Operacije s matricama

Cvitković, Anamarija

Sveučilište J.J. Strossmayera u Osijeku
Odjel za matematiku
Sveučilišni preddiplomski studij matematike

Anamarija Cvitković

Operacije s matricama

Završni rad

Osijek, 2020.

Sveučilište J.J. Strossmayera u Osijeku
Odjel za matematiku
Sveučilišni preddiplomski studij matematike

Anamarija Cvitković

Operacije s matricama

Završni rad

Mentor: izv.prof.dr.sc. Darija Marković

Osijek, 2020.

Sažetak

U ovom radu proučavat ćemo neke operacije s matricama. Navest ćemo svojstva operacija i primjene. Osim poznatijih operacija kao što su zbrajanje matrica, množenje skalarom te standardno množenje matrica, razlikujemo još neke vrste množenja: Kroneckerov produkt, Hadamard-Schurov te Khatri-Raov produkt. Upoznat ćemo se i sa pojmom i svojstvima vektorizacije te matričnih derivacija.

Ključne riječi: Kroneckerov produkt, vektorizacija, Hadamard-Schurov produkt, Khatri-Raov produkt, matrične derivacije, svojstvene vrijednosti

Abstract

In this paper we will study some operations on matrices. We will state properties and usefulness. Exept for known operations like addition, scalar multiplication and standard multiplication, we distinguish some other operations: Kronecker product, Hadamard- Schur product and Khatri-Rao product. We will introduce the vec operation and matrix derivatives and state some of theirs properties.

Key words: Kronecker product, the vec operation, Hadamard-Schur product, Khatri-Rao product, matrix derivatives, eigenvalues

Sadržaj

1 Uvod

Tema ovog rada su operacije s matricama. Matrica tipa (m, n) s koeficijentima iz polja F je preslikavanje $A : \{1, 2, \dots, m\} \times \{1, 2, \dots, n\} \to F$. Obično se djelovanje takve funkcije zapisuje tablično, u m redaka i n stupaca. Za danu matricu A koeficijent u i-tom retku i j-tom stupcu uobičajno označavamo s $(A)_{ij}$ ili jednostavnije a_{ij}, Dakle, takvu matricu zapisujemo u obliku:

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

Skup svih matrica s m redaka i n stupaca s koeficijentima iz polja F označavamo s $M_{mn}(F)$. Ako matrica ima jednak broja stupaca i redaka, to jest ako je $m = n$, takvu matricu zovemo kvadratnom, a skup svih kvadratnih matrica reda m označavamo sa $M_n(F)$. U ovom radu navest ćemo osnovne operacije s matricama: zbrajanje, množenje matrice skalarom i klasično množenje te njihova svojstva . Uz klasično množenje pokazat ćemo neka složenija množenja i njihovu primjenu u računanju. Produkti koje ćemo promatrati su:

- Kroneckerov produkt
- Hadamard-Schurov produkt
- Khatri-Raov produkt

Također ćemo pokazati vezu između produkata i svojstvenih vrijednosti matrica. Osim produkata, koristan nam je i postupak vektorizacije koji olakšava rješavanje sustava matričnih jednadžbi te nam olakšava i razumijevanje Kroneckerovog produkta. Za kraj, upoznat ćemo se i sa pojmom matričnih derivacija. Pokazat ćemo svojstva tih derivacija, njihovu primjenu te veze matričnih derivacija sa ostalim produktima.
Svi teoremi i dokazi u ovom radu pisani su prema [5]

2 Standardne operacije s matricama

Dokaze teorema u ovom poglavlju možemo pronaći u [1]
Osnovna matematička operacija s kojom se prvom susrećemo je zbrajanje. Prema tome, krenimo s opisom takve operacije nad matricama. Ova operacija zahtjeva da imamo matrice istog tipa.

Definicija 2.1. *Neka su $A = [a_{ij}]$ i $B = [b_{ij}]$ matrice s m redaka i n stupaca. Zbroj matrica A i B definiramo kao:*

$$\begin{bmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} + \begin{bmatrix} b_{11} & \dots & b_{1n} \\ b_{21} & \dots & b_{2n} \\ \vdots & & \vdots \\ b_{m1} & \dots & b_{mn} \end{bmatrix} = \begin{bmatrix} a_{11}+b_{11} & \dots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & \dots & a_{2n}+b_{2n} \\ \vdots & & \vdots \\ a_{m1}+b_{m1} & \dots & a_{mn}+b_{mn} \end{bmatrix}$$

Navedimo osnovna svojstva zbrajanja matrica.

Teorem 2.2. *Neka su A,B i C matrice istog tipa. Za zbrajanje matrica vrijedi sljedeće:*

- *komutativnost:* $A + B = B + A$;
- *asocijativnost:* $A + (B + C) = (A + B) + C$;
- *postojanje neutralnog elementa (nulmatrica):* $A + 0 = 0 + A = A$;
- *postojanje suprotnog elementa:* $A + (-A) = -A + A = 0$.

Definicija 2.3. *Umnožak matrice $A = [a_{ij}]$ skalarom α je matrica oblika:*

$$\alpha \begin{bmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} = \begin{bmatrix} \alpha a_{11} & \dots & \alpha a_{1n} \\ \alpha a_{21} & \dots & \alpha a_{2n} \\ \vdots & & \vdots \\ \alpha a_{m1} & \dots & \alpha a_{mn} \end{bmatrix}$$

Zbrajanje i množenje su dvije binarne operacije, no množenje matrica je operacija u kojoj ni faktori ni umnožak nisu matrice istog tipa. Matrice koje možemo množiti su ulančane matrice. Matrice A i B su ulančane ako je broj stupaca matrice A jednak broju redaka matrice B.

