SVEUČILIŠTE U ZAGREBU

PRIRODOSLOVNO–MATEMATIČKI FAKULTET

MATEMATIČKI ODSJEK

Anamarija Zibar

GENERALIZIRANI INVERZI

Diplomski rad

Voditelj rada:
prof. dr. sc. Sanja Singer

Zagreb, rujan, 2018.

Ovaj diplomski rad obranjen je dana ____________________ pred ispitnim povjerenstvom u sastavu:

1. ____________________________, predsjednik
2. ____________________________, član
3. ____________________________, član

Povjerenstvo je rad ocijenilo ocjenom ________________.

Potpisi članova povjerenstva:

1. ____________________________
2. ____________________________
3. ____________________________

Sadržaj

Uvod

Matrice i matrični inverzi se javljaju u mnogim primjenama. Da bi matrica imala inverz, mora zadovoljavati dosta stroga ograničenja – mora biti kvadratna i mora biti regularna. U slučaju da ne vrijedi jedan od tih uvjeta, može se definirati generalizirani inverz koji će na neki način biti sličan matričnom inverzu u klasičnom smislu. Time se dobije veća klasa matrica na kojoj je definiran pojam inverza. Ovisno o primjeni, uzet ćemo neka svojstva matričnog inverza koja su potrebna, a odbaciti ona koja nisu. U slučaju da je matrica regularna, generalizirani inverz će biti upravo njen inverz.

Promatrat ćemo dvije vrste inverza, Moore–Penroseov i Drazinov inverz. Oba generalizirana inverza uvodimo promatrajući svojstva matričnog inverza koja zadovoljavaju. Vidjet ćemo da ako želimo jača svojstva, smanjujemo skup matrica na kojima je takav generalizirani inverz definiran. Moore–Penroseov inverz je od koristi u rješavanju linearnog sustava jednadžbi, dok je Drazinov inverz koristan kod rješavanja sustava linearnih diferencijalnih jednadžbi a njegov poseban slučaj se pojavljuje u Markovljevim lancima.

Radi ilustracije postojanja oba generalizirana inverza promatramo i dvije matrične dekompozicije, dekompoziciju singularnih vrijednosti i Jordanovu formu, ali neće obje biti jednako korisne pri računanju.

Poglavlje 1

Moore–Penroseov inverz

1.1 Definicija i neka svojstva

Promatramo matricu $A \in \mathbb{C}^{m\times n}$ i pripadni sustav $A\boldsymbol{x} = \boldsymbol{b}$ za neki $\boldsymbol{b} \in \mathbb{C}^m$. Ako je matrica A kvadratna i regularna, jedinstveno rješenje prethodnog sustava dano s $\boldsymbol{x} = A^{-1}\boldsymbol{b}$. Što možemo reći o rješenju danog sustava ako matrica A nije regularna ili ako je $m \neq n$? Koja su to svojstva inverza koja su potrebna kako bismo bili u mogućnosti promatrati rješenje danog sustava? Što možemo reći o jedinstvenosti takvog rješenja? Ta pitanja su motivacija za naredno poglavlje.

Definicija 1.1.1. *Neka je $A \in \mathbb{C}^{m\times n}$ matrica. Matrica $X \in \mathbb{C}^{n\times m}$, koja zadovoljava Penroseove uvjete,*

$$AXA = A, \tag{1}$$
$$XAX = X, \tag{2}$$
$$(AX)^* = AX, \tag{3}$$
$$(XA)^* = XA, \tag{4}$$

zove se Moore–Penroseov inverz. Ako je matrica A regularna, matrica X je upravo jednaka njezinom inverzu.

Pokažimo prvo da je takva matrica, ako postoji, jedinstvena. Neka su X i Y matrice koje zadovoljavaju sva četiri uvjeta.

$$\begin{aligned} X &= XAX = XAYAX = (XA)^*(YA)^*X = A^*X^*A^*Y^*X \\ &= A^*Y^*X = YAYAX = YY^*A^*X^*A^* = YAY = Y. \end{aligned}$$

Dakle, ako postoji, jedinstveni Moore–Penroseov inverz matrice A označavat ćemo s $A^\dagger$.

Naredne tri tvrdnje o inverzima jednostavnih matrica lako se provjeravaju uvrštavanjem u jednadžbe (1)–(4). Generalizirani inverz skalara λ (1×1 matrice) je

$$\lambda^\dagger = \begin{cases} \lambda^{-1}, & \text{za } \lambda \neq 0, \\ 0, & \text{za } \lambda = 0. \end{cases}$$

Generalizirani inverz vektora $\boldsymbol{x}$ je

$$\boldsymbol{x}^\dagger = \begin{cases} \boldsymbol{x}^*/\|\boldsymbol{x}\|^2, & \text{za } \boldsymbol{x} \neq \boldsymbol{0}, \\ \boldsymbol{0}, & \text{za } \boldsymbol{x} = \boldsymbol{0}, \end{cases}$$

gdje $\boldsymbol{0}$ označava nul-vektor.

Neka je $A = \operatorname{diag}(a_1, a_2, \ldots, a_{\min(m,n)}) \in \mathbb{C}^{m\times n}$ dijagonalna matrica, onda je

$$A^\dagger = \operatorname{diag}\left(a_1^\dagger, a_2^\dagger, \ldots, a_{\min(m,n)}^\dagger\right) \in \mathbb{C}^{n\times m}.$$

Postojanje Moore–Penroseovog inverza može se pokazati na nekoliko načina. Pokažimo njegovo postojanje korištenjem singularne dekompozicije matrice A koja eksplicitno formira $A^\dagger$. Za računanje singularne dekompozicije postoje numerički stabilni algoritmi, pa je to jedan od načina na koji možemo izračunati Moore–Penroseov inverz.

Teorem 1.1.2 (Singularna dekompozicija matrice, [1, str. 205–206])**.** *Ako je $A \in \mathbb{C}^{m\times n}$, tada postoje unitarne matrice $U \in \mathbb{C}^{m\times m}$ i $V \in \mathbb{C}^{n\times n}$ takve da je*

$$U^*AV = \Sigma \qquad \Sigma = \operatorname{diag}\left(\sigma_1, \sigma_2, \ldots, \sigma_{\min(m,n)}\right),$$

pri čemu vrijedi

$$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_{\min(n,m)} \geq 0.$$

Skalari $\sigma_1, \sigma_2, \ldots, \sigma_{\min(m,n)}$ su singularne vrijednosti matrice A, stupci matrice U su lijevi, a stupci matrice V desni singularni vektori matrice A.

Sada lako dobijemo izraz za Moore–Penroseov inverz.

Korolar 1.1.3. *Neka je $A \in \mathbb{C}^{m\times n}$ i neka je njezina singularna dekompozicija $A = U\Sigma V^*$. Tada je*

$$A^\dagger = V\Sigma^\dagger U^*.$$

Dokaz. Provjerimo Penroseove uvjete (1)–(4).

$$\begin{aligned} AA^\dagger A &= U\Sigma V^* V\Sigma^\dagger U^* U\Sigma V^* = U\Sigma\Sigma^\dagger\Sigma V^* = A, \\ A^\dagger AA^\dagger &= V\Sigma^\dagger U^* U\Sigma V^* V\Sigma^\dagger U^* = V\Sigma^\dagger\Sigma\Sigma^\dagger U^* = A^\dagger, \\ (AA^\dagger)^* &= (U\Sigma V^* V\Sigma^\dagger U^*)^* = U(\Sigma\Sigma^\dagger)^* U^* = U\Sigma\Sigma^\dagger U^* = AA^\dagger, \\ (A^\dagger A)^* &= (V\Sigma^\dagger U^* U\Sigma V^*)^* = V(\Sigma^\dagger\Sigma)^* V^* = V\Sigma^\dagger\Sigma V^* = A^\dagger A. \end{aligned}$$

□

Postoji li još neka faktorizacija koja bi nam dala lako izračunljiv izraz za Moore–Penroseov inverz? Primijetimo da ne mora uvijek vrijediti

$$(AB)^\dagger = B^\dagger A^\dagger \tag{1.1}$$

kao kod običnog inverza. Na primjer, Moore–Penrosov inverz matrice A,

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

je

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

Ako uzmemo $B = A$ u tvrdnji (1.1), onda za zadanu matricu A vrijedi $AB = A^2 = A$. Generalizirani inverz lijeve strane je $(AB)^\dagger = (AA)^\dagger = A^\dagger$. Desna strana u (1.1) jednaka je $B^\dagger A^\dagger = (A^\dagger)^2 = \frac{1}{2}A^\dagger$, pa je odmah jasno da ne vrijedi jednakost lijeve i desne strane u (1.1).

Ovaj jednostavan primjer daje naslutiti da za Moore–Penroseov inverz neće vrijediti jednostavna pravila kao za obične inverze.

Teorem 1.1.4 ([1, str. 26])**.** *Ako je $A \in \mathbb{C}^{m\times n}$ ranga $r > 0$ tada postoje matrice $F \in \mathbb{C}^{m\times r}$ i $G \in \mathbb{C}^{r\times n}$ obje ranga r takve da je*

$$A = FG.$$

Takvu faktorizaciju zovemo faktorizacija punog ranga.

Teorem 1.1.5 ([1, str. 48])**.** *Ako je $A \in \mathbb{C}^{m\times n}$ ranga $r > 0$ i matrice $F \in \mathbb{C}^{m\times r}$ i $G \in \mathbb{C}^{r\times n}$ obje ranga r takve da je $A = FG$ tada je*

$$A^\dagger = G^*(F^*AG^*)^{-1}F^*.$$

Dokaz. Prvo pokazujemo da možemo izračunati izraz $(F^*AG^*)^{-1}$, odnosno da je (F^*AG^*) regularna matrica. U jednakosti $(F^*AG^*) = (F^*F)(GG^*)$ oba produkta s desne strane, F^*F i GG^* su dimenzija $r \times r$ i ranga r, pa je matrica s lijeve strane regularna, jer je produkt dvaju regularnih matrica. Prvo izračunamo X korištenjem matrica F i G,

$$X = G^*(F^*AG^*)^{-1}F^* = G^*(F^*FGG^*)^{-1}F^* = G^*(GG^*)^{-1}(F^*F)^{-1}F^*,$$

a onda provjerimo Penroseove uvjete (1)–(4)

$$\begin{aligned}
AXA &= FGG^*(GG^*)^{-1}(F^*F)^{-1}F^*FG = FG = A, \\
XAX &= G^*(GG^*)^{-1}(F^*F)^{-1}F^*FGG^*(GG^*)^{-1}(F^*F)^{-1}F^* \\
&= G^*(GG^*)^{-1}(F^*F)^{-1}F^* = X, \\
(AX)^* &= F(F^*F)^{-1}F^* = FGG^*(GG^*)^{-1}(F^*F)^{-1}F^* = AX, \\
(XA)^* &= G^*(GG^*)^{-1}G = G^*(GG^*)^{-1}(F^*F)^{-1}F^*FG = XA.
\end{aligned}$$

□

Sad kad znamo da Moore–Penroseov inverz svake matrice postoji i jedinstven je, zapišimo neka njegova zanimljiva svojstva.

Teorem 1.1.6. *Neka je $A \in \mathbb{C}^{m \times n}$. Vrijedi:*

1. $(A^\dagger)^\dagger = A$;
2. $(A^\dagger)^* = (A^*)^\dagger$;
3. *ako je* $\lambda \in \mathbb{C}$, *onda je* $(\lambda A)^\dagger = \lambda^\dagger A^\dagger$;
4. $A^* = A^*AA^\dagger = A^\dagger AA^*$;
5. $(A^*A)^\dagger = A^\dagger A^{*\dagger}$;
6. $(AA^*)^\dagger = A^{*^\dagger} A^\dagger$;
7. $A^\dagger = (A^*A)^\dagger A^* = A^*(AA^*)^\dagger$;
8. $(UAV)^\dagger = V^*A^\dagger U^*$, *gdje su U i V unitarne matrice.*

1.2 $\{i, j, k\}$–inverzi

Sad kad znamo da za svaku matricu $A \in \mathbb{C}^{m \times n}$ možemo pronaći odgovarajuću matricu $A^\dagger$ koja zadovoljava uvjete (1)–(4), pogledajmo pobliže što svaki od tih uvjeta znači. Spomenuto je da je motivacija za ovakvu definiciju generaliziranog inverza vezana uz rješavanje sustava. Što koji uvjet tu predstavlja? Postoji li mogućnost da bi za neke potrebe matrica koja zadovoljava samo neke uvjete bila dovoljna?

Definicija 1.2.1. *Za matricu $A \in \mathbb{C}^{m \times n}$, matrica $X \in \mathbb{C}^{n \times m}$ se zove $\{i, j, k\}$–inverz matrice A ako zadovoljava i-ti, j-ti i k-ti Penroseov uvjet. Takvu matrixu označavamo s $X = A^{(i,j,k)}$, pišemo $X \in A\{i, j, k\}$, gdje $A\{i, j, k\}$ označava skup svih $\{i, j, k\}$–inverza matrice A.*

$\{1\}$–inverz

Promatramo sustav

$$A\boldsymbol{x} = \boldsymbol{b}.$$

Definicija 1.2.2. *Za sustav $A\boldsymbol{x} = \boldsymbol{b}$ kažemo da je konzistentan ako je $\boldsymbol{b} \in \operatorname{Im} A$, inače je nekonzistentan.*

Koja svojstva zadovoljava matrica X za koju vrijedi da je $X\boldsymbol{b}$ rješenje gornjeg sustava za svaki $\boldsymbol{b} \in \mathbb{C}^m$ za koji je sustav konzistentan, odnosno vrijedi da je $\boldsymbol{b} \in \operatorname{Im} A$?

