Kibővített előadásjegyzet a VIK A2 Matematika tárgyhoz

Pitrik József

8. oktatási hét

Ez a oktatási segédlet az A2 előadásaim kommentekkel ellátott, szerkesztett változata, ami azzal a céllal íródott, hogy a távoktatás keretein belül elősegítse a Koronavírus járvány miatt sajnálatosan elmaradt előadások pótlását. Mivel sietve készült, ezért számos hibát, elírást tartalmazhat és a nyelvi megfogalmazások is inkább közelebb állnak a beszélt nyelvhez, mint a szokásos jegyzeteké. Kiegészítésképpen a jegyzetek végén ajánlott irodalom található, mely a további elmélyedést segítheti. Bármilyen észrevételt szívesen fogadok a pitrik@math.bme.hu címre. A subjectbe legyenek szívesek beírni, hogy „A2 jegyzet". Mindenkinek jó egészséget, türelmet, kitartást és sikeres felkészülést kívánok! A mihamarabbi viszontlátás reményében, üdvözlettel: PJ.

1. Lineáris leképezések skalárszorosa, összege és kompozíciója

Emlékezzünk vissza, hogy ha V és W két vektortér ugyanazon $\mathbb{K}$ számtest felett, míg $T : V \to W$ egy lineáris leképezés, akkor ezt a $T \in \mathcal{L}(V, W)$ módon jelöltük. Legyen $T, U \in \mathcal{L}(V, W)$ és $c \in \mathbb{K}$ tetszőleges. Definiáljuk T és U leképezések összegét, illetve egy lineáris leképezés skalárszorosát a következőképpen:

$$(T + U)(\mathbf{v}) := T(\mathbf{v}) + U(\mathbf{v}), \quad \text{minden } \mathbf{v} \in V \quad \text{esetén},$$

illetve

$$(cT)(\mathbf{v}) := cT(\mathbf{v}), \quad \text{minden } \mathbf{v} \in V \quad \text{esetén}.$$

Könnyű látni, hogy ezekkel a definíciókkal $cT, T+U \in \mathcal{L}(V, W)$, vagyis $\mathcal{L}(V, W)$ vektortér $\mathbb{K}$ számtest felett ezekkel a műveletekkel. Legyen V egy n-dimenziós vektortér, benne az $\alpha = (\mathbf{v}_1, \mathbf{v}_2 \ldots, \mathbf{v}_n)$ bázissal, W pedig m-dimenziós $\beta = (\mathbf{w}_1, \mathbf{w}_2 \ldots, \mathbf{w}_m)$ bázissal. Az előző fejezet alapján az (α, β) bázispárra vonatkozólag meg tudjuk adni T és U mátrixát, jelölje ezt szokásos módon: $[T]_\alpha^\beta$, illetve $[U]_\alpha^\beta$. Könnyű meggondolni, hogy mik lesznek a $T+U$ illetve cT lineáris leképezések mátrixai az (α,β) bázispárra vonatkozólag:

$$[T+U]_\alpha^\beta = [T]_\alpha^\beta + [U]_\alpha^\beta \quad \text{illetve} \quad [cT]_\alpha^\beta = c[T]_\alpha^\beta,$$

vagyis egyszerűen a reprezentáló mátrixok összegét, illetve skalárszorosát kell vennünk. A fenti műveletekkel tudjuk értelmezni véges sok $\mathcal{L}(V,W)$-beli elem lineáris kombinációját is.

Ha ezentúl valamilyen probléma, vagy feladat megoldásakor rögzítjük az α és β bázisokat és a lineáris leképezések ezen bázispárra vonatkozó mátrixaival dolgozunk, visszatérünk a mátrixos jelöléshez, azaz $[T]_\alpha^\beta$ helyett $\mathbf{T}$-t, $[T(\mathbf{v})]^\beta = [T]_\alpha^\beta[\mathbf{v}]^\alpha$ helyett $\mathbf{T}\,\mathbf{v}$-t, $[T]_\alpha^\beta + [U]_\alpha^\beta$ helyett $\mathbf{T}+\mathbf{U}$-t, $c[T]_\alpha^\beta$ helyett $c\,\mathbf{T}$-t írunk. A következőkben megvizsgáljuk lineáris leképezések kompozícióját.

Definíció. *Legyenek V_1, V_2, V_3 vektorterek ugyanazon számtest felett. Legyen $T \in \mathcal{L}(V_1, V_2)$ és $U \in \mathcal{L}(V_2, V_3)$. Ekkor a T és U lineáris leképezések* ***kompozícióján*** *vagy* ***szorzatán*** *az*

$$U \circ T \equiv (UT) \in \mathcal{L}(V_1, V_3), \quad (U \circ T)(\mathbf{v}) \equiv (UT)(\mathbf{v}) := U(T(\mathbf{v}))$$

leképezést értjük.

Könnyű meggondolni, hogy az így definiált leképezés valóban lineáris. Látható, hogy lineáris leképezések szorzatát nem mindegy, hogy milyen sorrendben végezzük el. Sőt, azért mert UT értelmes, egyáltalán nem biztos, hogy TU is értelmes. Ha viszont speciálisan $V_1 = V_2 = V_3$, akkor természetesen UT és TU is értelmes, de általában ekkor sem igaz, hogy $UT = TU$. Fennáll viszont az alábbi

Tétel. *Legyen $c \in \mathbb{K}$ egy skalár és legyenek A, B, C olyan lineáris leképezések, melyekre ha az alábbi egyenlőségek egyik oldala értelmezve van, akkor a másik oldal is. Ekkor fennállnak a következők:*

1. *$A(BC) = (AB)C$ (asszociatívitás),*
2. *$A(B+C) = AB + AC$ és $(A+B)C = AC + BC$ (disztributívitás),*
3. *$c(AB) = (cA)B = A(cB)$.*

Ha most a V_1, V_2, V_3 vektorterekben rendre rögzítünk egy α, β, γ bázist, és a $T : V_1 \to V_2$ leképezés mátrixa az (α, β) bázisra vonatkozólag $[T]_\alpha^\beta = \mathbf{T}$, míg az $U : V_2 \to V_3$ leképezés mátrixa az (β, γ) bázisra vonatkozólag $[U]_\beta^\gamma = \mathbf{U}$, akkor az $(UT) : V_1 \to V_3$ leképezés mátrixa az (α, γ) bázispárra vonatkozólag

$$[UT]_\alpha^\gamma = [U]_\beta^\gamma [T]_\alpha^\beta = \mathbf{U}\,\mathbf{T}\,.$$

Nincs tehát más dolgunk, mint az egyes reprezentáló mátrixokat megfelelő sorrendben összeszorozni. Nézzünk egy példát!

Példa 1. Írjuk fel $\mathbb{R}^3$ standard bázisában annak a lineáris leképezésnek a mátrixát, ami minden $\mathbb{R}^3$-beli vektort először $30°$-kal elforgat az x-tengely körül, majd az elforgatott vektort levetíti az $S: x-y+2z=0$ egyenletű síkra! Hova képezi le a

$$\mathbf{v}=\begin{pmatrix}1\\2\\0\end{pmatrix}$$

vektort a transzformáció?

Emlékezzünk, hogy az x-tengely körüli θ szögű forgatás mátrixa a standard bázisban:

$$R^x_\theta=\begin{pmatrix}1&0&0\\0&\cos\theta&-\sin\theta\\0&\sin\theta&\cos\theta\end{pmatrix}.$$

A mi esetünkben tehát $\theta=30°$ helyettesítés után:

$$\mathbf{R}\equiv R^x_{30°}=\begin{pmatrix}1&0&0\\0&\frac{\sqrt{3}}{2}&-\frac{1}{2}\\0&\frac{1}{2}&\frac{\sqrt{3}}{2}\end{pmatrix}.$$

Az origón átmenő $\mathbf{n}$ normálvektorú síkra való vetítés $\mathbf{P}$ mátrixa

$$\mathbf{P}=\mathbf{I}-\mathbf{n}\circ\mathbf{n},$$

(lásd: Wettl Ferenc, Lineáris algebra, 305. o, 7.32. Állítás, http://tankonyvtar.ttk.bme.hu/pdf/14.pdf). A mi esetünkben az S sík normálvektora $\mathbf{n}=(1,-1,2)^T$, így

$$\mathbf{n}\circ\mathbf{n}=\begin{pmatrix}1\\-1\\2\end{pmatrix}\begin{pmatrix}1&-1&2\end{pmatrix}=\begin{pmatrix}1&-1&2\\-1&1&-2\\2&-2&4\end{pmatrix},$$

azaz

$$\mathbf{P}=\mathbf{I}-\mathbf{n}\circ\mathbf{n}=\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}-\begin{pmatrix}1&-1&2\\-1&1&-2\\2&-2&4\end{pmatrix}=\begin{pmatrix}0&1&-2\\1&0&2\\-2&2&-3\end{pmatrix}.$$