Definicija 2.4. *Neka su $A = [a_{ij}] \in M_{mn}$ i $B = [b_{ij}] \in M_{ns}$ ulančane matrice. Tada je produkt AB definiran kao:*

$$AB = [c_{ij}] \in M_{ms}$$

pri čemu je

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}, \forall i = 1, 2, \dots, m, \forall j = 1, 2, \dots, s.$$

Koeficijent c_{ij} je umnožak i-tog retka od A i j-tog stupca od B pa je jasno zašto matrice moraju biti ulančane. Primjetimo kako komutativnost množenja matrica općenito ne vrijedi. Navedimo u teoremu osnovna svojstva matričnog množenja.

Teorem 2.5. *Ukoliko su navedeni produkti definirani, vrijedi sljedeće:*

1. $A(B+C) = AB + AC$;
2. $(A+B)C = AC + BC$;
3. $(\alpha A)B = A(\alpha B) = \alpha(AB)$;
4. $(AB)C = A(BC)$;
5. $IA = A, AI = A$.

3 Kroneckerov produkt

Kroneckerov produkt definira se za dvije matrice proizvoljnog tipa.

Definicija 3.1. *Neka je $A = [a_{ij}]$ matrica tipa $m \times n$ te $B = [b_{ij}]$ matrica tipa $p \times q$. Kroneckerov produkt matrica A i B označavamo i definiramo na sljedeći način:*

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

Općenito je Kroneckerov produkt matrica tipa $mp \times nq$. Primjetimo kako jednakost $A \otimes B = B \otimes A$ općenito ne vrijedi baš kao i kod standardnog produkta matrica. U sljedećem teoremu navest ćemo neka svojstva Kroneckerovog produkta.

Teorem 3.2.
1. *Kroneckerov produkt je asocijativna operacija. Odnosno, za matrice A, B, C vrijedi: $(A \otimes B) \otimes C = A \otimes (B \otimes C)$*
2. *Neka su A, B, C matrice takve da su B i C istog tipa. Tada vrijedi:* $A \otimes (B + C) = A \otimes B + A \otimes C$.
3. *Ako je α skalar i A bilo koja matrica, tada je: $\alpha \otimes A = \alpha A = A \otimes \alpha$ (α smatramo matricom reda 1).*
4. *Ako su A, B, C i D takve matrice da su A i C te B i D ulančane, onda je:* $(A \otimes B)(C \otimes D) = AC \otimes BD$.
5. *Ako su A i B bilo koje dvije matrice. Vrijedi: $(A \otimes B)^T = A^T \otimes B^T$.*
6. *Ako su A i B bilo koje dvije matrice. Vrijedi: $(A \otimes B)^* = A^* \otimes B^*$.*
7. *Neka su A i B kvadratne matrice ne nužno istog reda. Vrijedi: $tr(A \otimes B) = [tr(A)][tr(B)]$.*
8. *Neka su A i B regularne matrice ne nužno istog reda. Vrijedi $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$.*

Dokaze ovih svojstava možemo pronaći u [5]. Navedimo još neka zanimljiva svojstva Kroneckerovog produkta.

Teorem 3.3. *1. Ako su A i B nenegativno definitne matrice, tada je i $A \otimes B$ također nenegativno definitna.*

2. Ako je A matrica reda m te B matrica reda n tada je:

$$det(A \otimes B) = [det(A)]^n [det(B)]^m.$$

3. Neka su A i B dvije ne nužno kvadratne matrice, tada je:

$$r(A \otimes B) = [r(A)][r(B)].$$

3.1 Svojstvene vrijednosti i svojstveni vektori Kroneckerovog produkta

Također možemo promatrati vezu svojstvenih vrijednosti Kroneckerovog produkta i matrica koje čine produkt.

Teorem 3.4. *Neka su A i B kvadratne matrice sa svojstvenim vrijednostima $\lambda_1, \dots, \lambda_m$ odnosno $\mu_1, \dots, \mu_n$. Tada su $\lambda_i \mu_j$, za $i = 1, \dots, m$ te $j = 1, \dots, n$, svojstvene vrijednosti pripadnog Kroneckerovog produkta.*

Dokaz. Ako je λ svojstvena vrijednost od A te x odgovarajući svojstveni vektor i μ svojstvena vrijednost od B te y odgovarajući svojstveni vektor, lako se pokaže da je $\lambda\mu$ svojstvena vrijednost od $A \otimes B$ sa svojstvenim vektorom $x \otimes y$. Tvrdnja je sadržana u sljedećoj jednakosti:

$$(A \otimes B)(x \otimes y) = (Ax) \otimes (By) = (\lambda x) \otimes (\mu y) = \lambda\mu(x \otimes y).$$

Dokažimo tvrdnju pozivajući se na teorem o dekompoziciji matrica: Postoje unitarne matrice U reda m te V reda n takve da je $UAU^* = \Delta_1$ i $VBV^* = \Delta_2$, gdje su Δ_1 i Δ_2 gornjetrokutaste a elementi na njihovim dijagonalama su svojstvene vrijednosti od A odnosno B. Sada promotrimo jednakost:

$$(U \otimes V)(A \otimes B)(U \otimes V)^* = (U \otimes V)(A \otimes B)(U^* \otimes V^*) = (UAU^*) \otimes (VBV^*) = \Delta_1 \otimes \Delta_2$$

$U \otimes V$ je unitarna, a $\Delta_1 \otimes \Delta_2$ gornjetrokutasta matrica. Dijagonalni elementi od $\Delta_1 \otimes \Delta_2$ sadrže svojstvene vrijednosti od $A \otimes B$ čime je tvrdnja dokazana.

