Eötvös Loránd Tudományegyetem

Természettudományi Kar

Maros Alexandra

Perron–Frobenius tételkör és alkalmazásai

Matematika BSc szakdolgozat

Témavezető:

Kunszenti-Kovács Dávid

Alkalmazott Analízis és Számításmatematikai Tanszék

Budapest, 2016

Köszönetnyilvánítás

Ezúton szeretném megköszönni témavezetőmnek, Kunszenti-Kovács Dávidnak, hogy elvállalta a konzulensi teendőket, mindig a rendelkezésemre állt, és hogy mindig segítőkészen elmagyarázta a témához kapcsolódó fogalmakat és tételeket, hogy azok szemléletes jelentését is megértsem.

Külön köszönöm Édesanyámnak és Édesapámnak, hogy végigolvasták a szakdolgozatomat, és észrevételeikkel, valamint tanácsaikkal hozzájárultak, hogy a dolgozat végső formája a lehető legjobb legyen.

Szeretnék továbbá köszönetet mondani barátaimnak, akik mindig mellettem álltak és támogattak a dolgozat írásakor. Külön köszönettel tartozom barátnőmnek, Borbély Zsófiának, aki végig nyomon követte a szakdolgozat alakulását, minden egyes verziót végigolvasott, és az esetleges hibákra felhívta a figyelmemet.

Budapest, 2016. május 27.

Maros Alexandra

Tartalomjegyzék

Bevezetés

A Perron–Frobenius tételkör egy olyan témakör, amely adott tulajdonságú mátrixok sajátértékeiről mond ki bizonyos állításokat. Bár a tétel alapvetően a lineáris algebra témakörébe tartozik, számos egyéb területen is fontos szerepet tölt be. Dolgozatom célja, hogy részletesen ismertessem a tételt, és hogy bemutassam alkalmazásait több, egymástól eltérő területen.

Az első három fejezetben ismertetem a tételhez kapcsolódó alapvető fogalmakat, a tétel eredeti verzióját, amelyet Perron dolgozott ki, illetve a kiterjesztett változatot, amely Frobenius nevéhez fűződik. Ezeket a fejezeteket alapvetően az [1] illetve a [2] irodalom alapján dolgoztam ki. A további megértéshez elolvastam még a [3], [4], [5], [6] forrásokat, de ezeket végül közvetlenül nem használtam fel a szakdolgozatomban.

A tétellel viszonylag részletesen foglalkozik egy korábbi szakdolgozat [7], azonban ott csak a nemnegatív mátrixok szemszögéből kerül elő a tétel. Ezzel ellentétben én a tételt vizsgáltam általánosan, de mivel ez a korábbi dolgozat is ismerteti a tétel egy alkalmazását (versenyzők rangsorolása), így azt én nem fejtettem ki a dolgozatomban.

A negyedik fejezetben szerepel a Perron–Frobenius tételkörnek négy különböző alkalmazása. Ezekkel azt próbáltam szemléltetni, hogy milyen sokféle területen lehet alkalmazni ezt a tételt, ráadásul olyan problémák megoldására, amelyek első ránézésre nem kapcsolódnak közvetlenül hozzá. Ezt a fejezetet az [1] irodalom alapján dolgoztam ki, de a fogalmak részletesebb megértéséhez elolvastam még néhány további forrást: a primitív mátrixok fejezetéhez a [8] irodalmat, a Leontief-féle input-output modellhez a [9], a Leslie-modellhez a [10], illetve a Markov-láncok fejezetéhez a [11] forrásanyagot, valamint egy másik korábbi szakdolgozatot [12].

Az ötödik fejezetben bemutatom, hogy hogyan működik a Google kereső algoritmusa, és hogy ebben is milyen fontos szerepet tölt be a Perron–Frobenius tétel. Ennek a fejezetnek a kidolgozásához a [13] irodalmat használtam fel.

Az utolsó fejezet egy rövid kitekintést nyújt a tétel végtelen dimenziós általánosítására. Ez azért lényeges, mert a megelőző fejezetekben csak véges dimenziós esetekről olvashatunk, azonban a tétel végtelen dimenziós általánosítása fontos szerepet tölt be a parciális differenciálegyenletek témakörében. Ezt a fejezetet a [14] irodalom alapján dolgoztam ki, de itt sok segítséget kaptam a témavezetőmtől is, hogy megértsem a témához kapcsolódó fogalmak jelentését. Ebben a fejezetben számos gondolatot az ő magyarázatai alapján fejtettem ki. Erről a témakörről részletesebben a [15] irodalomban olvashatunk.

A szakdolgozatom írásakor az eredeti szakirodalmakban szereplő bizonyításokat sokszor kiegészítettem önálló magyarázatokkal, azért, hogy érthetőbb legyen a bizonyítások menete. A dolgozatban szereplő ábrákat önállóan hoztam létre a LaTeX programban. A fejezetekben említett példákat a megjelölt források alapján önállóan dolgoztam ki úgy, hogy az eredeti példákat részben módosítottam. Erre azért volt szükség, hogy lássam, hogy ha apró módosításokat végzek, hogyan változik az adott modell. Az ehhez szükséges számításokat pedig MATLAB programmal végeztem el.

Összességében úgy gondolom, hogy a Perron–Frobenius tételkör egy nagyon érdekes téma, amelyet tulajdonképpen az emberek nap mint nap használnak. Például ennek a tételnek köszönhetjük, hogy amikor az interneten rákeresünk egy adott kifejezésre, akkor a kapott találatok sorrendje jól tükrözi azt, hogy melyik oldal mennyire hasznos számunkra a keresett kifejezéssel kapcsolatban.

Ahogy Meyer írja [1, 662. oldal 4. sor]: *„A Perron–Frobenius tételkör elegáns. Annak a ténynek a megtestesülése, hogy a gyönyörű matematika néha hasznos, és a hasznos matematika néha gyönyörű.”*.

1. fejezet

Algebrai összefoglaló

Ebben a fejezetben tisztázzuk azokat az algebrai fogalmakat, illetve jelöléseiket, amelyeket a későbbi fejezetek során használni fogunk. Egy A mátrix és egy v vektor koordinátáit a szokásos módon a_{ij} és v_i jelöli, a szorzatuk koordinátáit pedig $[Av]_i$ fogja jelölni. Továbbá $|A|$ az A mátrix abszolút értékét jelöli, tehát a koordinátái az eredeti mátrix $|a_{ij}|$ elemei. A mátrix rangját $\mathrm{rang}(A)$ fogja jelölni.

1.0.1. Definíció. Egy $A \in \mathbb{R}^{m\times n}$ mátrix *pozitív*, ha minden i-re és j-re $a_{ij} > 0$. Jelölés: $A > 0$. Ha $B \in \mathbb{R}^{m\times n}$, akkor $A > B$ azt jelenti, hogy minden i-re és j-re $a_{ij} > b_{ij}$. Hasonlóan, A *nemnegatív*, ha minden $a_{ij} \geq 0$ (jelölés: $A \geq 0$) és $A \geq B$, ha minden $a_{ij} \geq b_{ij}$.

1.0.2. Definíció. Egy $A \in \mathbb{R}^{n\times n}$ mátrixnak egy $\lambda \in \mathbb{C}$ szám a *sajátértéke*, ha létezik olyan $v \in \mathbb{R}^{n\times 1}$, $v \neq 0$, hogy $Av = \lambda v$. Ekkor v a λ-hoz tartozó *jobb oldali sajátvektor*. Ha pedig $w \in \mathbb{R}^{1\times n}$, $w \neq 0$ és $wA = \lambda w$, akkor w a λ-hoz tartozó *bal oldali sajátvektor*.

1.0.3. Definíció. Egy A mátrix *spektrumának* nevezzük a sajátértékeinek a halmazát. Jelölés: $\sigma(A)$.

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

1.0.5. Definíció. Egy A mátrix *hasonló* egy B mátrixhoz, ha létezik olyan C invertálható mátrix, hogy $B = C^{-1}AC$.

1.0.6. Definíció. Egy A mátrix *magterén* a $\mathrm{Ker}(A) = \{x \in \mathbb{R}^n \mid Ax = 0\}$ halmazt értjük. A mátrix *képtere* pedig $\mathrm{Im}(A) = \{Ax \mid x \in \mathbb{R}^n\}$.

1.0.7. Állítás. *Egy A mátrix rangja megegyezik a képterének a dimenziójával, azaz* $\operatorname{rang}(A) = \dim \operatorname{Im}(A)$.

Bizonyítás: Legyen $x \in \mathbb{R}^n$ tetszőleges és $\operatorname{rang}(A) = m$, azaz A oszlopaiból m lineárisan függetlent tudunk kiválasztani. Jelölje az A mátrix i-edik oszlopát a_i, ekkor az Ax szorzat a következőképpen írható fel: $Ax = a_1x_1 + \ldots + a_nx_n$. Ekkor az összes lehetséges Ax szorzatot az A oszlopai közül kiválasztható m független oszlop generálja, tehát az $\{Ax \mid x \in \mathbb{R}^n\}$ halmaz m dimenziós, azaz $\dim \operatorname{Im}(A) = m = \operatorname{rang}(A)$. $\square$

1.0.8. Definíció. (Jordan-féle normálalak) Legyen $A \in \mathbb{R}^{n\times n}$, és legyen J_i egy olyan négyzetes mátrix, amelynek a főátlójában végig az A mátrix λ_i sajátértéke szerepel, a főátló feletti átlóban végig 1-esek szerepelnek, az összes többi eleme pedig nulla, és legyen J a következő mátrix:

$$J = \begin{pmatrix} J_1 & & 0 \\ & \ddots & \\ 0 & & J_m \end{pmatrix}, \quad \text{ahol} \quad J_i = \begin{pmatrix} \lambda_i & 1 & & 0 \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_i \end{pmatrix}$$

Ekkor J főátlójában az úgynevezett *Jordan-blokkok* szerepelnek, azon kívül pedig nullák. Ha létezik J-hez egy olyan P invertálható mátrix, amelyre A felírható $A = PJP^{-1}$ alakban (azaz, ha A hasonló J-hez), akkor azt mondjuk, hogy J az A mátrix *Jordan-féle normálalakja.*

1.0.9. Definíció. Egy $A \in \mathbb{R}^{n\times n}$ mátrix *nilpotens*, ha létezik olyan $k \in \mathbb{N}$ szám, amelyre $A^k = 0$.

1.0.10. Definíció. Egy $A \in \mathbb{R}^{n\times n}$ mátrix *reducibilis*, ha létezik egy olyan $P \in \mathbb{R}^{n\times n}$ permutációs mátrix, illetve B, C, D mátrixok úgy, hogy

$$P^T AP = \begin{pmatrix} B & C \\ 0_{n-r,r} & D \end{pmatrix}, \quad \text{ahol} \quad 1 \le r \le n-1.$$

Itt $0_{n-r,r}$ egy olyan $(n-r)\times r$ méretű mátrixot jelöl, amelynek minden eleme nulla, $B \in \mathbb{R}^{r\times r}$ és $C \in \mathbb{R}^{(n-r)\times(n-r)}$ pedig négyzetes mátrixok. Vegyük észre, hogy a B, C és D mátrixokról csak annyit teszünk fel, hogy van legalább egy nemnulla elemük, továbbá fontos, hogy B és D legalább 1×1-es mátrixok legyenek. Így például az 1×1-es mátrixok nem reducibilisek.

Ezt a P^TAP mátrixot az A mátrix *szimmetrikus permutációjának* nevezzük, ugyanis pontosan azokat a sorokat cseréli fel az eredeti A mátrixban, amely oszlopokat is.

1.0.11. Definíció. Egy A mátrix *irreducibilis*, ha nem reducibilis.

1.0.12. Definíció. Az $A \in \mathbb{R}^{n \times n}$ mátrixhoz tartozó $\mathcal{G}(A)$ *gráf* egy olyan irányított gráf, amelynek n csúcsa van, legyenek ezek $\{V_1, \ldots, V_n\}$, és egy V_i csúcsból irányított él vezet a V_j csúcsba akkor és csak akkor, ha $a_{ij} \neq 0$.

Vegyük észre, hogy ha P egy permutációs mátrix, akkor $\mathcal{G}(P^T AP) = \mathcal{G}(A)$, hiszen az A mátrixhoz tartozó gráf csúcsait átcímkézve könnyen megkaphatjuk a $P^T AP$ mátrixhoz tartozó gráfot.

1.0.13. Definíció. Egy irányított $\mathcal{G}$ gráf *erősen összefüggő*, ha bármely pontjából bármely más pontjába vezet irányított út.

1.0.14. Tétel. *Egy A mátrix akkor és csak akkor irreducibilis, ha a hozzá tartozó $\mathcal{G}(A)$ gráf erősen összefüggő.*

Ennek a tételnek a bizonyítása megtalálható az [1] irodalomban, a 4.4.20. feladat megoldásánál a 36. oldalon.

1.0.15. Definíció. Ha létezik olyan $k > 1$ szám, amelyre az A^k mátrix főátlójában minden elem pozitív, akkor azt mondjuk, hogy A *periodikus*. Ha ez teljesül $k = 1$ esetén (azaz ha minden i esetén $a_{ii} > 0$), akkor azt mondjuk, hogy A *aperiodikus*.

2. fejezet

Perron–Frobenius tétel pozitív mátrixokra

Ez a tétel két német matematikus, Oskar Perron és Ferdinand Georg Frobenius nevéhez fűződik. A tételt Perron 1907-ben eredetileg csak pozitív mátrixok esetére dolgozta ki, később pedig Frobenius terjesztette ki nemnegatív mátrixokra 1912-ben. Ebben a fejezetben a pozitív mátrixokról szóló Perron tétellel ismerkedhetünk meg.

2.0.1. Tétel. (Perron) *Legyen $A \in \mathbb{R}^{n \times n}$ pozitív mátrix, $\rho(A) = r$. Ekkor a következő állítások teljesülnek.*

- *$r > 0$, azaz A spektrálsugara pozitív.* (2.1)
- *$r \in \sigma(A)$, azaz r sajátértéke A-nak és ekkor azt mondjuk, hogy r az A-hoz tartozó **Perron–Frobenius sajátérték**.* (2.2)
- *Létezik egy $x > 0$ sajátvektor, amelyre $Ax = rx$.* (2.3)
- *r egyszeres sajátértéke A-nak.* (2.4)
- *A-nak egyértelműen létezik egy olyan $p > 0$ sajátvektora, amelyre $Ap = rp$ és $\sum_{i=1}^{n} p_i = 1$, továbbá A-nak p pozitív skalárszorosain kívül nem létezik más nemnegatív sajátvektora. Ez az úgynevezett **Perron vektor**.* (2.5)
- *Nem létezik A-nak másik olyan sajátértéke, amelynek az abszolút értéke egyenlő r-rel.* (2.6)

- ***Collatz-Wielandt formula:*** $r = \max_{x \in \mathcal{N}} f(x)$,

  ahol $f(x) = \min_{\substack{1 \le i \le n \\ x_i \ne 0}} \frac{[Ax]_i}{x_i}$ *és* $\mathcal{N} = \{x \mid x \ge 0 \text{ és } x \ne 0\}$ (2.7)

- *Ha A minden sorának (vagy oszlopának) az összege egyenlő egy ρ számmal, akkor $r = \rho$.* (2.8)

- $\min_i \sum_{j=1}^{n} a_{ij} \le r \le \max_i \sum_{j=1}^{n} a_{ij}$ (2.9)

Bizonyítás:

- $r > 0$*, azaz A spektrálsugara pozitív.*

  Indirekt tegyük fel, hogy $\rho(A) = 0$, azaz $\sigma(A) = \{0\}$. Ez azt jelenti, hogy A-nak az egyetlen sajátértéke a 0, vagyis A Jordan normálalakjában a főátlóban végig nullák szerepelnek, tehát egy J_i Jordan-blokk a következőképpen néz ki:

$$J_i = \begin{pmatrix} 0 & 1 & & \text{\Large 0} \\ & 0 & \ddots & \\ & & \ddots & 1 \\ \text{\Large 0} & & & 0 \end{pmatrix}$$

  Ha egy ilyen mátrixot elégszer hatványozunk, végül nullmátrixot kapunk, tehát létezik egy olyan k_i szám, amelyre $J_i^{k_i} = 0$. Tudjuk, hogy a J mátrix k-adik hatványa úgy néz ki, hogy a főátlóban a Jordan-blokkok k-adik hatványai szerepelnek, az összes többi eleme pedig nulla:

$$J = \begin{pmatrix} J_1 & & \text{\Large 0} \\ & \ddots & \\ \text{\Large 0} & & J_m \end{pmatrix} \quad \Longrightarrow \quad J^k = \begin{pmatrix} J_1^k & & \text{\Large 0} \\ & \ddots & \\ \text{\Large 0} & & J_m^k \end{pmatrix}$$

  Így, ha ezt a J mátrixot hatványozzuk, egy idő után nullmátrixot kapunk, tehát létezik egy olyan k szám ($k = \max_i k_i$), amelyre $J^k = 0$, vagyis A Jordan normálalakja nilpotens, és így A is nilpotens, mivel $A^k = PJ^kP^{-1} = 0$. Ez viszont ellentmondás, ugyanis A pozitív.