A keresett transzformáció mátrixa, tehát

$$\mathbf{T}=\mathbf{P}\,\mathbf{R}=\begin{pmatrix}0&1&-2\\1&0&2\\-2&2&-3\end{pmatrix}\begin{pmatrix}1&0&0\\0&\frac{\sqrt{3}}{2}&-\frac{1}{2}\\0&\frac{1}{2}&\frac{\sqrt{3}}{2}\end{pmatrix}=\begin{pmatrix}0&\frac{\sqrt{3}}{2}-1&-\frac{1}{2}-\sqrt{3}\\1&1&\sqrt{3}\\-2&\sqrt{3}-\frac{3}{2}&-1-\frac{3\sqrt{3}}{2}\end{pmatrix}.$$

A $\mathbf{v}$ vektor $\mathbf{T}$ általi képe:

$$\mathbf{v}'=\mathbf{T}\,\mathbf{v}=\begin{pmatrix}0&\frac{\sqrt{3}}{2}-1&-\frac{1}{2}-\sqrt{3}\\1&1&\sqrt{3}\\-2&\sqrt{3}-\frac{3}{2}&-1-\frac{3\sqrt{3}}{2}\end{pmatrix}\begin{pmatrix}1\\2\\0\end{pmatrix}=\begin{pmatrix}\sqrt{3}-2\\3\\-5+2\sqrt{3}\end{pmatrix}.$$

2. A báziscsere transzformáció

A következőkben megvizsgáljuk, hogy mi történik a koordinátás leírásban, ha egy bázisról egy másik bázisra térünk át. Az áttekinthetőség kedvéért 2-dimenziós vektortéren dolgozunk, a bizonyítások ugyanígy mennek magasabb dimenzió esetén is. Legyen tehát V egy 2-dimenziós vektortér, és tekintsünk két (rendezett) bázist V-ben:

$$\alpha = (\mathbf{e}_1, \mathbf{e}_2), \quad \alpha' := (\mathbf{e}_1', \mathbf{e}_2').$$

Tekintsünk egy tetszőleges $\mathbf{v} \in V$ vektort és fejtsük ki az α illetve α' bázis szerint:

$$\mathbf{v} = v_1\,\mathbf{e}_1 + v_2\,\mathbf{e}_2, \quad \text{azaz} \quad [\mathbf{v}]^{\alpha} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix},$$

illetve

$$\mathbf{v} = v_1'\,\mathbf{e}_1' + v_2'\,\mathbf{e}_2', \quad \text{azaz} \quad [\mathbf{v}]^{\alpha'} = \begin{pmatrix} v_1' \\ v_2' \end{pmatrix}.$$

A célunk, hogy megtaláljuk az összefüggést $[\mathbf{v}]^{\alpha}$ és $[\mathbf{v}]^{\alpha'}$ között. E célból írjuk fel α bázis vektorait az α' bázisban:

$$\begin{aligned} \mathbf{e}_1 &= a_1\,\mathbf{e}_1' + a_2\,\mathbf{e}_2' \\ \mathbf{e}_2 &= b_1\,\mathbf{e}_1' + b_2\,\mathbf{e}_2', \end{aligned}$$

vagyis az eddigi jelöléseinkkel:

$$[\mathbf{e}_1]^{\alpha'} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \quad \text{és} \quad [\mathbf{e}_2]^{\alpha'} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}.$$

Ezeket visszaírva $\mathbf{v}$ kifejezésébe:

$$\begin{aligned} \mathbf{v} &= v_1\,\mathbf{e}_1 + v_2\,\mathbf{e}_2 \\ &= v_1(a_1\,\mathbf{e}_1' + a_2\,\mathbf{e}_2') + v_2(b_1\,\mathbf{e}_1' + b_2\,\mathbf{e}_2') \\ &= (v_1 a_1 + v_2 b_1)\,\mathbf{e}_1' + (v_1 a_2 + v_2 b_2)\,\mathbf{e}_2', \end{aligned}$$

vagyis a $\mathbf{v}$ vektor koordinátás alakja az α' bázisban

$$[\mathbf{v}]^{\alpha'} = \begin{pmatrix} v_1 a_1 + v_2 b_1 \\ v_1 a_2 + v_2 b_2 \end{pmatrix}.$$

Vegyük észre, hogy ezt mátrixos alakban is írhatjuk:

$$[\mathbf{v}]^{\alpha'} = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} [\mathbf{v}]^{\alpha}.$$

Az ezekben a kifejezésekben megjelenő mátrixot az $\alpha \to \alpha'$ **báziscsere transzformáció mátrixának** hívjuk, melyre egy közkeletű jelölés:

$$[I]_\alpha^{\alpha'} = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix},$$

mellyel tehát

$$[\mathbf{v}]^{\alpha'} = [I]_\alpha^{\alpha'} [\mathbf{v}]^\alpha.$$

Látható, hogy hogyan kapható meg ez a mátrix: α bázisvektorainak α' bázisban felírt koordinátavektorait kell oszlopvektorokkén egy mátrixban felsorolni, azaz

$$[I]_\alpha^{\alpha'} = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} = \left(\ [\mathbf{e}_1]^{\alpha'} \quad | \quad [\mathbf{e}_2]^{\alpha'} \ \right).$$

A fenti módon igazolható az alábbi általános, n-dimenziós estre vonatkozó tétel.

Tétel. *Legyen V egy n-dimenziós vektortér, benne két (rendezett) bázissal:*

$$\alpha = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n), \quad \alpha' := (\mathbf{e}'_1, \mathbf{e}'_2, \ldots, \mathbf{e}'_n).$$

Ekkor az $\alpha \to \alpha'$ báziscsere transzformáció mátrixa:

$$[I]_\alpha^{\alpha'} = \left(\ [\mathbf{e}_1]^{\alpha'} \quad | \quad [\mathbf{e}_2]^{\alpha'} \quad | \quad \ldots \quad | \quad [\mathbf{e}_n]^{\alpha'} \ \right),$$

mellyel fennáll, hogy

$$[\mathbf{v}]^{\alpha'} = [I]_\alpha^{\alpha'} [\mathbf{v}]^\alpha.$$

Nézzünk egy egyszerű példát!

Példa 2. Legyen $V = \mathbb{R}^2$ és tekintsünk két bázist V-ben:

$$\alpha = \left(\mathbf{e}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathbf{e}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right) \quad \text{és} \quad \alpha' = \left(\mathbf{e}'_1 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \mathbf{e}'_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right).$$

Tudjuk, hogy a $\mathbf{v} \in V$ vektor az α bázisban

$$[\mathbf{v}]^\alpha = \begin{bmatrix} 3 \\ -2 \end{bmatrix},$$

a cél, hogy megadjuk $\mathbf{v}$-t az α' bázisban is. Ehhez ki kell fejtenünk az α bázis vektorait az α' bázisban, azaz meg kell keresnünk azon a_1, a_2, b_1, b_2 együtthatókat, melyekkel

$$a_1 \begin{bmatrix} 2 \\ -1 \end{bmatrix} + a_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \text{és} \quad b_1 \begin{bmatrix} 2 \\ -1 \end{bmatrix} + b_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Könnyű számolás után azt kapjuk, hogy $a_1 = 2, a_2 = 3$ illetve $b_1 = 0, b_2 = -1$. Ez azt jelenti, hogy

$$[\mathbf{e}_1]^{\alpha'} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \quad \text{és} \quad [\mathbf{e}_2]^{\alpha'} = \begin{bmatrix} 0 \\ -1 \end{bmatrix}.$$

Innen már könnyen felírhatjuk az $\alpha \to \alpha'$ báziscsere transzformációs mátrixát:

$$[I]_{\alpha}^{\alpha'} = \begin{bmatrix} 2 & 0 \\ 3 & -1 \end{bmatrix},$$

ahonnan a $\mathbf{v}$ vektor az α' bázisban:

$$[\mathbf{v}]^{\alpha'} = \begin{bmatrix} 2 & 0 \\ 3 & -1 \end{bmatrix} \begin{bmatrix} 3 \\ -2 \end{bmatrix} = \begin{bmatrix} 6 \\ 11 \end{bmatrix}.$$

Mivel az $[I]_{\alpha}^{\alpha'}$ bázistranszformációs mátrix bázist bázisba visz, ezért természetesen kölcsönösen egyértelmű, azaz invertálható is. Ez azt jelenti, hogy a

$$[\mathbf{v}]^{\alpha'} = [I]_{\alpha}^{\alpha'} [\mathbf{v}]^{\alpha}$$

egyenletet balról az inverzmátrixszal megszorozva, azt kapjuk, hogy

$$[\mathbf{v}]^{\alpha} = \left([I]_{\alpha}^{\alpha'}\right)^{-1} [\mathbf{v}]^{\alpha'}.$$

Ez azt is mutatja, hogy az $\alpha' \to \alpha$ báziscsere transzformációs mátrixa, az $\alpha \to \alpha'$ báziscsere transzformáció mátrixának az inverze:

$$[I]_{\alpha'}^{\alpha} = \left([I]_{\alpha}^{\alpha'}\right)^{-1}. \tag{1}$$

Ezt az észrevételt a továbbiakban messzemenőleg használni fogjuk. Most megvizsgáljuk, hogy mi történik egy lineáris leképezés mátrixával báziscsere esetén. A következő tételt az egyszerűség kedvéért csak abban az esetben mondjuk ki, mikor a lineáris leképezés indulási- és érkezési tere azonos. Az általános eset megtalálható az ajánlott irodalomban.

