SVEUČILIŠTE U ZAGREBU

Vjeran Hari

MATRIČNA TEORIJA PERTURBACIJE

Zagreb, 1996.

Sadržaj:

1. OSNOVNI REZULTATI

1.1 Oznake

Na početku navedimo osnovnu notaciju koju ćemo koristiti.

Polje realnih (komplesnih) brojeva ćemo označavati s $\mathbf{R}$ ($\mathbf{C}$). Brojeve i druge skalarne veličine ćemo najčešće označavati malim Grčkim slovima. Vektorski prostor svih matrica s elementima iz $\mathbf{R}$ ($\mathbf{C}$) tipa $m \times n$ ćemo označavati s $\mathbf{R}^{m\times n}$ ($\mathbf{C}^{m\times n}$). Za prostor svih n-dimenzionalnih vektora (jednostupčastih matrica) s komponentama iz skupa $\mathbf{R}$ ($\mathbf{C}$) ćemo koristiti oznaku $\mathbf{R}^n$ ($\mathbf{C}^n$). Vektore i matrice ćemo uglavnom označavati pomoću malih i velikih Latinskih slova, respektivno. Lijepo pisana (kaligrafska) velika slova koristit će se za označavanje potprostora (npr. $\mathcal{Z} \subset \mathbf{C}^n$) i linearnih operatora (npr. operator $\mathcal{P}$ ima u standardoj bazi zapis P).

Nul matricu, nul vektor ili skalar nulu ćemo označavati s 0 (ili 0_n gdje je n red nul matrice). Jediničnu matricu ćemo označavati s I (I_n ako iz konteksta nije jasno o kojoj je dimenziji riječ). Stupce jedinične matrice označavat ćemo s e_1, e_2, $\ldots$ Transponiranu matricu matrice A ćemo označavati s A^T, dok će A^* biti hermitski adjungirana (još kažemo kompleksno transponirana) matrica A. Oznaka A^{-1} će predstavljati inverznu matricu matrice A, a A^{-*} će označavati $(A^{-1})^*$.

Trag kvadratne matrice A označavat ćemo s $tr(A) = \sum_{k=1}^{n} a_{kk}$. Oznaka $|A|$ će predstavljati matricu čiji su elementi jednaki $|a_{ij}|$, a $\bar{A}$ matricu čiji su elementi jednaki $\bar{a}_{ij}$. Singularne vrijednosti matrice A ćemo označiti sa $\sigma_i(A)$, a skup svih singularnih vrijednosti sa $\sigma(A)$. Slično, spektar kvadratne matrice A (skup svih vlastitih vrijednosti $\lambda_i(A)$) označavat će se s $\lambda(A)$. Koristit ćemo i spektralni radijus kvadratne matrice A, $\rho(A) = \max\{|\lambda| : \lambda \in \lambda(A)\}$.

Oznaku $\|\cdot\|$ ili $\|\cdot\|_2$ćemo koristiti za Euklidsku vektorsku ili spektralnu matričnu normu (tzv. matrična 2-norma). Od koristi će biti i općenitije $\|\cdot\|_p$ vektorske i matrične norme (posebno za $p = 1$ i $p = \infty$). Još je vrlo važna Frobeniuousova matrična norma $\|A\|_F = (\sum_{i=1}^m \sum_{i=1}^n |a_{ij}|^2)^{\frac{1}{2}}$.

Dijagonalnu matricu označavamo obično sa D, Δ ili sa $\text{diag}(d_1, d_2, \ldots, d_p)$ ako želimo istaknuti dijagonalne elemente. Pritom je općenito

$$\text{diag}(d_1, d_2, \ldots, d_p) \in \mathbf{C}^{m\times n}$$

ako je $p = \min\{m, n\}$ i d_{ii} se nalazi na poziciji (i, i), $1 \leq i \leq p$. Ostali elementi su naravno nula.

Oznaka $\bigoplus_{i=1}^{r} A_i$ će predstavljati direktnu sumu matrica $A_1, ..., A_r$, tj.

$$\bigoplus_{i=1}^{r} A_i = \begin{bmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_r \end{bmatrix}.$$

Ovu blok-dijagonalnu matricu ćemo još označavati s $\text{diag}(A_1, \ldots, A_r)$. Također, oznaka $A \otimes B$, gdje je $A = (a_{ij})$ i $B = (b_{ij})$, će predstavljati tenzorski (Kroneckerov) produkt matrica A i B, tj.

$$A \otimes B = \begin{bmatrix} a_{11}\,B & a_{12}B & a_{13}B & \cdots \\ a_{21}\,B & a_{22}\,B & a_{23}\,B & \cdots \\ a_{31}\,B & a_{32}\,B & a_{33}\,B & \cdots \\ \vdots & \vdots & \vdots & \end{bmatrix}.$$

Ako je $A \in \mathbf{C}^{m\times n}$, potprostor $\mathcal{R}(A) = \{Ax : x \in \mathbf{C}^n\} \subset \mathbf{C}^m$ razapet stupcima od A, ćemo zvati slikom matrice A. Za razliku od slike matrice, nul potprostor $\mathcal{N}(A) = \{x : Ax = 0\}$ će biti iz $\mathbf{C}^n$. Sa $\mathcal{Y} = \mathcal{X}_1 \oplus \mathcal{X}_2$ označavamo direktnu sumu potprostora $\mathcal{X}_1$ i $\mathcal{X}_2$. To znači da se svaki $y \in \mathcal{Y}$ može prikazati na jednoznačan način kao $y = x_1 + x_2$, gdje je $x_1 \in \mathcal{X}_1$, $x_2 \in \mathcal{X}_2$. Ako je potprostor $\mathcal{X}_2$ ortogonalan na potprostor $\mathcal{X}_1$ tada tu sumu zovemo ortogonalna suma.

1.2 QR dekompozicija

Rješavanje nekih praktičnih problema često je povezano s pronalaženjem ortonormiranih baza za $\mathcal{R}(A)$ i $\mathcal{R}(A)^\perp$ zadane matrice A. Taj problem se uspješno rješava pomoću QR-dekompozicije. Pojam QR dekompozicije uvest ćemo koristeći se slijedećom lemom (Householderova lema), koja sama po sebi predstavlja zanimljiv rezultat.

Lema 1.1 [15, Householder] *Neka je* $\|x\|_2 = 1$ *i neka je prva komponenta vektora* x *realna i nenegativna. Ako je*

$$u = \frac{x + e_1}{\|x + e_1\|_2},$$

tada je matrica $H = I - 2uu^*$ *hermitska i unitarna. Štoviše*

$$Hx = -e_1. \tag{1.1}$$

Dokaz. Neka je $x = (\zeta_1, x_2, ..., x_n)^T$. Tada je ${\zeta_1}^2 + \mid x_2 \mid^2 + ... + \mid x_n \mid^2 = 1$. Matrica $H = I - 2uu^*$ je očito hermitska, a kako vrijedi

$$H^*H = H^2 = (I - 2uu^*)^2 = I - 2uu^* - 2uu^* + 4uu^*uu^* = I,$$

ona je i unitarna. Budući da je

$$\|x + e_1\|_2^2 = (\zeta_1 + 1)^2 + |x_2|^2 + ... + |x_n|^2 = 1 + 2\zeta_1 + 1 = 2(1 + \zeta_1),$$

i $(x + e_1)^* x = 1 + \zeta_1$, imamo

$$Hx = \left(I - 2\frac{(x + e_1)(x + e_1)^*}{{\|x + e_1\|_2}^2} \right) x = x - 2\,\frac{(x + e_1)(1 + \zeta_1)}{2(1 + \zeta_1)} = -e_1.$$

■

Potpuno analogno može se dokazati da uz izbor vektora $u = \dfrac{x - e_1}{\|x - e_1\|_2}$ dobivamo $He_1 = x$.

Ako je $x \approx e_1$ (dakle je $\zeta_1 > 0$), izbor $u = (x - e_1)/\|x - e_1\|_2$ može u računanju s konačnom aritmetikom dovesti do netočnog vektora u. Naime, ako je $\|x - e_1\|_2$ blizu $\sqrt{\eta}$, gdje je η najmanji prikaziv pozitivni broj u računalu, tada se nepažljivim računanjem može $\|x - e_1\|_2$ izračunati s velikom relativnom greškom, a to znači da izračunata H neće biti ortogonalna. U takvom slučaju lijek je ili pažljivo računati normu (npr. koristeći gotove rutine iz paketa BLAS1) ili koristiti vektor $u = (x + e_1)/(\|x + e_1\|_2)$.

Ako je $\zeta_1 < 0$ (dakle još uvijek realno) u se može definirati pomoću vektora $-x$. Zbog numeričke stabilnosti, za realno ζ_1 najsigurnije je definirati $u = (x + \text{sign}(\zeta_1)e_1)/\|x + \text{sign}(\zeta_1)e_1\|_2$ i za njega će vrijediti $Hx = -\text{sign}(\zeta_1)e_1$. Ako baš želimo $Hx = e_1$ možemo jednostavno zamijeniti H sa $-\text{sign}(\zeta_1)\, H$.

Ako je $\|x\|_2 = \alpha \neq 1$, tada se u računa iz vektora x/α, pa se dobije $u = (x \pm \alpha e_1)/(\|x \pm \alpha e_1\|_2)$. Uočimo da će H još uvijek biti unitarna i hermitska matrica za koju vrijedi $Hx = \mp\alpha e_1$.

Ako je prva komponenta od x kompleksni broj x_1, tada imamo dva izbora: ili ćemo zahtijevati da je H i unitarna i hermitska pa dobiti $Hx = ze_1$ uz kompleksni broj z, ili odustati of hermitičnosti matrice H pa dobiti $Hx = \beta e_1$ uz realni β. Prvi slučaj ide preko vektora $u = (x \pm \alpha\Phi e_1)/(\|x \pm \alpha\Phi e_1\|_2)$, gdje je $\Phi = \text{diag}(x_1/|x_1|, 1, \ldots, 1)$, $\alpha = \|x\|_2$ i on daje $Hx = \mp\alpha\Phi e_1$. Drugi slučaj ide preko vektora $u = (\Phi^* x \pm \alpha\ e_1)/(\|\Phi^* x \pm \alpha e_1\|_2)$ i unitarne matrice $H\Phi^*$, $H = I - 2uu^*$. Ovdje su Φ i α definirani kao i prije. Pritom vrijedi $H\Phi^* x = \mp\alpha e_1$.

Od sve ove priče za dokaz slijedećeg važnog teorema dostatno je znati da za dani kompleksni ili realni vektor x, postoji unitarna matrica H, takva da je $H^* x = \|x\|_2\, e_1$, tj. $H^*(x/\|x\|_2) = e_1$. Za nju vrijedi $x/\|x\|_2 = He_1$, tj. njen prvi stupac je baš $x/\|x\|_2$.

Teorem 1.2 [15] *Neka je $A \in \mathbf{C}^{m\times n}$, $m \geq n$, tada postoji unitarana matrica Q takva da je*

$$Q^* A = \begin{bmatrix} R \\ O \end{bmatrix},$$

gdje je R gornje-trokutasta matrica s nenegativnim elementima na dijagonali.

Dokaz. Dokaz provodimo matematičkom indukcijom po broju stupaca matrice A. Neka je $n = 1$, tako da je $A = [a]$ jednostupčana. Neka je Q unitarna matrica čiji je prvi stupac vektor $a/\|a\|_2$ (ako je $a = 0$, onda je $Q = I$).
Tada je $Q^*(a/\|a\|_2) = e_1$ pa je

$$Q^* A = \begin{bmatrix} \|a\|_2 \\ 0 \end{bmatrix},$$

što je trebalo dokazati.
Pretpostavimo da je tvrdnja teorema istinita za sve matrice sa $n-1$ stupaca, $n > 1$

i odaberimo $A \in \mathbf{C}^{m\times n}$, $m \geq n$. Matricu A zapišimo u obliku $A = [a\ A_*]$. Neka je H unitarna matrica čiji je prvi stupac $a/\|a\|_2$ (ako je $a = 0$, onda je $H = I$). Tada je

$$H^*\left(\frac{1}{\|a\|}A\right) = \begin{bmatrix} e_1 & b_0^* \\ 0 & C_0 \end{bmatrix} \quad \text{pa je} \quad H^*A = \begin{bmatrix} \|a\| & b^* \\ 0 & C \end{bmatrix}.$$

Prema pretpostavci indukcije, postoji unitarna matrica V takva da je

$$V^*C = \begin{bmatrix} S \\ 0 \end{bmatrix},$$

gdje je S gornja trokutasta matrica s nenegativnim elementima na dijagonali. Ako za matricu Q uzmemo $Q = H\,\mathrm{diag}(1, V)$, onda dobivamo

$$Q^*A = \begin{bmatrix} \|a\| & b^* \\ 0 & S \\ 0 & \mathrm{O} \end{bmatrix},$$

što je traženi oblik. ■

Ako je A ranga n, tj. ima nezavisne stupce, onda i matrica Q^*A ima nezavisne stupce, a to znači da je R regularna. Kako je R regularna, trokutasta i ima nenegativne dijagonalne elemente, zaključujemo da su dijagonalni elementi pozitivni. Ako izvršimo particiju $Q = [Q_A\,,\ Q_A^{\perp}]$, gdje Q_A ima n stupaca, dobivamo

$$A = [Q_A\,,\ Q_A^{\perp}] \begin{bmatrix} R \\ 0 \end{bmatrix} = Q_A R.$$

Dekompoziciju $Q_A R$ nazivat ćemo skraćena QR dekompozicija. Prostor razapet stupcima matrice Q_A jednak je prostoru razapetom stupcima matrice A, tj. $\mathcal{R}\,(Q_A) = \mathcal{R}(A)$. Dakle stupci matrice Q_A čine ortonormiranu bazu za prostor $\mathcal{R}(A)$, dok stupci matrice $Q_A^{\perp}$ čine ortonormiranu bazu za ortogonalni komplement $\mathcal{R}\,(A)^{\perp}$.

Jedna od važnijih posljedica QR ddekompozicije je i Hadamardova nejednakost. Ona neposredno slijedi iz sljedećeg teorema.

Teorem 1.3 *Neka je $A \in \mathbf{C}^{n\times n}$ i neka je $\{u_1, \ldots, u_n\}$ ortonormirana baza u $\mathbf{C}^n$. Tada je*

$$|\det(A)| \leq \prod_{i=1}^{n} \|Au_i\|_2.$$

Pritom nejednakost prelazi u jednakost ako i samo ako su vektori Au_i međusobno ortogonalni ili je neki Au_i nulvektor.

Dokaz. Neka je $Q^*A = R = [r_1, \ldots, r_n]$, QR dekompozicija matrice A. Označimo dijagonalne elemente trokutaste matrice R s r_{ii}, a stupce od A s a_i, $1 \leq i \leq n$. Jer je Q unitarna imamo $|\det(A)| = |\det(R)|$ i $\|a_i\|_2 = \|r_i\|_2$, $1 \leq i \leq n$. Stoga vrijedi

$$| \det(A) |=| \det(R) |= \prod_{i=1}^{n} |r_{ii}| \leq \prod_{i=1}^{n} \|r_i\|_2 = \prod_{i=1}^{n} \|a_i\|_2. \tag{1.2}$$

U slučaju regularne matrice A, nejednakost prelazi u jednakost ako i samo ako je R dijagonalna, tj. ako i samo ako su stupci od A ortogonalni. U slučaju singularne matrice A, nejednakost prelazi u jednakost ako i samo ako A ima nul stupac. Tvrdnja teorema slijedi primjenom relacije (1.2) na matricu AU, gdje je $U = [u_1, \ldots, u_n]$ unitarna matrica. ■

Nejednakost (1.2) se spominje pod imenom Hadamardova nejednakost. Teorem 1.3 daje prirodno poopćenje te nejednakosti.

1.3 Vlastite vrijednosti i vlastiti vektori

U ovoj točki A će biti kvadratna matrica reda n.

Definicija 1.4 *Par (x, λ) je* vlastiti par *matrice A, ako je $x \neq 0$ i*

$$Ax = \lambda x.$$

Vektor x je vlastiti vektor *matrice A, a skalar λ je* pripadna vlastita vrijednost. *Skup svih vlastitih vrijednosti čini spektar matrice.*

Spektar matrice A ćemo označiti sa $\lambda(A)$. Jednadžbu $Ax = \lambda x$ možemo pisati u obliku $(\lambda I - A)x = 0$. Dakle je λ vlastita vrijednost matrice A onda i samo onda ako je $\lambda I - A$ singularna matrica, a to je onda i samo onda ako je λ nultočka karakterističnog polinoma $\Phi_A(\lambda) = det(\lambda I - A)$ matrice A. Kako je $det(\lambda I - A)$ polinom n-tog stupnja, matrica A ima točno n vlastitih vrijednosti (ako ih brojimo uvažavajući višestrukosti nultočaka polinoma). Ako je λ nultočka karakterističnog polinoma višestrukosti k, ona ima algebarsku višestrukost (ili kratnost) k.

Uočimo da je $\Phi_A(\lambda) = 0$ onda i samo onda ako je $\Phi_{A^*}(\bar{\lambda}) = 0$. Dakle svakoj vlastitoj vrijednosti λ matrice A možemo pridružiti vlastitu vrijednost $\bar{\lambda}$ matrice A^* iste kratnosti. Relacija $A^*y = \bar{\lambda}y$ je ekvivalentna s $y^*A = \lambda y^*$. Vektor y zovemo lijevi vlastiti vektor matrice A, a x još zovemo desni vlastiti vektor matrice A (x općenito nećemo zvati desni vlastiti vektor, osim kada ga moramo razlikovati od lijevog vlastitog vektora y). Skup svih vlastitih vektora koji su pridruženi određenoj vlastitoj vrijednosti λ čine vektorski potprostor koji je upravo nul potprostor matrice $\lambda I - A$, a zove se vlastiti potprostor matrice pridružen λ. Dimenzija potprostora $\mathcal{N}(\lambda I - A)$ naziva se geometrijska višestrukost vlastite vrijednosti λ. Sjetimo se da vrijedi $\dim\,[\mathcal{N}(\lambda I - A)] = n - \mathrm{rang}(\lambda I - A)$. Geometrijska višestrukost vlastite vrijednosti λ uvijek je manja ili jednaka algebarskoj višestrukosti. Sljedeći primjer pokazuje da stroga nejednakost između dviju višestrukosti vlastite vrijednsti λ ne mora biti izuzetak

Primjer 1.5 Neka je

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

Tada je $\Phi_A(\lambda) = (\lambda - 1)^2$, što znači da je 1 vlastita vrijednost algebarske višestrukosti dva. S druge strane, jedini vlastiti vektor matrice A je višekratnik vektora $e_1 = [1\ ,\ 0]^T$. To znači da je geometrijska višestrukost te vlastite vrijednosti jednaka jedan.

1.4 Unitarne transformacije sličnosti

U ovoj točki ćemo dokazati važan Schurov teorem i njegove posljedice. Potsjetimo se, A je oznaka za kvadratnu matricu.

Teorem 1.6 [15, Schur] *Za proizvoljnu matricu A postoji unitarna matrica U takva da je $T = U^*AU$ gornja trokutasta matrica. Matrica U može biti izabrana tako da se na dijagonali matrice T pojave vlastite vrijednosti od A u bilo kojem zadanom poretku.*

Dokaz. Dokaz provodimo indukcijom po redu matrice n. Za $n = 1$ tvrdnja je trivijalna. Pretpostavimo sada da tvrdnja teorema vrijedi za sve matrice reda manjeg od

n, $n > 1$, i neka je A matrica reda n. Neka je λ prva vlastita vrijednost u zadanom poretku vlastitih vrijednosti matrice A. Tada postoji vektor x takav da je $Ax = \lambda x$ i $\|x\| = 1$. Nadalje, neka je $H = \begin{bmatrix} x & X \end{bmatrix}$ unitarna matrica, tada matrica H^*AH ima oblik

$$H^*AH = \begin{bmatrix} x^*Ax & x^*AX \\ X^*Ax & X^*AX \end{bmatrix} \equiv \begin{bmatrix} \lambda & b^* \\ 0 & M \end{bmatrix}.$$

Budući da je H^*AH gornja blok-trokutasta matrica, vlastite vrijednosti matrice A su λ i vlastite vrijednosti od M. Po pretpostavci indukcije postoji unitarna matrica V, takva da je V^*MV gornja trokutasta matrica s vlastitim vrijednostima na dijagonali u danom poretku. Za $U = [x\,,\, XV]$ imamo

$$T = U^*A\,U = \begin{bmatrix} \lambda & b^*V \\ 0 & V^*MV \end{bmatrix}.$$

Matrica V^*MV je gornje-trokutasta pa je i matrica T gornja trokutasta, a njene vlastite vrijednosti su smještene na dijagonali u zadanom poretku. ■

Prethodni teorem ima neke vrlo važne posljedice, kao što je naredni teorem koji kaže da je Schurova dekompozija normalnih matrica dijagonalna. Sjetimo se matrica A je normalna ako je $A^*A = AA^*$.

Teorem 1.7 *Ako je A normalna matrica, onda je njena Schurova dekompozicija dijagonalna matrica.*

Dokaz. Treba pokazati da postoji unitarna matrica U takva da je U^*AU dijagonalna matrica. Dokaz ćemo provesti indukcijom. Za $n = 1$ tvrdnja je trivijalna, stoga pretpostavimo da je tvrdnja istinita za sve normalne matrice reda manjeg od n, $n > 1$. Neka je A matrica reda n i $T = U^*AU$ njena Schurova dekompozicija. Promotrimo sljedeću particiju od T

$$T = \begin{bmatrix} \tau & t^* \\ 0 & T_0 \end{bmatrix}.$$

Jer je T unitarno slična normalnoj matrici ona je normalna, pa iz $T^*T = TT^*$ dobivamo

$$\begin{bmatrix} \bar{\tau} & 0 \\ t & T_0^* \end{bmatrix} \begin{bmatrix} \tau & t^* \\ 0 & T_0 \end{bmatrix} = \begin{bmatrix} |\,\tau\,|^2 & \bar{\tau}t^* \\ \tau\, t & t\,t^* + T_0^*T_0 \end{bmatrix}$$

$$\begin{bmatrix} \tau & t^* \\ 0 & T_0 \end{bmatrix} \begin{bmatrix} \bar{\tau} & 0 \\ t & T_0^* \end{bmatrix} = \begin{bmatrix} |\tau|^2 + t^*t & t^*T_0^* \\ T_0 t & T_0 T_0^* \end{bmatrix}.$$

Gledajući elemente na pozicijama $(1,1)$ i $(2,2)$, dobivamo

$$|\tau|^2 = |\tau|^2 + t^*t,$$

i

$$tt^* + T_0^* T_0 = T_0 T_0^*.$$

Prva jednadžba povlači $t = 0$, dok iz druge slijedi da je T_0 normalna matrica. Jer je T Schurova dekompozicija od A, T_0 je gornje-trokutasta, a kako je normalna i reda manjeg od n ona je po pretpostavci indukcije dijagonalna. To znači da je T dijagonalna matrica. ■

Jedna od posljedica ovog teorema jest tvrdnja da normalna matrica reda n ima sustav od n ortonormiranih vlastitih vektora koji razapinju $\mathbf{C}$.

Dvije najznačajnije klase normalnih matrica su unitarne i hermitske matrice. Slijedeći korolar govori o karakterizaciji tih matrica pomoću njihovih vlastitih vrijednosti.

Korolar 1.8 *Unitarna matrica je normalna matrica čije vlastite vrijednosti imaju modul jedan. Hermitska matrica je normalna matrica koja ima realne vlastite vrijednosti.*

Dokaz. Neka je U unitarna matrica. Po teoremu 1.7 postoji unitarna matrica V, takva da je

$$\Lambda \equiv \mathrm{diag}(\lambda_1, ..., \lambda_n) = V^* U V,$$

gdje su λ_i, $1 \leq i \leq n$ vlastite vrijednosti matrice U. Jer je Λ unitarna vijedi $\bar{\Lambda}\Lambda = \Lambda\bar{\Lambda} = I$, što povlači

$$|\lambda_i|^2 = 1 \quad 1 \leq i \leq n.$$

Obratno, ako je U normalna matrica s vlastitim vrijednostima modula jedan, tada je prema teoremu 1.7 $U = V\Lambda V^*$, za neku unitarnu matricu V i $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. Jer su sve vlastite vrijednosti modula jedan, Λ je unitarna, a jer je produkt unitarnih matrica unitarna matrica, U je unitarna.

Pretpostavimo sada da je H hermitska matrica. U tom slučaju po teoremu 1.7 postoji unitarna matrica W takva da je

$$M \equiv \mathrm{diag}(\mu_1, ..., \mu_n) = W^* H\, W,$$

gdje su μ_i, $i = 1, ..., n$ vlastite vrijednosti matrice H. Zbog $H^* = H$, slijedi $M^* = M$, odnosno $\bar{\mu}_i = \mu_i$, $1 \leq i \leq n$. To znači da su svi μ_i realni brojevi.

Obratno, ako je H normalna i $H = WDW^*$ uz unitarnu W i dijagonalnu realnu D, tada je $H = H^*$, tj. H je hermitska. ■

Iz korolara 1.8 odmah se vidi da se hermitska matrica H može prikazati u obliku

$$H = U\Lambda U^*, \tag{1.3}$$

gdje je U unitarna matrica, a Λ dijagonalna matrica s realnim brojevima na dijagonali. Taj oblik zovemo *spektralna dekompozicija* od H. Također ponekad ćemo koristiti i zapis

$$H = \sum_i \lambda_i u_i u_i^*, \tag{1.4}$$

gdje je u_i i-ti stupac matrice U. Pomoću formule (1.4) možemo proširiti pojam skalarne funkcije na hermitske matrice. To se čini na slijedeći način

$$\varphi(H) = \sum_i \varphi(\lambda_i) u_i u_i^*. \tag{1.5}$$

Uočimo da formula (1.5) vrijedi za polinome (pokažite običnim matričnim množenjem da je za polinom φ, $\varphi(H) = U\varphi(\Lambda)U^*$), a kako oni čine gust skup u prostoru neprekidnih funkcija na $\mathbf{R}$, zbog neprekidnosti matričnog množenja (1.5) vrijedi i za neprekidne funkcije realnog argumenta. Slični argument se koristi u pokazivanju da (1.5) vrijedi za neprekidne funkcije definirane u okolini spektra matrice. Jedna od načešće korištenih funkcija je drugi korijen. Ako je H pozitivno definitna matrica, onda se drugi korijen matrice H dobiva formulom

$$H^{\frac{1}{2}} = \sum_i \lambda_i^{\frac{1}{2}} u_i u_i^*.$$

1.5 Invarijantni potprostori

Prirodna generalizacija vlastitih potprostora su invarijantni potprostori. Za normalne matrice oni su i razapeti vlastitim vektorima.

Definicija 1.9 *Za potprostor $\mathcal{X}$ kažemo da je* invarijantan potprostor *kvadratne matrice A ako vrijedi*

$$A\mathcal{X} \subset \mathcal{X}.$$

Neka osnovna svojstva invarijantnih potprostora sadržana su u slijedećem teoremu.

Teorem 1.10 *Neka je $\mathcal{X}$ invarijantni potprostor od A i neka stupci matrice X čine bazu za $\mathcal{X}$. Tada postoji jedinstvena matrica L takva da je*

$$AX = XL. \tag{1.6}$$

Matricu L zovemo reprezentacija matrice A na $\mathcal{X}$ s obzirom na bazu X. Specijalno, (v, λ) je vlastiti par od L onda i samo onda ako je (Xv, λ) vlastiti par od A.

Dokaz. Neka je $X = [x_1\,, x_2\,, \ldots, x_k]$. Kako je za svaki i, $Ax_i \in \mathcal{X}$ to znači da se Ax_i može na jedinstven način prikazati kao linearna kombinacija vektora x_j. Neka je $Ax_i = Xl_i$. Radi jedinstvenosti prikaza vektora Ax_i u bazi $X = [x_1\,, x_2\,, \ldots, x_k]$, vektor l_i je određen na jedinstven način. To vrijedi za svako $1 \leq i \leq k$. Dakle je matrica $L = [l_1\,, l_2\,, \ldots, l_k]$ jedinstveno određena matricom A i za nju vrijedi relacija (1.6). Ako je $u = Xv$, tj. ako je v reprezentacija vektora u u bazi X, onda je zbog $AXv = XLv$, L reprezentacija matrice A u bazi X.

Konačno, ako je (v, λ) vlastiti par od L, onda vrijedi

$$AXv = XLv = \lambda Xv,$$

pa je (Xv, λ) vlastiti par od A. Obratno, ako je (Xv, λ) vlastiti par za A, imamo

$$XLv = AXv = \lambda Xv \quad \text{pa je} \quad X\,(Lv - \lambda v) = 0.$$

Jer su stupci od X linearno nezavisni mora biti $Lv = \lambda v$. ∎

2. Singularna dekompozicija matrice

Ovdje ćemo prikazati osnovne rezultate vezane uz singularnu dekompoziciju matrice. Najprije ćemo dokazati teoreme egzistencije i jedinstvenosti dekompozicije, a onda ćemo prikazati neke primjene.

2.1 Definicija i osnovni teoremi

Ako je H hermitska pozitivno definitna (semidefinitna) matrica onda su njene vlastite vrijednosti pozitivne (nenegativne). Ako je $C \in \mathbf{C}^{m\times n}$ onda su obje matrice C^*C i CC^* hermitske i pozitivno semidefinitne. Ako je $m = n$, vlastite vrijednosti matrica C^*C i CC^* se podudaraju. Ako je $m \neq n$, matrica C^*C (CC^*) ima n (m) vlastitih vrijednosti. Pritom su netrivijalne vlastite vrijednosti ovih matrica su jednake. Kvadratni korijeni vlastitih vrijednosti matrice C^*C odnosno CC^* često se zovu singularne vrijednosti matrice C. Uskoro ćemo pokazati vezu vlastitih vrijednosti ovih matrica sa singularnim vrijednostima od C. Sljedeći teorem o egzistenciji singularne dekompozicije služi ujedno i kao definicija singularnih vrijednosti i vektora.

Teorem 2.1 (Singularna dekompozicija matrice) *Ako je $C \in \mathbf{C}^{m\times n}$, tada postoje unitarne matrice $U \in \mathbf{C}^{m\times m}$ i $V \in \mathbf{C}^{n\times n}$ tako da je*

$$U^*CV = \Sigma, \quad \Sigma = \operatorname{diag}(\sigma_1, \sigma_2, \ldots, \sigma_{\min\{m,n\}}),$$

pri čemu vrijedi

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

Brojeve $\sigma_1, \sigma_2, \ldots, \sigma_{\min\{m,n\}}$ zovemo singularne vrijednosti *matrice C. Stupce matrice U zovemo lijevi, a stupce matrice V desni singularni vektori matrice C.*

Dokaz. Pošto je jedinična sfera u $\mathbf{C}^n$ kompaktni skup i funkcija $f(x) = \|Cx\|_2$ neprekidna, postoji jedinični vektor $v \in \mathbf{C}^n$ takav da je

$$\|Cv\|_2 = \max\{\|Cx\|_2 : \ \|x\|_2 = 1, \ x \in \mathbf{C}^n\}.$$

Ako je $\|Cv\| = 0$, onda je $C = 0$ i faktorizacija u iskazu teorema je trivijalna uz $\Sigma = 0$ i s proizvoljnim unitarnim matricama U i V reda m i n, respektivno.

Ako je $\|Cv\|_2 \neq 0$ stavimo $\sigma_1 = \|Cv\|_2$ i formirajmo jedinični vektor

$$u_1 = \frac{Cv}{\sigma_1} \in \mathbf{C}^m.$$

Za u_1 postoji $m-1$ ortonormiranih vektora $u_2, u_3, \ldots, u_m \in \mathbf{C}^m$, tako da matrica $U_1 = [u_1, u_2, \ldots, u_m]$ bude unitarna. Slično, za $v_1 = v$ postoji $n-1$ ortonormiranih vektora v_2, v_3, $\ldots$, $v_n \in \mathbf{C}^n$ tako da matrica $V_1 = [v_1, v_2, \ldots, v_n]$ bude unitarna. Tada je

$$\begin{aligned} C_1 = U_1^* C V_1 \ &= \begin{bmatrix} u_1^* \\ u_2^* \\ \vdots \\ u_m \end{bmatrix} [Cv_1 \ Cv_2 \ \ldots \ Cv_n] = \begin{bmatrix} u_1^* \\ u_2^* \\ \vdots \\ u_m \end{bmatrix} [\sigma_1 u_1 \ Cv_2 \ \ldots \ Cv_n] \\ &= \begin{bmatrix} \sigma_1 & u_1^* C v_2 & \ldots & u_1^* C v_n \\ 0 & u_2^* C v_2 & \ldots & u_2^* C v_n \\ \vdots & \vdots & \ldots & \vdots \\ 0 & u_m^* C v_2 & \ldots & u_m^* C v_n \end{bmatrix} = \begin{bmatrix} \sigma_1 & z^* \\ 0 & C_2 \end{bmatrix}, \end{aligned}$$

gdje je $z \in \mathbf{C}^{n-1}$, $C_2 \in \mathbf{C}^{(m-1)\times(n-1)}$. Neka je

$$y = \frac{1}{\sqrt{\sigma_1^2 + z^* z}} \begin{bmatrix} \sigma_1 \\ z \end{bmatrix}. \tag{2.1}$$

Za jedinični vektor $V_1 y$ imamo

$$\|C\,(V_1 y)\|_2^2 = \|(U_1^* C V_1) y\|_2^2 = \|C_1 y\|_2^2 = \frac{(\sigma_1^2 + z^* z)^2 + \|C_2 z\|_2^2}{\sigma_1^2 + z^* z} \geq \sigma_1^2 + z^* z,$$

a ovo je striktno veće od σ_1^2 ako je $z \neq 0$. Pošto je to u suprotnosti sa maksimalnošću od σ_1, zaključujemo da je $z = 0$. Stoga je

$$C_1 = U_1^* C V_1 = \begin{bmatrix} \sigma_1 & 0 \\ 0 & C_2 \end{bmatrix} \quad . \tag{2.2}$$

Sada ponavljamo isti argument za matricu $C_2 \in \mathbf{C}^{(m-1)\times(n-1)}$. Na taj taj način dobivamo unitarne matrice U i V kao produkt unitarnih matrica koje su dobijene nakon svakog koraka. Ako je $m \geq n$ taj postupak vodi do dijagonalne matrice Σ. Ako je $m \leq n$, u zadnjem koraku radimo s matricom $C_m \in \mathbf{C}^{1\times(n-m+1)}$. Kako za C_m postoji unitarna matrica, tako da je

$$C_m V_m = [\|C_m\|_2, 0, \cdots, 0] \quad ,$$

očito će lijeve i desne komponentne unitarne matrice u zadnjem koraku biti $U_m = I_1$ i V_m, respektivno. ∎

Primjer 2.2 *Evo primjera singularne dekompozicije matrice reda dva:*

$$\begin{bmatrix} 0.96 & 1.72 \\ 2.28 & 0.96 \end{bmatrix} = C = U\Sigma V^T = \begin{bmatrix} 0.6 & -0.8 \\ 0.8 & 0.6 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0.8 & 0.6 \\ 0.6 & -0.8 \end{bmatrix}^T .$$

∎

Singularne vrijednosti su unitarno invarijantne jer ako je $C = U\Sigma V^*$ singularna dekompozicija od C, tada je za proizvoljne unitarne matrice $W_1 \in \mathbf{C}^{m\times m}$ i $W_2 \in \mathbf{C}^{n\times n}$, $(W_1U)\Sigma(W_2V)^*$ singularna dekompozicija od W_1CW_2.

