Eötvös Loránd Tudományegyetem
Természettudományi Kar

Lineáris algebra alkalmazásai

Szakdolgozat

Készítette:
Nováky Csaba
Matematika BSc
Matematikai elemző szakirány

Témavezető:
Dr. Károlyi Gyula
Egyetemi tanár

Budapest,
2015

Tartalomjegyzék

1. Bevezetés

Szakdolgozatom célja a lineáris algebra néhány alkalmazásának szemléletes, példákon keresztül vett bemutatása. Számos területen szükség van arra, hogy egy feladatból adódó algebrai egyenletrendszert megoldjunk, így elsősorban ezek megoldási módszereiről írok. A XX. század második felében a számítógépek megjelenésével a matematikának ezen területe újra fejlődésnek indult, igény volt arra, hogy egy adott feladatra minél gyorsabb megoldási módszert találjunk. Először olyan lineáris algebrai egyenletrendszerek direkt megoldásával foglalkozom, melyeknek a megoldása egyértelmű. Két módszert ismertetek, melyek hatékonyságát és számítási gyorsaságát konkrét példákon keresztül hasonlítom össze.

Gyakran előfordul azonban, hogy nincs ilyen szerencsénk, az egyenletrendszer túlhatározott, azaz több egyenletünk van, mint ismeretlenünk. Gondoljunk csak egy fizikai mérésre, amikor egy adott rugó rugóállandóját szeretnénk meghatározni. Nagy valószínűséggel nem tudunk két egyforma adatot mérni, hiszen elég csak az emberi tényezőre gondolni, nem tudjuk a stopperórát teljesen pontosan megállítani. Ekkor már 2 mérés eredménye is matematikai értelemben ellentmond egymásnak, így egy ellentmondásos lineáris algebrai egyenletrendszerhez jutunk. Az általánosított inverzek segítségével feloldható ez az ellentmondás, és meg tudjuk mondani a pontos megoldás egy jó közelítését. Egy példán keresztül szemléltetem a megoldás menetét és a pontos eredmény meghatározását.

A harmadik részben sztochasztikus mátrixokról írok, melyet egy lakások állapoti eloszlásával kapcsolatos feladat inspirált. A természetben számos olyan folyamat játszódik le, melynek matematikai leírásában szerepet kapnak sztochasztikus mátrixok. Általában ekkor a mátrix elemei eloszlások, de gondolhatunk átmenet-valószínűségekre vagy permutáció mátrixokra is. A feladat kapcsán megemlítek néhány fontosabb tételt a sztochasztikus mátrixok alkalmazásával kapcsolatban.

2. Lineáris egyenletrendszerek megoldása

Ahhoz, hogy numerikus megoldást kapjunk egy lineáris modellből, szükségünk van arra, hogy megoldjuk a feladatból adódó lineáris egyenletrendszert. Fontos tehát, hogy ezt a feladatot hatékonyan végezzük el, és hogy a megoldás során milyen kerekítési hibák lépnek fel. A számítógépek elterjedésével nagyon gyors algoritmusokat találtak az alapvető mátrixműveletekre, így a lineáris algebrai egyenletrendszerek megoldási módszerei rengeteget fejlődtek az 1940-es évek óta. Hogy egy adott vektor mennyire pontosan közelíti meg a pontos megoldást, szükség van a norma fogalmának bevezetésére.

2.1. Alapfogalmak

2.1.1. Definíció A $\|.\|: \mathbb{R}^n \to \mathbb{R},\ n \in \mathbb{N}$ többváltozós függvényt vektornormának nevezzük, ha teljesülnek a következő feltételek:

- $\|x\| \geq 0\ \ \forall x \in \mathbb{R}^n$-re.
- $\|x\| = 0\ \ \Leftrightarrow x = 0$.
- $\|\lambda x\| = |\lambda|\|x\|\ \ \forall \lambda \in \mathbb{R}, \forall x \in \mathbb{R}^n$-re.
- $\|x + y\| \leq \|x\| + \|y\|\ \ \forall x, y \in \mathbb{R}^n$-re.

A következő formulák vektornormát definiálnak:

$$\|x\|_1 = \sum_{i=1}^{n} |x_i|$$

$$\|x\|_2 = \left(\sum_{i=1}^{n} |x_i|^2\right)^{\frac{1}{2}} = \sqrt{(x, x)}$$

$$\|x\|_\infty = \max_i |x_i|.$$

2.1.2. Definíció Legyen $\|.\|$ egy tetszőleges vektornorma. Ekkor minden $A \in \mathbb{R}^{n \times n}$ mátrix vektornorma által indukált mátrixnormája:

$$\|A\| = \sup \frac{\|Ax\|}{\|x\|}.$$

Ebben a fejezetben olyan lineáris egyenletrendszerekkel foglalkozom, amelyeknek pontosan egy megoldása létezik. Ha az ismeretlenek és az egyenletek száma megegyezik, akkor egy ilyen rendszer felírható

$$Ax = b$$

alakban, ahol A egy invertálható négyzetes mátrix, a b egy adott vektor, x pedig a meghatározandó ismeretlenekből álló vektor. Fontos megtudni, hogy milyen gyorsan és pontosan működik egy ennek megoldására szolgáló algoritmus, ha minden aritmetikai művelet elvégzése pontos, továbbá az, hogy mi a kerekítés hatása. Ez elkerülhetetlen azon számítógépek esetén, melynek aritmetikája véges számjegyű, hiszen ez egy abszolút határt szab a megoldás pontosságának. Képzeljük el a hogy a jobb oldalon álló b vektoron egy δb változtatást. Ekkor:

$$A(x + \delta x) = b + \delta b,$$

ahol δx az x vektor megfelelő változtatását jelenti. Mivel tudjuk, hogy $Ax = b$, ezért

$$A\delta x = \delta b.$$

Hasonlítsuk össze az x relatív megváltozását a b relatív megváltozásával, azaz tekintsük a

$$\frac{\frac{|\delta x|}{|x|}}{\frac{|\delta b|}{|b|}}$$

hányadost. A fenti egyenleteket felhasználva ez a következő alakban írható fel:

$$\frac{|b|}{|x|}\frac{|\delta x|}{|\delta b|} = \frac{|Ax|}{|x|}\frac{|A^{-1}\delta b|}{|\delta b|}.$$

Az $Ax = b$ lineáris egyenletrendszer megoldásának érzékenysége b megváltozására becsülhető ennek maximumával az összes lehetséges x és δb vektoron. A jobb oldalon szereplő első tényezőnek a maximuma $\|A\|$, azaz A euklideszi (kettes) normája, a második tag maximuma pedig az A^{-1} mátrix $\|A^{-1}\|$ euklideszi normája. Ezzel azt kapjuk, hogy az x megoldás relatív hibájának és a b relatív hibájának hányadosa nem lehet nagyobb, mint:

$$\kappa(A) = \|A\|\|A^{-1}\|.$$

2.1.3. Definíció A $\kappa(A)$ számot az A mátrix kondíciószámának nevezzük.

Ha a mátrix kondíciószáma nagy, akkor érzékeny a kerekítési hibákra, így kis változtatás esetén az $Ax = b$ megoldása nagymértékben változhat.

Lássunk erre egy példát, tekintsük a következő lineáris egyenletrendszert:

$$0{,}01x_1 + 0{,}1x_2 = 0{,}32$$
$$0{,}1x_1 + 1{,}01x_2 = 3{,}23$$

Az együtthatókból álló mátrix:

$$A = \begin{pmatrix} 0{,}01 & 0{,}1 \\ 0{,}1 & 1{,}01 \end{pmatrix}.$$

Vegyük az A mátrix kettes normáját:

$$\|A\|_2 = \sqrt{\lambda_{\max}(A^*A)} = 1{,}0199,$$

ahol A^* az A mátrix konjugált transzponáltját, $\lambda_{\max}$ pedig a legnagyobb sajátértéket jelenti.

Most számítsuk ki az A^{-1} mátrixot. Ismert, hogy 2×2-es mátrixok inverzét kiszámíthatjuk a következő módon:

$$A^{-1} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{\det(A)} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Mivel az A mátrix determinánsa 10^{-4}, ezért:

$$A^{-1} = 10^4 \begin{pmatrix} 1{,}01 & -0{,}1 \\ -0{,}1 & 0{,}01 \end{pmatrix}.$$

Ennek kettes normája:

$$\|A^{-1}\|_2 = \sqrt{\lambda_{\max}(A^*A)} = 10199.$$

A kettes normában vett kondíciószám tehát meglehetősen nagy:

$$\kappa(A) = \|A\|_2 \|A^{-1}\|_2 = 1{,}0199 * 10199 = 10402.$$

Most oldjuk meg a feladatot! Az A^{-1} ismeretével kifejezhető az x megoldásvektor:

$$x = A^{-1}b = 10^4 \begin{pmatrix} 1,01 & -0,1 \\ -0,1 & 0,01 \end{pmatrix} * \begin{pmatrix} 0,32 \\ 3,23 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}.$$

Tehát az egyenletrendszer megoldása:

$$x_1 = 2, \qquad x_2 = 3.$$

Most nézzük meg azt, hogy mi történik akkor, ha a második egyenletben egy kis változtatást végzünk:

$$\begin{gathered} 0,01x_1 + 0,1x_2 = 0,32 \\ 0,1x_1 + 1,02x_2 = 3,23. \end{gathered}$$

Jogosan várhatnánk azt, hogy a második egyenlet második változójának együtthatóján végzett egyszázados változtatás kismértékben változtatja meg a megoldást, de ezt felírva a következőt kapjuk:

$$\tilde{A} = \begin{pmatrix} 0,01 & 0,1 \\ 0,1 & 1,02 \end{pmatrix}, \quad \tilde{A}^{-1} = 10^3 \begin{pmatrix} 5,1 & -0,5 \\ -0,5 & 0,05 \end{pmatrix}.$$

Mivel $x = \tilde{A}^{-1}b$, ezért a megoldás a következő:

$$x_1 = 17 \qquad x_2 = 1,5.$$

Látható tehát, hogy ha a kondíciószám értéke nagy, akkor az egyenleten végzett kis változtatás a megoldást nagymértékben befolyásolja.

