Szemléletes lineáris algebra - összefoglaló I.
informatikusoknak

Segédanyag a
SZ13MI és SZ14MI tárgyakhoz

összeállította:
Dr. Szörényi Miklós
főisk. docens
2005.

Tartalom:

A lineáris algebra a vektoralgebra fogalomkörét (vektorok összege, nyújtása, abszolút értéke, skaláris szorzata stb.) általánosítja és különböző problémákra (pl. egyenletrendszerek iteratív megoldása) alkalmazza. Így teszi szemléletessé a geometriától eredetileg távol eső problémákat.

A legegyszerűbb általánosítás a lineáris tér, másképpen vektortér, amelyben két művelet az *összeadás* és a skalárral való szorzás – *nyújtás* - van értelmezve. Skalár alatt itt valós, vagy komplex számot értünk, ezek összességét Φ-vel, és ennek elemeit pedig többnyire görög kisbetűkkel (pl. α, β) jelöljük.

Def. : **Lineáris tér**

Az $\boldsymbol{X}$ halmazt lineáris térnek mondjuk, ha minden $\mathbf{x}, \mathbf{y} \in \boldsymbol{X}$-hez hozzá van rendelve egy $\mathbf{x}+\mathbf{y} \in \boldsymbol{X}$ elem, amelyet **x**, **y** összegének nevezünk, és minden $\alpha \in \Phi$, $\mathbf{x} \in \boldsymbol{X}$ párhoz hozzá van rendelve egy $\alpha\mathbf{x} \in \boldsymbol{X}$, amelyet **x**-nek az α skalárral való szorzatának nevezünk, és amely hozzárendelések (mi mondjuk meg hogyan kell értelmezni ill. kiszámolni ezeket!) a következő tulajdonságokkal rendelkeznek:

I. összeadás tulajdonságai

a./ kommutativitás:

$$\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$$

b./ asszociativitás:

$$(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$$

c./ van $\varnothing$-val jelölt zéruselem, mellyel minden $\mathbf{x} \in \boldsymbol{X}$-re:

$$\mathbf{x} + \varnothing = \mathbf{x}$$

II. nyújtás tulajdonságai

minden $\mathbf{x},\mathbf{y} \in \boldsymbol{X}$ és $\alpha,\beta \in \Phi$-re

a./ disztributivitás:

$$\alpha(\mathbf{x} + \mathbf{y}) = \alpha\mathbf{x} + \alpha\mathbf{y}$$
$$(\alpha + \beta)\mathbf{x} = \alpha\mathbf{x} + \beta\mathbf{x}$$

b./ asszociativitás:

$$(\alpha\beta)\mathbf{x} = \alpha(\beta\mathbf{x})$$

c./ spec. nyújtások:

$$0\mathbf{x} = \varnothing \quad \text{és} \quad 1\mathbf{x} = \mathbf{x}$$

A lineáris tér elemeit vektoroknak nevezzük.

A $(-1)\mathbf{x}$ vektorra a $-\mathbf{x}$ jelölést használjuk, és az $\mathbf{x}+(-\mathbf{y})$ összeget röviden $\mathbf{x}-\mathbf{y}$ alakban írjuk.

1. példa:

A valós szám n-esek lineáris teret alkotnak a következő műveletekkel:

$$\mathbf{x}=\begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ x_n \end{bmatrix} \quad \text{és} \quad \mathbf{y}=\begin{bmatrix} y_1 \\ y_2 \\ \cdot \\ y_n \end{bmatrix} \quad \text{akkor} \quad \mathbf{x}+\mathbf{y}=\begin{bmatrix} x_1+y_1 \\ x_2+y_2 \\ \cdot \\ x_n+y_n \end{bmatrix} \quad \text{és} \quad \lambda\mathbf{x}=\begin{bmatrix} \lambda x_1 \\ \lambda x_2 \\ \cdot \\ \lambda x_n \end{bmatrix}$$

2. példa:

Valamely [a,b] véges zárt intervallumon folytonos függvények lineáris teret alkotnak az

$(f+g)(t) = f(t)+g(t) \qquad t \in [a,b]$
$(\alpha f)(t) = \alpha f(t)$

szokásos összeadással és számmal való szorzással.

Egy ***X*** lineáris tér olyan ***Y*** részhalmazát, melyre a lineáris tér axiómái teljesülnek (a lineáris műveletek nem vezetnek ki az ***Y*** részhalmazból) **altér-**nek nevezzük.

Tájékozódás a lineáris térben

Def.: Lineáris kombináció

A lineáris tér két művelete gyakran kombinációban jelenik meg:

$\alpha\mathbf{x} + \beta\mathbf{y} \qquad \mathbf{x},\mathbf{y} \in \boldsymbol{X} \qquad \alpha,\beta \in \Phi$

ezt **x** és **y** lineáris kombinációjának nevezzük. Több elem esetén:

$\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + \alpha_3\mathbf{x}_3 + \ .... + \alpha_n\mathbf{x}_n = \Sigma\ \alpha_k\mathbf{x}_k$ ahol $\alpha_k \in \Phi$, $\mathbf{x}_k \in \boldsymbol{X}$, k=1,2, ...,n

Def.: **Lineáris függetlenség/függőség**

Az $\mathbf{x}_k$ k=1,2,..,n elemek lineárisan függetlenek, ha egyikük sem fejezhető ki a többi elem lineáris kombinációjaként (ilyenek például a geometriai térben az **i, j, k** vektorok). Más megfogalmazásban a tér $\varnothing$ eleme csak triviális módon állítható elő belőlük, azaz a

$$\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + \alpha_3\mathbf{x}_3 + \ .... + \alpha_n\mathbf{x}_n = \varnothing$$

egyenlőség csak a triviális $\alpha_k = 0$; k=1,2,..,n módon teljesülhet.

A lineárisan nem független elemeket lineárisan összefüggőknek nevezzük. Erre egyszerű példát találunk a geometriai vektorok terében, amikor három egy síkba eső nem párhuzamos vektort vizsgálunk. Bármelyikük előállítható a másik kettő lineáris kombinációjaként (ld. a vektorfelbontás "paralelogramma szabálya").

Def.: Bázis

Az $\boldsymbol{X}$ lineáris tér véges sok $\mathbf{e}_1, \mathbf{e}_2,..., \mathbf{e}_n$ lineárisan független vektoraiból álló $\boldsymbol{B}=\{ \mathbf{e}_1, \mathbf{e}_2,..., \mathbf{e}_n \}$ részhalmazt $\boldsymbol{X}$ egy bázisának nevezzük, ha $\boldsymbol{B}$ előállítja (kifeszíti) az $\boldsymbol{X}$ lineáris teret, azaz $\boldsymbol{X}$ minden vektora $\boldsymbol{B}$ elemeinek lineáris kombinációja. Ekkor az

$$\mathbf{x} = \alpha_1\mathbf{e}_1 + \alpha_2\mathbf{e}_2 + \alpha_3\mathbf{e}_3 + ... + \alpha_n\mathbf{e}_n$$

előállításában szereplő α_k együtthatókat az $\mathbf{x}$ vektor $\boldsymbol{B}$ bázisra vonatkozó koordinátáinak nevezzük. A $\boldsymbol{B}$ bázis vektorainak száma pedig az $\boldsymbol{X}$ lineáris tér **dimenziószáma**.

A példáink közül az 1. példában szereplő szám n-esek n-dimenziós teret alkotnak, a 2. példabeli C[a,b] halmaz pedig végtelen dimenziósat. A második esetben elég csak arra gondolni, hogy az akárhányszor differenciálható függvények Taylor sor szerinti előállításában szereplő x^n függvények képezik a tér bázisát.

Mi itt nem fogunk foglalkozni végtelen dimenziós terekkel, így a speciális konvergencia-problémák (pl. térbeli elemek konvergens sorozatának határértéke térbeli-e?) fel sem fognak merülni.

A távolságfogalom általánosítása: norma

Def.: Norma

Az $\boldsymbol{X}$ lineáris teret normált térnek nevezzük, ha minden eleméhez (vektorához) pontosan egy $||\mathbf{x}||$ valós szám tartozik a következő tulajdonságokkal:

I. $||\mathbf{x}|| \geq 0$, és $||\mathbf{x}|| = 0$ pontosan akkor, ha $\mathbf{x} = \varnothing$ (a távolság nem lehet negatív!)

II. $||\mathbf{x} + \mathbf{y}|| \leq ||\mathbf{x}|| + ||\mathbf{y}||$; $\mathbf{x},\mathbf{y} \in \boldsymbol{X}$ (háromszög egyenlőtlenség)

III. $|| \lambda\mathbf{x} || = |\lambda| \cdot ||\mathbf{x}||$; $\mathbf{x} \in \boldsymbol{X}$, $\lambda \in \Phi$

3. példa: A szám n-esek lineáris terében az

$$|| \mathbf{x} ||_p = \left(\sum_{k=1}^{n} | x_k |^p \right)^{\frac{1}{p}}$$ - val definiált mennyiség

normának tekinthető, mert mind a három axiómát kielégíti. Az I. és III. axióma teljesülése minden nehézség nélkül, a háromszög-egyenlőtlenség teljesülése pedig az ún. Minkowski egyenlőtlenség segítségével bizonyítható.

Itt p speciális értékei mellett különböző normákhoz jutunk:

a/ $\| \mathbf{x} \|_1 = \sum_{i=1}^{n} | x_i |$ i=1,2, ... n (oktaéder norma)

b/ $\| \mathbf{x} \|_\infty = \max | x_i |$ i=1,2, ... n (kocka norma)

c/ $\| \mathbf{x} \|_2 = \left(\sum_{k=1}^{n} | x_k |^2 \right)^{\frac{1}{2}}$ (euklideszi ill. gömb norma)

Ezen kívül még más távolságfogalmakkal is találkozhatunk.

A távolságfogalom általánosítása után már könnyű a lineáris térbeli konvergenciát definiálni.

Def.: Konvergencia

Az $\boldsymbol{X}$ lineáris tér $\{\mathbf{x}_n\}$ sorozata konvergens, ha van olyan $\mathbf{x} \in \boldsymbol{X}$, hogy az $\|\mathbf{x} - \mathbf{x}_0\|$ számsorozat nullához konvergál.

Egy $\{\mathbf{x}_n\}$ sorozatot önmagában konvergensnek, ill. Cauchy sorozatnak nevezünk, ha minden $\in>0$ -hoz létezik $N=N(\in)$, hogy $n,m>N(\in)$ esetén:

$$\|\mathbf{x} - \mathbf{x}_0\| < \in$$

Minden konvergens sorozat önmagában is konvergens, de megfordítás nem minden lineáris térben lesz igaz. Viszont a számunkra fontos szám n-esek lineáris terében igaz!

A lineáris teret zártnak mondjuk, ha minden Cauchy sorozatának létezik e térbeli határeleme, a lineáris zárt tereket **Banach tér**-nek is hívjuk.

A skaláris szorzat általánosítása

A geometriai vektortérben a skaláris szorzat a legjellegzetesebb fogalom, hiszen ezzel fejezhető ki a vektorok merőlegessége, vetülete, távolsága, sőt maga a vektor hossza, a norma is. Az itt tapasztalt geometriai tulajdonságokat követeljük meg bármelyik, lineáris térben bevezetett skaláris szorzattól.

Def.: Skaláris szorzat

Ha az $\boldsymbol{X}$ lineáris tér minden $\mathbf{x},\mathbf{y} \in \boldsymbol{X}$ elempárjához $(\mathbf{x} \mid \mathbf{y})$ -al jelölt valós(komplex) szám van hozzá rendelve úgy, hogy:

I. $(\mathbf{x} \mid \mathbf{y}) = \overline{(\mathbf{y} \mid \mathbf{x})}$ (konjugált kommutativitás)

II. $(\mathbf{x} + \mathbf{y} \mid \mathbf{z}) = (\mathbf{x} \mid \mathbf{z}) + (\mathbf{y} \mid \mathbf{z})$ (disztributivitás)

III. $(\alpha \cdot \mathbf{x} \mid \mathbf{y}) = \alpha \cdot (\mathbf{x} \mid \mathbf{y})$

IV. $(\mathbf{x} \mid \mathbf{x}) \geq 0$, és $(\mathbf{x} \mid \mathbf{x}) = 0$ pontosan akkor, ha $\mathbf{x} = \varnothing$

Néhány következmény:

1. $(\mathbf{x} \mid \mathbf{y} + \mathbf{z}) = (\mathbf{x} \mid \mathbf{y}) + (\mathbf{x} \mid \mathbf{z})$

Biz:

$$(\mathbf{x} \mid \mathbf{y} + \mathbf{z}) = \overline{(\mathbf{y} + \mathbf{z} \mid \mathbf{x})} = \overline{(\mathbf{y} \mid \mathbf{x}) + (\mathbf{z} \mid \mathbf{x})} = \overline{(\mathbf{y} \mid \mathbf{x})} + \overline{(\mathbf{z} \mid \mathbf{x})} = (\mathbf{x} \mid \mathbf{y}) + (\mathbf{x} \mid \mathbf{z})$$

2. $(\mathbf{x} \mid \alpha \cdot \mathbf{y}) = \overline{\alpha} \cdot (\mathbf{x} \mid \mathbf{y})$

Biz.:

$$(\mathbf{x} \mid \alpha \mathbf{y}) = \overline{(\alpha \cdot \mathbf{y} \mid \mathbf{x})} = \overline{\alpha \cdot (\mathbf{y} \mid \mathbf{x})} = \overline{\alpha} \cdot \overline{(\mathbf{y} \mid \mathbf{x})} = \overline{\overline{\alpha}} \cdot (\mathbf{x} \mid \mathbf{y})$$

3. $|(\mathbf{x} \mid \mathbf{y})|^2 \leq (\mathbf{x} \mid \mathbf{x}) \cdot (\mathbf{y} \mid \mathbf{y})$ (Cauchy-Bunyakovszkij-Schwartz egyenlőtlenség)

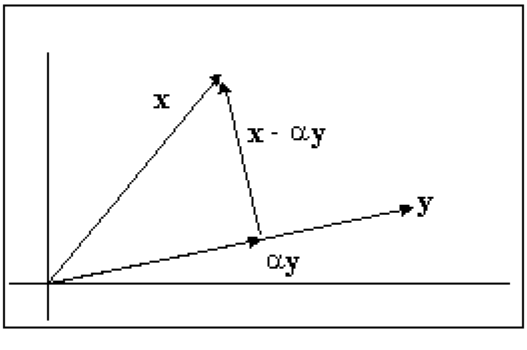
Ez az optimális közelítések elméletében fontos egyenlőtlenség a geometriai vektorok terében a $\cos(\mathbf{x},\mathbf{y}) \mid \leq 1$ egyenlőtlenség miatt triviálisan teljesül.

A bizonyításhoz előbb egy vektorgeometriából ismert segédtételt általánosítunk.

Adott két lineárisan független vektor: **x, y** .

