Pázmány Péter Katolikus Egyetem
Bölcsészettudományi Kar
Szociológia Intézet

Matematikai statisztika 1

(BBNSZ01201)

LINEÁRIS ALGEBRA

MARX KÁROLY KÖZGAZDASÁGTUDOMÁNYI EGYETEM

Dr. Halmai Erzsébet — Dr. Krekó Béla

LINEÁRIS ALGEBRA

KÉZIRAT
17. változatlan kiadás

TANKÖNYVKIADÓ, BUDAPEST, 1989

LINEÁRIS ALGEBRA

3.1. Az elemi bázistranszformáció

Az n-elemü vektorok teréről kimutattuk, hogy n-dimenziós és az

$$\underline{e}_1, \underline{e}_2, \dots, \underline{e}_n$$

vektorrendszer a tér triviális bázisát képezi. A következőkben megvizsgáljuk, hogyan konstruálható e térben a triviálistól különböző bázis, illetve egy adott bázisból hogyan térhetünk át egy uj bázisba.

Ha az L_n valamely adott bázisából egy másik bázisába térünk át, akkor bázis transzformációról beszélünk. A bázistranszformációnak azt a legegyszerübb esetét, amikor az adott bázisnak csak egyik vektorát változtatjuk meg, elemi transzformációnak nevezzük.

Az itt definiált elemi bázistranszformáció alapján keressünk uj bázist az L_n-ben.

Az L_n bármely $\underline{c}$ vektora - a tanultak szerint - felirható

$$\underline{c} = c_1 \underline{e}_1 + c_2 \underline{e}_2 + \dots + c_n \underline{e}_n$$

alakban, ahol a $c_1, c_2, \dots, c_n$ skalárok a bázisvektorokra vonatkozó koordináták. A $\underline{c} \neq \underline{0}$ esetén ezen skalárok közt van zérustól különböző is. Ha történetesen $c_1 \neq 0$, akkor az egyenlőségből $\underline{e}_1$ vektor kifejezhető, vagyis

$$\underline{e}_1 = \frac{1}{c_1} \cdot \underline{c} - \frac{c_2}{c_1} \underline{e}_2 - \dots - \frac{c_n}{c_1} \underline{e}_n .$$

Az $\underline{e}_1 \in L_n$ vektort tehát felirtuk a

$$\underline{c}, \underline{e}_2, \dots, \underline{e}_n$$

vektorok lineáris kombinációjaként. Állitjuk, hogy ezen vektorok az L_n-re nézve uj bázist jelentenek.

Mivel a bázis definiciója szerint az n-elemü vektorok terében n lineárisan független vektorból álló rendszer bázist képez, csupán azt kell kimutatnunk, hogy a

$$\underline{c},\ \underline{e}_2,\ \ldots,\ \underline{e}_n$$

vektorok lineárisan függetlenek. E végből irjuk fel az

$$x_1\,\underline{c} + x_2\,\underline{e}_2 + \ldots + x_n\,\underline{e}_n = \underline{0}$$

egyenletet. Ha ennek volna a triviálistól különböző megoldása is, akkor a $\underline{c}$ vektort az $\underline{e}_2, \ldots, \underline{e}_n$ vektorok is előállitanák, ez pedig ellentétben áll azzal a kikötéssel, hogy a $\underline{c}$-nek az $\underline{e}_1$-re vonatkozó koordinátája 0-tól különböző.

Ezzel az L_n-nek egy - a triviálistól különböző - bázisához jutottunk.

Mivel a gondolatmenetben sehol sem használtuk fel, hogy az induló bázisvektorok egységvektorok, a fenti megfontolásokkal az L_n lineáris tér bármely adott

$$\underline{b}_1,\ \underline{b}_2,\ \ldots,\ \underline{b}_n$$

bázisáról áttérhetünk uj bázisra.

Megjegyzések: a) Az L_n lineáris tér egy tetszőleges bázisát jelöli a

$$\underline{b}_1,\ \underline{b}_2,\ \ldots,\ \underline{b}_n$$

vektorrendszer. (Ez lehet a triviális vagy attól különböző is!)

b) Annak a feltétele, hogy a $\underline{b}_j$ bázisvektort kicserélhessük az L_n egy adott $\underline{c} \neq \underline{0}$ vektorára, az, hogy a $\underline{c}$-nek a $\underline{b}_j$ bázisvektorra vonatkozó koordinátája 0-tól különböző legyen, azaz teljesüljön a $c_j \neq 0$ követelmény.

c) Ha a

$$\underline{c} = c_1\,\underline{b}_1 + c_2\,\underline{b}_2 + \ldots + c_n\,\underline{b}_n$$

lineáris kombinációban a $\underline{c} \neq \underline{0}$ vektornak a bázisvektorokra vonatkozó koordinátái között több zérustól különböző skalár szerepel, a $\underline{c}$-vel kapcsolatos elemi bázistranszformációt annyi különböző módon lehet végrehajtani, ahány 0-tól különböző koordinátája van a $\underline{c}$-nek.

3.2. A kompatibilitás

Az előző pontban tárgyalt példában a feladat igy is megfogalmazható: Vizsgáljuk meg, hogy $\underline{a}$ vektor benne fekszik-e a $\underline{c}_1$, $\underline{c}_2$, $\underline{c}_3$ vektorok által meghatározott altérben.

A számitások szerint $\underline{a}$ benne fekszik a kérdéses altérben, hiszen $\underline{a}$ előállitható a $\underline{c}_1$, $\underline{c}_2$, $\underline{c}_3$ vektorok lineáris kombinációjaként.

Ha az $\underline{a}$ vektor előállitható a $\underline{c}_1, \underline{c}_2, \ldots, \underline{c}_k$ vektorok lineáris kombinációjaként, akkor azt mondjuk, hogy $\underline{a}$ vektor kompatibilis a $\underline{c}_1, \underline{c}_2, \ldots, \underline{c}_k$ vektorok által generált altérre nézve.

Azt a kifejezést is szoktuk használni, hogy $\underline{a}$ vektor eleget tesz a kompatibilitás követelményének.

Ha az $\underline{a}$ vektor nem állitható elő az alteret generáló $\underline{c}_1, \underline{c}_2, \ldots, \underline{c}_k$ vektorok lineáris kombinációjaként, akkor az $\underline{a}$ vektor inkompatibilis a $\underline{c}_1, \underline{c}_2, \ldots, \underline{c}_k$ vektorok által kifeszitett altérre nézve.

Előfordulhat azonban, hogy az alteret generáló $\underline{c}_1, \underline{c}_2, \ldots, \underline{c}_k$ vektorrendszer független vektorainak részhalmaza is elegendő az $\underline{a}$ vektor előállitásához.

3.4. Egy speciális faktorizáció

Azt az eljárást, amelynek révén egy adott mátrixot mátrixok szorzatára bontunk fel, faktorizációnak nevezzük.

A legegyszerübb ilyen tényezőkre bontás az

$$\underline{A} = \underline{A} \cdot \underline{E}$$

alaku felbontás, amellyel már találkoztunk.

A mátrixok rangjának numerikus meghatározásánál felhasznált módszer alapján egy speciális faktorizációhoz juthatunk.

Tekintsük ugyanis az előző pontban vizsgált

$$\underline{A} = \begin{bmatrix} 1 & 0 & 2 & 1 & 5 \\ 3 & 1 & 0 & 1 & 4 \\ 5 & 1 & 4 & 3 & 14 \\ 7 & 2 & 2 & 3 & 13 \end{bmatrix}$$

mátrixot. A számitások szerint tudjunk, hogy a mátrix rangja 2 és a mátrix oszlopvektorterében az $\underline{a}_1$ és $\underline{a}_2$ vektorok bázist képeznek. Ennek alapján az $\underline{A}$ mátrix oszlopvektorterének bármely vektora előállitható ezek lineáris kombinációjaként, igy az $\underline{A}$ oszlopai is. Ha bevezetjük az

$$\underline{A}_1 = [\,\underline{a}_1, \underline{a}_2\,] = \begin{bmatrix} 1 & 0 \\ 3 & 1 \\ 5 & 1 \\ 7 & 2 \end{bmatrix}$$

jelölést, akkor felirhatjuk, hogy

$$\underline{a}_1 = \underline{A}_1 \cdot \underline{e}_1$$

$$\underline{a}_2 = \underline{A}_1 \cdot \underline{e}_2 \,,$$

$$\begin{bmatrix} 1 & 0 & 2 & 1 & 5 \\ 3 & 1 & 0 & 1 & 4 \\ 5 & 1 & 4 & 3 & 14 \\ 7 & 2 & 2 & 3 & 13 \end{bmatrix} = \begin{bmatrix} 1 & 5 \\ 1 & 4 \\ 3 & 14 \\ 3 & 13 \end{bmatrix} \cdot \begin{bmatrix} 11 & 5 & -8 & 1 & 0 \\ -2 & -1 & 2 & 0 & 1 \end{bmatrix}$$

felbontáshoz jutunk.

A felbontás helyességét ellenőrizzük itt is!

Ha az $\underline{A}$ mátrix előző faktorizációiban a második tényezőt $\underline{A}_2$ szimbólummal jelöljük, akkor

$$\underline{A} = \underline{A}_1 \cdot \underline{A}_2$$

egyenlőséggel irhatjuk fel a tényezőkre bontást. Amint már megállapitottuk, az első tényező oszlopvektorai bázisát alkotják az $\underline{A}$ oszlopvektorterének. Ugyanakkor - a mátrixok rangjára adott definició szerint -, a második tényező sorvektorai meg az $\underline{A}$ sorvektorterének alkotják egy bázisát. Ezért ezt a faktorizációt <u>bázisfaktorizációnak</u>, vagy <u>bázisfelbontásnak</u> nevezzük.

A következő példában

$$\begin{bmatrix} 1 & 0 & 2 & 1 & 5 \\ 3 & 1 & 0 & 1 & 4 \\ 5 & 1 & 4 & 3 & 14 \\ 7 & 2 & 2 & 3 & 13 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 5 \\ 0 & 1 & 4 \\ 4 & 3 & 14 \\ 2 & 3 & 13 \end{bmatrix} \cdot \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 11 & 5 & 0 & 1 & 0 \\ -2 & -1 & 0 & 0 & 1 \end{bmatrix}$$

felbontás nem bázisfelbontás, ezért nem tekintjük speciális faktorizációnak.

Az elmondottakkal kapcsolatban állitjuk:

1. <u>Tétel:</u> Ha az $\underline{A}$ mátrix rangja r és történetesen az $\underline{A}$ első r oszlopvektorából alkotott $\underline{A}_1 = [\underline{a}_1, \ldots, \underline{a}_r]$ mátrix oszlopvektorai lineárisan függetlenek, akkor az $\underline{A}$ az

$$\underline{A} = \underline{A}_1 \, [\underline{E}_r \ , \ \underline{D}]$$

összefüggés mintájára faktorizálható, ahol $\underline{E}_r$ egy r-rendü egységmátrixot jelent. Maga az $[\underline{E}_r, \underline{D}]$ szimbólum olyan mátrixot jelent, amelynek első r oszlopvektora megegyezik $\underline{E}_r$ oszlopvektoraival, további oszlopvektorai pedig azonosak a $\underline{D}$ oszlopvektoraival.

<u>Bizonyitás:</u> A bemutatott numerikus példák alapján az állítás kézenfekvő. Ha ugyanis

$$\underline{A} = [\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_r, \underline{a}_{r+1}, \ldots, \underline{a}_n],$$

akkor nyilvánvaló, hogy

$$\underline{a}_1 = \underline{A}_1 \cdot \underline{e}_1$$

$$\underline{a}_2 = \underline{A}_1 \cdot \underline{e}_2$$

$$\vdots$$

$$\underline{a}_r = \underline{A}_1 \cdot \underline{e}_r \ .$$

S mivel az $\underline{A}_1$ oszlopvektorai bázisát alkotják $\underline{A}$ oszlopvektorterének, kell lenni olyan

$$\underline{d}_{r+1}\ ,\ \underline{d}_{r+2},\ \dots\ ,\ \underline{d}_n$$

vektoroknak, hogy teljesüljenek az

$$\underline{a}_{r+1} = \underline{A}_1 \cdot \underline{d}_{r+1}$$

$$\underline{a}_{r+2} = \underline{A}_1 \cdot \underline{d}_{r+2}$$

$$\vdots$$

$$\underline{a}_n = \underline{A}_1 \cdot \underline{d}_n$$

egyenlőségek is.

