MARX KÁROLY KÖZGAZDASÁGTUDOMÁNYI EGYETEM

Dr. Halmai Erzsébet — Dr. Krekó Béla

LINEÁRIS ALGEBRA

KÉZIRAT
17. változatlan kiadás

TANKÖNYVKIADÓ, BUDAPEST, 1989

1. Mátrixokról

1.1. Alapfogalmak

A gazdasági élet információs és ellenőrzési rendszerében igen fontos szerepet játszik az adatok gyors áttekinthetősége, egyszerü kezelhetősége és megbizható értékelése. E célkitüzések elérésében messzemenő segitséget jelent a szakember számára a matematikai módszerek alkalmazása, amelyek néha már az adatok célszerü elrendezésével, azok egyszerü matematikai eszközökkel való kezelésével célravezetnek, máskor pedig csak komoly matematikai felkészültséget igénylő eljárások utján adnak felvilágositást.

Feladatunk tehát azoknak az alapfogalmaknak az elsajátitása, amelyek birtokában gazdasági problémák megoldását elősegitő matematikai módszerekkel ismerkedhetünk meg.

Ilyen alapfogalom - a lineáris algebra tárgykörön belül - a mátrix fogalma. A mátrix tulajdonképpen egy téglalap alaku számtáblázatot jelent. Pontosabb megfogalmazásban:
Az

$$\begin{bmatrix} a_{11} & \dots & a_{1j} & \dots & a_{1n} \\ \vdots & & & & \\ a_{i1} & \dots & a_{ij} & \dots & a_{in} \\ \vdots & & & & \\ a_{m1} & \dots & a_{mj} & \dots & a_{mn} \end{bmatrix}$$

alaku számtáblázatot m.n. tipusu mátrixnak nevezzük. Az a_{ij} szimbólum a mátrix i-edik sorának j-edik elemét, illetve j-edik oszlopának i-edik elemét jelöli. Ezért az első index az un. sorindex, a második pedig az un. oszlopindex. (Az a_{ij}-t a mátrix általános elemének szoktuk hivni.) A szögletes zárójel az elemek összetartozására utal.

A mátrix jelölésére sokszor csak szimbólumot irunk az elemek részletezése helyett:

Pl.

$$\underline{A} = \left[a_{ij}\right] \qquad \begin{matrix} i = 1, 2, \dots, m \\ j = 1, 2, \dots, n\,. \end{matrix}$$

Jelölhetjük tehát a mátrixokat egyetlen betüvel is. Erre a célra nyomtatásban a "félkövéren szedett" latin nagybetüket, kézirásban pedig az aláhuzott nagybetüket szoktuk felhasználni. Igy: $\underline{A}$, $\underline{B}$, $\underline{C}$ stb., esetleg feltüntetjük a mátrix sorainak és oszlopainak számát:

$$\underset{(m\ .\ n)}{\underline{A}}\ ,$$

amit igy olvasunk: "m . n tipusu mátrix."

A gazdasági tevékenységek vizsgálatát sok esetben megkönnyithetjük, ha a rendelkezésre álló adatokat egy-egy ilyen táblázatba, mátrixba foglaljuk össze. Pl. egy kereskedelmi vállalat évi forgalmáról jól áttekinthető képet nyerhetünk, ha a forgalmi adatokból olyan táblázatot készitünk, amely az oszlopok szerint havonkénti bontásban, a sorok szerint pedig cikkcsoportonkénti bontásban mutatja az áruforgalom alakulását. Igy, ha a vállalat 30 cikkcsoportot forgalmaz, akkor egy évi tevékenységét egy 30.12-es tipusu mátrixszal reprezentálhatja, amelyben az a_{23} elem a második cikkcsoport március havi forgalmát mutatja.

A statisztikai vizsgálatok eredménye is sokszor egy-egy mátrixban ölt testet.

<u>Ha egy mátrixban felcseréljük a sorok és oszlopok szerepét, akkor az igy nyert uj mátrixot az eredeti transzponáltjának nevezzük.</u>
<u>Az $\underline{A}$ mátrix transzponáltját az $\underline{A}^*$ szimbólummal jelöljük.</u>

Ha történetesen az előbbi kereskedelmi vállalatnál a táblázatot ugy készitjük, hogy az oszlopok cikkcsoportonkénti bontásban, a sorok pedig havonkénti bontásban mutatják az áruforgalmat, akkor a vállalat tevékenységét jellemző mátrix 12.30-as tipusu, mert a sorok és oszlopok szerepet cseréltek.

Legyen $\underline{A}$ egy 2.3-as tipusu mátrix:

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

akkor transzponáltja egy 3.2-es tipusu mátrix:

$$\underline{A}^{*} = \begin{bmatrix} 1 & 2 \\ 4 & 5 \\ 3 & 0 \end{bmatrix}.$$

Általában egy m.n tipusu mátrix transzponáltja n.m tipusu mátrix, miközben az i-edik sor j-edik eleme a j-edik sor i-edik elemévé válik. Ennek megfelelően a transzponálás folyamatát jól érzékelteti az

$$\left[a_{ij}\right]^{*} = \left[a_{ji}\right]$$

összefüggés.

<u>Az 1.1 tipusu, egyetlen számból álló mátrixot skalárnak nevezzük.</u>

Ebből látható, hogy a mátrix-fogalom felfogható a közönséges értelemben vett számfogalom általánositásaként is. A gazdasági gyakorlatban előadódó mátrixok elemei rendszerint racionális számok. A mi megállapitásaink azonban minden olyan esetre érvényesek lesznek, amikor ezek az elemek valós számok. Azt a kifejezést is használhatjuk, hogy mi a valós számok halmazán értelmezett mátrixokkal foglalkozunk. Ezért a skalár, számunkra mindig valamilyen valós számot jelent. A skalárokkal kapcsolatban a mátrix - irásmódot nem alkalmazzuk. Az ötödik természetes számot tehát továbbra is az 5, nem pedig az $[5]$ szimbólummal fogjuk jelölni. Ez is arra utal, hogy a skalároknak kitüntetett szerepük van a mátrixok között.

<u>Az egyetlen oszlopból álló mátrixot, (vagyis az m.1 tipusu mátrixot) oszlopmátrixnak, vagy oszlopvektornak, az egyetlen sorból álló mátrixot (vagyis az 1.n tipusu mátrixot) sormátrixnak, vagy sorvektornak szoktuk nevezni. Az oszlopvektort és sorvektort közös néven vektornak hivjuk.</u>

A vektorok jelölésére különben külön szimbólumot is szoktunk alkalmazni, nyomtatásban félkövéren szedett latin kisbetüt, kézirásban (ill. irógépen történő irás alkalmával) pedig aláhuzott latin kisbetüt.

A sorvektor jelét ugy különböztetjük meg az oszlopvektorétól, hogy a sorvektor szimbólumát ellátjuk a transzponálás jelével is. Igy az $\underline{a}$ szimbólum mindig oszlopvektort, az $\underline{a}^{*}$ pedig mindig sorvektort jelent.

Ha például

$$\underline{a} = \begin{bmatrix} 5 \\ -4 \\ 1 \end{bmatrix}, \text{ akkor } \quad \underline{a}^{*} = [5, \quad -4, \quad 1].$$

A jobb helykihasználás kedvéért gyakran alkalmazzuk az

$$\underline{a} = \begin{bmatrix} 5, & -4, & 1 \end{bmatrix}^*$$

irásmódot is, ami az $\begin{bmatrix} 5, & -4, & 1 \end{bmatrix}$ sorvektor transzponáltját az $\underline{a}$ oszlopvektort szimbolizálja. A felirt vektor háromelemü vektor. A vektor elemeit komponenseknek is szoktuk nevezni. A sorvektorban a komponenseket vesszővel választjuk el. (A mátrixoknál az elemek közé nem teszünk vesszőt!)

A gyakorlatban sokszor előforduló speciális vektorokra külön elnevezéseket vezetünk be.

A nullvektor vagy zérusvektor olyan vektor, amelynek mindegyik komponense zérus. A nullvektor szimbóluma: $\underline{o}$ (aláhuzott kis o betü!)

Pl. a $\underline{o} = \begin{bmatrix} o, & o, & o, & o, \end{bmatrix}^*$ vektor egy 4 elemü nullvektor, oszlopvektor.

Az egységvektor olyan vektor, amelynek az egyik komponense 1, a többi nulla. Az egységvektor jelölésére az $\underline{e}_i$ szimbólumot használjuk, ahol az i index azt mutatja, hogy az 1 hányadik komponense a vektornak.

Példaként felirjuk a 3 elemü egységvektorokat oszlopvektor alakban:

$$\underline{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}; \quad \underline{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}; \quad \underline{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},$$

és a nekik megfelelő sorvektorokat:

$$\underline{e}_1^* = \begin{bmatrix} 1, & 0, & 0 \end{bmatrix}; \quad \underline{e}_2^* = \begin{bmatrix} 0, & 1, & 0 \end{bmatrix}; \quad \underline{e}_3^* = \begin{bmatrix} 0, & 0, & 1 \end{bmatrix}.$$

Az összegező vektor olyan vektor, amelynek minden komponense 1. Ezt a vektort az $\underline{1}$ szimbólummal jelöljük.

Példaként felirjuk a 3 elemü összegező vektort:

$$\underline{1} = \begin{bmatrix} 1, & 1, & 1 \end{bmatrix}^*.$$

A gyakorlatban sokszor előforduló speciális mátrixokat is külön névvel, szimbólummal látjuk el.

A nullmátrix vagy zérusmátrix olyan mátrix, amelynek minden eleme zérus. Az ilyen mátrixot a $\underline{0}$ (aláhuzott nagy 0 betü) szimbólummal jelöljük.

Ezek szerint a

$$\underline{0} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

mátrix egy 2.3 tipusu nullmátrix.

Az $n.n$ tipusu mátrixot kvadratikus mátrixnak nevezzük. Az n jelenti a kvadratikus mátrix rendjét. (A nem kvadratikus mátrixoknál nem beszélhetünk rendről!)

Az

$$\underline{A} = \begin{bmatrix} 2 & 1 \\ 3 & 5 \end{bmatrix} \quad \text{és} \quad \underline{B} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

mátrixok másodrendü kvadratikus mátrixok, ahol $\underline{B}$ egyben nullmátrix is. A kvadratikus mátrixoknál az a_{ii} elemeket diagonális elemeknek nevezzük, ezek alkotják az un. főátlót.

A diagonális mátrix olyan kvadratikus mátrix, amelynek minden főátlón kivüli eleme zérus. A diagonális mátrixok jelölésére az $\langle a_{11}, a_{22}, \ldots, a_{nn} \rangle$ szimbólumot is használhatjuk.

Diagonális mátrixok pl. a következők:

$$\underline{A} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 4 \end{bmatrix}; \quad \underline{B} = \begin{bmatrix} 5 & 0 \\ 0 & 0 \end{bmatrix}; \quad \underline{C} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

ahol $\underline{A}$ és $\underline{C}$ harmadrendü diagonális mátrix, - $\underline{C}$ egyben nullmátrix is -, $\underline{B}$ pedig másodrendü diagonális mátrix.

A külön szimbólummal jelölve:

$$\underline{A} = \langle 2, \ -3, \ 4 \rangle; \quad \underline{B} = \langle 5, \ 0 \rangle; \quad \underline{C} = \langle 0, \ 0, \ 0 \rangle.$$

Az egységmátrix olyan diagonális mátrix, amelynek diagonális elemei mind 1-gyel egyenlők.
Az egységmátrixokat általánosan az $\underline{E}$ szimbólummal jelöljük.

Harmadrendü egységmátrix pl. a következő:

$$\underline{E} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \langle 1, \quad 1, \quad 1 \rangle .$$

Néha az $\underline{E}$ szimbólumhoz egy indexet is csatolunk annak feltüntetésére, hogy hányadrendü egységmátrixról van szó. Az n-ed rendü egységmátrixot tehát $\underline{E}_n$ jelöli.

A háromszögmátrix (vagy trianguláris mátrix) olyan kvadratikus mátrix, amelynek vagy a főátló feletti vagy a főátló alatti elemei mind nullák. Az első esetben alsó háromszögmátrixról, a másodikban pedig felső háromszögmátrixról beszélünk.

Tekintsük az alábbi példákat:

$$\underline{A} = \begin{bmatrix} 1 & 3 & 5 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix}, \quad \underline{B} = \begin{bmatrix} 0 & 0 & 0 \\ 3 & 5 & 0 \\ 0 & 15 & 10 \end{bmatrix}, \quad \underline{C} = \begin{bmatrix} 0 & 2 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix},$$

ahol $\underline{A}$ és $\underline{C}$ mátrixok felső háromszögmátrixok, a $\underline{B}$ pedig alsó háromszögmátrix. Természetesen a diagonális mátrixok egyben trianguláris mátrixok is.

A gyakorlatban előforduló számtáblázatok, mátrixok általában nagyméretüek. Az ilyen mátrixokat sokszor kisebb részekre, un. blokkokra (vagy minormatrixokra) bontjuk vizszintes és függőleges egyenesek segitségével.

A mátrixoknak blokkokra való bontását particionálásnak nevezzük. A blokkokból felépitett mátrix elemei is mátrixok. Erre utal a hipermátrix elnevezés is.

Tekintsünk egy 5.5-ös mátrixot, amit a következőképpen bontunk blokkokra:

$$\underline{A} = \left[\begin{array}{ccc|cc} 5 & 4 & -3 & 5 & 0 \\ 3 & 0 & -6 & 2 & 1 \\ 0 & -4 & 5 & 8 & 0 \\ \hline 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{array}\right]$$

Particionálás után mátrixunk uj alakja:

$$\underline{A} = \begin{bmatrix} \underline{A}_{11} & \underline{A}_{12} \\ \underline{A}_{21} & \underline{A}_{22} \end{bmatrix},$$

ahol a mátrix elemei a blokkok, amelyek rendre

$$\underline{A}_{11} = \begin{bmatrix} 5 & 4 & -3 \\ 3 & 0 & -6 \\ 0 & -4 & 5 \end{bmatrix}$$

egy 3.3 tipusu mátrix,

$$\underline{A}_{12} = \begin{bmatrix} 5 & 0 \\ 2 & 1 \\ 8 & 0 \end{bmatrix}$$

egy 3.2 tipusu mátrix,

$$\underline{A}_{21} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

egy 2.3 tipusu mátrix, és

$$\underline{A}_{22} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

egy 2.2 tipusu mátrix.

Ez az irásmód sok esetben megkönnyitheti a tárgyalást.

A particionálásnak egyik leggyakrabban alkalmazott módja a mátrixoknak oszlopvektorokra vagy sorvektorokra való bontása. Gondoljunk pl. a már emlitett kereskedelmi vállalat áruforgalmát reprezentáló 30.12-es mátrixra, ahol az oszlopokra való bontás - lényegében - a vállalat január, február, ..., december havi forgalmát mutatja cikkcsoportonként; a sorokra való bontás pedig az egyes cikkcsoportokban lebonyolitott forgalmat adja meg havi részletezéssel. Mindkét felbontásra szükség van az adatok elemzésénél, értékelésénél, esetleg a jövőre vonatkozó tervezésnél.

Ha az m.n tipusu $\underline{A}$ mátrix oszlopvektorait rendre az $\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_n$ szimbólumokkal jelöljük, akkor az

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

felbontáshoz jutunk. Az $\underline{a}_j$ oszlopvektorok m-elemü vektorok. Ha pedig az m.n tipusu $\underline{A}$ mátrix sorvektorait rendre a $\underline{d}_1^*, \underline{d}_2^*, \ldots, \underline{d}_m^*$ szimbólumokkal jelöljük, akkor

$$\underline{A} = \begin{bmatrix} \underline{d}_1^* \\ \underline{d}_2^* \\ \cdot \\ \cdot \\ \cdot \\ \underline{d}_m^* \end{bmatrix}$$

felbontáshoz jutunk. A $\underline{d}_i^*$ sorvektorok n elemü vektorok.

Igy pl. minden n-edrendü egységmátrix n oszlopvektorra bontható, amelyek mindegyike egységvektor:

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

vagy n sorvektorra, amelyek ugyancsak egységvektorok:

$$\underline{E}_n = \begin{bmatrix} \underline{e}_1^* \\ \underline{e}_2^* \\ \cdot \\ \cdot \\ \underline{e}_n^* \end{bmatrix}$$

Az egységmátrixoknak oszlopvektorokra és sorvektorokra való particionálásánál a megfelelő blokkok egymásnak transzponáltjai.

1.2. Nagyságrendi relációk és műveleti szabályok

Mivel a mátrix fogalom felfogható a közönséges értelemben vett számfogalom általánositásaként is, természetes, hogy vele kapcsolatban nagyságrendi relációkról és müveleti szabályokról beszéljünk.

Két azonos tipusu mátrix között az alábbi relációk értelmezhetők:

egyenlő	$(=)$
nem-egyenlő	$(\neq)$
nagyobb	$(>)$
kisebb	$(<)$
nem-nagyobb	$(\leqq)$
nem-kisebb	$(\geqq)$

A fenti relációk valamelyike akkor teljesül a vizsgált mátrixokra, ha a kérdéses reláció elemről elemre érvényes.

Tekintsük az alábbi mátrixokat:

$$\underline{A} = \begin{bmatrix} 1 & -2 \\ 0 & 5 \end{bmatrix}, \quad \underline{B} = \begin{bmatrix} 1 & -2 \\ 0 & 5 \end{bmatrix}, \quad \underline{C} = \begin{bmatrix} 3 & -2 \\ 0 & 5 \end{bmatrix} \quad \text{és} \quad \underline{D} = \begin{bmatrix} 3 & 7 \\ 5 & 8 \end{bmatrix}.$$

Vizsgáljuk meg, hogy a fenti relációk közül melyek érvényesek az $\underline{A}$, $\underline{B}$, $\underline{C}$ és $\underline{D}$ mátrixokra. Könnyen belátható az alábbi állitások helyessége:

$\underline{A} = \underline{B}$ (Megfelelő elemeik ugyanis egyenlők.)

$\underline{A} \neq \underline{C}$ (Az egyenlőség ugyanis az első sor első elemére nem áll fenn.)

$\underline{D} > \underline{A}$ ($\underline{D}$ minden eleme nagyobb $\underline{A}$ megfelelő eleménél.)

$\underline{A} < \underline{D}$ ($\underline{A}$ elemei ugyanis mind kisebbek $\underline{D}$ megfelelő elemeinél.)

$\underline{A} \leqq \underline{C}$ (Az $\underline{A}$ első sorának első eleme kisebb ugyanis $\underline{C}$ megfelelő eleménél.)

$\underline{C} \geqq \underline{A}$ (A $\underline{C}$ első sorának első eleme nagyobb mint $\underline{A}$ megfelelő eleme.)

Mivel a vektorok speciális mátrixok az előbbi megállapitások vektorokra is érvényesek. Igy pl. az

$$\underline{x} \geqq \underline{0}$$

reláció olyan vektort értelmez, amelynek nincs negativ komponense.

Megjegyzés: A skalár-aritmetikában értelmezett nagyságrendi relációkkal ellentétben itt nem áll fenn, hogy két azonos tipusu mátrix között a " > ", " = ", " < " szimbólumok meghatározta relációk egyike, de csakis egyike fennáll.

Például, ha

$$\underline{A} = \begin{bmatrix} 3 & 1 \\ 4 & -2 \end{bmatrix} \quad \text{és} \quad \underline{B} = \begin{bmatrix} -2 & 5 \\ 4 & 0 \end{bmatrix},$$

akkor a fenti relációk egyike sem érvényes a két mátrixra.
Ezek után rátérünk a müveletek tárgyalására.

1.2.1. Összeadás

A közönséges számokra értelmezett müveletekhez hasonlóan először az összeadást vezetjük be.

Az összeadást csak azonos tipusu mátrixokra értelmezzük. Az adott m.n tipusu $\underline{A} = [a_{ij}]$ és az ugyanilyen tipusu $\underline{B} = [b_{ij}]$ mátrixok összegén azt az $\underline{A} + \underline{B}$ szimbólummal jelölt m.n tipusu mátrixot értjük, amelynek egyes elemeit ugy nyerjük, hogy $\underline{A}$ és $\underline{B}$ megfelelő elemeit összeadjuk.

Jelekben:

$$[a_{ij}] + [b_{ij}] = [a_{ij} + b_{ij}].$$

Például, ha

$$\underline{A} = \begin{bmatrix} 1 & 2 \\ 0 & -3 \end{bmatrix} \quad \text{és} \quad \underline{B} = \begin{bmatrix} 5 & -1 \\ -2 & 4 \end{bmatrix},$$

akkor

$$\underline{A} + \underline{B} = \begin{bmatrix} 6 & 1 \\ -2 & 1 \end{bmatrix}.$$

A meghatározásból következik, hogy a mátrixok összeadását a közönséges értelemben vett számok az un. skalárok összeadására vezettük vissza. Ebből következik, hogy

a) a mátrixok összegére is érvényes a "skalár-aritmetikai összeadásra" jellemző kommutativitás (tagok felcserélhetősége):

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

és

asszociativitás (tagok csoportosithatósága): $(\underline{A} + \underline{B}) + \underline{C} = \underline{A} + (\underline{B} + \underline{C})$, aminek alapján a mátrixok összeadását több, de csak véges sok tagra is kiterjeszthetjük;

b) Mivel skalár transzponáltja önmagával egyenlő, ezért összeg transzponáltja egyenlő a tagok transzponáltjának összegével:

$$(\underline{A} + \underline{B})^* = \underline{A}^* + \underline{B}^*$$

Az elmondottakra tekintsük a következő példákat:

$$\underline{A} = \begin{bmatrix} 1 & 2 \\ 0 & -3 \end{bmatrix}, \quad \underline{B} = \begin{bmatrix} 5 & -1 \\ -2 & 4 \end{bmatrix}, \quad \underline{C} = \begin{bmatrix} -4 & 5 \\ 0 & -1 \end{bmatrix}$$

mátrixokra az összeadás definiciója szerint

$$\underline{A} + \underline{B} = \begin{bmatrix} 6 & 1 \\ -2 & 1 \end{bmatrix} \quad \text{és} \quad \underline{B} + \underline{A} = \begin{bmatrix} 6 & 1 \\ -2 & 1 \end{bmatrix}.$$

Valóban igaz tehát, hogy

$$\underline{A} + \underline{B} = \underline{B} + \underline{A}.$$

Hasonló módon nyerjük, hogy

$$(\underline{A} + \underline{B}) + \underline{C} = \begin{bmatrix} 6 & 1 \\ -2 & 1 \end{bmatrix} + \begin{bmatrix} -4 & 5 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 2 & 6 \\ -2 & 0 \end{bmatrix}$$

és

$$\underline{A} + (\underline{B} + \underline{C}) = \begin{bmatrix} 1 & 2 \\ 0 & -3 \end{bmatrix} + \begin{bmatrix} 1 & 4 \\ -2 & 3 \end{bmatrix} = \begin{bmatrix} 2 & 6 \\ -2 & 0 \end{bmatrix}.$$

Érvényes tehát az $(\underline{A} + \underline{B}) + \underline{C} = \underline{A} + (\underline{B} + \underline{C})$ összefüggés is.
Végül pedig, ha

$$(\underline{A} + \underline{B}) = \begin{bmatrix} 6 & 1 \\ -2 & 1 \end{bmatrix}, \text{ akkor } (\underline{A} + \underline{B})^* = \begin{bmatrix} 6 & -2 \\ 1 & 1 \end{bmatrix}$$

és

$$\underline{A}^* = \begin{bmatrix} 1 & 0 \\ 2 & -3 \end{bmatrix}, \quad \underline{B}^* = \begin{bmatrix} 5 & -2 \\ -1 & 4 \end{bmatrix}, \text{ igy } \underline{A}^* + \underline{B}^* = \begin{bmatrix} 6 & -2 \\ 1 & 1 \end{bmatrix},$$

amiből következik az

$$(\underline{A} + \underline{B})^* = \underline{A}^* + \underline{B}^*$$

azonosság helyessége.

1.2.2. Szorzás

Az összeadás után értelmezzük a szorzás müveletét. Mátrixok összeadását csak azonos tipusu mátrixokra definiáltuk. A szorzás - bár itt is vannak megkötések - szélesebb körben végezhető el. A lehetséges eseteket tekintsük külön-külön. Elsőnek a skalárral való szorzást vegyük:

Adott $\underline{A} = [a_{ij}]$ mátrixnak egy λ skalárral alkotott szorzatán azt a $\lambda\underline{A}$ szimbólummal jelölt mátrixot értjük, amelynek minden eleme az $\underline{A}$ megfelelő elemének λ -szorosa, azaz

$$\lambda [a_{ij}] = [\lambda a_{ij}].$$

Például:

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

A meghatározásból következik, hogy mátrixnak skalárral való szorzását közönséges értelemben vett számok szorzatára vezettük vissza. Továbbá, hogy skalárral bármilyen tipusu mátrix megszorozható.