Ako je $AX\boldsymbol{b} = \boldsymbol{b}$ za svaki $\boldsymbol{b} \in \operatorname{Im} A$, onda vrijedi da je $AXA = A$, odnosno zadovoljen je prvi Penroseov uvjet, (1).

Vrijedi li i obratno? Neka matrica $X \in \mathbb{C}^{n\times m}$ zadovoljava $AXA = A$. Tada za svaki $\boldsymbol{b} \in \operatorname{Im} A$ postoji $\boldsymbol{x} \in \mathbb{C}^n$ takav da vrijedi $A\boldsymbol{x} = \boldsymbol{b}$, tj. vrijedi

$$\boldsymbol{b} = A\boldsymbol{x} = AXA\boldsymbol{x} = AX\boldsymbol{b}.$$

Time smo pokazali sljedeći teorem:

Teorem 1.2.3. *Za $A \in \mathbb{C}^{m\times n}$, matrica $X \in \mathbb{C}^{n\times m}$ ima svojstvo da je $X\boldsymbol{b}$ rješenje sustava $A\boldsymbol{x} = \boldsymbol{b}$ za svaki $\boldsymbol{b} \in \operatorname{Im} A$ ako i samo ako je $X \in A\{1\}$.*

Ovime je pokazano da je za rješavanje konzistentnog sustava dovoljna matrica koja zadovoljava samo svojstvo (1). Primijetimo da, ako vrijede uvjeti prethodnog teorema, rješenje sustava je

$$\boldsymbol{x} = A^{(1)}\boldsymbol{b} + (I - A^{(1)}A)\boldsymbol{y}$$

za proizvoljni $\boldsymbol{y} \in \mathbb{C}^n$, odnosno, rješenje nije jedinstveno! Možemo li dodavanjem nekog uvjeta dobiti "najbolje" rješenje u nekom smislu?

{1, 4}–inverz

Za rješenje $\boldsymbol{x}$ sustava $A\boldsymbol{x} = \boldsymbol{b}$ kažemo da je rješenje s minimalnom normom ako vrijedi

$$\|\boldsymbol{x}\| < \|\boldsymbol{y}\|, \quad \forall \boldsymbol{y} \in \{\boldsymbol{x} : A\boldsymbol{x} = \boldsymbol{b}\} \quad \text{i} \quad \boldsymbol{y} \neq \boldsymbol{x}.$$

Promotrimo vektor $\boldsymbol{x} = A^{(1,4)}\boldsymbol{b} + (I - A^{(1,4)}A)\boldsymbol{y}$. Primijetimo da je

$$(I - A^{(1,4)}A)\boldsymbol{y} \in \operatorname{Ker} A,$$

što odmah izlazi korištenjem (1). Također, znamo da je $\operatorname{Ker} A = (\operatorname{Im} A^*)^\perp$. Nadalje, za svaki $\boldsymbol{b} \in \operatorname{Im} A$ postoji $\boldsymbol{z} \in \mathbb{C}^n$ takav da je $A\boldsymbol{z} = \boldsymbol{b}$, pa vrijedi

$$A^{(1,4)}\boldsymbol{b} = A^{(1,4)}A\boldsymbol{z} = (A^{(1,4)}A)^*\boldsymbol{z} = A^*(A^{(1,4)})^*\boldsymbol{z} \in \operatorname{Im} A^*.$$

Dobili smo da je

$$(A^{(1,4)}\boldsymbol{b}, (I - A^{(1,4)}A)\boldsymbol{y}) = 0,$$

gdje $(\cdot, \cdot)$ označava skalarni produkt.

Korištenjem Pitagorinog teorema slijedi

$$\|A^{(1,4)}\boldsymbol{b}\|^2 \leq \|A^{(1,4)}\boldsymbol{b}\|^2 + \|(I - A^{(1,4)}A)\boldsymbol{y}\|^2 = \|A^{(1,4)}\boldsymbol{b} + (I - A^{(1,4)}A)\boldsymbol{y}\|^2.$$

Iz ove nejednakosti slijedi da je za $\boldsymbol{b} \in \operatorname{Im} A$ rješenje sustava $A\boldsymbol{x} = \boldsymbol{b}$ s minimalnom normom, dano s $A^{(1,4)}\boldsymbol{b}$.

Želimo provjeriti vrijedi li i obrat. Ako za $\boldsymbol{b} \in \operatorname{Im} A$ vrijedi da je $X\boldsymbol{b} = A^{(1,4)}b$ onda je

$$XA = A^{(1,4)}A,$$

što je upravo ekvivalentno tome da su ispunjeni Penroseovi uvjeti (1) i (4):

$$AXA = A \qquad \text{i} \qquad (XA)^* = XA.$$

Time je dokazan sljedeći teorem.

Teorem 1.2.4. *Za $A \in \mathbb{C}^{m\times n}$, matrica $X \in \mathbb{C}^{n\times m}$ ima svojstvo da je $X\boldsymbol{b}$ rješenje sustava $A\boldsymbol{x} = \boldsymbol{b}$ za svaki $\boldsymbol{b} \in \operatorname{Im} A$ i $\|X\boldsymbol{b}\| < \|z\|$ za svaki $z \neq X\boldsymbol{b}$ i $z \in \{\boldsymbol{x} : A\boldsymbol{x} = \boldsymbol{b}\}$ ako i samo ako je $X \in A\{1, 4\}$.*

Ovime smo odgovorili na pitanje kakvo rješenje možemo očekivati u slučaju da se desna strana sustava zbilja može zapisati kao linearna kombinacija stupaca matrice sustava i kako ga dobiti. Što ako sustav nije konzistentan?

$\{1, 3\}$–inverz

Neka je $A\boldsymbol{x} = \boldsymbol{b}$ za neki $\boldsymbol{b} \notin \operatorname{Im} A$. Tražimo rješenje problema najmanjih kvadrata, odnosno takav $\boldsymbol{x} \in \mathbb{C}^n$ da vrijedi

$$\|A\boldsymbol{x} - \boldsymbol{b}\| \leq \|A\boldsymbol{y} - \boldsymbol{b}\| \quad \forall \boldsymbol{y} \in \mathbb{C}^n.$$

Zapisujemo vektor $\boldsymbol{b} = \boldsymbol{b_1} + \boldsymbol{b_2}$, gdje je $\boldsymbol{b_1} \in \operatorname{Im} A$, $\boldsymbol{b_2} \in \operatorname{Ker} A^*$. Postoji $z \in \mathbb{C}^n$ takav da $Az = \boldsymbol{b_1}$, a korištenjem svojstava $\{ 1, 3\}$–inverza dobivamo da vrijedi i

$$AA^{(1,3)}\boldsymbol{b_1} = AA^{(1,3)}Az = Az = \boldsymbol{b_1},$$
$$AA^{(1,3)}\boldsymbol{b_2} = (A^{(1,3)})^*A^*\boldsymbol{b_2} = 0.$$

Iz te dvije jednakosti slijedi

$$AA^{(1,3)}\boldsymbol{b} = AA^{(1,3)}\boldsymbol{b_1} + AA^{(1,3)}\boldsymbol{b_2} = \boldsymbol{b_1}.$$

Budući da je $AX\boldsymbol{b} - \boldsymbol{b_2} \in \operatorname{Im} A$, prema Pitagorinom teoremu vrijedi sljedeća jednakost

$$\|AX\boldsymbol{b} - \boldsymbol{b}\|^2 = \|(AX\boldsymbol{b} - \boldsymbol{b_1}) + (-\boldsymbol{b_2})\|^2 = \|AX\boldsymbol{b} - \boldsymbol{b_1}\|^2 + \|\boldsymbol{b_2}\|^2,$$

koja se minimizira za $AX\boldsymbol{b} = \boldsymbol{b_1} = AA^{(1,3)}\boldsymbol{b}$ za sve $\boldsymbol{b} \in \mathbb{C}^m$. Odnosno, mora biti zadovoljen uvjet

$$AX = AA^{(1,3)},$$

koji je ekvivalentan Penroseovima uvjetima (1) i (3),

$$AXA = A \quad \text{i} \quad (AX)^* = AX.$$

Formalno to zapisujemo kao sljedeći teorem.

Teorem 1.2.5. *Za $A \in \mathbb{C}^{m\times n}$, matrica $X \in \mathbb{C}^{n\times m}$ ima svojstvo da je $X\boldsymbol{b}$ rješenje problema najmanjih kvadrata za $A\boldsymbol{x} = \boldsymbol{b}$ za svaki $\boldsymbol{b} \notin \operatorname{Im} A$ ako i samo ako je $X \in A\{1, 3\}$.*

Sljedeće pitanje na koje ćemo odgovoriti je možemo li okarakterizirati rješenje problema najmanjih kvadrata koje ima minimalnu normu?

U slučaju da je $\boldsymbol{b} \in \operatorname{Im} A$, rješenje problema najmanjih kvadrata se podudara s rješenjem sustava $A\boldsymbol{x} = \boldsymbol{b}$, tako da možemo promatrati problem najmanjih kvadrata za proizvoljni $\boldsymbol{b} \in \mathbb{C}^m$.

Teorem 1.2.6. *Neka je $A \in \mathbb{C}^{m\times n}$ i $\boldsymbol{b} \in \mathbb{C}^m$. $A^\dagger\boldsymbol{b}$ je rješenje problema najmanjih kvadrata s minimalnom normom za $A\boldsymbol{x} = \boldsymbol{b}$.*

Dokaz. Zapisujemo $\boldsymbol{b} = \boldsymbol{b_1} + \boldsymbol{b_2}$, gdje je

$$\boldsymbol{b_1} = AA^\dagger\boldsymbol{b} \in \operatorname{Im} A \quad \text{i} \quad \boldsymbol{b_2} = (I - AA^\dagger)\boldsymbol{b} \in (\operatorname{Im} A)^\perp.$$

Nadalje, vrijedi

$$\|Ax - \boldsymbol{b}\|^2 = \|Ax - \boldsymbol{b_1} - \boldsymbol{b_2}\|^2 = \|Ax - \boldsymbol{b_1}\|^2 + \|\boldsymbol{b_2}\|^2.$$

Prethodni izraz se minimizira za $\boldsymbol{x}$ koji je rješenje konzistentnog sustava $A\boldsymbol{x} = AA^\dagger\boldsymbol{b}$.

Očito, s

$$\boldsymbol{x} = A^\dagger\boldsymbol{b} + (I - A^\dagger A)\boldsymbol{y}, \qquad \forall \boldsymbol{y} \in \mathbb{C}^n$$

su dana rješenja tog sustava. Za $\|A^\dagger\boldsymbol{x}\|$ vrijedi

$$\|A^\dagger\boldsymbol{x}\|^2 < \|A^\dagger\boldsymbol{x}\|^2 + \|(I - A^\dagger A)\boldsymbol{y}\|^2 = \|A^\dagger\boldsymbol{x} + (I - A^\dagger A)\boldsymbol{y}\|^2, \qquad (I - A^\dagger A)\boldsymbol{y} \neq 0,$$

pa je s $\boldsymbol{x} = A^\dagger\boldsymbol{b}$ rješenje problema najmanjih kvadrata s minimalnom normom. □

Svojstva i upotreba {1}–inverza

Svi $\{i, j, k\}$–inverzi dosad spomenuti su bili iz $A\{1\}$ pa je fokus ovog dijela stavljen na proučavanje upravo te grupe inverza. Prvo su navedena neka svojstva $\{1\}$–inverza.

Teorem 1.2.7. *Za $A \in \mathbb{C}^{m\times n}$ vrijedi:*

1. $(A^{(1)})^* \in A^*\{1\}$;
2. $\lambda^\dagger A^{(1)} \in (\lambda A)\{1\}$, $\lambda \in \mathbb{C}$;
3. $\operatorname{rang}(A^{(1)}) \geq \operatorname{rang}(A)$;
4. *za nesingularne P i Q vrijedi $Q^{-1}A^{(1)}P^{-1} \in (PAQ)\{1\}$;*
5. *ako je P punog stupčanog ranga, a Q punog retčanog ranga, onda je $Q^{(1)}A^{(1)}P^{(1)} \in (PAQ)\{1\}$;*

6. *ako je A punog stupčanog ranga, onda je* $A^{(1)}A = I_m$, *a ako je punog retčanog ranga, onda je* $AA^{(1)} = I_n$;
7. $AA^{(1)}$ *i* $A^{(1)}A$ *su idempotentne matrice i* $\mathrm{rang}(AA^{(1)}) = \mathrm{rang}(A) = \mathrm{rang}(A^{(1)}A)$;
8. *ako je A regularna matrica, onda je* $A^{(1)} = A^{-1}$;
9. *ako* $A^* = A$ *onda postoji matrica* $X \in A\{1\}$ *takva da je* $X^* = X$;
10. $\mathrm{Im}(AA^{(1)}) = \mathrm{Im}\, A$, $\mathrm{Ker}(A^{(1)}A) = \mathrm{Ker}\, A$, $\mathrm{Im}((A^{(1)}A)^*) = \mathrm{Im}(A^*)$.

Nakon rješenja sustava, logični korak je promatrati rješenje matrične jednadžbe $AXB = D$.

Teorem 1.2.8. *Neka su zadane matrice* $A \in \mathbb{C}^{m\times n}$, $B \in \mathbb{C}^{p\times q}$, $D \in \mathbb{C}^{m\times q}$. *Matrična jednadžba*

$$AXB = D \tag{1.2}$$

je konzistentna ako i samo ako za neke $A^{(1)}$ *i* $B^{(1)}$ *vrijedi*

$$AA^{(1)}DB^{(1)}B = D \tag{1.3}$$

i u tom je slučaju rješenje dano s

$$X = A^{(1)}DB^{(1)} + Y - A^{(1)}AYBB^{(1)}, \tag{1.4}$$

za proizvoljan $Y \in \mathbb{C}^{n\times p}$.

