

Tartalomjegyzék

A nulltér altér	2
Minden véges dimenziós lineáris leképezéshez létezik egy mátrix, ami azt a leképezést generálja	3
Merőleges vetítés $\mathbb{R}^n$ egy alterére	4
$\text{rang} A + \dim \mathcal{N} A = n$	4
$S(A) \perp \mathcal{N}(A)$	5
Megoldható egyenletrendszer egyértelmű megoldása a sortérben és e megoldás abszolút értékének minimalizálása	6
Páronként ortogonális vektorok függetlensége	8
Gram-Schmidt ortogonalizáció	8
A QR felbontás létezése	9
Egyenletrendszer optimális megoldása QR felbontással	10
Hasonló mátrixok karakterisztikus polinomja megegyezik	10
Diagonizálhatósággal ekvivalens állítások	11
Különböző sajátértékhez tartozó sajátvektorok függetlensége	13
Pontosan akkor diagonizálható unitéren, ha normális	14
SVD létezése	15

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

$$\mathbf{a}_1 x_1 + \mathbf{a}_2 x_2 + \dots + \mathbf{a}_n x_n = \mathbf{0}$$

$$\mathbf{a}_1 y_1 + \mathbf{a}_2 y_2 + \dots + \mathbf{a}_n y_n = \mathbf{0},$$

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

$$\begin{aligned} & \mathbf{a}_1(cx_1 + dy_1) + \mathbf{a}_2(cx_2 + dy_2) + \dots + \mathbf{a}_n(cx_n + dy_n) \\ &= (c\mathbf{a}_1x_1 + d\mathbf{a}_1y_1) + (c\mathbf{a}_2x_2 + d\mathbf{a}_2y_2) + \dots + (c\mathbf{a}_nx_n + d\mathbf{a}_ny_n) \\ &= c(\mathbf{a}_1x_1 + \mathbf{a}_2x_2 + \dots + \mathbf{a}_nx_n) + d(\mathbf{a}_1y_1 + \mathbf{a}_2y_2 + \dots + \mathbf{a}_ny_n) \\ &= \mathbf{0} + \mathbf{0} = \mathbf{0}. \end{aligned}$$

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). $\square$

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

7.9. TÉTEL (Az $\mathbb{R}^n \rightarrow \mathbb{R}^m$ LINEÁRIS LEKÉPEZÉSEK MÁTRIXLEKÉPEZÉSEK). Legyen $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ 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

$$\mathbf{A} = [\mathbf{Ae}_1 | \mathbf{Ae}_2 | \dots | \mathbf{Ae}_n],$$

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

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

$$\mathbf{A} = [\mathbf{Ae}_1 | \mathbf{Ae}_2 | \dots | \mathbf{Ae}_n] \tag{7.2}$$

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

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

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 $\mathbf{Ae}_i$ lehet. $\square$

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

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

BIZONYÍTÁS. A $\mathbf{P} = \mathbf{P}^2$ feltétel szükségessége szemléletesen világos, hisz minden P lineáris leképezés, mely az egész $\mathbb{R}^n$ teret egy alterre – nevezetesen $\text{Im } P$ -re – vetíti, az alter 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.

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

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

másrészt

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

($\Leftarrow$) Tegyük fel, hogy $\mathbf{P} = \mathbf{P}^\top = \mathbf{P}^2$. 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 $\mathbf{P} = \mathbf{P}^\top$, így $\mathbf{x} - \mathbf{P}\mathbf{x} \in \mathcal{N}(\mathbf{P}^\top)$. Ez épp azt jelenti, hogy $\mathbf{x} - \mathbf{P}\mathbf{x}$ merőleges $\mathcal{O}(\mathbf{P})$ -re, és ezt akartuk belátni. $\square$

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

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

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

BIZONYÍTÁS. A mátrix sortérének dimenziója megegyezik a mátrix rangjával, azaz az $[A|0]$ 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 kifesztő 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 kifesztő 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 kifesztik a nullteret, számuk megadja a nulltér dimenzióját. $\square$

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

3.38. ÁLLÍTÁS (A SORTÉR ÉS A NULLTÉR MERŐLEGESSÉGE). A valós A mátrix sortérének bármely s vektora és nullterének tetszőleges x vektora merőleges egymásra, azaz $s \cdot x = 0$.