A skalárral alkotott szorzatra érvényesek az alábbi azonosságok:

$$\lambda \cdot \underline{A} = \underline{A} \cdot \lambda \qquad \text{(kommutativitás)}$$

$$(\lambda_1 \cdot \lambda_2) \cdot \underline{A} = \lambda_1 (\lambda_2 \underline{A}) \qquad \text{(asszociativitás)}$$

$$\left.\begin{aligned} (\lambda_1 + \lambda_2) \cdot \underline{A} &= \lambda_1 \underline{A} + \lambda_2 \underline{A} \\ \lambda (\underline{A} + \underline{B}) &= \lambda \underline{A} + \lambda \underline{B} \end{aligned}\right\} \text{(disztributivitás, vagyis a szorzás és az összeadás sorrendjének felcserelése)}$$

A skalárral alkotott szorzat tulajdonságainak illusztrálására legyen ismét

$$\underline{A} = \begin{bmatrix} 1 & 2 \\ 0 & -3 \end{bmatrix} \quad \text{és} \quad \underline{B} = \begin{bmatrix} 5 & -1 \\ -2 & 4 \end{bmatrix}.$$

Legyen továbbá $\lambda_1 = 5$ és $\lambda_2 = -2,5$. Ekkor a skalárral való szorzásra adott definició értelmében

$$(\lambda_1 \ \lambda_2) \underline{A} = -12,5 \begin{bmatrix} 1 & 2 \\ 0 & -3 \end{bmatrix} = \begin{bmatrix} -12,5 & -25 \\ 0 & 37,5 \end{bmatrix}$$

és

$$\lambda_1 (\lambda_2 \underline{A}) = 5 \begin{bmatrix} -2,5 & -5 \\ 0 & 7,5 \end{bmatrix} = \begin{bmatrix} -12,5 & -25 \\ 0 & 37,5 \end{bmatrix},$$

vagyis

$$(\lambda_1 \ \lambda_2) \underline{A} = \lambda_1 (\lambda_2 \underline{A}).$$

Továbbá

$$(\lambda_1 + \lambda_2) \underline{A} = 2,5 \begin{bmatrix} 1 & 2 \\ 0 & -3 \end{bmatrix} = \begin{bmatrix} 2,5 & 5 \\ 0 & -7,5 \end{bmatrix}$$

és

$$\lambda_1 \underline{A} + \lambda_2 \underline{A} = \begin{bmatrix} 5 & 10 \\ 0 & -15 \end{bmatrix} + \begin{bmatrix} -2,5 & -5 \\ 0 & 7,5 \end{bmatrix} = \begin{bmatrix} 2,5 & 5 \\ 0 & -7,5 \end{bmatrix},$$

azaz

$$(\lambda_1 + \lambda_2)\underline{A} = \lambda_1 \underline{A} + \lambda_2 \underline{A}.$$

S végül a

$$\lambda_1 (\underline{A} + \underline{B}) = 5 \begin{bmatrix} 6 & 1 \\ -2 & 1 \end{bmatrix} = \begin{bmatrix} 30 & 5 \\ -10 & 5 \end{bmatrix}$$

és a

$$\lambda_1 \underline{A} + \lambda_1 \underline{B} = \begin{bmatrix} 5 & 10 \\ 0 & -15 \end{bmatrix} + \begin{bmatrix} 25 & -5 \\ -10 & 20 \end{bmatrix} = \begin{bmatrix} 30 & 5 \\ -10 & 5 \end{bmatrix}$$

egyenlőség azt tanusitja, hogy

$$\lambda_1 (\underline{A} + \underline{B}) = \lambda_1 \underline{A} + \lambda_1 \underline{B}.$$

Megjegyzések: a) Egy adott $\underline{A} = [a_{ij}]$ mátrixnak a $\lambda = 1$ skalárral alkotott szorzata

$$1 \,.\, \underline{A} = \underline{A},$$

vagyis az adott mátrix,

b) $\lambda = -1$ skalárral alkotott szorzata

$$(-1)\,\underline{A} = -\underline{A}$$

olyan mátrix, amelynek minden eleme ellentettje az $\underline{A}$ mátrix megfelelő elemeinek, ezért az $\underline{A}$ mátrix ellentettjének nevezzük. Az $\underline{A}$ mátrixnak a (-1) skalárral való szorzása helyett tehát az $\underline{A}$ ellentettje "$-\underline{A}$" irható.

c) Ha az $\underline{A}$ mátrixhoz hozzáadjuk ellentettjét, akkor az összeadás definiciója szerint

$$\underline{A} + (-\underline{A}) = \underline{0}.$$

d) Az $\underline{A}$ mátrixnak a $\lambda = 0$ skalárral alkotott szorzata:

$$0 \cdot \underline{A} = \underline{0}.$$

Az egyenlőség definiciója és a skalárral való szorzás értelmezése lehetővé teszi a szimmetrikus és ferdén szimmetrikus mátrix fogalmának bevezetését.

Az $\underline{A}$ mátrixot szimmetrikusnak nevezzük, ha

$$\underline{A}^* = \underline{A},$$

azaz, ha $\underline{A}$ mátrix megegyezik saját transzponáltjával.

Például: ha

$$\underline{A} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 7 \\ 3 & 7 & 6 \end{bmatrix}, \text{ akkor } \underline{A}^* = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 7 \\ 3 & 7 & 6 \end{bmatrix}$$

és igy

$$\underline{A} = \underline{A}^*.$$

Megjegyzések: a) A szimmetrikus mátrix minden esetben kvadratikus mátrix, amelyre

$$a_{ij} = a_{ji},$$

minden i és j értékre.

b) Minden diagonális mátrix szimmetrikus mátrix.

c) Kimutatható, hogy szimmetrikus mátrixok összege is szimmetrikus.

A szimmetrikus mátrix meghatározása után definiáljuk a ferdén szimmetrikus mátrixot.

A $\underline{B}$ mátrixot akkor nevezzük ferdén szimmetrikusnak, ha

$$\underline{B}^* = -\underline{B},$$

azaz, ha a $\underline{B}$ mátrix transzponáltja a $\underline{B}$ mátrix ellentettjével egyenlő.

Például, ha

$$\underline{B} = \begin{bmatrix} 0 & -1 & 2 \\ 1 & 0 & -3 \\ -2 & 3 & 0 \end{bmatrix}, \text{ amelyre } \underline{B}^* = \begin{bmatrix} 0 & 1 & -2 \\ -1 & 0 & 3 \\ 2 & -3 & 0 \end{bmatrix}$$

és

$$-\underline{B} = \begin{bmatrix} 0 & 1 & -2 \\ -1 & 0 & 3 \\ 2 & -3 & 0 \end{bmatrix}, \text{ akkor } \underline{B}^* = -\underline{B},$$

ezért a $\underline{B}$ mátrix ferdén szimmetrikus.

Megjegyzések: a) A ferdén szimmetrikus mátrix is kvadratikus mátrix, amelyre

$$a_{ij} = -a_{ji}$$

minden i és j értékre.

b) Az $a_{ji} = -a_{ji}$ egyenlőségből következik, hogy a ferdén szimmetrikus mátrixok diagonális elemei mindig zérussal egyenlők.

c) A ferdén szimmetrikus mátrixra adott meghatározásból következik, hogy

$$\underline{B}^* + \underline{B} = \underline{0}.$$

Az eddigi müveletekkel kapcsolatban bevezetjük a lineáris algebrában fontos szerepet játszó lineáris kombináció fogalmát.

Ha az adott m.n tipusu

$$\underline{A}_1, \underline{A}_2, \ldots, \underline{A}_k$$

mátrixokat rendre megszorozzuk a

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

skalárokkal és az igy nyert szorzatokat összeadjuk, akkor a

$$\lambda_1 \underline{A}_1 + \lambda_2 \underline{A}_2 + \ldots + \lambda_k \underline{A}_k$$

kifejezésnek megfelelő m.n tipusu mátrixot az adott mártixok lineáris kombinációjának nevezzük.

Igy pl. az

$$\underline{A}_1 = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}, \quad \underline{A}_2 = \begin{bmatrix} 10 & 5 \\ 3 & 7 \end{bmatrix} \text{ és } \underline{A}_3 = \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}$$

mátrixoknak a

$$\lambda_1 = 2, \quad \lambda_2 = 3 \quad \text{és} \quad \lambda_3 = 10$$

skalárakkal képzett lineáris kombinációja a

$$2 \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} + 3 \begin{bmatrix} 10 & 5 \\ 3 & 7 \end{bmatrix} + 10 \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 52 & 31 \\ 13 & 59 \end{bmatrix}$$

mátrix.

Megjegyzések: a) A definiált lineáris kombinációban a "k" értéke 1 is lehet. Ez azt jelenti, hogy speciális esetben egyetlen mátrix lineáris kombinációjáról is beszélhetünk.

b) Az azonos tipusu $\underline{A}$ és $\underline{B}$ mátrixoknak a $\lambda_1 = 1$ és $\lambda_2 = -1$ skalárokkal vett lineáris kombinációja, az

$$\underline{A} - \underline{B}$$

szimbólummal jelölt mártix felfogható ugy is, mint különbség, $\underline{A}$ és $\underline{B}$ mátrixok különbsége. Ez olyan mátrixot jelent, amelynek elemeit ugy nyerjük, hogy vesszük az $\underline{A}$ és $\underline{B}$ megfelelő elemeinek különbségét.
Szimbólikusan:

$$\left[a_{ij}\right] - \left[b_{ij}\right] = \left[a_{ij} - b_{ij}\right].$$

Például:

$$\begin{bmatrix} 2 & -1 & 0 \\ 3 & 2 & -5 \end{bmatrix} - \begin{bmatrix} 5 & 4 & -3 \\ -2 & 3 & 5 \end{bmatrix} = \begin{bmatrix} -3 & -5 & 3 \\ 5 & -1 & -10 \end{bmatrix}.$$

Az $\underline{A}_1$, $\underline{A}_2$, ..., $\underline{A}_k$ mátrixokra értelmezett lineáris kombinációk közül kiemelünk néhányat, amelyeknek a későbbiekben fontos szerepük lesz. Igy,

ha egy lineáris kombinációban a skalár szorzók között nincs negativ, akkor nem-negativ lineáris kombinációról beszélünk.

Például, a

$$2\underline{A}_1 + 3\underline{A}_2 + 0 \cdot \underline{A}_3 + 5\underline{A}_4$$

lineáris kombináció nem-negativ lineáris kombináció, mert a

$$\lambda_1 = 2, \quad \lambda_2 = 3, \quad \lambda_3 = 0 \quad \text{és} \quad \lambda_4 = 5$$

skalárok között nem szerepel negativ skalár.

Az olyan nem-negativ lineáris kombinációt pedig, amelyben a skalár szorzók összege egy, konvex lineáris kombinációnak nevezzük.

Például:

1. a $0{,}5\,\underline{A}_1 + 0{,}2\,\underline{A}_2 + 0{,}3\,\underline{A}_3$

   lineáris kombináció konvex lineáris kombináció, mert a

   $$\lambda_1 = 0{,}5, \quad \lambda_2 = 0{,}2 \quad \text{és} \quad \lambda_3 = 0{,}3$$

   skalárokra teljesül

a) a nem-negativitás: $\lambda_1, \lambda_2, \lambda_3 \geqq 0$

b) $\lambda_1 + \lambda_2 + \lambda_3 = 1$ feltétel.

2. az

$$\frac{1}{2}\underline{a}_1 + \frac{1}{6}\underline{a}_2 + \frac{1}{3}\underline{a}_3,$$

speciális mátrixokra, oszlopvektorokra felirt lineáris kombináció is konvex lineáris kombináció, mert

$$\frac{1}{2}, \frac{1}{6}, \frac{1}{3}. > 0 \quad \text{és} \qquad \frac{1}{2} + \frac{1}{6} + \frac{1}{3} = 1.$$

A skalárral való szorzás meghatározása után értelmezzük az un. skaláris szorzatot.

Az

$\underline{a} = [a_1, a_2, \dots, a_i, \dots a_n]^*$

és

$\underline{b} = [b_1, b_2, \dots b_i, \dots b_n]^*$

vektorok skaláris szorzatán az

$$a_1 b_1 + a_2 b_2 + \dots + a_i b_i + \dots + a_n b_n$$

skalárt értjük, azaz $\underline{a}$ és $\underline{b}$ vektorok szorzata olyan skalárt jelent, amely megegyezik az azonos indexü komponensek szorzatának összegével. Jelölésére az

$$\underline{a}^* \underline{b} = \alpha$$

szimbólumot használjuk.

Például, ha

$$\underline{a} = [2, \ 3, \ 5]^* \quad \text{és} \quad \underline{b} = [1, \ 0, \ 2]^*,$$

akkor

$$\underline{a}^* \underline{b} = 2 \ . \ 1 + 3 \ . \ 0 + 5 \ . \ 2 = 12.$$

A skaláris szorzattal értelmezett

$$a_1 b_1 + a_2 b_2 + \dots + a_i b_i + \dots + a_n b_n$$

skalár felirására, az alábbi rövid szimbólumot is szokták használni:

$$\sum_{i=1}^{n} a_i b_i,$$

amit igy olvasunk: "szumma i = 1-től n-ig a_i b_i, vagyis a_i b_i szorzatokat adjuk össze i = 1, 2, ... n esetén.

Megjegyzések: a) Csak olyan vektorok skaláris szorzata van értelmezve, amelyekben a komponensek száma azonos.

b) A skaláris szorzat kommutativ és disztributiv kifejezés az alábbi azonosságoknak megfelelően:

$$\underline{a}^* \underline{b} = \underline{b}^* . \underline{a} \quad \text{(kommutativitás)}$$

$$\left.\begin{aligned} &\underline{a}^* (\underline{b} + \underline{c}) = \underline{a}^* \underline{b} + \underline{a}^* \underline{c} \\ &(\underline{a}^* + \underline{b}^*) . \underline{c} = \underline{a}^* \underline{c} + \underline{b}^* \underline{c} \end{aligned}\right\} \text{(disztributivitás)}$$

Példaképpen tekintsük az alábbi vektorokat:

$$\underline{a} = [3,\ 1,\ 0,\ -2]^*, \quad \underline{b} = [5,\ -2,\ 4,\ 0]^* \quad \text{és} \quad \underline{c} = [4,\ 0,\ -2,\ 3]^*$$

A skaláris szorzatra adott definició alapján könnyen ellenőrizhetjük az alábbi egyenlőségek helyességét:

$$\underline{a}^* \underline{b} = 3.5 + 1.(-2) + 0.4 + (-2).0 = 13.$$

$$\underline{b}^* . \underline{a} = 5.3 + (-2).1 + 4.0 + 0.(-2) = 13.$$

$$\underline{a}^* (\underline{b} + \underline{c}) = 3.(5+4) + 1.(-2+0) + 0.(4-2) + (-2)(0+3) = 19.$$

$$\underline{a}^* \underline{b} + \underline{a}^* \underline{c} = [3.5 + 1.(-2) + 0.4 + (-2).0] + [3.4 + 1.0 + 0(-2) + (-2).3] = 13+6 = 19.$$

$(\underline{a}^* + \underline{b}^*) \cdot \underline{c} = (3+5) \cdot 4 + (1-2) \cdot 0 + (0+4) \cdot (-2) + (-2+0) \cdot 3 = 18.$

$$\underline{a}^* \underline{c} + \underline{b}^* \underline{c} = [3.4 + 1.0 + 0.(-2) + (-2).3] + [5.4 + (-2).0 + 4.(-2) + 0.3] = 6+12 = 18.$$

A nyert eredmények szerint az adott vektorok valóban eleget tesznek az előbb megfogalmazott müveleti szabályoknak.

Megjegyzések: a skaláris szorzat értelmezéséből következik, hogy

a) egy vektornak önmagával képzett skaláris szorzata nem lehet negativ skalár, ugyanis ha $\underline{a} = [\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_n]$, akkor $\underline{a}^* \cdot \underline{a} = \underline{a}_1^2 + \underline{a}_2^2 + \ldots + \underline{a}_n^2$.

b) Egy vektornak önmagával képzett skaláris szorzata, csak akkor lehet nulla, ha a vektor nullvektor, azaz

$$\underline{a}^* \cdot \underline{a} = 0, \quad \text{ha} \quad \underline{a} = \underline{0}.$$

c) Egy vektornak a megfelelő összegező vektorral alkotott skaláris szorzata a kérdéses vektor komponenseinek összegét adja:

$\underline{a}^* \cdot \underline{1} = a_1 + a_2 + \ldots + a_n$, vagy

$\underline{1}^* \cdot \underline{a} = a_1 + a_2 + \ldots + a_n$.

Ennek alapján világos, hogy egy n elemü összegező vektornak önmagával képzett skaláris szorzata n-et ad eredményül:

$$\underline{1}^* \underline{1} = 1 + 1 + \ldots + 1 = n.$$

d) Bármely n elemü vektornak az n-elemü nullvektorral alkotott skaláris szorzata zérus:

$$\underline{a}^* \cdot \underline{0} = 0$$

e) A skalárokra értelmezett szorzással ellentétben abból, hogy

$$\underline{a}^* \cdot \underline{b} = 0$$

nem következik, hogy a tényezők valamelyike zérus vektor, ugyanis, ha

$$\underline{a} = [3,\ 5]^* \quad \text{és} \quad \underline{b} = [-5,\ 3]^*,$$

akkor

$$\underline{a}^* \cdot \underline{b} = 3(-5) + 5 \cdot 3 = 0.$$

Az $\underline{a}^*$ $\underline{b}$ skaláris szorzat bevezetése után értelmezzük az un. <u>diadikus szorzatot</u>.

<u>Az</u>

$$\underline{a} = [a_1,\ a_2,\ \ldots,\ a_m]^*$$

<u>és</u>

$$\underline{b} = [b_1,\ b_2,\ \ldots,\ b_n]^*$$

<u>vektorok diadikus szorzatát a következő előírással definiál-juk:</u>

$$\begin{bmatrix} a_1 b_1 & a_1 b_2 & \ldots & a_1 b_n \\ a_2 b_1 & a_2 b_2 & \ldots & a_2 b_n \\ \vdots & & & \\ a_m b_1 & a_m b_2 & \ldots & a_m b_n \end{bmatrix},$$

azaz egy m elemü $\underline{a}$ vektor és egy n elemü $\underline{b}$ vektor diadikus szorzata olyan mátrixot jelent, amelynek i-edik sora

$$a_i\,[b_1,\ b_2,\ \ldots\ b_n] = [a_i b_1,\ a_i b_2,\ \ldots,\ a_i b_n]$$

vektor. A mátrix m.n tipusu, vagyis sorainak száma megegyezik az $\underline{a}$ vektor komponenseinek számával, az oszlopok száma pedig azonos a $\underline{b}$ komponenseinek számával.

Jelekben:

$$\underset{(m.1)}{\underline{a}} \cdot \underset{(1.n)}{\underline{b}^*} = \underset{(m.n)}{\underline{C}}$$

Például:

$$\begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix} \begin{bmatrix} 1, & 3, & -2, & 4 \end{bmatrix} = \begin{bmatrix} 2 & 6 & -4 & 8 \\ 0 & 0 & 0 & 0 \\ -1 & -3 & 2 & -4 \end{bmatrix}.$$

Megjegyzés: a diadikus szorzatot röviden diádnak nevezzük.

A mátrixoknak blokkokra való bontásánál emlitettük, hogy a particionálásnak egyik leggyakrabban alkalmazott módja a mátrixoknak sorvektorokra, vagy oszlopvektorokra való szétbontása. Ezt használjuk fel a mátrixoknak mátrixszal való szorzásának értelmezésénél.

Egy $p \cdot q$ tipusu mátrix és egy $q \cdot r$ tipusu mátrix szorzatát a következő előírással definiáljuk:

$$\underset{(p\,\cdot\,q)}{\underline{A}} \cdot \underset{(q\,\cdot\,r)}{\underline{B}} = \begin{bmatrix} \underline{a}_1^* \\ \underline{a}_2^* \\ \cdot \\ \cdot \\ \cdot \\ \underline{a}_p^* \end{bmatrix} \cdot \begin{bmatrix} \underline{b}_1, & \underline{b}_2, & \dots, & \underline{b}_r \end{bmatrix} =$$

$$= \begin{bmatrix} \underline{a}_1^* \underline{b}_1 & \underline{a}_1^* \underline{b}_2 \dots \underline{a}_1^* \underline{b}_r \\ \underline{a}_2^* \underline{b}_1 & \underline{a}_2^* \underline{b}_2 \dots \underline{a}_2^* \underline{b}_r \\ \cdot & \\ \cdot & \\ \cdot & \\ \underline{a}_p^* \underline{b}_1 & \underline{a}_p^* \cdot \underline{b}_2 \dots \underline{a}_p^* \cdot \underline{b}_r \end{bmatrix} = \underset{(p,\ r)}{\underline{C}},$$

azaz az $\underset{(p.q)}{\underline{A}}$ mátrixnak a $\underset{(q.r)}{\underline{B}}$ mátrixszal képzett szorzata egy olyan mátrix, amelyben a sorok száma megegyezik az első tényező sorainak számával, az oszlopok száma pedig a második tényező oszlopainak számával.

Szimbólikusan:

$$\underset{(p.q)}{\underline{A}} \cdot \underset{(q.r)}{\underline{B}} = \underset{(p.r)}{\underline{C}}.$$

A szorzat i-edik sorának j-edik eleme nem más, mint az első tényező i-edik sorvektorának és a második tényező j-edik oszlopvektorának a skaláris szorzata, vagyis

$$c_{ij} = \underline{a}_i^* \cdot \underline{b}_j .$$

Például: ha

$$\underline{A} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 3 & -2 \end{bmatrix} \quad \text{és} \quad \underline{B} = \begin{bmatrix} 1 & -3 \\ 2 & 1 \end{bmatrix},$$

akkor

$$\underline{A} = \begin{bmatrix} \underline{a}_1^* \\ \underline{a}_2^* \\ \underline{a}_3^* \end{bmatrix}, \text{ ahol } \underline{a}_1^* = \begin{bmatrix} 1, & 2 \end{bmatrix} \; \underline{a}_2^* = \begin{bmatrix} 0, & 1 \end{bmatrix} \; \underline{a}_3^* = \begin{bmatrix} 3, & -2 \end{bmatrix},$$

$$\underline{B} = \begin{bmatrix} \underline{b}_1, & \underline{b}_2 \end{bmatrix}, \text{ ahol } \underline{b}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad \underline{b}_2 = \begin{bmatrix} -3 \\ 1 \end{bmatrix},$$

ahonnan

$$\underline{a}_1^* \cdot \underline{b}_1 = 5 \qquad \underline{a}_1^* \cdot \underline{b}_2 = -1$$

$$\underline{a}_2^* \cdot \underline{b}_1 = 2 \qquad \underline{a}_2^* \cdot \underline{b}_2 = 1$$

$$\underline{a}_3^* \cdot \underline{b}_1 = -1 \qquad \underline{a}_3^* \cdot \underline{b}_2 = -11,$$

ezért

$$\underline{A} \cdot \underline{B} = \begin{bmatrix} 5 & -1 \\ 2 & 1 \\ -1 & -11 \end{bmatrix}.$$

Megjegyzések: a) Két mátrix szorzata csak abban az esetben van értelmezve, ha az első tényező oszlopvektorainak száma megegyezik a második tényező sorvektorainak számával, mert ez biztositja az

$$\underline{a}_i^* \cdot \underline{b}_j = [a_{i1}, a_{i2}, \ldots a_{iq}] \cdot \begin{bmatrix} b_{1j} \\ b_{2j} \\ \cdot \\ \cdot \\ \cdot \\ b_{qj} \end{bmatrix}$$

skaláris szorzatok létezését.

b) A mátrixok szorzata általában nem kommutativ kifejezés. Az $\underline{A} \cdot \underline{B}$ szorzat tehát a legtöbb esetben nem egyezik meg a $\underline{B} \cdot \underline{A}$ szorzattal. A kommutativitás minden további megszoritás nélkül már csak azért sem állhat fenn, mert eleve az sem biztos, hogy a tényezők felcserélésével nyert uj szorzat egyáltalában értelmezve van-e.

Pl. $\underset{(3.2)}{\underline{A}} \cdot \underset{(2.2)}{\underline{B}} = \underset{(3.2)}{\underline{C}}$, de $\underset{(2.2)}{\underline{B}}$ mátrix nem szorozható meg a $\underset{(3.2)}{\underline{A}}$ mátrixszal, mivel a $\underline{b}_i^*$ sorvektor és $\underline{a}_j$ oszlopvektor között nincs értelme a skaláris szorzatnak. Kommutativitásról ezért csak kvadratikus mátrixoknál lehet szó, de ezeknél is csak speciális esetekben. Ilyen speciális eset - amelyeknek helyessége kimutatható - például

$$\left.\begin{matrix} \underline{A} \cdot \underline{0} = \underline{0} \\ \underline{0} \cdot \underline{A} = \underline{0} \end{matrix}\right\}, \text{ ahol } \underline{A} \cdot \underline{0} = \underline{0} \cdot \underline{A}$$

és

$$\left.\begin{matrix} \underline{A} \cdot \underline{E} = \underline{A} \\ \underline{E} \cdot \underline{A} = \underline{A} \end{matrix}\right\}, \text{ ahol } \underline{A} \cdot \underline{E} = \underline{E}\,\underline{A}.$$

Egyéb speciális eseteket később még emliteni fogunk.

c) A kommutativitással ellentétben az asszociativitás és a disztributivitás a mátrixok szorzatára is érvényes. Nevezetesen:

$$(\underline{A} \cdot \underline{B})\,\underline{C} = \underline{A}\,(\underline{B} \cdot \underline{C}) \qquad \textbf{(asszociativitás)}$$

$$\left.\begin{aligned} \underline{A}\,(\underline{B} + \underline{C}) &= \underline{A}\,\underline{B} + \underline{A}\,\underline{C} \\ (\underline{A} + \underline{B}) \cdot \underline{C} &= \underline{A}\,\underline{C} + \underline{B}\,\underline{C} \end{aligned}\right\} \textbf{(disztributivitás)}$$

Példaképpen tekintsük a következő mátrixokat:

$$\underline{A} = \begin{bmatrix} 1 & 2 \\ 1 & -1 \\ 0 & 2 \end{bmatrix}, \quad \underline{B} = \begin{bmatrix} -1 & 0 \\ 1 & 2 \end{bmatrix} \quad \text{és} \quad \underline{C} = \begin{bmatrix} 0 & 3 \\ 1 & -1 \end{bmatrix}.$$

Ezek között értelmezve van az

$$\underline{A}\,\underline{B}\,; \quad \underline{A}\,\underline{C}\,; \quad \underline{B}\,\underline{C} \text{ és } \underline{C}\,\underline{B}$$

szorzat, ahol

$$\underline{A}\,\underline{B} = \begin{bmatrix} 1 & 4 \\ -2 & -2 \\ 2 & 4 \end{bmatrix}; \quad \underline{A}\,\underline{C} = \begin{bmatrix} 2 & 1 \\ -1 & 4 \\ 2 & -2 \end{bmatrix}; \quad \underline{B}\,\underline{C} = \begin{bmatrix} 0 & -3 \\ 2 & 1 \end{bmatrix} \text{ és } \underline{C}\,\underline{B} = \begin{bmatrix} 3 & 6 \\ -2 & -2 \end{bmatrix}.$$

A szorzási szabály alapján könnyen ellenőrizhetjük az alábbi egyenlőségeket:

$$(\underline{A}\,\underline{B})\,\underline{C} = \begin{bmatrix} 1 & 4 \\ -2 & -2 \\ 2 & 4 \end{bmatrix} \cdot \begin{bmatrix} 0 & 3 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 4 & -1 \\ -2 & -4 \\ 4 & 2 \end{bmatrix}.$$

$$\underline{A}\,(\underline{B}\,\underline{C}) = \begin{bmatrix} 1 & 2 \\ 1 & -1 \\ 0 & 2 \end{bmatrix} \cdot \begin{bmatrix} 0 & -3 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 4 & -1 \\ -2 & -4 \\ 4 & 2 \end{bmatrix}.$$

$$\underline{A}\,(\underline{B}+\underline{C}) = \begin{bmatrix} 1 & 2 \\ 1 & -1 \\ 0 & 2 \end{bmatrix} \cdot \begin{bmatrix} -1 & 3 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 5 \\ -3 & 2 \\ 4 & 2 \end{bmatrix}.$$

$$\underline{AB} + \underline{AC} = \begin{bmatrix} 1 & 4 \\ -2 & -2 \\ 2 & 4 \end{bmatrix} + \begin{bmatrix} 2 & 1 \\ -1 & 4 \\ 2 & -2 \end{bmatrix} = \begin{bmatrix} 3 & 5 \\ -3 & 2 \\ 4 & 2 \end{bmatrix}.$$

Ezek szerint a fenti mátrixokra nézve érvényes mind az

$$(\underline{A}\,\underline{B}) \cdot \underline{C} = \underline{A}\,(\underline{B}\,\underline{C}),$$

mind pedig az

$$\underline{A}\,(\underline{B}+\underline{C}) = \underline{A}\,\underline{B} + \underline{A}\,\underline{C}$$

összefüggés.

