5.

Matrične norme, spektar i spektralni radijus matrice

5.1. Operatorske norme i konvergencija matrica

Vektorski prostor linearnih operatora. Ako su zadani normirani vektorski prostori X i Y onda se, kao što smo vidjeli, na njima može uvesti pojam konvergencije. To omogućava da definiramo i pojam neprekinute funkcije među normiranim prostorima. Za neku funkciju $F : X \to Y$ kažemo da je neprekinuta u točki $x \in X$ ako iz $x_k \to x$ slijedi $F(x_k) \to F(x)$, ili kraće: $\lim_{k\to\infty} F(x_k) = F(x)$. Ako je $F = A : X \to Y$ linearni operator, onda je zbog linearnosti neprekinutost linearnog operatora dovoljno provjeriti samo za $x = 0$. Svaki linearni operator definiran među konačno-dimenzionalnim vektorskim prostorima X i Y je neprekinut.

Za zadane konačno-dimenzionalne vektorske prostore X i Y skup svih linearnih operatora $A : X \to Y$ označavamo sa $\mathcal{L}(X,Y)$. Taj skup je vektorski prostor, ako zbrajanje operatora A i $B : X \to Y$ definiramo sa $(A+B)(x) = Ax + Bx$, a množenje sa skalarom λA sa $(\lambda A)(x) = \lambda \cdot Ax$. Ako su X i Y konačno-dimenzionalni prostori, lako se vidi da je $\dim \mathcal{L}(X,Y) = \dim X \cdot \dim Y$. Na pr. $\dim \mathcal{L}(\mathbf{R}^n, \mathbf{R}^m) = mn$, jer $\mathcal{L}(\mathbf{R}^n, \mathbf{R}^m)$ možemo poistovjetiti s vektorskim prostorom matrica tipa $m \times n$, tj. sa M_{mn}. Vektorski prostor $\mathcal{L}(X,X)$ svih linearnih operatora iz X u samog sebe označavamo kraće sa $\mathcal{L}(X)$.

Operatorska norma linearnih operatora i matrica. Ako su X i Y normirani vektorski prostori (ne nužno konačno-dimenzionalni), onda se i u vektorski prostor $\mathcal{L}(X,Y)$ svih neprekinutih linearnih operatora iz X u Y može na prirodan način uvesti tzv. **operatorska norma**:

$$\|A\| := \sup_{x\neq 0} \frac{\|Ax\|}{\|x\|}, \quad \text{tj.} \quad \|A\| = \sup_{\|x\|=1} \|Ax\|.$$

Pritom je neprekinutost operatora A ekvivalentna s uvjetom $\|A\| < \infty$, pa kažemo još da je A **omeđen linearan operator** (engl. "bounded linear operator"). Ako su X i Y konačno-dimenzionalni prostori, onda je svaki linearni operator omeđen, a $\sup$ se u definiciji operatorske norme može zamijeniti s $\max$ (tj. supremum je dostignut za neki x).

Ako je $\|A\| < \infty$, onda je A doista neprekinut operator: iz $x_k \to x$ kad $k \to \infty$, slijedi $\|Ax_k - Ax\| = \|A(x_k - x)\| \leqslant \|A\|\,\|x_k - x\| \to 0$.

Normu od A možemo opisati na ekvivalentan način s ova dva zahtjeva:

1. $\|Ax\| \leqslant \|A\| \cdot \|x\|$
2. $\|A\|$ ima svojstvo minimalnosti: $\|A\|$ je najmanji broj $M \geqslant 0$ za koji vrijedi $\|Ax\| \leqslant M\|x\|$ za sve x.

S tom normom vektorski prostor $\mathcal{L}(X,Y)$ postaje normiran prostor. Doista,

a) $\|A\| \geqslant 0$ je jasno; ako je $\|A\| = 0$ onda je $\|Ax\| = 0$ za sve x, dakle $A = 0$ (pozitivnost);
b) $\|\lambda A\| = \sup_{x\neq 0} \frac{\|\lambda Ax\|}{\|x\|} = \sup_{x\neq 0} |\lambda| \cdot \frac{\|Ax\|}{\|x\|} = |\lambda| \cdot \sup_{x\neq 0} \frac{\|Ax\|}{\|x\|} = |\lambda| \cdot \|A\|$ (homogenost);
c) $\|(A+B)x\| = \|Ax+Bx\| \leqslant \|Ax\| + \|Bx\| \leqslant (\|A\| + \|B\|)\|x\|$, dakle zbog svojstva minimalnosti je onda $\|A+B\| \leqslant \|A\| + \|B\|$ (nejednakost trokuta).

Za operatorsku normu vrijedi i još jedno vrlo važno svojstvo:

d) ako su $A, B \in \mathcal{L}(X)$, onda je

$$\|AB\| \leqslant \|A\| \cdot \|B\|.$$

Posebno, vrijedi $\|A^k\| \leqslant \|A\|^k$ za svaki prirodan broj k.

Dokaz je lagan. Vrijedi $\|ABx\| \leqslant \|A\| \cdot \|Bx\| \leqslant \|A\| \cdot \|B\| \cdot \|x\|$, dakle zbog svojstva minimalnosti broja $\|AB\|$ je onda $\|AB\| \leqslant \|A\| \cdot \|B\|$.

O toj važnoj normi bit će više riječi u odjeljku o matričnim normama.

Neka je X konačno-dimenzionalni vektorski prostor nad poljem $K = \mathbf{R}$ ili $\mathbf{C}$. Neprekinuti linearni operatori iz X u K zovu se **linearni funkcionali**. Vektorski prostor $\mathcal{L}(X,K)$ svih neprekinutih linearnih funkcionala iz X u K zove se **dualni prostor** od X, i označava sa X'. Na pr. funkcional $f : \mathbf{R}^n \to \mathbf{R}$ zadan sa

$$f(x) = a_1x_1 + \ldots + a_nx_n,$$

je linearan funkcional, tj. sadržan u $(\mathbf{R}^n)'$. To su zapravo i jedini funkcionali na $\mathbf{R}^n$, tj. svi su oblika $f_a : \mathbf{R}^n \to \mathbf{R}$, gdje je $f_a(x) = (x\,|\,a)$, $a \in \mathbf{R}^n$ (**Rieszov teorem o reprezentaciji linearnog funkcionala**). Vrlo lako je vidjeti da je pridruživanje $f \mapsto a$ linearni operator iz $(\mathbf{R}^n)'$ u $\mathbf{R}^n$, i to bijektivan (tj. izomorfizam), pa se dualni prostor $(\mathbf{R}^n)'$ može poistovjetiti sa $\mathbf{R}^n$.

Ako je X beskonačno-dimenzionalan prostor, točan opis dualnog prostora nije sasvim jednostavan. Na pr. ako je $X = L^p(0,1)$, $p \in (1,\infty)$, onda se pokazuje da je prostor X' izomorfan sa $L^q(a,b)$, gdje je q dualni eksponent od p, tj. $q = p/(p-1)$. To znači da svaki linearni funkcional $F \in L^p(0,1)'$, tj. $F : L^p(0,1) \to \mathbf{R}$, ima oblik

$$F(x) = \int_0^1 a(t)x(t)\,dt$$

gdje je $a(t)$ neka funkcija iz $L^q(0,1)$. Funkcional $F \in L^p(0,1)'$ poistovjećujemo s funkcijom $a \in L^q(0,1)$. Za $p = 2$ je $q = 2/(2-1) = 2$, pa dual prostora $L^2(0,1)$ možemo poistovjetiti sa samim sobom.

Neka je M_{mn} vektorski prostor svih realnih (ili kompleksnih) matrica $\mathbf{A}$ tipa $m \times n$. Gledajući matricu $\mathbf{A}$ kao linearni operator iz $\mathbf{R}^n$ u $\mathbf{R}^m$ ($\mathbf{R}^n \ni x \mapsto \mathbf{A}x \in \mathbf{R}^m$), možemo na prirodan način definirati operatorsku normu od $\mathbf{A}$, generiranu s nekom zadanom normom na $\mathbf{R}^n$ i još jednom normom na $\mathbf{R}^m$: $\|\mathbf{A}\| = \sup_{x\neq 0} \|\mathbf{Ax}\|/\|\mathbf{x}\|$. S tom matričnom normom vektorski prostor M_{mn} postaje normirani prostor. Budući da je on konačno-dimenzionalan (dimenzije mn) onda je to Banachov prostor, tj. svaki Cauchyjev slijed matrica u njemu konvergira: ako za slijed matrica $(\mathbf{A}_k)$ vrijedi $\|\mathbf{A}_k - \mathbf{A}_l\| \to 0$ kada $k, l \to \infty$, onda slijed matrica $\mathbf{A}_k$ konvergira k nekoj matrici $\mathbf{A}$.

Na pr., uzimajući ∞-normu na $\mathbf{R}^n$ i $\mathbf{R}^m$ dobivamo operatorsku matričnu normu $\|\mathbf{A}\|_\infty$. Vrijedi

$$\|\mathbf{A}\|_\infty = \max_{1\leqslant i\leqslant m} \sum_{j=1}^n |a_{ij}|.$$

Računamo maksimum zbrojeva po redcima u matrici, pa se zato ta norma nekad zove i **redčana norma matrice**.

DOKAZ. Uzmimo bilo koji $x \in \mathbf{R}^n$. Za $y = \mathbf{A}x$ je

$$\begin{aligned}\|\mathbf{A}x\|_\infty &= \max_i |y_i| = \max_i \Big|\sum_{j=1}^n a_{ij}x_j\Big| \\ &\leqslant \max_i \sum_{j=1}^n |a_{ij}| \cdot |x_j| \leqslant \max_j |x_j| \cdot \max_i \sum_{j=1}^n |a_{ij}| = M\|x\|_\infty.\end{aligned}$$

gdje smo označili $M = \max_i \sum_{j=1}^n |a_{ij}|$. Dakle, $\|\mathbf{A}\|_\infty \leqslant M$. Treba još samo vidjeti da je broj M doista najmanji za koji vrijedi ova nejednakost. Dovoljno je pronaći vektor x norme 1 za koji je $\|\mathbf{A}x\|_\infty = M$. Neka je i_0 indeks (retka) na kojem se dostiže maksimum redčane norme, tj. $M = \sum_{j=1}^n |a_{i_0j}|$. Stavimo $x = (x_1, \ldots, x_n)^\top$, $x_j = \operatorname{sgn} a_{i_0j}$. Onda je $a_{i_0j}x_j = |a_{i_0j}|$, pa za $y = Ax$ vrijedi $y_{i_0} = M$. Budući da za sve i vrijedi $|y_i| = |\sum_j a_{ij}x_j| \leqslant \sum_j |a_{ij}| \leqslant |y_{i_0}|$, onda je $\|Ax\|_\infty = \|y\|_\infty = |y_{i_0}| = M$. ☺

Na sličan način, uzimajući 1-normu na $\mathbf{R}^n$ i $\mathbf{R}^m$ dobivamo operatorsku normu $\|\mathbf{A}\|_1$. Vrijedi

$$\|\mathbf{A}\|_1 = \max_{1\leqslant j\leqslant n} \sum_{i=1}^m |a_{ij}|.$$

Računamo maksimum zbrojeva po stupcima u matrici, pa ju zato zovemo i **stupčanom normom**. Primijetite da je $\|\mathbf{A}\|_1 = \|\mathbf{A}^\top\|_\infty$, jer transponiranjem redci matrice $\mathbf{A}$ postaju stupci matrice $\mathbf{A}^\top$. Za obje ove operatorske matrične norme vrijedi važna nejednakost

$$\|\mathbf{AB}\| \leqslant \|\mathbf{A}\| \cdot \|\mathbf{B}\|,$$

pod uvjetom da su $\mathbf{A}$ i $\mathbf{B}$ ulančane.

Kad imamo matričnu normu onda možemo definirati i udaljenost među matricama $\mathbf{A}$ i $\mathbf{B}$ istoga tipa: $d(\mathbf{A},\mathbf{B}) = \|\mathbf{A}-\mathbf{B}\|$. To nam omogućava da uvedemo i pojam kugle u prostoru matrica. Pojam udaljenosti omogućava da uvedemo konvergenciju slijeda matrica.

Konvergencija slijeda matrica i redovi matrica. Pretpostavimo da je zadan slijed matrica $\mathbf{A}_k = (a_{ij}^{(k)})$, $k = 1, 2, \ldots$, istog tipa $m \times n$. Kažemo da **slijed matrica** $\mathbf{A}_k$ **konvergira** k matrici $\mathbf{A} = (a_{ij})$, i pišemo $\mathbf{A}_k \to \mathbf{A}$ kad $k \to \infty$, ili $\lim_{k\to\infty} \mathbf{A}_k = \mathbf{A}$, ako $\|\mathbf{A}_k - \mathbf{A}\| \to 0$ kad $k \to \infty$, gdje je $\|\cdot\|$ bilo koja matrična norma. Budući da je prostor matrica M_{mn} konačno-dimenzionalan (dimenzije mn), sve matrične norme na njemu su međusobno ekvivalentne, pa definicija konvergencije ne ovisi o izboru norme.

Uvjet $\mathbf{A}_k \to \mathbf{A}$ je ekvivalentan s uvjetom da $a_{ij}^{(k)} \to a_{ij}$ kad $k \to \infty$, za sve i, j. Drugim riječima, slijed matrica $\mathbf{A}_k$ konvergira matrici $\mathbf{A}$ onda i samo onda ako odgovarajući matrični koeficijenti konvergiraju na uobičajen način.

Doista, ako $a_{ij}^{(k)} \to a_{ij}$ kad $k \to \infty$, za sve ij, onda pišući $\mathbf{A}_k - \mathbf{A} = \sum_{i,j}(a_{ij}^{(k)} - a_{ij})\mathbf{E}_{ij}$, gdje je $\mathbf{E}_{ij}$ matrica koja je nula svuda osim na mjestu (i,j) na koje stavljamo jedinicu, dobivamo zbog nejednakosti trokuta $\|\mathbf{A}_k - \mathbf{A}\| \leqslant \sum_{i,j} |a_{ij}^{(k)} - a_{ij}| \cdot \|\mathbf{E}_{ij}\| \to 0$ kada $k \to \infty$. Obratno, pretpostavimo da $\|\mathbf{A}_k - \mathbf{A}\| \to 0$ kad $k \to \infty$. Zbog ekvivalentnosti norma na konačno-dimenzionalnom prostoru M_{mn} zaključujemo da onda $\|\mathbf{A}_k - \mathbf{A}\|_\infty \to 0$, pa slijedi da $|a_{ij}^{(k)} - a_{ij}| \leqslant \|\mathbf{A}_k - \mathbf{A}\| \to 0$ kad $k \to \infty$.

Primjer 1. Slijed matrica

$$\mathbf{A}_k = \begin{bmatrix} (1+\frac{1}{k})^k & \frac{k+1}{k} \\ \frac{1}{k} & k\sin\frac{1}{k} \end{bmatrix}$$

konvergira prema matrici

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

Na sličan način može se uvesti konvergencija beskonačnog reda matrica određenog slijedom matrica $\mathbf{A}_k$ istog tipa. Najprije gledamo parcijalne sume $\mathbf{S}_k := \mathbf{A}_1 + \ldots + \mathbf{A}_k$. Kažemo da **red matrica** $\sum_{k=1}^\infty A_k$ **konvergira** ako slijed parcijalnih suma $\mathbf{S}_k$ konvergira k matrici $\mathbf{S}$, tj. $\|\mathbf{S} - \mathbf{S}_k\| \to 0$. U tom slučaju pišemo $\sum_{k=1}^\infty \mathbf{A}_k = \mathbf{S}$. Red matrica $\sum_{k=1}^\infty \mathbf{A}_k$ konvergira onda i samo onda ako redovi brojeva $\sum_{k=1}^\infty a_{ij}^{(k)}$ konvergiraju k nekom broju s_{ij} za sve i, j.

Primjer 2. Vrijedi

$$\sum_{k=0}^\infty \begin{bmatrix} \frac{x^k}{k!} & (-1)^k\frac{x^{2k}}{(2k)!} \\ (-1)^k\frac{x^{2k+1}}{(2k+1)!} & \frac{1}{(k+1)^2} \end{bmatrix} = \begin{bmatrix} e^x & \cos x \\ \sin x & \frac{\pi^2}{6} \end{bmatrix}$$

za sve $x \in \mathbf{R}$.

Eksponencijalna funkcija matrice. Važan primjer je matrica $e^{\mathbf{A}}$, koju zovemo **eksponencijalna funkcija matrice**. Za bilo koju zadanu kvadratnu matricu $\mathbf{A}$ ona se definira s pomoću MacLaurinova reda funkcije $e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots$ ovako:

$$e^{\mathbf{A}} := \mathbf{I} + \mathbf{A} + \frac{\mathbf{A}^2}{2!} + \frac{\mathbf{A}^3}{3!} + \ldots$$

Dokažimo da taj red konvergira prema nekoj matrici za bilo koju odabranu kvadratnu matricu $\mathbf{A}$. Dovoljno je pokazati da slijed parcijalnih suma $\mathbf{S}_k = \mathbf{I} + \mathbf{A} + \ldots + \frac{\mathbf{A}^k}{k!}$ čini Cauchyjev slijed. Neka je $l < k$ i pokažimo da $\mathbf{S}_k - \mathbf{S}_l = \frac{\mathbf{A}^{l+1}}{(l+1)!} + \ldots + \frac{\mathbf{A}^k}{k!}$ teži prema nuli (nul matrici) kada k i l teže u ∞. Rabeći bilo koju operatorsku matričnu normu i činjenicu da je $\|\mathbf{A}^i\| \leqslant \|\mathbf{A}\|^i$, dobivamo

$$\|\mathbf{S}_k - \mathbf{S}_l\| \leqslant \frac{\|\mathbf{A}^{l+1}\|}{(l+1)!} + \ldots + \frac{\|\mathbf{A}^k\|}{k!} \leqslant \frac{\|\mathbf{A}\|^{l+1}}{(l+1)!} + \ldots + \frac{\|\mathbf{A}\|^k}{k!} \to 0$$

kada $k, l \to \infty$. Naime numerički red $e^{\|\mathbf{A}\|} = 1 + \|\mathbf{A}\| + \frac{\|\mathbf{A}\|^2}{2!} + \ldots$ konvergira (na pr. po D'Alambertovu kriteriju), pa njegove parcijalne sume konvergiraju.

Na sličan način mogu se definirati matrica $\sin\mathbf{A}$, $\cos\mathbf{A}$, pa i općenito $f(\mathbf{A})$ uz određene uvjete na f, koje ćemo opisati u sljedećem odjeljku.

Ako je $\mathbf{A}$ dijagonalna matrica,

$$\mathbf{A} = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix},$$

onda je i $P(\mathbf{A})$ dijagonalna matrica

$$P(\mathbf{A}) = \begin{bmatrix} P(\lambda_1) & & & \\ & P(\lambda_2) & & \\ & & \ddots & \\ & & & P(\lambda_n) \end{bmatrix}.$$

Lako se provjeri da je i $e^{\mathbf{A}}$ dijagonalna matrica:

$$e^{\mathbf{A}} = \begin{bmatrix} e^{\lambda_1} & & & \\ & e^{\lambda_2} & & \\ & & \ddots & \\ & & & e^{\lambda_n} \end{bmatrix}.$$

Ako je $\mathbf{A}$ gornja trokutasta matrica, onda je $P(\mathbf{A})$ gornja trokutasta, kao i $e^{\mathbf{A}}$. Na dijagonali su isti brojevi kao u dijagonalnom slučaju.