Iz teorema 2.1 odmah slijede važne relacije

$$\begin{aligned} Cv_i &= \sigma_i u_i, \\ C^* u_i &= \sigma_i v_i. \end{aligned}$$

koje otkrivaju istaknutu ulogu ortonormiranih baza $\{u_i\}$ i $\{v_j\}$ u odnosu na matricu C. Iz relacije $Cx = U\Sigma V^*$ vidimo da se djelovanje matrice na vektor može opisati kao niz od tri jednostavnije transformacije: rotacije, produljivanje/skraćivanje komponenata vektora i opet jedne rotacije.

Polazeći od definicije singularnih vrijednosti iz teorema 2.1, pokažimo da su σ_i, $1 \leq i \leq \min\{m, n\}$ kvadratni korijeni vlastitih vrijednosti matrica C^*C i CC^*. Neka je $U^*CV = \Sigma$, $\Sigma = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_{\min\{m,n\}})$ singularna dekompozicija matrice C. Tada je $\Sigma^* = V^*C^*U$ pa je

$$\begin{aligned} \Sigma^*\Sigma &= V^*C^*\underbrace{UU^*}_{I}CV = V^*C^*CV \in \mathbf{C}^{n\times n} \\ \Sigma\Sigma^* &= U^*C\underbrace{VV^*}_{I}C^*U = U^*CC^*U \in \mathbf{C}^{m\times m}. \end{aligned}$$

Dakle su $\sigma_1^2, \sigma_2^2, \ldots, \sigma_{\min\{m,n\}}^2$ vlastite vrijednosti od C^*C i CC^*. Ako je $m > n$ $(n > m)$ onda CC^* (C^*C) ima još dodatnih $m - n$ $(n - m)$ vlastitih vrijednosti koje su jednake nuli. U tom smislu kažemo da su singularne vrijednosti matrice C korijeni vlastitih vrijednosti od C^*C i CC^*.

Sveza između singularnih vrijednosti od C i vlastitih vrijednosti od C^*C pokazuje da su singularne vrijednosti jedinstvene, tj. da je matrica Σ određena na jedinstven način matricom C . Sljedeći teorem govori o tome koliko su U i V jedinstvene.

Teorem 2.3 *Neka je C kao u teoremu 2.1. Pretpostavimo da za netrivijalne singularne vrijednosti matrice C i za pripadne mnogostrukosti vrijedi $\sigma_1 > \sigma_2 > \ldots > \sigma_k > 0$ i $\mu_1 + \mu_2 + \ldots + \mu_k = r$, respektivno. Neka je*

$$C = U\Sigma V^*, \quad \Sigma = \operatorname{diag}(\sigma_1 I_{\mu_1}, \sigma_2 I_{\mu_2}, \ldots, \sigma_k I_{\mu_k}, 0_{\min\{m,n\}-r})$$

singularna dekompozicija od C i neka su $\tilde{U} \in \mathbf{C}^{m\times m}$ i $\tilde{V} \in \mathbf{C}^{n\times n}$ unitarne matrice. Tada je $C = \tilde{U}\Sigma\tilde{V}^*$ onda i samo onda ako postoje unitarne matrice $U_0 \in \mathbf{C}^{(m-r)\times(m-r)}$, $V_0 \in \mathbf{C}^{(n-r)\times(n-r)}$ i $U_i \in \mathbf{C}^{\mu_i\times\mu_i}$, $1 \le i \le k$, tako da je

$$\begin{aligned} \tilde{U} &= U(U_1 \oplus U_2 \oplus \ldots \oplus U_k \oplus U_0) \\ \tilde{V} &= V(U_1 \oplus U_2 \oplus \ldots \oplus U_k \oplus V_0). \end{aligned}$$

Dokaz. Iz singularne dekompozicije $C = U\Sigma V^*$ slijedi

$$\begin{aligned} CC^* &= U\Sigma\Sigma^*U^* = U[\sigma_1^2 I_{\mu_1} \oplus \sigma_2^2 I_{\mu_2} \oplus \ldots \oplus \sigma_k^2 I_{\mu_k} \oplus 0_{m-r}]U^* \equiv US_1U^*, \\ C^*C &= V\Sigma^*\Sigma V^* = V[\sigma_1^2 I_{\mu_1} \oplus \sigma_2^2 I_{\mu_2} \oplus \ldots \oplus \sigma_k^2 I_{\mu_k} \oplus 0_{n-r}]V^* \equiv VS_2V^*, \end{aligned}$$

gdje su $I_{\mu_j} \in \mathbf{C}^{\mu_j\times\mu_j}$, $1 \le j \le k$ jedinične matrice, a $0_{m-r} \in \mathbf{C}^{(m-r)\times(m-r)}$ i $0_{n-r} \in \mathbf{C}^{(n-r)\times(n-r)}$ nul-matrice. Pritom su S_1 i S_2 dijagonalne. Ako su $\tilde{U} \in \mathbf{C}^{m\times m}$ i $\tilde{V} \in \mathbf{C}^{n\times n}$ unitarne matrice za koje je $C = \tilde{U}\Sigma\tilde{V}^*$, tada vrijedi $CC^* = \tilde{U}S_1\tilde{U}^*$ pa je $US_1U^* = \tilde{U}S_1\tilde{U}^*$. Odavdje slijedi

$$S_1(U^*\tilde{U}) = (U^*\tilde{U})S_1 \quad .$$

Na slični način, relacija $C = \tilde{U}\Sigma\tilde{V}^*$ povlači $C^*C = \tilde{V}S_2\tilde{V}^*$ odnosno $VS_2V^* = \tilde{V}S_2\tilde{V}^*$ pa je

$$S_2(V^*\tilde{V}) = (V^*\tilde{V})S_2 \quad .$$

Dakle $U^*\tilde{U}$ i $V^*\tilde{V}$ komutiraju s dijagonalnim matricama S_1 i S_2, respektivno. To znači da su one unitarne blok dijagonalne s dimenzijama dijagonalnih blokova $\mu_1, \mu_2, \ldots, \mu_k, m-r$ i $\mu_1, \mu_2, \ldots, \mu_k, n-r$, respektivno. Stoga vrijedi

$$\begin{aligned} \tilde{U} &= U[U_1 \oplus U_2 \oplus \ldots \oplus U_k \oplus U_0], \\ \tilde{V} &= V[V_1 \oplus V_2 \oplus \ldots \oplus V_k \oplus V_0], \end{aligned}$$

gdje su $U_0 \in \mathbf{C}^{(m-r)\times(m-r)}$, $V_0 \in \mathbf{C}^{(n-r)\times(n-r)}$ i U_i, $V_i \in \mathbf{C}^{\mu_i\times\mu_i}$,$1 \le i \le k$, unitarne. Sada iz relacije $\tilde{U}\Sigma\tilde{V}^* = U\Sigma V^*$ slijedi nakon množenja s lijeva s U^* i s desna s V

$$U_i\sigma_i V_i^* = \sigma_i I_{\mu_i}, \quad 1 \le i \le k.$$

Kako su $\sigma_i > 0$, slijedi $U_i = V_i$, $1 \le i \le k$. ■

Ako je polazna matrica realna, singularna dekompozicija se zapisuje kao $C = U\Sigma V^T$, pri čemu je Σ kao i dosad, a $U \in \mathbf{R}^{m\times m}$ i $V \in \mathbf{R}^{n\times n}$ su ortogonalne. Pažljivo čitajući dokaze prethodnih teorema, ta tvrdnja se lako dokaže. Na svim mjestima gdje se koristi znak kompleksnog transponiranja treba staviti znak transponiranja, riječi unitarna zamijeniti s ortogonalna, i treba koristiti prostore $\mathbf{R}^m$, $\mathbf{R}^n$, $\mathbf{R}^{m\times n}$ umjesto $\mathbf{C}^m$, $\mathbf{C}^n$, $\mathbf{C}^{m\times n}$, respektivno.

Zanimljiv je odnos između vlastitih i singularnih vrijednosti normalne matrica. Neka je $N \in \mathbf{C}^{n\times n}$ normalna matrica i neka je $N = V\Lambda V^*$ njena spektralna dekompozicija. Pritom je V unitarna matrica, a Λ dijagonalna,

$$\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n), \quad |\lambda_1| \ge |\lambda_2| \ge \ldots \ge |\lambda_n|.$$

Ako definiramo $\Sigma = \mathrm{diag}(|\lambda_1|, |\lambda_2|, \ldots, |\lambda_n|)$ i $\Phi = \mathrm{diag}(e^{i\phi_1}, e^{i\phi_2}, \ldots, e^{i\phi_n})$, onda je $\Lambda = \Phi\Sigma$. Dakle, vrijedi

$$N = V\Lambda V^* = V\Phi\Sigma V^* = \underbrace{V\Phi}_{U}\Sigma V^* = U\Sigma V^* \quad ,$$

što je singularna dekompozicija matrice od N. Zaključujemo da su singularne vrijednosti normalne matrice apsolutne vrijednosti njenih vlastitih vrijednosti.

Sljedeći korolar pokazuje neke od pogodnosti koje pruža singularna dekompozicija matrice. Pritom ćemo koristiti ove oznake: $\sigma_i(C)$ je i-ta najveća singularna vrijednost matrice C, $\sigma_{\max}(C) = \sigma_1(C)$ je najveća, a $\sigma_{\min}(C) = \sigma_{\min\{m,n\}}(C)$ najmanja singularna vrijednost matrice C. Ako se matrica C podrazumijeva, koristit ćemo umjesto $\sigma_i(C)$ kraću oznaku σ_i. To znači, ako nije drukčije naglašeno, pretpostavljamo nerastući uređaj singularnih vrijednosti.

Korolar 2.4 *Neka je $C = U\Sigma V^*$ singularna dekompozicija matrice $C \in \mathbf{C}^{m\times n}$ za koju vrijedi*

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \cdots = \sigma_{\min} = 0.$$

Neka su

$$\begin{aligned} \Sigma_r &= \operatorname{diag}(\sigma_1, \sigma_2, \ldots, \sigma_r) \in \mathbf{C}^{r\times r}, \\ U_r &= [u_1, u_2, \ldots, u_r] \in C^{m\times r} \quad i \\ V_r &= [v_1, v_2, \ldots, v_r] \in \mathbf{C}^{n\times r} \end{aligned}$$

gdje su u_i i v_i stupci od U i V, respektivno. Tada je $\operatorname{rang}(C) = r$ *i*

(i)	$\mathcal{N}(C) = span\{v_{r+1}, \ldots, v_n\}$	*(ii)*	$\mathcal{N}(C^*) = span\{u_{r+1}, \ldots, u_m\}$
(iii)	$\mathcal{R}(C) = span\{u_1, u_2, \ldots, u_r\}$	*(iv)*	$\mathcal{R}(C^*) = span\{v_1, v_2, \ldots, v_r\}$
(v)	$\mathcal{R}(C^*) = \mathcal{N}(C)^{\perp}$	*(vi)*	$\mathcal{N}(C^*) = \mathcal{R}(C)^{\perp}$
(vii)	$\mathcal{R}(C^*C) = \mathcal{R}(C^*),\ \mathcal{R}(CC^*) = \mathcal{R}(C)$	*(viii)*	$C = \sum_{i=1}^{r} \sigma_i u_i v_i^* = U_r \Sigma_r V_r^*$
(ix)	$\|C\|_F^2 = \sigma_1^2 + \sigma_2^2 + \cdots + \sigma_{\min}^2$	*(x)*	$\|C\|_2 = \sigma_1$

Dokaz. Rang matrice se ne mijenja ako pomnožimo matricu s lijeva ili s desna s nesingularnom matricom. Stoga je $\operatorname{rang}(C) = \operatorname{rang}(\Sigma)$. Kako je rang broj linearno nezavisnih stupaca (ili redaka) matrice, očito je $\operatorname{rang}(\Sigma) = r$.

(i) Pošto je $Cv_i = 0$, $r + 1 \leq i \leq n$, zaključujemo da je

$$\operatorname{span}\{v_{r+1}, \ldots, v_n\} \subseteq \mathcal{N}(C).$$

Jer je $\operatorname{defekt}(C) = n - \operatorname{rang}(C) = n - r$, dimenzija potprostora $\operatorname{span}\{v_{r+1}, \ldots, v_n\}$ je ista kao i dimenzija od $\mathcal{N}(C)$. Stoga vrijedi (i).

(ii) Jer je $C^* = V\Sigma U^*$ singularna dekompozicija od C^* i jer C i C^* imaju isti rang, tvrdnja (ii) je tvrdnja (i) za C^*.

(iii) Znamo da je $\mathcal{R}(C) = \text{span}(Cv_1, \ldots, Cv_n)$ jer je $v_1, \ldots, v_n$ baza. Stoga je $\text{span}(Cv_1, \ldots, Cv_r) \subseteq \mathcal{R}(C)$. Međutim, za $1 \leq i \leq r$ vrijedi

$$Cv_i = U\Sigma V^* v_i = U\Sigma e_i = \sigma_i U e_i = \sigma_i u_i \quad \text{i} \quad \sigma_i > 0\,.$$

Jer su u_i linearno nezavisni,

$$\text{span}(\sigma_1 u_1, \ldots, \sigma_r u_r) = \text{span}(u_1, \ldots, u_r)$$

ima dimenziju r kao i $\mathcal{R}(C)$, pa je $\mathcal{R}(C) = \text{span}(u_1, \ldots, u_r)$.

(iv) Dokaz ove tvrdnje koristi isti argument kao i dokaz tvrdnje (ii).

(v) Koristeći tvrdnje (iv) i (i), imamo

$$\mathcal{R}(C^*) = \text{span}\{v_1, \ldots, v_r\} = \mathcal{N}(C)^\perp\,.$$

(vi) Koristeći tvrdnje (ii) i (iii), imamo

$$\mathcal{N}(C^*) = \text{span}\{u_1, \ldots, u_r\} = \mathcal{R}(C)^\perp\,.$$

(vii) Jer su

$$\mathcal{R}(C^*C) = \mathcal{R}(V\Sigma^*\Sigma V^*) \quad \text{i} \quad \mathcal{R}(CC^*) = \mathcal{R}(U\Sigma\Sigma^* U^*)$$

singularnae dekompozicije matrica C^*C i CC^*, respektivno, tvrdnja (vii) slijedi iz tvrdnji (iv) i (iii).

(viii) Tvrdnja slijedi izravno iz matričnog množenja:

$$\begin{aligned} C &= [U_r,\ U_{m-r}] \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} [V_r,\ V_{n-r}]^* = [U_r\Sigma_r, 0] \begin{bmatrix} V_r^* \\ V_{n-r}^* \end{bmatrix} \\ &= U_r \Sigma_r V_r^* = \sum_{i=1}^{r} \sigma_i u_i v_i^* \quad . \end{aligned}$$

(ix) i (x) Tvrdnje slijede iz činjenice da su obje norme unitarno invarijantne. Specijalno vrijedi

$$\begin{aligned} \|C\|_F^2 &= \|U\Sigma V^*\|_F^2 = \|\Sigma\|_F^2 = \sum_{i=1}^{r} \sigma_i^2, \\ \|C\|_2 &= \|U\Sigma V^*\|_2 = \|\Sigma\|_2 = \sigma_{\max}(C). \end{aligned}$$

■

Dekompoziciju opisanu u tvrdnji (viii) korolara 2.4 ćemo zvati *skraćena singularna dekompozicija* matrice C.

Jedna od važnijih posljedica singularne dekompozicije matrice je i polarna dekompozicija. Ona se dobije izravno od singularne dekompozicije matrice.

Teorem 2.5 *Svaka kvadratna matrica C dopušta prikaz*

$$C = Q\,H_1 = H_2\,Q,$$

gdje je Q unitarna, a H_1 i H_2 su hermitske pozitivno semidefinitne matrice cije su vlastite vrijednosti upravo singularne vrijednosti od C.

Dokaz. Iz singularne dekompozicije matrice C odmah slijedi

$$\begin{aligned} C &= U\Sigma V^* = UV^*V\Sigma V^* = (UV^*)\Big(V\Sigma V^*\Big) = Q\,H_1 \\ &= U\Sigma V^* = U\Sigma U^*UV^* = (U\Sigma U^*)\Big(UV^*\Big) = H_2\,Q. \end{aligned}$$

■

Jedna upotreba singularne dekompozicije matrice je određivanje ranga matrice. Brojni teoremi u linearnoj algebri imaju oblik: "ako je ta i ta matrica punog ranga, onda vrijedi to i to svojstvo". Zbog pogrešaka zaokruživanja i eventualnih pogrešaka ulaznih podataka, određivanje ranga postaje netrivijalan zadatak. Za singularnu dekompoziciju matrice postoje pouzdane i prilično točne metode. Stoga se, iako postoje brže metode, rang matrice najsigurnije računa iz singularne dekompozicije.

Slijedeći teorem nam govori o aproksimaciji matrice pomoću matrica manjeg ranga.

Teorem 2.6 (Ekhard, Young, Mirsky) *Neka je $C = U\Sigma V^*$ singularna dekompozicija matrice $C \in \mathbf{C}^{m\times n}$ ranga r. Neka je $k < r$ i*

$$C_k = \sum_{i=1}^{k} \sigma_i u_i v_i^*.$$

Tada je

$$\min_{rang(K)=k} \|C - K\|_2 = \|C - C_k\|_2 = \sigma_{k+1}.$$

Dokaz. Pošto je

$$U^* C_k V = \operatorname{diag}(\sigma_1, \sigma_2, \ldots, \sigma_k, 0, \ldots, 0),$$

lako se dobije

$$\|C - C_k\|_2 = \|U^*(C - C_k)V\|_2 = \|\operatorname{diag}(0, \ldots, 0, \sigma_{k+1}, \ldots, \sigma_{\min\{m,n\}})\|_2 = \sigma_{k+1}.$$

Neka je $K \in \mathbf{C}^{m\times n}$ proizvoljna matrica ranga k. To znači da možemo naći ortonormirane vektore $x_1, \ldots, x_{n-k}$ takve da je

$$\mathcal{N}(K) = \operatorname{span}\{x_1, \ldots, x_{n-k}\}.$$

Jer je

$$\operatorname{span}\{x_1, \ldots, x_{n-k}\} \bigcap \operatorname{span}\{v_1, \ldots, v_{k+1}\} \neq \{0\},$$

možemo pronaći jedinični vektor z koji pripada ovom presjeku. Pošto je $Kz = 0$ i

$$Cz = \sum_{i=1}^{k+1} \sigma_i (v_i^* z) u_i,$$

imamo

$$\|C - K\|_2^2 \geq \|(C - K)z\|_2^2 = \|Cz\|_2^2 = \sum_{i=1}^{k+1} \sigma_i^2 |v_i^* z|^2 \geq \sigma_{k+1}^2,$$

Zadnja nejednakost vrijedi jer se radi o konveksnoj sumi brojeva σ_i, $1 \leq i \leq k+1$. To slijedi iz činjenice da je $0 \leq |v_i^* z|^2 \leq 1$ i

$$\sum_{i=1}^{k+1} |v_i^* z|^2 = \|z\|_2^2 = 1.$$

■

Prethodni teorem nam specijalno kaže da je najmanja singularna vrijednost matrice jednaka udaljenosti (u spektralnoj normi) između matrice i skupa singularnih matrica.

Slijedeći primjer ilustrira tvrdnju teorema 2.6.

Primjer 2.7 a) *Neka su*

$$C = \begin{bmatrix} 6.4 & 1.8 & 2.4 \\ -4.8 & 2.4 & 3.2 \\ 0 & -2.4 & 1.8 \end{bmatrix}, \quad K = \begin{bmatrix} 6.4 & 0 & 0 \\ -4.8 & 0 & 0 \\ 0 & -2.4 & 1.8 \end{bmatrix}.$$

Iz singularne dekompozicije od C

$$\begin{bmatrix} 6.4 & 1.8 & 2.4 \\ -4.8 & 2.4 & 3.2 \\ 0 & -2.4 & 1.8 \end{bmatrix} = \begin{bmatrix} -0.8 & 0.6 & 0 \\ 0.6 & 0.8 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} 8 & & \\ & 5 & \\ & & 3 \end{bmatrix} \begin{bmatrix} -1 & 0 & 0 \\ 0 & 0.6 & 0.8 \\ 0 & 0.8 & -0.8 \end{bmatrix},$$

zaključujemo da je $\sigma_1(C) = 8$, $\sigma_2(C) = 5$, $\sigma_3(C) = 3$ *pa je* $\operatorname{rang}(C) = 3$. *Uočimo da je* $\operatorname{rang}(K) = 2$. *Spektralnu normu od* $C - K$ *dobijemo opet preko njene singularne dekompozicije*

$$\begin{bmatrix} 0 & 1.8 & 2.4 \\ 0 & 2.4 & 3.2 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0.6 & -0.8 & 0 \\ 0.8 & 0.6 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 5 & & \\ & 0 & \\ & & 0 \end{bmatrix} \begin{bmatrix} 0 & 0.6 & 0.8 \\ 0 & -0.8 & 0.6 \\ 1 & 0 & 0 \end{bmatrix}.$$

Dakle imamo $\|C - K\|_2 = 5 > 3 = \sigma_3(C)$.

b) *Neka su*

$$C = \begin{bmatrix} 2.4 & 0 & 2.4 & 0 \\ 1.92 & 0.96 & 1.08 & 1.28 \\ 2.56 & 0.72 & 1.44 & 0.96 \\ 0 & 0.8 & 0 & 0.6 \end{bmatrix}, \quad K = \begin{bmatrix} 2.4 & 0 & 2.4 & 0 \\ 1.92 & 0.96 & 1.08 & 1.28 \\ 2.56 & 0.72 & 1.44 & 0.96 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Slično kao u slučaju a) *singularnom dekompozicijom matrice* C *dobivamo* $\sigma_1(C) = 4$, $\sigma_2(C) = 3$, $\sigma_3(C) = 2$, $\sigma_4(C) = 1$. *Dakle vrijedi* $\operatorname{rang}(C) = 4$. *Očito je* $\operatorname{rang}(K) = 3$, *a spektralna norma od* $C - K$ *je*

$$\|C - K\|_2 = \left\| \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0.8 & 0 & 0.6 \end{bmatrix} \right\|_2 = 1$$

Dakle imamo $\|C - K\|_2 = \sigma_4(C)$, *pa se udaljenost matrice* C *do skupa singularnih matrica dostiže na matrici* K.

Sljedeći teorem je jedan od najznačajnih rezultata vezanih uz singularnu dekompoziciju matrice. On nam daje svezu između singularnih vrijednosti (i vektora) matrice i vlastitih vrijednosti (i vektora) jedne hermitske matrice koja se usko povezana uz danu matricu.

Teorem 2.8 (Jordan–Wielandt) *Neka je*

$$C = U\Sigma V^*, \quad \Sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_n),$$

singularna dekompozicija matrice $C \in \mathbf{C}^{m\times n}$*. Neka je*

$$A = \begin{bmatrix} 0 & C^* \\ C & 0 \end{bmatrix} \in \mathbf{C}^{(m+n)\times(m+n)}.$$

Tada postoji unitarna matrica Q*, takva da vrijedi*

$$Q^*AQ = \mathrm{diag}\{\sigma_1, \ldots, \sigma_{\min}, -\sigma_1, \ldots, -\sigma_{\min}, \underbrace{0, \ldots, 0}_{|m-n|}\}.$$

Dokaz. Možemo pretpostaviti da je $m \geq n$. U protivnom radimo s matricom C^*. Neka je

$$U^*CV = \begin{bmatrix} \Sigma \\ 0 \end{bmatrix}$$

singularna dekompozicija matrice C. Konjugiranjem i transponiranjem dobivamo

$$V^*C^*U = [\Sigma\ ,\ 0]\,.$$

Množenjem lijevih i desnih strana zadnjih jednakosti slijedi

$$V^*(C^*C)V = \mathrm{diag}\{\sigma_1^2, \ldots, \sigma_n^2\} \in \mathbf{C}^{n\times n},$$

odnosno

$$U^*(CC^*)U = \mathrm{diag}\{\sigma_1^2, \ldots, \sigma_n^2, \underbrace{0, \ldots, 0}_{m-n}\} \in \mathbf{C}^{m\times m}.$$

Označimo matricu prvih n stupaca matrice U s U_1, a matricu preostalih $m-n$ stupaca s U_2. Dakle $U = [U_1,\ U_2]$. Neka je

$$Q = \frac{1}{\sqrt{2}} \begin{bmatrix} V & V & 0 \\ U_1 & -U_1 & \sqrt{2}U_2 \end{bmatrix}.$$

Q je unitarna matrica jer vrijedi

$$\begin{aligned}
Q^*Q &= \frac{1}{2}\begin{bmatrix} V^* & U_1^* \\ V^* & -U_1^* \\ 0 & \sqrt{2}U_2^* \end{bmatrix}\begin{bmatrix} V & V & 0 \\ U_1 & -U_1 & \sqrt{2}U_2 \end{bmatrix} \\
&= \frac{1}{2}\begin{bmatrix} V^*V + U_1^*U_1 & V^*V - U_1^*U_1 & \sqrt{2}U_1^*U_2 \\ V^*V - U_1^*U_1 & V^*V + U_1^*U_1 & -\sqrt{2}U_1^*U_2 \\ \sqrt{2}U_2^*U_1 & -\sqrt{2}U_2^*U_1 & 2U_2^*U_2 \end{bmatrix} \\
&= \frac{1}{2}\begin{bmatrix} 2I_n & 0 & 0 \\ 0 & 2I_n & 0 \\ 0 & 0 & 2I_{m-n} \end{bmatrix} = I.
\end{aligned}$$

Izračunajmo još

$$\begin{aligned}
Q^*AQ &= \frac{1}{2}\begin{bmatrix} V^* & U_1^* \\ V^* & -U_1^* \\ 0 & \sqrt{2}U_2^* \end{bmatrix}\begin{bmatrix} 0 & C^* \\ C & 0 \end{bmatrix}\begin{bmatrix} V & V & 0 \\ U_1 & -U_1 & \sqrt{2}U_2 \end{bmatrix} \\
&= \frac{1}{2}\begin{bmatrix} U_1^*C & V^*C^* \\ -U_1^*C & V^*C^* \\ \sqrt{2}U_2^*C & 0 \end{bmatrix}\begin{bmatrix} V & V & 0 \\ U_1 & -U_1 & \sqrt{2}U_2 \end{bmatrix} \\
&= \frac{1}{2}\begin{bmatrix} U_1^*CV + V^*C^*U_1 & U_1^*CV - V^*C^*U_1 & \sqrt{2}V^*C^*U_2 \\ -U_1^*CV + V^*C^*U_1 & -U_1^*CV - V^*C^*U_1 & \sqrt{2}V^*C^*U_2 \\ \sqrt{2}U_2^*CV & \sqrt{2}U_2^*CV & 0 \end{bmatrix}.
\end{aligned}$$

Produkt $V^*C^*U_2$ je nul-matrica jer je

$$V^*C^*U_2 = V^*C^*\underbrace{UU^*}_{I}U_2 = \begin{bmatrix} \Sigma \, , & 0 \end{bmatrix}\begin{bmatrix} 0 \\ I_{m-n} \end{bmatrix} = 0 \quad .$$

Zbog toga vrijedi $\quad U_2^*CV = (V^*C^*U_2)^* = 0$. Pošto su $U_1^*CV = \Sigma$ i $V^*C^*U_1 = \Sigma^*$ kvadratne realne i dijagonalne, one su i jednake. Stoga dobivamo

$$\begin{aligned}
Q^*AQ &= \frac{1}{2}\begin{bmatrix} \Sigma + \Sigma & 0 & 0 \\ 0 & -\Sigma - \Sigma & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} \Sigma & 0 & 0 \\ 0 & -\Sigma & 0 \\ 0 & 0 & 0 \end{bmatrix} \\
&= \operatorname{diag}\{\sigma_1, \sigma_2, \ldots, \sigma_n, -\sigma_1, -\sigma_2, \ldots, -\sigma_n, \underbrace{0, \ldots, 0}_{m-n}\}.
\end{aligned}$$

■

Dakle, netrivijalne vlastite vrijednosti matrice A su pozitivne i negativne netrivijalne singularne vrijednosti matrice C, dok se iz matrice Q se lako odrede lijevi i desni singularni vektori od C i obratno. Prethodni teorem omogućava da mnoge rezultate o vlastitim vrijednostima hermitske matrice formuliramo, uz prirodne modifikacije, za singularne vrijednosti proizvoljne matrice.

2.2 Weylov teorem i neke posljedice

U ovoj točki ćemo prikazati tri teorema čiji će se rezultati koristiti kasnije. Rezultat prvog teorema je jedna od mnogobrojnih posljedica Fischerove karakterizacije vlastitih vrijednosti hermitskih matrica.

Teorem 2.9 *Neka su H i M hermitske matrice i neka je $\tilde{H} = H + M$. Neka su vlastite vrijednosti matrica H, M, i $\tilde{H}$ $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_n$, $\mu_1 \geq \mu_2 \geq ... \geq \mu_n$ i $\tilde{\lambda}_1 \geq \tilde{\lambda}_2 \geq ... \geq \tilde{\lambda}_n$, redom. Tada vrijedi*

$$\lambda_i + \mu_n \leq \tilde{\lambda}_i \leq \lambda_i + \mu_1, \qquad i = 1, ..., n.$$

Dokaz. Vidi [15, teorem 3.14 (str 25-26)].

Teorem 2.9 ima mnogo važnih posljedica. Dokazat ćemo dva teorema koja će nam koristiti u daljnjem radu.

Teorem 2.10 *Neka matrica $C \in \mathbf{C}^{m \times n}$ ima particiju*

$$C = \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} \begin{array}{l} \}\, k \\ \}\, m-k \end{array}.$$

Tada vrijedi

$$\begin{aligned} \sigma_i(C_1) &\leq \sigma_i(C), \quad 1 \leq i \leq \min\{n\,,\, k\} \\ \sigma_j(C_2) &\leq \sigma_j(C), \quad 1 \leq j \leq \min\{n\,,\, m-k\} \end{aligned}$$

Dokaz. Kvadrati singularnih vrijednosti od C jednaki su vlastitim vrijednostima od C^*C, a kvadrati singularnih vrijednosti od C_1 (C_2) su vlastite vrijednosti od

$C_1^*C_1$ ($C_2^*C_2$). Uočimo da je $C^*C = C_1^*C_1 + C_2^*C_2$. Matrica $C_2^*C_2$ ($C_1^*C_1$) je pozitivno semidefinitna. Prema teoremu 2.8 slijedi da su vlastite vrijednosti od C^*C, uzete u nepadajućem poretku, veće ili jednake od odgovarajućih vlastitih vrijednosti matrice $C_1^*C_1$ ($C_2^*C_2$). ■

Zanimljivo je vidjeti koji je odnos između singularnih vrijednosti produkta dviju matrica i singularnih vrijednosti samih matrica. O tome govori slijedeći teorem.

Teorem 2.11 *Neka su* $A \in \mathbf{C}^{m\times k}$, $B \in \mathbf{C}^{k\times n}$ *i* $C = AB$. *Tada vrijedi*

$$\sigma_i(AB) \le \sigma_1(A)\sigma_i(B), \quad \sigma_i(AB) \le \sigma_i(A)\sigma_1(B), \quad 1 \le i \le \min\{m, n, k\}.$$

Dokaz. Druga tvrdnja slijedi iz prve zbog

$$\sigma_i(C) = \sigma_i(C^*) = \sigma_i(B^*A^*) \le \sigma_1(B^*)\sigma_i(A^*) = \sigma_1(B)\sigma_i(A), \quad 1 \le i \le \min\{m, n, k\}.$$

Prvu tvrdnju ćemo dokazati uspoređujući vlastite vrijednosti od C^*C s vlastitim vrijednostima od $\|A\|_2^2 B^*B$. Neka je $A^*A = Q\Lambda^2Q^*$ spektralna dekompozicija od A^*A. Tada je $M = Q(\|A\|_2^2\, I - \Lambda^2)Q^*$ hermitska pozitivno semidefinitna matrica za koju je $A^*A + M = \|A\|_2^2 I$. Vrijedi

$$\|A\|_2^2 B^*B = B^*(A^*A + M)B = C^*C + B^*MB.$$

Uočimo da je $x^*B^*MBx = (Bx)^*M(Bx) \ge 0$ pa je B^*MB također pozitivno semidefinitna. Na osnovu teorema 2.9 zaključujemo da su vlastite vrijednosti od $\|A\|^2 B^*B = \sigma_1(A)^2 B^*B$ veće ili jednake od odgovarajućih vlastitih vrijednosti od C^*C. ■

2.3 Generalizirani inverz

Singularna dekompozicija ima direktnu primjenu u računanju generaliziranog inverza opće matrice. Generalizirani inverz je osnovni alat za jedno važno područje linearne algebre koje se opisuje nazivom linearni problem najmanjih kvadrata. Na ovom mjestu ćemo uvesti pojam Generaliziranog inverza $A^\dagger$ matrice A i navesti neka njegova osnovna svojstva.

Dobro je poznato, da za regularnu matricu $A \in \mathbf{C}^{n\times n}$, postoji jedinstvena matrica X koja zadovoljava

$$A\,X = X\,A = I. \tag{2.3}$$

Matrica X je inverz matrice matrice A i označava se sa A^{-1}.