Ha egy szorzatban az egyik tényező skalár, akkor a szorzat független attól, hogy a skalárt hányadik tényezőként irjuk fel. Ennek megfelelően

$$\lambda\,(\underline{A}\,\underline{B}) = (\underline{A}\,\underline{B})\,\lambda = \underline{A}\,(\lambda\underline{B}) = (\lambda\,\underline{A}) \cdot \underline{B}.$$

A továbbiakban megfogalmazunk néhány állitást a mátrixok szorzására vonatkozóan:

1. Tétel: Ha a szorzat egyik tényezője nullmátrix, akkor maga a szorzat is nullmátrixot jelent - amint ezt már láttuk az emlitett példánál.

Bizonyitás: Az állitás helyessége a szorzatra adott definicióból következik, ugyanis az $\underline{a}^{*}\,\underline{b}$ skaláris szorzata nulla, ha egyik tényező $\underline{0}$.

Ez a nyilvánvaló állitás megegyezik a skalár-arimetrikából már ismert, azonos tétellel. Van azonban egy lényeges különbség is. Amig a skalároritmetikában ez a tétel megfordítható, addig a mátrixaritmetikában már nem ez a helyzet. Két skalár szorzata ugyanis csak akkor lehet zérus, ha közülük legalább az egyik zérus. Ezzel szemben két mátrix szorzata akkor is szolgáltathat nullmátrixot, amikor egyik tényezője sem nullmátrix.

Pl.

$$\begin{bmatrix} 2 & -1 \\ -4 & 2 \\ 8 & -4 \end{bmatrix} \cdot \begin{bmatrix} 1 & -2 \\ 2 & -4 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

2. Tétel: Az $\underline{A}\ \underline{x}$ szorzat - amennyiben értelmezve van - az $\underline{A}$ mátrix oszlopvektorainak azt a lineáris kombinációját jelenti, amelyben skalárokként az $\underline{x}$ vektor megfelelő komponensei szerepelnek. Ha tehát a p.q tipusu $\underline{A}$ mátrixban

$$\underline{A} = \begin{bmatrix} a_{11} & \cdots & a_{1q} \\ a_{21} & \cdots & a_{2q} \\ \vdots & & \\ a_{p1} & \cdots & a_{pq} \end{bmatrix} = [\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_q]$$

és a kérdéses $\underline{x}$ vektorra nézve

$$\underline{x} = [x_1, x_2, \ldots, x_q]^*$$

akkor

$$\underline{A}\ \underline{x} = x_1 \underline{a}_1 + x_2 \underline{a}_2 + \ldots + x_q \underline{a}_q.$$

Bizonyitás: Ennek az állitásnak a helyessége ugy látható be, hogy felirjuk az $\underline{A}\ \underline{x}$ szorzatnak megfelelő oszlopvektort, vagyis

$$\begin{bmatrix} a_{11} & a_{12} \ldots & a_{1j} \ldots & a_{1q} \\ a_{21} & a_{22} \ldots & a_{2j} \ldots & a_{2q} \\ \vdots & & & \\ a_{i1} & a_{i2} \ldots & a_{ij} \ldots & a_{iq} \\ \vdots & & & \\ a_{p1} & a_{p2} \ldots & a_{pj} \ldots & a_{pq} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_j \\ \vdots \\ x_q \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1j}x_j + \ldots + a_{1q}x_q \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2j}x_j + \ldots + a_{2q}x_q \\ \vdots \\ a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{ij}x_j + \ldots + a_{iq}x_q \\ \vdots \\ a_{p1}x_1 + a_{p2}x_2 + \ldots + a_{pj}x_j + \ldots + a_{pq}x_q \end{bmatrix},$$

ahonnan az összeadás definiciója alapján:

$$\begin{bmatrix} a_{11}x_1+a_{12}x_2+\ldots+a_{1j}x_j+\ldots+a_{1q}x_q \\ a_{21}x_1+a_{22}x_2+\ldots+a_{2j}x_j+\ldots+a_{2q}x_q \\ \cdot \\ \cdot \\ \cdot \\ a_{i1}x_1+a_{i2}x_2+\ldots+a_{ij}x_j+\ldots+a_{iq}x_q \\ \cdot \\ \cdot \\ \cdot \\ a_{p1}x_1+a_{p2}x_2+\ldots+a_{pj}x_j+\ldots+a_{pq}x_q \end{bmatrix} = \begin{bmatrix} a_{11}x_1 \\ a_{21}x_1 \\ \cdot \\ \cdot \\ \cdot \\ a_{i1}x_1 \\ \cdot \\ \cdot \\ \cdot \\ a_{p1}x_1 \end{bmatrix} + \begin{bmatrix} a_{12}x_2 \\ a_{22}x_2 \\ \cdot \\ \cdot \\ \cdot \\ a_{i2}x_2 \\ \cdot \\ \cdot \\ \cdot \\ a_{p2}x_2 \end{bmatrix} + \ldots + \begin{bmatrix} a_{1j}x_j \\ a_{2j}x_j \\ \cdot \\ \cdot \\ \cdot \\ a_{ij}x_j \\ \cdot \\ \cdot \\ \cdot \\ a_{pj}x_j \end{bmatrix} + \ldots + \begin{bmatrix} a_{1q}x_q \\ a_{2q}x_q \\ \cdot \\ \cdot \\ \cdot \\ a_{iq}x_q \\ \cdot \\ \cdot \\ \cdot \\ a_{pq}x_q \end{bmatrix}$$

Majd az egyes oszlopvektorokból az azonos skalár szorzót kiemelve (a disztributiv törvény megforditása!):

$$\begin{bmatrix} a_{11} \\ a_{21} \\ \cdot \\ \cdot \\ \cdot \\ a_{i1} \\ \cdot \\ \cdot \\ \cdot \\ a_{p1} \end{bmatrix} x_1 + \begin{bmatrix} a_{12} \\ a_{22} \\ \cdot \\ \cdot \\ \cdot \\ a_{i2} \\ \cdot \\ \cdot \\ \cdot \\ a_{p2} \end{bmatrix} x_2 + \ldots + \begin{bmatrix} a_{1j} \\ a_{2j} \\ \cdot \\ \cdot \\ \cdot \\ a_{ij} \\ \cdot \\ \cdot \\ \cdot \\ a_{pj} \end{bmatrix} x_j + \ldots + \begin{bmatrix} a_{1q} \\ a_{2q} \\ \cdot \\ \cdot \\ \cdot \\ a_{iq} \\ \cdot \\ \cdot \\ \cdot \\ a_{pq} \end{bmatrix} x_q$$

adódik.

Az oszlopvektorok jelölésére az $\underline{a}_j$ szimbólumot bevezetve valóban

$$\underline{A}\,\underline{x} = x_1\,\underline{a}_1 + x_2\,\underline{a}_2 + \ldots + x_q\,\underline{a}_q,$$

ahogy állitottuk.

Legyen például:

$$\underline{A} = \begin{bmatrix} 1 & 0 & 5 \\ 2 & -1 & 1 \\ 3 & 2 & -3 \end{bmatrix} \quad \text{és} \quad \underline{x} = \begin{bmatrix} 3 \\ 5 \\ -1 \end{bmatrix}.$$

Az $\underline{A}\ \underline{x}$ szorzat a mátrixok szorzására adott definiciónak megfelelően:

$$\underline{A}\ \underline{x} = \begin{bmatrix} -2 \\ 0 \\ 22 \end{bmatrix},$$

ha pedig a fenti tétel alapján járunk el, akkor

$$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \cdot 3 + \begin{bmatrix} 0 \\ -1 \\ 2 \end{bmatrix} \cdot 5 + \begin{bmatrix} 5 \\ 1 \\ -3 \end{bmatrix} (-1) = \begin{bmatrix} -2 \\ 0 \\ 22 \end{bmatrix}.$$

A két eredmény valóban megegyezik.

Következmény:

A fenti állítás értelmében az $\underline{A} \cdot \underline{B}$ szorzat oszlopvektorai nem mások, mint az $\underline{A}$ mátrix - az első tényező - oszlopvektorainak lineáris kombinációi, mégpedig olyan lineáris kombinációi, amelyekben skalárokként a $\underline{B}$ mátrix - második tényező - oszlopvektorainak megfelelő komponensei szerepelnek, vagyis ha

$$\underline{B} = [\underline{b}_1, \underline{b}_2, \dots \underline{b}_r]$$

akkor az $\underline{A}\ \underline{B}$ szorzatra nézve érvényes az

$$\underset{(p.q)}{\underline{A}}\ \underset{(q.r)}{\underline{B}} = \underline{A}[\underline{b}_1, \underline{b}_2, \dots, \underline{b}_r] = [\underline{A}\ \underline{b}_1, \underline{A}\ \underline{b}_2, \dots, \underline{A}\ \underline{b}_r]$$

azonosság is.
Például:

$$\underline{A} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 3 & -2 \end{bmatrix} \quad \text{és} \quad \underline{B} = \begin{bmatrix} 1 & -3 \\ 2 & 1 \end{bmatrix}$$

mátrixok szorzatában az első oszlop:

$$\begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 3 & -2 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix} \cdot 1 + \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix} \cdot 2 = \begin{bmatrix} 5 \\ 2 \\ -1 \end{bmatrix},$$

a második oszlop pedig:

$$\begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 3 & -2 \end{bmatrix} \cdot \begin{bmatrix} -3 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ -11 \end{bmatrix}$$

és igy

$$\underline{A} \cdot \underline{B} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 3 & -2 \end{bmatrix} \begin{bmatrix} 1 & -3 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 5 & -1 \\ 2 & 1 \\ -1 & -11 \end{bmatrix}.$$

Ugyanazt az eredményt kaptuk, mint a szorzás eredeti definiciója alapján történt szorzás alkalmával.

3. Tétel: Az $\underline{y}^* \underline{B}$ szorzat a $\underline{B}$ mátrix sorvektorainak azt a lineáris kombinációját jelenti, amelyben skalárokként az $\underline{y}$ vektor komponensei szerepelnek.
Ha tehát

$$\underline{y}^* = [y_1, \; y_2, \ldots, \; y_q]$$

és

$$\underline{B} = \begin{bmatrix} b_{11} & b_{12} & \ldots & b_{1r} \\ b_{21} & b_{22} & \ldots & b_{2r} \\ \vdots & & & \\ b_{q1} & b_{q2} & \ldots & b_{qr} \end{bmatrix} = \begin{bmatrix} \underline{d}_1^* \\ \underline{d}_2^* \\ \vdots \\ \underline{d}_q^* \end{bmatrix},$$

akkor

$$\underline{y}^* \underline{B} = y_1 \underline{d}_1^* + y_2 \underline{d}_2^* + \ldots + y_q \underline{d}_q^* .$$

Az állitás helyessége a 2. tétel bizonyitásánál követett eljáráshoz hasonló módon látható be.
Például, ha

$$\underline{y}^* = [4, \quad -1, \quad 2] \quad \text{és} \quad \underline{B} = \begin{bmatrix} 5 & -1 \\ 0 & 3 \\ -1 & 2 \end{bmatrix} .$$

akkor

$$\underline{y}^* . \underline{B} = 4 [5, \ -1] + (-1)[0, \ 3] + 2[-1, \ 2] = [18, \ -3].$$

Ugyanerre az eredményre jutunk, ha a szorzás eredeti definiciója alapján járunk el.

Következmény:

Az $\underline{A}$. $\underline{B}$ szorzat sorvektorai - a fentiek értelmében - nem mások, mint a $\underline{B}$ mátrix, a második tényező sorvektorainak lineáris kombinációi. Az egyes lineáris kombinációkban skalárokként az $\underline{A}$, az első tényező, megfelelő sorvektorainak komponensei szerepelnek, vagyis ha

$$\underset{(p.q)}{\underline{A}} = \begin{bmatrix} \underline{g}_1^* \\ \underline{g}_2^* \\ . \\ . \\ . \\ \underline{g}_p^* \end{bmatrix} \quad \text{és} \quad \underset{(q\ r)}{\underline{B}} \text{, akkor } \underline{A} . \underline{B} = \begin{bmatrix} \underline{g}_1^* \ \underline{B} \\ \underline{g}_2^* \ \underline{B} \\ . \\ . \\ . \\ \underline{g}_p^* \ \underline{B} \end{bmatrix} .$$

Például:

$$\begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 3 & -2 \end{bmatrix} . \begin{bmatrix} 1 & -3 \\ 2 & 1 \end{bmatrix}$$

szorzat első sorvektora:

$$[1,\quad 2]\ .\begin{bmatrix}1 & -3\\ 2 & 1\end{bmatrix} = 1[1,\quad -3] + 2[2,\ 1] = [5,\quad -1],$$

második sorvektora:

$$[0,\quad 1]\ .\begin{bmatrix}1 & -3\\ 2 & 1\end{bmatrix} = 0\,.\,[1,\quad -3] + 1[2,\quad 1] = [2,\quad 1]\ ,$$

harmadik sorvektora:

$$[3,\ -2]\ .\begin{bmatrix}1 & -3\\ 2 & 1\end{bmatrix} = 3[1,\quad -3] + (-2)[2,\quad 1] = [-1, -11],$$

és igy

$$\begin{bmatrix}1 & 2\\ 0 & 1\\ 3 & -2\end{bmatrix} . \begin{bmatrix}1 & -3\\ 2 & 1\end{bmatrix} = \begin{bmatrix}5 & -1\\ 2 & 1\\ -1 & -11\end{bmatrix} .$$

Most is az ismert eredményre jutottunk.

4. <u>Tétel:</u> Két mátrix szorzatára mindig érvényes az

$$(\underline{A}\ \underline{B})^* = \underline{B}^* \cdot \underline{A}^*$$

azonosság.

<u>Bizonyitás:</u> Az állitás helyességének belátására tekintsük az

$$\underline{A}\ .\ \underline{B} = \underline{C}$$

szorzatot, ahol a szorzat mátrix általános eleme

$$c_{ij} = \underline{a}_i^* \ .\ \underline{b}_j$$

A transzponálás után ugyanez lesz a $\underline{C}^*$ mátrix j-edik sorának i-edik eleme:

$$[c_{ij}]^* = [c_{ji}] .$$

Másrészt a

$$\underline{B}^* . \underline{A}^*$$

mátrix j-edik sorának i-edik eleme viszont

$$\underline{b}_j^* \quad \underline{a}_i .$$

A skaláris szorzatra megismert kommutativ tulajdonság alapján azonban

$$\underline{a}_i^* . \underline{b}_j = \underline{b}_j^* . \underline{a}_i .$$

Ez azt jelenti, hogy az $(\underline{A}\ \underline{B})^*$ mátrix és a $\underline{B}^* . \underline{A}^*$ mátrix megfelelő elemei rendre megegyeznek egymással, vagyis

$$(\underline{A} . \underline{B})^* = \underline{B}^* . \underline{A}^* .$$

Ezzel állitásunk helyességét bebizonyitottuk.
Példaképpen legyen

$$\underline{A} = \begin{bmatrix} 1 & 2 \\ 0 & -1 \\ -1 & 0 \end{bmatrix} \quad \text{és} \quad \underline{B} = \begin{bmatrix} -1 & 0 & 2 \\ 1 & 1 & -1 \end{bmatrix} .$$

Ekkor

$$(\underline{A} . \underline{B}) = \begin{bmatrix} 1 & 2 & 0 \\ -1 & -1 & 1 \\ 1 & 0 & -2 \end{bmatrix} \quad (\underline{A}\ \underline{B})^* = \begin{bmatrix} 1 & -1 & 1 \\ 2 & -1 & 0 \\ 0 & 1 & -2 \end{bmatrix}$$

és

$$\underline{B}^* . \underline{A}^* = \begin{bmatrix} -1 & 1 \\ 0 & 1 \\ 2 & -1 \end{bmatrix} . \begin{bmatrix} 1 & 0 & -1 \\ 2 & -1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 1 \\ 2 & -1 & 0 \\ 0 & 1 & -2 \end{bmatrix} .$$

Az eredmények valóban eleget tesznek a

$$(\underline{A}\ \underline{B})^* = \underline{B}^* . \underline{A}^*$$

azonosságnak.

Következmény:

Az állítást többtényezős szorzatokra is ki lehet terjeszteni, azaz

$$(\underline{A}_1 . \underline{A}_2 . . . \underline{A}_k)^* = \underline{A}_k^* . . . \underline{A}_2^* . \underline{A}_1^* .$$

Ennek helyessége ugyancsak a szorzás definiciója és annak tulajdonságai alapján látható be.

A mátrixok szorzásával kapcsolatban érdemes megjegyezni, hogyan alakul a szorzat, ha a tényezők, vagy azok valamelyike speciális mátrix:

a) Két diagonális mátrix szorzata olyan diagonális mátrix, amelynek diagonális elemei megegyeznek a megfelelő diagonális elemek szorzatával.

Például:

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix} . \begin{bmatrix} 4 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 9 \end{bmatrix} = \begin{bmatrix} 2.4 & 0 & 0 \\ 0 & 3.7 & 0 \\ 0 & 0 & 5.9 \end{bmatrix} = \langle 6,\ 21,\ 45 \rangle .$$

Ez a megállapitás arra is rávilágit, hogy a diagonális mátrixok szorzatára nézve érvényes a kommutativitás.

b) Ha egy mátrixot jobbról szorzunk meg egy diagonális mátrixszal, akkor a szorzatnak megfelelő mátrix egyes oszlopvektorait legegyszerübben ugy kaphatjuk meg, hogy az első tényező oszlopvektorait rendre megszorozzuk a megfelelő diagonális elemekkel.

Például:

$$\begin{bmatrix} 5 & 0 & -1 \\ -2 & 3 & 4 \end{bmatrix} \cdot \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix} = \begin{bmatrix} 10 & 0 & -5 \\ -4 & 9 & 20 \end{bmatrix} .$$

c) Ha egy mátrixot balról szorzunk meg egy diagonális mátrixszal, akkor a szorzat egyes sorvektorait legegyszerübben ugy határozhatjuk meg, hogy a második tényező sorvektorait rendre megszorozzuk a megfelelő diagonális elemekkel.

Például:

$$\begin{bmatrix} 4 & 0 \\ 0 & -6 \end{bmatrix} \cdot \begin{bmatrix} 5 & 0 & -1 \\ -2 & 3 & 4 \end{bmatrix} = \begin{bmatrix} 20 & 0 & -4 \\ 12 & -18 & -24 \end{bmatrix} .$$

d) Felső háromszögmátrixok szorzata felső háromszögmátrix.

Például:

$$\begin{bmatrix} 2 & 3 & 1 \\ 0 & -3 & 0 \\ 0 & 0 & 5 \end{bmatrix} \cdot \begin{bmatrix} 3 & 0 & 4 \\ 0 & -2 & 8 \\ 0 & 0 & 4 \end{bmatrix} = \begin{bmatrix} 6 & -6 & 36 \\ 0 & 6 & -24 \\ 0 & 0 & 20 \end{bmatrix} .$$

Figyeljük meg, hogy a szorzat diagonális elemei megegyeznek a tényezők megfelelő diagonális elemeinek szorzatával. Hasonlóképpen az alsó háromszögmátrixok szorzata alsó háromszögmártix.

e) Ha egy mátrixot egységmátrixszal szorzunk, akkor a szorzat - amint azt már korábban láttuk -, maga a mátrix.

f) A különböző rendü egységvektorokból felépitett nem diagonális mátrixokkal való szorzásra is nézzünk példát:

$$\begin{bmatrix} 5 & -1 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 5 & 1 & -1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix} .$$

Ennek alapján észrevehetjük, hogy ha egy mátrixot megszorzunk egy ilyen mátrixszal, akkor az adott mátrixban csak az oszlopvektorok sorrendje változik meg. Az ilyen mátrixokat "permutáló" mátrixnak nevezzük. Ezekre jellemző, hogy oszlopainak megfelelő átrendezésével egységmátrixszá alakithatók át.

Az alábbi példa pedig arra mutat, ha egy mátrixot balról szorzunk meg egy permutáló mátrixszal, akkor az adott mátrixban csak a sorvektorok sorrendje változik meg:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 5 & -1 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 5 & -1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} .$$

1.2.3. Hatványozás

A speciális mátrixok szorzásánál külön foglalkozunk a kvadratikus mátrixok szorzásával abban az esetben, ha a tényezők azonosak, mert ez lehetővé teszi a hatványozás fogalmának bevezetését.

Az n-ed rendü kvadratikus mátrixok nem-negativ egész-kitevőjü hatványait a következő előírással értelmezzük:

$$\underline{A}^{o} = \underline{E}$$

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

$$\underline{A}^{2} = \underline{A} \cdot \underline{A}$$

$$\underline{A}^{3} = \underline{A}^{2} \cdot \underline{A}$$

$$\vdots$$

$$\underline{A}^{k} = \underline{A}^{k-1} \cdot \underline{A}$$

A nulladik hatványt tehát azonosnak tekintjük az egységmátrixszal, az első hatvány megegyezik magával a hatványozandó mátrixszal, az egynél nagyobb egész kitevőjü hatvány pedig olyan szorzatot jelent, amelyben az alapul vett mátrix annyiszor szerepel tényezőként, ahányszor azt a kitevő mutatja.
Ha például

$$\underline{A} = \begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix} ,$$

akkor

$$\underline{A}^{o} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}; \quad \underline{A}^{1} = \begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix}; \quad \underline{A}^{2} = \begin{bmatrix} -1 & 4 \\ -2 & -1 \end{bmatrix}; \quad \underline{A}^{3} = \begin{bmatrix} -5 & 2 \\ -1 & -5 \end{bmatrix} \quad \text{stb.}$$

A mátrixok hatványozására adott meghatározás és a szorzás asszociativ tulajdonsága értelmében az azonos alapu hatványok szorzatára érvényes a kommutativitás, azaz

$$\underline{A}^{p} \cdot \underline{A}^{q} = \underline{A}^{q} \cdot \underline{A}^{p}.$$

Például:

$$\underline{A}^{3} \cdot \underline{A}^{2} = \underline{A}^{2} \cdot \underline{A}^{3} \quad \text{mivel} \quad (\underline{A}\ \underline{A}\ \underline{A}) \cdot (\underline{A}\ \underline{A}) = (\underline{A}\ \underline{A}) \cdot (\underline{A}\ \underline{A}\ \underline{A})$$

<u>Megjegyzések:</u> a) a kvadratikus nullmátrix minden hatványa nullmátrix, azaz

$$\underline{0}^{k} = \underline{0}$$

b) az egységmátrixra pedig

$$\underline{E}^{k} = \underline{E}$$

azonosság érvényes, ahol k nem-negativ egész szám.

A mátrixok szorzatának értelmezéséből következik, hogy egy nullmátrixtól különböző mátrix hatványa is lehet null-mátrix.
Ha például

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

akkor

$$\underline{A}^{o} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}; \quad \underline{A}^{1} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}; \quad \underline{A}^{2} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{és} \quad \underline{A}^{3} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Az olyan mátrixot, amely valamely k kitevő mellett kielégiti az

$$\underline{A}^k = \underline{0}$$

egyenlőséget, nilpotens mátrixnak nevezzük.

E szerint a kvadratikus nullmátrix nilpotens, de nemcsak a nullmátrixok tartoznak a nilpotens mátrixok halmazához.

Érdekes megjegyezni, hogy a skaláraritmetikától eltérően nemcsak a null- és egységmátrixok k-adik hatványa lehet egyenlő első hatványával. Igy például, ha

$$\underline{A} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}, \text{ akkor } \underline{A}^2 = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \text{ is fennáll.}$$

Ha valamely $\underline{P}$ mátrix eleget tesz a

$$\underline{P}^2 = \underline{P}$$

követelménynek, akkor a $\underline{P}$ mátrixot projektor mátrixnak nevezzük.

E definicióból következik, hogy az egységmátrix is és a nullmátrix is projektor mátrix.

1.3. Mátrixpolinomok

A lineáris algebrában fontos szerepet játszanak az n-edrendü kvadratikus mátrixok

$$\underline{A}^0, \ \underline{A}^1, \ \underline{A}^2, \ldots, \ \underline{A}^k$$

hatványainak egymásutánja, illetve ezek lineáris kombinációja, amit a következőkben - a skalárpolinomok mintájára - mátrixpolinomként értelmezünk.

A

$$c_o \underline{A}^o + c_1 \underline{A}^1 + c_2 \underline{A}^2 + \ldots + c_k \underline{A}^k$$

lineáris kombinációt az $\underline{A}$ mátrix k-adfoku polinomjának nevezzük, feltéve, hogy $c_k \neq 0$.

