Eötvös Loránd Tudományegyetem
Természettudományi Kar

Nemnegatív mátrixok

SZAKDOLGOZAT

Bakonyi Eszter

Matematika BSc, matematikai elemző szakirány

Témavezető: Ágoston István

egyetemi docens

Algebra és Számelmélet Tanszék

Budapest

2014

Tartalomjegyzék

Bevezető

A következőkben a nemnegatív mátrixok speciális tulajdonságairól, és az ezek alkalmazásairól lesz szó. A nemnegatív mátrixok sok helyen megjelenhetnek, és fontos szerepet játszanak egyes területeken. Ilyen például a valószínűségszámítás. Valaminek a valószínűsége csak 0 és 1 között lehet, így nyilvánvaló, hogy egy valószínűségeket leíró mátrix is nemnegatív. Továbbá a gráfelméletben is hasznosíthatjuk a nemnegatív mátrixok tulajdonságait, ugyanis minden véges gráfnak egyszerűen meg lehet feleltetni egy nemnegatív mátrixot. Ezeknek pedig fontos gyakorlati alkalmazásai lehetnek, amiről majd a későbbiekben szó lesz. A nemnegatív mátrixok előnye az alkalmazásukkor sokszor abban rejlik, hogy a nemnegatív mátrixok spektrumának és sajátvektorainak sajátságos tulajdonságai vannak, továbbá az általános esethez képest sokkal kevesebb számítással megkaphatjuk a számunkra fontos információkat róluk.

A dolgozatot több forrás alapján készítettem, melyeknek a részletes leírása az utolsó oldalon található. A gráfelmélettel foglalkozó részhez Freud Róbert könyvét és Hajnal Péter előadásjegyzetét használtam. A Perron–Frobenius tételkör bizonyításaihoz Lax Péter és Praszolov könyveit használtam. A gyakorlati alkalmazásról szóló fejezethez az irodalomjegyzékben említett amerikai cikket dolgoztam fel. A példákat saját magam dolgoztam ki. A számításokat a MATLAB programmal végeztem el, az ábrákat pedig a GeoGebra programmal készítettem, kivéve a Hoffman–Singleton-gráfot ábrázoló gráfot, amit az internetről töltöttem le.

1. fejezet

Lineáris algebrai alapfogalmak

Általában a lineáris algebra alapfogalmai ismertnek tételezettek fel. De emlékeztetőként néhány később is előforduló fogalom most bevezetésre kerül.

1.0.1. Definíció. Egy valós mátrixot vagy vektort *pozitívnak* nevezünk, ha minden eleme pozitív. Jelölés: $A > 0$ ill. $\vec{v} > 0$ Ugyanígy, egy mátrix vagy vektor *nemnegatív*, ha minden eleme nagyobb vagy egyenlő, mint 0. Jelölése: $A \geq 0$, ill. $\vec{v} \geq 0$.

A továbbiakban „mátrix" alatt főleg nemnegatív négyzetes mátrixot kell érteni, ha nincs utalás ennek ellenkezőjére. A kimondott definíciók és tételek ettől függetlenül általánosabb mátrixokra is érvényesek lehetnek.

1.0.2. Definíció. Egy A mátrixnak $\vec{v} \neq \vec{0}$ vektor (jobb oldali) *sajátvektora*, ha $\exists \lambda \in \mathbb{C}$, melyre $A\vec{v} = \lambda\vec{v}$. Ekkor λ az A mátrix *sajátértéke.*

Egy $n \times n$-es mátrixnak legfeljebb n különböző λ_i sajátértéke lehet.

1.0.3. Definíció. Egy A négyzetes mátrix *spektrumának* nevezzük a sajátértékeinek a listáját. (Egy sajátérték többszörös is lehet.) jelölés: $\sigma(A)$

Egy mátrix spektruma megegyezik a transzponáltjának a spektrumával.

1.0.4. Definíció. Egy A négyzetes mátrix *spektrálsugarának* nevezzük a sajátértékeinek abszolút értékeinek a maximumát. Jelölés: $\rho(A)$. Tehát

$$\rho(A) = \max\{|\lambda_i| \mid \lambda_i \text{ sajátértéke } A\text{-nak}\}.$$

1.0.5. Definíció. Egy $\{\vec{v_1}, \vec{v_2} \dots \vec{v_k}\} \in V$ vektorhalmaz lineáris kombinációjaként előállítható vektorok halmazát a vektorok által kifeszített altérnek nevezzük. Az altér is egy vektortér. Jelölés: $S = \langle \vec{v_1}\vec{v_2} \dots \vec{v_k} \rangle$.

1.0.6. Definíció. Egy A négyzetes mátrix egy λ sajátértékéhez tartozó *sajátaltere* a λ-hoz tartozó sajátvektorok által kifeszített vektortér.

1.0.7. Definíció. Egy A négyzetes mátrix *karakterisztikus polinomja* az a $k_A(t)$-ként jelölt polinom, amely $tI - A$ mátrix determinánsaként áll elő, ahol I az A mátrix méretével megegyező egységmátrix. Tehát $k_A(t) = \det(tI - A)$.

A karakterisztikus polinom gyökei a mátrix sajátértékei. Ez tehát egy lehetséges módszert ad egy mátrix sajátértékeinek kiszámítására.

1.0.8. Definíció. Egy J mátrixot az A mátrix *Jordan-féle normálalakjának* nevezzük, ha A felírható az $A = PJP^{-1}$ alakban, ahol P invertálható mátrix és J főátlója körül ún. *Jordan-blokkok* vannak, azon kívül pedig nullák.

Egy J_k Jordan-blokk egy négyzetes mátrix, amelynek főátlójának minden elemében az A mátrix egy λ_k sajátértéke szerepel, a főátlótól eggyel feljebb lévő elemek pedig 1-esek. A Jordan-blokk többi eleme pedig 0. Tehát

$$J_k = \begin{pmatrix} \lambda_k & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_k & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \lambda_k & 1 & \ddots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & 0 & \ddots & \lambda_k & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & \lambda_k \end{pmatrix}$$

Megjegyzés: A J_k blokk lehet 1×1-es mátrix is, ekkor $J_k = \begin{pmatrix} \lambda_k \end{pmatrix}$. A J mátrix tehát ilyen J_k blokkokból áll össze a főátló mentén, felettük, alattuk nullákkal:

$$J = \begin{pmatrix} [J_1] & & & & \\ & [J_2] & & 0 & \\ & & [J_3] & & \\ & 0 & & \ddots & \\ & & & & [J_l] \end{pmatrix}$$

A J mátrix főátlójának elemei megfelelnek az A mátrix sajátértékeinek.

1.0.9. Definíció. Ha az A mátrix Jordan-normálalakja diagonális, tehát minden J_k blokk 1×1-es, akkor A mátrixot *diagonalizálhatónak* mondjuk.

Állítás:

1. $\forall$ komplex $n \times n$-es mátrixnak $\exists$ Jordan-normálalakja és ez lényegében egyértelmű.

2. Ha A valós szimmetrikus mátrix, akkor a Jordan-alakja valós diagonális mátrix, azaz A diagonalizálható.

3. Minden szimmetrikus mátrix diagonalizálható.

2. fejezet

Nemnegatív mátrixok és gráfok

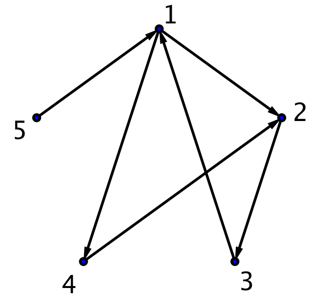
Ebben a fejezetben arról lesz szó, hogy hogyan lehet gráfokat nemnegatív mátrixként reprezentálni. Ezzel pedig egyes gráfokra érvényes tételeket ki lehet mondani nemnegatív mátrixokra, továbbá a gráfokról szóló tételek bizonyításához is lehet a lineáris algebra eszközeit használni. A fejezet végén lesznek erre példák.

2.1. Gráfok mátrix-reprezentációja

2.1.1. Szomszédsági mátrix

2.1.1. Definíció. Egy n csúcsú G gráf *szomszédsági mátrixa* egy olyan $n \times n$-es A_G mátrix, aminek i-edik sorának j-edik eleme megfelel a gráf i-edik csúcsából a j-edik csúcsába induló éleknek.

Ha a gráfban nincs hurokél, akkor a szomszédsági mátrixának a főátlójában csak nullák szerepelhetnek. Ha a gráf irányítatlan, akkor a szomszédsági mátrixa szimmetrikus. Ha nincsenek a gráfban többszörös élek, akkor a szomszédsági mátrix bináris, azaz csak 0-k és 1-esek az elemei.



$$\Leftrightarrow \begin{pmatrix} 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

A szomszédsági mátrix tehát minden esetben egy nemnegatív, négyzetes mátrix. A mátrix négyzetessége azért is előny, mert így lehet hatványozni. A hatványozásának pedig van is értelme.

Jelöljük A_{ij}-vel az A mátrix i. sorának j. elemét. Továbbá $A_{(i)}$-vel az A mátrix i. sorát (mint sorvektor) és $_{(j)}A$-val az A mátrix j-edik oszlopát, mint oszlopvektor. Ekkor

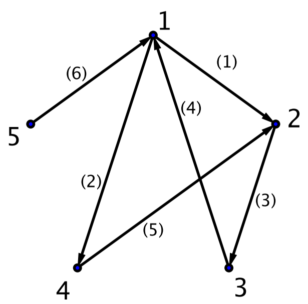
$$(A^2)_{ij} = A_{(i)} \cdot {}_{(j)}A \quad \forall i, j$$

Az A mátrix i-edik sora azt jelzi, hogy a hozzá tartozó gráf i-edik pontjából hová megy él. A j-edik oszlop pedig azt, hogy a j-edik pontba honnan megy él. A két vektor skaláris összeszorzásánál tehát a vektorok k-adik elemének a szorzata 1, ha van olyan k csúcs, amelyre van $i \to j \to k$ út. Ha nincs, akkor 0. Tehát $A^2_{ij} = 0$, ha i és j között nincs (pontosan) 2 hosszúságú út. Ha van, akkor $A^2_{ij} > 0$. (A pontos szám attól függ, hogy hányféleképpen lehet i-ből j-be eljutni.

Ez pedig általánosítható további hatványokra. Tehát A^n_{ij} azt jelzi, hogy az A-nak megfelelő gráfon hányféle különböző (de nem feltétlenül diszjunkt) útvonalon lehet eljutni i csúcsból j-be. Ez későbbi bizonyításokhoz fel lesz használva.

2.1.2. Illeszkedési mátrix

A gráfok reprezentációjának egy másik módja az *illeszkedési mátrix.* Ennek a mátrixnak a sorai feleltethetők meg a gráf csúcsainak, az oszlopai pedig a gráf éleinek. Ha i. csúcsra illeszkedik j. él, akkor $A_{ij} = 1$, egyébként 0. Irányított gráfok esetén meg kell különböztetni a kezdő- és végpontokat. Például ha j. él végpontja az i. csúcson van, akkor $A_{ij} = -1$ ha a kezdőpontja van ott, akkor pedig $A_{ij} = 1$. Ha nincs rajta, akkor természetesen továbbra is 0. Emiatt hurokélekkel rendelkező gráfot nem lehet ezzel a módszerrel reprezentálni, hiszen a mátrix eleme nem lehet egyszerre 1 és -1. Többszörös éleket viszont lehet, egyszerűen ilyenkor ugyanaz az oszlop többször is ismétlődik a mátrixban. Az illeszkedési mátrix természetesen nem feltétlenül négyzetes, csak akkor, ha a csúcsok száma megegyezik az élek számával. Példa:



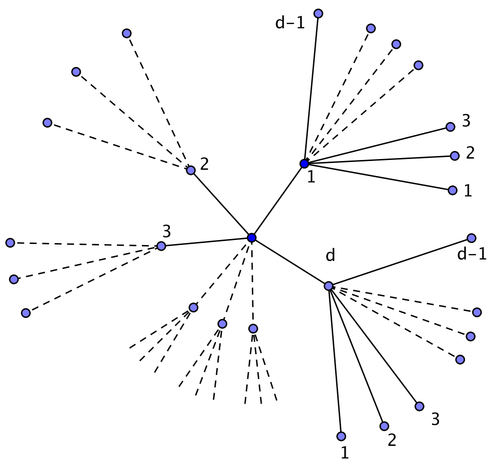
$$\Leftrightarrow \begin{pmatrix} 1 & 1 & 0 & -1 & 0 & -1 \\ -1 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Mivel egy irányított gráf illeszkedési mátrixában vannak -1-ek is, ezért a mátrix nem lesz nemnegatív. A továbbiakban szomszédsági mátrixról lesz szó, és így a „gráf mátrixa” alatt is a szomszédsági mátrixát kell érteni.