- *$r \in \sigma(A)$ és létezik egy $x > 0$ sajátvektor, amelyre $Ax = rx$.*
  $\rho(A) = r$ azt jelenti, hogy A-nak létezik egy olyan λ sajátértéke, amire $|\lambda| = r$. Azt fogjuk belátni, hogy $\lambda = r$ és a hozzá tartozó sajátvektor pozitív. Feltehető, hogy $\rho(A) = 1$, ugyanis $A > 0$ akkor és csak akkor, ha $A/\rho(A) > 0$, és $\rho(A) = r$ ekvivalens azzal, hogy $\rho(A/r) = 1$. Tehát tegyük fel, hogy $\rho(A) = 1$, ekkor $|\lambda| = 1$ és legyen a λ-hoz tartozó sajátvektor x, azaz $Ax = \lambda x$. Ekkor:

  $$|x| = |\lambda||x| = |\lambda x| = |Ax| \leq |A||x| = A|x| \quad \Longrightarrow \quad |x| \leq A|x| \tag{2.10}$$

  Azt fogjuk megmutatni, hogy egyenlőség áll fenn. Legyen $z = A|x|$ és $y = z - |x|$, ekkor a (2.10) egyenlőtlenség a következő módon írható át:

  $$|x| \leq A|x| = z \quad \Longrightarrow \quad 0 \leq z - |x| = y$$

  Tegyük fel, hogy $y \neq 0$, ekkor (mivel y nemnegatív) létezik egy olyan i index, amelyre $y_i > 0$. Abból, hogy $A > 0$ és $y \geq 0$, az következik, hogy $Ay > 0$ és $z > 0$. Ekkor létezik $\epsilon > 0$ úgy, hogy:

  $$\begin{aligned} \epsilon z < Ay =& A(z - |x|) = Az - A|x| = Az - z \\ &\Longrightarrow \quad Az - z > \epsilon z \\ &\Longrightarrow \quad Az > z + \epsilon z = (1+\epsilon)z \\ &\Longrightarrow \quad \frac{A}{1+\epsilon} z > z \end{aligned}$$

  Vezessük be a $B = \dfrac{A}{1+\epsilon}$ jelölést, így a fenti egyenlőtlenség alapján $Bz > z$, amit ha beszorzunk B-vel balról, a következőt kapjuk:

  $$B^2 z > Bz > z \implies B^3 z > B^2 z > z \implies \ldots \implies B^k z > z \quad \forall k \in \mathbb{N}^+ \tag{2.11}$$

  Vegyük azonban figyelembe a következőket:

  $$\rho(B) = \rho\left(\frac{A}{1+\epsilon}\right) = \frac{1}{1+\epsilon} < 1 \quad \Longrightarrow \quad \lim_{k\to\infty} B^k = 0 \tag{2.12}$$

  Ahhoz, hogy ezt belássuk, bizonyítás nélkül felhasználjuk a következő egyenlőséget, ez az úgynevezett *Gelfand formula*. Az állítás bizonyítása megtalálható az [1] irodalomban, a 7.10.1 példánál a 619. oldalon.

  $$\rho(A) = \lim_{k\to\infty} \left\|A^k\right\|^{1/k} \tag{2.13}$$

Ez teljesül tetszőleges mátrix norma esetén, tehát speciálisan a végtelen normára is igaz. Így tehát azt írhatjuk fel, hogy:

$$1 > \rho(B) = \lim_{k\to\infty} \left\|B^k\right\|_\infty^{1/k}$$

Ekkor létezik egy olyan $\widetilde{\epsilon} \in (0,1)$ és $N \in \mathbb{N}$, hogy minden $k \geq N$ esetén:

$$\left\|B^k\right\|_\infty^{1/k} < 1 - \widetilde{\epsilon}$$

Így, ha mindkét oldalt a k-adik hatványra emeljük:

$$0 < \left\|B^k\right\|_\infty < (1-\widetilde{\epsilon})^k < 1$$

Ekkor $(1-\widetilde{\epsilon})^k \longrightarrow 0$, ha $k \longrightarrow \infty$, és így mivel az egyenlőtlenség jobb oldala nullához tart, a rendőr elv alapján $\left\|B^k\right\|_\infty \longrightarrow 0$, amiből pedig az következik, hogy $B^k \longrightarrow 0$. Ezzel tehát beláttuk a (2.12) állítást.

Ekkor a (2.11) egyenlőtlenségből határátmenettel azt kapjuk, hogy $0 \geq z$, ami ellentmondás, hiszen z pozitív. Tehát helytelen volt az a feltevésünk, hogy $y \neq 0$, amiből az következik, hogy $0 = y = A|x| - |x|$. Ezzel bebizonyítottuk, hogy $r = 1$ sajátértéke az A mátrixnak, és a hozzá tartozó sajátvektor pedig $|x| > 0$.

Ezzel a bizonyítással tulajdonképpen a következő fontos állítást is beláttuk:

$$\textit{Ha } Ax = \rho(A)x \textit{ és } x \neq 0, \textit{ akkor } A|x| = \rho(A)|x| \textit{ és } |x| > 0. \tag{2.14}$$

- *r egyszeres sajátértéke A-nak.*
  Ahhoz, hogy ezt belássuk, szükségünk lesz a következő állításra, tehát először ezt fogjuk igazolni:

  Az A mátrix Jordan normálalakjában a $\rho(A)$-hoz tartozó legnagyobb Jordan-blokk 1×1-es méretű. (2.15)

  Feltehető, hogy $\rho(A) = 1$, ugyanis $A > 0$ akkor és csak akkor, ha $A/\rho(A) > 0$, és $\rho(A) = r$ ekvivalens azzal, hogy $\rho(A/r) = 1$. Tehát tegyük fel, hogy $\rho(A) = 1$ és tegyük fel, hogy az $r = 1$-hez tartozó legnagyobb Jordan-blokk mérete $\widetilde{m} > 1$, legyen ez az i-edik Jordan-blokk, azaz $J_i \in \mathbb{R}^{\widetilde{m}\times\widetilde{m}}$, ahol:

$$J_i = \begin{pmatrix} 1 & 1 & & 0 \\ & 1 & \ddots & \\ & & \ddots & 1 \\ 0 & & & 1 \end{pmatrix}$$

Ha ezt a mátrixot hatványozzuk, a mátrix elemei egyre nagyobbak lesznek, tehát $\left\|J_i^k\right\|_\infty \longrightarrow \infty$, ha $k \longrightarrow \infty$, amiből pedig az következik, hogy $\left\|J^k\right\|_\infty \longrightarrow \infty$. Mivel $A = PJP^{-1}$, ezért $J = P^{-1}AP$, tehát $J^k = P^{-1}A^kP$, aminek ha vesszük a végtelen normáját, a következőt kapjuk:

$$\left\|J^k\right\|_\infty = \left\|P^{-1}A^kP\right\|_\infty \leq \left\|P^{-1}\right\|_\infty \left\|A^k\right\|_\infty \|P\|_\infty$$

Oszzuk el az egyenlőtlenség két oldalát $\|P^{-1}\|_\infty \|P\|_\infty$-val:

$$\left\|A^k\right\|_\infty \geq \frac{\left\|J^k\right\|_\infty}{\|P^{-1}\|_\infty \|P\|_\infty} \longrightarrow \infty$$

A rendőr-elv alapján tehát $\left\|A^k\right\|_\infty \longrightarrow \infty$ ha $k \longrightarrow \infty$. Jelölje az A^k mátrix i-edik sorának j-edik elemét $a_{ij}^{(k)}$, és legyen i_k az a sor A^k-ban, amelynek a legnagyobb a sorösszege, azaz:

$$\left\|A^k\right\|_\infty = \sum_j a_{i_kj}^{(k)}$$

A tétel (2.3) pontja alapján létezik egy olyan $x > 0$, amelyre $x = Ax$. Nézzük a következő egyenlőtlenséget:

$$\|x\|_\infty \geq x_{i_k} = [Ax]_{i_k} = \sum_j a_{i_kj}^{(k)} x_j \geq \left(\sum_j a_{i_kj}^{(k)}\right)(\min_i x_i) \longrightarrow \infty$$

Ebből az következik, hogy $\|x\|_\infty \longrightarrow \infty$, ami viszont ellentmondás, hiszen x egy konstans vektor. Így tehát hibás volt az a feltevésünk, miszerint $\widetilde{m} > 1$, tehát $\widetilde{m} = 1$, azaz beláttuk a (2.15) állítást. A tétel (2.4) pontjának igazolásához szükségünk lesz a következő állításra:

Ha az A mátrixnak r egy m-szeres sajátértéke, akkor az r-hez tartozó lineárisan független sajátvektorok száma éppen m. (2.16)

Ahhoz, hogy ezt bebizonyítsuk, vizsgáljuk meg, hogy hogyan is néz ki A Jordan-normálalakja. A (2.15) állítás alapján az r-hez tartozó összes Jordan-blokk 1×1-es méretű, tehát:

$$P^{-1}AP = J = \begin{pmatrix} rI_{m,m} & 0 \\ 0 & B \end{pmatrix}$$

Itt $I_{m,m}$ az $m \times m$ méretű egységmátrixot jelöli, B pedig egy olyan mátrix, amelyre $r \notin \sigma(B)$. Vizsgáljuk meg az $(A - rI)$ mátrix rangját:

$$\operatorname{rang}(A - rI) = \operatorname{rang}\left(P \begin{pmatrix} 0 & 0 \\ 0 & B - rI \end{pmatrix} P^{-1}\right) = \operatorname{rang}(B - rI) = n - m$$

Most használjuk fel a dimenziótételt:

$$\dim \operatorname{Im}(A - rI) + \dim \operatorname{Ker}(A - rI) = n \tag{2.17}$$

A 1.0.7 Állítás alapján $\dim \operatorname{Im}(A - rI) = \operatorname{rang}(A - rI)$, és mivel kiszámoltuk, hogy $\operatorname{rang}(A - rI) = n - m$, így a (2.17) egyenlőségből azt kapjuk, hogy:

$$n - m + \dim \operatorname{Ker}(A - rI) = n \quad \Longrightarrow \quad \dim \operatorname{Ker}(A - rI) = m$$

Vegyük észre, hogy $\dim \operatorname{Ker}(A - rI) = m$ azt jelenti, hogy az r-hez tartozó lineárisan független sajátvektorok száma éppen m, ezzel tehát beláttuk a (2.16) állítást. Nézzük tehát a tétel (2.4) pontjának bizonyítását, azaz lássuk be, hogy r egyszeres sajátértéke A-nak.

Tegyük fel, hogy $r = 1$ egy olyan sajátértéke A-nak, amelynek a multiplicitása $m > 1$. Ez tehát a (2.16) állítás alapján azt jelenti, hogy az r-hez tartozó lineárisan független sajátvektorok száma éppen m. Ha x és y két olyan sajátvektor, amelyek az $r = 1$ sajátértékhez tartoznak, és lineárisan függetlenek, akkor minden $\alpha \in \mathbb{C}$ esetén $x \neq \alpha y$. Ekkor tehát $Ax = x$ és $Ay = y$, azaz:

$$Ax - x = 0 = Ay - y \tag{2.18}$$

Mivel az egyenlet jobb és bal oldala is a nullvektorral egyenlő, ezért megtehetjük, hogy az egyenlet jobb oldalát beszorozzuk egy konstanssal, hiszen akkor is ugyanúgy a nullvektort kapjuk eredményül. Válasszunk ki y-ból egy nemnulla komponenst, legyen ez $y_i \neq 0$, és szorozzuk be az x_i/y_i számmal a (2.18) jobb oldalát:

$$\begin{aligned} Ax - x &= \left(\frac{x_i}{y_i}\right) Ay - \left(\frac{x_i}{y_i}\right) y \\ Ax - \left(\frac{x_i}{y_i}\right) Ay &= x - \left(\frac{x_i}{y_i}\right) y \\ A\left(x - \left(\frac{x_i}{y_i}\right) y\right) &= x - \left(\frac{x_i}{y_i}\right) y \end{aligned} \tag{2.19}$$

Legyen $z = x - (x_i/y_i)y$, ekkor $z \neq 0$ mivel x és y függetlenek. A (2.19) egyenlőség alapján $Az = z$, azaz z sajátvektora A-nak, ekkor a (2.14) állítás alapján $A|z| = |z| > 0$, de:

$$z_i = x_i - \left(\frac{x_i}{y_i}\right) y_i = x_i - x_i = 0$$

Ez pedig ellentmondás, tehát hibás volt a feltevésünk, miszerint $m > 1$, tehát $m = 1$, azaz r valóban egyszeres sajátértéke A-nak.

- *A-nak egyértelműen létezik egy $p > 0$ sajátvektora, amelyre $Ap = rp$ és $\sum\limits_{i=1}^{n} p_i = 1$.*
  Azt már beláttuk a tétel (2.3) pontjánál, hogy létezik A-nak olyan pozitív sajátvektora, amely az r sajátértékhez tartozik, tehát csak annyit kell igazolni, hogy ez előállítható úgy, hogy az elemeinek összege 1-gyel egyenlő, és hogy ez egyértelmű. Legyen x olyan pozitív vektor, hogy $Ax = rx$, és legyen p a következő vektor:

$$p = \frac{x}{\|x\|_1}$$

Ekkor p elemeinek az összege pont 1 lesz, hiszen:

$$\sum_{i=1}^{n} p_i = \sum_{i=1}^{n} \frac{x_i}{\|x\|_1} = \frac{\sum\limits_{i=1}^{n} x_i}{\|x\|_1} = \frac{\|x\|_1}{\|x\|_1} = 1$$

Így már csak az egyértelműséget kell igazolni. Tegyük fel, hogy léteznek p_1 és p_2 sajátvektorai A-nak úgy, hogy $Ap_1 = rp_1$, $p_1 > 0$, $\|p_1\|_1 = 1$, illetve $Ap_2 = rp_2$, $p_2 > 0$, $\|p_2\|_1 = 1$. Ez azt jelenti, hogy p_1 és p_2 is az r sajátértékhez tartozó sajátvektorai A-nak. Mivel r csak egyszeres sajátérték, ezért a hozzá tartozó lineárisan független sajátvektorok száma is 1, amiből pedig az következik, hogy valamilyen $\alpha \in \mathbb{R}$ esetén:

$$p_1 = \alpha p_2$$

Vegyük mindkét oldal 1-es normáját:

$$1 = \|p_1\|_1 = \|\alpha p_2\|_1 = |\alpha| \, \|p_2\|_1 = |\alpha| \quad \Longrightarrow \quad \alpha = \pm 1$$

Mivel p_1 és p_2 is pozitívak, így biztos, hogy $\alpha \neq -1$, tehát $\alpha = 1$, és így $p_1 = p_2$.

- *A-nak p skalárszorosain kívül nem létezik más nemnegatív sajátvektora.* Legyen λ olyan sajátértéke és $y \geq 0$ olyan sajátvektora A-nak, amelyre $Ay = \lambda y$. Legyen $\widetilde{p} > 0$ az A^T-hoz tartozó Perron vektor, azaz:

$$A^T \widetilde{p} = r\widetilde{p}$$

Ha ezt transzponáljuk, majd beszorozzuk mindkét oldalt jobbról y-nal, akkor azt kapjuk, hogy:

$$\widetilde{p}^T Ay = r\widetilde{p}^T y$$

Felhasználva, hogy $Ay = \lambda y$, azt kapjuk, hogy:

$$\widetilde{p}^T \lambda y = r\widetilde{p}^T y \quad \Longrightarrow \quad \lambda\widetilde{p}^T y = r\widetilde{p}^T y$$

Mivel $\widetilde{p}^T y > 0$, ez pontosan akkor teljesül, ha $\lambda = r$, és mivel az előbb már beláttuk, hogy A-nak egyértelműen létezik r-hez tartozó sajátvektora, így létezik egy olyan $\alpha \geq 0$ szám, amelyre $y = \alpha p$.

- *Nem létezik A-nak másik olyan sajátértéke, amelynek az abszolút értéke egyenlő r-rel.*
Ismét feltehető, hogy $r = 1$. Legyen $|\lambda| = 1$ sajátértéke A-nak, és legyen a hozzá tartozó sajátvektor x, azaz:

$$\sum_{j=1}^{n} a_{kj}x_j = \lambda x_k \quad \forall k$$

A (2.14) állítás alapján $0 < |x| = A|x|$, tehát:

$$0 < |x_k| = [A\,|x|]_k = \sum_{j=1}^{n} a_{kj}\,|x_j|$$

Ugyanakkor vegyük figyelembe a következőeket:

$$|x_k| = |\lambda||x_k| = |\lambda x_k| = |[\lambda x]_k| = |\,[Ax]_k\,| = \left|\sum_{j=1}^{n} a_{kj}x_j\right|$$

Ha egyenlővé tesszük ezt a két szummát, amit $|x_k|$-ra felírtunk, akkor azt kapjuk, hogy:

$$\left|\sum_{j=1}^{n} a_{kj}x_j\right| = \sum_{j=1}^{n} a_{kj}\,|x_j| = \sum_{j=1}^{n} |a_{kj}x_j| \qquad (2.20)$$

Itt a második egyenlőség azért teljesül, mert A minden eleme pozitív. Tudjuk, hogy ha a $\{z_1, \ldots, z_n\} \subset \mathbb{C}^n$ nemnulla vektoroknak egy halmaza, akkor $\| \sum_j z_j \|_2 = \sum_j \|z_j\|_2$ akkor és csak akkor, ha minden j esetén létezik olyan $\alpha_j > 0$, hogy $z_j = \alpha_j z_1$. Ez azért igaz, mert a kettes norma az euklideszi távolságot jelöli, és a vektorok összegének a távolsága csak akkor lehet egyenlő a távolságaik összegével, ha minden vektor egy irányba mutat, azaz ha egymásnak a pozitív többszörösei. Ezt alkalmazhatjuk vektorok helyett skalárokra is, és így a (2.20) egyenlőség alapján léteznek olyan $\alpha_j > 0$ skalárok, hogy:

$$a_{kj}x_j = \alpha_j(a_{k1}x_1), \text{ azaz } x_j = \pi_j x_1, \text{ ahol } \pi_j = \frac{\alpha_j a_{k1}}{a_{kj}} > 0$$

Ekkor tehát ha $|\lambda| = 1$, akkor $x = x_1 p$, ahol $p = \begin{pmatrix} 1 & \pi_2 & \dots & \pi_n \end{pmatrix}^T > 0$, így:

$$\lambda x = Ax \implies \lambda p = Ap = |Ap| = |\lambda p| = |\lambda||p| = |\lambda| p = p \implies \lambda = 1$$

Itt $Ap = |Ap|$ azért teljesül, mert A is és p is pozitívak, tehát Ap is. Így tehát azt kaptuk, hogy $\lambda = 1 = r$, tehát valóban a Perron–Frobenius sajátérték az egyetlen r abszolút értékű sajátértéke A-nak.

- $r = \max\limits_{x \in \mathcal{N}} f(x)$

  Legyenek A r-hez tartozó jobb és baloldali sajátvektorai p és q^T. Legyen $x \in \mathcal{N}$ és $\xi = f(x)$. Ekkor $0 \leq \xi x \leq Ax$, ugyanis ha indirekt feltesszük, hogy létezik egy olyan j index, amelyre $[Ax]_j < \xi x_j$, akkor:

$$\frac{[Ax]_j}{x_j} < \xi = \min_{\substack{1 \leq i \leq n \\ x_i \neq 0}} \frac{[Ax]_i}{x_i}$$

  Ez pedig ellentmondás, tehát valóban $\xi x \leq Ax$. Ha az egyenlőtlenség mindkét oldalát beszorozzuk balról q^T-tal (ezt megtehetjük, hiszen q nemnegatív) és felhasználjuk, hogy $q^T A = r q^T$, akkor azt kapjuk, hogy:

$$\xi q^T x \leq q^T A x = r q^T x$$

  Mivel $q^T x > 0$, így ebből következik, hogy $\xi \leq r$, azaz $f(x) \leq r$ minden $x \in \mathcal{N}$ esetén. Mivel p az A-hoz tartozó Perron vektor, így $Ap = rp$, amiből következik, hogy minden i index esetén $[Ap]_i = r p_i$, tehát:

$$\frac{[Ap]_i}{p_i} = r$$

  Ez teljesül minden i-re, így a minimum is egyenlő r-rel, tehát $f(p) = r$. Mivel $p \in \mathcal{N}$, így azt kapjuk, hogy $r = \max\limits_{x \in \mathcal{N}} f(x)$.