Prirodno je pokušati proširiti pojam inverza i na matrice koje nisu regularne ili čak nisu ni kvadratne. To je moguće postići zahtijevajući da matrica X zadovoljava ponešto drugačije (oslabljene) uvjete nego što je (2.3). Najpoznatiji generalizirani inverz je tzv. Moore-Penroseov inverz koji je određen sa slijedeća četiri uvjeta (tzv. *Penroseovi uvjeti* [15]):

$$\begin{array}{ll} 1. & A\,X\,A = A, \\ 2. & X\,A\,X = X, \\ 3. & (A\,X)^* = A\,X, \\ 4. & (X\,A)^* = X\,A. \end{array} \tag{2.4}$$

Pokazat ćemo da ovi uvjeti na jedinstven način određuju matricu X koju onda označavamo sa $A^\dagger$. U slijedećem teoremu ćemo odrediti eksplicitnu formulu pomoću koje se određuje generalizirani inverz $A^\dagger$.

Teorem 2.12 *Neka je $A \in \mathbf{C}^{m\times n}$. Tada postoji jedinstvena matrica $X \in \mathbf{C}^{n\times m}$, koja zadovoljava Penroseove uvjete (2.4). Ta matrica ima oblik*

$$A^\dagger = V \begin{bmatrix} \Sigma_+^{-1} & 0 \\ 0 & 0 \end{bmatrix} U^*, \quad \textit{pri čemu je} \quad A = U \begin{bmatrix} \Sigma_+ & 0 \\ 0 & 0 \end{bmatrix} V^*$$

singularna dekompozicija matrice A.

Dokaz. Za matricu X koja zadovoljava, npr. i-ti i j-ti Penroseov uvjet koristit ćemo oznaku $A^{(i,j)}$ i zvati ga (i,j)–inverz. Tako će na primjer (1)–inverz $A^{(1)}$ zadovoljavati samo prvi Penroseov uvjet, a (1,2,4)–inverz, $A^{(1,2,4)}$, će zadovoljavati 1., 2. i 4. uvjet. Dakle $A^{(1,2,3,4)}$ će zadovoljavati sva četiri uvjeta, pa će ta matrica biti generalizirani inverz matrice A.

Jer je $A^{(1)} \in \mathbf{C}^{n\times m}$ matricu $A^{(1)}$ možemo zapisati u obliku

$$A^{(1)} = V \begin{bmatrix} T & K \\ L & M \end{bmatrix} U^*.$$

Uvrštavajući ovaj izraz za $A^{(1)}$ u prvi Penroseov uvjet dobivamo

$$A\,A^{(1)}\,A = U \begin{bmatrix} \Sigma_+ T \Sigma_+ & 0 \\ 0 & 0 \end{bmatrix} V^* = A = U \begin{bmatrix} \Sigma_+ & 0 \\ 0 & 0 \end{bmatrix} V^*.$$

Dakle je $T = \Sigma_+^{-1}$, tj. svaki (1)–inverz ima oblik

$$A^{(1)} = V \begin{bmatrix} \Sigma_+^{-1} & K \\ L & M \end{bmatrix} U^*,$$

gdje su K, L i M proizvoljne. Izračunajmo sada (1,2)–inverz. Jer je on i (1)–inverz možemo poći od te matrice. Iz drugog Penroseovog uvjeta slijedi

$$A^{(1,2)} = V \begin{bmatrix} \Sigma_+^{-1} & K \\ L & M \end{bmatrix} U^* = A^{(1,2)}\,A\,A^{(1,2)} = V \begin{bmatrix} \Sigma_+^{-1} & K \\ L & L\Sigma_+ K \end{bmatrix} U^*,$$

pa zaključujemo da je $M = L\Sigma_+ K$, dok su K i L još uvijek proizvoljne matrice.

Na isti način ćemo odrediti i (1,2,3)–inverz. Polazimo od

$$A^{(1,2,3)} = V \begin{bmatrix} \Sigma_+^{-1} & K \\ L & L\Sigma_+ K \end{bmatrix} U^*.$$

Treći Penroseov uvjet se zapisuje u obliku

$$A^{(1,2,3)} A = U \begin{bmatrix} I & 0 \\ K^* \Sigma_+ & 0 \end{bmatrix} U^* = \left(A^{(1,2,3)} A\right)^* = U \begin{bmatrix} I & \Sigma_+ K \\ 0 & 0 \end{bmatrix} U^*,$$

odakle zaključujemo da svaki (1,2,3)–inverz ima oblik

$$A^{(1,2,3)} = V \begin{bmatrix} \Sigma_+^{-1} & 0 \\ L & 0 \end{bmatrix} U^*.$$

Pritom je L proizvoljna podmatrica.

Na isti način, iz četvrtog Penroseovog uvjeta zaključujemo da svaki (1,2,4)–inverz ima oblik

$$A^{(1,2,4)} = V \begin{bmatrix} \Sigma_+^{-1} & K \\ 0 & 0 \end{bmatrix} U^*,$$

gdje je K proizvoljna podmatrica.

Konačno, iz općeg oblika (1,2,3)–inverza i (1,2,4)–inverza zaključujemo da je

$$A^{\dagger} = A^{(1,2,3,4)} = V \begin{bmatrix} \Sigma_{+}^{-1} & 0 \\ 0 & 0 \end{bmatrix} U^{*}. \tag{2.5}$$

Dakle, ako $A^{\dagger}$ postoji nužno ima oblik kao u relaciji (2.5). Iz dokaza međutim slijedi da $A^{(1,2,3,4)}$ zadovoljava sve sve Penroseove uvjete pa je $A^{\dagger} = A^{(1,2,3,4)}$. Kako u dokazu ništa nije proizvoljno, slijedi da je $A^{\dagger}$ relacijom (2.5) na jedinstven način određena iz singularne dekompozicije matrice A. Sama singularna dekompozicija je prema teoremu 2.3 određena do na množenje matrica U i V s desna s blok-dijagonalnim unitarnim matricama. Međutim, kao što A ne ovisi o toj slobodi u matricama U i V, tako ne ovisi ni $A^{\dagger}$. ■

Sada ćemo navesti neka osnovna svojstva generaliziranog inverza koja se lako dokazuju pomoću relacije (2.4) ili (2.5).

Teorem 2.13 *[15] Za proizvoljnu matricu A vrijedi :*

1. $\left(A^{\dagger}\right)^{\dagger} = A.$
2. $\left(\bar{A}\right)^{\dagger} = \overline{\left(A^{\dagger}\right)}.$
3. $\left(A^{T}\right)^{\dagger} = \left(A^{\dagger}\right)^{T}.$
4. $\mathrm{rang}(A) = \mathrm{rang}(A^{\dagger}) = \mathrm{rang}(A\,A^{\dagger}) = \mathrm{rang}(A^{\dagger}\,A).$
5. $(A\,A^{*})^{\dagger} = (A^{*})^{\dagger}\,A^{\dagger}, \quad (A^{*}\,A)^{\dagger} = A^{\dagger}\,(A^{*})^{\dagger}.$
6. $(A\,A^{*})^{\dagger}\,A\,A^{*} = A\,A^{\dagger}, \quad (A^{*}\,A)^{\dagger}\,A^{*}\,A = A^{\dagger}\,A.$
7. *Ako matrica $A \in \mathbf{C}^{m\times n}$ ima rang n, tada je $A^{\dagger} = (A^{*}A)^{-1}\,A^{*}$ i $A^{\dagger}A = I_n$.*
8. *Ako matrica $A \in \mathbf{C}^{m\times n}$ ima rang m, tada je $A^{\dagger} = A^{*}\,(A\,A^{*})^{-1}$ i $A\,A^{\dagger} = I_m$.*
9. *Ako je $A = F\,G^{*}$ i $\mathrm{rang}(A) = \mathrm{rang}(F) = \mathrm{rang}(G)$, tada je*
$$A^{\dagger} = G\,(F^{*}A\,G)^{-1}\,F^{*} \quad i \quad A^{\dagger} = \left(G^{\dagger}\right)^{*} F^{\dagger}.$$
10. *Ako su U i V unitarne matrice, tada je*
$$(U\,A\,V)^{\dagger} = V^{*}\,A^{\dagger}\,U^{*}.$$

3. PAROVI POTPROSTORA I PAROVI PROJEKTORA

U ovom poglavlju ćemo proučavati odnos dvaju potprostora u $\mathbf{C}^n$. Prvo ćemo dokazati teorem o CS dekompoziciji unitarne matrice. To je osnovni alat za definiranje kanonskih kutova između potprostora jednakih dimenzija. Poteškoće koje mogu nastati primijenom tih rezultata vezane su za izbor baza u kojima se prikazuju potprostori. U mnogim prilikama je bolje uspoređivati potprostore pomoću njihovih pripadnih ortogonalnih projektora, jer oni ne ovise o izboru baze. Stoga ćemo dati neka osnovna svojstva ortogonalnih projektora, a zatim ćemo prikazati rezultate Wedina [18] o dekompoziciji para ortogonalnih projektora. To je osnovni alat za proučavanje kutova između dva potprostora različite dimenzije.

3.1 CS dekompozicija

U ovoj točki ćemo dokazati teorem o CS dekompoziciji unitarne matrice. Nakon toga ćemo teorem ilustrirati sa dva numerička primjera.

Teorem 3.1 **[15]** *Neka je $W \in \mathbf{C}^{n\times n}$ unitarna matrica i neka je dana particija*

$$W = \begin{array}{cc} \left[\begin{array}{cc} W_{11} & W_{12} \\ W_{21} & W_{22} \end{array}\right] & \begin{array}{c} l \\ n-l \end{array} \\ \begin{array}{cc} l & \quad n-l \end{array} & \end{array} , \quad 2l \le n.$$

Tada postoje unitarne matrice $U = \mathrm{diag}(U_{11}, U_{22})$ i $V = \mathrm{diag}(V_{11}, V_{22})$, gdje su $U_{11}, V_{11} \in \mathbf{C}^{l\times l}$, takve da je

$$U^* W V = \begin{array}{cc} \left[\begin{array}{ccc} \Gamma & -\Sigma & 0 \\ \Sigma & \Gamma & 0 \\ 0 & 0 & I \end{array}\right] & \begin{array}{c} l \\ l \\ n-2l \end{array} \\ \begin{array}{ccc} l & l & n-2l \end{array} & \end{array} . \tag{3.1}$$

Pritom vrijedi

$$\begin{aligned} \Gamma &= \operatorname{diag}(\gamma_1, ..., \gamma_l) \geq 0, \\ \Sigma &= \operatorname{diag}(\sigma_1, ..., \sigma_l) \geq 0, \\ \Gamma^2 + \Sigma^2 &= I, \end{aligned}$$

pri čemu dijagonalni elementi od Γ *mogu biti u bilo kojem poretku.*

Dokaz. Kako je W unitarna matrica, njene su singularne vrijednosti jednake 1. Dvostrukom primjenom teorema 2.10 slijedi da su singularne vrijednosti matrice W_{11} manje ili jednake 1. Neka je $U_{11}^* W_{11} V_{11} = \Gamma$ singularna dekompozicija od W_{11}. Istovremenom permutacijom stupaca od U_{11} i V_{11} možemo postići

$$\Gamma = \operatorname{diag}(\Gamma_1, I_{l-k}), \quad \Gamma_1 = \operatorname{diag}(\gamma_1, ..., \gamma_k), \quad 0 \leq \gamma_1 \leq \cdots \leq \gamma_k < 1. \tag{3.2}$$

Uočimo da su stupci matrice

$$\begin{bmatrix} W_{11} \\ W_{21} \end{bmatrix} V_{11}$$

ortonormirani. Ako uz to uzmemo u obzir da je

$$V_{11}^* W_{11}^* W_{11} V_{11} = (U_{11}^* W_{11} V_{11})^* \, U_{11}^* W_{11} V_{11} = \Gamma^2,$$

lako zaključimo da vrijedi

$$I = V_{11}^* \begin{bmatrix} W_{11}^* & W_{21}^* \end{bmatrix} \begin{bmatrix} W_{11} \\ W_{21} \end{bmatrix} V_{11} = \Gamma^2 + V_{11}^* W_{21}^* W_{21} V_{11},$$

odnosno

$$(W_{21} V_{11})^* \, (W_{21} V_{11}) = \operatorname{diag}(I - \Gamma_1^2 \, , \, 0_{l-k}).$$

To znači da su prvih k stupaca matrice $W_{21} V_{11}$ ortogonalni, a zadnjih $l - k$ su nul-stupci. Stoga je u QR dekompoziciji matrice $W_{21} V_{11}$, trokutasta matrica R dijagonalna. Znači da postoji unitarna matrica $\hat{U}_{22} \in \mathbf{C}^{(n-l) \times (n-l)}$, takva da je

$$\hat{U}_{22}^* W_{21} V_{11} = \begin{array}{c} \begin{bmatrix} \Sigma \\ 0 \end{bmatrix} \end{array} \begin{array}{c} l \\ n - 2l \end{array}$$

gdje je Σ dijagonalna matrica reda l s nenegativnim elementima. Uočimo da je

$$\begin{bmatrix} U_{11} & 0 \\ 0 & \hat{U}_{22} \end{bmatrix}^* \begin{bmatrix} W_{11} \\ W_{21} \end{bmatrix} V_{11} = \begin{bmatrix} \Gamma \\ \Sigma \\ 0 \end{bmatrix}$$

matrica s ortonormiranim stupcima. Stoga je $\Gamma^2 + \Sigma^2 = I$, pa je

$$\gamma_i^2 + \sigma_i^2 = 1, \qquad 1 \leq i \leq l. \tag{3.3}$$

To opet zbog relacije (3.2) znači da je

$$\Sigma = \text{diag}\,(\ \Sigma_1\,,\, 0)\,, \quad \Sigma_1 = \text{diag}\,(\sigma_1, ..., \sigma_k)\,, \quad 1 \geq \sigma_1 \geq \cdots \geq \sigma_k > 0. \tag{3.4}$$

Na sličan način se radi i sa prvim blok-retkom matrice W. Matrica $U_{11}^*[W_{11}V_{11}\,,\, W_{12}]$ $= [\Gamma\,,\, U_{11}^* W_{12}]$ ima ortonormirane retke pa je

$$I = [\Gamma \;\; U_{11}^* W_{12}] \begin{bmatrix} \Gamma \\ W_{12}^* U_{11} \end{bmatrix} = \Gamma^2 + U_{11}^* W_{12} W_{12}^* U_{11}.$$

Zaključujemo da $U_{11}^* W_{12}$ ima ortogonalne retke, pa je u QR dekompoziciji matrice $-W_{12}^* U_{11}$ trokutasta matrica R dijagonalna. Dakle postoji unitarna matrica $V_{22} \in \mathbf{C}^{(n-l)\times(n-l)}$, takva da je

$$U_{11}^* W_{12} V_{22} = [T\,,\, 0]\,, \quad T = \text{diag}(\tau_1, ..., \tau_l), \quad \tau_i \leq 0, \quad 1 \leq i \leq l.$$

Kako je

$$U_{11}^*\,[W_{11}\,,\, W_{12}]\ \text{diag}(V_{11}\,,\, V_{22}) = U_{11}^*\,[W_{11}V_{11}\,,\, W_{12}V_{22}] = [\Gamma\,,\, T\,,\, 0]$$

matrica s ortonormiranim recima, imamo $\Gamma^2 + T^2 = I$, pa je $T = -\Sigma$. Neka su $\hat{U} = \text{diag}(U_{11}, \hat{U}_{22})$ i $V = \text{diag}(V_{11}, V_{22})$. Tada matrica $X = \hat{U}^* W V$ nosi particiju

$$X \;=\; \begin{array}{c} \begin{bmatrix} \Gamma_1 & 0 & -\Sigma_1 & 0 & 0 \\ 0 & I & 0 & 0 & 0 \\ \Sigma_1 & 0 & X_{33} & X_{34} & X_{35} \\ 0 & 0 & X_{43} & X_{44} & X_{45} \\ 0 & 0 & X_{53} & X_{54} & X_{55} \end{bmatrix} \begin{array}{l} k \\ l-k \\ k \\ l-k \\ n-2l \end{array} \\ \begin{array}{ccccc} k & l-k & k & l-k & n-2l \end{array} \end{array}.$$

Matrica X je unitarna, a kvadratna dijagonalna matrica Σ_1 ima pozitivne dijagonalne elemente. Zbog

$$0 = (X^*X)_{13} = -\Gamma_1\Sigma_1 + \Sigma_1 X_{33}$$

slijedi $X_{33} = \Gamma_1$. Sada

$$I = (X^*X)_{33} = {\Sigma_1}^2 + X_{33}^*X_{33} + X_{43}^*X_{43} + X_{53}^*X_{53} = {\Sigma_1}^2 + \Gamma_1^2 + X_{43}^*X_{43} + X_{53}^*X_{53}$$

povlači $X_{43}^*X_{43} + X_{53}^*X_{53} = 0$. Primijenimo li funkciju trag na obje strane ove matrične jednakosti dobivamo

$$0 = \mathrm{trag}(X_{43}^*X_{43}) + \mathrm{trag}(X_{53}^*X_{53}) = \|X_{43}\|_F^2 + \|X_{53}\|_F^2,$$

što znači da je $X_{43} = 0$ i $X_{53} = 0$. Potpuno analogno se iz relacije $I = (XX^*)_{33}$ pokazuje da je $X_{34} = 0$ i $X_{35} = 0$. Dakle matrica X ima oblik

$$X = \begin{bmatrix} \Gamma_1 & 0 & -\Sigma_1 & 0 & 0 \\ 0 & I & 0 & 0 & 0 \\ \Sigma_1 & 0 & \Gamma_1 & 0 & 0 \\ 0 & 0 & 0 & X_{44} & X_{45} \\ 0 & 0 & 0 & X_{54} & X_{55} \end{bmatrix} \quad \text{gdje je} \quad \begin{bmatrix} X_{44} & X_{45} \\ X_{54} & X_{55} \end{bmatrix}$$

unitarna matrica. Označimo ju s U_{33}, tj. neka je

$$U_{33} = \begin{bmatrix} X_{44} & X_{45} \\ X_{54} & X_{55} \end{bmatrix} \in \mathbf{C}^{(n-l-k)\times(n-l-k)}.$$

Sada imamo

$$\mathrm{diag}(I_{l+k},\ U_{33}^*)X = \begin{bmatrix} \Gamma_1 & 0 & -\Sigma_1 & 0 & 0 \\ 0 & I & 0 & 0 & 0 \\ \Sigma_1 & 0 & \Gamma_1 & 0 & 0 \\ 0 & 0 & 0 & I & 0 \\ 0 & 0 & 0 & 0 & I \end{bmatrix},$$

što je upravo jednakost (3.1). S obzirom da je

$$\mathrm{diag}(I_{l+k},\ U_{33}^*)X = \mathrm{diag}(I_{l+k},\ U_{33}^*)\hat{U}^*WV,$$

možemo U odnosno U_{22} definirati pomoću relacije

$$\begin{aligned} U &= \hat{U}\,\mathrm{diag}\,(I_{l+k},\quad U_{33}) = \mathrm{diag}\left(U_{11}, \hat{U}_{22}\right)\mathrm{diag}\,(I_l,\quad \mathrm{diag}\,(I_k, U_{33})) \\ &= \mathrm{diag}\left(U_{11}, \hat{U}_{22}\,\mathrm{diag}\,(I_k,\quad U_{33})\right) = \mathrm{diag}(U_{11},\quad U_{22}). \end{aligned}$$

Vidimo da matrica U^*WV ima upravo traženu particiju (3.1), gdje su U i V unitarne blok dijagonalne matrice.

Istovremenom permutacijom stupaca od U_{11} i V_{11}, kao i odgovarajućih stupaca matrica U_{22} i V_{22} možemo postići proizvoljni poredak dijagonalnih elemenata od Γ.

■

Primjer 3.2 Neka je W unitarna matrica (do na apsolutnu grešku elemenata od $5 \cdot 10^{-5}$)

$$W = \begin{bmatrix} 0.6828 & 0.3253 & -0.4635 & 0.4357 & 0.1527 \\ -0.3730 & 0.1332 & -0.3998 & 0.3261 & -0.7596 \\ 0.2608 & 0.2077 & 0.7878 & 0.3926 & -0.3377 \\ -0.0050 & -0.8188 & -0.0164 & 0.5634 & 0.1093 \\ -0.5715 & 0.4037 & 0.0668 & 0.4820 & 0.5231 \end{bmatrix}.$$

Uzmimo za $l = 2$, tada prema teoremu 3.1, postoje unitarne blok dijagonalne matrice U i V

$$U = \begin{bmatrix} 0.3869 & 0.9211 & 0 & 0 & 0 \\ 0.9221 & -0.3869 & 0 & 0 & 0 \\ 0 & 0 & 0.1223 & 0.5353 & 0.8358 \\ 0 & 0 & -0.8060 & -0.4378 & 0.3983 \\ 0 & 0 & 0.5791 & -0.7224 & 0.3780 \end{bmatrix},$$

$$V = \begin{bmatrix} -0.3056 & 0.9522 & 0 & 0 & 0 \\ 0.9522 & 0.3056 & 0 & 0 & 0 \\ 0 & 0 & 0.5677 & 0.4682 & 0.6771 \\ 0 & 0 & -0.4861 & -0.4732 & 0.7347 \\ 0 & 0 & 0.6644 & -0.7462 & -0.0410 \end{bmatrix},$$

takve da je

$$U^*W\ V = \begin{bmatrix} 0.2612 & 0 & -0.9653 & 0 & 0 \\ 0 & 0.8128 & 0 & -0.5825 & 0 \\ 0.9653 & 0 & 0.2612 & 0 & 0 \\ 0 & 0.5825 & 0 & 0.8128 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Ovdje je

$$\Gamma = \begin{bmatrix} 0.2612 & 0 \\ 0 & 0.8128 \end{bmatrix}, \qquad \Sigma = \begin{bmatrix} 0.9653 & 0 \\ 0 & 0.5825 \end{bmatrix}.$$

Korolar 3.3 *Pretpostavimo da su ispunjeni uvjeti teorema 3.1, osim što je W_{11} reda l, $2l > n$. Tada postoje unitarne matrice $U = \mathrm{diag}(U_{11}, U_{22})$ i $V = \mathrm{diag}(V_{11}, V_{22})$, gdje su $U_{11}, V_{11} \in \mathbf{C}^{l\times l}$, takve da je*

$$\begin{bmatrix} U_{11}^* & \\ & U_{22}^* \end{bmatrix} \begin{bmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{bmatrix} \begin{bmatrix} V_{11} & \\ & V_{22} \end{bmatrix} = \begin{array}{c} \begin{bmatrix} \Gamma & 0 & \Sigma \\ 0 & I & 0 \\ -\Sigma & 0 & \Gamma \end{bmatrix} \begin{array}{l} n-l \\ 2l-n \\ n-l \end{array} \\ \begin{array}{ccc} n-l & 2l-n & n-l \end{array} \end{array} \tag{3.5}$$

Pritom vrijedi

$$\begin{aligned} \Gamma &= \mathrm{diag}(\gamma_1, ..., \gamma_{n-l}) \geq 0, \\ \Sigma &= \mathrm{diag}(\sigma_1, ..., \sigma_{n-l}) \geq 0, \\ \Gamma^2 + \Sigma^2 &= I, \end{aligned}$$

a dijagonalni elementi od Γ mogu biti u bilo kojem poretku.

Dokaz. Jer je $2l > n$, mora biti $n - l < l$, odnosno $2(n - l) < n$, pa po teoremu 3.1, postoje unitarne matrice U_{22}, $V_{22} \in \mathbf{C}^{(n-l)\times(n-l)}$ i U_{11}, $V_{11} \in \mathbf{C}^{l\times l}$, takve da je

$$\begin{bmatrix} U_{22}^* & \\ & U_{11}^* \end{bmatrix} \begin{bmatrix} W_{22} & W_{21} \\ W_{12} & W_{11} \end{bmatrix} \begin{bmatrix} V_{22} & \\ & V_{11} \end{bmatrix} = \begin{bmatrix} \Gamma & -\Sigma & 0 \\ \Sigma & \Gamma & 0 \\ 0 & 0 & I \end{bmatrix}.$$

Ovdje su Γ i Σ kvadratne matrice reda $n-l$, a I jedinična matrica reda $2l-n$. Uočimo da smo smjeli primijeniti teorem 3.1 jer je srednja matrica na lijevoj strani zadnje jednakosti permutacijski slična sa W, pa je unitarna. Koristeći istu permutaciju još jedamput dobivamo

$$\begin{bmatrix} U_{11}^* & \\ & U_{22}^* \end{bmatrix} \begin{bmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{bmatrix} \begin{bmatrix} V_{11} & \\ & V_{22} \end{bmatrix} = \begin{bmatrix} \Gamma & 0 & \Sigma \\ 0 & I & 0 \\ -\Sigma & 0 & \Gamma \end{bmatrix},$$

čime je korolar dokazan. ■

Uočimo da u dokazu korolara3.3, zamjena matrica U_{22} i V_{22} s $-U_{22}$ i V_{-22}, respektivno, premješta blok $-\Sigma$ na poziciju $(1,3)$, a blok Σ na poziciju $(3,1)$.

Primjer 3.4 Neka je W (kao i prije, do na red veličine $5 \cdot 10^{-5}$) unitarna matrica

$$W = \begin{bmatrix} -0.4809 & -0.6648 & -0.3399 & 0.1333 & 0.4399 \\ -0.4023 & 0.3854 & -0.1110 & -0.7701 & 0.2903 \\ 0.7755 & -0.2675 & -0.2352 & -0.3647 & 0.3723 \\ 0.0042 & -0.0213 & 0.8135 & 0.0866 & 0.5747 \\ 0.0739 & 0.5809 & -0.3938 & 0.4986 & 0.5033 \end{bmatrix}.$$

Uzmimo za l = 3, tada iz prošlog korolara slijedi da postoje unitarne blok dijagonalne matrice U i V

$$U = \begin{bmatrix} 0.1146 & 0.9153 & 0.3861 & 0 & 0 \\ 0.8437 & -0.2949 & 0.4486 & 0 & 0 \\ 0.5245 & 0.2743 & -0.8060 & 0 & 0 \\ 0 & 0 & 0 & 0.7873 & 0.6166 \\ 0 & 0 & 0 & -0.6166 & 0.7873 \end{bmatrix},$$

$$V = \begin{bmatrix} 0.0440 & -0.1246 & -0.9912 & 0 & 0 \\ 0.3904 & -0.9112 & 0.1319 & 0 & 0 \\ -0.9196 & -0.3928 & 0.0086 & 0 & 0 \\ 0 & 0 & 0 & -0.8597 & 0.5108 \\ 0 & 0 & 0 & 0.5108 & 0.8597 \end{bmatrix},$$

takve da je

$$U^*WV = \begin{bmatrix} 0.2783 & 0 & 0 & 0.9605 & 0 \\ 0 & 0.8731 & 0 & 0 & 0.4875 \\ 0 & 0 & 1 & 0 & 0 \\ -0.9605 & 0 & 0 & 0.2783 & 0 \\ 0 & -0.4875 & 0 & 0 & 0.8731 \end{bmatrix}.$$

Ovdje je

$$\Gamma = \begin{bmatrix} 0.2783 & 0 \\ 0 & 0.8731 \end{bmatrix}, \qquad \Sigma = \begin{bmatrix} 0.9605 & 0 \\ 0 & 0.4875 \end{bmatrix}.$$

3.2 Parovi potprostora

Koristeći CS dekompoziciju možemo se posvetiti problemu određivanja međusobnog odnosa dvaju potprostora iste dimenzije. Slijedeći teorem nam govori o postojanju prirodnih baza za dva potprostora.

Teorem 3.5 *Neka su* X, $Y \in \mathbf{C}^{n\times l}$ *takve da je* $X^*X = I$ *i* $Y^*Y = I$.

(i) Ako je $2l \le n$*, onda postoje unitarne matrice* Q, U *i* V*, takve da je*

$$Q^*XU = \begin{array}{c} \begin{bmatrix} I \\ 0 \\ 0 \end{bmatrix} \\ l \end{array} \begin{array}{c} l \\ l \\ n-2l \\ \end{array} \qquad Q^*YV = \begin{array}{c} \begin{bmatrix} \Gamma \\ \Sigma \\ 0 \end{bmatrix} \\ l \end{array} \begin{array}{c} l \\ l \\ n-2l \\ \end{array}. \tag{3.6}$$

Pritom za

$$\Gamma = \mathrm{diag}(\gamma_1, ..., \gamma_l) \qquad \text{i} \qquad \Sigma = \mathrm{diag}(\sigma_1, ..., \sigma_l)$$

vrijedi

$$\begin{aligned} 0 &\le \gamma_1 \le \cdots \le \gamma_l, \\ \sigma_1 &\ge \cdots \ge \sigma_l \ge 0, \\ \gamma_i^2 &+ \sigma_i^2 = 1, \quad 1 \le i \le l. \end{aligned} \tag{3.7}$$

(ii) Ako je $2l > n$, onda postoje unitarne matrice Q, U i V, takve da je

$$Q^*XU = \begin{array}{c} \left[\begin{array}{cc} I & 0 \\ 0 & I \\ 0 & 0 \end{array}\right] \begin{array}{c} n-l \\ 2l-n \\ n-l \end{array} \\ \begin{array}{cc} n-l & 2l-n \end{array} \end{array} , \tag{3.8}$$

$$Q^*YV = \begin{array}{c} \left[\begin{array}{cc} \Gamma & 0 \\ 0 & I \\ \Sigma & 0 \end{array}\right] \begin{array}{c} n-l \\ 2l-n \\ n-l \end{array} \\ \begin{array}{cc} n-l & 2l-n \end{array} \end{array} . \tag{3.9}$$

Pritom za

$$\Gamma = \mathrm{diag}(\gamma_1, ..., \gamma_{n-l}) \quad i \quad \Sigma = \mathrm{diag}(\sigma_1, ..., \sigma_{n-l})$$

vrijedi

$$\begin{aligned} 0 &\leq \gamma_1 \leq \cdots \leq \gamma_{n-l}, \\ \sigma_1 &\geq \cdots \geq \sigma_{n-l} \geq 0, \\ \gamma_i^2 &+ \sigma_i^2 = 1, \quad 1 \leq i \leq n-l. \end{aligned} \tag{3.10}$$

Dokaz. (i) Neka je $2l \leq n$. Odaberimo $X_\perp$ i $Y_\perp$ tako da matrice $\tilde{X} = [X\,,\, X_\perp]$ i $\tilde{Y} = [Y\,,\, Y_\perp]$ budu unitarne. Prema teoremu 3.1, za unitarnu matricu

$$W = \tilde{X}^*\tilde{Y} = \begin{bmatrix} X^*Y & X^*Y_\perp \\ X_\perp^*Y & X_\perp^*Y_\perp \end{bmatrix}$$

postoje unitarne matrice U_{11}, U_{22}, V_{11} i V_{22} takve da vrijedi relacija (3.1), odnosno

$$\begin{bmatrix} U_{11}^* & \\ & U_{22}^* \end{bmatrix} \begin{bmatrix} X^*Y & X^*Y_\perp \\ X_\perp^*Y & X_\perp^*Y_\perp \end{bmatrix} \begin{bmatrix} V_{11} & \\ & V_{22} \end{bmatrix} = \begin{array}{c} \left[\begin{array}{ccc} \Gamma & -\Sigma & 0 \\ \Sigma & \Gamma & 0 \\ 0 & 0 & I \end{array}\right] \begin{array}{c} l \\ l \\ n-2l \end{array} \\ \begin{array}{ccc} l & l & n-2l \end{array} \end{array}$$

Stavimo li

$$\hat{X} = [XU_{11}\,,\, X_\perp U_{22}], \quad \hat{Y} = [YV_{11}\,,\, Y_\perp V_{22}],$$

imamo

$$\hat{X}^*\hat{Y} \quad = \quad \begin{array}{cc} \begin{bmatrix} \Gamma & -\Sigma & 0 \\ \Sigma & \Gamma & 0 \\ 0 & 0 & I \end{bmatrix} & \begin{array}{c} l \\ l \\ n-2l \end{array} \\ \begin{array}{ccc} l & l & n-2l \end{array} & \end{array} .$$

Ako je potrebno, možemo permutirati stupce matrica U_{ii} i V_{ii} tako da vrijedi (3.7). Uz oznake $Q = \hat{X}$, $U = U_{11}$ i $V = V_{11}$, odmah slijedi

$$\begin{aligned} Q^*XU &= \begin{bmatrix} U^*X^* \\ U_{22}^*X_\perp^* \end{bmatrix} XU = \begin{array}{cc} \begin{bmatrix} I \\ 0 \end{bmatrix} & \begin{array}{c} l \\ n-l \end{array} \end{array} , \\ Q^*YV &= \begin{bmatrix} U^*X^* \\ U_{22}^*X_\perp^* \end{bmatrix} YV = \begin{array}{cc} \begin{bmatrix} \Gamma \\ \Sigma \\ 0 \end{bmatrix} & \begin{array}{c} l \\ l \\ n-2l \end{array} \end{array} , \end{aligned}$$

čime je tvrdnja (i) dokazana.

(ii) U ovom slučaju treba ponoviti dokaz tvrdnje (i) uz korištenje relacije (3.5). Također umjesto matrice U_{22} treba koristiti $-U_{22}$. ■

Teorem 3.5 ima slijedeće geometrijsko značenje. Neka su $\mathcal{X}$ i $\mathcal{Y}$ l–dimenzionalni potprostori od $\mathbf{C}^n$. Ako je $2l \leq n$, onda postoji unitarna matrica $Q = [q_1, \ldots q_n]$, takva da stupci matrica

$$Q \begin{bmatrix} I \\ 0 \\ 0 \end{bmatrix} \qquad \text{i} \qquad Q \begin{bmatrix} \Gamma \\ \Sigma \\ 0 \end{bmatrix} \tag{3.11}$$

čine ortonormirane baze od $\mathcal{X}$ i $\mathcal{Y}$, respektivno. Dakle, $\mathcal{X} = \operatorname{span}\{q_1, \ldots, q_l\}$ i $\mathcal{Y} = \operatorname{span}\{\gamma_1 q_1 + \sigma_1 q_{l+1}, \ldots, \gamma_l q_l + \sigma_l q_{2l}\}$. Prostor razapet onim stupcima za koje je $\gamma_i = 0$ je potprostor od $\mathcal{Y}$ ortogonalan na $\mathcal{X}$.
Za $2l > n$ postoji unitarna matrica $Q = [q_1, \ldots q_n]$, takva da stupci matrica

$$Q \begin{bmatrix} I & 0 \\ 0 & I \\ 0 & 0 \end{bmatrix} \qquad \text{i} \qquad Q \begin{bmatrix} \Gamma & 0 \\ 0 & I \\ \Sigma & 0 \end{bmatrix} \tag{3.12}$$

čine ortonormirane baze od $\mathcal{X}$ i $\mathcal{Y}$, respektivno. Dakle je $\mathcal{X} = \operatorname{span}\{q_1, \ldots, q_l\}$ i $\mathcal{Y} = \operatorname{span}\{\gamma_1 q_1 + \sigma_1 q_{l+1}, \ldots, \gamma_{n-l} q_{n-l} + \sigma_{n-l} q_n, q_{n-l+1}, \ldots, q_l\}$. Prostor razapet onim stupcima za koje je $\gamma_i = 0$ je potprostor od $\mathcal{Y}$ ortogonalan na $\mathcal{X}$, dok je $\mathrm{span}\{q_{n-l+1}, \ldots, q_l\}$ najmanji mogući presjek potprostora $\mathcal{X}$ i $\mathcal{Y}$.