2.2. Kapcsolat a sajátértékek és a kondíciószám között

Legyen β az A mátrix sajátértékeinek abszolút értékei közül a legnagyobb. Ekkor:

$$\beta \leq \|A\|.$$

Legyen α az A mátrix sajátértékeinek abszolút értékei közül a legkisebb. Ha egy λ sajátértéke egy invertálható mátrixnak, akkor ezen mátrix inverzének $\frac{1}{\lambda}$ a sajátértéke. Ez alapján a következő becslés írható fel:

$$\frac{1}{\alpha} \leq |A^{-1}|$$

A két becslést összerakva a következő egyenlőtlenség adható a kondíciószámra:

$$\frac{|\beta|}{|\alpha|} \leq \kappa(A)$$

Legyen az A mátrix egy valós, négyzetes, önadjungált, pozitív definit mátrix, melynek a legnagyobb sajátértéke β, a legkisebb sajátértéke α. Mivel az A mátrix pozitív definit, ezért minden sajátértéke pozitív, továbbá egy pozitív definit mátrix normája az euklideszi normára nézve megegyezik a legnagyobb sajátértékével:

$$\|A\| = \beta$$

Mivel A^{-1} szintén pozitív definit, ezért

$$\|A^{-1}\| = \alpha^{-1}$$

Tehát a pozitív definit önadjungált mátrixok kondíciószámára a következő egyenlőség adható:

$$\kappa(A) = \frac{\beta}{\alpha}.$$

2.3. Megoldási módszerek

Egy lineáris algebrai egyenletrendszer megoldására számos megoldási módszer létezik. Ebben a fejezetben csak azokkal a módszerekkel foglalkozunk, melyek pontos megoldást adnak.

Egy olyan algoritmust, amely véges sok lépésben elvégzett pontos számításokkal pontos megoldást ad, direkt módszernek nevezzük.

2.3.1. Gauss-elimináció

A lineáris algebrai egyenletrendszer direkt megoldásának legalapvetőbb megoldási módszere a Gauss-elimináció, melyet a következő példán mutatok be:

$$\begin{aligned} x_1 + 2x_2 + 3x_3 - x_4 \quad &= -2 \\ 2x_1 + 5x_2 + 4x_3 - 3x_4 &= 1 \\ 2x_1 + 3x_2 + 4x_3 + x_4 \quad &= 1 \end{aligned}$$

$$x_1 + 4x_2 + 2x_3 - 2x_4 \quad = \quad 3$$

A feladat meghatározni x_1, x_2, x_3, x_4 ismeretleneket. Úgy oldjuk meg az egyenletrendszert, hogy egymás után kiküszöböljük az ismeretleneket. Az első egyenletet használjuk arra, hogy elimináljuk az x_1 ismeretlent a többi egyenletből. Ehhez vonjuk ki az első egyenlet kétszeresét a második és harmadik egyenletből, majd a negyedik egyenletből vonjuk ki az első egyenletet. Ekkor kapjuk a következőt:

$$\begin{aligned} x_2 - 2x_3 - x_4 \quad &= 5 \\ -x_2 - 2x_3 + 3x_4 &= 5 \\ 2x_2 - x_3 - x_4 \quad &= 5 \end{aligned}$$

Most az első egyenletet hozzáadva a másodikhoz és a kétszeresét a harmadikból kivonva kapjuk:

$$\begin{aligned} -4x_3 + 2x_4 &= 10 \\ 3x_3 + x_4 \quad &= -5 \end{aligned}$$

Végül az x_3 változót küszöböljük ki úgy, hogy az első egyenlet ¾-ét adjuk a második egyenlethez, ekkor kapjuk:

$$\frac{5}{2}x_4 = \frac{5}{2},$$

ebből következik, hogy $x_4 = 1$. Most visszafelé haladva x_4 értékét behelyettesítve az előző egyenletbe:

$$x_3 = -2.$$

Most x_3, x_4 ismeretében meghatározható x_2 értéke is, azaz

$$x_2 + 4 - 1 = 5$$

$$x_2 \quad = 2.$$

Végül az x_2, x_3, x_4 értékek segítségével meghatározható x_1 értéke:

$$x_1 + 4 - 6 - 1 = -2$$

$$x_1 = 1.$$

A Gauss-elimináció egy más megközelítésben:

Alapesetben az első egyenletet használjuk x_1 kiküszöbölésére, azaz kifejezzük a következő alakban:

$$x_1 = b_1 + l_1(x_2, \dots, x_n),$$

ahol $l_1(x_2, \dots, x_n)$ lineáris funkcionál. Behelyettesítjük x_1 helyére a $b_1 + l_1$ funkcionált a többi egyenletbe, és kifejezzük az x_2 ismeretlent a többi x_i lineáris funkcionáljaként:

$$x_2 = b_2 + l_2(x_3, \dots, x_n).$$

Ezt addig folytatjuk, míg az $(n-1)$-edik lépés után elérjük az x_n ismeretlent. Ezután sorra meghatározzuk az $x_{n-1}, \dots, x_1$ értékeket.

Ez az eljárás megakadhat az első lépésben, ha az x_1 ismeretlen együtthatója az első egyenletben nulla. Gyakorlatilag az is megakasztja a folyamatot, ha ez az együttható nem nulla, de egy nagyon kicsi szám, hiszen ekkor ezzel az együtthatóval leosztva a többi együttható értéke óriási lesz. Ez azért jelent problémát, mert a számítógépek véges számjegyű lebegőpontos aritmetikával számolnak, és amikor a többi egyenletbe helyettesítjük x_1-et, az együtthatók túlcsordulnak. Erre megoldást jelent egy másik kiküszöbölendő x_j ismeretlen és egy másik egyenlet választása úgy, hogy annak együtthatója ne legyen túl kicsi a többihez képest. Ezt a stratégiát hívják teljes főelem kiválasztásnak. Egy másik stratégia lehet az ismeretlenek sorrendjének megtartása, de másik egyenlet választása a kiküszöböléshez, ahol az együttható nem túl kicsi a többi együtthatóhoz képest. Ezt részleges főelem kiválasztásnak nevezik, és jól működik a gyakorlatban.

2.3.2. Cramer-szabály

2.3.2.1. Tétel Ha $A \in \mathbb{R}^{n\times n}, b \in \mathbb{R}^n, x \in \mathbb{R}^n$ és $\det A \neq 0$, akkor az $Ax = b$ lineáris egyenletrendszernek pontosan egy megoldása létezik. Ekkor a megoldásvektor komponensei a következő egyenletből számíthatók ki:

$$x_i = \frac{\det A_i}{\det A},$$

ahol A_i-vel jelöljük azon A-ból képzett mátrixokat, ahol az i-edik oszlop helyett a b vektor áll.

2.3.2.2. Definíció A lineáris algebrában egy négyzetes mátrix adjungáltjának $(\mathrm{adj}(A))$ nevezzük a mátrix előjeles aldeterminánsaiból alkotott mátrix transzponáltját. Egy $A \in \mathbb{R}^{n\times n}$ mátrix adjungáltján a következő eljárással elkészített mátrixot értjük:

- Felírjuk az A mátrix aldeterminánsmátrixát, vagyis azt a mátrixot, melynek i,j-edik eleme annak a mátrixnak a determinánsa, melyet az A i-edik sorának és j-edik oszlopának törlésével keletkezik.
- Ezen mátrix elemeinek előjelét a „sakktáblaszabály" szerint megváltoztatjuk, azaz az i,j-edik elemnek a $(-1)^{i+j}$ értéket adjuk, ez az előjeles aldetermináns-mátrix,
- Végül ezt a mátrixot transzponáljuk, azaz tükrözzük a főátlóra.

2.3.2.3. Lemma Egy mátrix inverzén azt a mátrixot értjük, amelyre $A^{-1}A = AA^{-1} = I$, azaz egy mátrix inverzének saját magával vett szorzata az egységmátrix. Igazolható, hogy egy négyzetes mátrix jobb és bal oldali inverze megegyezik. Egy invertálható A mátrix esetén:

$$A^{-1} = \frac{\mathrm{adj}(A)}{\det(A)}.$$

Bizonyítás (lemma) Elegendő belátni, hogy:

$$A * \mathrm{adj}(A) = \det(A) * I,$$

ahol $I \in R^{n\times n}$ az egységmátrix. Mivel az $\mathrm{adj}(A)$ mátrix pont olyan, hogy illeszkedjék a determináns kifejtési tételéhez, ezért ha az A mátrix i-edik sorát az adjungált mátrix i-edik oszlopával szorozzuk, pont az A mátrix determinánsát kapjuk eredményül. Ezért a szorzat főátlóbeli eleme A determinánsa lesz. A ferde kifejtési tétel szerint a determinánst úgy fejtve ki, hogy egy sort a nem hozzá tartozó „aldetermináns-oszloppal" szorzunk be, mindig 0-t kapunk, azaz a főátlón kívül csupa 0 lesz, ezzel a lemmát igazoltuk.

Bizonyítás (Cramer-szabály): Felhasználva, hogy az A invertálható mátrix, az $Ax = b$ egyenlet a következő alakban írható fel:

$$x = A^{-1}b = \frac{\mathrm{adj}(A)}{\det(A)}b.$$

Ebből az x vektor komponensei kiszámíthatóak a következő módon:

$$x_i = \frac{A_{1i}b_1 + A_{2i}b_2 + \cdots + A_{ni}b_n}{\det(A)},$$

ahol A_{ij} az A mátrix i-edik sorához és j-edik oszlopához tartozó előjeles aldetermináns értéke. Látható, hogy a számláló értéke éppen az A_i mátrix determinánsa az i-edik oszlopa szerint kifejtve, ezzel igazoltuk a tételt.