- *Ha A minden sorának (vagy oszlopának) az összege egyenlő egy ρ számmal, akkor $r = \rho$.*
  Legyen e a csupa 1-esekből álló vektor, ekkor $Ae = \rho e$, ami azt jelenti, hogy az A mátrixnak e egy sajátvektora. Mivel e pozitív, így a tétel (2.5) pontja miatt biztos, hogy ez az A-hoz tartozó p Perron vektornak valamely skalárszorosa, azaz létezik egy olyan α szám, amelyre $p = \alpha e$. Helyettesítsük ezt be az $Ap = rp$ egyenletbe:

$$A\alpha e = r\alpha e$$

  Mivel $Ae = \rho e$, így azt kapjuk, hogy:

$$\alpha\rho e = r\alpha e \quad \Longrightarrow \quad \rho = r$$

- $\min\limits_i \sum\limits_{j=1}^n a_{ij} \leq r \leq \max\limits_i \sum\limits_{j=1}^n a_{ij}$
  Ehhez először belátjuk az alábbi állítást, amit majd később felhasználunk a bizonyítás során.

$$|A| \leq B \quad \Longrightarrow \quad \rho(A) \leq \rho(|A|) \leq \rho(B) \tag{2.21}$$

  A háromszög egyenlőtlenségből következik, hogy $|A^k| \leq |A|^k$ minden k pozitív egész esetén. Ezenkívül abból, hogy $0 \leq |A| \leq B$, az következik, hogy $|A|^k \leq B^k$. A (2.13) egyenlőség (Gelfand formula), illetve az előbbi két egyenlőtlenség alapján:

$$\begin{gathered}
\left\|A^k\right\|_\infty = \left\||A^k|\right\|_\infty \leq \left\||A|^k\right\|_\infty \leq \left\|B^k\right\|_\infty \\
\Downarrow \\
\left\|A^k\right\|_\infty^{1/k} \leq \left\||A|^k\right\|_\infty^{1/k} \leq \left\|B^k\right\|_\infty^{1/k} \\
\Downarrow \\
\lim_{k\to\infty} \left\|A^k\right\|_\infty^{1/k} \leq \lim_{k\to\infty} \left\||A|^k\right\|_\infty^{1/k} \leq \lim_{k\to\infty} \left\|B^k\right\|_\infty^{1/k} \\
\Downarrow \\
\rho(A) \leq \rho(|A|) \leq \rho(B)
\end{gathered}$$

  Ezzel beláttuk a (2.21) állítást. Most nézzük tehát a tételben szereplő egyenlőtlenség bizonyítását. Legyen ρ a legnagyobb sorösszeg A-ban, azaz:

$$\rho = \max_i \sum_{j=1}^n a_{ij}$$

Ekkor létezik egy olyan $E \geq 0$ mátrix, amelyre teljesül, hogy $B = A+E$ minden sorának összege egyenlő ρ-val. A tétel (2.7) pontja alapján ekkor $\rho(B) = \rho$, tehát (2.21)-et felhasználva azt kapjuk, hogy:

$$r = \rho(A) \leq \rho(B) = \rho = \max_i \sum_{j=1}^{n} a_{ij}$$

Ezzel beláttuk a jobb oldali egyenlőtlenséget. A bal oldali bizonyításához legyen e a csupa 1-esekből álló vektor, ekkor $e \in \mathcal{N}$, így a Collatz-Wielandt formula alapján:

$$r = \max_{x \in \mathcal{N}} f(x) \geq f(e) = \min_i \frac{[Ae]_i}{e_i} = \min_i \sum_{j=1}^{n} a_{ij}$$

□

2.1. A Collatz-Wielandt formula másik alakja

A tételben szereplő Collatz-Wielandt formulával úgy kaphatjuk meg a Perron–Frobenius sajátértéket, hogy először nézzük az Ax és x vektorok elem-hányadosainak a minimumát, majd megkeressük, hogy ez milyen x vektor esetén lesz a legnagyobb, és ez a maximum lesz r. Ezt úgy is kiszámolhatjuk, hogy először vesszük a hányadosok maximumát, és utána nézzük ezeknek a minimumát, azaz, ha A egy pozitív mátrix, akkor $r = \min_{x \in \mathcal{P}} g(x)$, ahol:

$$g(x) = \max_{1 \leq i \leq n} \frac{[Ax]_i}{x_i} \quad \text{és} \quad \mathcal{P} = \{x \mid x > 0\}$$

Ezt az eredeti formulához hasonlóan tudjuk belátni. Legyenek tehát az A mátrix r-hez tartozó jobb és baloldali sajátvektorai p és q^T. Legyen $x \in \mathcal{P}$ tetszőleges és $\xi = g(x)$. Ekkor $0 < Ax \leq \xi x$, ezt szorozzuk be balról a nemnegatív q^T-tal:

$$q^T Ax \leq q^T \xi x$$

Mivel q^T az A mátrix baloldali sajátvektora, így az egyenlőtlenség bal oldala átírható $rq^T x$-re:

$$rq^T x \leq q^T \xi x = \xi q^T x$$

Felhasználva, hogy $q^T x > 0$, ebből az következik, hogy $r \leq \xi$, tehát $r \leq g(x)$ minden $x \in \mathcal{P}$ esetén. Mivel $p \in \mathcal{P}$ az A-hoz tartozó Perron vektor, így minden i index esetén $[Ap]_i / p_i = r$, tehát a maximumuk is egyenlő r-rel, azaz $g(p) = r$, így:

$$r = \min_{x \in \mathcal{P}} g(x)$$

Vegyük észre, hogy a tételben szereplő Collatz-Wielandt formulában $\mathcal{N} = \{x \mid x \geq 0 \text{ és } x \neq 0\}$ szerepel, míg ebben a másik alakban a $\mathcal{P} = \{x \mid x > 0\}$ halmazt alkalmazzuk. Most nézzük meg, hogy mi történne, ha az eredeti $\mathcal{N}$ halmazon vennénk a Collatz-Wielandt formulának ezt a másik alakját. Definiáljuk $g(x)$-et a következőképpen:

$$g(x) = \max_{\substack{1 \leq i \leq n \\ x_i \neq 0}} \frac{[Ax]_i}{x_i}$$

Most mutatunk példát egy olyan A mátrixra, amely pozitív, de $\rho(A) \neq \min\limits_{x \in \mathcal{N}} g(x)$.

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$$

Ennek a mátrixnak a két sajátértéke a 0 és az 5, tehát $\rho(A) = 5$. Nézzük az alábbi vektorokat:

$$x = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad Ax = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

Ekkor $x \in \mathcal{N}$, de $g(x) = 1$, tehát:

$$\min_{x \in \mathcal{N}} g(x) \leq 1 < 5 = \rho(A)$$

$$\Downarrow$$

$$\min_{x \in \mathcal{N}} g(x) < \rho(A)$$

2.2. Jobb és bal oldali Perron vektor

Tudjuk, hogy $A > 0$ ekvivalens azzal, hogy $A^T > 0$, továbbá ha $\rho(A) = r$, akkor $\rho(A^T) = r$. Így ha A pozitív, és az r-hez tartozó Perron vektora p, akkor létezik egy hozzá tartozó q Perron vektora az A^T mátrixnak, amit a következőképpen kapunk meg. Legyen az A^T mátrixnak az r-hez tartozó Perron vektora $\widetilde{p}$, azaz:

$$A^T \widetilde{p} = r \widetilde{p}$$

Transzponáljuk ezt az egyenletet:

$$\widetilde{p}^T A = r \widetilde{p}^T$$

Ekkor tehát $\widetilde{p}^T$ egy A-hoz tartozó bal oldali sajátvektor, amely a Perron–Frobenius sajátértékhez tartozik. Normáljuk ezt a $\widetilde{p}$ vektort úgy, hogy a normált vektor $q = \alpha\widetilde{p}$ legyen és teljesüljön az, hogy:

$$q^T p = \sum_{i=1}^{n} q_i p_i = 1$$

Ekkor tehát $q^T A = \alpha\widetilde{p}^T A = \alpha r\widetilde{p}^T = r(\alpha\widetilde{p}^T) = rq^T$, azaz q^T egy bal oldali sajátvektora A-nak, amely a Perron–Frobenius sajátértékhez tartozik. Ezt a q^T vektort nevezzük az A mátrix *bal oldali Perron vektorának*. Mivel $A^T > 0$, így a Perron tétel alapján egyértelműen létezik egy ilyen $\widetilde{p} > 0$, amiből az következik, hogy q^T is pozitív és egyértelmű.

Az alábbiakban láthatunk két fontos összefüggést az A mátrix jobb és bal oldali Perron vektoraira vonatkozóan.

2.2.1. Tétel. *Ha $A \in \mathbb{R}^{n\times n}$ egy pozitív mátrix és $\rho(A) = r$, továbbá az A-hoz tartozó jobb és bal oldali Perron vektorok p és q^T, akkor létezik egy olyan S invertálható mátrix, amelyre*

$$A = S \begin{pmatrix} r & 0 \\ 0 & B \end{pmatrix} S^{-1},$$

ahol B egy $(n-1)\times(n-1)$ méretű mátrix, S első oszlopában a p vektor szerepel (ezt jelölje $S = \begin{pmatrix} p & S_1 \end{pmatrix}$), S^{-1} első sorában pedig q^T áll, (azaz $S^{-1} = \begin{pmatrix} q^T \\ Z_1 \end{pmatrix}$).

Bizonyítás: Válasszunk ki egy $s_1, \ldots, s_{n-1}$ bázist q ortogonális kiegészítő alteréből, és legyenek ezek az S_1 mátrix oszlopai, ekkor tehát $q^T S_1 = 0$. Legyen $S = \begin{pmatrix} p & S_1 \end{pmatrix}$ és legyen z egy olyan vektor, amelyre $z = \begin{pmatrix} z_1 & c \end{pmatrix}^T$, ahol c egy $n-1$ hosszú vektor. Tegyük fel, hogy $Sz = 0$, ekkor $q^T(Sz) = q^T \cdot 0 = 0$, tehát:

$$0 = q^T(Sz) = q^T(pz_1 + S_1 c) = (q^T p)z_1 + (q^T S_1)c = 1 \cdot z_1 + 0 \cdot c = z_1$$

Tehát $z_1 = 0$, ekkor $0 = Sz = pz_1 + S_1 c = p \cdot 0 + S_1 c = S_1 c$, ami csak akkor lehetséges, ha $c = 0$, hiszen S_1 egy teljes rangú mátrix (mert így állítottuk elő). Így tehát z minden koordinátája nulla, és ezzel azt kaptuk, hogy $Sz = 0 \Longleftrightarrow z = 0$, tehát S invertálható.

Legyen $S^{-1} = \begin{pmatrix} d \\ Z_1 \end{pmatrix}$, ahol d egy n hosszúságú sorvektor és számítsuk ki az $S^{-1}S$ mátrixot:

$$I_n = S^{-1}S = \begin{pmatrix} d \\ Z_1 \end{pmatrix} \begin{pmatrix} p & S_1 \end{pmatrix} = \begin{pmatrix} dp & dS_1 \\ Z_1 p & Z_1 S_1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & I_{n-1} \end{pmatrix} \tag{2.22}$$

Itt I_n az n dimenziós egységmátrixot jelöli. Abból, hogy $dS_1 = 0$, az következik, hogy d merőleges S_1-re, azaz a q ortogonális kiegészítő alterére, tehát létezik egy olyan α skalár, hogy $d^T = \alpha q$, azaz $d = \alpha q^T$. A (2.22) egyenlet alapján $dp = 1$, így:

$$1 = dp = (\alpha q^T)p = \alpha(q^T p) = \alpha \cdot 1 = \alpha \quad \Longrightarrow \quad d = q^T$$

Továbbá a (2.22) egyenlet alapján $dS_1 = q^T S_1 = 0$ és $Z_1 p = 0$, így:

$$\begin{aligned} S^{-1}AS &= \begin{pmatrix} q^T \\ Z_1 \end{pmatrix} A \begin{pmatrix} p & S_1 \end{pmatrix} = \begin{pmatrix} q^T Ap & q^T AS_1 \\ Z_1 Ap & Z_1 AS_1 \end{pmatrix} \\ &= \begin{pmatrix} (rq^T)p & (rq^T)S_1 \\ Z_1(rp) & Z_1 AS_1 \end{pmatrix} = \begin{pmatrix} r(q^T p) & r(q^T S_1) \\ r(Z_1 p) & Z_1 AS_1 \end{pmatrix} \\ &= \begin{pmatrix} r \cdot 1 & r \cdot 0 \\ r \cdot 0 & Z_1 AS_1 \end{pmatrix} = \begin{pmatrix} r & 0 \\ 0 & Z_1 AS_1 \end{pmatrix} \end{aligned}$$

Ezek alapján ha bevezetjük a $B = Z_1 AS_1$ jelölést és beszorzunk balról S-sel, jobbról pedig S^{-1}-zel, akkor éppen a tételben szereplő alakját kapjuk A-nak.

□

2.2.2. Tétel. *Legyen $A \in \mathbb{R}^{n \times n}$ egy pozitív mátrix és $\rho(A) = r$, illetve legyenek az A-hoz tartozó jobb és bal oldali Perron vektorok p és q^T. Ekkor*

$$\lim_{k \to \infty} \left(\frac{A}{r} \right)^k = pq^T$$

és a pq^T egy olyan mátrix, amely pozitív, és $\operatorname{rang}(pq^T) = 1$.

Bizonyítás: A 2.2.1 Tétel alapján léteznek olyan $S = \begin{pmatrix} p & S_1 \end{pmatrix}$ és $S^{-1} = \begin{pmatrix} q^T \\ Z_1 \end{pmatrix}$ mátrixok, hogy:

$$S^{-1}AS = \begin{pmatrix} r & 0 \\ 0 & B \end{pmatrix} =: C$$

Ekkor az A mátrix hasonló a C mátrixhoz, és mivel A-nak r egyszeres sajátértéke és r-en kívül minden más sajátértéke abszolút értékben kisebb, mint $r = \rho(A)$, így C-nek is egyszeres sajátértéke r, és így $\rho(B) < r$, tehát:

$$\frac{\rho(B)}{r} = \rho(\frac{B}{r}) < 1$$

Így a (2.12) állítás alapján:

$$\lim_{k\to\infty}\left(\frac{B}{r}\right)^k = 0$$

Ebből pedig következik a tétel első állítása, ugyanis:

$$\begin{aligned}\lim_{k\to\infty}\left(\frac{A}{r}\right)^k &= \lim_{k\to\infty}\left(S\left(\frac{C}{r}\right)S^{-1}\right)^k = \lim_{k\to\infty} S\left(\frac{C}{r}\right)^k S^{-1} \\ &= \lim_{k\to\infty} S\begin{pmatrix}\frac{r}{r} & 0 \\ 0 & \frac{B}{r}\end{pmatrix}^k S^{-1} = \lim_{k\to\infty} S\begin{pmatrix}1^k & 0 \\ 0 & \left(\frac{B}{r}\right)^k\end{pmatrix} S^{-1} \\ &= S\begin{pmatrix}1 & 0 \\ 0 & 0_{n-1}\end{pmatrix}S^{-1} = \begin{pmatrix}p & S_1\end{pmatrix}\begin{pmatrix}1 & 0 \\ 0 & 0_{n-1}\end{pmatrix}\begin{pmatrix}q^T \\ Z_1\end{pmatrix} = pq^T\end{aligned}$$

Itt 0_{n-1} egy olyan $(n-1)\times(n-1)$-es négyzetes mátrixot jelöl, amelynek minden eleme nulla. Mivel p és q is pozitívak, így szükségképpen pq^T is pozitív. Az pedig, hogy $\mathrm{rang}(pq^T) = 1$ könnyen látszik, hiszen pq^T hasonló egy olyan mátrixhoz, amelynek az első eleme 1 és minden más eleme nulla, így csak egyetlen lineárisan független sort vagy oszlopot tudunk kiválasztani belőle (az elsőt), tehát a rangja 1.

□

2.3. A pozitivitás szükségessége

Most megmutatjuk, hogy mi történne, ha a Perron tételben szereplő feltételt elhagynánk, miszerint A pozitív. Ehhez tehát olyan A négyzetes mátrixokat láthatunk majd, amelyekre teljesül, hogy $A \geq 0$, de $A \not> 0$ (azaz A-nak van legalább egy nulleleme) és $r = \rho(A) \in \sigma(A)$. Vegyük először az alábbi mátrixot:

$$A_1 = \begin{pmatrix}0 & 1 \\ 0 & 0\end{pmatrix}$$

Ekkor $r = \rho(A_1) = 0$, mivel A_1 egyetlen sajátértéke a 0, így nem teljesülne a tétel (2.1) pontja, azaz, hogy $r > 0$.

$$A_2 = \begin{pmatrix}3 & 2 \\ 0 & 3\end{pmatrix}$$

Ennek a mátrixnak a 3 kétszeres sajátértéke, tehát ha elhagynánk azt a feltételt, hogy A legyen pozitív, akkor nem teljesülne a tétel (2.4) pontja, miszerint $r = \rho(A_2) = 3$ csak egyszeres sajátértéke A-nak. Sőt, módosítsuk a mátrixot úgy, hogy a jobb felső eleme is nulla legyen:

$$\widetilde{A_2} = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix}$$

Így $\widetilde{A_2}$-nak szintén kétszeres sajátértéke a 3, és tartozik hozzá két lineárisan független sajátvektor:

$$x = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad y = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Így tehát az $r = \rho(\widetilde{A_2}) = 3$-hoz tartozó lineárisan független sajátvektorok száma sem egy, ami ellentmond Perron tételének.

$$A_3 = \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}$$

Ennek a mátrixnak a sajátértékei a $\lambda_1 = 1$ és $\lambda_2 = 0$. Tehát $r = \rho(A_3) = 1$, amelyhez az alábbi sajátvektor tartozik:

$$x = \begin{pmatrix} a \\ 0 \end{pmatrix} \quad a \in \mathbb{R}\backslash\{0\}$$

Ekkor akárhogyan is választjuk meg a-t, x biztosan nem lesz pozitív, tehát nem teljesülne a tétel (2.3) pontja, azaz, hogy létezik A_3-nak egy x pozitív sajátvektora, amelyre $A_3x = rx$.

$$A_4 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Ennek a mátrixnak a spektrálsugara $r = \rho(A) = 1$, sajátértékei pedig 1 és -1. Mindkét sajátértékének az abszolút értéke $1 = \rho(A)$, tehát nem teljesül a tétel (2.6) pontja, miszerint r-en kívül nincs A-nak másik olyan sajátértéke, amelynek az abszolút értéke egyenlő r-rel.

