A nulltér altér

3.9. definíció (Altér). *Az* $\mathbb{R}^n$ *tér vektorainak olyan nem üres részhalmazát, mely zárt a vektorok skalárral való szorzásának és a vektorok összeadásának műveletére, az* $\mathbb{R}^n$ alterének *nevezzük. Képlettel kifejezve: az* $\mathcal{A} \subseteq \mathbb{R}^n$ *vektorhalmaz* $\mathbb{R}^n$ altere, *ha*

1. $\mathbf{u}, \mathbf{v} \in \mathcal{A}$ *esetén* $\mathbf{u} + \mathbf{v} \in \mathcal{A}$,
2. $\mathbf{u} \in \mathcal{A}$, *és* $c \in \mathbb{R}$ *esetén* $c\mathbf{u} \in \mathcal{A}$.

3.12. állítás (Megoldások altere). *Egy n-ismeretlenes* homogén *lineáris egyenletrendszer megoldáshalmaza alteret alkot* $\mathbb{R}^n$*-ben.*

Ezzel ekvivalens, hogy:

3.8. állítás (Megoldások lineáris kombinációja). *Egy homogén lineáris egyenletrendszer megoldásainak bármely lineáris kombinációja is megoldás.*

Bizonyítás. Elég az állítást két megoldásra bizonyítani. Jelölje $\mathbf{a}_1$, $\mathbf{a}_2, \ldots, \mathbf{a}_n$ az egyenletrendszer együtthatómátrixának oszlopvektorait. Legyen $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ és $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ két tetszőleges megoldás, azaz

$$\begin{aligned} \mathbf{a}_1 x_1 + \mathbf{a}_2 x_2 + \ldots + \mathbf{a}_n x_n &= \mathbf{0} \\ \mathbf{a}_1 y_1 + \mathbf{a}_2 y_2 + \ldots + \mathbf{a}_n y_n &= \mathbf{0}, \end{aligned}$$

és c, d legyen két tetszőleges skalár. Megmutatjuk, hogy ekkor $c\mathbf{x} + d\mathbf{y}$ is megoldás, ugyanis

$$\begin{aligned} &\mathbf{a}_1(cx_1 + dy_1) + \mathbf{a}_2(cx_2 + dy_2) + \ldots + \mathbf{a}_n(cx_n + dy_n) \\ =&(c\mathbf{a}_1 x_1 + d\mathbf{a}_1 y_1) + (c\mathbf{a}_2 x_2 + d\mathbf{a}_2 y_2) + \ldots + (c\mathbf{a}_n x_n + d\mathbf{a}_n y_n) \\ =&c(\mathbf{a}_1 x_1 + \mathbf{a}_2 x_2 + \ldots + \mathbf{a}_n x_n) + d(\mathbf{a}_1 y_1 + \mathbf{a}_2 y_2 + \ldots + \mathbf{a}_n y_n) \\ =&\mathbf{0} + \mathbf{0} = \mathbf{0}. \end{aligned}$$

azaz $c\mathbf{x} + d\mathbf{y}$ is megoldás. Ez bizonyítja állításunkat.

E bizonyítás az oszlopmodellre épült, de hasonlóan egyszerű bizonyítás adható a sormmodellben is (ld. 3.9. feladat). □

Minden véges dimenziós lineáris leképezéshez létezik egy mátrix, ami azt a leképezést generálja

7.9. TÉTEL (Az $\mathbb{R}^n \to \mathbb{R}^m$ LINEÁRIS LEKÉPEZÉSEK MÁTRIXLEKÉPEZÉSEK). *Legyen $A : \mathbb{R}^n \to \mathbb{R}^m$ egy tetszőleges függvény. Az A pontosan akkor lineáris, ha létezik egy olyan $m \times n$-es $\mathbf{A}$ mátrix, hogy az A függvény megegyezik az $\mathbf{x} \mapsto \mathbf{Ax}$ leképezéssel. Ekkor*

$$\mathbf{A} = [A\mathbf{e}_1|A\mathbf{e}_2|\ldots|A\mathbf{e}_n],$$

ahol $\mathbf{e}_i$ az i-edik standard egységvektor ($i = 1, 2, \ldots, n$).

BIZONYÍTÁS. Minden mátrixleképezés lineáris, ez bizonyítja az állítás egyik felét. A állítás másik felének bizonyításához tekintsük $\mathbb{R}^n$ standard bázisát és az $A\mathbf{e}_i$ vektorokból képzett

$$\mathbf{A} = [A\mathbf{e}_1|A\mathbf{e}_2|\ldots|A\mathbf{e}_n] \tag{7.2}$$

mátrixot, valamint legyen $\mathbf{x} \in \mathbb{R}^n$ egy tetszőleges vektor. Ha A lineáris leképezés, azaz megőrzi a lineáris kombinációt, akkor

$$\begin{aligned}
A\mathbf{x} &= A(x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \ldots + x_n\mathbf{e}_n) \\
&= x_1 A\mathbf{e}_1 + x_2 A\mathbf{e}_2 + \ldots + x_n A\mathbf{e}_n \\
&= \begin{bmatrix} A\mathbf{e}_1 & A\mathbf{e}_2 & \ldots & A\mathbf{e}_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \\
&= \mathbf{Ax}
\end{aligned}$$

Tehát valóban létezik olyan $\mathbf{A}$ mátrix, hogy $A\mathbf{x} = \mathbf{Ax}$. Ráadásul ilyen mátrix csak ez az egy van, mert bármely $\mathbf{e}_i$ bázisvektorra és bármely $\mathbf{A}$ mátrixra $\mathbf{Ae}_i = \mathbf{A}_{*i}$, tehát az $\mathbf{A}_{*i}$ oszlopvektor csak $A\mathbf{e}_i$ lehet. □

Merőleges vetítés $\mathbb{R}^n$ egy alterére

7.40. TÉTEL (MERŐLEGES VETÍTÉS MÁTRIXAI). *Egy* $\mathbf{P}$ *mátrix pontosan akkor merőleges vetítés mátrixa, ha* $\mathbf{P} = \mathbf{P}^\mathsf{T} = \mathbf{P}^2$.