Uočimo da transformacijom

$$Q \longmapsto Q \begin{bmatrix} \Gamma & \Sigma & 0 \\ \Sigma & -\Gamma & 0 \\ 0 & 0 & I \end{bmatrix}, \quad 2l \leq n$$
$$Q \longmapsto Q \begin{bmatrix} \Gamma & 0 & \Sigma \\ 0 & I & 0 \\ \Sigma & 0 & -\Gamma \end{bmatrix}, \quad 2l \leq n$$

kanonske baze zamijenjuju potprostore. Dakle brojevi σ_i i γ_i ne ovise o redoslijedu potprostora u iskazu teorema 3.5. Kako brojevi σ_i i γ_i zadovoljavaju $\sigma_i^2 + \gamma_i^2 = 1$, možemo ih smatrati sinusima i kosinusima kutova između baza. Štoviše, $\mathcal{X} = \mathcal{Y}$ onda i samo onda ako je $\Sigma = 0$. To znači da veličina elemenata matrice Σ mjeri razliku između $\mathcal{X}$ i $\mathcal{Y}$. Kako Σ ovisi samo o potprostorima $\mathcal{X}$ i $\mathcal{Y}$ možemo definirati pojam kanonskih kutova između dva potprostora.

Definicija 3.6 *Neka su $\mathcal{X}$ i $\mathcal{Y}$ potprostori iste dimenzije i neka je Σ kao u teoremu 3.5. Dijagonalne elemente matrice*

$$\Theta(\mathcal{X},\ \mathcal{Y}) = \arcsin \Sigma,$$

zovemo kanonski kutovi *između $\mathcal{X}$ i $\mathcal{Y}$.*

Slijedeći korolar govori nam kako se računaju kanonski kutovi.

Korolar 3.7 *Neka su $\mathcal{X}$ i $\mathcal{Y}$ potprostori od $\mathbf{C}^n$ jednake dimenzije. Neka stupci matrica $X_\perp$ i Y čine ortonormirane baze za $\mathcal{X}^\perp$ i $\mathcal{Y}$, respektivno. Tada su singularne vrijednosti matrice $X_\perp^* Y$ sinusi kanonskih kutova između $\mathcal{X}$ i $\mathcal{Y}$.*

Dokaz. Neka su $\tilde{X} = [X\ ,\ X_\perp]$ i $\tilde{Y} = [Y\ ,\ Y_\perp]$ unitarne matrice takve da je $\mathcal{R}(X) = \mathcal{X}$ i $\mathcal{R}(Y) = \mathcal{Y}$. Slično kao i u dokazu teorema 3.5, možemo na matricu

$$W = \tilde{X}^* \tilde{Y} = \begin{bmatrix} X^* Y & X^* Y_\perp \\ X_\perp^* Y & X_\perp^* Y_\perp \end{bmatrix}$$

primijeniti teorem 3.1. Dakle, postoje unitarne matrice U_{11}, U_{22}, V_{11} i V_{22} takve da vrijedi relacija (3.1). To znači da je

$$\begin{aligned} X_\perp^* Y &= U_{22} \begin{bmatrix} \Sigma \\ 0 \end{bmatrix} V_{11}^*, \quad \text{ako je } 2l \leq n, \\ X_\perp^* Y &= U_{22}\, [\Sigma\, , \, 0]\, V_{11}^*, \quad \text{ako je } 2l > n. \end{aligned}$$

Iz posljednje jednakosti vidimo da su singularne vrijednosti matrice $X_\perp^* Y$ upravo dijagonalni elementi matrice Σ. ∎

3.3 Ortogonalni projektori

U ovoj točki ćemo upoznati ortogonalne projektore i njihova osnovna svojstva. Pokazat ćemo kako se ortogonalni projektor može definirati pomoću ortonormiranih baza i pomoću generaliziranog inverza.

Neka je $\mathcal{X}$ potprostor od $\mathbf{C}^n$ i neka stupci od $Q_{\mathcal{X}}$ čine ortonormiranu bazu za $\mathcal{X}$. To znači, ako je $\mathcal{X}$ k – dimenzionalan, $Q_{\mathcal{X}} \in \mathbf{C}^{n\times k}$ je ortonormirana matrica. Tada matricu definiranu s

$$P_{\mathcal{X}} = Q_{\mathcal{X}} Q_{\mathcal{X}}^*$$

nazivamo *ortogonalni projektor na* $\mathcal{X}$. Iako se čini da $P_{\mathcal{X}}$ ovisi o ortonormiranoj bazi, to nije tako. Neka stupci matrice $\hat{Q}_{\mathcal{X}}$ također čine ortonormiranu bazu za $\mathcal{X}$. Onda postoji regularna matrica U takva da je $\hat{Q}_{\mathcal{X}} = Q_{\mathcal{X}} U$. Budući da je

$$I = \hat{Q}_{\mathcal{X}}^* \hat{Q}_{\mathcal{X}} = U^* Q_{\mathcal{X}}^* Q_{\mathcal{X}} U = U^* U,$$

vidimo da je U unitarna matrica. No sad je

$$\hat{Q}_{\mathcal{X}} \hat{Q}_{\mathcal{X}}^* = Q_{\mathcal{X}} U U^* Q_{\mathcal{X}}^* = Q_{\mathcal{X}} Q_{\mathcal{X}}^* = P_{\mathcal{X}}.$$

Matrica $P_{\mathcal{X}}$ je očito hermitska, a zbog

$$P_{\mathcal{X}}^2 = Q_{\mathcal{X}} Q_{\mathcal{X}}^* Q_{\mathcal{X}} Q_{\mathcal{X}}^* = Q_{\mathcal{X}} Q_{\mathcal{X}}^* = P_{\mathcal{X}},$$

ona je i idempotentna.

Samo svojstvo idempotentnosti ima za posljedicu da je $P_{\mathcal{X}}x = x$ za $x \in \mathcal{X}$. Zaista, kako je $\mathcal{R}(P_{\mathcal{X}}) = \mathcal{X}$, za svaki $x \in \mathcal{X}$ mora postojati vektor $w \in \mathbf{C}^n$, takav da je $x = P_{\mathcal{X}}w$. Sada je $P_{\mathcal{X}}x = P_{\mathcal{X}}^2 w = P_{\mathcal{X}}w = x$.

S druge strane, hermitičnost matrice $P_{\mathcal{X}}$ povlači $P_{\mathcal{X}}y = 0$ za $y \in \mathcal{X}^{\perp}$. Doista, za proizvoljno $z \in \mathbf{C}^n$ i $y \in \mathcal{X}^{\perp}$, imamo $P_{\mathcal{X}}z \in \mathcal{X}$ i $0 = (\, y^* P_{\mathcal{X}} z\,)^* = z^*\, P_{\mathcal{X}}\, y$. Stoga je $P_{\mathcal{X}}y = 0$.

Dakle svojstvo $P^2 = P$ osigurava da $P_{\mathcal{X}}$ djeluje na vektore iz $\mathcal{X}$ kao identiteta, dok svojstvo $P^* = P$ osigurava da se $\mathcal{X}^{\perp}$ preslikava u $\{0\}$. Od linearnih transformacija, takva svojstva ima samo ortogonalna projekcija.

Jer je $\mathcal{X} \oplus \mathcal{X}^{\perp}$ ortogonalna dekompozicija od $\mathbf{C}^n$, za svaki vektor $z \in \mathbf{C}^n$ postoji jedinstveni rastav

$$z = x + y, \qquad x \in \mathcal{X},\;\; y \in \mathcal{X}^{\perp}.$$

Primijenimo li ortogonalni projektor $P_{\mathcal{X}}$ lijevu i desnu stranu jednakosti, odmah dobijemo $x = P_{\mathcal{X}}z$ i $y = z - x = (I - P_{\mathcal{X}})z$. Lako je provjeriti da je $I - P_{\mathcal{X}}$ idempotentna i hermitska matrica, pa zaključujemo da je to ortogonalni projektor na $\mathcal{X}^{\perp}$.

Sada ćemo ukratkoopisati i "neortogonalne" projektore. Neka je $\mathbf{C}^n$ direktna (ne nužno ortogonalna) suma potprostora $\mathcal{X}_1$ i $\mathcal{X}_2$, $\mathbf{C}^n = \mathcal{X}_1 \dot{+} \mathcal{X}_2$. Tada se svaki $x \in \mathbf{C}^n$ na jedinstven način može prikazati kao suma $x = x_1 + x_2$, pri čemu su $x_1 \in \mathcal{X}_1$ i $x_2 \in \mathcal{X}_2$. To znači da relacije $x_i = \mathcal{P}_i x$ $i = 1, 2$ definiraju linearne operatore $\mathcal{P}_1$ i $\mathcal{P}_2$. Uzmimo bazu od $\mathcal{X}_1$ pa ju nadopunimo do baze od $\mathbf{C}^n$, ali tako da preostali vektori čine bazu za $\mathcal{X}_2$. Pripadna matrica P_1 (P_2) operatora $\mathcal{P}_1$ ($\mathcal{P}_2$) je dijagonalna s prvih (zadnjih) dim $(\mathcal{X}_1)$ (dim $(\mathcal{X}_2)$) jedinica na dijagonali, a s ostalim dijagonalnim elementima jednakim nuli. Ove matrice ne ovise o izboru baza u $\mathcal{X}_1$ i $\mathcal{X}_2$. Nadalje matrice P_1 i P_2 imaju slijedeća svojstva

$$\begin{aligned} &P_1 + P_2 = I \\ &P_1^2 = P_1, \quad P_2^2 = P_2, \quad P_1\, P_2 = P_2\, P_1 = 0. \end{aligned} \tag{3.13}$$

Vrijedi i obrat: matrice koje zadovoljavaju (3.13) na jedinstven način određuju dekompoziciju prostora $\mathbf{C}^n$ na direktnu sumu $\mathbf{C}^n = \mathcal{X}_1 \oplus \mathcal{X}_2$ (vidi [6, str. 322-323]). Tom direktnom sumom su zadani linearni operatori $\mathcal{P}_1$ i $\mathcal{P}_2$ koji opet opisanim izborom baza dovode do matrica P_1 i P_2. U opisanoj situaciji matricu P_1 (P_2) zovemo *projektor prostora* $\mathbf{C}^n$ *na* $\mathcal{X}_1$ ($\mathcal{X}_2$) *u smjeru*[1] $\mathcal{X}_2$ ($\mathcal{X}_1$).

Promotrimo sada slučaj kad imamo na raspolaganju samo jednu idempotentnu matricu P.

Propozicija 3.8 *Matrica P je projektor na $\mathcal{X}_1 = \mathcal{R}(P)$ u smjeru $\mathcal{X}_2 = \mathcal{N}(P)$ onda i samo onda ako je P idempotentna matrica, tj. ako je $P^2 = P$.*

Dokaz. Ako je P projektor na $\mathcal{X}_1$ u smjeru $\mathcal{X}_2$, onda za bilo koji vektor rastavljen na $x = x_1 + x_2$, gdje je $x_1 \in \mathcal{X}_1$ i $x_2 \in \mathcal{X}_2$ imamo $P\,x = x_1$. Jer vrijedi rastav $x_1 = x_1 + 0$ imamo $P\,x_1 = x_1$. Stoga je

$$P^2\,x = P\,x_1 = x_1 = Px \quad \text{za svako } x,$$

a to znači $P^2 = P$.

Obratno, neka je P idempotentna matrica i $\mathcal{X}_1 = \mathcal{R}(P)$ i $\mathcal{X}_2 = \mathcal{N}(P)$. Pokažimo prvo da je $\mathbf{C}^n$ direktna suma potprostora $\mathcal{X}_1$ i $\mathcal{X}_2$. Jer je $\operatorname{rang}(P) + \operatorname{defekt}(P) = n$, dovoljno je pokazati da je $\mathcal{X}_1 \bigcap \mathcal{X}_2 = \{0\}$. Zaista, kad bi postojao $0 \neq z \in \mathcal{X}_1 \bigcap \mathcal{X}_2$, imali bi s jedne strane $Pz = 0$, a s druge strane $Pw = z$ za neko $w \neq 0$. Dakle bilo bi $P^2w = 0$, pa bi zbog idempotentnosti matrice P bilo $z = Pw = P^2w = 0$, što se protivi pretpostavci o vektoru z. Dakle je $C^n = \mathcal{X}_1 \dot{+} \mathcal{X}_2$. Sada pokažimo da je P projektor na $\mathcal{X}_1$ u smjeru $\mathcal{X}_2$, tj. $Px_1 = x_1$ za proizvoljno $x_1 \in \mathcal{X}_1$ i $Px_2 = 0$ za proizvoljno $x_2 \in \mathcal{X}_2$. Za $x_1 \in \mathcal{X}_1$ postoji vektor x takav da je $P\,x = x_1$. Sada imamo $x_1 = P\,x = P^2\,x = P\,x_1$. Druga tvrdnja $P\,x_2 = 0$ za $x_2 \in \mathcal{X}_2$ slijedi direktno iz definicije potprostora $\mathcal{X}_2$. ■

Slijedeća propozicija nas vraća na pojam ortogonalog projektora.

Propozicija 3.9 Ako je P hermitska idempotentna matrica, tada je P ortogonalni projektor.

Dokaz. Zaista, iz propozicije 3.8 slijedi da je P projektor na $\mathcal{X}_1 = \mathcal{R}(P)$ duž $\mathcal{X}_2 = \mathcal{N}(P)$. Jer je $P = P^*$ i $P\,y = 0$ za $y \in \mathcal{X}_2$, imamo

$$y^*P\,x = (\,P\,y\,)^*x = 0, \qquad x \in \mathbf{C}^n,\ \ y \in \mathcal{X}_2,$$

[1] još se kaže *duž* potprostora $\mathcal{X}_2$ ($\mathcal{X}_1$)

pa je $y \perp \mathcal{R}(P) = \mathcal{X}_1$. Zbog proizvoljnosti od y imamo $\mathcal{X}_1 \perp \mathcal{X}_2$. ■

U slijedećem teoremu ćemo prikazati kako se ortogonalni projektori na $\mathcal{R}(A)$, $\mathcal{R}(A^*)$ i $\mathcal{N}(A)$ mogu izraziti pomoću generaliziranog inverza.

Teorem 3.10 *Za proizvoljnu matricu A vrijedi:*

(i) $A\,A^\dagger$ *je ortogonalni projektor na* $\mathcal{R}(A)$

(ii) $A^\dagger A$ *je ortogonalni projektor na* $\mathcal{R}(A^*)$

(iii) $I - A^\dagger A$ *je ortogonalni projektor na* $\mathcal{N}(A)$.

Dokaz. (i) Iz Penroseovih uvjeta vidimo da je $A\,A^\dagger$ hermitska idempotentna matrica, a to znači da predstavlja ortogonalni projektor na

$$\mathcal{R}(A\,A^\dagger) \subseteq \mathcal{R}(A).$$

Prema svojstvu *4.* teorema 2.13 vidimo da je $\operatorname{rang}(A\,A^\dagger) = \operatorname{rang}(A)$. To znači da je $\mathcal{R}(A\,A^\dagger) = \mathcal{R}(A)$, pa je prva tvrdnja dokazana.

(ii) Iz drugog Penroseovog uvjeta slijedi idempotentnost

$$A^\dagger A\; A^\dagger A = A^\dagger A,$$

a iz četvrtog uvjeta hermitičnost matrice $A^\dagger A$. Preostaje pokazati $\mathcal{R}(A^\dagger A) = \mathcal{R}(A^*)$. Iz šeste tvrdnje teorema 2.13 slijedi

$$A^*(A^*)^\dagger = (A^*A)^\dagger\, A^*A = A^\dagger A,$$

pa koristeći opet 4. tvrdnju teorema 2.13 dobivamo $\mathcal{R}(A^\dagger A) = \mathcal{R}(A^*(A^*)^\dagger) = \mathcal{R}(A^*)$.

(iii) Lako se provjeri da je $I - P$ ortogonalni projektor ako je to P. Pritom je $\mathcal{R}(I-P) = \mathcal{N}(P)$, pa tvrdnja (iii) slijedi neposredno iz tvrdnje (ii). ■

3.4 Dekompozicija para ortogonalnih projektora

U ovoj točki ćemo dokazati vrlo koristan Wedinov rezultat [18] o dekompoziciji para ortogonalnih projektora. Važnost tog rezultata proizlazi iz činjenice da pripadni potprostori $\mathcal{X}$ i $\mathcal{Y}$ ne moraju biti iste dimenzije.

Teorem 3.11 *Neka su $\mathcal{X}$ i $\mathcal{Y}$ potprostori od $\mathbf{C}^n$. Pretpostavimo da za pripadne ortogonalne projektore vrijedi*

(i) $\operatorname{rang}(P_{\mathcal{X}}P_{\mathcal{Y}}) = k + r$

(ii) $P_{\mathcal{X}}P_{\mathcal{Y}}$ *ima* k *singularnih vrijednosti jednakih* 1.

Tada postoji unitarna matrica $Z \in \mathbf{C}^{n\times n}$, takva da su matrice $P_1 = Z^*P_{\mathcal{X}}Z$ i $P_2 = Z^*P_{\mathcal{Y}}Z$ oblika

$$P_1 = \begin{bmatrix} I_k & & \\ & \bigoplus_{i=1}^{r} \Phi_i & \\ & & D_1 \end{bmatrix}, \qquad P_2 = \begin{bmatrix} I_k & & \\ & \bigoplus_{i=1}^{r} \Psi_{\theta_i} & \\ & & D_2 \end{bmatrix}, \tag{3.14}$$

pri čemu je I_k jedinična matrica reda k,

$$\Phi_i = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix}, \qquad \Psi_{\theta_i} = \begin{bmatrix} \cos\theta_i \\ \sin\theta_i \end{bmatrix} \begin{bmatrix} \cos\theta_i & \sin\theta_i \end{bmatrix}, \tag{3.15}$$

a D_1, $D_2 \in \mathbf{C}^{(n-k-2r)\times(n-k-2r)}$ su dijagonalne matrice s dijagonalnim elementima iz skupa $\{0,1\}$, za koje vrijedi

$$D_1 D_2 = 0. \tag{3.16}$$

Pritom za kutove θ_i vrijedi: $\pi/2 > \theta_1 \geq \theta_2 \geq \cdots \geq \theta_r > 0$,

Dokaz. Uočimo da su P_1 i P_2 zapisi (tj. matrice) operatora projekcije određenih polaznim matricama $P_{\mathcal{X}}$ i $P_{\mathcal{Y}}$.

Pretpostavimo da su potprostori $\mathcal{X}$ i $\mathcal{Y}$ dimenzija m_1 i m_2, respektivno. Neka stupci matrica $X \in \mathbf{C}^{n\times m_1}$ i $Y \in \mathbf{C}^{n\times m_2}$ čine baze za $\mathcal{X}$ i $\mathcal{Y}$. Načinimo skraćenu QR-dekompoziciju matrica X i Y,

$$X = Q_{\mathcal{X}}R_{\mathcal{X}} \qquad \text{i} \qquad Y = Q_{\mathcal{Y}}R_{\mathcal{Y}}.$$

Matrice $Q_{\mathcal{X}}$ i $Q_{\mathcal{Y}}$ su istih dimenzija kao X i Y, ali imaju ortonormirane stupce. Tako smo došli do ortonormiranih baza za $\mathcal{X}$ i $\mathcal{Y}$. Stoga su

$$P_{\mathcal{X}} = Q_{\mathcal{X}} Q_{\mathcal{X}}^*, \qquad P_{\mathcal{Y}} = Q_{\mathcal{Y}} Q_{\mathcal{Y}}^* \tag{3.17}$$

upravo polazni ortogonalni projektori. Načinit ćemo skraćenu singularnu dekompoziciju matrice

$$P_{\mathcal{X}} P_{\mathcal{Y}} = Q_{\mathcal{X}} (Q_{\mathcal{X}}^* Q_{\mathcal{Y}}) Q_{\mathcal{Y}}^*. \tag{3.18}$$

U tu svrhu načinimo singularnu dekompoziciju matrice $Q_{\mathcal{X}}^* Q_{\mathcal{Y}} \in \mathbf{C}^{m_1 \times m_2}$,

$$Q_{\mathcal{X}}^* Q_{\mathcal{Y}} = [S\,,\,S'] \begin{bmatrix} \Gamma & 0 \\ 0 & 0 \end{bmatrix} [T\,,\,T']^*\,. \tag{3.19}$$

Ovdje je Γ kvadratna dijagonalna matrica reda $k + r$ s pozitivnim dijagonalnim elementima, a S i T imaju broj stupaca koji je jednak redu od Γ. Uočimo da smo ovdje iskoristili pretpostavku (i) teorema. Dakle je

$$Q_{\mathcal{X}}^* Q_{\mathcal{Y}} = S\ \Gamma\ T^*, \tag{3.20}$$

pa ako definiramo

$$U = Q_{\mathcal{X}} S \qquad \text{i} \qquad V = Q_{\mathcal{Y}} T,$$

dolazimo do skraćene singularne dekompozicije matrice $P_{\mathcal{X}} P_{\mathcal{Y}}$,

$$P_{\mathcal{X}} P_{\mathcal{Y}} = U\ \Gamma\ V^*. \tag{3.21}$$

Neka je

$$\Gamma = diag(\gamma_1, ..., \gamma_{k+r}) \qquad \text{i} \qquad \gamma_{\max} = \max\{\gamma_1, ..., \gamma_{k+r}\}.$$

Tada vrijedi

$$\gamma_{\max} = \|P_{\mathcal{X}} P_{\mathcal{Y}}\|_2 \leq \|P_{\mathcal{X}}\|_2\ \|P_{\mathcal{Y}}\|_2 = 1,$$

što znači da su sve singularne vrijednosti od $P_{\mathcal{X}} P_{\mathcal{Y}}$ manje ili jednake 1. Permutirajmo istovremeno stupce od U i V, tako da (3.21) vrijedi uz uvjet

$$0 < \gamma_1 \leq \gamma_2 \leq ... \leq \gamma_r < \gamma_{r+1} = ... = \gamma_{r+k} = 1. \tag{3.22}$$

Ovdje smo iskoristili pretpostavku (ii) teorema o broju singularnih vrijednosti produkta $P_{\mathcal{X}} P_{\mathcal{Y}}$ jednakih jedan. Ako označimo s $u_1, \ldots, u_{r+k}$ i $v_1, \ldots, v_{r+k}$ stupce matrica

U i V, respektivno, tada relacija (3.21) daje

$$\begin{aligned} P_{\mathcal{X}} P_{\mathcal{Y}} v_i &= \gamma_i \; u_i \\ (P_{\mathcal{X}} P_{\mathcal{Y}})^* u_i = P_{\mathcal{Y}} P_{\mathcal{X}} u_i &= \gamma_i \; v_i . \end{aligned}$$

S obzirom da su svi u_i iz $\mathcal{X}$ i svi v_i iz $\mathcal{Y}$, odmah dobivamo za $1 \leq i \leq k+r$,

$$\begin{aligned} P_{\mathcal{X}} v_i &= \gamma_i \; u_i \\ P_{\mathcal{Y}} u_i &= \gamma_i \; v_i \end{aligned} . \tag{3.23}$$

Brojevi

$$\theta_i = \arccos \gamma_i, \quad 1 \leq i \leq k+r$$

zadovoljavaju poredak naveden u iskazu teorema (vidi i relaciju (3.28). Pokažimo da (3.23) povlači relaciju

$$u_i = v_i \;\text{ ako i samo ako je }\; \gamma_i = 1. \tag{3.24}$$

Zaista, $u_i = v_i$ i relacija (3.23) daju $P_{\mathcal{X}} u_i = \gamma_i u_i$, a kako je $P_{\mathcal{X}} u_i = u_i$, imamo $u_i = \gamma_i u_i$, $u_i \neq 0$. Dakle je $\gamma_i = 1$. Obratno, ako je $\gamma_i = 1$, tada (3.23) povlači $u_i = P_{\mathcal{X}} v_i$, pa je

$$v_i = P_{\mathcal{X}} v_i + (I - P_{\mathcal{X}})\, v_i = u_i + (I - P_{\mathcal{X}})\, v_i \, .$$

S obzirom da je $(I - P_{\mathcal{X}})\, v_i \perp P_{\mathcal{X}} v_i = u_i$ imamo,

$$1 = \|v_i\|^2 = \|u_i\|^2 + \|(I - P_{\mathcal{X}})\, v_i\|^2 = 1 + \|(I - P_{\mathcal{X}})\, v_i\|^2,$$

odnosno $0 = (I - P_{\mathcal{X}})\, v_i = v_i - u_i$, što je i trebalo pokazati.

Uvažimo li (3.22) i (3.24) dobijemo

$$\begin{aligned} u_i &= v_i, \quad r+1 \leq i \leq r+k \\ u_i &\neq v_i, \quad 1 \leq i \leq r. \end{aligned}$$

Promotrimo niz dvodimenzionalnih potprostora $\mathcal{W}_i = \text{span}\{u_i \, , \, v_i\}$, $1 \leq i \leq r$. Svaki $\mathcal{W}_i$ ima slijedeća dva važna svojstva.

(a) invarijantan je u odnosu na $P_{\mathcal{X}}$ i $P_{\mathcal{Y}}$

(b) $\mathcal{W}_i \perp \mathcal{W}_j$ za sve $1 \leq j \leq r$, $j \neq i$.

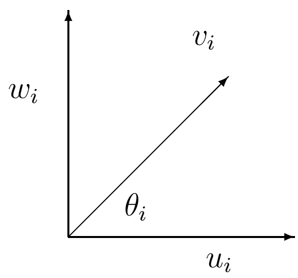
Prvo svojstvo je direktna posljedica jednakosti (3.23). Drugo svojstvo slijedi iz ortonormiranosti skupova vektora $\{ u_1,\ldots, u_r \}$ i $\{ v_1,\ldots, v_r \}$, te zbog relacije

$$v_i^* u_j = (P_{\mathcal{Y}} v_i)^* u_j = v_i^* P_{\mathcal{Y}} u_j \overset{(3.23)}{=} \gamma_j v_i^* v_j = \begin{cases} 0, & za \;\; i \neq j \\ \gamma_j = \cos\theta_j, & za \;\; i = j \end{cases} .$$

Primijetimo da je $\mathcal{W}_i$ okomit i na vektore $u_i = v_i$, $r+1 \leq i \leq r+k$. Sada ćemo konstruirati ortonormiranu bazu za svaki $\mathcal{W}_i$. Definirajmo vektore w_i relacijom

$$v_i = u_i \cos\theta_i + w_i \sin\theta_i, \quad 1 \leq i \leq r. \tag{3.25}$$

Jasno, $w_i \in \mathcal{W}_i$ jer je razapet vektorima u_i i v_i. Također iz slike koja leži u (u_i, v_i) ravnini



zaključujemo da je $\|w_i\| = 1$ i $w_i \perp u_i$, $1 \leq i \leq r$. Ove tvrdnje ćemo sada dokazati i analitički. Dokažimo prvo ortogonalnost. Zbog relacije (3.23) imamo

$$\begin{aligned} \gamma_i\, u_i &= P_{\mathcal{X}} v_i = P_{\mathcal{X}} u_i \cos\theta_i + P_{\mathcal{X}} w_i \sin\theta_i = \gamma_i P_{\mathcal{X}} u_i + P_{\mathcal{X}} w_i \sin\theta_i \\ &= \gamma_i u_i + P_{\mathcal{X}} w_i \sin\theta_i. \end{aligned}$$

Kako je $\cos\theta_i = \gamma_i < 1$, imamo $\sin\theta_i > 0$, pa je $P_{\mathcal{X}} w_i = 0$. To znači, $w_i \perp \mathcal{X}$, $1 \leq i \leq r$, odnosno

$$w_i^* u_j = 0 \;\; \text{za sve} \;\; 1 \leq j \leq r+k, \; 1 \leq i \leq r. \tag{3.26}$$

Iz relacije (3.25) također slijedi $\|w_i\| = 1$. Naime, zbog ortogonalnosti vektora u_i i w_i imamo

$$1 = \|v_i\|^2 = \|u_i\|^2 \cos^2\theta_i + \|w_i\|^2 \sin^2\theta_i = \cos^2\theta_i + \|w_i\|^2 \sin^2\theta_i$$

pa je zbog $\sin\theta_i > 0$, $\|w_i\| = \sin\theta_i / \sin\theta_i = 1$.

Množenjem relacije (3.25) s $P_{\mathcal{Y}}$ i korištenjem relacije (3.23) dobivamo na sličan način i

$$P_{\mathcal{Y}} w_i = \sin\theta_i v_i. \tag{3.27}$$

Pokazali smo da je skup vektora $\{u_{r+1}, ..., u_{r+k}, u_1, w_1, ..., u_r, w_r\}$ ortonormiran. Definirajmo ortonormiranu matricu

$$Z_1 = [u_{r+1} \ldots, u_{r+k}, u_1, w_1, \ldots, u_r, w_r] \in \mathbf{C}^{n\times(k+2r)}.$$

Iz konstrukcije slijedi $\mathcal{R}(Z_1)$ =span$\{\mathcal{R}(U), \mathcal{R}(V)\}$. Pokazat ćemo da je

$$Z_1^* P_{\mathcal{X}} Z_1 = \begin{bmatrix} I_k & \\ & \bigoplus_{i=1}^r \Phi_i \end{bmatrix}, \qquad Z_1^* P_{\mathcal{Y}} Z_1 = \begin{bmatrix} I_k & \\ & \bigoplus_{i=1}^r \Psi_{\theta_i} \end{bmatrix}.$$

Zaista, jer je $P_{\mathcal{X}} w_i = 0$, $P_{\mathcal{X}} u_i = u_i$, $1 \le i \le r$, imamo

$$\begin{aligned} P_{\mathcal{X}} Z_1 &= [u_{r+1}, \ldots, u_{r+k}, u_1, 0, \ldots, u_r, 0] \\ &= [u_{r+1}, \ldots, u_{r+k}, u_1, w_1, \ldots, u_r, w_r] \begin{bmatrix} I_k & & & \\ & \Phi_1 & & \\ & & \ddots & \\ & & & \Phi_r \end{bmatrix} \\ &= Z_1 \begin{bmatrix} I_k & \\ & \bigoplus_{i=1}^r \Phi_i \end{bmatrix}, \end{aligned}$$

pa množenjem s lijeva sa Z_1^* dobivamo prvu tvrdnju.

Da bi dokazali drugu tvrdnju uočimo da je

$$v_i = u_i \cos\theta_i + w_i \sin\theta_i = [u_i\,,\ w_i] \begin{bmatrix} \cos\theta_i \\ \sin\theta_i \end{bmatrix},$$

pa je

$$\begin{aligned} [v_i \cos\theta_i\,,\ v_i \sin\theta_i] &= [u_i\,,\ w_i] \begin{bmatrix} \cos\theta_i \\ \sin\theta_i \end{bmatrix} [\cos\theta_i\,,\ \sin\theta_i] \\ &= [u_i\,,\ w_i] \Psi_{\theta_i}\,. \end{aligned}$$

Koristeći (3.27) i zadnju relaciju imamo

$$\begin{aligned} P_{\mathcal{Y}} Z_1 &= [u_{r+1}, \dots, u_{r+k}, \cos\theta_1 v_1\,,\, \sin\theta_1 v_1\,, \dots\,,\, \cos\theta_r v_r\,,\, \sin\theta_r v_r] \\ &= [u_{r+1}, \dots, u_{r+k}, [u_1, w_1]\Psi_{\theta_1}, \dots, [u_r, w_r]\Psi_{\theta_r}] \\ &= Z_1 \begin{bmatrix} I_k & \\ & \bigoplus_{i=1}^{r} \Psi_{\theta_i} \end{bmatrix}, \end{aligned}$$

pa množenjem sa Z_1^* s lijeva, dobivamo drugu tvrdnju.

Preostaje pokazati postojanje matrica D_1 i D_2. Iz dosadašnjeg razmatranja zaključujemo da k i r moraju biti takovi da je $k + 2r \leq n$. Ako je $k + 2r = n$, onda je $Z = Z_1$, D_1 i D_2 ne postoje i teorem je dokazan. Pretpostavimo zato $k + 2r < n$. Iz relacije (3.19) vidimo da je $k + r \leq \min\{m_1\,,\, m_2\}$.

Ako je $k + r < m_1$, tada se u singularnoj dekompoziciji matrice $Q_{\mathcal{X}}^* Q_{\mathcal{Y}}$ pojavljuje ortonormirana matrice $S' \in \mathbf{C}^{m_1 \times (m_1 - k - r)}$. Definirajmo ortonormiranu matricu $Z_2' = Q_{\mathcal{X}} S'$. Pokažimo da je $[Z_1\,,\, Z_2']$ ortonormirana. Za to je dovoljno pokazati $Z_1^* Z_2' = 0$, a za to je opet dovoljno pokazati $U^* Z_2' = 0$ i $V^* Z_2' = 0$. Korištenjem relacija (3.19) i (3.20) dobivamo

$$\begin{aligned} U^* Z_2' &= S^* Q_{\mathcal{X}}^* Q_{\mathcal{X}} S' = S^* S' = 0, \\ V^* Z_2' &= T^* Q_{\mathcal{Y}}^* Q_{\mathcal{X}} S' = T^* (T \Gamma S^*) S' = \Gamma S^* S' = 0. \end{aligned}$$

Također, zbog

$$\begin{aligned} P_{\mathcal{X}} Z_2' &= Q_{\mathcal{X}} Q_{\mathcal{X}}^* Q_{\mathcal{X}} S' = Q_{\mathcal{X}} S' = Z_2', \\ P_{\mathcal{Y}} Z_2' &= Q_{\mathcal{Y}} Q_{\mathcal{Y}}^* Q_{\mathcal{X}} S' = Q_{\mathcal{Y}} (S \Gamma T^*)^* S' = Q_{\mathcal{Y}} T \Gamma S^* S' = 0, \end{aligned}$$

odmah slijedi $[Z_1\,,\, Z_2']^* P_{\mathcal{X}}\, Z_2' = [0\,,\, I_{m_1-k-r}]^*$ i $[Z_1\,,\, Z_2']^* P_{\mathcal{Y}}\, Z_2' = 0$. Zbog hermitičnosti matrica $P_{\mathcal{X}}$ i $P_{\mathcal{Y}}$ još vrijedi $(Z_2')^* P_{\mathcal{X}}\, [Z_1\,,\, Z_2'] = [0\,,\, I_{m_1-k-r}]$ odnosno $(Z_2')^* P_{\mathcal{Y}}\, [Z_1\,,\, Z_2'] = 0$.