Az $\mathbf{x} - \alpha\mathbf{y}$ vektor akkor lesz merőleges az **y** vektorra (ld. ábra), ha az $\alpha\mathbf{y}$ vektor éppen az **x** vektornak az **y** vektorra vonatkozó vetülete, azaz ha

$$\alpha = \frac{(\mathbf{x} \mid \mathbf{y})}{(\mathbf{y} \mid \mathbf{y})}$$



Mi most ezen ötlet alapján először az **y** és (**x** - α**y**) vektorok skaláris szorzatát nézzük meg a fenti α érték mellett (bármely skaláris szorzattal rendelkező lineáris térben!):

$$(\mathbf{x} - \alpha\mathbf{y} \mid \mathbf{y}) = (\mathbf{x} - \frac{(\mathbf{x}\mid\mathbf{y})}{(\mathbf{y}\mid\mathbf{y})}\mathbf{y} \mid \mathbf{y}) = (\mathbf{x}\mid\mathbf{y}) - \frac{(\mathbf{x}\mid\mathbf{y})}{(\mathbf{y}\mid\mathbf{y})}(\mathbf{y}\mid\mathbf{y}) = (\mathbf{x}\mid\mathbf{y}) - (\mathbf{x}\mid\mathbf{y}) = 0$$

Itt csak az előbbi axiómákat használtuk fel, tehát ez bármilyen skaláris szorzattal ellátott lineáris térben igaz.

Ezek után már csak a triviális $0 \le (\mathbf{x} - \alpha\mathbf{y} \mid \mathbf{x} - \alpha\mathbf{y})$ egyenlőtlenség jobb oldalát kell megvizsgálnunk a fenti α értéke mellett:

$$(\mathbf{x} - \alpha\mathbf{y} \mid \mathbf{x} - \alpha\mathbf{y}) = (\mathbf{x} - \alpha\mathbf{y} \mid \mathbf{x}) - (\mathbf{x} - \alpha\mathbf{y} \mid \alpha\mathbf{y}) = (\mathbf{x} - \alpha\mathbf{y} \mid \mathbf{x}) - \overline{\alpha}(\mathbf{x} - \alpha\mathbf{y} \mid \mathbf{y}) = (\mathbf{x} - \alpha\mathbf{y} \mid \mathbf{x}) =$$

$$= (\mathbf{x} - \frac{(\mathbf{x}\mid\mathbf{y})}{(\mathbf{y}\mid\mathbf{y})}\mathbf{y} \mid \mathbf{x}) = (\mathbf{x}\mid\mathbf{x}) - \frac{(\mathbf{x}\mid\mathbf{y})}{(\mathbf{y}\mid\mathbf{y})}(\mathbf{y}\mid\mathbf{x}) = \frac{(\mathbf{x}\mid\mathbf{x})(\mathbf{y}\mid\mathbf{y}) - (\mathbf{x}\mid\mathbf{y})(\mathbf{y}\mid\mathbf{x})}{(\mathbf{y}\mid\mathbf{y})} =$$

$$= \frac{(\mathbf{x}\mid\mathbf{x})(\mathbf{y}\mid\mathbf{y}) - (\mathbf{x}\mid\mathbf{y})\overline{(\mathbf{x}\mid\mathbf{y})}}{(\mathbf{y}\mid\mathbf{y})} = \frac{(\mathbf{x}\mid\mathbf{x})(\mathbf{y}\mid\mathbf{y}) - |(\mathbf{x}\mid\mathbf{y})|^2}{(\mathbf{y}\mid\mathbf{y})}$$

Mivel az egyenlőségsorozat bal oldala nem negatív, ezért a jobb oldala sem az, így a nem negatív jobboldali nevező miatt a jobboldali tört számlálója sem lehet negatív szám, amiből a nevezetes egyenlőtlenségünk fennállása bármely skaláris szorzattal ellátott lineáris térben már következik.

4. Minden skaláris szorzattal ellátott tér egyben normált tér is a $\|\mathbf{x}\| = (\mathbf{x}\mid\mathbf{x})^{\frac{1}{2}}$ normával, azaz először minden lineáris térben ha jól akarunk tájékozódni, akkor először a skaláris szorzat számítási módját definiáljuk, majd a fenti formulával vezetjük be a norma fogalmát.

A norma axiómáinak teljesülése könnyen ellenőrizhető, mi csak a legnehezebben beláthatóval, az ún. háromszög egyenlőtlenséggel foglalkozunk. Itt fel fogjuk használni az előbb bizonyított Cauchy-Bunyakovszkij-Schwartz egyenlőtlenséget.

$$\|\mathbf{x}+\mathbf{y}\|^2 = (\mathbf{x}+\mathbf{y}\mid\mathbf{x}+\mathbf{y}) = (\mathbf{x}\mid\mathbf{x}) + (\mathbf{y}\mid\mathbf{x}) + (\mathbf{x}\mid\mathbf{y}) + (\mathbf{y}\mid\mathbf{y}) = (\mathbf{x}\mid\mathbf{x}) + (\mathbf{y}\mid\mathbf{x}) + \overline{(\mathbf{y}\mid\mathbf{x})} + (\mathbf{y}\mid\mathbf{y}) =$$

$$= \|\mathbf{x}\|^2 + 2\cdot\mathrm{Re}(\mathbf{x}\mid\mathbf{y}) + \|\mathbf{y}\|^2 \le \|\mathbf{x}\|^2 + 2\cdot|(\mathbf{x}\mid\mathbf{y})| + \|\mathbf{y}\|^2 \le \|\mathbf{x}\|^2 + 2\cdot\|\mathbf{x}\|\cdot\|\mathbf{y}\| + \|\mathbf{y}\|^2 =$$

$$= (\|\mathbf{x}\| + \|\mathbf{y}\|)^2$$

4. példa: Legyen $x(t), y(t) \in C[a,b]$ (azaz x(t) és y(t) az [a,b] véges zárt intervallumon folytonos függvények körébe tartozik, melyek lineáris teret alkotnak), ekkor a

$$(\mathbf{x}\mid\mathbf{y}) = \int_a^b x(t)\cdot\overline{y(t)}\,dt$$

definícióval számított érték könnyen ellenőrizhető módon teljesíti a skaláris szorzat axiómáit, és a $\|\mathbf{x}\| = (\mathbf{x}\,|\,\mathbf{x})^{\frac{1}{2}} = \left(\int_a^b x(t)\cdot \overline{x(t)}dt \right)^{\frac{1}{2}}$ módon számított érték pedig így normának tekinthető.

Def.: **Euklideszi tér**

A skaláris szorzatból származtatott normával - $\|\mathbf{x}\| = (\mathbf{x}|\mathbf{x})^{\frac{1}{2}}$ - ellátott lineáris teret Euklideszi térnek nevezzük.

5. Minden Euklideszi térben:

$$\|\mathbf{x}+\mathbf{y}\|^2 + \|\mathbf{x}-\mathbf{y}\|^2 = 2\cdot\|\mathbf{x}\|^2 + 2\cdot\|\mathbf{y}\|^2$$

A bizonyítás a baloldali skaláris szorzatok szétejtésével történhet.

A skaláris szorzat a vektorok merőlegességének általánosítására is alkalmas:

Def.: **Ortogonalitás**

Az $\boldsymbol{X}$ Euklideszi tér **x,y** vektorát ortogonálisnak ("merőlegesnek") nevezzük, ha $(\mathbf{x}\,|\,\mathbf{y}) = 0$

Ortogonális vektorpárra érvényes a Pithagorasz-tétel:

$$\|\mathbf{x}+\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$$

Bizonyítása $(\mathbf{x}+\mathbf{y}\,|\,\mathbf{x}+\mathbf{y})$ felbontásával és az ortogonalitás felhasználásával történhet.

Def.: **Ortonormált rendszer, ortonormált bázis**

Az $\{\mathbf{e}_k;\ k=1,2,..,n\}$ vektorrendszert ortonormált rendszernek nevezzük, ha:

$(\mathbf{e}_i\,|\mathbf{e}_j) = \delta_{i,j}$, ahol $\delta_{i,j}$ a Kronecker féle szimbólum, ami csak akkor 1, ha $i = j$, különben 0:

$$\delta_{i,j} = \begin{cases} 1, & ha \quad i = j \\ 0, & ha \quad i \neq j \end{cases}$$

Az ortonormált rendszer elemei lineárisan függetlenek.

Az ortonormált rendszer teljes (bázist alkot), ha az $\mathbf{x} = \varnothing$ vektoron kívül nincs olyan másik vektor, amely a rendszer $\mathbf{e}_k$; $k=1,2,..,n$ vektorainak mindegyikére ortogonális lenne.

Ortonormált rendszer képzése (a Gram-Schmidt eljárás)

Legyenek $\{\mathbf{a}_k;\ k=1, 2, ..,n\}$ lineárisan független vektorok. Mivel az ortonormált rendszer minden elemének normája 1, ezért

$$\mathbf{e}_1 = \frac{\mathbf{a}_1}{||\mathbf{a}_1||}$$

A második vektornak ortogonálisnak kell lenni $\mathbf{e}_1$ -re, ezért az $\mathbf{a}_2$ vektor figyelembe vételével $\mathbf{e}_1$ -re ortogonális vektort gyártunk:

$$\mathbf{b}_2 = \mathbf{a}_2 - \lambda_{21}\mathbf{e}_1$$

és itt $(\mathbf{a}_2 - \lambda_{21}\mathbf{e}_1 \,|\, \mathbf{e}_1) = (\mathbf{a}_2 \,|\, \mathbf{e}_1) - \lambda_{21}(\mathbf{e}_1 \,|\, \mathbf{e}_1) = (\mathbf{a}_2 \,|\, \mathbf{e}_1) - \lambda_{21} = 0$ miatt

$\lambda_{21} = (\mathbf{a}_2 \,|\, \mathbf{e}_1)$ adódik.

Ezután $\mathbf{e}_2 = \frac{\mathbf{b}_2}{||\mathbf{b}_2||}$ választással $\{\mathbf{e}_1, \mathbf{e}_2\}$ már kéttagú ortonormált rendszer.

Ha az $\mathbf{a}_3$ vektor figyelembe vételével készítjük el az $\mathbf{e}_3$ vektort, hasonlóan járunk el, csak most

$$\mathbf{b}_3 = \mathbf{a}_3 - (\lambda_{31}\mathbf{e}_1 + \lambda_{32}\mathbf{e}_2)$$

segédvektornak kell ortogonálisnak lenni az $\mathbf{e}_1$ -re, és az $\mathbf{e}_2$ -re is:

$$(\mathbf{a}_3 - (\lambda_{31}\mathbf{e}_1 + \lambda_{32}\mathbf{e}_2) \,|\, \mathbf{e}_1) = (\mathbf{a}_3 | \mathbf{e}_1) - \lambda_{31}(\mathbf{e}_1 \,|\, \mathbf{e}_1) - \lambda_{32}(\mathbf{e}_2 \,|\, \mathbf{e}_1) = (\mathbf{a}_3 | \mathbf{e}_1) - \lambda_{31} = 0$$
$$(\mathbf{a}_3 - (\lambda_{31}\mathbf{e}_1 + \lambda_{32}\mathbf{e}_2) \,|\, \mathbf{e}_2) = (\mathbf{a}_3 | \mathbf{e}_2) - \lambda_{31}(\mathbf{e}_1 \,|\, \mathbf{e}_2) - \lambda_{32}(\mathbf{e}_2 \,|\, \mathbf{e}_{12}) = (\mathbf{a}_3 | \mathbf{e}_2) - \lambda_{32} = 0$$

ahonnan

$$\lambda_{31} = (\mathbf{a}_3 | \mathbf{e}_1) \qquad \lambda_{32} = (\mathbf{a}_3 | \mathbf{e}_2)$$

adódik, majd ezután $\mathbf{e}_3 = \frac{\mathbf{b}_3}{||\mathbf{b}_3||}$ választással $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ már háromtagú ortonormált rendszer lesz, és így tovább...

5. példa: Szám n-esek Euklideszi tere

Egy $\mathbf{x}$ vektor elemeit az x_k számokat a triviális bázisra vonatkozó koordinátáknak nevezzük:

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3 + \ldots + x_n\mathbf{e}_n = x_1\begin{bmatrix}1\\0\\0\\.\\0\end{bmatrix} + x_2\begin{bmatrix}0\\1\\0\\.\\0\end{bmatrix} + x_3\begin{bmatrix}0\\0\\1\\.\\0\end{bmatrix} + \ldots + x_n\begin{bmatrix}0\\0\\0\\.\\1\end{bmatrix}$$

Két vektor skaláris szorzatát a geometriai vektortérhez hasonlóan a koordináták páronkénti szorzatainak összegeként definiáljuk:

$$(\mathbf{x} \,|\, \mathbf{y}) = x_1y_1 + x_2y_2 + x_3y_3 + \ldots + x_ny_n$$

Az axiómák teljesülése könnyen ellenőrizhető.

Ezzel a definícióval a triviális $\{\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3, \dots ,\mathbf{e}_n\}$ bázis ortonormált lesz.

Ha az $|(\mathbf{x} \mid \mathbf{y})|^2 \le (\mathbf{x} \mid \mathbf{x})\cdot(\mathbf{y} \mid \mathbf{y})$ Cauchy-Bunyakovszkij-Schwartz egyenlőtlenséget most konkrétan kifejtjük, akkor a

$$(\sum_{k=1}^{n} x_k y_k)^2 \le (\sum_{k=1}^{n} x_k^2)\cdot(\sum_{k=1}^{n} y_k^2)$$

egyenlőtlenséghez jutunk. Ezt az egyenlőtlenséget elemi eszközökkel bizony elég nehéz lett volna igazolni.

6. példa: Ha a Gram-Schmidt eljárást a [-1,1] intervallumban folytonos (és így négyzetesen is integrálható) $P_k(t) = t^k$, k=0,1,2,3, ... polinomokra alkalmazzuk, akkor az így kapott ortonormált rendszer a *Legendre polinomok:*

$$L_0(t) = \sqrt{\frac{1}{2}}\cdot 1 \quad \text{mert} \quad (L_0(t)|L_0(t)) = \int_{-1}^{1}\left(\sqrt{\frac{1}{2}}\cdot 1\right)^2 dt = 1$$

$$L_1(t) = \sqrt{\frac{3}{2}}\cdot t \quad \text{mert} \quad (L_1(t)|L_1(t)) = \int_{-1}^{1}\left(\sqrt{\frac{3}{2}}\cdot t\right)^2 dt = 1 \quad \text{és } (L_0(t)|L_1(t)) = \int_{-1}^{1}\left(\sqrt{\frac{1}{2}}\cdot\sqrt{\frac{3}{2}}\cdot t\right)^2 dt = 0$$

hasonlóan kapjuk meg a következőket is.:

$$L_2(t) = \sqrt{\frac{5}{2}}\cdot(\frac{3}{2}t^2 - \frac{1}{2}) \qquad \text{mert} \quad (L_2(t)|L_2(t)) = 1 \quad (L_2(t)|L_1(t)) = 0 \quad (L_2(t)|L_0(t)) = 0$$

$$L_3(t) = \sqrt{\frac{7}{2}}\cdot(\frac{5}{2}t^3 - \frac{3}{2}t) \qquad \text{és így tovább...}$$

Az ortonormáltságot a megfelelő integrálok kiszámításával ellenőriztük és ellenőrizhetjük.

