha_{ik} = \sum_{k=1}^{n}\alpha_{ik}b_k \\ &= e_i^T\mathrm{adj}\,(A)b\,, \quad 1 \leq i \leq n\,. \end{aligned} \tag{10.18}$$

Jer je A regularna, vrijedi $x = A^{-1}b$, pa je $x_i = e_i^T A^{-1}b$, $1 \leq i \leq n$. Po korolaru 10.31 i relaciji (10.18), imamo

$$\begin{aligned} \det(A)x_i &= \det(A)e_i^T A^{-1}b \\ &= \det(A)e_i^T\frac{\mathrm{adj}\,(A)}{\det(A)}b = e_i^T\mathrm{adj}\,(A)b \\ &= \det(A_i)\,, \quad 1 \leq i \leq n\,, \end{aligned}$$

odakle odmah slijedi (10.17). ∎

Primjer 10.33 *Provjerimo formule (10.17) kad je* $n = 2$*. Ako linearne jednadžbe glase*

$$\begin{aligned} a\,x_1 + b\,x_2 &= e \\ c\,x_1 + d\,x_2 &= f \end{aligned}$$

tada množenjem prve jednadžbe sa d *i druge jednadžbe sa* $-b$ *i sumiranjem dobivenih jednadžbi (eliminacija nepoznanice* x_2*), dolazimo do*

$$x_1 = \frac{ed - bf}{ad - bc} = \frac{\begin{vmatrix} e & b \\ f & d \end{vmatrix}}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}} = \frac{\det(A_1)}{\det(A)}.$$

Slično, množenjem prve jednadžbe sa $-c$ *i druge jednadžbe sa* a *i sumiranjem dobivenih jednadžbi (eliminacija nepoznanice* x_1*), dolazimo do*

$$x_2 = \frac{af - ce}{ad - bc} = \frac{\begin{vmatrix} a & e \\ c & f \end{vmatrix}}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}} = \frac{\det(A_2)}{\det(A)}.$$

■

Lekcija 11

Linearni operatori

Linearni operatori su specijalna preslikavanja između vektorskih prostora. Prisjetimo se pojma preslikavanja.

11.1 Funkcije

Neka su $\mathcal{S}$ i $\mathcal{T}$ dva prizvoljna neprazna skupa. Pridruži li se svakom elementu $x \in \mathcal{S}$ po jedan element y iz $\mathcal{T}$ dobije se **preslikavanje** skupa $\mathcal{S}$ **u** $\mathcal{T}$. Preslikavanje se još naziva **funkcija** i označava se najčešće malim slovima, npr. f, g, h, ... Ako spomenuto preslikavanje označimo sa f, tada se piše $f : \mathcal{S} \longrightarrow \mathcal{T}$. Skup $\mathcal{S}$ se zove **područje definicije funkcije** f, a skup $R(f) = \{f(x) : x \in \mathcal{S}\}$ je **područje vrijednosti funkcije** f. Ako je $y = f(x)$, x je original od y, a y je slika od x. Stoga se još kaže da je $R(f)$ slika cijelog skupa $\mathcal{S}$.

Ako je $R(f) = \mathcal{T}$ kaže se da je f preslikavanje ili funkcija skupa $\mathcal{S}$ **na** skup $\mathcal{T}$. U tom se slučaju još kaže da je f **surjekcija**. Ako za svaka dva različita elementa x_1, x_2 skupa $\mathcal{S}$ vrijedi $f(x_1) \neq f(x_2)$, preslikavanje ili funkcija je $1-1$ ili **injekcija**. Ako je je f injekcija i surjekcija tada se f zove **bijekcija**. Ako je f bijekcija, postoji funkcija $g : \mathcal{T} \longrightarrow \mathcal{S}$ sa svojstvom $g(f(x)) = x$ za sve $x \in X$ i $f(g(y)) = y$ za sve $y \in Y$. Pritom se g označava s f^{-1} i zove inverzno preslikavanje ili inverzna funkcija od f.

Dvije su funkcije jednake, ako imaju iste vrijednosti na svakom elementu iz $\mathcal{S}$. Funkcija f je **proširenje funkcije** g, a g je **restrikcija**

funkcije f, ako područje definicije funkcije f sadrži područje definicije funkcije g. Restrikcija funkcije f na podskup $\mathcal{S}'$ skupa $\mathcal{S}$ se označava sa $f\,|_{\mathcal{S}'}$. Dakle, $f\,|_{\mathcal{S}'}$ je definirano samo na $\mathcal{S}'$ i za sve $x' \in \mathcal{S}'$ vrijedi $(f\,|_{\mathcal{S}'})(x') = f(x')$.

Neka je $\mathcal{F}$ kolekcija svih preslikavanja iz $\mathcal{S}$ u $\mathcal{T}$. Ako je u $\mathcal{T}$ definirana operacija zbrajanja, npr. ako je $(\mathcal{T},+)$ grupa, tada se u skupu $\mathcal{F}$ lako definira operacija zbrajanja funkcija. Za $f,g \in \mathcal{F}$ definiramo $f+g$ kao funkciju h za koju vrijedi

$$h(x) = f(x) + g(x) \quad \text{za sve } x \in \mathcal{S}. \tag{11.1}$$

Odmah se vidi da je $h \in \mathcal{F}$ čim je $\mathcal{T}$ zatvoren u odnosu na zbrajanje. Jednostavno je pokazati da se svojstva asocijativnosti i komutativnosti zbrajanja u $\mathcal{T}$ prenose na operaciju zbrajanja funkcija. Specijalno, ako je $(\mathcal{T},+)$ grupa, tada se lako pokaže da je i $(\mathcal{F},+)$ grupa (neutralni element je preslikavanje koje $\mathcal{S}$ prevodi u neutralni element u grupi $(\mathcal{T},+)$, a suprotni element od f je $-f$ koja je definirana s $(-f)(x) = -f(x)$ za sve $x \in \mathcal{S}$. Ako je u $\mathcal{T}$ definirana operacija množenja sa skalarom, tada se formulom

$$(\alpha f)(x) = \alpha f(x) \quad \text{za sve } x \in \mathcal{S} \tag{11.2}$$

definira množenje funkcije f skalarom α. Tom formulom se glavna svojstva množenja prenose sa $\mathcal{T}$ na $\mathcal{F}$. Specijalno, ako je $(\mathcal{T},+,\cdot)$ vektorski prostor jednostavno je provjeriti da uz definirane operacije zbrajanja funkcija i množenja funkcija skalarom, $(\mathcal{F},+,\cdot)$ postaje vektorski prostor.

Ako su $\mathcal{S}$, $\mathcal{T}$ i $\mathcal{Z}$ skupovi i $f : \mathcal{S} \longrightarrow \mathcal{T}$, $g : \mathcal{T} \longrightarrow \mathcal{Z}$ preslikavanja, tada se može definirati kompozicija $g \circ f$ preslikavanja f i g kao preslikavanje $h : \mathcal{S} \longrightarrow \mathcal{Z}$, za koje vrijedi

$$h(x) = g(f(x)), \quad x \in \mathcal{S}\,.$$

Ako su f i g injekcije (surjekcije, bijekcije) takva je i kompozicija $g \circ f$. Ako je $\mathcal{S} = \mathcal{T} = \mathcal{Z}$, onda se u $\mathcal{F}$ pored zbrajanja funkcija i množenja funkcija skalarom, uvodi i množenje funkcija kao njihova kompozicija. Tada se obično umjesto $g \circ f$ piše gf.

11.2 Linearna preslikavanja

Posebno je važan slučaj kada su skupovi $\mathcal{S}$ i $\mathcal{T}$ vektorski prostori. U toj situaciji možemo u skupu $\mathcal{F}$ izdvojiti preslikavanja koja linearne kombinacije vektora originala prevode u linearne kombinacije vektora slika.

Definicija 11.1 *Neka su X i Y vektorski prostori. Preslikavanje $\mathcal{A}: X \longrightarrow Y$ je* **linearni operator** *ako vrijedi*

(i) $\mathcal{A}(x+y) = \mathcal{A}(x) + \mathcal{A}(y)$ *za sve $x, y \in X$ (aditivnost)*

(ii) $\mathcal{A}(\alpha x) = \alpha \mathcal{A}(x)$ *za sve $x \in X$ i sve skalare α (homogenost).*

Ako je $\mathcal{A}$ bijekcija, tada se još zove **izomorfizam** *vektorskih prostora X i Y.* ■

Iz svojstva (ii) se vidi da X i Y moraju biti definirani nad istim poljem. Ako su X i Y realni vektorski prostori, onda α može biti samo realni broj. Ako su X i Y kompleksni vektorski prostori, onda α može biti i realni i kompleksni broj. Kad god ćemo koristiti linearni operator implicitno ćemo pretpostavljati da su polazni i dolazni vektorski prostori definirani nad istim poljem.

Iz svojstva (ii) odmah slijedi, uzimanjem $\alpha = 0$, $\mathcal{A}(0) = 0$.

Primjer 11.2 *Neka je $X = \mathbf{R}^n$, $Y = \mathbf{R}^m$ i neka je $A \in \mathbf{R}^{m\times n}$ proizvoljna matrica. Preslikavanje $\mathcal{A}: X \longrightarrow Y$ definirano formulom*

$$\mathcal{A}(x) = Ax, \quad x \in \mathbf{R}^n$$

je linearni operator. Provjerimo prvo da je $\mathcal{A}(x)$ definirano za svako $x \in \mathbf{R}^n$ i da je $\mathcal{A}(x) \in \mathbf{R}^m$. To je jasno jer je produkt $m \times n$ i $n \times 1$ matrica uvijek definiran i daje $m \times 1$ matricu kao rezultat. $\mathcal{A}$ je linearan operator jer je

$$\begin{aligned} \mathcal{A}(x+y) &= A(x+y) = Ax + Ay = \mathcal{A}(x) + \mathcal{A}(y) \quad \textit{za sve } x, y \in \mathbf{R}^n \\ \mathcal{A}(\alpha x) &= A(\alpha x) = \alpha Ax = \alpha \mathcal{A}(x) \quad \textit{za sve } x \in \mathbf{R}^n,\ \alpha \in \mathbf{R}\,. \end{aligned}$$

■

Uvjeti (i) i (ii) iz definicije 11.1 mogu se zamijeniti jednim uvjetom

(l) $\mathcal{A}(\alpha x + \beta y) = \alpha \mathcal{A}(x) + \beta \mathcal{A}(y)$ za sve $x, y \in X$ i sve skalare α, β,

koji zovemo *linearnost*.
Pokažimo prvo da aditivnost i homogenost povlače linearnost.

$$\begin{aligned} \mathcal{A}(\alpha x + \beta y) &= \mathcal{A}(\alpha x) + \mathcal{A}(\beta y) \qquad \text{jer vrijedi (i)} \\ &= \alpha \mathcal{A}(x) + \beta \mathcal{A}(y) \qquad \text{jer vrijedi (ii).} \end{aligned}$$

Pokažimo da linearnost preslikavanja povlači aditivnost i homogenost. Doista, uvrštavajući u (l) $\alpha = \beta = 1$ dobivamo (i). Uzimajući u (l) $\beta = 0$, dobivamo

$$\begin{aligned} \mathcal{A}(\alpha x) = \mathcal{A}(\alpha x + 0y) &= \alpha \mathcal{A}(x) + 0\mathcal{A}(y) = \alpha \mathcal{A}(x) + 0 \\ &= \alpha \mathcal{A}(x)\,, \end{aligned}$$

za sve x i sve skalare α. Time je pokazano da (l) povlači (ii). Dakle su uvjeti (i) i (ii) ekvivalentni (tj. znače isto) uvjetu (l). Iz uvjeta (l) se vidi da je linearni operator ono preslikavanje koje linearni spoj vektora iz X prevodi u isti linearni spoj slika u Y. Sljedeća propozicija pokazuje da linearni operator pevodi linearnu kombinaciju proizvoljnog broja vektora u istu linearnu kombinaciju slika. Dokaz je tek vježba za korištenje matematičke indukcije.

Propozicija 11.3

$$\mathcal{A}(\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_k x_k) = \alpha_1 \mathcal{A}(x_1) + \alpha_2 \mathcal{A}(x_2) + \cdots + \alpha_k \mathcal{A}(x_k)\,.$$

Dokaz: Dokaz koristi matematičku indukciju po broju vektora u linearnoj kombinaciji k. Za $k = 2$ je tvrdnja dokazana. Pretpostavimo da je istinita za neko $k \geq 2$ pa pokažimo datada vrijedi za $k + 1$. Imamo

$$\begin{aligned} &\mathcal{A}(\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_{k+1} x_{k+1}) = \\ &\quad \mathcal{A}[(\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_k x_k) + \alpha_{k+1} x_{k+1}] \\ &\quad = \mathcal{A}(\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_k x_k) + \mathcal{A}(\alpha_{k+1} x_{k+1}) \\ &\quad = [\alpha_1 \mathcal{A}(x_1) + \alpha_2 \mathcal{A}(x_2) + \cdots + \alpha_k \mathcal{A}(x_k)] + \alpha_{k+1} \mathcal{A}(x_{k+1}) \\ &\quad = \alpha_1 \mathcal{A}(x_1) + \alpha_2 \mathcal{A}(x_2) + \cdots + \alpha_k \mathcal{A}(x_k) + \alpha_{k+1} \mathcal{A}(x_{k+1})\,, \end{aligned}$$

pri čemu smo prvo iskoristili asocijativnost zbrajanja u Y, zatim aditivnost operatora $\mathcal{A}$, pa indukcijsku pretpostavku, te homogenost operatora i opet asocijativnost zbrajanja u Y. ■

Uz svaki linearni operator $\mathcal{A}: X \mapsto Y$, vezana su dva važna potprostora,

$$\mathcal{N}(\mathcal{A}) = \{x \in X; \mathcal{A}(x) = 0\} \quad \text{i} \quad \mathcal{R}(\mathcal{A}) = \{\mathcal{A}(x); x \in X\}.$$

Dokažite za vježbu da su $\mathcal{N}(\mathcal{A})$ i $\mathcal{R}(\mathcal{A})$ potprostori. $\mathcal{N}(\mathcal{A})$ je **nul-potprostor** (ili **jezgra**), a $\mathcal{R}(\mathcal{A})$ je **slika** (ili **područje vrijednosti**) operatora $\mathcal{A}$. Dimenzija jezgre se zove **defekt**, a dimenzija slike **rang** operatora $\mathcal{A}$.

Sa $\mathcal{L}(X, Y)$ označavamo skup svih linearnih preslikavanja vektorskog prostora X u vektorski prostor Y.

11.3 Izomorfizmi

Posebno važnu klasu linearnih operatora čine izomorfizmi.To su linearni operatori koji su bijekcije. Ako je $\mathcal{A} \in \mathcal{L}(X, Y)$ izomorfizam, tada postoji inverzno preslikavanje $\mathcal{A}^{-1}$. Sljedeća propozicija pokazuje da je $\mathcal{A}^{-1}$ također linearni operator (tj. izomorfizam), a također da izomorfizam preslikava bazu polaznog prostora u bazu dolaznog prostora.

Propozicija 11.4 *Ako je $\mathcal{A} \in \mathcal{L}(X, Y)$ izomorfizam, tada vrijede sljedeće tvrdnje*

(i) $\mathcal{A}^{-1} \in \mathcal{L}(Y, X)$ *je izomorfizam,*

(ii) *skup vektora $\{u_1, \ldots, u_k\}$ je linearno nezavisan ako i samo ako je skup slika $\{\mathcal{A}(u_1), \ldots, \mathcal{A}(u_k)\}$ linearno nezavisan.*

Dokaz: (i) Jer je $\mathcal{A}$ bijekcija, postoji inverz $\mathcal{A}^{-1}$ koji je također bijekcija. Preostaje pokazati da je $\mathcal{A}^{-1}$ linearan operator. Neka su $y, z \in Y$ te α, β skalari. Neka je $u = \mathcal{A}^{-1}(y)$, $v = \mathcal{A}^{-1}(z)$ i $x = \mathcal{A}^{-1}(\alpha y + \beta z)$. Iz zadnje jednadžbe slijedi

$$\begin{aligned}
\mathcal{A}(x) &= \mathcal{A}\left(\mathcal{A}^{-1}(\alpha y + \beta z)\right) = \alpha y + \beta z \\
&= \alpha \mathcal{A}(u) + \beta \mathcal{A}(v) \\
&= \mathcal{A}(\alpha u + \beta v)\,.
\end{aligned}$$

Jer je $\mathcal{A}$ injekcija, zaključujemo da je $x = \alpha u + \beta v$, tj.

$$\mathcal{A}^{-1}(\alpha y + \beta z) = \alpha \mathcal{A}^{-1}(y) + \beta \mathcal{A}^{-1}(z)\,.$$

(ii) Neka su $u_1, \ldots, u_k$ linearno nezavisni. Pođimo od jednadžbe

$$\alpha_1 \mathcal{A}(u_1) + \ldots + \alpha_k \mathcal{A}(u_k) = 0$$

i pokažimo da tada mora vrijediti $\alpha_1 = \cdots = \alpha_k = 0$. Zbog linearnosti operatora $\mathcal{A}$, ona se može zapisati u obliku

$$\mathcal{A}\left(\alpha_1 u_1 + \ldots + \alpha_k u_k\right) = 0\,.$$

Jer je $\mathcal{A}$ injekcija, mora vrijediti $\alpha_1 u_1 + \ldots + \alpha_k u_k = 0$. Sada zbog pretpostavljene linearne nezavisnosti vektora $u_1, \ldots, u_k$ zaključujemo da je $\alpha_1 = \cdots = \alpha_k = 0$.

Obrnuti smjer se dokazuje još jednostavnije. Sada je pretpostavka da su vektori $\mathcal{A}(u_1), \ldots, \mathcal{A}(u_k)$ linearno nezavisni, a pokazuje se da $\alpha_1 u_1 + \ldots + \alpha_k u_k = 0$ povlači $\alpha_1 = \cdots = \alpha_k = 0$.

Primijenimo na obje strane jednadžbe $\alpha_1 u_1 + \ldots + \alpha_k u_k = 0$ operator $\mathcal{A}$. Dobivamo

$$0 = \mathcal{A}(0) = \mathcal{A}(\alpha_1 u_1 + \ldots + \alpha_k u_k) = \alpha_1 \mathcal{A}(u_1) + \cdots + \alpha_k \mathcal{A}(u_k)\,.$$

Jer su $\mathcal{A}(u_1), \ldots, \mathcal{A}(u_k)$ linearno nezavisni, mora biti $\alpha_1 = \cdots = \alpha_k = 0$. ■

Kako razlikovati izomorfizam od općeg linearnog operatora? Koje svojstvo ga izdvaja? Sljedeća propozicija daje jedno takvo svojstvo.