Primjedba 1. Pokazuje se da ako kvadratne matrice $\mathbf{A}$ i $\mathbf{B}$ komutiraju, tj. $\mathbf{AB} = \mathbf{BA}$, onda je

$$e^{\mathbf{A}+\mathbf{B}} = e^{\mathbf{A}}e^{\mathbf{B}}.$$

Primjer 3. Pogledajmo za ilustraciju gornju trokutastu matricu

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

i nađimo $e^{\mathbf{A}t} = I + \mathbf{A}t + \frac{(\mathbf{A}t)^2}{2!} + \frac{(\mathbf{A}t)^3}{3!} + \ldots$, gdje je t bilo koji realni broj. Kao što znamo, vidi Primjer 3.18,

$$\mathbf{A}^k = \begin{bmatrix} \lambda^k & k\lambda^{k-1} \\ 0 & \lambda^k \end{bmatrix},$$

dakle

$$e^{\mathbf{A}t} = \begin{bmatrix} \sum_{k=0}^\infty \frac{(\lambda t)^k}{k!} & \sum_{k=1}^\infty \frac{k\lambda^{k-1}t^k}{k!} \\ 0 & \sum_{k=0}^\infty \frac{(\lambda t)^k}{k!}, \end{bmatrix}.$$

Zbog $\sum_{k=1}^\infty \frac{k\lambda^{k-1}t^k}{k!} = t\sum_{k=1}^\infty \frac{\lambda^{k-1}t^{k-1}}{(k-1)!} = [j := k-1] = t\sum_{j=0}^\infty \frac{(\lambda t)^j}{j!} = te^{\lambda t}$ dobivamo

$$e^{\mathbf{A}t} = \begin{bmatrix} e^{\lambda t} & te^{\lambda t} \\ 0 & e^{\lambda t} \end{bmatrix}.$$

Primijetimo da je λ vlastita vrijednost od $\mathbf{A}$, te ako je λ negativan realan broj, tj. $\lambda < 0$, onda vrijedi da $e^{\mathbf{A}t} \to 0$ kada $t \to +\infty$. Gornja matrica $\mathbf{A}$ s $\lambda < 0$ je specijalan slučaj tzv. stabilne matrice, koje ćemo definirati malo kasnije, vidi Odjeljak 5.5. Matrica $\mathbf{A}$ je u ovom primjeru matrica linearnog sustava diferencijalnih jednačaba

$$\dot{x} = \lambda x + y, \;\; \dot{y} = \lambda y,$$

jer uz oznaku $\mathbf{x} = (x,y)^\top$ ovaj sustav možemo kratko zapisati kao

$$\dot{\mathbf{x}} = \mathbf{Ax}.$$

Sa zadanim početnim uvjetima $x(0) = x_0$, $y(0) = y_0$ dobivamo Cauchyjev problem čije rješenje je

$$\mathbf{x}(t) = e^{\mathbf{A}t}\mathbf{x}_0,$$

gdje je $\mathbf{x}_0 = (x_0, y_0)^\top$. Matrica $e^{\mathbf{A}t}$ zove se *evolucijska matrica sustava*. Više o njoj vidi u Odjeljku 5.9. U ovom slučaju iz gornjeg rješenja vidimo da je $x(t) = e^{\lambda t}x_0 + te^{\lambda t}y_0$, $y(t) = e^{\lambda t}y_0$. Ako je $\lambda < 0$ onda $\mathbf{x}(t) \to 0$ kada $t \to +\infty$, tj. trajektorija $\mathbf{x}(t)$ je asimptotski stabilna (teži k nuli).

Primjer 4. Pokažimo da za matricu $\mathbf{A}$ operatora rotacije ravnine za kut $\pi/2$ (oko ishodišta), tj. za matricu

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

vrijedi

$$e^{\mathbf{A}t} = \begin{bmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{bmatrix},$$

pri čemu je t bilo koji realan broj. Matrica $e^{\mathbf{A}t}$ zove se *evolucijska* ili *fundamentalna matrica* matrice $\mathbf{A}$, o čemu će više riječi biti u Odjeljku 5.9. Kao što vidimo, matrica $e^{\mathbf{A}t}$ je u ovom slučaju matrica operatora *rotacije ravnine* za kut t oko ishodišta.

Pokažimo da je na mjestu $(1,1)$ u matrici $e^{\mathbf{A}t} := I + \mathbf{A}t + \mathbf{A}^2\frac{t^2}{2!} + \ldots + \mathbf{A}^k\frac{t^k}{k!} + \ldots$ stvarno broj $\cos t$. Doista, budući da je $\mathbf{A}^2 = -\mathbf{I}$ (tj. $\mathbf{A}$ je matrična imaginarna jedinica), $\mathbf{A}^3 = -\mathbf{A}$, $\mathbf{A}^4 = \mathbf{I}$, itd., onda na mjestu $(1,1)$ matrice $e^{\mathbf{A}t}$ imamo $1 - \frac{t^2}{2!} + \frac{t^4}{4!} + \ldots$, što je dobro poznati MacLaurinov razvoj funkcije $\cos t$. Slično i za ostala mjesta evolucijske matrice $e^{\mathbf{A}t}$.

Matrica $\mathbf{A}$ je u ovom primjeru matrica linearnog sustava diferencijalnih jednačaba

$$\dot{x} = -y, \;\; \dot{y} = x,$$

gdje je $x = x(t)$ i $y = y(t)$, jer uz oznaku $\mathbf{x} = (x,y)^\top$ ovaj sustav možemo kratko zapisati kao

$$\dot{\mathbf{x}} = \mathbf{Ax}.$$

Sa zadanim početnim uvjetima $x(0) = x_0$, $y(0) = y_0$ dobivamo Cauchyjev problem čije rješenje je

$$\mathbf{x}(t) = e^{\mathbf{A}t}\mathbf{x}_0,$$

pri čemu je $\mathbf{x}_0 = (x_0, y_0)^\top$. Drugim riječima, vektor $\mathbf{x}(t)$ nastaje iz $\mathbf{x}_0$ rotacijom za kut t oko ishodišta. Primijetite da je $\ddot{x} = -\dot{y} = -x$, tj. $\ddot{x} + x = 0$, pa funkcija $x = x(t)$ predstavlja rješenje dobro poznate jednadžbe *harmonijskog oscilatora*.

5.2. Funkcija matrice i teorem o preslikavanju spektra

Umjesto polinoma matrice $P(\mathbf{A})$ moguće je definirati čak i funkciju matrice $f(\mathbf{A})$. O tome govori sljedeći važan teorem, koji navodimo bez dokaza.

Teorem 1. *Neka je* $\mathbf{A}$ *kvadratna matrica s kompleksnim koeficijentima. Pretpostavimo da je zadana funkcija* $f : K_r(z_0) \to \mathbf{C}$ *kompleksne varijable, definirana na otvorenom krugu* $K_r(z_0) = \{z \in \mathbf{C} : |z - z_0| < r\}$*, takva da posjeduje Taylorov red*

$$f(z) = \sum_{k=0}^{\infty} \alpha_k (z - z_0)^k$$

s radijusom konvergencije (barem) r *oko* z_0*. Ako je* $\sigma(\mathbf{A}) \subset K_r(z_0)$*, tj. sve vlastite vrijednosti od* $\mathbf{A}$ *su u nutrini kruga konvergencije reda, onda matrični red*

$$f(\mathbf{A}) := \sum_{k=0}^{\infty} \alpha_k (\mathbf{A} - z_0 \mathbf{I})^k$$

konvergira.

Ako je $\lambda \in K_r(z_0)$, onda za matricu $\mathbf{J} = \lambda \mathbf{I} + \mathbf{N}$, gdje je $\mathbf{N}$ nilpotentna matrica s jedinicama na gornjoj sporednoj dijagonali i nulama inače, vrijedi

$$f(\mathbf{J}) = \begin{bmatrix} f(\lambda) & f'(\lambda) & \frac{f''(\lambda)}{2!} & \cdots & \frac{f^{(n-1)}(\lambda)}{(n-1)!} \\ & f(\lambda) & f'(\lambda) & \cdots & \frac{f^{(n-2)}(\lambda)}{(n-2)!} \\ & & & & \vdots \\ & & & \ddots & f'(\lambda) \\ & & & & f(\lambda) \end{bmatrix}.$$

Ako je $f(z)$ funkcija kao u Teoremu 1 i $\mathbf{J}$ Jordanova matrica sa Jordanovim blokovima $\mathbf{J}_1(\lambda_1), \ldots, \mathbf{J}_r(\lambda_r)$, tako da su svi λ_i u nutrini kruga $K_r(z_0)$, onda je

$$f(\mathbf{J}) = \begin{bmatrix} f(\mathbf{J}_1(\lambda_1)) & & & \\ & f(\mathbf{J}_2(\lambda_2)) & & \\ & & \ddots & \\ & & & f(\mathbf{J}_r(\lambda_r)) \end{bmatrix}.$$

Matrica $f(\mathbf{J})$ je gornja trokutasta i na dijagonali se nalaze brojevi $f(\lambda_i)$, gdje su $\lambda_i \in \sigma(\mathbf{J})$.

Nije teško vidjeti da i u Teoremu 1 vrijedi **Teorem o preslikavanju spektra** kao i za funkcije od polinoma: ako je $\sigma(\mathbf{A}) = \{\lambda_1, \ldots, \lambda_n\}$ onda je

$$\sigma(f(\mathbf{A})) = \{f(\lambda_1), \ldots, f(\lambda_n)\},$$

ili kraće zapisano,

$$\sigma\big(f(\mathbf{A})\big) = f(\sigma(\mathbf{A})).$$

Drugim riječima, vlastite vrijednosti matrice $f(A)$ su kompleksni brojevi oblika $f(\lambda)$, gdje je λ vlastita vrijednost od $\mathbf{A}$. Dokaz je sličan kao i u slučaju polinoma. Naime, budući da je matrica $\mathbf{A}$ slična gornjoj trokutastoj $\mathbf{J}$, onda je $f(\mathbf{A})$ slična gornjoj trokutastoj matrici $f(\mathbf{J})$ koja na dijagonali ima brojeve oblika $f(\lambda)$, gdje je $\lambda \in \sigma(\mathbf{A})$. To su upravo vlastite vrijednosti od $f(\mathbf{J})$, dakle i od $f(\mathbf{A})$.

Ako odgovarajući Taylorov red od $f(z)$ konvergira za $z_0 = 0$ (tzv. MacLaurinov red od $f(z)$) s *beskonačnim* radijusom konvergencije, tj.

$$f(z) = \sum_{k=0}^{\infty} \alpha_k z^k$$

konvergira na cijelom $\mathbf{C}$, onda takvu funkciju kompleksne varijable zovemo **cijelom funkcijom**. U tom slučaju je funkcija $f(\mathbf{A})$ definirana za *svaku* kvadratnu matricu $\mathbf{A}$:

$$f(\mathbf{A}) := \sum_{k=0}^{\infty} \alpha_k \mathbf{A}^k.$$

Kao što je poznato, funkcije $f(z) = e^z$, $\sin z$, $\cos z$ su cijele funkcije, pa prema tome redovi

$$e^{\mathbf{A}} := \sum_{k=0}^{\infty} \frac{\mathbf{A}^k}{k!}, \quad \sin \mathbf{A} := \sum_{k=0}^{\infty} (-1)^k \frac{\mathbf{A}^{2k+1}}{(2k+1)!}, \quad \cos \mathbf{A} := \sum_{k=0}^{\infty} (-1)^k \frac{\mathbf{A}^{2k}}{(2k)!}$$

konvergiraju za *svaku* kvadratnu matricu A s kompleksnim koeficijentima.

Primjer 5. Vrijedi:

$$\sin\left(\begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}\right) = \begin{bmatrix} \sin\lambda & \cos\lambda \\ 0 & \sin\lambda \end{bmatrix}.$$

Poznato je također da funkcija $(1-z)^{-1}$ ima MacLaurinov red

$$(1-z)^{-1} = 1 + z + z^2 + \ldots$$

koji konvergira za $|z| < 1$, tj. na $K_1(0)$. Rabeći Teorem 1 dobivamo da ako je $\mathbf{A}$ kvadratna matrica čije vlastite vrijednosti su unutar jediničnog kruga, tj. $|\lambda| < 1$ za sve $\lambda \in \sigma(\mathbf{A})$ (tj. spektralni radijus od $\mathbf{A}$ je manji od 1), onda vrijedi

$$(\mathbf{I} - \mathbf{A})^{-1} = \mathbf{I} + \mathbf{A} + \mathbf{A}^2 + \ldots$$

Izravan dokaz te tvrdnje, bez pozivanja na Teorem 1, vidjet ćemo u sljedećem odjeljku.

5.3. Spektralni radijus i Neumannov red za $(\mathbf{I}-\mathbf{A})^{-1}$

Dobro je poznato da za neki kompleksan broj a vrijedi da $a^k \to 0$ kad $k \to \infty$ onda i samo onda ako je $|a| < 1$. Poopćenje ovog rezultata na slučaj kvadratnih matrica $\mathbf{A}$ daje sljedeći teorem. Podsjetimo se definicije spektralnog radiusa matrice: $r(\mathbf{A}) := \max\{|\lambda| : \lambda \in \sigma(\mathbf{A})\}$.

Teorem 1. *Neka je $\mathbf{A}$ kvadratna matrica s kompleksnim koeficijentima. Ako je $r(\mathbf{A}) < 1$ onda $\mathbf{A}^k \to 0$ kad $k \to \infty$, i obratno.*

DOKAZ. Matrica $\mathbf{A}$ je slična svojoj Jordanovoj formi $\mathbf{J}$. Matrica $\mathbf{A}^k$ je onda slična gornjoj trokutastoj matrici $\mathbf{J}^k$ koja na dijagonali ima λ_l^k, gdje su λ_l vlastite vrijednosti, a na mjestu (i,j), $i<j$ (tj. iznad glavne dijagonale), stoji element oblika $\binom{k}{j-i}\lambda^{k-j+i}$.

Neka je $r(\mathbf{A}) < 1$, tj. $|\lambda_l| < 1$. Očevidno $|\lambda_l^k| = |\lambda_l|^k \to 0$ kada $k \to \infty$. Za čvrste i, j, gdje je $i \leqslant j$, vrijedi da je $\binom{k}{j-i} = \frac{k(k-1)\ldots(k-(j-i)+1)}{(j-i)!}$ polinom u k konstantnog stupnja $j-i$, pa s pomoću L'Hospitalova pravila lako dobijemo da $\binom{k}{j-i}\lambda_l^k$ konvergira k 0 kad $k \to \infty$, jer je $|\lambda_l| < 1$. Prema tome $\mathbf{J}^k$ konvergira k nuli kad $k \to \infty$, dakle i $\mathbf{A}^k$.

Obratno, ako $\mathbf{A}^k$ konvergira k 0, onda i $\mathbf{J}^k$ konvergira k nuli. Budući da su na dijagonali gornje trokutaste matrice $\mathbf{J}^k$ brojevi λ_l^k, gdje su λ_l vlastite vrijednosti od $\mathbf{A}$, onda $\lambda_l^k \to 0$ kad $k \to \infty$, pa mora biti $|\lambda_l| < 1$. Odatle je $r(\mathbf{A}) < 1$. ☺

Iz Matematike 2 znamo da za svaki realan broj x takav da je $|x| < 1$, vrijedi $\frac{1}{1-x} = 1 + x + x^2 + \ldots$. Slično vrijedi i za matrice.

Teorem 2. *Neka je $\mathbf{A}$ kvadratna matrica reda n s kompleksnim koeficijentima i $r(\mathbf{A}) < 1$ (tj. sve vlastite vrijednosti od $\mathbf{A}$ su unutar jediničnog kruga u $\mathbf{C}$). Onda je matrica $\mathbf{I}-\mathbf{A}$ regularna i vrijedi*

$$(\mathbf{I}-\mathbf{A})^{-1} = \mathbf{I} + \mathbf{A} + \mathbf{A}^2 + \ldots$$

pri čemu matrični red konvergira. Obratno, ako red na desnoj strani konvergira, onda je $r(\mathbf{A}) < 1$.

DOKAZ. Neka je $r(\mathbf{A}) < 1$. Ako je $\lambda \in \sigma(\mathbf{A})$ onda je prema teoremu o preslikavanju spektra $1-\lambda \in \sigma(\mathbf{I}-\mathbf{A})$ (uzmi $f(x) = 1-x$). Zbog $|\lambda| < 1$ je $1-\lambda \neq 0$, pa je matrica $\mathbf{I}-\mathbf{A}$ regularna. Iz jednakosti $(\mathbf{I}+\mathbf{A}+\mathbf{A}^2+\ldots+\mathbf{A}^k) - (\mathbf{A}+\mathbf{A}^2+\ldots+\mathbf{A}^k+\mathbf{A}^{k+1}) = \mathbf{I}-\mathbf{A}^{k+1}$ slijedi $(\mathbf{I}-\mathbf{A})(\mathbf{I}+\mathbf{A}+\mathbf{A}^2+\ldots+\mathbf{A}^k) = \mathbf{I}-\mathbf{A}^{k+1}$, tj.

$$\mathbf{I}+\mathbf{A}+\mathbf{A}^2+\ldots+\mathbf{A}^k = (\mathbf{I}-\mathbf{A})^{-1}(\mathbf{I}-\mathbf{A}^{k+1}).$$

Ako na desnoj strani pustimo $k \to \infty$, zbog Teorem 1 dobivamo na desnoj strani $(\mathbf{I}-\mathbf{A})^{-1}$. Time smo dokazali da je $\sum_{k=0}^{\infty}\mathbf{A}^k = (\mathbf{I}-\mathbf{A})^{-1}$.

Obratno, pretpostavimo da red $\sum_{k=0}^{\infty}\mathbf{A}^k$ konvergira. Onda vrijedi $\mathbf{A}^k \to 0$ kad $k \to \infty$ (nuždan uvjet konvergencije reda), pa iz Teorema 1 zaključujemo da je $r(\mathbf{A}) < 1$. ☺

Tu je činjenicu dokazao njemački matematičar F. Neumann (1798.–1895.) koncem 19. st. Matrični red u prethodnom teoremu zove se **Neumannov red** (ne miješati to ime s Johnom von Neumannom).

Sljedeći korolar pokazuje da se za svaki μ izvan kruga polumjera $r(\mathbf{A})$ u kompleksnoj ravnini matrica $(\mu\mathbf{I}-\mathbf{A})^{-1}$ može razviti u konvergentan red potencija od $\mathbf{A}$.

Korolar 3. *Neka je $\mathbf{A}$ zadana matrica i μ kompleksan broj takav da je $|\mu| > r(\mathbf{A})$. Onda je*

$$(\mu\mathbf{I}-\mathbf{A})^{-1} = \frac{1}{\mu}\sum_{k=0}^{\infty}\frac{\mathbf{A}^k}{\mu^k}.$$

DOKAZ. Zbog $r(\mathbf{A}) < |\mu|$ je $r(\mathbf{A}/\mu) < 1$, jer iz $\lambda \in \sigma(\mathbf{A})$ slijedi $\lambda/\mu \in \sigma(\mathbf{A}/\mu)$, pa zbog $|\lambda| < |\mu|$ imamo $|\lambda/\mu| < 1$. Stoga je $I - A/\mu$ regularna matrica, pa invertiranjem matrice $\mu\mathbf{I}-\mathbf{A} = \mu(\mathbf{I}-\mathbf{A}/\mu)$ dobivamo

$$(\mu\mathbf{I}-\mathbf{A})^{-1} = \frac{1}{\mu}\left(\mathbf{I}-\frac{\mathbf{A}}{\mu}\right)^{-1} = \frac{1}{\mu}\sum_{k=0}^{\infty}\left(\frac{\mathbf{A}}{\mu}\right)^k. \quad ☺$$

Može se pokazati da red $\sum_{k=0}\frac{\mathbf{A}^k}{\mu^k}$ divergira ako je μ unutar kruga polumjera $r(\mathbf{A})$ oko ishodišta, tj. za $|\mu| < r(\mathbf{A})$.

Kao što znamo, skup svih kompleksnih brojeva μ za koje je matrica $\mu\mathbf{I}-\mathbf{A}$ regularna, tj. postoji $(\mu\mathbf{I}-\mathbf{A})^{-1}$, zove se **rezolventni skup** matrice $\mathbf{A}$, i označava sa $\rho(\mathbf{A})$. Taj skup je kompleksna ravnina bez vlastitih vrijednosti od $\mathbf{A}$: $\rho(\mathbf{A}) = \mathbf{C}\setminus\sigma(\mathbf{A}) = \mathbf{C}\setminus\{\lambda_1,\ldots,\lambda_n\}$. Funkcija $R(\cdot) : \rho(\mathbf{A}) \subset \mathbf{C} \to M_n$ koja kompleksnom broju μ pridružuje matricu $R(\mu) = (\mu\mathbf{I}-\mathbf{A})^{-1}$ zove se **rezolventa matrica $\mathbf{A}$**. Ako je u nehomogenoj matričnoj jednadžbi

$$(\mu\mathbf{I}-\mathbf{A})\mathbf{x} = \mathbf{h}$$

vektor $\mathbf{h}$ zadan, $\mathbf{x}$ nepoznanica, i $\mu \in \rho(\mathbf{A})$, onda je rješenje jednako $\mathbf{x} = R(\mu)\mathbf{h} = (\mu\mathbf{I}-\mathbf{A})^{-1}\mathbf{h}$. U slučaju kad je $\mu \in \sigma(\mathbf{A})$ imamo **Fredholmovu alternativu**: ako je $\mathbf{h} \perp N(\overline{\mu}\mathbf{I}-\mathbf{A}^*)$ onda postoji rješenje, a u suprotnom rješenja nema.