BIZONYÍTÁS. A $\mathbf{P} = \mathbf{P}^2$ feltétel szükségessége szemléletesen világos, hisz minden P lineáris leképezés, mely az egész $\mathbb{R}^n$ teret egy altérre – nevezetesen $\operatorname{Im} P$-re – vetíti, az altér vektorait helyben hagyja. Tehát $P^2\mathbf{x} = P\mathbf{x}$ minden $\mathbf{x}$-re fennáll, így ennek az összefüggésnek P minden mátrixára is igaznak kell lennie.

($\Longrightarrow$) Tegyük fel, hogy $\mathbf{P}$ egy P merőleges vetítés mátrixa $\mathbb{R}^n$ standard bázisában. Tekintsük $\operatorname{Im}(P) = \mathcal{O}(\mathbf{P})$ egy tetszőleges bázisát, és legyen $\mathbf{A}$ az a mátrix, melynek e bázis elemei az oszlopai. A 7.38. tétel szerint ekkor $\mathbf{P} = \mathbf{A}(\mathbf{A}^\mathsf{T}\mathbf{A})^{-1}\mathbf{A}^\mathsf{T}$. Erre viszont könnyen ellenőrizhető, a tételbeli feltétel.

$$\begin{aligned}\mathbf{P}^2 &= \left(\mathbf{A}(\mathbf{A}^\mathsf{T}\mathbf{A})^{-1}\mathbf{A}^\mathsf{T}\right)^2 = \mathbf{A}(\mathbf{A}^\mathsf{T}\mathbf{A})^{-1}\mathbf{A}^\mathsf{T}\mathbf{A}(\mathbf{A}^\mathsf{T}\mathbf{A})^{-1}\mathbf{A}^\mathsf{T}\\ &= \mathbf{A}(\mathbf{A}^\mathsf{T}\mathbf{A})^{-1}\mathbf{A}^\mathsf{T} = \mathbf{P},\end{aligned}$$

másrészt

$$\begin{aligned}\mathbf{P}^\mathsf{T} &= \left(\mathbf{A}(\mathbf{A}^\mathsf{T}\mathbf{A})^{-1}\mathbf{A}^\mathsf{T}\right)^\mathsf{T} = \mathbf{A}\left((\mathbf{A}^\mathsf{T}\mathbf{A})^{-1}\right)^\mathsf{T}\mathbf{A}^\mathsf{T}\\ &= \mathbf{A}(\mathbf{A}^\mathsf{T}\mathbf{A})^{-1}\mathbf{A}^\mathsf{T} = \mathbf{P}.\end{aligned}$$

($\Longleftarrow$) Tegyük fel, hogy $\mathbf{P} = \mathbf{P}^\mathsf{T} = \mathbf{P}^2$. Megmutatjuk, hogy $\mathbf{P}$ az $\mathcal{O}(\mathbf{P})$-re való merőleges vetítés mátrixa. Ehhez elég megmutatnunk, hogy az $\mathbf{x} - \mathbf{P}\mathbf{x}$ vektor merőleges $\mathcal{O}(\mathbf{P})$-re bármely $\mathbf{x}$ vektor esetén. A $\mathbf{P}^2 = \mathbf{P}$ feltétel miatt $\mathbf{P}(\mathbf{x} - \mathbf{P}\mathbf{x}) = \mathbf{P}\mathbf{x} - \mathbf{P}^2\mathbf{x} = \mathbf{0}$, tehát $\mathbf{x} - \mathbf{P}\mathbf{x} \in \mathcal{N}(\mathbf{P})$, de $\mathbf{P} = \mathbf{P}^\mathsf{T}$, így $\mathbf{x} - \mathbf{P}\mathbf{x} \in \mathcal{N}(\mathbf{P}^\mathsf{T})$. Ez épp azt jelenti, hogy $\mathbf{x} - \mathbf{P}\mathbf{x}$ merőleges $\mathcal{O}(\mathbf{P})$-re, és ezt akartuk belátni. □

$\boldsymbol{rang(A) + \dim(\mathcal{N}(A)) = n}$

3.36. TÉTEL (DIMENZIÓTÉTEL). *Bármely* $\mathbf{A} \in \mathbb{R}^{m\times n}$ *mátrix esetén a sortér dimenziójának és a nulltér dimenziójának összege* n. *Képlettel:*

$$\dim(\mathcal{S}(\mathbf{A})) + \dim(\mathcal{N}(\mathbf{A})) = n.$$

Bizonyítás. A mátrix sorterének dimenziója megegyezik a mátrix rangjával, azaz az $[\mathbf{A}|\mathbf{0}]$ mátrixú egyenletrendszerben a kötött változók számával. Megmutatjuk, hogy a nulltér dimenziója megegyezik a szabad változók számával, így e két szám összege valóban n, ami bizonyítja az állítást (ld. még a 3.4. állítást).

Elég tehát megmutatnunk, hogy egy homogén lineáris egyenletrendszer redukált lépcsős alakkal előállított megoldásában a szabad változók száma megegyezik a nulltérből kiválasztható bázis elemszámával. Először lássunk egy ilyen megoldást konkrétan. Például a 2.37. példabeli homogén lineáris egyenletrendszer megoldása

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -2s - \frac{3}{2}t - u \\ s \\ -\frac{1}{2}t \\ t \\ u \end{bmatrix} = s \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} -\frac{3}{2} \\ 0 \\ -\frac{1}{2} \\ 1 \\ 0 \end{bmatrix} + u \begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix},$$

ahol $x_2 = s$, $x_4 = t$ és $x_5 = u$ a három szabad változó. A nullteret kifeszítő három vektor közül az elsőben $x_2 = 1$, de az összes többiben $x_2 = 0$, így az első vektor független a többitől. Hasonlóképp általában is igaz, hogy a redukált lépcsős alakból való származtatás következtében a nullteret kifeszítő minden megoldásvektorban az összes szabad változóhoz tartozó koordináta 0, azt az egy koordinátát kivéve, amelyikhez a vektor tartozik. Így viszont mindegyik vektor független a többitől, vagyis e vektorok függetlenek, és mivel kifeszítik a nullteret, számuk megadja a nulltér dimenzióját. □

$\mathbf{S}(\mathbf{A}) \perp \mathcal{N}(\mathbf{A})$

3.38. Állítás (A sortér és a nulltér merőlegessége). *A valós* $\mathbf{A}$ *mátrix sorterének bármely* $\mathbf{s}$ *vektora és nullterének tetszőleges* $\mathbf{x}$ *vektora merőleges egymásra, azaz* $\mathbf{s} \cdot \mathbf{x} = 0$.