2.3.2.4. Példa Adott a következő lineáris algebrai egyenletrendszer. Oldjuk meg a Cramer-szabály segítségével!

$$\begin{aligned} 2x_1 - x_2 - x_3 \quad &= 4 \\ 3x_1 + 4x_2 - 2x_3 &= 11 \\ 3x_1 - 2x_2 + 4x_3 &= 11 \end{aligned}$$

Ekkor:

$$A = \begin{pmatrix} 2 & -1 & -1 \\ 3 & 4 & -2 \\ 3 & -2 & 4 \end{pmatrix}, \qquad b = \begin{pmatrix} 4 \\ 11 \\ 11 \end{pmatrix}$$

$$x_1 = \frac{\begin{vmatrix} 4 & -1 & -1 \\ 11 & 4 & -2 \\ 11 & -2 & 4 \end{vmatrix}}{\begin{vmatrix} 2 & -1 & -1 \\ 3 & 4 & -2 \\ 3 & -2 & 4 \end{vmatrix}} = \frac{180}{60} = 3$$

$$x_2 = \frac{\begin{vmatrix} 2 & 4 & -1 \\ 3 & 11 & -2 \\ 3 & 11 & 4 \end{vmatrix}}{\begin{vmatrix} 2 & -1 & -1 \\ 3 & 4 & -2 \\ 3 & -2 & 4 \end{vmatrix}} = \frac{60}{60} = 1$$

$$x_3 = \frac{\begin{vmatrix} 2 & -1 & 4 \\ 3 & 4 & 11 \\ 3 & -2 & 11 \end{vmatrix}}{\begin{vmatrix} 2 & -1 & -1 \\ 3 & 4 & -2 \\ 3 & -2 & 4 \end{vmatrix}} = \frac{60}{60} = 1.$$

A Cramer-szabály alkalmazása során a lineáris egyenletrendszerből képzett mátrixok determinánsainak hányadosából adódik a megoldás. Fontos megemlíteni, hogy míg a Gauss-elimináció során nem volt feltétel a megoldhatóságra, itt fel kell tennünk, hogy az együtthatókból képzett mátrix determinánsa nem nulla. Továbbá azt is fontos megjegyezni, hogy a Gauss-elimináció a Cramer-szabállyal ellentétben kiterjeszthető arra az esetre is, amikor a meghatározandó ismeretlenek és az egyenletek száma nem egyezik meg, azaz az együtthatókból képzett mátrix nem négyzetes. A Cramer-szabály műveleti igénye is jóval nagyobb, mint a Gauss-elimináció műveleti igénye, így ennek alkalmazása kevésbé hatékony.

3. Általánosított inverz

3.1. Alapfogalmak

Lineáris egyenletrendszerek megoldása során gyakran szükségünk van mátrixok különböző értelemben vett inverzeire. Egy n egyenletből álló és m ismeretlent tartalmazó lineáris egyenletrendszer általánosan a következőképp írható fel:

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m = b_1$$

$$a_{21}x_1 + a_{22}x_2 + \cdots + a_{2m}x_m = b_2$$

$$\vdots$$

$$a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nm}x_m = b_n$$

Ennek mátrixos alakja a következő: $Ax = b$

Az együtthatókból álló mátrix: $A = \begin{pmatrix} a_{11}a_{12} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1}a_{n2} & \cdots & a_{nm} \end{pmatrix}$

ahol $A \in \mathbb{R}^{n \times m}, b \in \mathbb{R}^n, x \in \mathbb{R}^m$.

A korábbi fejezetből már láttuk, hogy ha $n = m$, akkor az egyenlet megoldása: $x = A^{-1}b$, ahol A^{-1} az A mátrix inverzét jelöli. Általában viszont egy rendszer leírása során nem ugyanannyi egyenletünk van, mint ismeretlenünk, ekkor $n \neq m$. Ezért érdemes definiálni a jobb oldali és bal oldali inverzet. Ha a mátrix nem négyzetes, akkor nem invertálható, de létezhet bal- vagy jobbinverze:

3.1.1. Definíció Ha az A $n \times m$-es mátrix rangja m, akkor létezik egy B mátrix, hogy $BA = I$. Ez a B mátrix A balinverze.

3.1.2. Definíció Ha az A $n \times m$-es mátrix rangja n, akkor létezik egy B mátrix, hogy $AB = I$. Ez a B mátrix A jobbinverze.

3.1.3. Rangtartó átalakítások tétele $a_1, \dots, a_k \in V$ vektortér $\mathbb{R}$ felett, $\lambda \in \mathbb{R}$, $\lambda \neq 0$, $i < j \in \mathbb{N}$ esetén:

- két tetszőleges vektor cseréje során a rang nem változik, azaz:

$$r(a_1, a_2, \dots, a_i, \dots, a_j, \dots, a_k) = r(a_1, a_2, \dots, a_j, \dots, a_i, \dots, a_k)$$

- egy tetszőleges vektort szorozva egy nemnulla valós számmal:

$$r(a_1, a_2, \dots, a_i, \dots, a_j, \dots, a_k) = r(a_1, a_2, \dots, \lambda a_i, \dots, a_j, \dots, a_k)$$

- egy tetszőleges vektorhoz egy másik vektortérbeli vektort adva sem változik meg a rang:

$$r(a_1, a_2, \dots, a_i, \dots, a_j, \dots, a_k) = r(a_1, a_2, \dots, a_i + a_j, \dots, a_j, \dots, a_k)$$

- az előző kettő kombinációjából adódik tehát:

$$r(a_1, a_2, \dots, a_i, \dots, a_j, \dots, a_k) = r(a_1, a_2, \dots, a_i + \lambda a_j, \dots, a_j, \dots, a_k)$$

Tegyük fel, hogy az $A \in \mathbb{R}^{n \times m}$ nem nullmátrix. Ekkor keressünk egy nemnulla elemet, majd ezt sorcserével és oszlopcserével a bal felső sarokba visszük. Szorozzuk meg az első oszlopot ennek a számnak a reciprokával, hogy a bal felső elem 1 legyen, majd az első oszlop számszorosát adjuk hozzá a többi oszlophoz úgy, hogy az első sor második elemétől kezdve minden elem 0 legyen. Ugyanezt hajtsuk végre az első oszlopon, majd az első sor és első oszlop elhagyásával kapott mátrixon is. Ez mindaddig végrehajtható, amíg nemnulla mártixot kapunk a sorok és oszlopok elhagyása után. Ekkor a bal felső része a mátrixnak $r \times r$-es egységmátrix lesz, ahol $r = r(A)$ a mátrix rangja, jelöljük ezt I_r-rel. A fenti tétel miatt a mátrix rangja nem változott az átalakítás során.

3.1.4. Tétel $A \in \mathbb{R}^{n\times m}, r \geq 1$ esetén A átalakítható $\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$ alakú mátrixszá.

Vegyünk egy I_n egységmátrixot a soros átalakítások kezdőmátrixaként, egy I_m egységmátrixot az oszlopos átalakításokhoz, és hajtsuk végre rajtuk a transzformációkat, melyeket az A mátrixon hajtottunk végre, hogy megkapjuk a fent látható alakot. Az első soros átalakítást az I_n-en, a következő sorosat a már belőle kapott mátrixon, és így tovább, az utolsó átalakítás legyen S; az első oszlopos átalakítást I_m-en hajtsuk végre, a következő oszloposat az I_m transzformáltján, stb., a végül kapott mátrix legyen P. Ekkor S és P is invertálható, mivel az egységmátrixból indultunk ki, és rangtartó átalakításokat végeztünk rajtuk, melyekre

$$SAP = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}.$$

3.2. Általánosított inverz kiszámítása

3.2.1. Definíció $A \in \mathbb{R}^{n\times m}$ esetén az $A^{(g)}$ egy általánosított inverze A-nak, ha $A^{(g)} \in \mathbb{R}^{m\times n}$ és $AA^{(g)}A = A$.

Nyilván ha $A = 0$, akkor az általánosított inverzeinek halmaza $\{0^{(g)}\} = \mathbb{R}^{m\times n}$.

Ha $A \neq 0, A \in \mathbb{R}^{n\times m}$ esetén $r \geq 1$, így létezik olyan S és P invertálható mátrix, hogy $SAP = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$. Ha S és P rögzített, akkor tetszőleges $X \in \mathbb{R}^{m\times n}$-re:

$$AXA = A \Leftrightarrow SAXAP = SAP \Leftrightarrow SAPP^{-1}XS^{-1}SAP = SAP.$$

Legyen $Y = P^{-1}XS^{-1} = \begin{bmatrix} U & V \\ W & Z \end{bmatrix}$, ahol a felső U blokk $r \times r$-es legyen, a többi pedig az $m \times n$-es méretnek megfelelően. Ekkor:

$$AXA = A \Leftrightarrow SAPYSAP = SAP \Leftrightarrow \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} U & V \\ W & Z \end{bmatrix}\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} \Leftrightarrow$$
$$\begin{bmatrix} U & V \\ 0 & 0 \end{bmatrix}\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} \Leftrightarrow \begin{bmatrix} U & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} \Leftrightarrow U = I_r.$$

Így rögzített S és P esetén meghatároztuk a fenti A mátrix összes általánosított inverzét:

$$\left\{A^{(g)}\right\} = \left\{P \begin{bmatrix} I_r & V \\ W & Z \end{bmatrix} S\right\},$$

ahol V, W, Z tetszőleges, valós elemű mátrixok, amit az előírt alak megenged.