U žargonu linearne algebre se nerijetko umjesto vlastitih vrijednosti λ matrice $\mathbf{A}$ govori o njenim recipročnim vrijednostima (pod uvjetom $\lambda \neq 0$), koje se zovu **karakteristične vrijednosti matrice**. Prema tome, kompleksni broj μ je karakteristična vrijednost od $\mathbf{A}$ ako postoji vektor $\mathbf{x} \neq 0$ takav da je $\mu\mathbf{A}\mathbf{x} = \mathbf{x}$.

U svjetlu Teorema 1 korisno je znati sljedeći teorem, jer on daje jednostavan dovoljni uvjet da bi spektralni radijus bio < 1.

Propozicija 2. *(a) Za svaku kvadratnu matricu $\mathbf{A}$ vrijedi $r(\mathbf{A}) \leqslant \|\mathbf{A}\|$, gdje je $\|\mathbf{A}\|$*

5.4. Spektralni radijus i spektralna norma matrice

Podsjetimo se da za matrice s koeficijentima iz polja realnih brojeva imamo pojmove *ortogonalnih matrica* i *simetričnih matrica*. U slučaju polja kompleksnih brojeva, njima odgovaraju pojmovi *unitarnih matrica* i *hermitskih matrica*.

Jasno je da je realan broj b nenegativan (pozitivan) onda i samo onda ako je $bx^2 \geqslant 0$ za sve realne brojeve x ($bx^2 > 0$ za sve $x \neq 0$).

Za hermitsku matricu $\mathbf{B}$ reda n (tj. takvu da je $\mathbf{B}^{*}=\mathbf{B}$) kažemo da je **pozitivno semidefinitna matrica** ako je $(\mathbf{Bx} \mid \mathbf{x}) \geqslant 0$ za sve $\mathbf{x} \in \mathbf{C}^n$. Ako je $(\mathbf{Bx} \mid \mathbf{x}) > 0$ za sve $\mathbf{x} \neq 0$, kažemo da je $\mathbf{B}$ **pozitivno definitna matrica**.

Lako se vidi da matrica je $\mathbf{B}$ pozitivno semidefinitna (definitna) onda i samo onda ako su sve njene vlastite vrijednosti nenegativne (pozitivne). Doista, ako je $\mathbf{B}$ pozitivno semidefinitna matrica i $\lambda_i \in \sigma(\mathbf{B})$, onda za pripadajući jedinični vlastiti vektor $\mathbf{x}_i$ vrijedi $0 \leqslant (\mathbf{Bx}_i \mid \mathbf{x}_i) = (\lambda \mathbf{x}_i \mid \mathbf{x}_i) = \lambda_i$. Obratno, ako su sve vlastite vrijednosti λ_i neke hermitske matrice pozitivne, onda je $\mathbf{B}$ pozitivno semidefinitna. Doista, za svaki $\mathbf{x} = x_1\mathbf{e}_1 + \ldots + x_n\mathbf{e}_n$, gdje $(\mathbf{e}_i)$ čine ortonormiranu bazu vlastitih vektora hermitske matrice $\mathbf{B}$, vrijedi

$$
\begin{aligned}
(\mathbf{Bx} \mid \mathbf{x}) &= (x_1\mathbf{Be}_1 + \ldots + x_n\mathbf{Be}_n \,| x_1\mathbf{e}_1 + \ldots + x_n\mathbf{e}_n) \\
&= \lambda_1|x_1|^2 + \ldots + \lambda_n|x_n|^2 \geqslant 0.
\end{aligned}
$$

Zanimljivo je da se iz pozitivno semidefinitne matrice $\mathbf{B}$ može 'izvaditi korijen', tj. postoji jednoznačno određena pozitivno semidefinitna matrica $\mathbf{C}$ takva da je $\mathbf{C}^2 = \mathbf{B}$. Pišemo $\mathbf{C} = \sqrt{\mathbf{B}}$.

Spektralna norma bilo koje *pravokutne* matrice $\mathbf{A}$ tipa $m \times n$ s kompleksnim koeficijentima definira se kao operatorska norma $\|\mathbf{A}\|_2$ generirana Euklidskom normom $\|\mathbf{x}\|_2$, tj.

$$
\|\mathbf{A}\|_2 = \sup_{\mathbf{x} \neq 0} \frac{\|\mathbf{Ax}\|_2}{\|\mathbf{x}\|_2}, \qquad \|\mathbf{x}\|_2 = \sqrt{|x_1|^2 + \ldots + |x_n|^2}.
$$

Sljedeći teorem pokazuje da se spektralna norma matrice $\mathbf{A}$ može vrlo lako izračunati znajući spektar kvadratne matrice $\mathbf{A}^*\mathbf{A}$. Time je naziv spektralne norme opravdan. (U uporabi je i oznaka $\|\mathbf{A}\|_S = \|\mathbf{A}\|_2$.)

Teorem 1. *Neka je $\mathbf{A}$ pravokutna matrica s kompleksnim koeficijentima. Onda je*

$$
\|\mathbf{A}\|_2 = r(\mathbf{A}^*\mathbf{A})^{1/2},
$$

gdje je $\mathbf{A}^*$ hermitski adjungirana matrica od $\mathbf{A}$, tj. $\mathbf{A}^* := \overline{\mathbf{A}}^{\top}$.

DOKAZ. Normu $\|\cdot\|_2$ na $\mathbf{C}^n$ označavat ćemo u dokazu radi kratkoće samo sa $\|\cdot\|$. Matrica $\mathbf{A}^*\mathbf{A}$ je hermitska, jer je $(\mathbf{A}^*\mathbf{A})^* = \mathbf{A}^*\mathbf{A}^{**} = \mathbf{A}^*\mathbf{A}$. Odatle slijedi da su njene vlastite vrijednosti λ_k realne, a pripadni vlastiti vektori mogu se uzeti tako da tvore ortonormiranu bazu $\mathbf{e}_k$ u $\mathbf{C}^n$:

$$
\mathbf{A}^*\mathbf{Ae}_k = \lambda_k\mathbf{e}_k, \qquad (\mathbf{e}_k \mid \mathbf{e}_j) = \delta_{kj}.
$$

Nadalje, matrica $\mathbf{A}^*\mathbf{A}$ je pozitivno semidefinitna:

$$
(\mathbf{A}^*\mathbf{Ax} \mid \mathbf{x}) = (\mathbf{Ax} \mid \mathbf{Ax}) = \|\mathbf{A}x\|^2 \geqslant 0,
$$

pa su vlastite vrijednosti od $\mathbf{A}^*\mathbf{A}$ nenegativne: $0 \leqslant (\mathbf{A}^*\mathbf{Ae}_k \,|\mathbf{e}_k) = (\lambda_k\mathbf{e}_k \mid \mathbf{e}_k) = \lambda_k\|\mathbf{e}_k\|^2 = \lambda_k$. Poredajmo vlastite vrijednosti po veličini:

$$
(0 \leqslant)\ \lambda_1 \leqslant \lambda_2 \leqslant \ldots \leqslant \lambda_n.
$$

Onda je $r(\mathbf{A}^*\mathbf{A}) = \max\{|\lambda_1|, |\lambda_2|, \ldots, |\lambda_n|\} = \lambda_n$.

Odaberimo bilo koji $\mathbf{x} = \sum_{k=1}^n x_k\mathbf{e}_k$ takav da je $\|\mathbf{x}\| = 1$. Vrijedi

$$
\begin{aligned}
\|\mathbf{Ax}\|^2 &= (\mathbf{Ax} \mid \mathbf{Ax}) = (\mathbf{A}^*\mathbf{Ax} \mid \mathbf{x}) \\
&= \Big(\sum_k x_k\mathbf{A}^*\mathbf{Ae}_k \mid \sum_k x_k\mathbf{e}_k\Big) = \Big(\sum_k \lambda_k x_k\mathbf{e}_k \mid \sum_k x_k\mathbf{e}_k\Big) = [\text{zbog } e_i \perp e_j] \\
&= \sum_k |x_k|^2(\lambda_k\mathbf{e}_k \mid \mathbf{e}_k) = \sum_k \lambda_k|x_k|^2 \leqslant [\text{zbog } 0 \leqslant \lambda_k \leqslant \lambda_n] \\
&\leqslant \lambda_n \sum_k |x_k|^2 = \lambda_n
\end{aligned}
$$

Uzimajući supremum po svim $\mathbf{x}$ takvim da je $\|\mathbf{x}\| = 1$ dobivamo $\|\mathbf{A}\|^2 \leqslant \lambda_n$. S druge strane za jedinični vektor $\mathbf{x} = \mathbf{e}_n$ vrijedi $\|\mathbf{Ae}_n\|^2 = (\mathbf{Ae}_n \mid \mathbf{Ae}_n) = (\mathbf{A}^*\mathbf{Ae}_n \mid \mathbf{e}_n) = \lambda_n(\mathbf{e}_n \mid \mathbf{e}_n) = \lambda_n$, pa dobivamo $\|A\|^2 \geqslant \lambda_n$. Dakle, $\|\mathbf{A}\| = \lambda_n^{1/2} = r(\mathbf{A}^*\mathbf{A})^{1/2}$. ☺

Korolar 2. *Neka je $\mathbf{A}$ hermitska matrica, tj. $\mathbf{A}^* = \mathbf{A}$. (a) spektralni radijus od $\mathbf{A}$ jednak je spektralnoj normi, tj. $r(\mathbf{A}) = \|\mathbf{A}\|_2$. (b) Štoviše, za svaki polinom $p(x)$ s realnim koeficijentima je $r(p(\mathbf{A})) = \|p(\mathbf{A})\|_2$.*

DOKAZ. (a) Prema teoremu o preslikavanju spektra je $\sigma(\mathbf{A}^2) = \sigma(\mathbf{A})^2$, tj. spektar matrice $\mathbf{A}^2$ dobiva se kvadriranjem svih elemenata spektra matrice $\mathbf{A}$. Budući da su za hermitsku matricu vlastite vrijednosti λ realne, vrijedi $\lambda^2 = |\lambda|^2$. Prema tome je $r(\mathbf{A}^2) = \max\{\lambda^2 : \lambda \in \sigma(\mathbf{A})\} = \max\{|\lambda|^2 : \lambda \in \sigma(\mathbf{A})\} = \max\{|\lambda| : \lambda \in \sigma(\mathbf{A})\}^2 = r(\mathbf{A})^2$. Iz prethodnog teorema je onda $\|\mathbf{A}\|_2^2 = r(\mathbf{A}^*\mathbf{A}) = r(\mathbf{A}^2) = r(\mathbf{A})^2$, dakle $\|\mathbf{A}\|_2 = r(\mathbf{A})$.

(b) Budući da je $\mathbf{A}$ hermitska matrica, onda je i $p(\mathbf{A})$ hermitska: $p(\mathbf{A})^* = [\text{jer } A \text{ ima realne koeficijente}] = p(\mathbf{A}^*) = p(\mathbf{A})$. Tvrdnja slijedi iz (a). ☺

Primjer 7. Tvrdnja korolara ne vrijedi za matrice koje nisu hermitske. Na pr. za matricu

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

je očevidno $\sigma(\mathbf{A}) = \{0\}$, tj. $r(\mathbf{A}) = 0$. S druge strane je

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

dakle $\sigma(\mathbf{A}^*\mathbf{A}) = \{0, 1\}$. Prema Teoremu 1 je $\|\mathbf{A}\|_2 = r(\mathbf{A}^*\mathbf{A})^{1/2} = 1$, i odatle $r(\mathbf{A}) < \|\mathbf{A}\|_2$.

Primjer 8. Nađimo spektralnu normu $\|\mathbf{A}\|_2$ matrice $\mathbf{A} = \begin{bmatrix} i & 1 \\ 0 & 1+i \end{bmatrix}$. Vrijedi $\mathbf{A}^* = \begin{bmatrix} -i & 1 \\ 0 & 1-i \end{bmatrix}^{\top} = \begin{bmatrix} -i & 0 \\ 1 & 1-i \end{bmatrix}$, pa je

$$
\mathbf{A}^*\mathbf{A} = \begin{bmatrix} 1 & -i \\ i & 3 \end{bmatrix}
$$

Karakteristični polinom ove matrice je $k(\lambda) = \lambda^2 - 4\lambda + 2$. To znači da je $\sigma(\mathbf{A}^*\mathbf{A}) = \{2 - \sqrt{2}, 2 + \sqrt{2}\}$. Budući da je $r(\mathbf{A}^*\mathbf{A}) = 2 + \sqrt{2}$, prema Teoremu 1 je $\|\mathbf{A}\|_2 = \sqrt{2 + \sqrt{2}}$.

5.5. Stabilne matrice

Za kvadratnu matricu $\mathbf{A}$ s kompleksnim koeficijentima kažemo da je **stabilna matrica** ako je $\operatorname{Re}\lambda < 0$ za svaki $\lambda \in \sigma(\mathbf{A})$. Drugim riječima, spektar matrice $\mathbf{A}$ nalazi se u lijevoj poluravnini u odnosu na imaginarnu os kompleksne ravnine. Razlog zašto se takve matrice zovu stabilnima vidljiv je iz sljedećeg teorema.

Teorem 1. *Matrica* $\mathbf{A}$ *je stabilna onda i samo onda ako* $e^{\mathbf{A}t} \to 0$ *kad* $t \to +\infty$.

DOKAZ. Prema Jordanovu teoremu matrica $\mathbf{A}$ je slična svojoj Jordanovoj matrici $\mathbf{J}$, koja je dijagonalna blok matrica sa Jordanovim blokovima $\mathbf{J}(\lambda_1),\dots,\mathbf{J}(\lambda_r)$ na glavnoj dijagonali. Budući da je $\mathbf{A}t = \mathbf{T}(\mathbf{J}t)\mathbf{T}^{-1}$, onda je

$$e^{\mathbf{A}t} = \mathbf{T}e^{\mathbf{J}t}\mathbf{T}^{-1}.$$

Jasno je da $e^{\mathbf{A}t} \to 0$ onda i samo onda ako $e^{\mathbf{J}t} \to 0$ kada $t \to +\infty$.

Neka je t na trenutak fiksiran. Definirajmo funkciju $f(x) = e^{xt}$. Onda je

$$e^{\mathbf{J}t} = f(\mathbf{J}) = \begin{bmatrix} f(\mathbf{J}_1(\lambda_1)) & & & \\ & f(\mathbf{J}_2(\lambda_2)) & & \\ & & \ddots & \\ & & & f(\mathbf{J}_r(\lambda_r)) \end{bmatrix}.$$

Označimo radi kratkoće bilo koji od brojeva λ_i sa λ. Svaka od matrica $f(\mathbf{J}(\lambda))$ je gornja trokutasta:

$$f(\mathbf{J}(\lambda)) = \begin{bmatrix} f(\lambda) & f'(\lambda) & \frac{f''(\lambda)}{2!} & \dots & \\ & f(\lambda) & f'(\lambda) & \dots & \\ & & & & \vdots \\ & & & \ddots & f'(\lambda) \\ & & & & f(\lambda) \end{bmatrix} = \begin{bmatrix} e^{\lambda t} & te^{\lambda t} & \frac{t^2}{2!}e^{\lambda t} & \dots & \\ & e^{\lambda t} & te^{\lambda t} & \dots & \\ & & & & \vdots \\ & & & \ddots & te^{\lambda t} \\ & & & & e^{\lambda t} \end{bmatrix}.$$

Tvrdnja će biti dokazana ako se pokaže da je $\operatorname{Re}\lambda < 0$ onda i samo onda ako $t^k e^{\lambda t} \to 0$ kad $t \to +\infty$ (za sve $k = 0,1,2,\dots$).

($\Rightarrow$) Neka je $\lambda = c + di \in \sigma(\mathbf{A})$ takav da je $\operatorname{Re}\lambda < 0$, tj. $c < 0$. Onda iz $|e^{dti}| = 1$ slijedi da $|e^{\lambda t}| = |e^{ct}e^{dti}| = e^{ct} \to 0$ kad $t \to +\infty$. Isto tako $|te^{\lambda t}| = te^{ct} \to 0$ primjenom L'Hospitalova pravila:

$$\lim_{t\to+\infty} te^{ct} = [\infty\cdot 0] = \lim_{t\to+\infty}\frac{t}{e^{-ct}} = \left[\frac{\infty}{\infty}\right] = \lim_{t\to+\infty}\frac{1}{-ce^{-ct}} = 0.$$

Na isti način k-strukom primjenom L'Hospitalova pravila dobivamo da i $t^k e^{ct} \to 0$ kad $t \to +\infty$.

($\Leftarrow$) Pretpostavimo da $t^k e^{\lambda t} \to 0$ kad $t \to +\infty$, za bilo koji $k = 0,1,2,\dots$. Za $k = 0$ imamo $e^{\lambda t} \to 0$, tj. za $\lambda = c + di$ vrijedi $e^{ct} \to 0$ kad $t \to +\infty$. To je moguće jedino kad je $c < 0$. Dakle $\operatorname{Re}\lambda < 0$. ☺

Ako je matrica $\mathbf{A}$ stabilna, onda $e^{\mathbf{A}t} \to 0$ čak eksponencijalnom brzinom. Točnije ako odaberemo bilo koji *negativan* broj α takav da je $\alpha > \max\{\operatorname{Re}\lambda : \lambda \in \sigma(\mathbf{A})\}$, onda postoji konstanta $M > 0$ neovisna od t takva da je

$$\|e^{\mathbf{A}t}\| \leqslant Me^{\alpha t}.$$

Doista, za $\lambda = c + di \in \sigma(\mathbf{A})$ je $|t^k e^{\lambda t}| = t^k e^{ct} = t^k e^{(c-\alpha)t}e^{\alpha t}$, a izraz $t^k e^{(c-\alpha)t}$ je odozgor omeđen nekim brojem $M > 0$, jer on zbog $c < \alpha$ teži u nulu kad $t \to +\infty$.

Primjer 9. Matrica

$$\mathbf{A} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$

gdje su a i b bilo koja dva realna broja, je stabilna onda i samo onda ako je $a < 0$. Doista, pišući

$$\mathbf{A} = a\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + b\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = a\mathbf{I} + b\mathbf{J}$$

koristeći teorem o preslikavanju spektra od $\mathbf{J}$ (koji se sastoji od $\pm i$) s funkcijom $f(x) = a + bx$, dobivamo da je $\sigma(\mathbf{A}) = \sigma(f(\mathbf{J})) = f(\sigma(\mathbf{J})) = \{f(i), f(-i)\} = \{a+bi, a-bi\}$, jer je $\sigma(\mathbf{J}) = \{i,-i\}$ (isti rezultat se lako dobije i izravnim računom, preko karakterističnog polinoma matrice $\mathbf{A}$). Dakle, matrica $\mathbf{A}$ stabilna onda i samo onda ako je $a < 0$. Naravno, spektar matrice $\mathbf{A}$ može se lako izračunati i izravno, preko njenog karakterističnog polinoma.

Nije teško vidjeti da za matricu $\mathbf{A} = a\mathbf{I} + b\mathbf{J}$ iz ovog primjera vrijedi da je

$$e^{\mathbf{A}} = e^a\begin{bmatrix} \cos b & -\sin b \\ \sin b & \cos b \end{bmatrix}.$$

Doista, budući da matrice $a\mathbf{I}$ i $b\mathbf{J}$ komutiraju (s obzirom na množenje), onda je (vidi Primjedbu 1 i Primjer 4):

$$e^{\mathbf{A}} = e^{a\mathbf{I}+b\mathbf{J}} = e^{a\mathbf{I}}e^{b\mathbf{J}} = e^a Ie^{b\mathbf{J}} = e^a e^{b\mathbf{J}} = e^a\begin{bmatrix} \cos b & -\sin b \\ \sin b & \cos b \end{bmatrix}.$$

Stabilne matrice zovu se još i Hurwitzove matrice.