Bizonyítás. A sortér minden vektora az $\mathbf{A}$ sorvektorainak valamely $c_1, \dots, c_m$ skalárokkal vett lineáris kombinációja. Ezt felhasználva

$$\begin{aligned}\mathbf{s}\cdot\mathbf{x} &= (c_1\mathbf{a}_{1*} + c_2\mathbf{a}_{2*} + \cdots + c_m\mathbf{a}_{m*})\cdot\mathbf{x}\\ &= c_1\mathbf{a}_{1*}\cdot\mathbf{x} + c_2\mathbf{a}_{2*}\cdot\mathbf{x} + \cdots + c_m\mathbf{a}_{m*}\cdot\mathbf{x}\\ &= c_1 0 + c_2 0 + \cdots + c_m 0 = 0.\end{aligned}$$

□

Megoldható egyenletrendszer egyértelmű megoldása a sortérben és e megoldás abszolút értékének minimalizálása

3.41. tétel (Lineáris egyenletrendszer megoldásai). *Minden valós együtthatós megoldható (konzisztens) lineáris egyenletrendszerre igazak a következő állítások:*

a) egyetlen megoldása esik az együtthatómátrix sorterébe;

b) a sortérbe eső megoldás az összes megoldás közül a legkisebb abszolút értékű;

c) az összes megoldás előáll úgy, hogy a sortérbe eső megoldáshoz hozzáadjuk a homogén rész összes megoldását.

Bizonyítás. A tétel a homogén lineáris egyenletrendszerekre semmitmondó, hisz ekkor a megoldások a nullteret adják, és mivel annak metszete a sortérrel csak a nullvektorból áll, ezért csak a nullvektor esik a sortérbe, mely természetesen a legkisebb abszolút értékű megoldás. Ráadásul a nullvektort hozzáadva a nulltérhez, valóban a nullteret kapjuk, vagyis az összes megoldások terét. Így ezután csak az inhomogén esettel foglalkozunk.

a) Tegyük fel, hogy $\mathbf{x}_1$ és $\mathbf{x}_2$ két megoldása az $[\mathbf{A}|\mathbf{b}]$ mátrixú egyenletrendszernek, és mindkettő a sortérbe esik. Az i-edik egyenlet alakja $\mathbf{a}_{i*}\cdot\mathbf{x} = b_i$, így $\mathbf{a}_{i*}\cdot\mathbf{x}_1 = b_i$ és $\mathbf{a}_{i*}\cdot\mathbf{x}_2 = b_i$ is fönnáll minden $i = 1, 2, \dots m$ értékre. A két megoldás különbsége is a sortérbe esik, hisz sortérbeli vektorok lineáris kombinációja a sortérbe esik. Ekkor viszont minden i esetén

$$\mathbf{a}_{i*}\cdot(\mathbf{x}_1 - \mathbf{x}_2) = b_i - b_i = 0,$$

vagyis $\mathbf{x}_1 - \mathbf{x}_2$ megoldása a homogén egyenletrendszernek, tehát a nulltérbe esik. Annak metszete a sortérrel csak a nullvektort tartalmazza, így $\mathbf{x}_1 - \mathbf{x}_2 = \mathbf{0}$, vagyis $\mathbf{x}_1 = \mathbf{x}_2$.

Meg kell még mutatnunk, hogy mindig van a sortérbe eső megoldás. Legyen $\mathbf{x}$ egy tetszőleges megoldás, és tekintsük az egyértelműen létező felbontását egy sortérbeli és egy nulltérbeli vektor összegére, azaz legyen

$$\mathbf{x} = \mathbf{x}_{\mathcal{S}} + \mathbf{x}_{\mathcal{N}}.$$

E megoldásvektort beírva az i-edik egyenletbe kapjuk, hogy

$$b_i = \mathbf{a}_{i*} \cdot \mathbf{x} = \mathbf{a}_{i*} \cdot (\mathbf{x}_{\mathcal{S}} + \mathbf{x}_{\mathcal{N}}) = \mathbf{a}_{i*} \cdot \mathbf{x}_{\mathcal{S}} + \mathbf{a}_{i*} \cdot \mathbf{x}_{\mathcal{N}} = \mathbf{a}_{i*} \cdot \mathbf{x}_{\mathcal{S}}.$$

Ez azt jelenti, hogy bármely megoldás sortérbeli összetevője is megoldása az egyenletrendszernek! Találtunk tehát egy megoldást a sortérben. Egyúttal azt is beláttuk, hogy az összes megoldás e sortérbeli megoldás és a homogén egy megoldásának összege. Az előző egyenlőségekből az is kiolvasható, hogy az $\mathbf{x}_{\mathcal{S}}$ megoldáshoz bármely nulltérbeli vektort adva az egyenletrendszer egy megoldását kapjuk, igazoltuk tehát a *c)* állítást is.