3.2.2. Példa a kiszámításhoz

Nézzük meg egy példán keresztül, hogy pontosan milyen átalakításokra van szükség az általánosított inverz kiszámításához. Legyen:

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 2 & 1 & 0 \\ 1 & 2 & 1 \end{pmatrix}.$$

Látható, hogy az első sor első eleme az egyes, így az első lépésünk legyen az, hogy az első oszlop segítségével kinullázzuk az első sor elemeit. Ehhez a harmadik oszlopból ki kell vonni az első oszlopot:

$$A^{(1)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 1 & -3 \\ 1 & 2 & -2 \end{pmatrix}.$$

Most az első oszlopot nullázzuk ki. Ehhez az első sor kétszeresét vonjuk ki a harmadik sorból, majd vonjuk ki a negyedik sorból egyszer:

$$A^{(2)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & -2 \\ 0 & 2 & 0 \end{pmatrix}.$$

Most a második sort nullázzuk ki. Ehhez ki kell vonni a harmadik oszlopból a másodikat:

$$A^{(3)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & -3 \\ 0 & 2 & -2 \end{pmatrix}.$$

Majd harmadik sorból egyszer-, a negyedik sorból kétszer vonjuk ki a második sort:

$$A^{(4)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -3 \\ 0 & 0 & -2 \end{pmatrix}.$$

Osszuk el a harmadik oszlopot (-3)-mal, így a következő mátrixot kapjuk:

$$A^{(5)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Tehát „szép” alakra hoztuk a mátrixot. Most ezeket az átalakításokat kell elvégezni az I_n, I_m egységmátrixokon. Az átalakítások sorrendben a következők voltak:

- Soros: $S_3 - 2S_1, S_4 - S_1, S_3 - S_2, S_4 - 2S_2, S_4 - \frac{2}{3}S_3$.

$$I_n = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow{S_3-2S_1} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow{S_4-S_1} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -2 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix}$$

$$\xrightarrow{S_3-S_2} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -2 & -1 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix} \xrightarrow{S_4-2S_2} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -2 & -1 & 1 & 0 \\ -1 & -2 & 0 & 1 \end{pmatrix} \xrightarrow{S_4-\frac{2}{3}S_3} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -2 & -1 & 1 & 0 \\ \frac{1}{3} & -\frac{4}{3} & -\frac{2}{3} & 1 \end{pmatrix} = S.$$

- Oszlopos: $O_3 - O_1, O_3 - O_2, O_3 * \left(-\frac{1}{3}\right)$.

$$I_m = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{O_3 - O_1} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{O_3 - O_2} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{O_3 * \left(-\frac{1}{3}\right)} \begin{pmatrix} 1 & 0 & \frac{1}{3} \\ 0 & 1 & \frac{1}{3} \\ 0 & 0 & -\frac{1}{3} \end{pmatrix} = P.$$

Ellenőrizhető, hogy ekkor valóban $SAP = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$ alakú, ahol I_r a 3×3-as egységmátrix:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -2 & -1 & 1 & 0 \\ \frac{1}{3} & -\frac{4}{3} & -\frac{2}{3} & 1 \end{pmatrix} * \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 2 & 1 & 0 \\ 1 & 2 & 1 \end{pmatrix} * \begin{pmatrix} 1 & 0 & \frac{1}{3} \\ 0 & 1 & \frac{1}{3} \\ 0 & 0 & -\frac{1}{3} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Tudjuk, hogy $A^{(g)}$-nek 3×4-es mátrixnak kell lenni, hiszen A 4×3-as, és $AA^{(g)}A = A$. Mivel $\{A^{(g)}\} = \left\{P \begin{bmatrix} I_r & V \\ W & Z \end{bmatrix} S\right\}$, ezért az előírt alak miatt csak a V komponens szerepel az I_r mellett, ahol $V = \begin{pmatrix} p \\ q \\ r \end{pmatrix}$. Ellenőrizhető, hogy tetszőleges $p, q, r \in \mathbb{R}$ számok választása esetén:

$$AA^{(g)}A = A.$$

Az általánosított inverz tehát paraméteres alakban a következő:

$$\{A^{(g)}\} = \left\{ \begin{pmatrix} 1 & 0 & \frac{1}{3} \\ 0 & 1 & \frac{1}{3} \\ 0 & 0 & -\frac{1}{3} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & p \\ 0 & 1 & 0 & q \\ 0 & 0 & 1 & r \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -2 & -1 & 1 & 0 \\ \frac{1}{3} & -\frac{4}{3} & -\frac{2}{3} & 1 \end{pmatrix} \right\},$$

$$\{A^{(g)}\} = \begin{pmatrix} \frac{1}{3} + \frac{1}{3}p + \frac{1}{3}r & -\frac{1}{3} - \frac{4}{3}p - \frac{4}{9}r & \frac{1}{3} - \frac{2}{3}p - \frac{2}{9}r & p + \frac{1}{3}r \\ -\frac{2}{3} + \frac{1}{3}q + \frac{1}{9}r & \frac{2}{3} - \frac{4}{3}q - \frac{4}{9}r & \frac{1}{3} - \frac{2}{3}q - \frac{2}{9}r & q + \frac{1}{3}r \\ \frac{2}{3} - \frac{1}{9}r & \frac{1}{3} + \frac{4}{9}r & -\frac{1}{3} + \frac{2}{9}r & -\frac{1}{3}r \end{pmatrix}.$$

3.3. Általánosított inverz alkalmazása

Most már ki tudjuk számolni egy mátrix általánosított inverzeit, de mire lehet használni? A következő feladatot ennek megválaszolásához mutatom be:

Adott az alábbi egyenletrendszer:

$$\begin{aligned} x_1 + x_3 \qquad &= 1 \\ x_2 + x_3 \qquad &= 2 \\ 2x_1 + x_2 \qquad &= 0 \\ x_1 + 2x_2 + x_3 &= 1 \end{aligned}$$

Látható, hogy az előbb számolt A mátrix az egyenletrendszer együtthatóiból álló mátrix. Ez egy túlhatározott lineáris egyenletrendszer, mivel 4 sora van 3 ismeretlennel, és a sorok lineárisan nem függenek össze. Gyakorlatban ilyen gyakran előfordul, mert nincs tökéletes mérés, így egyik egyenlettől sem várhatjuk, hogy pontosan teljesüljön.

A-nak előbb kiszámoltuk már az általánosított inverzeit, és most ezek segítségével szeretnénk meghatározni az ellentmondásos egyenletrendszert valamilyen értelemben legjobban közelítő megoldást. Ilyenkor egy olyan x vektort keresünk, ami a legközelebb áll ahhoz, hogy teljesítse az összes egyenletet abban az értelemben, hogy $\|Ax - b\|^2$ értéke a lehető legkisebb legyen.

Az ortogonalizációs eljárás segítségével egy n-dimenziós V euklideszi tér W alterének ortonormált bázisát kiegészíthetjük a tér ortonormált bázisává. Így egy tetszőleges $b \in V$ de $b \notin W$ vektor felírható $b = b_1 + b_2$ alakban, ahol $b_1 \in W$, a b_2 pedig merőleges a W minden elemére. Ezt alkalmazhatjuk ellentmondásos lineáris algebrai egyenletrendszerre:

$$Ax = b,\ A = \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix} \in \mathbb{C}^{n\times m}, \qquad b \in \mathbb{C}^n, \qquad x \in \mathbb{C}^m, \text{ amikor } b \notin \langle a_1, \dots, a_m \rangle.$$

Mivel $Ax, b_1 \in W$, ezért felírható:

$$\|Ax - b\|^2 = \|(Ax - b_1) + (-b_2)\|^2 = \|Ax - b_1\|^2 + \|-b_2\|^2 \geq \|b_2\|^2.$$

Az egyenlőség teljesülésének feltétele az, hogy $\|Ax - b_1\| = 0$, azaz $Ax = b_1$. Ezt a lineáris egyenletrendszert meg lehet oldani, hiszen $b_1 \in W = \langle a_1, \dots, a_m \rangle$. Ennek tetszőleges megoldása adja az eredeti $Ax = b$ egyenletrendszernél az euklideszi normában vett legjobb közelítést.

3.3.1. Tétel $x = A^{(g)} b_1$ megoldás.

Mivel végtelen sok általánosított inverz létezik, ezért válasszunk egy olyat, melynek létezése egyértelmű.

3.3.2. Definíció $A \in \mathbb{C}^{n \times m}$-es mátrixnak Moore-Penrose-féle pszeudoinverze egy olyan $A^+ \in \mathbb{C}^{m \times n}$-es mátrix, melyre:

- $AA^+A = A$
- $A^+AA^+ = A^+$
- AA^+ önadjungált
- A^+A önadjungált.

3.3.4. Tétel Ha az $A^+ \in \mathbb{C}^{m \times n}$ mátrix rendelkezik a Moore–Penrose-féle pszeudoinverz tulajdonságokkal, akkor $x = A^+b$ jó megoldás, azaz $A^+b_1 = A^+b$.

Bizonyítás Nyilván a tétel bizonyításához az is elegendő, ha azt megmutatjuk, hogy $A^+b_2 = 0$, hiszen $A^+b = A^+b_1 + A^+b_2$. A Moore–Penrose-féle pszeudoivenz tulajdonságok miatt az A^+ mátrixra a következő négy tulajdonságnak kell teljesülni:

1. $AA^+A = A$
2. $A^+AA^+ = A^+$
3. $(A^+A)^* = A^+A$
4. $(AA^+)^* = AA^+$

Ekkor a második tulajdonság miatt $A^+b_2 = A^+AA^+b_2$.
A negyedik tulajdonságot felhasználva $A^+(AA^+)b_2 = A^+(AA^+)^*b_2 = A^+(A^+)^*A^*b_2$.

$A^* b_2 = \begin{bmatrix} a_1^* \\ \vdots \\ a_n^* \end{bmatrix} b_2 = \begin{bmatrix} a_1^* b_2 \\ \vdots \\ a_n^* b_2 \end{bmatrix} = \begin{bmatrix} \langle a_1, b_2 \rangle \\ \vdots \\ \langle a_n, b_2 \rangle \end{bmatrix}$. Mivel a b_2 vektor merőleges az $a_1, \dots, a_n$ vektorok mindegyikére, így igazoltuk, hogy $A^+ b_2 = 0$, tehát $A^+ b_1 = A^+ b$.

A pszeudoinverz kiszámításához először szükséges bevezetni néhány fogalmat:

3.3.5. Definíció Legyen $x = (x_1, x_2, \dots, x_n)^T$ és $y = (y_1, y_2, \dots, y_n)^T$ két $\mathbb{R}^n$-beli vektor. Ekkor az xy^T szorzatot a két vektor diadikus szorzatának (vagy röviden diádnak) nevezzük, melynek eredménye egy $\mathbb{R}^{n \times n}$-es mátrix.

Mivel minden mátrix felírható diádok összegeként, ez sugallja azt az ötletet, hogy beszéljünk egy mátrix minimális diadikus előállításáról.

3.3.6. Definíció Ha egy mátrixot a lehető legkevesebb diád összegeként állítunk elő, akkor azt a mátrix minimális diadikus előállításának nevezzük.