□

Promotrimo i što se događa sa svojstvenim vektorima.

Teorem 3.5. *Neka je x svojstveni vektor koji odgovara nekoj svojstvenoj vrijednosti matrice A te y svojstveni vektor koji odgovara nekoj svojstvenoj vrijednosti matrice B. Tada je $x \otimes y$ svojstveni vektor matrice $A \otimes B$*

Dokaz ovog teorema uključen je u dokaz prethodnog teorema.

3.2 Primjena Kroneckerovog produkta

Neka je dana linearna jednadžba:

$$AX = B \tag{3.1}$$

gdje je A matrica tipa $m \times n$ i B matrica tipa $m \times p$ te X nepoznata matrica tipa $n \times p$. Za primjer uzmimo $A, B, X \in M_2$:

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \tag{3.2}$$

Postoje dva načina kako možemo zapisati jednadžbu (3.2). Jedan je:

$$\begin{bmatrix} a_{11} & a_{12} & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 \\ 0 & 0 & a_{11} & a_{12} \\ 0 & 0 & a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_3 \\ x_2 \\ x_4 \end{bmatrix} = \begin{bmatrix} b_{11} \\ b_{12} \\ b_{21} \\ b_{22} \end{bmatrix} \tag{3.3}$$

Neka je $x^T = (x_1, x_2, x_3, x_4)$ i $b^T = (b_{11}, b_{21}, b_{12}, b_{22})$. Sustav (3.3) možemo zapisati u sljedećem obliku:

$$(I_2 \otimes A)x = b. \tag{3.4}$$

Drugi način zapisivanja jednadžbe (3.2) je:

$$\begin{bmatrix} a_{11} & 0 & a_{12} & 0 \\ 0 & a_{11} & 0 & a_{12} \\ a_{21} & 0 & a_{22} & 0 \\ 0 & a_{21} & 0 & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} b_{11} \\ b_{12} \\ b_{21} \\ b_{22} \end{bmatrix} \tag{3.5}$$

Neka je $y^T = (x_1, x_2, x_3, x_4)$ i $c^T = (b_{11}, b_{21}, b_{12}, b_{22})$. Sustav 3.5 možemo zapisati u sljedećem obliku:

$$(A \otimes I_2)y = c. \tag{3.6}$$

Vratimo se na općeniti slučaj i jednadžbu (3.1), x je vektor stupac tipa $np \times 1$ dobiven iz matrice X slaganjem redova jedan po jedan te b također vektor stupac tipa $mp \times 1$ dobiven iz B na isti način kao i x. Sada sustav (3.1) zapisujemo

$$(A \otimes I_p)x = b \tag{3.7}$$

Pretpostavimo li da je $m = n$, odnosno da je A kvadratna matrica. Tada sustav (3.7) ima jedinstveno rješenje ako je A regularna matrica.

4 Vektorizacija

Pomoću neke matrice možemo napravit vektor stupac koji sadrži sve elemente matrice A. Postupak možemo provesti na dva načina. Prvi način je da složimo sve elemente redaka matrice A po redu, počevši od prvog. Drugi način je slaganje stupaca. Ovaj način je više korišten u praksi pa ćemo ga i opisati.

Definicija 4.1. *Neka je $A = [a_{ij}]$ matrica tipa $m \times n$ i neka je a_i i-ti stupac matrice A, $i = 1, 2, \dots, n$. Tada sa $vec\, A$ označavamo vektor stupac tipa $mn \times 1$ definiran kao:*

$$vec\, A = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}.$$

Sada sustav oblika $AX = B$ gdje su: $A \in M_{mn}$, $X \in M_{np}$ te $B \in M_{mp}$ s možemo zapisati na sljedeći način:

$$(I_p \otimes A) vec\, X = vec\, B. \tag{4.1}$$

4.1 Vektorizacija

Navedimo sad nekoliko svojstva vektorizacije i napomena vezanih uz postupak.

Napomena 4.2. *Vektorizaciju možemo provesti za sve matrice a ne nužno kvadratne.*

Napomena 4.3. *Ako za neke dvije matrice vrijedi $vec\, A = vec\, B$ to ne znači da je $A = B$. Pogledajmo sljedeći dvije matrice:*

$$A = \begin{bmatrix} 2 & 7 \\ 4 & 12 \\ 8 & 5 \end{bmatrix} \quad B = \begin{bmatrix} 2 & 8 & 12 \\ 4 & 7 & 5 \end{bmatrix}$$

Njihova vektorizacija je istog oblika:

$$vec\, A = vec\, B = \begin{bmatrix} 2 \\ 4 \\ 8 \\ 7 \\ 12 \\ 5 \end{bmatrix}$$

Teorem 4.4. *1. Ako su x i y dva vektor stupca, ne nužno istog tipa, onda je*

$$vec\, xy^T = y \otimes x.$$

2. Ako su A i B dvije matrice istog reda, onda je

$$tr(A^T B) = [vec\, A]^T vec\, B).$$

3. Ako su A i B dvije matrice istog reda, onda je

$$vec\, A + B = vec\, A + vec\, B.$$

4. Ako su A, B, C tri matrice takve da produkt ABC ima smisla, onda je

$$vec\, ABC = (C^T \otimes A) vec\, B.$$

5. Ako je $A \in M_{mn}$ i $B \in M_{np}$, onda je

$$vec\, AB = (B^T \otimes I_m) vec\, A = (I_p \otimes A) vec\, B.$$

Dokaz. Tvrdnje 1.,2.,3. lako se provjere pa pokažimo tvrdnju 4.
Neka su $A = [a_{ij}] \in M_{mn}$, $B = [b_{ij}] \in M_{np}$ i $C = [c_{ij}] \in M_{pq}$. Neka su $b_1, b_2, \ldots, b_p$ stupci matrice B te $e_1, e_2, \ldots, e_p$ stupci jedinične matrice I_p. Matricu B možemo zapisati na sljedeći način.