A

$$c_0, c_1, \dots, c_k$$

skalárok a polinom együtthatói.
Jelölésére a p ($\underline{A}$) szimbólumot használjuk.

Ha a p($\underline{A}$) mátrixpolinomban az $\underline{A}$ mátrixnak megfelelő számtáblázatot behelyettesitjük, s a kijelölt müveleteket elvégezzük, megkapjuk a mátrixpolinom helyettesitési értékét az $\underline{A}$ helyen.
Például a

$$p(\underline{A}) = -6\,\underline{A}^0 + 5\,\underline{A} + \underline{A}^2$$

mátrixpolinom helyettesitési értéke az

$$\underline{A} = \begin{bmatrix} 1 & 3 \\ 2 & 0 \end{bmatrix} \quad \text{helyen} \quad p(\underline{A}) = \begin{bmatrix} 6 & 18 \\ 12 & 0 \end{bmatrix}.$$

Mivel a mátrixpolinomok mátrixot értelmeznek a mátrixokra definiált összeadás és szorzás müveletét ezekre is kiterjeszthetjük. Legyen az $\underline{A}$ mátrixnak két polinomja:

$$p_1(\underline{A}) = c_o\,\underline{A}^o + c_1\underline{A}^1 + c_2\,\underline{A}^2 + \dots + c_k\,\underline{A}^k$$

és

$$p_2(\underline{A}) = u_o\,\underline{A}^o + u_1\,\underline{A}^1 + u_2\,\underline{A}^2 + \dots + u_\ell\,\underline{A}^\ell\,,$$

ahol $\ell > k$ és $u_\ell \neq 0$, $c_k \neq 0$.

E két polinom összege nyilván ℓ-edfoku polinom, azaz

$$p_1(\underline{A}) + p_2(\underline{A}) = (c_o + u_o)\,\underline{A}^o + (c_1 + u_1)\,\underline{A}^1 + \dots + u_\ell\,\underline{A}^\ell\,.$$

A két polinom szorzata pedig $k + \ell$ -edfoku polinomot ad, azaz

$$p_1(\underline{A}) \cdot p_2(\underline{A}) = v_o \underline{A}^o + \ldots + v_{k+\ell} \underline{A}^{k+\ell}$$

Egy mártix hatványainak szorzatára vonatkozó kommutatitivitási tételből következik, hogy két polinom szorzata független a tényezők sorrendjétől, azaz

$$p_1(\underline{A}) \cdot p_2(\underline{A}) = p_2(\underline{A}) \cdot p_1(\underline{A}).$$

Ha például

$$p_1(\underline{A}) = 3\,\underline{A}^o - 2\,\underline{A} + 5\,\underline{A}^2$$

és

$$p_2(\underline{A}) = \underline{A}^o + 4\,\underline{A}$$

akkor

$$p_1(\underline{A}) \cdot p_2(\underline{A}) = (3\underline{A}^o - 2\underline{A} + 5\underline{A}^2)(\underline{A}^o + 4\underline{A}) = 3\underline{A}^o + 10\underline{A} - 3\underline{A}^2 + 20\,\underline{A}^3,$$

$$p_2(\underline{A}) \cdot p_1(\underline{A}) = (\underline{A}^o + 4\underline{A})\,(3\underline{A}^o - 2\underline{A} + 5\underline{A}^2) = 3\,\underline{A}^o + 10\,\underline{A} - 3\underline{A}^2 + 20\,\underline{A}^3,$$

ahonnan valóban

$$p_1(\underline{A}) \cdot p_2(\underline{A}) = p_2(\underline{A}) \cdot p_1(\underline{A}).$$

Ennek alapján belátható, hogy ugyanazon mátrixra értelmezett mátrixpolinomokra nézve az összeadásra és szorzásra támaszkodó átalakitási szabályok megegyeznek a skalárpolinomokra vonatkozó átalakitási szabályokkal.

Ha a mátrixpolinomban az $\underline{A}$ mátrix helyébe egy skalárt irunk, akkor a mátrixpolinomhoz tartozó skalárpolinomhoz jutunk.

Amennyiben a kérdéses skalárt x-szel jelöljük, akkor a

$$p(\underline{A}) = c_o \underline{A}^o + c_1 \underline{A}^1 + c_2 \underline{A}^2 + \ldots + c_k \underline{A}^k$$

mátrixpolinomnak megfelelő skalárpolinom

$$p(x) = c_o + c_1\, x + c_2\, x^2 + \ldots + c_k\, x^k.$$

Az elmondottakból következik, hogy ha a $p(\underline{A})$ mátrixpolinomhoz tartozó $p(x)$ skalárpolinomot a

$$p(x) = c_k\ (x-x_1)\ (x-x_2)\ldots(x-x_k)$$

összefüggésnek megfelelően szorzattá tudjuk átalakitani, akkor érvényes a

$$p(\underline{A}) = c_k\ (\underline{A} - x_1\,\underline{A}^o)\ (\underline{A} - x_2\,\underline{A}^o)\ldots(\underline{A} - x_k\,\underline{A}^o)$$

összefüggés is.
Tekintsük például a

$$p(\underline{A}) = -24\ \underline{A}^o + 2\ \underline{A} + 2\ \underline{A}^2$$

mátrixpolinomot. Ennek a

$$p(x) = -24 + 2x + 2\ x^2$$

skalárpolinom felel meg, ami

$$p(x) = 2(x - 3)\ (x + 4)$$

alakban is felirható. Ebből viszont következik, hogy

$$p(\underline{A}) = 2\ (\underline{A} - 3\ \underline{A}^o)\ (\underline{A} + 4\ \underline{A}^o).$$

A mátrixpolinomnak megfelelő skalárpolinom felirásánál ügyelni kell a nulladfoku tag felirására.

Ha egy mátrixpolinom együtthatói adottak, akkor felmerülhet az a kérdés, milyen mátrix behelyettesitése révén ad ez a polinom nullmátrixot. Ezzel a kérdéssel nem foglalkozunk behatóbban. Itt csupán annyit emlitünk meg, ha $p(\underline{A}) = \underline{0}$, akkor azt mondjuk, hogy az $\underline{A}$ mátrix kielégiti a $p(x) = 0$ egyenletnek megfelelő $p(\underline{A}) = \underline{0}$ egyenletet.
Például, ha

$$p(\underline{A}) = -24\ \underline{A}^o + 2\ \underline{A} + 2\ \underline{A}^2$$

a szóbanforgó mátrixpolinom, akkor közvetlen számolással meggyőződhetünk, hogy az

$$\underline{A} = \begin{bmatrix} 3 & 0 \\ 0 & -4 \end{bmatrix}$$

mátrixra $p(\underline{A}) = \underline{0}$, ugyanis

$$p(\underline{A}) = -24 \begin{bmatrix} 3 & 0 \\ 0 & -4 \end{bmatrix}^0 + 2 \begin{bmatrix} 3 & 0 \\ 0 & -4 \end{bmatrix}^1 + 2 \begin{bmatrix} 3 & 0 \\ 0 & -4 \end{bmatrix}^2 =$$

$$= -24 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + 2 \begin{bmatrix} 3 & 0 \\ 0 & -4 \end{bmatrix} + 2 \begin{bmatrix} 9 & 0 \\ 0 & 16 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Itt a $p(\underline{A}) = \underline{0}$ egyenletet kielégitő $\underline{A}$ mátrix diagonális mártix, amelynek diagonális elemei a mátrix-polinomnak megfelelő $p(x)$ skalárpolinom gyökei (vagyis a $p(x) = -24 + 2x+2x^2 = 2(x-3)(x+4)$ polinom az $x = 3$ és $x = -4$ helyen nullává válik!)

E szerint a $p(\underline{A}) = \underline{0}$ feltételt minden olyan diagonális mátrix kielégiti, amelynek a diagonális elemei megegyeznek a megfelelő skalárpolinom gyökeivel.

1.4. Számolás blokkokra bontott mátrixokkal

A mátrixok particionálásánál emlitettük, hogy a fogalom bevezetése sok esetben leegyszerüsiti a tárgyalásmódot. Ezt tükrözi a blokkokra bontott mátrixokkal való számolás is.

A blokkokra bontott mátrixokra nézve az eddig megismert müveleti szabályok azzal a kiegészitéssel érvényesülnek, hogy a megfelelő blokkok között a történetesen elvégzendő müveletek értelmezve legyenek.

Az elmondottak megvilágitására tekintsük a következő két mátrixot, amelynek mindegyike négy blokkot tartalmaz:

$$\underline{A} = \begin{bmatrix} \underline{A}_{11} & \underline{A}_{12} \\ \underline{A}_{21} & \underline{A}_{22} \end{bmatrix} \quad \text{és} \quad \underline{B} = \begin{bmatrix} \underline{B}_{11} & \underline{B}_{12} \\ \underline{B}_{21} & \underline{B}_{22} \end{bmatrix}.$$

Az $\underline{A}$ és $\underline{B}$ particionált mátrixok összege csak akkor van értelmezve, ha az $\underline{A}_{ij}$ és $\underline{B}_{ij}$ mátrixok rendre azonos tipusuak. Ez esetben

$$\underline{A} + \underline{B} = \begin{bmatrix} \underline{A}_{11} + \underline{B}_{11} & \underline{A}_{12} + \underline{B}_{12} \\ \\ \underline{A}_{21} + \underline{B}_{21} & \underline{A}_{22} + \underline{B}_{22} \end{bmatrix}.$$

Az $\underline{A}$ és $\underline{B}$ particionált mátrixok esetén a szorzás csak abban az esetben van értelmezve, ha a szorzásra adott definició szerint előállitott elemekben a kijelölt müveletek elvégezhetők, vagyis az

$$\underline{A}\ \underline{B} = \begin{bmatrix} \underline{A}_{11} \cdot \underline{B}_{11} + \underline{A}_{12} \cdot \underline{B}_{21} & \underline{A}_{11} \cdot \underline{B}_{12} + \underline{A}_{12} \cdot \underline{B}_{22} \\ \\ \underline{A}_{21} \cdot \underline{B}_{11} + \underline{A}_{22} \cdot \underline{B}_{21} & \underline{A}_{21} \cdot \underline{B}_{12} + \underline{A}_{22} \cdot \underline{B}_{22} \end{bmatrix}$$

szorzatban pl. az

$$\underline{A}_{11}\ \underline{B}_{11} + \underline{A}_{12}\ \underline{B}_{21}$$

blokkot tekintve

a) $\underline{A}_{11}$ oszlopainak száma megegyezik $\underline{B}_{11}$ sorainak számával, azaz az $\underline{A}_{11} \cdot \underline{B}_{11}$ szorzat értelmezve van,

b) hasonlóképpen az $\underline{A}_{12} \cdot \underline{B}_{21}$ szorzat is előállitható, továbbá

c) az $\underline{A}_{11}\ \underline{B}_{11}$ és $\underline{A}_{12}\ \underline{B}_{21}$ szorzatok azonos tipusuak s igy összegezhetők.

E feltételek teljesülését blokkonként ellenőrizve, ha azok teljesülnek, a szorzás elvégezhető blokkokra bontott mátrixok esetén.

Példaként tekintsük az alábbi particionált mátrixokat:

$$\underline{A} = \begin{bmatrix} \underline{A}_{11} & \underline{A}_{12} \\ \\ \underline{A}_{21} & \underline{A}_{22} \end{bmatrix} = \left[\begin{array}{ccc:cc} 5 & 4 & -3 & 5 & 0 \\ 3 & 0 & -6 & 2 & 1 \\ 0 & -4 & 5 & 8 & 0 \\ \hdashline 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{array}\right],$$

$$\underline{B} = \begin{bmatrix} \underline{B}_{11} & \underline{B}_{12} \\ \underline{B}_{21} & \underline{B}_{22} \end{bmatrix} = \left[\begin{array}{cc:ccc} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ \hdashline 1 & 0 & 2 & 0 & 1 \\ 0 & 1 & 0 & 2 & 0 \end{array}\right]$$

Mind az $\underline{A}$, mind a $\underline{B}$ mátrix 5.5 tipusu mátrix, amelyek mindegyike négy blokkot tartalmaz.

A blokkokra bontott mártixokra figyeljük meg, hogy

a) az összeadás az igy particionált mátrixok között nincs értelmezve, mert pl. a 3.3-as tipusu $\underline{A}_{11}$ blokkhoz nem adható hozzá a 3.2-es tipusu $\underline{B}_{11}$ blokk, jóllehet az $\underline{A} + \underline{B}$ összeadás, ha magát az egész számtáblázatot tekintjük, elvégezhető,

b) az $\underline{A} \cdot \underline{B}$ szorzat az igy particionált formában értelmezve van, mert a blokkokra kijelölt müveletek a kivánt formában elvégezhetők, ugyanis:

$$\underline{A}_{11} \cdot \underline{B}_{11} + \underline{A}_{12} \cdot \underline{B}_{21} = \begin{bmatrix} 5 & 4 & -3 \\ 3 & 0 & -6 \\ 0 & -4 & 5 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 5 & 0 \\ 2 & 1 \\ 8 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 10 & 4 \\ 5 & 1 \\ 8 & -4 \end{bmatrix}$$

$$\underline{A}_{11} \underline{B}_{12} + \underline{A}_{12} \underline{B}_{22} = \begin{bmatrix} 5 & 4 & -3 \\ 3 & 0 & -6 \\ 0 & -4 & 5 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 5 & 0 \\ 2 & 1 \\ 8 & 0 \end{bmatrix} \cdot \begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & 0 \end{bmatrix} = \begin{bmatrix} 16 & 0 & 7 \\ 1 & 2 & -1 \\ 17 & 0 & 13 \end{bmatrix}$$

$$\underline{A}_{21} \cdot \underline{B}_{11} + \underline{A}_{22} \cdot \underline{B}_{21} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$$

$$\underline{A}_{21}\underline{B}_{12}+\underline{A}_{22}\underline{B}_{22} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & 0 \end{bmatrix} = \begin{bmatrix} 5 & 0 & 3 \\ 2 & 2 & 1 \end{bmatrix}.$$

Ezért

$$\underline{A} \cdot \underline{B} = \left[\begin{array}{cc|ccc} 10 & 4 & 16 & 0 & 7 \\ 5 & 1 & -1 & 2 & -1 \\ 8 & -4 & 17 & 0 & 13 \\ \hline 2 & 1 & 5 & 0 & 3 \\ 1 & 1 & 2 & 2 & 1 \end{array}\right]$$

A megoldás helyességét ellenőrizhetjük azzal, hogy a szorzást blokkokra való felbontás nélkül is elvégezzük.

Az elmondottak értelemszerüen kiterjeszthetők arra az esetre is, amikor az adott mátrixokat négynél több blokkra bontjuk fel.

A számolás blokkokra bontott mátrixok esetén igen egyszerüvé válik, ha a blokkokban speciális mátrixok vannak.

Tekintsük pl. az alábbi mátrixokat:

$$\underline{A} = \left[\begin{array}{ccc|cc} 1 & 2 & 1 & 0 & 0 \\ 0 & 3 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] \quad \text{és} \quad \underline{B} = \left[\begin{array}{ccc|cc} 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 1 \\ \hline 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{array}\right].$$

Képezzük ezek szorzatát:

$$\underline{A}_{11}\underline{B}_{11} + \underline{A}_{12}\underline{B}_{21} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad \underline{A}_{11}\underline{B}_{12} + \underline{A}_{12}\underline{B}_{22} = \begin{bmatrix} 5 & 2 \\ 3 & 1 \\ 3 & 1 \end{bmatrix}$$

$$\underline{A}_{21} \cdot \underline{B}_{11} + \underline{A}_{22} \cdot \underline{B}_{21} = \begin{bmatrix} 0 & 3 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$\underline{A}_{21} \cdot \underline{B}_{12} + \underline{A}_{22}\underline{B}_{22} = \begin{bmatrix} 0 & 0 \\ & 0 \\ 0 & 0 \end{bmatrix},$$

ahonnan

$$\underline{A} \cdot \underline{B} = \left[\begin{array}{ccc|cc} 0 & 0 & 0 & 5 & 2 \\ 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 3 & 1 \\ \hline 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

A blokkokra bontott mátrixok hatványait a szorzásra mondottak figyelembevételével értelmezhetjük. Igy az előbb felirt $\underline{A}$ mátrix négyzete:

$$\underline{A}^2 = \begin{bmatrix} \underline{A}_{11} & \underline{A}_{12} \\ \underline{A}_{21} & \underline{A}_{22} \end{bmatrix} \cdot \begin{bmatrix} \underline{A}_{11} & \underline{A}_{12} \\ \underline{A}_{21} & \underline{A}_{22} \end{bmatrix},$$

ahol

$$\underline{A}_{11} \cdot \underline{A}_{11} + \underline{A}_{12} \cdot \underline{A}_{21} = \underline{A}_{11}^2$$

$$\underline{A}_{11} \cdot \underline{A}_{12} + \underline{A}_{12} \cdot \underline{A}_{22} = \underline{0}$$

$$\underline{A}_{21} \cdot \underline{A}_{11} + \underline{A}_{22} \cdot \underline{A}_{21} = \underline{0}$$

$$\underline{A}_{21} \cdot \underline{A}_{12} + \underline{A}_{22} \cdot \underline{A}_{22} = \underline{A}_{22}^2$$

vagyis

$$\underline{A}^2 = \left[\begin{array}{ccc|cc} 1 & 8 & 4 & 0 & 0 \\ 0 & 9 & 4 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Mivel a blokkokra bontott mátrix diagonális blokkjain kivüli blokkok zérus blokkok voltak, a szorzás nagyon leegyszerüsödött, lényegében a diagonális blokkokat kellett a megfelelő hatványra emelni.

3.6 Determinánsok

A K számtest felett értelmezett $(n \cdot n)$-es $\mathbf{A}$ mátrixok halmazának a K számtestbe való egyik lehetséges leképezését *determinánsnak* nevezzük.

A leképezést így definiáljuk:

minden $(n \cdot n)$*-es mátrixnak megfeleltethetünk egy* n*-ed rendű determinánsnak nevezett* $n!$ *tagból álló algebrai összeget, amelynek tagjai a mátrix elemeiből képzett összes lehetséges* n *tényezős szorzatok. A tagok tényezői között a mátrix minden sorából és minden oszlopából egy-egy elem szerepel. A tagok előjele pozitív, ha a kiválasztott elemek sor- és oszlopindexei páros permutációt (a permutációban az előzések száma páros) alkotnak és negatív az ellenkező esetben.*

Az $(n \cdot n)$-es

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

mátrix determinánsának jelölésére az

$$|\mathbf{A}|, \text{ illetve az } \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & & \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} \tag{85}$$

szimbólumot használjuk. Szokásos még a det $(\mathbf{A})$ jelölés is.

Az $\mathbf{A}$ mátrix determinánsának megfelelő összeg felírását megkönnyíti az a megállapodás, hogy a sorokból az $1, 2, \dots, n$ sorrendben történik az elemek kiválasztása, az oszlopokból pedig az n elem összes lehetséges permutációinak megfelelően. Ez esetben csak az oszlopindexekben előforduló előzések számát kell meghatározni, abból már megállapítható a kérdéses tag előjele. Eszerint

$$|\mathbf{A}| = \sum (-1)^{\mathscr{I}} a_{1\alpha_1} a_{2\alpha_2} \dots a_{n\alpha_n} \tag{86}$$

ahol $\mathscr{I}$ az $1, 2, \dots, n$ elemek $\alpha_1 \alpha_2 \dots \alpha_n$ permutációjában az előzések (inverziók) számát jelenti. Az összeg $n!$ tagból áll, ahány lehetséges permutációja van az $1, 2, \dots, n$ elemeknek.

$n=1$ esetén, azaz egy elemből álló mátrixokra, a determináns egyenlő magával az elemmel.

$n=2$ esetén az

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

mátrixnak megfelelő determináns

$$|\mathbf{A}| = a_{11}a_{22} - a_{12}a_{21}$$

$2! = 2$ tagból álló összeg, amelyben az első tag előjele pozitív (az oszlopindexek permutációjában az előzések száma 0), a második tag előjele negatív (az oszlopindexek permutációjában az előzések száma egy).

$n=3$ esetén az

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

mátrixnak megfelelő determináns $3! = 6$ tagból álló összeg.

Az $(n \cdot n)$-es

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

mátrix determinánsának jelölésére az

$$|\mathbf{A}|, \text{ illetve az } \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & & \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} \tag{85}$$

szimbólumot használjuk. Szokásos még a det $(\mathbf{A})$ jelölés is.

Az $\mathbf{A}$ mátrix determinánsának megfelelő összeg felírását megkönnyíti az a megállapodás, hogy a sorokból az $1, 2, \dots, n$ sorrendben történik az elemek kiválasztása, az oszlopokból pedig az n elem összes lehetséges permutációinak megfelelően. Ez esetben csak az oszlopindexekben előforduló előzések számát kell meghatározni, abból már megállapítható a kérdéses tag előjele. Eszerint

$$|\mathbf{A}| = \sum (-1)^{\mathscr{I}} a_{1\alpha_1} a_{2\alpha_2} \dots a_{n\alpha_n} \tag{86}$$

ahol $\mathscr{I}$ az $1, 2, \dots, n$ elemek $\alpha_1 \alpha_2 \dots \alpha_n$ permutációjában az előzések (inverziók) számát jelenti. Az összeg $n!$ tagból áll, ahány lehetséges permutációja van az $1, 2, \dots, n$ elemeknek.

$n = 1$ esetén, azaz egy elemből álló mátrixokra, a determináns egyenlő magával az elemmel.

$n = 2$ esetén az

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

mátrixnak megfelelő determináns

$$|\mathbf{A}| = a_{11}a_{22} - a_{12}a_{21}$$

$2! = 2$ tagból álló összeg, amelyben az első tag előjele pozitív (az oszlopindexek permutációjában az előzések száma 0), a második tag előjele negatív (az oszlopindexek permutációjában az előzések száma egy).

$n = 3$ esetén az

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

mátrixnak megfelelő determináns $3! = 6$ tagból álló összeg.

A tételből következik, hogy minden, a determináns soraira mondott állítás érvényes az oszlopaira is és megfordítva. A determinánsban így a sorok és oszlopok egyenrangúak.

2. *Ha egy determináns egyik sora csupa zérus, akkor a determináns értéke is zérus.*

Bizonyítás. Legyen a determináns i-edik sorának minden eleme zérus. A determináns minden tagjában szerepel egy tényező az i-edik sorból, ezért a determináns minden tagja és így ezek összege is zérus.

3. *Ha egy determináns két sorát felcseréljük, akkor a determináns előjelet vált.*

Bizonyítás. Legyen

$$\sum(-1)^{\mathscr{I}} a_{1\alpha_1} a_{2\alpha_2} \dots a_{i\alpha_i} \dots a_{j\alpha_j} \dots a_{n\alpha_n}$$

az eredeti determináns értéke. Ha az i-edik és j-edik sort a determinánsban felcseréljük és az összeg képzését az előzővel azonos módon végezzük, akkor

$$\sum(-1)^{\mathscr{I}'} a_{1\alpha_1} a_{2\alpha_2} \dots a_{j\alpha_i} \dots a_{i\alpha_j} \dots a_{n\alpha_n}$$

összegben a sorindexekben végrehajtott egyetlen transzpozíció miatt az előjel ellenkezőjére fordul.

4. *Ha egy determinánsban két sor elemenként megegyezik, akkor a determináns értéke zérus.*

Bizonyítás. Cseréljük fel az azonos sorokat. Az előző tétel szerint ekkor a determináns előjele megváltozik. A két sor megegyezése miatt azonban a determináns változatlan. Így

$$|\mathbf{A}| = -|\mathbf{A}|.$$

Innen következik, hogy $|\mathbf{A}|=0$.

5. *Ha egy determináns valamely sorának minden egyes elemét megszorozzuk λ-val, akkor a determináns értéke is szorzódik λ-val.*

Bizonyítás. Legyen

$$|\mathbf{A}| = \sum(-1)^{\mathscr{I}} a_{1\alpha_1} a_{2\alpha_2} \dots a_{i\alpha_i} \dots a_{n\alpha_n}$$

az eredeti determináns értéke. Mivel a determináns minden tagjában szerepel minden sorból és minden oszlopból egy elem tényezőként, ezért a szorzás utáni determináns, ha annak i-edik sorát szoroztuk λ-val:

$$\sum(-1)^{\mathscr{I}} a_{1\alpha_1} a_{2\alpha_2} \dots \lambda a_{i\alpha_i} \dots a_{n\alpha_n}$$

alakú. Eszerint a λ az összeg minden tagjában szorzótényező, így kiemelhető:

$$\lambda \sum(-1)^{\mathscr{I}} a_{1\alpha_1} a_{2\alpha_2} \dots a_{i\alpha_i} \dots a_{n\alpha_n},$$

és ez éppen

$$\lambda |\mathbf{A}|,$$

ahogy állítottuk.

A tételből következik, ha az $|\mathbf{A}|$ determináns minden elemét λ-val szorozzuk, akkor a determináns értéke λ^n-szeresére változik, azaz

$$|\lambda \mathbf{A}| = \lambda^n |\mathbf{A}|.$$

6. *Az egymással arányos sorokat tartalmazó determináns értéke zérus.*

Bizonyítás. Legyenek az $|\mathbf{A}|$ determináns i-edik sorának elemei a j-edik $(i \neq j)$ sor elemeinek λ-szorosai. Ha ezt a λ tényezőt kiemeljük a determináns i-edik sorából, akkor olyan determinánst kapunk, amelyben két sor azonos. Innen már következik, hogy $|\mathbf{A}|=0$.

7. *Ha egy determináns i-edik sorának minden eleme két tag összegeként áll elő, azaz*

$$a_{ij} = a'_{ij} + a''_{ij} \qquad (j = 1, 2, \ldots, n),$$

akkor a determináns egyenlő két determináns összegével, melyekben az i-edik sor kivételével az összes többi azonos az adott determináns megfelelő sorával, az i-edik sor pedig az egyik összeadandóban az a'_{ij}, a másikban pedig az a''_{ij} elemekből áll.