2.2. Tételek

Ebben a fejezetben gráfokról szóló tételekről lesz szó, amelyek bizonyításában a szomszédsági mátrixoknak fontos szerepe lesz. Ha gráfokkal dolgozunk, akkor sokszor van valamiféle feltétel, aminek meg kell felelnie a gráfnak. Például hogy legyen összefüggő, legyen reguláris. Páros gráfok esetén nem lehet benne páratlan kör. Síkba rajzolható gráf esetén nem szerepelhet benne teljes ötpontú részgráf vagy teljes $3+3$ pontú páros részgráf, illetve ezek meghosszabbításai. Ezekhez hasonló „tiltott alakzat” pédául az ötnél rövidebb kör is. Vagyis hogy egy gráfban nem lehet háromszög vagy négyszög.

Vegyünk egy olyan összefüggő gráfot, amelynek minden csúcsának d a foka (vagyis reguláris) és nem szerepel benne 5 hosszúságúnál rövidebb kör. Nézzük meg, hogy néz ki egy ilyen gráf. Kezdjük el rajzolni:



A képen látható befejezetlen gráf egyelőre megfelel a feltételeknek. Kérdés, hogy a gráfot be lehet-e fejezni anélkül, hogy újabb pontokat vennénk fel. Az látszik, hogy az ábrán felvett pontok mind szükségesek, hiszen különben nem jöhetne ki a d fok ötnél rövidebb kör nélkül.

Vegyük a középső csúcsot, annak d szomszédját. A d szomszéd csúcsai között nem mehet él, mert akkor háromszög lenne. Továbbá a d szomszédnak nem lehet más közös szomszédja a középső csúcson kívül, különben négyszöget kapnánk. Tehát a d „szomszéd" mindegyikéből megy él még $d-1$ különböző csúcsba, ami összesen $d(d-1)$ csúcs. A középső csúcsot és a szomszédjait is hozzászámolva $d(d-1)+d+1$, vagyis d^2+1. Ennél kevesebb csúcsból tehát biztosan nem tudjuk megoldani. Kérdés, hogy ennyiből meg tudjuk-e. Erről szól a következő tétel.

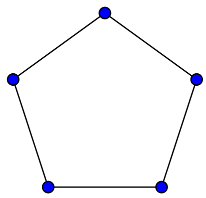
2.2.1. Hoffman–Singleton-tétel

2.2.1. Tétel. *Vegyünk egy G gráfot, aminek minden csúcsából d számú él indul ki, a gráfnak d^2+1 csúcsa van, és nincs benne 5 hosszúságúnál kisebb kör. Ekkor d értéke nem lehet más, csak 1, 2, 3, 7 vagy 57.*

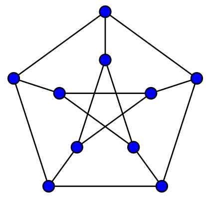
Példák:

– $d = 1$-re

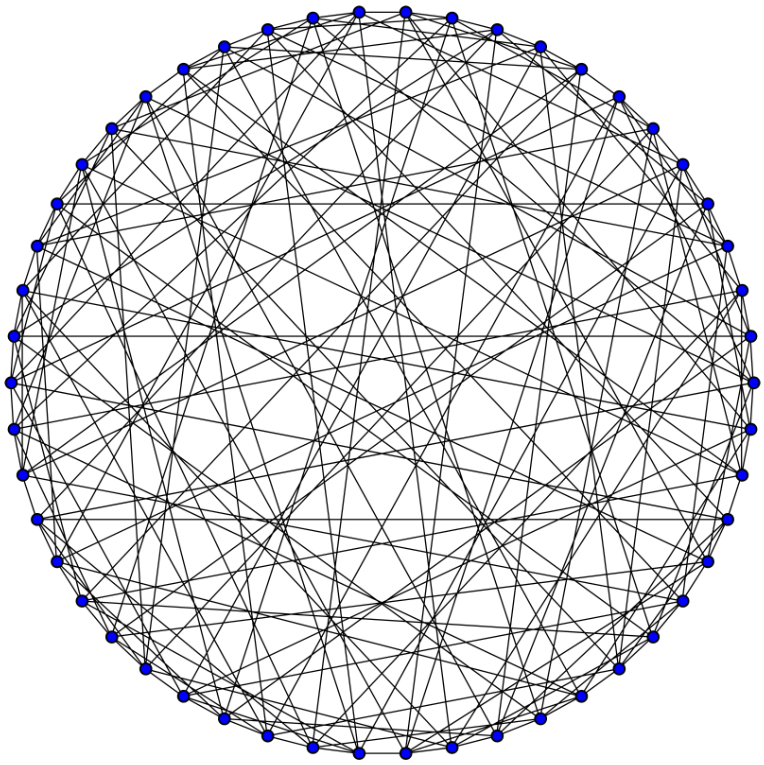
– $d = 2$-re



– $d = 3$-ra: Petersen-gráf



– $d = 7$-re: Hoffman-Singleton-gráf:



A $d = 57$ esetre nem lehet tudni, hogy valójában létezik-e ilyen gráf. A tétel nem bizonyítja, hogy nincs.

Bizonyítás.

Vegyük a G gráf $A \in \mathbb{R}^{n \times n}$ szomszédsági mátrixát.

Legyen $B = A^2$. Tehát $b_{ij} = \sum\limits_{k=1}^{n} a_{ik} \cdot a_{kj}$.

b_{ij} így az i és a j csúcs közös szomszédjainak számát jelöli. $b_{ii} = d$.

Ha $n = d^2 + 1$, akkor két pont vagy szomszédos, vagy pontosan egy közös szomszédjuk van. Ez a gráf csúcsszámának minimalitása miatt van, ahogy az a korábbi ábrán is látszik, a középső pontból legfeljebb 2 távolságra van az összes többi pont. (És bármelyik csúcs kinevezhető „középsőnek.") Tehát $i \neq j$-re $a_{ij} + b_{ij} = 1$.

$$A^2 + A = \begin{pmatrix} d & 1 & 1 & \cdots & 1 \\ 1 & d & 1 & \cdots & 1 \\ 1 & 1 & d & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \cdots & d \end{pmatrix}$$

Tekintsük a csupa 1-esből álló J mátrixot:

$$J = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix}$$

A fentiek alapján $J = A^2 + A - (d - 1 \cdot E)$. ($E$ az egységmátrix.)

A szimmetrikus, ezért $\exists\ n$ db lineárisan független sajátvektora, tehát A diagonalizálható.

Tudjuk, hogy A minden sajátvektora sajátvektora A^2-nek is, hiszen ha $A\vec{v} = \lambda\vec{v}$, akkor

$$A^2\vec{v} = A(A\vec{v}) = A(\lambda\vec{v}) = \lambda A\vec{v} = \lambda^2\vec{v}.$$

Ugyanakkor:

$$J\vec{v} = (A^2 + A - (d-1) \cdot E)\vec{v} = \lambda^2\vec{v} + \lambda\vec{v} - (d-1)\vec{v} = (\lambda^2 + \lambda - (d-1))\vec{v}$$

Tehát ha $\vec{v}$ sajátvektora A-nak, akkor sajátvektora J-nek is. Ha λ sajátértéke A-nak, akkor $\lambda^2 + \lambda - (d-1)$ sajátértéke J-nek. A J sajátértékeit viszont ismerjük:

$$J \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} n \\ n \\ n \\ \vdots \\ n \end{pmatrix} \Rightarrow n \text{ sajátérték.}$$

Mivel J rangja 1, ezért $\dim(\operatorname{Ker} J) = n-1$, tehát 0 $n-1$-szeres sajátérték.

Mivel G d-reguláris, ezért A $\forall$ sorában pontosan d darab 1-es van. Ezért

$$A \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} d \\ d \\ d \\ \vdots \\ d \end{pmatrix} \Rightarrow d \text{ sajátértéke.}$$

$$\lambda^2 + \lambda - (d-1) = d^2 + d - d + 1 = d^2 + 1 = n.$$

Tehát a G d sajátértéke a J n sajátértékével feleltethető meg. Tehát d 1-szeres sajátértéke A-nak. Az A többi λ sajátértékéhez J-ben a $\mu = \lambda^2 + \lambda - (d-1) = 0$ fog tartozni. $\lambda^2 + \lambda - (d-1) = 0$ gyökei pedig

$$\lambda_1 = \frac{-1+\sqrt{4d-3}}{2}, \quad \text{és} \quad \lambda_2 = \frac{-1-\sqrt{4d-3}}{2}.$$

A λ_1 és λ_2 sajátértékek multiplicitása egyelőre nem ismert. Jelöljük m_1-gyel és m_2-vel. Mivel d az egyik sajátérték, ezért

$$\begin{aligned} m_1 + m_2 + 1 = n = d^2 + 1, \quad \text{tehát} \\ m_1 + m_2 = d^2. \end{aligned} \tag{2.1}$$

Az A mátrix nyoma 0.

Legyen D az A mátrix diagonális alakja. Ekkor D főátlójában a sajátértékei szerepelnek. D nyoma tehát a sajátértékek összege, vagyis

$$\mathrm{tr}(D) = m_1 \cdot \lambda_1 + m_2 \cdot \lambda_2 + d.$$

Mivel a *hasonló* (bázistranszformációval egymásba vihető) mátrixoknak a nyoma megegyezik, ezért

$$\begin{aligned} 0 &= \mathrm{tr}(D) = \mathrm{tr}(A) \\ 0 &= m_1 \cdot \lambda_1 + m_2 \cdot \lambda_2 + d \\ 0 &= \frac{m_1 - m_2}{2} \cdot \sqrt{4d-3} - \frac{m_1 + m_2}{2} + d. \end{aligned}$$

$4d-3$ egész szám, tehát a négyzetgyöke vagy egész, vagy irracionális. Mivel m_1 és m_2 is egész szám, ezért $\frac{m_1-m_2}{2}$ racionális. Racionális és irracionális szám szorzata csak akkor lehet racionális, ha a racionális szám 0. $\frac{m_1+m_2}{2} + d$ is racionális. Egy irracionális és egy racionális szám különbsége nem lehet racionális, így nem lehet 0 sem. Tehát két eset lehetséges:

1. $m_1 - m_2 = 0$, tehát $m_1 = m_2$
2. $\sqrt{4d-3} = s \in \mathbb{N}$

1. eset: $m_1 = m_2$, tehát

$$\begin{aligned} \frac{m_1 + m_2}{2} &= d \\ m_1 + m_2 &= 2d = d^2 \end{aligned}$$

Mivel $d = 0$ nem lehet, ezért ebben az esetben $d = 2$

2. eset: $d = \frac{s^2+3}{4}$. Ekkor (2.1)-t is felhasználva:

$$\begin{aligned}
0 &= \frac{m_1 - m_2}{2} \cdot s - \frac{(\frac{s^2+3}{4})^2}{2} + \frac{s^2+3}{4} \quad / \cdot 32 \\
0 &= 16(m_1 - m_2)s - (s^2+3)^2 + 8(s^2+3) \\
0 &= -s^4 + 2s^2 + 16(m_1 - m_2)s + 15 \\
15 &= s \cdot (s^3 - 2s - 16(m_1 - m_2))
\end{aligned}$$

Mivel $s^3 - 2s - 16(m_1 - m_2)$ egész, ezért $s \mid 15$.
Tehát s lehet 1, 3, 5 és 15.