Altérbeli legjobb közelítés (projekciótétel)

Vektorgeometriából jól ismert tény, hogy egy origón átmenő $\boldsymbol{M}$ síkbeli $\mathbf{m}_0$ vektor akkor közelíti meg legjobban a térbeli $\mathbf{x}$ vektort, ha az $\mathbf{x} - \mathbf{m}_0$ eltérésvektor merőleges az $\boldsymbol{M}$ síkra, azaz merőleges annak bármelyik $\mathbf{m}$ vektorára:

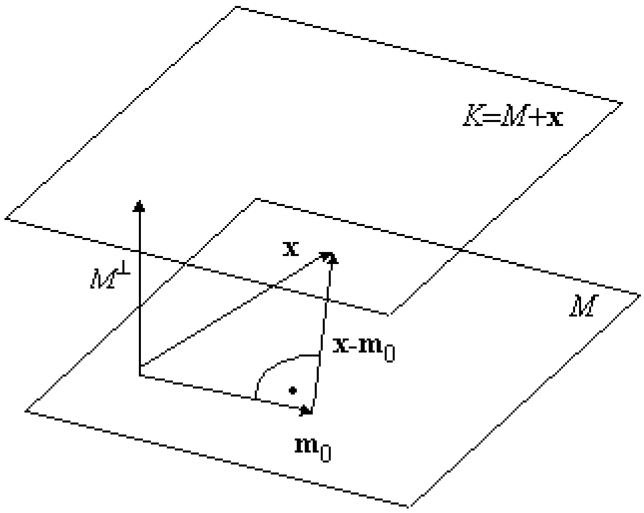
$(\mathbf{x} - \mathbf{m}_0 \mid \mathbf{m}) = 0, \qquad \mathbf{m} \in \boldsymbol{S}$

Ez minden $\boldsymbol{X}$ Euklideszi-térben is hasonlóképp van, sőt $\boldsymbol{M}$-ről csak annyit kell mondanunk, hogy altere az $\boldsymbol{X}$-nek, hiszen ha $\mathbf{m}_0$ olyan, hogy $\mathbf{x} - \mathbf{m}_0$ ortogonális az $\boldsymbol{M}$ altérre, akkor

$(\mathbf{x} - \mathbf{m}_0 \mid \mathbf{m} - \mathbf{m}_0) = 0$ miatt

$\|\mathbf{x} - \mathbf{m}\|^2 = \|(\mathbf{x} - \mathbf{m}_0) + (\mathbf{m}_0 - \mathbf{m})\|^2 = \|\mathbf{x} - \mathbf{m}_0\|^2 + \|\mathbf{m}_0 - \mathbf{m}\|^2$ (ld. Pithagorasz tétel általánosítás)

tehát $\|\mathbf{x} - \mathbf{m}_0\| \leq \|\mathbf{x} - \mathbf{m}\|$, minden $\mathbf{m} \in \boldsymbol{M}$ vektorra.



Ennek folyományaként a következők is beláthatók:

Ha $\mathbf{m}_0 \in \boldsymbol{M}$ az $\mathbf{x} \in \boldsymbol{X}$ minimalizáló vektora és $\boldsymbol{K} = \{\ \mathbf{x}+\mathbf{m};\ \ \mathbf{m} \in \boldsymbol{M}\ \}$, akkor $\mathbf{x} - \mathbf{m}_0$ a $\boldsymbol{K}$ minimális normájú vektora, ilyen csak egyetlen egy van, és ez egyben ortogonális minden $\mathbf{m} \in \boldsymbol{M}$ vektorra.

Ezt úgy is kifejezhetjük, hogy ha $\boldsymbol{M}^{\perp} = \{\ \mathbf{x};\ \ (\mathbf{x} \mid \mathbf{m}) = 0;\ \ \mathbf{m} \in \boldsymbol{M}\ \}$, akkor $\boldsymbol{M}^{\perp} \cap \boldsymbol{K}$ halmazban pontosan csak egy vektor van, és ez éppen a $\boldsymbol{K}$ minimális normájú vektora.

Legyen adott az n-dimenziós $\boldsymbol{M} \subset \boldsymbol{X}$ altér egy bázisa $\{\mathbf{y}_1, \ldots, \mathbf{y}_n\}$. Felmerül a kérdés, hogyan lehet konkrétan meghatározni egy adott $\mathbf{x} \in \boldsymbol{X}$ elem $\boldsymbol{M}$ altérbeli legjobb közelítését? Jelölje a legjobb közelítés altérbeli koordinátáit $\alpha_1, \ldots, \alpha_n$:

$$\mathbf{y} = \alpha_1\mathbf{y}_1 + \ldots + \alpha_n\mathbf{y}_n$$

Mivel a legjobban közelítő $\mathbf{y}$ vektorra igaz, hogy az $\mathbf{x} - \mathbf{y}$ vektor ortogonális az $\boldsymbol{M}$ altérre és ezzel együtt annak minden vektorára és így az $\mathbf{y}_1, \ldots, \mathbf{y}_n$ bázisvektorokra is, ezért

$$(\ \mathbf{x} - (\alpha_1\mathbf{y}_1 + \ldots + \alpha_n\mathbf{y}_n) \mid \mathbf{y}_k) = 0$$

minden k = 1, …, n-re. Átrendezve az egyenleteket a következő ún. normálegyenletekhez jutunk:

$$\begin{aligned}
(\mathbf{y}_1 \mid \mathbf{y}_1)\,\alpha_1 + (\mathbf{y}_2 \mid \mathbf{y}_1)\,\alpha_2 + \ldots + (\mathbf{y}_n \mid \mathbf{y}_1)\,\alpha_n &= (\mathbf{x} \mid \mathbf{y}_1)\\
(\mathbf{y}_1 \mid \mathbf{y}_2)\,\alpha_1 + (\mathbf{y}_2 \mid \mathbf{y}_2)\,\alpha_2 + \ldots + (\mathbf{y}_n \mid \mathbf{y}_2)\,\alpha_n &= (\mathbf{x} \mid \mathbf{y}_2)\\
&\vdots\\
(\mathbf{y}_1 \mid \mathbf{y}_n)\,\alpha_1 + (\mathbf{y}_2 \mid \mathbf{y}_n)\,\alpha_2 + \ldots + (\mathbf{y}_n \mid \mathbf{y}_n)\,\alpha_n &= (\mathbf{x} \mid \mathbf{y}_n)
\end{aligned}$$

Ez a lineáris egyenletrendszer – mint a későbbiekben látni fogjuk – egyértelműen megoldható. Ha az $\{\mathbf{y}_1, \ldots, \mathbf{y}_n\}$ bázis még ortonormált is, akkor a baloldali skaláris szorzatok közül csak a $(\mathbf{y}_k \mid \mathbf{y}_k)$ alakúak nem nullák, sőt

$$(\mathbf{y}_k \mid \mathbf{y}_k) = 1 \qquad k = 1, \ldots, n$$

miatt

$$\alpha_k = (\mathbf{x} \mid \mathbf{y}_k) \qquad k = 1, \ldots, n$$

És ezzel az x-et legjobban közelítő altérbeli **y** vektorra:

$$\mathbf{y} = (\mathbf{x} \mid \mathbf{y}_1)\, \mathbf{y}_1 + (\mathbf{x} \mid \mathbf{y}_2)\, \mathbf{y}_2 + \ldots + (\mathbf{x} \mid \mathbf{y}_n)\, \mathbf{y}_n$$

Transzformációk a lineáris térben

A transzformáció a függvényfogalom általánosítása.

Def.: **Transzformáció**

A T: $\boldsymbol{X}$->$\boldsymbol{Y}$ transzformáció az $\boldsymbol{X}$ lineáris tér $\boldsymbol{D}_T$ részhalmazának (értelmezési tartomány) egy $\mathbf{x}$ eleméhez az $\boldsymbol{Y}$ lineáris tér T($\mathbf{x}$)-el jelölt egy elemét rendeli.

Egy transzformáció lineáris, ha sorrendben felcserélhető a lineáris tér műveleteivel:

$$T(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha T(\mathbf{x}) + \beta T(\mathbf{y}) \quad \alpha, \beta \in \Phi \qquad \mathbf{x},\mathbf{y} \in \boldsymbol{D}_T$$

A transzformációk összegét $(T + S)(\mathbf{x}) = T(\mathbf{x}) + S(\mathbf{x})$

és skalárral vett szorzatát $(\alpha T)(\mathbf{x}) = \alpha(T(\mathbf{x}))$ $\qquad \mathbf{x} \in \boldsymbol{X}$ módon értelmezzük.

A T transzformáció inverze T^{-1}, ha $T^{-1}(T(\mathbf{x})) = \mathbf{x}$ minden $\mathbf{x} \in \boldsymbol{D}_T$ -re.

Lineáris transzformációk normája

Konvergencia-vizsgálatoknál szükségünk lesz a T lineáris transzformáció normájára. Itt szűkítünk a folytonos lineáris transzformációkra, amelyek $\boldsymbol{X}$-beli konvergens sorozathoz konvergens képsorozatot rendelnek.

Def.: **Folytonos transzformáció**

A T lineáris transzformáció folytonos az $\mathbf{x}_0 \in \boldsymbol{X}$-ben, ha minden $\in > 0$-hoz létezik $\delta=\delta(\in) > 0$ úgy, hogy $\|\mathbf{x} - \mathbf{x}_0\| < \delta(\in)$ esetén $\|T(\mathbf{x}) - T(\mathbf{x}_0)\| < \in$. Ez ekvivalens azzal, hogy bármely $\mathbf{x}_n \in \boldsymbol{X}$ pontsorozatra, amely konvergál $\mathbf{x}_0 \in \boldsymbol{X}$-hez, $T(\mathbf{x}_n) \to T(\mathbf{x}_0)$, ha $n \to \infty$.

Ha egy T lineáris transzformáció a $\varnothing$ pontban folytonos, akkor mindenütt folytonos. Ez könnyen belátható, mivel a T transzformációra $T\varnothing = T(\varnothing+\varnothing) = T\varnothing + T\varnothing$, azaz $T\varnothing = \varnothing$. Legyen $\mathbf{x}_n \to \mathbf{x}_0$ ha $n \to \infty$. Ekkor $T\mathbf{x}_n - T\mathbf{x}_0 = T(\mathbf{x}_n - \mathbf{x}_0) \to \varnothing$ akkor és csak akkor, ha T folytonos a $\varnothing$ pontban

Def.: **Korlátos transzformáció**

A T lineáris transzformáció korlátos, ha van olyan pozitív M szám, amelyre:

$$\|T(\mathbf{x})\| \le M\cdot\|\mathbf{x}\| \qquad \text{minden } \mathbf{x} \in \boldsymbol{X} \text{ esetén.}$$

Az analízisből közismert tételhez hasonlóan itt is kimondhatjuk, hogy a T lineáris transzformáció pontosan akkor folytonos, ha korlátos. A bizonyítás az analízisbelivel analóg módon történik, ezért mellőzzük.

Def.: **Folytonos lineáris transzformáció normája**

A folytonos lineáris transzformáció legkisebb korlátját a transzformáció normájának nevezzük.

$$\|T\| = \sup_{\|\mathbf{x}\|=1} \|T(\mathbf{x})\|$$

azaz az egységgömb T($\mathbf{x}$) képhalmazában a leghosszabb vektor hossza. Szemléletesen a $\|T\|$ annak a mértéke, hogy a T: $\mathbf{x}\to T(\mathbf{x})$ lineáris transzformáció mennyire nyújtja meg a vektorokat.

A definíció alapján az alábbi tulajdonságok nyilvánvalóak:

1./ $\|\lambda T\| = |\lambda|\cdot\|T\|$

2./ $\|T + S\| \le \|T\| + \|S\|$

3./ $\|T\cdot S\| \le \|T\|\cdot\|S\|$

4./ $\|T(\mathbf{x})\| \le \|T\|\cdot\|\mathbf{x}\|$, ezt a későbbiekben becslésekre fogjuk használni.

Lineáris transzformációk a szám n-esek lineáris terében, mátrixok

Mivel a lineáris transzformáció egyetlen vektort sem visz ki a térből, ezért az $\mathbf{e}_k$ vektort a

$$T(\mathbf{e}_k) = t_{1,k}\mathbf{e}_1 + t_{2,k}\mathbf{e}_2 + \dots + t_{n,k}\mathbf{e}_n = \sum_{i=1}^{n} t_{i,k}\mathbf{e}_i = \begin{bmatrix} t_{1,k} \\ t_{2,k} \\ \cdot \\ t_{n,k} \end{bmatrix}$$ vektorba transzformálja.

Végezzük el ezt a T transzformációt mindegyik bázisvektorra! A kapott $t_{i,k}$ koordinátákat a sorban következő számoszlopok egymás mellé írásával egy táblázatba foglalhatjuk, és ezt a T transzformáció mátrixának nevezzük, amit **T**-vel jelölünk.

$$\mathbf{T} = \begin{bmatrix} t_{1,1} & t_{1,2} & . & t_{1,n} \\ t_{2,1} & t_{2,2} & . & t_{2,n} \\ . & . & . & . \\ t_{n,1} & t_{n,2} & & t_{n,n} \end{bmatrix} = \left[t_{i,k}\right]_{n,n} \quad \text{ahol } \mathbf{T}_k = \mathrm{T}(\mathbf{e}_k) = \begin{bmatrix} t_{1,k} \\ t_{2,k} \\ . \\ t_{n,k} \end{bmatrix} \text{ a } \mathbf{T} \text{ mátrix k-adik oszlopa}$$

(oszlopvektora) az $\mathbf{e}_k$ bázisvektor transzformáltjának koordinátáit tartalmazza.

Mátrixok nem csak négyzetes fazonúak (ugyanannyi sor, ahány oszlop) hanem téglalap alakúak is lehetnek.

Itt bevezetjük a transzponált mátrix fogalmát.

Def.: **Mátrix transzponáltja**

Egy n sorral és m oszloppal rendelkező **T** mátrix transzponáltja $\mathbf{T}^*$, ha $t^*_{i,k} = t_{k,i}$ minden i,k párra (i=1,2,..,m ; k=1,2,.. ,n), azaz $\mathbf{T}^*$ nem más, mint a **T** mátrix <u>főátló</u>ra (*azonos indexű elemek által kijelölt átlóra*) való tükrözöttje.

Ebben az értelemben $\mathbf{T}^*_{\mathbf{i}}$ a **T** mátrix transzponáltjának i-edik oszlopvektora, azaz az eredeti **T** mátrix i-edik sorából képezett (oszlop)vektor.

Megjegyezzük, hogy a $\mathbf{T}^*$ jelölést a szakirodalomban a transzponált konjugáltjára (az adjungált transzformáció mátrixára ld. később) használják, de mi itt most csak valós komponensű mátrixokkal foglalkozunk, ezért itt $\mathbf{T}^*$ egyben a transzponáltat is jelenti.

Hogyan kell számolni mátrixokkal? Erre a kérdésre a lineáris tér eddig megismert szabályai adják meg a választ.