$$B = BI_p = (b_1, b_2, \ldots, b_p)(e_1, e_2, \ldots, e_p)^T = \sum_{j=1}^{p} b_j e_j^T.$$

Imamo sljedeće:

$$\begin{aligned} vec(ABC) &= vec(A(\sum_{j=1}^{p} b_j e_j^T)C) = \sum_{j=1}^{p} vec(A b_j e_j^T C) \\ &= \sum_{j=1}^{p} vec((Ab_j)C^T e_j)^T) = \sum_{j=1}^{p} (C^T e_j) \otimes (Ab_j) \\ &= \sum_{j=1}^{p} C^T \otimes A(e_j \otimes b_j) = (C^T \otimes A) \sum_{j=1}^{p} e_j \otimes b_j \\ &= (C^T \otimes A) \sum_{j=1}^{p} vec(b_j e_j^T) = (C^T \otimes A) vec(\sum_{j=1}^{p} b_j e_j^T) \\ &= (C^T \otimes A) vec B \end{aligned}$$

Tvrdnja 5. slijedi iz 4. ako primjetimo da AB možemo zapisati kao $AB = I_m AB = ABI_p$.

□

5 Hadamard-Schurov produkt

Definicija 5.1. *Neka su $A = [a_{ij}]$ i $B = [b_{ij}]$ dvije matrice istog tipa. Hadamard-Schurov produkt matrica A i B je matrica istog tipa čiji je (i, j)-ti element oblika $a_{ij} b_{ij}$.*

Hadamard-Schurov produkt, radi jednostavnosti, pišemo HS produkt i označavamo ga $A \cdot B$. Često se u literaturi ovakvo množenje naziva entry-wise produkt. Navedimo osnovna svojstva ovakvog množenja. Kako je HS produkt definiran samo za matrice istog reda, tvrdnje iz sljedećeg teorema vrijede za takve matrice.

5.1 Svojstva Hadamar-Schurovog produkta

Teorem 5.2. *1. HS množenje je komutativno:*

$$A \cdot B = B \cdot A.$$

2. HS množenje je asocijativno:

$$A \cdot (B \cdot C) = (A \cdot B) \cdot C.$$

3. $(A + B) \cdot (C + D) = A \cdot C + A \cdot D + B \cdot C + B \cdot D.$

4. $A \cdot 0 = 0$ *gdje je* 0 *nul-matrica istog tipa.*

5. $A \cdot J = A$ *gdje je* J *matrica istog tipa čiji su svi elementi 1.*

6. Ako je $m = n, A = (a_{ij})$, I_m *jedinična matrica:*

$$A \cdot I_m = diag(a_{11}, a_{22}, \ldots, a_{mm}).$$

7. $(A \cdot B)^T = A^T \cdot B^T.$

8. Ako je $A = (a_{ij})$ *takva da je* $a_{ij} \neq 0$ *i* $B = (1/a_{ij})$*, onda je* $A \cdot B = J$ *(*B *je tada HS multiplikativni inverz od* A*).*

Teorem 5.3 (Schurov teorem)**.** *Ako su* A *i* B *dvije nenegativno definitne matrice istog reda, onda je* $A \cdot B$ *također nenegativno definitna. Ako su* A *i* B *pozitivno definitne, tada je takva i* $A \cdot B$

Schurov teorem možemo dokazati na nekoliko načina. Jedan od njih dan je u [5]. Također, vrijedi i obrat teorema: Ako je A nenegativno definitna matrica, možemo ju zapisati kao HS produkt dvije nenegativno definitne matrice. Isto vrijedi i za pozitivno definitne matrice.
Idući teorem iskazuje vezu između ranga matrica i HS produkta matrica.

Teorem 5.4. *Neka su* A *i* B *dvije matrice tipa* $m \times n$*. Tada je:*

$$r(A \cdot B) \leq [r(A)][r(B)].$$

Navedimo jedan jednostavan primjer kojim ćemo ilustrirati prethodni teorem.

Primjer 5.5. $A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} B = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} A \cdot B = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

Primjetimo da vrijedi: $r(A) = r(B) = 2$ *te* $r(A \cdot B) = 4.$

5.2 HS produkt i definitnost matrica

Teorem 5.6. *Ako je A pozitivno definitna matrica te B nenegativno definitna matrica sa r nenul elemenata na dijagonali, tada je:*

$$r(A \cdot B) = r.$$

Ovaj teorem omogućava nam još neke zanimljive rezultate. Vrijede sljedeći rezultati:

1. Ako je A pozitivno definitna i B nenegativno definitna matrica sa dijagonalnim elementima različitim od nule, tada je $A \cdot B$ nenegativno definitna.
2. Ako je A pozitivno definitna matrica i B negativno definitna matrica, onda je $A \cdot B$ negativno definitna matrica.
3. Moguće je da matrica $A \cdot B$ bude punog ranga iako to nisu matrice A i B.

Navedimo primjer za posljednju tvrdnju.