5.6. Geršgorinov teorem o krugovima

Geršgorinov teorem omogućava lokalizaciju vlastitih vrijednosti matrice $\mathbf{A}$, i vrlo je koristan ako je $\mathbf{A}$ velikog reda, kada ne možemo lako pronaći sve vlastite vrijednosti. Za matricu $\mathbf{A}$ reda n vrlo je teško računati sve nultočke karakterističnog polinoma ako je n velik. Taj zanimljiv teorem otkrio je ruski matematičar S. Geršgorin godine 1931. Za kvadratnu matricu $\mathbf{A} = (a_{ij})$ reda n označimo

$$\begin{aligned} r_1 &= |a_{12}| + |a_{13}| + \ldots + |a_{1n}|, \\ r_2 &= |a_{21}| + |a_{23}| + \ldots + |a_{2n}|, \\ &\ldots \\ r_n &= |a_{n1}| + |a_{n2}| + \ldots + |a_{n\,n-1}|. \end{aligned}$$

Drugim riječima, r_i dobiva se kao zbroj apsolutnih vrijednosti brojeva u i-tome retku matrice $\mathbf{A}$, ali bez dijagonalnog elementa $|a_{ii}|$:

$$r_i = \sum_{\substack{j=1\\ j\neq i}}^{n} |a_{ij}|.$$

Definirajmo Geršgorinove krugove u kompleksnoj ravnini:

$$D_i = \{z \in \mathbf{C} : |z - a_{ii}| \leqslant r_i\}.$$

Za matricu $\mathbf{A}$ kažemo da je **dijagonalno dominantna matrica** ako je u svakom njenom retku apsolutna vrijednost broja na dijagonali veća od zbroja apsolutnih vrijednosti ostalih koeficijenata u tom retku, tj.

$$|a_{ii}| > \sum_{\substack{j=1\\ j\neq i}}^{n} |a_{ij}|.$$

Primijetimo da ako je matrica $\mathbf{A}$ dijagonalno dominantna, tj. $|a_{ii}| > r_i$ za sve i, onda niti jedan krug D_i ne sadrži $z = 0$.

Teorem 1. (Geršgorin, 1931.) *Svaka vlastita vrijednost matrice $\mathbf{A}$ sadržana je u barem jednom krugu D_i, tj.*

$$\sigma(\mathbf{A}) \subseteq \bigcup_{i=1}^{n} D_i.$$

Drugim riječima, $\{\lambda_1, \ldots, \lambda_n\} \subseteq D_1 \cup \ldots \cup D_n$.

Prije dokaza teorema pogledajmo jedan važan posljedak tog teorema.

Korolar 2. *Neka je $\mathbf{A}$ dijagonalno dominantna matrica.*

(a) Onda je $\mathbf{A}$ regularna.

(b) Ako su svi dijagonalni elementi pozitivni, tj. $a_{ii} > 0$, onda sve vlastite vrijednosti leže u desnoj poluravnini, tj. $\mathrm{Re}\,\lambda > 0$ za sve $\lambda \in \sigma(\mathbf{A})$.

(c) Ako su svi dijagonalni elementi negativni, tj. $a_{ii} < 0$, onda sve vlastite vrijednosti leže u lijevoj poluravnini, tj. $\mathrm{Re}\,\lambda < 0$ za sve $\lambda \in \sigma(\mathbf{A})$. Drugim riječima, matrica $\mathbf{A}$ je stabilna.

Dokaz. (a) Stroga dijagonalna dominantnost znači da je $|a_{ii}| > r_i$, pa D_i ne sadrži $z = 0$. Prema Geršgorinovu teoremu onda niti $\sigma(\mathbf{A})$ ne sadrži 0. Odatle slijedi da su svi $\lambda_i \neq 0$, pa je $\mathbf{A}$ regularna.

(b) Svaki Geršgorinov krug ima središte na pozitivnom dijelu x-osi, i sadržan je u desnoj poluravnini.

(c) Svaki Geršgorinov krug ima središte na negativnom dijelu x-osi, i sadržan je u lijevoj poluravnini. ☺

Dokaz Geršgorinova teorema. Neka je $\lambda \in \sigma(\mathbf{A})$, tj. λ je vlastita vrijednost od $\mathbf{A}$. Treba dokazati da je λ u nekom od krugova D_i. Vrijedi $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$ za neki $\mathbf{x} = [x_1, \ldots, x_n]^\top \neq 0$. Neka je m maksimum apsolutnih vrijednosti komponenata od $\mathbf{x}$. Vektor $\mathbf{y} = \frac{1}{m}\mathbf{x}$ je također vlastiti vektor za λ. Maksimum apsolutnih vrijednosti komponenta od y iznosi 1. Neka je taj maksimum dostignut na i-toj komponenti, tj. $|y_i| = 1$. U jednakosti $\mathbf{A}\mathbf{y} = \lambda\mathbf{y}$ pogledajmo i-komponentu s lijeva (to je umnožak i-tog retka matrice $\mathbf{A}$ i stupca $\mathbf{y}$) i s desna:

$$a_{i1}y_1 + a_{i2}y_2 + \ldots + a_{in}y_n = \lambda y_i.$$

Oduzimajući $a_{ii}y_i$ lijevo i desno, dobivamo

$$\sum_{\substack{j=1\\ j\neq i}}^{n} a_{ij}y_j = (\lambda - a_{ii})y_i.$$

Rabeći $|y_i| = 1$ i nejednakost trokuta, slijedi

$$|\lambda - a_{ii}| = |\sum_{\substack{j=1\\ j\neq i}}^{n} a_{ij}y_j| \leqslant \sum_{\substack{j=1\\ j\neq i}}^{n} |a_{ij}y_j| \leqslant \sum_{\substack{j=1\\ j\neq i}}^{n} |a_{ij}| = r_i. \quad ☺$$

Primjedba 2. Vrijedi i ovakvo profinjenje Geršgorinova teorema. Ako se unija svih Geršgorinovih krugova sastoji od više komponenata povezanosti, onda svaka komponenta sadrži barem jednu vlastitu vrijednost. Štoviše, ako je neka komponenta povezanosti dobivena kao unija m krugova, onda ona sadrži točno m vlastitih vrijednosti računajući i njihovu kratnost.

Primjedba 3. Ako su *svi* dijagonalni elementi kvadratne matrice $\mathbf{A}$ jednaki 0, tj. $a_{ii} = 0$ za sve i, onda su Geršgorinovi krugovi koncentrični sa zajedničkim središtem u ishodištu Gaussove ravnine, pa je njihova unija krug maksimalnog radijusa $\max_i r_i$. Posebno, ako su svi $r_i < 1$, onda je $r(\mathbf{A}) < 1$, pa odgovarajući Neumannov red $\sum_{k=1}^{\infty} \mathbf{A}^k$ konvergira k matrici $(\mathbf{I} - \mathbf{A})^{-1}$. Vidi Teorem 2 u Odjeljku 5.3 gore. Tu ćemo situaciju imati i u Odjeljku 5.8 niže, posvećenom nekim iterativnim metodama u Linearnoj algebri.

5.7. Bauerov teorem o Cassinijevim ovalima

Uz pomoć Cassinijevih ovala njemački matematičar Bauer je godine 1946. dobio zanimljivo poopćenje Geršgorinova teorema. Cassinijev oval definira se kao skup svih točaka u kompleksnoj ravnini takav da je umnožak njihovih udaljenosti od dvije čvrste točke jednak zadanoj konstanti. Ako su zadana dva kompleksna broja a_1 i a_2, onda Cassinijev oval čine svi kompleksni brojeva z za koje je

$$|z-a_1|\cdot|z-a_2| = r^2$$

gdje je r zadani pozitivan broj. Ako je $a_1 = a_2$, dobivamo kružnicu polumjera r. Ako je $|a_1-a_2| < 2r$, dobivamo povezanu zatvorenu krivulju u obliku kikirikija. Ako je $|a_1-a_2| = 2r$, dobivamo krivulju koja se zove lemniskata, koja sliči na znak ∞. Za $|a_1-a_2| > 2r$ Cassinijev oval se sastoji od dvije "jajaste" krivulje od kojih jedna sadrži a_1 a druga a_2. Pogledajte odgovarajuće izvore na mreži s pomoću ključnih riječi *Cassini ovals*.

`etna.mcs.kent.edu/vol.8.1999/pp15-20.dir/gershini.html`

Skup svih $z \in \mathbf{C}$ za koje vrijedi nejednakost

$$|z-a_1|\cdot|z-a_2| \leqslant r^2$$

također zovemo **Cassinijevim ovalom**.

Uz zadanu matricu $\mathbf{A}$, za naše potrebe definiramo sljedeće Cassinijeve ovale za svaki par $i \neq j$:

$$D_{ij} = \{z \in \mathbf{C} : |z-a_{ii}|\cdot|z-a_{jj}| \leqslant r_i r_j\}.$$

Primijetimo da su $i, j \in \{1,\ldots,n\}$, pa je broj tih ovala jednak najviše $\binom{n}{2} = \frac{1}{2}n(n-1)$.

Teorem 1. (Bauer, 1946) *Neka je* $\mathbf{A}$ *kompleksna matrica reda* n, $n \geqslant 2$. *Onda je*

$$\sigma(\mathbf{A}) \subset \bigcup_{i\neq j} D_{ij}.$$

Drugim riječima, za svaku vlastitu vrijednost λ *matrice* $\mathbf{A}$ *postoji neki Cassinijev oval* D_{ij} *u kojem je* λ *sadržan, tj.*

$$|\lambda-a_{ii}|\cdot|\lambda-a_{jj}| \leqslant r_i r_j.$$

DOKAZ. Primijenit ćemo sličan postupak kao u dokazu Geršgorinova teorema, uz malu izmjenu. Neka $\lambda \in \sigma(\mathbf{A})$ i $\mathbf{A}x = \lambda x$, $x \neq 0$. Možemo pretpostaviti da je $\|x\|_\infty = 1$, tj. $|x_j| \leqslant 1$ i postoji k takav daje $x_k = 1$ (u suprotnom pronađemo komponentu x_k s najvećom apsolutnom vrijednošću među komponentama od x, i zatim gledamo vlastiti vektor $\frac{1}{x_k}x$ umjesto x).

Ako su svi $x_j = 0$ za $j \neq k$, onda gledajući jednadžbu $\mathbf{A}x = \lambda x$ na k-toj komponenti dobivamo $a_{kk} = \lambda$. Prema tome je $\lambda \in D_{kl}$ za bilo koji $l \neq k$, jer je $|\lambda - a_{kk}| = 0$.

Pretpostavimo da nisu svi $x_j = 0$ za $j \neq k$, i neka je onda x_l druga najveća komponenta po apsolutnoj vrijednosti:

$$1 = x_k \geqslant |x_l| > 0, \qquad |x_l| \geqslant |x_j|,$$

za sve $j \neq k$. Gledajući k-tu koponentu u $\mathbf{A}x = \lambda x$ dobivamo $\sum_{j=1}^{n} a_{kj}x_j = \lambda x_k$. Oduzimajući $a_{kk}x_k$ s lijeva i desna dobivamo

$$(\lambda - a_{kk})x_k = \sum_{\substack{j=1\\ j\neq k}}^{n} a_{kj}x_j. \tag{1}$$

Na isti način, gledajući l-tu komponentu u $Ax = \lambda x$ dobivamo

$$(\lambda - a_{ll})x_l = \sum_{\substack{j=1\\ j\neq l}}^{n} a_{lj}x_j. \tag{2}$$

Uzimajući apsolutnu vrijednost u (1) i rabeći $x_k = 1$ dobivamo

$$|\lambda - a_{kk}| \leqslant \sum_{\substack{j=1\\ j\neq k}}^{n} |a_{kj}x_j| \leqslant \sum_{\substack{j=1\\ j\neq k}}^{n} |a_{kj}||x_l| = r_k|x_l|.$$

Iz (2) slijedi zbog $|x_j| \leqslant 1$ za sve j,

$$|\lambda - a_{ll}|\cdot|x_l| \leqslant \sum_{\substack{j=1\\ j\neq l}}^{n} |a_{lj}x_j| \leqslant \sum_{\substack{j=1\\ j\neq l}}^{n} |a_{lj}| = r_l.$$

Množenjem prethodnih dviju nejednakosti, i zatim skraćivanjem s $|x_l| > 0$, dobivamo $|\lambda - a_{kk}|\cdot|\lambda - a_{ll}| \leqslant r_k r_l$. ☺

Nije teško vidjeti da je Bauerov teorem doista bolji od Geršgorinova. Vrijedi naime

$$D_{ij} \subseteq D_i \cup D_j,$$

odakle slijedi da je $\sigma(\mathbf{A}) \subseteq \cup_{i,j} D_{ij} \subseteq \cup_i D_i$, tj. Bauerov teorem daje bolju lokalizaciju spektra nego Geršgorinov.

Da bi dokazali gornju inkluziju, neka je $z \in D_{ij}$, tj. $|z-a_{ii}|\cdot|z-a_{jj}| \leqslant r_i r_j$. Pretpostavimo da $z \notin D_i$ (u suprotnom nemamo što dokazivati), tj. $|z-a_{ii}| > r_i$. Onda je

$$r_i|z-a_{jj}| \leqslant |z-a_{ii}|\cdot|z-a_{jj}| \leqslant r_i r_j,$$

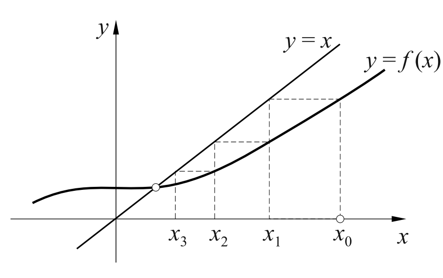
odakle (uz pretpostavku $r_i \neq 0$) slijedi $|z-a_{jj}| \leqslant r_j$, tj. $z \in D_j$. Ako je $r_i = 0$, onda je $z = a_{jj} \in D_j$.

5.8. Jacobijeva i Gauss–Seidelova iterativna metoda

Da bismo ukratko motivirali iterativnu metodu (uzastopnih aproksimacija), neka je zadana bilo koja diferencijabilna funkcija $f : \mathbf{R} \to \mathbf{R}$ (ne nužno linearna), za koju postoji broj $q \in (0,1)$ takav da je ispunjen *uvjet kontrakcije* (ili sažimanja):

$$|f'(x)| \leqslant q \quad \text{za sve } x \in \mathbf{R}.$$

Vrlo je lako uvjeriti se (crtanjem grafova funkcija $y = f(x)$ i $y = x$) da *jednadžba fiksne točke*, $f(x) = x$, posjeduje točno jedno rješenje. To rješenje zovemo *fiksnom točkom* funkcije f.



Sl. 5.1. Funkcija $f : \mathbf{R} \to \mathbf{R}$ koja je kontrakcija.

Doista, budući da f ne raste brže od pravca $y = x$, onda se njihovi grafovi sigurno sijeku u fiksnoj točki od f. (Taj rezultat je specijalan slučaj puno općenitijeg *Banachova teorema o fiksnoj točki*.) Štoviše, do tog se rješenja može doći tako da odaberemo *bilo koju* početnu iteraciju $x_0 \in \mathbf{R}$, te zatim konstruiramo slijed realnih brojeva $(x_k)_{k\geqslant 0}$ definiran sa

$$x_{k+1} = f(x_k), \quad k = 0, 1, 2, \ldots$$

Vrlo lako je zaključiti da svaki takav slijed konvergira prema fiksnoj točki. Iz tog razloga članove navedenog slijeda zovemo *sukcesivnim* (uzastopnim) *aproksimacijama* fiksne točke. Razlog zašto se gore navedeni uvjet zove uvjetom kontrakcije, vidljiv je odmah iz Lagrangeova teorema o srednjoj vrijednosti: za sve realne brojeve a i b vrijedi $|f(b) - f(a)| = |f'(\xi)(b-a)| \leqslant q|b-a|$, tj. udaljenost između $f(a)$ i $f(b)$ se smanjila u odnosu na udaljenost između a i b.

Prijeđimo sada na jedan problem iz Linearne algebre.

Pretpostavimo da nam imamo linearni sustav $\mathbf{Ax} = \mathbf{b}$ s regularnom kvadratnom matricom $\mathbf{A}$, gdje je vektor $\mathbf{b}$ na desnoj strani zadan, a $\mathbf{x}$ je nepoznanica. Pretpostavimo da su svi dijagonalni elementi a_{ii} različiti od nula. Prva jednadžba u sistemu $\mathbf{Ax} = \mathbf{b}$ glasi $a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1$. Iz nje izračunamo x_1:

$$x_1 = -\frac{a_{12}}{a_{11}}x_2 - \frac{a_{13}}{a_{11}}x_3 - \ldots - \frac{a_{1n}}{a_{11}}x_n + \frac{b_1}{a_{11}}.$$

Iz druge jednadžbe izračunamo x_2,

$$x_2 = -\frac{a_{21}}{a_{22}}x_1 - \frac{a_{23}}{a_{22}}x_3 - \ldots - \frac{a_{2n}}{a_{22}}x_n + \frac{b_1}{a_{22}}.$$

itd. do x_n. Na taj način dolazimo do sistema oblika

$$\mathbf{x} = \mathbf{Tx} + \mathbf{b}'$$

koji je ekvivalentan s $\mathbf{Ax} = \mathbf{b}$. Novi sistem je zapisan kao problem nalaženja fiksne točke funkcije $F : \mathbf{R}^n \to \mathbf{R}^n$ zadane sa $F(\mathbf{x}) = \mathbf{Tx} + \mathbf{b}'$, tj. u obliku $\mathbf{x} = F(\mathbf{x})$. Rješenje jednadžbe $\mathbf{x} = \mathbf{Tx} + \mathbf{b}'$ nalazimo **Jacobijevom iterativnom metodom**:

$$\mathbf{x}^{(k+1)} = \mathbf{Tx}^{(k)} + \mathbf{b}'$$
$$\mathbf{x}^{(0)} = 0.$$

Zapravo, za početnu vrijednost $\mathbf{x}^{(0)}$ možemo uzeti bilo koji vektor iz $\mathbf{C}^n$, ne nužno 0. Drugim riječima, računamo slijed vektora u $\mathbf{C}^n$: $\mathbf{x}^{(0)}$, $\mathbf{x}^{(1)} = F(\mathbf{x}^{(0)})$, $\mathbf{x}^{(2)} = F(\mathbf{x}^{(1)})$ itd. Uz neke vrlo općenite uvjete može se postići da taj slijed konvergira točno prema rješenju jednadžbe $\mathbf{x} = F(\mathbf{x})$, tj. $\mathbf{x} = \mathbf{Tx} + \mathbf{b}'$, tj. $\mathbf{Ax} = \mathbf{b}$.

Mala varijacija Jacobijeve metode jest **Gauss-Seidelova iterativna metoda**. Opišimo ju na primjeru rješavanja sistema tri jednadžbe s tri nepoznanice. Najprije, Jacobijeva metoda opisana je sa

$$\begin{aligned}
x_1^{(k+1)} &= t_{11}x_1^{(k)} + t_{12}x_2^{(k)} + t_{13}x_3^{(k)} + b_1' \\
x_2^{(k+1)} &= t_{21}x_1^{(k)} + t_{22}x_2^{(k)} + t_{23}x_3^{(k)} + b_2' \\
x_3^{(k+1)} &= t_{31}x_1^{(k)} + t_{32}x_2^{(k)} + t_{33}x_3^{(k)} + b_3',
\end{aligned}$$

gdje su svi $t_{ii} = 0$. Kao što vidimo, u trenutku dok prelazimo na računanje $x_2^{(k+1)}$, vrijednost $x_1^{(k+1)}$ je već poznata. Stoga nju uvrštavamo u desnu stranu druge jednadžbe umjesto $x_1^{(k)}$. Na sličan način, prije nego prijeđemo na računanje $x_3^{(k+1)}$, vrijednosti $x_1^{(k+1)}$ i $x_2^{(k+1)}$ su poznate, pa u desnu stranu uvrštavamo te dvije vrijednosti:

$$\begin{aligned}
x_1^{(k+1)} &= t_{11}x_1^{(k)} + t_{12}x_2^{(k)} + t_{13}x_3^{(k)} + b_1' \\
x_2^{(k+1)} &= t_{21}x_1^{(k+1)} + t_{22}x_2^{(k)} + t_{23}x_3^{(k)} + b_2' \\
x_3^{(k+1)} &= t_{31}x_1^{(k+1)} + t_{32}x_2^{(k+1)} + t_{33}x_3^{(k)} + b_3'.
\end{aligned}$$

Često je konvergencija Gauss-Seidelove metode brža od Jacobijeve, ali ne uvijek. Može se dogoditi da za neki sistem Jacobijeva metoda konvergira, a Gauss-Seidelova divergira, i obratno.