1. $(\lambda \mathrm{T})(\mathbf{e}_k) = \lambda \cdot \mathrm{T}(\mathbf{e}_k) = \lambda \cdot \sum_{i=1}^{n} t_{i,k}\mathbf{e}_i = \sum_{i=1}^{n} \lambda \cdot t_{i,k}\mathbf{e}_i$

tehát egy transzformáció "nyújtásakor" a mátrixának minden elemét meg kell szorozni a nyújtási tényezővel, azaz egy skalárral úgy szorzunk egy mátrixot, hogy minden elemét megszorozzuk.

2. $(\mathrm{T} + \mathrm{S})\,(\mathbf{e}_k) = \mathrm{T}(\mathbf{e}_k) + \mathrm{S}(\mathbf{e}_k) = \sum_{i=1}^{n} t_{i,k}\mathbf{e}_i + \sum_{i=1}^{n} s_{i,k}\mathbf{e}_i = \sum_{i=1}^{n} (t_{i,k} + s_{i,k})\mathbf{e}_i$

azaz két mátrix összegének elemei az elemek páronkénti összege.

3. $(\mathrm{ST})\,(\mathbf{e}_k) = \mathrm{S}(\mathrm{T}(\mathbf{e}_k)) = \mathrm{S}(\sum_{j=1}^{n} t_{j,k}\mathbf{e}_j) = \sum_{j=1}^{n} t_{j,k}\mathrm{S}(\mathbf{e}_j) = \sum_{j=1}^{n} t_{j,k}\sum_{i=1}^{n} s_{i,j}\mathbf{e}_i = \sum_{i=1}^{n}(\sum_{j=1}^{n} s_{i,j}t_{j,k})\mathbf{e}_i$

itt a zárójelen belül az **U** = **S·T** szorzatmátrix i-edik sorának k-adik elemét látjuk, amit az **S** mátrix i-edik sorának és a **T** mátrix k-adik oszlopának kompozíciójával (elempáronkénti szorzatok összege) kapunk. Ezt "skaláris szorzatként" is felírhatjuk.

$$u_{i,k} = \sum_{j=1}^{n} s_{i,j} t_{j,k} = (\mathbf{S}_i^* | \mathbf{T}_k)$$

A mátrixok szorzata asszociatív művelet, azaz

$$\mathbf{A}\cdot\mathbf{B}\cdot\mathbf{C} = (\mathbf{A}\cdot\mathbf{B})\cdot\mathbf{C} = \mathbf{A}\cdot(\mathbf{B}\cdot\mathbf{C})$$

Itt jegyezzük meg, hogy két mátrix szorzása általában nem kommutatív művelet!

Mivel $T(\mathbf{x}) = T(\sum_{k=1}^{n} x_k \mathbf{e}_k) = \sum_{k=1}^{n} x_k T(\mathbf{e}_k) = \sum_{k=1}^{n} x_k \sum_{j=1}^{n} t_{j,k} \mathbf{e}_j = \sum_{j=1}^{n} (\sum_{k=1}^{n} t_{j,k} x_k) \mathbf{e}_j$

ezért a transzformált vektort a $\mathbf{T}$ mátrix és az $\mathbf{x}$ vektor (mint 1 oszlopos mátrix) szorzataként kapjuk, amit az eddig megismert szabályok segítségével több alakban is felírhatunk:

$$T(\mathbf{x}) = \mathbf{T}\mathbf{x} = \begin{bmatrix} (\mathbf{T}_1^* | \mathbf{x}) \\ (\mathbf{T}_2^* | \mathbf{x}) \\ . \\ (\mathbf{T}_n^* | \mathbf{x}) \end{bmatrix} = \mathbf{T}_1 x_1 + \mathbf{T}_2 x_2 + ... + \mathbf{T}_n x_n$$

ahol az utolsó alak szerint a transzformált vektor a transzformációs mátrix oszlopvektorainak lineáris kombinációjaként is felírható.

Az $E(\mathbf{x}) = \mathbf{x}$ egységtranszformáció (helybenhagyás) mátrixa az **egységmátrix**:

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

Az egységmátrix főátlójában minden elem 1, azon kívül pedig 0

Nyilván $\mathbf{AE}=\mathbf{A}$ és $\mathbf{EA}=\mathbf{A}$

A négyzetes mátrixok szorzatának transzponáltját a következőképp számítjuk:

ha $\mathbf{U} = \mathbf{S}\cdot\mathbf{T}$, akkor $\mathbf{U}^* = (\mathbf{S}\cdot\mathbf{T})^* = \mathbf{T}^*\cdot\mathbf{S}^*$

Ez egyszerűen belátható, hiszen

$$u_{i,k} = \sum_{j=1}^{n} s_{i,j} t_{j,k} = (\mathbf{S}_i^* | \mathbf{T}_k) \text{ miatt } \quad u^*_{i,k} = u_{k,i} = (\mathbf{S}^*_k \,|\mathbf{T}_i\,) = (\mathbf{T}_i \,|\mathbf{S}^*_k\,) = ((\mathbf{T}^*)^*_i \,|\, \mathbf{S}^*_k)$$

A transzponált mátrixokra vonatkozó további szabályok triviálisak:

$$(\mathbf{T}^*)^* = \mathbf{T}; \qquad (\mathbf{S} + \mathbf{T})^* = \mathbf{S}^* + \mathbf{T}^*; \qquad (\lambda\mathbf{T})^* = \lambda\mathbf{T}^*; \quad \mathbf{E}^* = \mathbf{E};$$

A négyzetes mátrixokkal kapcsolatban szükségünk lesz még a mátrix nyoma fogalomra

Def.: A négyzetes mátrix főátlóbeli elemeinek szorzatát a **mátrix nyomának** nevezzük, és tr(**A**)-val (<u>trace</u> – nyom) jelöljük. (Német eredetű szóval a mátrix <u>spur</u>-járól is beszélhetünk)

$$\mathrm{tr}(\mathbf{A}) = \prod_{i=1}^{n} a_{i,i}$$

Ezután néhány megjegyzés következik.

a./ Az $(\mathbf{x} \mid \mathbf{y}) = \sum_{k=1}^{n} x_k y_k$ skaláris szorzatot szokásos módon $\mathbf{y}^*\mathbf{x}$ mátrixszorzásos alakba is írhatjuk.

b./ Felmerül a kérdés, hogy az A mátrixszal való szorzás hogyan vihető át a skaláris szorzat második tényezőjébe? Ehhez tekintsük a következő $(\mathbf{Ax} \mid \mathbf{y})$ skaláris szorzatot, ahol **A** valós, amit az a./ pont szerint mátrixszorzásos alakba is írhatunk:

$$(\mathbf{Ax} \mid \mathbf{y}) = \mathbf{y}^*(\mathbf{Ax}),$$

majd a jobboldali szorzat asszociativitása miatt

$$\mathbf{y}^*(\mathbf{Ax}) = \mathbf{y}^*\mathbf{Ax}$$

alakra, illetve a szorzat transzponáltjára vonatkozó $\mathbf{y}^*\mathbf{A} = (\mathbf{A}^*\mathbf{y})^*$ szabály miatt

$$\mathbf{y}^*\mathbf{Ax} = (\mathbf{A}^*\mathbf{y})^*\mathbf{x}$$

alakra hozhatunk. Ez utóbbit a $(\mathbf{x} \mid \mathbf{A}^*\mathbf{y})$ skaláris szorzat fazonra visszaírva kapjuk az

$$(\mathbf{Ax} \mid \mathbf{y}) = (\mathbf{x} \mid \mathbf{A}^*\mathbf{y})$$

összefüggést.

Mátrix normája

Def.: Egy lineáris transzformáció mátrixának normáját a transzformáció normájával azonosítjuk.

Nézzük meg ezt például Euklideszi-norma esetén:

A transzformált vektor normája a koordinátáinak négyzetösszegéből vont négyzetgyök:

$$\| T(x) \| = \| T(\sum_{k=1}^{n} x_k \mathbf{e}_k) \| = \| \sum_{k=1}^{n} x_k T(\mathbf{e}_k) \| = \| \sum_{k=1}^{n} x_k \sum_{i=1}^{n} t_{i,k} \mathbf{e}_k \| = \| \sum_{i=1}^{n} (\sum_{k=1}^{n} t_{i,k} x_k) \mathbf{e}_i \| =$$

$$= \left\{ \sum_{i=1}^{n} \left(\sum_{k=1}^{n} t_{i,k} x_k \right)^2 \right\}^{\frac{1}{2}}$$

A Cauchy-Bunyakovszkij-Schwartz egyenlőtlenség - lásd a szám-nesekre felírt alakját - felhasználásával pedig:

$$||T(\mathbf{x})|| \le \left\{ \sum_{i=1}^{n} \left(\sum_{k=1}^{n} t_{i,k}^2 \right) \left(\sum_{k=1}^{n} x_k^2 \right) \right\}^{\frac{1}{2}} = \left\{ \sum_{i=1}^{n} \sum_{k=1}^{n} t_{i,k}^2 \right\}^{\frac{1}{2}} ||\mathbf{x}||$$

Tehát ekkor

$$||\mathbf{T}||_F = \left\{ \sum_{i=1}^{n} \sum_{k=1}^{n} t_{i,k}^2 \right\}^{\frac{1}{2}}$$

a mátrix elemeinek négyzetösszegéből vont négyzetgyök értékét tekinthetnénk a **T** mátrix 2-es normájának. Ez valóban norma tulajdonságú, de nem tekinthetjük a 2-es vektornormából származtatottnak, mert

$$||\mathbf{E}||_F = \sqrt{n}$$

a várt 1 helyett, hiszen az $||\mathbf{Ex}|| = \mathbf{x}$ miatt $||\mathbf{E}|| = 1$ –nek kellene teljesülnie.

Megjegyezzük, hogy a 2-es normaként a **T*T** szorzatmátrix legnagyobb sajátértékéből vont négyzetgyök a megfelelő norma. (lásd később a sajátérték fejezetnél)

A korábban megismert másik vektornormákhoz pedig az alábbi mátrixnormák tartoznak:

$$||\mathbf{T}||_\infty = \max_i \sum_{j=1}^{n} |t_{i,j}| \quad \text{(max. sorösszeg)}$$

$$||\mathbf{T}||_1 = \max_j \sum_{i=1}^{n} |t_{i,j}| \quad \text{(max. oszlopösszeg)}$$

A mátrixnormák általános tulajdonságait a lineáris transzformációk normája c. részben leírt norma-tulajdonságok mátrixokra való kimondásával kapjuk meg (ld. ott)

Itt jegyezzük meg azt a fontos tételt, mely szerint ha egy normált térben egy sorozat konvergens, akkor más normát használva is az lesz. Ezt a tényt a konvergencia-vizsgálatoknál tudjuk jól hasznosítani.

Bázistranszformáció, mátrix inverze

Azt láttuk, hogy egy T transzformáció mátrixát úgy írjuk fel, hogy az oszlopaiba rendre a bázisvektorok transzformáltjainak koordinátái írjuk, azaz az első oszlopba $T(\mathbf{e}_1)$ koordinátáit és így tovább...

Most arra a kérdésre keressük a választ, hogy ha a bázisvektorok mindegyikét transzformáljuk, az így kapott vektorrendszer bázis lesz-e, és ha igen, akkor ebben az új bázisban egy régiben koordinátáival ismert vektor új bázisbeli koordinátáit hogyan kell kiszámolni.

Def : A T transzformáció nem szinguláris (más szóval <u>reguláris</u>), ha a lineáris tér minden bázisát bázisba transzformálja.

Ez azt jelenti, hogy a transzformáció nem csökkenti a dimenziószámot. Ekkor a transzformáció mátrixának oszlopvektorai lineárisan függetlenek. Természetesen vannak olyan transzformációk, amelyek csökkentik. Ilyen például a 3 dimenziós geometriai vektortérben egy adott ***S*** síkra (2 dimenziós altér) vetítés, vagy egy sík pontjainak vetítése annak egy egyenesére.

7. példa:

A $\mathbf{P} = \begin{bmatrix} \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \end{bmatrix}$ mátrixszal leírt transzformáció a háromdimenziós vektortérben bármely pontot az

origón átmenő $\mathbf{n}^* = (1,1,1)$ normálvektorú síkra vetít merőlegesen, hiszen az $\mathbf{r}_1 = (x_1,y_1,z_1)^*$ vektort $\mathbf{r}_2 = (x_2,y_2,z_2)^*$ -be viszi, ahol:

$$x_2 = (2x_1 - y_1 - z_1) / 3; \qquad y_2 = (- x_1 + 2y_1 - z_1) / 3; \qquad z_2 = (- x_1 - y_1 + 2z_1) / 3$$

Így $x_2 + y_2 + z_2 = 0$ adódik bármely pontra, azaz a transzformált valóban a $\mathbf{n}^* = (1,1,1)$ normálvektorú síkra vetít. A vetítés merőleges voltát az $(\mathbf{r}_2 - \mathbf{r}_1 \mid \mathbf{n}) = 0$ ortogonalitási feltétel teljesülésével ellenőrizhetjük.

Az is látható, hogy **P** oszlopvektorai nem függetlenek, hiszen összegük a **0** vektort adja. Az is ellenőrizhető, hogy **P** bármelyik hatványa **P**, ami nyilvánvaló, hiszen ha a síkra már rávetítettünk, akkor a további ugyanilyen vetítés a pont helyzetén már nem fog változtatni.

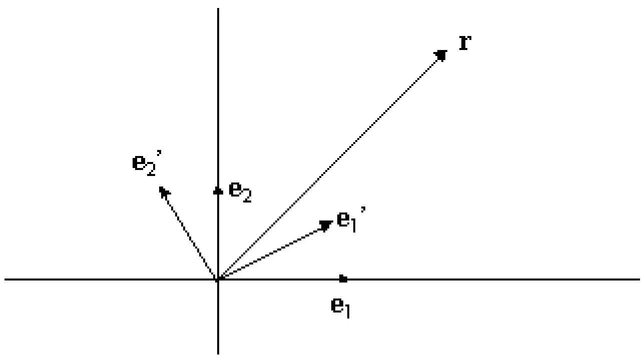
A bázistranszformációs problémánkat egy konkrét példával világítjuk meg.

Legyen ***X*** 2-dimenziós euklideszi tér és T_α az a transzformáció, mely minden elemet α szöggel elforgat. A $\boldsymbol{B} = \{\mathbf{e}_1, \mathbf{e}_2\}$ bázis legyen a triviális bázis. Ekkor az ábra alapján

$\mathbf{e}_1' = T_\alpha(\mathbf{e}_1) = \begin{bmatrix} \cos\alpha \\ \sin\alpha \end{bmatrix}$ $\qquad$ $\mathbf{e}_2' = T_\alpha(\mathbf{e}_2) = \begin{bmatrix} \text{-}\sin\alpha \\ \cos\alpha \end{bmatrix}$ $\qquad$ és így $\boldsymbol{B}' = \{\mathbf{e}_1', \mathbf{e}_2'\}$ is bázis lesz. (sőt a forgatás miatt az ortonormált tulajdonság is megmarad)

Tehát a T_α transzformáció mátrixa: $\mathbf{T}_\alpha = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}$

Az **r** vektor zárjon be az $\mathbf{e}_1$ irányával β szöget, ekkor ***B***-beli koordinátái: r·sinβ, r·cosβ
Ha a ***B***' bázisra térünk át, ott az **r** vektor az $\mathbf{e}_1$' iránnyal már csak (β - α) szöget fog bezárni.