Ahhoz, hogy előállítsuk egy mátrix minimális diadikus felbontását, vezessük be az A mátrix $a_{ij} \neq 0$ eleme által generált diád fogalmát: Ezen a mátrix j-edik oszlopának és i-edik sorának szorzatát értjük, osztva az a_{ij} elemmel, azaz:

$$\frac{a_j a_i}{a_{ij}} = \frac{A e_j e_i^T A}{e_i^T A e_j}; \qquad a_{ij} = e_i^T A e_j \neq 0.$$

Tegyük fel, hogy ennek a mátrixnak a bal felső eleme nem nulla, hiszen ha nulla, akkor rangtartó átalakítások sorozatával átrendezhető a mátrix. Vonjuk ki az A mátrixból az a_{11} által generált diádot:

$$A^{(2)} = A - \frac{A e_1 e_1^T A}{e_1^T A e_1}.$$

Észrevehető, hogy az $A^{(2)}$ mátrix első sora és első oszlopa is csupa zérus elemeket tartalmaz. Most tegyük fel, hogy az $A^{(2)}$ mátrix első eleme nem nulla, majd vonjuk ki az $A^{(2)}$ mátrixból az első eleme által generált diádot:

$$A^{(3)} = A^{(2)} - \frac{A^{(2)} e_2 e_2^T A^{(2)}}{e_2^T A^{(2)} e_2}.$$

Ennek a mátrixnak már az első két sora és első két oszlopa is csupa nulla elemet tartalmaz. Ezt az eljárást addig folytathatjuk, amíg a nullmátrixot meg nem kapjuk. Tegyük fel, hogy ezt s lépés után érjük el, ekkor:

$$A^{(s)} - \frac{A^{(s)} e_s e_s^T A^{(s)}}{e_s^T A^{(s)} e_s} = 0.$$

Az eljárás az alábbi rekurzióval írható fel:

$$A^{(k+1)} = A^{(k)} - \frac{A^{(k)} e_k e_k^T A^{(k)}}{e_k^T A^{(k)} e_k}; \qquad A^{(1)} = A; \qquad A^{(s+1)} = 0.$$

Összeadva az egyenlőségeket $k = 1,2, \ldots, s$-re és rendezve a következő egyenlet írható fel:

$$A = \sum_{k=1}^{s} \frac{A^{(k)} e_k e_k^T A^{(k)}}{e_k^T A^{(k)} e_k},$$

tehát az A mátrixot s diád összegére bontottuk. Meggondolható, hogy a minimális felbontás esetén s pont a mátrix rangjával lesz egyenlő. A továbbiakban jelöljük ezeket a diádokat

$$u_k v_k^T, \qquad k = 1, \ldots, s$$

alakban. Mivel tudjuk, hogy s diád összege mindig felírható a diádok oszlopvektoraiból és sorvektoraiból alkotott s oszlopos, illetve s soros mátrix szorzataként, ezzel tehát az A mátrixot faktorizáltuk:

$$A = \sum_{k=1}^{s} u_k v_k^T = UV^T.$$

3.3.7. Tétel Bármely A mátrix egy minimális diadikus előállítását megkapjuk, ha azt a fenti rekurzív összeg segítségével diádösszegként írjuk fel.

3.3.8. Tétel Legyen A egy $m \times n$ típusú valós elemű mátrix, amelynek a rangja r. Ha a mátrix egy minimális diadikus előállítása:

$$A = UV^T,$$

akkor a Moore–Penrose-féle inverze:

$$A^+ = V(V^TV)^{-1}(U^TU)^{-1}U^T.$$

Bizonyítás A Moore–Penrose-féle inverz mind a négy tulajdonságot kielégíti A és A^+, beszorzással meggyőződhetünk erről.

3.3.9. A Moore–Penrose-féle pszeudoinverz kiszámítása

Visszatérve a feladat megválaszolásához, számítsuk ki a feladatban szereplő A mátrix pszeudoinverzét. Kezdetben az A mátrix:

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 2 & 1 & 0 \\ 1 & 2 & 1 \end{pmatrix}.$$

Most vegyük az a_{11} elem által generált diádot, majd vonjuk ki ezt az A mátrixból:

$$A^{(2)} = A - \frac{Ae_1e_1^TA}{e_1^TAe_1} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & -2 \\ 0 & 2 & 0 \end{pmatrix}.$$

Ez után az vegyük az $A^{(2)}$ mátrix a_{22} eleme által generált diádot, és vonjuk ki ezt $A^{(2)}$-ből:

$$A^{(3)} = A^{(2)} - \frac{A^{(2)}e_2e_2^TA^{(2)}}{e_2^TA^{(2)}e_2} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -3 \\ 0 & 0 & -2 \end{pmatrix}.$$

Végül az $A^{(3)}$ mátrix a_{33} eleme által generált diádot kivonva az $A^{(3)}$ mátrixból valóban a nullmátrixot kapjuk. Tehát az A mátrix faktorizációja:

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 1 & 1 \\ 0 & 2 & \frac{2}{3} \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & -3 \end{pmatrix}.$$

Ebből a Matlab program segítségével a Moore–Penrose-féle pszudoinverz 4 értékes jegyre kerekítve:

$$A^{+} = \begin{pmatrix} 0,3 & -0,2 & 0,4 & -0,1 \\ -0,5333 & 0,1333 & 0,0667 & 0,4 \\ 0,6667 & 0,3333 & -0,3333 & 0 \end{pmatrix}.$$

Tehát a tétel szerint $x = A^{+}b$ megoldás, azaz:

$$x = \begin{pmatrix} 0,3 & -0,2 & 0,4 & -0,1 \\ -0,5333 & 0,1333 & 0,0667 & 0,4 \\ 0,6667 & 0,3333 & -0,3333 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -0.2 \\ 0.1333 \\ 1.3333 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

4. Sztochasztikus mátrixok

4.1. Példa

Nézzük először a következő feladatot:

Egy város lakásait három kategóriába sorolják állapotuk szerint: elhanyagolt, átlagos, és kitűnő. A jelenlegi finanszírozási lehetőségek mellett az éves statisztikák azt mutatják, hogy az elhanyagolt lakások 60 százaléka elhanyagolt marad az év végére is, 30 százalékukat felújítják átlagos állapotúra, a maradék 10 százalékot pedig kitűnő állapotúra. Az átlagos állapotú lakások 20 százaléka leromlik elhanyagoltra, 70 százaléka marad átlagos állapotú, a maradék 10 százalékot pedig felújítják. Végül a kitűnő állapotú lakások 80 százaléka kitűnő állapotban marad, a maradék 20 százalék leromlik átlagosra az év végére. Hosszú távon mi lesz a lakások állapotának eloszlása?

Ezen a feladaton a mátrix diagonalizálhatóságának hasznosságát szeretném megmutatni. A feladat érdekessége abban rejlik, hogy nem tudjuk, mi a kiindulási állapot, de mégis lehet mondani valamit arra, hogy a lakások állapotára milyen eloszlás várható hosszú távon. A feladat tehát a következő alakban írható:

$$\begin{pmatrix} e_{n+1} \\ a_{n+1} \\ k_{n+1} \end{pmatrix} = \begin{pmatrix} 0,6 & 0,2 & 0 \\ 0,3 & 0,7 & 0,2 \\ 0,1 & 0,1 & 0,8 \end{pmatrix} * \begin{pmatrix} e_n \\ a_n \\ k_n \end{pmatrix},$$

ahol $e_n,\ a_n,\ k_n$ rendre a n-edik év után lévő elhanyagolt, átlagos, és kitűnő állapotban lévő házak számát jelenti, nyilván $e_1 + a_1 + k_1 = 1$, $e_{n+1},\ a_{n+1},\ k_{n+1}$ pedig a $n + 1$-edik év után lévő állapotot írja le n segítségével, nyilván $n \in \mathbb{N}$. Legyen

$$v_1 = \begin{pmatrix} e_1 \\ a_1 \\ k_1 \end{pmatrix}, \qquad v_n = \begin{pmatrix} e_n \\ a_n \\ k_n \end{pmatrix}, \qquad v_{n+1} = \begin{pmatrix} e_{n+1} \\ a_{n+1} \\ k_{n+1} \end{pmatrix}, \qquad A = \begin{pmatrix} 0,6 & 0,2 & 0 \\ 0,3 & 0,7 & 0,2 \\ 0,1 & 0,1 & 0,8 \end{pmatrix},$$

ekkor a fenti egyenlet a következő alakban írható fel:

$$v_{n+1} = Av_n.$$

Ebből a következő összefüggés adódik:

$$v_{n+1} = Av_n = A^2 v_{n-1} = \cdots = A^n v_1.$$

Így elegendő az A mátrix hatványaira explicit képletet adni, melyhez az A mátrixot bázistranszformációval diagonális alakra hozom. Az A karakterisztikus polinomja:

$$k(x) = \begin{vmatrix} 0,6 - x & 0,2 & 0 \\ 0,3 & 0,7 - x & 0,2 \\ 0,1 & 0,1 & 0,8 - x \end{vmatrix} = -x^3 + 2,1x^2 - 1,38x + 0,28.$$

Mivel $-x^3 + 2,1x^2 - 1,38x + 0,28 = -(x-1)(x-0,4)(x-0,7)$, ezért ennek gyökei, vagyis az A sajátértékei:

$$\lambda_1 = 1, \qquad \lambda_2 = 0,4, \qquad \lambda_3 = 0,7.$$

Legyen w_{λ_i} az i-edik sajátértékhez tartozó sajátvektor. Ekkor w_{λ_i} $(i = 1,2,3)$ kielégíti a következő egyenletet:

$$Aw_{\lambda_i} = \lambda_i w_{\lambda_i}.$$

λ_1-re a következő egyenletrendszer adódik:

$$\begin{aligned} 0{,}6w_1 + 0{,}2w_2 \qquad\qquad &= w_1 \\ 0{,}3w_1 + 0{,}7w_2 + 0{,}2w_3 &= w_2 \\ 0{,}1w_1 + 0{,}1w_2 + 0{,}8w_3 &= w_3 \end{aligned}$$

melynek megoldása: $w_{\lambda_1} = \mu_1 \begin{pmatrix} 2 \\ 4 \\ 3 \end{pmatrix}$, ahol $\mu_1 \in \mathbb{R}$.