Da bismo točno opisali matricu $\mathbf{T}$ u ovisnosti od $\mathbf{A}$, rastavimo najprije matricu $\mathbf{A}$ na zbroj $\mathbf{L} + \mathbf{D} + \mathbf{U}$, gdje je

(a) $\mathbf{D}$ dijagonalni dio od $\mathbf{A}$, tj. matrica s brojevima a_{ii} na dijagonali,
(b) matrica $\mathbf{L}$ je dolnji trokutasti dio od $\mathbf{A}$, tj. elementi su joj a_{ij} za $j < i$, ostali nula,
(c) matrica $\mathbf{U}$ je gornji trokutasti dio od $\mathbf{A}$, tj. elementi su joj a_{ij} za $j > i$, ostali nula.

Pogledajmo primjer rastava $\mathbf{A} = \mathbf{L} + \mathbf{D} + \mathbf{U}$:

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

Teorem 1. *Neka je $\mathbf{A}$ kvadratna matrica. (a) Jacobijeva iterativna metoda konvergira k rješenju jednadžbe $\mathbf{Ax} = \mathbf{b}$ za svaki $\mathbf{b}$ onda i samo onda ako je*

$$r(\mathbf{D}^{-1}(\mathbf{L} + \mathbf{U})) < 1,$$

gdje je $r(\cdot)$ spektralni radijus matrice.

(b) Gauss-Seidelova iterativna metoda konvergira k rješenju jednadžbe $\mathbf{Ax} = \mathbf{b}$ za svaki $\mathbf{b}$ onda i samo onda ako je

$$r((\mathbf{D} + \mathbf{L})^{-1}\mathbf{U}) < 1.$$

DOKAZ. (a) Jednadžbu $\mathbf{Ax} = \mathbf{b}$ možemo zapisati kao $(\mathbf{L} + \mathbf{D} + \mathbf{U})\mathbf{x} = \mathbf{b}$, tj. $\mathbf{Đ} = -(\mathbf{L} + \mathbf{U})\mathbf{x} + \mathbf{b}$, tj. $\mathbf{x} = -\mathbf{D}^{-1}(\mathbf{L} + \mathbf{U})\mathbf{x} + \mathbf{D}^{-1}\mathbf{b}$. Uz oznake

$$\mathbf{T} = -\mathbf{D}^{-1}(\mathbf{L} + \mathbf{U}), \quad \mathbf{b}' = \mathbf{D}^{-1}\mathbf{b}$$

ova jednadžba postaje $\mathbf{x} = \mathbf{Tx} + \mathbf{b}'$. Iz $\mathbf{x}^{(0)} = 0$ dobivamo $\mathbf{x}^{(1)} = \mathbf{b}'$, $\mathbf{x}^{(2)} = \mathbf{Tb}' + \mathbf{b}' = (\mathbf{T} + \mathbf{I})\mathbf{b}'$, $\mathbf{x}^{(3)} = \mathbf{Tx}^{(2)} + \mathbf{b}' = \mathbf{T}(\mathbf{T} + \mathbf{I})\mathbf{b}' + \mathbf{b}' = (\mathbf{T}^2 + \mathbf{T} + \mathbf{I})\mathbf{b}'$, i dalje induktivno, $\mathbf{x}^{(k)} = (\mathbf{T}^{k-1} + \ldots + \mathbf{T} + \mathbf{I})\mathbf{b}'$. Vidimo da slijed $\mathbf{x}^{(k)}$ konvergira za sve $\mathbf{b}$ onda i samo onda ako red $\mathbf{I} + \mathbf{T} + \mathbf{T}^2 + \ldots$ konvergira, a to će biti onda i samo onda ako je $r(\mathbf{T}) < 1$.

(b) Promatrajući jednadžbu $\mathbf{x} = -\mathbf{D}^{-1}(\mathbf{L} + \mathbf{U})\mathbf{x} + \mathbf{b}'$, tj. $\mathbf{x} = -\mathbf{D}^{-1}\mathbf{Lx} - \mathbf{D}^{-1}\mathbf{Ux} + \mathbf{b}'$, vidimo da Gauss-Seidelova metoda varijablama koje odgovaraju dolnjem trokutu u $\mathbf{L}$ pridjeljuje već izračunate vrijednosti. Prema tome je $\mathbf{x}^{(k+1)} = -\mathbf{D}^{-1}\mathbf{Lx}^{(k+1)} - \mathbf{D}^{-1}\mathbf{Ux}^{(k)} + \mathbf{b}'$, tj. $(\mathbf{I} + \mathbf{D}^{-1}\mathbf{L})\mathbf{x}^{(k+1)} = -\mathbf{D}^{-1}\mathbf{Ux}^{(k)} + \mathbf{b}'$. Da bi izračunali $\mathbf{x}^{(k+1)}$, rabimo $(\mathbf{I} + \mathbf{D}^{-1}\mathbf{L})^{-1}\mathbf{D}^{-1} = [\mathbf{D}(\mathbf{I} + \mathbf{D}^{-1}\mathbf{L})]^{-1} = (\mathbf{D} + \mathbf{L})^{-1}$. Prema tome je $\mathbf{x} = \mathbf{Tx} + \mathbf{b}''$ gdje je $\mathbf{T} = -(\mathbf{D} + \mathbf{L})^{-1}\mathbf{U}$ i $\mathbf{b}'' = (\mathbf{D} + \mathbf{L})^{-1}\mathbf{b}'$. Slično kao u (a), Gauss-Seidelova metoda konvergira onda i samo onda ako je $r(\mathbf{T}) < 1$. ☺

Dijagonalna dominantnost matrice $\mathbf{A}$ osigurava konvergenciju obje metode.

Korolar 2. *Neka je $\mathbf{A}$ dijagonalno dominantna matrica. Onda Jacobijeva i Gauss-Seidelova metoda konvergiraju k rješenju problema $\mathbf{Ax} = \mathbf{b}$.*

DOKAZ. Zbog $t_{ij} = -a_{ij}/a_{ii}$ za $i \neq j$ i $t_{ii} = 0$, zbroj apsolutnih vrijednosti elemenata u i-tom retku matrice $\mathbf{T}$ jednak je

$$\sum_j |t_{ij}| = \frac{1}{|a_{ii}|} \sum_{\substack{j=1 \\ j\neq i}} |a_{ij}|.$$

Iz dijagonalne dominantnosti od $\mathbf{A}$ dobivamo $\sum_j |t_{ij}| < 1$. Odatle slijedi da je i $\|\mathbf{T}\|_\infty = \max_i \sum_j |t_{ij}| < 1$. Budući da je $r(\mathbf{T}) \leqslant \|\mathbf{T}\|_\infty < 1$, vidi Propoziciju 1 u Odjeljku 5.1, tvrdnja slijedi iz Teorema 1. ☺

Opišimo još i kriterij za zaustavljanje iterativne metode za rješavanje sistema $\mathbf{Ax} = \mathbf{b}$.

U tu svrhu uvedimo pojam apsolutne i relativne pogreške za izračunavanje točne vrijednosti $\mathbf{x}$ rješenja nekog problema. Pretpostavimo da smo nekom iterativnom metodom našli približnu vrijednost $\mathbf{x}^{(k)}$. Onda apsolutnom pogreškom zovemo iznos $\|\mathbf{x} - \mathbf{x}^{(k)}\|$. **Relativnom pogreškom** zovemo broj $\|\mathbf{x} - \mathbf{x}^{(k)}\|/\|\mathbf{x}\|$.

Za ilustraciju, ako imamo realan broj $x = 2$ i $x^{(k)} = 2.1$ za neku iteraciju k, onda je apsolutna pogreška 0.1, a relativna $0.1/2 = 0.05$. Ako je $x = 2000$ i $x^{(k)} = 2000.1$, onda je apsolutna pogreška opet 0.1, ali relativna pogreška (tj. pogreška *s obzirom* na točan iznos) je mnogo manja: $0.1/2000 = 0.00005$.

Razumije se, u konkretnom problemu relativnu pogrešku ne možemo izračunati jer unaprijed ne znamo točnu vrijednost $\mathbf{x}$. Stoga se relativna pogreška zamjenjuje s izrazom

$$\frac{\|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|}{\|\mathbf{x}^{(k)}\|}$$

koji ne ovisi o $\mathbf{x}$. Slijed iteracija $\mathbf{x}^{(k)}$ računamo sve dok taj kvocjent ne bude manji od nekog zadanog (malog) broja $\varepsilon > 0$. Za taj k ne možemo na žalost reći koliko je $\mathbf{x}^{(k)}$ udaljen od točnog rješenja $\mathbf{x}$.

Ukoliko je matrica $\mathbf{A}$ u problemu $\mathbf{Ax} = \mathbf{b}$ dijagonalno dominantna, onda možemo procijeniti apsolutnu pogrešku k-te iteracije Jacobijeve metode, tj. $\|\mathbf{x}^{(k)} - \mathbf{x}\|_\infty$. Primijetite da je taj iznos maksimum odgovarajućih apsolutnih pogrešaka po komponentama, tj. $\max_i |x_i^{(k)} - x_i|$, gdje je $\mathbf{x} = (x_1, \ldots, x_n)^\top$ točno rješenje problema.

Teorem 3. (Točna procjena pogreške aproksimacije) *Neka je matrica $\mathbf{A}$ u problemu $\mathbf{Ax} = \mathbf{b}$ dijagonalno dominantna. Onda za k-tu iteraciju $\mathbf{x}^{(k)}$ Jacobijeve metode (1) vrijedi ocjena*

$$\|\mathbf{x}^{(k)} - \mathbf{x}\|_\infty \leqslant \frac{\|\mathbf{T}\|_\infty}{1 - \|\mathbf{T}\|_\infty}\|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|_\infty,$$

gdje je $\mathbf{T} = -\mathbf{D}^{-1}(\mathbf{L} + \mathbf{U})$. Ako je zadana točnost $\varepsilon > 0$, onda slijed $\mathbf{x}^{(k)}$ računamo sve dok ne bude $\|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|_\infty \leqslant (1 - \|\mathbf{T}\|_\infty)\varepsilon/\|\mathbf{T}\|_\infty$. Za taj k je $\|\mathbf{x}^{(k)} - \mathbf{x}\|_\infty \leqslant \varepsilon$, gdje je x točno rješenje problema $\mathbf{A}x = b$.

DOKAZ. Pokažimo da je $\mathbf{x}^{(k)}$ Cauchyjev slijed u $\mathbf{R}^n$. Računamo normu od $\mathbf{x}^{(k+p)} - \mathbf{x}^{(k)} = (\mathbf{x}^{(k+p)} - \mathbf{x}^{(k+p-1)}) + (\mathbf{x}^{(k+p-1)} - \mathbf{x}^{(k+p-2)}) + \ldots + (\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)})$. Radi kratkoće pišemo samo $\|\mathbf{T}\|$ umjesto $\|\mathbf{T}\|_\infty$. Rabeći nejednakost trokuta dobivamo da je

$$\begin{aligned}
\|\mathbf{x}^{(k+p)} - \mathbf{x}^{(k)}\| \leqslant{} & \|\mathbf{x}^{(k+p)} - \mathbf{x}^{(k+p-1)}\| \\
& + \|\mathbf{x}^{(k+p-1)} - \mathbf{x}^{(k+p-2)}\| + \ldots + \|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\|.
\end{aligned}$$

Zbog $\mathbf{x}^{(m+1)} = \mathbf{Tx}^{(m)} + \mathbf{b}'$ je najprije $\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\| = \|\mathbf{T}(\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)})\| \leqslant \|\mathbf{T}\|\|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|$. Za sljedeći član je $\|\mathbf{x}^{(k+2)} - \mathbf{x}^{(k+1)}\| = \|\mathbf{T}(\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\| \leqslant \|\mathbf{T}\|\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\| \leqslant \|\mathbf{T}\|^2\|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|$, i slično dalje.

Prema tome je

$$\begin{aligned}
\|\mathbf{x}^{(k+p)} - \mathbf{x}^{(k)}\| &\leqslant (\|\mathbf{T}\|^p + \ldots + \|\mathbf{T}\|^2 + \|\mathbf{T}\|)\,\|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\| \\
&= \|\mathbf{T}\|(1 + \|\mathbf{T}\| + \|\mathbf{T}\|^2 + \ldots + \|\mathbf{T}\|^{p-1})\,\|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\| \\
&\leqslant \|\mathbf{T}\|(1 + \|\mathbf{T}\| + \|\mathbf{T}\|^2 + \ldots)\,\|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\| \qquad (2) \\
&\leqslant \frac{\|\mathbf{T}\|}{1 - \|\mathbf{T}\|}\|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|.
\end{aligned}$$

Zbog $\|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\| \leqslant \|\mathbf{T}\|^{k-1}\|\mathbf{x}^{(1)} - \mathbf{x}^{(0)}\|$ je

$$\|\mathbf{x}^{(k+p)} - \mathbf{x}^{(k)}\| \leqslant \frac{\|\mathbf{T}\|}{1 - \|\mathbf{T}\|}\|\mathbf{T}\|^{k-1}\|\mathbf{x}^{(1)} - \mathbf{x}^{(0)}\|$$

Time je pokazano da je $\mathbf{x}^{(k)}$ Cauchyjev slijed vektora u $\mathbf{R}^n$, jer zbog $\|\mathbf{T}\| < 1$ desna strana teži u nulu kad $k \to \infty$, neovisno o p. Budući da je $\mathbf{R}^n$ potpun vektorski prostor (tj. Banachov prostor), onda taj slijed konvergira. Tražena ocjena u teoremu dobiva se iz (2) tako da pustimo $p \to \infty$. ☺

Primijetimo da je jednadžba $\mathbf{x} = \mathbf{Tx} + \mathbf{b}'$ ekvivalentna s $(\mathbf{I} - \mathbf{T})x = \mathbf{b}'$, tj. $\mathbf{x} = (\mathbf{I} - \mathbf{T})^{-1}\mathbf{b}'$, tj. $\mathbf{x} = (\mathbf{I} + \mathbf{T} + \mathbf{T}^2 + \ldots)\mathbf{b}'$, što smo već vidjeli iterativnim postupkom. Dobiveni matrični red konvergira jer mu je spektralni radijus manji od 1, tj. $r(\mathbf{T}) < 1$.

Matrica $\mathbf{A}$ može biti takva da pripadne Jacobijeva i Gauss-Seidelova metoda za problem $\mathbf{Ax} = \mathbf{b}$ obje konvergiraju, a da sama matrica $\mathbf{A}$ nije dijagonalno dominantna.

Primjer 10. Matrica

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

nije dijagonalno dominantna. S druge strane, za nju je $\mathbf{A} = \mathbf{L} + \mathbf{D} + \mathbf{U}$ uz

$$\mathbf{L} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad \mathbf{D} = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}, \quad \mathbf{U} = \begin{bmatrix} 0 & 3 \\ 0 & 0 \end{bmatrix}.$$

Zbog

$$\mathbf{T} = \mathbf{D}^{-1}(\mathbf{L} + \mathbf{U}) = \begin{bmatrix} 0 & 3/2 \\ 1/4 & 0 \end{bmatrix}, \quad (\mathbf{D} + \mathbf{L})^{-1}\mathbf{U} = \begin{bmatrix} 0 & 3/2 \\ 0 & -3/8 \end{bmatrix},$$

je spektar prve matrice $\{\sqrt{3/8}, -\sqrt{3/8}\}$, a druge $\{0, -3/8\}$. Prema tome je $r(\mathbf{D}^{-1}(\mathbf{L} + \mathbf{U})) = \sqrt{3/8} < 1$ i $r((\mathbf{D} + \mathbf{L})^{-1}\mathbf{U}) = 3/8 < 1$, pa Jacobijeva i Gauss-Seidelova metoda za problem $\mathbf{Ax} = \mathbf{b}$ konvergiraju za svaki $\mathbf{b}$. Primijetite da je u oba slučaja $\|\mathbf{T}\|_\infty = 3/2 > 1$.

Primjer 11. Pogledajmo primjer matrice $\mathbf{A}$ za koju Jacobijeva metoda za problem $\mathbf{Ax} = \mathbf{b}$ konvergira za svaki b, a Gauss-Seidelova metoda ne konvergira za barem jedan $\mathbf{b}$, jest

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

Za nju je $A = L + D + U$ uz

$$\mathbf{L} = \begin{bmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ 1 & 2 & 0 \end{bmatrix}, \quad \mathbf{D} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -3 \end{bmatrix}, \quad \mathbf{U} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Nakon malog računa dobivamo da matrica

$$\mathbf{T} = -\mathbf{D}^{-1}(\mathbf{L}+\mathbf{U}) = -\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1/3 \end{bmatrix} \cdot \begin{bmatrix} 0 & 0 & 1 \\ -1 & 0 & 0 \\ 1 & 2 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & -1 \\ 1 & 0 & 0 \\ 1/3 & 2/3 & 0 \end{bmatrix}$$

koja odgovara Jacobijevoj metodi ima karakteristični polinom $\kappa(\lambda) = \lambda^3 - \frac{1}{3}\lambda + \frac{2}{3}$. Vlastite vrijednosti su $\lambda_1 \approx -0.747$, $\lambda_2 \approx -0.374 - 0.867i$ i $\lambda_3 \approx -0.374 + 0.867i$, pa je spektralni radijus jednak $r(T) = \max\{|\lambda_1|, |\lambda_2|, |\lambda_3|\} < 1$. Odatle slijedi da Jacobijeva metoda konvergira.

S druge strane, za matricu

$$\mathbf{T} = -(\mathbf{D} + \mathbf{L})^{-1}\mathbf{U} = \ldots = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & -1 \\ 0 & 0 & -1 \end{bmatrix},$$

koja odgovara Gauss-Seidelovoj metodi, dobivamo vlastite vrijednosti $\lambda_{1,2} = 0$, $\lambda_3 = -1$. Prema tome je $r(\mathbf{T}) = 1$, pa postoji barem jedan $\mathbf{b}' \in \mathbf{R}^n$ za koji Gauss-Seidelov slijed iteracija za jednadžbu $\mathbf{x} = \mathbf{Tx} + \mathbf{b}'$ ne konvergira.

5.9. Matrična analiza za linearne diferencijalne jednadžbe

Linearna nehomogena (homogena) jednadžba višeg reda može se uvijek svesti na vektorsku linearnu nehomogenu (homogenu) jednadžbu prvog reda. Pogledajmo jedan primjer diferencijalne jednadžbe četvrtog reda:

$$\frac{d^4u}{dt^4} + e^t\frac{du}{dt} - tu = \sin t.$$

U njoj je nepoznanica funkcija $u = u(t)$ realne varijable t, tj. traže se funkcije $u(t)$ koje tu jednadžbu ispunjavaju za sve t na nekom zadanom intervalu u $\mathbf{R}$. Uvodimo nove funkcije od t, njih četiri:

$$x_1 := u, \quad x_2 = \frac{du}{dt}, \quad x_3 = \frac{d^2u}{dt^2}, \quad x_4 = \frac{d^3u}{dt^3}.$$

Drugim riječima, broj novih funkcija jednak je maksimalnoj derivaciji u početnoj jednadžbi. Zadnja funkcija x_4 odgovara maksimalnoj derivaciji umanjenoj za jedan. Deriviranjem novih funkcija dobivamo linearni sistem u kojem se pojavljuju samo prve derivacije:

$$\begin{cases} \frac{dx_1}{dt} &= x_2 \\ \frac{dx_2}{dt} &= x_3 \\ \frac{dx_3}{dt} &= x_4 \\ \frac{dx_4}{dt} &= tx_1 - e^t x_2 + \sin t, \end{cases}$$

jer je $\frac{dx_4}{dt} = \frac{d^4u}{dt^4} = \sin t - e^t\frac{du}{dt} + tu$. Taj sistem možemo napisati sažeto u matričnom obliku

$$\frac{d\mathbf{x}}{dt} = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{f}(t),$$

tj. kao vektorsku diferencijalnu jednadžbu *prvog* stupnja, gdje je

$$\mathbf{x} = \mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \\ x_4(t) \end{bmatrix}, \quad \mathbf{A}(t) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ t & -e^t & 0 & 0 \end{bmatrix}, \quad \mathbf{f}(t) = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \sin t \end{bmatrix}.$$

Pogledajmo sada kako riješiti nehomogenu vektorsku jednadžbu prvog stupnja $\mathbf{x}' = \mathbf{A}(t)x(t) + \mathbf{f}(t)$, gdje je $\mathbf{x}(t) \in \mathbf{R}^n$ tražena funkcija, s početnim uvjetom $\mathbf{x}(0) = \mathbf{x}_0 \in \mathbf{R}^n$. Problem nalaženja vektorske funkcije $\mathbf{x}(t)$ iz zadane diferencijalne jednadžbe prvog reda po $\mathbf{x}(t)$, i početnog uvjeta za $t = 0$, zove se **Cauchyjev problem**.