Primjer 5.7. $A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ $B = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$ $A \cdot B = \begin{bmatrix} 4 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$ *Matrice A i B su ranga 2, ali matrica $A \cdot B$ je ranga 3.*

Teorem 5.8 (Fejerov teorem)**.** *Neka je $A = (a_{ij})$ matrica reda n. Matrica A je nenegativno definitna ako i samo ako je $tr(A \cdot B) \leq 0$ za sve nenegativno definitne matrice B reda n.*

5.3 HS i svojstvene vrijednosti

Pogledajmo sada vezu HS produkta i svojstvenih vrijednosti. Za bilo koju hermitsku matricu A reda m neka su $\lambda_1(A) \geq \lambda_2(A) \geq \cdots \geq \lambda_m(A)$ svojstvene vrijednosti matrice A poredane od najveće prema najmanjoj.

Teorem 5.9. *Neka su A i B dvije nenegativno definitne matrice reda m. Neka su b_1 i b_m najveći i najmanji dijagonalni element matrice B. Tada je:*

$$b_m \lambda_m \leq \lambda_j(A \cdot B) \leq b_1 \lambda_1, j = 1, 2, \ldots, m.$$

Uz ovaj teorem vežemo i sličan korolar.

Korolar 5.10. *Ako su A i B dvije nenegativno definitne matrice reda m, onda za sve j vrijedi:*

$$\lambda_m(A)\lambda_m(B) \leq \lambda_j(A \cdot B) \leq \lambda_1(A)\lambda_1(B)$$

6 Khatri-Raov produkt

Definicija 6.1. *Neka je A matrica tipa $p \times n$ i B matrica tipa $m \times n$. Neka je a_i i-ti stupac matrice A te $b_i i$ -ti stupac matrice B, $i = 1, 2, \ldots, n$. Khatri-Raov produkt $A \odot B$ je blok-matrica tipa $pm \times n$ oblika:*

$$A \odot B = (a_1 \otimes b_1 | a_2 \otimes b_2 | \ldots | a_n \otimes b_n |) \tag{6.1}$$

6.1 Veza Khatri-Raovog produkta sa Kroneckerovim i HS produktom

Teorem 6.2. *Neka su: A, B, C, D četiri matrice tipova: $p \times n, m \times n, m \times p, n \times m$. Tada je:*

$$(C \otimes D)(A \odot B) = (CA) \odot (DB).$$

Dokaz. Neka je α_i i-ti stupac matrice A te β_i i-ti stupac matrice B, $i = 1, 2, \dots, n$. Tada je i-ti stupac matrice CA $C\alpha_i$, a matrice DB $D\beta_i$. Prema tome, i-ti stupac matrice $(CA) \odot (DB)$ je $C\alpha_i \otimes D\beta_i = (C \otimes D)(\alpha_i \otimes \beta_i)$ što je zapravo i-ti stupac matrice $(C \otimes D)(A \odot B)$. $\square$

Teorem 6.3. *Neka su A i B dvije nenegativno definitne matrice reda n. Neka su $A = \Gamma^T\Gamma$ i $B = \Omega^T\Omega$ Gramovi prikazi matrica A i B, za neke Γ reda $r \times n$ i Ω reda $s \times n$. Tada su HS umnožak i Khatri-Raov umnožak povezani na sljedeći način:*

$$A \cdot B = (\Gamma \odot \Omega)^T(\Gamma \odot \Omega).$$

Dokaz. Neka su $\alpha_1, \alpha_2, \dots, \alpha_n$ stupci matrice Γ i $\beta_1, \beta_2, \dots, \beta_n$ stupci matrice Ω. Ako su $A = (a_{ij})$ i $B = (b_{ij})$ onda vrijedi: $a_{ij} = \alpha_i^T\alpha_j$ i $b_{ij} = \beta i^T\beta_j$ $\forall i, j$. (i, j)-ti element matrice $(\Gamma \odot \Omega)^T(\Gamma \odot \Omega)$ je dan kao: $(\alpha_i \otimes \beta_i)^T(\alpha_j \otimes \beta_j) = \alpha_i^T\alpha_j \otimes \beta_i^T\beta_j = (\alpha_i^T\alpha_j)(\beta_i^T\beta_j) = a_{ij}b_{ij}$ što je zapravo (i, j)-ti element HS produkta. $\square$

Korolar 6.4. *Uz prepostavke teorema 6.3 donosimo sljedeći zaključak. Ako matrica dobivena kao HS produkt $A \cdot B$ nije punog ranga, onda postoji nenul dijagonalna matrica Δ takva da je $\Gamma\Delta\Omega^T = 0$.*

Dokaz. Prema teoremu 6.3 možemo pisati: $A \cdot B = (\Gamma \odot \Omega)^T(\Gamma \odot \Omega)$. Kako matrica $A \cdot B$ nije punog ranga, postoji nenul vektor x takav da je $(A \cdot B)x = 0$. Iz ovoga slijedi: $(\Gamma \odot \Omega) = 0$. Neka je $\Gamma = (\gamma_{ij}), \Omega = (\omega_{ij})$ i $x^T = (x_1, x_2, \dots, x_n)$. Neka je $\Delta = diag(x_1, x_2, \dots, x_n)$ Sada iz $(\Gamma \odot \Omega) = 0$ slijedi:

$$x_1\gamma_{i1}\omega_{j1} + x_2\gamma_{i2}\omega_{j2} + \cdots + x_n\gamma_{in}\omega_{jn} = 0$$

za $1 \leq i \leq r$ i $1 \leq j \leq s$. Ovo je ekvivalentno jednakosti $\Gamma\Delta\Omega^T = 0$. $\square$

7 Matrične derivacije

Neka je f funkcija realnih varijabli, $x_{ij}, i = 1, 2, \dots, m$ i $j = 1, 2, \dots, n$. Pretpostavimo da su ove varijable elementi matrice $X = (x_{ij})$ tipa $m \times n$. Pretpostavimo da je f parcijalno derivabilna po svakoj varijabli.

