EÖTVÖS LORÁND TUDOMÁNYEGYETEM
TERMÉSZETTUDOMÁNYI KAR

Mátrixok és hálózati keresés

Szakdolgozat

Kovács Zsófia
Matematika Bsc
Matematikai elemző szakirány

Témavezető: Ágoston István
egyetemi docens
Algebra és Számelmélet Tanszék

ELTE
2017

Tartalomjegyzék

1. Bevezetés **4**

2. Algebrai bevezető **5**
2.1. Gersgorin-körök 8
2.2. Hatványmódszer 8

3. Pozitív mátrixok **10**
3.1. Perron-tétel 10
3.2. A Perron-tétel alkalmazása 14
3.3. Általánosítás nemnegatív mátrixokra 16

4. A Page Rank algoritmus **22**
4.1. Keresőmotorok a Page Rank előtt 22
4.2. Motiváció 22
4.3. A Page Rank alapötlete 23
4.4. Az egyszerű Page Rank algoritmus működése 23
4.5. Példa 24
4.6. Problémák az egyszerű működéssel 26
4.7. Kibővített Page Rank algoritmus 27
4.8. Markov-láncok és a Page Rank 28

5. Egyéb rangsoroló algoritmusok **30**
5.1. A HITS algoritmus 30

Irodalomjegyzék **31**

Köszönetnyilvánítás

Ezúton szeretnék köszönetet mondani témavezetőmnek, Ágoston Istvánnak, aki tanácsaival, észrevételeivel és szakértelmével nagyban hozzájárult a dolgozatom elkészítéséhez. Köszönet illeti, amiért mindig bizalommal fordulhattam hozzá kérdéseimmel. Továbbá köszönöm Lukács Erzsébetnek, hogy szakértelmével hozzájárult a dolgozatom elkészítéséhez.

Hálával tartozom családomnak, akik tanulmányaim alatt végig támogattak és nyugodt hátteret biztosítottak számomra.

Szeretném megköszönni barátaimnak, hogy mellettem álltak és mindvégig biztattak, Dudás Eszternek és Karácsonyi Csabának külön köszönöm, hogy figyelmes észrevételeikkel segítették munkámat.

1. Bevezetés

Dolgozatom célja a keresési algoritmusok matematikájának bemutatása. A keresőmotorok azon részére fektetek nagy hangsúlyt, ami a találatok sorbarendezéséért felelős. Erre azért van nagy szükség, mert egy keresés során a Google akár több százmillió oldalt is találhat.

Az iteratív eljárás során nemnegatív mátrixok gyors hatványozására van szükség, ezért az algoritmus matematikai hátterének bemutatása során pozitív, majd általánosabban, nemnegatív mátrixokról lesz szó. Az alkalmazás szempontjából fontos megvizsgálni ezen mátrixok sajátértékeit és sajátvektorait, a mátrixok hatványozásában ugyanis döntő szerepet játszanak a sajátértékek.

A második fejezetben bemutatom a Perron-tételt és annak bizonyítását. Ezt követően pedig egy általánosabb formában ismertetem a tétel nemnegatív mátrixokra való kiterjesztését, az úgynevezett Frobenius-tételt. Később, a kereső algoritmus bemutatása során láthatjuk, hogy az iterációban nemnegatív, azaz nem feltétlenül pozitív mátrixokkal számolunk, ezért van szükség az általánosítás részletesebb megismerésére. A fejezetben szereplő definíciók és tételek áttekintéséhez Peter D. Lax Lineáris Algebra és alkalmazásai című könyvet [6] vettem alapul, noha az általános tétel bizonyításában kissé más utat követtem.

A harmadik fejezetben nagy vonalakban bemutatom a Google keresőmotor által használt rangsoroló algoritmust, majd egy példa segítségével szemléltetem a működését.

2. Algebrai bevezető

Ebben a fejezetben mátrixok sajátértékeivel és sajátvektoraival kapcsolatos definíciókat és tételeket fogok ismertetni, amikre később is támaszkodni fogunk. Néhányat közülük ismertnek tekintek, ezért ezeket csak vázlatosan fogom leírni. A továbbiakban az $A \in \mathbb{C}^{n\times n}$ mátrixot, $\mathbf{x} \in \mathbb{C}^n$ vektort és $\lambda \in \mathbb{C}$ skalárt fog jelölni.