3. fejezet

Általánosítás nemnegatív mátrixokra

Láthattuk, hogy Perron tételét nem tudjuk közvetlenül alkalmazni a nemnegatív mátrixokra. Frobeniusnak azonban volt egy fontos észrevétele ezekkel a mátrixokkal kapcsolatban: nem egyszerűen az a probléma, hogy egy adott mátrixnak nullelemei is vannak, hanem az, hogy hol helyezkednek el ezek a nullelemek. Nézzük például az alábbi két mátrixot:

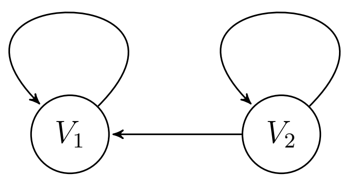
$$A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \qquad \widetilde{A} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$$

Itt az A mátrixra nem teljesülnek a Perron tétel (2.3) és (2.4) pontjai, de az $\widetilde{A}$ mátrixra igen. Továbbá A reducibilis, ugyanis:

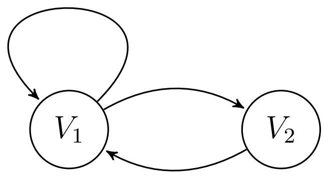
$$P^T A P = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \text{ ahol } P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Most vizsgáljuk meg a reducibilitást illetve az irreducibilitást a mátrixokhoz tartozó gráfokból kiindulva.

A 3.1 ábrán láthatjuk, hogy az A-hoz tartozó gráf nem erősen összefüggő, hiszen nem tudunk irányított úton eljutni V_1-ből V_2-be. Így a 1.0.14 Tétel alapján láthatjuk, hogy A valóban reducibilis.



3.1. ábra. $\mathcal{G}(A)$



3.2. ábra. $\mathcal{G}(\widetilde{A})$

Ugyanakkor a 3.2 ábrán jól látható, hogy az $\widetilde{A}$-hoz tartozó gráfban már el tudunk jutni a V_1 csúcsból a V_2 csúcsba irányított úton, így bármely csúcsból bármely másik csúcsba vezet irányított út, azaz $\mathcal{G}(\widetilde{A})$ erősen összefüggő, tehát a 1.0.14 Tétel alapján $\widetilde{A}$ irreducibilis.

Ezek alapján azt remélhetjük, hogy a Perron tétel állításai kiterjeszthetők nemnegatív mátrixokra is, amennyiben a mátrix irreducibilis. Ennek igazolásához szükségünk lesz az alábbi lemmára, amely megmutatja, hogy hogyan alakíthatunk át egy nemnegatív mátrixot pozitívvá.

3.0.1. Lemma. *Ha $A \in \mathbb{R}^{n\times n}$ egy nemnegatív mátrix és I az n-dimenziós egységmátrix, akkor $(I+A)^{n-1} > 0$.*

Bizonyítás: Bár a mátrixszorzás nem kommutatív művelet, de vegyük észre, hogy mivel $IA = AI = A$, így az I és A mátrixok esetén teljesül a binomiális tétel, azaz:

$$(I+A)^{n-1} = \sum_{k=0}^{n-1} \binom{n-1}{k} A^k$$

Elég tehát azt megmutatnunk, hogy az egyenlőség jobb oldalán szereplő mátrix minden eleme pozitív. Jelölje az A^k mátrix i-edik sorának j-edik elemét $a_{ij}^{(k)}$, amely a következőképpen néz ki:

$$a_{ij}^{(k)} = \sum_{1\le h_1,\ldots,h_{k-1}\le n} a_{ih_1}a_{h_1h_2}\cdots a_{h_{k-1}j}$$

Vegyük észre, hogy $a_{ij}^{(k)} > 0$ akkor és csak akkor, ha létezik az indexnek egy olyan $h_1,\ldots,h_{k-1}$ sorozata, hogy:

$$a_{ih_1} > 0 \quad \text{és} \quad a_{h_1h_2} > 0 \quad \text{és} \quad \ldots \quad \text{és} \quad a_{h_{k-1}j} > 0$$

Ez azt jelenti, hogy létezik az éleknek egy olyan $V_i \to V_{h_1} \to \cdots \to V_{h_{k-1}} \to V_j$ k hosszú sorozata $\mathcal{G}(A)$-ban, amely a V_i csúcsból a V_h csúcsba vezet akkor és csak akkor, ha $a_{ij}^{(k)} > 0$. Mivel A irreducibilis, így $\mathcal{G}(A)$ erősen összefüggő, tehát bármely (V_i, V_j) csúcspárra létezik egy k élhosszúságú út, amely V_i-ből V_j-be vezet ($k < n$). Ez azt jelenti, hogy minden (i,j) esetén létezik

$0 \leq k \leq n-1$ úgy, hogy $a_{ij}^{(k)} > 0$. Ezek alapján:

$$\left[(I+A)^{n-1}\right]_{ij} = \left[\sum_{k=0}^{n-1} \binom{n-1}{k} A^k\right]_{ij} = \sum_{k=0}^{n-1} \binom{n-1}{k} a_{ij}^{(k)} > 0$$

□

A Perron tétel tehát kiterjeszthető nemnegatív mátrixokra is abban az esetben, ha feltesszük, hogy a mátrix irreducibilis.

3.0.2. Tétel. (Perron–Frobenius) *Ha $A \in \mathbb{R}^{n \times n}$ egy nemnegatív, irreducibilis mátrix, és $r = \rho(A)$, akkor a következő állítások teljesülnek:*

- *$r \in \sigma(A)$.* (3.1)
- *r egyszeres sajátértéke A-nak.* (3.2)
- *Létezik egy $x > 0$ sajátvektora A-nak, amelyre $Ax = rx$.* (3.3)
- *$r > 0$, azaz A spektrálsugara pozitív.* (3.4)
- *A-nak egyértelműen létezik egy olyan $p > 0$ sajátvektora, amelyre $Ap = rp$ és $\sum_{i=1}^{n} p_i = 1$, továbbá A-nak p pozitív skalárszorosain kívül nem létezik más nemnegatív saját-vektora.* (3.5)
- ***Collatz-Wielandt formula**: $r = \max_{x \in \mathcal{N}} f(x)$,*

  ahol $f(x) = \min_{\substack{1 \leq i \leq n \\ x_i \neq 0}} \frac{[Ax]_i}{x_i}$ *és* $\mathcal{N} = \{x \mid x \geq 0 \text{ és } x \neq 0\}$ (3.6)

Bizonyítás:

- $r \in \sigma(A)$.

  Azt, hogy A spektrálsugara egyúttal sajátértéke is A-nak, a következő állítással együtt fogjuk belátni:

  $$A\text{-nak létezik olyan } z \geq 0,\ z \neq 0 \text{ sajátvektora, amelyre } Az = rz. \tag{3.7}$$

  Legyen $A_k = A + (1/k)E > 0$, ahol E a csupa 1-esekből álló mátrix. Jelölje az A_k-hoz tartozó Perron–Frobenius sajátértéket $r_k > 0$, a hozzá tartozó Perron vektort pedig $p_k > 0$. Vegyük észre, hogy $\{p_k\}_{k=1}^{\infty}$ egy korlátos halmaz, ugyanis minden i esetén p_i elemeinek az összege 1, ami nem teljesülne, ha a halmaz nem lenne korlátos.

A Bolzano-Weierstrass tétel szerint minden korlátos $\mathbb{R}$-ben haladó sorozatnak van konvergens részsorozata, amiből az következik, hogy minden korlátos $\mathbb{R}^n$-ben haladó sorozatnak is van konvergens részsorozata. Ugyanis ha vesszük először a vektorok első koordinátáját, az egy korlátos $\mathbb{R}$-ben haladó sorozat, tehát van konvergens részsorozata. Ezután ha vesszük ezt a konvergens részsorozatot, akkor a vektorok első koordinátáiban minden további részsorozat is konvergens marad, és a maradék $n-1$ koordinátában nézzük a további konvergenciát: ennek az első (tehát a vektornak eredetileg a második) koordinátája is egy $\mathbb{R}$-ben haladó sorozatot ad, tehát létezik konvergens részsorozata. Így tehát találunk egy olyan részsorozatot, amely az első két koordinátájában konvergens. Iterálva találunk olyan részsorozatot, amely már minden koordinátában konvergens, ez pedig épp a vektorok konvergenciáját jelenti.

Így tehát $\{p_k\}_{k=1}^{\infty}$ is tartalmaz egy konvergens részsorozatot, azaz létezik egy olyan z vektor és léteznek olyan k_i indexek, hogy:

$$\{p_{k_i}\}_{i=1}^{\infty} \longrightarrow z$$

Mivel bármely i esetén $p_{k_i} > 0$, így $z \geq 0$, és mivel $\|p_{k_i}\|_1 = 1$, így $z \neq 0$. A (2.21) állítás alapján mivel $A_1 > A_2 > \ldots > A \geq 0$, így $r_1 \geq r_2 \geq \ldots \geq r$, azaz $\{r_k\}_{k=1}^{\infty}$ egy korlátos monoton csökkenő sorozat, amelynek alsó korlátja r. Ekkor létezik olyan $\widetilde{r} \geq r$, amelyre:

$$\lim_{k\to\infty} r_k = \widetilde{r}$$

Ekkor az r_k-knak tetszőleges részsorozata is tart $\widetilde{r}$-hoz, tehát speciálisan a k_i indexekből álló részsorozat is:

$$\lim_{i\to\infty} r_{k_i} = \widetilde{r} \geq r$$

Ugyanakkor vegyük figyelembe a következőeket:

$$\lim_{k\to\infty} A_k = A \quad \Longrightarrow \quad \lim_{i\to\infty} A_{k_i} = A$$

Ezek alapján, ha felhasználjuk, hogy véges határértékek esetén a szorzatok határértéke egyenlő a határértékek szorzatával, azt kapjuk, hogy:

$$Az = \lim_{i\to\infty} A_{k_i} p_{k_i} = \lim_{i\to\infty} r_{k_i} p_{k_i} = \widetilde{r} z \quad \Longrightarrow \quad \widetilde{r} \in \sigma(A) \quad \Longrightarrow \quad \widetilde{r} \leq r$$

Így tehát $\widetilde{r} = r$ és $Az = rz$ úgy, hogy $z \geq 0$ és $z \neq 0$. Ezzel tehát beláttuk a (3.7) állítást és a tétel első állítását is, azaz, hogy $r \in \sigma(A)$.

- *r egyszeres sajátértéke A-nak.*
  Ha egy C mátrixra $Cx = \lambda x$ akkor minden $n \in \mathbb{N}$ esetén $C^n x = \lambda^n x$. Ez teljes indukcióval könnyen belátható, hiszen az $n = 1$ eset triviális és ha feltesszük, hogy $n-1$-re már igaz az állítás, akkor:

  $$C^n x = (CC^{n-1}x) = C(C^{n-1}x) = C(\lambda^{n-1}x) = \lambda^{n-1}Cx = \lambda^{n-1}\lambda x = \lambda^n x$$

  Legyen $B = (I+A)^{n-1} > 0$, ekkor az előzőek alapján λ pontosan akkor (m-szeres) sajátértéke A-nak, ha $(1+\lambda)^{n-1}$ is (m-szeres) sajátértéke B-nek. Sőt, ha x a λ-hoz tartozó sajátvektor, az ekvivalens azzal, hogy az $(1+\lambda)^{n-1}$-hez tartozó sajátvektor szintén x. Így tehát ha $\mu = \rho(B)$, akkor:

  $$\mu = \max_{\lambda \in \sigma(A)} \left|(1+\lambda)^{n-1}\right| = \max_{\lambda \in \sigma(A)} |(1+\lambda)|^{n-1} = \left\{\max_{\lambda \in \sigma(A)} |(1+\lambda)|\right\}^{n-1} \tag{3.8}$$

  A háromszög egyenlőtlenségből tudjuk, hogy $|1+\lambda| \leq 1 + |\lambda|$, amiből az következik, hogy:

  $$\max_{\lambda \in \sigma(A)} |(1+\lambda)| \leq \max_{\lambda \in \sigma(A)} (1+|\lambda|) = 1 + \max_{\lambda \in \sigma(A)} |\lambda| = 1 + r$$

  Ugyanakkor mivel $\rho(A) = r$ és $r \in \sigma(A)$, így a fenti egyenlőtlenségben $\lambda = r$ választással egyenlőség teljesül, amit ha felhasználunk a (3.8) egyenletnél, azt kapjuk, hogy:

  $$\mu = \cdots = \left\{\max_{\lambda \in \sigma(A)} |(1+\lambda)|\right\}^{n-1} = (1+r)^{n-1}$$

  Mivel B pozitív, így alkalmazhatjuk rá a Perron tételt, miszerint μ egyszeres sajátértéke B-nek, amiből pedig az következik, hogy r egyszeres sajátértéke A-nak.

- *Létezik egy $x > 0$ sajátvektor, amelyre $Ax = rx$.*
  Ahhoz, hogy ezt belássuk, felhasználjuk a (3.7) állítást, azaz, hogy A-nak létezik olyan x nemnegatív sajátvektora, amely az r sajátértékhez tartozik. Mivel $Ax = rx \iff Bx = (I+A)^{n-1}x = (1+r)^{n-1}x = \mu x$, így ez az x sajátvektor tartozik $\mu = \rho(B)$-hez is. Viszont mivel B pozitív, így a Perron tétel alapján biztos, hogy x a B-hez tartozó Perron vektornak pozitív többszöröse, amiből az következik, hogy x pozitív.

- *$r > 0$, azaz A spektrálsugara pozitív.*
  Mivel $A \geq 0$ és $x > 0$, így $Ax > 0$. Ha indirekt feltesszük, hogy $r = 0$, akkor $Ax = rx = 0$, ami ellentmondás, hiszen $Ax > 0$. Így tehát biztos, hogy $r > 0$.

- A tétel (3.5) pontja a Perron vektorról szól, amely ugyanazt mondja ki nemnegatív irreducibilis mátrixokra, mint a Perron tétel a pozitív mátrixokra. A Perron tétel bizonyításánál a Perron vektorról szóló állítások igazolásához nem használtuk ki, hogy A pozitív, így azok a bizonyítások itt is érvényesek.

- $r = \max\limits_{x \in \mathcal{N}} f(x)$
  Ahogy a tétel (3.1) pontjának bizonyításánál is, legyen $A_k = A + (\frac{1}{k}E)$, ahol E a csupa 1-esekből álló mátrix, ekkor bármely k esetén A_k pozitív. Legyen minden k esetén az A_k-hoz tartozó Perron-Frobenius sajátérték r_k, és a hozzá tartozó bal oldali Perron vektor $q_k^T > 0$. A Perron tétel alapján minden $x \in \mathcal{N}$ és $k > 0$ esetén $q_k^T x > 0$, és mivel $0 \leq f(x)x \leq Ax$ (ezt láthattuk a Perron tétel esetén a Collatz-Wielandt formula bizonyításánál), így:

  $$0 \leq f(x)x \leq Ax \leq A_k x \quad \Longrightarrow \quad f(x)x \leq A_k x$$

  Szorozzuk be az egyenlőtlenséget balról q_k^T-tal:

  $$q_k^T f(x)x = f(x)q_k^T x \leq q_k^T A_k x = r_k q^T x$$

  Így tehát azt kaptuk, hogy $f(x) \leq r_k$. A tétel (3.1) pontjának bizonyításánál már beláttuk, hogy $r_k \longrightarrow r$ amint $k \longrightarrow \infty$, tehát $f(x) \leq r$. Legyen $z \in \mathcal{N}$ egy olyan vektor, amelyre $Az = rz$. Ilyen vektor biztos hogy létezik a (3.7) állítás alapján. Ekkor $f(z) = r$, ugyanis:

  $$f(z) = \min_{\substack{1 \leq i \leq n \\ z_i \neq 0}} \frac{[Az]_i}{z_i} = \min_{\substack{1 \leq i \leq n \\ z_i \neq 0}} \frac{[rz]_i}{z_i} = \min_{\substack{1 \leq i \leq n \\ z_i \neq 0}} \frac{rz_i}{z_i} = r$$

  Így tehát $f(z) = r$ és $z \in \mathcal{N}$, amiből az következik, hogy $\max\limits_{x \in \mathcal{N}} f(x) = r$.

$\square$

4. fejezet

Alkalmazások

A Perron–Frobenius tételt számos területen lehet alkalmazni, ezek közül ebben a fejezetben megismerkedhetünk néhány algebrai (primitív mátrixok), illetve sztochasztikai (Markov-láncok), egy közgazdaságtani (Leontief-féle input-output modell), illetve egy demográfiai (Leslie-modell egy populáció méretének leírására) alkalmazással.

A tételt rangsorok készítésére is kiválóan lehet alkalmazni, például labdajátékoknál a csapatok páronkénti összehasonlításával kialakuló rangsorhoz, vagy akár egy internetes kereséskor a találatok sorrendjének felállításához. Utóbbira a következő fejezetben láthatunk egy példát, amelyben megismerhetjük a Google kereső működését és kapcsolatát a Perron–Frobenius tétellel.

4.1. Primitív mátrixok

A primitív mátrixokat a Perron–Frobenius tételben szereplő egyik fontos tulajdonság alapján definiáljuk.

4.1.1. Definíció. Legyen $A \in \mathbb{R}^{n \times n}$ egy nemnegatív, irreducibilis mátrix. Ha $r = \rho(A)$ sajátértéke A-nak, de ezen kívül A-nak nincs másik olyan sajátértéke, amelynek az abszolút értéke egyenlő r-rel, akkor azt mondjuk, hogy A *primitív.*

A következő tétel ugyanazt mondja ki primitív mátrixokra, mint amit a 2.2.2 Tétel állít pozitív mátrixokra. Ennek a két tételnek a bizonyítása ugyanúgy zajlik, éppen ezért az alábbi tétel bizonyítása nem tartalmaz minden lépést, mert azokat már láthattuk a korábbi tétel bizonyítása során.