Így a ***B***'-beli koordináták: $r_1' = r\cdot\cos(\beta - \alpha)$ és $r_2' = r\cdot\sin(\beta - \alpha)$ lesznek.

Ezeket kifejtve :

$r_1' = r\cdot\cos(\beta - \alpha) = r\cdot[\cos(\beta)\cos(\alpha) + \sin(\beta)\sin(\alpha)] = r\cdot[\cos(\beta)\cos(-\alpha) - \sin(\beta)\sin(-\alpha)]$

$r_2' = r\cdot\sin(\beta - \alpha) = r\cdot[\sin(\beta)\cos(\alpha) - \cos(\beta)\sin(\alpha)] = r\cdot[\sin(\beta)\cos(-\alpha) + \cos(\beta)\sin(-\alpha)]$

Mivel $\mathbf{T}_{-\alpha} = \begin{bmatrix} \cos(-\alpha) & -\sin(-\alpha) \\ \sin(-\alpha) & \cos(-\alpha) \end{bmatrix}$ ezért az **r**' a $\mathbf{T}_{-\alpha}$ mátrix és az **r** vektor szorzataként is felírható:

$$\mathbf{r}' = \mathbf{T}_{-\alpha}\,\mathbf{r}$$

Az pedig nyilvánvaló, hogy a -α szögű forgatás (a $T_{-\alpha}$ transzformáció) az α szöggel elforgatott bármely vektort visszaviszi a kiindulási helyzetébe, ezért a T_α transzformáció inverze: $T_\alpha^{-1} = T_{-\alpha}$ és ez a transzformációk mátrixaira is igaz lesz, tehát most írhatjuk, hogy ha egy bázistranszformáció **T** mátrixát (*ún. átmeneti mátrix*) ismerjük, akkor a transzformált új bázisban egy régiben adott vektor újbeli koordinátáit

$$\mathbf{r}' = \mathbf{T}^{-1}\cdot\mathbf{r}$$

módon számíthatjuk.

A bizonyításhoz csak azt kell meggondolnunk, hogy a **T** mátrix oszlopvektorai a bázisvektorok transzformáltjai: $\mathbf{T} = [T(\mathbf{e}_1) \;\; T(\mathbf{e}_2) \;\; . \;\; . \;\; T(\mathbf{e}_n)]$, és ezért az **r**' (az új bázisban felírt **r** vektor) az új bázisvektorok lineáris kombinációjaként adódik:

$$\mathbf{r} = T(\mathbf{e}_1)\mathbf{r}'_1 + T(\mathbf{e}_2)\mathbf{r}'_2 + \ldots + T(\mathbf{e}_n)\,\mathbf{r}'_n = \mathbf{T}\cdot\mathbf{r}'$$

ahonnét $\mathbf{T}^{-1}$-el szorozva kapjuk a bizonyítandó állítást.

Mivel a **T** mátrix oszlopvektorai az eredeti bázisvektorok transzformáltjai, ezért ezen oszlopvektorokra a T^{-1} transzformációt alkalmazva $\mathbf{e}_1$, $\mathbf{e}_2$,..., $\mathbf{e}_n$ vektorokat kapjuk vissza, amelyeket ha mátrixba foglalunk, akkor az egységmátrix adódik:

$$\mathbf{E} = [\ \mathbf{e}_1\ \mathbf{e}_2 \ldots \mathbf{e}_n] = \mathbf{T}^{-1} \cdot [T(\mathbf{e}_1)\ T(\mathbf{e}_2)\ \ldots\ T(\mathbf{e}_n)] = \mathbf{T}^{-1} \cdot \mathbf{T}$$

Azt is kimondhatjuk, hogy ha egy transzformáció nem szinguláris, azaz bázist bázisba visz, akkor bármely vektor ebben az új bázisban is felírható lesz a bázisvektorok lineáris kombinációjaként, és mivel a $T^{-1} \cdot T$ transzformáció az egységtranszformáció, ezért mátrixaik szorzata is az egységmátrix lesz. Ez azt is jelenti, hogy ha a

$$\boldsymbol{B} = \{\ T(\mathbf{e}_1)\quad T(\mathbf{e}_2)\quad .\quad .\quad T(\mathbf{e}_n)\,\}$$

bázisban rendre felírjuk az $\mathbf{e}_1$, $\mathbf{e}_2$, ... $\mathbf{e}_n$ vektorok koordinátáit, akkor éppen a $\mathbf{T}^{-1}$ mátrix oszlopvektoraihoz jutunk.

Tehát egy mátrix inverzét a transzformáció inverzének (ha létezik) mátrixával azonosítjuk, és többek között egy vektornak a transzformált új bázisbeli koordinátáinak kiszámítására is használhatjuk.

A mátrix inverzét elemi bázistranszformációs lépésekkel a **T** mátrixból kiindulva számíthatjuk ki. Ugyanis ha a **T** mátrixot a $T(\mathbf{e}_k)$ k=1,2, ...,n oszlopvektorokból egymás mellé írott alaknak tekintjük, melyek koordinátáit a $\{\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n\}$ bázisban látjuk, akkor ha rendre kicseréljük az $\mathbf{e}_k$ bázisvektorokat a $T(\mathbf{e}_k)$ vektorokra, az előbb elmondottak szerint éppen a $\mathbf{T}^{-1}$ mátrix oszlopvektoraihoz jutunk.

Egy elemi bázistranszformációnál pedig - a korábbi tanulmányaikból ismert módon - ha a generáló elem $t_{k,k}$:

1. a generáló elem sorát osztjuk a generáló elemmel:
$$t_{k,j}' = t_{k,j} / t_{k,k} \quad j = 1,2,...,n \quad j \neq k$$

2. a generáló elem oszlopát osztjuk a generáló elem -1 –szeresével:
$$t_{i,k}' = - t_{i,k} / t_{k,k} \quad i = 1,2,...,n \quad i \neq k$$

3. minden további új $t_{i,j}'$ elemet a
$$t_{i,j}' = t_{i,j} - (t_{i,k} \cdot t_{k,j}) / t_{k,k} \qquad i = 1,2,...,n \quad i \neq k \text{ és } j = 1,2,...,n \quad j \neq k$$

képlettel számolunk ki,

4. Végezetül a generáló elem helyére a reciprokát írjuk:
$$t_{k,k}' = 1 / t_{k,k}$$

Ha a generáló elemet (elemeket) nem a főátlóból választjuk, akkor amikor már minden vektorcserét végrehajtottunk, a mátrixunk sorainak sorrendjét a
$$\boldsymbol{B} = \{\ T(\mathbf{e}_1)\quad T(\mathbf{e}_2)\quad .\quad .\quad T(\mathbf{e}_n)\,\}$$
bázisnak megfelelőre változtatjuk, és az oszlopok sorrendjét is az eredeti $\{\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n\}$ bázisnak megfelelőre változtatjuk annak érdekében, hogy a mátrix inverzét megkapjuk.

A két nem szinguláris mátrix szorzatának inverzét úgy számítjuk ki, hogy a tényezők inverzét fordított sorrendben szorozzuk össze:

ha $\quad \mathbf{U} = \mathbf{S}\cdot\mathbf{T}$, akkor $\mathbf{U}^{-1} = (\mathbf{S}\cdot\mathbf{T})^{-1} = \mathbf{T}^{-1}\cdot\mathbf{S}^{-1}$

Valóban, hiszen: $\quad \mathbf{U}\cdot\mathbf{U}^{-1} = \mathbf{S}\cdot\mathbf{T}\cdot\mathbf{T}^{-1}\cdot\mathbf{S}^{-1} = \mathbf{S}\cdot\mathbf{E}\cdot\mathbf{S}^{-1} = \mathbf{S}\cdot\mathbf{S}^{-1} = \mathbf{E}$

és $\quad \mathbf{U}^{-1}\cdot\mathbf{U} = \mathbf{T}^{-1}\cdot\mathbf{S}^{-1}\cdot\mathbf{S}\cdot\mathbf{T} = \mathbf{T}^{-1}\cdot\mathbf{E}\cdot\mathbf{T} = \mathbf{T}^{-1}\cdot\mathbf{T} = \mathbf{E}$

Itt kihasználtuk azt a tényt, hogy a nem szinguláris mátrixok inverze egyértelműen létezik.

Felmerül a kérdés: mit tehetünk akkor, ha a mátrixunk szinguláris, sőt nem is négyzetes alakú?

Erre a kérdésre a **pszeudóinverz** *(kváziinverz*) fogalmának a bevezetése adja meg a választ.

Alapproblémaként tekintsük azt az n ismeretlenes lineáris egyenletrendszert, melyben az együtthatómátrix sorainak száma kevesebb, mint az ismeretlenek száma:

Legyenek $\mathbf{a}^*_1$, $\mathbf{a}^*_2$, ..., $\mathbf{a}^*_m$ lineárisan független sorvektorok ($m<n$) , az együtthatómátrix sorvektorai és **b** az egyenletrendszer jobboldali vektora. Ekkor az $\mathbf{A}\mathbf{x} = \mathbf{b}$ egyenletrendszer

$$(\mathbf{x} \mid \mathbf{a}^*_i) = b_i \quad i = 1,2, \ldots ,m$$

alakba írható. Azt tudjuk, hogy ilyenkor végtelen sok megoldása létezik, de mi válasszuk ki ezek közül azt, melynek a normája minimális. A projekciótétel alapján bizonyítható, hogy a minimalizáló megoldás

$\mathbf{x} = \sum_{k=1}^{m} \xi_k \mathbf{a}^*_k = \mathbf{A}\mathbf{x}'$ alakú lesz, azaz benne van az **A** sorvektorainak lineáris alterében.

Itt a ξ_k értékek meghatározásához helyettesítsünk be az eredeti egyenletbe:

$\sum_{k=1}^{m} \xi_k (\mathbf{a}^*_k | \mathbf{a}^*_i) = b_i$ Ennek az egyenletrendszernek a mátrixát Gram-mátrixnak hívjuk.

Az $\mathbf{A}\cdot\mathbf{A}^*$ Gram mátrix-szal felírt $\mathbf{A}\cdot\mathbf{A}^*\mathbf{x}' = \mathbf{b}$ egyenletrendszer $\mathbf{x}' = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \cdot \\ \xi_m \end{bmatrix}$ megoldásával

$\mathbf{x} = \mathbf{A}^*\mathbf{x}'$ az eredeti egyenlet megoldása lesz, ami a projekció tétel miatt minimális normájú. Így a keresett megoldás:

$$\mathbf{x} = \mathbf{A}^*(\mathbf{A}\mathbf{A}^*)^{-1}\,\mathbf{b}$$

Ebből leolvashatjuk, hogy az **A** mátrix - melynek több oszlopa van, mint sora - kváziinverze:

$$\mathbf{A}^*(\mathbf{A}\mathbf{A}^*)^{-1}$$

Hasonlóan a több sorral, mint oszloppal (m >n) rendelkező **A** mátrixnak a kváziinverze:

$$(\mathbf{A}^*\mathbf{A})^{-1}\mathbf{A}^*$$

Ez abból következik, hogy ekkor - ha a **b** vektor nincs az **A** mátrix oszlopvektorainak alterében, és ez többnyire így is van - az **Ax** = **b** egyenletrendszernek nincs is megoldása, de mégis olyan **x'** vektort keresünk, melyre ||**Ax'** - **b**|| minimális. A projekciótétel miatt erre az **x'** vektorra igaz az, hogy ortogonális az **A** oszlopvektorainak alterére, azaz annak minden egyes vektorára is:

$$(\mathbf{A}\mathbf{x}' - \mathbf{b} \mid \mathbf{a}_k) = 0 \quad k = 1,2, ..., n$$

Mátrixos írásmóddal ez pedig a következő egyenletrendszert jelenti:

$$(\mathbf{A}^*\mathbf{A})\,\mathbf{x}' = \mathbf{A}^*\mathbf{b}$$

Mivel lineárisan független $\mathbf{a}_k$ oszlopvektorok esetén az $\mathbf{A}^*\mathbf{A}$ Gram mátrixnak van inverze, ezért:

$$\mathbf{x}' = (\mathbf{A}^*\mathbf{A})^{-1}\mathbf{A}^*\mathbf{b}$$

Jelöljük az **A** mátrix kváziinverzét **X**-el. Könnyen ellenőrizhetők a következő tulajdonságok:

AXA = A, **XAX = X**, és
AX mátrix és **XA** mátrix is szimmetrikus.

Speciális mátrixok

1. Az **E** egységmátrix sorainak, ill. oszlopainak felcserélésével kapott mátrixot **permutáló mátrixnak** hívjuk:

pl.:

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

a **P·A** szorzat eredményében az **A** mátrix sorai más sorrendben szerepelnek (itt az 1. és 2. sor fel fog cserélődni), az **A· P** szorzat eredményében az oszlopok cserélődnek fel (itt szintén az 1. és a 2. oszlop cserélődik)

Azokat a mátrixokat, amelyek egy mátrix sorait, illetve oszlopait eggyel tovább tolják, **ciklikus permutáló mátrixoknak** hívjuk.

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

A ciklikus permutáló mátrixok n-edik hatványa **E**.

2. Idempotens mátrixok, projektorok

A 7. példában közölt **P** mátrix síkba vetítést írt le:

$$\mathbf{P} = \begin{bmatrix} \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \end{bmatrix}$$

Kiemeltük azt a tulajdonságát, hogy $\mathbf{P}^2 = \mathbf{P}$, azaz akárhányadik hatványa saját maga. Az ilyen tulajdonságú mátrixokat projektoroknak nevezzük.

Az egységmátrix kivételével minden projektor alacsonyabb dimenziós térbe transzformál.

3. **Háromszögmátrixok**:

Ha egy mátrix főátló alatti elemei mind zérusok, akkor felső háromszögmátrixnak, ha pedig a főátló feletti elemei zérusok, akkor alsó háromszögmátrixnak nevezzük.

Nézzük meg, az alábbi alsó háromszögmátrix milyen transzformációt valósít meg?

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

A szorzás végrehajtásával ellenőrizhető, hogy a **L·A** szorzatban, annyi a változás az **A** mátrixhoz képest, hogy a harmadik sorhoz az első sor 2-szerese adódott, a negyedik sor pedig megháromszorozódott. Ilyen transzformációkat hajtunk végre egy lineáris egyenletrendszer mátrixán a korábbi tanulmányainkból ismert ún. ismeretlenek kiküszöbölési eljárása közben.

Nem nehéz belátni, hogy alsó háromszögmátrixok szorzata is alsó háromszögmátrix. Hasonlót mondhatunk a felső háromszögmátrixok szorzatára.

Ha egy háromszögmátrix nyoma nem 0, akkor invertálható, és az inverze is háromszögmátrix. A példánkbeli **L** mátrix inverze:

$$\mathbf{L}^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -2 & 0 & 1 & 0 \\ 0 & 0 & 0 & \frac{1}{3} \end{bmatrix}$$

Ez nyilvánvaló, hiszen a harmadik sorból az első kétszeresét kell elvenni és a negyedik sort harmadolni kell, hogy az eredeti mátrixhoz jussunk vissza.