Bizonyítás. Legyen $|\mathbf{A}|$ a tétel szerinti determináns, akkor a tétel állítása a következőképpen írható:

$$|\mathbf{A}| = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a'_{i1}+a''_{i1} & a'_{i2}+a''_{i2} & \ldots & a'_{in}+a''_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix} =$$

$$= \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & & & \\ a'_{i1} & a'_{i2} & \ldots & a'_{in} \\ \vdots & & & \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & & & \\ a''_{i1} & a''_{i2} & \ldots & a''_{in} \\ \vdots & & & \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix}$$

Az egyenlőség bal oldalán levő determináns értéke

$$\sum (-1)^{\mathscr{I}} a_{1\alpha_1} a_{2\alpha_2} \ldots (a'_{i\alpha_i} + a''_{i\alpha_i}) \ldots a_{n\alpha_n}$$

alakú, amely a beszorzás és rendezés után

$$\sum (-1)^{\mathscr{I}} a_{1\alpha_1} a_{2\alpha_2} \ldots a'_{i\alpha_i} \ldots a_{n\alpha_n} + \sum (-1)^{\mathscr{I}} a_{1\alpha_1} a_{2\alpha_2} \ldots a''_{i\alpha_i} \ldots a_{n\alpha_n}$$

kifejezést adja. Ez a két összeg pedig éppen a fenti egyenlőség jobb oldalán álló két determináns. Ezzel állításunkat bizonyítottuk.

8. *Ha egy determináns valamelyik sora két másik sorának lineáris kombinációja, akkor a determináns értéke zérus.*

Bizonyítás. Legyen az **A**-ban a determináns i-edik sora a j-edik és a k-adik sor lineáris kombinációja, azaz

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & & & \\ a_{i-1\,1} & a_{i-1\,2} & \dots & a_{i-1\,n} \\ \alpha a_{j1}+\beta a_{k1} & \alpha a_{j2}+\beta a_{k2} & \dots & \alpha a_{jn}+\beta a_{kn} \\ a_{i+1\,1} & a_{i+1\,2} & \dots & a_{i+1\,n} \\ \vdots & & & \\ a_{j1} & a_{j2} & \dots & a_{jn} \\ \vdots & & & \\ a_{k1} & a_{k2} & \dots & a_{kn} \\ \vdots & & & \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}.$$

Ez az előző tétel szerint felírható két determináns összegeként:

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & & & \\ a_{i-1\,1} & a_{i-1\,2} & & a_{i-1\,n} \\ \alpha a_{j1} & \alpha a_{j2} & & \alpha a_{jn} \\ \vdots & & & \\ a_{j1} & a_{j2} & & a_{jn} \\ \vdots & & & \\ a_{k1} & a_{k2} & & a_{kn} \\ \vdots & & & \\ a_{n1} & a_{n2} & & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & & & \\ a_{i-1\,1} & a_{i-1\,2} & & a_{i-1\,n} \\ \beta a_{k1} & \beta a_{k2} & & \beta a_{kn} \\ \vdots & & & \\ a_{j1} & a_{j2} & & a_{jn} \\ \vdots & & & \\ a_{k1} & a_{k2} & & a_{kn} \\ \vdots & & & \\ a_{n1} & a_{n2} & & a_{nn} \end{vmatrix}$$

Mind a két tag i-edik sorából kiemelve a megfelelő szorzót:

$$\alpha\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & & & \\ a_{i-1\,1} & a_{i-1\,2} & & a_{i-1\,n} \\ a_{j1} & a_{j2} & & a_{jn} \\ \vdots & & & \\ a_{j1} & a_{j2} & & a_{jn} \\ \vdots & & & \\ a_{k1} & a_{k2} & & a_{kn} \\ \vdots & & & \\ a_{n1} & a_{n2} & & a_{nn} \end{vmatrix} + \beta\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & & & \\ a_{i-1\,1} & a_{i-1\,2} & & a_{i-1\,n} \\ a_{k1} & a_{k2} & & a_{kn} \\ \vdots & & & \\ a_{j1} & a_{j2} & & a_{jn} \\ \vdots & & & \\ a_{k1} & a_{k2} & & a_{kn} \\ \vdots & & & \\ a_{n1} & a_{n2} & & a_{nn} \end{vmatrix}$$

adódik. A 4. tulajdonság szerint mind a két determináns értéke zérus.

Ezzel állításunkat igazoltuk.

Az állítás kiterjeszthető arra az esetre is, amelyben a determináns valamelyik sora összes többi sorának lineáris kombinációja.

9. *A determináns értéke nem változik meg, ha valamely sorának minden eleméhez hozzáadjuk egy másik sora megfelelő elemeinek λ-szorosát.*

Bizonyítás. Az $|\mathbf{A}|$-nak i-edik sorához adjuk hozzá a k-adik sor λ-szorosát, azaz írjuk fel az

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & & & \\ a_{i1}+\lambda a_{k1} & a_{i2}+\lambda a_{k2} & \dots & a_{in}+\lambda a_{kn} \\ \vdots & & & \\ a_{k1} & a_{k2} & \dots & a_{kn} \\ \vdots & & & \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

determinánst. Ez az előző tételek alapján felbontható két tag összegére:

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & & & \\ a_{i1} & a_{i2} & \dots & a_{in} \\ \vdots & & & \\ a_{k1} & a_{k2} & \dots & a_{kn} \\ \vdots & & & \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & & & \\ \lambda a_{k1} & \lambda a_{k2} & \dots & \lambda a_{kn} \\ \vdots & & & \\ a_{k1} & a_{k2} & \dots & a_{kn} \\ \vdots & & & \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

ahol a második tagban a λ kiemelése után két azonos sor van, ezért értéke zérus. Így állításunkat bizonyítottuk.

A 9. tétel általánosítása: a determináns értéke nem változik, ha valamely sorához hozzáadjuk a többi sor tetszőleges kombinációját.

10. *Két $(n \cdot n)$-es mátrix szorzatának determinánsa megegyezik a tényezők determinánsának szorzatával.*

Jelekben: $|\mathbf{AB}| = |\mathbf{A}|\,|\mathbf{B}|$.

Az állítás helyessége az előző tételek felhasználásával bizonyítható.

$n=2$ esetén a tételt így igazoljuk: legyen

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \quad \text{és} \quad \mathbf{B} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

adott, akkor az

$$\mathbf{AB} = \begin{bmatrix} a_{11}b_{11}+a_{12}b_{21} & a_{11}b_{12}+a_{12}b_{22} \\ a_{21}b_{11}+a_{22}b_{21} & a_{21}b_{12}+a_{22}b_{22} \end{bmatrix}$$

mátrix determinánsa a 3., a 4., az 5. és a 7. tétel felhasználása után így írható.

$$
\begin{aligned}
|\mathbf{AB}| &= \begin{vmatrix} a_{11}b_{11} & a_{11}b_{12} \\ a_{21}b_{11} & a_{21}b_{12} \end{vmatrix} + \begin{vmatrix} a_{12}b_{21} & a_{11}b_{12} \\ a_{22}b_{21} & a_{21}b_{12} \end{vmatrix} + \\
&\quad + \begin{vmatrix} a_{11}b_{11} & a_{12}b_{22} \\ a_{21}b_{11} & a_{22}b_{22} \end{vmatrix} + \begin{vmatrix} a_{12}b_{21} & a_{12}b_{22} \\ a_{22}b_{21} & a_{22}b_{22} \end{vmatrix} = \\
&= b_{11}b_{12}\begin{vmatrix} a_{11} & a_{11} \\ a_{21} & a_{21} \end{vmatrix} + b_{12}b_{21}\begin{vmatrix} a_{12} & a_{11} \\ a_{22} & a_{21} \end{vmatrix} + \\
&\quad + b_{11}b_{22}\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} + b_{21}b_{22}\begin{vmatrix} a_{12} & a_{12} \\ a_{22} & a_{22} \end{vmatrix} = \\
&= \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}(b_{11}b_{22} - b_{12}b_{21}) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}\begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix} = \\
&= |\mathbf{A}|\,|\mathbf{B}|,
\end{aligned}
$$

ahogy állítottuk.

2. A lineáris térről

2.1. Az n-elemű vektorok halmaza

A következőkben a speciális mátrixok közt bevezetett n-elemü vektorokkal foglalkozunk.

Mivel egy vektor komponensei bármely valós értéket felvehetnek, nyilvánvaló, hogy végtelen sok n-elemü vektor van. Ezek szerint az n elemü vektorok összessége végtelen halmaz.

Az n-elemü vektorok halmazát a következőkben L_n szimbólummal jelöljük, ahol n a vektorok komponenseinek számára utal.

A halmazokról tanultak szerint ha

$$\underline{a} \in L_n,$$

akkor az $\underline{a}$ egy n-elemü vektort jelent.

Az előzőekben megismert összefüggések alapján az L_n halmaz az alábbi tulajdonságokkal rendelkezik:

a) Az L_n halmaz elemei között értelmezve van mind az összeadás, mind pedig a skalárral való szorzás müvelete.

Ez annyit jelent, hogy két n-elemü vektor összege, n-elemü vektornak bármely skalárral alkotott szorzata ugyancsak n-elemü vektort ad eredményül.

A most megfogalmazott két állítást matematikai szimbólumokkal igy rögzithetjük:

$$\underline{a}_1, \underline{a}_2 \in L_n \Longrightarrow \underline{a}_1 + \underline{a}_2 \in L_n$$

$$\underline{a}_1 \in L_n \Longrightarrow \lambda\, \underline{a}_1 \in L_n.$$

Ezekből az is következik, hogy az L_n elemeinek bármely lineáris kombinációja ugyancsak eleme az L_n halmaznak, amit röviden igy irhatunk fel:

$$\underline{a}_1, \underline{a}_2, \dots, \underline{a}_k \in L_n \Rightarrow \lambda_1 \underline{a}_1 + \lambda_2 \underline{a}_2 + \dots + \lambda_k \underline{a}_k \in L_n\ .$$

b) Az n-elemü vektorok összeadására és skalárral való szorzására érvényes mind a kommutativitás mind az asszociativitás, mind pedig a disztributivitás.

Közelebbről vizsgálva ezeket a tulajdonságokat ez azt jelenti, hogy a tetszőlegesen kiválasztott

$$\underline{a}_1, \underline{a}_2, \underline{a}_3 \in L_n$$

vektorokra nézve érvényesek az alábbi relációk:

$$\left.\begin{aligned} \underline{a}_1 + \underline{a}_2 &= \underline{a}_2 + \underline{a}_1 \\ \alpha\, \underline{a}_1 &= \underline{a}_1\, \alpha \end{aligned}\right\} \text{kommutativitás}$$

$$\left.\begin{aligned} (\underline{a}_1 + \underline{a}_2) + \underline{a}_3 &= \underline{a}_1 + (\underline{a}_2 + \underline{a}_3) \\ \alpha(\beta\, \underline{a}_1) &= (\alpha\beta)\, \underline{a}_1 \end{aligned}\right\} \text{asszociativitás}$$

$$\left.\begin{aligned} (\alpha + \beta)\, \underline{a}_1 &= \alpha \underline{a}_1 + \beta \underline{a}_1 \\ \alpha(\underline{a}_1 + \underline{a}_2) &= \alpha \underline{a}_1 + \alpha \underline{a}_2 \end{aligned}\right\} \text{disztributivitás}$$

c) Az L_n-nek van nulla eleme is: az n-elemü nullvektor, amelyre nézve bármely $\underline{a}_1 \in L_n$ vektor esetén érvényes az

$$\underline{a}_1 + \underline{0} = \underline{a}_1$$

azonosság.

d) Minden $\underline{a}_1 \in L_n$ elem eleget tesz az

$$1\, \underline{a}_1 + (-1)\, \underline{a}_1 = \underline{0}$$

egyenlőségnek, vagyis az L_n halmaz bármely eleméhez tartalmazza annak ellentettjét.

Ha a fenti tulajdonságokat tekintjük, könnyü belátni, hogy ezeknek nemcsak az n-elemü vektorok tesznek eleget, hanem pl. az n-edrendü mátrixok is.

E közös tulajdonságok teljesitésének kifejezésére uj fogalmat, a lineáris tér fogalmát vezetjük be.

Valamely adott L halmazt lineáris térnek nevezünk, ha azokra a következő feltételek teljesülnek:

a) Az L halmaz elemei között értelmezve van mind az összeadás, mind a skalárral való szorzás müvelete.

b) Az L-ben értelmezett két müvelet rendelkezik mind a kommutativitás, mind az asszociativitás, mind pedig a disztributivitás tulajdonságával.

c) Az L halmaznak van olyan null eleme, amelyet az L halmaz bármely eleméhez hozzáadva, azt változatlanul hagyja.

d) Az L halmaz bármely eleméhez az L halmaznak van olyan eleme, amelyre ezek összege a null elemet adja, vagyis az L halmaz minden elemének van ellentettje.

Példák a lineáris térre:

a) az n-elemü vektorok halmaza,

b) az n-edrendü kvadratikus mátrixok halmaza

c) az azonos tipusu pl. m.n-es mátrixok összessége stb.

A következőkben nem kivánjuk teljes általánosságukban vizsgálni a lineáris tereket, csupán az n-elemü vektorok terére irányitjuk figyelmünket, mert annak ismerete elégséges számunkra a gazdasági problémák vizsgálatához.

Az n-elemü vektorok L_n lineáris terével kapcsolatban n értékétől függően beszélünk az egy-elemü vektorok lineáris teréről, az L_1 lineáris térről, a kételemü vektorok L_2 lineáris teréről stb. Az L_1 lineáris tér nem más, mint a valós számok halmaza, amit szokás a számegyenes pontjaival szemléltetni. Ilyen "geometriai megfeleltetés" lehetséges az L_2 lineáris térre is, amelyet a sik pontjai reprezentálhatnak, vagy az L_3 lineáris térre, amelyet a mindennapi szóhasználat szerinti pontok szemléltethetnek. Mivel a többelemü vektorok ilyen értelmü ábrázolása lehetetlen, látható, hogy az n-elemü vektorok lineáris terénél a "tér" kifejezés csupán átvitt értelemben használatos.

A lineáris térrel kapcsolatban bevezetjük az "altér" fogalmát.

Ha az n-elemü vektorok L_n lineáris teréből az

$$\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_k$$

vektorokat kiemeljük és azok összes lehetséges

$$x_1 \underline{a}_1 + x_2 \underline{a}_2 + \ldots + x_k \underline{a}_k = \underline{s}$$

lineáris kombinációit tekintjük, ahol $x_1, x_2, \ldots, x_k$ valós számok, ismét n-elemü vektorok végtelen halmazát nyerjük. E halmazt L'-vel jelöljük és az L_n lineáris tér alterének nevezzük.

Az L' halmaz egyszerübb leirására az $\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_k$ vektorokat tekintsük egy $\underline{A}$ mátrix oszlopvektorainak, vagyis:

$$\underline{A} = [\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_k],$$

az $x_1, x_2, \ldots, x_k$ skalárokat pedig egy $\underline{x}$ vektor komponenseinek.

Igy az

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

kifejezés éppen az $\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_k$ vektorok lineáris kombinációt szimbolizálja. Az összes lehetséges $\underline{s}$ vektor alkotta L' halmazt ezek szerint a következő formában irhatjuk:

$$\underline{L'} = \{\underline{s} / \underline{s} = \underline{A}\,\underline{x}\},$$

s igy olvassuk: az L' mindazon $\underline{s}$ vektorok halmazát jelenti, amelyek az $\underline{A}$ mátrix oszlopvektorainak összes lehetséges lineáris kombinációival állithatók elő. (Egy-egy halmaz definiálására a későbbiekben is gyakran fogunk ehhez hasonló irásmódhoz folyamodni. A jelölésnél először feltüntetjük a halmaz elemeinek általános szimbólumát, azután függőleges vonást huzunk, és végül megadjuk a halmaz elemeinek képzési módját.)

1. Tétel: Az L' altér részhalmaza az L_n lineáris térnek, vagyis $L' \subset L_n$.

Bizonyitás: Az állitás szerint az L' halmaz minden eleme egyuttal eleme az L_n lineáris térnek is. Az L_n-re adott definició alapján, ha

$$\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_k \in L_n,$$

akkor

$$\underline{s} = x_1 \underline{a}_1 + x_2 \underline{a}_2 + \ldots + x_k \underline{a}_k \in L_n$$

is fennáll, ami éppen a tételt bizonyitja.

2. Tétel: Az L' altér önmagában is lineáris teret képez.

Bizonyitás: Annak belátására, hogy L' altér lineáris tér, ki kell mutatnunk, hogy L' halmazra teljesülnek a lineáris térre kikötött feltételek. E végből

a) legyen $\underline{s}_1$ és $\underline{s}_2$ két tetszőleges eleme az L'-nek, azaz teljesüljön az $\underline{s}_1, \underline{s}_2 \in L'$ feltétel. E feltétel szerint található olyan $\underline{x}_1$ és $\underline{x}_2$ vektor, amelyre nézve

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

ill.

$$\underline{s}_2 = \underline{A}\,\underline{x}_2.$$

A müveleti szabályok értelmében

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

ami azt jelenti, hogy

$$\underline{s}_1 + \underline{s}_2 \in L'.$$

Másrészt

$$\alpha\,\underline{s}_1 = \alpha(\underline{A}\,\underline{x}_1) = \underline{A}\,(\alpha\,\underline{x}_1),$$

ami szerint

$$\alpha\,\underline{s}_1 \in L'.$$

Igy az L' halmaz elemei között értelmezve van mind az összeadás, mind pedig a skalárral való szorzás.

b) Abból a tényből, hogy L' részhalmaza L_n-nek, következik, hogy az L' vektorainak összegére és skalárral való szorzatára nézve érvényes a kommutativitás, az asszociativitás és a disztributivitás.

c) S mivel $\underline{x} = \underline{o}$ esetén

$$\underline{s} = \underline{A}\,.\,\underline{o} = \underline{o},$$

világos, hogy az L' is tartalmaz nullelemet, azaz $\underline{o} \in L'$.

d) Végül pedig bármely $\underline{s} \in L'$ vektorra nézve érvényes az

$$1 \cdot \underline{s} = 1(\underline{A}\,\underline{x}) = (1\,\underline{A}) \cdot \underline{x} = \underline{A}\,\underline{x} = \underline{s},$$

azaz

$$1\,\underline{s} = \underline{s}$$

illetve

$$1 \cdot \underline{s} - \underline{s} = \underline{o}$$

kikötés is.

Mindezekből látható, hogy az L' valóban eleget tesz a lineáris térrel szemben támasztott négy követelménynek. Az L' tér tehát valóban lineáris tér.

Az L'-t az $\underline{a}_1, \underline{a}_2, \dots, \underline{a}_k$ vektorok által kifeszitett (vagy generált) altérnek is szoktuk nevezni.

Megjegyzések:

a) Ha az $\underline{a}_1, \underline{a}_2, \dots, \underline{a}_k$ vektorok által meghatározott $\underline{A}$ mátrix nullmátrix, akkor az L' egyetlen vektort tartalmaz, a nullvektort. Ezt igy jelöljük:

$$L' = \{\underline{0}\}.$$

Az igy nyert altér az un. zérusaltér. A zérus-alteret nem valódi altérnek tekintjük.

b) Nem valódi altérről beszélünk akkor is, ha az L_n minden vektora L'-nek is vektora és megforditva.

c) Az L'-t az L_n valódi alterének hivjuk, ha az L' olyan lineáris tér, amelyre nézve

$$L' \subset L_n,$$

de

$$L' \neq L_n \quad \text{és} \quad L' \neq \{\underline{0}\}.$$

d) Az L' mindig részhalmaza L_n-nek, de L_n-nek nem minden részhalmaza altér.

e) Az L' altér generálásának nem egyetlen módja az itt adott eljárás.

Ezek után tekintsük pl. az L_3 lineáris teret, a három-elemü vektorok terét.

1. Példa: Határozzuk meg az $\underline{e}_1$, $\underline{e}_2 \in L_3$ vektorok által generált alteret. A generáló vektorokhoz tartozó $\underline{A}$ mátrix

$$[\underline{e}_1, \underline{e}_2] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix},$$

igy

$$\underline{s} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ 0 \end{bmatrix},$$

ahonnan

$$L' = \{\underline{s} \mid \underline{s} = \underline{A}\ \underline{x}\}.$$

Az L', tehát mindazon három-elemü vektorokat tartalmazza, amelyek harmadik komponense zérus.

2. Példa: Határozzuk meg az $\underline{e}_1 = [1, 0, 0]^*$ vektor által kifeszitett alteret!
Mivel az alteret generáló mátrix itt egyetlen oszlopot tartalmaz, az L' altér elemeit az $\underline{s} = x\,\underline{e}_1$ vektorok képezik, amelyeknek második és harmadik komponense zérus.

Emlitettük, hogy az L_3 tér vektorait a közönséges értelemben vett tér pontjai szemléltethetik. Igy az

$$\underline{a} = [a_1, a_2, a_3]^*$$

vektor geometriai képét a térbeli koordináta-rendszer felvétele esetén a következő módon ábrázolhatjuk:

1. Példa: Határozzuk meg az $\underline{e}_1$, $\underline{e}_2 \in L_3$ vektorok által generált alteret. A generáló vektorokhoz tartozó $\underline{A}$ mátrix

$$[\underline{e}_1, \underline{e}_2] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix},$$

igy

$$\underline{s} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ 0 \end{bmatrix},$$

ahonnan

$$L' = \{ \underline{s} \mid \underline{s} = \underline{A}\ \underline{x} \}.$$

Az L', tehát mindazon három-elemü vektorokat tartalmazza, amelyek harmadik komponense zérus.

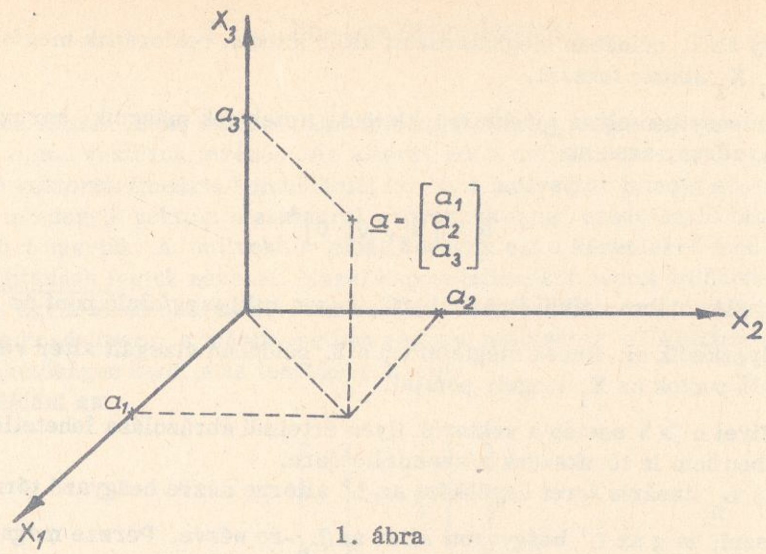
2. Példa: Határozzuk meg az $\underline{e}_1 = [1, 0, 0]^*$ vektor által kifeszitett alteret!

Mivel az alteret generáló mátrix itt egyetlen oszlopot tartalmaz, az L' altér elemeit az $\underline{s} = x\ \underline{e}_1$ vektorok képezik, amelyeknek második és harmadik komponense zérus.

Emlitettük, hogy az L_3 tér vektorait a közönséges értelemben vett tér pontjai szemléltethetik. Igy az

$$\underline{a} = [a_1, a_2, a_3]^*$$

vektor geometriai képét a térbeli koordináta-rendszer felvétele esetén a következő módon ábrázolhatjuk:

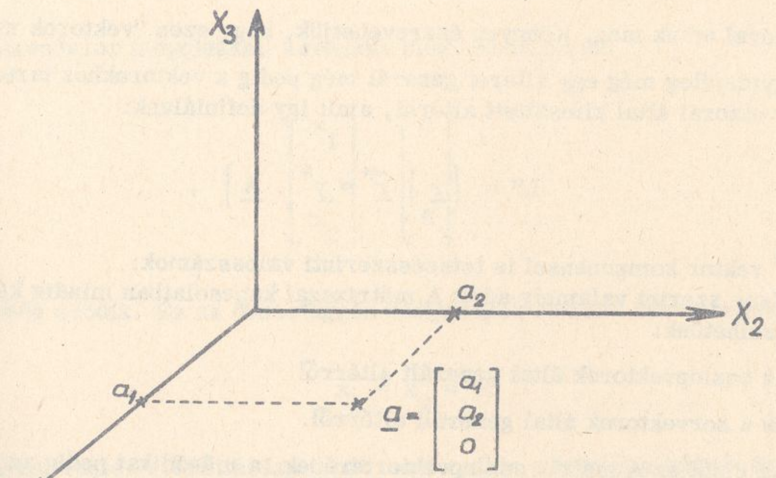


1. ábra

Ha történetesen az $\underline{a}$ vektor harmadik komponense zérus, azaz

$$\underline{a} = [a_1, a_2, 0]^*,$$

akkor a vektornak megfelelő pont az alábbi ábra szerint az X_1X_2 tengelyek által meghatározott síkban van.



2. ábra

Igy az 1. példában meghatározott altér minden vektorának megfelelő pont is az X_1 X_2 sikban fekszik.

Amennyiben olyan $\underline{a}$ vektort tekintünk, amelynek második, harmadik komponense zérus, azaz az

$$\underline{a} = [a_1, 0, 0]^*$$

vektort, akkor ábra nélkül is belátható, hogy a neki megfelelő pont az X_1 tengelyen helyezkedik el. Ennek megfelelően a 2. példában vizsgált altér vektorainak megfelelő pontok az X_1 tengely pontjai.

Mivel $n > 3$ esetén a vektorok ilyen értelmü ábrázolása lehetetlen, a későbbiekben nem is törekszünk a szemléltetésre.

Az L_n lineáris teret egyébként az L' altérre nézve beágyazó térnek is szoktuk nevezni, mig az L' beágyazott altér az L_n-re nézve. Persze maga a beágyazó tér is lehet altere, még pedig valódi altere is egy harmadik lineáris térnek. Amikor tehát altérről beszélünk, valójában mindig eleve feltételezünk egy beágyazó lineáris teret.