A λ_2-höz tartozó sajátvektor $w_{\lambda_2} = \mu_2 \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$, $\mu_1 \in \mathbb{R}$, a A λ_3-hoz tartozó sajátvektor pedig $w_{\lambda_3} = \mu_3 \begin{pmatrix} 2 \\ 1 \\ -3 \end{pmatrix}$, $\mu_3 \in \mathbb{R}$. A megfelelő sajátvektorokat egy mátrix oszlopaiba írva kapjuk a diagonalizálást elvégző bázistranszformáció mátrixát:

$$S = \begin{pmatrix} 2 & 1 & 2 \\ 4 & -1 & 1 \\ 3 & 0 & -3 \end{pmatrix}$$

Mivel $S^{-1}AS = \mathrm{diag}(\lambda_i)$, ezért:

$$\begin{pmatrix} 2 & 1 & 2 \\ 4 & -1 & 1 \\ 3 & 0 & -3 \end{pmatrix}^{-1} * \begin{pmatrix} 0{,}6 & 0{,}2 & 0 \\ 0{,}3 & 0{,}7 & 0{,}2 \\ 0{,}1 & 0{,}1 & 0{,}8 \end{pmatrix} * \begin{pmatrix} 2 & 1 & 2 \\ 4 & -1 & 1 \\ 3 & 0 & -3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0{,}4 & 0 \\ 0 & 0 & 0{,}7 \end{pmatrix} = D.$$

Innen $D^n = S^{-1}ASS^{-1}AS \dots S^{-1}AS = S^{-1}A^nS$. Mivel a mátrixszorzás asszociatív művelet, ezért $SS^{-1} = E$ egységmátrix, így kifejezhető:

$$A^n = SD^nS^{-1}.$$

S^{-1} meghatározásához a már korábban használt $S^{-1} = \frac{1}{\det(S)} * \mathrm{adj}(S)$ képletet használva a következő adódik:

$$\det(S) = 27, \ \mathrm{adj}(S) = \begin{pmatrix} 3 & 3 & 3 \\ 15 & -12 & 6 \\ 3 & 3 & -6 \end{pmatrix}, \text{ tehát:}$$

$$S^{-1} = \frac{1}{27} * \begin{pmatrix} 3 & 3 & 3 \\ 15 & -12 & 6 \\ 3 & 3 & -6 \end{pmatrix} = \frac{1}{9}\begin{pmatrix} 1 & 1 & 1 \\ 5 & -4 & 2 \\ 1 & 1 & -2 \end{pmatrix}.$$

Ismert, hogy diagonális mátrixot elemenként lehet hatványozni, így:

$$A^n = \frac{1}{9}\begin{pmatrix} 2 & 1 & 2 \\ 4 & -1 & 1 \\ 3 & 0 & -3 \end{pmatrix} * \begin{pmatrix} {\lambda_1}^n & 0 & 0 \\ 0 & {\lambda_2}^n & 0 \\ 0 & 0 & {\lambda_3}^n \end{pmatrix} * \begin{pmatrix} 1 & 1 & 1 \\ 5 & -4 & 2 \\ 1 & 1 & -2 \end{pmatrix}.$$

Elvégezve a mátrixszorzást a következő formula adódik:

$$A^n = \frac{1}{9}\begin{pmatrix} 2\lambda_1^n + 5\lambda_2^n + 2\lambda_3^n & 2\lambda_1^n - 4\lambda_2^n + 2\lambda_3^n & 2\lambda_1^n + 2\lambda_2^n - 4\lambda_3^n \\ 4\lambda_1^n - 5\lambda_2^n + \lambda_3^n & 4\lambda_1^n + 4\lambda_2^n + \lambda_3^n & 4\lambda_1^n - 2\lambda_2^n - 2\lambda_3^n \\ 3\lambda_1^n - 3\lambda_3^n & 3\lambda_1^n - 3\lambda_3^n & 3\lambda_1^n + 6\lambda_3^n \end{pmatrix}$$

Mivel $v_{n+1} = A^n v_1$, ezért a lakások állapotára n év elteltével a következő képletek írhatóak fel:

$$e_{n+1} = \frac{1}{9}[(2\lambda_1^n + 5\lambda_2^n + 2\lambda_3^n)e_1 + (2\lambda_1^n - 4\lambda_2^n + 2\lambda_3^n)a_1 + (2\lambda_1^n + 2\lambda_2^n - 4\lambda_3^n)k_1]$$

$$a_{n+1} = \frac{1}{9}[(4\lambda_1^n - 5\lambda_2^n + \lambda_3^n)e_1 + (4\lambda_1^n + 4\lambda_2^n + \lambda_3^n)a_1 + (4\lambda_1^n - 2\lambda_2^n - 2\lambda_3^n)k_1]$$

$$k_{n+1} = \frac{1}{9}[(3\lambda_1^n - 3\lambda_3^n)e_1 + (3\lambda_1^n - 3\lambda_3^n)a_1 + (3\lambda_1^n + 6\lambda_3^n)k_1]$$

Azt szeretnénk kiszámolni, hogy a lakások végső megoszlása milyen érték felé fog tartani. Ehhez nézzük meg először az A^n sajátértékeinek határértékét, ha a n-nel tartunk a végtelenbe:

$$\lim_{n\to\infty} \lambda_1^n = \lim_{n\to\infty} 1^n = 1$$

$$\lim_{n\to\infty} \lambda_2^n = \lim_{n\to\infty} 0{,}4^n = 0$$

$$\lim_{n\to\infty} \lambda_3^n = \lim_{n\to\infty} 0{,}7^n = 0.$$

Ebből látható, hogy a λ_1^n sajátérték lesz a domináns, a többi sajátérték a nullához konvergál. Ebből következik, hogy:

$$\lim_{n\to\infty} e_n = \frac{2}{9}(e_1 + a_1 + k_1)$$

$$\lim_{n\to\infty} a_n = \frac{4}{9}(e_1 + a_1 + k_1)$$

$$\lim_{n\to\infty} k_n = \frac{3}{9}(e_1 + a_1 + k_1).$$

Mivel $e_1 + a_1 + k_1 = 1$, ezért a lakások eloszlására a következő mondható hosszú távon:

- Az elhanyagolt lakások száma a lakások számának $^2/_9$-szerese felé fog tartani
- Az átlagos lakások száma a lakások számának $^4/_9$-szerese felé fog tartani
- A kitűnő állapotban lévő lakások száma a lakások számának $^1/_3$-szorosa felé fog tartani.

Vegyük észre, hogy a lakások eloszlása nem függ a kezdeti értéktől, azaz bármennyi elhanyagolt, átlagos, és kitűnő állapotú lakás van kezdetben, az eloszlások mindig ezekhez a számolt értékekhez fognak tartani. Az a tény, hogy az 1 sajátértéke a feladatnak, könnyen belátható, hiszen ekkor:

$$Av = v, \qquad v \neq 0.$$

Azaz:

$$\begin{pmatrix} a_{11} - 1 & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} - 1 \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = 0.$$

Ez teljesül is, hiszen ennek a mátrixnak a determinánsa nulla, mivel a sorok összege 0. A feladat jellegéből adódóan a többi sajátérték abszolútértéke pedig nem lehet nagyobb 1-nél, különben a határértékek divergálnának, és a feladatnak nem létezne megoldása. Most tehát nézzük meg az emögött rejlő általános elméletet.

4.2. Sztochasztikus mátrixok elmélete

4.2.1. Definíció Egy valós $n \times n$-es P mátrixot elemenként pozitívnak (nemnegatívnak) nevezünk, ha annak minden p_{ij} eleme pozitív (nemnegatív) valós szám.

4.2.2. Definíció Egy $n \times n$-es A mátrix sztochasztikus, ha elemei nemnegatívak, és az egy oszlopban álló elemeinek összege 1, azaz:

$$\sum_{i=1}^{n} a_{ij} = 1.$$

A természetben gyakran játszódnak le olyan folyamatok, melyeket sztochasztikus mátrixokkal lehet leírni. Ilyen lehet például a biológiában n különböző faj vizsgálata, melyeknek lehetőségük van átváltozni egy másikká. Az a_{ij} számokat átmenet valószínűségnek nevezzük; a populáció azon részét reprezentálják, melyek a j-edik fajból az i-edik fajba mennek át. Ebben az interpretációban a feltétel, miszerint az egy oszlopban álló elemek összege 1, azt jelenti, hogy a teljes populáció megmarad.

4.2.3. Tétel (Perron) Minden pozitív *P* mátrixnak van egy domináns sajátértéke, amelyet $\lambda(P)$ jelöl, és a következő tulajdonságokkal rendelkezik:

a) $\lambda(P)$ pozitív, és a hozzátartozó v sajátvektor komponensei is pozitívak:

$$Pv = \lambda(P)v, \qquad v > 0.$$

b) $\lambda(P)$ egyszeres sajátérték.
c) P minden más κ sajátértéke abszolút értékben kisebb, mint $\lambda(P)$:

$$|\kappa| < \lambda(P).$$

d) A P mátrixnak nincs másik nemnegatív komponensű sajátvektora.

Bizonyítás Jelölje $p(P)$ azon λ nemnegatív számok halmazát, amelyek esetén létezik olyan $x \neq 0$ vektor, hogy:

$$Px \geq \lambda x, \qquad x \geq 0.$$

4.2.3.1. Lemma Pozitív P esetén:

i. $p(P)$ nem üres, és tartalmaz pozitív számot,
ii. $p(P)$ korlátos,
iii. $p(P)$ zárt.

Bizonyítás Legyen x egy tetszőleges, pozitív vektor. Mivel P pozitív, ezért Px is pozitív vektor. Ha λ elég kicsi, akkor $Px \geq \lambda x$ teljesül, ezzel igazoltuk a lemma i. részét. Mivel az egyenlőtlenség mindkét oldala lineáris függvénye az x vektornak, ezért x normálható úgy, hogy:

$$\xi x = \sum x_i = 1, \qquad \xi = (1, \dots, 1).$$

Ha az egyenlőtlenséget balról szorozzuk ξ vektorral, akkor a következőt kapjuk:

$$\xi P x \geq \lambda \xi x = \lambda.$$

Jelölje b a ξP vektor legnagyobb komponensét, ekkor:

$$b\xi \geq \xi P.$$

Ezt behelyettesítve a fentebbi egyenlőtlenségbe az adódik, hogy $b \geq \lambda,$ ezzel a lemma ii. részét is igazoltuk.