4.1.2. Tétel. *Legyen A egy primitív mátrix, $\rho(A) = r$, és legyenek az A-hoz tartozó jobb és bal oldali Perron vektorok p és q^T. Ekkor*

$$\lim_{k\to\infty} \left(\frac{A}{r}\right)^k = pq^T,$$

ahol pq^T egy pozitív mátrix és $\mathrm{rang}(pq^T) = 1$.

Bizonyítás: Ha A primitív és $\rho(A) = r$, akkor r egyszeres sajátértéke A-nak, és a hozzá tartozó p és q^T sajátvektorok pozitívak úgy, hogy $p^Tq = 1$. Járjunk el ugyanúgy, mint a 2.2.2 Tétel bizonyításánál, és írjuk fel A-t a következő alakban:

$$A = S \begin{pmatrix} r & 0 \\ 0 & B \end{pmatrix} S^{-1}$$

Ekkor a korábbi bizonyításhoz hasonlóan B-nek minden sajátértéke abszolút értékben kisebb, mint r, tehát itt is elmondható, hogy:

$$\lim_{k\to\infty} \left(\frac{B}{r}\right)^k = 0$$

Ebből pedig (ahogy korábban is láthattuk) már következik a tétel állítása. Az, hogy $pq^T > 0$ és rangja 1, ugyanúgy látható be, mint a 2.2.2 Tétel bizonyításánál. □

Frobenius nem csak Perron tételét általánosította a nemnegatív mátrixokra, hanem azt is megmutatta, hogy egy nemnegatív mátrix hatványai meghatározzák, hogy a mátrix primitív-e. Erről szól az alábbi tétel.

4.1.3. Tétel. *Egy nemnegatív A mátrix akkor és csak akkor primitív, ha létezik olyan $m > 0$, amelyre $A^m > 0$.*

Bizonyítás: Először tegyük fel, hogy A primitív és $r = \rho(A)$. Ekkor a 4.1.2 Tétel alapján:

$$\lim_{k\to\infty} \left(\frac{A}{r}\right)^k > 0$$

Így biztos, hogy létezik olyan $m > 0$, amelyre $\left(\dfrac{A}{r}\right)^m > 0$, amiből pedig következik, hogy $A^m > 0$.

Most nézzük a másik irány bizonyítását, tehát tegyük fel, hogy valamely $m > 0$-ra $A^m > 0$. Ekkor A^m irreducibilis, ugyanis ha reducibilis lenne, az azt jelentené, hogy létezik olyan P permutációs mátrix, amelyre:

$$A = P \begin{pmatrix} B & C \\ 0 & D \end{pmatrix} P^T \quad\Longrightarrow\quad A^m = P \begin{pmatrix} B^m & \widetilde{C} \\ 0 & D^m \end{pmatrix} P^T$$

Ez viszont nem lehetséges, hiszen A^m-nek minden eleme nagyobb, mint nulla. Tegyük fel, hogy A-nak van h darab olyan sajátértéke $(\lambda_1, \dots, \lambda_h)$ amelyekre teljesül, hogy:

$$r = |\lambda_1| = \cdots = |\lambda_h| > |\lambda_{h+1}| > \cdots > |\lambda_n|$$

Mivel $\lambda \in \sigma(A)$-ból következik, hogy $\lambda^m \in \sigma(A^m)$, így ha λ l-szeres sajátértéke A-nak, akkor λ^m is l-szeres sajátértéke A^m-nek. (Ez A Jordan-normálalakjából látszik.) Ezek alapján minden $1 \leq i \leq h$ esetén λ_i és λ_i^m multiplicitása megegyezik, továbbá:

$$r^m = \sigma(A^m) = |\lambda_1^m| = \cdots = |\lambda_h^m|$$

Perron tétele alapján $\rho(A^m) = r^m$ az egyetlen olyan sajátértéke A^m-nek, amelynek az abszolút értéke r^m és ennek egyszeres a multiplicitása, tehát $h = 1$ és $r^m = \lambda_1^m$. λ_1-hez tartozik A-nak egy x sajátvektora, amely ekkor A^m-nek az r^m-hez tartozó sajátvektora. A Perron tétel alapján feltehető, hogy ez pozitív, így $0 \leq Ax = \lambda_1 x$, ami csak akkor lehet, ha λ_1 nemnegatív, és ekkor $r = \lambda_1$. $\square$

4.2. Leontief-féle input-output modell

Tegyük fel, hogy van egy zárt gazdasági rendszer, amely n különböző iparágból áll úgy, hogy minden iparág pontosan egy terméket állít elő, és mindegyik terméket pontosan egy iparág állítja elő, tehát a termékek és a termelési folyamatok közötti kapcsolat kölcsönösen egyértelmű. Nevezzük J-*egységnek* azt, amit a J iparág gyárt és 100 Ft-ért adja el. Például ha az A ipar autókat gyárt, a B pedig ceruzákat, akkor egy A-*egységnek* megfelel az autó egy töredékének, de egy B-*egység* egy csomag ceruzának felel meg. Vezessük be a következő jelöléseket:

$$\begin{aligned} 0 \leq s_j &:= \text{ a } J \text{ iparág által évente gyártott } J\text{-}\textit{egységek} \text{ száma} \\ 0 \leq a_{ij} &:= \text{ egy } J\text{-}\textit{egység} \text{ előállításához szükséges } I\text{-}\textit{egységek} \text{ száma} \end{aligned}$$

Így tehát:

$$\begin{aligned} a_{ij}s_j &= \text{a } J \text{ iparág által évente felhasznált } I\text{-}\textit{egységek} \text{ száma} \\ \sum_{j=1}^{n} a_{ij}s_j &= \text{az összes iparág által évente felhasznált } I\text{-}\textit{egységek} \text{ száma} \\ d_i &:= \text{a nem ipari célokra évente fennmaradó } I\text{-}\textit{egységek} \text{ száma} \end{aligned}$$

Ekkor tehát:

$$d_i = s_i - \sum_{j=1}^{n} a_{ij} s_j \tag{4.1}$$

Jelölje $s = \left(s_1 \dots s_n\right)^T$ a *készlet vektort*, azaz, hogy az egyes termékekből évente mennyit állítanak elő, és jelölje az iparok által nem felhasznált (kereskedelmi célokra megmaradó) termékek mennyiségét a $d = \left(d_1 \dots d_n\right)^T$ *kereskedelmi vektor*. Azt szeretnénk kiszámolni, hogy egy adott kereskedelmi vektorhoz milyen készlet vektorra van szükségünk. Tehát a feladatunk az, hogy egy adott $d \geq 0$ esetén kiszámítsuk az $s \geq 0$ vektort.

Első ránézésre a feladat egyszerűnek tűnhet, hiszen a (4.1) egyenletet felírhatjuk $(I - A)s = d$ alakban, és így ha $(I - A)$ invertálható, akkor $s = (I - A)^{-1}d$. Ezzel azonban az a probléma, hogy annak ellenére, hogy A egy nemnegatív mátrix, a kapott s vektornak lehetnek negatív elemei is, pedig mi olyan készlet vektort szeretnénk kiszámolni, amely nemnegatív.

Ebben a modellben fontos azonban, hogy az egyes iparágak szoros összefüggésben állnak egymással, pontosabban: minden egyes termékre – közvetve vagy közvetlenül ugyan – de mindenképp szükségünk van ahhoz, hogy a gazdasági rendszer összes termékét előállítsuk. Ez matematikailag azt jelenti, hogy ha az a_{ij} számokat felírjuk egy közös A mátrixba, akkor a hozzá tartozó $\mathcal{G}(A)$ gráf erősen összefüggő lesz, és így a 1.0.14 Tétel alapján az A mátrix irreducibilis.

Ha az A mátrix j-edik oszlopának elemeit összegezzük, akkor megkapjuk azt a mennyiséget, amely megmutatja, hogy összesen hány egységre van szükségünk egy *J-egység* előállításához. Azonban vegyük figyelembe, hogy az egyes egységek 100 Ft-ba kerülnek, tehát:

$$\begin{aligned} c_{ij} = \sum_{i=1}^{n} a_{ij} &= \text{ hány egység kell egy } J\text{-}egység \text{ előállításához} \\ &= \text{ hányszor 100 Ft-ot költünk el egy } J\text{-}egység \text{ előállítására} \end{aligned}$$

Mivel az előállított *J-egységet* 100 Ft-ért tudjuk eladni, így értelemszerűen azt szeretnénk, hogy az előállítása ne kerüljön többe, mint amennyiért eladjuk. Tehát egy jó gazdasági rendszerben minden j-re $c_j \leq 1$ és létezik legalább egy olyan iparág (legyen ez az I), amelyre $c_i < 1$ (különben az egész rendszer nem termelne hasznot). Ekkor A oszlopait ki tudjuk úgy egészíteni, hogy mindegyik oszlopösszeg pontosan 1 legyen, azaz létezik egy olyan nemnegatív E mátrix, hogy $E \neq 0$ és az $A + E$ mátrix minden oszlopának összege egyenlő 1-gyel. Ekkor ha e egy olyan vektor, amelynek mindene eleme 1, akkor a következő teljesül:

$$e^T(A+E) = e^T$$

Tegyük fel, hogy $\rho(A) \geq 1$, és legyen az A-hoz tartozó jobb oldali Perron vektor p, ekkor az $e^T p$ szorzat a p vektor elemeinek az összegét jelenti, ami a Perron–Frobenius tétel alapján 1-gyel egyenlő, így az előző egyenlet alapján felírhatjuk a következő egyenletet:

$$\begin{aligned} 1 = e^T p = e^T(A+E)p = e^T(Ap) + e^T(Ep) = e^T(\rho(A)p) + e^T(Ep) \\ = \rho(A)(e^T p) + e^T(Ep) = \rho(A) + e^T(Ep) \end{aligned} \tag{4.2}$$

Mivel $E \geq 0$, $E \neq 0$ és $p > 0$, így $Ep > 0$, tehát az Ep vektor elemeinek az összege is pozitív, azaz $e^T(Ep) > 0$, amit ha felhasználunk a (4.2) egyenletben, akkor a következőt kapjuk:

$$1 = \ldots = \rho(A) + e^T(Ep) \geq 1 + e^T(Ep) > 1$$

Ez pedig ellentmondás, tehát nem igaz, hogy $\rho(A) \geq 1$, tehát biztos, hogy $\rho(A) < 1$. Most pedig bizonyítás nélkül fel fogjuk használni a következő tételt.

4.2.1. Tétel. *Ha $A \in \mathbb{C}^{n \times n}$, akkor a következő állítások ekvivalensek:*

- $\sum\limits_{k=0}^{\infty} A^k$ *konvergens*
- $\rho(A) < 1$
- $\lim\limits_{k \to \infty} A^k = 0$

Ekkor létezik az $(I-A)^{-1}$ mátrix és $\sum\limits_{k=0}^{\infty} A^k = (I-A)^{-1}$.

Emlékezzünk vissza, hogy a 3.0.1 Lemma bizonyításakor beláttuk, hogy ha az A^k mátrix i-edik sorának j-edik elemét $a_{ij}^{(k)}$ jelöli, akkor ha A irreducibilis, biztos, hogy minden i és j indexhez létezik olyan $1 \leq k \leq n-1$, amelyre $a_{ij}^{(k)} > 0$, ebből pedig az következik, hogy $\sum_k A^k > 0$, tehát $(I-A)^{-1}$ egy pozitív mátrix.

Ezek alapján pedig minden $d \geq 0$ kereskedelmi vektorhoz egyértelműen létezik egy $s > 0$ készlet vektor, amelyet az $s = (I-A)^{-1}d$ egyenlettel tudunk kiszámolni, így tehát megszűnt az eredetileg felvetett probléma.

Van még egy érdekes következménye is annak, hogy az $(I-A)^{-1}$ mátrix és az s vektor is pozitív: ha a d vektornak egyetlen elemét megnöveljük, az azt eredményezi, hogy s-nek minden eleme növekedni fog. Ez pedig azt jelenti, hogy ha akár csak egyetlen iparág egyetlen termékéből többet szeretnénk a kereskedelemben forgalmazni, az hatással lesz az egész gazdaságra, és szükségképpen minden iparág termékéből többet kell gyártani.

4.3. Leslie-modell

Osszuk fel a népesség női populációját n darab korosztályra (jelölje ezeket $G_1, \dots, G_n$) úgy, hogy az egyes korosztályok 10 éves időintervallumokat alkossanak, tehát:

$$\begin{aligned} G_1 = &\ \text{a 10 év alatti lányok csoportja} \\ G_2 = &\ \text{a 10-19 év közötti lányok csoportja} \\ G_3 = &\ \text{a 20-29 év közötti nők csoportja} \\ \text{stb.} & \end{aligned}$$

Tekintsünk az időben diszkrét pontokat, legyenek ezek $t = 0, 1, 2, \dots$ évtizedek. (A modellben az intervallumok egységénél nem fontos, hogy évtizedeket vegyünk, csak arra kell figyelni, hogy az egyes korosztályok azonos nagyságú időintervallumokat alkossanak, és hogy t is az intervallumok nagyságával azonos értékeket vegyen fel. Például ha a korosztályokat 15 éves időintervallumokra osztjuk, akkor $t = 0 \cdot 15,\ 1 \cdot 15,\ \dots$ éveket jelent.)

Jelölje b_k a születési-, s_k pedig a túlélési rátát a G_k korcsoportban, azaz:

$$\begin{aligned} b_k = &\ \text{a } G_k \text{ csoportbeli nőktől időegység alatt született lányok} \\ &\ \text{várható száma} \\ s_k = &\ \text{azon nők száma, akik a } t\text{-edik évtizedben a } G_k \text{ csoportban} \\ &\ \text{vannak, a } (t+1)\text{-edik évtizedben pedig } G_{k+1}\text{-ben} \end{aligned}$$

A túlélési ráta tehát azt méri, hogy hány olyan nő van, aki megéli azt, hogy egy adott korosztályból átkerül a következő korosztályba.

Jelölje $f_k(t)$ a G_k-beli nők számát a t-edik évtizedben, ekkor a következő egyenlőségek teljesülnek:

$$\begin{aligned} f_1(t+1) &= f_1(t)b_1 + f_2(t)b_2 + \dots + f_n(t)b_n \\ f_k(t+1) &= f_{k-1}(t)s_{k-1} \qquad k = 2, 3, \dots, n \end{aligned} \tag{4.3}$$

Jelöljük $F_k(t)$-vel azt, hogy a t-edik évtizedben a G_k csoport hány százalékát teszi ki a teljes populációnak:

$$F_k(t) = \frac{f_k(t)}{f_1(t) + \dots + f_n(t)}$$

Ekkor az $F(t) = \begin{pmatrix} F_1(t) & \dots & F_n(t) \end{pmatrix}^T$ vektor fejezi ki a *populáció életkor-eloszlását* a t-edik évtizedben, és (amennyiben létezik) $\widetilde{F} := \lim_{t\to\infty} F(t)$ pedig a *hosszútávú életkor-eloszlást.*

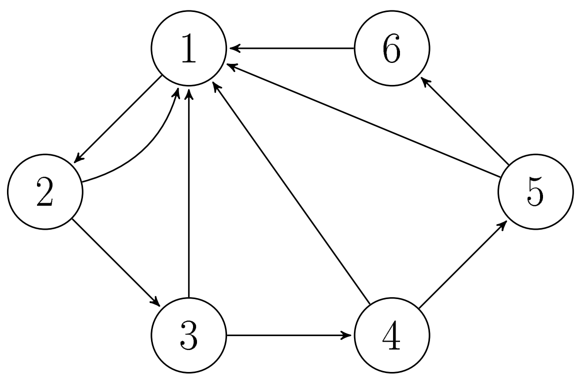
Most azt fogjuk megmutatni, hogy ha $s_1, \dots, s_{n-1}$ illetve $b_2, \dots, b_n$ is pozitívak, akkor létezik ez az $\widetilde{F}$ határérték, és ki is számítjuk az értékét. A (4.3) egyenletek egy lineáris egyenletrendszert alkotnak, amely az $f(t) = \big(f_1(t) \quad \cdots \quad f_n(t)\big)^T$ jelölés bevezetésével felírható a következő mátrixos alakban:

$$f(t+1) = Lf(t), \quad \text{ahol} \quad L = \begin{pmatrix} b_1 & b_2 & \cdots & b_{n-1} & b_n \\ s_1 & 0 & \cdots & \cdots & 0 \\ 0 & s_2 & 0 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & s_{n-1} & 0 \end{pmatrix} \in \mathbb{R}^{n \times n}$$

Ezt az L mátrixot P. H. Leslie tiszteletére *Leslie-mátrixnak* nevezzük, mivel ezt a modellt ő dolgozta ki 1945-ben.

Gondoljuk végig, hogy hogyan is néz ki ehhez az L-hez tartozó $\mathcal{G}(L)$ gráf: az első csúcsból minden további csúcsba vezet él, továbbá minden i-edik csúcsból vezet él az $i-1$-edik csúcsba. Így bármely csúcsból bármely másik csúcsba el tudunk jutni irányított úton, tehát $\mathcal{G}(L)$ erősen összefüggő, így a 1.0.14 Tétel alapján L irreducibilis. Viszont ez a $\mathcal{G}(L)$ gráf szemléletesen azt mutatja nekünk, hogy egy adott korosztályba melyik másik korosztályokból kerültek át az emberek, tehát a forrásokat ábrázolja. Éppen ezért, ha szemléletesen szeretnénk ábrázolni (azaz hogyha azt szeretnénk mutatni, hogy egy adott korcsoport mely más csoportok növekedéséhez járul hozzá), akkor először transzponálni kell a Leslie-mátrixot, és az így kapott L^T-hoz tartozó gráfot érdemes tekinteni.

Az alábbiakban láthatunk egy példát, hogy $n = 6$ esetén hogyan néz ki a $\mathcal{G}(L^T)$ gráf, amennyiben $b_1 = 0$ (ha b_1 is pozitív, akkor az első csúcsból vezet még egy hurokél is önmagába).