Definicija 7.1. *Matrična derivacija $\partial f/\partial X$ od f s obzirom na X je matrica tipa $m \times n$ dana kao:*

$$\frac{\partial f}{\partial X} = \left(\frac{\partial f}{\partial x_{ij}}\right).$$

Napomena 7.2. *Ako je n=1, X je vektor stupac sa elementima: $x_1, x_2, \dots, x_m$. Odgovarajuća derivacija $\partial f/\partial x$ se naziva vektor derivacija funkcije f s obzirom na x.*

Općenito, pretpostavimo da je $F = (f_{ij})$ matrična funkcija matrične varijable X. To znači da je svaki element f_{ij} funkcija realnih varijabli iz X. Definirajmo sada derivaciju takve matrice.

Definicija 7.3. *Neka je F matrica tipa $p \times q$ i X matrica tipa $m \times n$. Pretpostavimo da svaki element od F ima parcijalne derivacije svih varijabli iz X. Matrična derivacija $\partial F/\partial X$ s obzirom na X je definirana kao matrica*

$$\frac{\partial F}{\partial X} = \left(\frac{\partial f_{ij}}{\partial X}\right). \tag{7.1}$$

Napomena 7.4. *Matrica (7.1) je tipa $pm \times qn$ sastavljena od pq blokova. Svaki blok matrice je tipa $m \times n$.*

Pretpostavimo sada da je matrična funkcija identiteta, to jest $F(X) = X$, ili $f_{ij}(X) = x_{ij}$ za sve i, j. Kada bi željeli koristiti matricu parcijalnih derivacija za razvijanje Jacobijeve matrice transformacija, ovakav način derivacije nije dobro definiran.
Neka je npr. X tipa 2×3 i $F(X) = X$. Tada je $\partial f/\partial X = ((vecI_2) \otimes I_3)^T$, što je tipa 4×9. Jasno je kako je matrica $\partial f/\partial X$ ranga 1. Jakobijan matrica transformacije $F(X) = X$ je I_0. Derivacija $\partial F/\partial X$ nije ni približna Jakobijanu. Kako bi poboljšali standardnu definiciju 7.1 definirajmo ju na sljedeći način:

$$\frac{*\partial F}{\partial X} = \frac{\partial vecF}{\partial (vecX)^T}. \tag{7.2}$$

Ovako definirana matrica je tipa $pq \times mn$. Opišimo postupak dobivanja nove matrice parcijalnih derivacija:
Potrebno je složiti varijable u X stupac po stupac u jedan vektor, dobiveni vektor transponirati. Zatim je potrebno posložiti elemente iz F stupac po stupac u jedan vektor, a zatim parcijalno derivirati svaki element s obzirom na $(vecX)^T$.
Primjetimo kako je tip od $vecF$ $pq \times 1$ i $(vecX)^T$ tipa $1 \times mn$. Prema tome, tip matrice 7.2 je $pq \times mn$.

Primjer 7.5. *Matrica F tipa 2×4 i X tipa 3×2 nam daju sljedeću matricu derivacija:*

$$\frac{^*\partial F}{\partial X} = \begin{bmatrix} \frac{\partial f_{11}}{\partial x_{11}} & \frac{\partial f_{11}}{\partial x_{21}} & \frac{\partial f_{11}}{\partial x_{31}} & \frac{\partial f_{11}}{\partial x_{12}} & \frac{\partial f_{11}}{\partial x_{22}} & \frac{\partial f_{11}}{\partial x_{32}} \\ \frac{\partial f_{21}}{\partial x_{11}} & \frac{\partial f_{21}}{\partial x_{21}} & \frac{\partial f_{21}}{\partial x_{31}} & \frac{\partial f_{21}}{\partial x_{12}} & \frac{\partial f_{21}}{\partial x_{22}} & \frac{\partial f_{21}}{\partial x_{32}} \\ \frac{\partial f_{12}}{\partial x_{11}} & \frac{\partial f_{12}}{\partial x_{21}} & \frac{\partial f_{12}}{\partial x_{31}} & \frac{\partial f_{12}}{\partial x_{12}} & \frac{\partial f_{12}}{\partial x_{22}} & \frac{\partial f_{12}}{\partial x_{32}} \\ \frac{\partial f_{22}}{\partial x_{11}} & \frac{\partial f_{22}}{\partial x_{21}} & \frac{\partial f_{22}}{\partial x_{31}} & \frac{\partial f_{22}}{\partial x_{12}} & \frac{\partial f_{22}}{\partial x_{22}} & \frac{\partial f_{22}}{\partial x_{32}} \\ \frac{\partial f_{13}}{\partial x_{11}} & \frac{\partial f_{13}}{\partial x_{21}} & \frac{\partial f_{13}}{\partial x_{31}} & \frac{\partial f_{13}}{\partial x_{12}} & \frac{\partial f_{13}}{\partial x_{22}} & \frac{\partial f_{13}}{\partial x_{32}} \\ \frac{\partial f_{23}}{\partial x_{11}} & \frac{\partial f_{23}}{\partial x_{21}} & \frac{\partial f_{23}}{\partial x_{31}} & \frac{\partial f_{23}}{\partial x_{12}} & \frac{\partial f_{23}}{\partial x_{22}} & \frac{\partial f_{23}}{\partial x_{32}} \\ \frac{\partial f_{14}}{\partial x_{11}} & \frac{\partial f_{14}}{\partial x_{21}} & \frac{\partial f_{14}}{\partial x_{31}} & \frac{\partial f_{14}}{\partial x_{12}} & \frac{\partial f_{14}}{\partial x_{22}} & \frac{\partial f_{14}}{\partial x_{32}} \\ \frac{\partial f_{24}}{\partial x_{11}} & \frac{\partial f_{24}}{\partial x_{21}} & \frac{\partial f_{24}}{\partial x_{31}} & \frac{\partial f_{24}}{\partial x_{12}} & \frac{\partial f_{24}}{\partial x_{22}} & \frac{\partial f_{24}}{\partial x_{32}} \end{bmatrix}$$