Most tekintsünk egy λ_n sorozatot a $p(P)$ halmazból. A definíció szerint létezik olyan $x_n \neq 0$ sorozat, amelyre teljesül:

$$Px_n \geq \lambda_n x_n, \qquad x_n \geq 0.$$

Ismét éljünk azzal a feltevéssel, hogy az x_n vektorok normálva vannak, azaz

$$\xi x_n = 1.$$

Mivel a normált x_n vektorok halmaza az $\mathbb{R}^n$ tér zárt korlátos részhalmazát alkotják, ezért ez a halmaz kompakt. Így egy x_n részsorozat egy nemnegatív x vektorhoz konvergál, ami szintén normálva van, és λ_n a λ értékhez tart. Emiatt λ és x kielégíti a $Px \geq \lambda x$ egyenlőtlenséget, ezért $p(P)$ zárt. Ezzel beláttuk a lemma iii. részét.

Megmutattuk tehát a lemmánkkal, hogy $p(P)$ zárt és korlátos, tehát létezik maximuma, legyen ez $\lambda_{\max}$. Az i. rész miatt tudjuk, hogy $\lambda_{\max} > 0$. Mutassuk meg, hogy ez lesz a domináns sajátérték! Tudjuk, hogy $Px \geq \lambda_{\max}x$, ezért létezik egy olyan v vektor, melyre:

$$Pv \geq \lambda_{\max}v, \qquad v > 0.$$

4.2.3.2 Állítás $Pv = \lambda_{\max}v.$

Bizonyítás Mivel $Pv \geq \lambda_{\max}v,$ ezért a k-adik komponensre a következők teljesülnek:

$$\sum p_{ij}v_j \geq \lambda_{\max}v_i, \qquad ha\ i \neq k$$
$$\sum p_{kj}v_j > \lambda_{\max}v_k.$$

Legyen az x vektor a következő:

$$x = v + \varepsilon e_k,$$

ahol $\varepsilon > 0$ és az e_k vektor k-adik komponense 1, a többi 0. Mivel P pozitív, ezért teljesül a $Px > Ph,$ hiszen a bal oldal minden komponensét növeltük, viszont a jobb oldalnak csak a k-adik komponense nő. Így következik, hogy:

$$Px > \lambda_{\max}x.$$

Tehát egy szigorú egyenlőtlenséget kaptunk x-re, így elég kis $\delta > 0$ számra $\lambda_{\max}$ helyére $\lambda_{\max} + \delta$ írható, így $\lambda_{\max}$ nem maximális, tehát ellentmondásra jutottunk. Továbbá észrevehető az is, hogy $\lambda_{\max}$ sajátértéke a P mátrixnak egy megfelelő nemnegatív v sajátvektorral. Sőt, belátható, hogy v pozitív, hiszen P pozitív, és $v \geq 0$, ezáltal $Pv > 0$. Mivel $Pv = \lambda_{\max} v$, ezért $v > 0$. Ezzel tehát a Perron-tétel *a*) részét beláttuk.

Most mutassuk meg, hogy $\lambda_{\max}$ egyszeres sajátérték. Mivel $Pv = \lambda_{\max} v$, ezért P minden $\lambda_{\max}$ sajátértékhez tartozó sajátvektora v többszöröse. Tegyük fel indirekt, hogy y egy sajátvektor, ami nem v többszöröse, akkor valamilyen $c \neq 0$ választással lenne olyan $v + cy$ vektor, amelyre $v + cy \geq 0$, de $v + cy$ valamelyik komponense 0. Ez a fenti gondolatmenetnek ellentmond, miszerint P nemnegatív sajátvektorai valójában pozitívak.

Most igazoljuk, hogy a P mátrixnak nem létezik a $\lambda_{\max}$ sajátértékhez tartozó általánosított sajátvektora, azaz olyan y vektor, melyre

$$Py = \lambda_{\max} y + cv.$$

Hogy biztosítsuk y pozitivitását, y helyére írható $y + bv$, valamint ha $c < 0$, akkor y helyére $-y$ írható. Ekkor $y = \lambda_{\max} y + cv$ és $v > 0$ miatt $Py > \lambda_{\max} y$. Ekkor egy elég kicsi $\delta > 0$-ra:

$$Py > (\lambda_{\max} + \delta) y,$$

ami ellentmond annak, hogy $\lambda_{\max}$ a $p(P)$ legnagyobb eleme. Ezzel beláttuk a Perron-tétel *b*) részét.

A c) bizonyításához tegyük fel indirekt, hogy $\kappa \in \mathbb{C}$ a P mátrix egy sajátértéke, ami nem egyenlő a $\lambda_{\max}$ sajátértékkel, továbbá legyen $y \in \mathbb{C}^l$ a hozzá tartozó sajátvektor. Ekkor:

$Py = \kappa y,$ azaz koordinátánként $\sum_j p_{kj} y_j = \kappa y_k$.

A komplex számokra és abszolút értékükre alkalmazva a háromszög-egyenlőtlenséget a következőt kapjuk:

$$\sum_j p_{kj} |y_j| \geq \left| \sum_j p_{kj} y_j \right| = |\kappa| |y_k|.$$

Felhasználva a $Px \geq \lambda x$ egyenlőtlenséget látjuk, hogy $|\kappa| \in p(P)$. Ha $|\kappa| = \lambda_{\max}$ teljesülne, akkor az

$$\begin{pmatrix} |y_1| \\ \vdots \\ |y_l| \end{pmatrix}$$

vektor a $\lambda_{\max}$ sajátértékhez tartozó sajátvektor, ezért v többszöröse lenne:

$$|y_i| = cv_i,$$

továbbá teljesülne, hogy:

$$\sum_j p_{kj}|y_j| = \left|\sum_j p_{kj}y_j\right|.$$

Tudjuk, hogy ez komplex számok esetén csak akkor teljesül, ha minden y_k argumentuma megegyezik, tehát:

$$y_k = |y_k|(\cos(\varphi) + i\sin(\varphi)), \qquad k = 1, \dots, l.$$

Ebből következik tehát, hogy:

$$y_k = cv_k(\cos(\varphi) + i\sin(\varphi)),$$

azaz:

$$y = c(\cos(\varphi) + i\sin(\varphi))v.$$

Ez azt jelenti, hogy $\kappa = \lambda_{\max}$, ezáltal a *c*) részt beláttuk.

A *d*) részt bizonyítás nélkül közlöm.

Láttuk tehát, hogy egy pozitív mátrixnak van egy domináns, egyszeres, pozitív sajátértéke, mely a sajátértékek közül a legnagyobb. A korábban megoldott feladatnál észrevettük, hogy a domináns sajátérték 1.

4.2.4. Tétel Legyen $S \in \mathbb{R}_{+}^{n \times n}$ egy pozitív sztochasztikus mátrix. Ekkor:

- A domináns sajátérték $\lambda(S) = 1$.
- Minden nemnegatív x vektorra:

$$\lim_{N \to \infty} S^N x = cv,$$

  ahol $c > 0$ és v a domináns sajátvektor.

Bizonyítás Mivel $S > 0$, ezért $S^T > 0$, ahol S^T az S mátrix transzponáltját jelenti. A sztochasztikus mátrixra igaz, hogy ha:

$$Sx = y\lambda,$$

ahol $x \in \mathbb{R}^n$, $x \neq [0, \dots, 0]$, akkor x az S jobb oldali sajátvektora $\lambda \in \mathbb{R}$, $\lambda \neq 0$ sajátértékkel, és:

$$yS = \mu y,$$

ahol $y \in \mathbb{R}^n$, $y \neq [0, \dots, 0]$, akkor y az S bal oldali sajátvektora $\mu \in \mathbb{R}$, $\mu \neq 0$ sajátértékkel, akkor:

$$yS = \mu y \Leftrightarrow S^T y^T = \mu y^T,$$

tehát a bal oldali sajátvektor és sajátérték keresése visszavezethető a jobb oldali sajátvektor és sajátérték keresésére. Mivel az S sztochasztikus mátrixra teljesül, hogy $\sum_{i=1}^{n} s_{ij} = 1$, ezért az egységvektor nyilván ennek bal oldali sajátvektora lesz 1 sajátértékkel, hiszen:

$$eS = 1 * S,$$

ahol $e \in \mathbb{R}^n$, $e = [1, \dots, 1]$, ahol $eS = S$, mivel a sztochasztikus mátrix bármely oszlopában lévő elemeinek összege 1. Tehát e bal oldali sajátvektor lesz 1 sajátértékkel, így e^T jobb oldali sajátvektor lesz ugyancsak 1 sajátértékkel. A Perron-tétel miatt az egységvektor lesz a domináns sajátvektor, és 1 lesz a hozzá tartozó domináns sajátérték.

A második rész igazolásához fejtsük ki az x vektort S sajátvektorainak összegeként:

$$x = \sum_j c_j v_j.$$

Tegyük fel, hogy S sajátvektorai nem általánosított sajátvektorok. Az S mátrix λ sajátértékéhez tartozó x vektor *általánosított sajátvektor*, ha valamilyen k természetes számra:

$$(S - \lambda I)^k x = 0.$$

Ekkor a következő írható:

$$S^N x = \sum_j c_j \lambda_j^N v_j.$$

Legyen λ_1 a domináns sajátérték, azaz $\lambda_1 = 1, \left|\lambda_j\right| < 1$, ha $j \neq 0$. Ebből tehát adódik, hogy:

$$S^N x \to cv,$$

ahol v a domináns sajátvektor. Továbbá belátható az is, hogy $c > 0$.

Ennek a tételnek alkalmazásait olyan rendszerekre alkalmazhatjuk leginkább, ahol a rendszer megváltozását az átmenetvalószínűségek adják. A fejezet elején megoldott feladatra alkalmazzuk ezt a tételt, azt kapjuk, hogy a lakások új megoszlása a régi megoszlás szerint a következőképpen alakul:

$$v_i = \sum_j s_{ij} x_j, \qquad i, j = 1,2,3$$

ahol jelöljük most v_i-vel az átlagos, elhanyagolt, és kitűnő állapotban lévő lakások új eloszlását, x_j az előző évi megoszlást jelenti. Ez felírható a mátrixok nyelvén:

$$v = Sx,$$

ahol $v \in \mathbb{R}^3,\ S \in \mathbb{R}^{3\times 3},\ x \in \mathbb{R}^3$. N év elteltével a lakások megoszlásának vektora $S^N x$ lesz. A tétel jelentősége ebben az alkalmazásban azt mutatja, hogy ahogy $N \to \infty$, a lakások eloszlása egy állandó értékhez tart, ami nem függ attól, hogy kezdetben milyen volt a lakások eloszlása.