Propozicija 11.5 *Neka je* $\dim(X) = \dim(Y)$ *i* $\mathcal{A} \in \mathcal{L}(X, Y)$. $\mathcal{A}$ *je izomorfizam onda i samo onda ako je* $\mathcal{N}(\mathcal{A}) = \{0\}$.

Dokaz: Ako je $\mathcal{A}$ izomorfizam, on je injekcija. Znamo da je uvijek $\mathcal{A}(0) = 0$. Jer je $\mathcal{A}$ injekcija, nema drugih vektora koje bi $\mathcal{A}$ prebacivao u nulu, pa je $\mathcal{N}(\mathcal{A}) = \{0\}$.

Obratno, pretpostavimo da je $\mathcal{N}(\mathcal{A}) = \{0\}$. Kad $\mathcal{A}$ ne bi bio izomorfizam, ne bi bio bijekcija, pa ne bi bio injekcija ili ne bi bio surjekcija.

Pretpostavimo da $\mathcal{A}$ nije injekcija. Tada postoje vektori $u, v \in X$, takvi da je $u \neq v$ i $\mathcal{A}(u) = \mathcal{A}(v)$. Zbog linearnosti operatora imamo $\mathcal{A}(u-v) = 0$, pa je $u - v \neq 0$ u jezgri od $\mathcal{A}$, a to ne može biti. Dakle, moramo odbaciti pretpostavku da $\mathcal{A}$ nije injekcija.

Pokažimo da je $\mathcal{A}$ surjekcija. Već smo pokazali da $\mathcal{A}$ mora biti injekcija. Stoga je on izomorfizam vektorskih prostora X i $\mathcal{R}(\mathcal{A})$. Kako po prethodnoj propoziciji izomorfizam prebacuje svaku bazu u X u neku bazu u $\mathcal{R}(\mathcal{A})$, mora dimenzija od $\mathcal{R}(\mathcal{A})$ biti n. Jer je $\mathcal{R}(\mathcal{A}) \subseteq Y$ i jer $\mathcal{R}(\mathcal{A})$ i Y imaju istu dimenziju n, mora biti $\mathcal{R}(\mathcal{A}) = Y$. Dakle je $\mathcal{A}$ surjekcija.

Pokazano je da je $\mathcal{A}$ injekcija i surjekcija pa je bijekcija. ∎

Izomorfizam je relacija ekvivalencije (ili klasifikacije) u skupu vektorskih prostora. Kasnije ćemo konstruirati važnu klasu izomorfizama između vektorskog prostora dimenzije n i vektorskog prostora jednostupčanih matrica $\mathbf{R}^n$ (ili $\mathbf{C}^n$). To će omogućiti svađanje mnogih problema vezanih uz opće vektorske prostore i operatore između njih, na matrične probleme.

11.4 Vektorski prostor $\mathcal{L}(X,Y)$

Označimo sa $\mathcal{L}(X,Y)$ skup svih linearnih operatora koji vektorski prostor X preslikavaju u vektorski prostor Y. Jer je Y vektorski prostor u $\mathcal{L}(X,Y)$ se pomoću relacija (11.1) i (11.2) uvodi operacija zbrajanja linearnih operatora i operacija množenja linearnog operatora sa skalarom. Sada se relacije (11.1) i (11.2) zapisuju kao

$$(\mathcal{A}+\mathcal{B})(x) = \mathcal{A}(x) + \mathcal{B}(x), \quad x \in X. \tag{11.3}$$

i

$$(\gamma\mathcal{A})(x) = \gamma\mathcal{A}(x), \quad x \in X \tag{11.4}$$

Sljedeće dvije relacije pokazuju da su $\mathcal{A}+\mathcal{B}$ i $\gamma\mathcal{A}$ linearni operatori. Za u, $v \in X$ i α, $\beta \in \Phi$ (Φ polje nad kojim su građeni X i Y), imamo

$$\begin{aligned}(\mathcal{A}+\mathcal{B})\,(\alpha u + \beta v) &= \mathcal{A}(\alpha u + \beta v) + \mathcal{B}(\alpha u + \beta v) \\ &= \alpha\mathcal{A}(u) + \beta\mathcal{A}(v) + (\alpha\mathcal{B}(u) + \beta\mathcal{B}(v)\end{aligned}$$

$$
\begin{aligned}
&= \alpha\,(\mathcal{A}+\mathcal{B})\,(u)+\beta\,(\mathcal{A}+\mathcal{B})\,(v)\,,\\
(\gamma\mathcal{A})\,(\alpha u+\beta v) &= \gamma\mathcal{A}(\alpha u+\beta v)\\
&= \gamma\,(\alpha\mathcal{A}(u)+\beta\mathcal{A}(v))\\
&= \alpha\gamma\mathcal{A}(u)+\beta\gamma\mathcal{A}(v)\\
&= \alpha\,(\gamma\mathcal{A})\,(u)+\beta\,(\gamma\mathcal{A})\,(v)\,.
\end{aligned}
$$

Teorem 11.6 $(\mathcal{L}(X,Y),+,\cdot)$ *je vektorski prostor.*

Dokaz: Pokažimo prvo da je $\mathcal{L}(X,Y)$ zatvoren s obzirom na spomenute operacije. Neka su $\mathcal{A}$ i $\mathcal{B}$ iz $\mathcal{L}(X,Y)$. Tada je za svako $x\in X$, $\mathcal{A}(x)\in Y$ i $\mathcal{B}(x)\in Y$, pa je zbog zatvorenosti skupa Y s obzirom na zbrajanje vektora, $\mathcal{A}(x)+\mathcal{B}(x)\in Y$. Stoga je zbrajanje linearnih operatora $\mathcal{A}+\mathcal{B}$ dobro definirano formulom (11.3).

Isto tako, za proizvoljni $\gamma\in\Phi$ i proizvoljni $\mathcal{A}\in\mathcal{L}(X,Y)$ je $\gamma\mathcal{A}(x)\in Y$ jer je Y vektorski prostor. Stoga je formulom (11.4) dobro definiran produkt linearnog operatora sa skalarom.

Između svih linearnih operatora iz $\mathcal{L}(X,Y)$ posebno se ističe onaj koji cijeli X prebacuje u nul-vektor od Y, $\mathcal{O}(x)=0\in Y$ za sve $x\in X$. Za njega vrijedi $\mathcal{O}+\mathcal{A}=\mathcal{A}+\mathcal{O}=\mathcal{A}$, za sve $\mathcal{A}\in\mathcal{L}(X,Y)$. Zaključujemo da je $\mathcal{O}$ neutralni element u $(\mathcal{L}(X,Y),+)$.

Svojstva komutativnosti i asocijativnosti proizlaze iz istih svojstava komutativne grupe $(Y,+)$,

$$
\begin{aligned}
(\mathcal{A}+\mathcal{B})(x) &= \mathcal{A}(x)+\mathcal{B}(x)=\mathcal{B}(x)+\mathcal{A}(x)\\
&= (\mathcal{B}+\mathcal{A})(x), \quad x\in X\\
[(\mathcal{A}+\mathcal{B})+\mathcal{C}](x) &= (\mathcal{A}+\mathcal{B})(x)+\mathcal{C}(x)=(\mathcal{A}(x)+\mathcal{B}(x))+\mathcal{C}(x)\\
&= \mathcal{A}(x)+(\mathcal{B}(x)+\mathcal{C}(x))=\mathcal{A}(x)+(\mathcal{B}+\mathcal{C})(x)\\
&= [\mathcal{A}+(\mathcal{B}+\mathcal{C})](x), \quad x\in X\,.
\end{aligned}
$$

Konačno, za $\mathcal{A}\in\mathcal{L}(X,Y)$ postoji linearni operator $-\mathcal{A}$, koji je definiran s $(-\mathcal{A})(x)=-\mathcal{A}(x)$ za sve $x\in X$. On ima svostvo da je

$$
[(-\mathcal{A})+\mathcal{A}](x)=-\mathcal{A}(x)+\mathcal{A}(x)=0, \quad x\in X\,,
$$

pa je $-\mathcal{A}$ inverzni element za $\mathcal{A}$ s obzirom na operaciju zbrajanja. Time je pokazano da je $(\mathcal{L}(X,Y),+)$ aditivna Abelova grupa.

Ostala svojstva koja se tiču množenja skalarom, dokazuju se jednako lako. Npr.,

$$[(\alpha\beta)\mathcal{A}](x) = (\alpha\beta)\mathcal{A}(x) = \alpha[\beta\mathcal{A}(x)] = [\alpha(\beta\mathcal{A})](x),\ x \in X$$

pokazuje da vrijedi $(\alpha\beta)\mathcal{A} = \alpha(\beta\mathcal{A})$ za $\alpha, \beta \in \Phi$ i $\mathcal{A} \in \mathcal{L}(X, Y)$. Spomenimo na kraju da $1 \in \Phi$ ima svojstvo $1 \cdot \mathcal{A} = \mathcal{A}$. ■

Uočimo da je $(\mathcal{L}(X, Y), +, \cdot)$ potprostor vektorskog prostora svih preslikavanja iz X u Y. Ako je $X = Y$ u $(\mathcal{L}(X, X), +, \cdot)$ se može uvesti i produkt linearnih preslikavanja kao njihova kompozicija.

11.5 Koordinatizacija

Neka je X n-dimenzionalan vektorski prostor. Odaberimo neku bazu $e = (e_1, \ldots, e_n)$ u X. Tada se svaki vektor $x \in X$ može na jednoznačan način napisati u obliku

$$x = x_1 e_1 + x_2 e_2 + \cdots + x_n e_n\,, \tag{11.5}$$

pri čemu su $x_1, \ldots, x_n$ skalari iz polja Φ nad kojim je definiran X. Neka je npr. $\Phi = \mathbf{R}$. Tada je relacijom (11.5) definirano preslikavanje koje svakom vektoru $x \in X$ pridružuje $x(e) \in \mathbf{R}^n$,

$$\mathcal{J} : x \longrightarrow x(e) = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Teorem 11.7 *$\mathcal{J}$ je izomorfizam prostora X i $\mathbf{R}^n$.*

Dokaz: Pokažimo da je preslikavanje $\mathcal{J}$ linearno. Neka su $x, y \in X$ i $\alpha, \beta \in \mathbf{R}$ proizvoljni. Jer je

$$\begin{aligned} \alpha x + \beta y &= \alpha(x_1 e_1 + \cdots + x_n e_n) + \beta(y_1 e_1 + \cdots + y_n e_n) \\ &= (\alpha x_1 + \beta y_1)e_1 + \cdots + (\alpha x_n + \beta y_n)e_n\,, \end{aligned}$$

po definiciji od $\mathcal{J}$, imamo

$$\begin{aligned}
\mathcal{J}(\alpha x+\beta y) &= \begin{bmatrix} \alpha x_1+\beta y_1 \\ \vdots \\ \alpha x_n+\beta y_n \end{bmatrix} = \begin{bmatrix} \alpha x_1 \\ \vdots \\ \alpha x_n \end{bmatrix} + \begin{bmatrix} \beta y_1 \\ \vdots \\ \beta y_n \end{bmatrix} \\
&= \alpha \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + \beta \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \alpha x(e)+\beta y(e) \\
&= \alpha \mathcal{J}(x)+\beta \mathcal{J}(y),
\end{aligned}$$

pa je $\mathcal{J}$ linearni operator iz X u $\mathbf{R}^n$. Ovdje smo iskoristili svojstva zbrajanja i množenja skalarom vektora u $\mathbf{R}^n$.

Pokažimo da je $\mathcal{J}$ surjekcija. Uzmimo proizvoljnu jednostupčanu matricu $[x_1,\ldots,x_n]^T$iz $\mathbf{R}^n$. Iz skalara x_i, $1\leq i\leq n$ formirajmo vektor $x_1e_1+\cdots+x_ne_n=x$ iz X. Iz definicije od $\mathcal{J}$ vidimo da je $\mathcal{J}(x)$ upravo polazna jednostupčana matrica, pa je $\mathcal{J}$ preslikavanje na $\mathbf{R}^n$.

Pokažimo da je $\mathcal{J}$ injekcija, tj. da međusobno različite vektore iz X prebacuje u međusobno različite jednostupčane matrice. Neka su x i y iz X različiti. Tada u razvoju po bazi e (kao u relaciji (11.5)), postoji barem jedna vrijednost od i za koju su skalari x_i i y_i međusobno različiti. To onda znači i da su $x(e)$ i $y(e)$ međusobno različiti. Dakle je $\mathcal{J}$ bijektivno linearno preslikavanje pa je izomorfizam vektorskih prostora X i $\mathbf{R}^n$. ■

Isti zaključak vrijedi i u slučaju kompleksnog vektorskog prostora X, s tim da je tada $\mathcal{J}$ izomorfizam prostora X i $\mathbf{C}^n$.

Funkciju $\mathcal{J}$ ćemo zvati **koordinatizacija** vektorskog prostora X, a vektor $x(e)$ ćemo zvati **koordinatni vektor** od $x\in X$. Jasno je da koordinatizacija ovisi isključivo o (izabranoj) bazi, a kako svaki vektorski prostor ima beskonačno mnogo baza, toliko ima koordinatizacija.

11.6 Matrica kao zapis operatora

Neka je X n-dimenzionalni, Y m-dimenzionalni vektorski prostor i $\mathcal{A}: X\to Y$ linearni operator. Neka su $e=(e_1,\ldots,e_n)$ i $f=(f_1,\ldots,f_m)$ baze u X i Y, respektivno. Uz njih su vezane koordinatizacije $\mathcal{J}_X$ i

$\mathcal{J}_Y$ prostora X i Y. Ako su prostori realni, to znači da su $x(e) = \mathcal{J}_X(x) \in \mathbf{R}^n$ i $y(f) = \mathcal{J}_Y(y) \in \mathbf{R}^m$ za $x \in X$ i $y \in Y$. Postavlja se pitanje koji je odnos između $x(e)$ i $y(f)$ ako je $y = \mathcal{A}(x)$?

Jer je $\mathcal{A}$ linearan operator, imamo

$$\mathcal{A}(x) = \mathcal{A}(x_1 e_1 + \cdots + x_n e_n) = x_1 \mathcal{A}(e_1) + \cdots + x_n \mathcal{A}(e_n). \tag{11.6}$$

Specijalno, iz relacije (11.6) se vidi da je operator $\mathcal{A}$ zadan ako znamo njegove vrijednosti na vektorima (bilo koje) baze.

Svaki $\mathcal{A}(e_j)$ leži u Y pa se može razviti po bazi f. Dakle, za svako j postoje skalari $\alpha_{1j}, \ldots, \alpha_{mj}$, takvi da je

$$\mathcal{A}(e_j) = \alpha_{1j} f_1 + \alpha_{2j} f_2 + \cdots + \alpha_{mj} f_m\,, \quad 1 \leq j \leq n\,. \tag{11.7}$$

Vidimo da smo skalare α_{ij} tako numerirali da vrijedi

$$\mathcal{J}_Y\left(\mathcal{A}(e_j)\right) = \begin{bmatrix} \alpha_{1j} \\ \alpha_{2j} \\ \vdots \\ \alpha_{mj} \end{bmatrix}, \quad 1 \leq j \leq n\,. \tag{11.8}$$

Polazeći od vektora $\mathcal{A}(x)$ i koristeći linearnost operatora $\mathcal{A}$, uz pomoć relacija (11.6)–(11.7) dobivamo

$$\begin{aligned} \mathcal{A}(x) &= \sum_{j=1}^{n} x_j \mathcal{A}(e_j) = \sum_{j=1}^{n} x_j \sum_{k=1}^{m} \alpha_{kj} f_k \\ &= \sum_{j=1}^{n} \sum_{k=1}^{m} \alpha_{kj} x_j f_k = \sum_{k=1}^{m} \sum_{j=1}^{n} \alpha_{kj} x_j f_k \\ &= \sum_{k=1}^{m} \left(\sum_{j=1}^{n} \alpha_{kj} x_j \right) f_k\,. \end{aligned} \tag{11.9}$$

Iz jednadžbe $y = \mathcal{A}(x)$ slijedi $y(f) = \mathcal{J}_Y(y) = \mathcal{J}_Y(\mathcal{A}(x))$, pa koristeći relaciju (11.9), dobivamo

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} = \begin{bmatrix} \alpha_{11} x_1 + \alpha_{12} x_2 + \cdots + \alpha_{1n} x_n \\ \alpha_{21} x_1 + \alpha_{22} x_2 + \cdots + \alpha_{2n} x_n \\ \cdots \\ \alpha_{m1} x_1 + \alpha_{m2} x_2 + \cdots + \alpha_{mn} x_n \end{bmatrix}$$

$$= \begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ & & \cdots & \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

ili

$$y(f) = \mathcal{A}(f,e)x(e)\,, \tag{11.10}$$

gdje je

$$\mathcal{A}(f,e) = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ & & \cdots & \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} \end{bmatrix} \in \mathbf{R}^{m\times n} \tag{11.11}$$

matrica čiji je j-ti stupac upravo $\mathcal{J}_Y\left(\mathcal{A}(e_j)\right)$ (vidi relaciju (11.8)). Svakom linearnom operatoru $\mathcal{A} : X \to Y$ u paru baza e, f pripada matrica $\mathcal{A}(f,e)$, koju zovemo **zapis** ili **reprezentacija** operatora $\mathcal{A}$ u paru baza e, f. Pritom vrijedi dijagram na strani 184.

$$\begin{tikzcd} X \arrow[r, "\mathcal{A}"] \arrow[d, leftrightarrow, "\mathcal{J}_X"'] & Y \arrow[d, leftrightarrow, "\mathcal{J}_Y"] \\ \mathbf{R}^n \arrow[r, "\mathcal{A}(f,e)"] & \mathbf{R}^m \end{tikzcd}$$

Fiksirajmo sada baze e i f i promotrimo skup linearnih operatora $\mathcal{L}(X,Y)$. Svakom linearnom operatoru $\mathcal{A} \in \mathcal{L}(X,Y)$ pripada matrica $A = \mathcal{A}(f,e)$, pa je time definirano preslikavanje $\mathcal{J}_{\mathcal{L}} : \mathcal{L}(X,Y) \to \mathbf{R}^{m\times n}$.

Teorem 11.8 *$\mathcal{J}_{\mathcal{L}}$ je izomorfizam vektorskih prostora $\mathcal{L}(X,Y)$ i $\mathbf{R}^{m\times n}$.*

Dokaz: Već smo pokazali da je $\mathcal{J}_{\mathcal{L}}$ definiran za svaki linearni operator i da ima vrijednosti u $\mathbf{R}^{m\times n}$. Moramo pokazati linearnost i bijektivnost tog preslikavanja.