Primjetimo kako su elementi matrice $\frac{^*\partial F}{\partial X}$ zapravo elementi matrice $\frac{\partial F}{\partial X}$.

7.1 Svojstva matričnih derivacija

Navedimo prvo novu definiciju koja nam je korisna u dobivanju nekih formula vezanih za deriviranje matrica.

Definicija 7.6. *Neka je f skalarna funkcija matrične varijabe X tipa $m \times n$. Neka je Y konstantna matrica tipa $m \times n$. Pretpostavimo da je f diferencijabilna, to jest postoje sve parcijalne derivacije i one su neprekidne. Tada postoji usmjerena derivacija funkcije f u smjeru Y i definirana je kao:*

$$\lim_{t \to \infty} \frac{f(X + tY) - f(X)}{t} = tr(Y^T \frac{\partial f}{\partial X}). \tag{7.3}$$

Teorem 7.7. *Neka su f i g dvije diferencijabilne funkcije realne matrične varijable X. Tada vrijede sljedeće tvrdnje:*

1. $\frac{\partial fg}{\partial X} = f\frac{\partial g}{\partial X} + g\frac{\partial f}{\partial X}$.
2. $\frac{\partial (f/g)}{\partial X} = \frac{1}{g}\frac{\partial f}{\partial X} - \frac{f}{g^2}\frac{\partial g}{\partial X}, g \neq 0$.

Fokusirajmo se sada na derivacije vektora. Sve funkcije f su funkcije realne varijable definirane na vektorskom prostoru $\mathbb{R}^m$.

Teorem 7.8.

1. *Ako je $f(x) = a^T x$ za neki konstantni vektor $a \in \mathbb{R}^m$, onda je $\frac{\partial f}{\partial x} = a$.*
2. *Ako je $f(x) = x^T x$, onda je $\frac{\partial f}{\partial x} = 2x$.*
3. *Ako je $f(x) = x^T Ax$ za neku konstantnu matricu A reda m sa realnim elementima, onda je $\frac{\partial f}{\partial x} = (A + A^T)x$.*
4. *Ako je $f(x) = x^T Ax$ za neku konstantnu simetričnu matricu A reda $m \times m$ sa realnim elementima, onda je $\frac{\partial f}{\partial x} = 2Ax$.*

7.2 Primjena matričnih derivacija

Neka je A simetrična matrica i B pozitivno definitna matrica sa realnim elementima. Neka je $f(x) = x^T Ax$ i $g(x) = x^T Bx$, $x \in \mathbb{R}^m$. Želimo odrediti stacionarne točke funkcije $f(x)/g(x), x \in \mathbb{R}^m \backslash \{0\}$. Koristeći teorem 7.8 imamo sljedeću jednakost:

$$\frac{\partial (f/g)}{\partial x} = \frac{2}{x^T Bx} Ax - \frac{2x^T Ax}{(x^T Bx)^2} Bx = 0.$$

Što vodi do jednadžbe:

$$(A - \lambda B)x = 0$$

ako uzmemo u obzir da je $\lambda = x^T Ax / x^T Bx$. Nenul rješenje jednadžbe će postojati ako je determinanta $|A - \lambda B| = 0$.

Pogledajmo sada vezu s matričnim derivacijama. Funkcija f je realna funkcija matrične varijable X reda m. Neka je $M_m(ns)$ skup svih nesingularnih matrica reda m sa realnim elementima kojeg označavamo. Ovaj skup je otvoren podskup skupa svih matrica reda m. Za domenu ove funkcije promatrat ćemo skup $\{X \in M_m(ns) : |X| > 0\}$ na kojem je dobro definirana diferencijabilnost funkcije determinante, $|X|$ matrice X.

Teorem 7.9.

1. *Ako je $f(X) = |X|$, $X \in M_m(ns)$, onda je $\partial f / \partial X = |X|(X^{-1})^T$.*
2. *Ako je $f(X) = log|X|, |X| > 0$, onda je $\frac{\partial f}{\partial X} = (X^{-1})^T$.*

3. *Ako je* $f(X) = |X|^r, |X| > 0$, *onda je* $\partial f/\partial X = r|X|^r|(X^{-1})^T|$.

Obratimo pažnju na simetrične matrice. Prostor $M_m(s)$ svih simetričnih matrica reda m je $m(m+1)/2$-dimenzionalan vektorski prostor. Pogledajmo podskup $M_m(s, ns)$ svih nesingularnih matrica u $M_m(s)$. Na ovom otvorenom skupu, determinanta je definirana.