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

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

□

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

3.41. TÉTEL (LINEÁRIS EGYENLETRENDSZER MEGOLDÁSAI). Minden valószínűleg együtthetős megoldható (konzisztens) lineáris egyenletrendszerre igazak a következő állítások:

- egyetlen megoldása esik az együtthetős mátrix sortérébe;
- a sortérbe eső megoldás az összes megoldás közül a legkisebb abszolút értékű;
- 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 fennáll 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

$$\mathbf{x} = \mathbf{x}_S + \mathbf{x}_N.$$

E megoldásvektort beírva az i -edik egyenletbe kapjuk, hogy

$$b_i = \mathbf{a}_{i*} \cdot \mathbf{x} = \mathbf{a}_{i*} \cdot (\mathbf{x}_S + \mathbf{x}_N) = \mathbf{a}_{i*} \cdot \mathbf{x}_S + \mathbf{a}_{i*} \cdot \mathbf{x}_N = \mathbf{a}_{i*} \cdot \mathbf{x}_S.$$

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 $\mathbf{x}_S$ megoldáshoz bármely nulltérbeli vektort adva az egyenletrendszer egy megoldását kapjuk, igazoltuk tehát a c) állítást is.

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

$$x^2 = x_S^2 + x_N^2 \geq x_S^2, \text{ azaz } |x| \geq |x_S|.$$

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

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 $c_1\mathbf{a}_1 + \dots + c_k\mathbf{a}_k = \mathbf{0}$ egyenletet. Be kell látnunk, hogy ez csak a $c_1 = \dots = c_k = 0$ esetben áll fenn. 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:

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

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*

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

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 $\text{span}(\mathbf{a}_1) = \text{span}(\mathbf{v}_1)$ összefüggés teljesül, ha

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

A $\text{span}(\mathbf{a}_1, \mathbf{a}_2) = \text{span}(\mathbf{v}_1, \mathbf{v}_2)$ 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 $\text{span}(\mathbf{a}_1, \mathbf{a}_2) = \text{span}(\mathbf{v}_1, \mathbf{v}_2)$ fennáll. 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 $\text{span}(\frac{\mathbf{v}_1}{|\mathbf{v}_1|}, \frac{\mathbf{v}_2}{|\mathbf{v}_2|}, \dots, \frac{\mathbf{v}_i}{|\mathbf{v}_i|})$ 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. $\square$

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 $\mathbf{A}$ sorvektoraai 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 $\hat{\mathbf{x}} = \mathbf{R}^{-1}\mathbf{Q}^T\mathbf{b}$, ami megkapható az

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

egyenletrendszerből egyszerű visszahelyettesítéssel is.

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

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

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

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 fölteszük, hogy valamely invertálható $\mathbf{C}$ mátrixszal $\mathbf{A} = \mathbf{C}^{-1}\mathbf{B}\mathbf{C}$. Ekkor

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

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. $\square$

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 $\mathbf{\Lambda} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$, akkor

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

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. $\square$

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

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

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. $\square$

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 legfeljebb 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. $\square$

Különböző sajátértékhez tartozó sajátvektorok függetlensége

8.30. TÉTEL (KÜLÖNBÖZŐ SAJÁTÉRTÉKEK SAJÁTVEKTORAI). *Ha $\lambda_1, \lambda_2, \dots, \lambda_k$ különböző sajátértékei az $n \times n$ -es $\mathbf{A}$ mátrixnak, akkor a hozzájuk tartozó $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$ 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

$$\mathbf{x}_i = c_1 \mathbf{x}_1 + \dots + c_{i-1} \mathbf{x}_{i-1}, \quad (8.11)$$

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

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

majd használjuk ki, hogy e vektorok sajátvektorok:

$$\lambda_i \mathbf{x}_i = c_1 \lambda_1 \mathbf{x}_1 + \dots + c_{i-1} \lambda_{i-1} \mathbf{x}_{i-1}. \quad (8.12)$$

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

$$\lambda_i \mathbf{x}_i = c_1 \lambda_i \mathbf{x}_1 + \dots + c_{i-1} \lambda_i \mathbf{x}_{i-1}. \quad (8.13)$$

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

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

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

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

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

► 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 diagonalizálható unitéren, ha normális

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

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

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

($\Leftarrow$) A **Schur-felbontás** szerint minden komplex négyzetes A mátrix előáll

$$A = UTU^H$$

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

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

A T mátrix alakja

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

ezért $[T^H T]_{11} = |t_{11}|^2$, $[T T^H]_{11} = |t_{11}|^2 + |t_{12}|^2 + \cdots + |t_{1n}|^2$, amiből $t_{12} = \cdots = t_{1n} = 0$ adódik. Hasonlóan fölírva a $[T^H T]_{22}$ és a $[T T^H]_{22}$ elemeket kapjuk, hogy $t_{23} = \cdots = t_{2n} = 0$, stb. Tehát 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 $\mathbf{A}^\top \mathbf{A}$, illetve az $\mathbf{A}\mathbf{A}^\top$ (komplex esetben az $\mathbf{A}^H \mathbf{A}$, illetve az $\mathbf{A}\mathbf{A}^H$) pozitív sajátértékeinek négyzetgyökeivel. A szinguláris értékek monoton csökkenő sorozata egyértelmű.

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