2.1. Definíció. (Sajátvektor, sajátérték) Ha $\mathbf{v} \neq \mathbf{0}$ olyan vektor, melyre

$$A\mathbf{v} = \lambda\mathbf{v}, \tag{1}$$

akkor azt mondjuk, hogy $\mathbf{v}$ az A mátrix *sajátvektora*, és λ az $\mathbf{v}$-hez tartozó *sajátértéke.*

2.2. Tétel. *Az A mátrixnak pontosan n darab sajátértéke van a multiplicitásukat is figyelembe véve. A sajátértékek a $det(A-\lambda I) = 0$ egyenlet megoldásai. A λ sajátértékhez tartozó sajátvektorok pedig a $(A - \lambda I)\mathbf{v} = 0$ egyenletrendszer nullvektortól különböző megoldásai.*

2.3. Definíció. (Karakterisztikus egyenlet, karakterisztikus polinom) A $det(A - \lambda I) = 0$ egyenletet az A mátrix *karakterisztikus egyenletének* nevezzük, és ha A egy $n \times n$-es mátrix, akkor az egyenlet bal oldala a determináns kifejtése után egy n-edfokú polinom. Ezt az A mátrix *karakterisztikus polinomjának* nevezzük és, $k_A(\lambda)$-val jelöljük.

2.4. Tétel. *Ha k_A az A mátrix karakterisztikus polinomja, akkor igazak az alábbi állítások.*

(i) A k_A n-edfokú polinom $(-1)^n$ főegyütthatóval.

(ii) A k_A karakterisztikus polinom gyökei az A sajátértékei.

(iii) A sajátértékek összege az A mátrix nyoma, amit a főátlóban álló elemek összegeként definiálunk. Ez pedig megegyezik a karakterisztikus polinom $(n-1)$-edfokú tagjához tartozó együttható (-1)-szeresével.

(iv) A sajátértékek szorzata egyenlő az A mátrix determinánsával. Ez a k_A polinom konstans tagja.

2.5. Tétel. *Az A mátrixnak és transzponáltjának karakterisztikus polinomja megegyezik, ezért a sajátértékeik is megegyeznek.*

2.6. Tétel. (Sajátalterek) *Az A mátrix egy λ sajátértékéhez tartozó sajátvektorai és a nullvektor együtt a $\mathbb{C}^n$ egy alterét alkotják. Ezt a λ sajátértékhez tartozó sajátaltérnek nevezzük és általában S_λ-val jelöljük.*

2.7. Definíció. (Multiplicitás) Ha a λ k-szoros gyöke a karakterisztikus egyenletnek, akkor azt mondjuk, hogy λ az A mátrix *k-szoros sajátértéke*, vagy hogy λ *algebrai multiplicitása k*. A λ *geometriai multiplicitása* pedig a λ sajátértékhez tartozó altér dimenziója.

2.8. Tétel. *Egy sajátérték geometriai multiplicitása nem lehet nagyobb, mint az algebrai multiplicitása.*

Bizonyítás. Legyen az A mátrix egy λ_0 sajátértékének geometriai multiplicitása g. Ez azt jelenti, hogy az S_{λ_0} sajátaltérnek van g vektorból álló bázisa. Ezek legyenek a $\mathbf{v_1}, \ldots, \mathbf{v_g}$ vektorok. Egészítsük ki ezeket a $\mathbb{C}^n$ egy bázisára a $\mathbf{v_{g+1}}, \ldots, \mathbf{v_n}$ vektorokkal. Ha ezekből a független vektorokból írunk fel egy $B = [\mathbf{v_1}, \ldots, \mathbf{v_g}, \mathbf{v_{g+1}}, \ldots, \mathbf{v_n}]$ mátrixot, akkor ez invertálható lesz. Írjuk fel ezt a mátrixot $B = (X|Y)$ alakban úgy, hogy X a $\mathbf{v_1}, \ldots, \mathbf{v_g}$ vektorokból álló blokk. Mivel ezek a vektorok a λ_0 sajátértékhez tartozó sajátvektorok, ezért $AX = \lambda_0 X$. Ezután a B inverzét is blokkokra bontjuk $\begin{pmatrix} Z \\ W \end{pmatrix}$ alakban, ahol Z az első g sorból áll. Ekkor a $B^{-1}B = I$ összefüggést felírhatjuk ezekkel a blokkmátrixokkal.