Ako je $k + r < m_2$, tada se u singularnoj dekompoziciji matrice $Q_{\mathcal{X}}^* Q_{\mathcal{Y}}$ pojavljuje ortonormirana matrice $T' \in \mathbf{C}^{m_2 \times (m_2 - k - r)}$. Definirajmo ortonormiranu matricu $Z_2'' = Q_{\mathcal{Y}} T'$. Pokažimo da je $[Z_1\,,\, Z_2',\,,\, Z_2'']$ ortonormirana. Za to je dovoljno pokazati $Z_1^* Z_2'' = 0$ i $(Z_2')^* Z_2'' = 0$. Mi ćemo pokazati $U^* Z_2'' = 0$, $V^* Z_2'' = 0$ i $(Z_2')^* Z_2''$

$= 0$. Korištenjem relacija (3.19) i (3.20), dobivamo

$$\begin{aligned} U^* Z_2'' &= S^* Q_{\mathcal{X}}^* Q_{\mathcal{Y}} T' = S^*(S\Gamma T^*)T' = 0, \\ V^* Z_2'' &= T^* Q_{\mathcal{Y}}^* Q_{\mathcal{Y}} T' = T^* T' = 0, \\ (Z_2')^* Z'' *_2 &= (S')^* Q_{\mathcal{X}}^* Q_{\mathcal{Y}} T' = (S')^*(S\Gamma T^*)T' = 0. \end{aligned}$$

Također, zbog

$$\begin{aligned} P_{\mathcal{X}} Z_2'' &= Q_{\mathcal{X}} Q_{\mathcal{X}}^* Q_{\mathcal{Y}} T' = Q_{\mathcal{X}}(S\Gamma T^*)T' = 0, \\ P_{\mathcal{Y}} Z_2'' &= Q_{\mathcal{Y}} Q_{\mathcal{Y}}^* Q_{\mathcal{Y}} T' = Q_{\mathcal{Y}} T' = Z_2'', \end{aligned}$$

odmah slijedi $[Z_1\,, Z_2'\,, Z_2'']^* P_{\mathcal{X}} Z_2'' = 0$ i $[Z_1\,, Z_2'\,, Z_2'']^* P_{\mathcal{Y}} Z_2'' = [0\,, 0\,, I_{m_2-k-r}]^*$, a zbog hermitičnosti ortogonalnih projektora i $(Z_2'')^* P_{\mathcal{X}}[Z_1\,, Z_2'\,, Z_2''] = 0$ odnosno $(Z_2'')^* P_{\mathcal{Y}}[Z_1\,, Z_2'\,, Z_2''] = [0\,, 0\,, I_{m_2-k-r}]$.

Matrica $[Z_1\,, Z_2'\,, Z_2'']$ ima $k + 2r + m_1 - (k + r) + m_2 - (k + r) = m_1 + m_2 - k$ ortonormiranih stupaca. Uočimo da je $m_1 + m_2 - k = \dim(\mathcal{X}) + \dim(\mathcal{Y}) - \dim(\mathcal{X} \bigcap \mathcal{Y}) = \dim(\mathcal{X} + \mathcal{Y}) \leq n$. Ako je $m_1 + m_2 - k < n$, tada možemo matricu $[Z_1\,, Z_2'\,, Z_2'']$ nadopuniti s ortonormiranom matricom Z_2''' do unitarne matrice. Stupci matrice Z_2''' su ortogonalni na stupce Z_1, dakle i na stupce od $Q_{\mathcal{X}}$ i $Q_{\mathcal{Y}}$. Stoga je je $P_{\mathcal{X}} Z_2''' = 0$ i $P_{\mathcal{Y}} Z_2''' = 0$. Zaključujemo da za matricu $Z_2 = [Z_2'\,, Z_2''\,, Z_2''']$ vrijedi

$$P_{\mathcal{X}} Z_2 = Z_2 \begin{bmatrix} I & & \\ & 0 & \\ & & 0 \end{bmatrix} \qquad P_{\mathcal{Y}} Z_2 = Z_2 \begin{bmatrix} 0 & & \\ & I & \\ & & 0 \end{bmatrix},$$

pa matrica $Z = [Z_1\,, Z_2]$ ima tražena svojstva. Time je teorem dokazan. ■

Dokaz teorema 3.11 baca svjetlo na geometriju potprostora koji su vezani uz dekompoziciju ortogonalnih projektora. U tabeli 1 smo naveli glavne potprostore koje smo susreli u dokazu i naveli pripadne ortonormirane baze (Pritom se zadnji redak tabele dobije po analogiji s prethodnim retkom).

Brojevi

$$\frac{\pi}{2} > \theta_1 \geq \theta_2 \geq \cdots \geq \theta_r > 0 \tag{3.28}$$

se zovu glavni oštri (šiljasti) kutovi između potprostora $\mathcal{X}$ i $\mathcal{Y}$. Obično se piše $\theta_{r+1} = ... = \theta_{r+k} = 0$, pa ćemo i "nula kutove" θ_i, $r + 1 \leq i \leq r + k$, smatrati glavnim oštrim kutovima.

potprostor	baza (stupci matrice)
$\mathcal{X}$	$[U, Z_2']$
$\mathcal{Y}$	$[V, Z_2'']$
$\mathcal{X} \bigcap \mathcal{Y}$	$[u_{r+1}, \ldots, u_{r+k}]$
$P_{\mathcal{X}}\mathcal{Y}$	$[P_{\mathcal{X}}V, P_{\mathcal{X}}Z_2''] \to U$
$P_{\mathcal{Y}}\mathcal{X}$	$[P_{\mathcal{Y}}U, P_{\mathcal{Y}}Z_2'] \to V$
$\mathcal{X} \bigcap \mathcal{Y}^{\perp}$	Z_2'
$\mathcal{X}^{\perp} \bigcap \mathcal{Y}$	Z_2''
$\mathcal{X}^{\perp} \bigcap \mathcal{Y}^{\perp}$	Z_2'''
$\mathcal{X}^{\perp}$	$[[w_1, \ldots, w_r]\,,\, Z_2''\,,\, Z_2''']$
$\mathcal{Y}^{\perp}$	$[[w_1', \ldots, w_r']\,,\, Z_2'\,,\, Z_2''']$
	where $w_i' = (u_i - v_i \cos\theta_i)/\sin\theta_i$

Tabela 1: Baze potprostora

Uočimo da unitarna transformacija

$$Z \mapsto Z\,R, \quad R = \begin{bmatrix} I_k & & \\ & \bigoplus_{i=1}^r R_i & \\ & & I_{n-k-2r} \end{bmatrix}, \quad R_i = \begin{bmatrix} \cos\theta_i & \sin\theta_i \\ \sin\theta_i & -\cos\theta_i \end{bmatrix},$$

zamijenjuje blokove $\bigoplus_{i=1}^r \Phi_i$ i $\bigoplus_{i=1}^r \Psi_{\theta_i}$ u matricama P_1 i P_2 iz relacije (3.14). To ukazuje da glavni šiljasti kutovi ne ovise o poretku u kojem se navode potprostori.

Teoremom 3.11 su definirani glavni oštri kutovi između $\mathcal{X}$ i $\mathcal{Y}$. U slučaju potprostora iste dimenzije, možemo ih povezati sa kanonskim kutovima.

Korolar 3.12 *Neka su $\mathcal{X}$, $\mathcal{Y} \subset \mathbf{C}^n$ potprostori iste dimenzije. Tada su netrivijalni glavni šiljasti kutovi između $\mathcal{X}$ i $\mathcal{Y}$ oni netrivijalni kanonski kutovi između $\mathcal{X}$ i $\mathcal{Y}$ koji su manji od $\pi/2$.*

Dokaz. Neka je $\dim(\mathcal{X}) = \dim(\mathcal{Y}) = l$ i neka su $X_\perp$ i Y proizvoljne ortonormirane baze od $\mathcal{X}^\perp$ i $\mathcal{Y}$, respektivno. Prema korolaru 3.7, sinusi kanonskih kutova između $\mathcal{X}$ i $\mathcal{Y}$ su singularne vrijednosti matrice $X_\perp^* Y$. Prema tabeli 3.4, kao baze $X_\perp$ i Y

možemo uzeti

$$
\begin{aligned}
X_{\perp} &= [W_r\,,\, Z_2''\,,\, Z_2''']\,, \quad W_r = [w_1, \ldots, w_r]\,, \\
Y &= [V\,,\, Z_2'']
\end{aligned}
$$

pri čemu su $W_r \in \mathbf{C}^{n\times r}$, $V \in \mathbf{C}^{n\times(r+k)}$, $Z_2'' \in \mathbf{C}^{n\times(l-r-k)}$ i $Z_2''' \in \mathbf{C}^{n\times(n+k-2l)}$ kao u dokazu teorema 3.11, a r i k su nenegativni brojevi za koje vrijedi $r+k \leq l$. Sada je

$$
X_{\perp}Y = \begin{bmatrix} W_r^* \\ (Z_2'')^* \\ (Z_2''')^* \end{bmatrix} [V\,,\, Z_2''] \;=\; \begin{array}{cc} \begin{bmatrix} W_r^*V & 0 \\ 0 & I \\ 0 & 0 \end{bmatrix} & \begin{array}{l} r \\ l-r-k \\ n+k-2l \end{array} \\ \begin{array}{cc} r+k & \quad l-r-k \end{array} & \end{array}
$$

Kako je za $1 \leq i \leq r$

$$
w_i^* v_j = \delta_{ij} \sin\theta_i, \quad 1 \leq j \leq r+k,
$$

zaključujemo da je

$$
X_{\perp}Y \;=\; \begin{array}{cc} \begin{bmatrix} \Sigma_r & 0 & 0 \\ 0 & 0 & I \\ 0 & 0 & 0 \end{bmatrix} & \begin{array}{l} r \\ l-r-k \\ n+k-2l \end{array} \\ \begin{array}{ccc} r & \quad k & \text{l-r-k} \end{array} & \end{array}, \quad \Sigma_r = \begin{bmatrix} \sin\theta_1 & & \\ & \ddots & \\ & & \sin\theta_r \end{bmatrix}.
$$

Dakle su netrivijalni kanonski kutovi svi glavni šiljasti kutovi θ_i, $1 \leq i \leq r$ i eventualno pravi kutovi (ako je $\mathcal{X}^{\perp} \bigcap \mathcal{Y} = \{0\}$, tj. ako Z_2'' postoji). ■

Zadatak: *Dokažite korolar 3.12 polazeći od baza X i Y opisanim u relacijama (3.6) – (3.9). Konstruirajte $P_{\mathcal{X}} = XX^*$ i $P_{\mathcal{Y}} = YY^*$, te pokažite da postoji permutacija S, takva da su $P_1 = S^*W^*P_{\mathcal{X}}WS$ i $P_2 = S^*W^*P_{\mathcal{Y}}WS$ kao u teoremu 3.11.*

Korolar 3.13 *Neka su ispunjene pretpostavke teorema 3.11 i neka su $\sigma_1 \geq \ldots \geq \sigma_r > 0$ sinusi od nule različitih glavnih oštrih kutova između $\mathcal{X}$ i $\mathcal{Y}$. Tada vrijedi:*

(i) Singularne vrijednosti matrice $P_{\mathcal{X}}(I-P_{\mathcal{Y}})$ (ili matrice $(I-P_{\mathcal{Y}})P_{\mathcal{X}}$) su

$$
\underbrace{1\,,\,1\,,\,\ldots\,,\,1}_{dim(\mathcal{X})-k-r}\,,\; \sigma_1\,,\ldots\,,\,\sigma_r\,,\;0\,,\ldots\,,\,0\,.
$$

(ii) Singularne vrijednosti matrice $P_{\mathcal{Y}}(I - P_{\mathcal{X}})$ (ili matrice $(I - P_{\mathcal{X}})P_{\mathcal{Y}}$) su

$$\underbrace{1\ ,\ 1\ ,\ \ldots\ ,\ 1}_{dim(\mathcal{Y})-k-r}\ ,\ \sigma_1\ ,\ldots\ ,\ \sigma_r\ ,\ 0\ ,\ldots\ ,\ 0\ .$$

(iii) Singularne vrijednosti od $P_{\mathcal{X}} - P_{\mathcal{Y}}$ su

$$\underbrace{1\ ,\ 1\ ,\ \ldots\ ,\ 1}_{dim(\mathcal{X})+dim(\mathcal{Y})-2(k+r)}\ ,\ \sigma_1\ ,\ \sigma_1\ ,\ \ldots\ ,\ \sigma_r\ ,\ \sigma_r\ ,\ 0\ ,\ \ldots\ ,\ 0\ .$$

Dokaz. (i) Hermitski adjungirane matrice imaju iste singularne vrijednosti. Stoga $P_{\mathcal{X}}(I - P_{\mathcal{Y}})$ i $(I - P_{\mathcal{Y}})P_{\mathcal{X}} = (I - P_{\mathcal{Y}})^* P_{\mathcal{X}}^* = [P_{\mathcal{X}}(I - P_{\mathcal{Y}})]^*$ imaju iste singularne vrijednosti.

U bazi Z ortogonalni projektori $P_{\mathcal{X}}$ i $P_{\mathcal{Y}}$ imaju zapise P_1 i P_2, respektivno. Pritom je $Z^* P_{\mathcal{X}}(I - P_{\mathcal{Y}})Z = P_1(I - P_2)$, a P_1 i P_2 su određeni relacijom (3.14). Označimo li $c_i = \cos\theta_i$, $s_i = \sin\theta_i$, imamo

$$\begin{aligned}
\Phi_i(I_2 - \Psi_{\theta_i}) &= \begin{bmatrix} 1 \\ 0 \end{bmatrix} [1 \;\; 0] \begin{bmatrix} 1-c_i^2 & -c_i s_i \\ -c_i s_i & 1-s_i^2 \end{bmatrix} \\
&= \begin{bmatrix} s_i^2 & -c_i s_i \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \sigma_i & \\ & 0 \end{bmatrix} \begin{bmatrix} s_i & -c_i \\ c_i & s_i \end{bmatrix}
\end{aligned}$$

Odavdje slijedi

$$P_1(I - P_2) = \begin{bmatrix} 0 & & \\ & \bigoplus_{i=1}^{r} \begin{bmatrix} \sigma_i & \\ & 0 \end{bmatrix} & \\ & & D_1 \end{bmatrix} \begin{bmatrix} I_k & & \\ & \bigoplus_{i=1}^{r} \begin{bmatrix} s_i & -c_i \\ c_i & s_i \end{bmatrix} & \\ & & I_{n-k-2r} \end{bmatrix},$$

pri čemu jedinica (u D_1) ima onoliko koliko ima stupaca u matrici Z_2' iz dokaza teorema 3.11, a to je $\dim(\mathcal{X}) - k - r$. Kako je druga matrica na desnoj strani zadnje relacije unitarna, dokaz je završen.

(ii) Na sličan način kao u (i) se pokazuje da $P_{\mathcal{Y}}(I - P_{\mathcal{X}})$ i $(I - P_{\mathcal{X}})P_{\mathcal{Y}}$ imaju iste singularne vrijednosti. Ostatak dokaza slijedi iz tvrdnje (i) jednostavnom zamjenom prostora $\mathcal{X}$ i $\mathcal{Y}$.

(iii) Koristeći iste oznake kao u (i) i relaciju

$$\Phi_i - \Psi_{\theta_i} = \begin{bmatrix} 1-c_i^2 & -c_is_i \\ -c_is_i & -s_i^2 \end{bmatrix} = \begin{bmatrix} s_i^2 & -c_is_i \\ -c_is_i & -s_i^2 \end{bmatrix} = \begin{bmatrix} s_i & -c_i \\ -c_i & -s_i \end{bmatrix} \begin{bmatrix} \sigma_i & 0 \\ 0 & \sigma_i \end{bmatrix},$$

lako dobivamo

$$P_1 - P_2 = Q_1 \begin{bmatrix} 0 & & \\ & \bigoplus_{i=1}^{r} \begin{bmatrix} \sigma_i & 0 \\ 0 & \sigma_i \end{bmatrix} & \\ & & D_1 - D_2 \end{bmatrix},$$

pri čemu je Q_1 unitarna matrica. U računu za broj jedinica kao singularnih vrijednosti matrice $P_1 - P_2$, treba uzeti u obzir da $D_1 - D_2$ ima $\dim(Z_2') = \dim(\mathcal{X}) - k - r$ jedinica i $\dim(Z_2'') = \dim(\mathcal{Y}) - k - r$ negativnih jedinica. ■

Iz dokaza zadnje tvrdnje korolara 3.13 zaključujemo da vrijedi

$$\|P_{\mathcal{Y}} - P_{\mathcal{X}}\| = \begin{cases} \sin\theta_1, & \textit{ako su } D_1 = D_2 = 0 \textit{ ili ne postoje} \\ 1, & \textit{ako je } D_1 \neq 0 \textit{ ili } D_2 \neq 0 \end{cases} \tag{3.29}$$

Primijetimo, u slučaju $\dim(\mathcal{X}) = \dim(\mathcal{Y}) = k + r$, nema matrica Z_2' i Z_2'' u dokazu teorema 3.11 pa je ili $D_1 = D_2 = 0$ ili D_1 i D_2 ne postoje. Stoga nema prvih $\dim(\mathcal{X}) - (k + r)$ jedinica u iskazu (i), nema prvih $\dim(\mathcal{Y}) - (k + r)$ jedinica u iskazu (ii), pa nema ni prvih $\dim(\mathcal{X}) + \dim(\mathcal{Y}) - 2(k + r)$ jedinica u iskazu (iii) korolara 3.13. To znači da su svi netrivijalni kanonski kutovi manji od $\pi/2$, odnosno da su u ovom slučaju svi kanonski kutovi jednaki glavnim šiljastim kutovima.

Međutim, ako je $\dim(\mathcal{X}) = \dim(\mathcal{Y}) > k + r$, tada postoje te jedinice (jer postoje obje matrice Z_2' i Z_2'' iz dokaza teorema 3.11). Te jedinice su sinusi pravih (jednakih $\pi/2$) kanonskih kutova.

Primijetimo još da u slučaju $\dim(\mathcal{X}) \neq \dim(\mathcal{Y})$ postoji najmanje jedna jedinica u matrici D_1 ili D_2, a isto tako i u matrici $D_1 - D_2$.

Ova razmatranja uz relaciju (3.29) i korolar 3.13 pokazuju da vrijedi

Korolar 3.14 *Neka su $\mathcal{X}$ i $\mathcal{Y}$ potprostori od $\mathbf{C}^n$.*

(i) Ako je $\dim(\mathcal{X}) \neq \dim(\mathcal{Y})$, *onda je* $\|P_{\mathcal{X}} - P_{\mathcal{Y}}\| = 1$.

(ii) Ako je $dim(\mathcal{X}) > dim(\mathcal{Y})$ *(* $dim(\mathcal{X}) < dim(\mathcal{Y})$ *), tada je*

$$\|(I - P_{\mathcal{Y}})P_{\mathcal{X}}\|_2 = 1 \qquad (\ \|(I - P_{\mathcal{X}})P_{\mathcal{Y}}\|_2 = 1 \ ) \ .$$

(ii) Ako je $dim(\mathcal{X}) = dim(\mathcal{Y})$*, onda je*

$$\|(I - P_{\mathcal{X}})P_{\mathcal{Y}}\|_2 = \|(I - P_{\mathcal{Y}})P_{\mathcal{X}}\|_2 = \|P_{\mathcal{Y}} - P_{\mathcal{X}}\|_2.$$

Za potprostore jednakih dimenzija vrijedi

Korolar 3.15 *Neka su* $\mathcal{X}$ *i* $\mathcal{Y}$ l *dimenzionalni potprostori iz* $\mathbf{C}^n$*, a* $P_{\mathcal{X}}$ *i* $P_{\mathcal{Y}}$ *pripadni ortogonalni projektori. Ako je* $2l > n$ *onda matrica* $P_{\mathcal{X}} P_{\mathcal{Y}}$ *ima najmanje* $2l - n$ *singularnih vrijednosti jednakih jedan.*

Dokaz. Zbog $2l > n$, potprostori $\mathcal{X}$ i $\mathcal{Y}$ imaju zajednički potprostor $\mathcal{Z} = \mathcal{X} \bigcap \mathcal{Y}$, kojem je dimenzija najmanje $2l - n$. Za $z \in \mathcal{Z}$ vrijedi

$$(P_{\mathcal{X}} P_{\mathcal{Y}}) (P_{\mathcal{X}} P_{\mathcal{Y}})^* \ z = P_{\mathcal{X}} P_{\mathcal{Y}} P_{\mathcal{X}} \ z = P_{\mathcal{X}} P_{\mathcal{Y}} z = P_{\mathcal{X}} z = z.$$

Dakle je $\mathcal{Z}$ sadržan u vlastitom potprostoru matrice $(P_{\mathcal{X}} P_{\mathcal{Y}}) (P_{\mathcal{X}} P_{\mathcal{Y}})^*$ koji pripada vlastitoj vrijednosti jedan. To znači da je barem $\dim(\mathcal{Z}) \geq 2l - n$ singularnih vrijednosti matrice $P_{\mathcal{X}} P_{\mathcal{Y}}$ jednako jedan. ∎

Zadatak: *Ako su* $\mathcal{X}$ *i* $\mathcal{Y}$ *potprostori jednakih dimenzija, onda su prirodne baze za potprostore* $\mathcal{X}$ *i* $\mathcal{Y}$ *određene teoremom 3.5. Pokažite da je matrica* Q *iz relacije (3.6), odnosno (3.8), do na permutaciju stupaca jednaka matrici* Z *iz teorema 3.11. To znači da je teorem 3.5 specijalni slučaj teorema 3.11. Ova tvrdnja je ilustrirana slijedećim primjerom.*

Primjer 3.16 Neka stupci matrica X i Y čine ortonormirane baze za potprostore $\mathcal{X}$ i $\mathcal{Y}$,

$$X = \begin{bmatrix} -0.3913 & -0.2755 \\ 0.6633 & -0.5708 \\ 0.3891 & 0.4800 \\ -0.2951 & 0.3766 \\ 0.4104 & 0.4755 \end{bmatrix}, \qquad Y = \begin{bmatrix} -0.1630 & 0.3218 \\ 0.4179 & 0.4920 \\ 0.3549 & 0.6570 \\ -0.8153 & 0.4715 \\ -0.0899 & 0.0205 \end{bmatrix}.$$

Tada postoji unitarna matrica Q čiji stupci razapinju cijeli prostor

$$Q = \begin{bmatrix} -0.2543 & -0.4055 & -0.2572 & 0.4787 & 0.6896 \\ 0.8325 & -0.2697 & 0.2204 & 0.4259 & -0.0651 \\ 0.1731 & 0.5931 & -0.6833 & 0.3752 & -0.1028 \\ -0.4178 & 0.2330 & 0.5023 & 0.6602 & -0.2880 \\ 0.1944 & 0.5973 & 0.4075 & -0.1129 & 0.6532 \end{bmatrix},$$

i unitarne matrice U i V, tako da je

$$X\,U = Q \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \qquad Y\,V = Q \begin{bmatrix} 0.2612 & 0 \\ 0 & 0.8128 \\ 0.9653 & 0 \\ 0 & 0.5825 \\ 0 & 0 \end{bmatrix}.$$

Odmah zaključujemo da je

$$\Gamma = \begin{bmatrix} 0.2612 & 0 \\ 0 & 0.8128 \end{bmatrix}, \qquad \Sigma = \begin{bmatrix} 0.9653 & 0 \\ 0 & 0.5825 \end{bmatrix}.$$

Također zaključujemo da je

$$P_{\mathcal{X}} = Q \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} Q^*,$$

odnosno

$$P_{\mathcal{Y}} = Q \begin{bmatrix} 0.0682 & 0 & 0.2521 & 0 & 0 \\ 0 & 0.6607 & 0 & 0.4735 & 0 \\ 0.2521 & 0 & 0.9318 & 0 & 0 \\ 0 & 0.4735 & 0 & 0.3393 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} Q^*.$$

Nakon permutacije drugog i trećeg stupca matrice Q dobijamo matricu Z, a također prepoznajemo i matrice P_1 i P_2 iz teorema 3.11.

Kanonske kutove θ_i možemo izračunati pomoću singularnih vrijednosti matrice Y^*X, (singularne vrijednosti matrice $Y_1^*X_1$ su kosinusi kanonskih kutova). Dobije se

$$\begin{bmatrix} \cos\theta_1 \\ \cos\theta_2 \end{bmatrix} = \begin{bmatrix} 0.2612 \\ 0.8128 \end{bmatrix},$$

a iz njih se također mogu dobiti i matrice P_1 i P_2. Ovim pristupom se ne dobije direktna veza matrica P_1 i P_2 sa polaznim ortogonalnim projektorima XX^* i YY^*.

3.5 Udaljenost potprostora

Uz danu matricu A vezani su razni potprostori. Osim $\mathcal{R}(A)$ i $\mathcal{N}(A)$ zanimljivi su i vlastiti potprostori, invarijantni potprostori, te potprostori razapeti određenim lijevim odnosno desnim singularnim vektorima. Ako se matrica malo promijeni promijenit će se i ti potprostori, pa nam treba neka mjera za udaljenost potprostora.

Ako su $\mathcal{X}$ i $\mathcal{Y}$ potprostori iste dimenzije jedna dobra mjera za udaljenost između njih je $\|P_{\mathcal{X}} - P_{\mathcal{Y}}\|_2$. Naime, odmah se vidi da je $\|P_{\mathcal{X}} - P_{\mathcal{Y}}\|_2$ metrika na skupu potprostora od $\mathbf{C}^n$ iste dimanzije. Ova metrika je invarijantna na unitarne transformacije prostora $\mathbf{C}^n$. Također, korolar 3.13(iii), korolar 3.14(iii) i relacija (3.29) pokazuju da $\|P_{\mathcal{X}} - P_{\mathcal{Y}}\|_2$ poprima maksimalnu vrijednost jedan, ako i samo ako postoji vektor iz $\mathcal{X}$ ($\mathcal{Y}$) koji je okomit na cijeli $\mathcal{Y}$ ($\mathcal{X}$). U protivnom (sada se kanonski i glavni oštri kutovi poklapaju), $\|P_{\mathcal{X}} - P_{\mathcal{Y}}\|_2$ ima vrijednost najvećeg kanonskog kuta.

Ako su $\mathcal{X}$ i $\mathcal{Y}$ potprostori različitih dimenzija, $\|P_{\mathcal{X}} - P_{\mathcal{Y}}\|_2 = 1$, pa ova mjera nema smisla. Zato ćemo definirati novu mjeru udaljenosti između potprostora, a to je kut između $\mathcal{X}$ i $\mathcal{Y}$. Razmatranje počinjemo definiranjem kuta između dva vektora i između vektora i potprostora.

Kut između vektora x i y iz $\mathbf{C}^n$ je broj kojeg označavamo sa $\angle(x,y)$, a definiran je sa

$$\angle(x,y) = \arccos\left(\frac{|\, x^*y \,|}{\|x\|_2\|y\|_2}\right), \quad x \neq 0,\ y \neq 0. \tag{3.30}$$

Na bazi kuta između vektora, definiran je i kut između vektora i potprostora

$$\angle(x,\mathcal{Y}) = \inf\{\, \angle(x,y)\, ;\ y \in \mathcal{Y}\,\}, \quad x \neq 0,\ \mathcal{Y} \neq \{0\}. \tag{3.31}$$

Iz definicije (3.31) slijedi $\angle(x, \mathcal{Y}) = \pi/2$ ako i samo ako je x okomit na $\mathcal{Y}$, tj, ako je $P_{\mathcal{Y}}x = 0$.

Propozicija 3.17 *Ako je $P_{\mathcal{Y}}$ ortogonalni projektor na $\mathcal{Y} \subset \mathbf{C}^n$, tada je*

$$\angle(x, \mathcal{Y}) = \arcsin(\frac{\|(I - P_{\mathcal{Y}})x\|_2}{\|x\|_2}).$$

Dokaz. Za $y \in \mathcal{Y}$ vrijedi

$$| y^* x | = | (P_{\mathcal{Y}} y)^* x | = | y^* P_{\mathcal{Y}} x | \leq \|P_{\mathcal{Y}} x\|_2 \|y\|_2,$$

pa imamo

$$\frac{| y^* x |}{\|x\|_2 \|y\|_2} \leq \frac{\|P_{\mathcal{Y}} x\|_2}{\|x\|_2}, \quad x \neq 0,\ y \neq 0,\ y \in \mathcal{Y}. \tag{3.32}$$

Za $y = P_{\mathcal{Y}}x$ vrijedi

$$| y^* x | = | x^* P_{\mathcal{Y}} x | = | x^* P_{\mathcal{Y}}^2 x | = | (P_{\mathcal{Y}} x)^* P_{\mathcal{Y}} x | = \|P_{\mathcal{Y}} x\|_2^2 = \|y\|_2 \|P_{\mathcal{Y}} x\|_2,$$

pa postižemo jednakost u (3.32). Dakle je

$$\max_{0 \neq y \in \mathcal{Y}} \frac{| y^* x |}{\|x\|_2 \|y\|_2} = \frac{\|P_{\mathcal{Y}} x\|_2}{\|x\|_2}, \quad x \neq 0.$$

S obzirom da je $\arccos(\phi)$ padajuća funkcija za $0 \leq \phi \leq \pi/2$, imamo

$$\begin{aligned} \angle(x, \mathcal{Y}) &= \inf_{0 \neq y \in \mathcal{Y}} \arccos(\frac{| x^* y |}{\|x\|_2 \|y\|_2}) = \arccos(\sup_{0 \neq y \in \mathcal{Y}} \frac{| x^* y |}{\|x\|_2 \|y\|_2}) \\ &= \arccos(\frac{\|P_{\mathcal{Y}} x\|_2}{\|x\|_2}). \end{aligned}$$

Koristeći zadnju relaciju i jednakost

$$\|x\|_2^2 = \|P_{\mathcal{Y}} x\|_2^2 + \|(I - P_{\mathcal{Y}})x\|_2^2$$

lako slijedi

$$\sin \angle(x, \mathcal{Y}) = \sqrt{1 - \cos^2 \angle(x, \mathcal{Y})} = \sqrt{1 - \frac{\|P_{\mathcal{Y}} x\|_2^2}{\|x\|_2^2}} = \frac{\|(I - P_{\mathcal{Y}})x\|_2}{\|x\|_2}.$$

Time je propozicija dokazana. ∎

Iz propozicije 3.18 odmah slijedi da je kut između x i $\mathcal{Y}$ zapravo kut između vektora x i $P_{\mathcal{Y}}x$.

Da bismo dobro definirali kut između potprostora $\mathcal{X}$ i $\mathcal{Y}$, trebat će nam slijedeća

Propozicija 3.18

$$\sup_{0\neq x\in\mathcal{X}} \angle\,(x,\mathcal{Y}) = \arcsin\left(\,\|(I-P_{\mathcal{Y}})P_{\mathcal{X}}\|_2\,\right).$$

Dokaz. Neka je

$$\Theta_{\mathcal{X}}(\mathcal{Y}) = \sup_{\substack{x\in\mathcal{X}\\ x\neq 0}} \angle\,(x,\mathcal{Y}).$$

Tada je zbog propozicije 3.17 i monotonosti funkcije arcsin,

$$\Theta_{\mathcal{X}}(\mathcal{Y}) = \sup_{0\neq x\in\mathcal{X}} \arcsin\left(\frac{\|(I-P_{\mathcal{Y}})x\|_2}{\|x\|_2}\right) = \arcsin\left(\sup_{0\neq x\in\mathcal{X}} \frac{\|(I-P_{\mathcal{Y}})x\|_2}{\|x\|_2}\right).$$

Uočimo da je

$$\sup_{0\neq x\in\mathcal{X}} \frac{\|(I-P_{\mathcal{Y}})x\|_2}{\|x\|_2} = \sup_{\substack{z\in\mathbf{C}^n\\ P_{\mathcal{X}}z\neq 0}} \frac{\|(I-P_{\mathcal{Y}})P_{\mathcal{X}}z\|_2}{\|P_{\mathcal{X}}z\|_2}.$$

Koristeći teorem 3.11 i unitarnu invarijantnost spektralne norme, imamo

$$\frac{\|(I-P_{\mathcal{Y}})P_{\mathcal{X}}z\|_2}{\|P_{\mathcal{X}}z\|_2} = \frac{\|Z(I-P_2)Z^*ZP_1Z^*z\|_2}{\|ZP_1Z^*z\|_2} = \frac{\|(I-P_2)P_1y\|_2}{\|P_1y\|_2}, \quad y = Z^*z.$$

Uočimo da je $y=0$ ako i ako ako je $z=0$, i $P_{\mathcal{X}}z=0$ ako i samo ako je $P_1y=0$. Dakle je

$$\sin\left(\Theta_{\mathcal{X}}(\mathcal{Y})\right) = \sup_{\substack{y\in\mathbf{C}^n\\ P_1y\neq 0}} \frac{\|(I-P_2)P_1y\|_2}{\|P_1y\|_2}.$$

Koristeći oznake iz relacije (3.14) teorema 3.11, na sličan način kao u dokazu korolara 3.13(i), lako se pokaže da je

$$(I-P_2)P_1 = \begin{bmatrix} 0_k & & \\ & \bigoplus_{i=1}^{r}(I_2-\Psi_{\theta_i})\Phi_i & \\ & & D_1 \end{bmatrix} = R\begin{bmatrix} 0_k & & \\ & \bigoplus_{i=1}^{r}\Phi_i\cdot\sigma_i & \\ & & D_1 \end{bmatrix},$$

gdje je $\sigma_i = sin\theta_i$, $1\leq i\leq r$, a R je unitarna matrica. Načinimo particiju vektora y u suglasnosti sa particijom matrice P_1, $y^T = [y_1^T, y_2^T, y_3^T]$. Tada je

$$\frac{\|(I-P_2)P_1y\|_2^2}{\|P_1y\|_2^2} = \frac{\|\bigoplus_{i=1}^{r}\Phi_i\cdot\sigma_i\; y_2\|_2^2 + \|D_1\,y_3\|_2^2}{\|\bigoplus_{i=1}^{r}\Phi_i\; y_2\|_2^2 + \|D_1\,y_3\|_2^2},$$

pa lako zaključujemo da je

$$\sin(\Theta_{\mathcal{X}}(\mathcal{Y})) = \begin{cases} 1\ , \text{ ako je } D_1 \neq 0, \\ \sigma_1\ , \text{ ako je } D_1 = 0 \end{cases} .$$

Čak i više, u slučaju $D_1 \neq 0$, supremum 1 se dostiže na vektoru $y = e_{2r+k+1}$ (tada je $y_3 = e_1$), odnosno na vektoru $x = z = Ze_{2r+k+1}$. U drugom slučaju se supremum σ_1 dostiže na vektoru $y = e_{k+1}$ (tada je $y_2 = e_1$), odnosno na vektoru $x = z = Ze_{k+1}$. Dakle, umjesto supremuma možemo pisati maksimum.