Neka su $\mathcal{A}, \mathcal{B} \in \mathcal{L}(X,Y)$ i $\alpha, \beta \in \mathbf{R}$ proizvoljni. Neka je $\mathcal{J}_{\mathcal{L}}(\mathcal{A}) = A = (a_{ij})$ i $\mathcal{J}_{\mathcal{L}}(\mathcal{B}) = B = (b_{ij})$. Uočimo da je $(\alpha\mathcal{A} + \beta\mathcal{B})(e_j) \in Y$. Stoga je prema relaciji (11.8) $\mathcal{J}_Y\left((\alpha\mathcal{A} + \beta\mathcal{B})(e_j)\right)$ j-ti stupac matrice $(\alpha\mathcal{A} + \beta\mathcal{B})(f,e)$. Koristeći relacije (11.3), (11.4) i (11.8), dobivamo

$$\begin{aligned}
\mathcal{J}_Y\left((\alpha\mathcal{A} + \beta\mathcal{B})(e_j)\right) &= \mathcal{J}_Y\left(\alpha\mathcal{A}(e_j) + \beta\mathcal{B}(e_j)\right) \\
&= \alpha\mathcal{J}_Y\left(\mathcal{A}(e_j)\right) + \beta\mathcal{J}_Y\left(\mathcal{B}(e_j)\right) \\
&= \alpha \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix} + \beta \begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{mj} \end{bmatrix} = \alpha A\mathbf{e}_j + \beta B\mathbf{e}_j \\
&= (\alpha A + \beta B)\mathbf{e}_j, , \quad 1 \le j \le n \,,
\end{aligned}$$

gdje su[1] $\mathbf{e}_j$, $1 \le j \le n$ stupci jedinične matrice I_n. Kako je prema matričnom množenju $(\alpha A + \beta B)\mathbf{e}_j$ upravo j-ti stupac od $\alpha A + \beta B$, zaključujemo da je

$$\mathcal{J}_{\mathcal{L}}\left(\alpha\mathcal{A} + \beta\mathcal{B}\right) = \alpha A + \beta B = \alpha\mathcal{J}_{\mathcal{L}}(\mathcal{A}) + \beta\mathcal{J}_{\mathcal{L}}(\mathcal{B})\,.$$

Za dokaz surjektivnosti, uzmimo proizvoljnu matricu $A \in \mathbf{R}^{m\times n}$, $A = (\alpha_{ij})$ i pokažimo da postoji linearni operator $\mathcal{A} \in \mathcal{L}(X,Y)$, za koji vrijedi (11.8). Vidjeli smo da je za definiranje linearnog operatora dovoljno znati njegove vrijednosti na vektorima baze. Definirajmo stoga linearni operator $\mathcal{A}$ pomoću elemenata matrice A na sljedeći način

$$\mathcal{A}(e_j) = \alpha_{1j}f_1 + \cdots + \alpha_{mj}f_m, \quad 1 \le j \le n\,. \tag{11.12}$$

Operator $\mathcal{A} \in \mathcal{L}(X,Y)$ je dobro definiran relacijama (11.6) i (11.12). Prema relacijama (11.8) i (11.11), vrijedi $A = \mathcal{A}(f,e) = \mathcal{J}_{\mathcal{L}}(\mathcal{A})$.

Konačno, pokažimo da je $\mathcal{J}_{\mathcal{L}}$ injekcija. Neka za $\mathcal{A}, \mathcal{B} \in \mathcal{L}(X,Y)$ vrijedi $\mathcal{A} \neq \mathcal{B}$. To znači da postoji e_j – vektor baze u X, za koji je $\mathcal{A}(e_j) \neq \mathcal{B}(e_j)$. Najme, u protivnom bi relacija (11.6) implicirala $\mathcal{A}(x) = \mathcal{B}(x)$ za sve $x \in X$, a to bi značilo $\mathcal{A} = \mathcal{B}$. Jer je $\mathcal{J}_Y$ bijekcija, $\mathcal{A}(e_j) \neq \mathcal{B}(e_j)$ povlači $\mathcal{J}_Y\left(\mathcal{A}(e_j)\right) \neq \mathcal{J}_Y\left(\mathcal{B}(e_j)\right)$. Koristeći relacije (11.8) i (11.11) odmah se vidi da se matrice $\mathcal{A}(f,e) = \mathcal{J}_{\mathcal{L}}(\mathcal{A})$ i $\mathcal{B}(f,e) = \mathcal{J}_{\mathcal{L}}(\mathcal{B})$ razlikuju u j-tom stupcu. ■

[1]u ovom poglavlju ćemo vektore kanonske baze u $\mathbf{R}^n$ (i $\mathbf{R}^m$) označiti masnim slovima kako bi ih razlikovali od vektora baze e u X.

Kako je dimenzija vektorskog prostora $\mathbf{R}^{m\times n}$ jednaka mn, iz teorema 11.8 i propozicije 11.4 zaključujemo da je dimenzija od $\mathcal{L}(X,Y)$ također mn. Lako se vidi da je npr. $\mathcal{E} = (\mathcal{E}_{11},\ldots,\mathcal{E}_{mn})$,

$$\mathcal{E}_{ij} = \mathcal{J}_{\mathcal{L}}^{-1}(\mathbf{e}_i\mathbf{e}_j^T), \quad \mathbf{e}_i \in \mathbf{R}^m\,,\ \mathbf{e}_j \in \mathbf{R}^n\,,$$

baza u $\mathcal{L}(X,Y)$. Razvoj u toj bazi $\mathcal{A} = \sum_{i=1}^m \sum_{j=1}^n a_{ij}\mathcal{E}_{ij}$ odgovara razvoju $\mathcal{A}(f,e) = \sum_{i=1}^m \sum_{j=1}^n a_{ij}\mathbf{e}_i\mathbf{e}_j^T$. Stoga su koeficijenti u razvoju upravo matrični elementi $\mathbf{e}_i^T\mathcal{A}(f,e)\mathbf{e}_j$.

11.7 Kompozicija linearnih operatora

Neka su X, Y i Z vektorski prostori nad istim poljem Φ (= $\mathbf{R}$ ili $\mathbf{C}$). Neka su $\mathcal{L}(X,Y)$, $\mathcal{L}(Y,Z)$ i $\mathcal{L}(X,Z)$ pripadajući vektorski prostori linearnih operatora. Za svaki par $\mathcal{A}\in\mathcal{L}(X,Y)$, $\mathcal{B}\in\mathcal{L}(Y,Z)$ definiramo produkt linearnih operatora $\mathcal{C} = \mathcal{B}\mathcal{A}$ kao kompoziciju preslikavanja $\mathcal{A}$ i $\mathcal{B}$,

$$\mathcal{C}(x) = \mathcal{B}(\mathcal{A}(x)), \quad x \in X\,.$$

Pokažimo da je $\mathcal{C}$ linearni operator. Za prizvoljne $\alpha,\beta\in\Phi$ i $u,v\in X$, imamo

$$\begin{aligned}
\mathcal{C}(\alpha u+\beta v) &= \mathcal{B}\left(\mathcal{A}(\alpha u+\beta v)\right) && \text{po definiciji od } \mathcal{C}\\
&= \mathcal{B}\left(\alpha\mathcal{A}(u)+\beta\mathcal{A}(v)\right) && \text{jer je } \mathcal{A} \text{ linearno}\\
&= \alpha\mathcal{B}\left(\mathcal{A}(u)\right)+\beta\mathcal{B}\left(\mathcal{A}(v)\right) && \text{jer je } \mathcal{B} \text{ linearno}\\
&= \alpha\mathcal{C}(u)+\beta\mathcal{C}(v) && \text{jer je } \mathcal{C}=\mathcal{B}\mathcal{A}\,.
\end{aligned}$$

Zaključujemo da je $\mathcal{C}\in\mathcal{L}(X,Z)$. Produkt linearnih operatora zadovoljava svojstvo obostrane distributivnosti, homogenosti i asocijativnosti

$$\begin{aligned}
(\mathcal{B}_1+\mathcal{B}_2)\,\mathcal{A} &= \mathcal{B}_1\mathcal{A}+\mathcal{B}_2\mathcal{A}\\
\mathcal{B}\,(\mathcal{A}_1+\mathcal{A}_2) &= \mathcal{B}\mathcal{A}_1+\mathcal{B}\mathcal{A}_2\\
\alpha\,(\mathcal{B}\mathcal{A}) &= (\alpha\mathcal{B})\,\mathcal{A} = \mathcal{B}\,(\alpha\mathcal{A})\,, \quad \alpha \text{ proizvoljni skalar}\\
(\mathcal{C}\mathcal{B})\,\mathcal{A} &= \mathcal{C}\,(\mathcal{B}\mathcal{A})
\end{aligned}$$

pri čemu su $\mathcal{A},\mathcal{A}_1,\mathcal{A}_2\in\mathcal{L}(X,Y)$, $\mathcal{B},\mathcal{B}_1,\mathcal{B}_2\in\mathcal{L}(Y,Z)$ i $\mathcal{C}\in\mathcal{L}(Z,W)$, W je vektorski prostor. Dokazi tih svojstava su jednostavni, pa ih ostavljamo za vježbu čitatelju.

Ako je polazni vektorski prostor jednak dolaznom, prostor $\mathcal{L}(X, X)$ se označava sa $\mathcal{L}(X)$. U tom prostoru postoji jedinični element $\mathcal{I} \in \mathcal{L}(X)$ s obzirom na množenje linearnih operatora. On je definiran formulom $\mathcal{I}(x) = x$ za $x \in X$. $\mathcal{I}$ se zove **jedinični** linearni operator (ili identiteta) u $\mathcal{L}(X)$ jer vrijedi

$$\mathcal{I}\mathcal{A} = \mathcal{A}\mathcal{I} = \mathcal{A} \quad \text{za svaki } \mathcal{A} \in \mathcal{L}(X).$$

Rezimirajući svojstva množenja operatora, možemo zaključiti da vrijedi

Teorem 11.9 $(\mathcal{L}(X), +, \cdot, \circ)$ *je algebra s jedinicom.*

Pritom $\cdot$ i $\circ$ označavaju operacije množenja operatora sa skalarom i množenje operatora sa operatorom. U toj lagebri je dobro definirana operacija potenciranja

$$\mathcal{A}^0 = \mathcal{I}, \quad \mathcal{A}^2 = \mathcal{A}\mathcal{A}, \quad \mathcal{A}^3 = \mathcal{A}^2\mathcal{A}$$

itd., pa se mogu proučavati polinomi od operatora, npr. $\mathcal{A}^3 - 3\mathcal{A}^2 + 2\mathcal{A} + 5\mathcal{J}$. Ako je $\mathcal{A} : X \to X$ izomorfizam, obično se naziva **automorfizam**. Kod koordinatizacije se najčešće koristi jedna baza u X, nazovimo ju e. Koordinatizacijom se $\mathcal{A} \in \mathcal{L}(X)$ karakterizira matricom $\mathcal{A}(e, e)$ koja se kraće označava s $\mathcal{A}(e)$.

Vratimo se opet trojki realnih vektorskih prostora. Neka je X n-dimenzionalan, Y m-dimenzionalan i Z p-dimenzionalan vektorski prostor i neka je $\mathcal{A} \in \mathcal{L}(X, Y)$, $\mathcal{B} \in \mathcal{L}(Y, Z)$ i $\mathcal{C} \in \mathcal{L}(X, Z)$. Odaberimo baze:

$$\begin{aligned} e &= (e_1, \ldots, e_n) \quad \text{u } X \\ f &= (f_1, \ldots, f_m) \quad \text{u } Y \\ g &= (g_1, \ldots, g_p) \quad \text{u } Z. \end{aligned}$$

Tada linearnim operatorima $\mathcal{A}$, $\mathcal{B}$ i $\mathcal{C}$ pripadaju, u odgovarajućim parovima baza, zapisi, odnosno matrice

$$A = \mathcal{A}(f, e) \in \mathbf{R}^{m \times n}, \quad B = \mathcal{B}(g, f) \in \mathbf{R}^{p \times m} \quad \text{i} \quad C = \mathcal{C}(g, e) \in \mathbf{R}^{p \times n}.$$

Postavlja se pitanje koja je veza između matrica A, B i C, ako je

$$\mathcal{C} = \mathcal{B}\mathcal{A}\,. \tag{11.13}$$

Pođimo od definicije matrica $A = (a_{ij})$, $B = (b_{ij})$ i $C = (c_{ij})$:

$$\mathcal{A}(e_j) = \sum_{i=1}^{m} a_{ij} f_i\,, \quad 1 \le j \le n, \tag{11.14}$$

$$\mathcal{B}(f_i) = \sum_{k=1}^{p} b_{ki} g_k\,, \quad 1 \le i \le m, \tag{11.15}$$

$$\mathcal{C}(e_j) = \sum_{k=1}^{p} c_{kj} g_k\,, \quad 1 \le j \le n. \tag{11.16}$$

Koristeći relacije (11.13), (11.14) i (11.15), imamo

$$\begin{aligned}
\mathcal{C}(e_j) &= (\mathcal{B}\mathcal{A})(e_j) = \mathcal{B}(\mathcal{A}(e_j)) = \mathcal{B}\left(\sum_{i=1}^{m} a_{ij} f_i\right) = \sum_{i=1}^{m} a_{ij}\mathcal{B}(f_i) \\
&= \sum_{i=1}^{m} a_{ij} \sum_{k=1}^{p} b_{ki} g_k = \sum_{i=1}^{m}\sum_{k=1}^{p} a_{ij} b_{ki} g_k \\
&= \sum_{k=1}^{p}\sum_{i=1}^{m} a_{ij} b_{ki} g_k = \sum_{k=1}^{p}\left(\sum_{i=1}^{m} b_{ki} a_{ij}\right) g_k.
\end{aligned}$$

Uspoređujući zadnju relaciju s relacijom (11.16), zaključujemo

$$\sum_{k=1}^{p}\left(c_{kj} - \sum_{i=1}^{m} b_{ki} a_{ij}\right) g_k = 0.$$

Dobili smo iščezavajuću linearnu kombinaciju baznih vektora prostora Z, pa moraju svi koeficijenti biti nula. Tako dobivamo

$$c_{kj} = \sum_{i=1}^{m} b_{ki} a_{ij}, \quad 1 \le k \le p,\ 1 \le j \le n,$$

a to znači

$$C = BA\,. \tag{11.17}$$

Time smo dokazali

Teorem 11.10 *Neka su X, Y, Z konačno-dimenzionalni realni vektorski prostori i neka su e, f, g pripadne baze, respektivno. Neka su $\mathcal{A} \in \mathcal{L}(X, Y)$, $\mathcal{B} \in \mathcal{L}(Y, Z)$ i $\mathcal{C} \in \mathcal{L}(X, Z)$ linearni operatori i $A \in \mathbf{R}^{m\times n}$, $B \in \mathbf{R}^{p\times m}$ i $C \in \mathbf{R}^{p\times n}$ njihovi zapisi u odgovarajućim parovima baza, respektivno. Ako je $\mathcal{C} = \mathcal{B}\mathcal{A}$, tada je $C = BA$.*

Relaciju (11.17) možemo i ovako zapisati

$$(\mathcal{B}\mathcal{A})\,(g,e) = \mathcal{B}(g,f)\mathcal{A}(f,e)\,. \tag{11.18}$$

Ako su X, Y i Z kompleksni vektorski prostori, tada teorem vrijedi uz napomenu da su $A \in \mathbf{C}^{m\times n}$, $B \in \mathbf{C}^{p\times m}$ i $C \in \mathbf{C}^{p\times n}$, respektivno.

Ako se vratimo definiciji produkta dviju matrica, tada postaje jasno da je *produkt matrica upravo tako definiran da bi vrijedio teorem 11.10, odnosno relacija (11.18).*

Koordinatizacijom vektorskih prostora, vektorska jednadžba $\mathcal{A}(x) = b$ i operatorska jednadžba $\mathcal{B}\mathcal{X} = \mathcal{C}$, prelaze pomoću relacija (11.11) i (11.18) u matrične jednadžbe $\mathcal{A}(f,e)x(e) = b(f)$ i $\mathcal{B}(g,f)\mathcal{X}(f,e) = \mathcal{C}(g,e)$. Obje matrične jednadžbe znamo riješiti svađanjem matrice na reducirani oblik ili pomoću Gaussove metode eliminacija s parcijalnim pivotiranjem. U slučaju jednog prostora i jedne koordinatizacije, ove matrične jednadžbe poprimaju oblik $\mathcal{A}(e)x(e) = b(e)$ i $\mathcal{B}(e)\mathcal{X}(e) = \mathcal{C}(e)$.

Primjer 11.11 *Neka su dani vektorski prostori X i Y dimenzije n. Neka je zadan linearni operator $\mathcal{A} : X \to Y$ npr. pomoću relacije (11.7), pri čemu su e i f baze u X i Y, respektivno. Neka je zadan vektor $y \in Y$. Kako ćete pronaći $x \in X$ za koji vrijedi $\mathcal{A}(x) = y$?*

Rješenje: *S bazama e i f posve su definirani izomorfizmi $\mathcal{J}_X$ i $\mathcal{J}_Y$, respektivno. Oni su definirani relacijama*

$$\mathcal{J}_X(e_i) = \mathbf{e}_i \in \mathbf{R}^n \quad i \quad \mathcal{J}_Y(f_i) = \mathbf{e}_i \in \mathbf{R}^n, \quad 1 \le i \le n$$

gdje je $\mathbf{e}_i$ i-ti stupac jedinične matrice I_n.

Neka je $y(f) = [y_1, \ldots, y_n]^T \in \mathbf{R}^n$ koordinatni vektor od y i neka je $A = \mathcal{A}(f,e)$ kvadratna matrica reda n. Svako rješenje sustava $A\mathbf{x} = y(f)$ je koordinatni vektor jednog rješenja vektorske jednadžbe $\mathcal{A}(x) = y$. I obrnuto, za svako rješenje x vektorske jednadžbe, pripadni koordinatni vektor $x(e)$ zadovoljava matrični sustav $A\mathbf{x} = y(f)$.

Riješimo matrični sustav $A\mathbf{x} = y(f)$ pomoću Gaussove metode eliminacija s parcijalnim pivotiranjem ili svađanjem matrice na reducirani oblik.

Ako postoji jedinstveno rješenje $[x_1, \ldots, x_n]^T$ matričnog sustava, postojat će jedinstveno rješenje x polazne vektorske jednadžbe u oblika $x = \mathcal{J}_X^{-1}(\mathbf{x}) = x_1e_1 + \cdots + x_ne_n$.