$$B^{-1}B = \begin{pmatrix} Z \\ W \end{pmatrix} \begin{pmatrix} X & Y \end{pmatrix} = \begin{pmatrix} ZX & ZY \\ WX & WY \end{pmatrix} = \begin{pmatrix} I_g & 0 \\ 0 & I_{n-g} \end{pmatrix} \tag{2}$$

Ebből pedig leolvasható, hogy $WX = 0$, $ZY = 0$, $ZX = I_g$, $WY = I_{n-g}$. Ezeket felhasználva

$$B^{-1}AB = \begin{pmatrix} Z \\ W \end{pmatrix} A \begin{pmatrix} X & Y \end{pmatrix} = \begin{pmatrix} ZAX & ZAY \\ WAX & WAY \end{pmatrix} = \begin{pmatrix} \lambda_0 I_g & ZAY \\ 0 & WAY \end{pmatrix}, \tag{3}$$

mert $ZAX = Z\lambda_0 X = \lambda_0 ZX = \lambda_0 I_g$ és $WAX = W\lambda_0 X = \lambda_0 WX = 0$. Az A és a $B^{-1}AB$ mátrix karakterisztikus polinomja megegyezik, ezért

$$\det A = \det(B^{-1}AB) = \begin{vmatrix} \lambda_0 I_g - \lambda I_g & ZAY \\ 0 & WAY - \lambda I_{n-g} \end{vmatrix} \tag{4}$$

Belátható, hogy a $\begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$ alakú mátrix determinánsa $\det A \cdot \det C$, ezért $k_A(\lambda) = (\lambda_0 - \lambda)^g \det(WAY - xI_{n-g})$. Ez pedig azt jelenti, hogy a karakterisztikus polinomnak λ legalább g-szeres gyöke, tehát az algebrai multiplicitása legalább g. $\square$

2.9. Definíció. (Jordan-blokk, Jordan-alakú mátrix) A $J_{\lambda,m}$ $m\times m$-es mátrixot *λ-hoz tartozó Jordan-blokknak* nevezzük, ha főátlójában λ-k, fölötte 1-esek, az összes többi helyen pedig 0-k szerepelnek, tehát az alakja

$$J_{\lambda,m} = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & \dots & 0 & 0 \\ 0 & 0 & \lambda & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{pmatrix}. \tag{5}$$

Egy mátrix *Jordan-alakú*, ha diagonálisában Jordan-blokkok szerepelnek, a többi elem pedig 0.

2.10. Tétel. (Jordan) *Minden $A \in \mathbb{C}^{n\times n}$ mátrixhoz létezik olyan $S \in \mathbb{C}^{n\times n}$ invertálható mátrix, hogy $J = S^{-1}AS$ Jordan-alakú, és ez a blokkok sorrendjétől eltekintve egyértelmű. Azaz minden λ és m esetén egyértelműen meghatározott, hogy hány darab $J_{\lambda,m}$ blokk szerepel a J mátrixban.*

Egy mátrix hatványozása akkor a legegyszerűbb, ha diagonális. Tudjuk, hogy nem minden mátrix hozható diagonális alakra, viszont Jordan-normálalakja minden mátrixnak van. Ezért fontos az alábbi teljes indukcióval bizonyítható tétel.

2.11. Tétel. *Ha k egy természetes szám, akkor a $J_{\lambda,m}$ k-adik hatványa a következő.*

$$\begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \dots & 0 \\ 0 & 0 & \lambda & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & \lambda \end{pmatrix}^k = \begin{pmatrix} \lambda^k & \binom{k}{1}\lambda^{k-1} & \binom{k}{2}\lambda^{k-2} & \dots & \binom{k}{m}\lambda^{k-m} \\ 0 & \lambda^k & \binom{k}{1}\lambda^{k-1} & \dots & \binom{k}{1}\lambda^{k-1} \\ 0 & 0 & \lambda^k & \dots & \binom{k}{m-2}\lambda^{k-(m-2)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \binom{k}{1}\lambda^{k-1} \\ 0 & 0 & 0 & \dots & \lambda^k \end{pmatrix} \tag{6}$$

2.12. Megjegyzés. Ha a λ sajátérték abszolút értéke kisebb mint 1, akkor $k \to \infty$ esetén a hatványok sorozata a $\mathbf{0}$ mátrixhoz tart.