Példánkból is látszik, és általánosan is igaz, hogy egy alsó háromszögmátrix inverze is alsó háromszögmátrix.

Egy nem szinguláris négyzetes mátrix mindig felbontható egy alsó és egy felső háromszögmátrix szorzatára (**LU** *felbontás)*

Ezt a következő példánkon illusztráljuk.

8. példa:

Megfelelő transzformációk segítségével állítsuk elő az **A** mátrix **LU** felbontását!

$$\mathbf{A} = \begin{bmatrix} 2 & 6 & 4 \\ 4 & 15 & 2 \\ -1 & -1 & -5 \end{bmatrix}$$

Az eljárás során az ún. Gauss kiküszöbölés lépéseit mátrixszorzások segítségével valósítjuk meg, és ezen mátrixok inverzeit is felhasználjuk.

Először az első sort $a_{1,1} = 2$ –vel elosztjuk, az ezt végrehajtó T_1 transzformáció mátrixa és az inverze:

$$\mathbf{T}_1 = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{és} \quad \mathbf{T}_1^{-1} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{így} \quad \mathbf{T}_1\mathbf{A} = \begin{bmatrix} 1 & 3 & 2 \\ 4 & 15 & 2 \\ -1 & -1 & -5 \end{bmatrix}$$

Ezután az első sor $a_{2,1} = -4$ –szeresét adjuk a második sorhoz, és az első sor $a_{3,1} = 1$ –szeresét adjuk a harmadik sorhoz, az ezt megvalósító T_2 transzformáció és inverzének mátrixa:

$$\mathbf{T}_2 = \begin{bmatrix} 1 & 0 & 0 \\ -4 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \quad \text{és} \quad \mathbf{T}_2^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \quad \text{így} \quad \mathbf{T}_2\mathbf{T}_1\mathbf{A} = \begin{bmatrix} 1 & 3 & 2 \\ 0 & 3 & -6 \\ 0 & 2 & -3 \end{bmatrix}$$

A következő lépésben a második sort $a_{2,2}' = 3$ –al elosztjuk, a megfelelő transzformáció:

$$\mathbf{T}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{és} \quad \mathbf{T}_3^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{így} \quad \mathbf{T}_3\mathbf{T}_2\mathbf{T}_1\mathbf{A} = \begin{bmatrix} 1 & 3 & 2 \\ 0 & 1 & -2 \\ 0 & 2 & -3 \end{bmatrix}$$

Végezetül a harmadik sorból a második $a_{3,2}' = 2$ –szeresét elvesszük:

$$\mathbf{T}_4 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix} \quad \text{és} \quad \mathbf{T}_4^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix} \quad \text{így} \quad \mathbf{T}_4\mathbf{T}_3\mathbf{T}_2\mathbf{T}_1\mathbf{A} = \begin{bmatrix} 1 & 3 & 2 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix}$$

Mivel $\mathbf{T}_1^{-1}\mathbf{T}_2^{-1}\mathbf{T}_3^{-1}\mathbf{T}_4^{-1}\mathbf{T}_4\mathbf{T}_3\mathbf{T}_2\mathbf{T}_1\mathbf{A} = \mathbf{A}$ ezért a keresett felbontást már meg is kaptuk:

$\mathbf{L} = \mathbf{T}_1^{-1}\mathbf{T}_2^{-1}\mathbf{T}_3^{-1}\mathbf{T}_4^{-1}$ és $\mathbf{U} = \mathbf{T}_4\mathbf{T}_3\mathbf{T}_2\mathbf{T}_1\mathbf{A}$

Valóban:

$$\mathbf{LU} = \begin{bmatrix} 2 & 0 & 0 \\ 4 & 3 & 0 \\ -1 & 2 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 3 & 2 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 6 & 4 \\ 4 & 15 & 2 \\ -1 & -1 & -5 \end{bmatrix}$$

Az **U** mátrix főátlójában most csupa 1-es szerepel. Az ilyen felső háromszögmátrixokat felső egység-háromszögmátrixnak nevezzük.

Egy **A** mátrix **LU** felbontása többféleképp is lehetséges, az előző példánkban a főátlóbeli elemekkel a saját soruk végigosztása eredményezte azt, hogy a felső háromszögmátrix főátlójába csupa 1-es került, ha ezeket a lépéseket kihagyjuk, akkor az alsó háromszögmátrix főátlójában lesz minden elem 1-es (alsó egység-háromszögmátrix). Természetesen ekkor a generáló elemek sorának a megfelelő számszorosát kell az alatta lévő sorokból kivonni a kiküszöbölés érdekében:

$$\mathbf{T}_1 = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ \frac{1}{2} & 0 & 1 \end{bmatrix} \quad \text{és} \quad \mathbf{T}_1^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -\frac{1}{2} & 0 & 1 \end{bmatrix} \quad \text{így} \quad \mathbf{T}_1\mathbf{A} = \begin{bmatrix} 2 & 6 & 4 \\ 0 & 3 & -6 \\ 0 & 2 & -3 \end{bmatrix}$$

$$\mathbf{T}_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\frac{2}{3} & 1 \end{bmatrix} \quad \text{és} \quad \mathbf{T}_2^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \frac{2}{3} & 1 \end{bmatrix} \quad \text{így} \quad \mathbf{T}_2\mathbf{T}_1\mathbf{A} = \begin{bmatrix} 2 & 6 & 4 \\ 0 & 3 & -6 \\ 0 & 0 & 1 \end{bmatrix}$$

és ezzel $\mathbf{T}_1^{-1}\mathbf{T}_2^{-1}\mathbf{T}_2\mathbf{T}_1\mathbf{A} = \mathbf{A}$ miatt

$$\mathbf{LU} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -\frac{1}{2} & \frac{2}{3} & 1 \end{bmatrix} \cdot \begin{bmatrix} 2 & 6 & 4 \\ 0 & 3 & -6 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 6 & 4 \\ 4 & 15 & 2 \\ -1 & -1 & -5 \end{bmatrix}$$

egy másik lehetséges felbontás adódik.

Az $\mathbf{T}_1$, $\mathbf{T}_2$ mátrixok mindegyike egy ún. eliminációs mátrix. Ezeket az jellemzi, hogy az egységmátrixtól csak annyiban térnek el, hogy valamelyik, pl. k-adik oszlopukban szereplő nullák helyett más érték is szerepelhetnek:

$$\mathbf{T}_k = \mathbf{E} + \mathbf{t}^{(k)}\mathbf{e}_k^* \quad ahol \quad \mathbf{t}_k^{(k)} = 0$$

A példáinkból látszik és egyszerűen ellenőrizhető, hogy egy eliminációs mátrix inverze a főátlón kívüli elemek előjelfordításával megkapható:

$$\mathbf{T}_k^{-1} = \mathbf{E} - \mathbf{t}^{(k)}\mathbf{e}_k^*$$

Az eliminációs (kiküszöböléses) technika egy másik alkalmazását nézzük meg ezután.

4. Mátrixok diadikus felbontása

Az A mátrix **LU** felbontásából a mátrixnak egy ún. diadikus felbontása is kiolvasható. A diadikus felbontás azt jelenti, hogy a mátrix diádok (oszlopvektor-sorvektor szorzatok, más szóval diadikus szorzatok) összegére bontható. Egy ilyen triviális felbontás, ha a mátrixot úgy tekintjük, mint az oszlopvektorainak sorvektorát, és megszorozzuk az egységmátrix sorvektorainak oszlopvektorával:

$$\mathbf{A} = \mathbf{AE} = [\mathbf{a}_1, \mathbf{a}_2 \ldots \mathbf{a}_n] \cdot \begin{bmatrix} \mathbf{e}_1^* \\ \mathbf{e}_2^* \\ \vdots \\ \mathbf{e}_n^* \end{bmatrix} = \sum_{k=1}^{n} \mathbf{a}_k \mathbf{e}_k^*$$

Eliminációs technikával is előállíthatunk (egy másik) diadikus felbontást. Vonjuk le az **A** mátrixból az első oszlop $1/a_{1,1}$-szeresének és az első sornak diadikus szorzatát!

Legyen $\overset{(1)}{\mathbf{A}} = \mathbf{A}$

$$\overset{(2)}{\mathbf{A}} = \overset{(1)}{\mathbf{A}} - \frac{1}{\mathbf{e}_1^* \overset{(1)}{\mathbf{A}} \mathbf{e}_1} \cdot (\overset{(1)}{\mathbf{A}} \mathbf{e}_1) \cdot (\mathbf{e}_1^* \overset{(1)}{\mathbf{A}}) = \overset{(1)}{\mathbf{A}} - \frac{1}{\mathbf{a}_{1,1}} \cdot (\overset{(1)}{\mathbf{A}} \mathbf{e}_1) \cdot (\mathbf{e}_1^* \overset{(1)}{\mathbf{A}}) \quad \text{ezzel}$$

$$\overset{(2)}{\mathbf{A}} = \begin{bmatrix} 2 & 6 & 4 \\ 4 & 15 & 2 \\ -1 & -1 & -5 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 2 \\ 4 \\ -1 \end{bmatrix} [2 \quad 6 \quad 4] = \begin{bmatrix} 2 & 6 & 4 \\ 4 & 15 & 2 \\ -1 & -1 & -5 \end{bmatrix} - \begin{bmatrix} 1 \\ 2 \\ -\frac{1}{2} \end{bmatrix} [2 \quad 6 \quad 4] =$$

$$= \begin{bmatrix} 2 & 6 & 4 \\ 4 & 15 & 2 \\ -1 & -1 & -5 \end{bmatrix} - \begin{bmatrix} 2 & 6 & 4 \\ 4 & 12 & 8 \\ -1 & -3 & -2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 3 & -6 \\ 0 & 2 & -3 \end{bmatrix}$$

Ezt is tovább bontható:

$$\overset{(3)}{\mathbf{A}} = \overset{(2)}{\mathbf{A}} - \frac{1}{\mathbf{e}_2^* \overset{(2)}{\mathbf{A}} \mathbf{e}_2} \cdot (\overset{(2)}{\mathbf{A}} \mathbf{e}_2) \cdot (\mathbf{e}_2^* \overset{(2)}{\mathbf{A}}) = \overset{(2)}{\mathbf{A}} - \frac{1}{\mathbf{a}_{2,2}} \cdot (\overset{(2)}{\mathbf{A}} \mathbf{e}_2) \cdot (\mathbf{e}_2^* \overset{(2)}{\mathbf{A}})$$

$$\overset{(3)}{\mathbf{A}} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 3 & -6 \\ 0 & 2 & -3 \end{bmatrix} - \frac{1}{3} \begin{bmatrix} 0 \\ 3 \\ 2 \end{bmatrix} \begin{bmatrix} 0 & 3 & -6 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 3 & -6 \\ 0 & 2 & -3 \end{bmatrix} - \begin{bmatrix} 0 \\ 1 \\ \frac{2}{3} \end{bmatrix} \begin{bmatrix} 0 & 3 & -6 \end{bmatrix} =$$

$$= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 3 & -6 \\ 0 & 2 & -3 \end{bmatrix} - \begin{bmatrix} 0 & 0 & 0 \\ 0 & 3 & -6 \\ 0 & 2 & -4 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}$$

Mindezek alapján az **A** mátrix diadikus felbontása:

$$\mathbf{A} = \begin{bmatrix} 1 \\ 2 \\ -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 2 & 6 & 4 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ \frac{2}{3} \end{bmatrix} \begin{bmatrix} 0 & 3 & -6 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}$$ amiből a diádok összevonásával éppen

a hozzátartozó és megfelelő **LU** felbontást kapjuk:

$$\mathbf{LU} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -\frac{1}{2} & \frac{2}{3} & 1 \end{bmatrix} \cdot \begin{bmatrix} 2 & 6 & 4 \\ 0 & 3 & -6 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 6 & 4 \\ 4 & 15 & 2 \\ -1 & -1 & -5 \end{bmatrix}$$

Egy **u** egységvektor sajátmagával vett diadikus szorzata a saját irányára vetítés mátrixát eredményezi, ugyanis ha

$\mathbf{T} = \mathbf{u}\mathbf{u}^*$ mátrixot megvizsgáljuk, ahol $\mathbf{u}^*\mathbf{u} = 1$

akkor $\mathbf{y} = \mathbf{T}\mathbf{x} = \mathbf{u}\mathbf{u}^*\mathbf{x}$. Vizsgáljuk meg az **y** képvektor i-edik koordinátáját!

$$y_i = \sum_{k=1}^{n} u_i u_k x_k = \left(\sum_{k=1}^{n} u_k x_k \right) u_i = (\mathbf{u}^*\mathbf{x}) u_i$$

Emiatt $\mathbf{u}\mathbf{u}^*\mathbf{x} = (\mathbf{u}^*\mathbf{x})\mathbf{u} = (\mathbf{u}|\mathbf{x})\mathbf{u}$, ahol mint tudjuk $(\mathbf{u}|\mathbf{x})\mathbf{u}$ az **x** vektornak az **u** irányára való vetülete.

5. Szalagmátrixok, sávmátrixok, Hessenberg-alakú mátrixok

Egy **A** mátrixot akkor nevezünk m sávszélességű szalagmátrixnak, ha létezik $1 \leq m < n-1$ úgy, hogy a főátlótól m-nél „messzebb" lévő elemek mind zérusok, azaz

$$a_{i,j} = 0 \text{ , ha } |i - j| > m$$

Ha a sávszélesség 1, akkor **tridiagonális mátrix**ról beszélünk. A tridiagonális mátrixok **kontinuáns** mátrix néven is előfordulnak. A főátló melletti átlókat alsó (i=j+1) és felső (i+1=j) mellékátlónak nevezzük.

Egy mátrixot **felső Hessenberg-mátrixnak** vagy felső Hessenberg-alakúnak nevezünk, ha az alsó mellékátló alatti elemei csupa 0-ák, azaz

$$a_{i,j} = 0 \text{ , ha } i > j+1$$

6. **Pozitív szemidefinit mátrix**: ha a komplex szám-nesek minden térbeli **x** vektorára eleget tesz a

$$(\mathbf{Ax} \mid \mathbf{x}) \geq 0$$

egyenlőtlenségnek Ha a szigorúbb egyenlőtlenség is teljesül minden $\mathbf{x} \neq \mathbf{0}$ vektorra, akkor a mátrix pozitív definit.