Ako matrična jednadžba ima beskonačno rješenja, odredi se baza $\mathbf{z}_1$*, ...,*$\mathbf{z}_n$ *nul-potprostora* $\mathcal{N}(A)$ *i jedno partikularno rješenje* $\mathbf{z}_0$*. Pomoću izomorfizma* $\mathcal{J}_X^{-1}$ *odrede se pripadni originali* $z_1 = \mathcal{J}_X^{-1}(\mathbf{z}_1)$*, ...,* $z_n = \mathcal{J}_X^{-1}(\mathbf{z}_n)$ *i* $z_0 = \mathcal{J}_X^{-1}(\mathbf{z}_0)$*. Opće rješenje operatorske jednadže je dano formulom* $z_0 + \alpha_1 z_1 + \cdots + \alpha_n z_n$*, gdje su* α_i *proizvoljni skalari.* ∎

U slučaju kad je $X = Y = Z$ dimenzije n, teorem 11.10 pokazuje da je $\mathcal{J}_{\mathcal{L}}$ izomorfizam algebri $(\mathcal{L}(X), +, \cdot, \circ)$ i $(\mathbf{R}^{n\times n}, +, \cdot, *)$ (odnosno $(\mathcal{L}(X), +, \cdot, \circ)$ i $(\mathbf{C}^{n\times n}, +, \cdot, *)$ ako je $\Phi = \mathbf{C}$), gdje smo s $\circ$ označili produkt (kompoziciju) linearnih operatora, a sa $*$ produkt matrica.

11.8 Promjena baza

Neka je X vektorski prostor i neka su $e = (e_1, \ldots, e_n)$ i $e' = (e'_1, \ldots, e'_n)$ dvije baze u X. Koristeći relaciju (11.10) zaključujemo da jednadžba $x = \mathcal{I}(x)$, gdje je $\mathcal{I}$ identiteta u $\mathcal{L}(X)$, prelazi u matričnu jednadžbu $x(e') = \mathcal{I}(e', e)x(e)$. Označimo matricu $\mathcal{I}(e', e)$ sa $S = (s_{ij})$, tako da je

$$x(e') = Sx(e)\,, \tag{11.19}$$

Elementi matrice S su koeficijenti koji povezuju bazu e s bazom e', pomoću relacije

$$e_j = \sum_{i=1}^{n} s_{ij} e'_i\,, \quad 1 \le j \le n\,. \tag{11.20}$$

Uvjerimo se u to. Pretpostavimo da vrijedi

$$e_j = \sum_{i=1}^{n} t_{ij} e'_i\,, \quad 1 \le j \le n\,.$$

i neka je T matrica koeficijenata, $T = (t_{ij})$. Usporedbom desnih strana u relacijama

$$x = \sum_{j=1}^{n} x_j e_j = \sum_{j=1}^{n} x_j \sum_{i=1}^{n} t_{ij} e'_i = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} t_{ij} x_j \right) e'_i$$

i

$$x = \sum_{i=1}^{n} x'_i e'_i$$

dobivamo

$$x'_i = \sum_{j=1}^{n} t_{ij} x_j, \quad 1 \le i \le n,$$

odnosno $x(e') = Tx(e)$. Oduzmimo od te jednadžbe jednadžbu (11.19). Dobivamo

$$(T - S)x(e) = 0 \quad \text{za svako x},$$

pa zaključujemoda je $T = S$. Time je dokazana relacija (11.20).

Lema 11.12 *Matrica S definirana relacijom (11.20) je regularna.*

Dokaz: Pokazali smo da S zadovoljava relaciju (11.19). Neka su $\mathcal{J}_e$ i $\mathcal{J}_{e'}$ koordinatizacije prostora X pomoću baza e i e', respektivno. Dakle, x(e)=$\mathcal{J}_e(x)$ i $x(e') = \mathcal{J}_{e'}(x)$. Za specijalni izbor vektora x imamo

$$\mathcal{J}_e(e_i) = \mathbf{e}_i, \quad \mathcal{J}_{e'}(e'_i) = \mathbf{e}_i, \quad 1 \le i \le n.$$

gdje je $\mathbf{e}_i$ i-ti stupac matrice I_n. Prema relaciji (11.19), za svako i, imamo

$$\mathcal{J}_{e'}(e_i) = S\mathcal{J}_e(e_i) = S\mathbf{e}_i = \text{ } i\text{-ti stupac od } S\,.$$

Jer je $\mathcal{J}_{e'}$ izomorfizam, iz propozicije 11.4 slijedi da su vektori $\mathcal{J}_{e'}(e_i)$ linearno nezavisni, a kako su to stupci od S, S je regularna. ■

Iz prethodnih izlaganja i uz pomoć leme 11.12 odmah slijedi

Teorem 11.13 *Neka je X vektorski prostor i e, e' dvije baze u X. Tada postoji regularna matrica S, takva da za sve $x \in X$ vrijedi (11.19), pri čemu su $x(e)$ i $x(e')$ koordinatni vektori za x u bazama e i e', respektivno.*

Matrica $S = \mathcal{I}(e', e)$ se obično zove **matrica tranzicije** s baze e na bazu e'. Odmah se vidi da je $S^{-1} = \mathcal{I}(e, e')$ matrica tranzicije s baze e' na bazu e. ■

Neka je X n-dimenzionalni, Y i m-dimenzionalni vektorski prostor i $\mathcal{A} \in \mathcal{L}(X, Y)$. Neka su e i e' baze u X, a f i f' baze u Y. U paru baza e, f, $\mathcal{A}$ se reprezentira matricom $\mathcal{A}(f, e)$, a u paru e', f', matricom $\mathcal{A}(f', e')$. Odredimo vezu između tih dviju matrica.

Polazimo od jednadžbe

$$y = \mathcal{A}(x)$$

Koordinatizacijom u paru baza e, f, ona prelazi u

$$y(f) = \mathcal{A}(f, e)x(e)\,, \tag{11.21}$$

dok u paru e', f', prelazi u

$$y(f') = \mathcal{A}(f', e')x(e')\,. \tag{11.22}$$

Pomoću teorema 11.13, relacija (11.22) prelazi u

$$Ty(f) = \mathcal{A}(f', e')Sx(e)\,, \tag{11.23}$$

gdje su $S = \mathcal{I}(e', e)$ i $T = \mathcal{I}(f', f)$. Po lemi 11.12, S i T su regularne matrice. Množenjem jednadžbe (11.23) sa T^{-1} s lijeva, dobijemo

$$y(f) = T^{-1}\mathcal{A}(f', e')Sx(e)\,. \tag{11.24}$$

Oduzimajući jednadžbu (11.24) od jednadžbe (11.21), dobijemo

$$\Big(\mathcal{A}(f, e) - T^{-1}\mathcal{A}(f', e')S\Big)\, x(e) = 0\,. \tag{11.25}$$

Kad x prolazi vektorskim prostorom X, $x(e)$ prolazi cijelim $\mathbf{R}^n$, pa iz relacije (11.25) slijedi

$$\mathcal{A}(f, e) = T^{-1}\mathcal{A}(f', e')S \quad \text{ili} \quad \mathcal{A}(f', e') = T\mathcal{A}(f, e)S^{-1}\,. \tag{11.26}$$

Ako je $X = Y$ i $e = f$, $e' = f'$, onda je $T = S$, pa (11.25) prelazi u

$$\mathcal{A}(e) = S^{-1}\mathcal{A}(e')S \quad \text{ili} \quad \mathcal{A}(e') = S\mathcal{A}(e)S^{-1}\,. \tag{11.27}$$

Definicija 11.14 *Ako vrijedi $B = SAT$ za neke regularne matrice S i T, tada je matrica B* **ekvivalentna** *matrici A. Ako vrijedi $B = S^{-1}AS$ za neku nesingularnu matricu S, B je* **slična** *matrici A.* ■

Uočimo da su ekvivalentne matrice uvijek istog tipa, tj. imaju isti broj redaka i isti broj stupaca. Za razliku od ekvivalentnih matrica, slične matrice su uvijek kvadratne istog reda. I ekvivalentnost i sličnost matrica su relacije ekvivalencije u skupu matrica.

Sada možemo rezimirati dobijene rezultate.

Teorem 11.15 *Neka su X i Y konačno-dimenzionalni vektorski prostori i $\mathcal{A} \in \mathcal{L}(X,Y)$ linearni operator. Neka su e i e' odnosno f, f', baze u X odnosno Y, respektivno, i neka su $A = \mathcal{A}(f,e)$ i $A' = \mathcal{A}(f',e')$ reprezentacije operatora u navedenim parovima baza. Tada su matrice A i A' ekvivalentne, tj. postoje regularne matrice S i T, takve da je $A' = SAT$. Pritom su $S = \mathcal{I}(e,e')$ i $T = \mathcal{I}(f',f)$ tranzicijske matrice.*

Ako je $X = Y$ i $e = f$, $e' = f'$, $A = \mathcal{A}(e)$ i $A' = \mathcal{A}(e')$, tada je matrica A' slična matrici A, tj. vrijedi $A' = T^{-1}AT$, pri čemu je $T = \mathcal{I}(e',e)$ tranzicijska matrica s baze e na bazu e'. ∎

Iz teorema 11.15 i Binet-Cauchyjeva teorema možemo zaključiti da *svi zapisi $\mathcal{A}(e)$ operatora $\mathcal{A} \in \mathcal{L}(X)$ imaju istu determinantu*. Naime, slične matrice imaju jednaku determinantu. Doista, ako je $B = S^{-1}AS$, dvostrukom upotrebom Binet-Cauchyjeva teorema dobijemo

$$\begin{aligned} \det(B) &= \det(S^{-1}AS) = \det(S^{-1})\det(A)\det(S) \\ &= \det(S^{-1})\det(S)\det(A) = \det(S^{-1}S)\det(A) \\ &= \det(I)\det(A) = \det(A)\,. \end{aligned}$$

To znači da se determinanta može definirati i za operatore iz $\mathcal{L}(X)$:

$$\det(\mathcal{A}) \stackrel{\text{def}}{=} \det(\mathcal{A}(e))\,, \quad e \text{ bilo koja baza u } X\,. \tag{11.28}$$

Određivanje jezgre i slike operatora $\mathcal{A}$ se preko koordinatizacija prostora X i Y lako svede na određivanje nul-potprostora i područja vrijednosti matrice $\mathcal{A}(f,e)$. Naime, jednadžbe $\mathcal{A}(x) = 0$ i $y = \mathcal{A}(x)$ prelaze u $\mathcal{A}(f,e)x(e) = 0$ i $y(f) = \mathcal{A}(f,e)x(e)$ pa se zaključi,

$$\mathcal{J}_X\left(\mathcal{N}(\mathcal{A})\right) = \mathcal{N}(A) \quad \text{i} \quad \mathcal{J}_Y\left(\mathcal{R}(\mathcal{A})\right) = \mathcal{R}(A)$$

gdje je $A = \mathcal{A}(f,e)$. To je u suglasnosti sa propozicijom 11.4 koja garantira da su dimenzije potprostora $\mathcal{N}(\mathcal{A})$ i $\mathcal{N}(A)$ jednake. Sada iz odgovarajućeg rezultata za matrice odmah slijedi da je suma ranga i defekta operatora $\mathcal{A}$ jednaka dimenziji polaznog prostora X.

Operator $\mathcal{A} \in \mathcal{L}(X)$ je regularan (invertibilan, nesingularan) ako je automorfizam, tj. ako za njega postoji inverz $\mathcal{A}^{-1} \in \mathcal{L}(X)$. U protivnom je singularan. Na osnovu propozicije 11.5 odmah slijedi da je $\mathcal{A}$

singularan ako i samo ako jezgra od $\mathcal{A}$ nije nul-vektor, tj. ako i samo ako postoji $x \neq 0$ takav da je $\mathcal{A}(x) = 0$. To je u potpunoj korespondenciji s analognim rezultatom za matrice. Zato je za svaku bazu e matrica $\mathcal{A}(e)$ regularna (singularna) ako i samo ako je takav operator $\mathcal{A}$. Do istog zaključka se lako dolazi i ako se pođe od činjenice da su algebre $(\mathcal{L}(X), +, \cdot, \circ)$ i $(\mathbf{R}^{n\times n}, +, \cdot, *)$ (odnosno $(\mathcal{L}(X), +, \cdot, \circ)$ i $(\mathbf{C}^{n\times n}, +, \cdot, *)$ izomorfne.

Sada se lako zaključi da singularni (regularni) operator ima determinantu jednaku nuli (različitu od nule).

Lekcija 12

Vlastite vrijednosti i vektori

Titranje tijela u prirodi, strojeva, brodova, aviona i raketa u tehnici, kuća, mostova u graditeljstvu i još mnogih drugih stvari (od sićušnih dijelova čipova, do velikih stadiona) opisuju se tzv. problemom vlastitih vrijednosti za razne diferencijalne operatore. Pritom su vlastite vrijednosti proporcionalne recipročnim vrijednostima vlastitih frekvencija sistema, a vlastiti vektori određuju smjerove u kojima sistem titra.

Kad se žele izračunati glavne frekvencije titranja, aproksimiraju se diferencijalni operatori manje složenim matematičkim objektima, pa se na kraju dobije problem vlastitih vrijednosti za matrice. Stoga je važno znati rješavati matrični problem vlastitih vrijednosti.

Definicija 12.1 *Neka je X vektorski prostor i $\mathcal{A} \in \mathcal{L}(X)$ linearni operator. Vektor $x \in X$, $x \neq 0$ je* **vlastiti vektor** *od $\mathcal{A}$ ako postoji skalar λ takav da je $\mathcal{A}x = \lambda x$. U tom slučaju λ je* **vlastita vrijednost***, a par (x, λ) je* **vlastiti par** *operatora $\mathcal{A}$.* ■

Osim termina vlastita vrijednost (vektor, par) još se koristi termin **svojstvena** vrijednost (vektor, par). Iz definicije odmah slijedi da je i αx vlastiti vektor za svaki skalar α. Dakle, traži se "smjer" (zato x mora biti netrivijalan) u kojem operator djeluje tako da ga ne mijenja, već vektore tog smjera samo skraćuje, produžuje i pritom možda mijenja orijentaciju. Ako je $\mathcal{N}(\mathcal{A}) \neq \{0\}$, tada su svi netrivijalni vektori iz $\mathcal{N}(\mathcal{A})$ vlastiti vektori koji pripadaju vlastitoj vrijednosti 0. Specijalno, za nul-operator $\mathcal{O}$ vrijedi $\mathcal{N}(\mathcal{A}) = X$, pa su svi netrivijalni vektori iz X

vlastiti vektori. Slično svojstvo ima operator identitete $\mathcal{I} : \mathbf{R}^n \to \mathbf{R}^n$ za koji vrijedi $\mathcal{I}(x) = x = 1 \cdot x$ za sve x, pa su svi netrivijalni vektori iz X vlastiti vektori za vlastitu vrijednost 1.

U drugoj krajnosti leže operatori koji nemaju vlastitih parova.

Primjer 12.2 *Neka je $\mathcal{R}$ operator rotacije u $\mathbf{R}^2$ za kut $0 < \phi < \pi$. Tada ne postoji niti jedan smjer (tj. radij-vektor) koji ne bi pod rotacijom promijenio smjer. Stoga za $\mathcal{R}$ ne postoji ni jedan vlastiti vektor, pa zato niti jedna vlastita vrijednost.* ■

Ipak, najčešći je slučaj kada postoje posve određeni smjerovi koji određuju vlastite vektore. Sljedeći primjer pokazuje da je taj slučaj prisutan i u beskonačno-dimenzionalnim prostorima.

Primjer 12.3 *Neka je C_∞ vektorski prostor svih beskonačno diferencijabilnih funkcija nad poljem realnih brojeva. Neka je $\mathcal{A} : C_\infty \to C_\infty$ operator diferenciranja. Tada je $f(t) = e^{\lambda t}$ vlastiti vektor za $\mathcal{A}$, a pripadna vlastita vrijednost je $\lambda \in \mathbf{R}$. To se vidi iz sljedeće relacije*

$$\mathcal{A}(f) = \frac{d}{dt} e^{\lambda t} = \lambda e^{\lambda t} = \lambda f \, , \; f \neq 0.$$

Primijetimo da je C_∞ beskonačno-dimenzionalni vektorski prostor. Kako skalara λ ima (neprebrojivo) beskonačno, toliko ima vlastitih vrijednosti odnosno (međusobno linearno nezavisnih) vlastitih vektora. ■

Neka je λ vlastita vrijednost operatora $\mathcal{A} \in \mathcal{L}(X)$. Po definiciji, postoji $x \in X$, $x \neq 0$, takav da je $\mathcal{A}x = \lambda x$. Da li osim multipla od x, ima i drugih vlastitih vektora za λ? Djelomični odgovor na to pitanje daje

Propozicija 12.4 *Neka je λ vlastita vrijednost operatora $\mathcal{A} \in \mathcal{L}(X)$. Tada je skup $X_\lambda = \{x \in X : \mathcal{A}(x) = \lambda x\}$ vektorski potprostor od X.*

Dokaz: Neka su $u, v \in X_\lambda$ i neka su $\alpha, \beta \in \Phi$ skalari iz polja Φ nad kojim je X definiran. Tada za u i v vrijedi

$$\mathcal{A}(u) = \lambda u, \quad \mathcal{A}(v) = \lambda v \, ,$$

pa je

$$\begin{aligned}
\mathcal{A}(\alpha u+\beta v) &= \alpha\mathcal{A}(u)+\beta\mathcal{A}(v), \qquad \text{jer je } \mathcal{A} \text{ linearan}\\
&= \alpha\lambda u+\beta\lambda v, \qquad \text{jer su } u \text{ i } v \text{ u } X_\lambda\\
&= \lambda(\alpha u+\beta v), \qquad \text{jer je } X \text{ vektorski prostor}\,.
\end{aligned}$$

Pokazali smo da $u, v \in X_\lambda$ i $\alpha, \beta \in \Phi$ povlače $\alpha u + \beta v \in X_\lambda$, pa je X_λ vektorski potprostor od X. ■

Ako su u i v vlastiti vektori koji pripadaju međusobno različitim vlastitim vrijednostima λ i μ, respektivno, onda $u+v$ nije vlastiti vektor, a nije to niti bilo koja linearna kombinacija $\alpha u + \beta v$ koja zadovoljava uvjet $\alpha \neq 0$ i $\beta \neq 0$. To osigurava

Teorem 12.5 *Neka je X vektorski prostor i $\mathcal{A} \in \mathcal{L}(X)$ linearni operator. Neka su $u_1, \ldots, u_r$ vlastiti vektori od $\mathcal{A}$ i $\lambda_1, \ldots, \lambda_r$ pripadne vlastite vrijednosti. Ako vrijedi*

$$\lambda_i \neq \lambda_j \qquad \text{čim je } i \neq j\,, \tag{12.1}$$

onda je skup vektora $\{u_1, \ldots, u_r\}$ linearno nezavisan.