2.13. Definíció. (Spektrálsugár) Az A mátrix $\rho(A)$ *spektrálsugara* a legnagyobb abszolútértékű sajátértékének abszolútértéke.

2.1. Gersgorin-körök

Ha a sajátértékeket direkt módon szeretnénk számolni, a karakterisztikus egyenletből, akkor felmerülhetnek problémák. Például, ha a karakterisztikus egyenlet négynél nagyobb fokú, akkor direkt megoldóképlet nem létezik. Szerencsére sok esetben nincs szükségünk a pontos értékre, elég, ha csak becsüljük őket, vagy iterációs lépésekkel közelítjük. Az alábbi módszerrel nagyon egyszerű számításokkal becslések adhatók a mátrix sajátértékeire nézve a komplex számsíkon.

2.14. Definíció. (Gersgorin-körök) Tekintsük az $A \in \mathbb{C}^{n\times n}$ mátrixot. Az a_{ii} középpontú, $\sum\limits_{j=1, j\neq i}^{n} |a_{ij}| (i = 1, \ldots, n)$ sugarú K_i köröket az A mátrix *Gersgorin-köreinek* nevezzük.

2.15. Tétel. *Az A mátrix sajátértékei az $\bigcup_{i=1,\ldots,n} K_i$ halmazban vannak.*

Ismeretes ennek a tételnek egy kibővített verziója is, az úgynevezett második Gersgorin-tétel, ami a sajátértékek elhelyezkezkedésére még pontosabb becslést adhat.

2.16. Tétel. *Ha a Gersgorin-körök közül s darab diszjunkt a többi körtől, akkor pontosan s darab sajátérték van ezen s darab kör uniójában.*

2.17. Megjegyzés. A tétel speciálisabb formája, valós elemű mátrix esetén, ha bármely két kör diszjunkt, akkor minden körbe pontosan egy sajátérték esik. Ebből pedig az következik, hogy ha egy mátrix bármely két Gersgorin-köre diszjunkt, akkor a mátrix minden sajátértéke valós, mert ha egy sajátérték komplex, akkor a konjugáltjának is a körön belül kellene lennie.

2.2. Hatványmódszer

Ahogy korábban említettem, sokszor elegendő, ha iterációs módszert alkalmazunk, amivel a sajátértékek tetszőlegesen pontos közelítését adhatjuk meg. Ebben a fejezetben az egyik legismertebb ilyen módszert mutatom be, a hatványmódszert.

Tegyük fel, hogy az $A \in \mathbb{R}^{n\times n}$ mátrix diagonalizálható, és létezik egy λ_1 sajátértéke, ami egyszeres, és abszolútértékben a legnagyobb. Az, hogy egy mátrixnak mikor létezik ilyen speciális sajátértéke, és milyen további

tulajdonságokkal rendelkezik, a későbbiek során kerül bemutatásra. A sajátértékeket a következőképpen fogom indexelni.

$$|\lambda_1| > |\lambda_2| \geq \ldots \geq |\lambda_n| \tag{7}$$

$\mathbf{v_i}$ pedig legyen a λ_i-hez tartozó sajátvektor. A diagonizálhatóság miatt föltehető, hogy a $\mathbf{v_i}$-k lineárisan függetlenek.

A módszer alapötlete a következő. Ha a mátrix diagonalizálható, akkor minden $\mathbf{x} \in \mathbb{R}^n$ vektor felírható $\mathbf{x} = \alpha_1\mathbf{v_1} + \alpha_2\mathbf{v_2} + \ldots + \alpha_n\mathbf{v_n}$ alakban. Ekkor tetszőleges $k \in \mathbb{Z}$, $k \geq 0$ esetén teljesül az, hogy

$$\begin{aligned} A^k\mathbf{x} &= \alpha_1 A^k\mathbf{v_1} + \alpha_2 A^k\mathbf{v_2} + \ldots + \alpha_n A^k\mathbf{v_n} = \\ &= \alpha_1\lambda_1^k\mathbf{v_1} + \alpha_2\lambda_2^k\mathbf{v_2} + \ldots + \alpha_n\lambda_n^k\mathbf{v_n} \end{aligned} \tag{8}$$