Prema korolaru 3.13(i), funkcija $\|(I - P_2)P_1\|_2$ postiže iste vrijednosti 1 i σ_1 u slučajevima kad je $D_1 \neq 0$ i $D_1 = 0$. Dakle je $\sin(\Theta_{\mathcal{X}}(\mathcal{Y})) = \|(I - P_2)P_1\|_2$. ■

Na osnovu propozicije 3.18 odmah zaključujemo da vrijedi

$$\Theta_{\mathcal{X}}(\mathcal{Y}) = \sup_{0 \neq x \in \mathcal{X}} \angle(x, \mathcal{Y}) = \arcsin(\|(I - P_{\mathcal{Y}})P_{\mathcal{X}}\|_2)\ , \tag{3.33}$$

$$\Theta_{\mathcal{Y}}(\mathcal{X}) = \sup_{0 \neq y \in \mathcal{Y}} \angle(y, \mathcal{X}) = \arcsin(\|(I - P_{\mathcal{X}})P_{\mathcal{Y}}\|_2)\ . \tag{3.34}$$

Definicija 3.19 *Neka su $\mathcal{X}$ i $\mathcal{Y}$ netrivijalni potprostori od $\mathbf{C}^n$. Tada je*

$$\Theta(\mathcal{X}\,,\,\mathcal{Y}) = \min\{\,\Theta_{\mathcal{X}}(\mathcal{Y})\,,\,\Theta_{\mathcal{Y}}(\mathcal{X})\}$$

kut između $\mathcal{X}$ i $\mathcal{Y}$.

S obzirom da ovako definiran kut ovisi o spektralnoj normi, još se koristi i oznaka $\angle_2(\mathcal{X}\,,\,\mathcal{Y})$.

Teorem 3.20 *Neka su $\mathcal{X}$ i $\mathcal{Y}$ netrivijalni potprostori od $\mathbf{C}^n$. Tada za kut $\Theta(\mathcal{X}\,,\,\mathcal{Y})$ vrijede slijedeće tvrdnje:*

(i) $\Theta(\mathcal{X}\,,\,\mathcal{Y}) \geq 0$.

(ii) $\Theta(\mathcal{X}\,,\,\mathcal{Y}) = 0$ *ako i samo ako je* $\mathcal{X} \subseteq \mathcal{Y}$ *ili* $\mathcal{Y} \subseteq \mathcal{X}$.

(iii) $\Theta(\mathcal{X}\,,\,\mathcal{Y}) = \Theta(\mathcal{Y}\,,\,\mathcal{X})$.

(iv) $\Theta(\mathcal{W}\mathcal{X}\,,\,\mathcal{W}\mathcal{Y}) = \Theta(\mathcal{X}\,,\,\mathcal{Y})$ *za svaki unitarni operator* $\mathcal{W}$.

(v) $\Theta(\mathcal{X}, \mathcal{Y}) \le \Theta(\mathcal{X}, \mathcal{Z}) + \Theta(\mathcal{Z}, \mathcal{Y})$ *ako je* $dim(\mathcal{X}) \le dim(\mathcal{Z}) \le dim(\mathcal{Y})$ *ili* $dim(\mathcal{X}) \ge dim(\mathcal{Z}) \ge dim(\mathcal{Y})$.

Dokaz. (i) Ovo svojstvo slijedi odmah iz definicije 3.19.
(ii) $\Theta(\mathcal{X}, \mathcal{Y}) = 0$ ako i samo ako je $\Theta_{\mathcal{X}}(\mathcal{Y}) = 0$ ili $\Theta_{\mathcal{Y}}(\mathcal{X}) = 0$. Promotrima lanac ekvivalentnih tvrdnji:

$$\begin{aligned}\Theta_{\mathcal{X}}(\mathcal{Y}) = 0 &\iff \arcsin(\|(I - P_{\mathcal{Y}})P_{\mathcal{X}}\|_2) = 0 \iff \|(I - P_{\mathcal{Y}})P_{\mathcal{X}}\|_2 = 0 \\ &\iff (I - P_{\mathcal{Y}})P_{\mathcal{X}} = 0 \iff P_{\mathcal{X}} = P_{\mathcal{Y}}P_{\mathcal{X}} \iff \mathcal{X} \subseteq \mathcal{Y}.\end{aligned}$$

Na slični način se pokaže da je $\Theta_{\mathcal{Y}}(\mathcal{X}) = 0$ ako i samo ako $\mathcal{Y} \subseteq \mathcal{X}$, pa je tvrdnja (ii) dokazana.
(iii) Ova tvrdnja odmah slijedi zbog $\min\{\Theta_{\mathcal{X}}(\mathcal{Y}), \Theta_{\mathcal{Y}}(\mathcal{X})\} = \min\{\Theta_{\mathcal{Y}}(\mathcal{Y}), \Theta_{\mathcal{X}}(\mathcal{X})\}$.
(iv) Uzmimo ortonormiranu bazu Z prostora $\mathbf{C}^n$ iz teorema 3.11. Ortonormirane baze za $\mathcal{X}$ i $\mathcal{Y}$ su prema tabeli 3.4 $X = [U, Z_2']$ i $Y = [V, Z_2'']$, a $P_1 = XX^*$ i $P_2 = YY^*$ su reprezentacije ortogonalnih projektora na $\mathcal{X}$ i $\mathcal{Y}$. Unitarni operator $\mathcal{W}$ reprezentiran je u bazi Z nekom unitarnom matricom W. Ortonormirane baze za $\mathcal{W}\mathcal{X}$ i $\mathcal{W}\mathcal{Y}$ su WX i WY, pa su ortogonalni projektori na $\mathcal{W}\mathcal{X}$ i $\mathcal{W}\mathcal{Y}$ u bazi Z matrice $P_{\mathcal{W}\mathcal{X}} = WXX^*W^* = WP_1W^*$ i $P_{\mathcal{W}\mathcal{Y}} = WYY^*W^* = WP_2W^*$. Koristeći unitarnu invarijantnost spektralne norme dobivamo

$$\begin{aligned}\Theta_{\mathcal{W}\mathcal{X}}(\mathcal{W}\mathcal{Y}) &= \arcsin(\|(I - P_{\mathcal{W}\mathcal{Y}})P_{\mathcal{W}\mathcal{X}}\|_2) = \arcsin(\|W(I - P_2)W^*WP_1W^*\|_2) \\ &= \arcsin(\|(I - P_2)P_1\|_2) = \Theta_{\mathcal{X}}(\mathcal{Y}).\end{aligned}$$

Na sličan način se dobije $\Theta_{\mathcal{W}\mathcal{Y}}(\mathcal{W}\mathcal{X}) = \Theta_{\mathcal{Y}}(\mathcal{X})$, pa je (iv) dokazano.
(v) Dokaz slijedi iz relacije za kutove vektora

$$\angle(x, y) \le \angle(x, z) + \angle(z, y)$$

Dokažite za vježbu zadnju relaciju i onda tvrdnju (v). ■

Pogledajmo još na što se svodi kut $\Theta(\mathcal{X}, \mathcal{Y})$ ako su potprostori $\mathcal{X}$ i $\mathcal{Y}$ iste dimenzije. Prema korolaru 3.14(ii) u tom slučaju je $\Theta_{\mathcal{X}}(\mathcal{Y}) = \Theta_{\mathcal{Y}}(\mathcal{X}) = \arcsin(\|P_{\mathcal{X}} - P_{\mathcal{Y}}\|_2)$ pa je $\sin(\Theta(\mathcal{X}, \mathcal{Y})) = \|P_{\mathcal{X}} - P_{\mathcal{Y}}\|_2$. Dakle je $\sin(\Theta(\mathcal{X}, \mathcal{Y}))$ (da li i $\Theta(\mathcal{X}, \mathcal{Y})$?) metrika na skupu potprostora iste dimenzije.

Zadatak: Dajte geometrijsku prikaz i interpretaciju kuta $\Theta(\mathcal{X}, \mathcal{Y})$ između potprostora te brojeva $\Theta_{\mathcal{X}}(\mathcal{Y})$, $\Theta_{\mathcal{Y}}(\mathcal{X})$ i udaljenosti $\|P_{\mathcal{X}} - P_{\mathcal{Y}}\|_2$, ako je $n = 2$ i $n = 3$.

4. PERTURBACIJA VLASTITIH VRIJEDNOSTI

U ovom poglavlju bavit ćemo se slijedećim problemom: ako su $A\,,\ E \in \mathbf{C}^{n\times n}$, treba odrediti udaljenost spektara matrica A i $A+E$. U ovom problemu A se često naziva polazna (neperturbirana) matrica, a E matrica perturbacije. Odgovor na postavljeno pitanje nije trivijalan jer se razne klase matrica, čak i sa istim spektrima, različito ponašaju pod istim pertubacijama.

Primjer 4.1 *Neka je $A = 0$. Tada je za svako E, $A + E = E$, pa je $\lambda\,(A) = \{0\}$ i $\lambda\,(A + E) = \lambda\,(E)$. Stoga za $\lambda_i \in \lambda\,(A + E)$ vrijedi $|\lambda_i| \leq \|E\|_2$.*

Uzmimo sada $A = [0\,,\, e_1\,,\, \ldots\,,\, e_{n-1}]$ i $E = [\varepsilon e_n\,,\, 0\,,\, \ldots\,,\, 0]$. Opet imamo $\lambda\,(A) = \{0\}$. Međutim

$$\lambda\,(A + E) = \{\varepsilon^{\frac{1}{n}}\omega_i\,;\ \omega_i = e^{\frac{2\pi i}{n}},\ 0 \leq i \leq n-1\}.$$

Npr. ako je $n = 32$ i $\varepsilon = 10^{-16}$, tada se zbog te perturbacije, čija je norma $\|E\|_2 = \varepsilon$ reda veličine osnovne relativne greške zaokruživanja (za tip dvostruke točnosti u konačnoj aritmetici) na većini modernih računala, sve vlastite vrijednosti od A pomaknu u prvom slučaju za najviše ε, a u drugom slučaju za točno $\sqrt{0.1} \approx 0.31623$.

Ovaj primjer pokazuje da generalna teorija perturbacije mora dati vrlo pesimistične rezultate. Lijek je razviti specijalne teorije koje će dati povoljne rezultate barem za neke klase matrica i neke klase perturbacija. Nas će zanimati koliko se promijene vlastite vrijednosti u apsolutnom i u relativnom smislu. Prvo ćemo promotriti opći slučaj, kad je proizvoljna i matrica A i perturbacija E. Zatim ćemo se pozabaviti sa klasom normalnih, a onda, važnom klasom hermitskih matrica.

Načinimo dogovor o oznakama. Uz A i E koristit ćemo oznaku $\tilde{A} = A + E$ za perturbiranu matricu A. Karakteristični polinomi matrica A i $\tilde{A}$ bit će označeni s $\Phi_A(\lambda)$ i $\Phi_{\tilde{A}}(\lambda)$, a spektri matrica sa $\lambda\,(A)$ i $\lambda\,(\tilde{A})$, respektivno. Ove oznake ćemo slobodno koristiti, osim u izrekama teorema gdje ćemo biti pažljiviji.

4.1 Opća teorija perturbacije

U ovoj točki pretpostavljamo proizvoljnu polaznu matricu i proizvoljnu matricu perturbacije. Dokazujemo osnovna svojstva vlastitih vrijednosti, u prvom redu njihovu neprekidnu ovisnost o matrici, a zatim njihovu diferencijabilnost u slučaju kratnosti jedan. Također dokazujemo perturbacijske teoreme Ostrowskog, Elsnera, Bauera i Fikea, Henricia te Gerschgorina.

Neprekidnost: teoremi Ostrowskog i Elsnera

Neprekidna ovisnost vlastitih vrijednosti o matrici slijedi iz činjenice da su nule polinoma neprekidne funkcije koeficijenata polinoma i da su koeficijenti karakterističnog polinoma (kao sume glavnih minora matrice) neprekidne funkcije matričnih elemenata. Dokaz prve tvrdnje ide preko Rouchéovog teorema za analitičke funkcije.

Teorem 4.2 (Rouché) *Neka su ϕ i η analitičke funkcije u jednostruko povezanom području Ω i neka je $\mathcal{D}$ krug koji je zajedno sa svojim rubom $\partial\mathcal{D}$ sadržan u Ω. Ako je*

$$\mid \eta(\zeta) \mid < \mid \phi(\zeta) \mid, \quad \textit{za sve } \zeta \in \partial\mathcal{D},$$

onda ϕ i $\phi + \eta$ imaju jednaki broj nultočaka u $\mathcal{D}$.

Dokaz teorema 4.2 se može naći u svakoj knjizi o analitičkim funkcijama. U Rouchéovu teoremu nule se broje prema njihovim algebarskim višestrukostima. Pokažimo da su nultočke polinoma neprekidne funkcije koeficijenata polinoma.

Teorem 4.3 *Neka su*

$$\begin{aligned} p(\lambda) &= \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_0\,, \ a_j \in \mathbf{C}, \ 0 \le j \le n-1 \\ q(\lambda) &= \lambda^n + b_{n-1}\lambda^{n-1} + \cdots + b_0\,, \ b_j \in \mathbf{C}, \ 0 \le j \le n-1 \end{aligned}$$

polinomi s nultočkama $\lambda_1, \ldots, \lambda_n$ i $\mu_1, \ldots, \mu_n$, respektivno. Tada za svako $\varepsilon > 0$, postoji $\delta > 0$, tako da

$$\max_{0\le j\le n-1} \mid a_j - b_j \mid < \delta \qquad \textit{povlači} \qquad \min_{\pi\in\mathcal{S}_n} \max_{1\le j\le n} \mid \lambda_j - \mu_{\pi(j)} \mid < \varepsilon.$$

Pritom je $\mathcal{S}_n$ grupa permutacija skupa $\{1, \ldots, n\}$.

Dokaz. Neka su $\lambda_{s_1}, \ldots, \lambda_{s_m}$ međusobno različite vrijednosti iz niza $\lambda_1, \ldots, \lambda_n$. Ako je $m > 1$, tada je dobro definiran broj

$$\nu = \frac{1}{2} \min_{i \neq j} \mid \lambda_{s_i} - \lambda_{s_j} \mid > 0.$$

Definirajmo otvorene krugove

$$\mathcal{K}_i(\varepsilon) = \{z \in \mathbf{C} \, ; \mid z - \lambda_{s_i} \mid < \varepsilon\}, \quad 1 \leq i \leq m.$$

Neka je $\varepsilon < \nu$. Tada su krugovi $\mathcal{K}_i(\varepsilon)$ međusobno disjunktni, a skup (unija kružnica)

$$\partial\mathcal{K}(\varepsilon) = \bigcup_{i=1}^{m} \partial\mathcal{K}_i(\varepsilon)$$

je omeđen i zatvoren, dakle kompaktan. Stoga su dobro definirani brojevi

$$\begin{aligned} m_\varepsilon &= \min_{z \in \partial\mathcal{K}(\varepsilon)} \mid p(z) \mid \\ M_\varepsilon &= \max_{z \in \partial\mathcal{K}(\varepsilon)} \{1 + \mid z \mid + \cdots + \mid z \mid^{n-1}\}. \end{aligned}$$

Očito je $m_\varepsilon > 0$ i $1 \leq M_\varepsilon < \infty$. Definirajmo

$$\delta = \frac{m_\varepsilon}{M_\varepsilon}.$$

Neka vrijedi pretpostavka $\mid a_j - b_j \mid < \delta$ za sve $0 \leq j \leq n-1$. Zbog

$$\begin{aligned} \mid q(z) - p(z) \mid &\leq \mid b_{n-1} - a_{n-1} \mid\mid z \mid^{n-1} + \cdots + \mid b_1 - a_1 \mid\mid z \mid + \mid b_0 - a_0 \mid \\ &\leq \max_{0 \leq j \leq n-1} \mid a_j - b_j \mid \ \{1 + \mid z \mid + \cdots + \mid z \mid^{n-1}\} \end{aligned}$$

imamo

$$\mid q(z) - p(z) \mid < \delta \, M_\varepsilon = m_\varepsilon \leq \mid p(z) \mid, \ z \in \partial\mathcal{K}(\varepsilon).$$

Dakle je $|q(z) - p(z)| < |p(z)|$ na svakoj kružnici $\partial\mathcal{K}_i(\varepsilon)$. Prema Rouchéovu teoremu polinom p ima u svakom krugu $\mathcal{K}_i(\varepsilon)$ isto toliko nula koliko i $p + (q - p) = q$, pa je teorem dokazan za $m > 1$ i $\varepsilon < \nu$. Ako je $\varepsilon \geq \nu$, tada postoji δ takav da implikacija vrijedi za $\nu/(\nu + 1)$ umjesto ε, pa će pogotovu vrijediti za ε.

Ako je $m = 1$, tada imamo samo jedan krug pa je $\partial\mathcal{K}(\varepsilon) = \partial\mathcal{K}_1(\varepsilon)$. Na isti način se definiraju m_ε, M_ε i δ, te zaključi da p i q imaju isti broj nula u $\mathcal{K}_1(\varepsilon)$. ■

Korolar 4.4 *Vlastite vrijednosti matrice su neprekidne funkcije matrice.*

Dokaz. Prema teoremu 4.3, vlastite vrijednosti matrice su (kao nule karakterističnog polinoma) neprekidne funkcije koeficijenata karakterističnog polinoma. Koeficijenti karakterističnog polinoma su određene sume produkata matričnih elemenata, pa su neprekidne funkcije matrice. Stoga je i svaka vlastita vrijednost matrice, kao kompozicija navedenih neprekidnih funkcija neprekidna funkcija matrice. Shematski to bi opisali dijagramom

$$A \xrightarrow{\Phi} (a_{n-1}, \ldots, a_0) \xrightarrow{\nu_i} \lambda_i$$

u kojem ν_i predstavlja funkciju koja polinomu pridružuje i-tu nultočku, za neki definirani poredak nultočaka. ■

Korolar 4.4 garantira opće kvalitativno ponašanje vlastitih vrijednosti. Kako se iz primjera 4.1 vidi, za dani ε veličina promjene matričnih elemenata δ, koja garantira promjenu vlastitih vrijednosti ne veću od ε, može jako varirati (npr. od ε^n do ε). Eksplicitne ograde za promjenu vlastitih vrijednosti, u ovisnosti od matrice bit će dokazane u Elsnerovu teoremu. Prvo uvodimo oznake.

Definicija 4.5 *Neka matrice A i B imaju vlastite vrijednosti $\lambda_1, \ldots, \lambda_n$ i $\mu_1, \ldots, \mu_n$, respektivno.*

- *Spektralna varijacija matrice A s obzirom na matricu B je broj*

$$sv_A(B) = \max_{1\le i\le n} \min_{1\le j\le n} \mid \mu_i - \lambda_j \mid \tag{4.1}$$

- *Hausdorffova udaljenost između spektara matrica A i B je broj*

$$hd(A, B) = \max\{sv_A(B), sv_B(A)\} \tag{4.2}$$

- *Slagajuća udaljenost između spektara matrica A i B je broj*

$$md(A, B) = \min_{\pi\in S_n} \max_{1\le i\le n} \mid \mu_{\pi(i)} - \lambda_i \mid \tag{4.3}$$

gdje S_n označava grupu svih permutacija skupa $\{1, \ldots, n\}$.

Funkcija $\mathrm{sv}_A(B)$ nije metrika. Ona može biti nula i kad A i B imaju različite spektre (kao na primjer u slučaju $n = 2$: $\lambda_1 = \mu_1 = \mu_2 \neq \lambda_2$). Geometrijski, $\mathrm{sv}_A(B)$

ima slijedeću interpretaciju: ako načinimo zatvorene krugove radijusa $\mathrm{sv}_A(B)$ oko spektra od A

$$\mathcal{D}_i = \{\zeta; \mid \zeta - \lambda_i \mid \leq \mathrm{sv}_A(B)\}, \quad 1 \leq i \leq n,$$

tada je

$$\lambda(B) \subset \bigcup_{i=1}^{n} \mathcal{D}_i.$$

To specijalno znači da $\mathrm{sv}_A(B) = 0$ povlači $\lambda(B) \subseteq \lambda(A)$.

Hausdorffova udaljenost je metrika. (vidi zadatak 1). Uočimo da je $\mathrm{hd}(A, B) = 0$ ispunjeno ako i samo ako vrijedi $\lambda(B) \subseteq \lambda(A)$ i $\lambda(A) \subseteq \lambda(B)$ tj. $\lambda(B) = \lambda(A)$.

Slagajuća udaljenost je također metrika. Ona nije manja od Hausdorffove udaljenosti. U perurbacijskoj analizi najljepši su rezultati kod kojih je ograda za $\mathrm{md}(A, B)$ mala.

Teorem 4.6 (Elsner) *Za bilo koje matrice A i E vrijedi*

$$hd(A, \tilde{A}) \leq \left(\|A\|_2 + \|\tilde{A}\|_2\right)^{1-\frac{1}{n}} \|E\|_2^{\frac{1}{n}}. \tag{4.4}$$

Dokaz. S obzirom da je desna strana u relaciji (4.4) simetrična u varijablama A i $\tilde{A}$, dovoljno je dokazati da je desna strana međa za $\mathrm{sv}_A(\tilde{A})$. Pretpostavimo da se maksimum u (4.1) dostiže za vlastitu vrijednost $\tilde{\lambda} \in \lambda(\tilde{A})$. Odaberimo ortonormiranu bazu $\{u_1, \ldots, u_n\}$ u $\mathbf{C}^n$, ali tako da je prvi vektor u_1 vlastiti za $\tilde{\lambda}$, tj. tako da vrijedi

$$\tilde{A}u_1 = \tilde{\lambda}u_1.$$

Tada je

$$\begin{aligned}
\mathrm{sv}_A^n(\tilde{A}) &\leq \prod_{i=1}^{n} \mid \lambda_i - \tilde{\lambda} \mid = \mid \det(A - \tilde{\lambda}I) \mid \\
&\leq \prod_{i=1}^{n} \|(A - \tilde{\lambda}I)u_i\|_2 \quad (\text{ Hadamardova nejednakost }) \\
&= \|(A - \tilde{A})u_1\|_2 \prod_{i=2}^{n} \|(A - \tilde{\lambda}I)u_i\|_2 \\
&\leq \|E\|_2 \left(\|A\|_2 + \|\tilde{A}\|_2\right)^{n-1},
\end{aligned} \tag{4.5}$$

pa dokaz slijedi uzimanjem n - tog korijena. ■

Uočimo da dokaz teorema dopušta i malo profinjenje. Ako je $\tilde{\lambda}$ vlastita vrijednost od $\tilde{A}$ geometrijske kratnosti $\tilde{n}$ i ako uzmemo bazu $\{u_1, \ldots, u_n\}$ tako da prvih $\tilde{n}$ vektora baze razapinje vlastiti potprostor od $\tilde{A}$, tada jednakost (4.5) vrijedi i ako izlučimo $\tilde{n}$ faktora. Stoga se dobije

$$\begin{aligned} \mathrm{sv}_A^n(\tilde{A}) &\leq \|(A-\tilde{A})u_1\|_2 \cdots \|(A-\tilde{A})u_{\tilde{n}}\|_2 \prod_{i=\tilde{n}+1}^{n} \|(A-\tilde{\lambda}I)u_i\|_2 \\ &\leq \|E\|_2^{\tilde{n}} \left(\|A\|_2 + \|\tilde{A}\|_2\right)^{n-\tilde{n}}. \end{aligned}$$

Uzimanjem n - tog korijena, i zamjenom $\mathrm{sv}_A(\tilde{A})$ s $\mathrm{hd}(A, \tilde{A})$, dobije se

$$\mathrm{hd}(A, \tilde{A}) \leq \left(\|A\|_2 + \|\tilde{A}\|_2\right)^{1-\frac{\tilde{n}}{n}} \|E\|_2^{\frac{\tilde{n}}{n}}. \tag{4.6}$$

S obzirom da je

$$\max\{\|A\|_2\,,\, \|\tilde{A}\|_2\} \leq \max\{\|A+\tau E\|_2\,;\, 0 \leq \tau \leq 1\},$$

iz Elsnerova teorema zaključujemo da je

$$\mathrm{sv}_A(\tilde{A}) \leq \mu \|E\|_2^{\frac{1}{n}} \stackrel{\mathrm{def}}{=} \varepsilon$$

gdje je

$$\mu = [2\max\{\|A+tE\|_2\,;\, 0 \leq t \leq 1\}]^{1-\frac{1}{n}}.$$

Definirajmo krugove

$$\mathcal{D}_i = \{\zeta\,;\, |\, \zeta - \lambda_i \,| \leq \varepsilon\}, \quad 1 \leq i \leq n.$$

Očito je $\lambda(\tilde{A}) \subset \bigcup_{i=1}^n \mathcal{D}_i$.

Propozicija 4.7 *Ako je m krugova $\mathcal{D}_i$ izolirano od preostalih $n-m$ krugova, tada njihova unija sadrži točno m vlastitih vrijednosti od $\tilde{A}$.*

Dokaz. Prenumeracijom vlastitih vrijednosti, ako je potrebno, možemo postići da su prvih m krugova izolirani od preostalih $n-m$ krugova, dakle $\mathcal{D}_i \bigcap \mathcal{D}_j$, $1 \leq i \leq m$, $m+1 \leq j \leq n$. Definirajmo

$$\begin{aligned} A(\tau) &= \tau\tilde{A} + (1-\tau)A = A + \tau E, \\ \mathcal{D}_i(\tau) &= \left\{\zeta; |\, \zeta - \lambda_i \,| \leq \tau^{\frac{1}{n}}\varepsilon\right\}, \quad 1 \leq i \leq n. \end{aligned}$$

Uočimo da je $A(0) = A$. Stoga je

$$\|A\|_2 + \|A(\tau)\|_2 \leq 2\max\{\|A+tE\|_2\,;\, 0 \leq t \leq 1\}, \quad 0 \leq \tau \leq 1,$$

pa je po teoremu 4.6

$$\mathrm{sv}_A(A(\tau)) \leq \mu\|A(\tau) - A\|_2^{\frac{1}{n}} = \mu\|\tau E\|_2^{\frac{1}{n}} = \tau^{\frac{1}{n}}\varepsilon.$$

Zaključujemo da za svako $0 \leq \tau \leq 1$, sve vlastite vrijednosti od $A(\tau)$ leže u uniji krugova $\bigcup_{i=1}^n \mathcal{D}_i(\tau)$. Jer je $\tau \mapsto \tau^{\frac{1}{n}}$ rastuća funkcija od τ, krugovi $\mathcal{D}_i(\tau)$, $1 \leq i \leq m$ su za svako $0 \leq \tau \leq 1$ odvojeni od preostalih krugova. Po korolaru 4.4 vlastite vrijednosti od $A(\tau)$ su neprekidne funkcije od τ. Označimo s $\lambda_i(\tau)$ vlastite vrijednosti od $A(\tau)$. Za $\tau = 0$ one se poklapaju sa središtima krugova $\mathcal{D}_i$: $\lambda_i(0) = \lambda_i$, $1 \leq i \leq n$. Kad bi za neko i, $1 \leq i \leq m$ i neko τ, $0 < \tau \leq 1$, $\lambda_i(\tau)$ bilo u nekom krugu $\mathcal{D}_j(\tau)$, $m+1 \leq j \leq n$, tada bi zbog neprekidnosti krivulje $\{\lambda_i(t);\, 0 \leq t \leq 1\}$ postojalo τ', $0 < \tau' < \tau$, tako da $\lambda_i(\tau')$ nije niti u $\bigcup_{i=1}^m \mathcal{D}_i$ niti u $\bigcup_{j=m+1}^n \mathcal{D}_j$. To bi onda značilo da $\lambda_i(\tau')$ nije u $\bigcup_{i=1}^n \mathcal{D}_i(\tau')$ što proturječi teoremu 4.6. ■

Sada možemo dokazati

Teorem 4.8 (Ostrowski-Elsner)

$$md(A, \tilde{A}) \leq 2(2n-1)\left(\|A\|_2 + \|\tilde{A}\|_2\right)^{1-\frac{1}{n}} \|E\|_2^{\frac{1}{n}}.$$

Dokaz. Označimo sa $\mathcal{C}_1,\ldots,\mathcal{C}_k$ komponente povezanosti skupa $\bigcup_{i=1}^n \mathcal{D}_i$. Ako je $\mathcal{C}_r$ unija od m_r krugova $\mathcal{D}_i$, tada ona sadrži m_r vlastitih vrijednosti od A i (prema propoziciji 4.7) m_r vlastitih vrijednosti od $\tilde{A}$. Odaberimo permutaciju π_r koja sparuje λ_i i $\tilde{\lambda}_{\pi_r(i)}$ za λ_i i $\tilde{\lambda}_{\pi_r(i)}$ iz $\mathcal{C}_r$. Očito je

$$|\ \lambda_i - \tilde{\lambda}_{\pi_r(i)}\ | \leq (2m_r - 1)\varepsilon \leq (2n-1)\varepsilon, \quad \lambda_i, \tilde{\lambda}_{\pi_r(i)} \in \mathcal{C}_r.$$

Ako taj postupak ponovimo za sve $1 \leq r \leq k$, dobit ćemo permutaciju π za koju će vrijediti

$$|\ \lambda_i - \tilde{\lambda}_{\pi(i)}\ | \leq (2n-1)\varepsilon, \quad 1 \leq i \leq n.$$

S obzirom da je $A + \tau E = \tau\tilde{A} + (1-\tau)A$ imamo

$$\|A + \tau E\|_2 \leq \tau\|\tilde{A}\|_2 + (1-\tau)\|A\|_2 \leq \max\{\|\tilde{A}\|_2, \|A\|_2\} \leq \|A\|_2 + \|\tilde{A}\|_2.$$

Dakle je

$$\mathrm{md}(A,\tilde{A}) \leq (2n-1)\left[2(\|A\|_2+\|\tilde{A}\|_2)\right]^{1-\frac{1}{n}}\|E\|_2^{\frac{1}{n}},$$

što povlači tvrdnju teorema. ■

Kasnije je Elsner poboljšo faktor $2(2n-1)$ na $2[n/2]$, gdje je $[n/2]$ cijeli dio od $n/2$.

Teoremi Bauera - Fikea i Henricia

Za tri norme ρ, μ i ν na prostorima $\mathbf{C}^{m\times n}$, $\mathbf{C}^{m\times k}$ i $\mathbf{C}^{k\times n}$ se kaže da su konzistentne ako vrijedi

$$\rho(AB) \leq \mu(A)\nu(B), \quad A \in \mathbf{C}^{m\times k},\ B \in \mathbf{C}^{k\times n}.$$

Npr. vektorska Euklidska norma je konzistentna sa matričnom spektralnom normom (i svakom koja majorizira spektralnu npr. Frobeniusovom normom). Specijalno, norma $\|\cdot\|$ na $\mathbf{C}^{n\times n}$ je konzistentna ako je $\|AB\| \leq \|A\|\,\|B\|$. Takve norme npr. majoriziraju spektralni radius matrice. Tipični predstavnici konzistentnih normi su operatorske norme koje su definirane za razne vektorske norme.

Teorem 4.9 (Bauer-Fike) *Neka je Q nesingularna matrica i $\|\cdot\|$ konzistentna norma. Ako $\tilde{\lambda} \in \lambda(\tilde{A})$ nije vlastita vrijednost od A, tada je*

$$\|Q^{-1}(\tilde{A}-\tilde{\lambda}I)^{-1}Q\|^{-1} \leq \|Q^{-1}EQ\|. \tag{4.7}$$

Dokaz. Vrijedi

$$\begin{aligned} Q^{-1}(\tilde{A}-\tilde{\lambda}I)Q &= Q^{-1}\left[(A-\tilde{\lambda}I)+E\right]Q \\ &= Q^{-1}(A-\tilde{\lambda}I)\left[I+(A-\tilde{\lambda}I)^{-1}E\right]Q \\ &= Q^{-1}(A-\tilde{\lambda}I)QQ^{-1}\left[I+(A-\tilde{\lambda}I)^{-1}QQ^{-1}E\right]Q \\ &= Q^{-1}(A-\tilde{\lambda}I)Q\left[I+Q^{-1}(A-\tilde{\lambda}I)^{-1}Q\left(Q^{-1}EQ\right)\right]. \end{aligned}$$

Jer je $\tilde{A}-\tilde{\lambda}I$ singularna, a $A-\tilde{\lambda}I$ nesingularna mora biti zadnji izraz u uglatim zagradama singularna matrica. Kako je $I+S$ regularna za $\|S\|<1$, mora biti

$$1 \leq \|Q^{-1}\left(A-\tilde{\lambda}I\right)^{-1}Q\left(Q^{-1}EQ\right)\| \leq \|Q^{-1}\left(A-\tilde{\lambda}I\right)^{-1}Q\|\|Q^{-1}EQ\|,$$

a to je ekvivalentno s (4.7). ■

Ako lijevu stranu u (4.7) smatramo nulom u slučaju da je $\tilde{\lambda} \in \lambda(A)$, tada nejednakost vrijedi za sve vlastite vrijednosti iz $\lambda(\tilde{A})$.