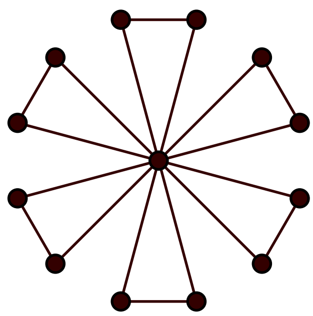
$S = 1 \quad \to \quad d = 1$
$S = 3 \quad \to \quad d = 3$
$S = 5 \quad \to \quad d = 7$
$S = 15 \quad \to \quad d = 57$

$\square$

2.2.2. Barátság-tétel

Most jöjjön egy másik állítás, ami szintén egy speciális gráftípusról szól, és aminek a bizonyításában nagyon hasonló módon felhasználjuk a gráf szomszédsági mátrixának sajátértékeit.

2.2.2. Tétel. *(Barátság-tétel) Legyen $G = (V, E)$ egy egyszerű, irányítatlan, véges gráf, amelyben baármely x és y különböző csúcsnak pontosan egy közös szomszédja van, azaz $\forall x, y \in \mathrm{V}(G), x \neq y \ \exists! z$, melyre $\{x, z\}, \{y, z\} \in \mathrm{E}(G)$. Ekkor G egy ún.* szélmalom*, vagyis olyan gráf, ami egyetlen közös csúcsból kiinduló háromszöggráfok uniója.*



Bizonyítás.

Először az alábbi segédállítást látjuk be:

Állítás: Ha a gráf két különböző pontja, x és y között nincs él, akkor a két él foka egyenlő.

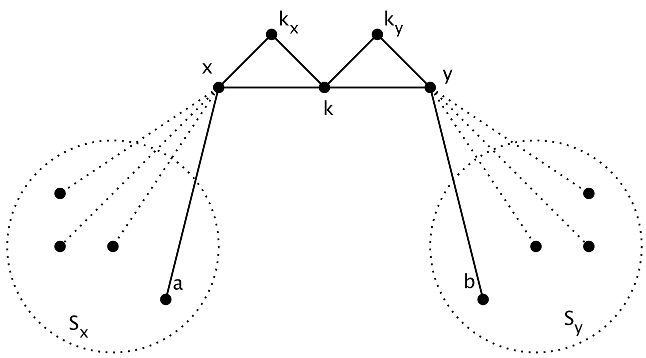
Bizonyítás:

– Legyen k az x és az y csúcs közös szomszédja.

– Legyen k_x az x és a k csúcs közös szomszédja.

– Legyen k_y az y és a k csúcs közös szomszédja.

Ekkor könnyen belátható, hogy az x, y, k, k_x, k_y csúcsok mind különbözőek. $k_x \neq k_y$, mivel máskülönben x-nek és y-nak két közös szomszédja lenne. A többi pont különbözősége a közös szomszédságból következik, mivel G irányítatlan gráf, ezért nem lehet egy pont saját maga szomszédja. És azt is tudjuk, hogy nincs az öt pont között más él a felsoroltakon kívül, mivel x és y között definíció szerint nincs él, x és k_y között nem lehet, mert akkor az is közös szomszédja lenne y-nal. Ugyanígy nem lehet él k_x és y között. k_x és k_y között mert azért nem lehet él, mert akkor két közös szomszédja lenne k_x-nak és y-nak. Ezzel kizártunk négy élt a maximális 10-ből (5 csúcsú teljes gráf éleinek száma). Így maradt 6, vagyis $\{x, k\}, \{y, k\}, \{x, k_x\}, \{y, k_y\}, \{k_x, k\}$ és $\{k_y, k\}$. (Lásd ábra)



Legyen S_x az x csúcs *többi* szomszédos csúcsainak halmaza (tehát k-t és k_x-et kivéve). Ugyanígy, S_y az y csúcs szomszédos csúcsainak halmaza, k-t és k_y-t kivéve.

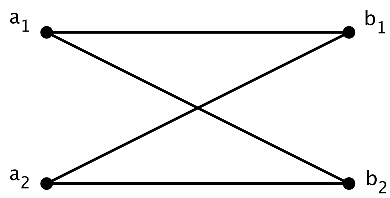
Ekkor minden $a_i \in S_x$ csúcsnak létezik pontosan egy b_i szomszédja S_y-ban. Ez az a_i és az y csúcsok közös szomszédja. Több ilyen csúcs a feltételek miatt nyilván nem lehet. Ez a csúcs muszáj, hogy S_y-ban legyen, hiszen ha $b_i = k$ lenne, akkor x-nek és k-nak két közös szomszédja is lenne, k_x és

a_i. Ha pedig k_y lenne, akkor x-nek és k_y-nak a közös szomszédjai a_i és k lennének. Mivel x és y szerepe felcserélhető, ezért ugyanígyminden $b_i \in S_y$ csúcsnak is pontosan egy a_i szomszédja van S_x-ben. Így egy kölcsönösen egyértelmű megfeleltetést kaphatunk S_x és S_y elemei között. Tehát x és y foka megegyezik.

Vegyük most a G gráf komplementerét, a G' gráfot.

Tegyük fel, hogy G' nem összefüggő, vagyis $\exists G_1, G_2 \subset \mathrm{V}(G') = \mathrm{V}(G)$ diszjunkt nemüres részhalmazok, hogy G_1 és G_2 között nem megy él G'-ben.

Tegyük fel, hogy a G_1 és G_2 halmazokban legalább 2-2 csúcs van. Ha $|G_1|, |G_2| \geq 2 \Rightarrow \exists a_1, a_2 \in G_1, \quad b_1, b_2 \in G_2$, ahol G-ben megy él a_1-ből b_1-be és b_2-be, és ugyanígy a_2-ből is megy b_1-be és b_2-be is.



Ez viszont ellentmond pl. annak, hogy a_1-nek és a_2-nek egyetlen közös szomszédja van.

Tehát az egyik osztályban csak egy csúcs lehet. Legyen ez a $|G_1| = x_0$. Ekkor az x_0 a G gráf minden más csúcsával össze van kötve, és bármely $y \neq x_0$-hoz $\exists! z \neq x_0 : \{y, z\} \in \mathrm{E}(G)$, hogy így z az x_0 és az y (egyetlen) közös szomszédja legyen. Így egy szélmalmot kapunk.

Most tegyük azt fel, hogy G' összefüggő. Ez azt jelenti, hogy G bármely két csúcsa között van G'-beli út, vagyis G-beli „nem-éleken” át vezető út. Ekkor viszont a korábban már bizonyított állítás szerint minden csúcsnak ugyanannyi a foka, azaz G (d-)reguláris gráf.

Legyen A a szomszédsági mátrixa G-nek. Mivel G irányítatlan, ezért A szimmetrikus. A^2-ben pedig a főátló elemei d-vel egyenlők, mivel a gráf d-reguláris, az A^2 többi eleme pedig 1, mert bármely csúcsnak pontosan 1 közös szomszédja van.

Ez azt jelenti, hogy $A^2 = J + (d-1) \cdot I$, azaz $A^2 - (d-1) \cdot I = J$. (A J itt is a csupa 1-es mátrixok jelöli.) Ekkor A sajátvektorai egyben J-nek is sajátvektorai, és ha a v sajátvektorhoz tartozó sajátérték A esetében λ, akkor J-nél $\mu = \lambda^2 - (d-1)$. Tudjuk, hogy a $v = \begin{pmatrix} 1 & 1 & \dots & 1 \end{pmatrix}^T$-hoz a $\lambda = d$ sajátérték tartozik A-nál, és ez azt jelenti, hogy a hozzá tartozó $\mu = d^2 - (d-1) = n$, ahol $n = |G|$. Tudjuk azt is, hogy A diagonalizálható, így n

darab lineárisan független sajátvektora van. A J-nek pedig a 0 sajátértékhez tartozó sajátaltere $n-1$ dimenziós, az n-hez tartozó pedig 1 dimenziós. Így (multiplicitással számolva) $n-1$ darab sajátértékre igaz, hogy $\lambda^2-(d-1)=0$, azaz $\lambda=\pm\sqrt{d-1}$. Legyen $\lambda_1=\sqrt{d-1}$ multiplicitása m_1, $\lambda_2=-\sqrt{d-1}$ multiplicitása pedig m_2. (Ahol természetesen $m_1+m_2+1=n$.)

Mivel A nyoma 0, ezért a diagonális alakjának is 0, tehát

$$0=\operatorname{tr}(A)=m_1\sqrt{d-1}-m_2\sqrt{d-1}+d,$$

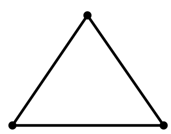
ezért

$$d=(m_2-m_1)\sqrt{d-1}.$$

Mivel az összefüggőség miatt $d\neq 0$, és m_2-m_1 racionális szám, ezért $\sqrt{d-1}$-nek is racionálisnak kell lennie. Tehát $d-1=s^2$, ahol $s\in\mathbb{N}$. Tehát $d=s^2+1$. Vagyis

$$\begin{aligned}
d&=s^2+1=(m_2-m_1)\cdot s\\
0&=s^2+(m_1-m_2)\cdot s+1\\
1&=s\cdot(m_2-m_1-s)
\end{aligned}$$

Ebből az következik, hogy $s\mid 1$. Mivel s pozitív, ezért ez csak úgy lehet, ha $s=1$. Tehát $d=2$ és $n=3$.



Ez a gráf pedig egy háromszöggráf lesz, ami a szélmalomgráfnak egy speciális esete. $\square$

3. fejezet

A Perron–Frobenius-tétel

Az előző fejezetben a tételek bizonyításaiban fontos szerepet játszottak a szomszédsági mátrixok sajátértékei. Kiderültek, hogy a fent említett speciális gráfoknak a szomszédsági mátrixai sajátértékeinek is speciális tulajdonságai vannak. Ebben a fejezetben azt mutatjuk meg, hogy mi mondható el általában a nemnegatív mátrixok sajátértékeiről és sajátvektorairól.

3.1. Perron-tétel pozitív mátrixokra

Először jöjjön a pozitív mátrixok esete. Tehát gráfokra vonatkoztatva, a teljes, irányított, súlyozott élű gráfok szomszédsági mátrixairól lesz szó.

3.1.1. Tétel. *(Perron)*
Minden P pozitív mátrixnak van egy λ_{max} domináns *sajátértéke, amelyre az alábbi állítások igazak:*
– *λ_{max} pozitív, és tartozik hozzá pozitív sajátvektor*
– *λ_{max} egyszeres sajátérték*
– *minden más sajátérték abszolút értékben kisebb, mint λ_{max}*
– *más sajátértékekhez nem tartozik nemnegatív sajátvektor*

Bizonyítás.
Legyen $\mathrm{p}(P)$ azoknak a nemnegatív λ számoknak a halmaza, amelyekre létezik olyan $\vec{x} \neq 0$ vektor, amire

$$P\vec{x} \geq \lambda\vec{x}, \quad \vec{x} \geq 0. \tag{3.1}$$

Állítás:
– $\mathrm{p}(P)$ nem üres, és tartalmaz pozitív számot
– $\mathrm{p}(P)$ korlátos és zárt.

Bizonyítás:

Legyen $\vec{x}$ tetszőleges pozitív vektor. Ekkor $P\vec{x}$ is pozitív. Ez elegendően kicsi λ esetén 3.1 nyilván teljesül. Tehát p(P) tartalmaz pozitív számot.

Mivel 3.1 mindkét oldala lineáris függvénye $\vec{x}$-nek, ezért $\vec{x}$-et lehet egy skalárszorzóval normalizálni úgy, hogy a komponenseinek összege 1 legyen.