Dokaz. Prvo pokažimo ekvivalenciju. Neka vrijedi da je sustav u (1.2) konzistentan, dakle postoji matrica X koja je njegovo rješenje,

$$D = AXB = AA^{(1)}AXBB^{(1)}B = AA^{(1)}DB^{(1)}B.$$

Time smo pokazali da vrijedi (1.3). Neka sad vrijedi (1.3). Definiramo X kao u relaciji (1.4). Tada vrijedi

$$AA^{(1)}DB^{(1)}B + AYB - AA^{(1)}AYBB^{(1)}B = D + AYB - AYB = D.$$

Neka je sada X neko rješenje (1.2). Vrijedi

$$X = A^{(1)}DB^{(1)} + X - A^{(1)}AXBB^{(1)},$$

što je, očito, oblika (1.4). □

1.3 Računanje Moore–Penroseovog inverza

Generalizirani inverz sume i particionirane matrice

Općenito, kao ni kod inverza regularnih matrica, ne postoji izraz za $(A + B)^\dagger$ za proizvoljne matrice $A, B \in \mathbb{C}^{m\times n}$.

Često je slučaj da nakon što smo izračunali $A^\dagger$ neke matrice A, želimo modificirati model koji je dan tom matricom tako da promijenimo jedan ili više njezinih elemenata te je korisno znati $\tilde{A}^\dagger$ tako promijenjene matrice $\tilde{A}$. Tada je moguće usporediti modificirani model s početnim, kako bismo vidjeli utjecaj modifikacija.

Također, ukoliko se javila greška na mjestu (i, j) pri unosu matrice A ili zbog nepreciznog mjerenja, htjeli bismo pri ponovnom računanju iskoristiti većinu napora koji smo uložili u računanje $A^\dagger$.

Općenito vrijedi da se matrica B ranga 1 može zapisati kao produkt dvaju vektora: $B = \boldsymbol{c}\boldsymbol{d}^*$, $\boldsymbol{c} \in \mathbb{C}^m$, $\boldsymbol{d} \in \mathbb{C}^n$. Postoji izraz koji povezuje $(A + \boldsymbol{c}\boldsymbol{d}^*)^\dagger$ s $A^\dagger$. Prije no što ga zapišemo, navodimo sve moguće slučajeve koji se mogu pojaviti, kako bi konačni izraz imao što manje članova.

1. $\boldsymbol{c} \notin \operatorname{Im} A$, $\boldsymbol{d} \notin \operatorname{Im}(A^*)$ i $1 + \boldsymbol{d}^* A^\dagger \boldsymbol{c}$ proizvoljan;
2. $\boldsymbol{c} \in \operatorname{Im} A$, $\boldsymbol{d} \notin \operatorname{Im}(A^*)$ i $1 + \boldsymbol{d}^* A^\dagger \boldsymbol{c} = 0$;
3. $\boldsymbol{c} \in \operatorname{Im} A$, $\boldsymbol{d}$ proizvoljan i $1 + \boldsymbol{d}^* A^\dagger \boldsymbol{c} \neq 0$;
4. $\boldsymbol{c} \notin \operatorname{Im} A$, $\boldsymbol{d} \in \operatorname{Im}(A^*)$ i $1 + \boldsymbol{d}^* A^\dagger \boldsymbol{c} = 0$;
5. $\boldsymbol{c}$ proizvoljan, $\boldsymbol{d} \in \operatorname{Im}(A^*)$ i $1 + \boldsymbol{d}^* A^\dagger \boldsymbol{c} \neq 0$;
6. $\boldsymbol{c} \in \operatorname{Im} A$, $\boldsymbol{d} \in \operatorname{Im}(A^*)$ i $1 + \boldsymbol{d}^* A^\dagger \boldsymbol{c} = 0$.

Navedenu podjelu koristimo u sljedećem teoremu, čiji dokaz se može pronaći u [2, str. 47–50].

Teorem 1.3.1. *Zadani su matrica $A \in \mathbb{C}^{m\times n}$, i vektori $\boldsymbol{c} \in \mathbb{C}^m$ i $\boldsymbol{d} \in \mathbb{C}^n$. Definiramo oznake za vektore stupce*

$$\boldsymbol{k} = A^\dagger \boldsymbol{c}, \qquad \boldsymbol{u} = (I - AA^\dagger)\boldsymbol{c},$$

vektore retke

$$\boldsymbol{h} = \boldsymbol{d}^* A^\dagger, \qquad \boldsymbol{v} = \boldsymbol{d}^*(I - A^\dagger A),$$

i skalar

$$\beta = 1 + \boldsymbol{d}^* A^\dagger \boldsymbol{c}.$$

Tvrdimo da vrijedi $\boldsymbol{c} \in \operatorname{Im} A$ ako i samo ako je $\boldsymbol{u} = \boldsymbol{0}$ i da vrijedi $\boldsymbol{d} \in \operatorname{Im}(A^*)$ ako i samo ako je $\boldsymbol{v} = \boldsymbol{0}$.

Generalizirani inverz matrice $(A + \boldsymbol{c}\boldsymbol{d}^*)$ dan je u ovisnosti o vektoru stupcu $\boldsymbol{u}$, vektoru retku $\boldsymbol{v}$ i skalaru β:

1. *ako je* $\boldsymbol{u} \neq \boldsymbol{0}$ *i* $\boldsymbol{v} \neq \boldsymbol{0}$, *onda je*
$$(A + \boldsymbol{c}\boldsymbol{d}^*)^\dagger = A^\dagger - \boldsymbol{k}\boldsymbol{u}^\dagger - \boldsymbol{v}^\dagger\boldsymbol{h} + \beta\boldsymbol{v}^\dagger\boldsymbol{u}^\dagger;$$

2. *ako je* $\boldsymbol{u} = \boldsymbol{0}$, $\boldsymbol{v} \neq \boldsymbol{0}$ *i* $\beta = 0$, *onda je*
$$(A + \boldsymbol{c}\boldsymbol{d}^*)^\dagger = A^\dagger - \boldsymbol{k}\boldsymbol{k}^\dagger A^\dagger - \boldsymbol{v}^\dagger\boldsymbol{h};$$

3. *ako* $\boldsymbol{u} = \boldsymbol{0}$ *i* $\beta \neq 0$, *onda je*
$$(A + \boldsymbol{c}\boldsymbol{d}^*)^\dagger = A^\dagger + \frac{1}{\beta}\boldsymbol{v}^*\boldsymbol{k}^*A^\dagger - \frac{\overline{\beta}}{\sigma_1}\boldsymbol{p_1}\boldsymbol{q_1^*},$$
gdje je $\boldsymbol{p_1} = -\left(\frac{\|k\|^2}{\overline{\beta}}\boldsymbol{v}^* + \boldsymbol{k}\right)$, $\boldsymbol{q_1^*} = -\left(\frac{\|v\|^2}{\overline{\beta}}\boldsymbol{k}^*A^\dagger + \boldsymbol{h}\right)$ *i* $\sigma_1 = \|k\|^2\|v\|^2 + |\beta|^2$;

4. *ako je* $\boldsymbol{u} \neq \boldsymbol{0}$, $\boldsymbol{v} = \boldsymbol{0}$ *i* $\beta = 0$, *onda je*
$$(A + \boldsymbol{c}\boldsymbol{d}^*)^\dagger = A^\dagger - A^\dagger\boldsymbol{h}^\dagger\boldsymbol{h} - \boldsymbol{k}\boldsymbol{u}^\dagger;$$

5. *ako je* $\boldsymbol{v} = \boldsymbol{0}$ *i* $\beta \neq 0$, *onda je*
$$(A + \boldsymbol{c}\boldsymbol{d}^*)^\dagger = A^\dagger + \frac{1}{\beta}A^\dagger\boldsymbol{h}^*\boldsymbol{u}^* - \frac{\overline{\beta}}{\sigma_2}\boldsymbol{p_2}\boldsymbol{q_2^*},$$
gdje je $\boldsymbol{p_1} = -\left(\frac{\|u\|^2}{\overline{\beta}}A^\dagger\boldsymbol{h}^* + \boldsymbol{k}\right)$, $\boldsymbol{q_2^*} = -\left(\frac{\|h\|^2}{\overline{\beta}}\boldsymbol{u}^* + \boldsymbol{h}\right)$ *i* $\sigma_2 = \|h\|^2\|u\|^2 + |\beta|^2$;

6. *ako* $\boldsymbol{u} = \boldsymbol{0}$, $\boldsymbol{v} = \boldsymbol{0}$ *i* $\beta = 0$, *onda je*
$$(A + \boldsymbol{c}\boldsymbol{d}^*)^\dagger = A^\dagger - \boldsymbol{k}\boldsymbol{k}^\dagger A^\dagger - A^\dagger\boldsymbol{h}^\dagger\boldsymbol{h} + (\boldsymbol{k}^\dagger A^\dagger\boldsymbol{h}^\dagger)\boldsymbol{k}\boldsymbol{h}.$$

Promotrimo sada slučaj matrice E koju možemo podijeliti na četiri bloka,
$$E = \begin{bmatrix} A & B \\ C & D \end{bmatrix}.$$
Ako znamo izračunati $A^\dagger, B^\dagger, C^\dagger, D^\dagger$, znamo li reći što o $E^\dagger$? Kada bi postojao izraz za $E^\dagger$, mogli bismo rekurzivno podijeliti matricu na male blokove kojima se lako računa generalizirani inverz.

Počnimo s matricom $P \in \mathbb{C}^{m\times(n+1)}$ koju particioniramo tako da joj odrežemo zadnji stupac i označimo ga s $\boldsymbol{c}$. Ostatak matrice označimo s B.

Teorem 1.3.2. *Neka je $B \in \mathbb{C}^{m\times n}$, $\boldsymbol{c} \in \mathbb{C}^m$,*

$$P = \begin{bmatrix} B & \boldsymbol{c} \end{bmatrix}$$

i neka su $\boldsymbol{k} = B^\dagger \boldsymbol{c}$, $\boldsymbol{u} = (I - BB^\dagger)\boldsymbol{c} = \boldsymbol{c} - B\boldsymbol{k}$. Tada vrijedi

$$P^\dagger = \begin{bmatrix} B^\dagger - \boldsymbol{k}\boldsymbol{y} \\ \boldsymbol{y} \end{bmatrix},$$

gdje je

$$\boldsymbol{y} = \begin{cases} \boldsymbol{u}^\dagger, & \text{ako je } \boldsymbol{u} \neq \boldsymbol{0}, \\ (1 + \boldsymbol{k}^*\boldsymbol{k})^{-1}\boldsymbol{k}^* B^\dagger, & \text{ako je } \boldsymbol{u} = \boldsymbol{0}. \end{cases}$$

Dokaz. Matricu P možemo zapisati kao

$$P = \begin{bmatrix} B & \boldsymbol{0_m} \end{bmatrix} + \boldsymbol{c} \begin{bmatrix} \boldsymbol{0_n}^* & 1 \end{bmatrix}.$$

Označimo s $\boldsymbol{d}^* = \begin{bmatrix} \boldsymbol{0_n}^* & 1 \end{bmatrix}$ i $A = \begin{bmatrix} B & \boldsymbol{0_m} \end{bmatrix}$. Uz te oznake, P ima oblik

$$P = A + \boldsymbol{c}\boldsymbol{d}^*,$$

pa na nju možemo primijeniti teorem 1.3.1. Dobivamo

$$A^\dagger = \begin{bmatrix} B^\dagger \\ \boldsymbol{0_m}^* \end{bmatrix} \implies \begin{matrix} \boldsymbol{h} = \boldsymbol{d}^* A^\dagger = \boldsymbol{0}, & \beta = 1 + \boldsymbol{d}^* A^\dagger \boldsymbol{c} = 1 \\ \boldsymbol{v} = \boldsymbol{d}^* \neq 0, & \boldsymbol{u} = (I - AA^\dagger)\boldsymbol{c} = (I - BB^\dagger)\boldsymbol{c}. \end{matrix}$$

U ovisnosti o $\boldsymbol{u}$, imamo dva slučaja.

Ako je $\boldsymbol{u} \neq \boldsymbol{0}$ primijenjujemo slučaj 1 i dobijemo izraz za $P^\dagger$

$$P^\dagger = \begin{bmatrix} B^\dagger \\ \boldsymbol{0_m}^* \end{bmatrix} - \begin{bmatrix} B^\dagger \boldsymbol{c}\boldsymbol{u}^\dagger \\ \boldsymbol{0} \end{bmatrix} + \begin{bmatrix} B^\dagger \\ \boldsymbol{0} \end{bmatrix} + \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{u}^\dagger \end{bmatrix} = \begin{bmatrix} B^\dagger - B^\dagger \boldsymbol{c}\boldsymbol{u}^\dagger \\ \boldsymbol{u}^\dagger \end{bmatrix}.$$

Ako $\boldsymbol{u} = 0$ koristimo slučaj 3 prethodnog teorema. Dobivamo

$$\boldsymbol{k} = B^\dagger \boldsymbol{c}, \quad \sigma_1 = 1 + \boldsymbol{c}B^{\dagger *} B^\dagger \boldsymbol{c} = 1 + \boldsymbol{k}^*\boldsymbol{k}, \quad \boldsymbol{p_1} = -\begin{bmatrix} \boldsymbol{k} \\ \boldsymbol{k}^*\boldsymbol{k} \end{bmatrix} \quad \text{i} \quad \boldsymbol{q_1} = -\boldsymbol{k}^* B^\dagger.$$

Stoga vrijedi:

$$P^\dagger = \begin{bmatrix} B^\dagger - \frac{\boldsymbol{k}\boldsymbol{k}^* B^\dagger}{1+\boldsymbol{k}^*\boldsymbol{k}} \\ \frac{\boldsymbol{k}^* B^\dagger}{1+\boldsymbol{k}^*\boldsymbol{k}} \end{bmatrix}.$$

□

Ovaj teorem se može koristiti i za matrice kojima su dodani retci, tako da sve radimo na transponiranoj matrici. Na početku istraživanja generaliziranih inverza, ti su teoremi korišteni kao rekurzivni algoritmi za traženje Moore–Penroseovog inverza matrica. Navedeni algoritmi su jako osjetljivi na loše uvjetovane matrice i korisni su samo kod dobro uvjetovanih matrica malih dimenzija. Njihova prava vrijednost leži u tome da bez mnogo dodatnog računa možemo dodati rezultat jednog ili više mjerenja u problemu najmanjih kvadrata ili novi parametar u promatranom modelu.