Tétel. *Legyen V egy véges dimenziós vektortér α illetve α' bázissal. Jelöljük az $\alpha \to \alpha'$ báziscsere mátrixát az egyszerűség kedvéért $\mathbf{S}$-sel, azaz $\mathbf{S} = [I]_{\alpha}^{\alpha'}$. Legyen $T : V \to V$ egy lineáris transzformáció, melynek az (α, α) bázispárra vonatkozó mátrixát az egyszerűség kedvéért jelölje $\mathbf{T}_{\alpha} = [T]_{\alpha}^{\alpha}$. Ekkor a T leképezés mátrixa az (α', α') bázispárra nézve*

$$\mathbf{T}_{\alpha'} \equiv [T]_{\alpha'}^{\alpha'} = \mathbf{S}\,\mathbf{T}_{\alpha}\,\mathbf{S}^{-1}. \tag{2}$$

(Írjuk le a biztonság kedvéért az eredeti jelöléseinkkel is, kihasználva (1)-t

$$[T]_{\alpha'}^{\alpha'} = [I]_{\alpha}^{\alpha'} [T]_{\alpha}^{\alpha} [I]_{\alpha'}^{\alpha}.)$$

Bizonyítás. Legyen $\mathbf{v} \in V$ egy tetszőleges vektor. Az α' bázisban dolgozva, tudjuk, hogy

$$[T(\mathbf{v})]^{\alpha'} = [T]^{\alpha'}_{\alpha'}[\mathbf{v}]^{\alpha'} = \mathbf{T}_{\alpha'}[\mathbf{v}]^{\alpha'}.$$

Hasonlóan, az α bázisban dolgozva:

$$[T(\mathbf{v})]^{\alpha} = [T]^{\alpha}_{\alpha}[\mathbf{v}]^{\alpha} = \mathbf{T}_{\alpha}[\mathbf{v}]^{\alpha}.$$

Tudjuk, hogy miután $\mathbf{S}$ írja le az $\alpha \to \alpha'$ báziscserét, ezért

$$[\mathbf{v}]^{\alpha'} = \mathbf{S}[\mathbf{v}]^{\alpha}, \quad \text{vagyis} \quad [\mathbf{v}]^{\alpha} = \mathbf{S}^{-1}[\mathbf{v}]^{\alpha'}.$$

Ez utóbbit behelyettesítve, azt kapjuk, hogy

$$[T(\mathbf{v})]^{\alpha} = \mathbf{T}_{\alpha}[\mathbf{v}]^{\alpha} = \mathbf{T}_{\alpha}\,\mathbf{S}^{-1}[\mathbf{v}]^{\alpha'}.$$

Másrészt $[T(\mathbf{v})]^{\alpha'}$-t is felírhatjuk $[T(\mathbf{v})]^{\alpha}$ transzformáltjaként:

$$[T(\mathbf{v})]^{\alpha'} = \mathbf{S}[T(\mathbf{v})]^{\alpha},$$

amibe behelyettesítve az előző kifejezést:

$$[T(\mathbf{v})]^{\alpha'} = \mathbf{S}[T(\mathbf{v})]^{\alpha} = \mathbf{S}\,\mathbf{T}_{\alpha}\,\mathbf{S}^{-1}[\mathbf{v}]^{\alpha'},$$

vagyis

$$\mathbf{T}_{\alpha'} = \mathbf{S}\,\mathbf{T}_{\alpha}\,\mathbf{S}^{-1},$$

ahogy állítottuk. □

Amennyiben az α bázist a standard bázisnak vesszük, a számolások jóval egyszerűbbé válnak. Nézzünk erre egy egyszerű példát, a későbbiekben több alkalmazást is látni fogunk.

Példa 3. Tekintsük a

$$T : \mathbb{R}^2 \to \mathbb{R}^2, \quad T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} -x + 2y \\ 3x + y \end{bmatrix}$$

lineáris leképezést. Vegyünk fel két bázist $\mathbb{R}^2$-ben, legyenek ezek

$$\alpha = \left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) \quad \text{és} \quad \alpha' = \left(\begin{bmatrix} 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}\right).$$

Látható, hogy α bázisnak a standard bázist választottuk. Írjuk fel T leképezés mátrixát az α bázisban (pontosabban az (α, α) bázispárra vonatkozólag). Ez standard bázisban nagyon egyszerű, hiszen csak az egyes bázisvektorok T általi képét kell oszlopvektorként egy mátrixba írnunk, hiszen a képvektorok egyből standard bázisban lesznek felírva. Tehát

$$T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} -1 \\ 3 \end{bmatrix}, \quad \text{és} \quad T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ 1 \end{bmatrix},$$

így T mátrixa az α bázisban

$$\mathbf{T}_\alpha = \begin{pmatrix} -1 & 2 \\ 3 & 1 \end{pmatrix}.$$

Most kiváncsiak vagyunk a T leképezés mátrixára az α' bázisban (pontosabban az (α', α') bázispárra vonatkozólag). Dolgozhatunk úgy is mint az előbb, vagyis megnézhetjük, hogy mik lesznek az α' bázisvektorainak T általi képei, de ezeket most az α' bázisban kell felírnunk és így egy mátrixba helyezni oszlopvektorként. A másik lehetőség, hogy a fenti tétel alapján, báziscsere transzformációval dolgozunk. Ehhez el kell készítenünk az $\alpha \to \alpha'$ báziscsere $\mathbf{S} = [I]_\alpha^{\alpha'}$ mátrixát. Ehhez írjuk fel az α bázis vektorait az α' bázisban:

$$\begin{aligned} \begin{bmatrix} 1 \\ 0 \end{bmatrix} &= a_1 \begin{bmatrix} 2 \\ -1 \end{bmatrix} + a_2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ \begin{bmatrix} 0 \\ 1 \end{bmatrix} &= b_1 \begin{bmatrix} 2 \\ -1 \end{bmatrix} + b_2 \begin{bmatrix} 1 \\ 0 \end{bmatrix}. \end{aligned}$$

Az egyenletrendszert megoldva azt kapjuk, hogy $a_1 = 0, a_2 = 1$ illetve $b_1 = -1, b_2 = 2$. Innen a báziscsere transzformáció mátrixa:

$$\mathbf{S} = [I]_\alpha^{\alpha'} = \begin{bmatrix} 0 & -1 \\ 1 & 2 \end{bmatrix}.$$

Az inverz mátrix egyszerű számolás után:

$$\mathbf{S}^{-1} = \begin{bmatrix} 2 & 1 \\ -1 & 0 \end{bmatrix}.$$

Most már mindenünk megvan, hogy a (2) képlet alapján megadjuk T mátrixát az α' bázisra vonatkozólag:

$$\mathbf{T}_{\alpha'} \equiv [T]_{\alpha'}^{\alpha'} = \mathbf{S}\,\mathbf{T}_\alpha\,\mathbf{S}^{-1} = \begin{bmatrix} 0 & -1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} -1 & 2 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} -5 & -3 \\ 6 & 5 \end{bmatrix}.$$

Fontos tehát megjegyeznünk, hogy egy lineáris transzformáció mátrixa mindig bázisfüggő, viszont ezek a mátrixok speciális kapcsolatban vannak egymással. Ezt a kapcsolatot hivatott kifejezni az alábbi definíció.

Definíció. *Legyen* $\mathbf{A}$ *és* $\mathbf{B}$ *két* $n \times n$*-es mátrix. Azt mondjuk, hogy* $\mathbf{A}$ ***hasonló*** $\mathbf{B}$ *mátrixhoz, ha létezik* $\mathbf{S}$ *invertálható mátrix, mellyel*

$$\mathbf{B} = \mathbf{S}\,\mathbf{A}\,\mathbf{S}^{-1}.$$

Világos, hogy az $\mathbf{R} = \mathbf{S}^{-1}$ jelöléssel

$$\mathbf{A} = \mathbf{R}\,\mathbf{B}\,\mathbf{R}^{-1},$$

vagyis ekkor $\mathbf{B}$ is hasonló $\mathbf{A}$-hoz: a hasonlóság szimmetrikus reláció. Könnyű meggondolni, hogy a hasonlóság nemcsak szimmetrikus, hanem reflexív és tranzitív is, vagyis ekvivalenciareláció. Ezért gyakran, azt, hogy $\mathbf{A}$ és $\mathbf{B}$ hasonlóak, az

$$\mathbf{A} \backsimeq \mathbf{B}$$

jelöléssel fejezzük ki. A fentiek alapján tehát azt mondhatjuk, hogy ha két mátrix hasonló, akkor ugyanazt a lineáris leképezést írják le, csak különböző bázisban.