Ezek szerint

$$\underline{A} = \left[\underline{A}_1 \underline{e}_1,\ \underline{A}_1 \cdot \underline{e}_2,\ \dots,\ \underline{A}_1 \underline{e}_r,\ \underline{A}_1 \cdot \underline{d}_{r+1},\ \underline{A}_1 \cdot \underline{d}_{r+2},\ \dots,\ \underline{A}_1 \cdot \underline{d}_n \right] ,$$

ahonnan

$$\underline{A} = \underline{A}_1 \left[\underline{e}_1,\ \underline{e}_2,\ \dots,\ \underline{e}_r,\ \underline{d}_{r+1},\ \underline{d}_{r+2},\ \dots,\ \underline{d}_n \right] .$$

Ha pedig bevezetjük az

$$\left[\underline{e}_1,\ \underline{e}_2,\ \dots,\ \underline{e}_r \right] = \underline{E}_r$$

és a

$$\left[\underline{d}_{r+1},\ \underline{d}_{r+2},\ \dots,\ \underline{d}_n \right] = \underline{D}$$

jelölést, akkor éppen a keresett

$$\underline{A} = \underline{A}_1 \left[\underline{E}_r , \underline{D} \right]$$

összefüggéshez jutottunk.

Ezzel a tételt bebizonyitottuk.

Az $\underline{A} = \underline{A}_1 \cdot \underline{A}_2$ jelölést véve az $\left[\underline{E}_r , \underline{D} \right]$ az $\underline{A}_2$ particionált formája, ahol az $\underline{E}_r$ blokk az $\underline{A}$ mátrixból kiválasztott független oszlopvektorokra vonatkozó koordinátákat tartalmazza, a $\underline{D}$ blokk pedig az $\underline{A}$ mátrix többi vektorának a kiválasztott bázisvektorokra vonatkozó koordinátáit foglalja magába. Az előző konkrét számitások szerint $\underline{D}$ elemei mindig az utolsó elemi transzformációs táblából olvashatók ki.

Megjegyezzük, hogy az

$$\underline{A} = \underline{A}_1 \cdot \left[\underline{E}_r , \underline{D} \right]$$

elrendezés sok esetben csak az $\underline{A}$ mátrix oszlopainak átrendezésével biztositható.

4. Lineáris egyenletrendszerek megoldása és mátrixok invertálása

4.1. Általános tudnivalók

Az $x_1, x_2, \dots, x_n$ ismeretleneket tartalmazó

$$
\begin{aligned}
a_{11} x_1 + a_{12} x_2 + \dots + a_{1n} x_n &= b_1 \\
a_{21} x_1 + a_{22} x_2 + \dots + a_{2n} x_n &= b_2 \\
&\ \vdots \\
a_{m1} x_1 + a_{m2} x_2 + \dots + a_{mn} x_n &= b_m
\end{aligned}
$$

egyenletek halmazát elsőfoku vagy lineáris egyenletrendszernek nevezzük.

A definicióban "m" az egyenletek számát, "n" az ismeretlenek számát jelenti. Mind az m, mind pedig az n bármilyen természetes szám lehet. (Középiskolában csak azokkal az egyenletrendszerekkel szoktak foglalkozni, amelyekben az egyenletek száma megegyezik az ismeretlenek számával. Mi ilyen megszoritást nem teszünk.)

Ha az egyenletrendszerben az egyenlőség jobboldalán álló $b_1, b_2, \dots, b_m$ skalárok mindegyike "0", akkor az egyenletrendszert homogénnek hivjuk, egyébként inhomogén egyenletrendszerről van szó.

Ha a fenti un. skaláregyenletrendszerben az adott skalárokra az

$$
\underline{a}_1 = \begin{bmatrix} a_{11} \\ a_{21} \\ \cdot \\ \cdot \\ \cdot \\ a_{m1} \end{bmatrix}, \quad
\underline{a}_2 = \begin{bmatrix} a_{12} \\ a_{22} \\ \cdot \\ \cdot \\ \cdot \\ a_{m2} \end{bmatrix}, \dots, \underline{a}_n = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \cdot \\ \cdot \\ \cdot \\ a_{mn} \end{bmatrix} \text{ és } \underline{b} = \begin{bmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ b_n \end{bmatrix}
$$

jelöléseket vezetjük be, az egyenletrendszer az alábbi un. vektoregyenlettel helyettesithető:

$$\underline{a}_1 x_1 + \underline{a}_2 x_2 + \dots + \underline{a}_n x_n = \underline{b} .$$

Ez a vektoregyenlet azonban felfogható ugyis, mint egy $\underline{A}$ oszlopvektorainak olyan lineáris kombinációja, amelyben a skalárok $x_1, x_2, \ldots, x_n$ ismeretlenek. Az egyenletrendszer azért igy is jelölhető:

$$\underline{A}\,\underline{x} = \underline{b}\,.$$

Az $\underline{A}$ mátrixot együttható mátrixnak nevezzük, mivel az egyenletrendszerben szereplő ismeretlenek együtthatóit reprezentálja.

Az $\underline{A}\,\underline{x} = \underline{b}$ egyenlőségnek megfelelő egyenletrendszert megoldani annyit jelent, mint meghatározni mindazon $\underline{x}$ vektorokat, amelyek eleget tesznek az $\underline{A}\,\underline{x} = \underline{b}$ követelménynek.

Ezeknek az $\underline{x}$ vektoroknak a meghatározása azonban nem jelent mást, mint $\underline{A}$ oszlopvektorainak azokat a lineáris kombinációit megkeresni, amelyek a $\underline{b}$ vektort állitják elő.

Amennyiben ilyen $\underline{x}$ vektor, vagy $\underline{x}$ vektorok léteznek, ezek egy halmazban az un. megoldás halmazban foglalhatók össze. Jelöljük ezt, az összes lehetséges megoldásokat tartalmazó halmazt az

$$M = \left\{ \underline{x} \;\middle|\; \underline{A}\,\underline{x} = \underline{b} \right\}$$

szimbólummal.

1. Tétel: Az M konvex halmaz.

Bizonyitás: A feltétel szerint ugyanis az M-nek van legalább egy eleme. Ha több történetesen nincs, akkor állitásunk máris bizonyitott, hiszen az egyetlen pontból álló halmazt is konvex halmaznak tekintjük.

Ha az M-nek több eleme is van, M konvex voltát a következőképpen láthatjuk be.

Legyen pl. $\underline{x}_1$ is és egy tőle különböző $\underline{x}_2$ is eleme az M-nek, azaz teljesüljön mind az

$$\underline{A}\,\underline{x}_1 = \underline{b}$$

mind az

$$\underline{A}\,\underline{x}_2 = \underline{b}$$

egyenlőség. Majd irjuk fel az $\underline{x}_1$ és $\underline{x}_2$ pontok által meghatározott $(\underline{x}_1, \underline{x}_2)$ szakasz egy tetszőleges

$$\underline{x} = \lambda\,\underline{x}_1 + (1-\lambda)\,\underline{x}_2,$$

ahol

$$0 \leqq \lambda \leqq 1$$

pontját.

A müveleti szabályok értelmében

$$\underline{A}\,\underline{x} = \underline{A}\left[\lambda \underline{x}_1 + (1-\lambda)\underline{x}_2\right] = \lambda(\underline{A}\,\underline{x}_1) + (1-\lambda)(\underline{A}\,\underline{x}_2)\ ,$$

ami pedig az $\underline{A}\,\underline{x}_1 = \underline{b}$ és $\underline{A}\,\underline{x}_2 = \underline{b}$ feltételek miatt azt adja, hogy

$$\underline{A}\,\underline{x} = \lambda\underline{b} + (1-\lambda)\,\underline{b} = \underline{b}\ .$$

Ez pedig azt jelenti, hogy $\underline{x}$ is eleme az M halmaznak. De az $\underline{x}$ az $(\underline{x}_1, \underline{x}_2)$ szakasz tetszőleges pontja, ezért ebből az következik, hogy az egész szakasz beletartozik az M-be, ami éppen azt mutatja, hogy az M konvex halmaz, ahogy állitottuk.

<u>Az M konvex halmaz dimenzióját az egyenletrendszer szabadságfokának nevezzük.</u>

A konvex halmazok dimenziójának definiálásakor megemlitettük, hogy az egyetlen pontból álló konvex halmaz dimenziója 0, ebből következik, ha M egyetlen elemből áll, dimenziója, vagyis szabadságfoka 0.

2. <u>Tétel</u>: ha a megoldáshalmaz szabadságfoka 0-nál nagyobb, akkor az M nem korlátos halmaz.

<u>Bizonyitás:</u> Ha M szabadságfoka 0-nál nagyobb, akkor az M-nek van legalább két különböző pontja. Ezek legyenek: $\underline{x}_1$, és $\underline{x}_2$, ahol $\underline{x}_1 \neq \underline{x}_2$ vagyis

$$\underline{h} = \underline{x}_1 - \underline{x}_2 \neq \underline{0}\ .$$

Ezek után tekintsük tetszőleges t skalár mellett az

$$\underline{x} = \underline{x}_1 + t\ \underline{h}$$

vektort és vizsgáljuk meg, hogy eleme-e az M halmaznak. E végből irjuk fel az

$$\underline{A}\,\underline{x} = \underline{A}\left[\underline{x}_1 + t\,\underline{h}\right] = \underline{A}\left[\underline{x}_1 + t\ (\underline{x}_1 - \underline{x}_2)\right]$$

összefüggést, ahol a müveleti szabályoknak megfelelő eljárás és az adott feltétel szerint

$$\underline{A}\left[\underline{x}_1 + t\,(\underline{x}_1 - \underline{x}_2)\right] = \underline{A}\,\underline{x}_1 + t\,(\underline{A}\,\underline{x}_1 - \underline{A}\,\underline{x}_2) = \underline{b} + t\,(\underline{b} - \underline{b})$$

eredményhez jutunk.

E szerint pedig

$$\underline{A}\,\underline{x} = \underline{b}\,,$$

ami éppen azt bizonyitja, hogy $\underline{x} \in M$-nek. S mivel ez az állitás bármilyen t mellett igaz, következik állitásunk helyessége, hogy az M nem-korlátos.

Megjegyzés: Az $\underline{A}\,\underline{x} = \underline{b}$ egyenletrendszer összes lehetséges megoldásainak - sok esetben - csak valamely részhalmaza érdekes a megoldó számára. Igy például a gazdasági problémák megoldása során általában csak az $\underline{x} \geqq \underline{0}$ megoldásokat veszik figyelembe.

4.2. A lineáris egyenletrendszerek megoldása

A lineáris egyenletrendszerrel kapcsolatos eddigi megállapitásaink során feltettük, hogy az M nem üres halmaz, de arra még nem adtunk választ, mi a feltétele annak, hogy egy egyenletrendszer megoldható legyen.

1. Tétel: Az $\underline{A}\,\underline{x} = \underline{b}$ alakban megadott lineáris egyenletrendszer akkor és csak akkor oldható meg, ha a $\underline{b}$ vektor kompatibilis az $\underline{A}$ mátrix oszlopvektorterére, azaz $\underline{b}$ benne fekszik $\underline{A}$ oszlopvektorterében.

Bizonyitás: A feladat megoldása, mint tudjuk abban áll, hogy meg kell keresni az $\underline{A}$ oszlopvektorainak minden olyan lineáris kombinációját, amely előállitja a $\underline{b}$ vektort. Az $\underline{A}$ mátrix oszlopvektorterére adott definició szerint

$$L' = \left\{ \underline{s} \;\middle|\; \underline{s} = \underline{A}\,\underline{x} \right\}$$

az $\underline{A}$ mátrix oszlopvektorainak összes lehetséges lineáris kombinációit tartalmazza. Igy, ha $\underline{b} \in L'$-nek, akkor az egyenletrendszer megoldható, s megfor-ditva, ha az egyenletrendszer megoldható, akkor $\underline{b} \in L'$-nek. Ebből pedig már következik az állitásunk helyessége.

Ezek szerint a lineáris egyenletrendszer megoldásánál a kiinduló feladat a kompatibilitás fennállásának eldöntése. Egy vektornak adott altérre vonatkozó kompatibilitását már vizsgáltuk. Igy az ott megismert eljárást az egyenletrendszerek megoldhatóságának vizsgálatára is felhasználhatjuk. Sőt, mint majd látni fogjuk, amennyiben a kompatibilitás fennáll, a módszer magát a megoldást is szolgáltatja.