Označimo sa $\mathbf{e}_1, \dots, \mathbf{e}_n$ kanonsku bazu u $\mathbf{R}^n$, i pogledajmo pomoćni, homogen Cauchyjev problem u $\mathbf{R}^n$:

$$\begin{aligned} \frac{d\mathbf{u}_j}{dt} &= \mathbf{A}(t)\,\mathbf{u}_j(t) \\ \mathbf{u}_j(0) &= \mathbf{e}_j. \end{aligned}$$

Zašto je od interesa znati rješenja $\mathbf{u}_j(t)$? U tom se slučaju svako rješenje $\mathbf{x}(t)$ vektorskog Cauchyjeva problema

$$\begin{aligned} \mathbf{x}'(t) &= \mathbf{A}(t)\mathbf{x}(t) \\ \mathbf{x}(0) &= c \in \mathbf{R}^n, \end{aligned}$$

gdje je $\mathbf{c} = c_1\mathbf{e}_1 + \ldots + c_n\mathbf{e}_n$, može napisati u obliku $\mathbf{x}(t) = c_1\mathbf{u}_1(t) + \ldots + c_n\mathbf{u}_n(t)$. Drugim riječima, rješenje Cauchyjeva problema dobiva se kao linearna kombinacija funkcija $\mathbf{u}_j(t)$ s koeficijentima koji predstavljaju komponente početnog uvjeta.

Formirajmo sada matricu reda n,

$$\mathbf{X}(t) := [\mathbf{u}_1(t), \dots, \mathbf{u}_n(t)],$$

čiji j-ti stupac je $\mathbf{u}_j(t)$. Gornjih ukupno n *vektorskih* homogenih Cauchyjevih problema po $\mathbf{u}_j(t)$ (za $j = 1, \dots, n$) može se sažeto zapisati kao jedan jedini *matrični* homogen Cauchyjev problem po $\mathbf{X}(t)$:

$$\begin{aligned} \mathbf{X}'(t) &= \mathbf{A}(t)\,\mathbf{X}(t) \\ \mathbf{X}(0) &= \mathbf{I}. \end{aligned}$$

Primijetite da je $\mathbf{I} = [\mathbf{e}_1, \dots, \mathbf{e}_n]$. Ako imamo rješenje $\mathbf{X}(t)$ ovog matričnog, homogenog Cauchyjeva problema, onda imamo odmah i funkcije $\mathbf{u}_j(t)$: to su upravo stupci od $\mathbf{X}(t)$. Štoviše, vrijedi ova propozicija.

Propozicija 1. *Neka je $\mathbf{x}(t)$ rješenje vektorskog Cauchyjeva problema $\mathbf{x}'(t) = \mathbf{A}(t)\mathbf{x}(t)$, $\mathbf{x}(0) = \mathbf{c} \in \mathbf{R}^n$. Onda je*

$$\mathbf{x}(t) = \mathbf{X}(t)\,\mathbf{c}.$$

DOKAZ. Doista, vrijedi $\mathbf{X}(t)\,c = c_1\mathbf{u}_1(t) + \ldots + c_n\mathbf{u}_n(t)$, gdje su $\mathbf{u}_j(t)$ stupci matrice $\mathbf{X}(t)$, a c_j komponente vektora $\mathbf{c} \in \mathbf{R}^n$.

Drugi, izravni način da se to pokaže je jednostavan račun. Za $\mathbf{x}(t) = \mathbf{X}(t)\,\mathbf{c}$ vrijedi

$$\mathbf{x}'(t) = \mathbf{X}'(t)\,\mathbf{c} = \mathbf{A}(t)\mathbf{X}(t)\,\mathbf{c} = \mathbf{A}(t)\mathbf{x}(t), \quad \mathbf{x}(0) = \mathbf{X}(0)\,\mathbf{c} = \mathbf{I}\mathbf{c} = \mathbf{c}.$$

Može se pokazati da su rješenja $\mathbf{X}(t)$, odnosno $\mathbf{x}(t)$, za odgovarajuće Cauchyjeve probleme jedincata. ☺

Matrica $\mathbf{X}(t)$ ima zbog svoje važnosti poseban naziv: zove se **fundamentalna matrica** (ili **fundamentalno rješenje**) Cauchyjeva problema $\mathbf{X}' = \mathbf{A}(t)\,\mathbf{X}(t)$, $\mathbf{X}(0) = \mathbf{I}$. U uporabi je još jedan sugestivan naziv: **evolucijska matrica**.

Što je zapravo matrica $\mathbf{X}(t)$? Na to pitanje možemo odgovoriti u slučaju kad je $\mathbf{A}(t)$ konstantna matrična funkcija, tj. $\mathbf{A}$. Najprije, u jednodimenzionalnom slučaju kad je $\mathbf{A} = a \in \mathbf{R}$, rješenje problema $x' = ax$, $x(0) = 1$, je $x(t) = e^{at}$.

Propozicija 2. *Ako je $\mathbf{A}$ konstantna matrica, onda je fundamentalno rješenje odgovarajućeg homogenog Cauchyjeva problema $\mathbf{X}'(t) = \mathbf{A}\mathbf{X}(t)$, $\mathbf{X}(0) = \mathbf{I}$, jednako eksponencijalnoj funkciji matrice $\mathbf{A}t$, tj.*

$$\mathbf{X}(t) = e^{\mathbf{A}t} = I + \mathbf{A}t + \frac{\mathbf{A}^2t^2}{2!} + \ldots.$$

Drugim riječima, Cauchyjev problem $\mathbf{x}'(t) = \mathbf{A}\mathbf{x}(t)$, $\mathbf{x}(0) = \mathbf{x}_0$, ima jednoznačno određeno rješenje $\mathbf{x}(t) = e^{\mathbf{A}t}\mathbf{x}_0$ za sve $t \in \mathbf{R}$.

DOKAZ. Za matričnu funkciju $t \mapsto e^{\mathbf{A}t}$ prema pravilu za derivaciju složene funkcije imamo $(e^{\mathbf{A}t})' = e^{\mathbf{A}t}\mathbf{A} = \mathbf{A}e^{\mathbf{A}t}$ (naime, matrice $\mathbf{A}$ i $e^{\mathbf{A}t}$ komutiraju radi $e^{\mathbf{A}t} = I + \mathbf{A}t + \frac{\mathbf{A}^2t^2}{2!} + \ldots$). Prema tome je $\mathbf{x}'(t) = (e^{\mathbf{A}t}\mathbf{x}_0)' = (e^{\mathbf{A}t})'\mathbf{x}_0 = \mathbf{A}e^{\mathbf{A}t}\mathbf{x}_0 = \mathbf{A}\mathbf{x}(t)$. ☺

U slučaju kad matrica $\mathbf{A}(t)$ nije konstantna, nalaženje odgovarajuće fundamentalne matrice $\mathbf{X}(t)$ nije jednostavno. Tim problemom se ovdje nećemo baviti.

U slučaju kada rješavamo običnu nehomogenu linearnu diferencijalnu jednadžbu prvog stupnja s početnim uvjetom (tj. Caućyjev problem)

$$\begin{aligned} y'(t) &= a(t)y(t) + f(t) \\ y(0) &= y_0, \end{aligned}$$

gdje je $y = y(t)$ nepoznata realna funkcija koju tražimo, $a = a(t)$ i $f = f(t)$ zadane funkcije i y_0 zadan realan broj (početni uvjet), onda je rješenje Caućyjeva problema jednako

$$y(t) = e^{\int_0^t a(s)\,ds}\left(y_0 + \int_0^t e^{-\int_0^s a(\sigma)\,d\sigma} f(s)\,ds\right).$$

Primijetimo da je fundamentalno rješenje za taj problem, tj. rješenje problema $\mathbf{X}'(t) = a(t)\mathbf{X}(t)$, $\mathbf{X}(0) = 1$, jednako $\mathbf{X}(t) = e^{\int_0^t a(s)\,ds}$. Zanimljivo je da sličan rezultat vrijedi i za vektorsku nehomogenu linearnu diferencijalnu jednadžbu prvog stupnja. U sljedećem teoremu radi jednostavnosti pretpostavljamo da su neke vektorske funkcije neprekinute, iako je jasno da se taj uvjet može znatno oslabiti.

Teorem 3. *Neka je $\mathbf{A}(t)$ zadana neprekinuta matrična funkcija, $\mathbf{f}(t) \in \mathbf{R}^n$ zadana neprekinuta vektorska funkcija, i $\mathbf{y}_0 \in \mathbf{R}^n$ zadana početna vrijednost. Rješenje odgovarajućeg nehomogenog linearnog Cauchyjeva problema*

$$\begin{aligned} \mathbf{y}'(t) &= \mathbf{A}(t)\mathbf{y}(t) + \mathbf{f}(t) \\ \mathbf{y}(0) &= \mathbf{y}_0 \end{aligned}$$

jednako je

$$\mathbf{y}(t) = \mathbf{X}(t)\left(\mathbf{y}_0 + \int_0^t \mathbf{X}^{-1}(s)\mathbf{f}(s)\,ds\right).$$

gdje je $\mathbf{X}(t)$ fundamentalna matrica homogenog Cauchyjeva problema $\mathbf{X}' = \mathbf{A}(t)\mathbf{X}(t)$, $\mathbf{X}(0) = \mathbf{I}$. Posebno, ako je $\mathbf{A}(t)$ konstantna matrica $\mathbf{A}$, onda je rješenje odgovarajućeg Cauchyjeva problema $\mathbf{y}' = \mathbf{A}\mathbf{y}(t) + \mathbf{f}(t)$ i $\mathbf{y}(0) = \mathbf{y}_0$, jednako

$$\mathbf{y}(t) = e^{\mathbf{A}t}\mathbf{y}_0 + e^{\mathbf{A}t}\int_0^t e^{-\mathbf{A}s}\mathbf{f}(s)\,ds.$$

tj.

$$\mathbf{y}(t) = e^{\mathbf{A}t}y_0 + \int_0^t e^{\mathbf{A}(t-s)}\mathbf{f}(s)\,ds.$$

DOKAZ. Može se pokazati da je matrica $\mathbf{X}(t)$ uvijek regularna. Uvedimo novu funkciju $\mathbf{z}(t)$ sa $\mathbf{y}(t) = \mathbf{X}(t)\mathbf{z}(t)$, tj. $\mathbf{z}(t) = \mathbf{X}(t)^{-1}\mathbf{y}(t)$. Onda je

$$\mathbf{y}' = \mathbf{X}'z + \mathbf{X}z'.$$

Doista, i-ta koponenta vektora $\mathbf{y}$ je $y_i = \sum_j X_{ij}z_j$, dakle $y_i' = \sum_j X_{ij}'z_j + \sum_j X_{ij}z_j'$. Zbog $\mathbf{y}' = \mathbf{A}(t)\mathbf{y} + \mathbf{f}(t)$ vrijedi $\mathbf{X}'z + \mathbf{X}z' = \mathbf{A}(t)\mathbf{X}\mathbf{z} + f$, dakle $\mathbf{X}\mathbf{z}' = \mathbf{f}$, tj. $\mathbf{z}' = \mathbf{X}^{-1}\mathbf{f}$. Integriranjem od 0 do t dobivamo $\mathbf{z}(t) - \mathbf{z}(0) = \int_0^t \mathbf{X}^{-1}(s)\mathbf{f}(s)\,ds$. Budući da je $\mathbf{y}(t) = \mathbf{X}(t)\mathbf{z}(t)$, onda je $\mathbf{y}_0 = \mathbf{X}(0)\mathbf{z}(0) = \mathbf{I}\mathbf{z}(0) = \mathbf{z}(0)$, pa imamo $\mathbf{z}(t) = \mathbf{y}_0 + \int_0^t \mathbf{X}^{-1}(s)\mathbf{f}(s)\,ds$. Tvrdnja slijedi množenjem ove jednakosti s $\mathbf{X}(t)$ s lijeva. ☺

Primjer 12. Pogledajmo Cauchyjev problem $\mathbf{y}' = \mathbf{A}\mathbf{y} + \mathbf{f}(t)$, $\mathbf{y}(0) = \mathbf{y}_0$, u slučaju kad je

$$\mathbf{A} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \quad \mathbf{f}(t) = \begin{bmatrix} 0 \\ e^{-t} \end{bmatrix}, \quad \mathbf{y}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Drugim riječima, radi se o Cauchyjevu problemu

$$\begin{bmatrix} y_1' \\ y_2' \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} + \begin{bmatrix} 0 \\ e^{-t} \end{bmatrix}, \quad \begin{bmatrix} y_1(0) \\ y_2(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

tj.

$$\begin{aligned} y_1' &= -y_2 \\ y_2' &= y_1 + e^{-t} \\ y_1(0) &= 1 \\ y_2(0) &= 0. \end{aligned}$$

Znajući da je

$$e^{\mathbf{A}t} = \begin{bmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{bmatrix}, \quad \text{dakle} \quad e^{A(t-s)} = \begin{bmatrix} \cos(t-s) & -\sin(t-s) \\ \sin(t-s) & \cos(t-s) \end{bmatrix},$$

iz $\mathbf{y}(t) = e^{\mathbf{A}t}y_0 + \int_0^t e^{\mathbf{A}(t-s)}\mathbf{f}(s)\,ds$ dobivamo

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} + \int_0^t \begin{bmatrix} \cos(t-s) & -\sin(t-s) \\ \sin(t-s) & \cos(t-s) \end{bmatrix}\begin{bmatrix} 0 \\ e^{-s} \end{bmatrix} ds,$$

tj.

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \cos t \\ \sin t \end{bmatrix} + \int_0^t \begin{bmatrix} -\sin(t-s)\cdot e^{-s} \\ \cos(t-s)\cdot e^{-s} \end{bmatrix} ds,$$

tj.

$$\begin{aligned} y_1(t) &= \cos t - \int_0^t \sin(t-s)\cdot e^{-s}ds = \frac{1}{2}(3\cos t - \sin t - e^{-t}) \\ y_2(t) &= \sin t + \int_0^t \cos(t-s)\cdot e^{-s}ds = \frac{1}{2}(3\sin t + \cos t - e^{-t}). \end{aligned}$$

U zadnja dva integrala smo dvaput rabili parcijalnu integraciju (ili se pomognemo s nekim od simboličkih matematičkih alata).

Još jednostavniji način da prijeđemo na računanje kompleksnog integrala $F(t) := \int_0^t e^{-i(t-s)}e^{-s}ds = \int_0^t e^{-i(t-s)-s}ds$, te nakon što (izravno!) izračunamo zadnji integral, u prvom integralu samo odvojimo realni i imaginarni dio od $F(t)$ s pomoću Eulerove formule: $e^{-i(t-s)} = \cos(t-s) + i\sin(t-s)$. Točnije,

$$F(t) = e^{-it}\int_0^t e^{-s(i+1)}ds = e^{-it}\frac{e^{-s(i+1)}}{-(i+1)}\Big|_{s=0}^{s=t} = \frac{e^{-it}}{i+1}(1 - e^{-t(1+i)}),$$

pa je $\int_0^t \cos(t-s)\cdot e^{-s}ds = \frac{F}{(}t)$, a $\int_0^t \sin(t-s)\cdot e^{-s}ds = F(t)$.

Zanimljivo je primijetiti da parametarski zadana krivulja $(y_1(t), y_2(t))$, gdje je $t > 0$, predstavlja spiralu koju se svi više približava krivulji $(z_1(t), z_2(t)) := (\frac{1}{2}(3\cos t - \sin t), \frac{1}{2}(3\sin t + \cos t))$, koja je najobičnija kružnica: $z_1^2 + z_2^2 = 5/2$. Ta kružnica na koju se spirala namata kada $t \to +\infty$ zove se *granični ciklus*.

Primjer 13. Nađimo fundamentalnu matricu $\mathbf{X}(t)$ matrice

$$\mathbf{A} = \begin{bmatrix} -7 & 4 \\ -8 & 1 \end{bmatrix}.$$

Nultočke karakterističnog polinoma $\kappa(\lambda) = \det(\lambda\mathbf{I} - \mathbf{A})$ su $\lambda_{1,2} = -3 \pm 4i$. Prema tome, matrica $\mathbf{A}$ je stabilna. vlastite vrijednosti su međusobno različite, pa je matrica A slična dijagonalnoj matrici $D = \operatorname{diag}(\lambda_1, \lambda_2)$. Rješavajući jednadžbu $(\lambda\mathbf{I} - \mathbf{A})\mathbf{v} = 0$ dobivamo da su vlastiti podprostori razapeti vektorima $\mathbf{v}_1 = [1, 1+i]^\top$ i $\mathbf{v}_2 = [1, 1-i]^\top$. Time dolazimo do matrice

$$\mathbf{T} = [\mathbf{v}_1, \mathbf{v}_2] = \begin{bmatrix} 1 & 1 \\ 1+i & 1-i \end{bmatrix}$$

koja dijagonalizira matricu $\mathbf{A}$, tj. $\mathbf{T}^{-1}\mathbf{A}\mathbf{T} = \mathbf{D}$. Zbog $\mathbf{A} = \mathbf{T}\mathbf{D}\mathbf{T}^{-1}$ i

$$\mathbf{T}^{-1} = \frac{1}{2}\begin{bmatrix} 1+i & -i \\ 1-i & i \end{bmatrix}$$

dobivamo da je

$$e^{\mathbf{A}t} = \mathbf{T}e^{\mathbf{D}t}\mathbf{T}^{-1} = \begin{bmatrix} 1 & 1 \\ 1+i & 1-i \end{bmatrix}\begin{bmatrix} e^{(-3+4i)t} & 0 \\ 0 & e^{(-3-4i)t} \end{bmatrix}\frac{1}{2}\begin{bmatrix} 1+i & -i \\ 1-i & i \end{bmatrix}$$

Rabeći Eulerov identitet $e^{i\varphi} = \cos\varphi + i\sin\varphi$ dobivamo nakon malog računa:

$$e^{\mathbf{A}t} = e^{-3t}\begin{bmatrix} \cos 4t - \sin 4t & \sin 4t \\ -2\sin 4t & \cos 4t + \sin 4t \end{bmatrix}.$$

Pretpostavimo da je zadan Cauchyjev problem oblika $x'(t) = \mathbf{A}x(t)$, $\mathbf{x}(0) = \mathbf{x}_0$, gdje je $\mathbf{x}(t) = [x_1(t), x_2(t)]^\top$, matrica $\mathbf{A}$ kao na početku ovog primjera, i $x_0 = [1, 0]^\top$. Drugim riječima, promatramo dinamički sistem opisan sustavom dviju autonomnih diferencijalnih jednačaba u ravnini,

$$\begin{aligned} x_1'(t) &= -7x_1(t) + 4x_2(t) \\ x_2'(t) &= -8x_1(t) + x_2(t) \end{aligned}$$

s početnim uvjetima $x_1(0) = 1$ i $x_2(0) = 0$. Rješenje tog Cauchyjevog problema jednako je $\mathbf{x}(t) = e^{\mathbf{A}t}\mathbf{x}_0$, tj.

$$\begin{aligned} x_1(t) &= e^{-3t}(\cos 4t - \sin 4t) \\ x_2(t) &= -2e^{-3t}\sin 4t. \end{aligned}$$

Budući da $e^{-3t} \to 0$ kad $t \to +\infty$, a sinus i kosinus su cijelo vrijeme omeđeni, onda $x_1(t) \to 0$ i $x_2(t) \to 0$ kad $t \to +\infty$. Drugim riječima, trajektorija $\mathbf{x}(t) = [x_1(t), x_2(t)]^\top$ je stabilna, u smislu da teži k 0 kad $t \to +\infty$. To je u vezi s činjenicom da je matrica A stabilna, jer obje vlastite vrijednosti $\lambda_{1,2}$ imaju negativne realne dijelove, tj. leže lijevo od imaginarne osi u Gaussovoj ravnini.

Primjer 14. Zadan je sustav diferencijskih (ili rekurzivnih) jednačaba

$$\begin{aligned} u_{k+1} &= -7u_k + 4v_k \\ v_{k+1} &= -8u_k + v_k, \end{aligned}$$

gdje je $k = 0, 1, 2, \ldots$, a sljedovi realnih brojeva (u_k) i (v_k) su nepoznanice, koje interpretiramo kao diskretne trajektorije, tj. (u_k, v_k) je diskretna trajektorija u ravnini. Indeks k interpretiramo kao (diskretno) vrijeme. Pretpostavimo da je početni položaj (u_0, v_0) zadan. Problem je naći u_k i v_k u zatvorenoj formi, tj. kao eksplicine funkcije od k.