Jelölje $\vec{\xi}$ az $\begin{pmatrix}1 & 1 & \dots & 1\end{pmatrix}$ sorvektort. Ekkor

$$\vec{\xi}P\vec{x} \geq \lambda\vec{\xi}\vec{x} = \lambda.$$

Legyen b a $\vec{\xi}P$ legnagyobb komponense. Ekkor

$$b\vec{\xi} \geq \vec{\xi}P.$$

Mivel $\vec{\xi}\vec{x} = 1$, ezért $b = b\vec{\xi}\vec{x} \geq \vec{\xi}P\vec{x} \geq \lambda$. Tehát p($P$) korlátos.

A zártság bizonyításához vegyünk egy λ_n sorozatot a p(P) halmazból. A definíció alapján van olyan $\vec{x}_n \neq 0$ sorozat, amelyre teljesül, hogy

$$P\vec{x}_n \geq \lambda_n\vec{x}_n \quad \text{, ahol } \vec{x}_n \geq 0.$$

Ismét feltehetjük, hogy az $\vec{x}_n$ vektorok normáltak, azaz $||\vec{x}_n||_1 = \vec{\xi}\vec{x}_n = 1$.

Az így normált $\vec{x}_n$ vektorok halmaza az $\mathbb{R}^n$ tér zárt korlátos részhalmazát alkotják, így egy $\vec{x}_n$ részsorozat egy nemnegatív $\vec{x}$ vektorhoz konvergál, ami szintén normált, és λ_n a λ értékhez tart. Mivel a $P\vec{x}_n \geq \lambda_n\vec{x}_n$ a határértékre is igaz, így a λ és az $\vec{x}$ kielégíti a 3.1 egyenlőtlenséget. Tehát p P zárt is.

Mivel p(P) zárt és korlátos, ezért van λ_{max} maximuma. És mivel tartalmaz pozitív számot, ezért $\lambda_{max} > 0$. Most azt fogjuk belátni, hogy λ_{max} lesz a domináns sajátérték. Mivel 3.1 teljesül a λ_{max} értékre, így létezik egy nemnegatív $\vec{h}$ vektor, amelyre igaz, hogy

$$P\vec{h} \geq \lambda_{max}\vec{h}, \qquad \vec{h} \geq 0, \vec{h} \neq 0.$$

Állítás: a fenti egyenlőtlenség valójában egyenlőség. Tegyük fel ugyanis, hogy ez nem így van. Ekkor pl. a k-adik komponensre

$$\begin{aligned} &\sum_j P_{ij}h_j \geq \lambda_{max}h_i, && \text{ha } i \neq k, \text{egyébként} \\ &\sum_j P_{kj}h_j > \lambda_{max}h_k. \end{aligned}$$

Legyen egy $\vec{x} := \vec{h} + \varepsilon\vec{e}_k$ vektor, ahol $\varepsilon > 0$. Az $\vec{e}_k$ egy olyan oszlopvektort jelöl, aminek a k-adik komponense 1, a többi pedig 0. Mivel P pozitív, ezért

$$P \cdot \varepsilon \vec{e}_k = \varepsilon \cdot P\vec{e}_k = \varepsilon P_{(k)} > 0,$$

ahol P_k a P k-adik oszlopa.

Így $P\vec{x} > P\vec{h}$. Ugyanakkor $\lambda_m ax\vec{x}$-nek csak a k-adik komponense nagyobb, mint $\lambda_m ax\vec{h}$-é ($\lambda_{max}\varepsilon - nal$), így elegendően kicsi ε-t választva $P\vec{x} \geq (\lambda_{max} + \delta)\vec{x}$, és mivel $0 \neq \vec{x} \geq 0$, ezért $\lambda_{max} + \delta \in \mathrm{p}(P)$. Így ellentmondásra jutottunk. Tehát λ_{max} sajátérték, és $\vec{h}$ sajátvektora P-nek.

Most azt mutatjuk meg, hogy $\vec{h}$ pozitív. Mivel $\vec{h} \geq 0, \vec{h} \neq 0$ és $P > 0$, ezért $P\vec{h} > 0$. És mivel $P\vec{h} = \lambda_{max} \cdot \vec{h}$, ezért $\vec{h} > 0$.

Ezzel tehát beláttuk a tétel 1. részét. Vagyis $\exists \lambda_{max} = \lambda(P) > 0$, amelynek létezik pozitív sajátvektora.

Most megmutatjuk, hogy λ_{max} geometriai multiplicitása 1, azaz a λ_{max}-hez tartozó $S_{\lambda_{max}}$ sajátaltér dimenziója 1, tehát a sajátalteret kigenerálja $\vec{h}$.

Ehhez tegyük fel, hogy $\exists \vec{y} \in S_{\lambda_{max}}$, ahol $\vec{y} \neq \vec{0}$ és lineárisan független $\vec{h}$-tól. Ekkor ugyanis $\vec{h} \pm c \cdot \vec{y} \neq \vec{0}$ semelyik $c \in \mathbb{R}$ konstansra. Másrészt pedig $\vec{h} \pm c\vec{y} > \vec{0}$, ha c elegendően kicsi. c értékét folytonosan növelve elérhető, hogy $\vec{h} + c\vec{y}$ vagy $\vec{h} - c\vec{y}$ egyik komponense 0-vá váljon, és a többi komponense ekkor nemnegatív. De mivel $\vec{h} + c\vec{y} \in S_{\lambda_{max}}$, azaz sajátvektorokról van szó, ezért, ahogy azt fentebb beláttuk, ha a sajátvektor nemnegatív, akkor pozitív. Ezzel ellentmondásra jutottunk.

Most azt látjuk be, hogy a λ_{max} egyszeres sajátértéke P-nek. Ehhez használjuk a következő észrevételt:

$(t-\lambda)^2 \mid k_A(t)$ valamely A mátrixra, (ahol $k_A(t)$ az A mátrix karakterisztikus polinomja,) és $\vec{v}$ egy λ-hoz tartozó sajátvektor, akkor van olyan $\vec{u} \neq \vec{0}$ vektor, melyre $\vec{u}$ és $\vec{v}$ lineárisan függetlenek, és $A\vec{u} = \lambda\vec{u} + c\vec{v}$ valamely c konstansra. Itt $A \in \mathbb{R}^{n \times n}$ tetszőleges test feletti mátrix, $\lambda \in \mathbb{R}$.

Bizonyítás: Legyen $\mathbb{R}^n = \langle \vec{v} \rangle \oplus U$. Ekkor $A \sim \begin{pmatrix} \lambda & * \\ \vec{0} & B \end{pmatrix}$, és $k_A(t) = (t-\lambda) \cdot k_B(t)$, így λ sajátértéke B-nek.

Ekkor $B \sim \begin{pmatrix} \lambda & * \\ \vec{0} & c \end{pmatrix}$, vagyis

$$A \sim \begin{pmatrix} \lambda & * & * \\ 0 & \lambda & * \\ 0 & 0 & c \end{pmatrix}.$$

Ez éppen azt jelenti, hogy $\exists \vec{v}, \vec{u}, \ldots$ bázis, melyre $A\vec{v} = \lambda\vec{v}$ és $A\vec{u} = \lambda\vec{u} + c\vec{v}$. Ezzel az észrevétel bizonyítását befejeztük.

Most megmutatjuk, hogy λ_{max}-hoz $\nexists$ ilyen $\vec{u}$ vektor.

Tegyük fel, hogy $\vec{u}$ egy ilyen vektor. Ekkor $P\vec{u} = \lambda_{max}\vec{u} + c\vec{h}$, ahol $\vec{h}$ a λ_{max}-hoz tartozó pozitív sajátvektor, és $\vec{u}, \vec{h}$ lineárisan függetlenek.

Tudjuk, hogy $c \neq 0$, mivel $\dim S_{\lambda_{max}} = 1$, így $\vec{u}$ nem sajátvektor. Szükség esetén $-\vec{u}$-t véve feltehető, hogy $c > 0$. Továbbá ha $\vec{u}$ helyett $\vec{u}' = (\vec{u} + b\vec{h})$-t veszünk, akkor $P(\vec{u} + b\vec{h}) = \lambda_{max}(\vec{u} + b\vec{h}) + c\vec{h}$-t is, tetszőleges b-vel. Ezzel viszont elérhető, hogy $\vec{u} > \vec{0}$ teljesüljön. Ekkor $P\vec{u} = \lambda_{max}\vec{u} + c\vec{h} > \lambda_{max}\vec{u}, \vec{u} > \vec{0}$-ra, és így $\exists \delta > 0 : P\vec{u} > (\lambda_{max} + \delta)\vec{u}$. Ebből viszont $\lambda_{max} + \delta \in$ p(P) következne. Így ellentmondásra jutottunk.

Ezzel beláttuk, hogy $(t - \lambda)^2 \nmid k_P(t)$, azaz λ egyszeres sajátérték.

P minden más κ sajátértéke $|\kappa| < \lambda_{max}$ bizonyítátához: legyen $\kappa \neq \lambda_{max}$ másik sajátértéke P-nek, és $\vec{y}$ a hozzá tartozó sajátvektor. (Itt most $\kappa, \vec{y}$ lehetnek komplexek is.) Ekkor $P\vec{y} = \kappa\vec{y}$, azaz

$$\sum_j P_{ij} y_j = \kappa y_j \quad \forall\ i\text{-re.}$$

Alkalmazva a háromszög-egyenlőtlenséget:

$$\sum_j P_{ij}|y_j| \geq |\sum_j P_{ij} y_j| = |\kappa| \cdot |y_i|$$

Ebből következik, hogy $\kappa \in$ p(P), így $|\kappa| \leq \lambda_{max}$. Ha $|\kappa| = \lambda_{max}$, akkor az $\big(|y_1| \dots |y_n|\big)^T$ vektor λ_{max}-hoz tartozó sajátvektor, így $\vec{h}$-nak többszöröse, vagyis $|y_i| = c \cdot h_i\ \forall\ i$-re. Továbbá akkor a fenti egyenlőtlenségben egyenlőség állna. Ez csak akkor lehetséges, ha $Py \cdot y_i = Py \cdot y_j \quad \forall i, j$-re, azaz $y = c\varepsilon h$, ahol $|\varepsilon| = 1, c \in \mathbb{R}$ pozitív. Ekkor viszont $\kappa = \lambda_{max}$.

Végezetül tegyük fel, hogy $\vec{f}$ egy nemnegatív sajátvektora P-nek, mely nem a λ_{max}-hoz tartozik. Legyen $\vec{\xi}$ a P^T-nak λ_{max}-hoz tartozó pozitív sajátvektora. Ilyen kell hogy legyen, hiszen ha $P > 0$, akkor P^T is pozitív mátrix, és P-nek és P^T-nek ugyanazok a sajátértékei. Ekkor

$$(\vec{\xi}, P\vec{f}) = (\vec{\xi}, \mu\vec{f}) = \mu(\vec{\xi}, \vec{f}),$$

másrészt

$$(\vec{\xi}, P\vec{f}) = (P^T\vec{\xi}, \vec{f}) = (\lambda_{max}\vec{\xi}, \vec{f}) = \lambda_{max}(\vec{\xi}, \vec{f}).$$

Mivel $\vec{\xi} > 0$ és $\vec{f} \geq 0$, ezért $(\vec{\xi}, \vec{f}) \neq 0$, és így $\mu = \lambda_{max}$ következne, ami ellentmondás. $\square$

3.1.1. Sztochasztikus mátrixok

3.1.2. Definíció. Egy mátrix *sztochasztikus,* ha nemnegatív, és minden oszlopában az elemek összege 1. Egy mátrix *duplán sztochasztikus,* ha nemnegatív, négyzetes, és minden sorának, és minden oszlopában az elemek összege 1.

A sztochasztikus mátrixokat főleg valószínűségek számítására lehet használni, ahol egy A mátrix i. sorának j. eleme azt mutatja, hogy az j. állapot milyen valószínűséggel fog átmenni i. állapotba egy lépés után. Nyilván ez azt jelenti, hogy diszkrét idővel, és véges sok lehetséges állapottal kell számolni.