4.1. ábra. $n = 6$ esetén a transzponált Leslie-mátrixhoz tartozó gráf

4.3.1. Állítás. *Ha L egy $n \times n$ méretű Leslie-mátrix, akkor $L^{n+2} > 0$.*

Bizonyítás: Vegyük észre, hogy L és a $\mathcal{G}(L)$ adjacencia mátrixa (jelölje ezt G) eléggé hasonlítanak egymásra abban az értelemben, hogy l_{ij} pontosan akkor nulla, ha g_{ij} is nulla, és ahol l_{ij} pozitív, ott $g_{ij} = 1$. Így tehát $L^{n+2} > 0$ pontosan akkor, ha $G^{n+2} > 0$. Viszont egy gráf adjacencia mátrixának hatványozása megadja, hogy tetszőleges csúcsból egy másik csúcsba hány különböző, adott hosszúságú úton (vagy sétán) tudunk eljutni.

Tehát $G^{n+2} > 0$ azzal ekvivalens, hogy $\mathcal{G}(L)$-ben bármely csúcsból bármely másik csúcsba el tudunk jutni pontosan $n+2$ élhosszúságú sétával. Ez könnyen látható, hiszen ha az i-edik csúcsból szeretnénk eljutni a j-edik csúcsba, akkor elindulunk i-ből a $\mathcal{G}(L)$-ben található irányított kör mentén az 1-es csúcshoz, majd onnan közvetlenül átmegyünk az $(j-i+2)$-edik csúcsba, ahonnan az irányított kör mentén ismét elmegyünk a j-edik csúcsba. Az alábbi ábra mutatja, hogy az említett csúcsok között milyen hosszúságú utakat teszünk meg:

$$i \xrightarrow{\;i-1\;} 1 \xrightarrow{\;1\;} j-i+2 \xrightarrow{\;n-i+2\;} j$$

4.2. ábra. A csúcsok közti utak élhosszúságai

Ilyen módon tehát bármely i csúcsból el tudunk jutni bármely j csúcsba. (Kivétel a $j = i-1$ eset, ilyenkor ugyanis $j-i+2 = 1$, ahová nem tudunk egy 1 élhosszúságú úton eljutni az első csúcsból, azonban ebben az esetben is belátható, hogy létezik $n+2$ élhosszúságú séta.) Így a séta élhossza összesen:

$$i - 1 + 1 + n - i + 2 = n + 2$$

Ezzel tehát beláttuk, hogy $G^{n+2} > 0$, és így $L^{n+2} > 0$. $\square$

A 4.1.3 Tétel alapján így L primitív, ezért alkalmazhatjuk a 4.1.2 tételt, tehát ha r a Leslie-mátrix Perron–Frobenius sajátértéke, és a hozzá tartozó jobb és bal oldali Perron vektorok p és q^T, akkor:

$$\lim_{t\to\infty} \frac{L^t}{r^t} = pq^T \tag{4.4}$$

Mivel $f(t+1) = Lf(t)$, így minden t esetén $f(t)$ kifejezhető az L mátrix és $f(0)$ segítségével:

$$f(t) = Lf(t-1) = L\left(Lf(t-2)\right) = \cdots = L^t f(0)$$

Feltehető, hogy $f(0) \neq 0$ (ugyanis $f(0) = 0$ azt jelentené, hogy a 0. évtizedben egyik korcsoportban se volt egyetlen nő sem), így ha beszorozzuk ezzel az $f(0)$-lal a (4.4) egyenletet jobbról, akkor a következőt kapjuk:

$$\lim_{t\to\infty} \frac{f(t)}{r^t} = \lim_{t\to\infty} \frac{L^t f(0)}{r^t} = pq^T f(0) \tag{4.5}$$

Fontos megjegyeznünk, hogy a $q^T f(0)$ szorzat egyenlő a q és $f(0)$ vektorok skaláris szorzatával, ami q pozitivitása miatt biztos, hogy egy pozitív szám, tehát $q^T f(0) = \alpha$ valamely pozitív α skalár esetén, és így a következőt írhatjuk fel:

$$\lim_{t\to\infty} \left\| \frac{f(t)}{r^t} \right\|_1 = \left\| \lim_{t\to\infty} \frac{f(t)}{r^t} \right\|_1 = \left\| pq^T f(0) \right\|_1 = \|p\alpha\|_1 = \alpha \|p\|_1 = \alpha \cdot 1 = \alpha > 0$$

Vegyük észre, hogy $F_k(t)$-t ki tudjuk fejezni $\|f(t)\|_1$ segítségével, és így $F(t)$ is felírható $\|f(t)\|_1$ segítségével:

$$F_k(t) = \frac{f_k(t)}{f_1(t) + \cdots + f_n(t)} = \frac{f_k(t)}{\|f(t)\|_1} \quad \Longrightarrow \quad F(t) = \frac{f(t)}{\|f(t)\|_1}$$

Ezek alapján pedig a következőeket írhatjuk fel:

$$\begin{aligned} \widetilde{F} = \lim_{t\to\infty} F(t) &= \lim_{t\to\infty} \frac{f(t)}{\|f(t)\|_1} = \lim_{t\to\infty} \frac{f(t)/r^t}{\|f(t)\|_1 / r^t} \\ &= \lim_{t\to\infty} \frac{f(t)/r^t}{\|f(t)/r^t\|_1} = \frac{\lim\limits_{t\to\infty} f(t)/r^t}{\lim\limits_{t\to\infty} \|f(t)/r^t\|_1} = \frac{p\alpha}{\alpha} = p \end{aligned}$$

Tehát $\widetilde{F}$ megegyezik a Leslie-mátrixhoz tartozó p Perron vektorral, és ez azt jelenti, hogy az idő előrehaladtával az egyes korosztályokban lévő nők létszámainak az aránya egy stabil állapothoz konvergál. A hosszútávú életkoreloszlást az L-hez tartozó Perron vektor adja meg, és mivel ez a p vektor pozitív, így minden korosztály valamely pozitív hányadosát teszi ki a teljes népességnek.

Érdekes észrevétel, hogy bár $t \longrightarrow \infty$ esetén a korcsoportok létszámának aránya állandó lesz, ennek ellenére az egyes csoportokban a nők létszáma növekedhet és csökkenhet is, r értékétől függően:

- Ha $r < 1$, akkor a (2.12) állítás alapján $f(t) = L^t f(0) \longrightarrow 0 \cdot f(0) = 0$.
- Ha $r = 1$, akkor ezt a (4.5) egyenletbe behelyettesítve azt kapjuk, hogy $f(t) \longrightarrow pq^T f(0) = p\alpha > 0$.

- Ha $r > 1$, akkor a (2.12) állítás bizonyításánál található gondolatmenet alapján létezik olyan $\epsilon > 0$ és $N \in \mathbb{N}$, hogy minden $t \geq N$ esetén $\|L^t\|_\infty > (1+\epsilon)^t \longrightarrow \infty$, és mivel $f(t) = L^t f(0)$, így minden k esetén $f_k(t) \longrightarrow \infty$.

4.4. Markov-láncok

Ahogy Meyer írja [1, 687. oldal 1. sor]: *„a Perron–Frobenius tétel egyik legelegánsabb alkalmazása a véges Markov-láncok algebrai vizsgálata”*. Most megismerkedhetünk a Markov-láncokhoz kapcsolódó alapvető fogalmakkal, és láthatjuk, hogy ezek milyen összefüggésben állnak a Perron–Frobenius tétellel.

4.4.1. Definíció. Legyen $P \in \mathbb{R}^{n\times n}$ egy nemnegatív mátrix, ekkor azt mondjuk, hogy P *sztochasztikus*, ha a mátrix minden sorában az elemek összege 1-gyel egyenlő.

4.4.2. Tétel. *Legyen $A \in \mathbb{R}^{n\times n}$ egy nemnegatív, irreducibilis mátrix, és legyen $\rho(A) = r$. Ekkor A hasonló egy rP mátrixhoz, ahol P egy irreducibilis sztochasztikus mátrix.*

Bizonyítás: Legyen $D \in \mathbb{R}^{n\times n}$ egy olyan mátrix, amelynek a főátlójában az A-hoz tartozó p Perron vektor elemei vannak, azon kívül pedig nullák:

$$D = \begin{pmatrix} p_1 & & 0 \\ & \ddots & \\ 0 & & p_n \end{pmatrix}$$

Legyen e egy csupa 1-esekből álló vektor, így mivel D^{-1} főátlójában a p_i elemek reciprokai vannak, ekkor a következő egyenlőség teljesül:

$$D^{-1}ADe = D^{-1}Ap = D^{-1}rp = rD^{-1}p = re$$

Ez azt jelenti, hogy a $D^{-1}AD$ mátrix minden sorában az elemek összege r-rel egyenlő. A Perron–Frobenius tétel alapján $r > 0$, így beszorozhatunk r^{-1}-el, tehát legyen $P = r^{-1}D^{-1}AD$. Ekkor P minden sorösszege 1-gyel egyenlő, tehát P sztochasztikus, továbbá $rP = D^{-1}AD$ hasonló A-hoz. Abból, hogy A irreducibilis, az következik, hogy P is irreducibilis, így tehát beláttuk a tétel állítását. □

4.4.3. Definíció. Legyen $X_0, X_1, \ldots$ valószínűségi változóknak egy sorozata úgy, hogy mindegyik X_t a lehetséges állapotoknak ugyanabból az $\{S_1, \ldots, S_n\}$ halmazából való. Ekkor azt mondjuk, hogy $\{X_t\}_{t=0}^{\infty}$ egy *sztochasztikus folyamat*.

4.4.4. Definíció. Egy folyamat *Markov-tulajdonságú*, ha a $(t+1)$-edik időpillanatbeli állapot valószínűsége csak a t-edik időpontbeli állapottól függ, az azelőttiektől pedig független, azaz bármely t esetén:

$$P(X_{t+1} = S_j \mid X_t = S_{i_t}, \dots, X_0 = S_{i_0}) = P(X_{t+1} = S_j \mid X_t = S_{i_t})$$

4.4.5. Definíció. A Markov-tulajdonságú sztochasztikus folyamatokat *Markov-láncoknak* nevezzük.

A Markov-láncokra tehát tekinthetünk úgy, mint véletlen állapotoknak egy sorozatára a $t = 0, 1, 2, \dots$ diszkrét időpontokban, ahol X_t írja le a t-edik időpontbeli állapotot. Például képzeljünk el egy udvart, amelyet kerítésekkel elkülönítettünk n különböző részre (úgy, hogy az egyes udvarrészek között vannak kapuk, amelyek biztosítják az átjárást), és ebben az udvarban egy kutya véletlenszerűen sétál. Ekkor S_i azt jelenti, hogy a kutya az i-edik udvarrészen tartózkodik éppen, és X_t írja le, hogy a kutya a t időpontban melyik udvarrészen van. Ekkor tehát a kutya mozgása Markov-tulajdonságú, ha a következő lépése csak az aktuális helyzetétől függ, és nem számít, hogy korábban az udvarnak melyik részein járt.

Minden Markov-lánc meghatároz egy sztochasztikus mátrixot, és ez fordítva is igaz.

4.4.6. Definíció. Jelölje $p_{ij}(t)$ azt a valószínűséget, hogy a t időpontban az S_j állapotban vagyunk, feltéve, hogy a $t-1$ időpontban az S_i állapotban voltunk, azaz: $p_{ij}(t) = P(X_t = S_j \mid X_{t-1} = S_i)$. Ekkor azt mondjuk, hogy p_{ij} az *átmenet-valószínűsége* annak, hogy a t időpontban az S_i állapotból az S_j állapotba kerülünk.

Ha ezeket az átmenet-valószínűségeket felírjuk egy $P(t) \in \mathbb{R}^{n \times n}$ mátrixba (ahol tehát a mátrix i-edik sorának j-edik eleme $p_{ij}(t)$-vel egyenlő), akkor $P(t)$ nyilván nemnegatív. Mivel a mátrix i-edik sora olyan számokat tartalmaz, amelyek megmutatják, hogy egy adott j állapotba mekkora valószínűségekkel kerülhetünk ebből az i állapotból, ezért ezeknek a valószínűségeknek az összege triviálisan 1-gyel egyenlő. Így $P(t)$ minden sorában az elemek összege 1, tehát $P(t)$ sztochasztikus.

4.4.7. Definíció. Azt mondjuk, hogy egy Markov-lánc *stacionárius átmenetvalószínűségű* (vagy más néven *homogén*), ha az átmenet-valószínűségek t-től függetlenek, azaz $p_{ij}(t) = p_{ij}$ minden t esetén. Ekkor $P(t)$ egy konstans mátrix, amelyet *átmenet-mátrixnak* nevezünk.

Így tehát minden Markov-lánc meghatároz egy P sztochasztikus mátrixot. Ugyanakkor, ha P egy $n \times n$ méretű sztochasztikus mátrix, akkor a p_{ij} elemekre tekinthetünk úgy, mint egy n elemű állapothalmazhoz tartozó átmenet-valószínűségekre, amelyek meghatároznak egy stacionárius Markov-láncot.

4.4.8. Definíció. Ha $p \in \mathbb{R}^n$ egy olyan nemnegatív vektor, amely elemeinek összege 1, akkor azt mondjuk, hogy p egy *valószínűségi eloszlás-vektor*. Például egy sztochasztikus mátrixnak minden sora ilyen vektor.

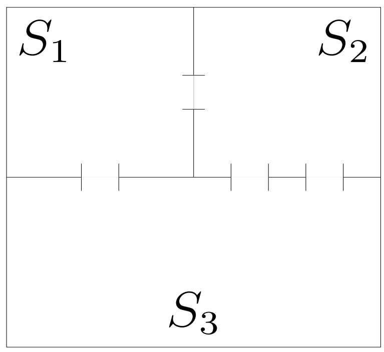
Egy olyan Markov-lánc esetén, amelynek az állapotterében n állapot van, a *k-adik lépés valószínűségi eloszlás-vektorán* a következő vektort értjük:

$$p(k) = \begin{pmatrix} p_1(k) & \dots & p_n(k) \end{pmatrix}^T \quad k = 1, 2, \dots, \quad \text{ahol} \quad p_j(k) = P(X_k = S_j)$$

A *kezdeti eloszlás-vektor* pedig a következő:

$$p(0) = \begin{pmatrix} p_1(0) & \dots & p_n(0) \end{pmatrix}^T, \quad \text{ahol} \quad p_j(0) = P(X_0 = S_j)$$

Ekkor tehát $p_j(k)$ annak a valószínűsége, hogy a k-adik lépés után (de még a $(k+1)$-edik lépés előtt) a j-edik állapotban vagyunk (azaz S_j-ben), $p_j(0)$ pedig annak a valószínűsége, hogy a folyamat az S_j állapotból indul.



4.3. ábra

Térjünk most vissza a kertben sétáló kutya példájához. Osszuk fel a kertet a 4.3 ábrán látható módon, és tegyük fel, hogy a kutyát minden t időpontban arra utasítjuk, hogy az aktuális udvarrészről menjen át egy másikra, de azt ő dönti el véletlenszerűen, hogy hová megy tovább.

Ha a kutyát kezdetben a harmadik udvarrészbe engedjük be, akkor a kezdeti eloszlás-vektor $p(0) = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix}^T$. Viszont ha például úgy indítjuk el a folyamatot, hogy kezdetben feldobunk egy labdát a kert fölött, majd ahová az leesik, a kutyát kezdetben arra az udvarrészbe engedjük be, akkor a kezdeti eloszlás-vektor $p(0) = \begin{pmatrix} 1/4 & 1/4 & 1/2 \end{pmatrix}^T$ lesz, mivel az egyes udvarrészek területei ekkora részeit teszik ki a teljes udvarnak. Ennél a példánál a Markov-lánchoz tartozó átmenet-mátrix a következő lesz:

$$M = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1/3 & 0 & 2/3 \\ 1/3 & 2/3 & 0 \end{pmatrix}$$

Ekkor $\sigma(M) = \{1, -1/3, -2/3\}$, tehát M egy olyan nemnegatív mátrix, amelyre $\rho(M) = 1$. Ez viszont nem véletlen, hiszen ha P egy tetszőleges sztochasztikus mátrix, akkor $\|P\|_\infty = 1$, és ha e pedig a csupa 1-esekből álló vektor, akkor $Pe = e$, tehát e sajátvektora P-nek és a hozzá tartozó sajátérték 1. Most használjuk fel azt, hogy az alábbi egyenlőség bármely A mátrixra igaz tetszőleges mátrixnorma esetén:

$$\forall \lambda \in \sigma(A) : \qquad |\lambda| \leq \|A\| \tag{4.6}$$

Ez könnyen látható, hiszen ha például a λ-hoz tartozó sajátvektor x, akkor legyen $X \in \mathbb{R}^{n \times n}$ egy olyan mátrix, amelynek az első osztlopában x áll, az összes többi eleme pedig nulla. Ekkor $\lambda X = AX$, amiből az következik, hogy $|\lambda| \|X\| = \|\lambda X\| = \|AX\| \leq \|A\| \|X\|$, ebből pedig már következik, hogy $|\lambda| \leq \|A\|$.

Így tehát a (4.6) egyenlőtlenség alapján:

$$1 \leq \rho(P) \leq \|P\|_\infty = 1 \quad \Longrightarrow \quad \rho(P) = 1$$

Tehát bármely P sztochasztikus mátrix spektrálsugara 1-gyel egyenlő, és $e > 0$ egy hozzá tartozó sajátvektor.

Most láthatunk egy lemmát, amely $p(0)$ segítségével megad egy explicit képletet $p(k)$-ra.