U sljedećem teoremu navodimo izraz za generalizirani inverz matrice kojoj je dodano nekoliko stupaca.

Teorem 1.3.3. *Neka je $A \in \mathbb{C}^{m\times n}$ i $C \in \mathbb{C}^{m\times p}$, tada vrijedi:*

$$\begin{bmatrix} A & C \end{bmatrix}^\dagger = \begin{bmatrix} (I - TT^*)^{-1}(A^\dagger - A^\dagger C B^\dagger) \\ T^*(I + TT^*)^{-1}(A^\dagger - A^\dagger C B^\dagger) + B^\dagger \end{bmatrix},$$

pri čemu je $B = (I - AA^\dagger)C$ i $T = A^\dagger C(I - B^\dagger B)$.

Za općenitu matricu particioniranu u blokove ne postoji korisna reprezentacija

$$\begin{bmatrix} A & C \\ R & D \end{bmatrix}^\dagger .$$

Samo u nekim posebnim slučajevima, kao što su blok-trokutaste matrice može se nešto reći, ali samo onda kad njihovi blokovi zadovoljavaju neke dodatne uvjete. Tim problemom se u nastavku rada nećemo baviti.

Neprekidnost Moore–Penroseovog inverza

U primjenama često nemamo točnu matricu A već neku njezinu aproksimaciju $\tilde{A}$. Ta pogreška se javlja zbog reprezentacije realnih brojeva u računalnoj memoriji, ali i zbog nepreciznih mjernih uređaja koji mjere vrijednosti koeficijenata matrice. Stoga je važno promotriti ponašanje $\tilde{A}^\dagger$ u odnosu na $A^\dagger$.

Za regularne matrice vrijedi ako je

$$\|E\| < \frac{1}{\|A^{-1}\|} \implies A + E \text{ je regularna matrica.}$$

Pa ako imamo niz matrica E_k takvih da za svaku vrijedi gornje svojstvo, i da niz matrica konvergira k nul matrici, onda zbog neprekidnosti funkcije matričnog inverza vrijedi

$$(A + E_k)^{-1} \to A^{-1}.$$

Vrijedi li slična tvrdnja i za Moore–Penroseov inverz? Promotrimo sljedeći primjer

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = A^\dagger \qquad E_k = \begin{bmatrix} 0 & 0 \\ 0 & 1/k \end{bmatrix} \qquad F_k = \begin{bmatrix} 0 & 1/k \\ 0 & 0 \end{bmatrix}.$$

I matrice $(E_k)_{k\in\mathbb{N}}$ i matrice $(F_k)_{k\in\mathbb{N}}$ zadovoljavaju gornje uvjete, ali za generalizirane inverze vrijedi

$$(A + E_k)^\dagger = \begin{bmatrix} 1 & 0 \\ 0 & k \end{bmatrix} \nrightarrow A^\dagger, \qquad (A + F_k)^\dagger = \frac{1}{1+k^2} \begin{bmatrix} k^2 & 0 \\ k & 0 \end{bmatrix} \to A^\dagger.$$

Ovaj jednostavan primjer pokazuje da će trebati jači uvjeti od ovih koji su potrebni za reglarne matrice kako bi bila osigurana neprekidnost generaliziranog inverza.

Primijetimo da rang matrica A i $A + E_k$ nije jednak, dok je rang matrice $A + F_k$ jednak rangu početne matrice A. Pokazuje se da je to upravo nužan i dovoljan uvjet za konvergenciju generaliziranog inverza malo perturbirane matrice prema generaliziranom inverzu neperturbirane matrice.

Propozicija 1.3.4. *Neka je $A \in \mathbb{C}^{m\times n}$ i neka je $\|\cdot\|$ je matrična norma. Ako je $E \in \mathbb{C}^{m\times n}$ i vrijedi $\|E\| < 1/\|A^\dagger\|$ onda je*

$$\operatorname{rang}(A + E) \geq \operatorname{rang}(A).$$

Korolar 1.3.5. *Ako za matrice $A, B \in \mathbb{C}^{m\times n}$ vrijedi*

$$\|A - B\| < \frac{1}{\max\{\|A^\dagger\|, \|B^\dagger\|\}},$$

onda je $\operatorname{rang}(A) = \operatorname{rang}(B)$.

Dokaz. Neka je $E = A - B$. Tada je $\|E\| < 1/\|A^\dagger\|$ i $\|E\| < 1/\|B^\dagger\|$ pa vrijedi

$$\begin{aligned} \operatorname{rang}(A) &= \operatorname{rang}(E + B) \geq \operatorname{rang}(B), \\ \operatorname{rang}(B) &= \operatorname{rang}(A - E) \geq \operatorname{rang}(A). \end{aligned}$$

□

Teorem 1.3.6. *Neka je $A \in \mathbb{C}^{m\times n}$ zadana matrica i neka je $(E_j) \subset \mathbb{C}^{m\times n}$ niz matrica takav da $E_j \to \mathbf{0}$. Tada vrijedi*

$$(A + E_j) \to A^\dagger$$

ako i samo ako je $\operatorname{rang}(A + E_j) = \operatorname{rang}(A)$ *za sve j veće od nekog fiksnog j_0.*

Dokaz. Vrijedi da je $P_{\mathrm{Im}(A)} = AA^\dagger$, gdje je s $P_{\mathrm{Im}(A)}$ označen projektor na $\mathrm{Im}(A)$. Pretpostavimo da je $E_j \to \mathbf{0}$ i $(A + E_j)^\dagger \to A^\dagger$. Tada je

$$P_{\mathrm{Im}(A+E_j)} = (A + E_j)(A + E_j)^\dagger \to AA^\dagger = P_{\mathrm{Im}(A)}.$$

Uzimamo limes s obzirom na spektralnu normu $\|\cdot\|_2$. Stoga postoji j_0 takav da za svaki $j \geq j_0$ vrijedi

$$\|P_{\mathrm{Im}(A+E_j)} - P_{\mathrm{Im}(A)}\|_2 < 1.$$

Za ortogonalne projektore $P_{\mathrm{Im}(A+E_j)}$ i $P_{\mathrm{Im}(A)}$ ispunjen je uvjet iz korolara 1.3.5, pa je $\mathrm{rang}(A) = \mathrm{rang}(A + E_j)$ za $j \geq j_0$.

Dokaz druge implikacije se može naći u [2, str. 219]. □

Promotrimo sada izraz za normu generaliziranog inverza sume ako uvjet prethodnog teorema ne vrijedi.

Teorem 1.3.7. *Neka su $A, E \in \mathbb{C}^{m\times n}$ takve da vrijedi* $\mathrm{rang}(A + E) > \mathrm{rang}(A)$. *Tada vrijedi*

$$\|(A + E)^{\dagger}\| \geq \frac{1}{\|E\|}$$

za sve operatorske norme.

Dakle, ako vrijedi $\mathrm{rang}(A + E_j) > \mathrm{rang}(A)$ niz matrica $(A + E_j)^{\dagger}$ nije odozgo ograničen u normi pa ne može ni konvergirati.

U [7, str. 273], je navedena sljedeća ocjena koja nam govori koliko su blizu generalizirani inverz početne matrice A i perturbirane matrice $A + E$.

Teorem 1.3.8. *Ako za $A, E \in \mathbb{C}^{m\times n}$ vrijedi* $\mathrm{rang}(A + E) = \mathrm{rang}(A)$ *i* $\Delta = \|A^{\dagger}\|_2\|E\|_2 < 1$, *onda vrijedi*

$$\frac{\|(A + E)^{\dagger} - A^{\dagger}\|_2}{\|A^{\dagger}\|_2} \leq C\frac{\Delta}{1 - \Delta},$$

gdje je

$$C = \begin{cases} (1 + \sqrt{5})/2, & \textit{ako } \mathrm{rang}(A) < \min\{m, n\}, \\ \sqrt{2}, & \textit{ako } \mathrm{rang}(A) = \min\{m, n\}, \\ 1, & \textit{ako } \mathrm{rang}(A) = m = n. \end{cases}$$

Računanje generaliziranog inverza

Za dobro uvjetovane probleme probleme male do srednje veličine možemo koristiti algoritam baziran na Teoremu 1.3.2. Koraci algoritma su:

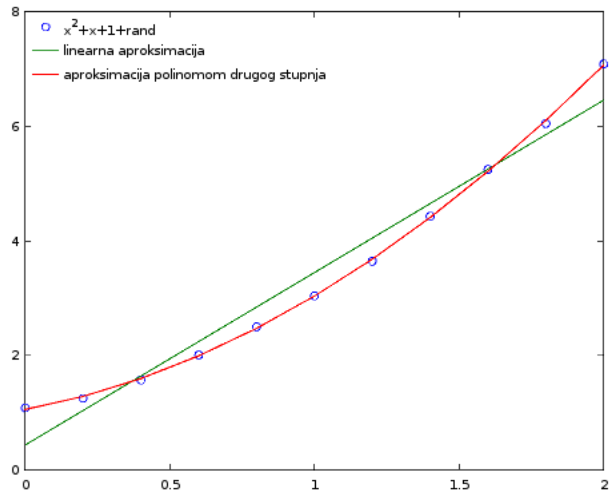
1. particionirati $A = \begin{bmatrix} \boldsymbol{c_1} & \dots & \boldsymbol{c_n} \end{bmatrix}$ i staviti $B_1 = \boldsymbol{c_1}$ i $B_1^{\dagger} = \boldsymbol{c_1}^{\dagger}$;
2. za $i \geq 2$, definirati $B_i = \begin{bmatrix} B_{i-1} & \boldsymbol{c_i} \end{bmatrix}$;
3. $\boldsymbol{k_i} = B_{i-1}^{\dagger}\boldsymbol{c_i}$;

4. $\boldsymbol{u_i} = \boldsymbol{c_i} - B_{i-1}\boldsymbol{k_i}\boldsymbol{y_i} = \begin{cases} \frac{\boldsymbol{u_i}^*}{\boldsymbol{u_i}^*\boldsymbol{u_i}}, & \text{ako } \boldsymbol{u_i} \neq \boldsymbol{0}, \\ \frac{\boldsymbol{k_i}^* B_{i-1}^\dagger}{1+\boldsymbol{k_i}^*\boldsymbol{k_i}}, & \text{ako } \boldsymbol{u_i} = \boldsymbol{0}; \end{cases}$
5. $B_i^\dagger = \begin{bmatrix} B_{i-1}^\dagger - \boldsymbol{k_i}\boldsymbol{y_i} \\ \boldsymbol{y_i} \end{bmatrix}$;
6. $A^\dagger = B_n^\dagger$.

Primjer 1.3.9. *Generiramo točke $x_i = i + 0.1$ za $i \in \{0, 1, \ldots, 10\}$. Definiramo uređeni par (x_i, y_i) tako da je $y_i = x_i^2 + x_i + 1$ polinom drugog stupnja. Na svaki y_i dodat ćemo "šum" – normalno distribuirane nasumične brojeve iz intervala $(0, 1)$. Primijenit ćemo generalizirani inverz na aproksimaciju podataka polinomom. Rješavamo sustav $A\boldsymbol{k} = \boldsymbol{y}$, gdje nepoznanica $\boldsymbol{k}$ označava koeficijente polinoma. Prvo aproksimiramo linearnom funkcijom, pa definiramo $A = \begin{bmatrix}\boldsymbol{1} & \boldsymbol{x}\end{bmatrix}$, gdje $\boldsymbol{1}$ označava vektor jedinica. Računamo njezin generalizirani inverz i rješenje sustava u obliku $\boldsymbol{k_1} = A^\dagger \boldsymbol{y}$. Sad je linearna aproksimacija točaka dana s*

$$lin_i = \boldsymbol{k_1}(1) + \boldsymbol{k_1}(2) * x_i.$$

Ako želimo aproksimirati polinomom drugog stupnja, dodajemo novi stupac u matricu pa je $A = \begin{bmatrix}\boldsymbol{1} & \boldsymbol{x} & \boldsymbol{x}^2\end{bmatrix}$, gdje $\boldsymbol{x}^2$ označava vektor kojemu su elementi kvadrirani elementi vektora $\boldsymbol{x}$. Sada, umjesto da računamo generalizirani inverz ispočetka, možemo iskoristiti teorem 1.3.2 i izračunati $A^\dagger$ tako definirane matrice.



Slika 1.1: Aproksimacija funkcije polinomom

Očito je polinom drugog stupnja bolja aproksimacija ovih točaka od polinoma prvog stupnja. Također, ako dodamo još jednu vrijednost (x_{11}, y_{11}), onda možemo algoritam iskoristiti na transponiranoj matrici.