Ha egy folyamat *Markov-tulajdonságú,* az azt jelenti, hogy a $n+1$-edik időpontbeli állapot valószínűsége csak az n-edik időpontbeli állapottól függ, az azelőttiektől nem. Ha ez teljesül, akkor ezt a folyamatot le lehet írni egy konstans sztochasztikus mátrixszal.

Példa:

Vegyük egy ország népességét aszerint, hogy melyik mobilszolgáltatónak az ügyfele. Tegyük fel, hogy valaki egyszerre csak egyik szolgáltatónak lehet az ügyfele. Továbbá azt is, hogy az, hogy valaki szolgáltatót vált-e, nem függ attól, hogy korábban is váltott szolgáltatót. Ekkor a folyamat felírható egy $A \geq 0, A \in \mathbb{R}^{(n+1)\times(n+1)}$-es sztochasztikus mátrixszal, ahol A_{ij} azt jelzi, hogy az j. szolgáltató ügyfelei a következő mért időpontban milyen valószínűséggel lesznek a i. szolgáltatónál. Azokat, akik nem tartoznak sehová, az $n+1$-es, virtuális szolgáltatóhoz rendeljük. Legyen tehát a következő mátrix:

$$\begin{pmatrix} 0.90 & 0.04 & 0.04 & 0.05 \\ 0.04 & 0.87 & 0.01 & 0.08 \\ 0.05 & 0.07 & 0.94 & 0.02 \\ 0.01 & 0.02 & 0.01 & 0.85 \end{pmatrix}$$

Ezt a mátrixot jobbról szorozhatjuk egy $\vec{v}_0$ vektorral, ami azt jelzi, hogy jelenleg melyik szolgáltatónak hány ügyfele van. Például a t_0 időpontban az 1. szolgáltatónak 4 millió, a másodiknak 3 millió, a harmadiknak 2 millió ügyfele van. A szolgáltató nélküliek pedig egymillióan vannak. Az egyszerűség kedvéért milliókban számolva tehát fel lehet írni a következő vektort:

$$\vec{v}_0 = \begin{pmatrix} 4 \\ 3 \\ 2 \\ 1 \end{pmatrix}$$

Így ki lehet számolni, hogy a t_1 időpontban kinek hány ügyfele lesz:

$$\vec{v}_1 = \begin{pmatrix} 0.90 & 0.04 & 0.04 & 0.05 \\ 0.04 & 0.87 & 0.01 & 0.08 \\ 0.05 & 0.07 & 0.94 & 0.02 \\ 0.01 & 0.02 & 0.01 & 0.85 \end{pmatrix} \cdot \begin{pmatrix} 4 \\ 3 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 3.85 \\ 2.87 \\ 2.31 \\ 0.97 \end{pmatrix}$$

Tehát a szolgáltatóváltásokból a 3. szolgáltató járt jól. További állapotokat is ki lehet számolni, hiszen ha $\vec{v}_1 = A\vec{v}_0$, akkor $\vec{v}_2 = A\vec{v}_1$, amit behelyettesítve

$\vec{v}_2 = A^2\vec{v}_0$. Ezt tovább általánosítva $\vec{v}_k = A^k\vec{v}_0$. Számoljuk tehát ki a t_{10} időpontbeli állapotot:

$$\vec{v}_{10} = \begin{pmatrix} 0.90 & 0.04 & 0.04 & 0.05 \\ 0.04 & 0.87 & 0.01 & 0.08 \\ 0.05 & 0.07 & 0.94 & 0.02 \\ 0.01 & 0.02 & 0.01 & 0.85 \end{pmatrix}^{10} \cdot \begin{pmatrix} 4 \\ 3 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 3.1588 \\ 2.1324 \\ 3.8935 \\ 0.8153 \end{pmatrix}$$

Tehát a 3-as a legnagyobb szolgáltatóvá lépett elő. Kérdés, hogy bármeddig tud-e növekedni? Van-e egy egyensúlyi állapot, amihez konvergál? Vagy ciklikusan változik? Az biztos, hogy A egy pozitív mátrix, és a korábban belátott Perron-tétel szerint kell lennie egy olyan $\vec{v} > 0$ vektornak, amelyre $A\vec{v} = \lambda\vec{v}$, ahol $\lambda > 0$. Tehát minden szolgáltató ügyfeleinek száma λ-szorosára változik. Mivel a szolgáltatóváltás nem vátoztat az ország populációján, ezért jogosan sejthetjük, hogy ez a $\lambda = 1$. Vagyis hogy van egy egyensúlyi állapot.

3.1.3. Tétel. *Vegyünk egy S pozitív sztochasztikus mátrixot. Ekkor*

1. *S domináns sajátértéke $\lambda(S) = 1$.*
2. *egy tetszőleges $x \geq 0, x \neq 0$ vektorra $\lim\limits_{N\to\infty} S^N x = c \cdot h$, ahol h a domináns sajátvektor, c pedig egy pozitív konstans.*

Bizonyítás.

1. S^T-hez az $(1, 1, \ldots 1)^T$ sajátvektor, és a hozzá tartozó sajátérték 1. S-nek a Perron-tétel miatt más nemnegatív sajátvektora nem létezhet, ezért 1 a domináns sajátérték.
2. Legyen $S \sim J$, ahol J az S Jordan-normálalakja, azaz

$$J = \begin{pmatrix} 1 & 0\cdots 0 \\ 0 & \\ \vdots & \boxed{\text{J'}} \\ 0 & \end{pmatrix}$$

és J' sajátértékei abszolút értékben 1-nél kisebbek. Ekkor a Jordan-blokkok hatványozásából

$$J^N = \begin{pmatrix} 1 & 0 \\ 0 & (J')^N \end{pmatrix} \to \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},$$

és

$$J^N \cdot x \to \begin{pmatrix} x_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Legyen T az a mátrix, melyre $T^{-1}ST = J$. Feltehető, hogy T első oszlopa éppen a $\vec{h}$ vektor. Ekkor

$$S = TJT^{-1}, \text{ és } S^N\vec{x} = TJ^NT^{-1}\vec{x}.$$

Legyen

$$T^{-1}\vec{x} = \begin{pmatrix} c \\ * \\ \vdots \\ * \end{pmatrix}.$$

Ekkor az előzőek szerint

$$J^NT^{-1}\vec{x} \to \begin{pmatrix} c \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

és így

$$TJ^NT^{-1}\vec{x} \to T \begin{pmatrix} c \\ 0 \\ \vdots \\ 0 \end{pmatrix} = c\vec{h}.$$

□

Most térjünk vissza a tétel előtti példához. A tétel alapján ha az A pozitív sztochasztikus mátrixot egyre nagyobb kitevővel hatványozzuk, akkor $A^n\vec{v}_0, \vec{v}_0 \geq 0, \vec{v}_0 \neq 0$ tartani fog a mátrix 1-es sajátértékéhez tartozó pozitív sajátvektorhoz. Ami a példánk esetében az egyensúlyi állapotot jelenti. A példafeladatban a kezdővektor $\vec{v}_0 = \begin{pmatrix} 4 & 3 & 2 & 1 \end{pmatrix}^T$ volt.

Sokadik fokú hatványozást elvégezve, négy tizedesjegy pontossággal a következő eredményt kapjuk: $\begin{pmatrix} 2.9094 & 1.7035 & 4.6557 & 0.7315 \end{pmatrix}^T$. Tehát ez lesz az egyensúlyi állapot, amihez a folyamat tartani fog, feltéve, hogy ha a kezdőállapotban az ügyfelek száma 10 (millió). Ha nem, akkor ennek a sajátvektornak a pozitív konstansszorosához fog tartani. Feltéve persze, hogy a szolgáltatók és a vásárlói szokások évszázadokig nem változnak...

3.2. Irreducibilis és primitív mátrixok

Kérdés, hogy a Perron-tételt lehet-e általánosítani nemnegatív mátrixokra. Ez ebben az esetben azt jelentené, hogy minden nemnegatív mátrixnak van egy domináns pozitív sajátértéke, ami egyszeres, és a hozzá tartozó sajátvektor nemnegatív, és más sajátértékhez nem tartozik nemnegatív sajátvektor.

Ez viszont biztosan nem igaz minden nemnegatív mátrixra, hiszen nem igaz például az egységmátrixra sem. Az $n \times n$-es egységmátrixnak ugyanis az 1 n-szeres sajátértéke, ami így véletlenül sem nevezhető dominánsnak. Továbbá az 1 sajátértékhez minden 0-tól különböző n-dimenziós vektor sajátvektor, hiszen ha egy vektort az egységmátrixszal szorzunk, azt éppúgy helyben hagyja, mint ha 1-gyel szorozzuk. Ennek ellenére az előző fejezetben szereplő Hoffman–Singleton-tételben és a Barátság-tételben szereplő

Tehát az biztos, hogy nem *minden* nemnegatív mátrixra lehet általánosítani Perron tételét.

Az előző fejezetben szó volt a Hoffmann-Singleton tételről és a Barátság-tételről. Ezek a tételek nagyon speciális gráfokra voltak csak érvényesek. A Hoffmann-Singleton esetében reguláris, irányítatlan gráf volt, aminek a szomszádsági mátrixa emiatt hasonló a duplán sztochasztikus mátrixokhoz. A barátság-tételnél pedig olyan irányítatlan gráfról volt szó, amiben két nemszomszédos csúcsnak pontosan egy közös szomszédja van. Egy ilyen gráf szomszédsági mátrixának a négyzete emiatt pozitív. Amikor a pozitív mátrixokra alkalmazott Perron-tételt néztük, az is egy speciális gráfra érvényes: mégpedig a teljes irányítatlan gráfra, amelynek az élei súlyozottak. Hiszen minden i és j csúcs között megy egy valamilyen pozitív súlyú él, ami nem feltétlenül egyenlő a j és i között menővel.

Most tehát egy általánosabb esetet vizsgálunk, amihez nem kell, hogy mindenhonnan mindenhová vezessen él. Az a korábban említett tételekre mind igaz volt, hogy összefüggő gráfokról szóltak. Most az összefüggőségnek kellene bevezetni egy definícióját szomszédsági mátrixokra.

3.2.1. Definíció. Egy $A \geq 0$ négyzetes mátrix *reducibilis*, ha a sorai és oszlopai permutálásával ún. blokk-(felső)háromszögmátrix kapható. Azaz

$$P^T A P = \left(\begin{array}{c|c} A & B \\ \hline 0 & C \end{array} \right)$$

ahol P permutációmátrix, A és C pedig négyzetes mátrix.

3.2.1. Mátrix reprezentációja gráffal

A szomszédsági mátrixok esetében, alapvető feltétel, hogy a mátrix négyzetes legyen. Ha egy A négyzetes mátrixban csak 0-k és 1-esek szerepelnek, és a főátlójában csak 0-k vannak, akkor az megfeleltethető egy irányított gráfnak, ahol ha $A_{ij} = 1$, akkor $(i, j) \in E(G)$. Tehát i csúcsból megy él j csúcsba. Az $i = j$ is megenfedett, ebben az esetben hurokélről beszélhetünk.

Ha egy mátrix elemei nemnegatív egész számok, akkor ha $A_{ij} = k$, akkor megfeleltethető egy (nem feltétlenül egyszerű) mátrixnak, melynek i csúcsából j csúcsába k darab él megy.

Ha egy mátrix elemei nemnegatív valós számok, akkor egy súlyozott irányított gráffal lehet reprezentálni, ahol ha $A_{ij} = c_{ij} > 0$, akkor a gráf i csúcsából a j csúcsába van egy c_{ij} súlyú él. Ha $A_{ij} = 0$, akkor i csúcsból j csúcs irányába nincs él. Ebben az esetben nem engedünk meg többszörös éleket.

Amennyiben a mátrix szimmetrikus, akkor az irányítatlan mátrixnak is megfeleltethető, függetlenül attól, hogy bináris, egész, vagy valós mátrixról van szó.

Szomszédsági mátrixok esetében a fent említett P^TAP transzformáció a gráf csúcsainak átszámozásának felel meg.