Ezek után vizsgáljuk meg, hogyan állitható elő a megoldáshalmaz valamely $\underline{x}$ eleme, feltéve, hogy az $\underline{A}\,\underline{x} = \underline{b}$ adott egyenletrendszernél a kompatibilitás teljesül.

Tegyük fel, hogy az $\underline{A}$ együttható mátrix rangja r és történetesen az első r oszlopvektor lineárisan független. Ez esetben a mátrixok speciális faktorizációjánál tanultak értelmében az $\underline{A}$ mátrix az

$$\underline{A} = \underline{A}_1 \left[\underline{E}_r \, , \, \underline{D} \right]$$

összefüggés mintájára faktorizálható, ahol $\underline{A}_1$ oszlopvektorai megegyeznek az $\underline{A}$ mátrix első r (lineárisan független) oszlopvektoraival.

Mivel $\underline{A}_1$ oszlopvektorai az $\underline{A}$ oszlopvektorterének egyik bázisát alkotják, a kompatibilitás következtében $\underline{b}$ felirható

$$\underline{b} = \underline{A}_1 \cdot \underline{d}$$

alakban. Meggondolásainkat felhasználva az $\underline{A}\ \underline{x} = \underline{b}$ lineáris egyenletrendszer uj formája:

$$\underline{A}_1 \left[\underline{E}_r , \ \underline{D} \right] \underline{x} = \underline{A}_1 \cdot \underline{d} \, .$$

Itt mind a bal oldal, mind a jobb oldal az $\underline{A}_1$ oszlopvektorainak olyan lineáris kombinációját jelenti, amely a $\underline{b}$ vektort állitja elő. Mivel pedig az $\underline{A}_1$ oszlopvektorai bázist alkotnak az $\underline{A}$ mátrix oszlopvektorterében, azért mind az $\left[\underline{E}_r , \ \underline{D} \right] \underline{x}$ szorzatnak megfelelő oszlopvektor komponensei, mind pedig a $\underline{d}$ komponensei a $\underline{b}$ vektornak az $\underline{A}_1$ által meghatározott bázisra vonatkozó koordinátáit jelentik. S mivel adott bázis mellett bármely vektor koordinátái egyértelmüen meghatározottak, azért azok és csakis azok az $\underline{x}$ vektorok jöhetnek számitásba, amelyek kielégitik az

$$\left[\underline{E}_r \, , \, \underline{D} \right] \underline{x} = \underline{d}$$

egyenletet. E szerint ez az egyenlet egyenértékü az $\underline{A}\ \underline{x} = \underline{b}$ lineáris egyenletrendszerrel, amely m darab n-ismeretlenes egyenletből áll.

Az uj egyenlet vizsgálatához az $\underline{x}$ vektornak az $\underline{A}$ független oszlopvektoraira vonatkozó koordinátáiból tehát az első r komponenséből képezzük az

$$\underline{x}_1 = \left[x_1 , \, x_2 , \, \ldots , \, x_r \right]^*$$

vektort, a további n-r = s komponenséből pedig az

$$\underline{x}_2 = \left[x_{r+1} \, , \, x_{r+2} , \, \ldots , \, x_n \right]^*$$

vektort. Ez azt jelenti, hogy az $\underline{x}$ vektort az

$$\underline{x} = \begin{bmatrix} \underline{x}_1 \\ \\ \underline{x}_2 \end{bmatrix}$$

egyenlőségnek megfelelően partícionáljuk. Ezt felhasználva egyenletünk uj alakja:

$$\left[\underline{E}_r , \underline{D}\right] \cdot \begin{bmatrix} \underline{x}_1 \\ \\ \underline{x}_2 \end{bmatrix} = \underline{d} \, .$$

S mivel a mátrixok szorzási szabályai szerint

$$\left[\underline{E}_r , \underline{D}\right] \cdot \begin{bmatrix} \underline{x}_1 \\ \\ \underline{x}_2 \end{bmatrix} = \underline{E}_r \cdot \underline{x}_1 + \underline{D}\,\underline{x}_2 = \underline{x}_1 + \underline{D}\,\underline{x}_2 \, ,$$

ezért az előbbi egyenlet igy alakitható át:

$$\underline{x}_1 + \underline{D}\,\underline{x}_2 = \underline{d}.$$

Ebből pedig

$$\underline{x}_1 = \underline{d} - \underline{D}\,\underline{x}_2 \, .$$

Ezt a formulát az egyenletrendszer általános megoldásának szoktuk nevezni.

Az átalakitásokból ugyanis kitünik, hogy azok az x_1 , x_2, ..., x_n ismeretlenek, amelyek kielégitik az

$$\underline{A}\,\underline{x} = \underline{b}$$

egyenletet, kielégitik az

$$\underline{x}_1 = \underline{d} - \underline{D}\,\underline{x}_2$$

összefüggést is és viszont. E szerint azonban az $\underline{x}_2$ vektor komponenseinek bármilyen értéket is adunk, $\underline{x}_2$ és a hozzátartozó $\underline{x}_1$ M valamely $\underline{x}$ elemét adja. Mivel $\underline{x}_2$ komponenseinek nagyságát egymástól függetlenül, szabadon választhatjuk meg, az $\underline{x}_2$ vektorban szereplő ismeretleneket szabad ismeretleneknek, az $\underline{x}_1$ komponenseit pedig kötött ismeretleneknek nevezzük. Az elnevezés arra utal, hogy $\underline{x}_1$ értéke már függ a szabad ismeretlenekétől.

2. Tétel: A homogén egyenletrendszernek mindig van megoldása.

Bizonyitás: A homogén egyenletrendszernél mindig teljesül a kompatibilitás, mivel a null-vektor minden altérnek pontja. Ez ugy is belátható, hogy az $\underline{A}\ \underline{x} = \underline{0}$ egyenletrendszert az $\underline{x} = \underline{0}$ vektor mindig kielégiti.

Az $\underline{A}\ \underline{x} = \underline{0}$ egyenlettel kapcsolatban azonban felmerülhet a kérdés, hogy az $\underline{x} = \underline{0}$ triviális megoldáson kivül van-e még más megoldása is. Ez azonban itt is csak a számitások során derül ki.

Mivel a homogén egyenletrendszereknél a kompatibilitás mindig teljesül, a $\underline{b}$-nek megfelelő $\underline{0}$ oszlopvektor kiirása felesleges (ha kiirnánk, a 0 elemek változatlanok maradnának a számitások során!)

A lineáris egyenletrendszer lehetséges megoldásainak halmazát az

$$M = \left\{ \underline{x} \mid \underline{A}\ \underline{x} = \underline{b} \right\}$$

szimbólummal jelöltük. Ennek megfelelően az általános megoldást is fel lehet irni $\underline{x}$ alakban.

Ha ugyanis az

$$\underline{x}_1 = \underline{d} - \underline{D} \cdot \underline{x}_2$$

összefüggéshez hozzácsatoljuk az

$$\underline{x}_2 = \underline{0} + \underline{E}_s \, \underline{x}_2$$

nyilvánvaló egyenlőséget, ahol az "$\underline{E}_s$" egy s-edrendü egységmátrix, a két egyenlőség alapján

$$\underline{x} = \begin{bmatrix} \underline{x}_1 \\ \underline{x}_2 \end{bmatrix} = \begin{bmatrix} \underline{d} \\ \underline{0} \end{bmatrix} + \begin{bmatrix} -\underline{D} \\ \underline{E}_s \end{bmatrix} \underline{x}_2 .$$

Az $\underline{x}_2$ vektort ugy vezettük be, mint egy n-r = s elemü vektort, amelynek komponensei tetszőleges valós számok. Ennek megfelelően a következőkben az egyenlőség jobb oldalán szereplő $\underline{x}_2$-t jelölhetjük az egyszerüség kedvéért $\underline{t}$ vektorral. Ekkor az általános megoldás az

$$\underline{x} = \begin{bmatrix} \underline{d} \\ \underline{0} \end{bmatrix} + \begin{bmatrix} -\underline{D} \\ \underline{E}_s \end{bmatrix} \cdot \underline{t}$$

alakban irható, ahol a $\underline{t}$ vektor egy tetszőleges s elemü vektor. Ezzel a jelölés móddal az 1. példa általános megoldása a következő:

$$\underline{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 5,5 \\ 4,5 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} -1,5 & 0 \\ 0,5 & -1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} t_1 \\ t_2 \end{bmatrix} ,$$

ahol t_1 és t_2 tetszőleges skalárok.

Megjegyzések: a) Az $\underline{A} \, \underline{x} = \underline{b}$ egyenletrendszer általános megoldását adó

$$\underline{x}_1 = \underline{d} - \underline{D} \, \underline{x}_2$$

összefüggésben a kötött ismeretleneket reprezentáló $\underline{x}_1$ vektor komponenseinek száma megegyezik az $\underline{A}$ mátrix rangjával (r -rel), a szabad ismeretlenek száma pedig az ismeretlenek számának és a rangnak a különbsége (n-r = s).

b) Az általános megoldás második formáját tekintve megállapithatjuk, hogy abban az első tag tulajdonképpen az egyenletrendszer egy partikuláris megoldása. Ha ugyanis a $\underline{t}$ vektor helyébe a $\underline{0}$ vektort irjuk, azt kapjuk, hogy

$$\underline{x} = \begin{bmatrix} \underline{d} \\ \\ \underline{0} \end{bmatrix} + \begin{bmatrix} \underline{-D} \\ \\ \underline{E}_s \end{bmatrix} \cdot \underline{0} = \begin{bmatrix} \underline{d} \\ \\ \underline{0} \end{bmatrix} .$$

c) Az általános megoldásból a

$$\begin{bmatrix} \underline{-D} \\ \\ \underline{E}_s \end{bmatrix} = \underline{G}$$

kifejezésről pedig könnyen belátható, hogy a $\underline{G}$ mátrix oszlopvektorai lineárisan függetlenek. (Az $\underline{E}_s$ blokk vektorai ugyanis lineárisan függetlenek!) Ebből az következik, hogy a $\underline{G}\ \underline{t}$ szorzat a $\underline{G}$ mátrix oszlopvektorai által kifeszitett s dimenziós altér egy tetszőleges vektorát jelenti.
Ezen az alapon az $\underline{A}\ \underline{x} = \underline{b}$ egyenletrendszer általános megoldása ugy fogható fel, mint egy partikuláris megoldásnak és egy "s" dimenziós altér vektorainak az összege.

d) Ha az

$$\underline{x} = \begin{bmatrix} \underline{d} \\ \\ \underline{0} \end{bmatrix} + \begin{bmatrix} \underline{-D} \\ \\ \underline{E}_s \end{bmatrix} \underline{t}$$

általános megoldásban az "s" történetesen nulla, akkor a megoldás egyértelmü, ellenkező esetben számtalan sok megoldás van. Ennek megfelelően "s" valójában az egyenletrendszer <u>szabadságfoka</u>.
Ezt ugyis megfogalmazhatjuk, ha az $\underline{A}\ \underline{x} = \underline{b}$ egyenletrendszer szabadságfoka nulla, akkor az M megoldáshalmaznak csak egy eleme van, ha pedig $s > 0$, akkor M végtelen sok elemü.

e) Az $\underline{A}\ \underline{x} = \underline{0}$ homogén egyenletrendszer megoldáshalmaza s dimenziós alteret képez, amely a zérus altérre redukálódik, ha csak az $\underline{x} = \underline{0}$ lehet megoldás.

f) A gazdasági életben - középiskolai gyakorlattól eltérően - főleg olyan egyenletrendszerekkel találkozunk, amelyeknek szabadságfoka különbözik a "0"-tól.

A lineáris egyenletrendszer megoldására bemutatott eljárás alkalmas az

$$\underline{A}\ \underline{X} = \underline{B}$$

un. mátrix egyenlet megoldására is, ahol $\underline{A}$ és $\underline{B}$ adott mátrixok, $\underline{X}$ pedig ismeretlen.

3. Tétel: Az $\underline{A}\ \underline{X} = \underline{B}$ egyenlet megoldásának szükséges és elégséges feltétele, hogy $\underline{B}$ minden oszlopvektora benne feküdjék az $\underline{A}$ oszlopvektorterében.