4.4.9. Lemma. *Legyen P egy irreducibilis átmenet-mátrixa egy olyan Markov-láncnak, amelynek a lehetséges állapotai $\{S_1, \ldots, S_n\}$. Ekkor bármely $p(0)$ kezdeti eloszlás-vektor esetén a k-adik lépés valószínűségi eloszlás-vektora az a $p(k)$ vektor, amelyre $p^T(k) = p^T(0)P^k$.*

Vezessük be a következő jelöléseket: legyen a P^k mátrix i-edik sorának j-edik eleme $p_{ij}^{(k)}$, jelölje e_i az i-edik egységvektort (azaz azt a vektort, amelynek i-edik eleme 1 és az összes többi 0), valamint legyen a $p(k)$ vektor j-edik eleme $p_j(k)$. Ekkor ha $p(0) = e_i$, akkor az előző tétel alapján $p(k) = e_i P^k$, azaz a $p(k)$ vektor megegyezik a P^k mátrix i-edik sorával. Ekkor tehát minden $1 \leq i \leq n$ esetén $p_j(k) = p_{ij}^{(k)}$, ami azt jelenti, hogy a P^k mátrix i-edik sorának j-edik eleme megmutatja, hogy mekkora a valószínűsége annak, hogy az S_i állapotból pontosan k lépéssel az S_j állapotba jutunk. Ezért azt mondjuk, hogy a P^k mátrix a *k-adik lépés átmenet-mátrixa.*

Most térjünk vissza a 4.3 ábrán látható kertben sétáló kutya példájához, és ahogy korábban feltettük, legyen véletlenszerű, hogy a kutya kezdetben honnan indul, tehát legyen $p(0) = \begin{pmatrix}1/4 & 1/4 & 1/2\end{pmatrix}^T$. Nézzük meg, hogy mi lesz például a 3., 8., 50., illetve 51. lépés eloszlás-vektora:

$$\begin{aligned} p(3) &= \begin{pmatrix}0,25 & 0,4120 & 0,3380\end{pmatrix}^T \\ p(8) &= \begin{pmatrix}0,25 & 0,3701 & 0,3799\end{pmatrix}^T \\ p(50) &= \begin{pmatrix}0,25 & 0,3750 & 0,3750\end{pmatrix}^T \\ p(51) &= \begin{pmatrix}0,25 & 0,3750 & 0,3750\end{pmatrix}^T \end{aligned}$$

Ez tehát azt jelenti, hogy például annak a valószínűsége, hogy a kutya a 8. lépés után a második udvarrészben van: $p_2(8) = 0,3701$.

Láthatjuk, hogy $p(50) = p(51)$, így mivel a 4.4.9 Lemma alapján $p(k) = p(k-1)M$, így minden további $k \geq 50$ esetén $p(k) = p(50)$ lesz, tehát jelen esetben az eloszlás-vektorok konvergensek. Ez azt jelenti, hogy ha hosszútávon vizsgáljuk a folyamatot, akkor az idő 25%-ában a kutya az első udvarrészben tartózkodik, míg a másodikban illetve harmadikban egyaránt az idő 37,5%-át tölti.

Az ilyen konvergencia általánosan is teljesül primitív mátrixok esetén, erről szól a következő tétel.

4.4.10. Tétel. *Ha egy Markov-lánchoz tartozó P átmeneti-mátrix primitív, amelynek a bal oldali Perron vektorát π^T jelöli és e-vel jelöljük a csupa 1-esekből álló vektort, akkor a következők teljesülnek:*

$$\lim_{k\to\infty} P^k = e\pi^T \quad \textit{és} \quad \lim_{k\to\infty} p^T(k) = \pi^T$$

Bizonyítás: Mivel P egy sztochasztikus mátrix, így a hozzá tartozó jobb oldali Perron vektor a csupa 1-esekből álló e vektor. Ekkor a 4.1.2 Tétel alapján mivel $r = \rho(P) = 1$, így:

$$\lim_{k\to\infty} P^k = \lim_{k\to\infty} \left(\frac{P}{r}\right)^k = e\pi^T$$

Az, hogy a $p(k)$ vektorok π-hez konvergálnak, könnyen látható, ha felhasználjuk a 4.4.9 Lemmát, miszerint $p^T(k) = p^T(0)P^k$:

$$\lim_{k\to\infty} p^T(k) = \lim_{k\to\infty} p^T(0)P^k = p^T(0)\lim_{k\to\infty} P^k = p^T(0)e\pi^T = \pi^T$$

Itt az utolsó egyenlőség azért teljesül, mert a $p^T(0)e$ szorzat azt jelenti, hogy a $p(0)$ vektor elemeit összeadjuk, és tudjuk, hogy ez az összeg 1, így $(p^T(0)e)\pi^T = 1 \cdot \pi^T = \pi^T$. □

A tételben szereplő π^T vektort *stacionárius eloszlás-vektornak*, vagy röviden *stacionárius vektornak* nevezzük, hiszen ez az a vektor, amelyre $\pi^T P = \pi^T$ teljesül.

5. fejezet

Google PageRank

A Perron–Frobenius tételnek fontos szerepe van a Google kereső működésében. A Google az úgynevezett PageRank algoritmussal határozza meg, hogy egy-egy kereséskor mely weboldalak jelenjenek meg elsőként a találatok között. Ezt az algoritmust Larry Page és Sergey Brin (a Google alapítói) dolgozták ki 1998-ban. Ahhoz, hogy megértsük a Perron–Frobenius tétel és a Google PageRank közötti összefüggést, érdemes először az algoritmus felépítésével foglalkozni.

5.1. A weboldalak reprezentációja

Ahhoz, hogy valamilyen fontossági sorrendet tudjunk felállítani az egyes honlapok között, ábrázolnunk kell őket matematikailag, ami a következő módon történik: képzeljük el, hogy az egyes weboldalak egy gráf csúcsai, és egy A csúcsból akkor megy (irányított) él egy B csúcsba, ha az A honlap hivatkozik a B-re. Ezáltal az interneten található weboldalak összessége felírható egy irányított gráffal.

A weboldalak sorrendbe állításához szükséges, hogy legyen egy olyan tulajdonságuk, ami alapján rendezni tudjuk őket: ez az adott honlap *rangja.* Ez az egyes oldalak hivatkozásaiból áll elő: ha egy weboldalra nagyon sok oldal hivatkozik, akkor nagyobb lesz a rangja, mint egy olyan oldalé, amelyikre kevés oldal hivatkozik. Fontos viszont, hogy a weboldal, ahonnan az adott honlapra érkezik a hivatkozás, milyen ranggal rendelkezik, ugyanis a nagyobb rangú oldalakról érkező hivatkozások fontosabbak, mint a kisebb rangúaké. Ugyanakkor az is számít, hogy egy-egy oldalból hány hivatkozás vezet más honlapokra: ha adott két ugyanolyan ranggal rendelkező weboldal, ahol az egyik jóval több oldalra hivatkozik, mint a másik, akkor ez utóbbi hivatkozásai számítanak értékesebbnek.

Röviden összefoglalva, a PageRank algoritmusban egy weboldal annál fontosabb, minél fontosabb oldalak hivatkoznak rá.

5.2. Matematikai felépítés

Brin és Page kezdetben az alábbi nagyon egyszerű egyenlettel határozták meg az egyes oldalak rangjait. Jelölje az i-edik weboldalt P_i, a rangját pedig $r(P_i)$, ami a P_i-re mutató összes honlap rangjainak az összege:

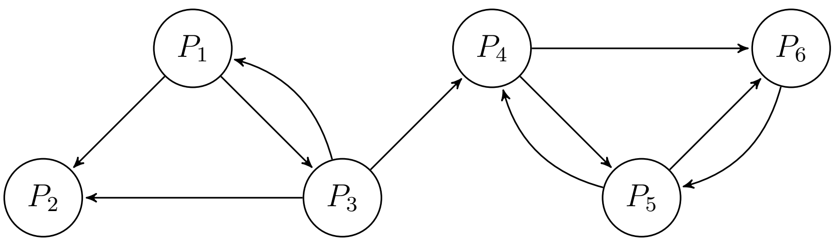
$$r(P_i) = \sum_{P_i \in B_{P_i}} \frac{r(P_j)}{|P_j|} \tag{5.1}$$

Itt B_{P_i} azt a halmazt jelöli, amelyben azok a honlapok találhatóak, amelyek hivatkoznak P_i-re, $|P_j|$ pedig a P_j oldalból kimenő hivatkozások száma. A probléma viszont ezzel az (5.1) egyenlettel az, hogy a P_i-re hivatkozó P_j oldalak rangjait nem tudjuk, így az $r(P_j)$-k ismeretlenek. Ezt viszont egy iterációval könnyen ki lehet küszöbölni: kezdetben minden oldal rangja $1/n$ (ahol n a weboldalak száma), és így alkalmazzuk az (5.1) egyenletet többször egymás után. Jelölje $r_{k+1}(P_i)$ a P_i oldal rangját a $k+1$. iterációnál. Így tehát:

$$r_{k+1}(P_i) = \sum_{P_i \in B_{P_i}} \frac{r_k(P_j)}{|P_j|} \tag{5.2}$$

Itt $r_0(P_i) = 1/n$ minden P_i esetén, és az iterációt addig ismételjük, amíg az oldalak rangjai valamely végső értékhez konvergálnak.

Az alábbiakban egy példát láthatunk egy olyan gráfra, amely 6 weboldalt ábrázol, majd láthatjuk, hogy az (5.2) egyenletet alkalmazva hogyan alakulnak az oldalak rangjai néhány iteráció után.



5.1. ábra. 6 weboldalt ábrázoló irányított gráf

Kezdeti érték	1. iteráció	2. iteráció	Rangsor a 2. iteráció után
$r_0(P_1) = 1/6$	$r_1(P_1) = 1/18$	$r_2(P_1) = 1/36$	5.
$r_0(P_2) = 1/6$	$r_1(P_2) = 5/36$	$r_2(P_2) = 1/18$	4.
$r_0(P_3) = 1/6$	$r_1(P_3) = 1/12$	$r_2(P_3) = 1/36$	5.
$r_0(P_4) = 1/6$	$r_1(P_4) = 5/36$	$r_2(P_4) = 11/72$	3.
$r_0(P_5) = 1/6$	$r_1(P_5) = 1/4$	$r_2(P_5) = 17/72$	1.
$r_0(P_6) = 1/6$	$r_1(P_6) = 1/6$	$r_2(P_6) = 14/72$	2.

5.1. táblázat. Az oldalak rangjai az első két iteráció során

5.2.1. A feladat mátrix reprezentációja

Az (5.1) és (5.2) egyenletekkel egyszerre csak egy oldal rangját tudjuk kiszámolni. Mátrixok segítségével viszont ki tudunk számítani egy n hosszú PageRank vektort, amelynek i-edik eleme a P_i oldal rangja. Ehhez bevezetünk egy $H \in \mathbb{R}^{n\times n}$ mátrixot és egy $\pi^T \in \mathbb{R}^{1\times n}$ sorvektort, ahol:

$$H_{ij} = \begin{cases} \dfrac{1}{|P_i|} & \text{ha } P_i\text{-ből } P_j\text{-be vezet él} \\ \\ 0 & \text{különben} \end{cases}$$

H nagyon hasonlít a gráf adjacencia mátrixához, csak itt a nemnulla elemek nem 1-esek, hanem valószínűségek: H_{ij} megmutatja, hogy mekkora valószínűséggel jutunk el P_i-ből P_j-be. Az 5.1 ábrán látható gráfhoz tartozó H mátrix a következőképpen néz ki:

$$H = \begin{array}{c} \\ P_1 \\ P_2 \\ P_3 \\ P_4 \\ P_5 \\ P_6 \end{array} \begin{array}{c} \begin{array}{cccccc} P_1 & P_2 & P_3 & P_4 & P_5 & P_6 \end{array} \\ \begin{pmatrix} 0 & 1/2 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1/3 & 1/3 & 0 & 1/3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \end{array}$$

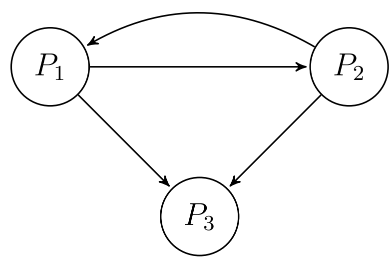
Legyen $\pi^{(k)T}$ a PageRank vektor a k-adik iteráció után, és legyen $\pi^{(0)T} = \begin{pmatrix} 1/n & \dots & 1/n \end{pmatrix}$. Ekkor az (5.2) egyenlet a következő módon írható fel:

$$\pi^{(k+1)T} = \pi^{(k)T} H \tag{5.3}$$

Ez az iteráció egy $n \times n$-es mátrix esetén $O(n^2)$ műveletet igényelne, de mivel H egy elég ritka mátrix, így a gyakorlatban sokkal könnyebben számolhatunk vele, ha csak a nemnulla elemeit és azok elhelyezkedését tároljuk. Becslések alapján egy weboldal átlagosan 10 másik oldalra hivatkozik, tehát egy $n \times n$-es mátrixnak körülbelül $10n$ nemnulla eleme van, és ezzel az iteráció műveletigénye $O(n)$-re csökkenthető.

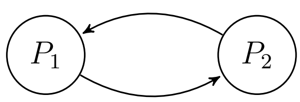
5.3. Az iteráció problémái

Az iteráció eddig felépített változatával még sok a probléma. Például vegyük azokat az oldalakat, amelyekből egyetlen él sem megy ki. Ezek az oldalak minden egyes iterációkor egyre több PageRanket gyűjtenek össze, és így „kisajátítják” a pontokat, hiszen mivel nem vezet ki belőlük él, a saját rangjukat nem osztják meg egyetlen oldallal sem. Az 5.2 ábrán láthatunk egy példát erre az esetre, ahol P_3-ból egyetlen él sem megy ki.



5.2. ábra

Az előző, azaz az 5.1 ábra is tartalmaz egy hasonló esetet: a P_4, P_5 és P_6 oldalak egy olyan halmazt alkotnak, ahonnan nem vezet ki él, így a kezdeti $\pi^{(0)T}$ vektor az (5.3) egyenletet alkalmazva már a 10. iteráció után a következőképpen néz ki: $\pi^{(10)T} = \begin{pmatrix} 0 & 0 & 0 & 2/15 & 4/15 & 1/5 \end{pmatrix}$. Az első három oldal rangja tehát nullára csökkent, pedig a cél az lenne, hogy minden weboldal egy pozitív ranggal rendelkezzen.



5.3. ábra

Egy másik komoly problémát a ciklusok jelentenek. Nézzük az 5.3 ábrán látható egyszerű példát, ahol a P_1 és P_2 oldalak csak egymásra hivatkoznak. Ennek a gráfnak a H mátrixa a következő:

$$H = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Ekkor ha az (5.3) egyenletet a $\pi^{(0)T} = \begin{pmatrix} 1 & 0 \end{pmatrix}$ kezdeti vektorra alkalmazzuk, az iteráció nem fog konvergálni, ugyanis $\pi^{(k)T}$ mindig az $\begin{pmatrix} 1 & 0 \end{pmatrix}$ vagy a $\begin{pmatrix} 0 & 1 \end{pmatrix}$ vektor lesz, attól függően, hogy k páros vagy páratlan.

5.4. Kapcsolat a Markov-láncokkal

Az (5.3) egyenletre tulajdonképpen úgy is tekinthetünk, hogy ez a hatvány iteráció alkalmazása egy Markov-lánc esetén, amelynek az átmeneti valószínűségi mátrixa H. Ez egy fontos észrevétel, ugyanis ahogy már korábban láthattuk, a Perron–Frobenius tétel jól alkalmazható a Markov-láncok körében, és így a PageRank algoritmusban is fontos szerepe van.

Ha P egy olyan Markov mátrix, amely sztochasztikus, irreducibilis és aperiodikus, akkor tudjuk, hogy tetszőleges kezdővektor esetén a hatvány módszer konvergál egy pozitív vektorhoz, amely egyértelmű, ez a *stacionárius vektor*. Így tehát a PageRank algoritmus már korábban említett konvergencia problémáira (csúcsok kimenő élek nélkül, illetve ciklusok) könnyen találhatunk megoldást, ha a H mátrixot úgy módosítjuk, hogy egy Markov mátrix legyen amely rendelkezik az előbbi három tulajdonsággal. Ez a végső mátrix lesz az úgynevezett *Google mátrix*.

Összefoglalva tehát a Perron–Frobenius tétel alapján egyértelműen létezik egy pozitív PageRank vektor, amennyiben a Google mátrix sztochasztikus és irreducibilis. Továbbá, ha a mátrix aperiodikus, akkor a hatvány iteráció tetszőleges kezdővektor esetén konvergál ehhez a PageRank vektorhoz.

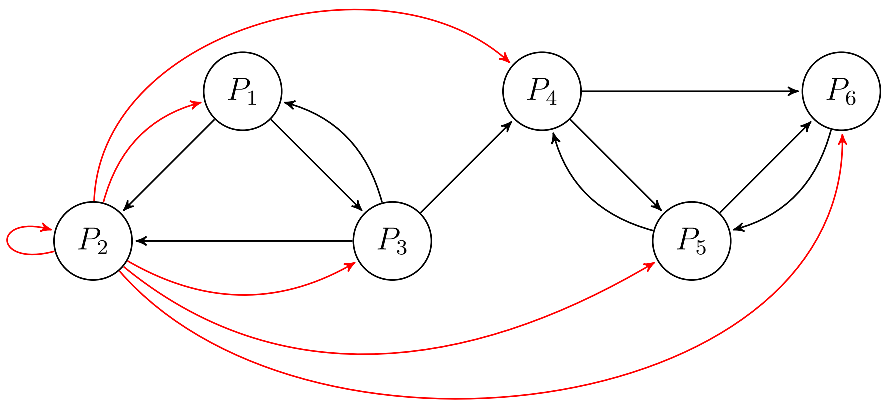
5.5. A Google mátrix felépítése

Brin és Page eredetileg nem hozták összefüggésbe az algoritmusukat a Markov-láncokkal, hanem egy „*random surfer*” nevű fogalmat használtak, ami egy olyan embert jelképez, aki véletlenszerűen böngészi az internetet, mégpedig úgy, hogy követi a hivatkozásokat. Tehát ha egy A weboldal több oldalra is hivatkozik, akkor az internetező kiválaszt egy véletlen oldalt ezek közül, így eljut egy B honlapra, majd ezt a folyamatot folytatva, megnézi, hogy B milyen oldalakra hivatkozik, és azok közül is választ egyet véletlenszerűen, és így halad tovább. Így azonban hamar egy olyan problémába ütközhet, amelyet korábban már említettünk: ha egy olyan oldalra jut, amely nem hivatkozik sehová sem (például egy pdf fájl).

Az ilyen oldalaknak a H mátrixban a csupa nulla sorok felelnek meg, így ezt a problémát úgy lehet megoldani, ha ezeket a sorokat lecseréljük csupa $1/n$-ből álló sorokra. Ezzel kapunk egy S mátrixot, amely már sztochasztikus, és így az internetező egy ilyen linkre érkezve egyenlő valószínűséggel bármelyik másik oldalra tovább haladhat. Ezt matematikailag a következőképpen írhatjuk fel: $S = H + a(\frac{1}{n}e^T)$, ahol e a csupa 1-esekből álló vektor, a pedig egy olyan vektor, amelyre $a_i = 1$, ha az i-edik oldalból nem vezet ki él, különben 0. Így például az 5.1 ábrához tartozó S mátrix a következőképpen néz ki:

$$S = \begin{pmatrix} 0 & 1/2 & 1/2 & 0 & 0 & 0 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/3 & 1/3 & 0 & 1/3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

Ha a weboldalak és a hivatkozások gráfos reprezentációjára gondolunk, ez a megoldás annak felel meg, mintha minden olyan pontba, amelyből nem vezet ki él, behúznánk egy-egy kimenő élt a gráf összes csúcsába, ami ebben az esetben így nézne ki:



5.4. ábra. Az S mátrix gráfos ábrázolása.
Az újonnan behúzott éleket piros szín jelöli.