Primjer 7.10. *Neka je* $X \in M_2(s, ns)$

$$X = \begin{bmatrix} x_{11} & x_{12} \\ x_{12} & x_{22} \end{bmatrix}$$

sa determinantom $x_{11}x_{22} - x_{12}^2 \neq 0$.

$$\frac{\partial |X|}{\partial X} = \begin{bmatrix} \frac{\partial |X|}{\partial x_{11}} & \frac{\partial |X|}{\partial x_{12}} \\ \frac{\partial |X|}{\partial x_{12}} & \frac{\partial |X|}{\partial x_{22}} \end{bmatrix} = \begin{bmatrix} x_{22} & -2x_{12} \\ -2x_{12} & x_{11} \end{bmatrix} = 2\begin{bmatrix} x_{22} & -x_{12} \\ -x_{12} & x_{11} \end{bmatrix} - \begin{bmatrix} x_{22} & 0 \\ 0 & x_{11} \end{bmatrix} =$$
$$|X|[2X^{-1} - diag(X^{-1})].$$

Navedimo još neke korisne rezultate:

Teorem 7.11. 1. *Ako je* $f(X) = |X|, X \in M_m(s, ns)$, *onda je*

$$\frac{\partial f}{\partial X} = |X|[2X^{-1} - diag(X^{-1})].$$

2. *Ako je* $f(X) = log|X|, X \in M_m(s, ns), |X| > 0$, *onda je*

$$\frac{\partial f}{\partial X} = [2X^{-1} - diag(X^{-1})].$$

3. *Ako je* $f(X) = |X|^r, X \in M_m(s, ns), |X| > 0$, *onda je*

$$\frac{\partial f}{\partial X} = r|X|^r[2X^{-1} - diag(X^{-1})].$$

7.3 Veza matričnih derivacija s ostalim produktima

Na početku ovog poglavlja, parcijalne matrične derivacije zapisali smo kao:

$$\frac{{}^*\partial F}{\partial X} = \frac{\partial vecF}{\partial (vecX)^T}.$$

Iako su elementi matrice ${}^*\partial F/\partial X$ zapravo elementi matrice $\partial F/\partial X$ u drugačijem poretku, ponekad je korisno računati s ${}^*\partial F/\partial X$. To i radimo u sljedećem teoremu.

Teorem 7.12. 1. *Neka je* $F(X) = AX$ *za neku* A *konstantnu matricu reda* $p \times m$ *i* X *tipa* $m \times n$. *Tada je*

$$\frac{{}^*\partial F}{\partial X} = I_n \otimes A.$$

2. *Neka je* $F(X) = XB$ *za neku* B *konstantnu matricu tipa* $n \times q$ *i* X *tipa* $m \times n$. *Tada je*

$$\frac{{}^*\partial F}{\partial X} = B^T \otimes I_m.$$

3. *Neka je* $F(X) = AXB$ *za neke* A *i* B *konstantne matrice tipa* $p \times m$ *i* $n \times q$ *i* X *tipa* $m \times n$. *Tada je*

$$\frac{{}^*\partial F}{\partial X} = B^T \otimes A.$$

4. *Neka je* $F(X) = AX^TB$ *za neke* A *i* B *konstantne matrice tipa* $p \times n$ *i* $m \times q$ *i* X *tipa* $m \times n$. *Tada je*

$$\frac{{}^*\partial F}{\partial X} = (A \otimes B^T)P,$$

gdje je P *matrica permutacija koja vektor vec(X) prevodi u* $vec(X^T)$, *to jest* $vec(X^T) = Pvec(X)$.

5. *Neka su U(X) i V(X) matrične funkcije matrične varijable X i neka je U(X) tipa* $p \times q$, *V(X) tipa* $q \times r$ *te X tipa* $m \times n$. *Tada je*

$$\frac{{}^*\partial}{\partial X}U(X)V(X) = (V(X) \otimes I_r)^T \frac{{}^*\partial}{\partial X}U(X) + (I \otimes U(X))\frac{{}^*\partial}{\partial X}V(X).$$

6. *Neka je* $F(X) = X^TAX$ *za neku* A *konstantnu matricu reda* m *i* X *tipa* $m \times n$. *Tada je*

$$\frac{{}^*\partial F}{\partial X} = (X^TA^T \otimes I_n)P + (I_n \otimes X^TA).$$

7. *Neka je* $F(X) = AX^{-1}B$ *za neke* A *i* B *konstantne matrice tipa* $p \times m$ *i* $m \times q$ *i* X *reda* m, *nesingularna. Tada je*

$$\frac{{}^*\partial F}{\partial X} = -(X^{-1}B)^T \otimes (AX^{-1}).$$

8. *Neka su U(X) i Z(X) matrične funkcije matrične varijable X i neka je U(·)tipap×q, Z(·)reda1teXtipam×n. Neka je* $f(X) = Z(X)U(X)$. *Tada je*

$$\frac{{}^*\partial f}{\partial X} = vec(U(X))\frac{{}^*\partial Z(X)}{\partial X} + Z(X)\frac{{}^*\partial U(X)}{\partial X}.$$

Literatura

[1] D. Bakić, Linearna algebra, Školska knjiga, Zagreb, 2008.

[2] I. Budimir, Matrični diferencijalni račun, diplomski rad, Odjel za matematiku, Split, 2016.

[3] S. Liu, G. Trenkler, Hadamard, Khatri-Rao, Kronecker and other matrix products, International Journal of Information and Systems Sciences, 4(2008), 160–177.

[4] K. Mercvajler, Kroneckerov produkt i operator vektorizacije, završni rad, Odjel za matematiku, Osijek, 2018.

[5] C. R. Rao, M. B. Rao, Matrix Algebra and Its Applications to Statistics and Econometrics, World Scientific Publishing, Singapore 2004.