Znamo da se normalne matrice mogu dijagonalizirati pomoću unitarnih transformacija sličnoati. Stoga je norma trokutaste matrice bez dijagonale jedna mjera koliko je matrica daleko od skupa normalnih matrica. Za očekivati je da će sa sve većim odstupanjem od skupa normalnih matrica, perturbacijski rezultati za vlastite vrijednosti biti sve slabiji. O tome postoji značajan rezultat od P. Henricija. Prije prikaza tog rezultata uvedimo pojam odstupanja (devijacije) od normalnosti.

Definicija 4.10 *Neka je $A \in \mathbf{C}^{n\times n}$ i neka je ν norma na $\mathbf{C}^{n\times n}$. Neka je $\mathcal{U}$ skup unitarnih matrica, takvih da je U^*AU gornje-trokutasta. Za svako $U \in \mathcal{U}$ neka je $U^*AU = \Lambda_U + R_U$, gdje je R_U strogo gornje-trokutasta. Tada je ν-*odstupanje od normalnosti *matrice A broj*

$$\delta_\nu(A) \;=\; \min_{U\in\mathcal{U}} \nu(R_U).$$

Lako se pokaže da je skup $\mathcal{U}$ zatvoren, a kako je sadržan u kompaktnom skupu svih unitarnih matrica reda n, on je kompaktan. Isto tako funkcija $U \mapsto U^*AU$ je nprekidna, a takvo je i preslikavanje $A \mapsto A - \mathrm{diag}\,(A)$. Stoga je $\nu(R_U)$ kao kompozicija od tri neprekidne funkcije, neprekidna fukcija na kompaktnom skupu $\mathcal{U}$. To pokazuje da je broj $\delta_\nu(A)$ u definiciji4.10 dobro definiran. Taj broj je nula ako i samo ako je A normalna matrica. Zaista, ako je $\delta_\nu(A) = 0$, tada je $U^*AU = \Lambda_U$ za neko $U \in \mathcal{U}$, pa je A normalna. Obratnu tvrdnju je još lakše dokazati jer je za normalnu matricu Schurova forma dijagonalna pa je $R_U = 0$. To znači da je $\delta_\nu(A)$ $= 0$.

Odstupanje od normalnosti nije lako izračunati jer Schurova forma nije jedinstvena. Međutim za Frobeniusovu normu situacija se bitno pojednostavljuje.

Propozicija 4.11 *Za kvadratnu matricu A sa vlastitim vrijednostima λ_i, $1 \le i \le n$, vrijedi*

$$\delta_F(A) = \sqrt{\|A\|_F^2 - \sum_{i=1}^{n} \mid \lambda_i \mid^2}.$$

Dokaz. Za svaku matricu $U \in \mathcal{U}$ vrijedi

$$\|A\|_F^2 = \|U^* AU\|_F^2 = \sum_{i=1}^{n} \mid \lambda_i \mid^2 + \|R_U\|_F^2,$$

pa $\|R_U\|_F$ ne ovisi o U. Iz zadnje jednadžbe odmah slijedi tvrdnja propozicije. ∎

U slijedećem teoremu se pojavljuje funkcija

$$\psi(\eta) = \frac{\eta^n}{1 + \eta + \cdots + \eta^{n-1}}, \quad \eta \geq 0. \tag{4.8}$$

i koristi

Lema 4.12 *Neka je S kvadratna matrica za koju vrijedi $S^n = 0$ i pritom je $I + S$ nesingularna matrica. Tada je*

$$(I + S)^{-1} = I - S + +S^2 - S^3 + \cdots + (-1)^{n-1} S^{n-1}.$$

Dokaz. Neka je $B = I - S + S^2 - S^3 + \cdots + (-1)^{n-1} S^{n-1}$. Množenjem matrice B sa $I + S$ dobivamo $B(I + S) = (I + S)B = I + (-1)^{n-1} S^n$. Iskoristimo li oba uvjeta na matricu S, dobivamo $(I + S)^{-1} = B$. ∎

Teorem 4.13 (Henrici) *Neka je ν norma na $\mathbf{C}^{n \times n}$, takva da je $\nu(C) \geq \|C\|_2$ za sve $C \in \mathbf{C}^{n \times n}$. Neka je su A i E matrice reda n i $\tilde{A} = A + E$. Ako je $\delta_\nu(A) > 0$, tada je*

$$\eta = \frac{sv_A(\tilde{A})}{\delta_\nu(A)} \tag{4.9}$$

vrijedi

$$\psi(\eta) \leq \frac{\|E\|_2}{\delta_\nu(A)}. \tag{4.10}$$

Dokaz. Ako je $\mathrm{sv}_A(\tilde{A}) = 0$, tvrdnja (4.10) vrijedi. Neka je $\mathrm{sv}_A(\tilde{A}) > 0$ i neka je $\tilde{\lambda} \in \lambda(\tilde{A})$ ona vlastita vrijednost kod koje se dostiže $\mathrm{sv}_A(\tilde{A})$. Dakle postoji $\lambda \in \lambda(A)$ tako da je $\delta = \mid \tilde{\lambda} - \lambda \mid = \mathrm{sv}_A(\tilde{A})$. Neka je $U^* AU = \Lambda + R$ Schurova forma od A. Tada je $U^*(A - \tilde{\lambda} I)U = \Lambda - \tilde{\lambda} I + R$. Primijenimo li tvrdnju (4.7) Bauer-Fikeovog teorema uz $Q = U$, dobit ćemo

$$\| \left(\Lambda - \tilde{\lambda} I + R\right)^{-1} \|_2^{-1} \leq \|E\|_2. \tag{4.11}$$

Neka je $S = (\Lambda - \tilde{\lambda}I)^{-1}R$. Jer je R, pa zato i S, strogo gornje-trokutasta imamo $S^n = 0$. Uočimo da je $\Lambda - \tilde{\lambda}I + R = \left(\Lambda - \tilde{\lambda}I\right)(I+S)$, odakle slijedi da je $I+S$ nesingularna. Primjenom lemme 4.12 dobivamo

$$\left(\Lambda - \tilde{\lambda}I + R\right)^{-1} = \left(I - S + S^2 - S^3 + \cdots + (-1)^{n-1}S^{n-1}\right)\left(\Lambda - \tilde{\lambda}I\right)^{-1}.$$

Uočimo da je $\|\Lambda - \tilde{\lambda}I\|_2^{-1} = \delta^{-1}$ i $\|R\|_2 \leq \nu(R) = \delta_\nu(A)$. Stoga je

$$\|\left(\Lambda - \tilde{\lambda}I + R\right)^{-1}\|_2 \leq \delta^{-1}\left\{1 + \delta^{-1}\delta_\nu(A) + \cdots + [\delta^{-1}\delta_\nu(A)]^{n-1}\right\}$$

pa je

$$\|\left(\Lambda - \tilde{\lambda}I + R\right)^{-1}\|_2^{-1} \geq \frac{\delta}{1 + \frac{\delta_\nu(A)}{\delta} + \cdots + \frac{\delta_\nu^{n-1}(A)}{\delta^{n-1}}} \tag{4.12}$$

Povezujući relacije (4.11) i (4.12), nakon dijeljenja sa $\delta_\nu(A)$ dobivamo

$$\frac{\frac{\delta}{\delta_\nu(A)}}{1 + \left(\frac{\delta}{\delta_\nu(A)}\right)^{-1} + \cdots + \left(\frac{\delta}{\delta_\nu(A)}\right)^{-(n-1)}} \leq \frac{\|\left(\Lambda - \tilde{\lambda}I + R\right)^{-1}\|_2^{-1}}{\delta_\nu(A)} \leq \frac{\|E\|_2}{\delta_\nu(A)}.$$

S obzirom da je razlomak na lijevoj strani prve nejednakosti jednak $\psi(\eta)$, teorem slijedi. ■

Henricijev rezultat je važan jer pretstavlja neprekinutu sponu između dva primjera na početku ovog poglavlja. Ako je η iz dokaza teorema 4.13 malen, tada je $\psi(\eta) \approx \eta^n$, odakle asimptotski vrijedi

$$\frac{\mathrm{sv}_A(\tilde{A})}{\delta_\nu(A)} \mathrel{\tilde{<}} \left(\frac{\|E\|_2}{\delta_\nu(A)}\right)^{\frac{1}{n}}.$$

Ako je η veliko, tada je $\psi(\eta) \approx \eta$ pa tvrdnja (4.10) poprima asimptotski oblik

$$\mathrm{sv}_A(\tilde{A}) \mathrel{\tilde{<}} \|E\|_2.$$

Egzaktne ocjene su dokazane u slijedećem korolaru.

Korolar 4.14 **(i)** *Ako je* $\dfrac{\|E\|_2}{\delta_\nu(A)} < \dfrac{1}{n}$, *tada je*

$$\frac{sv_A(\tilde{A})}{\delta_\nu(A)} \leq \left(\frac{n\|E\|_2}{\delta_\nu(A)}\right)^{\frac{1}{n}} \tag{4.13}$$

(ii) *Ako je* $\dfrac{\|E\|_2}{\delta_\nu(A)} > 1$, *tada je*

$$sv_A(\tilde{A}) \le \|E\|_2 + \delta_\nu(A). \tag{4.14}$$

Dokaz. (i) Uočimo da $\psi(\eta) < 1/n$ povlači $\eta < 1$. Doista, u protivnom bi $\eta \ge 1$ dalo

$$\frac{\eta}{n} = \frac{\eta^n}{n\eta^{n-1}} < \psi(\eta) < \frac{1}{n}$$

tj. $\eta < 1$. Dakle $\eta \ge 1$ vodi u kntradikciju pa mora biti $\eta < 1$. S obzirom da uvjet $\|E\|_2/\delta_\nu(A) < 1/n$ povlači $\psi(\mathrm{sv}_A(\tilde{A})/\delta_\nu(A)) < 1/n$, zaključujemo da je $\mathrm{sv}_A(\tilde{A})/\delta_\nu(A) < 1$. Sada je

$$\frac{\left(\frac{\mathrm{sv}_A(\tilde{A})}{\delta_\nu(A)}\right)^n}{n} \le \psi\left(\frac{\mathrm{sv}_A(\tilde{A})}{\delta_\nu(A)}\right) \le \frac{\|E\|_2}{\delta_\nu(A)}$$

odakle slijedi tvrdnja (4.13).

(ii) Uočimo da je $\psi(\eta) = \eta^n \frac{\eta-1}{\eta^n-1}$ rastuća funkcija. Stoga je i inverzna funkcija $\psi^{-1}(\eta)$ rastuća. Funkcija ψ je rasuća funkcija čiji graf polazi iz ishodišta, prolazi točkom $(1, 1/n)$, a između tih točaka ponaša se slično funkciji $y = \eta^n$. Za veće vrijednosti varijable η ima asimptotu $y = \eta - 1$ kojoj prilazi odozgo. To slijedi zbog ocjene

$$\psi(\eta) = \frac{\eta^n}{1 + \eta + \cdots + \eta^{n-1}} = \frac{\eta}{1 + \eta^{-1} + \cdots + \eta^{-(n-1)}} > \frac{\eta}{\frac{1}{1-\eta^{-1}}} = \eta(1 - \eta^{-1}) = \eta - 1,$$

koja vrijedi za $\eta > 1$. Dakle, graf funkcije ψ posve je sadržan unutar pruge određene pravcima $y = \eta$, $y = \eta - 1$ i $y = 0$. Zato zaključujemo da je graf funkcije ψ^{-1} sadržan unutar pruge određene pravcima $y = \eta$, $y = \eta + 1$ i $\eta = 0$. Stoga vrijedi $\psi^{-1}(\eta) < \eta + 1$, $\eta \ge 0$.

Pođimo od Henricijevog teorema, tj. od $\psi(\mathrm{sv}_A(\tilde{A})/\delta_\nu(A)) \le \|E\|_2/\delta_\nu(A)$. Primijenjujući ψ^{-1} na lijevu i desnu stranu te nejednakosti i koristeći monotonost te funkcije dobivamo

$$\frac{\mathrm{sv}_A(\tilde{A})}{\delta_\nu(A)} \le \psi^{-1}\left(\frac{\|E\|_2}{\delta_\nu(A)}\right) \le \frac{\|E\|_2}{\delta_\nu(A)} + 1.$$

Dakle smo dobili

$$\mathrm{sv}_A(\tilde{A}) \le \|E\|_2 + \delta_\nu(A).$$

Time je dokazana i druga tvrdnja. ∎

Korolar 4.15 *Neka je ψ definirana kao u relaciji (4.8). Tada je*

$$md(A, \tilde{A}) \leq (2n-1)\delta_\nu(A)\psi^{-1}\left(\frac{\|E\|_2}{\delta_\nu(A)}\right).$$

Dokaz. Uočimo da $\delta_\nu(A)\psi^{-1}\left(\frac{\|\tau E\|_2}{\delta_\nu(A)}\right)$ kao je ograda za $\mathrm{sv}_A(A+\tau E)$ monotona u τ. Stoga dokaz slijedi isti argument kao i dokaz teorema 4.8. ■

Nezgoda u dosadašnjim rezultatima je pojavljivanje n-tog korijena u perturbacijskim ogradama. Slijedeći teorem pokazuje da za matricu sa manjim elementarnim Jordanovim blokovima ocjena sadrži i manji korijen. Tim teoremom se odmičemo od unitarnih transformacija sličnosti i Schurove forme.

Teorem 4.16 *Neka je $Q^{-1}AQ = J$ Jordanova forma od A i neka je m red najvećeg elementarnog Jordanovog bloka u J. Tada za svaku vlastitu vrijednost $\tilde{\lambda} \in \lambda(\tilde{A})$ postoji vlastita vrijednost $\lambda \in \lambda(A)$, takva da je*

$$\frac{|\tilde{\lambda}-\lambda|^m}{1+|\tilde{\lambda}-\lambda|+\cdots+|\tilde{\lambda}-\lambda|^{m-1}} \leq \|Q^{-1}EQ\|_2. \tag{4.15}$$

Dokaz. Dokaz je sličan dokazu Henricijevog teorema. Polazimo od relacije $Q^{-1}(A-\tilde{\lambda}I)Q = \Lambda - \tilde{\lambda}I + J$, gdje je Q nesingularna matrica, a $\Lambda + J$ je direktna suma elementarnih Jordanovih blokova. Na sličan način kao prije, dolazimo do relacije

$$\|\left(\Lambda - \tilde{\lambda}I + J\right)^{-1}\|_2^{-1} \leq \|Q^{-1}EQ\|_2.$$

Uz oznake $\delta = |\tilde{\lambda}-\lambda| = \min\{|\tilde{\lambda}-\lambda_i|;\ 1 \leq i \leq n\}$ i $S = (\Lambda - \tilde{\lambda}I)^{-1}J$, slično kao prije dobijemo

$$\left(\Lambda - \tilde{\lambda}I + J\right)^{-1} = \left(I - S + S^2 - S^3 + \cdots + (-1)^{m-1}S^{m-1}\right)\left(\Lambda - \tilde{\lambda}I\right)^{-1}.$$

Ovdje smo uvažili pretpostavku od maksimalnom redu elementarnog Jordanovog bloka. S obzirom da je J nul-matrica izuzev jedinica na sporednoj dijagonali imamo $\|J\|_2 = 1$. Stoga uzimanjem spektralne norme lijeve i desne strane zadnje relacije, dobivamo

$$\|\left(\Lambda - \tilde{\lambda}I + J\right)^{-1}\|_2 \leq \delta^{-1}\left[1 + \delta^{-1} + \cdots + \delta^{-m+1}\right]$$

Uzimanjem inverza lijeve i desne strane, slijedi

$$\frac{\delta^m}{1+\delta+\cdots+\delta^{m-1}} \leq \|\left(\Lambda - \tilde{\lambda}I + J\right)^{-1}\|_2^{-1} \leq \|Q^{-1}EQ\|_2,$$

što je i trebalo dokazati. ■

Specijalno, ako se matrica A dade dijagonalizirati pomoću transformacije sličnosti tada vrijedi

Korolar 4.17 *Neka je Jordanova forma matrice A dijagonalna, dakle $Q^{-1}AQ = \Lambda$. Tada za svaku vlastitu vrijednost $\tilde{\lambda}$ matrice $\tilde{A}$ postoji vlastita vrijednost λ matrice A, takva da je*

$$\mid \tilde{\lambda} - \lambda \mid \leq \|Q^{-1}EQ\|_2.$$

Dokaz. Dovoljno je u iskazu teorema 4.16 dodatno pretpostaviti $m = 1$. ■

U teoremu 4.16 i korolaru 4.17 javlja se ograda $\|Q^{-1}EQ\|_2$. Uočimo da je $\|Q^{-1}EQ\|_2 \leq \|Q\|_2\|Q^{-1}\|_2\|E\|_2$. U praksi se rijetko susreću matrice koje nisu hermitske, a još mnogo rjeđe matrice koje imaju netrivijalnu Jordanovu formu. Stoga se za matrice koje se daju dijagonalizirati uvodi *broj uvjetovanosti* za problem vlastitih vrijednosti: $\kappa_2(Q) = \|Q\|_2\|Q^{-1}\|_2$. Ako je matrica normalna tada je taj broj 1, a korolar 4.17 se svodi na

Korolar 4.18 *Neka je A normalna matrica. Tada za svaku vlastitu vrijednost $\tilde{\lambda}$ matrice $\tilde{A}$ postoji vlastita vrijednost λ matrice A, takva da je*

$$\mid \tilde{\lambda} - \lambda \mid \leq \|E\|_2.$$

Dokaz. Dovoljno je u iskazu teorema 4.17 dodatno pretpostaviti $Q = U$ i iskoristiti unitarnu invarijantnost spektralne norme. ■

Uočimo da do istog rezultata vodi i dokaz Henricijevog teorema. Dakle za normalne matrice Henricijev teorem prelazi u korolar 4.18.

4.2 Hermitske matrice

U ovom poglavlju često ćemo koristiti slijedeće oznake: $A = A^* \in \mathbf{C}^{vn\times n}, \tilde{A} = A + E, E = E^* \in \mathbf{C}^{vn\times n}$, dok ćemo svojstvene vrijednosti označavati respektivno s $\lambda_1 \geq \cdots \geq \lambda_n, \tilde{\lambda}_1 \geq \cdots \geq \tilde{\lambda}_n$.

4.2.1. Inercija i djeljenje spektra

Transformacija kongruencije $A \to X^*AX$ čuva hermitičnost, ali ne i spektar matrica. Ipak vrijedi

Teorem 4.19 (Sylvester, Jacobi) *Neka je A hermitska matrica i neka je inercija od A uređena trojka $i(A) = [\pi(A), \nu(A), \rho(A)]$, gdje su $\nu(A), \pi(A)$ i $\rho(A)$ respektivno broj negativnih, broj pozitivnih i broj nula svojstvenih vrijednosti od A. Tada za svaku nesingularnu matricu X vrijedi $i(X^*AX) = i(A)$.*

Dokaz. Dokaz ide od suprotnog. Pretpostavimo da A ima više pozitivnih svojstvenih vrijednsti od X^*AX. Neka je $\mathcal{Y}$ potprostor razapet svojstvenim vektorima koji pripadaju pozitivnim svojstvenim vrijednostima. To znači da za $y \in \mathcal{Y} \setminus \{0\}$ vrijedi $y = \sum_i \alpha_i x_i, Ax_i = \lambda_i x_i, \lambda_i > 0$, pa je

$$y^*Ay = (\sum_i \alpha_i Ax_i | \sum_j \alpha_j x_j) = \sum_i \sum_j \alpha_i \bar{\alpha}_j \lambda_i \delta_{ij} = \sum_i \lambda_i \alpha_i^2 > 0$$

jer je bar jedan $\alpha_i \neq 0$. Dakle vrijedi implikacija

$$y \in \mathcal{Y} \setminus \{0\} \Rightarrow y^*Ay > 0.$$

Neka je $\mathcal{Z}$ potprostor razapet vektorima oblika Xz_i, gdje je z_i svojstveni vektor koji odgovara svojstvenim vrijednostima $\lambda_i \leq 0$ matrice X^*AX. Tada je

$$\begin{aligned} z \in \mathcal{Z} \quad &\Rightarrow \quad z = X\sum_i \alpha_i z_i \Rightarrow z^*Az = \\ &= \quad \sum_i \bar{\alpha}_i z_i^* X^*AX \sum_i \alpha_i z_i = \sum_i \bar{\alpha}_i z_i^* \alpha_i \lambda_i z_i = \sum_i \lambda_i |\alpha_i|^2 \leq 0 \end{aligned}$$

Dakle je $\mathcal{Y} \cap \mathcal{Z} = 0$ pa je $dim\mathcal{Y} + dim\mathcal{Z} \leq n$. Međutim po hipotezi je $dim\mathcal{Y} + dim\mathcal{Z} > n$. Dakle,treba odbaciti hipotezu da A ima više pozitivnih svojstvenih vrijednosti od X^*AX. Jer je $A = Y^*(X^*AX)Y$ uz $Y = X^{-1}$ zaključujemo da A i X^*AX imaju isti broj pozitivnih svojstvenih vrijednosti. One moraju imati i isti broj negativnih svojstvenih vrijednosti jer je $\pi(-X^*AX) = \nu(X^*AX)$ i $X^*AX = X^*(-A)X$, pa automatski vrijedi i $\rho(A) = \rho(X^*AX)$. ■

Za posljedicu imamo

Teorem 4.20 (Cauchy) *Neka je B glavna podmatrica od A reda $n-1$ sa svojstvenim vrijednostima $\mu_1 \geq \cdots \geq \mu_{n-1}$. Tada je*

$$\lambda_1 \geq \mu_1 \geq \lambda_2 \geq \mu_2 \geq \cdots \geq \mu_{n-1} \geq \lambda_n.$$

Dokaz. Bez gubitka općenitosti možemo pretpostaviti da je B vodeća glavna podmatrica od A, tako da je $A = \begin{bmatrix} B & a \\ a^* & \alpha \end{bmatrix}$. Pretpostavimo da teorem ne vrijedi. Tada je za neko i $\mu_i > \lambda_i$ ili $\lambda_{i+1} > \mu_i$. Promotrimo slučaj $\mu_i > \lambda_i$.
Neka je $\mu_i > \tau > \lambda_i$. Tada je $B - \tau I$ nesingularna i matrica

$$\begin{aligned} H &= \begin{bmatrix} B - \tau I & 0 \\ 0 & \alpha - \tau - a^*(B-\tau I)^{-1}a \end{bmatrix} \\ &= \begin{bmatrix} I & 0 \\ -a^*(B-\tau I)^{-1} & 1 \end{bmatrix} \begin{bmatrix} B - \tau I & a \\ a^* & \alpha - \tau \end{bmatrix} \begin{bmatrix} I & -(B-\tau I)^{-1}a \\ 0 & 1 \end{bmatrix} \end{aligned}$$

je kongruentna s $A - \tau I$. Dakle je $\pi(H) = \pi(A - \tau I) \leq i - 1$. Međutim, H ima barem toliko pozitivnih svojstvenih vrijednosti koliko ima $B - \tau I$ a to je i. Dakle slučaj da je $\mu_i > \lambda_i$ vodi u kontradikciju.
U slučaju $\lambda_{i+1} > \mu_i$ opet se uzme τ tako da $\lambda_{i+1} > \tau > \mu_i$ pa se zaključi da je $\pi(H) = \pi(A - \tau I) \geq i + 1$ odnosno $\pi(H) \leq \pi(B - \tau I) + 1 \leq i - 1 + 1 = i$. ■

Ako je u teoremu 4.20 C glavna podmatrica od A reda $n-2$ tada svojstvene vrijednosti od C, $\nu_1 \geq \cdots \geq \nu_{n-2}$ zadovoljavaju

$$\mu_1 \geq \nu_1 \geq \mu_2 \geq \nu_2 \geq \cdots \geq \nu_{n-2},$$

stoga je

$$\lambda_i \geq \mu_i \geq \nu_i \geq \mu_{i+1} \geq \lambda_{i+2}\,,\; i = 1, \ldots, n-2.$$

Nastavimo li ovako zaključivanje dolazimo do

Korolar 4.21 *Neka je B glavna podmatrica od A reda $n-k$ sa svojstvenim vrijednostima $\mu_1 \geq \cdots \geq \mu_{n-k}$. Tada je*

$$\lambda_i \geq \mu_i \geq \lambda_{i+k}\,,\ i=1,\ldots,n-k.$$

Dokaz. Dokazujemo indukcijom. Za $k=1$ imamo teorem 4.20. Ako tvrdnja vrijedi za k, tad za $k+1$ mora vrijediti $\lambda_i \geq \mu_i \geq \nu_i \geq \mu_{i+1} \geq \lambda_{i+k+1}$, to jest $\lambda_i \geq \nu_i \geq \lambda_{i+(k+1)}\,,\ i=1,\ldots,n-(k+1)$. ■ ■

Primjetimo da rezultat o preplitanju spektara možemo poopćiti. Neka $U \in \mathbf{C}^{n\times(n-k)}$ ima ortonormirane stupce. Neka je V takva da je (U,V) unitarna matrica. Primjenjujući korolar 4.21 na matricu $(U,V)^*A(U,V)$ dobivamo

Korolar 4.22 *Neka $U \in \mathbf{C}^{n\times(n-k)}$ ima ortonormirane stupce. Neka su svojstvene vrijednosti od U^*AU $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_{n-k}$. Tada je*

$$\lambda_i \geq \mu_i \geq \lambda_{i+k}\,,\ i=1,\ldots,n-k+1.$$

Dokaz.

$$\acute{A} = \begin{bmatrix} U^* \\ V^* \end{bmatrix} A \begin{bmatrix} U & V \end{bmatrix} = \begin{bmatrix} U^*AU & U^*AV \\ V^*AU & V^*AV \end{bmatrix}$$

Matrica $\acute{A}$ je slična s A i ima svojstvene vrijednosti $\lambda_1 \geq \cdots \geq \lambda_n$. Matrica U^*AU je glavna podmatrica od $\acute{A}$ pa se na nju primjeni korolar 4.22. ■ ■

4.2.2. Wielandtov teorem i posljedice

Znamo da je

$$\lambda_1 = \max_{x^*x=1} x^*Ax.$$

Poopćenje je Fischerov teorem koji kaže da je

$$\lambda_i = \max_{\substack{\chi \\ dim(\chi)=i}} \ \min_{\substack{x\in\chi \\ x^*x=1}} x^*Ax\ , 1 \leq i \leq n.$$

Daljnje poopćenje je Wielandtov teorem. Prije iskaza samog teorema dokazat ćemo jednu lemu.

Lema 4.23 *Neka su $A, X \in \mathbf{C}^{n\times n}$, A hermitska matrica, te neka vrijedi $X^*X = I$. Tada je $tr(X^*AX) = tr(A) = \sum \lambda$, gdje su λ svojstvene vrijednosti od A.*

Dokaz. Neka je $X = (x_{ij})$,$A = (a_{ij})$.Tada je $(AX)_{ik} = \sum_{j=1}^{n} a_{ij}x_{jk}$, i $(X^*AX)_{lk} = \sum_{i=1}^{n} x^*_{li}(AX)_{ik} = \sum_{i=1}^{n} \overline{x}_{il}(AX)_{ik}$.Računamo sada

$$\begin{aligned} tr(X^*AX) &= \sum_{l=1}^{n}(X^*AX)_{ll} = \sum_{l=1}^{n}\sum_{i=1}^{n}\overline{x}_{il}(AX)_{il} \\ &= \sum_{l=1}^{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\overline{x}_{il}a_{ij}x_{jl} = \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{l=1}^{n}\overline{x}_{il}x_{jl}a_{ij} \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}\sum_{l=1}^{n}x_{jl}x^*_{li} = \sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}\delta_{ij} \\ &= \sum_{i=1}^{n}a_{ii} = tr(A). \end{aligned}$$

Time smo dokazali prvu jednakost.Druga slijedi iz činjenice da je A hermitska pa postoji $X \in \mathbf{C}^{n\times n}$, $X^*X = I$ takva da je X^*AX dijagonalna matrica kojoj su na dijagonali svojstvene vrijednosti. ■ ■

Teorem 4.24 *(Wielandt)*

Neka je A hermitska matrica reda n i neka su njene svojstvene vrijednosti $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. Nadalje neka je $1 \leq i_1 < i_2 < \ldots < i_k \leq n$. Tada je

$$\lambda_{i_1} + \cdots + \lambda_{i_k} = \max_{\substack{\chi_{i_1} \subset \cdots \subset \chi_{i_k} \\ dim(\chi_{i_j}) = i_j}} \; \min_{\substack{X = (x_{i_1}, \ldots, x_{i_k}) \\ x_{i_j} \in \chi_{i_j} \\ X^*X = I}} tr(X^*AX) \tag{4.16}$$

i

$$\lambda_{i_1} + \cdots + \lambda_{i_k} = \min_{\substack{\chi_{i_1} \supset \cdots \supset \chi_{i_k} \\ dim(\chi_{i_j}) = n - i_j + 1}} \; \max_{\substack{X = (x_{i_1}, \ldots, x_{i_k}) \\ x_{i_j} \in \chi_{i_j} \\ X^*X = I}} tr(X^*AX). \tag{4.17}$$

Napomena 4.25 *Skup* $\mathcal{S} = \{Z \in \mathbf{C}^{n\times k}, Z = (z_1, \ldots, z_k), \ z_j \in \chi_{i_j}, Z^*Z = I\}$ *je neprazan za neki niz potprostora* $\chi_{i_1} \subset \cdots \subset \chi_{i_k}$ *koji ćemo nadalje zvati lancem. Naime, matrica čije stupce biramo na slijedeći način pripada skupu* $\mathcal{S}$*:* x_1 *neka je neki normirani vektor iz* χ_{i_1}*,*x_2 *je neki normirani vektor iz* $\chi_{i_2} \ominus \chi_{i_1}$ *,itd. Na taj način dolazimo do tražene matrice.*

Napomena 4.26 *Skup* $\mathcal{S} = \{Z \in \mathbf{C}^{n\times k}, Z = (z_1, \ldots, z_k), z_j \in \chi_{i_j}, Z^*Z = I\}$ *je kompaktan u* $\mathbf{C}^{n\times k}$*. Zaista :* $\mathcal{S}$ *je omeđen, jer je* $\|Z\|_F = \sqrt{k}, z \in \mathcal{S}$*;* $\mathcal{S}$ *je zatvoren jer* $Z_t \in \mathcal{S}, Z_t \to Z$ *povlači za sve* j $Z_te_j \to Ze_j \in \chi_{i_j}$ *jer je* χ_{i_j} *potprostor i* ${Z_t}^*Z_t \to Z^*Z$ *jer je* $X \to X^*X$ *neprekidna funkcija, pošto je* $I = {Z_t}^*Z_t$ *mora biti* $I = Z^*Z$*.*

Funkcije $B \to tr(B)$ *i* $X \to X^*AX$ *su neprekidne pa je* $X \to tr(X^*AX)$ *neprekidna funkcija.*

Stoga se $\inf\limits_{X\in\mathcal{S}} tr(X^*AX)$ *i* $\sup\limits_{X\in\mathcal{S}} tr(X^*AX)$ *u (4.16) i (4.17) mogu zamjeniti sa* $\min\limits_{X\in\mathcal{S}} tr(X^*AX)$ *i* $\max\limits_{X\in\mathcal{S}} tr(X^*AX)$*.*

Da se $\sup\limits_{\substack{\chi_{i_1}\subset\ldots\subset\chi_{i_k}\\ dim(\chi_{i_j})=i_j}}$ *odnosno* $\inf\limits_{\substack{\chi_{i_1}\supset\ldots\supset\chi_{i_k}\\ dim(\chi_{i_j})=n-i_j+1}}$ *mogu zamjeniti s* $\max\limits_{\substack{\chi_{i_1}\subset\ldots\subset\chi_{i_k}\\ dim(\chi_{i_j})=i_j}}$ *odnosno* $\min\limits_{\substack{\chi_{i_1}\supset\ldots\supset\chi_{i_k}\\ dim(\chi_{i_j})=n-i_j+1}}$ *slijedit će iz činjenice da se jednakost dostiže za barem jedan izbor potprostora* $\chi_{i_1} \subset \ldots \subset \chi_{i_k}$ *odnosno* $\chi_{i_1} \supset \ldots \supset \chi_{i_k}$*.*

Dokaz. Sa u_i označimo stupci unitarne matrice U za koju vrijedi $U^*AU = \Lambda$, gdje je Λ dijagonalna matrica.u_i su zapravo svojstveni vektori od A. Prvo dokazujemo (4.16). Dokažimo da postoji određeni niz potprostora $\chi_{i_1} \subset \ldots \subset \chi_{i_k}, dim(\chi_{i_j}) = i_j$ takav da vrijedi

$$tr(X^*AX) \geq \lambda_{i_1} + \lambda_{i_2} + \cdots + \lambda_{i_k} \tag{4.18}$$

čim je $X = (x_{i_1}, \ldots, x_{i_k}), x_{i_j} \in \chi_{i_j}, 1 \leq j \leq k$ i $X^*X = I$.

Zaista, neka je $\chi_{i_j} = \mathcal{L}(u_1, \ldots, u_{i_j}), Au_l = \lambda_l u_l, \|u_l\|_2 = 1$,$l = 1, \ldots, i_j$. Tada

$X = (x_{i_1}, \ldots, x_{i_k}), X^*X = I$ povlači

$$x_{i_j} = \sum_{r=1}^{i_j} \alpha_r^{(i_j)} u_r, \quad \sum_{r=1}^{i_j} |\alpha_r^{(i_j)}|^2 = 1, \quad 1 \leq j \leq k,$$

pa je

$$\begin{aligned} x_{i_j}^* A x_{i_j} &= x_{i_j}^* \sum_{r=1}^{i_j} \alpha_r^{(i_j)} \lambda_r u_r = \sum_{r=1}^{i_j} \lambda_r |\alpha_r^{(i_j)}|^2 \geq \\ &\geq \min_{1 \leq r \leq i_j} \lambda_r = \lambda_{i_j}, 1 \leq j \leq k. \end{aligned}$$

Dakle,

$$tr(X^*AX) = \sum_{j=1}^{k} x_{i_j}^* A x_{i_j} \geq \sum_{j=1}^{k} \lambda_{i_j}$$

vrijedi za sve $X \in \mathcal{S}$ gdje je $\mathcal{S}$ kao u Primjedbi 4.26, ali za ovaj specijalni izbor potprostora. Jer je $\mathcal{S}$ kompaktan imamo

$$\min_{\substack{X = (x_{i_1}, \ldots, x_{i_k}) \\ x_{i_j} \in \chi_{i_j} \\ X^*X = I}} tr(X^*AX) \geq \lambda_{i_1} + \cdots + \lambda_{i_k} \tag{4.19}$$

Ako odaberemo $x_{i_j} = u_{i_j} \in \chi_{i_j}$ i $X = (x_{i_1}, \ldots, x_{i_k})$ bit će $tr(X^*AX) = \lambda_{i_1} + \cdots + \lambda_{i_k}$, pa se minimum dostiže, što znači da ako dokažemo i drugu nejednakost umjesto *sup* možemo pisati *min* u relaciji (4.16).