Primjer 1.3.10. *Iskoristimo prethodni algoritam za računanje Moore–Penroseovog inverza loše uvjetovane matrice:*

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

Primijetimo da je treći stupac matrice zbroj prva dva, pa matrica nije punog stupčanog ranga. Pošto je broj singularnih vrijednosti koje su veće od nule upravo rang matrice, a uvjetovanost dobijamo kao omjer najveće i najmanje singularne vrijednosti, koja je u ovom slučaju jednaka nuli. Kao što je i očekivano, izračunata matrica nije ni blizu Moore–Penroseovom inverzu, ako provjeravamo Penroseove uvjete (1)–(4) *za tako dobivenu matricu, od lijeve strane oduzmemo desnu i izračunamo normu, dobijemo*

$\|\|AXA - A\|\|$	4.3284
$\|\|XAX - X\|\|$	2.5713e+15
$\|\|(AX)^* - AX\|\|$	0.86594
$\|\|(XA)^* - XA\|\|$	7.2967e+15

Za loše uvjetovane probleme koristimo dekompoziciju singularnih vrijednosti, SVD. Računanje SVD-a matrice A može se svesti na implicirno računanje svojstvenih vrijednosti simetrične matrice $A^T A$. Na primjer, za to računanje može se korititi Jacobijev algoritam, opisan na simetričnoj matrici B.

Definicija 1.3.11. *Ravninska rotacija je matrica $R(\theta, i, j)$ u ravnini (i, j) jednaka je jediničnoj matrici, osim na četiri mjesta: $r_{i,i} = \cos\theta, r_{i,j} = -\sin\theta, r_{j,i} = \sin\theta, r_{j,j} = \cos\theta$.*

Ideja je da djelovanjem ortogonalnih transformacija dobijemo niz sličnih simetričnih matrica

$$B_k = Q_{k-1}^T \cdots Q_1^T B Q_1 \cdots Q_{k-1} = (Q_1 \cdots Q_{k-1})^T B Q_1 \cdots Q_{k-1}$$

koje konvergiraju k fiksnoj dijagonalnoj matrici D.

Simetrična matrica reda 2,

$$M = \begin{bmatrix} \alpha & \gamma \\ \gamma & \beta \end{bmatrix}$$

dijagonalizira se pomoću ravninske rotacije

$$R = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

i to tako da je matrica

$$R^T MR = \begin{bmatrix} \alpha\cos^2\theta - 2\gamma\sin\theta\cos\theta + \beta\sin^2 & \gamma(\cos^2\theta - \sin^2\theta) + (\alpha-\beta)\sin\theta\cos\theta \\ \gamma(\cos^2\theta - \sin^2\theta) + (\alpha-\beta)\sin\theta\cos\theta & \alpha\sin^2\theta + 2\gamma\sin\theta\cos\theta + \beta\cos^2\theta \end{bmatrix}$$

dijagonalna. Najjednostavniji način za računanje kuta rotacije je napisati izvandijagonalne elemente kao funkciju tangensa ili kotangensa kuta 2θ. Ako definiramo

$$\zeta = \mathrm{ctg}(2\theta) = \frac{\beta-\alpha}{2\gamma},$$

onda je

$$\mathrm{ctg}(2\theta) = \frac{1-t^2}{2t},$$

pri čemu je $t := \mathrm{tg}\,\theta$ pa se radi izbjegavanja katastrofalnog kraćenja, tangens kuta može stabilno izračunati formulom

$$t = \frac{\mathrm{sign}(\zeta)}{|\zeta| + \sqrt{\zeta^2+1}}.$$

Iz osnovnih trigonometrijskih identiteta onda slijedi da je

$$\cos\theta = \frac{1}{\sqrt{1+t^2}} \quad \text{i} \quad \sin\theta = \frac{t}{\sqrt{1+t^2}}.$$

Sad su dijagonalni elementi dani s $\alpha' = \alpha - \gamma t$ i $\beta' = \beta + \gamma t$.

Za matricu B dimenzija $m \times m$ koristimo ravninsku rotaciju $R(\theta, i, j)$ koja mijenja i-ti i j-ti redak i stupac tako da dijagonalizira 2×2 podmatricu koja se nalazi na presjeku i-tog i j-tog retka i stupca. Promijenjeni elementi su dani izrazom $k \neq i, j$

$$\begin{aligned} b'_{i,k} &= \cos\theta b_{i,k} - \sin\theta b_{j,k} \\ b'_{k,i} &= \cos\theta b_{k,i} - \sin\theta b_{k,j} \\ b'_{j,k} &= \sin\theta b_{i,k} + \cos\theta b_{j,k} \\ b'_{k,j} &= \sin\theta b_{k,i} + \cos\theta b_{k,j}. \end{aligned}$$

Već u sljedećem koraku element koji je poništen će se, u većini slučajeva, pokvariti – tj. ponovno će postati ne-nula. Kako znamo da će algoritam uopće konvergirati i kako biramo elemente koje poništavamo? Postoji nekoliko strategija za koje je dokazana konvergencija.

Može se birati najveći po apsolutnoj vrijednosti izvandijagonalni element, ali to u svakom koraku zahtijeva pretraživanje matrice kako bismo ga pronašli pa nije efikasno. Nadalje, postoje cikličke strategije, koje redom poništavaju elemente po retcima ili po stupcima, također, konvergiraju.

Ako rotacije koje koristimo za dijagonalizaciju primjenjujemo na matrici koja je na početku bila jedinična, na kraju procesa na tom mjestu akumulirat će se matrica svojstvenih vektora.

Kako taj algoritam koristimo za nalaženje SVD-a? Nećemo eksplicitno računati matricu $A^T A$, već definiramo $U = A$ i $V = I$. U svakom koraku k izračunamo iz matrice U_k samo četiri elementa matrice $A_k^T A_k$ koji su potrebni za dobivanje kuta θ i djelujemo rotacijom na stupce matrica U_k i, ako je potrebno računati desne singularne vektore, i na stupce matrice V_k. Na kraju algoritma u U_n je spremljena matrica lijevih singularnih vektora, pomnožena dijagonalnom matricom singularnih vrijednosti. Da bismo dobili lijeve singularne vektore, stupace treba normirati, a pripadne norme su singularne vrijednosti. Na kraju algoritma, u matrici V_n nalaze se desni singularni vektori.

Poglavlje 2

Drazinov inverz

2.1 Definicija i svojstva

Poznato je da je λ, $\lambda \neq 0$, svojstvena vrijednost matrice A i $\boldsymbol{x}$ je pripadni svojstveni vektor ako i samo ako vrijedi da je $\boldsymbol{x}$ svojstveni vektor matrice A^{-1} pridružen svojstvenoj vrijednosti $\frac{1}{\lambda}$. Cilj je pronaći matricu za koju bi to vrijedilo čak i ako matrica A nije regularna, odnosno ako je 0 jedna od njezinih svojstvenih vrijednosti.

Neka je $A \in \mathbb{C}^{n\times n}$. Matrica $X \in \mathbb{C}^{n\times n}$ koja zadovoljava:

$$A^{k+1}X = A^k \tag{1^k}$$

$$XAX = X \tag{2}$$

$$AX = XA, \tag{5}$$

zove Drazinov inverz matrice A i označava se s A^D.

Oznaka je opravdana zbog jedinstvenosti takve matrice. Pretpostavimo da postoje matrice X i Y koje obje zadovoljavaju sva tri svojstva.

$$\begin{aligned} X &= XAX = XA^kX^k = A^kX^{k+1} = A^{k+1}YX^{k+1} = YA^{k+1}X^{k+1} \\ &= YAX = Y^{k+1}A^{k+1}X = Y^{k+1}A^k = YYA = YAY = Y. \end{aligned}$$

Sljedeće nas zanima postojanje Drazinovog inverza. Ne želimo se ograničiti samo na dijagonalizabilne matrice pa promatramo Jordanovu formu matrice.

Definicija 2.1.1. *Neka je $A \in \mathbb{C}^{n\times n}$ i $\boldsymbol{x}$ vektor različit od nul-vektora takav da postoji pozitivan $p \in \mathbb{N}$ i skalar $\lambda \in \sigma(A)$ za koji vrijedi*

$$(A - \lambda I)^p\boldsymbol{x} = \boldsymbol{0} \quad i \quad (A - \lambda I)^{p-1}\boldsymbol{x} \neq \boldsymbol{0}.$$

Tada $\boldsymbol{x}$ zovemo generalizirani svojstveni vektor matrice A ranga p.

Definicija 2.1.2. *Neka je $A \in \mathbb{C}^{n\times n}$. Najmanji nenegativni cijeli broj k takav da je*

$$\mathbb{C}^n = \mathrm{Im}(A^k) \oplus \mathrm{Ker}(A^k),$$

ili ekvivalentno, najmanji nenegativni cijeli broj takav da je

$$\mathrm{rang}(A^k) = \mathrm{rang}(A^{k+1})$$

zove se indeks matrice A i označava s $\mathrm{Ind}(A)$.

Očito vrijedi da je indeks regularne matrice jednak 0.

Teorem 2.1.3 (Jordanova forma)**.** *Neka je $A \in \mathbb{C}^{n\times n}$. Matrica A je slična blok-dijagonalnoj matrici kojoj su dijagonalni blokovi jednaki Jordanovim blokovima,*

$$P^{-1}AP = J = \begin{bmatrix} J_{k_1}(\lambda_1) & 0 & \dots & 0 \\ 0 & J_{k_2}(\lambda_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & J_{k_p}(\lambda_p) \end{bmatrix},$$

gdje su $\lambda_1, \dots, \lambda_p$ svojstvene vrijednosti matrice A. Jordanovi blokovi su bidijagonalne matrice

$$J_{k_i}(\lambda_i) = \begin{bmatrix} \lambda_i & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda_i & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \lambda_i & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda_i \end{bmatrix}.$$

Neka je dana Jordanova forma matrice A, pri čemu je $AP = PJ(\lambda)$. Ako P particioniramo po stupcima, dobijemo da je

$$\begin{aligned} A\boldsymbol{p_1} &= \lambda \boldsymbol{p_1}, \\ A\boldsymbol{p_i} &= \boldsymbol{p_{i-1}} + \lambda \boldsymbol{p_i}, \quad \text{za } i > 1, \end{aligned}$$

odnosno,

$$(A - \lambda I)\boldsymbol{p_i} = \boldsymbol{p_{i-1}},$$

pa množeći slijeva s $A - \lambda I$ dobivamo

$$(A - \lambda I)^{i-1}\boldsymbol{p_i} = \boldsymbol{p_1}$$

i

$$(A - \lambda I)^{i}\boldsymbol{p_i} = \boldsymbol{0}.$$

Time smo dobili vezu između Jordanove forme i generaliziranih svojstvenih vektora. Promotrimo reprezentaciju Drazinovog inverza.

Teorem 2.1.4. *Neka je $A \in \mathbb{C}^{n\times n}$ i neka je njezina Jordanova forma dana s*

$$A = PJP^{-1} = P\begin{bmatrix} J_1 & 0 \\ 0 & J_0 \end{bmatrix} P^{-1},$$

gdje je J_0 dio matrice J koji odgovara svojstvenim vrijednostima jednakim nula, a J_1 dio matrice koji odgovara ne-nul svojstvenim vrijednostima. Tada je

$$A^D = P\begin{bmatrix} J_1^{-1} & 0 \\ 0 & 0 \end{bmatrix} P^{-1}.$$

Dokaz. Jasno je da matrica J_1 ima inverz, a J_0 je nilpotentna matrica. Ako je indeks matrice A jednak k, onda je $J_0^l = 0$ za svaki $l \geq k$. Sad provjeravamo svojstva Drazinovog inverza,

$$\begin{aligned} A^{k+1}A^D &= P\begin{bmatrix} J_1^{k+1} & 0 \\ 0 & 0 \end{bmatrix} P^{-1}P\begin{bmatrix} J_1^{-1} & 0 \\ 0 & 0 \end{bmatrix} P^{-1} = P\begin{bmatrix} J_1^{k} & 0 \\ 0 & 0 \end{bmatrix} P^{-1} = A^k, \\ A^DAA^D &= P\begin{bmatrix} J_1^{-1} & 0 \\ 0 & 0 \end{bmatrix} P^{-1}P\begin{bmatrix} J_1 & 0 \\ 0 & J_0 \end{bmatrix} P^{-1}P\begin{bmatrix} J_1^{-1} & 0 \\ 0 & 0 \end{bmatrix} P^{-1} = P\begin{bmatrix} J_1^{-1} & 0 \\ 0 & 0 \end{bmatrix} P^{-1} = A^D, \\ AA^D &= P\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} P^{-1} = A^DA. \end{aligned}$$

□

Za regularnu matricu A i svojstvenu vrijednost λ vrijedi da je $\boldsymbol{x}$ generalizirani svojstveni vektor ranga p pridružen svojstvenoj vrijednosti λ ako i samo ako je generalizirani svojstveni vektor ranga p za matricu A^{-1} i svojstvenu vrijednost $\frac{1}{\lambda}$. Slično svojstvo vrijedi i za Drazinov inverz.

Teorem 2.1.5. *Za matricu $A \in \mathbb{C}^{n\times n}$ takvu da je $\operatorname{Ind}(A) = k$, vrijedi da je $\lambda \in \sigma(A)$ ako i samo ako je $\lambda^\dagger \in \sigma(A^D)$. $\boldsymbol{x}$ je generalizirani svojstveni vektor matrice A ranga p pridružen svojstvenoj vrijednosti $\lambda \in \sigma(A)$, $\lambda \neq 0$ ako i samo ako je $\boldsymbol{x}$ generalizirani svojstveni vektor matrice A^D ranga p pridružen svojstvenoj vrijednosti $\lambda^{-1} \in \sigma(A^D)$. Nadalje, $\boldsymbol{x}$ je generalizirani svojstveni vektor matrice A pridružen svojstvenoj vrijednosti $\lambda = 0$ ako i samo ako je $\boldsymbol{x} \in \operatorname{Ker}(A^k) = \operatorname{Ker}(A^D)$.*

Dokaz. Primijetimo da drugi dio tvrdnje posebno kaže da je $\boldsymbol{x}$ generalizirani svojstveni vektor matrice A^D. Ako je $\operatorname{Ind}(A) = 0$, dokaz je gotov jer je A regularna matrica.