A sortér és a nulltér merőlegessége miatt az $\mathbf{x} = \mathbf{x}_{\mathcal{S}} + \mathbf{x}_{\mathcal{N}}$ felbontás vektorai merőlegesek, azaz $\mathbf{x}_{\mathcal{S}} \perp \mathbf{x}_{\mathcal{N}}$. Használhatjuk tehát Pithagorász-tételét:

$$\mathbf{x}^2 = \mathbf{x}_{\mathcal{S}}^2 + \mathbf{x}_{\mathcal{N}}^2 \geq \mathbf{x}_{\mathcal{S}}^2, \text{ azaz } |\mathbf{x}| \geq |\mathbf{x}_{\mathcal{S}}|.$$

Így tehát minden megoldás abszolút értéke nagyobb vagy egyenlő a sortérbeli megoldás abszolút értékénél, ami bizonyítja a *b)* állítást is. □

Páronként ortogonális vektorok függetlensége

7.63. TÉTEL (ORTOGONÁLIS VEKTOROK FÜGGETLENSÉGE). *Ha a nullvektortól különböző $\mathbf{a}_1$, $\mathbf{a}_2, \dots$, $\mathbf{a}_k$ vektorok páronként ortogonálisak, akkor függetlenek is.*

BIZONYÍTÁS. Tekintsük a $c_1\mathbf{a}_1 + \cdots + c_k\mathbf{a}_k = \mathbf{0}$ egyenletet. Be kell látnunk, hogy ez csak a $c_1 = \cdots = c_k = 0$ esetben áll fönn. Szorozzuk be az egyenlőség mindkét oldalát az $\mathbf{a}_i$ vektorral ($i = 1, 2, \dots, k$). Ekkor a jobb oldal 0, a bal oldalon pedig egy tag kivételével mindegyik 0 lesz:

$$\begin{aligned}(c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + \cdots + c_k\mathbf{a}_k) \cdot \mathbf{a}_i &= \mathbf{0} \cdot \mathbf{a}_i \\ c_i\mathbf{a}_i \cdot \mathbf{a}_i &= 0.\end{aligned}$$

Mivel $\mathbf{a}_i \cdot \mathbf{a}_i \neq 0$, ezért $c_i = 0$, és ez igaz minden i-re. □

Gram-Schmidt ortogonalizáció

7.77. TÉTEL (GRAM–SCHMIDT-ORTOGONALIZÁCIÓ). *Ha $\mathcal{A} = \{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_k\}$ egy független vektorrendszer, akkor létezik olyan ortogonális $\mathcal{V} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ vektorrendszer, hogy minden $i = 1, 2, \dots, k$ esetén*

$$\operatorname{span}(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_i) = \operatorname{span}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_i). \tag{7.20}$$

Az ortogonális $\mathcal{V}$ rendszerből a vektorok normálásával kapott

$$\left\{ \frac{\mathbf{v}_1}{|\mathbf{v}_1|}, \frac{\mathbf{v}_2}{|\mathbf{v}_2|}, \dots, \frac{\mathbf{v}_k}{|\mathbf{v}_k|} \right\}$$

rendszer ortonormált.

BIZONYÍTÁS. A $\operatorname{span}(\mathbf{a}_1) = \operatorname{span}(\mathbf{v}_1)$ összefüggés teljesül, ha

$$\mathbf{v}_1 = \mathbf{a}_1.$$

A $\operatorname{span}(\mathbf{a}_1, \mathbf{a}_2) = \operatorname{span}(\mathbf{v}_1, \mathbf{v}_2)$ teljesülése érdekében olyan $\mathbf{v}_2$ vektort kell választani, mely az $\mathbf{a}_1$ és $\mathbf{a}_2$ síkjában van, másrészt $\mathbf{v}_2$-nek merőlegesnek kell lennie $\mathbf{v}_1$-re. E feltételeket teljesíti az $\mathbf{a}_2$-nek a $\mathbf{v}_1$ által kifeszített altérre merőleges összetevője, azaz a

$$\mathbf{v}_2 = \mathbf{a}_2 - \left(\mathbf{a}_2 \cdot \frac{\mathbf{v}_1}{|\mathbf{v}_1|}\right) \frac{\mathbf{v}_1}{|\mathbf{v}_1|} = \mathbf{a}_2 - \frac{\mathbf{a}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1$$

vektor. Látható, hogy e vektor nem lehet a 0-vektor, hisz $\mathbf{v}_2 = 0$ esetén $\mathbf{a}_2 = \frac{\mathbf{a}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 = \frac{\mathbf{a}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{a}_1$ lenne, azaz $\mathbf{a}_1$ és $\mathbf{a}_2$ nem lenne független, ami ellentmond annak, hogy $\mathcal{A}$ független. Az előző képletekből látható, hogy $\mathbf{v}_1$ és $\mathbf{v}_2$ előállítható az $\mathbf{a}_1$ és $\mathbf{a}_2$ lineáris kombinációjaként, és viszont, így $\operatorname{span}(\mathbf{a}_1, \mathbf{a}_2) = \operatorname{span}(\mathbf{v}_1, \mathbf{v}_2)$ fönnáll. Az eljárás hasonlóképp folytatható. Ha már megkonstruáltuk $\mathbf{v}_i$-t, akkor a 7.64. tétel szerint kiszámoljuk az $\mathbf{a}_{i+1}$ vektornak a $\operatorname{span}(\frac{\mathbf{v}_1}{|\mathbf{v}_1|}, \frac{\mathbf{v}_2}{|\mathbf{v}_2|}, \dots, \frac{\mathbf{v}_i}{|\mathbf{v}_i|})$ altérre merőleges összetevőjét, és ezt választjuk $\mathbf{v}_{i+1}$-nek, azaz

$$\mathbf{v}_{i+1} = \mathbf{a}_{i+1} - \frac{\mathbf{a}_{i+1} \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{a}_{i+1} \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 - \dots - \frac{\mathbf{a}_{i+1} \cdot \mathbf{v}_i}{\mathbf{v}_i \cdot \mathbf{v}_i}\mathbf{v}_i$$