Dokaz: Dokaz koristi princip matematičke indukcije po r. Za $r = 1$, vektor u_1 je sigurno linearno nezavisan, jer je kao vlastiti vektor netrivijalan. Pretpostavimo da je tvrdnja teorema istinita za $r-1$. To znači da je svaki skup od $r-1$ vlastitih vektora linearno nezavisan ako za pripadajuće vlastite vrijednosti vrijedi (12.1).

Neka je zadan skup od r vlastitih vektora za čije vlastite vrijednosti vrijedi relacija (12.1). Moramo pokazati da uvjet

$$\alpha_1 u_1 + \cdots + \alpha_r u_r = 0 \tag{12.2}$$

povlači $\alpha_1 = \cdots = \alpha_r = 0$.

Primijenimo na vektor desne i vektor lijeve strane jedandžbe (12.2), operator $\mathcal{A}$. Zbog linearnosti operatora, dobivamo

$$\begin{aligned}
\mathcal{A}(0)=0 &= \mathcal{A}(\alpha_1 u_1 + \cdots + \alpha_r u_r)\\
&= \alpha_1\mathcal{A}(u_1) + \cdots + \alpha_r\mathcal{A}(u_r)\\
&= \alpha_1\lambda_1 u_1 + \cdots + \alpha_r\lambda_r u_r. \qquad (12.3)
\end{aligned}$$

Pomnožimo vektore na lijevoj i desnoj strani jednadžbe (12.2) s λ_1,

$$\alpha_1\lambda_1 u_1 + \cdots + \alpha_r\lambda_1 u_r = 0\,. \tag{12.4}$$

Oduzmimo jednadžbu (12.4) od jednadžbe (12.3). Dobivamo

$$\alpha_2(\lambda_2 - \lambda_1)u_2 + \cdots + \alpha_r(\lambda_r - \lambda_1)u_r = 0$$

Iskoristimo sada indukcijsku pretpostavku, za skup $\{u_2, \ldots, u_r\}$ od $r-1$ vlastitih vektora. To smijemo jer su pripadajuće vlastite vrijednosti λ_i međusobno različite. Indukcijska pretpostavka odmah daje

$$\alpha_2(\lambda_2 - \lambda_1) = \cdots = \alpha_r(\lambda_r - \lambda_1) = 0.$$

Zbog uvjeta (12.1), zadnji niz jednadžbi je ekvivalentan s

$$\alpha_2 = \cdots = \alpha_r = 0\,. \tag{12.5}$$

Iskoristimo li informaciju (12.5) u jednadžbi (12.2), dobijemo $\alpha_1 u_1 = 0$, odakle zbog $u_1 \neq 0$ zaključujemo da je i $\alpha_1 = 0$. Dakle smo dobili $\alpha_1 = \cdots = \alpha_r = 0$, što je i trebalo dokazati. ■

Ako je u teoremu 12.5, $r = n = \dim(X)$, tvrdnja teorema se može interpretirati na sljedeći način

Korolar 12.6 *Ako linearni operator $\mathcal{A} \in \mathcal{L}(X)$ ima $n = \dim(X)$ međusobno različitih vlastitih vrijednosti, tada postoji baza prostora X koja se sastoji od vlastitih vektora operatora $\mathcal{A}$.* ■

Neka je dana baza e u n-dimenzionalnom realnom vektorskom prostoru X. Koordinatizacijom prostora X pomoću baze e, jednadžba

$$\mathcal{A}(x) = \lambda x$$

prelazi u matričnu jednadžbu

$$\mathcal{A}(e)x(e) = \lambda x(e), \quad x(e) \neq 0.$$

S obzirom da je koordinatizacija izomorfizam između vektorskih prostora X i $\mathbf{R}^n$ (ili $\mathbf{C}^n$), problem nalaženja vlastitih vektora i vrijednosti se reducira na odgovarajući matrični problem.

Matrični problem vlastitih vrijednosti se zapisuje u obliku

$$Ax = \lambda x, \quad x \neq 0, \qquad A \text{ kvadratna},$$

pri čemu je λ vlastita vrijednost, a x vlastiti vektor od A. Koordinatizacijom smo problem određivanja vlastitih vrijednosti i vektora operatora sveli na odgovarajući matrični problem.

12.1 Karakteristični polinom

Neka je λ vlastita vrijednost operatora $\mathcal{A} \in \mathcal{L}(X)$. Tada postoji $x \in X$, takav da je $\mathcal{A}(x) = \lambda x$. Zadnja jednadžba se može zapisati u obliku $\mathcal{A}(x) - \lambda x = 0$ ili $(\mathcal{A} - \lambda\mathcal{I})(x) = 0$. Dakle operator $\mathcal{A} - \lambda\mathcal{I}$ prebacuje netrivijalan vektor u nulu, pa je singularan. Stoga je njegova determinanta nula,

$$\det(\mathcal{A} - \lambda\mathcal{I}) = 0\,. \tag{12.6}$$

Izborom baze e i koordinatizacijom, dobivamo singularnu matricu $(\mathcal{A} - \lambda\mathcal{I})(e) = A - \lambda I$, gdje je $A = \mathcal{A}(e)$. Jer je $A - \lambda I$ singularna, imamo

$$\det(\lambda I - A) = (-1)^n \det(A - \lambda I) = \begin{vmatrix} \lambda - a_{11} & \cdots & -a_{nn} \\ \vdots & \ddots & \vdots \\ -a_{n1} & \cdots & \lambda - a_{nn} \end{vmatrix} = 0\,. \tag{12.7}$$

Determinanta matrice $zI - A$ je polinom n-tog stupnja u z,

$$\det(zI - A) = z^n - \sigma_1 z^{n-1} - \cdots - \sigma_{n-1} z - \sigma_n \tag{12.8}$$

koji se zove **karakteristični polinom** matrice A, a (12.7) je **karakteristična jednadžba** matrice A. S obzirom da $B = S^{-1}AS$ povlači $B - zI = S^{-1}(A - zI)S$, vidimo da slične matrice imaju isti karakteristični polinom. To znači da koeficijenti σ_i ne ovise o konkretnoj reprezentaciji operatora, već ovise o operatoru $\mathcal{A}$. Stoga se kaže da je $\det(z\mathcal{I} - \mathcal{A})$ karakteristični polinom, a $\det(z\mathcal{I}) - \mathcal{A} = 0$ karakteristična jednadžba operatora $\mathcal{A}$.

Jer je λ bila proizvoljna vlastita vrijednost, pokazali smo da je svaka vlastita vrijednost nultočka karakterističnog polinoma. Pokažimo da vrijedi i obrat. Neka je μ nultočka karakterističnog polinoma. Dakle je

$\det(\mu\mathcal{I}-\mathcal{A})=0$. Stoga je operator $\mu\mathcal{I}-\mathcal{A}$ singularan, pa mu jezgra nije samo nul-vektor. Dakle, postoji netrivijalan vektor $x\in X$, takav da je $(\mu\mathcal{I}-\mathcal{A})x=0$ ili $\mathcal{A}x=\mu x$, $x\neq 0$. Zaključujemo da je svaka nultočka karakterističnog polinoma vlastita vrijednost operatora $\mathcal{A}$. Time smo našli jednu karakterizaciju vlastitih vrijednosti i dokazali

Teorem 12.7 *Neka je X vektorski prostor i $\mathcal{A}\in\mathcal{L}(X)$. Skalar λ je vlastita vrijednost operatora $\mathcal{A}$ onda i samo onda ako je nultočka karakterističnog polinoma.* ■

Koeficijenti karakterističnog polinoma su određene sume produkata matričnih elemenata, pa neprekidno ovise o matrici. Kako su nultočke polinoma neprekidne funkcije koeficijenata polinoma, zaključujemo da su vlastite vrijednosti neprekidne funkcije matrice. Izomorfizam koji povezuje linearni operator $\mathcal{A}$ i matricu $\mathcal{A}(e)$ je neprekidna funkcija pa možemo zaključiti da su vlastite vrijednosti neprekidne funkcije operatora.

Nultočke polinoma mogu biti višestruke, npr. $p(z)=z^3(z-2)^2$ ima dvije nultočke: 0 kratnosti tri i 2 kratnosti dva. Pritom je zbroj kratnosti (ili višestrukosti) svih nultočaka jednak stupnju polinoma.

Definicija 12.8 **Algebarska višestrukost** *vlastite vrijednosti λ linearnog operatora $\mathcal{A}\in\mathcal{L}(X)$ (ili matrice A) je njena višestrukost kao nultočke karakterističnog polinoma.* **Geometrijska višestrukost** *vlastite vrijednosti λ je defekt operatora $\lambda\mathcal{I}-\mathcal{A}$ (matrice $\lambda I-A$).* ■

Vidjeli smo da se određivanje vlastitih vrijednosti operatora, koordinatizacijom svodi na određivanje vlastitih vrijednosti matrica. Ako je vektorski prostor definiran nad poljem kompeksnih brojeva, tada će zapis operatora biti matrica iz $\mathbf{C}^{n\times n}$, pa će pripadni karakteristični polinom imati kompleksne koeficijente. Koristeći fundamentalni teorem algebre, koji kaže da polinom stupnja n nad (zatvorenim poljem, a takvo je) $\mathbf{C}$ ima n nultočaka, brojeći ih s višestrukostima, zaključujemo da matrica iz $\mathbf{C}^{n\times n}$ ima točno n vlastitih vrijednosti. Ako pritom sve vlastite vrijednosti imaju algebarsku kratnost jedan, tada na osnovu korolara 12.6 znamo da postoji baza vektorskog prostora koja se sastoji od vlastitih vektora za $\mathcal{A}$. Ako postoje vlastite vrijednosti kratnosti veće od jedan, tada može, ali i ne mora postojati baza vlastitih vektora. Isti zaključak vrijedi naravno i za matrice iz $\mathbf{C}^{n\times n}$.

Primjer 12.9 *Neka je* $\nu \in \mathbf{C}$, $A = \operatorname{diag}(\nu, \ldots, \nu) = \nu I$ *i*

$$J_n(\nu) = \begin{bmatrix} \nu & 1 & 0 & \cdots & 0 & 0 \\ 0 & \nu & 1 & \cdots & 0 & 0 \\ 0 & 0 & \nu & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \nu & 1 \\ 0 & 0 & 0 & \cdots & 0 & \nu \end{bmatrix}.$$

Obje matrice imaju jednu vlastitu vrijednost: ν *algebarske kratnosti* n, *no prva ima* n *linearno nezavisnih vektora:* $e_1, \ldots e_n$, *dok druga ima samo jedan vlastiti vektor,* e_1. *Provjerimo to za* $n = 2$. *Matrica*

$$B(\nu, \nu_2) = \begin{bmatrix} \nu & 1 \\ 0 & \nu_2 \end{bmatrix}$$

ima vlastite parove

$$\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \nu \right) \quad i \quad \left(\begin{bmatrix} 1 \\ \nu_2 - \nu \end{bmatrix}, \nu_2 \right).$$

Puštajući $\nu_2 \to \nu$, *dobivamo* $B(\nu, \nu_2) \to J_2(\nu)$ *i drugi vlastiti vektor prelazi u prvi vlastiti vektor, pa* $J_2(\nu)$ *ima samo jedan vlastiti vektor. Slično ponašanje vrijedi i za opće* n. ■

Ako je operator definiran na realnom vektorskom prostoru, tada je i zapis operatora $\mathcal{A}$ realna matrica A. To znači da ćemo u razvoju determinante $\det(zI - A)$ imati uz potencije od z uvijek realne brojeve, tj. da će koeficijenti karakterističnog polinoma biti realni. Polinom stupnja n nad poljem realnih brojeva (koje nije zatvoreno) ne mora imati n nultočaka. Npr. $p(t) = t^2 + 1$ uopće nema realnih korijena. Stoga operatori (ili matrice) nad realnim poljem ne moraju imati vlastite vrijednosti, odnosno mogu ih imati manje od n.

S druge strane, svaka realna matrica A se može promatrati kao da je kompleksna, jer su realni brojevi također kompleksni brojevi. Sada je $A \in \mathbf{C}^{n \times n}$, pa karakteristični polinom ima prema osnovnom teoremu algebre n nultočaka koji su općenito kompleksni brojevi. Peciznije, ako karakteristični polinom ima realne koeficijente, njegove nultočke mogu biti ili realni brojevi ili kompleksni brojevi koji se javljaju kao parovi konjugirano kompleksnih brojeva. Stoga, kod računanja vlastitih vrijednosti realne matrice A, u pravilu se računaju sve vlastite vrijednosti A gledajući na A kao da je iz $\mathbf{C}^{n\times n}$. Nakon toga se odbace sve vlastite vrijednosti (i pripadni vlastiti vektori) koje nisu realni.

Primjer 12.10 *Neka je*

$$R = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix}$$

matrica rotacije za kut ϕ u pozitivnom smjeru. Karakteristični polinom od R je

$$\begin{aligned} \chi(\lambda) &= \det(A-\lambda I) = \begin{vmatrix} \lambda-\cos\phi & \sin\phi \\ -\sin\phi & \lambda-\cos\phi \end{vmatrix} \\ &= (\lambda-\cos\phi)^2 + \sin\phi^2 \,. \end{aligned}$$

Nultočke tog polinoma (promatranog nad poljem $\mathbf{C}$) su $\lambda_1 = \cos\phi + i\sin\phi$ i $\lambda_2 = \cos\phi - i\sin\phi$. Vidimo da su obje nultočke realne samo ako je kut ϕ višekratnik od π i tada je R jednako I ili $-I$. U ostalim slučajevima R (sada gledana kao element iz $\mathbf{R}^{2\times 2}$) nema vlastitih vrijednosti, pa zato nema niti vlastitih vektora. ■

Postavlja se pitanje kada kvadratna matrica ima bazu koja se sastoji od vlastitih vektora, a kad takovu bazu nema.

Teorem 12.11 *Neka je $A \in \mathbf{C}^{n\times n}$. Matrica A ima n linearno nezavisnih vektora ako i samo ako postoji regularna matrica G, takva da je*

$$A = G\Lambda G^{-1} \tag{12.9}$$

pri čemu je Λ dijagonalna matrica.

Dokaz: Neka vrijedi (12.9) s regularnom matricom G. Neka je $G = [g_1,\ldots,g_n]$ stupčana particija matrice G i neka je $\Lambda = \operatorname{diag}(\lambda_1,\ldots,\lambda_n)$. Množenjem jednadžbe (12.9) s desne strane matricom G, dobivamo $AG = G\Lambda$. Primijenjujući matricu na lijevoj i matricu na desnoj strani zadnje jednadžbe, na kanonski vektor e_i (i-ti stupac jedinične matrice), dobivamo $AGe_i = G\Lambda e_i$ odnosno $Ag_i = G\lambda_i e_i = \lambda_i g_i$. Jer je G regularna, svi stupci g_i su netrivijalni. Stoga iz $Ag_i = \lambda_i g_i$ zaključujemo da je (g_i, λ_i) vlastiti par za A. Ovo razmatranje vrijedi za sve i, pa su svi parovi (g_i, λ_i), $1 \leq i \leq n$ vlastiti parovi za A. Pritom su svi g_i kao stupci regularne matrice linearno nezavisni.

Neka sada A ima n linearno nezavisnih vlastitih vektora x_i, $1 \leq i \leq n$. Za svaki od njih vrijedi $Ax_i = \lambda_i x_i$, za neke skalare λ_i. Načinimo matricu $X = [x_1, \ldots, x_n]$. Zabog linearne nezavisnosti stupaca, X je regularna matrica. Iz skalara λ_i načinimo dijagonalnu matricu $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. Sada možemo pisati

$$\begin{aligned} AX &= A[x_1, \ldots, x_n] = [Ax_1, \ldots, Ax_n] = [\lambda_1 x_1, \ldots, \lambda_n x_n] \\ &= [x_1, \ldots, x_n] \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \\ &= X\Lambda\,. \end{aligned}$$

odakle, množenjem s desna s X^{-1} dobivamo $A = X\Lambda X^{-1}$. Dakle A dopušta rastav (12.9) s nesingularnom matricom $G = X$. ■

Matrice koje dopuštaju rastav (12.9) zovu se **dijagonalizibilne**. Sljedeći primjer osvjetljava jedan aspekt važnosti dijagonalizibilnih matrica.

Primjer 12.12 *Neka je*

$$F(t) = \begin{bmatrix} f_1(t) \\ \vdots \\ f_n(t) \end{bmatrix} \in \mathbf{R}^n\,, \quad t \in \mathbf{R}$$

Promotrimo sustav linearnih diferencijalnih jednadžbi

$$\frac{dF}{dt} = A\,F\,, \tag{12.10}$$

gdje je $A = (a_{ij})$ matrica i $F = F(t)$ funkcija. Relacija (12.10) se može zapisati u obliku

$$\frac{df_i}{dt}(t) = \sum_{j=1}^{n} a_{ij} f_j(t)\,, \quad 1 \leq i \leq n\,. \tag{12.11}$$

Pretpostavimo prvo da je A dijagonalna, $A = \mathrm{diag}\,(a_1, \ldots, a_n)$, gdje su $a_i \in \mathbf{R}$ dijagonalni elementi od A. Tada (12.11) poprima oblik

$$\frac{df_i}{dt}(t) = a_i\, f_i(t)\,, \quad 1 \leq i \leq n.$$

Lako je provjeriti da je rješenje i-te diferencijalne jednadžbe funkcija

$$f_i(t) = c_i e^{a_i t}\,,$$

pa je ukupno rješenje vektor funkcija

$$F(t) = \begin{bmatrix} c_1 e^{a_1 t} \\ \vdots \\ c_n e^{a_n t} \end{bmatrix}.$$

Lako se vidi da su vektori kanonske baze u $\mathbf{R}^n$, $e_1, \ldots, e_n$ vlastiti vektori od A, pa opće rješenje poprima oblik

$$F(t) = c_1 e^{a_1 t} e_1 + \cdots + c_n e^{a_n t} e_n\,.$$

Ako A nije dijagonalna matrica, ali je dijagonalizibilna,

$$A = U \Lambda U^{-1}, \quad \Lambda = \mathrm{diag}\,(\lambda_1, \ldots, \lambda_n),$$

tada polazna vektorska diferencijalna jednadžba (12.10) poprima oblik

$$\frac{d}{dt} F = U \Lambda U^{-1}\, F\,.$$

Množenjem s lijeva matricom U^{-1} daje

$$U^{-1} \frac{d}{dt} F = \frac{d}{dt} U^{-1} F = \Lambda U^{-1}\, F\,.$$

ili

$$\frac{d}{dt} G = \Lambda G\,, \quad G(t) = U^{-1} F(t),\ t \in \mathbf{R}\,.$$

Tako smo opet dobili vektorsku diferencijalnu jednadžbu s dijagonalnom matricom Λ. Taj slučaj smo već riješili, pa možemo odmah zapisati rješenje

$$G(t) = \begin{bmatrix} c_1 e^{\lambda_1 t} \\ \vdots \\ c_n e^{\lambda_n t} \end{bmatrix}.$$

Jer je $G(t) = U^{-1}F(t)$, *dobivamo*

$$\begin{aligned} F(t) &= UG(t) = U \begin{bmatrix} c_1e^{\lambda_1 t} \\ \vdots \\ c_ne^{\lambda_n t} \end{bmatrix} = [u_1, \ldots, u_n] \begin{bmatrix} c_1e^{\lambda_1 t} \\ \vdots \\ c_ne^{\lambda_n t} \end{bmatrix} \\ &= c_1e^{\lambda_1 t}u_1 + \cdots + c_ne^{\lambda_n t}u_n\,. \end{aligned}$$

Kako vrijedi $Au_i = \lambda_i u_i$, *vidimo da su* (u_i, λ_i) *vlastiti parovi matrice* A. *Dakle, rješenje sustava linearnih diferencijalnih jednadžbi može se dobiti pomoću rješenja matričnog problema vlastitih vrijednosti. Taj zaključak u određenom smislu vrijedi i za matrice koje nisu dijagonalizibilne.* ■

Ne mogu se sve matrice dijagonalizirati pomoću transformacije sličnosti. Prema teoremu 12.11 to mogu tek i samo one za koje postoji pun sistem vlastitih vektora. Kako možemo pokušati dijagonalizirati matricu? Možemo pokušati na sljedeći način.