Ha egy négyzetes mátrix nem reducibilis, akkor *irreducibilisnek* nevezzük.

A reducibilitás gráfokra vonatkoztatva azt jelenti, hogy a gráf csúcsainak van egy halmaza, amelyik halmazból kifelé nem megy él. Egy gráf szomszédsági mátrixának irreducibilitása pedig azt jelenti, hogy a gráf minden pontjából minden pontjába vezet út. Irányított gráf esetén irányított út. Ami azt jelenti, hogy az irreducibilis nemnegatív mátrix egy *erősen összefüggő* irányított gráfnak feleltethető meg.

3.2.2. Definíció. Egy $A \geq 0$ négyzetes mátrix *primitív*, ha $\exists$ olyan $k \in \mathbb{N}_+$, amelyre $A^k > 0$.

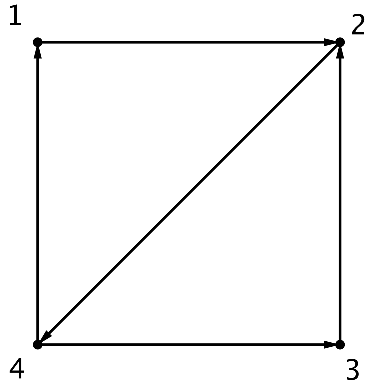
Nyilván minden pozitív mátrix primitív.

3.2.3. Tétel. *Minden primitív mátrix irreducibilis.*

Ahogy korábban szó volt róla, a szomszédsági mátrixok hatványozásával ki lehet számolni, hogy i és j pontok között hány lehetséges út van. És arról is, hogy minden nemnegatív mátrix értelmezhető egy gráf szomszédsági mátrixaként. (És definíció szerint minden primitív mátrix nemnegatív.)

Tehát a definícióban szereplő $A^k > 0$ pontosan azt jelenti, hogy az A mátrix által megadott gráfban minden pontból minden pontba el lehet jutni pontosan k hosszú irányított úton. Az irreducibilitáshoz pedig elég az is, hogy el lehet jutni, hiszen akkor az A gráf erősen összefüggő.

Ugyanakkor nem minden irreducibilis mátrix primitív. Ellenpélda a *periodikus* gráfok mátrixai. Egy gráfot periodikusnak nevezünk, ha az irányított köreinek a hosszai nem relatív prímek.



erősen összefüggő, periodikus gráf

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix}; A^2 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 2 & 0 & 0 \end{pmatrix}; A^3 = \left(\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 2 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 2 \end{array}\right); A^4 = \begin{pmatrix} 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 2 \\ 0 & 2 & 0 & 0 \\ 2 & 0 & 2 & 0 \end{pmatrix}$$

A fenti gráf mátrixának 3. hatványa már láthatóan reducibilis. Ha a gráf k-adik hatványa pozitív, akkor a $3k$-adik hatványának is pozitívnak kell lennie. Ez viszont nem lehet, hiszen egy irreducibilis mátrix hatványa nem lehet pozitív. Továbbá az is látszik, hogy a 4. hatvány a 2-szerese az eredeti mátrixnak, és ezért a további hatványok skalárszorzó erejéig ismétlődnek hármas ciklusokban. Ami azt jelenti, hogy a nullák megmaradnak, így nem lesz pozitív.

Az viszont már igaz, hogy ha egy irreducibilis nemnegatív mátrix nem periodikus (vagyis *aperiodikus*), akkor primitív is. Hiszen elég nagy hatvány esetén (azaz elegendően hosszú irányított utak esetén) egy aperiodikus erősen összefüggő gráfnak bármelyik két, relatív prím hosszú körébe el lehet jutni, és azon tetszőleges sokszor körbe lehet menni, és így egy elég nagy N szám felett már minden hosszúságú sétát ki lehet generálni.

3.2.4. Tétel. *Ha $A \geq 0$ irreducibilis, akkor $(A + I)^N > 0$.*

Ha a mátrixhoz hozzáadunk egy egységmátrixot, az a gráf esetén annak felel meg, hogy a a gráf minden csúcsához hozzárendeltünk egy hurokélet. A hurokél meg tulajdonképpen egy 1 hosszú kör. Mivel a mátrix irreducibilis, ezért a gráfja erősen összefüggő. Ez azt jelenti, hogy legfeljebb $n - 1$ hosszú úton el lehet jutni bármelyik csúcsból bármelyikbe. Az állítás viszont itt azt mondja ki, hogy *pontosan* $n - 1$ hosszú úton is el lehet jutni. Amihez csak az kell, hogy ha $n - k - 1$ hosszú úton is el lehet jutni i-ből j-be (k pozitív egész), akkor az i-ben lévő hurokélen végig lehet menni k-szor, majd elindulni az $n - k - 1$ hosszú úton.

3.2.2. Wielandt-tétel

Arra, hogy egy primitív mátrix hányadik hatványától lesz biztosan pozitív, Wielandt tétele ad egy felső becslést. Az alábbi tételt a teljesség kedvéért idézem Praszolov [5] könyvéből.

3.2.5. Tétel. *(Wielandt) Ha A n-edrendű nemnegatív primitív mátrix, akkor $A^{n^2-2n+2} > 0$.*

Bizonyítás. Minden n-ed rendű nemnegatív mátrix egyértelműen meghatároz egy n csúcspontú irányított gráfot, amelyben az i csúcsból pontosan akkor vezet él a j csúcsba, ha $A_{ij} > 0$ (nem zárjuk ki az $i = j$ esetet sem, ami a hurokélet jelentené). Az A^s mátrix B_{ij} eleme pontosan akkor pozitív, ha a szóban forgó gráfban létezik egy pontosan s hosszúságú út i csúcspontból a j-be. Így, $B_{ij} = \sum A_{ii_1}A_{i_1i_2}\dots A_{i_{s-1}j}$, ahol $A_{ii_1}A_{i_1i_2}\dots A_{i_{s-1}j} > 0$ pontosan akkor, ha $ii_1i_2\dots i_{s-1}j$ a gráf (irányított) éleinek sorozata. Így egy primitív mátrixnak olyan összefüggő gráfot lehet megfeleltetni meg, amelynek bármelyik csúcsából bármely másik csúcs elérhető egy irányított úton. Vegyünk egyet a gráf legrövidebb körútjai közül (ha $A_{ii} > 0$, az ii él ilyen). Tegyük fel, hogy a szóban forgó körút: $12\dots l1$. Az A^l mátrix $B_{11}, \dots B_{ll}$ elemei ezek szerint mind pozitívak.

Tetszőleges i csúcsból legfeljebb $n - l$ hosszú úton elérhető az $1, \dots l$ csúcsok valamelyike. Feltehetjük, hogy ez az út pontosan $n - l$ hosszú, mert ha rövidebb, akkor csak a körúton kell tenni még néhány lépést.

Vegyük most az A^l mátrixot, amely szintén primitív, és szintén meg lehet feleltetni neki egy irányított gráfot. Ebben a gráfban a $j \in \{1, \dots, l\}$ csúcsok bármelyikéből (amelyek mind különböznek az i csúcstól) el lehet érni a gráf bármelyik k csúcsát, méghozzá olyan útvonalon, amelynek hossza nem nagyobb, mint $n - 1$. Mivel a j csúcsból önmagába is vezet él, az út kiegészíthető úgy, hogy $n - 1$ hosszúságú legyen. Így az A-nak megfeleltetett gráf i és k csúcsai között létezik $n - l + l(n - 1) = l(n - 2) + n$ hosszú út. Ha $l = n$, akkor A-t felírhatjuk az alábbi alakban:

$$\begin{pmatrix} 0 & A_{12} & 0 & \cdots & 0 \\ 0 & 0 & A_{23} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & A_{n-1,n} \\ A_{n1} & 0 & 0 & \cdots & 0 \end{pmatrix}$$

Ez a mátrix nem primitív. Emiatt $l \le n - 1$, ebből pedig $l(n - 2) + n \le n^2 - 2n + 2$. Azt kell már csak belátni, hogy ha $A \ge 0$ és $A^p > 0$, akkor $A^{p+1} > 0$. Ez pedig azért igaz, mert ha $A^p > 0$, akkor A definíció szerint primitív, és ez azt jelenti, hogy irreducibilis is. Egy nemnegatív és egy pozitív mátrix szorzata csak akkor lehet nem-pozitív, ha a nemnegatív mátrixnak van csupa nulla sora. Ekkor viszont A reducibilis lenne, így ellentmondásra jutottunk. $\square$

A tételben szereplő becslés éles, mivel létezik olyan mátrix, amire eggyel kisebb hatvány már nem lenne pozitív. Vegyük ugyanis a következő $n \times n$-es mátrixot ($n \geq 3$):

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 1 & 1 & 0 & \cdots & 0 \end{pmatrix}.$$

Ez a mátrix azt a lineáris leképezést határozza meg, ahol

$$A\vec{e}_1 = A\vec{e}_n, A\vec{e}_2 = \vec{e}_1 + \vec{e}_n, A\vec{e}_3 = \vec{e}_2, \ldots, A\vec{e}_n = \vec{e}_n - 1.$$

Emlékeztetőül: $\vec{e}_i$ azt a vektort jelöli, amelynek az i-edik koordinátája 1, a többi pedig nulla. Legyen $B = A^{n-1}$. Ekkor

$$B\vec{e}_1 = \vec{e}_2, B\vec{e}_2 = \vec{e}_2 + \vec{e}_3, B\vec{e}_3 = \vec{e}_3 + \vec{e}_4, \ldots, B\vec{e}_n = \vec{e}_n + \vec{e}_1.$$

A B^{n-1} mátrixnak tehát már csak egy eleme nulla, mégpedig az $(1, 1)$ indexű helyen. Az $AB^{n-1} = A^{n^2-2n+2}$ mátrix viszont már pozitív.

Példa $n = 4$-re:

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{pmatrix}; A^3 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}; A^9 = \begin{pmatrix} 0 & 1 & 2 & 1 \\ 1 & 1 & 1 & 2 \\ 2 & 3 & 1 & 1 \\ 1 & 3 & 3 & 1 \end{pmatrix}; A^{10} = \begin{pmatrix} 1 & 1 & 1 & 2 \\ 2 & 3 & 1 & 1 \\ 1 & 3 & 3 & 1 \\ 1 & 2 & 3 & 3 \end{pmatrix}.$$

3.3. Frobenius tétele

Az előző fejezetben ellenpéldaként felhozott egységmátrix nyilvánvalóan nem irreducibilis (vagyis reducibilis), és ezért nem lehet primitív sem. A kérdés most az: lehet-e irreducibilis, ill. primitív mátrixokra általánosítani Perron tételét?

3.3.1. Tétel. *Ha P mátrix primitív,akkor ugyanazok elmondhatók róla, mint a Perron-tételben a pozitív mátrixokról.*

Bizonyítás. Ha felírjuk P Jordan-alakját, akkor láthatjuk, hogy P sajátértékeinek k-adik hatványai kiadják P^k sajátértékeit, és a P-hez tartozó sajátvektorok sajátvektorai P^k-nak is. Mivel P^k-nak a Perron-tétel miatt van egy egyszeres domináns sajátértéke, akkor ez a P valamely sajátértékének k-adik hatványa, ez a sajátérték egyszeres, és ez a sajátérték abszolút értékben nagyobb minden más sajátértéknél. És az is igaz, hogy az ehhez tartozó sajátvektor választható pozitívnak, mert ez a P^k domináns sajátértékéhez tartozó egydimenziós alteret generálja. Végezetül ebből az is következik, hogy ez a domináns sajátérték pozitív, mert egy nemnegatív mátrix pozitív vektorához tartozik. □

3.3.2. Tétel. *(Frobenius tétele) Legyen $A \geq 0$ irreducibilis mátrix. Ekkor*

1. *A mátrix spektrálsugara $\rho(A) > 0$, és $\rho(A)$ sajátérték, mégpedig olyan, ami egyszeres gyöke k_A-nak.*

2. *$\exists v > 0$, hogy v a $\rho(A)$-hoz tartozó sajátvektor.*

3. *Az A minden sajátértékére nyilván $\lambda \leq \rho(A)$. Ha viszont $|\lambda| = \rho(A)$, és ilyen λ-ból k különböző van, akkor*

$$\lambda_j = \rho(A) \cdot (cos(j \cdot \frac{2\pi}{k}) + i \cdot sin(j \cdot \frac{2 \cdot \pi}{k})) \quad \forall j = 0, \ldots k - 1.$$

vagyis $\rho(A)$ k-adik komplex gyökei.