Könnyen látható, hogy $\mathbf{v}_{i+1} \neq \mathbf{0}$, mert ellenkező esetben $\mathcal{A}$ nem volna független. Látható az is, hogy $\mathbf{v}_{i+1}$ kifejezhető az $\mathbf{a}_1$, $\mathbf{a}_2, \dots, \mathbf{a}_{i+1}$ vektorok lineáris kombinációjaként, és $\mathbf{a}_{i+1}$ kifejezhető az $\mathbf{v}_1$, $\mathbf{v}_2, \dots, \mathbf{v}_{i+1}$ vektorok lineáris kombinációjaként, tehát a tétel kifeszített alterekre vonatkozó állítása is fennáll. □

A QR felbontás létezése

7.81. TÉTEL (QR-FELBONTÁS LÉTEZÉSE ÉS EGYÉRTELMŰSÉGE). *Bármely valós, teljes oszloprangú* $\mathbf{A}$ *mátrixnak létezik QR-felbontása, azaz létezik egy szemiortogonális* $\mathbf{Q}$ *mátrix és egy* $\mathbf{R}$ *felső háromszögmátrix pozitív főátlóbeli elemekkel, hogy* $\mathbf{A} = \mathbf{QR}$. *Az így kapott felbontás egyértelmű.*

BIZONYÍTÁS. A felbontás létezését a Gram–Schmidt-ortogonalizációra alapozva az előzőekben megmutattuk. Mivel $\mathbf{A} = \mathbf{QR}$, és $\mathbf{A}$ sorvektorai az $\mathbf{R}$ sorvektorainak lineáris kombinációi, ezért ha $\mathbf{R}$ rangja kisebb lenne k-nál, az $\mathbf{A}$ rangja is az lenne, de $\mathbf{A}$ rangja k. Beláttuk tehát, hogy $\mathbf{R}$ invertálható, és mivel felső háromszögmátrix, ezért főátlójában nem lehetnek 0-elemek.

Egyenletrendszer optimális megoldása QR felbontással

7.84. TÉTEL (LEGKISEBB NÉGYZETEK QR-FELBONTÁSSAL). *Legyen* $\mathbf{A}$ *egy teljes oszloprangú* $m \times n$*-es valós mátrix,* $\mathbf{A} = \mathbf{QR}$ *egy QR-felbontása, és legyen* $\mathbf{b}$ *egy* $\mathbb{R}^m$*-beli vektor. Ekkor az* $\mathbf{Ax} = \mathbf{b}$ *egyenletrendszer egyetlen optimális megoldása* $\hat{\mathbf{x}} = \mathbf{R}^{-1}\mathbf{Q}^\mathsf{T}\mathbf{b}$*, ami megkapható az*

$$\mathbf{R}\hat{\mathbf{x}} = \mathbf{Q}^\mathsf{T}\mathbf{b}$$

egyenletrendszerből egyszerű visszahelyettesítéssel is.

BIZONYÍTÁS. Az egyenletrendszer optimális megoldásáról szóló 7.45. tétel szerint az optimális megoldás a normálegyenletből megkapható. Eszerint

$$\begin{aligned}
\mathbf{A}^\mathsf{T}\mathbf{A}\hat{\mathbf{x}} &= \mathbf{A}^\mathsf{T}\mathbf{b} && \mathbf{A} = \mathbf{QR} \text{ behelyettesítése után}\\
(\mathbf{QR})^\mathsf{T}\mathbf{QR}\hat{\mathbf{x}} &= (\mathbf{QR})^\mathsf{T}\mathbf{b} &&\\
\mathbf{R}^\mathsf{T}\mathbf{Q}^\mathsf{T}\mathbf{QR}\hat{\mathbf{x}} &= \mathbf{R}^\mathsf{T}\mathbf{Q}^\mathsf{T}\mathbf{b} && \mathbf{Q}^\mathsf{T}\mathbf{Q} = \mathbf{I}\\
\mathbf{R}^\mathsf{T}\mathbf{R}\hat{\mathbf{x}} &= \mathbf{R}^\mathsf{T}\mathbf{Q}^\mathsf{T}\mathbf{b} && \text{balról szorzás az } (\mathbf{R}^\mathsf{T})^{-1} \text{ mátrixszal}\\
\mathbf{R}\hat{\mathbf{x}} &= \mathbf{Q}^\mathsf{T}\mathbf{b}. &&
\end{aligned}$$

Az utolsó egyenlet visszahelyettesítésekkel is megoldható, mivel $\mathbf{R}$ felső háromszögmátrix. Mivel $\mathbf{R}$ főátlójában nincsenek zéruselemek, ezért $\mathbf{R}$ invertálható (ezt kihasználtuk, amikor $(\mathbf{R}^\mathsf{T})^{-1}$-gyel szoroztunk), tehát az egyenletből $\hat{\mathbf{x}}$ kifejezhető: $\hat{\mathbf{x}} = \mathbf{R}^{-1}\mathbf{Q}^\mathsf{T}\mathbf{b}$. □

Hasonló mátrixok karakterisztikus polinomja megegyezik

8.23. TÉTEL (SAJÁTÉRÉKHEZ KAPCSOLÓDÓ INVARIÁNSOK). *Ha* $\mathbf{A} \sim \mathbf{B}$*, akkor* $\mathbf{A}$ *és* $\mathbf{B}$ *karakterisztikus polinomja azonos, így sajátértékei, azok algebrai, sőt geometriai multiplicitásai is megegyeznek.*