1. odredi se $\chi(\lambda)$ matrice A;

2. odrede se nultočke $\lambda_1, \ldots, \lambda_n$ kao nultočke od $\chi(\lambda)$;

3. riješe se svi homogeni sustavi $(\lambda_i I - A)v = 0$ i ako se može pronaći ukupno n linearno nezavisnih rješenja, tada se matrica A dade dijagonalizirati transformacijom sličnosti pomoću matrice čiji su stupci ta rješenja.

Ovaj načelni postupak ima smisla za npr. $n = 2, 3, 4$. Za veće n svaka faza postupka postoje vrlo složeni i računski osjetljiv problem. Stoga su se razvile i još uvijek se razvijaju iterativne metode koje polaznu matricu A postepeno pomoću transformacija sličnosti s jednostavnim nesingularnim matricama svađaju na dijagonalni oblik. Jednu takvu metodu za dijagonalizaciju simetrične matrice ćemo upoznati u sljedećem poglavlju.

Važnu klasu dijagonalizibilnih matrica čine tzv. normalne matrice za koje vrijedi $A^*A = AA^*$. U klasu normalnih matrica spadaju sljedeće važne podklase:

Hermitske matrice, za koje vrijedi $A = A^*$. U realnom prostoru $\mathbf{R}^{n\times n}$ njima odgovaraju simetrične matrice za koje vrijedi $A = A^T$;

Antihermitske matrice, za koje vrijedi $A = -A^*$. U realnom prostoru $\mathbf{R}^{n\times n}$ njima odgovaraju antisimetrične matrice za koje vrijedi $A = -A^T$;

Unitarne matrice, za koje vrijedi $A^{-1} = A^*$. U realnom prostoru $\mathbf{R}^{n\times n}$ njima odgovaraju ortogonalne matrice za koje vrijedi $A^{-1} = A^T$.

Sve normalne matrice iz $\mathbf{C}^{n\times n}$ se mogu dijagonalizirati transformacijom sličnosti s unitarnom matricom G. U praksi su najvažnije simetrične matrice, pa ćemo im u sljedećem poglavlju posvetiti dužnu pažnju.

12.2 Matrični polinomi

Pretpostavimo da se matrica A dade dijagonalizirati, dakle da postoji regularna matrica S, takva da je

$$S^{-1}AS = \Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)\,.$$

Uočimo da je

$$\begin{aligned}
\Lambda^2 &= \Lambda\Lambda = \operatorname{diag}(\lambda_1^2, \ldots, \lambda_n^2)\\
\Lambda^3 &= \Lambda^2\Lambda = \operatorname{diag}(\lambda_1^3, \ldots, \lambda_n^3)\\
\cdots & \quad \cdots\\
\Lambda^p &= \Lambda^{p-1}\Lambda = \operatorname{diag}(\lambda_1^p, \ldots, \lambda_n^p)\,.
\end{aligned}$$

Općenitije, ako je $p(\lambda)$ bilo koji polinom, tada je

$$p(\Lambda) = \begin{bmatrix} p(\lambda_1) & & \\ & \ddots & \\ & & p(\lambda_n) \end{bmatrix}.$$

Ako je $A = S\Lambda S^{-1}$, imamo

$$\begin{aligned}
A^2 &= (S\Lambda S^{-1})(S\Lambda S^{-1}) = S\Lambda\Lambda S^{-1} = S\Lambda^2 S^{-1}\\
A^3 &= A^2 A = (S\Lambda^2 S^{-1})(S\Lambda S^{-1}) = S\Lambda^3 S^{-1}\\
\cdots & \quad \cdots\\
A^p &= A^2 A = (S\Lambda^{p-1} S^{-1})(S\Lambda S^{-1}) = S\Lambda^p S^{-1}\,, \qquad (12.12)
\end{aligned}$$

pa odmah slijedi

$$p(A) = Sp(\Lambda)S^{-1} = S \begin{bmatrix} p(\lambda_1) & & \\ & \ddots & \\ & & p(\lambda_n) \end{bmatrix} S^{-1}.$$

Primjer 12.13 *Izračunajte A^n ako je $A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$.*

Rješenje: *Iz karakteristične jednadžbe*

$$\chi(A) = \begin{vmatrix} 2 & -1 \\ -1 & 2 \end{vmatrix} = (\lambda - 2)^2 - 1 = 0\,,$$

slijedi $\lambda_1 = 1$ i $\lambda_2 = 3$. Pripadni vlastiti vektori se nađu kao rješenja sustava $(I - A)x = 0$ i $(3I - A)x = 0$. Rješenja su vektori $v_1 = [1, 1]^T$ i $v_2 = [-1, 1]^T$. Dakle je

$$S = [v_1, v_2] = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}, \text{ pa je } S^{-1} = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}.$$

Jednostavnim množenjem, lako se provjeri da je $A = S\Lambda S^{-1}$. Sada se prisjetimo formule (12.12) za potenciju matrice, pa jednostavno dobijemo

$$\begin{aligned} A^n &= S\Lambda^n S^{-1} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & \\ & 3 \end{bmatrix}^n \frac{1}{2} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \\ &= \frac{1}{2} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1^n & \\ & 3^n \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 + 3^n & 1 - 3^n \\ 1 - 3^n & 1 + 3^n \end{bmatrix} \end{aligned}$$

■

Na osnovu formule za polinom, funkcija matrice se definira na slični način. Npr. ako je : $\mathbf{R} \to \mathbf{R}$ beskonačno diferencijabilna funkcija i ako A ima realne vlastite vrijednosti, $f(A)$ se može ovako definirati

$$f(A) = Sf(\Lambda)S^{-1} = S \begin{bmatrix} f(\lambda_1) & & \\ & \ddots & \\ & & f(\lambda_n) \end{bmatrix} S^{-1}.$$

Primjer 12.14 *Izračunajte* $\sin(A)$ *za* A *iz prethodnog primjera.*

Rješenje: *Koristeće matricu* S *koja preko transformacije sličnosti dijagonalizira matricu* A *(i koja je izračunata u prethodnom primjeru) odmah dobijemo*

$$\begin{aligned}\sin(A) &= S\sin(\Lambda)S^{-1} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \sin(1) & \\ & \sin(3) \end{bmatrix} \frac{1}{2} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \\ &= \frac{1}{2} \begin{bmatrix} \sin(1)+\sin(3) & \sin(1)-\sin(3) \\ \sin(1)-\sin(3) & \sin(1)+\sin(3) \end{bmatrix}.\end{aligned}$$

■

Neka je $A \in \mathbf{C}^{n\times n}$. Tada se mogu izračunati potencije $A^0 = I$, $A^1 = A$, $A^2 = A \cdot A$, $\ldots$, $A^k = A^{k-1} \cdot A$, $\ldots$ Matrice I, A, A^2, $\ldots$ su elementi vektorskog prostora $\mathbf{C}^{n\times n}$ reda n^2, pa u nizu I, A, A^2, $\ldots$ nema više od n^2 linearno nezavisnih matrica. Stoga polazeći od I, A, A^2, $\ldots$ možemo naći prvu potenciju A^r, koja se može prikazati kao linearna kombinacija prethodnih potencija. Dakle, možemo pisati

$$A^r = \mu_1 A^{r-1} + \mu_2 A^{r-2} + \cdots + \mu_{r-1} A + \mu_r,$$

gdje su μ_i skalari. Ako definiramo polinom

$$\mu(z) = z^r - \mu_1 z^{r-1} - \mu_2 z^{r-2} - \cdots - \mu_{r-1} z - \mu_r \,,$$

on ima svojstvo da je polinom najmanjeg stupnja koji poništava matricu A, tj. vrijedi $\mu(A) = 0$ i da je koeficijent uz najvišu potenciju jedan. Ovi uvjeti najmanjeg stupnja i normiranosti vodećeg koeficijenta, osiguravaju jedinstvenost takvog polinoma, pa se on zove **minimalni polinom** matrice A. On ima važno svojstvo da dijeli svaki polinom koji poništava matricu A. Postavlja se pitanje gornje međe za stupanj minimalnog polinoma. Na to pitanje odgovaraju sljedeći primjer i teorem.

Primjer 12.15 *Neka je* $A = J_n(0)$, *gdje je* $J_n(\nu)$ *kao u primjeru 12.9.*

Tada je

$$A^2 \;=\; \begin{bmatrix} 0 & 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ & & & & \cdots & & & \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \end{bmatrix}.$$

Vidimo da se matrica A^2 od nul-matrice tazlikuje u tome što ima na trećoj sporednoj dijagonali jedinice. Na slični način se provjeri da A^3 ima samo na četvrtoj sporednoj dijegonali jedinice. Nastavljajući, lako nalazimo da će se A^{n-1} razlikovati od nul-matrice tek u jedinici na $(1,n)$ poziciji, a A^n će biti nul-matrica. Kako je $A = J_n(\nu) - \nu I_n$, iz $(J_n(\nu)-\nu I_n)^n = 0$ se dobije da polinom $\mu(z) = (z-\nu)^n$ poništava $J_n(\nu)$ za bilo koje ν. Lako se vidi da su potencije od $J_n(\nu)$ linearno nezavisne (jer $J_n^k(\nu)$ ima netrivijalne elemente točno na prvih $k+1$ sporednih dijagonala). Stoga je $\mu(z)$ minimalni polinom za $J_n(\nu)$ i on je stupnja n. ■

Iz primjera 12.15 vidimo da stupanj minimalnog polinoma može biti n, pa zaključujemo da je gornja međa za njega između n i n^2. Sljedeći teorem pokazuje da je upravo n gornja međa za stupanj minimalnog polinoma.

Teorem 12.16 (Hamilton-Cayley) *Neka je $A \in \mathbf{C}^{n\times n}$ i neka je $\chi(z)$ karakteristični polinom od A. Tada je $\chi(A) = 0$, tj. matrica poništava svoj karakteristični polinom.*

Dokaz: Neka je $B(z) = \operatorname{adj}(zI - A)$ adjunkta matrice $zI - A$. Ona zadovoljava

$$B(z)(zI - A) = \det(zI - A)I\,. \tag{12.13}$$

Elementi od $B(z)$ su algebarski komplementi elemenata matrice $zI-A$. Oni su dakle polinomi stupnja $\le n-1$ u z. Stoga $B(z)$ možemo napisati u obliku

$$B(z) = z^{n-1}B_0 + z^{n-2}B_1 + \cdots + zB_{n-2} + B_{n-1}\,.$$

Koristeći taj izraz u relaciji (12.13), dobijemo

$$\begin{aligned}\left(z^{n-1}B_0 + z^{n-2}B_1 + \cdots + zB_{n-2} + B_{n-1}\right)(zI - A)\\ = (z^n - \sigma_1 z^{n-1} - \cdots - \sigma_n)I\,,\end{aligned}$$

gdje je $\det(zI - A) = \chi(z) = z^n - \sigma_1 z^{n-1} - \cdots - \sigma_n$ karakteristični polinom. Izjednačavajući članove uz istu potenciju od z, dobivamo

$$\begin{aligned}B_0 &= I\\ B_1 - B_0A &= -\sigma_1 I\\ B_2 - B_1A &= -\sigma_2 I\\ &\vdots\\ B_{n-1} - B_{n-2}A &= -\sigma_{n-1} I\\ -B_{n-1}A &= -\sigma_n I\,.\end{aligned}$$

Pomnožimo li prvu jednadžbu s desna s A^{n-1}, drugu s A^{n-2}, itd. zadnju s $A^0 = I_n$ i zatim zbrojimo matrice na lijevoj i desnoj strani, dobijemo

$$0 = A^n - \sigma_1 A^{n-1} - \cdots - \sigma_{n-1}A - \sigma_n I = \chi(A)\,.$$

■

Kako karakteristični polinom poništava matricu, on je djeljiv (to znači djeljiv bez ostatka) s minimalnim polinomom. Dakle, minimalni polinom ima stupanj koji je manji ili jednak stupnju karakterističnog polinoma.

Iz Hamilton-Cayleyjeva teorema smo zaključili da minimalni polinom dijeli karakteristični polinom. To znači da nultočke minimalnog polinoma leže u skupu vlastitih vrijednosti matrice. Pomnija analiza pokazuje da je svaka vlastita vrijednost matrice nultočka minimalnog polinoma. To znači da možemo pisati

$$\begin{aligned}\mu(z) &= (z-\lambda_1)^{m_1}(z-\lambda_2)^{m_2}\cdots(z-\lambda_p)^{m_p},\\ \chi(z) &= (z-\lambda_1)^{n_1}(z-\lambda_2)^{n_2}\cdots(z-\lambda_p)^{n_p}\end{aligned}$$

pri čemu je p broj međusobno različitih vlastitih vrijednosti matrice. Pritom uvijek vrijedi $m_i \le n_i$ za sve $1 \le i \le p$, a za mnoge matrice vrijedi jednakost $m_i = n_i$ za sve $1 \leq i \leq p$. Npr. jednakost vrijedi za sve dijagonazibilne matrice, ali i za matricu $J_n(\nu)$. Nejednakost vrijedi npr. za blok dijagonalnu matricu

$$G = \begin{bmatrix} J_r(\nu) & \\ & J_{n-r}(\nu) \end{bmatrix}, \quad \nu \in \mathbf{C}\,,\ 2 \leq r \leq n-2\,.$$

Provjerite to npr. za slučaj $\nu = 1$, $n = 5$, $r = 3$.

Jordanova forma matrice

Iz primjera 12.9 se vidi da ima matrica koje nemaju puni sistem (tj. bazu) vlastitih vektora. Iz teorema 12.11 zaključujemo da se takove matrice ne mogu dijagonalizirati pomoću transformacije sličnosti. Da bi dobili uvid u doseg transformacija sličnosti kao alata za rješavnje problema vlastitih vrijednosti navest ćemo bez dokaza opći teorem o redukciji kvadratne matrice na oblik koji je onoliko blizak dijagonalnom obliku, koliko matrica dopušta.

Matrica oblika

$$J_k(\nu) = \begin{bmatrix} \nu & 1 & 0 & \cdots & 0 & 0 \\ 0 & \nu & 1 & \cdots & 0 & 0 \\ 0 & 0 & \nu & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \nu & 1 \\ 0 & 0 & 0 & \cdots & 0 & \nu \end{bmatrix} \in \mathbf{C}^{k\times k}$$

se zove elementarna Jordanova klijetka (ili blok) dimenzije k. Pomoću elementarnih Jordanovih klijetki je građena Jordanova klijetka (ili blok)

$$J(\nu) = \begin{bmatrix} J_{k_1}(\nu) & & & \\ & J_{k_2}(\nu) & & \\ & & \ddots & \\ & & & J_{k_r}(\nu) \end{bmatrix} \in \mathbf{C}^{l\times l}\,, \tag{12.14}$$

pri čemu je $k_1 \geq k_2 \geq \cdots \geq k_r \geq 1$ i $k_1 + k_2 + \cdots + k_r = l \leq n$.

Teorem 12.17 (Jordan) *Neka je $A \in \mathbf{C}^{n\times n}$ i neka je $\lambda_1, \ldots, \lambda_p$ zadani uređaj međusobno različitih vlastitih vrijednosti od A, čije su* *algebarske kratnosti* $n_1, \ldots, n_p$*, respektivno. Tada postoji regularna matrica* S*, takva da je*

$$S^{-1}AS = \begin{bmatrix} J(\lambda_1) & & & \\ & J(\lambda_2) & & \\ & & \ddots & \\ & & & J(\lambda_p) \end{bmatrix} \begin{array}{l} \}n_1 \\ \}n_1 \\ \vdots \\ \}n_p \end{array} . \tag{12.15}$$

Pritom je $n_1 + \cdots + n_p = n$*.* ■

Teorem 12.17 kaže da se svaka matrica može pomoću transformacije sličnosti svesti na oblik koji je blizak dijagonalnom. Naime, svaki Jordanov blok $J(\lambda_k)$ je oblika (12.14), pa za svako $1 \leq k \leq p$, postoji rastav $n_k = m_{k,1} + \cdots + m_{k,r_k}$, pri čemu je $m_{k,j}$ dimenzija j-te po redu elementarne Jordanove klijetke u $J(\lambda_k)$. U teoremu može biti $p = 1$, $r_1 = 1$ i tada je Jordanov oblik od A jedna elementarna Jordanova klijetka reda n. U drugoj krajnosti kad je $p = n$, mora za svako k biti $n_k = 1$, pa je $r_k = 1$ i $m_{k,1} = 1$, što znači da je Jordanova klijetka $J(\lambda_k)$ jedna elementarna Jordanova klijetka. Dakle, sve jordanove klijetke u Jordanovoj formi matrice su reda 1, pa je matrica dijagonalizibilna. U ostalim slučajevima će općenito vrijediti

$$\begin{aligned} n &= n_1 + \cdots + n_p \\ n_1 &= m_{1,1} + \cdots + m_{1,r_1} \\ \cdots & \quad \cdots \\ n_p &= m_{p,1} + \cdots + m_{p,r_p} . \end{aligned}$$

Za broj r_k elementarnih Jordanovih klijetki unutar Jordanove klijetke $J(\lambda_k)$ vrijedi ograničenje $1 \leq r_k \leq n$, a za dimenzije elementarnih Jordanovih klijetki $m_{k,j}$ vrijedi $0 \leq m_{k,j} \leq n_k \leq n$.