Označimo li sa $\mathbf{x}_k = (u_k, v_k)^\top$, onda gornji sistem diferencijskih jednačaba možemo zapisati kratko u vektorskom obliku:

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k, \quad k = 0, 1, 2, \ldots$$

gdje je $\mathbf{A}$ matrica iz prethodnog primjera. Uz zadani $\mathbf{x}_0 = (u_0, v_0)^\top \in \mathbf{R}^2$ vrijedi

$$\mathbf{x}_k = \mathbf{A}^k\mathbf{x}_0.$$

Doista, $\mathbf{x}_1 = \mathbf{A}\mathbf{x}_0$, $\mathbf{x}_2 = \mathbf{A}\mathbf{x}_1 = \mathbf{A}^2\mathbf{x}_0$, itd. Budući da je $\mathbf{A} = \mathbf{T}\mathbf{D}\mathbf{T}^{-1}$, onda je $\mathbf{A}^k = \mathbf{T}\mathbf{D}^k\mathbf{T}^{-1}$, pa slijedi $\mathbf{x}_k = \mathbf{A}^k\mathbf{x}_0 = \mathbf{T}\mathbf{D}^k\mathbf{T}^{-1}\mathbf{x}_0$, tj.

$$\begin{bmatrix} u_k \\ v_k \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1+i & 1-i \end{bmatrix}\begin{bmatrix} (-3+4i)^k & 0 \\ 0 & (-3-4i)^k \end{bmatrix}\frac{1}{2}\begin{bmatrix} 1+i & -i \\ 1-i & i \end{bmatrix}\begin{bmatrix} u_0 \\ v_0 \end{bmatrix}.$$

Nakon kratkog računa dobivamo

$$\begin{aligned} u_k &= \frac{1}{2}\left[\left((1-i)(-3+4i)^k + (1-i)(-3-4i)^k\right)u_0 + \left(-i(-3+4i)^k + i(-3-4i)^k\right)v_0\right] \\ v_k &= \frac{1}{2}\left[\left(2i(-3+4i)^k - 2i(-3-4i)^k\right)u_0 + \left((1-i)(-3+4i)^k + (1+i)(-3-4i)^k\right)v_0\right]. \end{aligned}$$

Primijetite da su u_k i v_k realni brojevi, iako su izraženi s pomoću kompleksnih brojeva. Budući da je $r(\mathbf{A}) = \sqrt{3^2 + 4^2} = 5 > 1$, onda neće vrijediti da $A^k \to 0$, vidi Teorem 2.3. Dakle postoji $\mathbf{x}_0 \in \mathbf{R}^2$ za koji $\mathbf{A}^k\mathbf{x}_0$ ne konvergira k 0.

5.10. Varijacijska karakterizacija vlastitih vrijednosti hermitske matrice

Pretpostavimo da je zadana kvadratna matrica $\mathbf{A}$ s kompleksnim koeficijentima. Podsjetimo se, matricu $\mathbf{A}^* = \overline{\mathbf{A}}^\top$ zovemo hermitski adjungiranom matricom od $\mathbf{A}$. Na pr.

$$\mathbf{A} = \begin{bmatrix} 1-i & i \\ 1 & 1+i \end{bmatrix} \quad \Longrightarrow \quad \mathbf{A}^* = \begin{bmatrix} 1+i & -i \\ 1 & 1-i \end{bmatrix}^\top = \begin{bmatrix} 1+i & 1 \\ -i & 1-i \end{bmatrix}.$$

Za matricu $\mathbf{A}$ kažemo da je **hermitska** ako je $\mathbf{A}^* = A$. Za matrice $\mathbf{A}$ s *realnim* koeficijentima pojam hermitske matrice se podudara s pojmom simetrične matrice.

Neka je $\mathbf{A}$ hermitska matrica reda n. Kao što znamo, onda su sve vlastite vrijednosti $\lambda_1, \dots, \lambda_n$ od $\mathbf{A}$ realne, i postoji ortonormirana baza (realnih) *vlastitih* vektora $\{e_1, \dots, e_n\}$ od $\mathbf{A}$. Možemo bez gubitka općenitosti uzeti da su vlastite vrijednosti poredane u opadajućem slijedu, tj.

$$\lambda_1 \geqslant \lambda_2 \geqslant \dots \geqslant \lambda_n.$$

Zbog $A\mathbf{e}_k = \lambda_k \mathbf{e}_k$ i $\|\mathbf{e}_k\| = 1$ je $\lambda_k = (\mathbf{A}\mathbf{e}_k \,|\, \mathbf{e}_k)$. Stoga je prirodno promatrati kvadratnu formu $f(\mathbf{x}) = (\mathbf{A}\mathbf{x} \,|\, \mathbf{x})$ definiranu za $\mathbf{x}$ takve da je $\|\mathbf{x}\| = 1$, tj. za $\mathbf{x}$ iz $(n-1)$-dimenzionalne jedinične sfere u $\mathbf{R}^n$ sa središtem u ishodištu.

Teorem 1. (Varijacijska karakterizacija vlastitih vrijednosti hermitske matrice) *Neka je $\mathbf{A}$ hermitska matrica, tj. $\mathbf{A}^* = \mathbf{A}$. (a) Onda je*

$$\lambda_1 = \max_{\substack{\mathbf{x} \in \mathbf{R}^n \\ \|\mathbf{x}\|=1}} (\mathbf{A}\mathbf{x} \,|\, \mathbf{x}),$$

i maksimum se dostiže točno kada je $\mathbf{x}$ vlastiti vektor matrice $\mathbf{A}$ koji pripada vlastitoj vrijednosti λ_1.

(b) Također,

$$\lambda_1 = \max_{\mathbf{x} \neq 0} \frac{(\mathbf{A}\mathbf{x} \,|\, \mathbf{x})}{\|\mathbf{x}\|^2}.$$

DOKAZ. (a) Odaberimo $\mathbf{x} \in \mathbf{R}^n$ takav da je $\|\mathbf{x}\| = 1$. Budući da vektori $\mathbf{e}_k$ čine ortonormiranu bazu u $\mathbf{R}^n$, vrijedi

$$\mathbf{x} = (\mathbf{x} \,|\, \mathbf{e}_1)\mathbf{e}_1 + \dots + (\mathbf{x} \,|\, \mathbf{e}_n)\mathbf{e}_n,$$

i $|(\mathbf{x} \,|\, \mathbf{e}_1)|^2 + \dots + |(\mathbf{x} \,|\, \mathbf{e}_n)|^2 = 1$ zbog $(\mathbf{x} \,|\, x) = 1$. Iz $\mathbf{A}\mathbf{e}_k = \lambda_k \mathbf{e}_k$ slijedi $\mathbf{A}\mathbf{x} = \lambda_1 (\mathbf{x} \,|\, \mathbf{e}_1)\mathbf{e}_1 + \dots + \lambda_n (\mathbf{x} \,|\, \mathbf{e}_n)\mathbf{e}_n$, dakle

$$(\mathbf{A}\mathbf{x} \,|\, \mathbf{x}) = \lambda_1 |(\mathbf{x} \,|\, \mathbf{e}_1)|^2 + \dots + \lambda_n |(\mathbf{x} \,|\, \mathbf{e}_n)|^2 \leqslant \lambda_1 |(\mathbf{x} \,|\, \mathbf{e}_1)|^2 + \dots + \lambda_1 |(\mathbf{x} \,|\, \mathbf{e}_n)|^2 = \lambda_1.$$

Time je dokazano da je $\lambda_1 \geqslant (\mathbf{A}\mathbf{x} \,|\, \mathbf{x})$ za sve $\mathbf{x} \in \mathbf{R}^n$ takve da je $\|\mathbf{x}\| = 1$. Još treba samo vidjeti da postoji neki $\mathbf{x}$ čija norma je jedan, za koji vrijedi $\lambda_1 = (\mathbf{A}\mathbf{x} \,|\, \mathbf{x})$. Dovoljno je uzeti $\mathbf{x} = \mathbf{e}_1$, jer onda imamo $(\mathbf{A}\mathbf{e}_1 \,|\, \mathbf{e}_1) = (\lambda_1 \mathbf{e}_1 \,|\, \mathbf{e}_1) = \lambda_1$.

(b) Tvrdnja slijedi iz (a), jer u izrazu

$$\frac{(\mathbf{A}\mathbf{x} \,|\, \mathbf{x})}{\|\mathbf{x}\|^2} = \left(\mathbf{A} \frac{\mathbf{x}}{\|\mathbf{x}\|} \,\Big|\, \frac{\mathbf{x}}{\|\mathbf{x}\|} \right)$$

je vektor $\mathbf{x}/\|\mathbf{x}\|$ jedinični. ☺

Razlomak

$$R(\mathbf{x}) := \frac{(\mathbf{A}\mathbf{x} \,|\, \mathbf{x})}{\|\mathbf{x}\|^2}$$

zove se **Rayleighov kvocijent**.

Ponekad je zgodno Teorem 1(b) sročiti u ovom obliku.

Korolar 2. *Ako je $\mathbf{A}$ hermitska matrica i λ_1 njena najveća vlastita vrijednost, onda za sve $x \in \mathbf{R}^n$ vrijedi*

$$(\mathbf{A}\mathbf{x} \,|\, \mathbf{x}) \leqslant \lambda_1 \|\mathbf{x}\|^2.$$

Konstanta $C = \lambda_1$ je najmanja konstanta za koju vrijedi $(\mathbf{A}\mathbf{x} \,|\, \mathbf{x}) \leqslant C\|\mathbf{x}\|^2$ za sve x.

Pozitivno semidefinitne i pozitivno definitne matrice imaju sličnu ulogu kao nenenegativni ($\geqslant 0$) i pozitivni brojevi (> 0) među svim realnim brojevima. Pokazuje se da za svaku pozitivno semidefinitnu matricu $\mathbf{A}$ postoji jednoznačno određena pozitivno semidefinitna matrica $\mathbf{B}$ takva da je $\mathbf{B}^2 = A$. Matrica $\mathbf{B}$ zove se **drugi korijen** iz $\mathbf{A}$, i označava sa $\mathbf{B} = \sqrt{A}$ ili $\mathbf{A}^{1/2}$.

Korolar 3. *Neka je $\mathbf{A}$ hermitska matrica.*

a) *$\mathbf{A}$ je pozitivno semidefinitna onda i samo onda ako je $\lambda_i \geqslant 0$ za sve i. U tom slučaju je spektralna norma jednaka najvećoj vlastitoj vrijednosti od $\mathbf{A}$, tj.*

$$\|\mathbf{A}\|_2 = \lambda_1,$$

gdje λ_1 najveća vlastita vrijednost od $\mathbf{A}$: $\lambda_1 \geqslant \lambda_2 \geqslant \dots \geqslant \lambda_n (\geqslant 0)$.

b) *$\mathbf{A}$ je pozitivno definitna onda i samo onda ako je $\lambda_i > 0$ za sve i. Ako je λ_n najmanja vlastita vrijednost od $\mathbf{A}$, onda za sve $x \in \mathbf{R}^n$ vrijedi*

$$(\mathbf{A}\mathbf{x} \,|\, \mathbf{x}) \geqslant \lambda_n \|\mathbf{x}\|^2.$$

Konstanta $C = \lambda_n$ je najveća konstanta sa svojstvom da za sve $\mathbf{x}$ vrijedi $(\mathbf{A}\mathbf{x} \,|\, \mathbf{x}) \geqslant C\|\mathbf{x}\|^2$, tj.

$$\lambda_n = \min_{\mathbf{x} \neq 0} \frac{(\mathbf{A}\mathbf{x} \,|\, \mathbf{x})}{\|\mathbf{x}\|^2}.$$

DOKAZ. a) Ako je $\mathbf{A} \geqslant 0$, onda je $\lambda_i = (\mathbf{A}\mathbf{e}_i \,|\, \mathbf{e}_i) \geqslant 0$ za sve i. Obratno, ako je $\lambda_i \geqslant 0$ za sve i, onda iz rastava $\mathbf{x} = x_1 \mathbf{e}_1 + \dots + x_n \mathbf{e}_n$ po volji odabranog vektora $\mathbf{x}$ dobivamo

$$(\mathbf{A}\mathbf{x} \,|\, \mathbf{x}) = (\sum_i \lambda_i x_i \mathbf{e}_i \,|\, \sum_i x_i \mathbf{e}_i) = \sum_i \lambda_i |x_i|^2 \geqslant 0.$$

Da bismo dokazali $\lambda_1 = \|A\|$, odaberimo bilo koji jedinični vektor $\mathbf{x}$. Onda je

$$(A\mathbf{x} \,|\, \mathbf{x}) \leqslant \|\mathbf{A}x\| \cdot \|\mathbf{x}\| \leqslant \|\mathbf{A}\| \cdot \|\mathbf{x}\|^2 = \|\mathbf{A}\|,$$

gdje smo rabili nejednakost CSB. Uzimajući u gornjoj nejednakosti maksimum po svim jediničnim vektorima x dobivamo prema Teoremu 1 da je $\|\mathbf{A}\| \leqslant \lambda_1$. Da bi dokazali obratnu nejednakost, za jedinični vektor x gledamo

$$\|\mathbf{A}x\|^2 = \|\mathbf{A}(\sum_i x_i e_i)\|^2 = \|\sum_i \lambda_i x_i e_i\|^2 = \sum_i \lambda_i^2 |x_i|^2 \leqslant \sum_i \lambda_1^2 |x_i|^2 = \lambda_1^2,$$

gdje smo rabili da je λ_1 najveća vlastita vrijednost. Prema tome je $\|\mathbf{A}x\| \leqslant \lambda_1$, pa uzimajući maksimum po svim jediničnim vektorima $\mathbf{x}$ dobivamo $\|\mathbf{A}\| \leqslant \lambda_1$.

b) Tvrdnja slijedi odmah iz dokaza provedenog u a). ☺

Pokazuje se da je hermitska matrica $\mathbf{A} = (a_{ij})$ pozitivno definitna onda i samo onda ako su sve glavne minore od $\mathbf{A}$ pozitivne, tj.,

$$a_{11} > 0, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \quad \dots \quad , \quad \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} > 0.$$

U primjenama (na pr. u teoriji antena) od koristi je sljedeći rezultat o tzv. **generaliziranim vlastitim vrijednostima** λ koja rješavaju problem $\mathbf{A}x = \lambda \mathbf{B}x$, $x \neq 0$ (zapravo, ovdje su nepoznanice parovi λ i $x \neq 0$). On predstavlja poopćenje Teorema 1. Naime, tvrdnja se svodi na Teorem 1 ako je matrica $\mathbf{B}$ jedinična. U tom slučaju se nalaženje generalizirane vlastite vrijednosti λ svodi na problem nalaženja klasične vlastite vrijednosti matrice $\mathbf{A}$: $\mathbf{A}x = \lambda x$, $x \neq 0$.

Teorem 4. *Neka su $\mathbf{A}$ i $\mathbf{B}$ hermitske matrice i $\mathbf{B}$ pozitivno definitna (tj. $(\mathbf{B}\mathbf{x} \,|\, \mathbf{x}) > 0$ za sve $x \neq 0$). Neka su $\lambda_1 \geqslant \dots \geqslant \lambda_n$ rješenja generalizirane karakteristične jednadžbe*

$$\det(\lambda \mathbf{B} - \mathbf{A}) = 0.$$

Onda postoje vektori $\mathbf{e}_1, \dots, \mathbf{e}_n \in \mathbf{R}^n$ takvi da je

$$\mathbf{A}e_i = \lambda_i \mathbf{B}e_i, \quad (Be_i \,|\, e_j) = \delta_{ij},$$

i vrijedi

$$\lambda_1 = \max_{\substack{x \in \mathbf{R}^n \\ x \neq 0}} \frac{(\mathbf{A}\mathbf{x} \,|\, \mathbf{x})}{(\mathbf{B}\mathbf{x} \,|\, \mathbf{x})}.$$

Maksimum se dostiže za $\mathbf{x} = \mathbf{e}_1$.

SKICA DOKAZA. (a) Pokažimo da postoji regularna matrica $\mathbf{T}$ takva da je

$$\mathbf{T}^*\mathbf{A}\mathbf{T} = \mathbf{D} := \operatorname{diag}(\lambda_1, \dots, \lambda_n), \quad \mathbf{T}^*\mathbf{B}\mathbf{T} = \mathbf{I}.$$

Doista, neka je $\mathbf{S}$ ortogonalna matrica koja dijagonalizira $\mathbf{B}$, tj. $\mathbf{S}^*\mathbf{B}\mathbf{S} = \mathbf{D}_1 := \operatorname{diag}(\mu_1, \dots, \mu_n)$. Pritom $\mathbf{A}$ prelazi u $\mathbf{A}_1 := \mathbf{S}^*\mathbf{A}\mathbf{S}$. Matrica $\mathbf{A}_1$ je također hermitska jer je $\mathbf{A}_1^* = (\mathbf{S}^*\mathbf{A}\mathbf{S})^* = \mathbf{S}^*\mathbf{A}\mathbf{S}^{**} = \mathbf{S}^*\mathbf{A}\mathbf{S} = \mathbf{A}_1$. Budući da su zbog pozitivne defnitnosti od $\mathbf{B}$ svi $\mu_i > 0$, matrica $\mathbf{D}_2 := \operatorname{diag}(1/\sqrt{\mu_1}, \dots, 1/\sqrt{\mu_n}) = D_2$ je dobro definirana. Ona prevodi $\mathbf{D}_1$ u jediničnu matricu (umnošku triju dijagonalnih matrica odgovara umnožak dijagonalnih elemenata):

$$\mathbf{D}_2^*(\mathbf{S}^*\mathbf{B}\mathbf{S})\mathbf{D}_2 = \mathbf{D}_2\mathbf{D}_1\mathbf{D}_2 = \mathbf{I}.$$

Matrica $\mathbf{A}_1$ prelazi u novu hermitsku matricu

$$\mathbf{D}_2^*(\mathbf{S}^*\mathbf{A}\mathbf{S})\mathbf{D}_2 = \mathbf{D}_2^*\mathbf{A}_1\mathbf{D}_2 := \mathbf{A}_2.$$

Neka je $\mathbf{S}_1$ ortogonalna matrica koja prevodi hermitsku matricu $\mathbf{A}_2$ u dijagonalnu:

$$\mathbf{S}_1^*(\mathbf{D}_2^*(\mathbf{S}^*\mathbf{A}\mathbf{S})\mathbf{D}_2)\mathbf{S}_1 = \mathbf{S}_1^*\mathbf{A}_2\mathbf{S}_1 = D := \operatorname{diag}(\lambda_1, \dots, \lambda_n).$$

Onda za $\mathbf{T} = \mathbf{S}\mathbf{D}_2\mathbf{S}_1$ imamo $\mathbf{T}^*\mathbf{A}\mathbf{T} = \mathbf{D}$ i

$$\mathbf{T}^*\mathbf{B}\mathbf{T} = \mathbf{S}_1^*(\mathbf{D}_2^*(\mathbf{S}^*\mathbf{B}\mathbf{S})\mathbf{D}_2)\mathbf{S}_1 = \mathbf{S}_1^*\mathbf{I}\mathbf{S}_1 = \mathbf{I}.$$

(b) Budući da je $(\mathbf{T}^*)^{-1} = \mathbf{B}\mathbf{T}$, onda je $\mathbf{A}\mathbf{T} = \mathbf{T}^{*-1}\mathbf{D} = \mathbf{B}\mathbf{T}\mathbf{D}$. Neka je $\mathbf{T} = [\mathbf{e}_1, \dots, \mathbf{e}_n]$, tj. $\mathbf{e}_i$ su stupci od $\mathbf{T}$. Budući da je $\mathbf{T}\mathbf{D} = [\lambda_1 e_1, \dots, \lambda_n e_n]$ (provjeri!), onda iz $\mathbf{A}\mathbf{T} = [\mathbf{A}\mathbf{e}_1, \dots, \mathbf{A}\mathbf{e}_n]$ i $\mathbf{B}\mathbf{T}\mathbf{D} = \mathbf{B}[\lambda_1 \mathbf{e}_1, \dots, \lambda_n \mathbf{e}_n] = [\lambda_1 \mathbf{B}\mathbf{e}_1, \dots, \lambda_n \mathbf{B}\mathbf{e}_n]$ dobivamo $\mathbf{A}\mathbf{e}_i = \lambda_i \mathbf{B}\mathbf{e}_i$. Vektori $\mathbf{e}_i$ su svi $\neq 0$ jer je $\mathbf{T} = [e_1, \dots, e_n]$ regularna matrica.