Inače, A zapisujemo u obliku

$$A = P\begin{bmatrix} J_1 & 0 \\ 0 & J_0 \end{bmatrix} P^{-1}$$

i promatramo vektor

$$\boldsymbol{x} = P\begin{bmatrix}\boldsymbol{u_1}\\ \boldsymbol{u_2}\end{bmatrix}.$$

Tada je $\boldsymbol{x}$ generalizirani svojstveni vektor matrice A ranga p pridružen svojstvenoj vrijednosti $\lambda \neq 0$ ako i samo ako vrijedi da je $\boldsymbol{u_2} = \boldsymbol{0}$ i $\boldsymbol{u_1}$ je generalizirani vektor ranga p matrice J_1. Pošto je J_1 invertibilna matrica, i

$$A^D = P\begin{bmatrix}J_1^{-1} & 0\\ 0 & 0\end{bmatrix}P^{-1},$$

pa vrijedi tvrdnja.

Zbog svojstva (1^k) vrijedi da je $\mathrm{Ker}(A^D) \subset \mathrm{Ker}(A^k)$, a zbog

$$A^D\boldsymbol{x} = A^DAA^D\boldsymbol{x} = A^D(A^k(A^D)^k)\boldsymbol{x},$$

što izlazi iz svojstava (2) i (5), vrijedi i drugi smjer, pa je tvrdnja očita. □

Još jedno svojstvo Drazinovog inverza A^D je da se može zapisati kao polinom u A. Prisjetimo se izraza za regularnu matricu: Hamilton–Cayleyev teorem kaže da matrica poništava svoj karakteristični polinom pa se stoga njen inverz može zapisati kao polinom u A:

$$A^{-1} = \frac{(-1)^{n-1}}{\det(A)}(A^{n-1} + c_{n-1}A^{n-2} + \cdots + c_1 I).$$

Slično vrijedi i za Drazinov inverz.

Teorem 2.1.6. *Neka je $A \in \mathbb{C}^{n\times n}$, tada postoji polinom $p(x)$ takav da vrijedi $A^D = p(A)$.*

Dokaz. Zapisujemo A kao

$$A = P\begin{bmatrix}J_1 & 0\\ 0 & J_0\end{bmatrix}P^{-1},$$

gdje su P i J_1 regularne matrice i J_0 je nilpotentna indeksa k. Zbog regularnosti matrice J_1, postoji polinom $q(x)$ takav da je $J_1^{-1} = q(J_1)$. Definiramo

$$p(x) = x^k(q(x))^{k+1}.$$

Tada vrijedi

$$\begin{aligned}p(A) = A^k(q(A))^{k+1} &= P\begin{bmatrix}J_1^k & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}q(J_1) & 0\\ 0 & q(J_0)\end{bmatrix}^{k+1}P^{-1}\\ &= P\begin{bmatrix}J_1^k(q(J_1))^{k+1} & 0\\ 0 & 0\end{bmatrix}P^{-1} = P\begin{bmatrix}J_1^{-1} & 0\\ 0 & 0\end{bmatrix}P^{-1} = A^D.\end{aligned}$$

□

Dalje navodimo neka lako dokaziva svojstva Drazinovog inverza.

Propozicija 2.1.7. *Neka je* $A \in \mathbb{C}^{n \times n}$. *Vrijede sljedeća svojstva:*

1. $(A^*)^D = (A^D)^*$ *i* $(A^T)^D = (A^D)^T$;
2. $(A^l)^D = (A^D)^l \quad \forall l \in \mathbb{N}$;
3. $(A^D)^D = A$ *ako i samo ako je* $\mathrm{Ind}(A) = 1$;
4. $((A^D)^D)^D = A^D$;
5. *ako je* $\mathrm{Ind}(A) = k$, *onda je* $\mathrm{Im}(A^D) = \mathrm{Im}(A^l)$ *i* $Ker(A^D) = \mathrm{Ker}(A^l)$ *za sve* $l \geq k$.

Zbog važnosti u primjenama, Drazinov inverz matrice A indeksa 1 ima posebno ime i oznaku.

Definicija 2.1.8. *Matrica* $X \in \mathbb{C}^{n \times n}$ *koja za matricu* $A \in \mathbb{C}^{n \times n}$, $\mathrm{Ind}(A) = 1$ *zadovoljava svojstva:*

$$\begin{aligned} AXA &= A, \\ XAX &= X, \\ AX &= XA, \end{aligned}$$

se zove grupni inverz i označava se $A^\#$.

Za $A^\#$ vrijede stroža svojstva vezana uz generalizirane svojstvene vektore dana u sljedećem teoremu.

Teorem 2.1.9. *Za zadanu matricu* $A \in \mathbb{C}^{n \times n}$, *kojoj je* $\mathrm{Ind}(A) = 1$, *vektor* $\boldsymbol{x}$ *je generalizirani svojstveni vektor ranga p pridružen svojstvenoj vrijednosti* λ *ako i samo ako je generalizirani svojstveni vektor matrice* $A^\#$ *ranga p pridružen svojstvenoj vrijednosti* $\lambda^\dagger$.

Dokaz. Ako A ima indeks 1, onda je njezina Jordanova forma jednaka

$$A = P \begin{bmatrix} J_1 & 0 \\ 0 & 0 \end{bmatrix} P^{-1}$$

i

$$A^\# = P \begin{bmatrix} J_1^{-1} & 0 \\ 0 & 0 \end{bmatrix} P^{-1},$$

pa tvrdnja vrijedi za $\lambda \neq 0$ zbog regularnosti matrice J_1, odnosno, ekvivalentne tvrdnje za regularne matrice. Za $\lambda = 0$ tvrdnja slijedi zato što je $\mathrm{Ker}(A^\#) = \mathrm{Ker}(A)$ i zato što su svi generalizirani svojstveni vektori matrice A pridruženi svojstvenoj vrijednosti 0 zbog $\mathrm{Ind}(A) = 1$ upravo svojstveni vektori. □

Propozicija 2.1.10. *Neka je $A \in \mathbb{C}^{n\times n}$ i neka vrijedi $\mathrm{Ind}(A) = 1$. Tada postoji $A^\#$ i vrijedi:*

1. $(A^\#)^\# = A$;
2. *ako je* $\mathrm{Ind}(A) = k$, *onda je* $\mathrm{Ind}(A^l) = 1$ *i* $(A^l)^\# = (A^D)^l$ *za sve* $l \geq k$;
3. *ako* $\mathrm{Ind}(A) = k$, *onda je* $\mathrm{Ind}(A^D) = 1$ *i* $(A^D)^\# = A^2A^D$;
4. *ako označimo* $A^{-j} := (A^\#)^j$ *za sve* $j \in \mathbb{N}$ *i* $A^0 := AA^\#$, *onda vrijedi* $A^lA^m = A^{l+m}$ *za sve* l, m.

Iz zadnjeg svojstva propozicije vidljivo je da tako definirano potenciranje matrice A čini Abelovu grupu s operacijom matričnog množenja kao binomnom operacijom, pa je opravdan naziv $A^\#$ kao grupnog inverza.

Sljedeći teorem daje uvjet egzistencije $A^\#$ i formulu za njegovo računanje.

Teorem 2.1.11. *Neka $A \in \mathbb{C}^{n\times n}$ ima faktorizaciju punog ranga*

$$A = FG.$$

Tada A ima grupni inverz $A^\#$ ako i samo ako je GF regularna matrica i vrijedi

$$A^\# = F(GF)^{-2}G.$$

Dokaz. Ako je $\mathrm{Ind}(A) = 1$ onda je $\mathrm{rang}(A^2) = \mathrm{rang}(A)$. Neka je $r = \mathrm{rang}(A)$, tada je $GF \in \mathbb{C}^{r\times r}$,

$$A^2 = FGFG,$$

pa je $\mathrm{rang}(A^2) = \mathrm{rang}(GF)$, jer je F punog stupčanog, a G punog retčanog ranga. Nadalje, $r = \mathrm{rang}(GF)$ ako i samo ako je GF regularna matrica. Provjerimo da vrijede uvjeti iz definicije grupnog inverza:

$$\begin{aligned} AA^\#A &= FGF(GF)^{-2}GFG = FG = A, \\ A^\#AA^\# &= F(GF)^{-2}GFGF(GF)^{-2}G = A^\#, \\ AA^\# &= FGF(GF)^{-2}G = A^\#A. \end{aligned}$$

□

Upotreba Drazinovog inverza

U prošlom poglavlju vidjeli smo da Moore–Penroseov inverz matrice daje rješenje linearnog sustava. Vrijedi li slično i za Drazinov inverz?

Promatramo sustav linearnih diferencijalnih jednadžbi

$$\boldsymbol{x}' + A\boldsymbol{x} = \boldsymbol{f}.$$

Rješenje tog sustava dano je s

$$\boldsymbol{x} = e^{-At}\left(\int e^{At}\boldsymbol{f}(t)dt\right).$$

Eksponencijalnu funkciju zapisujemo u obliku reda

$$e^{At} = \sum_{n=0}^{\infty}\frac{A^n t^n}{n!}.$$

U slučaju da je A regularna matrica, integral dobijemo iz

$$\int e^{At}dt = A^{-1}e^{At} + C.$$

Što možemo reći o rješenju ako je A singularna matrica? Neka je $\mathrm{Ind}(A) = k$. Tada je

$$\int e^{At}dt = A^D e^{At} + (I - AA^D)\left[tI + A\frac{t^2}{2!} + \ldots + A^{k-1}\frac{t^k}{k!}\right] + C,$$

što se lako vidi deriviranjem po t. Ako je $\boldsymbol{f}$ konstanta u vremenu t, onda je rješenje za singularnu matricu A dano s

$$\boldsymbol{x} = A^D\boldsymbol{f} + (I - AA^D)\left[tI - A\frac{t^2}{2!} + A^2\frac{t^3}{3!} - \ldots + (-1)^{k-1}A^{k-1}\frac{t^k}{k!}\right]\boldsymbol{f}.$$

Možemo li sada reći nešto o sustavu

$$A\boldsymbol{x}' + B\boldsymbol{x} = \boldsymbol{f}.$$

Kad bi A bila regularna, mogli bismo množiti slijeva s A^{-1}, što bi se svelo na prvi slučaj. Pretpostavimo stoga da je A singularna i da komutira s B, tj. da vrijedi $AB = BA$. Pošto se eksponencijalna funkcija može zapisati kao polinom, vrijedi da je rješenje homogene jednadžbe

$$A\boldsymbol{x}' + B\boldsymbol{x} = \boldsymbol{0}$$

dano s

$$\boldsymbol{x} = e^{-A^D Bt}A^D A\boldsymbol{v}$$

za proizvoljni vektor $\boldsymbol{v}$.

Upotreba grupnog inverza u Markovljevim lancima

Definicija 2.1.12. *Neka je S skup. Slučajan proces s diskretnim vremenom i prostorom stanja S je familija $X = (X_n : n \geq 0)$ slučajnih varijabli definiranih na nekom vjerojatnosnom prostoru (Ω, F, P) s vrijednostima u S. Dakle, za svaki $n \geq 0$ je $X_n : \Omega \to S$ slučajna varijabla.*

Definicija 2.1.13. *Neka je S prebrojiv skup. Slučajni proces $X = (X_n : n \geq 0)$ definiran na vjerojatnonom prostoru (Ω, F, P) s vrijednostima u skupu S je Markovljev lanac ako vrijedi*

$$P(X_{n+1} = j \mid X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) = P(X_{n+1} = j \mid X_n = i)$$

za svaki $n \geq 0$ i za svaki $i_0, i_1, \dots, i_{n-1}, i, j \in S$ za koje su obje uvjetne vjerojatnosti dobro definirane.

Elementi skupa S nazivaju se stanja Markovljevog lanca.

Definicija 2.1.14. *Matrica $P = (p_{ij} : i, j \in S)$ naziva se stohastičkom matricom ako je $p_{ij} \geq 0$ za svaki $i, j \in S$ i $\sum_{j \in S} p_{ij} = 1$ za svaki $i \in S$.*

Homogeni Markovljev lanac je Markovljev lanac u kojem vjerojatnost ne ovisi o trenutku n. Tada možemo označiti $p_{ij} = P(X_{n+1} = j \mid X_n = i)$ i $P = (p_{ij})$ i tako dobivenu matricu zovemo prijelazna matrica Markovljevog lanca.

Stanje j je dostižno iz stanja i, u oznaci $i \to j$ ako je $p_{ij}^{(n)} > 0$ za neki $n \in \mathbb{N}$, gdje je $P^n = (p_{ij}^{(n)} : i, j \in S)$. Kažemo da stanja i i j komuniciraju, označavamo $i \leftrightarrow j$ ako vrijedi $i \to j$ i $j \to i$.

Definicija 2.1.15. *Za matricu $A \in \mathbb{R}^{n \times n}$ kažemo da je reducibilna ako postoji permutacijska matrica $P \in \mathbb{R}^{n \times n}$ takva da je PAP^{-1} blok-gornjetrokutasta matrica. Matrica koja nije reducibilna je ireducibilna.*

Markovljev lanac je ireducibilan ako se prostor stanja S sastoji od samo jedne klase komuniciranja, odnosno ako sva njegova stanja komuniciraju. U promatranjima ćemo se ograničiti na konačne Markovljeve lance, tj. Markovljeve lance s konačnim skupom stanja.