Az $\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_k \in L_n$ vektorok által kifeszitett alteret

$$L' = \{\underline{s} \mid \underline{s} = \underline{A}\ \underline{x}\}$$

definicióval adtuk meg. Könnyen észrevehetjük, hogy ezen "vektorok rendszere"[1] egyidejüleg még egy alteret generál még pedig a vektorokhoz tartozó mátrix sorvektorai által kifeszitett alteret, amit így definiálunk:

$$L'' = \{\underline{r} \mid \underline{r}^* = \underline{y}^* \cdot \underline{A}\} ,$$

ahol $\underline{y}^*$ vektor komponensei is tetszésszerinti valósszámok.

Ezek szerint valamely adott $\underline{A}$ mátrixszal kapcsolatban mindig két altérről beszélhetünk:

az oszlopvektorok által generált altérről

és a sorvektorok által generált altérről.

Az elsőt az $\underline{A}$ mátrix <u>oszlopvektorterének</u>, a másodikat pedig az $\underline{A}$ mátrix <u>sorvektorterének</u> nevezzük.

---

[1] Vektorrendszeren mi mindig valamely, lineáris tér vektorainak véges halmazát fogjuk érteni.

2.2. A lineáris függetlenség

Mint láttuk, adott vektorok lineáris kombinációi fontos szerepet töltenek be az n-elemü vektorok terében. Az altérre adott definició szerint, az alteret generáló vektorok lineáris kombinációi között a nullvektor mindig szerepel. Ha ugyanis mindegyik vektort o skalárral szorozzuk meg, eredményül biztosan a nullvektort nyerjük. A nullvektor előállitásának ezt a kézenfekvő módját triviális előállitásnak fogjuk nevezni. Ezzel kapcsolatban két esetet különböztethetünk meg. Az egyik esetben a nullvektor előállitására csakis a triviális lehetőség áll rendelkezésünkre; a másik esetben pedig a nullvektor előállitására a triviális lehetőségen kivül más lehetőség is van.

Például az

$$x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

lineáris kombinációt vizsgálva, állapitsuk meg, hogy milyen (x_1, x_2) értékpár mellett szolgáltathat ez nullvektort, vagyis

$$x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

egyenletrendszer megoldásait keressük meg. Ahonnan az

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

egyenlőség adódik. Ez az összefüggés pedig akkor és csak akkor teljesülhet, ha

$$x_1 = x_2 = 0.$$

A nullvektor előállitására tehát most csakis a triviális lehetőség áll rendelkezésünkre.

Ha azonban az

$$x_1 \begin{bmatrix} 2 \\ -1 \end{bmatrix} + x_2 \begin{bmatrix} -4 \\ 2 \end{bmatrix}$$

lineáris kombinációt tekintjük, amely a müveleti szabályoknak megfelelően az

$$x_1 \begin{bmatrix} 2 \\ -1 \end{bmatrix} + x_2 \begin{bmatrix} -4 \\ 2 \end{bmatrix} = \begin{bmatrix} 2x_1 \\ -x_1 \end{bmatrix} + \begin{bmatrix} -4x_2 \\ 2x_2 \end{bmatrix} = \begin{bmatrix} -2(-x_1 + 2x_2) \\ -x_1 + 2x_2 \end{bmatrix}$$

vektorhoz vezet, már nemcsak a triviális megoldással nyerhetünk nullvektort. Mivel a kapott vektor első komponense a második komponens "-2"-szerese, azért ez a lineáris kombináció minden olyan esetben nullvektort ad, amikor teljesül a

$$-x_1 + 2x_2 = 0$$

egyenlőség. Ennek az egyenletnek azonban végtelen sok megoldása van. A triviálison kivül más lehetőség is van a nullvektor előállitására. Ilyen megoldás pl. az $x_1 = 8$ és $x_2 = 4$, amelyekre a kérdéses lineáris kombináció valóban nullvektort ad.

Ennek belátása alapján bevezetjük a lineárisan független vektorok fogalmát.

Az L_n lineáris tér $\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_k$ vektorait lineárisan független vektoroknak nevezzük, ha lineáris kombinációjuk révén a nullvektort csakis a triviális módon tudjuk előállitani.

Ezt a meghatározást igy is megfogalmazhatjuk: az

$$\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_k$$

vektorokat lineárisan független vektoroknak nevezzük, ha az

$$x_1\underline{a}_1 + x_2\underline{a}_2 + \ldots + x_k\underline{a}_k = \underline{0}$$

egyenlet akkor és csak akkor állhat fenn, ha az

$$x_1 = x_2 = \ldots = x_k = 0$$

feltétel teljesül. Ha az $\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_k$ vektorok csak zérus skalár szorzókkal képzett lineáris kombinációval állitják elő a nullvektort, az $\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_k$ vektorok lineárisan független rendszert képeznek, ellenkező esetben azt mondjuk, hogy az adott vektorok lineárisan függő rendszert alkotnak.

Az adott definició szerint a vizsgált

$$[1, \ 0] \quad \text{és} \quad [0, \ 1]$$

vektorok lineárisan független rendszert alkotnak, a

$$[2, \ -1] \quad \text{és} \quad [-4, \ 2]$$

vektorok pedig lineárisan függő vektorok.

A továbbiakban az L_n valamely $\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_k$ vektorrendszerének vizsgálatánál elsődleges lesz annak eldöntése, hogy a kérdéses vektorok lineárisan függetlenek-e vagy nem.

A bevezetett uj fogalommal kapcsolatban tekintsük a következő állitásokat:

1. Tétel: Ha az $\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_k$ vektorok között a nullvektor is szerepel, akkor ezek a vektorok lineárisan függő rendszert alkotnak.

Bizonyitás: Ha történetesen az $\underline{a}_1 = \underline{0}$, akkor az

$$x_1 \underline{a}_1 + x_2 \underline{a}_2 + \ldots + x_k \underline{a}_k = \underline{0}$$

egyenlőség bármely x_1 skalár mellett fennáll, mivel

$$x_1 \cdot \underline{0} + 0 \cdot \underline{a}_2 + \ldots + 0 \cdot \underline{a}_k = \underline{0}.$$

Igy az adott vektorok nem lehetnek lineárisan független vektorok. Az állitást ugyis megfogalmazhatjuk, hogy lineárisan független vektorok között a nullvektor nem szerepelhet.

2. Tétel: Lineárisan független vektorok minden nem-üres részhalmaza ugyancsak lineárisan független rendszert alkot.

Az állitás helyességét indirekt bizonyitással láthatjuk be.

Feltesszük, hogy az $\underline{a}_1, \underline{a}_2, \dots, \underline{a}_k$ lineárisan független vektorok közül egyiket pl. az $\underline{a}_1$ vektort elhagyva, az $\underline{a}_2, \underline{a}_3, \dots, \underline{a}_k$ un. maradékrendszer lineárisan függő, azaz az

$$x_2 \underline{a}_2 + x_3 \underline{a}_3 + \dots + x_k \underline{a}_k = \underline{0}$$

egyenletnek létezik a triviálistól különböző megoldása is. Aztán az elhagyott $\underline{a}_1$ vektor $x_1 = 0$-szorosát hozzácsatoljuk a fenti egyenlethez:

$$x_1 \underline{a}_1 + \left[x_2 \underline{a}_2 + \dots + x_k \underline{a}_k \right] = \underline{0}$$

Eszerint azonban az $\underline{a}_1, \dots, \underline{a}_k$ vektorok lineárisan függő rendszert adnának, ami ellentmond a tétel állításának.

Mivel a feltétel ellentmondáshoz vezetett, annak ellenkezője igaz, vagyis a maradékrendszer is lineárisan független rendszert alkot. Ugyanez vonatkozik az előbbi maradékrendszerre és az ebből nyerhető további maradék rendszerekre is. Igy nyilvánvaló állításunk helyessége.

Következmény:

Az egyetlen $\underline{a}$ vektorból álló rendszert lineárisan független rendszernek tekintjük, ha

$$\underline{a} \neq \underline{0}.$$

Ezért kellett a 2. tételben kikötnünk, hogy az állitás csak a visszamaradó "nem-üres" részhalmazra vonatkozik.

3. Tétel: Az $\underline{a}_1, \underline{a}_2, \dots, \underline{a}_k, \underline{b}$ vektorrendszer akkor és csak akkor lineárisan függőrendszer, ha legalább egyikük előállitható a többi lineáris kombinációjaként.

A tétel bizonyitását megkönnyiti, ha észrevesszük, hogy "megfordítható tételről" van szó. Az állitás helyessége ennek alapján két lépésben látható be.

a) Ha pl. a $\underline{b}$ vektor felirható, mint az $\underline{a}_1, \underline{a}_2, \dots, \underline{a}_k$ vektorok lineáris kombinációja, akkor az $\underline{a}_1, \underline{a}_2, \dots, \underline{a}_k, \underline{b}$ vektorok lineárisan függő rendszert képeznek.

A feltétel értelmében ugyanis

$$\underline{b} = x_1 \underline{a}_1 + x_2 \underline{a}_2 + \dots + x_k \underline{a}_k,$$

azaz

$$x_1 \underline{a}_1 + x_2 \underline{a}_2 + \ldots + x_k \underline{a}_k - 1 \cdot \underline{b} = \underline{0}.$$

Az utolsó egyenlőségben azonban az $\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_k, \underline{b}$ vektorok nem mind zérus skalárok segitségével állitották elő a nullvektort. Ezek szerint a kérdéses vektorok valóban lineárisan függő rendszert képeznek.

b) Másrészt, ha az $\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_k, \underline{b}$ vektorok lineárisan függő vektorok, akkor legalább az egyikük előállitható a többi lineáris kombinációjaként. Ebben az esetben ugyanis az

$$y_1 \underline{a}_1 + y_2 \underline{a}_2 + \ldots + y_k \underline{a}_k + y_{k+1} \cdot \underline{b} = \underline{0}$$

egyenletnek van a triviálistól különböző megoldása is. (Ha ilyen megoldás nem létezék, akkor az adott vektorok lineárisan független vektorok lennének, ami ellentmond a feltételünknek.)
Legyen pl. az $y_{k+1} \neq 0$, akkor a fenti egyenletből nyert

$$y_{k+1} \underline{b} = -y_1 \underline{a}_1 - y_2 \underline{a}_2 - \ldots - y_k \underline{a}_k$$

egyenletben az

$$y_{k+1} \neq 0$$

skalárral osztva, a

$$\underline{b} = -\frac{y_1}{y_{k+1}} \cdot \underline{a}_1 - \frac{y_2}{y_{k+1}} \cdot \underline{a}_2 - \ldots - \frac{y_k}{y_{k+1}} \underline{a}_k$$

összefüggéshez jutunk. Ha pedig bevezetjük a

$$-\frac{y_1}{y_{k+1}} = x_1, \quad -\frac{y_2}{y_{k+1}} = x_2, \quad \ldots, \quad -\frac{y_k}{y_{k+1}} = x_k$$

jelöléseket, akkor

$$\underline{b} = x_1 \underline{a}_1 + x_2 \underline{a}_2 + \ldots + x_k \underline{a}_k,$$

vagyis a $\underline{b}$ vektor kifejezhető a többi lineáris kombinációjaként, ahogy állitottuk.

Tételünk nem állitja, hogy egy lineárisan függő rendszer minden vektora kifejezhető a többi lineáris kombinációjaként, csupán azt mondja ki, hogy a lineárisan függő vektorok között legalább egy olyan vektor található, amelyre érvényes ez a tulajdonság.

2.3. Vektorrendszerek rangja

Az L' alteret generáló $\underline{a}_1, \underline{a}_2, \dots, \underline{a}_k$ vektorok alkotta vektorrendszert vizsgáljuk ujabb szempont szerint.

Mint már emlitettük:

Vektorrendszeren az adott $\underline{L_n}$ lineáris tér vektorainak valamely véges halmazát értjük.

A vektorrendszerek legfontosabb jellemzője a rang nagysága.

Egy vektorrendszer rangja r, ha a kérdéses vektorrendszerből feltétlenül kiválasztható r számu lineárisan független vektor, a vektorrendszer bármely r + 1 számu vektora azonban már lineárisan függő rendszert alkot.

Például az

$$\underline{a}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \underline{a}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \underline{a}_3 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

vektorrendszer rangja 2, mert ha az $\underline{a}_1$ és $\underline{a}_2$ vektorokat tekintjük, azok lineárisan függetlenek, ahogy ezt korábban már beláttuk, de az $\underline{a}_1, \underline{a}_2, \underline{a}_3$ már függő rendszert alkot, mivel

$$\underline{a}_3 = 1 \cdot \underline{a}_1 + 2\,\underline{a}_2.$$

Megjegyzések: a) A rang fogalma bizonyos maximum tulajdonsággal rendelkezik.

b) Ha történetesen a vektorrendszer minden vektora nullvektor, akkor a rangot 0-nak tekintjük, azaz r = 0. Minden egyéb esetben természetes számot jelent a rang.

Ezek után tekintsünk erre az uj fogalomra vonatkozóan néhány állitást.

1. Tétel: Ha az $\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_k$ vektorrendszer rangja r, és a belőle kiválasztott $\underline{a}_{j_1}, \underline{a}_{j_2}, \ldots, \underline{a}_{j_r}$ vektorok történetesen lineárisan függetlenek, akkor a vektorrendszer bármely vektora előállitható ennek az r lineárisan független vektornak lineáris kombinációjaként, méghozzá egyértelmüen.

Bizonyitás: Ennek belátására induljunk ki a következő meggondolásból: mivel az adott vektorok sorrendje a rang szempontjából teljesen közömbös, a vektorrendszer vektorai tetszés szerint átrendezhetők, azaz minden további nélkül feltehetjük, hogy

$$\underline{a}_{j_1} = \underline{a}_1 \; ; \; \underline{a}_{j_2} = \underline{a}_2, \ldots, \underline{a}_{j_r} = \underline{a}_r.$$

Ez annyit jelent, hogy a vektorrendszer átrendezése után, a kiválasztott r lineárisan független vektor éppen egybeesik az uj sorrendben felirt vektorrendszer első r vektorával.

Ilyen megállapodás mellett természetes, hogy az első r vektor előállitható az $\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_r$ vektorok lineáris kombinációjaként.

Igy

$$\underline{a}_1 = 1 \cdot \underline{a}_1 + 0 \cdot \underline{a}_2 + \ldots + 0 \cdot \underline{a}_r$$
$$\underline{a}_2 = 0 \cdot \underline{a}_1 + 1 \cdot \underline{a}_2 + \ldots + 0 \cdot \underline{a}_r$$
$$\vdots$$
$$\underline{a}_r = 0 \cdot \underline{a}_1 + 0 \cdot \underline{a}_2 + \ldots + 1 \cdot \underline{a}_r.$$

Könnyü azonban belátni, hogy az $\underline{a}_{r+1}, \underline{a}_{r+2}, \ldots, \underline{a}_k$ bármelyike is felirható az $\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_r$ vektorok lineáris kombinációjaként. Legyen $\underline{a}_j$ az $\underline{a}_{r+1}, \underline{a}_{r+2}, \ldots, \underline{a}_k$ vektorok valamelyike.

Mivel a vektorrendszer rangja r, az r+1 vektorból álló

$$\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_r, \underline{a}_j$$

rendszer feltétlenül lineárisan függő. Ez azt jelenti, hogy az

$$x_1 \underline{a}_1 + x_2 \underline{a}_2 + \ldots + x_r \underline{a}_r + x_j \underline{a}_j = \underline{0}$$

egyenletnek van nem-triviális megoldása. Ilyen nem triviális megoldásban az x_j skalár értéke biztosan különbözik a 0-tól. Az $x_j = 0$ esetén ugyanis az

$$x_1 \underline{a}_1 + x_2 \underline{a}_2 + \dots + x_r \underline{a}_r = \underline{0}$$

egyenletnek kellene rendelkeznie nem-triviális megoldással, ami ellentmond a kiválasztott vektorok lineáris függetlenségének.

Bármely nem-triviális megoldás esetén tehét $x_j \neq 0$ és az

$$x_1 \underline{a}_1 + x_2 \underline{a}_2 + \dots + x_r \underline{a}_r + x_j \underline{a}_j = \underline{0}$$

egyenlet $\underline{a}_j$-re megoldható, azaz

$$\underline{a}_j = \left(-\frac{x_1}{x_j}\right) \underline{a}_1 + \left(-\frac{x_2}{x_j}\right) \underline{a}_2 + \dots + \left(-\frac{x_r}{x_j}\right) \underline{a}_r .$$

Ezzel bebizonyitottuk tételünk első részét.

Az előállítás egyértelmüsége azt jelenti, hogy a kiválasztott r darab lineárisan független vektornak csak egyetlen olyan lineáris kombinációja van, amely az adott vektorrendszer valamely $\underline{a}_i$ vektorát előállitja.

Tegyük ugyanis fel, hogy $\underline{a}_i$ vektor az $\underline{a}_1, \underline{a}_2, \dots, \underline{a}_r$ vektorok lineáris kombinációjaként kétféleképpen is felirható:

$$\underline{a}_i = y_1 \underline{a}_1 + y_2 \underline{a}_2 + \dots + y_r \underline{a}_r$$

$$(i = 1, 2, \dots, r, r+1, \dots, k).$$

$$\underline{a}_i = z_1 \underline{a}_1 + z_2 \underline{a}_2 + \dots + z_r \underline{a}_r$$

Ha az első egyenletből kivonjuk a másodikat, azt kapjuk, hogy

$$\underline{0} = (y_1 - z_1) \underline{a}_1 + (y_2 - z_2) \underline{a}_2 + \dots + (y_r - z_r) \underline{a}_r .$$

Mivel a kiválasztott $\underline{a}_1, \underline{a}_2, \dots, \underline{a}_r$ vektorok lineárisan függetlenek, a nullvektort ezek csak a triviális módon állitják elő, vagyis

$$y_1 - z_1 = y_2 - z_2 = \dots = y_r - z_r = 0,$$

azaz

$$y_1 = z_1, \quad y_2 = z_2, \ldots, y_r = z_r.$$

Ez azt jelenti, hogy az $\underline{a}_i$ vektornak az adott feltételek mellett nem lehet két különböző előállitása, amivel tételünk második részét is bebizonyitottuk.

2. Tétel: Egy vektorrendszer rangja akkor és csak akkor nem változik meg, ha olyan vektort veszünk el belőle, vagy adunk hozzá, amely kifejezhető a többi vektor lineáris kombinációjaként.

Bizonyitás: Az állitás helyességét két lépésben láthatjuk be.

a) Először kimutatjuk, ha egy vektorrendszerből olyan vektort veszünk el, amely kifejezhető a többi vektor lineáris kombinációjaként, akkor a vektorrendszer rangja változatlan marad.

Legyen ugyanis az adott

$$\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_r, \underline{a}_{r+1}, \ldots, \underline{a}_k$$

vektorrendszer rangja r és az egyszerüség kedvéért azt is tegyük fel, hogy az első r vektor lineárisan független.

Igy ha az $\underline{a}_{r+1}, \underline{a}_{r+2}, \ldots, \underline{a}_k$ vektorok valamelyikét (amelyek az előző tétel értelmében előállithatók az első r vektor lineáris kombinációjaként) hagyjuk el, a rang nyilván változatlan marad.

Ezért a tételben foglalt állitást elég az első r vektorra nézve megvizsgálni.

Tegyük fel, hogy például az $\underline{a}_1$ kifejezhető a többi vektor lineáris kombinációjaként, azaz

$$\underline{a}_1 = x_2 \underline{a}_2 + \ldots + x_r \underline{a}_r + \ldots + x_k \underline{a}_k.$$

Ezek után azt kell kimutatni, hogy az

$$\underline{a}_2, \ldots, \underline{a}_r, \ldots, \underline{a}_k$$

maradékrendszer rangja is r, vagyis a rang nem változott. Itt is indirekt módon bizonyitunk.

Feltesszük, hogy az uj rendszer rangja r - 1. (A rang ugyanis vagy maradt r, vagy lecsökkent r-1-re.) Ez az előző tétel értelmében azt jelenti, hogy az

$$\underline{a}_2, \ldots, \underline{a}_r, \ldots, \underline{a}_k$$

vektorrendszer minden vektora előállitható a rendszer r-1 független vektorának lineáris kombinációjaként.

Tekintsük az

$$\underline{a}_2, \underline{a}_3, \ldots, \underline{a}_r$$

r-1 független vektort, akkor a maradékrendszer vektorait előállitó lineáris kombinációk:

$$\underline{a}_2 = 1 \cdot \underline{a}_2 + 0 \cdot \underline{a}_3 + \ldots + 0 \cdot \underline{a}_r$$

$$\vdots$$

$$\underline{a}_r = 0 \cdot \underline{a}_2 + 0 \cdot \underline{a}_3 + \ldots + 1 \cdot \underline{a}_r$$

$$\underline{a}_{r+1} = s_2 \underline{a}_2 + s_3 \underline{a}_3 + \ldots + s_r \underline{a}_r$$

$$\vdots$$

$$\underline{a}_k = y_2 \underline{a}_2 + y_3 \underline{a}_3 + \ldots + y_r \underline{a}_r$$

Ebből viszont az következik, hogy

$$\underline{a}_1 = x_2 \left[1 \cdot \underline{a}_2 + 0 \cdot \underline{a}_3 + \ldots + 0 \cdot \underline{a}_r\right] + \ldots + x_r \left[0 \cdot \underline{a}_2 + 0 \cdot \underline{a}_3 + \ldots + 1 \cdot \underline{a}_r\right] +$$

$$+ x_{r+1} \left[s_2 \underline{a}_2 + s_3 \underline{a}_3 + \ldots + s_r \underline{a}_r\right] + \ldots + x_k \left[y_2 \underline{a}_2 + y_3 \underline{a}_3 + \ldots + y_r \underline{a}_r\right],$$

ahonnan egyszerü átalakitással nyerhetjük az

$$\underline{a}_1 = (x_2 + x_{r+1} s_2 + \ldots + x_k y_2)\, \underline{a}_2 + (x_3 + x_{r+1} s_3 + \ldots + x_k y_3)\, \underline{a}_3 + \ldots +$$

$$+ (x_r + x_{r+1} s_r + \ldots + x_k y_r)\, \underline{a}_r$$

egyenlőséget. Ez azt jelenti, hogy az $\underline{a}_1$-et is ki lehet fejezni az $\underline{a}_2, \underline{a}_3, \ldots, \underline{a}_r$ vektorok lineáris kombinációjaként, ami ellentétben áll azzal a feltevésünkkel, hogy az

$$\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_r$$

vektorok, lineárisan független rendszert alkotnak.

Mivel ellentmondáshoz jutottunk, kiinduló feltételünk helytelen volt, tehát ellenkezője igaz, vagyis ha $\underline{a}_1$ előállitható, akkor annak elhagyása után a maradékrendszer rangja megegyezik az eredeti vektorok rangjával.

b) A bizonyitás második részében könnyen belátható az is, hogy ha az

$$\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_r, \ldots, \underline{a}_k$$

vektorrendszerből $\underline{a}_1$ vektort elhagyva az

$$\underline{a}_2, \ldots, \underline{a}_r, \ldots, \underline{a}_k$$

maradékrendszer rangja nem változik, akkor az $\underline{a}_1$ előállitható a többi lineáris kombinációjaként.

Legyen ugyanis adva az

$$\underline{a}_2, \ldots, \underline{a}_r, \ldots, \underline{a}_k$$

vektorrendszer, amelynek a rangja r, továbbá az

$$\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_r, \ldots, \underline{a}_k$$

un. bővitett vektorrendszer, amelynek a rangja ugyancsak r.

Ha az

$$\underline{a}_2, \ldots, \underline{a}_r, \ldots, \underline{a}_k$$

vektorrendszerben történetesen az

$$\underline{a}_2, \ldots, \underline{a}_r, \underline{a}_{r+1}$$

független rendszert képez, akkor ezek az

$$\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_r, \ldots, \underline{a}_k$$

rendszerben is függetlenek. Mivel pedig az uj rendszer rangja is r, ezért az 1. rang tétel értelmében az $\underline{a}_1$ vektor is felirható az

$$\underline{a}_2, \dots, \underline{a}_{r+1}$$

vektorok lineáris kombinációjaként, ami éppen azt jelenti, hogy ha a maradékrendszer rangja megegyezik az eredeti rendszer rangjával, akkor az elhagyott vektor kifejezhető a többi lineáris kombinációjával.

Az elmondottak egyben azt is bizonyitják, ha egy vektorrendszerhez olyan vektort csatolunk, amely kifejezhető a többi vektor lineáris kombinációjaként, akkor a vektorrendszer rangja változatlan marad.

Következmény: Ha az

$$\underline{a}_1, \underline{a}_2, \dots, \underline{a}_k$$

vektorrendszer vektorai mind előállithatók a

$$\underline{b}_1, \underline{b}_2, \dots, \underline{b}_p$$

vektorrendszer vektorainak lineáris kombinációjaként, akkor az $\underline{a}_1, \underline{a}_2, \dots, \underline{a}_k$ rendszer rangja legfeljebb p.

Bizonyitás:

Tekintsük ugyanis az

$$\underline{a}_1, \underline{a}_2, \dots, \underline{a}_k, \underline{b}_1, \underline{b}_2, \dots, \underline{b}_p$$

bővitett rendszert. Ennek rangja nyilván nem lehet kisebb mint az $\underline{a}_1, \underline{a}_2, \dots, \underline{a}_k$ rendszer rangja. Ha viszont a bővitett rendszerből egymásután elhagyjuk az $\underline{a}_1, \underline{a}_2, \dots, \underline{a}_k$ vektorokat, akkor a 2. rangtétel értelmében a rang változatlan marad. Igy a bővitett rendszer rangja megegyezik a $\underline{b}_1, \underline{b}_2, \dots, \underline{b}_p$ rendszer rangjával. S mivel ez utóbbi legfeljebb p, nyilvánvaló állitásunk helyessége.

2.4. Dimenzió és bázis

Az eddigiek során a lineáris függetlenség problémáit csupán a vektorrendszerekre vizsgáltuk. A továbbiakban ezt a vizsgálatot kiterjesztjük az egész lineáris térre, amelyben a vektorrendszerek csak részhalmazt képeznek.

A lineáris térben a vektorrendszerrel kapcsolatban értelmezett rang szerepét a dimenzió veszi át.

Egy adott lineáris teret n-dimenziósnak nevezünk, ha abban található n lineárisan független vektor, de bárhogyan is választunk ki belőle n + 1 vektort, azok mindig lineárisan függő rendszert alkotnak.