7. Unitér mátrix, ortogonális mátrix

Azokat a komplex elemű mátrixokat, amelyekre a mátrix inverze a transzponált konjugáltja, azaz $\mathbf{U}^{-1} = \overline{\mathbf{U}}^*$ teljesül, unitér mátrixoknak hívjuk, valós mátrixok esetében pedig

$$\mathbf{U}^{-1} = \mathbf{U}^* \qquad \text{azaz} \qquad \mathbf{U}^*\mathbf{U} = \mathbf{E} \qquad \text{fennállása esetén}$$

ortogonális (ortonormált) mátrixról beszélünk. Az ortonormált mátrixok oszlopvektorai páronként ortogonálisak egymásra és normájuk egységnyi, hiszen az $\mathbf{U}^*\mathbf{U} = \mathbf{E}$ összefüggés azt jelenti, hogy $\mathbf{u}^*_j \cdot \mathbf{u}_i = (\mathbf{u}_i \mid \mathbf{u}_j) = \delta_{i,j}$

Ha az $\mathbf{U}^*\mathbf{U} = \mathbf{E}$ összefüggést balról megszorozzuk az **U** mátrixszal és bal oldalra rendezzük, majd felhasználjuk az $\mathbf{UE} = \mathbf{EU}$ nyilvánvaló összefüggést:

$$(\mathbf{UU}^* - \mathbf{E})\,\mathbf{U} = 0$$

Emiatt $(\mathbf{UU}^* - \mathbf{E})$ sorvektorai mind ortogonálisak az $\mathbf{u}_j$ vektorokra, így ezek a sorvektorok mind **0** vektorok, tehát:

$$\mathbf{UU}^* = \mathbf{E}$$

Ezek szerint, ha **U** oszlopvektorai ortonormáltak, akkor a sorvektorai is azok.

Ortonormált mátrix esetén

$$||\mathbf{Ux}||^2 = (\mathbf{Ux} \mid \mathbf{Ux}) = (\mathbf{x} \mid \mathbf{U}^*\mathbf{Ux}) = (\mathbf{x} \mid \mathbf{Ex}) = (\mathbf{x} \mid \mathbf{x}) = ||\mathbf{x}||^2$$

látható, hogy az **U** mátrixhoz tartozó U leképezés nem változtatja meg a távolságot (normát), emiatt az **U** mátrixhoz tartozó U leképezéseket forgatásoknak (ill. tükrözéseknek) nevezzük.

Ortogonális mátrixok szorzata is ortogonális mátrix lesz.

A bázistranszformációt ismertető részben említett $\mathbf{T}_\alpha$ forgatómátrix is ortogonális.

Bármely ortogonális mátrix determinánsának négyzete egységnyi, mert

$$1 = |\mathbf{E}| = |\mathbf{Q}\mathbf{Q}^{-1}| = |\mathbf{Q}\mathbf{Q}^{*}| = |\mathbf{Q}|^2$$

Azokat az ortogonális transzformációkat, melyek mátrixának determinánsa +1, valódi ortogonális transzformációnak (forgatás), azokat melyeké –1, pedig nemvalódi ortogonális transzformációnak nevezzük.

Van egy bizonyos típusú ortogonális mátrix, amihez tartozó transzformációt nem forgatásnak, hanem elemi tükrözésnek (Householder transzformáció) hívunk. Ez minden vektort az origón átmenő **q** normálvektorú síkra (hipersíkra) tükröz, ahol

$$\|\mathbf{q}\|_2 = 1, \text{ azaz } \mathbf{q}^*\mathbf{q} = 1.$$

Mátrixa:

$$\mathbf{Q} = \mathbf{E} - 2\mathbf{q}\mathbf{q}^*$$

alakú, ami nyilván szimmetrikus:

$$\mathbf{Q}^* = (\mathbf{E} - 2\mathbf{q}\mathbf{q}^*)^* = \mathbf{E}^* - 2(\mathbf{q}^*)^*\mathbf{q}^* = \mathbf{E} - 2\mathbf{q}\mathbf{q}^* = \mathbf{Q} \quad \text{(emiatt } \mathbf{Q} \text{ inverze saját maga!)}$$

és ortogonális:

$$\mathbf{Q}^*\mathbf{Q} = \mathbf{Q}\mathbf{Q} = (\mathbf{E} - 2\mathbf{q}\mathbf{q}^*)(\mathbf{E} - 2\mathbf{q}\mathbf{q}^*) = \mathbf{E} - 2\mathbf{q}\mathbf{q}^* - 2\mathbf{q}\mathbf{q}^* + 4\mathbf{q}\mathbf{q}^*\mathbf{q}\mathbf{q}^* = \mathbf{E} - 4\mathbf{q}\mathbf{q}^* + 4\mathbf{q}(\mathbf{q}^*\mathbf{q})\mathbf{q}^* = \mathbf{E}$$

Azt, hogy a tükrözés mátrixa a fenti alakú, könnyen beláthatjuk. Legyen adott az **u** normált vektor és legyen **x** tetszőleges vektor. Ha az **x**-et egy **u** normálvektorú síkra tükrözzük, akkor **x** végpontjából az **u** irányában valamennyit el kell mozdulni.

Legyen tehát a tükrözött vektor $\mathbf{x} - \lambda\mathbf{u}$ alakú. Ennek a tükrözés miatt a normája nem változik:

$$(\mathbf{x} \mid \mathbf{x}) = (\mathbf{x} - \lambda\mathbf{u} \mid \mathbf{x} - \lambda\mathbf{u})$$

A skaláris szorzatot felbontva és felhasználva **u** normált voltát $\lambda = 2(\mathbf{u} \mid \mathbf{x}) = 2\mathbf{u}^*\mathbf{x}$ adódik. Tehát a tükörkép:

$$\mathbf{x} - 2(\mathbf{u}^*\mathbf{x})\mathbf{u} = \mathbf{x} - 2\mathbf{u}\mathbf{u}^*\mathbf{x} = \mathbf{E}\mathbf{x} - 2\mathbf{u}\mathbf{u}^*\mathbf{x} = (\mathbf{E} - 2\mathbf{u}\mathbf{u}^*)\mathbf{x} = \mathbf{Q}\mathbf{x}$$

Itt kihasználtuk az $\mathbf{u}\mathbf{u}^*$ diadikus szorzat (**u** irányvektorú egyenesre vetítési transzformáció) $\mathbf{u}\mathbf{u}^*\mathbf{x} = (\mathbf{u}^*\mathbf{x})\mathbf{u}$ tulajdonságát.

Példaként tekintsük a $\mathbf{q}^* = [\ 1/3\ \ 2/3\ \ 2/3\ ]$ normált vektort. A segítségével felírt

$$\mathbf{Q} = \begin{bmatrix} \frac{7}{9} & -\frac{4}{9} & -\frac{4}{9} \\ -\frac{4}{9} & \frac{1}{9} & -\frac{8}{9} \\ -\frac{4}{9} & -\frac{8}{9} & \frac{1}{9} \end{bmatrix}$$ mátrix elemi tükrözést valósít meg.

Ahogy a példánkból is kiszámítható, az elemi tükrözést végrehajtó ortogonális transzformáció mátrixának determinánsa: -1

8. Reguláris mátrixok QR felbontása

Az ortogonális mátrixok fontos szerepet kapnak a később tárgyalandó sajátérték-feladatokban. Itt most azt vizsgáljuk meg, hogyan lehet egy **A** mátrixot úgy felbontani, hogy az egyik tényezője ortogonális mátrix legyen.

Egy reguláris **A** mátrixnak mindig létezik **QR** alakú felbontása, ahol **Q** ortogonális mátrix és az **R** (right) mátrix pedig felső háromszögmátrix. A létezést az alább ismertetendő eljárás egyben bizonyítja is. Ötletként a Gram-Schmidt ortogonalizációs eljárás jut eszünkbe, ahol egy független vektorrendszerből (itt az **A** mátrix oszlopvektoraiból) egy ortonormált rendszert konstruálunk.

Írjuk fel a **QR** felbontást következőképp:

$$[\mathbf{a}_1 \quad \mathbf{a}_2 \quad \dots \quad \mathbf{a}_n] = [\mathbf{q}_1 \quad \mathbf{q}_2 \quad \dots \quad \mathbf{q}_n] \begin{bmatrix} r_{1,1} & r_{1,2} & \cdots & r_{1,n} \\ 0 & r_{2,2} & \cdots & r_{2,n} \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & r_{n,n} \end{bmatrix}$$

ami szerint az $\mathbf{a}_k$ vektorok a $\{\mathbf{q}_1, \mathbf{q}_2, \dots \mathbf{q}_n\}$ ortonormált vektorok lineáris kombinációi:

$$\begin{aligned} \mathbf{a}_1 &= r_{1,1}\mathbf{q}_1 \\ \mathbf{a}_2 &= r_{1,2}\mathbf{q}_1 + r_{2,2}\mathbf{q}_2 \\ &\ \vdots \\ \mathbf{a}_n &= r_{1,n}\mathbf{q}_1 + r_{2,n}\mathbf{q}_2 + \dots + r_{n,n}\mathbf{q}_n \end{aligned}$$

Az első egyenletből

$$\mathbf{a}_1^*\mathbf{a}_1 = r_{1,1}^2\mathbf{q}_1^*\mathbf{q}_1 = r_{1,1}^2 \quad \text{ahonnan} \quad r_{1,1} = \sqrt{\mathbf{a}_1^*\mathbf{a}_1} \quad \textit{és} \quad \mathbf{q}_1 = \frac{1}{r_{1,1}}\mathbf{a}_1$$

Itt volt egy választási lehetőségünk, éspedig $r_{1,1}$ előjele. Hasonló választási lehetőségeink lesznek a további lépéseink során is, ami azt mutatja, hogy a **QR** felbontás nem egyértelmű, pontosabban szólva, ha az $r_{k,k}$ elemek előjelét mindig pozitívnak választjuk, akkor a felbontás egyértelmű lesz.

Legyen most $\mathbf{b}_2 = \mathbf{a}_2 - r_{1,2}\mathbf{q}_1$ ortogonális vektor $\mathbf{q}_1$ - re.

$$0 = \mathbf{q}_1^*\mathbf{b}_2 = \mathbf{q}_1^*\mathbf{a}_2 - r_{1,2}\mathbf{q}_1^*\mathbf{q}_1 = \mathbf{q}_1^*\mathbf{a}_2 - r_{1,2} \quad \text{ahonnan} \quad r_{1,2} = \mathbf{q}_1^*\mathbf{a}_2 \quad \text{adódik.}$$

Ha ezután a $\mathbf{b}_2$ vektort normáljuk, megkapjuk $\mathbf{q}_2$ - t:

$$r_{2,2} = \sqrt{\mathbf{b}_2^*\mathbf{b}_2} \quad \textit{és} \quad \mathbf{q}_2 = \frac{1}{r_{2,2}}\mathbf{b}_2$$

Ezután $\mathbf{a}_3$ figyelembe vételével felírható $\mathbf{b}_3$ vektor legyen ortogonális $\mathbf{q}_1$ és $\mathbf{q}_2$ - re:

$$\mathbf{b}_3 = \mathbf{a}_3 - (r_{1,3}\mathbf{q}_1 + r_{2,3}\mathbf{q}_2)$$

Az ortogonalitási feltételek:

$$0 = \mathbf{q}_1^*\mathbf{b}_3 = \mathbf{q}_1^*\mathbf{a}_3 - r_{1,3}\mathbf{q}_1^*\mathbf{q}_1 - r_{2,3}\mathbf{q}_1^*\mathbf{q}_2 = \mathbf{q}_1^*\mathbf{a}_3 - r_{1,3} \quad \text{ahonnan} \quad r_{1,3} = \mathbf{q}_1^*\mathbf{a}_3 \quad \text{adódik, és}$$

$$0 = \mathbf{q}_2^*\mathbf{b}_3 = \mathbf{q}_2^*\mathbf{a}_3 - r_{1,3}\mathbf{q}_2^*\mathbf{q}_1 - r_{2,3}\mathbf{q}_2^*\mathbf{q}_2 = \mathbf{q}_2^*\mathbf{a}_3 - r_{2,3} \quad \text{ahonnan} \quad r_{2,3} = \mathbf{q}_2^*\mathbf{a}_3$$

Normálás után kapjuk a $\mathbf{Q}$ mátrix következő oszlopát $\mathbf{q}_3$ - at:

$$r_{3,3} = \sqrt{\mathbf{b}_3^*\mathbf{b}_3} \quad \textit{és} \quad \mathbf{q}_3 = \frac{1}{r_{3,3}}\mathbf{b}_3$$

Az eljárás nagyobb méretű mátrixokon hasonlóan folytatható.

Megjegyezzük, hogy a **QR** felbontás konstruálására más módszerek is léteznek, például elemi ortogonális transzformációk sorozatával az **A** mátrix felső háromszögmátrix alakúvá transzformálható.

9. Példa: Készítsük el a következő mátrix **QR** felbontását:

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

Megoldás:

$$r_{1,1}^2 = \mathbf{a}_1^*\mathbf{a}_1 = (1)^2 + (-2)^2 + (2)^2 = 9 \quad \text{azaz} \quad r_{1,1} = 3 \quad \text{és} \quad \mathbf{q}_1 = \frac{1}{3}\mathbf{a}_1 = \begin{bmatrix} \frac{1}{3} \\ -\frac{2}{3} \\ \frac{2}{3} \end{bmatrix}$$

$$r_{1,2} = \begin{bmatrix} \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix} = -1$$

$$\mathbf{b}_2 = \mathbf{a}_2 - r_{1,2}\mathbf{q}_1 = \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix} + \begin{bmatrix} \frac{1}{3} \\ -\frac{2}{3} \\ \frac{2}{3} \end{bmatrix} = \begin{bmatrix} \frac{4}{3} \\ -\frac{2}{3} \\ -\frac{4}{3} \end{bmatrix}$$

$$r_{2,2}^2 = \left(\tfrac{4}{3}\right)^2 + \left(-\tfrac{2}{3}\right)^2 + \left(-\tfrac{4}{3}\right)^2 = 4 \quad \text{ahonnan} \quad r_{2,2} = 2 \quad \textit{és} \quad \mathbf{q}_2 = \begin{bmatrix} \frac{2}{3} \\ -\frac{1}{3} \\ -\frac{2}{3} \end{bmatrix}$$

$$r_{1,3} = \begin{bmatrix} \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix} = -1 \quad \text{és} \quad r_{2,3} = \begin{bmatrix} \frac{2}{3} & -\frac{1}{3} & -\frac{2}{3} \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix} = -1$$

$$\mathbf{b}_3 = \mathbf{a}_3 - (r_{1,3}\mathbf{q}_1 + r_{2,3}\mathbf{q}_2) = \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix} + \begin{bmatrix} \frac{1}{3} \\ -\frac{2}{3} \\ \frac{2}{3} \end{bmatrix} + \begin{bmatrix} \frac{2}{3} \\ -\frac{1}{3} \\ -\frac{2}{3} \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}$$

$$r_{3,3} = 2^2 + 2^2 + 1^2 = 9 \quad \text{ahonnan} \quad r_{3,3} = 3 \quad \textit{és} \quad \mathbf{q}_3 = \begin{bmatrix} \frac{2}{3} \\ \frac{2}{3} \\ \frac{1}{3} \end{bmatrix}$$

Tehát a felbontás:

$$\begin{bmatrix} 1 & 1 & 1 \\ -2 & 0 & 3 \\ 2 & -2 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ -\frac{2}{3} & -\frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \end{bmatrix} \cdot \begin{bmatrix} 3 & -1 & -1 \\ 0 & 2 & -1 \\ 0 & 0 & 3 \end{bmatrix}$$

Mint említettük a **QR** felbontást elemi ortogonális transzformációk (tükrözések ill. forgatások) segítségével is előállíthatjuk.