Definicija 2.1.16. *Neka je $X = (X_n : n \geq 0)$ Markovljev lanac s prebrojivim skupom stanja S i prijelaznom matricom P. Vjerojatnosna distribucija $\boldsymbol{\pi} = (\pi_i : i \in S)$ na S je stacionarna distribucija Markovljevog lanca X ako vrijedi $\boldsymbol{\pi} = \boldsymbol{\pi} P$.*

Definicija 2.1.17. *Neka je $X = (X_n : n \geq 0)$ Markovljev lanac s prebrojivim skupom stanja S i prijelaznom matricom P. Vjerojatnosna distribucija $\boldsymbol{\pi} = (\pi_i : i \in S)$ je granična distribucija Markovljevog lanca X ako za sve $i, j \in S$ vrijedi $\lim_{n\to\infty} p_{ij}^{(n)} = \pi_j$.*

Granična distribucija Markovljevog lanca je i stacionarna distribucija tog Markovljevog lanca.

Teorem 2.1.18 ([2, str. 152–154])**.** *Neka je $P \in \mathbb{R}^{n\times n}$ prijelazna matrica i neka je $A = I - P$, tada je* $\mathrm{Ind}(A) = 1$*, odnosno postoji $A^{\#}$.*

Dokaz. Ako je P ireducibilna, onda je 1 po Perron–Frobeniusovom teoremu svojstvena vrijednost matrice P algebarske kratnosti 1. Slijedi da je 0 svojstvena vrijednost matrice $A = I - P$ algebarske kratnosti 1, odnosno $\mathrm{Ind}(A) = 1$ pa postoji $A^{\#}$.

Ako je P reducibilna matrica, možemo ju permutirati tako da je blok-gornjetrokutasta i da su joj dijagonalni blokovi ireducibilni.

Ako je B ireducibilna matrica, $B \geq 0$ (svi elementi matrice su ≥ 0) i ako postoji $M \geq 0$, ne-nul matrica takva da je $S = B + M$ stohastička, tada za spektralni radijus ρ matrice B vrijedi da je $\rho(B) < 1$.

Dijagonalni blokovi matrice P su, ili stohastički (kao što je to zadnji dijagonalni blok), ili zadovoljavaju uvjete prethodne tvrdnje, što znači da matricu P možemo permutirati tako da vrijedi

$$P = \begin{bmatrix} P_{11} & P_{12} & \dots & P_{1,r} & P_{1,r+1} & P_{1,r+2} & \dots & P_{1,k} \\ 0 & P_{22} & \dots & P_{2,r} & P_{2,r+1} & P_{2,r+2} & \dots & P_{2,k} \\ \vdots & \vdots & \ddots & P_{r,r} & P_{r,r+1} & P_{r,r+2} & \dots & P_{r,k} \\ 0 & 0 & 0 & 0 & P_{r+1,r+1} & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & 0 & P_{r+2,r+2} & \dots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & P_{k,k} \end{bmatrix},$$

gdje su $I - P_{ii}$ za $i = 1, \dots, r$ invertibilni, a blokovi P_{ii} za $i = r + 1, \dots, k$ su ireducibilni, pa po prvom dijelu dokaza postoji grupni inverz.

Neka je

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

gdje je A invertibilna matrica, a za C postoji $C^{\#}$, tada postoji $M^{\#}$ i jednak je

$$M^{\#} = \begin{bmatrix} A^{-1} & A^{-2}B(I - CC^{\#}) - A^{-1}BC^{\#} \\ 0 & C^{\#} \end{bmatrix}.$$

Stoga slijedi da postoji grupni inverz matrice $I - P$. □

Matrica $A = I - P$ i njen grupni inverz pokazat će se važnima u razmatranju svojstava Markovljevog lanca.

Stanje i je aperiodično ako postoji $m_0 \in N$ takav da je $p_{ii}^{(m)} > 0$ za svaki $m \geq m_0$. U suprotnom stanje zovemo periodičnim. Ako $i \leftrightarrow j$ i stanje i je aperiodično, onda je i stanje j aperiodično. Ireducibilan Markovljev lanac zovemo aperiodičnim ako ima aperiodično stanje i, pošto mu sva stanja komuniciraju, onda su sva stanja ireducibilnog Markovljevog lanca aperiodična.

Stanje $i \in S$ je apsorbirajuće ako se iz njega ne može izići, odnosno, vjerojatnost da se iz j dođe u bilo koje drugo stanje $i \in S$ je 0. Markovljev lanac nazivamo apsorbirajućim ako ima bar jedno apsorbirajuće stanje.

Teorem 2.1.19 ([2, str. 155–156]). *Neka je $P \in \mathbb{R}^{n\times n}$ prijelazna matrica konačnog Markovljevog lanca i neka je $A = I - P$. Tada je*

$$I - AA^{\#} = \begin{cases} \lim\limits_{m\to\infty} \dfrac{I + P + \ldots + P^{m-1}}{m}, & \textit{za svaku prijelaznu matricu}, \\ \lim\limits_{m\to\infty} (\alpha I + (1-\alpha)P)^m, & \textit{za svaku prijelaznu matricu } P \textit{ i } 0 < \alpha < 1, \\ \lim\limits_{m\to\infty} P^m, & \textit{za svaki regularni lanac}, \\ \lim\limits_{m\to\infty} P^m, & \textit{za svaki apsorbirajući lanac}. \end{cases}$$

Teorem 2.1.20 ([2, str. 156]). *Ako je P prijelazna matrica ireducibilnog, aperiodičnog Markovljevog lanca X tada je svaki redak matrice $I - AA^{\#}$ jednak vektoru*

$$\boldsymbol{w}^* = \left[w_1, w_2, \ldots, w_n\right],$$

koji je upravo jednak stacionarnoj distribuciji tog Markovljevog lanca X.

Dokaz. Ireducibilan, aperiodičan Markovljev lanac ima graničnu distribuciju, koja je, po prethodnom teoremu, upravo jednaka retku matrice $I - AA^{\#}$, a granična distribucija je ujedno i stacionarna distribucija Markovljevog lanca. □

Korolar 2.1.21 ([2, str. 159]). *Neka je X ireducibilan Markovljev lanac sa stacionarnom distribucijom $\boldsymbol{w}^*$. Ako lanac kreće iz stanja $i \in S$ onda je očekivani broj koraka prije prvog povratka u stanje $i \in S$ dan s $1/(I - AA^{\#})_{ii}$.*

Također, vrijedi da je očekivani broj posjeta stanju j prije prvog povratka u stanje i dano s

$$\frac{1/(I - AA^{\#})_{jj}}{1/(I - AA^{\#})_{ii}}.$$

Pogledajmo sada jedan jednostavan zadatak iz [3, str. 21].

Primjer 2.1.22. *Baca se simetričan novčić. Koliko bacanja očekujemo (u prosjeku) između dva pojavljivanja kombinacije pismo-glava?*

Prostor stanja je dan s $S = PP, PG, GP, GG$, gdje P označava da je palo pismo, a G glava, X_n označava što je palo u $(n-1)$-vom i n-tom bacanju. Prijelazna matrica je

$$P = \begin{bmatrix} 0.5 & 0.5 & 0 & 0 \\ 0 & 0 & 0.5 & 0.5 \\ 0.5 & 0.5 & 0 & 0 \\ 0 & 0 & 0.5 & 0.5 \end{bmatrix}.$$

Računamo $A^\#$ matrice $A = I - P$ i dobivamo

$$A^\# = \begin{bmatrix} 1 & 0 & -0.5 & -0.5 \\ -0.5 & 0.5 & 0 & 0 \\ 0 & 0 & 0.5 & -0.5 \\ -0.5 & -0.5 & 0 & 1 \end{bmatrix}.$$

Sad je stacionarna distribucija dana retkom matrice $I - AA^\#$:

$$\begin{bmatrix} 0.25 & 0.25 & 0.25 & 0.25 \end{bmatrix}$$

i rješenje zadatka je dano prethodnim korolarom kao $1/(I - AA^\#)_{22} = 4$.

2.2 Računanje Drazinovog inverza

Za razliku od Moore–Penroseovog inverza, kod Drazinovog inverza teoretska razmatranja nisu dala korisnu metodu za računanje. Zato promatramo iterativne metode za računanje inverza. Navodimo tri algoritma, njihovu brzinu konvergencije te kako izabrati prikladnu početnu aproksimaciju kako bismo osigurali konvergenciju.

Ideja iz [5] je koristiti iterativne metode za računanje inverza matrice i pokazati konvergenciju prema Drazinovom inverzu ukoliko je matrica singularna.

Prva metoda koju promatramo je

$$V_{n+1} = V_n(2I - AV_n). \tag{i_1}$$

Pokazat ćemo da tako definiran niz (V_n) za neku početnu aproksimaciju V_0 konvergira prema A^D. Prvo su potrebne dvije tvrdnje.

Propozicija 2.2.1. *Za matricu $A \in \mathbb{C}^{n\times n}$ kojoj je* $\mathrm{Ind}(A) = k$ *vrijedi*

$$AA^D = A^D A = \boldsymbol{P}_{\mathrm{Im}(A^k),\mathrm{Ker}(A^k)}.$$

Ova tvrdnja slijedi direktno iz toga da je $\mathrm{Ind}(A) = k$ i (1^k) i (5).

Propozicija 2.2.2. *Neka su L i M potprostori od $\mathbb{C}^n$ takvi da je $L \oplus M = C^n$. Vrijedi*

- $\boldsymbol{P}_{L,M} Q = Q$ *ako i samo ako* $\operatorname{Im} Q \subseteq L$,
- $\boldsymbol{P}_{L,M} = Q$ *ako i samo ako* $M \subseteq \operatorname{Ker}(Q)$.

Dokaz. Vektor $\boldsymbol{x} \in \mathbb{C}^n$ možemo zapisati kao sumu

$$\boldsymbol{x} = \boldsymbol{y} + \boldsymbol{z}, \quad \boldsymbol{y} \in L, \quad \boldsymbol{z} \in M.$$

Tada je $\boldsymbol{P}_{L,M} \boldsymbol{x} = \boldsymbol{y}$.

Nadalje,

$$Q\boldsymbol{x} = \boldsymbol{P}_{L,M} Q\boldsymbol{x}$$

ako i samo ako je $Q\boldsymbol{x} \in L$, odnosno ako je $\operatorname{Im}(Q) \subseteq L$.

Konačno,

$$Q\boldsymbol{y} = Q\boldsymbol{P}_{L,M}\boldsymbol{x} = Q\boldsymbol{x} = Q\boldsymbol{y} + Q\boldsymbol{z},$$

ako i samo ako je $Q\boldsymbol{z} = \boldsymbol{0}$, odnosno, $\boldsymbol{z} \in \operatorname{Ker}(Q)$. □

Teorem 2.2.3. *Za $A \in \mathbb{C}^{n \times n}$ i $V_{n+1} = V_n(2I - AV_n)$ vrijedi*

$$\frac{\|A^D - V_n\|}{\|A^D\|} \leq \|I - AV_0\|^{2^n}.$$

Dokaz. Označimo s $E_n = I - AV_n$ za $n \in \mathbb{N}$. Tada je

$$E_{n+1} = I - AV_{n+1} = I - AV_n(2I - AV_n) = (I - AV_n)(I - AV_n) = E_n^2.$$

Za konzistentnu normu vrijedi

$$\|E_{n+1}\| \leq \|E_n\|^2 \leq \cdots \leq \|E_0\|^{2^{n+1}}.$$

Promotrimo sad

$$\delta_n = A^D - V_n.$$

Iz definicije metode za V_n i odabira V_0 slijedi da je $\operatorname{Im}(V_n) \subseteq \operatorname{Im}(A^k)$ i $\operatorname{Ker}(A^k) \subseteq \operatorname{Ker}(V_n)$, za $n \geq 0$.

Iz prethodne dvije propozicije sad slijedi da je

$$A^D A V_n = V_n = V_n A A^D.$$

Time smo dobili da je

$$\delta_n = A^D - V_n = A^D(I - AV_n) = A^D E_n,$$

pa slijedi ocjena

$$\|A^D - V_n\| \leq \|A^D\| \, \|E_0\|^{2^n} = \|A^D\| \, \|I - AV_0\|^{2^n}.$$ □

Sad je jasno da, ako izaberemo početnu iteraciju takvu da je $\|I - AV_0\| < 1$, iterativna metoda konvergira upravo prema Drazinovom inverzu. Promatramo još dva algoritma, oba preuzeta iz [5]. Definiramo

$$V_{n+1} = V_n(3I - AV_n(3I - AV_n)). \tag{i_2}$$

Dokaz konvergencije oba algoritma je analogan dokazu konvergencije prve metode.

Konačno, treći algoritam definira se s

$$\begin{aligned} \chi_n &= -7I + \psi_n(9I + \psi_n(-5I + \psi_n)), \\ \vartheta_n &= \psi_n \chi_n, \\ V_{n+1} &= -\frac{1}{8} V_n \chi_n (12I + \vartheta_n(6I + \vartheta_n)), \end{aligned} \tag{i_3}$$

gdje je $\psi_n = AV_n$.

Za sve tri metode vrijedi da su konvergentne u slučaju da za početnu iteraciju $V_0 = \alpha A^k$ vrijedi $\|I - AV_0\| < 1$. To možemo osigurati tako da izaberemo

$$\alpha = \frac{2}{\mathrm{tr}(A^{k+1})},$$

pri čemu je tr oznaka za trag matrice. To se bazira na sljedećoj propoziciji i činjenici da je suma svojstvenih vrijednosti matrice upravo jednaka tragu te matrice.

Propozicija 2.2.4. *Neka je $M \in \mathbb{C}^{n \times n}$ i $\epsilon > 0$. Postoji matrična norma $\|\cdot\|$ takva da vrijedi*

$$\rho(M) \le \|M\| \le \rho(M) + \epsilon,$$

gdje je s $\rho(M)$ označen spektralni radijus matrice M.

Sva tri spomenuta algoritma se baziraju na matričnom množenju, što ih čini pogodnima za paralelizaciju.