3. Sajátérték, sajátvektor

Legyen V egy n-dimenziós vektortér a $\mathbb{K}$ számtest felett, és legyen $T : V \to V$ egy lineáris leképezés. Az alkalmazások szempontjából fontos kérdés, hogy vannak-e olyan nem $\mathbf{0}$ vektorok V-ben, melyeknek T csak a hosszát változtatja meg, de az irányát nem, azaz keressük azon $\mathbf{v} \in V$, $\mathbf{v} \neq \mathbf{0}$ vektorokat, melykhez létezik olyan $\lambda \in \mathbb{K}$ szám, melyre fennáll, hogy

$$T(\mathbf{v}) = \lambda\,\mathbf{v}\,.$$

Nézzünk néhány szemléltető példát:

1. Ha $R : \mathbb{R}^3 \to \mathbb{R}^3$ valamilyen origón átmenő $\mathbf{v}$ irányvektorú egyenes körüli $\theta \neq k \cdot 360°$ szögű forgatás ($k \in \mathbb{Z}$), akkor világos, hogy a forgástengellyel párhuzamos vektorokat R helyben hagyja, viszont az összes ettől eltérő vektornak megváltoztatja az irányát. Ezek szerint, minden $\mathbf{w} \parallel \mathbf{v}$ vektorra:

$$R(\mathbf{w}) = 1 \cdot \mathbf{w}\,.$$

2. Ha $P : \mathbb{R}^3 \to \mathbb{R}^3$ valamilyen origón átmenő $\mathbf{v}$ irányvektorú egyenesre való merőleges vetítés (projekció), akkor P az összes $\mathbf{v}$-vel párhuzamos vektort helyben hagyja, azaz minden $\mathbf{w} \parallel \mathbf{v}$ vektorra:

$$P(\mathbf{w}) = 1 \cdot \mathbf{w}\,.$$

   Ha most olyan vektorokat tekintünk, melyek merőlegesek $\mathbf{v}$-re, akkor ezeket P mind az origóba viszi, azaz ha $\mathbf{w} \perp \mathbf{v}$, akkor

$$P(\mathbf{w}) = \mathbf{0} = 0 \cdot \mathbf{w}\,.$$

   Könnyű meggondolni, hogy a felsoroltakon kívül minden egyéb vektornak a P projekció megváltoztatja az irányát.

3. Ha $P : \mathbb{R}^3 \to \mathbb{R}^3$ valamilyen origón átmenő $\mathbf{n}$ normálvektorú S síkra való merőleges vetítés (projekció), akkor P az S sík összes vektorát helyben hagyja, azaz minden $\mathbf{w} \perp \mathbf{n}$ vektorra:
$$P(\mathbf{w}) = 1 \cdot \mathbf{w}\,.$$
Ha most olyan vektorokat tekintünk, melyek merőlegesek S-re, akkor ezeket P mind az origóba viszi, azaz ha $\mathbf{w} \parallel \mathbf{n}$, akkor
$$P(\mathbf{w}) = \mathbf{0} = 0 \cdot \mathbf{w}\,.$$
Könnyű meggondolni, hogy a felsoroltakon kívül minden egyéb vektornak a P projekció megváltoztatja az irányát.

Szeretnénk a fenti problémát általánosan, tetszőleges T lineáris leképezés esetén tárgyalni. Ekkor érdemes egy bázist rögzíteni V-ben és végig ebben dolgozni. Ilyenkor tehát a $V \simeq \mathbb{K}^n$ izomorfizmus alapján V vektorait n-hosszúságú $\mathbf{v}$ oszlopvektorokkal reprezentáljuk, míg az általános T leképezés helyett annak $\mathbf{T}$ mátrixával dolgozunk. Így a $T(\mathbf{v})$ kifejezés helyett egyszerűen $\mathbf{T}\,\mathbf{v}$-t, vagyis a $\mathbf{T}$ mátrixszal való szorzást írhatunk. Ezután a megjegyzés után tudjuk definiálni a következőt.

Definíció. *Legyen* $\mathbf{T} \in \mathcal{M}_n(\mathbb{K})$ *egy mátrix. Azon* $\lambda \in \mathbb{K}$ *számokat, melyekhez létezik* $\mathbf{v} \neq \mathbf{0}$, $\mathbf{v} \in \mathbb{K}^n$ *vektor, amivel*
$$\mathbf{T}\,\mathbf{v} = \lambda\,\mathbf{v} \tag{3}$$
teljesül, a $\mathbf{T}$ *mátrix* $\mathbf{v}$ *vektorhoz tartozó* ***sajátértékének*** *nevezzük. A* $\mathbf{v}$ *vektort* $\mathbf{T}$ *mátrix* λ *sajátértékéhez tartozó* ***sajátvektorának*** *hívjuk. Az (3) egyenlet* $\mathbf{T}$ ***sajátérték egyenlete****.* $\mathbf{T}$ *mátrix sajátértékeinek összessége* $\mathbf{T}$ ***spektruma****, amit* Spec$(\mathbf{A})$*-val jelölünk.*

Megjegyzés 4. Nézzük meg a definíció néhány egyszerű következményét.

1. A definícióból a $\mathbf{v} = \mathbf{0}$ esetet azért kell kizárnunk, mert $\mathbf{T}\,\mathbf{0} = \lambda\,\mathbf{0} = \mathbf{0}$ teljesülne minden $\lambda \in \mathbb{K}$ számra, vagyis különben a $\mathbb{K}$ számtest minden eleme sajátérték lenne $\mathbf{0}$ sajátvektorral.

2. Látható, hogy a $\lambda = 0$ pontosan akkor sajátérték, ha $\mathbf{T}\,\mathbf{v} = \mathbf{0}$, azaz $\mathbf{v} \in \operatorname{Ker} \mathbf{T}$, vagyis pontosan a 0 sajátértékhez tartozó sajátvektorok alkotják $\mathbf{T}$ magterét.

3. Minden sajátvektorhoz pontosan egy sajátérték tartozik. Csakugyan, ha valamely $\mathbf{v} \neq \mathbf{0}$ sajátvektorhoz két sajátérték, λ és μ is tartozna, akkor
$$\mathbf{T}\,\mathbf{v} = \lambda\,\mathbf{v} = \mu\,\mathbf{v}$$
miatt $(\lambda - \mu)\,\mathbf{v} = \mathbf{0}$ lenne, ahol $\mathbf{v} \neq \mathbf{0}$, így $\lambda = \mu$ következne.

4. Az viszont előfordulhat, hogy egy adott sajátértékhez több sajátvektor is tartozik (lásd például a fenti 3. példát). Könnyű látni, hogy egy adott λ sajátértékhez tartozó vektorok a $\mathbf{0}$ vektorral együtt alteret alkotnak. Ehhez csak azt kell belátnunk, hogy ezen vektorok zártak az összeadásra és a skalárral való szorzásra nézve. Legyen tehát $\mathbf{v}$ és $\mathbf{w}$ két λ-hoz tartozó sajátvektor, azaz

$$\mathbf{T}\,\mathbf{v} = \lambda\,\mathbf{v}, \quad \text{és} \quad \mathbf{T}\,\mathbf{w} = \lambda\,\mathbf{w}\,.$$

Ekkor

$$\mathbf{T}(\mathbf{w}+\mathbf{w}) = \mathbf{T}\,\mathbf{v} + \mathbf{T}\,\mathbf{w} = \lambda\,\mathbf{v} + \lambda\,\mathbf{w} = \lambda(\mathbf{v}+\mathbf{w}),$$

vagyis $\mathbf{v}+\mathbf{w}$ csakugyan sajátvektor, ami λ-hoz tartozik. Hasonlóan, ha $c \in \mathbb{K}$, akkor

$$\mathbf{T}(c\,\mathbf{v}) = c\,\mathbf{T}\,\mathbf{v} = c\lambda\,\mathbf{v} = \lambda(c\,\mathbf{v}),$$

vagyis $c\,\mathbf{v}$ szintén λ-hoz tartozó sajátvektor. Ez utóbbi tehát azt mutatja, hogy ha $\mathbf{v}$ sajátvektor, akkor annak tetszőleges $\neq 0$ szorosa szintén sajátvektor ugyanazzal a sajátértékkel, így néha szokásos sajátvektorok helyett **sajátirányokról** beszélni. Egy adott λ sajátértékhez tartozó sajátvektorok (és a $\mathbf{0}$ vektor) által kifeszített teret a λ sajátértékhez tartozó **sajátaltérnek** nevezzük. A λ sajátértékhez tartozó sajátaltér dimenzióját λ **geometriai multiplicitásának** hívjuk.