Bizonyitás: A mátrixok szorzatára tanultak szerint az $\underline{A}\ \underline{X}$, vagyis $\underline{B}$ oszlopvektorai nem mások, mint az $\underline{A}$ oszlopvektorainak lineáris kombinációi, mégpedig olyan lineáris kombinációi, amelyekben skalárokként az $\underline{X}$ oszlopvektorainak megfelelő komponensek szerepelnek, azaz

$$\underline{A}\ \underline{x}_i = \underline{b}_i .$$

Igy az $\underline{A}\ \underline{X} = \underline{B}$ mátrixegyenlet megoldása tehát annyi lineáris egyenletrendszer megoldását igényli, ahány oszlopvektora van a $\underline{B}$ mátrixnak. Ha ezek mindegyikére teljesül a kompatibilitás, akkor $\underline{A}\ \underline{X} = \underline{B}$ megoldható, ahogy állitottuk.

A mátrixegyenletnél a kompatibilitás eldöntése a szokott módon történik. A különbség az eddigiekkel szemben csupán annyi, hogy az indulásnál az $\underline{A}$ oszlopvektorai mellé nem egyetlen vektort irunk, hanem odairjuk a $\underline{B}$ minden oszlopvektorát.

A következőkben az $\underline{A}\ \underline{X} = \underline{B}$ alaku mátrixegyenletek közül azokkal foglalkozunk, amelyekben az $\underline{A}$ mátrix kvadratikus és a $\underline{B}$ mátrix pedig az $\underline{A}$ rendjének megfelelő egységmátrix, vagyis

$$\underline{A}\ \underline{X} = \underline{E}$$

egyenleteket tekintjük. Ez speciális esetben az

$$a \cdot x = 1$$

skaláregyenletre redukálódik. Itt, ha $a \neq 0$, akkor

$$x = \frac{1}{a}.$$

Ezt az x számot az a szám reciprokának más szóval "inverzének" nevezzük. Ha az $a = 0$, akkor az a skalárnak nincs inverze.

A skalárarithmetikai meggondolásokhoz hasonlóan a fenti matrixegyenlet alapján kvadratikus mátrixokkal kapcsolatban is beszélhetünk inverzről.

<u>Valamely n-edrendü $\underline{A}$ mátrix inverzén egy olyan n-edrendü mátrixot értünk, amelynek az $\underline{A}$ mátrixszal képzett szorzata az n-edrendü egységmátrixot adja.</u>

Az a skalár inverzét az $ax = 1$ egyenlet megoldása adta. Az $\underline{A}$ mátrix inverzét az $\underline{A}\ \underline{X} = \underline{E}$ mátrixegyenlet megoldása adja, amennyiben az létezik.

1. <u>Tétel:</u> Az $\underline{A}\ \underline{X} = \underline{E}$ mátrixegyenlet megoldásának szükséges és elégséges feltétele, hogy az $\underline{A}$ mátrix nem-szinguláris mátrix legyen.

<u>Bizonyitás</u>: Ha az $\underline{X} = [\ \underline{x}_1, \underline{x}_2, \dots, \underline{x}\ ]$, akkor a szorzás értelmezése szerint

$$\underline{A}\,\underline{X} = \left[\underline{A}\,\underline{x}_1, \underline{A}\,\underline{x}_2, \dots, \underline{A}\,\underline{x}_n\right].$$

S mivel az $\underline{E}$ felirható az

$$\underline{E} = \left[\underline{e}_1, \underline{e}_2, \dots, \underline{e}_n\right]$$

alakban is, azért az $\underline{A}\,\underline{X} = \underline{E}$ egyenlet igy alakitható át:

$$\left[\underline{A}\ \underline{x}_1, \underline{A}\,\underline{x}_2, \ldots, \underline{A}\,\underline{x}_n\right] = \left[\underline{e}_1, \underline{e}_2, \ldots, \underline{e}_n\right].$$

Ebből következik, hogy az inverz meghatározása az

$$\begin{aligned} \underline{A}\,\underline{x}_1 &= \underline{e}_1 \\ \underline{A}\,\underline{x}_2 &= \underline{e}_2 \\ &\vdots \\ \underline{A}\,\underline{x}_n &= \underline{e}_n \end{aligned}$$

egyenlőségek által meghatározott lineáris egyenletrendszerek megoldására vezethető vissza (amint azt az előző pontban az $\underline{A}\,\underline{X} = \underline{B}$ mátrixegyenletnél is láttuk). Az igy nyert n egyenletrendszer mindegyike n ismeretlent tartalmaz, de együtthatómátrixuk azonos. Ezért mind az n egyenletrendszer csak abban az esetben lehet kompatibilis, ha a jobboldalon álló egységvektorok mindegyike benne fekszik az $\underline{A}$ oszlopvektorterében. Ez azonban csak akkor következhet be, ha $\underline{A}$ rangja n, vagyis az $\underline{A}$ mátrix nem-szinguláris. Belátható, hogy az inverz létezéséhez az $\underline{A}$ nem-szinguláris volta nemcsak szükséges feltétel, hanem elégséges is. Ez esetben ugyanis az $\underline{A}\,\underline{x}_i = \underline{e}_i$ egyenletrendszerek mind megoldhatók, még pedig egyértelmüen, s a megoldások éppen az inverz oszlopvektorait szolgáltatják.

Ezzel állitásunkat bebizonyitottuk.

Mivel a mátrixok szorzata általában nem kommutativ, felmerülhet az a lehetőség, hogy az $\underline{A}$ mátrixnak két különböző inverze van: egy jobboldali és egy baloldali inverz. Jobboldali inverzen olyan $\underline{X}$ mátrixot értünk, amelyre az

$$\underline{A}\,\underline{X} = \underline{E}\,,$$

baloldali inverzen meg egy olyan $\underline{Y}$ mátrixot, amelyre az

$$\underline{Y}\,\underline{A} = \underline{E}$$

egyenlőség teljesül. Ez utóbbi egyenlettel kapcsolatban megállapithatjuk, hogy a megoldásnak ez esetben is az a szükséges és elégséges feltétele, hogy az $\underline{A}$ nem szinguláris mátrix legyen. Ugyanis a mátrixok szorzatánál tanultak szerint

$$(\underline{Y}\,\underline{A})^* = \underline{A}^*\ \underline{Y}^*$$

és

$$\underline{E}^* = \underline{E}\,.$$

Ezért $\underline{Y}\,\underline{A} = \underline{E}$ egyenlet megoldását visszavezethetjük az

$$\underline{A}^{*}\ \underline{Y}^{*} = \underline{E}$$

egyenletre. Ez pedig csak akkor oldható meg, ha $\underline{A}^{*}$ nem-szinguláris. Ha pedig az $\underline{A}^{*}$ nem-szinguláris, akkor az $\underline{A}$ is nem-szinguláris mátrix.

2. Tétel: Ha az $\underline{A}\,\underline{X} = \underline{E}$ és $\underline{Y}\,\underline{A} = \underline{E}$ egyenletekkel definiált inverzek léteznek, akkor nem különböznek egymástól.

Bizonyitás: Tegyük fel, hogy $\underline{A}$ nem-szinguláris. Ez esetben mind az

$$\underline{A}\,\underline{X} = \underline{E}$$

és mind az

$$\underline{Y}\,\underline{A} = \underline{E}$$

mátrixegyenlet egyértelmüen megoldható. Szorozzuk meg ezekután az első egyenletet balról $\underline{Y}$-nal, a másodikat jobbról $\underline{X}$-szel, amikor is az

$$\underline{Y}\,\underline{A}\,\underline{X} = \underline{Y}\,\underline{E}$$

és

$$\underline{Y}\,\underline{A}\,\underline{X} = \underline{E}\,\underline{X}$$

egyenleteket kapjuk.

Mivel azonban

$$\underline{Y}\,\underline{E} = \underline{Y} \text{ és } \underline{E}\,\underline{X} = \underline{X},$$

az előbbi egyenletek igy is irhatók:

$$\underline{Y}\,\underline{A}\,\underline{X} = \underline{Y}$$

$$\underline{Y}\,\underline{A}\,\underline{X} = \underline{X}\,.$$

Ebből pedig látható, hogy

$$\underline{Y} = \underline{X}.$$

Azt az n-edrendü kvadratikus mátrixot, amely mind az $\underline{A}\,\underline{X} = \underline{E}$, mind az $\underline{Y}\,\underline{A} = \underline{E}$ matrixegyenletet kielégiti, az $\underline{A}$ mátrix inverzének nevezzük, és jelölésére - skaláraritmetika mintájára - bevezetjük az

$$\underline{A}^{-1}$$

szimbólumot.

Megjegyzés: a) Csak nem-szinguláris mátrixoknak van egyértelmüen meghatározott inverze.

b) Az $\underline{A}$ mátrix és inverze, az $\underline{A}^{-1}$ mátrix, eleget tesz az alábbi követelménynek

$$\underline{A} \cdot \underline{A}^{-1} = \underline{A}^{-1} \cdot \underline{A} = \underline{E},$$

vagyis az $\underline{A}$ mátrixnak és inverzének szorzására érvényes a kommutativitás.

5. A lineáris transzformáció

5.1. A lineáris transzformációról általában

A bázistranszformáció fogalmának megismerése után a transzformáció fogalmát uj vonatkozásban vetjük fel.

Ebben az értelemben a transzformáció bizonyos halmazok között létesithető leképezés megjelölésére szolgál. Igy tulajdonképpen a függvényfogalom általánositását jelenti.

Az eddigi ismereteink szerint függvényről akkor beszéltünk, amikor valamely x változó minden lehetséges értékéhez valamilyen előirással hozzárendeltünk egy f(x) szimbólummal jelölt számot. Az x változó lehetséges értékei alkották az értelmezési tartományt, az f(x) értékek halmaza pedig a függvény értékkészletét. E keretek között mind az x mind pedig az f(x) valamilyen skalárt jelentett.

Ilyen természetü hozzárendelés azonban nemcsak skalárok közt lehetséges. Tekintsük ugyanis az L_n lineáris tér (n-elemü) vektorait, s valamilyen utasitásnak megfelelően, ezek mindegyikéhez rendeljünk egy skalárt. Ha a tér vektorait $\underline{x}$ szimbólummal jelöljük, akkor a skalárt az $f(\underline{x})$ jelölheti. E leképezésnél az értelmezési tartomány elemei vektorok, az értékkészlet elemei pedig skalárok.

A középiskolában megismert függvény-fogalom további általánositásához jutunk, ha olyan leképezést tekintünk, amelyben mind az értelmezési tartomány, mind pedig az értékkészlet elemeit az L_n vektorai alkotják. Ezek előre bocsátása után a transzformáció fogalmára az alábbi definiciót adjuk:

Ha az L_n lineáris tér minden $\underline{x}$ vektorához valamilyen előirással hozzárendeljük az L_n-nek valamilyen $T(\underline{x})$ szimbólummal jelölt vektorát, akkor azt mondjuk, hogy az L_n-ben egy T transzformáció van értelmezve.

A $T(\underline{x})$ szimbólum, vagy röviden csak a T szimbólum - az f(x)-hez hasonlóan - itt is a hozzárendelésnek valamely létező, de konkrétan meg nem adott formáját jelenti. Magát az $\underline{x}$ vektort tárgyvektornak az $\underline{x}$-hez rendelt $T(\underline{x})$ vektort pedig képvektornak nevezzük. A képvektort néha röviden $\underline{y}$-nal jelöljük. Ez esetben

$$\underline{y} = T(\underline{x}) .$$

Attól függően, hogy a $T(\underline{x})$ szimbólumnak milyen konkrét formát adunk, különböző transzformációkról beszélhetünk.

1. Példa:

$$T(\underline{x}) = 6\,\underline{x}$$

transzformáció azt fejezi ki, hogy az $\underline{x}$ tárgyvektorhoz annak 6-szorosát rendeltük képelemként. Ezt a transzformációt nyujtásnak szoktuk nevezni. Általános alakja: $T(x) = \lambda \underline{x}$.

2. Példa: $T(\underline{x}) = \underline{x} + \underline{a}$

transzformációnál, $\underline{a} \neq \underline{0}$ esetén az $\underline{x}$ tárgyvektor a hozzá tartozó képvektortól egy konstans $\underline{a}$ vektorban különbözik.