Most a példabeli mátrix Householder transzformációkon (elemi tükrözések) keresztüli **QR** felbontását mutatjuk meg. Számítási lépéseink alapötlete az, hogy minden egyes lépésben az **A** mátrixnak (ill. transzformáltjainak) főátló alatti elemeit oszloponként sorban kinullázzuk.

Jelölje $\mathbf{a}_1$ az **A** mátrix első oszlopvektorát. Keressük meg azt az $\mathbf{u}_1$ normált vektort, mely egy az origón átmenő sík (hipersík) normálvektora, és erre a síkra az $\mathbf{a}_1$ vektor végpontját tükrözve (Q_1 transzformáció), a tükörpont éppen a legelső koordinátatengely pozitív felére esik (azaz csak az első koordinátája lehet 0-tól különböző).

A feladat elemi vektorgeometriai ismeretekkel megoldható.

Az $\mathbf{a}_1$ vektor normája: $\quad \|\mathbf{a}_1\| = \sqrt{1+4+4} = 3 \quad$ ezért

$$\overset{(1)}{\mathbf{a}_1} = \mathbf{Q}_1\mathbf{a}_1 = \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}$$

Az $\mathbf{a}_1$ és $\mathbf{Q}_1\mathbf{a}_1$ vektorok „átlaga” a végpontjaikat összekötő szakasz felezőpontjába mutat:

$$\mathbf{f}_1 = \frac{1}{2}\left\{\begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix} + \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}\right\} = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} \quad ,$$

az innen az $\mathbf{a}_1$ végpontjába mutató vektor már párhuzamos $\mathbf{u}_1$-el:

$$\mathbf{v}_1 = \mathbf{a}_1 - \mathbf{f}_1 = \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix} - \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} \quad \text{ahonnan normálással} \quad \mathbf{u}_1 = \frac{1}{\sqrt{3}}\begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} \quad \text{tehát}$$

$$\mathbf{Q}_1 = \mathbf{E} - 2\mathbf{u}_1\mathbf{u}_1^* = \mathbf{E} - \frac{1}{3}\begin{bmatrix} 2 & 2 & -2 \\ 2 & 2 & -2 \\ -2 & -2 & 2 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 1 & -2 & 2 \\ -2 & 1 & 2 \\ 2 & 2 & 1 \end{bmatrix}$$

és ezzel:

$$\overset{(1)}{\mathbf{A}} = \mathbf{Q}_1\mathbf{A} = \frac{1}{3}\begin{bmatrix} 1 & -2 & 2 \\ -2 & 1 & 2 \\ 2 & 2 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 & 1 \\ -2 & 0 & 3 \\ 2 & -2 & 1 \end{bmatrix} = \begin{bmatrix} 3 & -1 & -1 \\ 0 & -2 & 1 \\ 0 & 0 & 3 \end{bmatrix}$$

ahol az első oszlop már megfelelő alakú.

A következő lépésben olyan elemi tükrözést kell végrehajtani, ami az oszlopvektorok első koordinátáját nem bántja, csak a második oszlopban a főátló alatti elemeket nullázza ki. Ez azt jelenti, hogy valójában az előző lépéssorozatot kell megismételni eggyel kisebb dimenzióban, azaz az $\overset{(1)}{\mathbf{A}}$ mátrix helyett annak első sora és első oszlopa elhagyásával kapott $\overset{(1)}{\mathbf{A}}_{1,1}$ minormátrixon. Ezt pedig úgy érjük el, hogy a tükröző sík normálvektorának első (a következő menetben első és második stb.) koordinátáját 0-nak választjuk.

Mivel a mintapéldánkban a második oszlopban a főátló alatti elem már 0, ezért csak arra kell ügyelni, hogy az $\overset{(2)}{\mathbf{a}}_2$ második koordinátája pozitív legyen, amit az $\mathbf{u}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ normálvektorú síkra való tükrözéssel érünk el.

$$\mathbf{Q}_2 = \mathbf{E} - 2\mathbf{u}_2\mathbf{u}_2^* = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$ és ezzel

$$\mathbf{R} = \overset{(2)}{\mathbf{A}} = \mathbf{Q}_2 \overset{(1)}{\mathbf{A}} = \begin{bmatrix} 3 & -1 & -1 \\ 0 & 2 & -1 \\ 0 & 0 & 3 \end{bmatrix}$$

A kívánt **R** alakot el is értük, már csak az van hátra, hogy az $\mathbf{R}=\mathbf{Q}_2\mathbf{Q}_1\mathbf{A}$ összefüggés és az $\mathbf{A}=\mathbf{QR}$ összevetéséből adódó

$$\mathbf{Q} = \mathbf{Q}_1^{-1}\mathbf{Q}_2^{-1} = \mathbf{Q}_1^*\mathbf{Q}_2^* = \mathbf{Q}_1\mathbf{Q}_2$$

összefüggés alapján az ortogonális **Q** mátrixot felírjuk:

$$\mathbf{Q} = \frac{1}{3}\begin{bmatrix} 1 & 2 & 2 \\ -2 & -1 & 2 \\ 2 & -2 & 1 \end{bmatrix}$$

A **Q** és **R** mátrix segítségével most is ugyanazt a felbontást kaptuk, mint korábban:

$$\begin{bmatrix} 1 & 1 & 1 \\ -2 & 0 & 3 \\ 2 & -2 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ -\frac{2}{3} & -\frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \end{bmatrix} \cdot \begin{bmatrix} 3 & -1 & -1 \\ 0 & 2 & -1 \\ 0 & 0 & 3 \end{bmatrix}$$

Megjegyezzük, hogy mindkét (Gram-Schmidt, Householder) felbontási eljárás kiválóan alkalmas az algoritmizálásra.

Hasonlósági transzformációk

Az n dimenziós lineáris tér egy új bázisának vektorait foglaljuk össze a **T** mátrixban (amely így nyilván reguláris mátrix). Az eredeti bázisban felírt **x** vektor új bázisbeli koordinátáit – mint korábban láttuk -

$$\mathbf{x}' = \mathbf{T}^{-1}\mathbf{x}$$

módon kaphatjuk meg (azaz $\mathbf{x} = \mathbf{T}\mathbf{x}'$). Most vizsgáljuk meg azt az A transzformációt, mely az eredeti bázisban az **x** vektort **y**-ba viszi:

$$\mathbf{y} = \mathbf{A}\mathbf{x}$$

Az **y** vektor új bázisbeli koordinátáit

$\mathbf{y}' = \mathbf{T}^{-1}\mathbf{y}$ szolgáltatja. Emiatt

$$\mathbf{y}' = \mathbf{T}^{-1}\mathbf{y} = \mathbf{T}^{-1}\mathbf{A}\mathbf{x} = \mathbf{T}^{-1}\mathbf{A}\mathbf{T}\mathbf{x}' = \mathbf{A}'\mathbf{x}'$$

tehát ugyanazt az A transzformációt, melynek az eredeti bázisbeli mátrixa **A**, a $\{t_1, t_2, \dots t_n\}$ bázisban már a

$$\mathbf{A}' = \mathbf{T}^{-1}\mathbf{A}\mathbf{T}$$

mátrix írja le. Emiatt a **T** reguláris mátrix segítségével felírt és az **A** mátrixon végrehajtott

$$\mathbf{T}^{-1}\mathbf{A}\mathbf{T}$$

transzformációt **hasonlósági transzformáció-**nak nevezzük.

A hasonlósági transzformáció

a./ *reflexív* (minden mátrix hasonló önmagához), azaz van olyan reguláris **T** mátrix (pl. **T**=**E**), mellyel

$$\mathbf{A} = \mathbf{T}^{-1}\mathbf{A}\mathbf{T}$$

b./ *szimmetrikus*, azaz ha **A** és **B** hasonló, akkor **B** és **A** is az, hiszen

$$\mathbf{A} = \mathbf{T}^{-1}\mathbf{B}\mathbf{T} \text{ esetén } \quad \mathbf{B} = (\mathbf{T}^{-1})^{-1}\mathbf{A}(\mathbf{T}^{-1})$$

c./ *tranzitív*, azaz ha **A** hasonló **B**-hez, és **B** hasonló **C**-hez, akkor **A** hasonló **C**-hez, mivel ha

$$\mathbf{A} = \mathbf{P}^{-1}\mathbf{B}\mathbf{P} \quad \text{és} \quad \mathbf{B} = \mathbf{Q}^{-1}\mathbf{C}\mathbf{Q}\,,$$

$$\text{akkor} \quad \mathbf{A} = \mathbf{P}^{-1}\,\mathbf{Q}^{-1}\mathbf{C}\mathbf{Q}\,\mathbf{P} = (\mathbf{Q}\mathbf{P})^{-1}\mathbf{C}(\mathbf{Q}\mathbf{P}) = \mathbf{T}^{-1}\mathbf{B}\mathbf{T}\,, \text{ ahol } \quad \mathbf{T} = \mathbf{Q}\mathbf{P}$$

d./ két további nevezetes tulajdonsággal bír:

$$\mathbf{T}^{-1}\mathbf{A}_1\mathbf{T} + \dots + \mathbf{T}^{-1}\mathbf{A}_n\mathbf{T} = \mathbf{T}^{-1}(\mathbf{A}_1 + \dots + \mathbf{A}_n)\mathbf{T}$$

$$(\mathbf{T}^{-1}\mathbf{A}\mathbf{T})^k = \mathbf{T}^{-1}\mathbf{A}^k\mathbf{T} \quad \text{ahol k pozitív egész.}$$

A hasonlósági transzformációk között fontos szerepet kapnak azok, melyeknél a **T** mátrix speciális tulajdonságú (háromszög-alakú, ortogonális, stb.).

Amikor a **T** mátrix ortogonális, akkor az új bázis is ortonormált. Az ortogonális mátrixokat többnyire $\mathbf{U}=[\mathbf{u}_1, \mathbf{u}_2, .. \mathbf{u}_n]$ -el ill. $\mathbf{Q}=[\mathbf{q}_1,\mathbf{q}_2, .. \mathbf{q}_n]$-el jelölik, melynek vektorai ortonormáltak. Az alkalmazásokban az ortogonális hasonlósági transzformációk következő tulajdonságait használják ki:

a./ ortogonális mátrixok szorzata is ortogonális, azaz ha egymás után ortogonális mátrixok segítségével hajtunk végre hasonlósági transzformációt, akkor az egész transzformációs sorozat egyetlen ortogonális mátrixszal végrehajtott hasonlósági transzformációval helyettesíthető. Tehát

$$\mathbf{Q}=\mathbf{Q}_1\mathbf{Q}_2\dots\mathbf{Q}_m \text{ esetén}$$

$$\mathbf{Q}_m^*\mathbf{Q}_{m-1}^*\ldots\mathbf{Q}_1^*\mathbf{A}\,\mathbf{Q}_1\mathbf{Q}_2\ldots\mathbf{Q}_m = \mathbf{Q}^*\mathbf{A}\mathbf{Q}$$

b./ ortogonális mátrix segítségével végrehajtott hasonlósági transzformáció az A szimmetrikus mátrix szimmetriáját megőrzi, mert

$$\mathbf{A}' = \mathbf{Q}^*\mathbf{A}\mathbf{Q} \text{ esetén } \mathbf{A}'^* = (\mathbf{Q}^*\mathbf{A}\mathbf{Q})^* = \mathbf{Q}^*\mathbf{A}^*\mathbf{Q}^{**} = \mathbf{Q}^*\mathbf{A}\mathbf{Q} = \mathbf{A}'$$

c./ ha **A** tetszőleges négyzetes mátrix, akkor található olyan **Q** ortogonális mátrix, hogy a

$\mathbf{Q}^*\mathbf{A}\mathbf{Q}$ mátrix felső Hessenberg-alakú, azaz

tetszőleges négyzetes mátrix hasonló egy felső Hessenberg mátrixhoz.

A **Q** ortogonális mátrix megkonstruálásához a **QR** felbontáshoz hasonlóan a Householder-transzformációkat (elemi tükrözések) használhatjuk fel.

Válasszuk meg a $\mathbf{v}^{(1)} = (0, v_2^{(1)}, \ldots, v_n^{(1)})^*$ normált vektort úgy, hogy a

$$\mathbf{Q}_1 = \mathbf{E} - 2\mathbf{v}^{(1)}\mathbf{v}^{(1)*}$$

Householder-mátrix segítségével felírt $\mathbf{Q}_1\mathbf{A}$ szorzatban az első oszlopban a 3. elemtől kezve minden elem 0 legyen (n-1 ismeretlen az n-1 egyenlethez). Ezután képezzük az

$\mathbf{A}^{(2)} = \mathbf{Q}_1\mathbf{A}^{(1)}\mathbf{Q}_1$ mátrixot, ahol $\mathbf{A}^{(1)} = \mathbf{A}$.

Ez hasonló **A**-hoz, és a $\mathbf{Q}_1$-el való jobbról szorzás az első oszlopot nem változtatja meg, mert $\mathbf{Q}_1$ első oszlopa az $\mathbf{e}_1$ egységvektor. Így $\mathbf{A}^{(2)}$ első oszlopa már megfelelő alakú lesz.

Ezután a $\mathbf{v}^{(2)} = (0, 0, v_3^{(2)}, \ldots, v_n^{(2)})^*$ vektor koordinátáit választjuk meg úgy, hogy a

$\mathbf{A}^{(3)} = \mathbf{Q}_2\mathbf{A}^{(2)}\mathbf{Q}_2$ mátrix második oszlopa legyen megfelelő alakú.

Ez a transzformáció az első oszlopot nem rontja el! Az eljárást hasonlóan folytatva n-2 lépésben a megfelelő felső Hessenberg-alakú mátrixhoz jutunk.

Ha alsó Hessenberg mátrixszá akarjuk a négyzetes mátrixunkat transzformálni, akkor ugyanezt az algoritmust kell végrehajtani azzal a különbséggel, hogy most az utolsó oszloppal kezdjük a felső mellékátló feletti elemek kinullázását.

d./ Ortogonális hasonlósági transzformációval egy szimmetrikus négyzetes mátrix tridiagonális alakra hozható.

Ez a b./ és c./ egyenes következményeként adódik.

Példaként tekintsük az

$$\mathbf{A} = \begin{bmatrix} 2 & -0{,}28 & -0{,}96 \\ -0{,}28 & 0{,}232 & 2{,}224 \\ -0{,}96 & 2{,}224 & -0{,}232 \end{bmatrix}$$ szimmetrikus mátrixot és a $\mathbf{v} = \begin{bmatrix} 0 \\ 0{,}8 \\ 0{,}6 \end{bmatrix}$ normált vektort.

Ekkor a

$$\mathbf{Q} = \mathbf{E} - 2\mathbf{v}\mathbf{v}^* = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -0{,}28 & -0{,}96 \\ 0 & -0{,}96 & 0{,}28 \end{bmatrix}$$ ortogonális tükröző mátrixszal végrehajtott

hasonlósági transzformációval az

$$\mathbf{A}' = \mathbf{QAQ} = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 1 & 2 \\ 0 & 2 & -1 \end{bmatrix}$$ mátrix már tridiagonális lesz

Lineáris egyenletrendszerek, mátrix kondíciószáma

Sajátérték, sajátvektor

jön .._