4. fejezet

Lehetséges alkalmazás

4.1. Versenyzők rangsorolása

A Perron–Frobenius tételkör egy lehetséges alkalmazása a rangsorok készítése. Pontosabban a páronkénti összehasonlításokból kialakított rangsoroké. Tehát például olyan játékoknál lehet használni, ahol két játékos vagy két csapat játszik egymás ellen (pl. sakk, labdajátékok, küzdősportok, egyes számítógépes játékok). De lehet a párosítás virtuális is, tehát a már meglévő adatokból generált „meccsek." Például ha egy választáson hárman indulnak, akkor az már nem párosítás. De azzá tehető. Legyen egy falu, ahol három jelölt indul egy választáson, jelöljük A, B, C-vel. Ez egy olyan választás, ahol preferenciával lehet szavazni. Pl. $B > C > A$ jelentené azt, hogy a szavazó a B jelöltet szeretné leginkább (vagy azt tartja a legkisebb rossznak), de ha B nem nyerhet, akkor inkább a C legyen, de semmiképpen sem az A. Tehát egy ilyen választáson a következő eredmény született. 30-an szavaztak úgy, hogy $A > B > C$, 10-en úgy, hogy $A > C > B$, 35-en úgy, hogy $B > A > C$, 5-en $B > C > A$, és 50-en úgy, hogy $C > A > B$. ($C > B > A$-val senki nem szavazott.) Ez azt jelenti, hogy $30 + 10 + 50 = 90$-en preferálták A-t a B-vel szemben, míg $35+5 = 40$-en a B-t az A-val szemben. Tehát a virtuális játék eredménye $A - B : 90 - 40$. Ugyanígy $A - C : 75 - 55$ és $B - C : 65 - 65$. Innen már nyilvánvalóan látszik, hogy A-nak kell nyernie, hiszen A-t jobban akarja a nép b-nél és C-nél is. Annak ellenére, hogy C kapta a legtöbb elsődleges szavazatot.

De ez az előbbi eset azért volt olyan egyszerű, mert kihasználtunk egy olyan speciális tulajdonságot, ami általában nem szokott meglenni. Mégpedig azt, hogy mindenkinek a szavazólapján ott volt minden jelölt. Vagyis minden szükséges adat megvolt, így teljes volt a kép az erősorrendekről. Az ilyen versenyeket körmérkőzésnek, vagy round-robinnak nevezik. A továbbiakban a cél az lesz, hogy minél általánosabb esetben is ki lehessen számolni a teljes rangsort. Tehát:

- Két versenyző nem feltétlenül játszik egymás ellen
- Két versenyző akármilyen sokszor játszhat egymás ellen
- A versenyzők nem feltétlenül ugyanannyi mérkőzést játszanak összesen
- A versenyzők nem feltétlenül véletlenszerűen választják az ellenfelüket
- A fenti lehetőségekkel a versenyzők akár szándékosan vissza is élhetnek, azért, hogy ezzel a verseny eredményét a saját javukra torzítsák

A feladat tehát egy olyan rangsorolási módszer kidolgozása, ami a legáltalánosabb körülmények között is garantál egy „igazságos" eredményt.

Az persze nem egyértelmű, hogy mi számít „igazságosnak," de itt valamiféle objektív, előre meghatározott feltétel teljesülését jelentené. Ilyenek lehetnek például:

1. Csak az lehessen a bajnok, aki veretlen.
2. Ha A legyőzte B-t, akkor A-nak meg kell előznie a ranglistán B-t.
3. Egy versenyző ranglistabeli eredménye ne függjön olyan mérkőzések eredményétől, amelyen ő nem játszott.
4. Ha valaki megnyer egy meccset, azzal ne kerülhessen rosszabb helyzetbe, mintha elveszítette volna, döntetlen lett volna, vagy le se játszották volna. Ugyanígy, ha valaki elveszít egy meccset, akkor ne kerüljön előnyösebb helyzetbe, mintha megnyerte volna, döntetlen lett volna, vagy le se játszották volna.
5. Nagyobb arányú győzelem érjen többet, mint a kisebb arányú
6. „Erősebb" ellenfél elleni győzelem érjen többet, mint egy „gyengébb" ellenfél elleni. Ugyanígy egy „gyengébb" ellenfél elleni vereség járjon nagyobb büntetéssel

Az nyilvánvaló, hogy az első feltétel csak akkor teljesülhet biztosan, ha a verseny egyenes kieséses. Vagyis ha egy játékos/csapat elveszít egy meccset, akkor nem játszhat többet olyanok ellen, akik még veretlenek. Máskülönben előfordulhat, hogy a verseny végére senki sem marad veretlen, így a ranglista első helye üresen maradna. Hasonló mondható el a 2. és a 3. feltételről is.

Ami a következőkben leginkább érdekes lesz, az az utolsó feltétel: vagyis, hogy a győzelem értéke függjön a legyőzött erősségétől. Itt ugyanis van egy látszólagos ellentmondás. Ha az erős ellenfelek elleni győzelem többet ér, akkor az azt jelenti, hogy már van egy ranglistánk, ami alapján eldöntjük, hogy melyik ellenfél milyen erős. De hát pont egy ilyen ranglista elkészítése a cél... Ennek a problémának a megoldására kell a későbbiekben felhasználni a sajátértékekkel kapcsolatos tételeket.

Jelezze $\vec{r} \geq 0$ vektor a versenyzők „rangját," ahol $\vec{r_i}$ jelzi az i. versenyző erősségét. Arról, hogy ezt hogyan kell kiszámolni, később esik majd szó.

Legyen $A \geq 0$ vektor az „eredmény-mátrix", ahol $A_{i,j}$ írja le azt, hogy az i. játékos milyen eredményeket ért el a j. játékos ellen. Ehhez természetesen először a „milyen eredmény" fogalmát számszerűsíteni kell egy függvénnyel. Nyilván egy ilyen függvénynek monotonnak kell lennie, vagyis jobb eredményért ne járhasson rosszabb pontszám. Például lehet: győzelem = 2 pont, döntetlen = 1 pont, vereség = 0 pont. De ahogy fentebb szó volt róla, bele lehet számítani a győzelem arányát is, vagyis több pont járna a nagyobb arányú győzelemért, és kevesebb pont a nagyobb arányú vereségért. De arra figyelni kell, hogy az A mátrixnak nemnegatívnak kell maradnia! Egy lehetséges pontozófüggvény például a $\mathrm{p}(x) = 2^x$, ahol az x a győzelem mértékét jelzi. (Értelemszerűen ez vereség esetén negatív, döntetlen esetén pedig nulla.) De az ilyen függvénnyel az a baj, hogy x növelésével tart a végtelenbe, vagyis egyetlen mérkőzés eredménye akármekkora mértékben eltorzíthatja a ranglista eredményét. Tehát célszerű a pontozófüggvénynek korlátosnak is lennie. Például lehet úgy, hogy

$$\mathrm{p}(x) = \begin{cases} 2^x & \text{ha } x \leq 0 \\ 2 - 2^{-x} & \text{ha } x > 0. \end{cases}$$

Ez azt jelenti, hogy például egy 2 különbségű győzelem $1,75$ pontot ér, egy 3 különbségű vereség $0,125$ pontot, egy döntetlen pedig 1 pontot.

Megjegyzendő, hogy monoton növekvő, pozitív, korlátos függvényekből még mindig végtelen sok van, és a p függvény megválasztása hatással lehet a ranglista sorrendjére. Ezért tehát nem feltétlenül létezik „a" helyes eredmény, csak egy lehetséges olyan eredmény, ami egyes igazságossági kritériumoknak megfelel.

Ennek a pontozófüggvénynek az eredménye kerülhet tehát bele az A mátrixba. Ha i és j többször is játszott egymás ellen, akkor a pontok összeadódnak. De ha összeadódnak, akkor az azt jelenti, hogy több lejátszott meccsel több pontot lehet szerezni. Ezt korrigálandó, le kell osztani a szerzett pontszámot a lejátszott meccsek számával. Ez viszont azt jelentené, hogy ha i és j nem játszott egymással, akkor $A_{i,j}$-nél nullával osztás történne. Ennek az elkerülése érdekében nem a mátrix elemei, hanem a sorai kerülnének leosztásra. Ez azt jelenti, hogy csak akkor lehet nullával osztás, ha valaki senkivel nem játszott. De nyilván aki senkivel nem játszott, arról nem tudunk mit mondani, ezért az ilyet ki lehet venni a ranglistáról. Tehát akkor képletekben leírva, az i. játékos végső pontszáma

$$\vec{s}_i = \frac{1}{\vec{n}_i} A_i \vec{r},$$

ahol n_i az i. játékos összes lejátszott meccseinek száma, A_i pedig az A mátrix i. sorát jelöli. $A_i\vec{r}$ pedig skalárszorzásként értendő.

Ebben az egyenletben A és $\vec{n}$ ismert, viszont $\vec{r}$ és $\vec{s}$ ismeretlen. $\vec{s}$ ugye itt az eredmény-ranglista pontjait írja le, az $\vec{r}$ pedig az eredmény kiszámításához felhasznált pontszámokat. Egyszerűbb esetben $\vec{r}$ lehet az előző idény végeredménye. Ekkor már könnyen ki lehetne számolni $\vec{s}$-t, feltéve, hogy az összes mostani versenyző részt vett az előző idényben is. De most jöjjön az az eset, amikor $\vec{r}$-et is az A mátrixból számoljuk ki.

Legegyszerűbb az lenne, ha feltennénk, hogy $\vec{v} = \vec{r}$. Ez azt jelentené, hogy

$$A^*\vec{r} = \vec{r}$$

ahol A^* az A mátrix egy olyan transzformált alakja, ahol minden $A_{i,j}$ eleme le van osztva n_i-vel. Továbbiakban ez lesz az A mátrix. Ennek az egyenletnek akkor és csak akkor van megoldása, ha vagy $\vec{r} = \vec{0}$, vagy A sajátértéke 1. Az feltehető, hogy $\vec{r}$ nem nulla. Az viszont nem garantálható, hogy A sajátértéke 1 legyen. Viszont egy $\lambda > 0$ skalárral való szorzás esetén megoldható, feltéve hogy A-nak létezik pozitív sajátértéke. Legyen tehát

$$A\vec{r} = \lambda\vec{r}$$

A fenti képlet tehát azt jelenti, hogy $\vec{r}$ sajátvektora A-nak, és λ pedig sajátértéke. Ahogy korábban szó volt róla, $\vec{r} \geq 0$ és $\lambda > 0$ kell hogy legyen.

Felmerül a kérdés, hogy létezik-e, mikor létezik ilyen pozitív sajátvektor és sajátérték, és hogy egyértelmű-e. Éppen erről szólt a Perron–Frobenius-tétel.

Ha $A > 0$ akkor létezik $\lambda \geq 0$ sajátérték, $\vec{r} \geq 0$ sajátvektorral. Ha pedig A nemcsak nemnegatív, hanem irreducibilis is, akkor $\vec{r} > 0$, a hozzá tartozó $\lambda > 0$ sajátérték domináns, és λ-n kívül nincs más olyan sajátérték, amihez pozitív sajátvektor tartozna.

Az irreducibilitás ennél a feladatnál a gyakorlatban azt jelentené, hogy nincs a versenyzőknek egy olyan halmaza, akik 0 pontot szereztek a halmazon kívüli versenyzők ellen. Tehát A két esetben lehet reducibilis. Egyik, ha van egy izolált csoport, aki ellen nem játszott senki a csoporton kívül.