A továbbiakban a transzformációknak csak azzal a speciális tipusával foglalkozunk, amelyet lineáris transzformációnak nevezünk.

Az L_n-ben értelmezett T transzformációt lineáris transzformációnak nevezzük, ha eleget tesz az alábbi két, un. lineáritási feltételnek:

$$1.\quad T(\underline{x}_1 + \underline{x}_2) = T(\underline{x}_1) + T(\underline{x}_2)$$

$$2.\quad T(\lambda \underline{x}) = \lambda T(\underline{x}) .$$

Az első feltétel azt jelenti, hogy bármely két tárgyvektor összegéhez tartozó képvektor megegyezik a két tárgyvektorhoz tartozó képvektor összegével. A második feltétel pedig azt jelenti, hogy bármely $\underline{x}$ tárgyvektor λ -szorosához tartozó képvektor egyenlő az $\underline{x}$ tárgyvektorhoz tartozó képvektor λ-szorosával.

Könnyen belátható, hogy az említett példák közül az első lineáris transzformáció, a második pedig nem.

A $T(\underline{x}) = 6\underline{x}$ transzformáció lineáris, mert kielégíti a linearitási feltételeket:

$$6(\underline{x}_1 + \underline{x}_2) = 6\underline{x}_1 + 6\underline{x}_2$$

és

$$6(\lambda \underline{x}) = \lambda\, 6\, \underline{x} .$$

Ezzel szemben a $T(\underline{x}) = \underline{x} + \underline{a}$ transzformáció nem lineáris, mert

$$T(\underline{x}_1) = \underline{x}_1 + \underline{a}$$

$$T(\underline{x}_2) = \underline{x}_2 + \underline{a}$$

ahonnan

$$T(\underline{x}_1) + T(\underline{x}_2) = \underline{x}_1 + \underline{x}_2 + 2\,\underline{a} ,$$

ami azonban nem egyenlő a

$$T(\underline{x}_1 + \underline{x}_2) = (\underline{x}_1 + \underline{x}_2) + \underline{a} = \underline{x}_1 + \underline{x}_2 + \underline{a}$$

kifejezéssel. Ha pedig a lineáritási feltételek valamelyike nem teljesül, a transzformáció már nem lineáris.

1. <u>Tétel</u>: Egy L_n-ben értelmezett $T(\underline{x})$ transzformációt akkor és csak akkor nevezünk **lineáris transzformáció**nak, ha az felirható

$$T(\underline{x}) = \underline{A}\ \underline{x}$$

alakban.

<u>Bizonyitás</u>: Ha egy $T(\underline{x})$ transzformáció felirható $T(\underline{x}) = \underline{A}\ \underline{x}$ alakban, akkor lineáris transzformáció, ugyanis a mátrixokra megismert müveleti szabályokból következik, hogy

a) $\underline{A}(\underline{x}_1 + \underline{x}_2) = \underline{A}\,\underline{x}_1 + \underline{A}\,\underline{x}_2$

b) $\underline{A}(\lambda\,\underline{x}) = \lambda(\underline{A}\,\underline{x})$.

Ez pedig azt jelenti, hogy a $T(\underline{x}) = \underline{A}\,\underline{x}$ transzformáció eleget tesz a két linearitási feltételnek. Ki kell még mutatnunk, ha a $T(\underline{x})$ transzformáció lineáris transzformáció, akkor felirható $T(\underline{x}) = \underline{A}\,\underline{x}$ alakban. Ennek érdekében tegyük fel, hogy a L_n-ben értelmezve van egy $T(\underline{x})$ lineáris transzformáció. Tegyük fel továbbá, hogy a

$$\underline{b}_1, \underline{b}_2, \ldots, \underline{b}_n$$

vektorok az L_n-ben bázist alkotnak. Ez esetben az L_n bármelyik $\underline{x}$ vektora kifejezhető, és pedig egyértelmüen az

$$\underline{x} = x_1\underline{b}_1 + x_2\underline{b}_2 + \ldots + x_n\underline{b}_n$$

alakban. Az $\underline{x}$-hez tartozó képvektor a linearitás miatt

$$T(\underline{x}) = x_1\,T(\underline{b}_1) + x_2\,T(\underline{b}_2) + \ldots + x_n\,T(b_n)$$

alakban irható.

A képvektor ezek szerint nem más, mint a bázisvektorokhoz tartozó képvektorok lineáris kombinációja.
Ha bevezetjük a

$$\begin{aligned} T(\underline{b}_1) &= \underline{a}_1 \\ T(\underline{b}_2) &= \underline{a}_2 \\ &\vdots \\ T(\underline{b}_n) &= \underline{a}_n \end{aligned}$$

jelöléseket, akkor a $T(\underline{x})$ a

$$T(\underline{x}) = x_1\,\underline{a}_1 + x_2\,\underline{a}_2 + \ldots + x_n\underline{a}_n = \left[\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_n\right]\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

alakra hozható.

A kérdéses n . n-es

$$\left[\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_n\right] = \underline{A}\cdot$$

mátrixot a transzformáció mátrixának nevezzük, még pedig a $\underline{b}_1, \underline{b}_2, \dots, \underline{b}_n$ bázisra vonatkozó mátrixának. E szerint a bázis megválasztása döntő a transzformáció mátrixának formájára.
Ezzel állitásunkat bebizonyitottuk.

A tétel szerint a példaként bemutatott $T(\underline{x}) = 6\underline{x}$ transzformációt is fel lehet irni a lineáris transzformációk un. általános alakjában a $T(\underline{x}) = \underline{A}\ \underline{x}$ alakban, ugyanis

$$T(\underline{x}) = 6(\underline{E}\ \underline{x}) = (6\ \underline{E})\ \underline{x}$$

ahol a $6\underline{E}$ mátrixnak egy olyan diagonális mátrix felel meg, amelynek diagonális elemei 6-tal egyenlők.

Ha a $T(\underline{x}) = \underline{A}\ \underline{x}$ lineáris transzformációban a transzformáció mátrixa $\underline{0}$, zérustranszformációról beszélünk.

A zérustranszformáció az L_n minden vektorának a $\underline{o}$ vektort felelteti meg.

Ha a $T(\underline{x})$ transzformáció a tér minden pontjának saját magát felelteti meg, egységtranszformációról van szó.
Jelölése $E(\underline{x}) = \underline{x}$.

2. Tétel: Bármely T lineáris transzformáció a $\underline{0}$ vektort a $\underline{o}$ vektorba viszi át, vagyis

$$T(\underline{o}) = \underline{o}\ .$$

Bizonyitás: Irjuk fel a lineáris kombinációt $T(\underline{x}) = \underline{A}\ \underline{x}$ alakban. Mivel a mátrixok szorzási szabálya értelmében $\underline{A}\ \underline{o} = \underline{o}$, ezért $T(\underline{o}) = \underline{o}$ valóban.

3. Tétel: A képvektorok halmaza mindig olyan alteret alkot, amelynek dimenziója megegyezik a transzformáció matrixának rangjával.

Bizonyitás: Legyen ugyanis $\underline{A}$ a transzformáció mátrixa. Ez esetben a képvektorok halmaza a $T(\underline{x}) = \underline{A}\ \underline{x}$ összefüggésnek eleget tevő vektorok összessége lesz. Ez a halmaz pedig nem más, mint az $\underline{A}$ oszlopvektortere. Az oszlopvektortérről pedig tudjuk, hogy annak dimenziója megegyezik az $\underline{A}$ mátrix rangjával.

5.2. Műveletek lineáris transzformációkkal

A lineáris transzformációkra müveletek is értelmezhetők. Az adott müveletek teljes összhangban állnak a mátrixaritmetikában értelmezett müveletekkel, mivel azokat azért alakitották ki, hogy megteremtsék a lineáris transzformációk numerikus eszközét. A mátrixokkal való müveleti szabályok tartalmi hátterét tehát a lineáris transzformációk adják meg.

A lineáris transzformációkra vonatkozó müveleti szabályok definiálása előtt állapodjunk meg abban, hogy a L_n-ben értelmezett T_i lineáris transzformációk általános egyenletében szereplő $\underline{A}_i$ kvadratikus mátrixok mindegyike ugyanabban a bázisban adott.

Ezekután a következő müveleteket definiáljuk:

a) Összeadás: A T_1 és a T_2 lineáris transzformációk összegén azt

$$T_1 + T_2 = T$$

transzformációt értjük, amelyre nézve bármely $\underline{x}$ vektor esetén

$$T_1(\underline{x}) + T_2(\underline{x}) = T(\underline{x}) .$$

Ha $T_1(\underline{x}) = \underline{A}_1\,\underline{x}$ és $T_2(\underline{x}) = \underline{A}_2\,\underline{x}$, akkor ezek összege

$$\underline{A}_1\,\underline{x} + \underline{A}_2\,\underline{x} = (\underline{A}_1 + \underline{A}_2)\,\underline{x}$$

összefüggés alapján az

$$(\underline{A}_1 + \underline{A}_2)\underline{x} = \underline{A}\,\underline{x}$$

alakban is irható. Ez pedig minden lehetséges $\underline{x}$-re csak akkor állhat fenn, ha

$$\underline{A}_1 + \underline{A}_2 = \underline{A} .$$

Ez pedig éppen azt mutatja, hogy két lineáris transzformáció összegéhez tartozó transzformáció mátrixa egyenlő a taghoz tartozó mátrixok összegével. E szerint a transzformációk mátrixával ugyanolyan értelmü müveletet kell végeznünk, mint magukkal a transzformációkkal.

b) Skalárral való szorzás: A T_1 lineáris transzformáció λ-szorosán azt a

$$\lambda T_1 = T$$

transzformációt értjük, amelyre nézve bármely $\underline{x}$ vektor esetén

$$\lambda T_1(\underline{x}) = T(\underline{x})$$

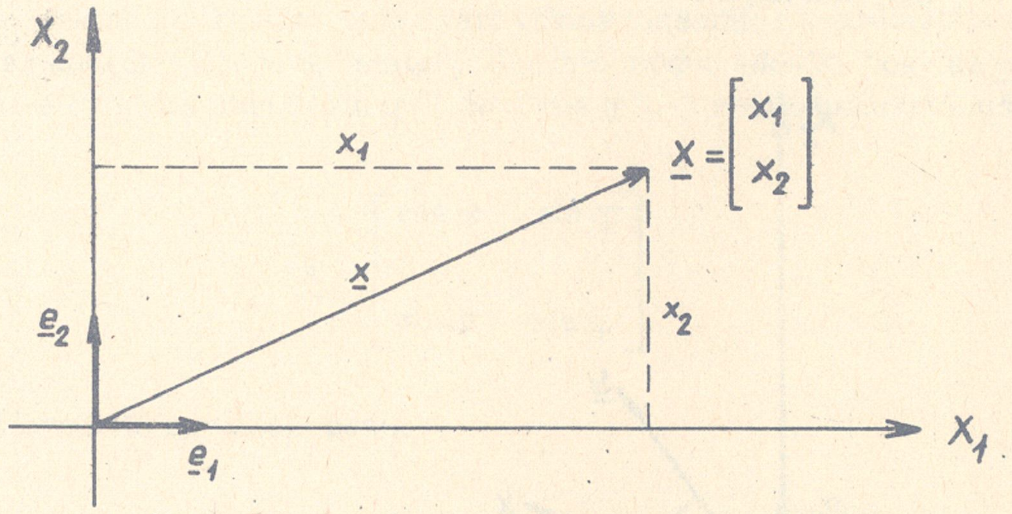
áris transzformáció. Mind a három esetben teljesülnek a linearitási feltételek. Ennek bizonyitására azonban nem térünk ki.

Az eddig tárgyalt müveletek kettőnél több lineáris transzformációra is kiterjeszthetők. Ezekre a müveletekre, azután ugyanazok a müveleti szabályok érvényesek, mint a mátrixokra, ezért ezeket sem vizsgáljuk meg részleteiben.

Megemlitjük még, hogy adott T transzformáció esetén nem negativ egész kitevőkre értelmezhetjük a hatványozást is az alábbi előirás szerint:

$$
\begin{aligned}
T^0 &= E \\
T^1 &= T \\
T^2 &= T\,T \\
T^3 &= T\,T\,T \\
&\vdots
\end{aligned}
$$

1. Példa: Tekintsük a sikot és abban vegyünk fel egy derékszögü koordináta rendszert. Rendeljük ezután a sik minden $\underline{x} = [\,x_1,\ x_2\,]^*$ vektorához képvektorként az $\underline{x}$ tárgyvektornak az x_1 tengelyre eső merőleges vetületét.