A következőkben megvizsgáljuk, hogy miként lehet meghatározni egy mátrix sajátértékeit és sajátvektorait. Induljunk ki a

$$\mathbf{T}\,\mathbf{v} = \lambda\,\mathbf{v}$$

sajátérték egyenletből. Ha ezt az egyenletet átrendezzük a bal oldalra, akkor az $\mathbf{I}$ $n\times n$-es egységmátrix segítségével a

$$(\mathbf{T} - \lambda\,\mathbf{I})\,\mathbf{v} = \mathbf{0} \tag{4}$$

egyenlethez jutunk, ami egy homogén lineáris egyenletrendszer mátrixos alakban. Tanultuk, hogy ennek mindig létezik a $\mathbf{v} = \mathbf{0}$ triviális megoldása, de bennünket ez most éppen nem érdekel. Arra vagyunk kiváncsiak, hogy mikor létezik a (4) egyenletnek nem triviális megoldása. A lineáris egyenletrendszerek elméletéből tudjuk, hogy (4)-nek pontosan akkor létezik nem triviális megoldása, ha

$$\det(\mathbf{T} - \lambda\,\mathbf{I}) = 0. \tag{5}$$

Az (5) egyenletről a determináns kifejtése után kiderül, hogy a bal oldalon egy λ-ban n-edfokú polinom áll, ezt a $\mathbf{T}$ mátrix **karakterisztikus polinomjának** nevezzük. A sajátértékek tehát a karakterisztikus polinom gyökei. A gyökök megkeresésénél már nem mindegy, hogy hogyan választjuk meg a $\mathbb{K}$ számtestet. Amennyiben $\mathbb{K} = \mathbb{C}$, akkor használhatjuk az Algebra alaptételét. Emlékezzünk vissza, hogy az **Algebra alaptétele** kimondja, hogy egy n-edfokú polinomnak (az együtthatók lehetnek valósak vagy komplexek), multiplicitásokkal együtt pontosan n darab $\mathbb{C}$-beli gyöke van. Ez azt jelenti, hogy a $\mathbb{K} = \mathbb{C}$ számtest esetén a karakterisztikus polinomnak mindig van legalább egy $\mathbb{C}$-beli gyöke, vagyis mindig van legalább egy sajátértékünk. Az Algebra alaptétele ellenben nem áll fenn $\mathbb{R}$-ben, azaz nem biztos, hogy van valós gyökünk. Elképzelhető tehát, hogy $\mathbb{K} = \mathbb{R}$ felett vektortér esetén $\mathbf{T}$-nek nincs sajátértéke. Azt, hogy egy adott λ sajátérték a karakterisztikus polinomnak hányszoros gyöke a $\mathbb{K}$ számtestben, λ **algebrai multiplicitásának** nevezzük. Látni fogjuk, hogy az algebrai- és a geometriai multiplicitás (vagyis a sajátaltér dimenziója) bizonyos esetekben megegyezik, és ennek messzemenő következményei vannak.

Példa 5. Tekintsük a

$$\mathbf{T} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

mátrixot. Ezt a mátrixot most felfoghatjuk egy $\mathbb{K} = \mathbb{C}$ feletti és egy $\mathbb{K} = \mathbb{R}$ feletti V vektortéren ható lineáris leképezés mátrixaként is. Az első esetben $V \simeq \mathbb{C}^2$ és $\mathbf{T} \in \mathcal{M}_2(\mathbb{C})$, a második esetben $V \simeq \mathbb{R}^2$ és $\mathbf{T} \in \mathcal{M}_2(\mathbb{R})$. Adjuk meg $\mathbf{T}$ sajátértékeit mindkét esetben. A karakterisztikus polinom mindkét esetben ugyanaz:

$$\det(\mathbf{T} - \lambda\,\mathbf{I}) = \det\left(\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right) = \det\begin{pmatrix} -\lambda & -1 \\ 1 & -\lambda \end{pmatrix} = \lambda^2 + 1 = 0.$$

A $\lambda^2 + 1 = 0$ polinomnak $\mathbb{C}$-ben két gyöke van: $\lambda_1 = i$ és $\lambda_2 = -i$. Tehát $\mathbb{K} = \mathbb{C}$ esetén két sajátértékünk is van, mindkettő algebrai multiplicitása 1. A karakterisztikus polinomnak viszont nincs valós gyöke, tehát ha $\mathbb{K} = \mathbb{R}$ felett nézzük a problémát, akkor $\mathbf{T}$-nek nincs egyetlen sajátértéke sem.

Példa 6. Most keressük meg a

$$\mathbf{T} = \begin{pmatrix} 1 & -1 & -1 \\ 1 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix}$$

mátrix sajátértékeit. A karakterisztikus polinom

$$p(\lambda) = \det(\mathbf{T} - \lambda\,\mathbf{I}) = \det\begin{pmatrix} 1-\lambda & -1 & -1 \\ 1 & 1-\lambda & 0 \\ 3 & 0 & 1-\lambda \end{pmatrix} = (1-\lambda)(\lambda^2 - 2\lambda + 5) = 0.$$

Kis számolással megkaphatjuk a gyököket: $\lambda_1 = 1$, $\lambda_2 = 1 + 2i$ és $\lambda_3 = 1 - 2i$. Látható, hogy $\mathbb{K} = \mathbb{C}$ felett nézve a problémát három különböző sajátértékünk van (vagyis mindegyik gyök algebrai multiplicitása 1), míg $\mathbb{K} = \mathbb{R}$ felett csak egy sajátértéke van a mátrixnak, szintén 1 algebrai multiplicitással.

Most térjünk vissza az általános problémához. Tegyük fel, hogy megtaláltuk a karakterisztikus egyenlet gyökeit, vagyis a sajátértékeket. Hogyan lehet meghatározni egy adott sajátértékhez tartozó sajátvektor(oka)t? A dolog egyszerű: az adott λ sajátértéket visszaírva a sajátérték egyenletbe, meg kell határozni azon $\mathbf{v} \neq \mathbf{0}$ vektorokat, melyekre

$$\mathbf{T}\,\mathbf{v} = \lambda\,\mathbf{v}, \quad \text{vagy másképp} \quad (\mathbf{T} - \lambda\,\mathbf{I})\,\mathbf{v} = \mathbf{0}$$

teljesül, vagyis meg kell oldanunk egy egyenletrendszert. Nézzük meg ezt egy példán:

Példa 7. Tekintsük a

$$\mathbf{T} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

mátrixot. A karakterisztikus polinom ezúttal

$$\det(\mathbf{T} - \lambda\,\mathbf{I}) = \det\left(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right) = \det\begin{pmatrix} -\lambda & 1 \\ 1 & -\lambda \end{pmatrix} = \lambda^2 - 1 = (\lambda - 1)(\lambda + 1) = 0,$$

melynek gyökei $\lambda_1 = 1$ és $\lambda_2 = -1$. Látható, hogy most akár $\mathbb{R}$, akér $\mathbb{C}$ felett vagyunk, mindig van két sajátértékünk. Számoljuk ki a $\lambda_1 = 1$-hez tartozó sajátvektort! A megoldandó egyenlet ekkor $(\mathbf{T} - 1 \cdot \mathbf{I})\,\mathbf{v} = \mathbf{0}$, azaz

$$\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Ezt könnyű megoldani, de gyakorlásképpen alkalmazzunk az együttható mátrixon Gauss-eliminációt:

$$\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \sim \begin{pmatrix} -1 & 1 \\ 0 & 0 \end{pmatrix},$$

ahol a második sorhoz hozzáadtuk az elsőt. Ezek szerint a megoldandó egyenlet: $-v_1 + v_2 = 0$, vagyis $v_1 = v_2$. A v_2-t p szabad paraméternek választva tehát a $\lambda_1 = 1$ sajátértékhez tartozó $\mathbf{v}_1$ sajátvektorok a következők:

$$\mathbf{v}_1 = \begin{pmatrix} p \\ p \end{pmatrix} = p \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Látható, hogy itt egyből találkozunk a már említett jelenséggel, inkább sajátirányaink vannak, mint sajátvektoraink, pl a $p = 1$ választással

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Most nézzük a $\lambda_2 = -1$ esetet! A megoldandó egyenlet ekkor $(\mathbf{T} - (-1) \cdot \mathbf{I})\,\mathbf{v} = \mathbf{0}$, azaz

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Ismét Gauss-eliminációval:

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix},$$

ahol a második sorhoz hozzáadtuk az elsőt. Ezek szerint a megoldandó egyenlet: $v_1 + v_2 = 0$, vagyis $v_1 = -v_2$. A v_2-t p szabad paraméternek választva tehát a $\lambda_2 = -1$ sajátértékhez tartozó $\mathbf{v}_2$ sajátvektorok a következők:

$$\mathbf{v}_2 = \begin{pmatrix} -p \\ p \end{pmatrix} = p \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

Ismét a $p = 1$ választással

$$\mathbf{v}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

Most megemlítjük a karakterisztikus polinom néhány nevezetes tulajdonságát. Vegyük egy általános $n \times n$-es mátrixot $\mathbb{K} = \mathbb{C}$ felett:

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \cdot & \cdot & \dots & \cdot \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}.$$

A karakterisztikus polinom, ekkor

$$p(\lambda) = \det(\mathbf{A} - \lambda\,\mathbf{I}) = \det \begin{pmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \cdot & \cdot & \dots & \cdot \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{pmatrix},$$

ami az első sora szerint kifejtve, a tagok rendezése után a következő alakra hozható (ellenőrizzük egy 3×3-as mátrixon!):

$$p(\lambda) = (-1)^n \lambda^n + (-1)^{n-1}(a_{11} + a_{22} + \dots + a_{nn})\lambda^{n-1} + \dots + \det \mathbf{A}\,.$$

Másrészről tudjuk, hogy ennek a polinomnak $\mathbb{C}$ felett pontosan n darab gyöke van, nevezetesen a $\lambda_1, \lambda_2, \dots, \lambda_n$ sajátértékek, ahol most úgy vesszük figyelembe a multiplicitásokat, hogy a felsorolt sajátértékek között lehetnek egyenlők. Mivel a polinom főegyütthatója $(-1)^n$, ezért a fentiek alapján felírhatjuk gyöktényezős alakban is, amit utána kifejtünk, vagyis:

$$p(\lambda) = (-1)^n \prod_{i=1}^{n} (\lambda - \lambda_i) = \prod_{i=1}^{n} (\lambda_i - \lambda) = (-1)^n \lambda^n + (-1)^{n-1}(\lambda_1 + \lambda_2 + \dots + \lambda_n)\lambda^{n-1} + \dots + \lambda_1 \lambda_2 \cdots \lambda_n.$$

Összehasonlítva a kétféle módon felírt karakterisztikus polinomban az egyes λ hatványok együtthatóit érdekes összefüggéseket kaphatunk. A konstans tagok összehasonlításából például azt kapjuk, hogy

$$\det \mathbf{A} = \lambda_1 \lambda_2 \cdots \lambda_n, \tag{6}$$

vagyis egy komplex test feletti mátrix determinánsa megegyezik a sajátértékeinek szorzatával! A λ^{n-1} hatvány együtthatóit összehasonlítva, azt kapjuk, hogy

$$a_{11} + a_{22} + \cdots + a_{nn} + = \lambda_1 + \lambda_2 + \cdots + \lambda_n,$$

vagyis a főátló beli elemek összege egyenlő a sajátértékek összegével. Ezt a közös értéket a **mátrix nyomának** hívjuk és $\operatorname{Tr} \mathbf{A}$-val jelöljük (nyom=trace angolul), de szokásos $\operatorname{Sp} \mathbf{A}$-val is jelölni (nyom=Spur németül), vagyis

$$\operatorname{Tr} \mathbf{A} = a_{11} + a_{22} + \cdots + a_{nn} + = \lambda_1 + \lambda_2 + \cdots + \lambda_n. \tag{7}$$

Könnyen ellenőrizhető a nyom egyik nevezetes tulajdonsága, miszerint, bármely $\mathbf{A}, \mathbf{B} \in \mathcal{M}_n(\mathbb{K})$ esetén

$$\operatorname{Tr}(\mathbf{A}\,\mathbf{B}) = \operatorname{Tr}(\mathbf{B}\,\mathbf{A}).$$

Egy lineáris leképezés sajátértékeinek és sajátvektorainak vizsgálatát úgy vezettük fel, hogy rögzítettünk a vektortérben egy bázist és a leképezés ezen bázisban felírt mátrixával definiáltuk a sajátértéket és sajátvektort. Joggal merül fel a kérdés, hogy az így kiszámolt sajátértékek és sajátvektorok hogyan függnek a bázistól. Megmutatjuk, hogy a sajátértékek függetlenek a bázistól. Igaz ugyanis a következő tétel.

Tétel. *Hasonló mátrixok karakterisztikus polinomja megegyezik.*

Bizonyítás. Tegyük fel, hogy $\mathbf{A} \simeq \mathbf{B}$, azaz létezik egy invertálható $\mathbf{S}$ mátrix, mellyel $\mathbf{B} = \mathbf{S}\,\mathbf{A}\,\mathbf{S}^{-1}$. Jelölje $p_{\mathbf{A}}(\lambda)$ illetve $p_{\mathbf{B}}(\lambda)$ $\mathbf{A}$ illetve $\mathbf{B}$ karakterisztikus polinomját. Ekkor

$$\begin{aligned} p_{\mathbf{B}}(\lambda) &= \det(\mathbf{B} - \lambda\,\mathbf{I}) = \det(\mathbf{S}\,\mathbf{A}\,\mathbf{S}^{-1} - \lambda\,\mathbf{S}\,\mathbf{I}\,\mathbf{S}^{-1}) = \det(\mathbf{S}(\mathbf{A} - \lambda\,\mathbf{I})\,\mathbf{S}^{-1}) \\ &= \det \mathbf{S} \det(\mathbf{A} - \lambda\,\mathbf{I}) \det \mathbf{S}^{-1} = \det(\mathbf{S}\,\mathbf{S}^{-1}) \det(\mathbf{A} - \lambda\,\mathbf{I}) \\ &= \det \mathbf{I} \det(\mathbf{A} - \lambda\,\mathbf{I}) = \det(\mathbf{A} - \lambda\,\mathbf{I}) = p_{\mathbf{A}}(\lambda), \end{aligned}$$

ahol felhasználtuk a determinánsok szorzattételét, miszerint $\det(\mathbf{A}\,\mathbf{B}) = \det \mathbf{A} \det \mathbf{B}$. $\square$

Miután a sajátértékek a karakterisztikus polinom gyökei, ezért a fenti tételből egyből következik, hogy hasonló mátrixok sajátértékei megegyeznek. Miután a báziscsere transzformáció éppen egy hasonlósági relációt jelent, azért megállapíthatjuk, hogy egy lineáris transzformáció sajátértékeit mindegy, hogy milyen bázisban számolom ki, ugyanazok lesznek. Ezt úgy szokás szépen mondani, hogy **a sajátértékek invariánsak a báziscsere transzformációra nézve**. Ezt a megfigyelést és néhány következményét külön is kimondjuk.

Következény 8. *Ha $\mathbf{A} \simeq \mathbf{B}$, akkor $\mathbf{A}$ és $\mathbf{B}$ sajátértékei megegyeznek, következésképpen*

$$\operatorname{Tr}\mathbf{A} = \operatorname{Tr}\mathbf{B}, \quad \textit{és} \quad \det\mathbf{A} = \det\mathbf{B}\,.$$

Az is világos, hogy a sajátvektorok a különböző bázisokban különbözőek lesznek, de egymásba transzformálódnak a báziscsere transzformáció során. Vagyis, ha $\mathbf{A}\,\mathbf{v} = \lambda\,\mathbf{v}$, és $\mathbf{B} = \mathbf{S}\,\mathbf{A}\,\mathbf{S}^{-1}$, akkor

$$\mathbf{B}(\mathbf{S}\,\mathbf{v}) = \mathbf{S}\,\mathbf{A}\,\mathbf{S}^{-1}(\mathbf{S}\,\mathbf{v}) = \mathbf{S}\,\mathbf{A}\,\mathbf{v} = \mathbf{S}(\lambda\,\mathbf{v}) = \lambda(\mathbf{S}\,\mathbf{v}),$$

vagyis $\mathbf{B}$ λ-hoz tartozó sajátvektora $\mathbf{S}\,\mathbf{v}$ lesz.

A következőkben megvizsgáljuk, hogyan érdemes bázisokat felvenni úgy, hogy a számolásaink könnyebbekké váljanak.

4. Mátrixok főtengelyre való transzformálása (diagonalizálás)

Először megmutatjuk a sajátvektoroknak egy nagyon fontos tulajdonságát.