Ebben az esetben nyilván nem lehet megmondani, hogy az izolált csoport tagjai jobbak-e vagy rosszabbak, mint a többi. Tehát ez esetben nem meglepő, hogy nem juthatunk egyértelmű eredményre. A másik eset pedig az, amikor van a versenyzőknek egy olyan halmaza, akik bár játszottak mások ellen, de 0 pontot szereztek minden mérkőzésen, ahol a halmazukon kívüli ellenféllel játszottak. Ez utóbbi esetet ki lehet küszöbölni egy olyan pontozási rendszerrel, amelyben minden mérkőzés után a vesztes fél is kap egy szigorúan pozitív pontszámot, még akkor is, ha a teljesítménye értékelhetetlenül rossz volt. Ez elvileg nem torzítaná el a ranglistát, hiszen az A mátrix elemei le vannak osztva az összes mérkőzés számával, így nem lehet sok rossz eredménnyel kiváltani egy jó eredményt, ahogy az iskolában se lehet sok 1-es összegyűjtésével javítani az átlagon.

Az $\vec{r}$ vektort ki lehet számolni a hatványiterációval, mivel az $\vec{r}$-hez tartozó λ a legnagyobb abszolút értékű sajátérték. Tehát

$$\lim_{n\to\infty} \frac{A^n\vec{v_0}}{\|A^n\vec{v_0}\|} = \vec{r}$$

ahol $\vec{v_0}$ tetszőleges nemnegatív vektor. Természetesen ha már lehet tudni $\vec{r}$ értékét, abból ki lehet számolni λ-t is, mintegy ellenőrzésképpen, hogy valóban pozitív.

4.1.1. Példa

A következőben szemléltetésképpen mutatok egy példát az előbb leírt módszer alkalmazására. Ezt a módszert akkor a leginkább célszerű használni, ha nagyon sok versenyző van nagyon sok eredménnyel. De most egy kisebb, kitalált példafeladatot mutatok, egyrészt hely- és időtakarékosságból, másrészt, hogy könnyebben át lehessen látni.

Legyen hét válogatott röplabdacsapat: Brazília, Egyiptom, Kenya, Németország, Oroszország, Szerbia és Uganda. Az egy földrészhez tartozó csapatok általában gyakrabban játszottak egymás ellen.

Egy röplabdamérkőzés három nyert szettig megy. Tehát nincs döntetlen, és a nyertes háromféleképpen nyerhet: 3 : 0-ra, 3 : 1-re és 3 : 2-re. Legyen a következő pontozási rendszer: a 3 : 0-ás győzelem 9 pontot ér, a 3 : 1-es 8 pontot, a 3 : 2-es 7 pontot. A vesztes 3 pontot kap, ha 3 : 2-re veszít, 2 pontot, ha 3 : 1-re és 1 pontot, ha 3 : 0-ra.

Legyenek a következő (fiktív) eredmények:

Németország–Brazília	3 : 2
Németország–Szerbia	3 : 1
Oroszország–Szerbia	1 : 3
Németország–Egyiptom	3 : 0
Szerbia–Brazília	1 : 3
Szerbia–Németország	3 : 2
Szerbia–Egyiptom	2 : 3
Szerbia–Németország (másik meccs)	2 : 3
Oroszország–Brazília	3 : 2
Németország–Oroszország	3 : 2
Brazília–Egyiptom	3 : 0
Egyiptom–Uganda	3 : 0
Uganda–Egyiptom	1 : 3
Uganda–Kenya	2 : 3
Kenya–Uganda	3 : 1
Uganda–Kenya	3 : 0
Kenya–Egyiptom	1 : 3
Oroszország–Egyiptom	3 : 0

Ebből az A mátrix (még nem normálva) így fog kinézni:

	B	E	K	N	O	Sz	U	összes mérk.
Brazília		9		3	3	8		4
Egyiptom	1		8	1	1	7	17	7
Kenya		2					16	4
Németország	7	9			7	18		6
Oroszország	7	9		3		2		4
Szerbia	2	3		12	8			6
Uganda		3	14					5

Ezt aztán soronként leosztjuk a mérkőzések számával, így a következő mátrixot kapjuk:

$$A = \begin{pmatrix} 0 & 2.25 & 0 & 0.75 & 0.75 & 2 & 0 \\ 0.14 & 0 & 1.14 & 0.14 & 0.14 & 1 & 2.43 \\ 0 & 0.50 & 0 & 0 & 0 & 0 & 4 \\ 1.17 & 1.50 & 0 & 0 & 1.17 & 3 & 0 \\ 1.75 & 2.25 & 0 & 0.75 & 0 & 0.50 & 0 \\ 0.33 & 0.5 & 0 & 2 & 1.33 & 0 & 0 \\ 0 & 0.60 & 2.80 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Állítás: Ez a mátrix primitív.

Az, hogy nemnegatív, az nyilvánvaló. Az, hogy irreducibilis, az is, hiszen nincsenek „elszigetelt” csapatok és a vereségért is pozitív pontszám jár, így a mátrixhoz rendelhető gráf erősen összefüggő. A Wielandt-tételből pedig tudjuk, hogy ha egy mátrix primitív, akkor az $n^2 - 2n + 2$-odik hatványa pozitív. Ez a jelenlegi 7×7-es mátrix esetében a 37. hatványt jelentené. Tehát a mátrix 37-edik hatványra emelésével bizonyítható a primitivitás. De valójában az A mátrixnak már a négyzete is szigorúan pozitív.

$$A^2 = \begin{pmatrix} 3.1756 & 3.8125 & 2.5714 & 4.8839 & 3.8631 & 4.8750 & 5.4643 \\ 0.7500 & 3.3857 & 6.8000 & 2.2143 & 1.6071 & 0.7857 & 4.5714 \\ 0.0714 & 2.4000 & 11.7714 & 0.0714 & 0.0714 & 0.5000 & 1.2143 \\ 3.2560 & 6.7500 & 1.7143 & 7.9643 & 5.0893 & 4.4167 & 3.6429 \\ 1.3631 & 5.3125 & 2.5714 & 2.6339 & 3.1756 & 8.0000 & 5.4643 \\ 4.7381 & 6.7500 & 0.5714 & 1.3214 & 2.6548 & 7.8333 & 1.2143 \\ 0.0857 & 1.4000 & 0.6857 & 0.0857 & 0.0857 & 0.6000 & 12.6571 \end{pmatrix}$$

amiből következik, hogy az A mátrix primitív. Így a mátrixra érvényesek kell, hogy legyenek a Perron-tétel állításai. Emlékeztetőül, a rangsorolás kiszámítási módja

$$\lim_{n\to\infty} \frac{A^n \vec{v_0}}{\|A^n \vec{v_0}\|} = \vec{r}.$$

Vektornormának használjuk az 1. hatványnormát, vagyis az összegnormát, ami jelen esetben a nemnegatív vektor elemeinek az összegét jelenti. Legyen a kezdővektor $\vec{v_0} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}^T$.

A MATLAB programmal a következő vektor jött ki eredménynek:

$$\vec{v} = \begin{pmatrix} 0.1993 \\ 0.0931 \\ 0.0448 \\ 0.2514 \\ 0.1837 \\ 0.1881 \\ 0.0396 \end{pmatrix}$$

Most ellenőrizzük, hogy $\vec{v}$ valóban sajátvektor, pozitív sajátértékhez: Ha $A\vec{v} = \lambda\vec{v}$, akkor ha $A\vec{v}$-t pontonként leosztjuk $\vec{v}$-vel, akkor $\lambda \cdot \begin{pmatrix} 1 & 1 & \dots & 1 \end{pmatrix}^T$-t kell kapnunk. És azt is kapjuk, az eredmény

$$\begin{pmatrix} 4.5767 \\ 4.5767 \\ 4.5767 \\ 4.5767 \\ 4.5767 \\ 4.5767 \\ 4.5767 \end{pmatrix}$$

tehát $\vec{v}$ valóban sajátvektor, és a sajátérték $\lambda = 4.5767$. Ezt a $A\vec{v} = \lambda\vec{v}$ képletbe visszahelyettesítve is ellenőrizhetjük.

Végezetül tehát a kiszámolt ranglista az országok között:

1.	Németország	0.2514
2.	Brazília	0.1993
3.	Szerbia	0.1881
4.	Oroszország	0.1837
5.	Egyiptom	0.0931
6.	Kenya	0.0448
7.	Uganda	0.0396

Az algoritmusnál alkalmazott elsőfokú normálás miatt nem meglepő, hogy a pontszámok összege 1-et ad ki. Mivel egy sajátvektor skalárszorosa is sajátvektor, azért ezen pontszámok pozitív skalárszorosa is ugyanúgy „helyes" eredmény lenne. Viszont mivel a Perron-tétel szerint a kiszámolt sajátértékhez tartozó sajátaltér egydimenziós, ezért nem létezhet a kiszámolttól lineárisan független helyes eredmény, tehát a csapatok közötti sorrend egyértelmű.

Észrevehetjük, hogy a példában bár Egyiptom hét meccsből négyet megnyert, mégis csak az 5. helyen végzett. Ugyanakkor a brazilok két győzelemmel és két vereséggel másodikok lettek. Innen láthatjuk, hogy a módszer mennyire figyelembe veszi, hogy ki ellen érték el az eredményeket. Egyiptom a pontjainak többségét a két utolsó helyezett ellen szerezte, így ezek a győzelmek nem értek sokat. Brazília viszont valószínűleg azért végzett jó helyen, mert a győzelmek és a vereségek mértékét is figyelembe vettük, és a brazilok vereségei kis mértékűek, a győzelmei nagyobbak voltak. Az viszont nem meglepő, hogy Németország végzett az első helyen, hiszen egy kivétellel minden meccsét megnyerte. És ezek a győzelmek értékesek is voltak, hiszen nem a két utolsó csapat ellen szerezték. Uganda és Kenya pedig azért került az utolsó két helyre, mert a szintén nem jól szereplő Egyiptomtól is vereséget szenvedtek, néluk erősebb csapattal pedig nem is játszottak.

Tehát elmondható, hogy összességében a módszer igazságosnak nevezhető rangsorolást eredményezett.

4.2. További lehetséges alkalmazás

Az előzőhöz hasonlóan működik a Google PageRank algoritmusa, amely a keresési találatokat rendezi sorba az oldalak „értéke” szerint. Az algoritmus azon az elven alapul, hogy egy oldal akkor értékes, ha sok más oldal hivatkozik rá. Viszont a hivatkozások nem egyenértékűek, hanem a „értékes” oldalakról jövő hivatkozások maguk is értékesebbek. Továbbá, egy oldal minél több oldalra hivatkozik, úgy az onnan érkező hivatkozások értéke fordítottan arányosan csökken. Tehát a feladat nagyon hasonló a versenyeztetéshez, azzal a különbséggel, hogy itt az oldalak nem egymás rovására szerzik a pontokat, mivel ha x oldal hivatkozis y oldalra, az nem csökkenti az y-ból x-be menő linkek számát. A módszer célja úgy, ahogy a versenyek esetében is, hogy egy „igazságos” rangsorolást adjon. Ezt pedig ugyanúgy a mátrix domináns sajátértékéhez tartozó pozitív sajátvektor kiszámításával kapjuk meg.

Irodalomjegyzék

[1] Freud Róbert: Lineáris algebra, ELTE Eötvös Kiadó, 2006

[2] Hajnal Péter előadásjegyzete,
http://www.math.u-szeged.hu/~hajnal/courses/MSc_Diszkret/MSc_kombi13/ea-sajatertek.pdf

[3] Keener, James P.: The Perron–Frobenius Theorem and the Ranking of Football Teams, 1993

[4] Lax Péter: Lineáris algebra és alkalmazásai, AKadémia Kiadó, 2008

[5] Praszolov, Viktor Vasziljevics: Lineáris algebra, Typotex kiadó, 2005