13. ábra

Állapitsuk meg a transzformáció mátrixát az

$$\underline{e}_1 = [\,1,\ 0\,]^* \qquad \underline{e}_2 = [\,0,\ 1\,]^*$$

bázisra vonatkozóan.

A transzformációról kimutatható, hogy lineáris transzformáció. Ennek megfelelően a transzformáció mátrixa:

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

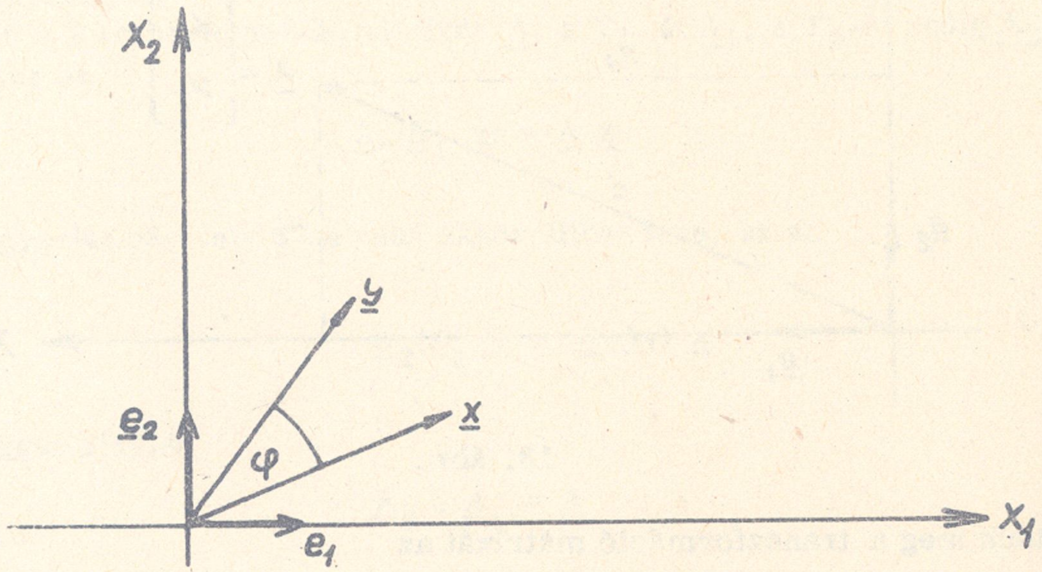
$$(\underline{a}_1 = T_1(\underline{e}_1) = [1,\ 0]^* \text{ és } \underline{a}_2 = T_1(\underline{e}_2) = [0,\ 0]^*),$$

azaz a transzformáció egyenlete:

$$T_1(\underline{x}) = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Az ilyen transzformációt vetitésnek nevezzük.

2. Példa: Tekintsük a sikot, s azon az egységvektorok által alkotott bázist. Rendeljük ezután a sik minden $\underline{x}$ vektorához képvektorként azt a vektort, amelyet ugy nyerünk, hogy az $\underline{x}$ vektornak megfelelő pontot az $\underline{x}$ abszolut értékével egyenlő sugárral huzott körön egy adott φ szöggel elforditjuk.

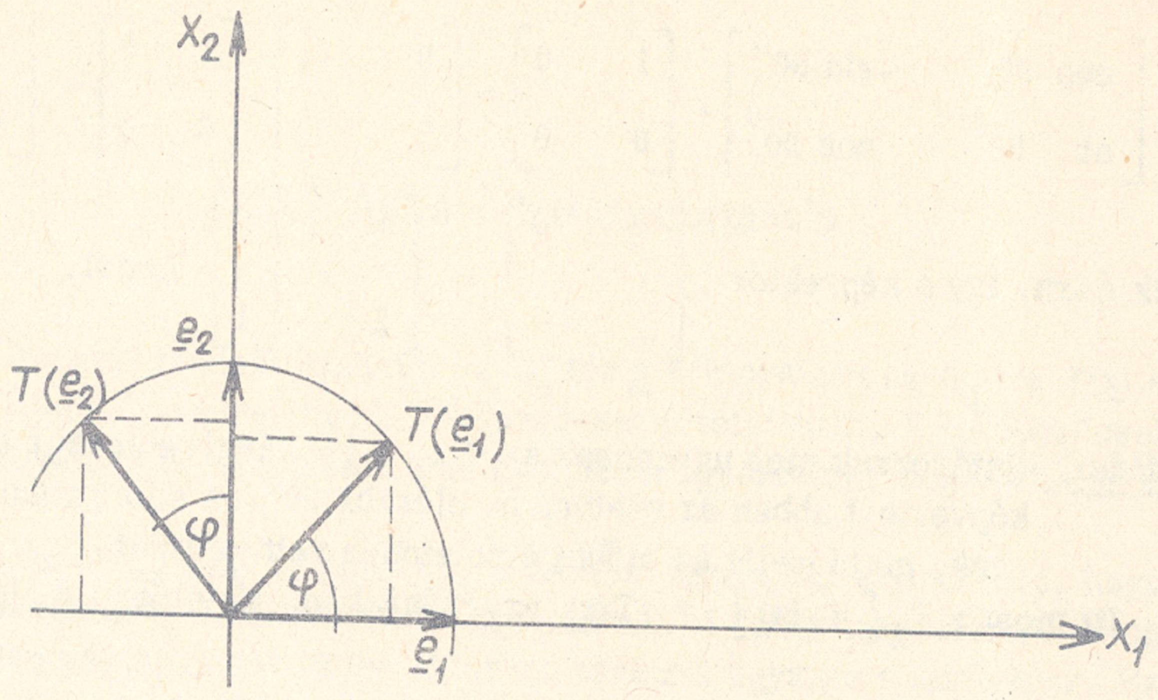


14. ábra

Irjuk fel a transzformáció egyenletét.

Itt is a bázisvektorokhoz tartozó képvektorokat kell először megállapitanunk, mert ennek alapján már felirható a transzformáció $\underline{A}$ mátrixa. Kimutatható ugyanis, hogy ez a transzformáció is lineáris transzformáció.

A képvektorok felirásában segit az alábbi ábra:



15. ábra

Ha ugyanis a bázisvektorokat elforgatjuk φ szöggel, akkor az ábra szerint az $\underline{e}_1$ vektor elforgatása után nyert vektor vizszintes koordinátája a "$\cos\varphi$" függőleges koordinátája meg "$\sin\varphi$". Hasonló módon adódik, hogy az $\underline{e}_2$ koordinátái az elforgatás után "$-\sin\varphi$" és "$\cos\varphi$". A transzformáció mátrixa igy

$$\underline{A} = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}$$

A transzformáció egyenlete pedig

$$T_2(\underline{x}) = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

5.3 A lineáris transzformáció invariáns alterei

Az L lineáris tér valamely L' lineáris alterét a tér $T(\mathbf{x})$ lineáris transzformációjára invariáns altérnek nevezzük, ha L' minden $\mathbf{x}$ elemére a $T(\mathbf{x})$ is benne van az L' altérben.

Az n elemű vektorok terében ez azt jelenti, hogy az L'-ből kiválasztott minden tárgyvektorra a hozzátartozó $\mathbf{y}=\mathbf{A}\mathbf{x}$ is eleme L'-nek.

Könnyű belátni, hogy a lineáris transzformációk magtere invariáns alteret képez. A magtér ugyanis a tárgyvektortér azon $\mathbf{x}$ elemeinek összessége, amelyekre $\mathbf{A}\mathbf{x}=\mathbf{0}$, azaz

$$M = \{\mathbf{x}/\mathbf{A}\mathbf{x} = \mathbf{0}\},$$

ahol az M halmaz a 4.2-ben részletezett

$$\mathbf{x} = \begin{bmatrix} -\mathbf{D} \\ \mathbf{E}_s \end{bmatrix} \mathbf{t}$$

összefüggéssel megadható s dimenziós altér. Az M-ből minden elem képe a zérusvektor, amely ugyancsak eleme az altérnek, és ez éppen az invariáns sajátosságra vall.

A komplex lineáris térben értelmezett lineáris transzformációk elméletében jelentős szerepük van az *egydimenziós* invariáns altereknek.

Legyen $L^{(1)}$ a zérustól különböző $\mathbf{x}$ n elemű vektor által generált altér. Mivel az $L^{(1)}$ alteret az $\mathbf{x}\neq\mathbf{0}$ vektor számszorosai alkotják, ahol skalárszorzók a komplex számok, az altér felírható $\alpha\mathbf{x}$ alakban. Az $L^{(1)}$ valódi altere az L_n lineáris térnek és akkor invariáns altér, ha $\mathbf{A}\mathbf{x}$ is eleme. Ez azt jelenti, hogy a generáló $\mathbf{x}\neq\mathbf{0}$ vektor skalárszorzói között van olyan, amellyel szorozva az $\mathbf{x}$ vektort $\mathbf{A}\mathbf{x}$ adódik. Ha ez a szám éppen λ, akkor

$$\mathbf{A}\mathbf{x}=\lambda\mathbf{x} \tag{140}$$

teljesül.

A következőkben ezekkel az egydimenziós invariáns altereket generáló vektorokkal és a hozzájuk tartozó skalárszorzókkal foglalkozunk.

5.4 Sajátérték, sajátvektor

Ha az

$$\mathbf{Ax} = \lambda \mathbf{x}$$

egyenletnek valamely rögzített λ mellett van a zérustól különböző megoldása, vagyis létezik olyan

$$\mathbf{x} \neq \mathbf{0}$$

vektor, amelyet a $T(\mathbf{x})$ lineáris transzformáció a λ-szorosába visz át, akkor a λ számot az $\mathbf{y} = T(\mathbf{x})$ lineáris transzformáció sajátértékének nevezzük, az egyenletet kielégítő $\mathbf{x}$ vektort pedig a lineáris transzformáció λ sajátértékhez tartozó sajátvektorának.

Az L_n-ben értelmezett $T(\mathbf{x})$ lineáris transzformációra invariáns $L^{(1)}$ egydimenziós alteret generáló $\mathbf{x} \neq \mathbf{0}$ vektor tehát sajátvektor. Mivel $L^{(1)}$ invariáns altér, azért

$$T(\mathbf{x}) = \mathbf{Ax}$$

esetén $\mathbf{Ax} \in L^{(1)}$, így

$$\mathbf{Ax} = \lambda \mathbf{x},$$

ahol az alkalmasan megválasztott λ skalár sajátérték.

33. tétel. *A $T(\mathbf{x})$ lineáris transzformáció sajátértékei között akkor és csak akkor fordul elő a zérus, ha a transzformáció mátrixa szinguláris.*

Bizonyítás. Ha $\mathbf{A}$ szinguláris mátrix, akkor $\mathbf{K} = \lambda\mathbf{E} - \mathbf{A}$ mátrix $\lambda = 0$ mellett szinguláris, így a zérus sajátérték. Megfordítva, ha $\lambda = 0$ sajátérték, $\mathbf{K} = 0\mathbf{E} - \mathbf{A}$ összefüggés alapján $\mathbf{A}$ is szinguláris. Ezzel az állítást igazoltuk.

A 4. példában $\mathbf{A}$ mindegyik sajátértéke zérus volt. Ennek a transzformációnak a mátrixa nilpotens mátrix, amelyre $\mathbf{A}^2 \neq \mathbf{0}$, de $\mathbf{A}^3 = \mathbf{0}$ teljesül. Bizonyítható, hogy a nilpotens mátrixok valamennyi sajátértéke 0.

36. tétel. *A $T(\mathbf{x})$ lineáris transzformáció különböző sajátértékeihez tartozó sajátvektorok lineárisan függetlenek.*

Bizonyítás. Legyenek

$$\lambda_1, \lambda_2, \ldots, \lambda_k$$

a $T(\mathbf{x})$ transzformáció egymástól különböző sajátértékei,

$$\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$$

pedig a hozzájuk tartozó sajátvektorok.