Tétel. *Legyenek $\mathbf{A} \in \mathcal{M}_n(\mathbb{K})$ mátrixnak $\lambda_1, \lambda_2, \ldots, \lambda_k$ különböző sajátértékei, rendre $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ sajátvektorokkal. Ekkor a $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ vektorrendszer lineárisan független.*

Bizonyítás. Teljes indukciót végzünk k-ban. A $k = 1$ eset triviális, hiszen csak egy sajátvektorunk van. Tegyük fel, hogy k vektorra már beláttuk az állítást. Nézzük most, mi történik $k+1$ vektor esetén. Indirekt bizonyítunk: tegyük fel, hogy a $\lambda_1, \lambda_2, \ldots, \lambda_k, \lambda_{k+1}$ különböző sajátértékek, ám a hozzájuk tartozó $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k, \mathbf{v}_{k+1}\}$ sajátvektorok lineárisan összefüggőek. Ez csak úgy lehet, ha

$$\mathbf{v}_{k+1} \in \operatorname{Span}(\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}), \quad \text{azaz} \quad \mathbf{v}_{k+1} = \sum_{i=1}^{k} a_i\,\mathbf{v}_i \quad (*),$$

valamilyen $a_i \in \mathbb{K}$ együtthatókkal. Ha a $(*)$ egyenlet mindkét oldalára hattatom az $\mathbf{A}$ mátrixot, akkor a linearitás miatt

$$\mathbf{A}\,\mathbf{v}_{k+1} = \sum_{i=1}^{k} a_i\,\mathbf{A}\,\mathbf{v}_i\,.$$

Míg, ha $(*)$ egyenlet mindkét oldalát megszorzom λ_{k+1}-gyel, akkor a

$$\lambda_{k+1}\,\mathbf{v}_{k+1} = \sum_{i=1}^{k} a_i\lambda_{k+1}\,\mathbf{v}_i$$

egyenlethez jutunk. Vegyük ez utóbbi két egyenlet különbségét, kihasználva, hogy $\mathbf{A}\,\mathbf{v}_{k+1} = \lambda_{k+1}\,\mathbf{v}_{k+1}$:

$$\begin{aligned} \mathbf{0} &= \mathbf{A}\,\mathbf{v}_{k+1} - \lambda_{k+1}\,\mathbf{v}_{k+1} = \sum_{i=1}^{k} a_i\,\mathbf{A}\,\mathbf{v}_i - \sum_{i=1}^{k} a_i \lambda_{k+1}\,\mathbf{v}_i \\ &= \sum_{i=1}^{k} a_i(\mathbf{A}\,\mathbf{v}_i - \lambda_{k+1}\,\mathbf{v}_i) = \sum_{i=1}^{k} a_i(\lambda_i - \lambda_{k+1})\,\mathbf{v}_i, \end{aligned}$$

ahol az utolsó lépésben kihasználtuk, hogy $\mathbf{A}\,\mathbf{v}_i = \lambda_i\,\mathbf{v}_i$. De az indukciós feltevésünk alapján a $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ vektorok lineárisan függetlenek, vagyis a fenti lineárkombináció csak úgy lehet $\mathbf{0}$, ha az összes együttható 0, azaz

$$a_i(\lambda_i - \lambda_{k+1}) = 0 \quad \text{minden } i = 1, \ldots, k \text{ esetén.}$$

Viszont a feltételeink alapján, a sajátértékek mind különbözőek, vagyis $\lambda_i \neq \lambda_{k+1}$, ami csak úgy teljesülhet, hogy $a_i = 0$ minden $i = 1, \ldots, k$ esetén, de ekkor $\mathbf{v}_{k+1} = \mathbf{0}$ következne, ami ellentmondás. □

A tételből kiolvasható, hogy ha $\mathbf{A} \in \mathcal{M}_n(\mathbb{K})$-nak n darab különböző sajátértéke van (vagyis minden sajátérték egyszeres algebrai multiplicitású), akkor van n darab lineárisan független sajátvektora $\mathbb{K}^n$-ben, vagyis $\mathbf{A}$ mátrix sajátvektorai bázist alkotnak. Ha valamelyik sajátérték többszörös $k \neq 1$ algebrai multipliplicitású, és ezen sajátérték algebrai multiplicitása megegyezik a geometriai multiplicitásával, akkor ezen a k dimenziós sajátaltéren szintén választhatunk egy k elemű bázist, és ezek a bázisvektorok természetesen szintén sajátvektorai lesznek $\mathbf{A}$-nak. Ha $\mathbf{A}$-nak van sajátvektoraiból álló bázisa, akkor azt **A sajátbázisának** szokás hívni. Szemléltessük ezt egy példán!

Példa 9. Legyen

$$\mathbf{A} = \begin{pmatrix} 2 & -1 & -1 \\ 3 & -2 & -3 \\ -1 & 1 & 2 \end{pmatrix}.$$

A sajátérték egyenlet ekkor

$$\det(\mathbf{A} - \lambda\,\mathbf{I}) = \det \begin{pmatrix} 2-\lambda & -1 & -1 \\ 3 & -2-\lambda & -3 \\ -1 & 1 & 2-\lambda \end{pmatrix} = -\lambda(\lambda-1)^2 = 0,$$

vagyis a sajátértékek: $\lambda_1 = 0$ (egyszeres algebrai multiplicitás), $\lambda_2 = \lambda_3 = 1$ (kétszeres algebrai multiplicitás). Számoljuk ki az egyes sajátértékekhez tartozó sajátvektorokat!

- $\underline{\lambda_1 = 0}$:
  A megoldandó egyenlet:
$$(\mathbf{A} - \lambda_1 \mathbf{I})\,\mathbf{v} = \begin{pmatrix} 2 & -1 & -1 \\ 3 & -2 & -3 \\ -1 & 1 & 2 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$
  ahonnan Gauss-elimináció után azt kapjuk, hogy $v_1 = p$ szabad paraméter választás után, $v_2 = 3p$, $v_3 = -p$. Innen a $p = 1$ választással, a $\lambda_1 = 0$ sajátértékhez tartozó $\mathbf{v}_1$ sajátvektor:
$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 3 \\ -1 \end{pmatrix}.$$

- $\underline{\lambda_2 = \lambda_3 = 1}$:
  A megoldandó egyenlet ezúttal:
$$(\mathbf{A} - \lambda_2 \mathbf{I})\,\mathbf{v} = \begin{pmatrix} 1 & -1 & -1 \\ 3 & -3 & -3 \\ -1 & 1 & 1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$
  ahonnan Gauss-elimináció után azt kapjuk, hogy most két szabad paraméterünk van, például $v_2 = p$ és $v_3 = q$ választás után, $v_1 = p + q$. Az 1 sajátértékhez tartozó sajátvektorok ezek szerint a következő alakúak:
$$\begin{pmatrix} p+q \\ p \\ p \end{pmatrix} = p \cdot \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + q \cdot \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$
  A jobb oldalon szereplő két vektor lineárisan független, így látható, hogy az 1 sajátértékhez most két lineárisan független sajátvektort is fel tudunk írni $p = q = 1$ választással, például:
$$\mathbf{v}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix},$$
  ami mutatja, hogy az 1 sajátérték geometriai multiplicitása 2. Összefoglalva azt látjuk tehát, hogy a mátrixnak van sajátbázisa $\mathbb{R}^3$-ben (vagy $\mathbb{C}^3$-ben), melynek vektorai:
$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 3 \\ -1 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

Tegyük fel, hogy az $\mathbf{A}$ mátrixnak van a sajátvektoraiból álló bázisa, legyen ez $\alpha = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$, ahol tehát fennáll, hogy

$$\mathbf{A}\,\mathbf{v}_i = \lambda_i\,\mathbf{v}_i, \quad \text{minden } i = 1, 2, \ldots, n \text{ esetén.}$$

Nézzük meg, hogy mi lesz ebben a bázisban ($\mathbf{A}$ sajátbázisában) $\mathbf{A}$ mátrixa, melyet most $\mathbf{A}_\alpha$-val jelölünk. Valójában itt az (α, α) bázispárra vonatkozó mátrixot kell felírni, amire ismerjük a módszert: $\mathbf{A}$-nak az egyes bázisvektorokon vett hatását kell a sajátbázisban kiszámolnunk és oszlopként egy mátrixba írni. Mivel azonban $\mathbf{A}\,\mathbf{v}_i = \lambda_i\,\mathbf{v}_i$, ezért

$$\begin{aligned}
\mathbf{A}_\alpha &= \left([\mathbf{A}\,\mathbf{v}_1]^\alpha \;\middle|\; [\mathbf{A}\,\mathbf{v}_2]^\alpha \;\middle|\; \cdots \;\middle|\; [\mathbf{A}\,\mathbf{v}_n]^\alpha \right) \\
&= \left([\lambda_1\,\mathbf{v}_1]^\alpha \;\middle|\; [\lambda_2\,\mathbf{v}_2]^\alpha \;\middle|\; \cdots \;\middle|\; [\lambda_n\,\mathbf{v}_n]^\alpha \right) \\
&= \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ . & . & . & \cdots & . \\ 0 & 0 & 0 & \cdots & \lambda_n \end{pmatrix} \equiv \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n),
\end{aligned}$$

vagyis azt kapjuk, hogy egy mátrix a sajátbázisában felírva (amennyiben létezik neki ilyen), mindig diagonális, ráadásul úgy, hogy a főátlójában (diagonálisában) éppen a sajátértékek állnak, olyan sorrendben felsorolva, ahogy a megfelelő sajátvektorok a bázisban rendezve voltak. Erre utal az utolsó sorban használt jelölés. Láttuk tehát, hogy ez az úgynevezett főtengelyre vett transzformáció egy báziscsere transzformációval (a sajátbázisra való áttéréssel) valósult meg. Ennek megfelelően élünk az alábbi definícióval.