(c) Vrijednosti λ_i dobivaju se iz jednadžbe $\det(\lambda \mathbf{I} - \mathbf{D}) = 0$, tj. $\det(\lambda \mathbf{T}^*\mathbf{B}\mathbf{T} - \mathbf{T}^*\mathbf{A}\mathbf{T}) = 0$, tj. $\det(\mathbf{T}^*(\lambda \mathbf{B} - \mathbf{A})\mathbf{T}) = 0$, tj. prema Binet-Cauchyjevu teoremu, iz jednadžbe $\det(\lambda \mathbf{B} - \mathbf{A}) = 0$.

(d) Iz svojstva $\mathbf{T}^*\mathbf{B}\mathbf{T} = \mathbf{I}$, rabeći definiciju množenja matrica, slijedi da na mjestu (i,j) vrijedi jednakost $\mathbf{e}_i^*\mathbf{B}\mathbf{e}_j = \delta_{ij}$, tj. $(\mathbf{e}_i \,|\, \mathbf{B}\mathbf{e}_j) = \delta_{ij}$, tj. vektori e_i su ortonormirani s obzirom na skalarni produkt $(\mathbf{x} \,|\, \mathbf{y})_b := (\mathbf{x} \,|\, \mathbf{B}\mathbf{y}) = (\mathbf{B}\mathbf{x} \,|\, \mathbf{y})$. Lako se provjeri da je to doista skalarni produkt, na pr. $(\mathbf{x} \,|\, \mathbf{x})_b = (\mathbf{B}\mathbf{x} \,|\, \mathbf{x}) > 0$ za $x \neq 0$. Primijetimo da je $(\mathbf{A}\varepsilon_i \,|\, \mathbf{e}_i) = (\lambda_i \mathbf{B}\mathbf{e}_i \,|\, e_i) = \lambda_i$. Preostali dio dokaza je sličan kao dokaz Teorema 1. ☺

Spomenimo na kraju bez dokaza sljedeći zanimljiv teorem koji daje varijacijski opis vlastitih vrijednosti hermitske matrice $\mathbf{A}$.

Teorem 6. (Courantov princip minimaksa) *Neka je $\mathbf{A}$ hermitska matrica reda n i neka su $\lambda_1 \geqslant \dots \geqslant \lambda_n$ sve vlastite vrijednosti od $\mathbf{A}$. Onda za sve $k < n$ vrijedi*

$$\lambda_{k+1} = \min_{x_1, \dots, x_k \in \mathbf{R}^n} \max_{\substack{\|\mathbf{x}\|=1 \\ \mathbf{x} \in L(\mathbf{x}_1, \dots, \mathbf{x}_k)^\perp}} (\mathbf{A}\mathbf{x} \,|\, \mathbf{x}).$$

Na pr. λ_2 dobiva se tako da za bilo koji $\mathbf{x}_1 \in \mathbf{R}^n$, $\mathbf{x}_1 \neq 0$ gledamo maksimum kvadratne forme $(\mathbf{A}\mathbf{x}_1 \,|\, \mathbf{x}_1)$ na jediničnoj sferi u hiperravnini $L(\mathbf{x}_1)^\perp$, i zatim minimum tih brojeva (maksimuma) po svim $\mathbf{x}_1$.

5.11. Fréchetova derivacija

Pretpostavimo da su zadani normirani vektorski prostori X i Y. Radi jednostavnosti pretpostavit ćemo da su oba konačno-dimenzionalni. Za funkciju $F : X \to Y$, općenito nelinearnu, kažemo da je **diferencijabilna u točki** $x \in X$ ako postoji linearni operator $A \in \mathcal{L}(X,Y)$ takav da je

$$F(x+h) = F(x) + Ah + o(h),$$

gdje $\frac{o(h)}{\|h\|} \to 0$ u Y kada $h \to 0$ u prostoru X. Drugim riječima, funkcija $F : X \to Y$ je diferencijabilna u točki x ako postoji linearni operator $A \in \mathcal{L}(X,Y)$ takav da vrijedi

$$\lim_{h \to 0} \frac{\|F(x+h) - F(x) - Ah\|}{\|h\|} = 0.$$

Lako se vidi da je gornjom jednakošću operator A određen jednoznačno. Označava se sa $A = F'(x) \in \mathcal{L}(X,Y)$, i zove se **Fréchetova derivacija** od F u točki x. Time smo dobili novu funkciju $F' : X \to \mathcal{L}(X,Y)$. Ako je funkcija F' neprekinuta (prostor $\mathcal{L}(X,Y)$ je također normiran), onda kažemo da je funkcija F Fréchet diferencijabilna, ili klase C^1, te u tom slučaju pišemo $F \in C^1(X,Y)$. Ovdje smo sa $C^1(X,Y)$ označili skup svih funkcija F iz X u Y koje su klase C^1. U literaturi se često umjesto $F'(x)\,h$ rabi i oznaka $dF(x)\,h$ (diferencijal od F u točki x).

Primjer 15. Neka je $\mathbf{A}$ zadana matrica tipa $m \times n$ i $F : \mathbf{R}^n \to \mathbf{R}^m$ linearni operator zadan sa $F(\mathbf{x}) = A\mathbf{x}$. Onda je $F'(\mathbf{x}) = \mathbf{A}$ za sve $\mathbf{x}$. Dakle $F'(\mathbf{x})$ je isti linearni operator za sve $\mathbf{x}$. Doista, za učvršćeni vektor $\mathbf{x}$ je $F(\mathbf{x}+\mathbf{h}) - F(\mathbf{x}) = \mathbf{A}(\mathbf{x}+\mathbf{h}) - \mathbf{A}\mathbf{x} = \mathbf{A}\mathbf{h}$, pa vidimo da je $F'(\mathbf{x})\mathbf{h} = \mathbf{A}\mathbf{h}$ za sve $\mathbf{h} \in \mathbf{R}^n$ (dakle $F'(\mathbf{x}) = \mathbf{A}$) i $o(\mathbf{h}) = 0$. Sjetimo se iz Matematike 1 da za učvršćeni $a \in \mathbf{R}$ i funkciju $F(x) = ax$, $x \in \mathbf{R}$, vrijedi $F'(x) = a$, što ovdje odgovara slučaju $m = n = 1$.

Nakon što smo dobili novu funkciju $F' : X \to \mathcal{L}(X,Y)$, općenito također nelinearnu, na analogan način možemo definirati i njenu derivaciju na X, tj. F'', jer je prostor $\mathcal{L}(X,Y)$ također normiran s operatorskom normom. Sada će prema prethodnom biti definiran ovakav linearni operator iz X u $\mathcal{L}(X,Y)$:

$$F''(x) \in \mathcal{L}(X, \mathcal{L}(X,Y))$$

Drugim riječima, u točki $x \in X$ za bilo koje vektore $h, k \in X$ vrijedi da je $F''(x)h \in \mathcal{L}(X,Y)$, dakle $[F''(x)h]k \in Y$. Ovaj izraz često kraće zapisujemo kao

$$[F''(x)h]k = F''(x)\,(h,k) = F''(x)\,hk \in Y.$$

Drugim riječima, $F''(x)$ shvaćamo kao *bilinearnu funkciju* dviju varijabla iz X. Bilinearnost znači da je funkcija $h \mapsto F''(x)\,hk$ linearna iz X u Y za bilo koji učvršćeni k (tj. $F''(x)\,(\lambda_1 h_1 + \lambda_2 h_2)k = \lambda_1 F''(x)\,h_1 k + \lambda_2 F''(x)\,h_2 k$), a funkcija $k \mapsto F''(x)\,hk$ je linearna za učvršćeni h (tj. $F''(x)\,h(\lambda_1 k_1 + \lambda_2 k_2) = \lambda_1 F''(x)\,hk_1 + \lambda_2 F''(x)\,hk_2$).

Primijetite da funkcija $(h,k) \mapsto [F''(x)h]k$ općenito nije linearna iz $X \times X$ u Y. Vektorski prostor bilinearnih operatora iz $X \times X$ u Y označavamo sa $\mathcal{L}^2(X \times X, Y)$, pa je $F''(x) \in \mathcal{L}^2(X \times X, Y)$.

Primjer 16. U uvjetima prethodnog primjera je $F'(x)k = \mathbf{A}\mathbf{k}$ (jednakost u $\mathbf{R}^m$), pa je $F'(\mathbf{x}+\mathbf{h})\mathbf{k} - F'(\mathbf{x})\mathbf{k} = \mathbf{A}\mathbf{k} - \mathbf{A}\mathbf{k} = 0$, za sve $\mathbf{k} \in \mathbf{R}^n$, dakle $F'(\mathbf{x}+\mathbf{h}) - F'(\mathbf{x}) = 0$ (jednakost u prostoru matrica tipa $m \times n$), dakle $F''(\mathbf{x})\mathbf{h} = 0$, tj. $F''(\mathbf{x}) = 0$ za svaki $\mathbf{x}$ ($F''(\mathbf{x})$ je linearni operator iz $\mathbf{R}^n$ u prostor $m \times n$ matrica, $F''(\mathbf{x}) \in \mathcal{L}(\mathbf{R}^n, \mathcal{L}(\mathbf{R}^n, eR^m))$).

Jednostavna bilinearna funkcija $B : \mathbf{R} \times \mathbf{R} \to \mathbf{R}$ je na pr. $B(h,k) = 3hk$. U slučaju $B : \mathbf{R}^2 \times \mathbf{R}^2 \to \mathbf{R}^2$ to je na pr. $B(\mathbf{h},\mathbf{k}) = (3h_1k_1 + h_2k_1 + h_2k_2, 2h_1k_1 - h_1k_2 + h_2k_1 - h_2k_2)^\top$, gdje je $\mathbf{h} = (h_1, h_2)$ i $\mathbf{k} = (k_1, k_2)$.

Na sličan način se dobiva i treća derivacije od F. To je funkcija $F'''(x) \in \mathcal{L}(X, \mathcal{L}(X, \mathcal{L}(X,Y)))$ koja je trilinearna: $X \times X \times X \ni (h,k,l) \mapsto F'''(x)\,hkl \in Y$, tj. $F'''(x) \in \mathcal{L}^3(X \times X \times X, Y)$, gdje je $\mathcal{L}^3(X \times X \times X, Y)$ prostor svih trilinearnih operatora iz $X \times X \times X$ u Y, itd.

Fréchetova derivacija kompozicije vektorskih funkcija (općenito nelinearnih), računa se kao kompozicija njihovih derivacija. Te su derivacije linearni operatori. Drugim riječima, ako su $F : X \to Y$ i $G : Y \to Z$ preslikavanja između triju vektorskih prostora, onda za $G \circ F : X \to Z$ vrijedi pravilo deriviranja kompozicije:

$$(G \circ F)'(x) = G'(F(x)) \circ F'(x).$$

Ovdje je $(G \circ F)'(x) \in \mathcal{L}(X,Z)$, $G'(F(x)) \in \mathcal{L}(Y,Z)$ i $F'(x) \in \mathcal{L}(X,Y)$.

Primjer 17. U slučaju realne funkcije realne varijable $F : \mathbf{R} \to \mathbf{R}$ je $F'(x)$ uobičajena derivacija.

Za $F : \mathbf{R}^2 \to \mathbf{R}$ je $F'(\mathbf{x})$ linearni operator iz $\mathbf{R}^2$ u $\mathbf{R}$ (za svaki $\mathbf{x}$ po jedan), definiran sa

$$F'(\mathbf{x})\mathbf{h} = \frac{\partial f(\mathbf{x})}{\partial x_1} h_1 + \frac{\partial f(\mathbf{x})}{\partial x_2} h_2,$$

tj. *totalni diferencijal* funkcije F, gdje je $\mathbf{h} = (h_1, h_2) \in \mathbf{R}^2$. Parcijalne derivacije izračunavamo u točki $\mathbf{x} = (x_1, x_2)$.

Ovdje je

$$F''(\mathbf{x}) = \begin{bmatrix} f''_{x_1x_1} & f''_{x_1x_2} \\ f''_{x_2x_1} & f''_{x_2x_2} \end{bmatrix} \in \mathcal{L}(\mathbf{R}^2, \mathcal{L}(\mathbf{R}^2, \mathbf{R})),$$

jer taj prostor linearnih operator možemo poistovjetiti s prostorom kvadratnih matrica M_{22}. Ova se matrica zove **Hesseova matrica** ili **Hessijan** preslikavanja F. Prema tome je

$$F''(x)\,hk = f''_{x_1x_1} h_1k_1 + f''_{x_1x_2}(h_1k_2 + h_2k_1) + f''_{x_2x_2} k_1k_2.$$

Posebno,

$$F''(\mathbf{x})\,\mathbf{hh} = f''_{x_1x_1} h_1^2 + 2f''_{x_1x_2} h_1h_2 + f''_{x_2x_2} h_2^2,$$

a to je upravo poznati drugi diferencijal $d^2F(\mathbf{x})$ realne funkcije dviju realnih varijabla F, izračunat u točki $\mathbf{x}$. Rezultat je kvadratni polinom (točnije, kvadratna forma) s varijablama h_1 i h_2.

Primjer 18. Zadana je vektorska funkcija vektorske varijable $F : \mathbf{R}^2 \to \mathbf{R}^2$ sa $F = (f,g)^\top$, tj. $F(x_1,x_2) = (f(x_1,x_2), g(x_1,x_2))^\top$, uz zadane realne funkcije dviju varijabla f i g. Onda za $\mathbf{h} = (h_1,h_2)^\top$, zbog $f(x_1+h_1, x_2+h_2) - f(x_1,x_2) = f'_{x_1}h_1 + f'_{x_2}h_2 + o_1(h)$, i slično za g, vrijedi

$$F(\mathbf{x}+\mathbf{h}) - F(\mathbf{x}) = (f'_{x_1}h_1 + f'_{x_2}h_2 + o_1(h), g'_{x_1}h_1 + g'_{x_2}h_2 + o_2(\mathbf{h}))^\top,$$

gdje se parcijalne derivacije izračunavaju u točki $\mathbf{x}$. Dakle, $F(\mathbf{x}+\mathbf{h}) = F(\mathbf{x}) + \mathbf{A}\mathbf{h} + O(\mathbf{h})$, gdje je

$$\mathbf{A} = F'(\mathbf{x}) = \begin{bmatrix} f'_{x_1} & f'_{x_2} \\ g'_{x_1} & g'_{x_2} \end{bmatrix},$$

tzv. **Jacobijeva matrica** preslikavanja F (ili Jacobijan od F) i $o(\mathbf{h}) = (o_1(h), o_2(h))$.

Primjer 19. Za $\mathbf{h} = (h_1, h_2)$ je

$$\begin{aligned} F'(\mathbf{x}{+}h) - F'(\mathbf{x}) &= \begin{bmatrix} f'_{x_1}(x{+}h) & f'_{x_2}(x{+}h) \\ g'_{x_1}(x{+}h) & g'_{x_2}(x{+}h) \end{bmatrix} - \begin{bmatrix} f'_{x_1}(x) & f'_{x_2}(x) \\ g'_{x_1}(x) & g'_{x_2}(x) \end{bmatrix} \\ &= \begin{bmatrix} f''_{x_1x_1}h_1{+}f''_{x_1x_2}h_2{+}o_{11}(h) & f''_{x_2x_1}h_1{+}f''_{x_2x_2}h_2{+}o_{12}(\mathbf{h}) \\ g''_{x_1x_1}h_1{+}g''_{x_1x_2}h_2{+}o_{21}(h) & g''_{x_2x_1}h_1{+}g''_{x_1x_2}h_2{+}o_{22}(\mathbf{h}) \end{bmatrix} \end{aligned}$$

tj.

$$F'(\mathbf{x}+\mathbf{h}) = F'(\mathbf{x}) + \begin{bmatrix} f''_{x_1x_1} & f''_{x_2x_1} \\ g''_{x_1x_1} & g''_{x_2x_1} \end{bmatrix} h_1 + \begin{bmatrix} f''_{x_1x_2} & f''_{x_2x_2} \\ g''_{x_1x_2} & g''_{x_1x_2} \end{bmatrix} h_2 + o(\mathbf{h}),$$

gdje je $o(\mathbf{h}) = (o_{ij}(\mathbf{h}))$ matrica takva da $\frac{o(\mathbf{h})}{\|\mathbf{h}\|}$ konvergira prema nul-matrici kada $\mathbf{h} \to 0$. Odavde čitamo da je

$$F''(\mathbf{x})\,h = \begin{bmatrix} f''_{x_1x_1} & f''_{x_2x_1} \\ g''_{x_1x_1} & g''_{x_2x_1} \end{bmatrix} h_1 + \begin{bmatrix} f''_{x_1x_2} & f''_{x_2x_2} \\ g''_{x_1x_2} & g''_{x_1x_2} \end{bmatrix} h_2 \in \mathcal{L}(\mathbf{R}^2, \mathbf{R}^2).$$

Drugim riječima, operator $F''(\mathbf{x}) \in L(\mathbf{R}^2, \mathcal{L}(\mathbf{R}^2, \mathbf{R}^2))$ možemo shvatiti kao par dviju matrica reda 2, tj. kao kvadratnu matricu tipa 4×2:

$$F''(\mathbf{x}) = \begin{bmatrix} f''_{x_1x_1} & f''_{x_2x_1} \\ g''_{x_1x_1} & g''_{x_2x_1} \\ f''_{x_1x_2} & f''_{x_2x_2} \\ g''_{x_1x_2} & g''_{x_1x_2} \end{bmatrix}. \tag{1}$$

Primijetite da je prostor $L(\mathbf{R}^2, \mathcal{L}(\mathbf{R}^2, \mathbf{R}^2))$ izomorfan s prostorom matrica $M_{4,2}$, jer je $L(\mathbf{R}^2, \mathbf{R}^2))$ izomorfan s $\mathbf{R}^4$.

Smisao druge derivacije (ili tzv. drugog diferencijala) je u ovom slučaju ovakav:

$$F''(\mathbf{x})\,\mathbf{hk} = \begin{bmatrix} f''_{x_1x_1} & f''_{x_2x_1} \\ g''_{x_1x_1} & g''_{x_2x_1} \end{bmatrix} h_1 \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} + \begin{bmatrix} f''_{x_1x_2} & f''_{x_2x_1} \\ g''_{x_1x_2} & g''_{x_2x_2} \end{bmatrix} h_2 \begin{bmatrix} k_1 \\ k_2 \end{bmatrix},$$

tj.

$$F''(x)\,\mathbf{hk} = \begin{bmatrix} f''_{x_1x_1}h_1k_1 + f''_{x_1x_2}(h_1k_2 + h_2k_1) + f''_{x_2x_2}h_2k_2 \\ g''_{x_1x_1}h_1k_1 + g''_{x_1x_2}(h_1k_2 + h_2k_1) + g''_{x_2x_2}h_2k_2 \end{bmatrix}.$$

Posebno,

$$F''(\mathbf{x})\,\mathbf{hh} = \begin{bmatrix} f''_{x_1x_1}h_1^2 + 2f''_{x_1x_2}h_1h_2 + f''_{x_2x_2}h_2^2 \\ g''_{x_1x_1}h_1^2 + 2g''_{x_1x_2}h_1h_2 + g''_{x_2x_2}h_2^2 \end{bmatrix} \in \mathbf{R}^2.$$

Prepoznajemo da su komponente od $F''(x)\,\mathbf{hh}$ upravo drugi diferencijali funkcija f i g.

Pokazuje se da je za vektorske funkcije $F : \mathbf{R}^2 \to \mathbf{R}^2$, kao u ovom primjeru, moguć prikaz u obliku Taylorova razvoja. Na pr. za funkciju F klase C^1 vrijedi

$$F(\mathbf{x}+\mathbf{h}) = F(\mathbf{x}) + F'(\mathbf{x})\,\mathbf{h} + o(\mathbf{h}),$$

za $\mathbf{h} \to 0$. Za funkciju klase C^2 vrijedi

$$F(\mathbf{x}+\mathbf{h}) = F(\mathbf{x}) + F'(\mathbf{x})\,\mathbf{h} + \frac{1}{2!}F''(\mathbf{x})\,\mathbf{hh} + o(\|\mathbf{h}\|^2),$$

gdje je $o(\|\mathbf{h}\|^2)$ vektor koji trne brže od $\|\mathbf{h}\|^2$, točnije, takav da $\frac{o(\|\mathbf{h}\|^2)}{\|\mathbf{h}\|^2} \to 0$ kad $\mathbf{h} \to 0$.