A tételt teljes indukcióval bizonyítjuk: $k=1$ esetén igaz az állítás, mert a sajátvektor nem a nullavektor, ezért egymagában lineárisan független. Tegyük fel, hogy $k-1$ különböző sajátérték esetén a sajátértékekhez tartozó sajátvektorok lineárisan függetlenek. A továbbiakban indirekt módon bizonyítunk. Feltesszük, hogy a k különböző sajátértékhez tartozó sajátvektorok lineárisan függőek, vagyis a

$$\xi_1 \mathbf{x}_1 + \xi_2 \mathbf{x}_2 + \ldots + \xi_k \mathbf{x}_k = \mathbf{0} \tag{145}$$

lineáris kombinációban zérustól különböző skalár szorzó is van, pl. a $\xi_1 \neq 0$. Alkalmazzuk a T transzformációt a (145) összefüggésre. Mivel a zérusvektort bármely lineáris transzformáció a zérusvektorba viszi át, ezért

$$\mathbf{A}(\xi_1 \mathbf{x}_1 + \xi_2 \mathbf{x}_2 + \ldots + \xi_k \mathbf{x}_k) = \mathbf{0}. \tag{146}$$

Felhasználva, hogy a $T(\mathbf{x}) = \mathbf{A}\mathbf{x}$ transzformáció összeg- és aránytartó, (146) így írható

$$\xi_1 \mathbf{A}\mathbf{x}_1 + \xi_2 \mathbf{A}\mathbf{x}_2 + \ldots + \xi_k \mathbf{A}\mathbf{x}_k = \mathbf{0},$$

illetve az

$$\mathbf{A}\mathbf{x}_i = \lambda_i \mathbf{x}_i \qquad (i=1, \ldots, k)$$

helyettesítések után:

$$\xi_1 \lambda_1 \mathbf{x}_1 + \xi_2 \lambda_2 \mathbf{x}_2 + \ldots + \xi_k \lambda_k \mathbf{x}_k = \mathbf{0}. \tag{147}$$

Vonjuk ki ebből az egyenlőségből az $\mathbf{x}_i$ vektorok felírt lineáris kombinációjának λ_k-szorosát. A különbséget a

$$\xi_1(\lambda_1 - \lambda_k)\mathbf{x}_1 + \xi_2(\lambda_2 - \lambda_k)\mathbf{x}_2 + \ldots + \xi_{k-1}(\lambda_{k-1} - \lambda_k)\mathbf{x}_{k-1} = \mathbf{0}$$

összefüggés mutatja, ahol a feltételek szerint

$$\xi_1 \neq 0 \qquad \text{és} \qquad \lambda_1 - \lambda_k \neq 0$$

teljesül. Ezek szerint az $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{k-1}$ vektorok lineárisan összefüggőek volnának és így ellentmondáshoz jutottunk, ami az állítást igazolja. Ezzel a tételt bizonyítottuk.

A tétel tehát kimondja, hogy ha egy lineáris transzformáció sajátértékei mind különbözőek, akkor a hozzájuk tartozó sajátvektorok lineárisan függetlenek.

8.1. Bilineáris formák

Az adott $\mathbf{A} \in \mathcal{M}(p, q)$ mátrix segítségével képzett

$$B(\mathbf{x}, \mathbf{y}) = \mathbf{x}^* \mathbf{A} \mathbf{y} \tag{8.1.1}$$

kifejezést (ahol $\mathbf{x} \in E_p$ és $\mathbf{y} \in E_q$) az $\mathbf{x}$ és $\mathbf{y}$ vektorra nézve *bilineáris formának nevezzük*. Az ismert műveleti szabályok értelmében a $B(\mathbf{x}, \mathbf{y})$ olyan skalárt jelent, amelynek értéke az $\mathbf{x}$ és az $\mathbf{y}$ megválasztásától függ. Ha az $\mathbf{y}$-t rögzítjük az E_q valamely pontjában, akkor a $B(\mathbf{x}, \mathbf{y})$-nak megfelelő skalár az $\mathbf{x}$ komponenseinek lineáris függvénye lesz. Éppen így: ha az $\mathbf{x}$-et rögzítjük egy pontban, akkor a $B(\mathbf{x}, \mathbf{y})$ az $\mathbf{y}$ komponenseinek lesz lineáris függvénye. Valójában innen ered a bilineáris elnevezés, ami magyarul kétszeresen lineárisat jelent.

A (8.1.1) alatt szereplő $\mathbf{A}$ mátrixot az adott *bilineáris forma mátrixának* nevezzük, az $\mathbf{A}$ rangját pedig *a bilineáris forma rangjának*.

Ha például

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

akkor a megfelelő bilineáris forma

$$B(\mathbf{x}, \mathbf{y}) = [x_1, x_2] \begin{bmatrix} 2 & -1 & -1 \\ 1 & 0 & 3 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}.$$

A műveleti szabályok értelmében ez a

$$B(\mathbf{x}, \mathbf{y}) = 2x_1y_1 - x_1y_2 - x_1y_3 + x_2y_1 + 3x_2y_3$$

skaláralakban is felírható.

Ha a (8.1.1) alatt szereplő $\mathbf{A}$ mátrix i-edik sorának j-edik elemét a_{ij}-vel jelöljük, akkor — az előbbi példa mintájára — az ott definiált bilineáris formának a

$$B(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{p} \sum_{j=1}^{q} a_{ij} x_i y_j$$

skalárkifejezést feleltetjük meg. S ha egy bilineáris forma történetesen skaláralakban van megadva, akkor az elmondottak alapján könnyen átírhatjuk a vele ekvivalens mátrixalakba is. Így pl. a

$$2x_1y_1 - 3x_1y_2 + 4x_2y_1 - 3x_3y_1 + x_3y_2$$

bilineáris forma ekvivalens az

$$[x_1, x_2, x_3] \begin{bmatrix} 2 & -3 \\ 4 & 0 \\ -3 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

mátrixkifejezéssel.

Mivel minden skalár azonos saját transzponáltjával, azért a (8.1.1) alatti reláció ekvivalens a

$$B(\mathbf{x}, \mathbf{y}) = \mathbf{y}^* \mathbf{A}^* \mathbf{x} \tag{8.1.2}$$

relációval. Itt említjük meg a *szimmetrikus bilineáris forma* fogalmát. Egy bilineáris forma akkor szimmetrikus, ha *a mátrixa szimmetrikus.*[1] Ha tehát a

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

bilineáris forma szimmetrikus, akkor $\mathbf{A}^* = \mathbf{A}$, s így a (8.1.2) alapján fennáll a

$$B(\mathbf{x}, \mathbf{y}) = B(\mathbf{y}, \mathbf{x}) \tag{8.1.3}$$

összefüggés.[2] Ha egy szimmetrikus bilineáris forma mátrixa történetesen az egységmátrixszal egyezik meg, akkor eredményül skaláris szorzatot nyerünk. Ha ugyanis $\mathbf{A} = \mathbf{E}$, akkor a (8.1.1) a

$$B(\mathbf{x}, \mathbf{y}) = \mathbf{x}^* \mathbf{y} \tag{8.1.4}$$

alakot ölti. Ez azt jelenti, hogy a skaláris szorzat valójában egy speciális szimmetrikus bilineáris forma. (A következő alfejezetben erre a kérdésre még visszatérünk.)

8.2. Kvadratikus formák

Ha a kvadratikus

$$B(\mathbf{x}, \mathbf{y}) = \mathbf{x}^*\mathbf{A}\mathbf{y} \tag{8.2.1}$$

bilineáris formában az $\mathbf{y}$ helyébe is az $\mathbf{x}$ vektort írjuk, akkor a

$$Q(\mathbf{x}) = \mathbf{x}^*\mathbf{A}\mathbf{x} = \mathbf{x}^*\mathbf{A}^*\mathbf{x} \tag{8.2.2}$$

kifejezést nyerjük, amelyet (az $\mathbf{x}$-re nézve) *kvadratikus formának* nevezünk. Ha az $\mathbf{A}=[a_{ij}]$ mátrix n-edrendű, akkor a megfelelő kvadratikus forma a

$$Q(\mathbf{x}) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}x_i x_j \tag{8.2.3}$$

skaláralakban is felírható.

Ha például

$$\mathbf{A} = \begin{bmatrix} 3 & -1 \\ 3 & 2 \end{bmatrix}, \tag{8.2.4}$$

akkor

$$Q(\mathbf{x}) = 3x_1^2 - x_1x_2 + 3x_2x_1 + 2x_2^2 = 3x_1^2 + 2x_1x_2 + 2x_2^2 \tag{8.2.5}$$

Az eddigiekben az $\mathbf{A}$-ra nézve nem kötöttük ki a szimmetricitást.[5] Megemlítjük azonban, hogy bármely kvadratikus forma szimmetrizálható. Pontosabban: bármely kvadratikus formához egyértelműen hozzárendelhető egy szimmetrikus mátrix. Világos ugyanis, hogy

$$\mathbf{x}^*\mathbf{A}\mathbf{x} = \mathbf{x}^* \frac{\mathbf{A}+\mathbf{A}^*}{2} \mathbf{x},$$

[5] Itt említjük meg, hogy a kézikönyvek a kvadratikus formák definiálásánál rendszerint eleve kikötik, hogy a kvadratikus forma mátrixa szimmetrikus legyen. Mi azért tértünk el a klasszikus definíciótól, mert az optimumszámítási feladatoknál gyakran szerepel a nemszimmetrikus eset is.

bármilyen kvadratikus mátrixot is jelent az **A**. S mivel az

$$\frac{\mathbf{A}+\mathbf{A}^*}{2}$$

mátrix már szimmetrikus, előbbi állításunkat ezzel igazoltuk is. Erre támaszkodva a következőkben feltesszük, hogy a kérdéses kvadratikus formát már eleve szimmetrizáltuk, vagyis azt, hogy az $\mathbf{x}^*\mathbf{A}\mathbf{x}$ formula kielégíti az $\mathbf{A}=\mathbf{A}^*$ kikötést. Az így előttünk álló szimmetrikus mátrixot nevezzük *a kvadratikus forma mátrixának* és e mátrix rangját *a kvadratikus forma rangjának*. A (8.2.5) alatt megadott kvadratikus forma mátrixa tehát nem a (8.2.4) alatti mátrix, hanem az ebből képezett

$$\frac{1}{2}\begin{bmatrix}3 & -1\\ 3 & 2\end{bmatrix}+\frac{1}{2}\begin{bmatrix}3 & -1\\ 3 & 2\end{bmatrix}^* = \begin{bmatrix}3 & 1\\ 1 & 2\end{bmatrix}$$

szimmetrikus mátrix.

Megjegyezzük, hogy az így nyert szimmetrikus mátrix a megfelelő kvadratikus formára nézve a

$$3x_1^2+x_1x_2+x_2x_1+2x_2^2 = 3x_1^2+2x_1x_2+2x_2^2$$

skaláralakot adja, s ez azonos a (8.2.5) alatt felírt kifejezéssel.

Az elmondottak alapján az is belátható, hogy a szimmetricitási kikötésre támaszkodva, a skaláralakból a kvadratikus forma mátrixa egyértelműen határozható meg. Nem ez a helyzet, ha lemondunk az **A** szimmetricitásáról, ekkor ugyanis a kérdéses mátrix számtalan különböző alakot ölthet.

Bármely $Q(\mathbf{x})$ kvadratikus formára nézve

$$Q(\mathbf{o}) = \mathbf{o}^*\mathbf{A}\mathbf{o} = 0.$$

Minden kvadratikus forma felveszi tehát a 0 értéket. Mindjárt megjegyezzük, hogy ez az $\mathbf{x}=\mathbf{o}$ ponttól különböző helyeken is bekövetkezhet. Annak megfelelően azután, hogy a további lehetséges értékek előjele hogyan alakul, két esetet különböztetünk meg:

a) *Ha a $Q(\mathbf{x})$ pozitív értéket is, meg negatívot is felvehet, akkor a $Q(\mathbf{x})$-et indefinit kvadratikus formának nevezzük.*

b) *Egyébként definit kvadratikus formával állunk szemben.*

Az utóbbiakat ismét két csoportra lehet osztani: a *pozitív definit* és a *negatív definit* kvadratikus formák csoportjába. Az első esetről akkor beszélünk, ha a kvadratikus forma *nem válik negatívvá,* a második eset pedig akkor áll fenn, ha a *kvadratikus forma nem válik pozitívvá* egyetlen $\mathbf{x}\in E_n$ pontban sem. A $Q(\mathbf{x})$ tehát pozitív definit, illetve negatív definit, ha bármely **x** mellett teljesül a

$$Q(\mathbf{x})\geqq 0,$$

illetve a

$$Q(\mathbf{x}) \leqq 0$$

reláció, ahol $\mathbf{x} \in E_n$.