Definíció. *Az $\mathbf{A} \in \mathcal{M}_n(\mathbb{K})$ mátrix* ***diagonalizálható*** *vagy* ***főtengelyre transzformálható*** *$\mathbb{K}$ felett, ha, létezik egy olyan invertálható $\mathbf{S} \in \mathcal{M}_n(\mathbb{K})$ mátrix, melyre $\mathbf{S}^{-1}\,\mathbf{A}\,\mathbf{S}$ diagonális.*

Ha most vissza szeretnénk a mátrixunkat transzformálni a standard bázisba, akkor a szokásos báziscserét kell végrehajtanunk:

$$\mathbf{A} = \mathbf{S}\,\mathbf{A}_\alpha\,\mathbf{S}^{-1} = \mathbf{S} \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ . & . & . & \cdots & . \\ 0 & 0 & 0 & \cdots & \lambda_n \end{pmatrix} \mathbf{S}^{-1},$$

ahol a szokott módon $\mathbf{S}$ azon mátrix, melynek oszlopaiban az eredeti bázisvektorok, vagyis a sajátvektorok állnak oszlopvektorként:

$$\mathbf{S} = \left(\mathbf{v}_1 \;\middle|\; \mathbf{v}_2 \;\; \ldots \;\middle|\; \mathbf{v}_n \right).$$

Az eddigiek alapján és még egy kis erőfeszítéssel könnyen bizonyítható az alábbi

Tétel. *Az* $\mathbf{A} \in \mathcal{M}_n(\mathbb{K})$ *mátrix akkor és csak akkor diagonalizálható* $\mathbb{K}$ *felett, ha létezik* $\mathbb{K}^n$*-ben az* $\mathbf{A}$*-nak sajátbázisa.*

Ez a tétel megfogalmazható a sajátértékek multiplicitásaival is. Láttuk ugyanis, hogy ha egy sajátérték algebrai multiplicitása k, akkor csak akkor tartozik hozzá k darab sajátvektor, ha a hozzá tartozó sajátaltéren találunk k darab lineárisan független sajátvektort, vagyis a geometriai multiplicitása is k. Természetesen annak is teljesülnie kell, hogy algebrai multiplicitásokkal együtt legyen annyi sajátértékünk, mint ahány a tér dimenziója. Jelöljük a $\lambda \in \mathrm{Spec}(\mathbf{A})$ sajátérték algebrai multiplicitását $a(\lambda)$-val, a geometriai multiplicitását $g(\lambda)$-val. Igaz tehát az alábbi

Tétel. *Az* $\mathbf{A} \in \mathcal{M}_n(\mathbb{K})$ *mátrix pontosan akkor diagonalizálható* $\mathbb{K}$ *felett, ha*

$$\sum_{\lambda \in \mathrm{Spec}(\mathbf{A})} a(\lambda) = n, \quad \textit{és } a(\lambda) = g(\lambda) \textit{ minden } \lambda \in \mathrm{Spec}(\mathbf{A}) \textit{ esetén.}$$

Nézzünk néhány szemléltető példát, a részletszámításokat végezzük el magunk!

Példa 10. Legyen

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 8 & 1 \end{pmatrix}.$$

A karakterisztikus egyenlet

$$\det(\mathbf{A} - \lambda\,\mathbf{I}) = \det \begin{pmatrix} 1-\lambda & 2 \\ 8 & 1-\lambda \end{pmatrix} = (1-\lambda)^2 - 16 = 0,$$

ahonnan a sajátértékek $\lambda_1 = 5$ és $\lambda_2 = -3$, melyekhez tartozó sajátvektorok

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad \text{illetve} \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}.$$

Ekkor tehát

$$\mathbf{S} = \begin{pmatrix} 1 & 1 \\ 2 & -2 \end{pmatrix},$$

amivel könnyen ellenőrizhető, hogy

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 8 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} 5 & 0 \\ 0 & -3 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 2 & -2 \end{pmatrix}^{-1}.$$

Példa 11. Legyen

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

A karakterisztikus egyenlet

$$\det(\mathbf{A} - \lambda\,\mathbf{I}) = \det\begin{pmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{pmatrix} = (1-\lambda)^2 = 0,$$

ahonnan a sajátérték $\lambda_1 = \lambda_2 = 1$ kétszeres algebrai multiplicitású. A sajátvektor megtalálásához most az

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

egyenletet kell megoldanunk, melynek a megoldása $v_1 = p$, $v_2 = 0$, ahol p szabad paraméter. A $p = 1$ választással tehát egyetlen sajátvektorunk van:

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Azt kapjuk tehát, hogy az 1 sajátérték geometriai multiplicitása 1, ami kevesebb, mint az algebrai multiplicitása, vagyis a mátrixunk nem diagonalizálható.

Miért jó, ha egy mátrix főtengelyre transzformálható. Az eddigi számolásainknál tapasztalhattuk, hogy a mátrixszorzás meglehetősen műveletigényes. Két $n \times n$-es mátrix összeszorzásakor például $2n^3$ műveletet kell elvégeznünk. Sokszor előfordul, hogy egy mátrixnak sokadik hatványát kell kiszámolnunk, például szükségünk van az $\mathbf{A}^N = \mathbf{A}\,\mathbf{A}\cdots\mathbf{A}$ szorzatra. Amennyiben $\mathbf{A}$ diagonalizálható, azaz létezik $\mathbf{S}$ invertálható mátrix, mellyel $\mathbf{A} = \mathbf{S}\,\mathbf{D}\,\mathbf{S}^{-1}$, ahol $\mathbf{D} = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ diagonális mátrix, akkor ez írható az alábbi módon:

$$\begin{aligned}
\mathbf{A}^N &= \underbrace{(\mathbf{S}\,\mathbf{D}\,\mathbf{S}^{-1})(\mathbf{S}\,\mathbf{D}\,\mathbf{S}^{-1})\ldots(\mathbf{S}\,\mathbf{D}\,\mathbf{S}^{-1})}_{N \text{ darab}} \\
&= \mathbf{S}\,\mathbf{D}\,\underbrace{(\mathbf{S}^{-1}\,\mathbf{S})}_{\mathbf{I}}\,\mathbf{D}\,\underbrace{(\mathbf{S}^{-1}\,\mathbf{S})}_{\mathbf{I}}\,\mathbf{D}\ldots\underbrace{(\mathbf{S}^{-1}\,\mathbf{S})}_{\mathbf{I}}\,\mathbf{D}\,\mathbf{S}^{-1} \\
&= \mathbf{S}\,\mathbf{D}^N\,\mathbf{S}^{-1}.
\end{aligned}$$

Könnyű azonban látni, hogy diagonális mátrixokat viszont élmény összeszorozni, hiszen csak az azonos helyen lévő diagonális elemeket kell összeszoroznunk, speciálisan például $\mathbf{D}^N = \mathrm{diag}(\lambda_1^N, \lambda_2^N, \ldots, \lambda_n^N)$. Így azt kaptuk tehát, hogy

$$\mathbf{A}^N = \mathbf{S}\begin{pmatrix} \lambda_1^N & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2^N & 0 & \cdots & 0 \\ . & . & . & \cdots & . \\ 0 & 0 & 0 & \cdots & \lambda_n^N \end{pmatrix}\mathbf{S}^{-1}.$$

Ez a megfigyelés nagymértékben általánosítható (később, mikor már tanultunk pár dolgot a Taylor-sorokról, indokolni is fogjuk, hogy miért igaz). Jelen esetben definícióként célszerű kimondanunk, hogy hogyan értelmezzük mátrixok függvényét.

Definíció. *Legyen* $\mathbf{A} \in \mathcal{M}_n(\mathbb{R})$ *egy diagonalizálható mátrix, azaz tegyük fel, hogy létezik egy* $\mathbf{S} \in \mathcal{M}_N(\mathbb{R})$ *invertálható mátrix, mellyel* $\mathbf{A} = \mathbf{S}\operatorname{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)\,\mathbf{S}^{-1}$. *Legyen* $f : \mathbb{R} \to \mathbb{R}$ *függvény, melynek az értelmezési tartománya tartalmazza* $\mathbf{A}$ *sajátértékeit. Ekkor értelmezzük az* f *függvényt az* $\mathbf{A}$ *helyen a következőképpen:*

$$f(\mathbf{A}) = \mathbf{S}\operatorname{diag}(f(\lambda_1), f(\lambda_2), \dots, f(\lambda_n))\,\mathbf{S}^{-1} = \mathbf{S}\begin{pmatrix} f(\lambda_1) & 0 & 0 & \cdots & 0 \\ 0 & f(\lambda_2) & 0 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & \cdots & f(\lambda_n) \end{pmatrix}\mathbf{S}^{-1}.$$

Példa 12. Számoljuk ki $e^{\mathbf{A}}$-t a 10. Példában szereplő mátrixra.

$$e^{\mathbf{A}} = e^{\begin{pmatrix} 1 & 2 \\ 8 & 1 \end{pmatrix}} = \begin{pmatrix} 1 & 1 \\ 2 & -2 \end{pmatrix}\begin{pmatrix} e^5 & 0 \\ 0 & e^{-3} \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 2 & -2 \end{pmatrix}^{-1}.$$