Da bi dokazali (4.16) sada je dovoljno dokazati

$$\sup_{\substack{\chi_{i_1} \subset \ldots \subset \chi_{i_k} \\ dim(\chi_{i_j}) = i_j}} \min_{\substack{X = (x_{i_1}, \ldots, x_{i_k}),\ x_{i_j} \in \chi_{i_j} \\ X^*X = I}} tr(X^*AX) \leq \lambda_{i_1} + \lambda_{i_2} + \cdots + \lambda_{i_k}. \tag{4.20}$$

Dokaz ćemo provesti indukcijom po n.Hipoteza indukcije je da za svaku matricu A reda n i za svaki lanac $\chi_{i_1} \subset \ldots \subset \chi_{i_k}$ postoji $X \in \mathcal{S}(\chi_{i_1}, \ldots, \chi_{i_k})$ tako da vrijedi $tr(X^*AX) \leq \lambda_{i_1} + \lambda_{i_2} + \cdots + \lambda_{i_k}$, gdje su $\lambda_1, \lambda_2, \ldots, \lambda_n$ svojstvene vrijednosti od A poredano u nerastućem poretku.

Ako je $k = n$, tada je X^*AX slično s A pa tvrdnja slijedi iz leme 4.23. Stoga vrijedi i za $n = 1$, pa bazu indukcije imamo.

Pretpostavimo da tvrdnja vrijedi za sve hermitske matrice reda $n-1$, $n > 1$.

Ako je $k = n$, tada je tvrdnja dokazana.

Neka je $k < n$ i neka je $\chi_{i_1} \subset \ldots \subset \chi_{i_k}, dim(\chi_{i_j}) = i_j$ proizvoljno odabrani niz potprosrora. Ako pokažemo da postoji matrica $X \in \mathcal{S}$ ($\mathcal{S}$ nezavisan uz ovaj izbor potprostora), takva da je $tr(X^*AX) \leq \lambda_{i_1} + \cdots + \lambda_{i_k}$, tada će pogotovo vrijediti $\min_{X\in\mathcal{S}} tr(X^*AX) \leq \lambda_{i_1}+\cdots+\lambda_{i_k}$, a zbog proizvoljnosti izbora niza potprostora vrijedit će (4.20). Promatramo dva slučaja.

1. **Slučaj** $i_k < n$.

   Neka je $\tilde{\chi}_{n-1}$ neki $(n-1)$-dimenzionalni potprostor koji sadrži χ_{i_k}. Neka je $Z = (z_1, \ldots, z_{n-1})$ ortonormirana matrica, takva da je $\mathcal{R}((z_1, \ldots, z_{i_j})) = \chi_{i_j}$, $1 \leq j \leq k$ i $\mathcal{R}(Z) = \tilde{\chi}_{n-1}$.

   Neka je $B = Z^*AZ \in \mathbf{C}^{(n-1)\times(n-1)}$. Ako su $\mu_1 \geq \ldots \geq \mu_{n-1}$ svojstvene vrijednosti od B, tada zbog posljedice Cauchyjevog teorema vrijedi

   $$\mu_i \leq \lambda_i, \quad i = 1, \ldots, n-1 \tag{4.21}$$

   Uz B vežemo potprostore $\mathcal{Y}_{i_j} = \{Z^*x, x \in \chi_{i_j}\}$, $1 \leq j \leq k$. Uočimo da je $Z\mathcal{Y}_{i_j} = ZZ^*\chi_{i_j} = P_{\tilde{\chi}_{n-1}}\chi_{i_j} = \chi_{i_j}$, $\quad 1 \leq j \leq k$. Ako je $\{y_1, \ldots, y_j\}$ ortonormirana baza za $\mathcal{Y}_{i_j}$ tada je $\{Zy_1, \ldots, Zy_j\}$ ortonormirana baza za χ_{i_j}. Dakle je $dim(\mathcal{Y}_{i_j}) = dim(\chi_{i_j}) = i_j, 1 \leq j \leq k$, pa iz definicije slijedi $\mathcal{Y}_{i_1} \subset \mathcal{Y}_{i_2} \subset \ldots \subset \mathcal{Y}_{i_k}$. Jer je $i_k \leq n-1$ možemo primjeniti indukcijsku hipotezu na B. Dobivamo vektore $y_{i_j} \in \mathcal{Y}_{i_j}, 1 \leq j \leq k$, odnosno matricu $Y = (y_{i_1}, \ldots, y_{i_k})$ tako da vrijedi

   $$tr(Y^*BY) = \sum_{j=1}^{k} y_{i_j}^* B y_{i_j} \leq \mu_{i_1} + \cdots + \mu_{i_k}.$$

   Neka je $X = ZY$ tj. $x_{i_j} = Zy_{i_j} \in \chi_{i_j}$, $1 \leq j \leq k$. Tada je $Y = Z^*X$ i za ortonormiranu matricu $X \in \mathcal{S}$ vrijedi

   $$tr(Y^*BY) = tr(Y^*Z^*AZ) = tr(X^*AX),$$

   pa je

   $$tr(X^*AX) \leq \mu_{i_1} + \cdots + \mu_{i_k} \leq \lambda_{i_1} + \cdots + \lambda_{i_k}$$

što je trebalo pokazati.

2. **Slučaj** $i_k = n$.

   Neka je l najveći indeks za koji vrijedi $i_l + 1 < i_{l+1}$. Kad takav ne bi postojao onda bi bilo $k = n$, što smo već obradili. Označimo $i_l = p$ i $i_{l+1} = q$. Dakle, polazimo od potprostora

$$\chi_{i_1} \subset \chi_{i_2} \subset \ldots \subset \chi_p \subset \chi_q \subset \chi_{q+1} \subset \ldots \subset \chi_n.$$

   Neka je $\tilde{\chi}_{n-1}$ $(n-1)$-dimenzionalni potprostor koji je razapet nekom bazom od χ_p, svojstvenim vektorima od $\lambda_q, \ldots, \lambda_n$ i možda još ponekim vektorom.

   Očito vrijedi

$$\chi_p \subset \chi_q \cap \tilde{\chi}_{n-1} \subseteq \chi_{q+1} \cap \tilde{\chi}_{n-1} \subseteq \ldots \subseteq \chi_{n-1} \cap \tilde{\chi}_{n-1} \subseteq \tilde{\chi}_{n-1} = \mathbf{C}^n \cap \tilde{\chi}_n. \quad (4.22)$$

$$\begin{aligned} dim(\chi_q \cap \tilde{\chi}_{n-1}) &= dim(\tilde{\chi}_{n-1}) + dim(\chi_q) - dim(\chi_q \cup \tilde{\chi}_{n-1}) \\ &= n - 1 + q - dim(\tilde{\chi}_{n-1} \cup \chi_q) \geq n - 1 + q - n \\ &= q - 1 \end{aligned}$$

   Na sličan način se dobije $dim(\chi_i \cap \tilde{\chi}_{n-1}) \geq i - 1$ za $q \leq i \leq n-1$. Također je jasno da vrijedi $dim(\chi_i \cap \tilde{\chi}_{n-1}) \leq i$ za $q \leq i \leq n-1$. Cilj nam je modificirati niz potprostora iz (4.22) tako da oni postanu lanac. Slijedeći primjer će pokazati da do jednakosti u (4.22) može doći.

   Primjer 4.27 $A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & -1 & 1 \end{pmatrix}$ *Svojstvene vrijednosti od* A *su* $\lambda_1 = 1, \lambda_2 = 1, \lambda_3 = 0, \lambda_4 = 2$, *a pripadni svojstveni vektori su* $\mu_1 = e_1, \mu_2 = e_2, \mu_3 = e_3 + e_4, \mu_4 = e_3 - e_4$. *Neka su* $\chi_{i_1} = \chi_1 = [e_1] = \chi_p$,
   $\chi_{i_2} = \chi_3 = [e_1, e_3, e_4] = \chi_q$,
   te $\chi_{i_3} = \chi_4 = [e_1, e_2, e_3, e_4] = \mathbf{C}^4$. *Tada će vrijediti* $\chi_q \cap \tilde{\chi}_{n-1} = \chi_{q+1} \cap \tilde{\chi}_{n-1}$.

   Tvrdnja je da do jednakosti može doći najviše jednom.Da bismo to dokazali razmotrimo slučaj kada vrijedi $\chi_q \cap \tilde{\chi}_{n-1} = \chi_{q+1} \cap \tilde{\chi}_{n-1}$. Tada mora biti

$\dim(\chi_q \cap \tilde{\chi}_{n-1}) = q = \dim(\chi_{q+1} \cap \tilde{\chi}_{n-1})$. Iz $\dim(\chi_q \cup \tilde{\chi}_{n-1}) = n-1$ slijedi $\chi_q \subset \tilde{\chi}_{n-1}$. Neka je $\tilde{x} \in \mathbf{C}^n$ takav da je $[\chi_q, \tilde{x}] = \chi_{q+1}$, tada slijedi $\tilde{x} \in \tilde{\chi}_{n-1}^{\perp}$. Sada je $\tilde{x} \in \chi_{q+j}$, pa imamo $q+j-1 \leq \dim(\chi_{q+j} \cap \tilde{\chi}_{n-1}) \leq q+j-1$, dakle $\dim(\chi_{q+j} \cap \tilde{\chi}_{n-1}) = q+j-1$, što vrijedi za $j = 2, \ldots, n-q-1$. Dakle, nadalje imamo prave inkluzije. Potpuno analogno vrijedi i za preostale slučajeve, kada je na nekom drugom mjestu jednakost. Stoga je konstrukcija novog lanca slijedeća :

Ako su sve inkluzije prave , tada definiramo

$$\begin{aligned} \tilde{\chi}_{q-1} &= \chi_q \cap \tilde{\chi}_{n-1} \\ \tilde{\chi}_q &= \chi_{q+1} \cap \tilde{\chi}_{n-1} \\ &\vdots \\ \tilde{\chi}_{n-2} &= \chi_{n-1} \cap \tilde{\chi}_{n-1} \end{aligned}$$

Ako vrijedi $\chi_{q+j} \cap \tilde{\chi}_{n-1} = \chi_{q+j+1} \cap \tilde{\chi}_{n-1}$, za neki j,

tada konstruiramo $\tilde{\chi}_{q+j-1}$ tako da vrijedi $\tilde{\chi}_{q+j-2} \subset \tilde{\chi}_{q+j-1} \subset \chi_{q+j}$ i $\dim(\tilde{\chi}_{q+j-1}) = q+j-1$, a preostale potprostore definiramo kao u prethodnom slučaju.

Za ove potprostore vrijedi

$$\tilde{\chi}_{q-1} \subset \chi_q \ , \ \tilde{\chi}_q \subset \chi_{q+1} \ , \ \ldots \ , \tilde{\chi}_{n-2} \subset \chi_{n-1} \tag{4.23}$$

i

$$\chi_{i_1} \subset \ldots \subset \chi_p \subset \tilde{\chi}_{q-1} \subset \ldots \subset \tilde{\chi}_{n-2} \subset \tilde{\chi}_{n-1}. \tag{4.24}$$

Ponovimo sada konstrukciju matrice B iz slučaja $i_k < n$ za svojstvene vrijednosti $\mu_{i_1} \geq \mu_{i_2} \geq \ldots \geq \mu_{n-1}$ za koje vrijedi (4.21). Pritom možemo kao $\tilde{\chi}_{n-1}$ uzeti upravo zadnji u nizu (4.24).

Konstrukcija daje ortonormiranu matricu

$$X = (x_{i_1}, \ldots, x_{i_{l-1}}, x_p, \ldots, x_{n-1})$$

za koju je

$$x_{i_j} \in \chi_{i_j},\ 1 \leq j \leq l$$

$$x_i \in \tilde{\chi}_i \subset \chi_{i+1},\ i = q-1, \ldots, n-1$$

i vrijedi

$$tr(X^*AX) \leq \sum_{j=1}^{l} \mu_{i_j} + \sum_{i=q-1}^{n-1} \mu_i \leq \sum_{j=1}^{l} \lambda_{i_j} + \sum_{i=q-1}^{n-1} \mu_i.$$

Po konstrukciji $\tilde{\chi}_{n-1}$ sadrži svojstvene vektore od A koji pripadaju $\lambda_q, \ldots, \lambda_n$. Po konstrukciji od Z je $ZZ^* = P_{\tilde{\chi}_{n-1}}$, pa $Ay_i = \lambda_i y_i, y_i \in \tilde{\chi}_{n-1}$ povlači $y_i = ZZ^* y_i$ tj. $AZZ^* y_i = \lambda_i ZZ^* y_i \Rightarrow B(Z^* y_i) = \lambda_i (Z^* y_i)$. Dakle B ima $\lambda_q, \ldots, \lambda_n$ kao svojstvene vrijednosti. Jer su $\mu_{q-1}, \mu_q, \ldots, \mu_{n-1}$ zadnjih $n-q+1$ vlastitih vrijednosti od B vrijedi $\mu_{q-1} + \cdots + \mu_{n-1} \leq \lambda_q + \cdots + \lambda_n$. Dakle smo dobili

$$tr(X^*AX) \leq \sum_{j=1}^{l} \lambda_{i_j} + \sum_{i=q}^{n} \lambda_{i_j} = \sum_{j=1}^{k} \lambda_{i_j}.$$

Time smo dokazali relaciju (4.16).Relacija (4.17) se dobije tako da se u relaciju (4.16) uvrsti matrica $-A$. ■ ■

Kada je $k = 1$, Wielandtov teorem daje Fischerovu karakterizaciju svojstvenih vrijednosti hermitske matrice.

Korolar 4.28 *(Fischer) Za svojstvene vrijednosti od A vrijedi*

$$\lambda_i = \max_{dim(\chi)=i} \min_{\substack{x \in \chi \\ x^*x=1}} x^*Ax$$

i

$$\lambda_i = \min_{dim(\chi)=n+1-i} \max_{\substack{x \in \chi \\ x^*x=1}} x^*Ax.$$

Dokaz. $k = 1$ daje $X = (x_{i_1}),\ x_{i_1} \in \chi_{i_1}, dim(\chi_{i_1}) = i_1$. Supstitucija $i_1 \rightarrow i$ daje obje relacije. ■ ■

Za $i = 1$ druga karakterizacija daje

$$\lambda_1 = \max_{x^*x=I} x^*Ax.$$

Neka su λ, ε i $\tilde{\lambda}$ najveće svojstvene vrijednosti hermitskih matrica A, E, $\tilde{A} = A + E$.Tada je

$$\tilde{\lambda}_1 = \max_{x^*x=1} x^*\tilde{A}x \leq \max_{x^*x=1} x^*Ax + \max_{x^*x=1} x^*Ex = \lambda_1 + \varepsilon_1$$

Jer je $|\varepsilon_1| \leq ||E||_2$ očito je $\tilde{\lambda}_1 \leq \lambda_1 + ||E||_2$.Uzimajući u obzir da je $A = \tilde{A} - E$ dobijemo $\lambda_1 \leq \tilde{\lambda}_1 + ||E||_2$, to jest $|\lambda_1 - \tilde{\lambda}_1| \leq ||E||_2$.Da bi generalizirali ovaj rezultat prvo dokazujemo

Teorem 4.29 *Neka su $A, E, \tilde{A} = A + E$ hermitske matrice sa svojstvenim vrijednostima $\lambda_1 \geq \cdots \geq \lambda_n$, $\varepsilon_1 \geq \varepsilon_2 \geq \cdots \geq \varepsilon_n$ i $\tilde{\lambda}_1 \geq \cdots \geq \tilde{\lambda}_n$ respektivno, i neka su $i_1, i_2, \ldots, i_k$ međusobno različiti indeksi između 1 i n. Tada je*

$$\lambda_{i_1} + \cdots + \lambda_{i_k} + \varepsilon_{n-k+1} + \cdots + \varepsilon_n \leq \tilde{\lambda}_{i_1} + \cdots + \tilde{\lambda}_{i_k} \leq \lambda_{i_1} + \cdots + \lambda_{i_k} + \varepsilon_1 + \cdots + \varepsilon_k$$

Dokaz. Bez gubitka općenitosti možemo pretpostaviti da je

$$i_1 < i_2 < \cdots < i_k.$$

Prvo ćemo dokazati drugu nejednakost.

Prema Wielandtovom teoremu postoje potprostori $\chi_{i_1} \subset \chi_{i_2} \subset \cdots \chi_{i_k}$ takvi da je

$$\tilde{\lambda}_{i_1} + \cdots + \tilde{\lambda}_{i_k} = \min_{\substack{X = (x_{i_1}, \ldots, x_{i_k}) \\ x_{i_j} \in \chi_{i_j} \\ X^*X = I}} tr(X^*\tilde{A}X)$$

Također postoje vektori $x_{i_j} \in \chi_{i_j}$ takvi da je $X = (x_{i_1}, \ldots, x_{i_k})$ ortonormirana matrica i $\lambda_{i_1} + \cdots + \lambda_{i_k} \geq tr(X^*\tilde{A}X)$. Tada imamo za taj izbor od X

$$\begin{aligned} \tilde{\lambda}_{i_1} + \cdots + \tilde{\lambda}_{i_k} &\leq tr(X^*(A + E)X) = tr(X^*AX) + tr(X^*EX) \\ &\leq \lambda_{i_1} + \cdots + \lambda_{i_k} + tr(X^*EX) \end{aligned}$$

Korolar 4.22 daje $\varepsilon_i \geq \mu_i \geq \varepsilon_{i+n-k}$ $i = 1, \ldots, k$, gdje su μ_i vlastite vrijednosti od $X^*EX \in \mathbf{C}^{k \times k}$, pa vrijedi $tr(X^*EX) = \mu_1 + \cdots + \mu_k \leq \varepsilon_1 + \cdots + \varepsilon_n$ što dokazuje drugu nejednakost.Prva se nejednakost može dobiti iz druge ako pišemo $A = \tilde{A} - E$ jer je sada

$$\begin{aligned}\lambda_{i_1} + \cdots + \lambda_{i_k} &\leq \tilde{\lambda}_{i_1} + \cdots + \tilde{\lambda}_{i_k} + tr(X^*(-E)X) \\ &\leq \tilde{\lambda}_{i_1} + \cdots + \tilde{\lambda}_{i_k} - \varepsilon_n - \cdots - \varepsilon_{n-(k-1)}.\end{aligned}$$

Ovdje je korišteno da su vlastite vrijednosti od $-E$ $\quad -\varepsilon_1 \leq -\varepsilon_2 \leq \cdots \leq -\varepsilon_n$. ■ ■

Korolar 4.30 *(Weyl) Za* $i = 1, 2, \ldots, n$ *vrijedi* $\tilde{\lambda}_i \in [\lambda_i + \varepsilon_n, \lambda_i + \varepsilon_1]$

Dokaz. Uzmemo $k = 1$ u prethodnom teoremu : $\tilde{\lambda}_{i_1} \leq \lambda_{i_1} + \varepsilon_1$, $\lambda_{i_1} \leq \tilde{\lambda}_{i_1} - \varepsilon_n$ i zamijenimo i_1 s i. ■ ■

Ako je E pozitivno definitna tada vlastite vrijednosti moraju porasti .

Korolar 4.31

$$\max_i |\tilde{\lambda}_i - \lambda_i| \leq ||E||_2 \tag{4.25}$$

Dokaz. Iz Weylovog teorema slijedi $\varepsilon_n \leq \tilde{\lambda}_i - \lambda_i \leq \varepsilon_1$, to jest $|\tilde{\lambda}_i - \lambda_i| \leq \max\{|\varepsilon_1|, |\varepsilon_n|\} = ||E||_2$. ■ ■

Primjer 4.32 *Dokažite da je jedna vlastita vrijednost od*

$$A = \begin{pmatrix} B & c \\ c^* & \delta \end{pmatrix}, B = B^* u \textit{ intervalu } [\delta - ||c||_2, \delta + ||c||_2]$$

Dokaz.

$$A = \begin{pmatrix} B & \\ & \delta \end{pmatrix} + \begin{pmatrix} & c \\ c^* & \end{pmatrix}, \delta \in \sigma(B \oplus [\delta])$$

Postoji $\lambda \in \sigma(A)$ tako da je

$$|\lambda - \delta| \leq \left\| \begin{bmatrix} & c \\ c^* & \end{bmatrix} \right\|_2.$$

Međutim postoji $H = H^*$ tako da vrijedi $Hc = ||c||_2 e_1$. Tada je

$$\begin{aligned}\left\| \begin{bmatrix} & c \\ c^* & \end{bmatrix} \right\|_2 &= \left\| \begin{bmatrix} H & \\ & 1 \end{bmatrix} \begin{bmatrix} & c \\ c^* & \end{bmatrix} \begin{bmatrix} H^* & \\ & 1 \end{bmatrix} \right\|_2 = \\ &= \left\| \begin{bmatrix} & & ||c||_2 \\ & 0 & \\ ||c||_2 & & \end{bmatrix} \right\|_2 = \left\| P \begin{bmatrix} ||c||_2 & & \\ & 0 & \\ & & ||c||_2 \end{bmatrix} \right\|_2 = ||c||_2,\end{aligned}$$

gdje je P neka permutacija. ■ ■

4.2.3. Teorem Mirskog

Relacija (4.30) se može napisati u obliku

$$\|\, diag(\tilde{\lambda}_i - \lambda_i)\|_2 \leq \|E\|_2. \tag{4.26}$$

To sugerira da $\|\cdot\|_2$ zamijenimo s nekom drugom normom. Dokazat ćemo da (4.26) vrijedi za svaku unitarno invarijantnu normu. Za to će nam trebati analogni rezultat za singularne vrijednosti.

Teorem 4.33 *(Mirsky) Neka su X i $\tilde{X}$ matrice iste dimenzije sa singularnim vrijednostima*

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_p$$

$$\tilde{\sigma}_1 \geq \tilde{\sigma}_2 \geq \cdots \geq \tilde{\sigma}_p.$$

Tada za svaku unitarno invarijantnu normu $\|\cdot\|$ vrijedi

$$\|\, diag(\tilde{\sigma}_i - \sigma_i)\| \leq \|\tilde{X} - X\|.$$

Dokaz. Bez gubitka općenitosti možemo pretpostaviti da su X i $\tilde{X}$ kvadratne. Ako nisu nađemo unitarnu matricu Q tako da su QX i $Q\tilde{X}$ ili XQ i $\tilde{X}Q$ kvadratne. Time se neće ništa promijeniti jer je norma unitarno invarijantna.
Prema teoremu 4.20 svojstvene vrijednosti matrice
$\begin{bmatrix} 0 & X \\ X^* & 0 \end{bmatrix}$ su $\pm\sigma_1, \ldots, \pm\sigma_p$ i slično vrijedi za $\tilde{X}$.
Ako su $\varepsilon_1 \geq \cdots \geq \varepsilon_p$ singularne vrijednosti od $\tilde{X} - X$, tada su svojstvene vrijednosti od
$\begin{bmatrix} 0 & \tilde{X} - X \\ \tilde{X} - X^* & 0 \end{bmatrix}$ $\pm\varepsilon_1, \ldots, \pm\varepsilon_p$.
U teoremu 4.29 uzmimo

$$i_k = \begin{cases} k & \tilde{\sigma}_k \geq \sigma_k \\ n + k & \tilde{\sigma}_k < \sigma_k. \end{cases}$$

Tada je

$$|\tilde{\sigma}_1 - \sigma_1| + \cdots + |\tilde{\sigma}_k - \sigma_k| \leq \varepsilon_1 + \cdots + \varepsilon_k, \; k = 1, \ldots, p.$$

Po Fatuovom teoremu

$$\Phi(\tilde{\sigma}_1 - \sigma_1, \ldots, \tilde{\sigma}_p - \sigma_p) \leq \Phi(\varepsilon_1, \ldots, \varepsilon_p)$$

vrijedi za bilo koju SBF. Teorem sada vrijedi iz von Neumannove karaktrizacije unitarno invarijantnih normi. ■ ■

Korolar 4.34 *Neka je Φ simetrična kvadratna funkcija i $||\cdot||_\Phi$ pripadna unitarno invarijantna norma. Tada je*

$$||\,diag(\tilde{\lambda}_i - \lambda_i)||_\Phi \leq ||E||_\Phi. \tag{4.27}$$

Dokaz. Neka je $\rho = \min\{\lambda_n, \tilde{\lambda}_n\}$. Tada su svojstvene vrijednosti matrica $A - \rho I$ i $\tilde{A} - \rho I$ nenegativne. Stoga su svojstvene vrijednosti i singularne vrijednosti jednake pa rezultat slijedi primjenom Mirskijevog teorema. ■ ■

Kada Φ generira Frobeniusovu normu dobijemo hermitski W.Hoffmanov teorem.

Korolar 4.35

$$\sum_{i=1}^{n} |\tilde{\lambda}_i - \lambda_i|^2 \leq ||E||_F^2.$$

Ovaj rezultat je nešto jači od W.Hoffmanovog rezultata jer ukazuje na uređene svojstvenih vrijednosti dok W.Hoffmanov teorem samo daje egzistenciju uređaja.

4.2.4. Rezidualne ograde

Imamo matricu X čiji stupčani prostor aproksimira neki invarijantni potprostor od A. Dakle, za neki izbor od M rezidual

$$R = AX - XM$$

je mali. Ako X ima ortonormirane stupce, tada je po teoremu ??? svaka unitarno invarijantna norma od R minimizirana kad je $M = X^*AX$.

S teoremom Mirskog, za hermitsku matricu A možemo ocijeniti svojstvene vrijednosti od M kao aproksimacije od A.

Teorem 4.36 *Neka je A hermitska i neka $X \in \mathbf{R}^{n\times k}$ zadovoljava $X^*X = I$. Neka je $M = X^*AX$ i $R = AX - XM$. Neka je Φ SBF na $\mathbf{R}^n$ i neka $||\cdot||_\Phi$ označava pripadnu familiju unitarno invarijantnih normi. Ako su $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ i $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_n$ svojstvene vrijednosti od A i M respektivno. Tada postoje indeksi $i_1 < i_2 < \ldots < i_k$ takvi da je*

$$||\, diag(\mu_j - \lambda_{i_j})||_\Phi \leq ||XR^* + RX^*|||_\Phi = \Phi(\rho_1, \rho_1, \rho_2, \rho_2, \ldots) \tag{4.28}$$

gdje su $\rho_1 \geq \rho_2 \geq \cdots$ singularne vrijednosti od R.

Dokaz. Dokazat ćemo (4.28) za slučaj $2k \leq n$, ostavljajući drugi slučaj kao vježbu. Za $M = X^*AX$, neka je

$$E = -(XR^* + RX^*).$$

Tada je E hermitska i vrijedi

$$(A + E)X = XM. \tag{4.29}$$

Zaista $(A + E)X = AX - XR^*X - RX^*X = AX - R - XR^*X = XM$ jer je $XR^*X = XX^*AX - XMX^*X = XX^*AX - XM = 0$. Dakle $\mathcal{R}(X)$ je invarijantan potprostor od $A+E$ i za svaku svojstvenu vrijednost μ_j od M postoji odgovarajuća svojstvena vrijednost $\tilde{\lambda}_{i_j}$ od $A+E$. Prema teoremu Mirskog (zapravo relaciji (4.27)) $||\, diag(\tilde{\lambda}_i - \lambda_i)||_\Phi \leq ||E||_\Phi$. Stoga je $||\, diag(mu_j - \lambda_{i_j})||_\Phi \leq ||E||_\Phi = ||XR^* + RX^*||_\Phi$ pa preostaje dokazati jednakost u (4.28), tj. da su singularne vrijednosti od E $\rho_1, \rho_1, \rho_2, \rho_2, \ldots$

Neka je $X_\perp$ takva da je $(X, X_\perp)$ unitarna. Tada je

$$(X, X_\perp)^* E (X, X_\perp) = \begin{bmatrix} X^* \\ X_\perp^* \end{bmatrix} E \begin{bmatrix} X & X_\perp \end{bmatrix} = \begin{bmatrix} X^*EX & X^*EX_\perp \\ X_\perp^*EX & X_\perp^*EX_\perp \end{bmatrix}.$$

Jer je $X^*EX = X^*XM - X^*AX = M - M = 0$,
$X_\perp^*EX_\perp = -X_\perp^*(XR^* + RX^*)X_\perp = -X_\perp^*XR^*X_\perp - X_\perp^*RX^*X_\perp = 0$,
$X^*EX_\perp = -X^*(XR^* + RX^*)X_\perp = -R^*X_\perp$,
$X_\perp^*EX = -X_\perp^*(XR^* + RX^*)X = -X_\perp^*R$,
vrijedi

$$(X, X_\perp)^* E (X, X_\perp) = \begin{bmatrix} 0 & -R^*X_\perp \\ -X_\perp^*R & 0 \end{bmatrix}$$

pa su singularne vrijednosti od E upravo singularne vrijednosti od $X_\perp^* R$ uzete dva puta. Jer su stupci od $X_\perp$ ortonormirani i $\mathcal{R}(R) \subset \mathcal{R}(X_\perp)$
$(R = AX - XM = AX - XX^*AX = (I - XX^*)AX = X_\perp(X_\perp^* AX))$ imamo $(X_\perp^* R)^*(X_\perp^* R) = R^* X_\perp X_\perp^* R = R^* R$ pa su singularne vrijednosti od $X_\perp^* R$ upravo singularne vrijednosti od R. ■

Korolar 4.37 *Za spektralnu normu vrijedi*

$$\max_j\{|\mu_j - \lambda_{i_j}|\} \leq ||R||_2,$$

a za Frobeniusovu normu vrijedi

$$\left[\sum_j |\mu_j - \lambda_{i_j}|^2\right]^{1/2} \leq \sqrt{2}||R||_F. \tag{4.30}$$

Napomena 4.38 *U slučaju 2-norme uvjet $M = X^*AX$ se može maknuti, tj. za proizvoljnu hermitsku matricu M vrijedi*

$$||\, diag(\mu_j - \lambda_{i_j})||_2 \leq ||R||_2.$$

U slučaju F-norme faktor $\sqrt{2}$ se može maknuti (relacija (4.30)).

Rezidualne ograde mogu biti jako dobre i jako loše.

Primjer 4.39 *Neka je $A = \begin{bmatrix} 0 & \varepsilon \\ \varepsilon & 0 \end{bmatrix}$, $X = e_1$ tako da je $M = \{0\}$ i $R = Ae_1 - e_1\, 0 = \begin{bmatrix} 0 \\ \varepsilon \end{bmatrix}$. Dakle je ograda $|\lambda - 0| \leq |\varepsilon|$, pa je oštra, tj. dostiže se.*

Neka je $A = \begin{bmatrix} 0 & \varepsilon \\ \varepsilon & 1 \end{bmatrix}$, $X = e_1$,$M = \{0\}$. Sada je ograda opet $|\varepsilon|$ ali su svojstvene vrijednosti $\approx -\varepsilon^2$ i $1 + \varepsilon^2$.

Svojstvene vrijednosti od $M = X^*AX$ se zovu Rayleigh-Ritzove aproksimacije svojstvenih vrijednosti od A.

Iako smo spominjali da se X bira tako da su stupci od X aproksimativno svojstveni vektori, zapravo sve što se treba zahtjevati je da je R mali.

Literatura

[1] J. Barlow, J. Demmel, *Computing Accurate Eigensystem of Scaled Diagonally Dominant Matrices,* SIAM Journal on Numerical Analysis Vol. 27, No. 3, (762-791) 1990.

[2] C. Davis, W. M. Kahan, *The Rotation of Eigenvectors by a Perturbation. III,* SIAM Journal on Numerical Analysis Vol. 7, No. 1. (1-46) 1970.

[3] J. Demmel, K. Veselić, *Jacobi's Method is More Accurate than QR,* SIAM Journal on Matrix Analysis and Applications, Vol. 13. No. 4, (1204-1244) 1992.

[4] G. H. Golub, C. F. Van Loan, *Matrix Computation,* Johns Hopkins U. P., Baltimore, 1989.

[5] H. Kraljević, S. Kurepa, *Matematička analiza. Četvrti dio. Funkcije kompleksne varijable,* Tehnička knjiga, Zagreb, 1986.

[6] S. Kurepa, *Konačno dimenzionalni vektorski prostori i primjene,* Tehnička knjiga, Zagreb, 1967.

[7] J. M. Ortega, W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations of Several Variables,* Academic Press, New York, 1970.

[8] S. Singer, *Računanje spektralne dekompozicije pozitivno definitne matrice,* magistarski rad, Zagreb, 1993.

[9] I. Slapničar, *Accurate Symmetric Eigenreduction by Jacobi Method,* doktorska disertacija, Hagen, 1992.

[10] I. Slapničar, K.Veselić, *Perturbations of the Eigenprojections of a Factorised Hermitian Matrix,* Linear Algebra Appl., 218 (273-280) 1995.

[11] A. van der Sluis, *Numbers and Equilibration of Matrices,* Numerische Matematik, 14. (14-23) 1969.

[12] G. W. Stewart, *Error Bounds for Approximate Invariant Subspaces of Closed Linear Operators,* SIAM Journal on Numerical Analysis Vol. 8, No. 4, (796-808) 1971.

[13] G. W. Stewart, *Error and Perturbation Bounds for Subspaces Associated with Certain Eigenvalue Problem,* SIAM Review Vol 15. No. 4. (727-764) 1973.

[14] G. W. Stewart, *Computing the CS Decomposition of Partitioned Orthonormal Matrix,* Numerische Matematik 40, (297-306), 1982.

[15] G. W. Stewart, Ji-Guang Sun, *Matrix Perturbation Theory,* Academic Press, 1990.

[16] K.Veselić, *A Jacobi Eigenreduction Algotithm for definite matrix pairs,* Numerische Matematik, 64. (241-269) 1993.

[17] K.Veselić, I. Slapničar, *Floating - Point perturbation of Hermitian matrices,* Linear Algebra Appl., 195 (81-116) 1993.

[18] P. Å. Wedin, *On Angles Between Subspaces of a Finite Dimensional Inner Product Space,* Matrix Pencils, Proceedings of Conference, Sweden 1982. Springer Verlag Berlin Heidelberg 1983.