BIZONYÍTÁS. A bizonyítás során föltesszük, hogy valamely invertálható $\mathbf{C}$ mátrixszal $\mathbf{A} = \mathbf{C}^{-1}\mathbf{BC}$. Ekkor

$$\begin{aligned}
\mathbf{A} - \lambda\mathbf{I} &= \mathbf{C}^{-1}\mathbf{BC} - \lambda\mathbf{C}^{-1}\mathbf{IC})\\
&= \mathbf{C}^{-1}(\mathbf{BC} - \lambda\mathbf{IC})\\
&= \mathbf{C}^{-1}(\mathbf{B} - \lambda\mathbf{I})\mathbf{C},
\end{aligned}$$

azaz $\mathbf{A} - \lambda\mathbf{I}$ és $\mathbf{B} - \lambda\mathbf{I}$ is hasonlóak. A 7.16. tétel szerint hasonló mátrixok determinánsa megegyezik, így $\det(\mathbf{A} - \lambda\mathbf{I}) = \det(\mathbf{B} - \lambda\mathbf{I})$, azaz megegyeznek $\mathbf{A}$ és $\mathbf{B}$ karakterisztikus polinomjai is. Ez maga után vonja, hogy megegyeznek sajátértékeik, és azok (algebrai) multiplicitásai. A geometriai multiplicitások egyenlőségéhez elég belátni, hogy $\mathbf{A} - \lambda\mathbf{I}$ és $\mathbf{B} - \lambda\mathbf{I}$ nullterének dimenziója megegyezik, azt viszont ugyancsak a 7.16. tételben igazoltuk. □

Diagonizálhatósággal ekvivalens állítások

8.25. TÉTEL (DIAGONALIZÁLHATÓSÁG SZÜKSÉGES ÉS ELÉGSÉGES FELTÉTELE). *Az $n \times n$-es $\mathbf{A}$ mátrix pontosan akkor diagonalizálható, azaz pontosan akkor létezik olyan $\mathbf{C}$ mátrix, melyre $\mathbf{C}^{-1}\mathbf{A}\mathbf{C}$ diagonális, ha $\mathbf{A}$-nak van n lineárisan független sajátvektora. Ekkor a diagonális mátrix az $\mathbf{A}$ sajátértékeiből, $\mathbf{C}$ a sajátvektoraiból áll.*

BIZONYÍTÁS. Ha $\mathbf{A}$ hasonló egy diagonális mátrixhoz, azaz van olyan $\mathbf{C}$ mátrix, hogy $\mathbf{\Lambda} = \mathbf{C}^{-1}\mathbf{A}\mathbf{C}$ diagonális, akkor $\mathbf{C}$-vel balról szorozva a $\mathbf{C}\mathbf{\Lambda} = \mathbf{A}\mathbf{C}$ egyenlőséget kapjuk. Ha $\mathbf{C} = [\mathbf{x}_1\ \mathbf{x}_2\ \dots\ \mathbf{x}_n]$ és $\mathbf{\Lambda} = \operatorname{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$, akkor

$$[\mathbf{x}_1\ \mathbf{x}_2\ \dots\ \mathbf{x}_n] \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix} = \mathbf{A}[\mathbf{x}_1\ \mathbf{x}_2\ \dots\ \mathbf{x}_n]. \tag{8.5}$$

Itt a bal oldali mátrix i-edik oszlopa $\lambda_i\mathbf{x}_i$, a jobb oldali mátrixé $\mathbf{A}\mathbf{x}_i$. Ezek megegyeznek, azaz $\mathbf{A}\mathbf{x}_i = \lambda_i\mathbf{x}_i$, tehát $\mathbf{x}_i$ a λ_i sajátértékhez tartozó sajátvektor. Mivel $\mathbf{C}$ invertálható, ezért oszlopvektorai függetlenek, ami bizonyítja az állításunk egyik felét. Tegyük most fel, hogy van $\mathbf{A}$-nak n független sajátvektora. Képezzünk a sajátértékekből egy $\mathbf{\Lambda}$ diagonális mátrixot, úgy hogy a $\mathbf{C}$ mátrix i-edik oszlopába kerülő $\mathbf{x}_i$ vektorhoz tartozó λ_i sajátérték a $\mathbf{\Lambda}$ mátrix i-edik oszlopába kerüljön. Mivel $\lambda_i\mathbf{x}_i = \mathbf{A}\mathbf{x}_i$, ezért fönnáll a (8.5) összefüggés, azaz $\mathbf{\Lambda}$ hasonló $\mathbf{A}$-hoz. □

8.28. ÁLLÍTÁS (DIAGONALIZÁLHATÓ MÁTRIX POLINOMJA). *Legyen* $\mathbf{A} = \mathbf{C}\mathbf{\Lambda}\mathbf{C}^{-1}$, *ahol* $\mathbf{\Lambda} = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$, *és* $p(x)$ *egy tetszőleges polinom. Ekkor*

$$p(\mathbf{A}) = \mathbf{C} \begin{bmatrix} p(\lambda_1) & 0 & \ldots & 0 \\ 0 & p(\lambda_2) & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & p(\lambda_n) \end{bmatrix} \mathbf{C}^{-1}.$$

8.31. KÖVETKEZMÉNY (KÜLÖNBÖZŐ SAJÁTÉRTÉKEK ÉS A DIAGONALIZÁLHATÓSÁG). *Ha az* n*-edrendű* $\mathbf{A}$ *mátrixnak* n *darab különböző sajátértéke van, akkor diagonalizálható.*

BIZONYÍTÁS. A 8.30. tétel szerint n különböző sajátértékhez n független sajátvektor tartozik, ami a 8.25. tétel szerint épp azt jelenti, hogy a mátrix diagonalizálható. □

8.34. TÉTEL (DIAGONALIZÁLHATÓSÁG ÉS A GEOMETRIAI MULTIPLICITÁS). *Egy* n*-edrendű négyzetes mátrix pontosan akkor diagonalizálható, ha a sajátértékeihez tartozó geometriai multiplicitások összege* n.

BIZONYÍTÁS. ($\Rightarrow$) Ha a mátrix diagonalizálható, akkor egy adott sajátértékhez tartozó sajátaltér dimenziója megegyezik e sajátérték geometriai multiplicitásával. A geometriai multiplicitások összege tehát épp n, hisz egyetlen sajátvektor sem lehet két sajátaltérben.