Mivel az altér is lineáris tér, a dimenzió fenti definiciója arra is érvényes.

A legkisebb dimenzióju lineáris tér a zérusaltér. Ebben az altérben nem vehető fel lineárisan független vektor, amit ugy is mondhatunk, hogy zérus számu független vektor található benne. Igy a zérusaltér dimenzióját 0-nak tekintjük. Tárgyalásainkban a zérusaltérnek a későbbiekben nem lesz szerepe.

Számunkra azok a lineáris terek fontosak, amelyeknek dimenziója természetes szám. Megjegyezzük azonban, hogy a lineáris terek elmélete kiterjed a végtelen dimenzióju lineáris terekre is, de mi ezekkel nem foglalkozunk.

A "dimenzió" kifejezés, geometriából, ill. fizikából már ismert, itt azonban értelmezése általánosabb.

A dimenzió egyébként a lineáris terek legfontosabb adata.

A dimenzió fogalmához szorosan kapcsolódik a bázis fogalma.

Az n-dimenziós lineáris tér bármely, n lineárisan független vektorból álló vektorrendszerét e tér bázisának nevezzük. A bázis vektorait bázisvektoroknak hivjuk.

Ennek alapján bárhogyan is veszünk fel az n-dimenziós lineáris térben n lineárisan független vektort, azok bázist alkotnak ebben a térben. Az elnevezés a bázis alapvető tulajdonságára utal.

1. Tétel: A lineáris tér bármelyik vektora kifejezhető a bázisvektorok lineáris kombinációjaként, méghozzá egyértelmüen.

Bizonyitás: tegyük fel, hogy a

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

vektorok bázist alkotnak az adott lineáris térben, a $\underline{c}$ vektor pedig ugyanennek a térnek egy tetszőleges vektora. Ilyen feltételek mellett a

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

n + 1 vektorból álló vektorrendszer "rangja" n (a $\underline{b}_1, \ldots, \underline{b}_n$ vektorok ugyanis lineárisan függetlenek, hiszen bázist képeznek, s az adott térben n-nél több lineárisan független vektor nem vehető fel). Igy az első rangtétel értelmében a $\underline{c}$ vektor kifejezhető a $\underline{b}_1, \underline{b}_2, \ldots, \underline{b}_n$ vektorok lineáris kombinációjaként, mégpedig egyértelmüen. Ezek szerint

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

ahol az $x_1, x_2, \ldots, x_n$ skalárok egyértelmüen meghatározottak.

Megjegyzés: A

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

lineáris kombinációban az

$$x_1, x_2, \ldots, x_n$$

skalárokat a $\underline{c}$ vektornak a

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

bázisvektorokra vonatkozó koordinátáinak nevezzük.

Ezek után irányitsuk figyelmünket az n-elemü vektorok terének ez uj szempont szerinti vizsgálatára.

2. Tétel: Az n-elemü vektorok tere n-dimenziós.

Bizonyitás: Először is azt mutatjuk meg, hogy az n-elemü vektorok közül ki tudunk választani n lineárisan független vektort. E végből tekintsük az n-edrendü egységvektorokat:

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

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

$$\vdots$$

$$\underline{e}_n = [0, 0, \ldots, 1]^*.$$

A belőlük képzett

$$x_1 \underline{e}_1 + x_2 \underline{e}_2 + \ldots + x_n \underline{e}_n = \underline{0}$$

egyenlet csakis triviális módon oldható meg. Ez az

$$x_1 \underline{e}_1 + x_2 \underline{e}_2 + \ldots + x_n \underline{e}_n = [x_1, x_2, \ldots, x_n]^* = \underline{x}$$

összefüggés alapján könnyen belátható. Az előbbi egyenlet ugyanis ilyen jelölésmód mellett

$$\underline{x} = \underline{o}$$

alakban is felirható, ami éppen azt jelenti, hogy valóban csak a triviális megoldás létezik, vagyis az n-elemü egységvektorok lineárisan független rendszert alkotnak.

Ez egyszersmind azt is jelenti, hogy az n-elemü vektorok tere legalább n-dimenziós.

Másrészt viszont bármely n-elemü vektor felirható az n-elemü egységvektorok lineáris kombinációjaként.

Legyen pl.

$$\underline{y} = [y_1, y_2, \ldots, y_n]^*$$

egy tetszőleges n-elemü vektor. A müveleti szabályokból folyik, hogy

$$\underline{y} = y_1 \underline{e}_1 + y_2 \underline{e}_2 + \ldots + y_n \underline{e}_n,$$

vagyis az $\underline{y}$ valóban felirható, mint az egység vektorok lineáris kombinációja.

Ezek után vegyük az

$$\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_k$$

n-elemü vektoroknak egy olyan rendszerét, amelyben a k egy tetszőlegesen nagyra választható természetes szám. Mivel mindezek a vektorok kifejezhetők mint az egységvektorok lineáris kombinációi, az $\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_k$ vektorrendszer rangja legfeljebb n.

Ebből következik, hogy az n-elemü vektorok terében n-nél több lineárisan független vektort nem lehet kiválasztani. Az n-elemü vektorok tere tehát valóban n-dimenziós.

Következmény: Az $\underline{e}_1, \underline{e}_2, \ldots, \underline{e}_n$ egységvektorok bázist alkotnak az L_n-ben.

Például az

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

vektorok bázisát képezik az L_2 lineáris térnek, azaz ezen vektorok lineáris kombinációjaként minden két elemü vektor előállitható. Igy az

$$5\,\underline{e}_1 + 12\,\underline{e}_3$$

lineáris kombináció az $[5,\ 12]^*$ vektort állitja elő.

Megjegyzések: a) Az egységvektorokból képzett bázisra jellemző, hogy a reá vonatkozó koordináták megegyeznek a kérdéses vektor komponenseivel. E szerint az

$$\underline{a} = [a_1,\ a_2,\ \ldots,\ a_n]^*$$

kifejezés valójában rövid irásmódja az

$$a_1\,\underline{e}_1 + a_2\,\underline{e}_2 + \ldots + a_n\,\underline{e}_n$$

lineáris kombinációnak.

b) Az elmondottak alapján az egységvektorok által meghatározott bázist triviális bázisnak is szoktuk nevezni.

Valamely altér dimenziójának meghatározásához, az alteret generáló vektorrendszer vizsgálatából indulunk ki.

3. Tétel: Az $\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_k$ vektorrendszer által generált altér dimenziója megegyezik a vektorrendszer rangjával.

Bizonyitás: Mivel a vektorrendszer minden vektora benne fekszik a kérdéses altérben, következik, hogy az altér dimenziója legalább akkora, mint a rang. De az altér minden vektora előállitható, az alteret generáló vektorok lineáris kombinációjaként, amiből viszont következik, hogy a kérdéses dimenzió legfeljebb akkora, mint a rang. E két tény pedig egyidejüleg csak akkor állhat fenn, ha az alteret generáló vektorrendszer rangja megegyezik az altér dimenziójával.

2.5. Mátrixok rangja

Az eddigiekben az

$$L' = \{\underline{s} \mid \underline{s} = \underline{A}\,\underline{x}\} \quad \text{és} \quad L'' = \{\underline{r} \mid \underline{r}^* = \underline{y}^*\,\underline{A}\}$$

alterekről volt szó, ahol L' valamely adott $\underline{A}$ mátrix oszlopvektorterét, L'' pedig sorvektorterét jelölte. Mivel a lineáris térnek s igy az altérnek is legfontosabb jellemzője a dimenzió, vizsgáljuk meg, milyen kapcsolat van a két altér dimenziója között.

1. Tétel: Az oszlopvektortér és a sorvektortér dimenziója bármely adott $\underline{A}$ mátrixra nézve megegyezik.

Bizonyitás: Legyen az $\underline{A}$ olyan n.k tipusu mátrix, amelynél az oszlopvektortér dimenziója r_1, a sorvektortér dimenziója meg r_2. Az előző pont utolsó tétele alapján tudjuk, hogy az r_1 egyuttal az oszlopvektorokból alkotott vektorrendszer rangját, az r_2 pedig a sorvektorokból alkotott vektorrendszer rangját is jelenti. Ezért

$$r_1 \leqq k$$

és

$$r_2 \leqq n.$$

Ha az oszlopvektorokból alkotott vektorrendszer rangja r_1, akkor az $\underline{A}$ oszlopvektoraiból ki tudunk választani r_1 lineáris független vektort, amelyek egyben bázist alkotnak az oszlopvektortérben. A belőlük felépithető $n.r_1$ tipusu mátrixot jelöljük $\underline{B}_1$-gyel. Az első rangtételből következik, hogy az $\underline{A}$ minden oszlopvektora előállitható a $\underline{B}_1$ mátrix oszlopvektorainak lineáris kombinációjaként. Igy, ha az $\underline{A}$ oszlopvektorait az $\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_k$ szimbólumokkal jelöljük, akkor

$$\underline{a}_1 = \underline{B}_1 \cdot \underline{x}_1$$

$$\underline{a}_2 = \underline{B}_1 \cdot \underline{x}_2$$

$$\vdots$$

$$\underline{a}_k = \underline{B}_1 \cdot \underline{x}_k,$$

ahol az $\underline{x}_i$ komponensei az $\underline{a}_i$ vektornak a $\underline{B}_1$ által meghatározott bázisra vonatkozó koordinátáit jelentik. $(i = 1, 2, \ldots, k)$

Ennek megfelelően

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

A mátrixok szorzatával kapcsolatos 2. tétel alapján a

$$\left[\underline{B}_1\,\underline{x}_1,\ \underline{B}_1\,\underline{x}_2,\ \ldots,\ \underline{B}_1\,\underline{x}_k\right] = \underline{B}_1\left[\underline{x}_1,\ \underline{x}_2,\ \ldots,\ \underline{x}_n\right]$$

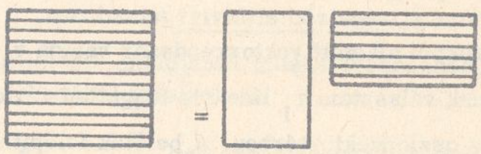
szorzattal. Ha az

$$\left[\underline{x}_1,\ \underline{x}_2,\ \ldots,\ \underline{x}_k\right] = \underline{B}_2$$

jelölést alkalmazzuk, akkor

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

alakban is irható, ahol $\underline{B}_2$ $(r_1 \cdot k)$ tipusu. Az $\underline{A} = \underline{B}_1 \cdot \underline{B}_2$ összefüggés azonban a szorzatra tanult 3. tétel alapján ugyis értelmezhető, hogy az $\underline{A}$ sorvektorai előállithatók a $\underline{B}_2$ sorvektorainak lineáris kombinációjaként.



$$\underset{(n,k)}{\underline{A}} = \underset{(n,r_1)}{\underline{B}_1} \cdot \underset{(r_1 \cdot k)}{\underline{B}_2}$$

Ez azt is jelenti, hogy az $\underline{A}$ sorvektortere megegyezik a $\underline{B}_2$ sorvektorteré-vel, amelynek dimenziója viszont nem lehet nagyobb sorvektorainak számánál az r_1-nél. Igy az $\underline{A}$ sorvektorterének dimenziója legfeljebb r_1 vagyis az elmondottak értelmében érvényes az

$$r_2 \leqq r_1$$

reláció.

Az $\underline{A}$ transzponálása után a fenti gondolatmenetet megismételve arra az eredményre jutunk, hogy az $\underline{A}$ oszlopvektorterének dimenziója legfeljebb r_2, vagyis

$$r_1 \leqq r_2.$$

A fenti két egyenlőtlenség azonban csak ugy állhat fenn egyidejüleg, ha

$$r_1 = r_2 .$$

Ezzel állitásunkat bebizonyitottuk.

Az $\underline{A}$ mátrix sorvektorterének ill. oszlopvektorterének dimenzióját meghatározó számot a mátrix rangjának nevezzük, s a

$$\rho(\underline{A})$$

szimbólummal jelöljük, amelyet igy olvasunk: "ró $\underline{A}$".

Tekintsük pl. a következő mátrixot:

$$\underline{A} = \begin{bmatrix} 1 & 0 & 5 \\ 0 & 1 & 12 \end{bmatrix}$$, amelyben az $\underline{e}_1$ és e_2

egységvektorok lineárisan függetlenek és az $[5,\ 12]^*$ ezek lineáris kombinációja, mégpedig: $5\,\underline{e}_1 + 12\,\underline{e}_2$. Igy $\rho(\underline{A}) = 2$.

Megjegyzések: a) Az $\underline{A}$ mátrix rangja megegyezik az oszlopvektoraiból, ill. sorvektoraiból alkotott vektorrendszerek rangjával.

b) Az $\underline{A}$ rangja nem haladhatja meg sem a sorvektorok számát, sem az oszlopvektorok számát.

A kvadratikus mátrixokat a rang szempontjából két csoportba szoktuk sorolni: szinguláris és nem-szinguláris mátrixok csoportjába.

Ha egy kvadratikus mátrix rangja kisebb, mint sorainak (ill. oszlopainak) száma, akkor a mátrixot szinguláris mátrixnak, ellenkező esetben pedig nem-szinguláris mátrixnak nevezzük.

Például: Tekintsük az

$$\underline{A} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{és} \quad \underline{B} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$$ mátrixokat.

Az $\underline{A}$ mátrix szinguláris, mert $\rho(\underline{A}) = 1$, mivel a nullvektor nem szerepelhet a független vektorok közt. A $\underline{B}$ mátrix viszont nem-szinguláris, mert kimutatható, hogy $\rho(\underline{B}) = 2$, ami megegyezik a sorok számával.

2.6. Az euklideszi tér

Az L_n lineáris teret, mint bizonyos követelményeknek eleget tevő elemek halmazát definiáltuk. A következőkben megvizsgáljuk a mérés lehetőségét is e lineáris térben.

Mivel a geometriai értelemben vett mérés a távolság és szög segitségével történik, a lineáris térben való "méréshez" is ezeket a fogalmakat értelmezzük. E fogalmaknak a lineáris térbe való átültetésénél természetesen el kell tekintenünk a közvetlen fizikai szemlélettől. (Ez csupán $n \leqq 3$ esetén adódhat.) A meghatározásoknak azonban olyanoknak kell lennie, hogy az elemi geometriai fogalmakat speciális esetként magukba foglalják.

Ahhoz, hogy a lineáris térben távolságról beszélhessünk a pontot is értelmeznünk kell. Az egy-, kettő- és három elemü vektorokat már kapcsolatba hoztuk az elemi geometriai ponttal. Most ezt általánositjuk és

a lineáris tér pontjain, a lineáris tér vektorait értjük.

Ezért a következőkben a pont és a vektor fogalmát egyenértékü fogalmaknak tekintjük. Ezek szerint a

$$\underline{a} = [a_1, a_2, \ldots, a_n]^*$$

vektor az L_n egyik pontja. Ha $n > 3$, akkor az uj pontfogalomnak természetesen nem tudunk szemléletes értelmet tulajdonitani.

Ezek után definiáljuk két pont távolságát a lineáris térben.

Ha $\underline{a}$ és $\underline{b}$ ugyanannak az L_n lineáris térnek egy-egy vektora, akkor az $\underline{a}$ és $\underline{b}$ vektorok távolságán azt a d szimbólummal jelölt nem-negativ valós számot értjük, amelyet

$$d = \sqrt{(\underline{a} - \underline{b})^* (\underline{a} - \underline{b})}$$

formula alapján számitunk ki.

Ha tehát

$$\underline{a} = [a_1, a_2, \ldots, a_n]^*$$

és

$$\underline{b} = [b_1, b_2, \ldots, b_n]^*$$

.

akkor

$$d = \sqrt{(a_1 - b_1)^2 + \ldots + (a_n - b_n)^2} .$$

Vegyük észre, hogy a két elemü vektorok terében formulánk éppen a két pont távolságának meghatározására szolgáló ismert összefüggést adja.
Igy

$$\underline{a} = [11,\ 8]^* \quad \text{és} \quad \underline{b} = [3,\ 2]^*$$

vektorok esetén

$$\underline{d} = \sqrt{(11 - 3)^2 + (8 - 2)^2} = 10$$

a pontok távolsága.

Egy vektornak a nullvektortól való távolságával definiáljuk a vektor hosszát.

Valamely adott $\underline{a}$ vektor hosszán (vagy abszolut értékén) azt az $|\underline{a}|$ szimbólummal jelölt nem-negativ valós számot értjük, amelyet az

$$|\underline{a}| = \sqrt{\underline{a}^* \cdot \underline{a}}$$

formula alapján határozunk meg.
Ha

$$\underline{a} = [a_1,\ a_2,\ \ldots,\ a_n]^*,$$

akkor

$$|\underline{a}| = \sqrt{a_1^2 + \ldots + a_n^2} .$$

Legyen pl. $\underline{a} = [3,\ 4]^*$ akkor $|\underline{a}| = \sqrt{25} = 5$. Ennek megfelelően az $\underline{a}$ és $\underline{b}$ vektorok távolságát jelölő d skalár igy is irható

$$d = |\underline{a} - \underline{b}| .$$

Ha az $\underline{a}$ két elemü vektor, akkor az $|\underline{a}|$ szimbólum az analitikus geometriában szerzett ismeretek szerint nem más, mint az $\underline{a}$ vektornak megfelelő pont távolsága az origótól. Ez a magyarázata annak, hogy $n = 2$, $n = 3$ esetén szokás a vektorokhoz rendelt pontot egy irányitott egyenes segitségével az origóval összekötni:

$$|\underline{r}|^2 + |\underline{s}|^2 + 2\,(|\underline{r}|\cdot|\underline{s}|) \geqq |\underline{r}|^2 + |\underline{s}|^2 + 2\,(\underline{r}^* \cdot \underline{s})$$

alakra hozható, amelyből egyszerü átalakitással nyerjük az

$$|\underline{r}| \cdot |\underline{s}| \geqq \underline{r}^* \underline{s}$$

összefüggést. Mivel okoskodásunk minden lépése, amellyel ez utóbbi relációhoz jutottunk, megforditható, azért

$$|\underline{a} - \underline{b}| + |\underline{b} - \underline{c}| \geqq |\underline{a} - \underline{c}|$$

egyenlőtlenség ekvivalens, az

$$|\underline{r}| \cdot |\underline{s}| \geqq \underline{r}^* \cdot \underline{s}$$

egyenlőtlenséggel. Igy ennek bizonyitása egyszersmind annak bizonyitását is jelenti, hogy távolságfogalmunkra érvényes a háromszögegyenlőtlenség.

A bizonyitásnál azt a tényt használjuk fel, hogy bármely λ skalár mellett igaz, hogy

$$(\underline{r} - \lambda\underline{s})^* (\underline{r} - \lambda\underline{s}) \geqq 0,$$

hiszen az $\underline{r} - \lambda\,\underline{s}$ kifejezés egy vektort jelent, és egy vektornak önmagával alkotott skaláris szorzata nem lehet negativ. Sőt azt is tudjuk, hogy az egyenlőtlenség csak akkor válik egyenlőséggé, ha $\underline{r} - \lambda\,\underline{s} = \underline{0}$, azaz $\underline{r} = \lambda\,\underline{s}$.

Mivel pedig

$$(\underline{r} - \lambda\,\underline{s})^* (\underline{r} - \lambda\underline{s}) = \underline{r}^* \cdot \underline{r} - \lambda(\underline{s}^* \cdot \underline{r}) - \lambda(\underline{r}^* \underline{s}) + \lambda^2(\underline{s}^* \cdot \underline{s}) =$$

$$= \underline{r}^* \cdot \underline{r} - 2\lambda(\underline{r}^* \underline{s}) + \lambda^2 \underline{s}^* \underline{s} = |\underline{s}|^2 \cdot \lambda^2 - 2\,(\underline{r}^* \cdot \underline{s})\lambda + |\underline{r}|^2,$$

az $(\underline{r} - \lambda\,\underline{s})^* (\underline{r} - \lambda\,\underline{s}) \geqq 0$ egyenlőtlenség igy is felirható:

$$|\underline{s}|^2 \lambda^2 - 2\,(\underline{r}^* \underline{s})\lambda + |\underline{r}|^2 \geqq 0.$$

Ez azt jelenti, hogy a bal oldalon álló, a λ skalárra nézve másodfoku kifejezés negativ értéket nem vehet fel. Ez csak ugy lehetséges, ha ennek a λ-ban másodfoku kifejezésnek a diszkriminánsa negativ vagy zérus.

A diszkrimináns most

$$\left[2\,(\underline{r}^{*}\,\underline{s})\right]^{2} - 4\,(|\underline{s}|^{2}\cdot|\underline{r}|^{2}) = 4\,(\underline{r}^{*}\,\underline{s})^{2} - 4\,(|\underline{s}|^{2}\cdot|\underline{r}|^{2}) =$$

$$= 4\left[(\underline{r}^{*}\,\underline{s})^{2} - (|\underline{s}|\cdot|\underline{r}|)^{2}\right]$$

alaku.

Ez pedig akkor nem pozitiv, ha a szögletes zárójelben nem-pozitiv szám áll, vagyis

$$(|\underline{r}|\cdot|\underline{s}|)^{2} \geqq (\underline{r}^{*}\,\underline{s})^{2}.$$

A kapott, un. Cauchy-féle egyenlőtlenségből már következik a vizsgált állitás helyessége.

Az L_n-ben értelmezett távolságfogalom tehát eleget tesz a távolság fogalmával szemben támasztott követelményeknek.

A mérés lehetőségének teljessé tételéhez meg kell adnunk még az L_n-ben értelmezett szög fogalmát is. Ezzel kapcsolatban először a sugarat definiáljuk.

Az $\underline{a} \neq \underline{0}$ vektorhoz tartozó sugáron mindazon

$$\lambda\,\underline{a}$$

alakban kifejezhető pontok halmazát értjük, ahol $\lambda \geqq 0$.

A két dimenziós lineáris térben egy adott ponthoz tartozó sugár ezek szerint egy olyan félegyenest jelent, amely az origóból indul ki és átmegy az adott ponton. Az $\underline{a} = \underline{0}$ ponthoz nem tartozik sugár.

Amikor két vektor hajlásszögéről beszélünk, akkor ezen mindig a két vektor által meghatározott sugár hajlásszögét értjük. Az előzőek szerint világos, hogy két vektor hajlásszöge csak akkor van értelmezve, ha egyik sem nullvektor. Ilyen feltételek mellett a következő definiciót adjuk:

Az L_n tér adott $\underline{a} \neq \underline{0}$ és $\underline{b} \neq \underline{0}$ vektorának hajlásszögén azt a φ szöget értjük, amely eleget tesz a

$$\cos\varphi = \frac{\underline{a}^{*}\cdot\underline{b}}{|\underline{a}|\cdot|\underline{b}|}$$

összefüggésnek.

Legyen például

$$\underline{a} = \left[1,\ -1,\ 1,\ 1\right]^{*}$$

és

$$\underline{b} = [0,\ -2,\ \ 2,\ -1]^*.$$

Ekkor

$$\cos\varphi = \frac{3}{2.3} = \frac{1}{2}$$

és igy

$$\varphi = 60^o + n\ .\ 360^o,$$

vagy radiánban kifejezve:

$$\varphi = \frac{\pi}{3} + n\ .\ 2\pi\ .$$

Megjegyzések: a) A trigonometriából tudjuk, hogy bármely φ mellett teljesülnie kell a

$$-1 \leqq \cos\varphi \leqq 1$$

összefüggésnek. Az L_n-ben értelmezett szögfogalom csak akkor jogos, ha arra is fennáll ez a kritérium.
Ennek belátására irjuk fel, az előbb megismert Cauchy-féle egyenlőtlenséget az L_n $\underline{a}$ és $\underline{b}$ vektorára, ami szerint

$$(|\underline{a}|\ .\ |\underline{b}|)^2 \geqq (\underline{a}^*\ .\ \underline{b})^2.$$

Ebből

$$1 \geqq \frac{(\underline{a}^*\ \underline{b})^2}{|\underline{a}|^2\ .\ |\underline{b}|^2}$$

és igy

$$-1 \leq \frac{\underline{a}^*\ \underline{b}}{|\underline{a}|\ .\ |\underline{b}|} \leq 1.$$

b) Mivel a szögre adott definició a φ értékét nem határozza meg egyértelmüen, megállapodunk abban, hogy a számtalan sok lehetőség közül mindig azt választjuk, amely nem nagyobb a derékszögnél.

c) Ha az L_n $\underline{a}$ és $\underline{b}$ vektorára nézve $\cos\varphi = 0$, azaz $\varphi = \frac{\pi}{2}$,

akkor azt mondjuk a két vektor ortogonális. Az ortogonális kifejezést ugyanolyan értelemben használjuk, mint a merőleges kifejezést az elemi geometriában.
Igy, ha pl.

$$\underline{a} = [2,\ 5]^{*} \quad \text{és} \quad \underline{b} = [5,\ -2]^{*},$$

akkor

$$\cos\varphi = \frac{\underline{a}^{*} \cdot \underline{b}}{|\underline{a}| \cdot |\underline{b}|} = \frac{0}{29} = 0.$$

Az L_n lineáris térben értelmezett távolság és szög fogalmat olyan formulákkal vezettük be, amelyekben minden esetben szerepelt két vektor skaláris szorzata. Ezek szerint a skaláris szorzat felhasználása és annak sajátságai tették lehetővé elemi euklideszi geometriai fogalmak általánositását az n-elemü vektorok körében. Végső soron tehát a skaláris szorzat teremti meg a mérés lehetőségét a lineáris térben.

Az előbbiekben értelmezett "metrikával" ellátott lineáris teret euklideszi térnek nevezzük. Ezt igy is szokták definiálni:

Az L_n lineáris térben értelmezve van a skaláris szorzat, ha bármely két $\underline{a}, \underline{b} \in L_n$ vektorának megfelel egy

$$\underline{a}^{*} \underline{b} = a_1 b_1 + a_2 b_2 + \ldots + a_n b_n$$

skalár ugy, hogy ez a megfelelés rendelkezik a következő tulajdonságokkal is:

1. $\underline{a}^{*} \cdot \underline{b} = \underline{b}^{*} \underline{a}$ (vagyis a skaláris szorzat szimmetrikus)

2. $\underline{a}^{*} (\lambda \underline{b}) = \lambda (\underline{a}^{*} \underline{b})$, ahol λ tetszőleges skalár,

3. $(\underline{a}^{*} + \underline{b}^{*})\, \underline{c} = \underline{a}^{*} \underline{c} + \underline{b}^{*} \underline{c}$.