Primjer 2.2.5. *Računamo Drazinov inverz matrice A*

$$A = \begin{bmatrix} 2 & 0.4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -2 & 0.4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & -1 & 1 & -1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ -1 & -1 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & -1 & -1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & -1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & -2 & 0.4 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0.4 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.4 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.4 & 2 \end{bmatrix}$$

preuzete iz [5]. *Njen indeks je jednak* 3, *pa biramo početnu aproksimaciju*

$$V_0 = \frac{2}{\mathrm{tr}(A^4)} A^3.$$

Uz toleranciju $1e{-}10$, *sve tri metode konvergiraju prema istoj matrici i vrijedi*

metoda	i_1	i_2	i_3
koraci	14	9	5
$\|\|A^3A^DA - A^3\|\|$	$2.0231e{-}13$	$4.5286e{-}11$	$2.2080e{-}11$
$\|\|A^DAA^D\|\|$	$3.8066e{-}13$	$3.0911e{-}12$	$2.0805e{-}12$
$\|\|AA^D - A^DA\|\|$	$3.8271e{-}13$	$2.5688e{-}10$	$5.1586e{-}11$

Primjer 2.2.6. *Želimo provjeriti konvergira li metoda i_3 najbrže prema rješenju. Uzimamo matrice sve veće dimenzije, uvijek indeksa tri, i promatramo broj potrebnih iteracija do konvergencije za svaku od metoda. Matrica je definirana s*

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

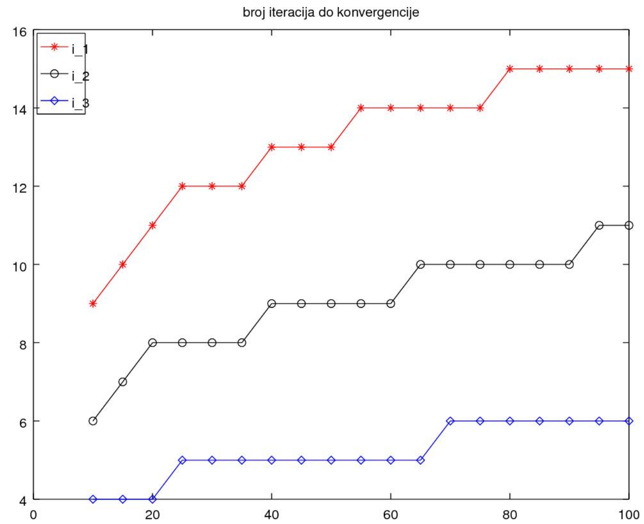
čime smo osigurali da će za svaki n matrica imati indeks 3. *Dimenzije matrice su u rasponu od* 10 *do* 100 *s korakom* 5 *i sve tri metode kreću s istom početnom aproksimacijom. Broj iteracija do konvergencije može se vidjeti na slici 2.1.*

Očito, metoda i_3 najbrže konvergira za svaki n. Primijetimo da je promatrana matrica jednostavna i skoro u Jordanovoj formi. Tako je izabrana zbog mogućnosti lakog određivanja indeksa matrice i početne aproksimacije.

Zašto je važno odrediti indeks matrice? Naposljetku, pokazali smo da će metode konvergirati za

$$V_0 = \frac{2}{\mathrm{tr}(A^{l+1})} A^l$$

za svaki $l \geq \mathrm{Ind}(A)$. U teoriji, to je istina i ništa nas ne sprječava da jednostavno uzmemo A^n. U praksi, što više potenciramo matricu A, dobiveno rješenje je sve dalje od stvarnog A^D.



Slika 2.1: Broj iteracija do konvergencije u primjeru 2.2.6

Primjer 2.2.7. *Pošto je prethodni primjer pokazao da upravo algoritam i_3 ima najbržu konvergenciju, promatramo samo rješenja dobivena tom metodom. Uzimamo i istu matricu, samo ovaj put fiksne dimenzije $n = 20$. Promatramo greške $\|A^3X_lA - A^3\|$, $\|X_lAX - X_l\|$ i $\|AX_l - X_lA\|$ za X_l dobiven kao rješenje metodom s početnom aproksimacijom*

$$V_0^l = \frac{2}{\mathrm{tr}(A^{l+1})}A^l, \quad l \in \{3, 4, 5, \dots, 20\}.$$

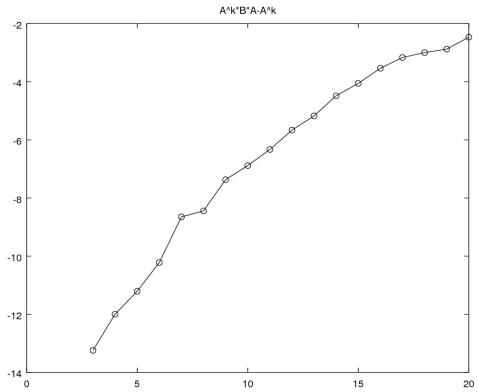
Pošto je indeks matrice poznat, uzet ćemo X_3 kao aproksimaciju Drazinovog inverza i pogledati normu razlike $\|X_l - X_3\|$. Na slikama 2.2–2.5 prikazane su razne mjere za pogrešku izračunatog inverza.

Kao što je i očekivano, izračunato rješenje je lošije za veće potencije matrice A.

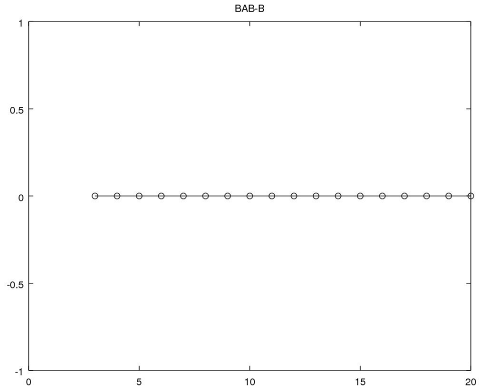
Ovaj primjer pokazuje da moramo imati neku ideju o indeksu matrice prije početka metode. Nažalost, u nekim primjerima ni to neće biti dovoljno.

Primjer 2.2.8. *Promatramo matricu indeksa 5, dimenzija 30×30 konstruiranu tako da na dijagonali ima 25 jedinica i 5 nula, dok su na gornjoj sporednoj dijagonali jedinice. Metoda i_3 s početnom iteracijom*

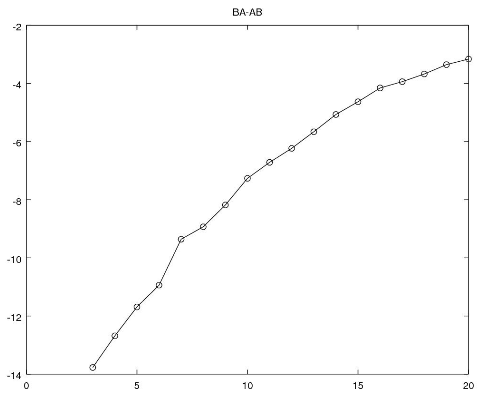
$$V_0 = \frac{2}{\mathrm{tr}(A^6)}A^5$$



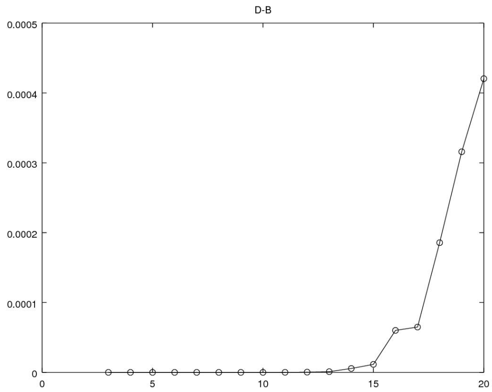
Slika 2.2: $\log_{10}(\|A^3X_lA - A^3\|)$



Slika 2.3: $\|X_lAX_l - X_l\|$



Slika 2.4: $\log_{10}(\|AX_l - X_lA\|)$



Slika 2.5: $\|X_l - X_3\|$

je iskonvergirala u 5 *koraka, ali dobiveno rješenje nije blizu Drazinovog inverza.*

$\|A^5XA - A^5\|$	0.012652
$\|XAX - X\|$	0
$\|AX - XA\|$	0.029138

Promotrimo sad stohastičke matrice. Dokazano je da ako je P stohastička da je onda $A = I - P$ indeksa 1. Time je odmah riješeno pitanje indeksa matrice i u računanju početne aproksimacije nema puno matričnog potenciranja.

Primjer 2.2.9. *Slično kao u prvom primjeru u ovom poglavlju, konstruiramo matrice sve većih dimenzija s tim da ovaj put počinjemo s matricama* 100×100 *i povećavamo im red* *za* 50*, sve dok ne dođemo od matrice dimenzija* 600×600. *Korištene test matrice nisu rijetko popunjene. Matrice konstruiramo pomoću Matlabove funkcije* rand() *koja vraća normalno distribuirane brojeve iz intervala* $(0, 1)$ *i nakon toga svaki redak podijelimo sa sumom elemenata retka kako bismo dobili stohastičku matricu. Norme različitih mjera grešaka dane su u narednoj tablici.*

n	$\|\|AXA - A\|\|$	$\|\|XAX - X\|\|$	$\|\|AX - XA\|\|$
100	$3.274692e{-}15$	$2.018293e{-}13$	$4.295226e{-}15$
150	$2.920582e{-}15$	$1.496199e{-}13$	$5.682733e{-}15$
200	$3.919980e{-}15$	$4.553684e{-}15$	$6.730610e{-}15$
250	$4.572605e{-}15$	$1.646839e{-}14$	$7.541633e{-}15$
300	$4.466758e{-}15$	$1.111230e{-}13$	$7.710160e{-}15$
350	$5.573431e{-}15$	$3.442800e{-}14$	$8.434284e{-}15$
400	$4.579168e{-}15$	$1.417037e{-}14$	$8.605520e{-}15$
450	$6.021053e{-}15$	$2.174383e{-}14$	$8.808872e{-}15$
500	$6.463138e{-}15$	$1.167561e{-}14$	$9.692232e{-}15$
550	$7.904070e{-}15$	$3.600841e{-}14$	$1.036156e{-}14$
600	$8.348713e{-}15$	$9.449247e{-}15$	$1.019290e{-}14$

Bibliografija

[1] A. Ben-Israel i T. N. E. Greville, *Generalized Inverses: Theory and Applications (second edition)*, Springer-Verlag, New York, 2004.

[2] S. L. Campbell i C. D. Meyer, *Generalized Incerses of Linear Transformations*, SIAM,Philadelphia, 2009.

[3] R. Mrazović i H. Planinić, *Markovljevi lanci-vježbe*, `https://web.math.pmf.unizg.hr/nastava/mala/vjezbe.pdf`, vježbe iz kolegija Markovljevi lanci na diplomskom studiju PMF–MO, posjećeno 31. 8. 2018.

[4] S. Singer i N. Bosner, *Numerička analiza 10. predavanje*, `https://web.math.pmf.unizg.hr/~nela/nadpredavanja/10nb.pdf`, predavanja iz kolegija Numerička analiza na doktorskom studiju PMF–MO, posjećeno 29. 8. 2018.

[5] F. Soleymani i P. S. Stanimirović, *A Higher Order Iterative Method for Computing the Drazin Inverse*, The Scientific World Journal **2013** (2013), Article ID 708647, `https://doi.org/10.1155/2013/708647`.

[6] Z. Vondraček, *Markovljevi lanci-predavanja*, `https://web.math.pmf.unizg.hr/~vondra/ml12-predavanja.html`, predavanja iz kolegija Markovljevi lanci na diplomskom studiju PMF–MO, posjećeno 31. 8. 2018.

[7] G. Wang, Y. Wei i S. Qiao, *Generalized Inverses: Theory and Computations*, Springer, Singapore, 2018.

Sažetak

Tema diplomskog rada je pojam generaliziranog inverza matrice. Prvo poglavlje se bavi Moore–Penroseovim inverzom i njegovom ulogom u rješavanju sustava linearnih jednadžbi. Nadalje, promotrili smo i $\{i, j, k\}$-inverze koji zadovoljavaju samo neke od Penroseovih uvjeta, svojstva koja zadovoljavaju i njihovu primjenu. Pokazane su i dvije metode za njegovo računanje, jedna koja se bazira na particioniranju matrice po stupcima i druga koja se bazira na dekompoziciji singularnih vrijednosti matrice.

U drugom poglavlju smo promotrili Drazinov inverz i njegov specijalan slučaj, grupni inverz. Spomenute su primjene u teoriji Markovljevih lanaca i u sustavima diferencijalnih jednadžbi. Detaljnije su proučena tri algoritma za njegovo računanje i njihova konvergencija, te njihovo ponašanje na odabranim primjerima.

Summary

This thesis' emphasis has been on the theory and application of generalized inverses. The first part is about the Moore–Penrose inverse and its role in finding solutions of linear systems. $\{i, j, k\}$-inverse satisfies some, but not all, of the Penrose equations. Its properties and applications are studied. Two methods of computing the Moore–Penrose inverse are mentioned, one relying on partitioning the matrix by columns, and the other on the singular value decomposition.

The second part was about the Drazin inverse and its properties. A special case of the Drazin inverse, the group inverse, was introduced and its role in the theory of finite Markov chains examined. An application to linear systems of differential equations was briefly discussed. Finally, three algorithms for its computing have been studied in some detail along with the convergence and some examples.

Životopis

Anamarija Zibar je rođena 7. 12. 1994. u Zagrebu. Osnovnu i srednju školu je pohađala u Glini. Nakon završetka srednje škole, 2013. godine upisuje preddiplomski studij Matematika na Matematičkom odsjeku Prirodoslovno–matematičkog fakulteta Sveučilišta u Zagrebu. Nakon završetka, 2016. godine upisuje diplomski studij Primijenjene matematike na istom fakultetu.