($\Leftarrow$) A geometriai multiplicitás nem nagyobb az algebrainál, az algebraiak összege pedig legföljebb n (komplex esetben pontosan n, valós esetben lehet n-nél kisebb is, ha a karakterisztikus polinomnak vannak nem valós gyökei). Így ha a geometriai multiplicitások összege n, akkor minden sajátaltérből kiválasztva egy bázist, és véve ezek egyesítését, egy n sajátvektorból álló független vektorrendszert kapunk (ld. a 8.30. tétel utáni megjegyzéseket). Így tehát a mátrix diagonalizálható. □

Különböző sajátértékhez tartozó sajátvektorok függetlensége

8.30. TÉTEL (KÜLÖNBÖZŐ SAJÁTÉRTÉKEK SAJÁTVEKTORAI). *Ha* λ_1, $\lambda_2, \ldots \lambda_k$ *különböző sajátértékei az* $n \times n$*-es* $\mathbf{A}$ *mátrixnak, akkor a hozzájuk tartozó* $\mathbf{x}_1, \mathbf{x}_2, \ldots \mathbf{x}_k$ *sajátvektorok lineárisan függetlenek.*

BIZONYÍTÁS. Indirekt módon bizonyítunk. Tegyük fel, hogy e vektorok lineárisan összefüggők. Ekkor van a vektorok közt olyan, amely csak a kisebb indexűek lineáris függvénye. Legyen ezek közül a legkisebb indexű $\mathbf{x}_i$, azaz

$$\mathbf{x}_i = c_1\mathbf{x}_1 + \ldots + c_{i-1}\mathbf{x}_{i-1}, \tag{8.11}$$

de az i-nél kisebb indexű vektorok már lineárisan függetlenek. Szorozzuk meg az egyenlőség mindkét oldalát balról az $\mathbf{A}$ mátrixszal:

$$\mathbf{A}\mathbf{x}_i = \mathbf{A}(c_1\mathbf{x}_1 + \ldots + c_{i-1}\mathbf{x}_{i-1}) = c_1\mathbf{A}\mathbf{x}_1 + \ldots + c_{i-1}\mathbf{A}\mathbf{x}_{i-1},$$

majd használjuk ki, hogy e vektorok sajátvektorok:

$$\lambda_i\mathbf{x}_i = c_1\lambda_1\mathbf{x}_1 + \ldots + c_{i-1}\lambda_{i-1}\mathbf{x}_{i-1}. \tag{8.12}$$

Ezután a (8.11) egyenlet mindkét oldalát λ_i-vel szorozva kapjuk, hogy

$$\lambda_i\mathbf{x}_i = c_1\lambda_i\mathbf{x}_1 + \ldots + c_{i-1}\lambda_i\mathbf{x}_{i-1}. \tag{8.13}$$

Végül a (8.13) egyenletből a (8.12) egyenletet kivonva kapjuk, hogy

$$\mathbf{0} = c_1(\lambda_i - \lambda_1)\mathbf{x}_1 + \ldots + c_{i-1}(\lambda_i - \lambda_{i-1})\mathbf{x}_{i-1},$$

Mivel az $\mathbf{x}_1, \ldots, \mathbf{x}_{i-1}$ vektorok már lineárisan függetlenek, és a $\lambda_1, \ldots,$ λ_i értékek különbözőek, ezért $c_1 = \cdots = c_{i-1} = 0$. Eszerint

$$\mathbf{x}_i = 0\mathbf{x}_1 + \cdots + 0\mathbf{x}_{i-1} = \mathbf{0},$$

ami ellentmondás, hisz $\mathbf{x}_i$ sajátvektor, tehát nem lehet a $\mathbf{0}$. Ez bizonyítja az indirekt feltevés helytelen voltát, azaz igazolja állításunkat. □

► Szokás úgy fogalmazni, hogy a különböző sajátértékekhez tartozó sajátalterek lineárisan függetlenek, hisz bárhogy választunk mindegyikükből egy-egy nemzérus vektort, azok lineárisan függetlenek lesznek.

Pontosan akkor diagonizálható unitéren, ha normális

9.9. TÉTEL (UNITÉR DIAGINALIZÁLHATÓSÁG). *Az* $\mathbf{A} \in \mathbb{C}^{n\times n}$ *mátrix pontosan akkor unitéren diagonalizálható, ha normális.*

BIZONYÍTÁS. ($\Rightarrow$) Tegyük fel, hogy $\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^{\mathsf{H}}$, azaz $\mathbf{A}$ unitéren diagonalizálható. Mivel bármely komplex z számra $\bar{z}z = z\bar{z}$, ezért minden komplex diagonális mátrix normális, így $\mathbf{\Lambda}^{\mathsf{H}}\mathbf{\Lambda} = \mathbf{\Lambda}\mathbf{\Lambda}^{\mathsf{H}}$. Eszerint

$$\begin{aligned}\mathbf{A}^{\mathsf{H}}\mathbf{A} &= (\mathbf{U}\mathbf{\Lambda}\mathbf{U}^{\mathsf{H}})^{\mathsf{H}}(\mathbf{U}\mathbf{\Lambda}\mathbf{U}^{\mathsf{H}}) = \mathbf{U}\mathbf{\Lambda}^{\mathsf{H}}\mathbf{U}^{\mathsf{H}}\mathbf{U}\mathbf{\Lambda}\mathbf{U}^{\mathsf{H}} = \mathbf{U}\mathbf{\Lambda}^{\mathsf{H}}\mathbf{\Lambda}\mathbf{U}^{\mathsf{H}}\\ &= \mathbf{U}\mathbf{\Lambda}\mathbf{\Lambda}^{\mathsf{H}}\mathbf{U}^{\mathsf{H}} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^{\mathsf{H}}\mathbf{U}\mathbf{\Lambda}^{\mathsf{H}}\mathbf{U}^{\mathsf{H}} = (\mathbf{U}\mathbf{\Lambda}\mathbf{U}^{\mathsf{H}})(\mathbf{U}\mathbf{\Lambda}\mathbf{U}^{\mathsf{H}})^{\mathsf{H}} = \mathbf{A}\mathbf{A}^{\mathsf{H}}.\end{aligned}$$