Ekkor a λ_1^k-nal való osztás után, ha $k \to \infty$, azt kapjuk, hogy

$$\frac{1}{\lambda_1^k}A^k\mathbf{x} = \alpha_1\mathbf{v_1} + \alpha_2\left(\frac{\lambda_2}{\lambda_1}\right)^k\mathbf{v_2} + \ldots + \alpha_n\left(\frac{\lambda_n}{\lambda_1}\right)^k\mathbf{v_n} \to \lambda_1\mathbf{v_1}, \tag{9}$$

mert $(\lambda_i/\lambda_1)^k \to 0$, ha $1 < i \leq n$. Tehát $A^k\mathbf{x}$ iránya a $\mathbf{v_1}$ sajátvektor irányához tart. Ha $\lambda_1 \neq \pm 1$, akkor a keletkező vektorok hossza vagy nullához, vagy végtelenhez fog tartani, ezért célszerű az iteráció minden lépésében normálni a kapott vektort. Így kapjuk a módszer végső alakját, amit a következő tétel foglal össze.

2.18. Tétel. (Hatványmódszer) *Ha $A \in \mathbb{R}^{n\times n}$ és az $\mathbf{y_0} \in \mathbb{R}^n$ kezdővektorra teljesül, hogy $\mathbf{v_1^T}\mathbf{y_0} \neq 0$ és $\|\mathbf{y_0}\| = 1$, ahol a $\|.\|$ az euklideszi normát jelöli, akkor a következő rekurzióval kapott ν_k számok a λ_1 sajátértéket közelítik.*

$$\begin{aligned} \mathbf{x_k} &= A\mathbf{y_{k-1}} \\ \mathbf{y_k} &= \frac{\mathbf{x_k}}{\|\mathbf{x_k}\|} \\ \nu_\mathbf{k} &= \mathbf{y_k^T}A\mathbf{y_k} \end{aligned} \tag{10}$$

3. Pozitív mátrixok

3.1. Definíció. (Pozitív és nemnegatív mátrix) Egy $A \in \mathbb{R}^{n\times n}$ mátrix *pozitív*, ha minden eleme pozitív, és *nemnegatív*, ha minden eleme nemnegatív.

Jelölés: $A > 0$ és $A \geq 0$.

Érdekes kérdés lehet, hogy mit tudunk ezeknek a speciális mátrixoknak a sajátértékeiről. A továbbiakban erről fogok írni, majd bemutatom, hogy a matematikának ezen része nemcsak szép, hanem hasznos is.

3.1. Perron-tétel

Először olyan mátrixokat fogok vizsgálni, ahol a szigorú egyenlőtlenség teljesül. Ezekre alkalmazható az a tétel, ami Oskar Perron nevéhez fűződik. Ő ugyanis 1907-ben csak pozitív mátrixokra vonatkozóan írta le a tételt, majd 1912-ben Ferdinand Georg Frobenius terjesztette ki nemnegatív mátrixokra is.

3.2. Tétel. (Perron) *Legyen $P \in \mathbb{R}^{n\times n}$ pozitív mátrix. Ekkor a P mátrixnak van egy $\lambda(P)$ kitüntetett sajátértéke amelyre igazak az alábbi állítások:*

(i) $\lambda(P)$ pozitív, és van olyan $\lambda(P)$-hez tartozó $\mathbf{h}$ sajátvektor, melynek komponensei pozitívak

$$P\mathbf{h} = \lambda(P)\mathbf{h}, \ \textit{ahol} \ \ \mathbf{h} > 0. \tag{11}$$

(ii) $\lambda(P)$ egyszeres sajátérték.

(iii) P minden más κ sajátértéke abszolútértékben kisebb, mint $\lambda(P)$. Így tehát $\lambda(P) = \rho(P)$.

(iv) Ha $\mathbf{v}$ nemnegatív sajátvektora P-nek, akkor $\mathbf{v} = \mu\mathbf{h}$, valamely $\mu > 0$ valós számra.