Zbog $\det(zI - S^{-1}AS) = \det(S^{-1}(zI - A)S) = \det(zI - A)$ zaključujemo da je n_i algebarska kratnost od λ_i. Kako elementarne Jordanove klijetke imaju samo jedan vlastiti vektor e_1, zaključujemo da je geometrijska kratnost od λ_i broj r_i. Iz teorema 12.17 također slijedi da je minimalni polinom od A,

$$\mu(z) = (z - (\lambda_1)^{m_{1,1}}(z - (\lambda_1)^{m_{2,1}} \cdots (z - (\lambda_1)^{m_{p,1}}$$

gdje je za svako k, $m_{k,1}$ najveći red elementarne klijetke unutar Jordanove klijetke $J(\lambda_k)$.

Lekcija 13

Dijagonalizacija simetrične matrice

Vidjeli smo da se transformacijom sličnosti matrica može svesti na Jordanovu formu, oblik koji je dosta blizak dijagonalnoj fromi. Zbog teoretskih, a posebno numeričkih razloga, posebno su pogodne transformacije sličnosti s unitarnom matricom. Međutim ako zahtijevamo da sličnost bude načinjena pomoću unitarne matrice, tada je doseg dijagonalizacije slabiji. Pomoću unitarne transformacije sličnosti matrica se može svesti na trokutasti oblik.

Teorem 13.1 (Schur) *Za proizvoljnu matricu A postoji unitarna matrica U, takva da je $T = U^*AU$ gornja trokutasta matrica. Matrica U može biti izabrana tako da se na dijagonali matrice T pojave vlastite vrijednosti od A u bilo kojem zadanom poretku.*

Dokaz: Dokaz se provodi indukcijom po redu matrice n, ali ga nećemo dokazivati. ∎

Prethodni teorem ima neke vrlo važne posljedice, kao što je naredni teorem koji kaže da je Schurova dekompozija normalnih matrica dijagonalna. Sjetimo se, matrica A je normalna ako je $A^*A = AA^*$. Pokažimo da je unitarno slična matrica normalnoj matrici opet normalna. Neka je A normalna i $B = U^*AU$ za neku unitarnu matricu U. Tada je

$$B^*B = U^*A^*UU^*AU = U^*A^*AU = U^*AA^*U = U^*AUU^*A^*U = BB^*,$$

pa je B normalna. Sljedeći teorem pokazuje da se normalne matrice mogu dijegonalizirati pomoću uniterne transformacije sličnosti.

Teorem 13.2 *Ako je A normalna matrica i $T = U^*AU$ Schurova dekompozicija od A, tada je T dijagonalna matrica.*

Dokaz: Dokaz ćemo provesti indukcijom. Za $n = 1$ tvrdnja je trivijalna, stoga pretpostavimo da je tvrdnja istinita za sve normalne matrice reda manjeg od n, $n > 1$. Neka je A matrica reda n i $T = U^*AU$ njena Schurova dekompozicija. Promotrimo sljedeću particiju od T

$$T = \begin{bmatrix} \tau & t^* \\ 0 & T_0 \end{bmatrix}.$$

Jer je T unitarno slična normalnoj matrici ona je normalna, pa iz $T^*T = TT^*$ dobivamo

$$\begin{bmatrix} \bar{\tau} & 0 \\ t & T_0^* \end{bmatrix} \begin{bmatrix} \tau & t^* \\ 0 & T_0 \end{bmatrix} = \begin{bmatrix} \mid \tau \mid^2 & \bar{\tau} t^* \\ \tau\, t & t\, t^* + T_0^* T_0 \end{bmatrix}$$
$$\begin{bmatrix} \tau & t^* \\ 0 & T_0 \end{bmatrix} \begin{bmatrix} \bar{\tau} & 0 \\ t & T_0^* \end{bmatrix} = \begin{bmatrix} \mid \tau \mid^2 + t^* t & t^* T_0^* \\ T_0 t & T_0 T_0^* \end{bmatrix}.$$

Gledajući elemente na pozicijama $(1,1)$ i $(2,2)$, dobivamo

$$|\tau|^2 = |\tau|^2 + t^*t,$$

i

$$tt^* + T_0^*T_0 = T_0T_0^*.$$

Prva jednadžba povlači $t = 0$, dok iz druge slijedi da je T_0 normalna matrica. Jer je T Schurova dekompozicija od A, T_0 je gornje-trokutasta, a kako je normalna i reda manjeg od n ona je po pretpostavci indukcije dijagonalna. To znači da je T dijagonalna matrica. ∎

Jedna od posljedica ovog teorema jest tvrdnja da normalna matrica reda n ima sustav od n ortonormiranih (ortogonalnih i normiranih tj. Euklidske norme jedan) vlastitih vektora koji razapinju $\mathbf{C}$.

Dvije najznačajnije klase normalnih matrica su unitarne i hermitske matrice. Sljedeći korolar govori o karakterizaciji tih matrica pomoću njihovih vlastitih vrijednosti.

Korolar 13.3 *Unitarna matrica je normalna matrica čije vlastite vrijednosti imaju modul jedan. Hermitska matrica je normalna matrica koja ima realne vlastite vrijednosti.*

Dokaz: Neka je U unitarna matrica. Jer je ona i normalna, po teoremu 13.2 postoji unitarna matrica V, takva da je

$$\Lambda \equiv \mathrm{diag}\,(\lambda_1, ..., \lambda_n) = V^* U\, V,$$

gdje su λ_i, $1 \leq i \leq n$ vlastite vrijednosti matrice U. Jer je U unitarna i jer je produkt unitarnih matrica unitarna matrica, Λ je unitarna, pa vijedi $\bar{\Lambda}\Lambda = \Lambda\bar{\Lambda} = I$. To povlači

$$|\lambda_i|^2 = 1 \quad 1 \leq i \leq n,$$

pa su vlastite vrijednosti od U jediničnog modula.

Obratno, ako je U normalna matrica s vlastitim vrijednostima koje imaju modul jedan, tada je prema teoremu 13.2, $U = V\Lambda V^*$, za neku unitarnu matricu V i $\Lambda = \mathrm{diag}\ (\lambda_1, \ldots, \lambda_n)$. Jer su sve vlastite vrijednosti modula jedan, Λ je unitarna, a jer je produkt unitarnih matrica unitarna matrica, U je unitarna.

Pretpostavimo sada da je H hermitska matrica. U tom slučaju, po teoremu 13.2 postoji unitarna matrica W takva da je

$$M \equiv \mathrm{diag}\,(\mu_1, ..., \mu_n) = W^* H\, W,$$

gdje su μ_i, $i = 1, ..., n$ vlastite vrijednosti matrice H. Zbog $H^* = H$, slijedi $M^* = M$, odnosno $\bar{\mu}_i = \mu_i$, $1 \leq i \leq n$. To znači da su sve vlastite vrijednosti μ_i normalne matrice H realni brojevi.

Obratno, neka je H normalna matrica s realnim vlastitim vrijednostima i neka je prema teoremu 13.2 $H = WDW^*$. Jer je W unitarna i D dijagonalna i realna, lako se provjeri $H^* = (W^*)^* DW^* = WDW^* = H$, tj. H je hermitska. ∎

Iz korolara 13.3 odmah se vidi da se hermitska matrica H može prikazati u obliku

$$H = U\Lambda U^*, \tag{13.1}$$

gdje je $U = [u_1, \ldots, u_n]$ unitarna matrica čiji stupci su vlastiti vektori od H, a Λ je dijagonalna matrica s realnim vlastitim vrijednostima od H na dijagonali. Taj oblik zovemo *spektralna dekompozicija* od H. Ponekad ćemo koristiti i zapis

$$H = \sum_i \lambda_i u_i u_i^*, \tag{13.2}$$

koji je ekvivalentan s (13.1). To odmah slijedi iz zadatka 4.10(v) u kojem uzmite $A = U\Lambda$, $B = U^*$. Pomoću formule (13.2) možemo proširiti pojam skalarne funkcije na hermitske matrice. To se čini na slijedeći način

$$\varphi(H) = \sum_i \varphi(\lambda_i) u_i u_i^*. \tag{13.3}$$

Uočimo da formula (13.3) vrijedi za polinome, a vrijedi i za neprekidne funkcije realnog argumenta. Jedna od načešće korištenih funkcija je drugi korijen. Ako je H pozitivno definitna matrica, onda se drugi korijen matrice H dobiva formulom

$$H^{\frac{1}{2}} = \sum_i \lambda_i^{\frac{1}{2}} u_i u_i^*.$$

Ako je Hermitska matrica realna, zove se simetrična. Ona se može dijagonalizirati pomoću transformacije sličnosti s ortogonalnom matricom, pa se njena spektralna dekompozicija zapisuje u obliku

$$A = Q \Lambda Q^T ,$$

gdje je Λ dijagonalna matrica vlastitih vrijednosti od A, a Q je ortogonalna matrica čiji su stupci vlastiti vektori od A.

13.1 Jacobijeva metoda

Postoji veći broj numeričkih metoda za rješavanje matričnog problema vlastitih vrijednosti. Izbor metode ovisi o karakteristikama matrice (normalnost, simetričnost, pozitivna definitnost, tridijagonalnost,...) i o tome da li se žele naći sve vlastite vrijednosti i vektori ili samo neke vlastite vrijednosti i samo neki vlastiti vektori (npr. ako je dimenzija matrice velika, obično se traži manji broj najvećih ili najmanjih vlastitih vrijednosti i pripadnih vektora). U praksi se najčešće pojavljuju simetrične matrice, pa ćemo mi opisati jednu "staru", ali vrlo pouzdanu metodu za računanje svih vlastitih vrijednosti i vektora simetrične matrice.

Još 1846. Jacobi je predložio metodu za dijagonalizaciju simetrične matrice pomoću niza rotacijskih transformacija sličnosti. Kako je metoda bila računski zahtjevna bila je zaboravljena sve do pojave modernih računala. Godine 1949. metodu su ponovo otkrili Goldstine, Murray i von Neumann. Kasnije, tijekom pedestih i šezdesetih godina metoda se počela sustavno proučavati. Prvi značajniji radovi su od Forsythea i Henricia (globalna konvergencija), Schönhagea i Wilkinsona (kvadratična konvergencija u slučaju jednostrukih vlastitih vrijednosti) i Rutishousera (sofisticirani numerički algoritam). Nakon 1970. metoda je pala u sjenu QR metode koja je nekoliko puta brža. Iza 1980., s razvojem paralelnih računala metoda se ponovo koristi za računanje vlastitih vrijednosti i vektora. U posljednje vrijeme je otkriveno da je metoda točnija od QR i ostalih metoda pa se koristi kad je točnost izračunatih vlastitih vrijednosti vektora važna. Postoje generalizacije i modifikacije metode za različite tipove matrica kao i za rješavanje nekih drugih matričnih problema.

Jacobijeva metoda je iterativni proces oblika

$$A^{(k+1)} = R_k^T A^{(k)} R_k, \quad k \geq 1; \quad A^{(1)} = A, \tag{13.4}$$

pri čemu su R_k ravninske rotacije. Ravninska rotacija R

$$R(p,q,\phi) = \left[\begin{array}{ccccccccccc}
1 & & & & & & & & & & \\
 & \ddots & & & & & & & & & \\
 & & 1 & & & & & & & & \\
 & & & \cos\phi & & & & -\sin\phi & & & \\
 & & & & 1 & & & & & & \\
 & & & & & \ddots & & & & & \\
 & & & & & & 1 & & & & \\
 & & & \sin\phi & & & & \cos\phi & & & \\
 & & & & & & & & 1 & & \\
 & & & & & & & & & \ddots & \\
 & & & & & & & & & & 1
\end{array}\right]$$

u (p,q) ravnini je matrica koja se od jedinične matrice I_n razlikuje tek u četiri elementa na pozicijama (p,p), (p,q), (q,p) i (q,q). Ti su elementi redom $\cos\phi$, $-\sin\phi$, $\sin\phi$, $\cos\phi$. Takvu ćemo matricu još označavati s $R(p,q;\phi)$. Dakle, Transformaciju $A^{(k)} \mapsto A^{(k+1)}$ zovemo k-ti korak metode. Svrha k-tog koraka je načiniti matricu $A^{(k+1)}$ više dijagonalnom nego što je $A^{(k)}$. Da bi odstupanje od dijagonalne forme mogli mjeriti definiramo mjeru

$$S(A) = \|\Omega(A)\|_F = \sqrt{\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} \mid a_{ij} \mid^2},$$

pri čemu sa $\Omega(A)$ označavamo izvandijagonalni dio matrice $\Omega(A) = A - \mathrm{diag}(A)$, a $\|\cdot\|_F$ je Frobeniusova norma (korijen iz sume kvadrata svih elemenata).

13.1.1 Dijagonalizacija simetrične matrice reda 2

Neka je $n = 2$. Tada se Jacobijeva metoda sastoji od jednog koraka koji možemo zapisati kao

$$\begin{bmatrix} a'_{11} & a'_{12} \\ a'_{12} & a'_{22} \end{bmatrix} = \begin{bmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix}. \tag{13.5}$$

Očito, maksimalna redukcija mjere $S(A)$ će se dogoditi ako izborom kuta ϕ uspijemo poništiti element a'_{12}. Izrazimo element a'_{12} pomoću elemenata matrice A i trigonometrijskih funkcija kuta ϕ. Korištenjem jediničnih vektora e_1 i e_2 (stupaca jedinične matrice I_2), dobivamo

$$\begin{aligned} a'_{12} &= e_1^T R^T A R e_2 = [\cos\phi\ ,\ \sin\phi] \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} \begin{bmatrix} -\sin\phi \\ \cos\phi \end{bmatrix} \\ &= [a_{11}\cos\phi + a_{12}\sin\phi\ ,\ a_{12}\cos\phi + a_{22}\sin\phi] \begin{bmatrix} -\sin\phi \\ \cos\phi \end{bmatrix} \\ &= a_{12}(\cos^2\phi - \sin^2\phi) - (a_{11} - a_{22})\cos\phi\sin\phi \\ &= a_{12}\cos 2\phi - \frac{1}{2}(a_{11} - a_{22})\sin 2\phi. \end{aligned}$$

Uvjet $a'_{12} = 0$ povlači

$$\tan 2\phi = \frac{2a_{12}}{a_{11} - a_{22}} \tag{13.6}$$

Kao interval za 2ϕ obično se uzima $[-\frac{\pi}{2}\,,\,\frac{\pi}{2}]$ pa se kut ϕ bira u intervalu $[-\frac{\pi}{4}\,,\,\frac{\pi}{4}]$. Jer je element a'_{12} poništen, jasno je da vrijedi

$$S^2(A') = S^2(A) - 2a_{12}^2. \tag{13.7}$$

Sada se postavlja problem kako na jednostavan i numerički točan način odrediti elemente matrice R, tj. $\cos\phi$ i $\sin\phi$. Pišemo li

$$t = \tan\phi, \quad \lambda = 2a_{12}\cdot \mathrm{sign}(a_{11} - a_{22}), \quad \mu = \mid a_{11} - a_{22} \mid,$$

i iskoristimo li formulu za tangens dvostrukog kuta,

$$\tan\, 2\phi = \frac{2t}{1-t^2}$$

dobivamo

$$\frac{2t}{1-t^2} = \frac{\lambda}{\mu},$$

dakle kvadratnu jednadžbu u t,

$$\lambda\, t^2 + 2\mu\, t - \lambda = 0.$$

Rješenja su dana formulom, koja nakon kraćenja sa dva ima oblik

$$t_{1,2} = \frac{-\mu \pm \sqrt{\mu^2 + \lambda^2}}{\lambda}.$$

Jer ϕ leži u intervalu $[-\frac{\pi}{4}\,,\,\frac{\pi}{4}]$, t i $\tan\,2\phi$ imaju isti predznak. Kako je μ nenegativan, $\tan\,2\phi$ ima isti predznak kao i λ. Dakle, polazeći od zahtjeva da t i λ imaju isti predznak dolazimo do formule

$$t = \frac{-\mu + \sqrt{\mu^2 + \lambda^2}}{\lambda}.$$

Ova formula je za male kvocijente λ/μ, pri korištenju konačne aritmetike (npr. na kalkulatoru ili računalu) nestabilna, jer vodi na oduzimanje gotovo jednakih brojeva u brojniku. Stoga dolazi do kraćenja vodećih znamenaka, pa netočnost u brojevima, koja dolazi od zaokruživanja međurezultata, dolazi bliže vodećim nulama. To znači da brojnik može biti netočno izračunat. Da bismo to izbjegli "racionalizirajmo" brojnik, tj. pomnožimo razlomak s $(\sqrt{\mu^2+\lambda^2}+\mu)/(\sqrt{\mu^2+\lambda^2}+\mu)$. Dobijemo algoritamski stabilnu formulu

$$t = \frac{\lambda}{\mu + \sqrt{\mu^2+\lambda^2}},$$

Iz tangensa se lako izračunaju cosinus i sinus

$$\cos\phi = \frac{1}{\sqrt{1+t^2}}, \quad \sin\phi = \frac{t}{\sqrt{1+t^2}} = \sin\phi = t * \cos\phi.$$

Koristeći formulu (13.6) mogu se dobiti jednostavne formule za dijagonalne elemente a'_{11} i a'_{22}. Ako je $a_{12} = 0$, kut $\phi = 0$ pa je $R = I_2$, a to znači $A' = A$. Ako je $a_{12} \neq 0$, za a'_{11} imamo

$$\begin{aligned}
a'_{11} &= e_1^T R^T A R e_1 = [\cos\phi\ ,\ \sin\phi] \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} \begin{bmatrix} \cos\phi \\ \sin\phi \end{bmatrix} \\
&= [a_{11}\cos\phi + a_{12}\sin\phi\ ,\ a_{12}\cos\phi + a_{22}\sin\phi] \begin{bmatrix} \cos\phi \\ \sin\phi \end{bmatrix} \\
&= a_{11}\cos^2\phi + a_{22}\sin^2\phi + 2a_{12}\cos\phi\sin\phi \\
&= a_{11} + (a_{22}-a_{11})\sin^2\phi + 2a_{12}\cos\phi\sin\phi \\
&= a_{11} + a_{12}\tan\phi(\frac{a_{22}-a_{11}}{a_{12}}\sin\phi\cos\phi + 2\cos^2\phi) \\
&= a_{11} + a_{12}\tan\phi(-2\text{cotan}\,2\phi \cdot \frac{1}{2}\sin\,2\phi + 2\cos^2\phi) \\
&= a_{11} + a_{12}\tan\phi(-\cos\,2\phi + 2\cos^2\phi) \\
&= a_{11} + a_{12}\tan\phi(\sin^2\phi + \cos^2\phi) = a_{11} + a_{12}\tan\phi.
\end{aligned}$$