($\Leftarrow$) A Schur-felbontás szerint minden komplex négyzetes $\mathbf{A}$ mátrix előáll

$$\mathbf{A} = \mathbf{U}\mathbf{T}\mathbf{U}^{\mathsf{H}}$$

alakban, ahol $\mathbf{U}$ unitér, $\mathbf{T}$ felsőháromszög-mátrix. Tegyük fel, hogy $\mathbf{A}$ normális. Ekkor $\mathbf{T}$ is normális, ugyanis a fenti levezetéshez hasonlóan

$$\begin{aligned}\mathbf{T}^{\mathsf{H}}\mathbf{T} &= (\mathbf{U}^{\mathsf{H}}\mathbf{A}\mathbf{U})^{\mathsf{H}}(\mathbf{U}^{\mathsf{H}}\mathbf{A}\mathbf{U}) = \mathbf{U}^{\mathsf{H}}\mathbf{A}^{\mathsf{H}}\mathbf{U}\mathbf{U}^{\mathsf{H}}\mathbf{A}\mathbf{U} = \mathbf{U}^{\mathsf{H}}\mathbf{A}^{\mathsf{H}}\mathbf{A}\mathbf{U}\\ &= \mathbf{U}^{\mathsf{H}}\mathbf{A}\mathbf{A}^{\mathsf{H}}\mathbf{U} = \mathbf{U}^{\mathsf{H}}\mathbf{A}\mathbf{U}\mathbf{U}^{\mathsf{H}}\mathbf{A}^{\mathsf{H}}\mathbf{U} = (\mathbf{U}^{\mathsf{H}}\mathbf{A}\mathbf{U})(\mathbf{U}^{\mathsf{H}}\mathbf{A}\mathbf{U})^{\mathsf{H}} = \mathbf{T}\mathbf{T}^{\mathsf{H}}.\end{aligned}$$

A $\mathbf{T}$ mátrix alakja

$$\begin{bmatrix} t_{11} & t_{12} & \dots & t_{1n}\\ 0 & t_{22} & \dots & t_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \dots & t_{nn}\end{bmatrix},$$

ezért $[\mathbf{T}^{\mathsf{H}}\mathbf{T}]_{11} = |t_{11}|^2$, $[\mathbf{T}\mathbf{T}^{\mathsf{H}}]_{11} = |t_{11}|^2 + |t_{12}|^2 + \cdots + |t_{1n}|^2$, amiből $t_{12} = \cdots = t_{1n} = 0$ adódik. Hasonlóan fölírva a $[\mathbf{T}^{\mathsf{H}}\mathbf{T}]_{22}$ és a $[\mathbf{T}\mathbf{T}^{\mathsf{H}}]_{22}$ elemeket kapjuk, hogy $t_{23} = \cdots = t_{2n} = 0$, stb. Tehát $\mathbf{T}$ diagonális. $\square$

SVD létezése

10.6. TÉTEL (A SZINGULÁRIS ÉRTÉKEK LÉTEZÉSE ÉS EGYÉRTELMŰSÉGE). *Minden r-rangú valós vagy komplex $\mathbf{A}$ mátrixnak létezik r szinguláris értéke. Ezek valós esetben megegyeznek az $\mathbf{A}^\mathsf{T}\mathbf{A}$, illetve az $\mathbf{A}\mathbf{A}^\mathsf{T}$ (komplex esetben az $\mathbf{A}^\mathsf{H}\mathbf{A}$, illetve az $\mathbf{A}\mathbf{A}^\mathsf{H}$) pozitív sajátértékeinek négyzetgyökeivel. A szinguláris értékek monoton csökkenő sorozata egyértelmű.*

BIZONYÍTÁS. A bizonyítást valós esetre írjuk le, komplexre lényegében azonos. Az $\mathbf{A}^\mathsf{T}\mathbf{A}$ mátrix szimmetrikus ($\mathbf{A}^\mathsf{H}\mathbf{A}$ önadjungált), mert $(\mathbf{A}^\mathsf{T}\mathbf{A})^\mathsf{T} = \mathbf{A}^\mathsf{T}(\mathbf{A}^\mathsf{T})^\mathsf{T} = \mathbf{A}^\mathsf{T}\mathbf{A}$. Ennek következtében minden sajátértéke valós, így e mátrix ortogonálisan diagonalizálható. Másrészt minden sajátértéke nemnegatív, másként fogalmazva $\mathbf{A}^\mathsf{T}\mathbf{A}$ pozitív szemidefinit, ugyanis tetszőleges $\mathbf{x}$ vektorra $\mathbf{x}^\mathsf{T}\mathbf{A}^\mathsf{T}\mathbf{A}\mathbf{x} = (\mathbf{A}\mathbf{x})^\mathsf{T}(\mathbf{A}\mathbf{x}) = |\mathbf{A}\mathbf{x}|^2 \geqslant 0$. A 0-tól különböző sajátértékek száma megegyezik $\mathbf{A}^\mathsf{T}\mathbf{A}$ rangjával, hisz az megegyezik diagonális alakja nemnulla elemeinek számával. Másrészt ???? szerint $\mathrm{r}(\mathbf{A}^\mathsf{T}\mathbf{A}) = \mathrm{r}(\mathbf{A}) = r$. Tehát, ha nagyság szerinti sorba rendezzük a sajátértékeket ($\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$), akkor $\lambda_i > 0$, ha $1 \leq i \leq r$, és $\lambda_i = 0$, ha $r < i \leq n$. Eszerint $\sigma_i = \sqrt{\lambda_i} > 0$, ha $1 \leq i \leq r$. Végül, mivel $\mathbf{A}^\mathsf{T}\mathbf{A}$ sajátértékei egyértelműek, ezért $\mathbf{A}$ szinguláris értékei is azok. Azt, hogy $\mathbf{A}^\mathsf{T}\mathbf{A}$ és $\mathbf{A}\mathbf{A}^\mathsf{T}$ 0-tól különböző sajátértékei megegyeznek, korábban beláttuk. □