4. Egy vektornak önmagával való skaláris szorzata nem-negativ:

$$\underline{a}^{*} \cdot \underline{a} \geqq 0,$$

s csak akkor nulla, ha $\underline{a} = \underline{0}$.

Az olyan lineáris tereket, amelyekben az 1-4. terjedő feltételeknek eleget tevő skaláris szorzat értelmezve van, euklideszi térnek nevezzük. Az n-elemü vektorok tere tehát euklideszi tér, amit az E_n szimbólummal jelölünk.

Megjegyzések: a) Mivel az euklideszi tér egyben lineáris tér, annak is legjellemzőbb adata a dimenzió. Igy pl. a két elemü vektorok által meghatározott euklideszi tér dimenziója kettő, ami megegyezik a vektoroknak megfelelő elemi geometriai pontok meghatározta sik dimenziójával. Természetesen az E_n dimenziója általánosabb fogalom, mint a geometriában értelmezett dimenzió, ahol ez, $n > 3$ esetén már nincs értelmezve.

b) Az euklideszi térben, a szög fogalmának bevezetése lehetővé teszi, hogy a tér bázisai közt az un. ortogonális bázist is értelmezzük.

Az E_n euklideszi tér valamely bázisát ortogonálisnak nevezzük, ha annak bázisvektorai páronként ortogonálisak.

Ha pl.

$$\underline{b}_1 = \begin{bmatrix} 6 \\ 8 \end{bmatrix} \quad \text{és} \quad \underline{b}_2 = \begin{bmatrix} -8 \\ 6 \end{bmatrix}$$

akkor $\underline{b}_1$ és $\underline{b}_2$ az E_2 ortogonális bázisát képezi.

c) Ha a páronként ortogonális vektorokból felépitett $\underline{B}$ mátrixra fennáll a

$$\underline{B} \cdot \underline{B}^* = \underline{B}^* \cdot \underline{B} = \underline{E}$$

összefüggés, a $\underline{B}$ mátrixot ortogonális mátrixnak nevezzük. Pl. ha

$$\underline{B} = \begin{bmatrix} 0,8 & 0,6 \\ -0,6 & 0,8 \end{bmatrix}, \text{ akkor}$$

$$\begin{bmatrix} 0,8 & 0,6 \\ -0,6 & 0,8 \end{bmatrix} \cdot \begin{bmatrix} 0,8 & -0,6 \\ 0,6 & 0,8 \end{bmatrix} = \begin{bmatrix} 0,8 & -0,6 \\ 0,6 & 0,8 \end{bmatrix} \cdot \begin{bmatrix} 0,8 & 0,6 \\ -0,6 & 0,8 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

2.7. Konvex halmazok

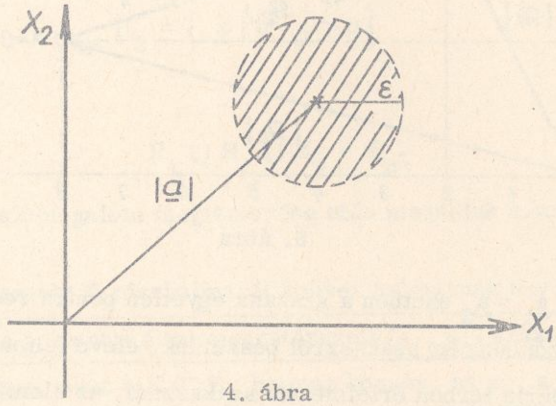
A lineáris algebra gyakorlati alkalmazásainál (igy például az optimumszámitási feladatoknál) fontos szerepet játszanak az E_n, euklideszi tér bizonyos részhalmazai. Mielőtt ezek vizsgálatára rátérnénk, bevezetünk néhány alapfogalmat:

Az $\underline{a} \in E_n$ pont ε sugaru környezetén az E_n mindazon $\underline{x}$ pontjainak halmazát értjük, amelyek eleget tesznek az

$$|\underline{a} - \underline{x}| < \varepsilon$$

követelménynek, ahol az ε tetszőleges pozitiv szám.

Ha az $\underline{a}$ történetesen két elemü vektor, akkor az ε-sugaru környezetét ábrával szemléltethetjük ugy, hogy az $\underline{a}$-nak megfelelő pont körül $\varepsilon > 0$ sugárral kört rajzolunk. A kör belső pontjai képezik az $\underline{a}$ környezetét.



4. ábra

A továbbiakban definiáljuk az egyenes és szakasz fogalmát.

Ha adva van az E_n-ben $\underline{x}_1$ és $\underline{x}_2$ pont, akkor e két pont által meghatározott egyenesen a következő pontok halmazát értjük:

$$X = \left\{ \underline{x} \mid \underline{x} = \lambda \underline{x}_1 + (1 - \lambda)\underline{x}_2 ; \lambda \text{skalár} \right\} .$$

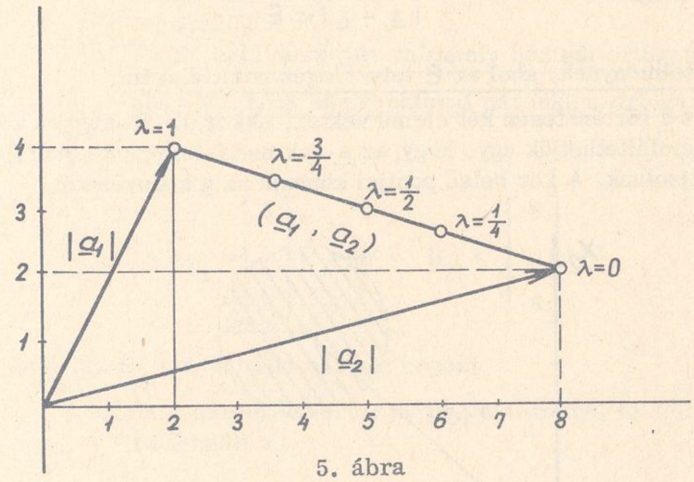
Vagyis a két pont összes olyan lineáris kombinációját, amelyekben a sulyok összege egy, a két pont által meghatározott egyenesnek nevezzük. Könnyen meggyőződhetünk arról, hogy a két-, illetve háromdimenziós vektorok esetében ez a definició az analitikus geometriában definiált egyenesfogalomhoz vezet.

Az E_n euklideszi tér $\underline{a}_1$ és $\underline{a}_2$ pontjai által meghatározott egyenesszaka-

szon pedig az $\underline{a}_1$ és $\underline{a}_2$ vektorok összes lehetséges konvex lineáris kombinációinak halmazát értjük.
A szakaszra adott definiciót igy foglalhatjuk össze:

$$(\underline{a}_1, \underline{a}_2) = \left\{ \underline{a} \,\middle|\, \underline{a} = \lambda \underline{a}_1 + (1 - \lambda)\underline{a}_2, \quad 0 \leqq \lambda \leqq 1 \right\} ,$$

ahol $(\underline{a}_1, \underline{a}_2)$ a szakasz fogalmát szimbólizálja.
Az $\underline{a}_1 = [2, 4]^*$ és $\underline{a}_2 = [8, 2]^*$ pontok által meghatározott szakaszt az alábbi ábra szemlélteti:



5. ábra

Megjegyzés: Az $\underline{a}_1 = \underline{a}_2$ esetben a szakasz egyetlen pontra redukálódik. A következőkben, ha szakaszról beszélünk, eleve feltesszük, hogy $\underline{a}_1 \neq \underline{a}_2$.

Az L_n lineáris térben értelmezett szakasznál, az elemi geometria analógiájára, beszélhetünk a szakasz osztópontjairól is.
Igy

az $\underline{a}_1$ és $\underline{a}_2$ pontok meghatározta $(\underline{a}_1, \underline{a}_2)$ szakasz felező pontja az

$$\frac{1}{2}\underline{a}_1 + \frac{1}{2}\underline{a}_2 = \frac{1}{2}(\underline{a}_1 + \underline{a}_2)$$

vektor.

Bevezetjük még az E_n euklideszi tér un. hipersikjainak fogalmát.
Mindazon $\underline{x} \in E_n$ pontok összességét, amelyekre a

$$\underline{c}^* \underline{x} = b$$

feltétel teljesül, ahol a $\underline{c}^* \neq \underline{0}$ és a $[c_1, c_2, \dots, c_n]$, b számértékek adottak az E_n euklideszi tér hipersikjának nevezzük.

A feltételekből látható, hogy az $\underline{x} = \underline{0}$ vektor csak akkor elégitheti ki az egyenletet, ha b = 0, vagyis

$$\underline{c}^* \underline{x} = 0.$$

Ez azonban azt jelenti, hogy az adott $\underline{c}^*$ vektor, "merőleges" a hipersik minden vektorára.

A $\underline{c}^* \underline{x} = b$ hipersik az n dimenziós euklideszi teret három diszjunkt halmazra bontja fel. Nevezetesen

$$H_1 = \{ \underline{x} \mid \underline{c}^* \cdot \underline{x} < b \}$$

$$H_2 = \{ \underline{x} \mid \underline{c}^* \underline{x} = b \}$$

$$H_3 = \{ \underline{x} \mid \underline{c}^* \underline{x} > b \}$$

amelyekre

$$H_1 \cup H_2 \cup H_3 = E_n.$$

E néhány alapfogalom megismerése után megadjuk a konvex halmazok definicióját.

Az E_n valamely K részhalmazát konvex halmaznak nevezzük, ha a K bármely $\underline{a}_1$ és $\underline{a}_2$ pontja által meghatározott $(\underline{a}_1, \underline{a}_2)$ szakasz benne fekszik a K halmazban, azaz a $K \subset E_n$ halmaz konvex, ha $\underline{a}_1, \underline{a}_2 \in K$ esetén mindig teljesül az $(\underline{a}_1, \underline{a}_2) \subset K$ kikötés. Egy pontot önmagában is konvex halmaznak nevezünk.

A konvex halmazokkal kapcsolatban beszélhetünk belső pontokról és határpontokról.

Az $\underline{a}$ pontot a K konvex halmaz un. belső pontjának nevezzük, ha találunk olyan $\varepsilon > 0$ sugarat, hogy az $\underline{a}$ pont ε sugaru környezetének minden pontja eleme a halmaznak.

Nyilvánvaló, hogy a halmaz belső pontja feltétlenül hozzátartozik a halmazhoz, annak szükségképpen eleme.

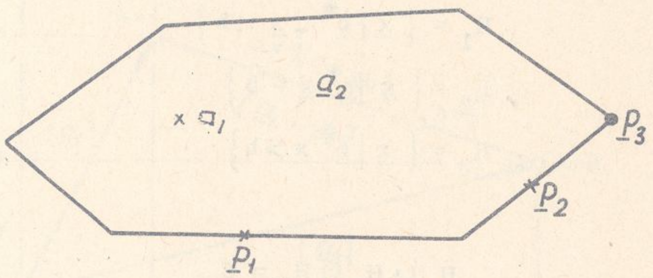
Egy $\underline{b}$ pontot a K konvex halmaz határpontjának nevezünk, ha a pont bármilyen kicsi $\varepsilon > 0$ sugaru környezetének meg van az a tulajdonsága, hogy tartalmaz a K halmazhoz tartozó és annak komplementeréhez tartozó pontokat.

A határpontok nem feltétlenül elemei magának a halmaznak, de tartozhatnak a halmazhoz is. A K konvex halmaz határpontjai között kitüntetett szerepük van az extremális pontoknak.

A K konvex halmazra nézve valamely $\underline{p}$ pontot extremális pontnak nevezzük, ha nincs a K-ban olyan szakasz, amelynek a $\underline{p}$ felező pontja volna.

Az egyetlen pontból álló konvex halmaz egyetlen elemét is extremális pontnak tekintjük.

Példaképpen bemutatjuk, hogy az alábbi ábrán látható hatszög pontjai konvex halmazt alkotnak, ahol az $\underline{a}_1$, és az $\underline{a}_2$ belső pont, a $\underline{p}_1$, a $\underline{p}_2$ és a $\underline{p}_3$ ezzel szemben határpont. A határpontok közül azonban csupán a $\underline{p}_3$ jelent extremális pontot.



6. ábra

Ha a K konvex halmazban kiválasztunk egy $\underline{a}_0$ vektort, aztán képezzük a K minden vektorának és az $\underline{a}_0$-nak különbségét, olyan vektorhalmazhoz jutunk, amely mindig beágyazható legalább egy altérbe. Jelölje r annak a legkisebb dimenziós altérnek a dimenzióját, amely az S halmazt teljes egészében tartalmazza. k S halmaz értelmezése lehetővé teszi, hogy K jellemzésére is felhasználjuk a dimenzió fogalmát.

A K konvex halmaz r-dimenziós, ha az S-ből ki lehet választani r lineárisan független vektort, de az S bármely r + 1 vektorból álló részhalmaza már lineárisan függő rendszer.

A definició szerint a pont pl. nulldimenziós konvex halmaz, a szakasz pedig egydimenziós.

A K halmazra azt mondjuk, hogy nem-korlátos, ha bármely $\underline{x} \in K$ ponthoz található olyan $\underline{h} \neq \underline{0}$ vektor, hogy az

$$\underline{x} + \lambda \underline{h}$$

pont bármely $\lambda \geqq 0$ mellett eleme a K-nak. Egyébként a K korlátos.

Ennek megfelelően, pl. a szakasz korlátos konvex halmaz, de a sugár már nem.

A konvex halmazokkal kapcsolatban bevezetjük még a konvex poliéder és a konvex kónusz fogalmát.

Egy konvex halmazt konvex poliédernek nevezünk, ha

a) korlátos, továbbá,

b) extremális pontjainak száma véges.

Igy például a szakasz konvex poliéder, de konvex poliéder a háromszög és a kocka is.

A konvex poliéderek extremális pontjait csucspontoknak is szoktuk nevezni.

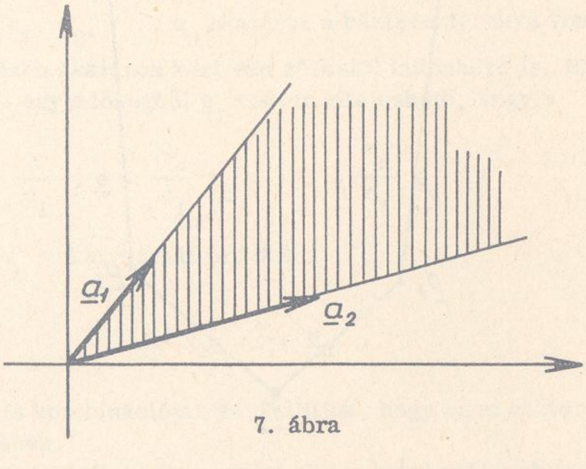
A konvex poliéderek között a legegyszerübbek az un. szimplexek.

Az n-dimenziós szimplex olyan konvex poliéder, amelynek n + 1 csucspontja van.

Ezek szerint a pont nulldimenziós szimplex, a szakasz egy-dimenziós szimplex, a háromszög két-dimenziós szimplex, a tetraéder pedig három-dimenziós szimplex. A magasabb dimenzióju szimplexeknek nincs külön nevük.

Egy C konvex halmazt konvex kónusznak nevezünk, ha C bármely $\underline{a}$ pontjára nézve igaz, hogy a $\lambda \underline{a}$ is pontja a C-nek bármely nem negativ λ esetén.

A konvex kónuszra legegyszerübb példa a $\{\underline{0}\}$, halmaz, ami azonban konvex poliéder is. Ez a kónus egyetlen $\underline{0}$ elemből áll, mert $\lambda \cdot \underline{0} = \underline{0}$ ezért ez a halmaz korlátos halmaz. Ha azonban a C-nek van a $\underline{0}$ vektortól különböző eleme is, akkor a C nem korlátos halmaz. Konvex kónust szemléltet a sikban az alábbi ábra szerint az $\underline{a}_1$ és $\underline{a}_2$ vektorok által meghatározott szögtartomány is.



7. ábra

1. Tétel: A konvex kónusoknak legfeljebb egy csucspontjuk lehet.

Bizonyitás: A definició szerint, ha $\underline{a} \in C$ halmaznak, akkor $\lambda \underline{a} \in C$ állitás is igaz.

Ha már most $\underline{a} \neq \underline{0}$ vektor, akkor az nem lehet extremális pont, mivel felező pontja a $(\underline{0}, 2\underline{a})$ szakasznak.

A tétel szemléltetésére tekintsük az $\underline{a}_1$ és $\underline{a}_2$ n-elemü vektorok nem-negatív lineáris kombinációit. Ezek nyilván konvex kónuszt képeznek mégpedig olyant amelynek a $\underline{0}$ az egyetlen csucspontja. Kételemü vektorok esetén éppen a 7. ábra mutatja e halmaz képét.

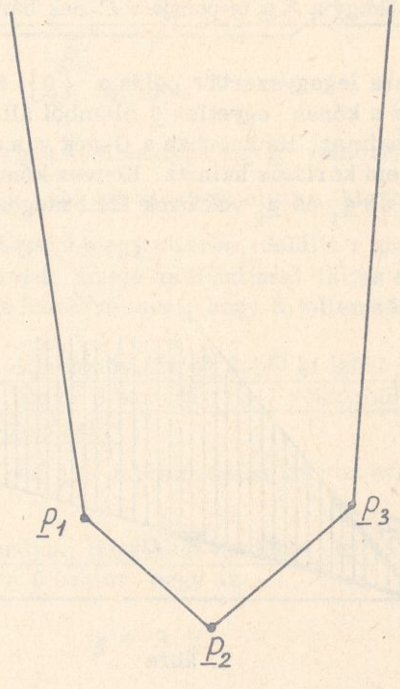
Konvex kónusznak tekinthető az $\underline{a}_1$, $\underline{a}_2 \in L_n$ vektorok összes lehetséges lineáris kombinációi is. A két-elemü vektorok esetén ez éppen a sik pontjait jelenti. Ebben az esetben a konvex kónusznak nincs extremális pontja.

Ez utóbbi példa alapján megállapithatjuk, hogy nemcsak a zérus-altér tesz eleget a konvex kónusz követelményeinek.

A konvex poliéder és konvex kónusz definiálása után megemlitjük még a poliédrikus halmaz fogalmát, amelyről később éppen az előbbi két halmazzal kapcsolatban lesz szó.

Egy konvex halmazt poliédrikus halmaznak nevezünk, ha annak véges sok csucspontja van.

Például:



8. ábra

3.1. Az elemi bázistranszformáció

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

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

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

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

alakban, ahol a $c_1, c_2, \ldots, c_n$ 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

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

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

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

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

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

vektorok lineárisan függetlenek. E végből irjuk fel az

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

egyenletet. Ha ennek volna a triviálistól különböző megoldása is, akkor a $\underline{c}$ vektort az $\underline{e}_2, \dots, \underline{e}_n$ 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

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

bázisáról áttérhetünk uj bázisra.

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

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

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

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

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 állitás 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

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

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

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

ahonnan

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

Ha pedig bevezetjük az

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

és a

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

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 biztosítható.

4. Lineáris egyenletrendszerek megoldása és mátrixok invertálása

4.1. Általános tudnivalók

Az x_1, x_2 ..., x_n ismeretleneket tartalmazó

$$a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1$$

$$a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2$$

$$\vdots$$

$$a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m$$

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, ..., b_m skalárok mindegyike "0", akkor az egyenletrendszert homogénnek hivjuk, egyébként inhomogén egyenletrendszerről van szó.

Ha a fenti un. skaláregyenletrendszerben az adott skalárokra az

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

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

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

Ez a vektoregyenlet azonban felfogható ugyis, mint egy $\underline{A}$ oszlopvektorainak olyan lineáris kombinációja, amelyben a skalárok $x_1, x_2, \dots, x_n$ 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.

Az M konvex halmaz dimenzióját az egyenletrendszer szabadságfokának nevezzük.

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. Tétel: ha a megoldáshalmaz szabadságfoka 0-nál nagyobb, akkor az M nem korlátos halmaz.

Bizonyitás: 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.

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.

Az M konvex halmaz dimenzióját az egyenletrendszer szabadságfokának nevezzük.

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. Tétel: ha a megoldáshalmaz szabadságfoka 0-nál nagyobb, akkor az M nem korlátos halmaz.

Bizonyitás: 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.

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.

Az M konvex halmaz dimenzióját az egyenletrendszer szabadságfokának nevezzük.

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. Tétel: ha a megoldáshalmaz szabadságfoka 0-nál nagyobb, akkor az M nem korlátos halmaz.

Bizonyitás: 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 $\underline{x} \geq \underline{0}$ 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 megforditva, 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

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

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

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

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

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

egyenlőségnek megfelelően particionáljuk. 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

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

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

<u>Megjegyzések:</u> 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.

4.3. Mátrixok inverze

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

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.

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

Bizonyitás: 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 - skalararitmetika 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. $T(\underline{x}_1 + \underline{x}_2) = T(\underline{x}_1) + T(\underline{x}_2)$

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

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 emlitett példák közül az első lineáris transzformáció, a második pedig nem.

A $T(\underline{x}) = 6\underline{x}$ transzformáció lineáris, mert kielégiti 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. Tétel: Egy L_n-ben értelmezett $T(\underline{x})$ transzformációt akkor és csak akkor nevezünk lineáris transzformációnak, ha az felirható

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

alakban.

Bizonyitás: 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 lineáritási 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 lineáritás miatt

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

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

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

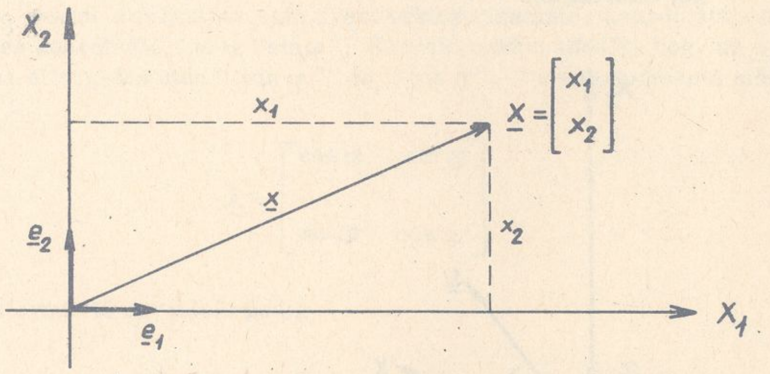
áris transzformáció. Mind a három esetben teljesülnek a linearitási feltételek. Ennek bizonyitására azonban nem térünk ki.

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 egy 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:

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

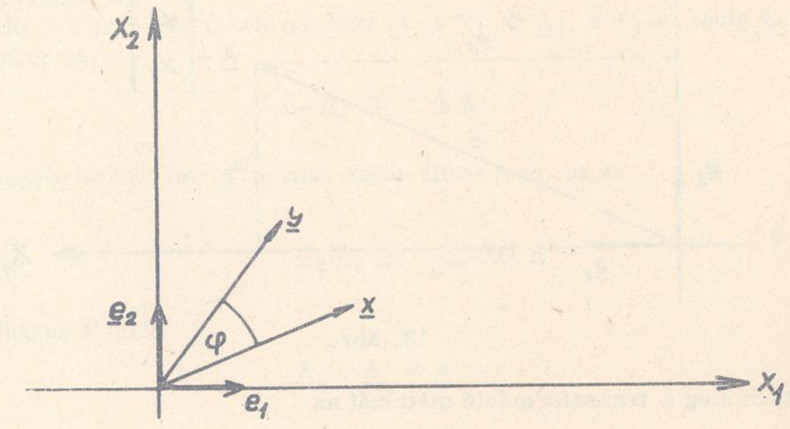
$$(\underline{a}_1 = T_1(\underline{e}_1) = [1,\ 0]^* \text{ és } \underline{a}_2 = T_1(\underline{e}_2) = [0,\ 0]^*),$$

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.

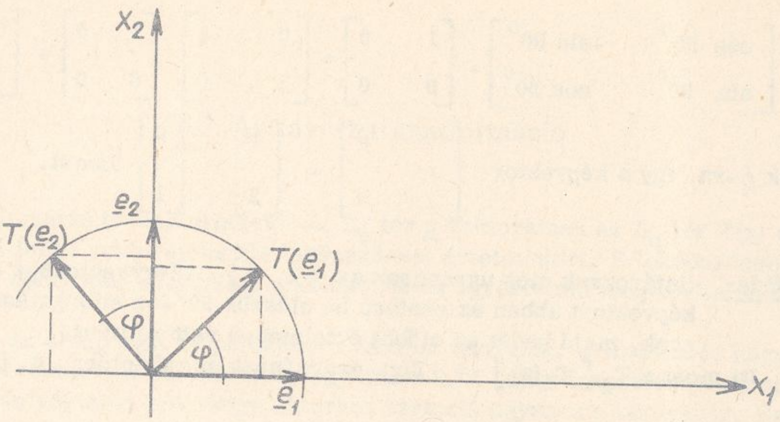


14. ábra

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(**x**) lineáris transzformációjára invariáns altérnek nevezzük, ha L' minden **x** elemére a T(**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 $\mathbf{A} \in \mathscr{M}(p, q)$ 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

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

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

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

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

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

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 $\mathbf{A}$. 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 $\mathbf{A}$ 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 $\mathbf{x}$ 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

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

akkor a definíció értelmében

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

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 $\mathbf{M}\mathbf{x}$ 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 **A**, és innen áttérünk a **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 **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 **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 **A** szimmetricitása miatt azonos az **A** oszlopvektor-terének ortogonális komplementerével is.

relációnak, amely az $s^*As=0$ feltevés miatt végül is az

$$y^*Ay+2ty^*As \geqq 0$$

egyenlőtlenségre redukálódik. Ez azonban minden y és minden t mellett valóban csak akkor állhat fenn, ha $As=o$.

Ezzel tételünk bizonyítását be is fejeztük.

Hasonló módon bizonyítható a

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

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

3. *tétel: Az* x^*Ax *definit kvadratikus forma akkor és csakis akkor szigorúan definit, ha az* 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

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

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 $\mathbf{a}$ és $\mathbf{b}$ 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