Na slični način (ili korištenjem činjenice da je trag (A') =trag A), lako se dobije $a'_{22} = a_{22} - a_{12}\tan\phi$. Povežimo dobivene formule u algoritam

Algoritam 13.4 *Dana je simetrična matrica A reda dva. Sljedeći algoritam računa elemente rotacije R i matrice $A' = R^T A R$ nakon jednog Jacobijevog koraka.*

$$\begin{aligned}
\textit{Ako je } a_{12} &= 0 \textit{ stavi se } R = I_2,\ A' = A \\
\textit{Ako je } a_{12} &\neq 0 \textit{ računa se} \\
\lambda &= 2a_{12} \cdot \textit{sign}(a_{11} - a_{22})
\end{aligned}$$

$$\begin{aligned}
\mu &= \mid a_{11} - a_{22} \mid \\
\nu &= \sqrt{\lambda^2 + \mu^2} \\
t &= \frac{\lambda}{\mu + \nu} \\
c &= \frac{1}{\sqrt{1 + t^2}} \\
s &= t * c \\
a'_{11} &= a_{11} + t \cdot a_{12} \\
a'_{22} &= a_{22} - t \cdot a_{12} \\
a'_{12} &= 0
\end{aligned}$$

$$\textit{Vlastitoj vrijednosti} \quad a'_{11} \quad \textit{pripada vlastiti vektor} \begin{bmatrix} c \\ s \end{bmatrix}$$

$$\textit{Vlastitoj vrijednosti} \quad a'_{22} \quad \textit{pripada vlastiti vektor} \begin{bmatrix} -s \\ c \end{bmatrix}$$

Primjer 13.5 *Neka je*

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

Algoritam 13.4 daje $\lambda = 2$*,* $\mu = 0$*,* $\nu = 2$*,* $t = 1$*,* $c = s = \sqrt{2}/2$*,* $a'_{11} = 1 + 1 \cdot 1 = 2$*,* $a'_{22} = 1 - 1 \cdot 1 = 0$ *i vlastitoj vrijednosti* 2 *odgovara vlastiti vektor* $c \cdot [1, 1]^T$ *dok vlastitoj vrijednosti* 0 *odgovara vlastiti vektor* $c \cdot [-1, 1]^T$*.*

13.1.2 Jedan korak Jacobijeve metode na matrici reda n

Neka je A simetrična matrica reda n i neka je (p, q) par indeksa, takav da je $1 \leq p < q \leq n$. Neka je $A' = R^T A R$ gdje je $R = R(p, q; \phi)$. Prikažimo matricu A kao zbroj

$$A = A_1 + A_2, \quad A_1 = \operatorname{diag}(A) + a_{pq} e_p e_q^T + a_{qp} e_q e_p^T \, ,$$

gdje su e_j, $1 \leq j \leq n$ stupci jedinične matrice I_n. Dakle,

$$A = \begin{bmatrix} a_{11} & & & & & & \\ & \ddots & & & & & \\ & & a_{pp} & & a_{pq} & & \\ & & & \ddots & & & \\ & & a_{qp} & & a_{qq} & & \\ & & & & & \ddots & \\ & & & & & & a_{nn} \end{bmatrix} + \begin{bmatrix} 0 & & & & & & \\ & \ddots & & & & & \\ & & 0 & & 0 & & \\ & & & \ddots & & & \\ & & 0 & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 0 \end{bmatrix}$$

Jer je

$$A' = R^T AR = R^T A_1 R + R^T A_2 R,$$

oduzimanjem dijagonalnog dijela matrice, dobivamo

$$A' - \mathrm{diag}(A') = R^T A_1 R - \mathrm{diag}(R^T A_1 R) + R^T A_2 R$$

Uzimanjem Frobeniusove norme matrice na lijevoj i desnoj strani, kvadriranjem tih normi, zaključujemo

$$\begin{aligned} \|A' - \mathrm{diag}(A')\|_F^2 &= \|R^T A_1 R - \mathrm{diag}(R^T A_1 R)\|_F^2 + \|R^T A_2 R\|_F^2 \\ &= |\ a'_{pq}\ |^2 + |\ a'_{qp}\ |^2 + \|R^T A_2 R\|_F^2 \\ &= 2\ |\ a'_{pq}\ |^2 + \|R^T A_2 R\|_F^2. \end{aligned}$$

Poznato je da je Frobeniusova norma invarijantna u odnosu na ortogonalne transformacije. To znači $\|Q_1 B Q_2\|_F = \|B\|_F$ za bilo koje ortogonalne matrice. Da bi to pokazali polazimo od definicije norme, $\|B\|_F^2 = \mathrm{trag}(B^* B)$. Kako za trag vrijedi $\mathrm{trag}(AB) = \mathrm{trag}(BA)$, imamo

$$\begin{aligned} \|Q_1 B Q_2\|_F^2 &= \mathrm{trag}(Q_2^* B^* Q_1^* Q_1 B Q_2) = \mathrm{trag}(Q_2^* B^* B Q_2) \\ &= \mathrm{trag}(Q_2 Q_2^* B^* B) = \mathrm{trag}(B^* B) = \|B\|_F^2. \end{aligned}$$

Primjenom na matricu$R^T A_2 R$, imamo

$$\begin{aligned} \|R^T A_2 R\|_F^2 &= \|A_2\|_F^2 = \|A - \mathrm{diag}(A)\|_F^2 - (|\ a_{pq}\ |^2 + |\ a_{qp}\ |^2) \\ &= S^2(A) - 2\ |\ a_{pq}\ |^2\ . \end{aligned}$$

Zbog $S^2(A') = \|A' - \mathrm{diag}(A')\|_F^2$, imamo

$$S^2(A') = S^2(A) + 2(|\ a'_{pq}\ |^2 - |\ a_{pq}\ |^2). \tag{13.8}$$

Vidimo da će se $S(A)$ maksimalno reducirati ako odaberemo kut ϕ tako da bude $a'_{pq} = 0$. Lako je vidjeti da i sada vrijedi relacija (13.5) uz uvjet da se indeksi 1 i 2 zamijene s p i q, respektivno. Odatle zaključujemo da vrijedi

$$a'_{pq} = a_{pq} \cos 2\phi - \frac{1}{2}(a_{pp} - a_{qq}) \sin 2\phi,$$

pa je kut ϕ određen formulom

$$\tan 2\phi = \frac{2a_{pq}}{a_{pp} - a_{qq}}, \tag{13.9}$$

a redukcija odstupanja od dijagonalne forme s

$$S^2(A') = S^2(A) - 2 \mid a_{pq} \mid^2 . \tag{13.10}$$

Za računanje $\cos\phi$ i $\sin\phi$ koristi se isti algoritam kao i u slučaju matrice reda dva, uz očitu zamjenu indeksa. Promotrimo sada kako se vrše množenja s R^T slijeva odnosno s R zdesna.

Za množenje slijeva koristi se relacija

$$\begin{bmatrix} b_p^T \\ b_q^T \end{bmatrix} = \begin{bmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} a_p^T \\ a_q^T \end{bmatrix},$$

gdje smo s B označili matricu $R^T A$, a sa a_i^T, b_i^T retke matrica A i B. Ako pišemo $A' = BR$, p ti i q ti stupac matrica A' i B zadovoljavat će relaciju

$$\begin{bmatrix} a'_p & a'_q \end{bmatrix} = \begin{bmatrix} b_p & b_q \end{bmatrix} \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix}.$$

U algoritamskom opisu ove relacije se prevode u

Za $i = 1, \ldots, n$ računaj	Za $i = 1, \ldots, n$ računaj
$b_{pi} = a_{pi} \cos\phi + a_{qi} \sin\phi$	$a'_{ip} = b_{ip} \cos\phi + b_{iq} \sin\phi$
$b_{qi} = -a_{pi} \sin\phi + a_{qi} \cos\phi$	$a'_{iq} = -b_{ip} \sin\phi + b_{iq} \cos\phi$

Da bi se uštedjelo na memorijskom prostoru, cijeli postupak se provodi na gornjem trokutu matrice, uključujući dijagonalu. To je moguće jer su obje matrice A i A' simetrične. Uz to, zbog računski vrlo pogodnih formula za dijagonalne elemente, poželjno je iskoristiti ih. Promatrajući samo gornji trokut matrica, vidimo da se od izvandijagonalnih elemenata samo onaj na poziciji (p, q) dvaput mijenja. Kako on ionako postaje nula, što se numeričkog računanja tiče, nema nikakva razloga za uvođenje pomoćne matrice B. U slijedećem algoritmu bitno se koristi svojstvo simetričnosti matrice A'.

Algoritam 13.6 *Dana je simetrična matrica A reda n. Sljedeći algoritam računa elemente rotacije R i matrice $A' = R^T AR$ nakon Jacobijevog koraka koji poništava element a_{pq}. Matrice A i A' koriste isto dvodimenzionalno polje koje poistovjećujemo s matricom A. Algoritam koristi samo gornji trokut i dijagonalu od A.*

$$
\begin{aligned}
\textit{Ako je } a_{pq} &= 0 \textit{ postupak je gotov.} \\
\textit{Ako je } a_{pq} &\neq 0 \textit{ računa se} \\
\lambda &= 2a_{pq} \cdot sign(a_{pp} - a_{qq}) \\
\mu &= |\ a_{pp} - a_{qq}\ | \\
\nu &= \sqrt{\lambda^2 + \mu^2} \\
t &= \frac{\lambda}{\mu + \nu} \\
c &= \frac{1}{\sqrt{1 + t^2}} \\
s &= t * c \\
a_{pp} &= a_{pp} + t \cdot a_{pq} \\
a_{qq} &= a_{qq} - t \cdot a_{pq} \\
a_{pq} &= 0 \\
\textit{Za } i &= 1, \ldots p - 1 \\
x &= c * a_{ip} + s * a_{iq} \\
a_{iq} &= -s * a_{ip} + c * a_{iq} \\
a_{ip} &= x \\
\textit{Za } i &= p + 1, \ldots q - 1 \\
x &= c * a_{pi} + s * a_{iq} \\
a_{iq} &= -s * a_{pi} + c * a_{iq} \\
a_{pi} &= x \\
\textit{Za } i &= q + 1, \ldots n
\end{aligned}
$$

$$\begin{aligned} x &= c * a_{pi} + s * a_{qi} \\ a_{qi} &= -s * a_{pi} + c * a_{qi} \\ a_{pi} &= x \end{aligned}$$

Algoritmom 13.6 riješili smo jedan korak Jacobijeve metode. Općenito, ako u jednom koraku poništimo neki element, on se može već u slijedećem koraku "pokvariti". To ukazuje da nije svrha jednog koraka poništavanje pivotnog elementa, već smanjivanje odstupanja od dijagonalne forme.

Da bismo dijagonalizirali simetričnu matricu rada n trebamo u načelu beskonačno koraka. Na sreću, niz dobivenih matrica konvergira prema dijagonalnoj matrici, pa se nakon, obično $3n$ do $5n$ koraka, stigne do matrice koja je toliko dijagonalna da ju računalo tretira kao dijagonalnu. To obično znači da smo vlastite vrijednosti (o vektorima ćemo govoriti kasnije) izračunali na 7 (u tzv. jednostrukoj točnosti računanja) ili 15 (u dvostrukoj točnosti računanja) znamenaka.

Stoga se postavlja pitanje kako izabirati pivotne elemente, dakle redoslijed poništavanja elemenata, pa da dijagonalizacija bude što brža. To nije jednostavni problem jer često brža konvergencija niza dobivenih matrica prema dijagonalnoj matrici zahtijeva više pretraživanja veličine elemenata pri izboru pivotnog elementa, pa mjereći brojem računskih i logičkih operacija u računalu, traži više vremena.

13.1.3 Pivotna strategija

Način ili pravilo izbora pivotnih parova $(p, q) = (p(k), q(k))$ u ovisnosti o k se zove pivotna strategija. Pivotna strategija se može poistovjetiti s funkcijom $\mathrm{I} : \mathbf{N} \to \mathbf{P}_n$, gdje je $\mathbf{N}$ skup prirodnih brojeva (kojima k prolazi), a $\mathbf{P}_n = \{(i, j); 1 \leq i < j \leq n\}$.

Povijesno je najvažnija ona koju je i sam Jacobi koristio, a poznata je pod nazivom *klasična pivotna strategija*. Ona odabire najvećeg po modulu izvandijagonalnog elementa kao pivotnog. Dakle, par (p, q) se u k tom koraku bira tako da zadovoljava

$$\mid a_{pq}^{(k)} \mid = \max_{i<j} \mid a_{ij}^{(k)} \mid, \quad k \geq 1.$$

Ova strategija daje možda najbržu konvergenciju niza $S(A^{(k)})$, $k \geq$, ali na svakom koraku mora pretražiti $N = n(n-1)/2$ izvandijagonalnih elemenata. Kako to pretraživanje pretstavlja priličan napor za računalo, klasična strategija se uglavnom ne koristi u praksi.

Posebno su važne periodičke (pivotne) strategije za koje je I periodička funkcija. To znači $\mathrm{I}(k+M) = \mathrm{I}(k)$ za neki period M. Između njih najvažnije su cikličke strategije za koje je $M = N$ i $\{\mathrm{I}(k); 1 \leq k \leq N\} = \mathbf{P}_n$. Dakle, unutar svakih N koraka ponište se svi izvandijagonalni elementi, svaki u svoje vrijeme i samo jedanput. Niz od prvih N koraka $1, \ldots, N$ obično se zove ciklus. Nakon N-tog koraka cijeli ciklus se ponavlja i kad završi opet se ponavlja. Između cikličkih strategija najčešće korištene su tzv. serijalne: retčana i stupčana strategija. Kod retčane strategije pivotni par neprestalno prolazi nizom parova,

$$(1,2), (1,3), \ldots, (1,n), (2,3), \ldots, (2,n), \ldots, (n-1,n)$$

dok kod stupčane, prolazi nizom

$$(1,2), (1,3), (2,3), (1,4), (2,4), (3,4) \ldots, (1,n), (2,n), \ldots, (n-1,n).$$

Može se pokazati da polazeći od iste simetrične matrice A, korištenjem retčane i stupčane trategije, nakon svakog ciklusa dolazimo do jednakih matrica. Zato se te dvije strategije zovu ekvivalentnim strategijama.

13.1.4 Kriteriji zaustavljanja

U praksi svaki iterativni proces se nakon nekog vremena (koraka) zaustavi. Kod Jacobijeve metode proces se zaustavlja kad je iterirana matrica dovoljno dijagonalna tako da su dijagonalni elementi dovoljno dobre aproksimacije vlastitih vrijednosti matrice, a stupci produkta rotacija dovoljno dobre aproksimacije vlastitih vektora. Pretpostavimo da smo zaustavili proces nakon K koraka. Tada smo dobili matricu $A^{(K+1)}$ za koju vrijedi

$$A^{(K+1)} = R_K^T \cdots R_2^T R_1^T \, A \, R_1 R_2 \cdots R_K = {U^{(K)}}^T A \, U^{(K)}$$

pri čemu je $U^{(K)} = R_1 R_2 \cdots R_K$. Jasno, ako je $A^{(K+1)}$ blizu dijagonalne matrice Λ, onda možemo pisati

$${U^{(K)}}^T A \, U^{(K)} \approx \Lambda, \quad \text{ili} \quad A \, U^{(K)} \approx U^{(K)} \Lambda,$$

pa su dijagonalni elementi od $A^{(K+1)}$ aproksimacije vlastitih vrijednosti, a stupci od $U^{(K)}$ aproksimacije vlastitih vektora matrice A.

Dakle, za pravilno određivanje kriterija zaustavljanja potrebni su teoretski rezultati koji za skoro dijagonalne matrice mjere udaljenost između dijagonalnih elemenata i vlastitih vrijednosti matrice. Kako se ti teoretski rezultati poboljšavaju, tako se mijenjaju i kriteriji zaustavljanja. Danas se za nesingularne simetrične matrice najčešće koristi uvjet

$$\max_{1\leq i<j\leq n} \frac{| a_{ij}^{(k)} |}{\sqrt{| a_{ii}^{(k)} a_{jj}^{(k)} |}} \leq \varepsilon, \tag{13.11}$$

gdje je ε točnost računanja odnosno točnost aritmetike koja se koristi u računalu. Kod cikličkih metoda taj uvjet se obično provjerava na kraju svakog ciklusa. Za te strategije obično je potrebno 6 do 10 ciklusa prije nego bude ispunjen. Pokazuje se da je spomenuti kriterij zaustavljanja primjeran za pozitivno definitne matrice, jer osigurava visoku relativnu točnost izračunatih vlastitih vrijednosti i vektora.

Postoji i Rutishouserov kriterij zaustavljanja koji je modifikacija spomenutog, nešto je stroži od spomenutog, dobro funkcionira i na singularnim matricama i koristi tehniku poništavanja pivotnih elemenata bez transformacija na matrici, kad je to moguće.

13.1.5 Jacobijev algoritam

U ovoj završnoj točki ćemo ukratko izložiti retčani Jacobijev algoritam za simetrične matrice. Uočimo da se u $k$tom koraku metode vlastiti vektori ažuriraju po formuli $U^{(k+1)} = U^{(k)}R_k$, pri čemu je $U^{(1)} = I_n$.

Algoritam 13.7 *Dana je simetrična matrica A reda n. Sljedeći algoritam dijagonalizira A pomoću retčane cikličke Jacobijeve metode. Algoritam koristi samo gornji trokut i dijagonalu matrice A. Ako je potrebno algoritam računa i vlastite vektore.*

1. *Ako se traže vlastiti vektori, definiraj $U = I_n$*

2. *Stavi $k = 1$*

3. *Za dani k odredi pivotni par (p, q) prema odabranoj pivotnoj strategiji*

4. *Odredi parametre rotacije prema Algoritmu 13.6*

5. *Primijeni transformaciju na A prema Algoritmu 13.6*

6. *Ako se žele i vlastiti vektori, računaj*
   Za $i = 1, \dots n$
   $x = c * u_{ip} + s * u_{iq}$
   $u_{iq} = -s * u_{ip} + c * u_{iq}$
   $u_{ip} = x$

7. *Ako je* k *dijeljiv s* N *provjeri uvjet (13.11). Ako je ispunjen zaustavi proces. U protivnom stavi* $k = k + 1$ *i prijeđi na 3.*

8. *Dijagonalni element* a_{ii} *je vlastita vrijednost od A. Pripadni vlastiti vektor je* $[u_{1i}, u_{2i}, \dots, u_{ni}]^T$ *(ako se matrica U iterirala).*