Ha a $Q(\mathbf{x})$ pozitív definit kvadratikus forma csakis az $\mathbf{x}=\mathbf{o}$ pontban válik 0-vá, s így minden egyéb pontban pozitív értéket vesz fel, akkor *szigorúan pozitív definit* kvadratikus formának nevezzük. Hasonló módon definiálhatjuk a *szigorúan negatív definit* kvadratikus formákat is. A $Q(\mathbf{x})$ szigorúan negatív definit, ha minden $\mathbf{x} \neq \mathbf{0}$ pontban $Q(\mathbf{x})<0$.[6]

Bizonyos esetekben az osztályozást könnyen végrehajthatjuk. Különösen egyszerű problémával állunk szemben, ha a kvadratikus forma mátrixa diagonális. Ha ugyanis

(8.2.6) $$\mathbf{A}=\langle \lambda_1, \lambda_2, \ldots, \lambda_n \rangle,$$

akkor a definíció értelmében

(8.2.7) $$Q(\mathbf{x}) = \lambda_1 x_1^2 + \lambda_2 x_2^2 + \ldots + \lambda_n x_n^2.$$

Könnyű belátni, hogy ekkor

a) a $Q(\mathbf{x})$ *pozitív definit*, ha a diagonális elemek között nincs negatív (0 lehet),

b) a $Q(\mathbf{x})$ *szigorúan pozitív definit*, ha minden diagonális elem pozitív,

c) a $Q(\mathbf{x})$ *negatív definit*, ha a diagonális elemek között nincs pozitív (0 lehet),

d) a $Q(\mathbf{x})$ *szigorúan negatív definit*, ha minden diagonális elem negatív, és végül

e) a $(Q\mathbf{x})$ *indefinit*, ha a diagonális elemek között pozitív is található és negatív is található.

A probléma akkor is könnyen eldönthető, ha a kvadratikus forma mátrixa *Gram*-féle mátrix. Ekkor ugyanis a kérdéses kvadratikus forma biztosan pozitív definit. Ha ugyanis a $Q(\mathbf{x})=\mathbf{x}^*\mathbf{A}\mathbf{x}$ kvadratikus formánál $\mathbf{A}=\mathbf{M}^*\mathbf{M}$, akkor érvényes a

$$Q(\mathbf{x}) = \mathbf{x}^*(\mathbf{M}^*\mathbf{M})\mathbf{x} = (\mathbf{M}\mathbf{x})^*(\mathbf{M}\mathbf{x})$$

reláció, amelynek jobb oldalán az $\mathbf{Mx}$ vektornak az önmagával alkotott skaláris szorzata áll. Ezek szerint bármely $\mathbf{x}$ vektor mellett

$$Q(\mathbf{x}) \geqq 0,$$

azaz a $Q(\mathbf{x})$ valóban pozitív definit. Könnyen belátható, hogy ebben az esetben a szigorúan pozitív definitség akkor és csakis akkor áll fenn, ha az $\mathbf{M}$ oszlopvektorai lineárisan függetlenek.[7]

[6] Megemlítjük, hogy pozitív definit, illetve a negatív definit kifejezés helyett szokták használni a pozitív szemidefinit illetve a negatív szemidefinit kifejezést is, ugyenekkor a szigorúan negatív definit kifejezéseket egyszerűen a pozitív definit, illetve negatív definit kifejezésekkel helyettesítik.

[7] A bizonyítást az olvasóra bízzuk.

Az általános esetben a definitség problémája nem olyan egyszerű, mint diagonális mátrixoknál és a *Gram*-féle mátrixoknál. Ezt a problémát éppen ezért a következő alfejezetben fogjuk részletesebben megvizsgálni. A vizsgálatot arra a szerencsés körülményre fogjuk alapozni, hogy bármely kvadratikus formát olyan alakra lehet transzformálni, amelyben a kvadratikus forma mátrixa diagonális alakot ölt. Ezzel kapcsolatban jegyezzük meg, hogy a kvadratikus forma mátrixa természetesen éppen úgy transzformálódik, mint a megfelelő bilineáris forma mátrixa. Ha tehát egy kvadratikus forma mátrixa az adott bázisban $\mathbf{A}$, és innen áttérünk a $\mathbf{C}$ átmeneti mátrix meghatározta bázisba, akkor ugyanennek a kvadratikus formának az új bázisra vonatkozó mátrixát a

$$\mathbf{C}^*\mathbf{A}\mathbf{C}$$

szorzat szolgáltatja. A későbbiekből ki fog tűnni, hogy e transzformációval szemben nemcsak a kvadratikus forma mátrixának a rangja marad invariáns, hanem a definitséggel, illetve az indefinitséggel kapcsolatos tulajdonságok is.

Egyébként a kvadratikus formák definit, illetve indefinit voltával kapcsolatos elnevezéseket szoktuk alkalmazni a kvadratikus forma mátrixára is. Ilyen értelemben beszélhetünk definit, illetve indefinit mátrixokról, továbbá pozitív definit, illetve negatív definit mátrixokról is, és így tovább.

Érdeklődésre tarthat számot az a kérdés, milyen $\mathbf{x}$ pontokban válik nullává a pozitív definit $\mathbf{x}^*\mathbf{A}\mathbf{x}$ kvadratikus forma.

1. *tétel: Ha az* $\mathbf{x}^*\mathbf{A}\mathbf{x}$ *kvadratikus forma pozitív definit, akkor az* $\mathbf{x}^*\mathbf{A}\mathbf{x}=0$ *egyenlőség akkor és csakis akkor teljesül, ha* $\mathbf{A}\mathbf{x}=\mathbf{o}$.

Bizonyítás: Mivel az $\mathbf{A}\mathbf{x}=\mathbf{o}$ egyenlőséget kielégítő pontok halmaza nem más, mint az $\mathbf{A}$ sorvektorterének ortogonális komplementere, azért a tételben foglalt állítás úgy is megfogalmazható, hogy az $\mathbf{x}^*\mathbf{A}\mathbf{x}=0$ egyenlőséget kielégítő pontok halmaza megegyezik az $\mathbf{A}$ sorvektorterének ortogonális komplementerével.[8]

S most nézzük a bizonyítást! Az állítás első része triviális, hiszen $\mathbf{A}\mathbf{x}=\mathbf{o}$ esetén nyilván igaz, hogy $\mathbf{x}^*\mathbf{A}\mathbf{x}=0$. A második rész bizonyítására tegyük fel, hogy valamely $\mathbf{s}$ vektor esetén $\mathbf{s}^*\mathbf{A}\mathbf{s}=0$. Azt kell megmutatnunk, hogy ebből már következik az $\mathbf{A}\mathbf{s}=\mathbf{o}$ egyenlőség is. Mivel pozitív definit kvadratikus formáról van szó, bármely lehetséges $\mathbf{y}$ vektor és bármilyen t skalár mellett teljesülni kell az

$$(\mathbf{y}+t\mathbf{s})^*\mathbf{A}(\mathbf{y}+t\mathbf{s}) \geqq 0,$$

azaz az ebből egyszerű átalakítás révén nyerhető

$$\mathbf{y}^*\mathbf{A}\mathbf{y}+2t\mathbf{y}^*\mathbf{A}\mathbf{s}+t^2\mathbf{s}^*\mathbf{A}\mathbf{s} \geqq 0$$

[8] Ez utóbbi persze az $\mathbf{A}$ szimmetricitása miatt azonos az $\mathbf{A}$ oszlopvektor-terének ortogonális komplementerével is.

relációnak, amely az $\mathbf{s}^*\mathbf{A}\mathbf{s}=0$ feltevés miatt végül is az

$$\mathbf{y}^*\mathbf{A}\mathbf{y}+2t\mathbf{y}^*\mathbf{A}\mathbf{s} \geqq 0$$

egyenlőtlenségre redukálódik. Ez azonban minden $\mathbf{y}$ és minden t mellett valóban csak akkor állhat fenn, ha $\mathbf{A}\mathbf{s}=\mathbf{o}$.

Ezzel tételünk bizonyítását be is fejeztük.

Hasonló módon bizonyítható a

2. *tétel: Ha az* $\mathbf{x}^*\mathbf{A}\mathbf{x}$ *kvadratikus forma negatív definit, akkor az* $\mathbf{x}^*\mathbf{A}\mathbf{x}=0$ *egyenlőség akkor és csakis akkor teljesülhet, ha* $\mathbf{A}\mathbf{x}=\mathbf{o}$.

Az 1. és a 2. tétel egyenes következménye az alábbi állítás.

3. *tétel: Az* $\mathbf{x}^*\mathbf{A}\mathbf{x}$ *definit kvadratikus forma akkor és csakis akkor szigorúan definit, ha az* $\mathbf{A}$ *nemszinguláris mátrix.*

Erre az állításra a következő alfejezetben még vissza fogunk térni.

Most néhány megjegyzést teszünk a skaláris szorzat fogalmára nézve. Emlékeztetünk arra, hogy a 6.1 alfejezet értelmében az $(\mathbf{a}, \mathbf{b})$ szimbólummal jelölt skalár akkor tekinthető az $\mathbf{a}$ és a $\mathbf{b}$ vektor skaláris szorzatának, ha eleget tesz az alábbi előírásoknak:

a) $(\mathbf{a}, \mathbf{b}) = (\mathbf{b}, \mathbf{a})$

b) $(\lambda\mathbf{a}, \mathbf{b}) = \lambda(\mathbf{a}, \mathbf{b})$

c) $(\mathbf{a}_1+\mathbf{a}_2, \mathbf{b}) = (\mathbf{a}_1, \mathbf{b})+(\mathbf{a}_2, \mathbf{b})$

d) $(\mathbf{a}, \mathbf{a}) \geqq 0$

e) $(\mathbf{a}, \mathbf{a}) = 0 \Leftrightarrow \mathbf{a}=\mathbf{o}$.

Könnyen belátható, hogy az

(8.2.8) $$(\mathbf{a}, \mathbf{b}) = \mathbf{a}^*\mathbf{A}\mathbf{b}$$

formulával értelmezett bilineáris forma, amelyben az $\mathbf{A}$ *szigorúan pozitív definit szimmetrikus mátrix*, kielégíti a fenti kikötéseket. Világos ugyanis, hogy ez esetben

a) $\mathbf{a}^*\mathbf{A}\mathbf{b} = \mathbf{b}^*\mathbf{A}\mathbf{a}$,

b) $(\lambda\mathbf{a})^*\mathbf{A}\mathbf{b} = \lambda(\mathbf{a}^*\mathbf{A}\mathbf{b})$,

c) $(\mathbf{a}_1+\mathbf{a}_2)^*\mathbf{A}\mathbf{b} = \mathbf{a}_1^*\mathbf{A}\mathbf{b}+\mathbf{a}_2^*\mathbf{A}\mathbf{b}$,

d) $\mathbf{a}^*\mathbf{A}\mathbf{a} \geqq 0$ és

e) az $\mathbf{A}$ szigorú definitsége miatt az $\mathbf{a}^*\mathbf{A}\mathbf{a}=0$ egyenlőség akkor és csakis akkor teljesülhet, ha $\mathbf{a}=\mathbf{o}$.

A (8.2.8) alatt megadott formula tehát valóban az **a** és **b** skaláris szorzatának tekinthető. Ez a formula speciális esetként a skaláris szorzatnak az 1.2 alfejezet szerinti definícióját is magában foglalja. Ha ugyanis $\mathbf{A}=\mathbf{E}$, akkor a (8.2.8) az

$$(\mathbf{a}, \mathbf{b}) = \mathbf{a}^*\mathbf{b}$$

képletre redukálódik. Annak ellenére azonban, hogy a skaláris szorzat fogalma az adott módon általánosítható, a továbbiakban mégis megmaradunk a skaláris szorzat 1. fejezetbeli definíciója mellett. Ezen némileg csak a komplex számtestre vonatkozó tárgyalásainkban fogunk változtatni.

Dr. Halmai Erzsébet

LINEÁRIS ALGEBRA

Tankönyvkiadó, Budapest, 1979

Krekó Béla

Lineáris algebra

Közgazdasági és Jogi Könyvkiadó
Budapest 1976