4.3. Nemnegatív mátrixok

Fontos megjegyezni, hogy a feladatban nem pozitív volt a mátrix, hiszen volt egy nulla eleme. Vizsgáljuk meg az alábbi nemnegatív mátrixokat, hogy létezik-e domináns sajátértékük!

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

$$B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

$$C = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix},$$

$$D = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0{,}9 \\ 0 & 0 & 0{,}1 \end{pmatrix}.$$

Az A mátrix nem sztochasztikus és van egy domináns sajátértéke, a B sztochasztikus mátrixnak sajátértékei ± 1. Így egyik sem domináns, a mátrix hatványának paritása dönti el, hogy melyik sajátérték lesz a domináns, tehát $B^N x$ alternáló viselkedésű. A C nem sztochasztikus mátrixnak az 1 kétszeres sajátértéke. A D sztochasztikus mátrixnak sajátértékei ± 1 és $0{,}1$, ennek sincs domináns sajátértéke. Hasonló tulajdonságot láthatunk a B és a D mátrix között, ám egy adott probléma által meghatározott D mátrix esetén egyszer a $+1$, egyszer a -1 sajátérték lesz a domináns, ám sosem fogják ezt elérni, nagy N esetén $D^N x$ aszimptotikusan periodikus viselkedésű. Látható, hogy a nemnegatív mátrixok esetén már nem feltétlenül létezik egyértelmű domináns sajátérték. A következő tételt bizonyítás nélkül közlöm:

4.3.1. Tétel (Frobenius): Minden nemnegatív $F \in \mathbb{R}^{l \times l}$ mátrixnak $F \neq 0$ esetén van egy $\lambda(F)$ sajátértéke a következő tulajdonságokkal:

a) $\lambda(F)$ nemnegatív, és a hozzá tartozó sajátvektor komponensei is nemnegatívak:
$$Fv = \lambda(F)v, \qquad v \geq 0.$$
b) Az összes többi κ sajátértékre teljesül, hogy:
$$|\kappa| \leq \lambda(F).$$
c) Ha $|\kappa| = \lambda(F)$, akkor:

$$\kappa = \lambda(F) * \left(\cos\left(\frac{2\pi k}{m}\right) + i * \sin\left(\frac{2\pi k}{m}\right)\right),$$

ahol k és m pozitív egész számok, és $m \leq l$.

Továbbá belátható az is, hogy ha egy A nemnegatív mátrix A^m hatványa pozitív valamilyen m természetes számra, akkor A mátrixnak létezik domináns sajátértéke. A feladatban szereplő mátrixnak a második hatványa:

$$A^2 = \begin{pmatrix} 0,6 & 0,2 & 0 \\ 0,3 & 0,7 & 0,2 \\ 0,1 & 0,1 & 0,8 \end{pmatrix}^2 = \begin{pmatrix} 0,42 & 0,26 & 0,04 \\ 0,41 & 0,57 & 0,3 \\ 0,17 & 0,17 & 0,66 \end{pmatrix},$$

tehát A^2 pozitív sztochasztikus mátrix, melynek domináns sajátértéke az 1. Érdekességképpen a mátrix néhány további hatványát is kiszámítottam:

$$A^5 = \begin{pmatrix} 0,2653 & 0,255 & 0,1498 \\ 0,4574 & 0,4677 & 0,4048 \\ 0,2773 & 0,2773 & 0,4454 \end{pmatrix}$$

$$A^{10} = \begin{pmatrix} 0,2286 & 0,2285 & 0,2097 \\ 0,4475 & 0,4476 & 0,4381 \\ 0,3239 & 0,3239 & 0,3522 \end{pmatrix}$$

$$A^{20} = \begin{pmatrix} 0,2224 & 0,2224 & 0,2219 \\ 0,4445 & 0,4445 & 0,4443 \\ 0,3331 & 0,3331 & 0,3339 \end{pmatrix}.$$

Látható tehát, hogy a mátrix oszlopai a $\begin{pmatrix} 2/9 \\ 4/9 \\ 1/3 \end{pmatrix}$ vektorhoz tartanak, ami épp az eloszlás lesz.

Sztochasztikus mátrixokkal tehát leírhatók azon folyamatok, melyek egy eloszlást írnak le, hiszen ekkor az oszlopokban szereplő elemek összege 1. A következő fejezetben röviden a duplán sztochasztikus mátrixokkal foglalkozom, azaz azon mátrixokkal, ahol nem csak a sorok összege, de az oszlopok összege is 1.

4.4. Duplán sztochasztikus mátrixok

4.4.1. Definíció Egy mátrixot duplán sztochasztikus (bisztochasztikus) mátrixnak nevezünk, ha négyzetes, nemnegatív, és az oszlopösszegek mellett minden sorösszeg is egy, azaz:

$$\sum_{i=1}^{n} a_{ij} = 1, \ \sum_{j=1}^{n} a_{ij} = 1.$$

Legegyszerűbb példa egy bisztochasztikus mátrixra a permutációmátrix, melynek minden sorában és minden oszlopában pontosan egy darab egyes szerepel, a többi érték nulla. Egy ilyen mátrix lehet például egy kártyakeverés mátrixa, ahol n darab kártyát keverünk össze. Legyen $a_{ij} = 1$, ha a keverés után az i-edik kártya a j-edik helyre került, egyébként 0.

4.4.2. Tétel Bisztochasztikus mátrixok szorzata is bisztochasztikus mátrix.

Bizonyítás Legyenek A és B $n \times n$-es bisztochasztikus mátrixok, azaz:

$$\sum_{i=1}^{n} a_{ij} = 1, \ \sum_{j=1}^{n} a_{ij} = 1, \qquad \sum_{i=1}^{n} b_{ij} = 1, \ \sum_{j=1}^{n} b_{ij} = 1.$$

Továbbá legyen $C = A * B$, azaz a mátrixszorzás definíciója szerint:

$$c_{ij} = \sum_{k=1}^{n} a_{ik} * b_{kj}.$$

Vizsgáljuk meg a C mátrix oszlopában lévő elemek összegét:

$$\sum_{i=1}^{n} c_{ij} = \sum_{i=1}^{n} \sum_{k=1}^{n} a_{ik} * b_{kj}.$$

A két szumma felcserélhető, így kapjuk:

$$\sum_{i=1}^{n} c_{ij} = \sum_{k=1}^{n} \sum_{i=1}^{n} a_{ik} * b_{kj}.$$

Mivel $b_{k,j}$ az i-től független, ezért kiemelhető a szumma elé:

$$\sum_{i=1}^{n} c_{ij} = \sum_{k=1}^{n} b_{kj} \sum_{i=1}^{n} a_{ik}.$$

Mivel az A mátrix sztochasztikus, ezért $\sum_{i=1}^{n} a_{ik} = 1$, tehát:

$$\sum_{i=1}^{n} c_{ij} = \sum_{k=1}^{n} b_{kj}.$$

De tudjuk, hogy a B mátrix is sztochasztikus, azaz $\sum_{k=1}^{n} b_{kj} = 1$, tehát azt kapjuk, hogy:

$$\sum_{i=1}^{n} c_{ij} = 1,$$

ami azt jelenti, hogy a C mátrix oszlopában lévő elemeinek összege egy. Ugyanígy belátható a sorokban lévő elemek összegéről, hogy egy, tehát a C mátrix bisztochasztikus mátrix.

4.4.3.Definíció Legyen $\lambda_i \in \mathbb{R}$ minden i-re, ekkor $\sum_i^n \lambda_i a_i$ lineáris kombinációt konvex kombinációnak nevezzük, ha $\sum_i^n \lambda_i = 1$ és $\lambda_i \geq 0$.

4.4.4. Tétel (Birkhoff és Neumann) Egy mátrix akkor és csak akkor bisztochasztikus, ha permutáció mátrixok konvex kombinációja.

A tétel érdekessége, hogy kapcsolatot teremt a bisztochasztikus mátrixok és egy teljes páros gráfban a párosítások között. Legyen a $G(A, B)$ teljes páros gráfhoz tartozó mátrix egy olyan M mátrix, melyre $\forall i \in \mathbb{N},\ i = 1, \dots, |A|,\ \forall j \in \mathbb{N},\ j = 1, \dots, |B|$-re $m_{ij} = 1$, ha az i-edik csúcs össze van kötve a j-edik csúccsal, egyébként 0. Ekkor egy párosításhoz tartozó mátrix minden sorában és minden oszlopában pontosan egy egyes áll, azaz ez egy permutáció mátrix.

5. Köszönetnyilvánítás

Elsősorban szeretném megköszönni témavezetőmnek, Dr. Károlyi Gyulának a segítségét, nélküle nem jöhetett volna létre ez a szakdolgozat. Ő ajánlotta figyelmembe a témákat és több szakirodalmat is ajánlott, mindenben számíthattam a segítségére. Hálás vagyok még Dr. Szalay Mihály Tanár Úrnak az általánosított inverzek témájában elmondott bizonyításáért. Továbbá köszönetet mondok Hackné Nyerges Rita tanárnőnek, aki a középiskolás éveim alatt megszerettette velem a matematikát. Nem utolsó sorban szeretném megköszönni családomnak, barátnőmnek, barátaimnak a támogatást.

6. Irodalomjegyzék

[1] Peter D. Lax: Lineáris algebra és alkalmazásai, Akadémiai Kiadó, 2008

[2] V.V. Praszolov: Lineáris algebra, Typotex Kiadó, 2005

[3] Rózsa Pál: Lineáris algebra és alkalmazásai, Műszaki Könyvkiadó, 1991

[4] Frank András: Kombinatorikus optimalizálás III, 2012
http://www.cs.elte.hu/~frank/jegyzet/polkomb/upoli.2012.2.pdf

[5] Szalay Mihály: Lineáris algebra előadásjegyzet, 2008
http://www.cs.elte.hu/~mszalay/bboard/2008o/index.html