Most azt szeretnénk, hogy a mátrix primitív is legyen, ami Markov-láncok esetén azt jelenti, hogy a stacionárius vektor (ami jelen esetben a PageRank vektor) egyértelműen létezik, és könnyen kiszámítható a hatvány módszerrel. Brin és Page a random surfer segítségével ezt úgy fogalmazták meg, hogy bár az internetező követi a hivatkozások struktúráját, de néha meggondolja magát, és függetlenül attól, hogy épp melyik oldalon van és az mely oldalakra hivatkozik, beír a böngésző URL sávjába egy tetszőleges weboldalt, és onnan böngészik tovább.

Ahhoz, hogy ezeket a véletlen ugrásokat matematikailag leírjuk, a következő mátrixot vezetjük be:

$$G = \alpha S + (1-\alpha)\frac{1}{n}ee^T$$

Itt $\alpha \in]0,1[$, e pedig a csupa 1-esekből álló vektor, és ezt a G mátrixot nevezzük *Google mátrixnak*. Az α paraméter azt modellezi, hogy az internetező mennyi időt töltött a hivatkozások struktúrájának követésével, és mennyit az oldalak közti ugrásokkal. Például $\alpha = 0,7$ azt jelenti, hogy a böngészéssel töltött idő 70%-ában úgy jutott el egy A oldalból B-be, hogy A hivatkozott B-re, és így rákattintott, az idő 30%-ában pedig beírt a böngészőjébe egy tetszőleges weboldalt. Az ilyen utóbbi ugrások valóban véletlenszerűek, ugyanis az ezeknek megfelelő $E = \frac{1}{n}ee^T$ mátrixnak minden eleme $\frac{1}{n}$-nel egyenlő, tehát az internetező valóban egyforma valószínűséggel jut bármelyik oldalra az egyes ugrások során.

5.5.1. A mátrix tulajdonságai

Az alábbiakban láthatjuk összefoglalva, hogy a felépített G Google mátrix milyen tulajdonságokkal rendelkezik.

- G **sztochasztikus**, mivel konvex kombinációja két sztochasztikus mátrixnak, S-nek és $E = \frac{1}{n}ee^T$-nak.
- G **irreducibilis**, ugyanis az E mátrix által minden oldal közvetlenül össze van kötve az összes többi oldallal, és így szükségképpen teljesül az irreducibilitás.
- G **aperiodikus**, ez következik abból, hogy minden i-re $G_{ii} > 0$.
- G **primitív**, hiszen már $k = 1$ esetén teljesül, hogy $G^k > 0$. Ebből következik az, hogy egyértelműen létezik egy pozitív π^T PageRank vektor, és hogy G-re alkalmazva a hatvány iteráció ehhez a vektorhoz konvergál.
- G **sűrű** abban az értelemben, hogy az elemeinek a nagy rész nem nulla, ez pedig a számítások szempontjából kifejezetten előnytelen. Azonban ki tudjuk fejezni G-t úgy, hogy csak az eredeti H mátrixunk szerepel az egyenletben, és ezáltal H-val már könnyen tudunk számolni:

$$\begin{aligned} G &= \alpha S + (1-\alpha)\frac{1}{n}ee^T \\ &= \alpha\left(H + \frac{1}{n}ae^T\right) + (1-\alpha)\frac{1}{n}ee^T \\ &= \alpha H + (\alpha a + (1-\alpha)e)\frac{1}{n}e^T \end{aligned} \tag{5.4}$$

A Google mátrix elég mesterséges abból a szempontból, hogy az eredeti H mátrixot kétszer is módosítottuk ahhoz, hogy végül egy olyan mátrixot kapjunk, amelyre teljesülnek a konvergenciához szükséges feltételek. Ugyanakkor ezen módosítások nélkül nem tudnánk kiszámolni a rangsort, hiszen a H mátrixhoz nem létezik stacionárius vektor, viszont G-hez igen, és az így kapott PageRank vektor jól mutatja az egyes weboldalak fontosságát.

5.6. A PageRank vektor kiszámítása

Összefoglalva, a Google PageRank algoritmusa az alábbi iterációból áll:

$$\pi^{(k+1)T} = \pi^{(k)T} G \tag{5.5}$$

Ez tulajdonképpen a hatvány módszer G-re alkalmazva. Így a PageRank vektor kiszámítása az alábbi sajátérték problémával írható fel:

$$\begin{aligned} \pi^T &= \pi^T G \\ \pi^T e &= 1 \end{aligned}$$

Ha ezt transzponáljuk, akkor a Perron–Frobenius tételben szereplő alakban láthatjuk a feladatot:

$$\begin{aligned} G^T \pi &= \pi \\ e^T \pi &= 1 \end{aligned}$$

A Perron tétel (2.0.1 Tétel) pont erre vonatkozik: G egy pozitív mátrix, tehát alkalmazható rá a tétel, és mivel G sztochasztikus, így a legnagyobb sajátértéke 1, tehát a tétel (2.5) pontja alapján egyértelműen létezik egy ilyen pozitív π vektor.

A PageRank vektor kiszámításához célszerű G-t az (5.4) egyenlet segítségével felírni, mert így csak a H-t kell használnunk a számítások során.

Ez azért hasznos, mert H egy ritka mátrix, így sokkal könnyebb tárolni is és számolni vele. Így az (5.5) egyenlet a következőképpen írható át:

$$\begin{aligned}\pi^{(k+1)T} &= \pi^{(k)T}G \\ &= \alpha\pi^{(k)T}S + \frac{1-\alpha}{n}\pi^{(k)T}ee^T \\ &= \alpha\pi^{(k)T}H + \left(\alpha\pi^{(k)T}a + 1 - \alpha\right)\frac{1}{n}e^T \end{aligned} \tag{5.6}$$

Most nézzük meg, hogy az 5.1 ábrán látható 6 weboldalnál mi is lesz a végső rangsor, azaz számoljuk ki a hozzá tartozó Google mátrixot, majd a PageRank vektort az (5.6) egyenlet segítségével. Ehhez el kell döntenünk, hogy milyen α paraméterrel dolgozzunk. Brin és Page eredetileg $\alpha = 0,85$-tel számoltak, és végül a Google kereső is ezt a számot használja. Ennek az az oka, hogy $\alpha^{50} = 0,85^{50} \approx 0,000296$, ami azt jelenti, hogy már az 50. iteráció után körülbelül 2-3 tizedesjegy pontossággal megadható a PageRank vektor. Ez a pontosság pedig elegendő egy jó rangsor felállításához, tehát legyen a mi esetünkben is $\alpha = 0,85$, és ekkor az 5.1 ábrához tartozó Google mátrix a következő:

$$G = 0,85H + (0,85\begin{pmatrix}0\\1\\0\\0\\0\\0\end{pmatrix} + 0,15\begin{pmatrix}1\\1\\1\\1\\1\\1\end{pmatrix})\frac{1}{6}\begin{pmatrix}1 & 1 & 1 & 1 & 1 & 1\end{pmatrix}$$

$$= \begin{pmatrix} 1/40 & 9/20 & 9/20 & 1/40 & 1/40 & 1/40 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 37/120 & 37/120 & 1/40 & 37/120 & 1/40 & 1/40 \\ 1/40 & 1/40 & 1/40 & 1/40 & 9/20 & 9/20 \\ 1/40 & 1/40 & 1/40 & 9/20 & 1/40 & 9/20 \\ 1/40 & 1/40 & 1/40 & 1/40 & 7/8 & 1/40 \end{pmatrix}$$

A hozzá tartozó PageRank vektor pedig a következő:

$$\begin{array}{cc} & \begin{array}{cccccc} P_1 & P_2 & P_3 & P_4 & P_5 & P_6 \end{array} \\ \pi^T = & \begin{pmatrix} 0,0365 & 0,0520 & 0,0405 & 0,1410 & 0,2460 & 0,1895 \end{pmatrix} \end{array}$$

Ennek a vektornak az i-edik koordinátája megmondja, hogy az internetező mennyi időt töltött a P_i oldal látogatásával, például a P_3 honlapot az idő 4,05%-ában látogatta meg. Ezek alapján a rangsort felírhatjuk a $\begin{pmatrix} P_5 & P_6 & P_4 & P_2 & P_3 & P_1 \end{pmatrix}$ vektorral, ami azt jelenti, hogy P_5 a legnagyobb ranggal rendelkező weboldal, P_1 pedig a legkisebbel.

6. fejezet

Általánosítás végtelen dimenzióban

A korábbi fejezetekben láthattuk, hogy véges dimenzióban a Perron–Frobenius tételkört számos helyen lehet alkalmazni, pozitív és nemnegatív mátrixok esetén egyaránt. Ebben a fejezetben arról olvashatunk, hogy hogyan lehetne a tételt általánosítani végtelen dimenziós esetre.

A véges dimenziós mátrixoknak a végtelen dimenzióban a korlátos lineáris operátorok felelnek meg (például a Banach-terek felett). Egy operátor sajátértékét a mátrixok sajátértékéhez hasonlóan definiáljuk, azonban az operátor spektruma nem a sajátértékeinek, hanem a szinguláris értékeinek a halmaza. Ezek minden Banach tér esetében értelmezhetők, a pozitivitás viszont csak bizonyos esetekben. Például L^p terekben, ahol a pozitivitás a pontonkénti kiértékelésekből örökíthető, azaz egy $f \in L^p$ függvény pozitív, ha $f \geq 0$ majdnem mindenütt. Vegyük észre, hogy amit a Banach-terek kontextusában pozitivitásnak nevezünk, az tulajdonképpen véges dimenzióban a nemnegativitásnak felel meg. Ilyenkor tehát egy T operátort akkor mondunk pozitívnak, ha a pozitív elemeknek az operátor által megadott képei is pozitívak, vagyis $f \geq 0 \Rightarrow T(f) \geq 0$.

Az ilyen, pozitivitás fogalommal ellátott Banach-tereket *Banach-hálóknak* nevezzük. Ekkor a Perron–Frobenius tételnek megfelelően a következő tétel mondható ki.

6.0.1. Tétel. *Legyen $T\colon \mathbb{B} \longrightarrow \mathbb{B}$ pozitív lineáris operátor, ahol $\mathbb{B}$ tetszőleges Banach-háló. Jelölje T spektrumát $\sigma(T)$ és legyen $r(T) = \lim\limits_{k\to\infty} \left\|T^k\right\|^{1/k}$. Ekkor $r(T) \in \sigma(T)$.*

Továbbá, ha feltesszük, hogy létezik olyan $M > 0$ *konstans, amelyre minden* $n \in \mathbb{N}$ *esetén teljesül, hogy* $\|T^n\| < Mr(T)^n$*, akkor:*

$$r(T)e^{i\varphi} \in \sigma(T) \quad \Longrightarrow \quad r(T)e^{in\varphi} \in \sigma(T) \quad \forall n \in \mathbb{Z}.$$

A Perron–Frobenius tétel általánosítása fontos szerepet tölt be a differenciálegyenletek témakörében. Azt mondjuk, hogy egy $(T(t))_{t\geq 0}$ operátorcsalád *erősen folytonos*, ha az adott Banach-tér minden x elemére teljesül, hogy bármely t_0 esetén $\lim\limits_{t\to t_0} \|T(t)x - T(t_0)x\| = 0$ (ekkor a vizsgált norma a Banach-tér normáját jelenti). Parciális differenciálegyenletek esetén ha egy x kezdeti feltétel mellett $T(t)x$ adja meg a megoldást a t időpontban, akkor általában teljesül az erősen folytonosság feltétele. Így ez a feltétel teljesül például a hővezetési egyenletre is.

Ha $(T(t))_{t\geq 0}$ egy olyan erősen folytonos operátorcsalád, amelyre $T(0) = I$, és bármely nemnegatív t_1 és t_2 esetén $T(t_1 + t_2) = T(t_1)T(t_2)$, akkor azt mondjuk, hogy $(T(t))_{t\geq 0}$ egy C_0-*félcsoport*. Ez szemléletesen azt jelenti, hogy ha például vizsgáljuk egy szoba hőmérsékletét (például $t_1 = 1$ és $t_2 = 2$ perc elteltével), akkor a szoba hőmérséklete a 0. pillanathoz képest $1 + 2$ perc múlva ugyanannyi lesz, mintha a kezdethez képest külön vizsgáltuk volna 1 majd újabb 2 perc elteltével. (Tehát a vizsgált idő szétbontása nem módosított az eredményen.)

Vezessük be a $D(A)$ jelölést azon kezdeti feltételek halmazára, amelyekre differenciálható módon változik a megoldás:

$$D(A) = \left\{ x \in X : \exists \lim_{t\to 0+} \frac{T(t)x - x}{t} \right\}$$

Ez a halmaz általában megegyezik a klasszikus kezdeti feltételekkel. Ha A egy olyan operátor, amely $D(A)$-ból képez az $\mathbb{X}$ Banach-térbe úgy, hogy Ax megegyezik a $D(A)$ halmazt meghatározó határértékkel, akkor szemléletesen Ax azt mutatja, hogy mi lesz $T(t)x$ deriváltja a 0-ban:

$$Ax(t) = \partial_t x(t) \quad \Longrightarrow \quad x(t) = T(t)x(0)$$

Ugyanakkor, ha tekintjük véges dimenzióban az $Ax(t) = \partial_t x(t)$ differenciálegyenlet-rendszert, ahol A egy négyzetes mátrix, akkor tudjuk, hogy annak megoldása felírható $x(t) = e^{At}x(0)$ alakban. Ha ehhez hozzávesszük az előző egyenletet, akkor azt kapjuk, hogy $T(t) = e^{At}$. Ez azt jelenti, hogy ez a $T(t)$ operátor félcsoport általánosítja a véges dimenziós e^{At} hatványt.

Így végül operátor-félcsoportokra is kimondható egy tétel, amely a Perron–Frobenius tételt általánosítja.

6.0.2. Tétel. *Legyen* $(T(t))_{t\geq 0}$ *egy pozitív* C_0*-félcsoport az* $\mathbb{X}$ *Banach-háló felett. Ekkor* $s(A) := \sup\{\mathrm{Re}\lambda\colon \lambda \in \sigma(A)\} \in \sigma(A)$.

Továbbá, ha feltesszük, hogy létezik olyan $M > 0$ *konstans, hogy minden* $t \geq 0$ *esetén* $\|T(t)\| < Me^{s(A)t}$*, akkor:*

$$s(A) + i\beta \in \sigma(A) \quad \Longrightarrow \quad s(A) + in\beta \in \sigma(A) \quad \forall n \in \mathbb{Z}.$$

Irodalomjegyzék

[1] Carl D. Meyer, *Matrix Analysis and Applied Linear Algebra*, SIAM, 2000.

[2] Roger A. Horn, Charles R. Johnson, *Matrix Analysis*, Cambridge university press, 2013.

[3] Peter D. Lax, *Linear Algebra and Its Applications*, Wiley-interscience, 2007.

[4] Thomas L. Saaty, *Rank According to Perron: A New Insight*, University of Pittsburgh.
https://galton.uchicago.edu/~lekheng/meetings/mathofranking/ref/saaty.pdf

[5] Shlomo Sternberg, *The Perron–Frobenius theorem*, Harvard University.
http://www.math.harvard.edu/library/sternberg/slides/1180912pf.pdf

[6] Hannah Cairns, *A short proof of Perron's theorem*, 2014.
http://www.math.cornell.edu/~web6720/Perron-Frobenius_Hannah%20Cairns.pdf

[7] Bakonyi Eszter, *Nemnegatív mátrixok* szakdolgozat, 2014.
https://www.cs.elte.hu/blobs/diplomamunkak/bsc_matelem/2014/bakonyi_eszter.pdf

[8] Lukács Erzsébet, *Alkalmazott algebra* előadás jegyzet, 2015.
http://math.bme.hu/~lukacs/bboard/alkalg/2015/ea12_15aa.pdf

[9] Bogya Norbert, *Leontief-féle input-output modell.*
http://www.math.u-szeged.hu/~nbogya/linalgk1011II/leontief.pdf

[10] *Leslie-modell*, ELTE biológia alapszak 2008/09 1. féléves matematika gyakorlat jegyzet.
http://people.inf.elte.hu/pazqaat/Matek/gy1002.pdf

[11] Csiszár Villő, *Diszkrét és folytonos paraméterű Markov láncok.*
http://www.cs.elte.hu/~villo/ml/ML.pdf

[12] Böjthy Barbara Adrienn, *Sztochasztikus mátrixok és Markov-láncok* szakdolgozat, 2013.
https://www.cs.elte.hu/blobs/diplomamunkak/bsc_matelem/2013/bojthy_barbara_adrienn.pdf

[13] Amy N. Langville, Carl D. Meyer, *Google's PageRank and Beyond: The Science of Search Engine Rankings*, Princeton University Press.
http://geza.kzoo.edu/~erdi/patent/langvillebook.pdf

[14] Jochen Glück, *Positive Semigroups and Perron Frobenius Theory*, Ulm University, 2013.
https://www.uni-ulm.de/fileadmin/website_uni_ulm/mawi.inst.020/glueck/Dateien_fuer_Homepage/Vortragsfolien/J_Glueck_Perron_Frobenius_Theory.pdf

[15] Adam Kanigowski, Wojciech Kryszewski, *Perron-Frobenius and Krein-Rutman theorems for tangentially positive operators* Central European Journal of Mathematics, 2012.

Nyilatkozat

Név: Maros Alexandra
ELTE Természettudományi Kar, szak: matematika BSc.
NEPTUN azonosító: N9T3HS
Szakdolgozat címe: Perron–Frobenius tételkör és alkalmazásai

A **szakdolgozat** szerzőjeként fegyelmi felelősségem tudatában kijelentem, hogy a dolgozatom önálló munkám eredménye, saját szellemi termékem, abban a hivatkozások és idézések standard szabályait következetesen alkalmaztam, mások által írt részeket a megfelelő idézés nélkül nem használtam fel.

Budapest, 2016. május 27.

a hallgató aláírása