Bizonyítás. Jelölje $p(P)$ azon $\lambda \in \mathbb{R}$, $\lambda \geq 0$ számok halmazát, amelyek esetén létezik olyan $\mathbf{x}$ vektor, amelyre $\mathbf{x} \geq 0$, $\mathbf{x} \neq \mathbf{0}$, és

$$P\mathbf{x} \geq \lambda\mathbf{x}. \tag{12}$$

Bizonyítsunk be először egy erre vonatkozó lemmát!

3.3. Lemma. *$P > 0$ esetén:*

(i) $p(P)$ nem üres, és tartalmaz pozitív számot.

(ii) $p(P)$ korlátos.

(iii) $p(P)$ zárt.

Bizonyítás. Legyen $\mathbf{x} > 0$ tetszőleges vektor. Mivel $P > 0$, így $P\mathbf{x} > 0$. Ebből látszik, hogy elég kis λ szám esetén a $p(P)$ definíciójában szereplő (12) egyenlőtlenség teljesül. Ez igazolja a lemma (i) részét.

Most normáljuk az $\mathbf{x}$ vektort úgy, hogy $\sum \mathbf{x}_i = 1$ teljesüljön. Jelölje ξ a csupa egyesből álló sorvektort. Ekkor, ha balról beszorozzuk a (12) egyenlőtlenséget a ξ vektorral, akkor a következőt kapjuk:

$$\xi P\mathbf{x} \geq \lambda\xi\mathbf{x} = \lambda. \tag{13}$$

Jelölje b a ξP vektor legnagyobb komponensét, b tehát nem más, mint P oszlopösszegei közül a legnagyobbik. Ekkor $b\xi \geq \xi P$. Ebből következik, hogy

$$b\xi\mathbf{x} \geq \xi P\mathbf{x} \geq \lambda, \tag{14}$$

és itt $\xi\mathbf{x} = 1$ miatt $b\xi\mathbf{x} = b$. Tehát $b \geq \lambda$. Így $p(P)$ korlátos, ez pedig az állítás (ii) része.

Az (iii) rész bizonyításához tegyük föl, hogy λ_k egy sorozat $p(P)$ elemeiből, amelyre teljesül az, hogy $\lambda_k \to \lambda$, és $\mathbf{x_k}$ olyan vektor, melyre

$$P\mathbf{x}_k \geq \lambda_k\mathbf{x}_k, \text{ ahol } \mathbf{x}_k \geq 0. \tag{15}$$

Feltehető, hogy az $\mathbf{x}_k$ vektorokra szintén igaz, hogy $\xi\mathbf{x}_k = 1$. Ekkor az $\mathbf{x}_k$ vektorok az $\mathbb{R}^n$ tér egy korlátos és zárt részhalmazát alkotják. Így egy $\mathbf{x_{ik}}$ részsorozat egy $\mathbf{x} \geq 0$ vektorhoz tart, ami szintén normált, és λ_{ik} pedig a λ értékhez tart. Ekkor $Px_{ik} \geq \lambda_{ik}x_k$ miatt $P\mathbf{x} \geq \lambda\mathbf{x}$. Ez azt jelenti, hogy $\lambda \in p(P)$, vagyis $p(P)$ zárt és ezzel beláttuk az (iii) részt is. $\square$

Térjünk vissza a Perron-tétel bizonyításához. A fenti lemmára azért van szükség, mert ha $p(P)$ zárt és korlátos, akkor van maximuma, ami az (i) rész miatt pozitív. Nevezzük el ezt az elemet $\rho'(P)$-nek. Most azt fogjuk megmutatni, hogy ez a P mátrix sajátértéke. A $p(P)$ halmaz definíciója szerint, létezik egy $\mathbf{h}$ nemnegatív vektor, amelyre

$$P\mathbf{h} \geq \rho'(P)\mathbf{h}, \text{ ahol } \mathbf{h} \geq 0. \tag{16}$$

A továbbiakban azt kell belátnunk, hogy itt egyenlőség teljesül. Indirekt módon tegyük fel, hogy létezik olyan k komponens, amelyre

$$\begin{aligned} &(P\mathbf{h})_k > \rho'(P)h_k, \\ &(P\mathbf{h})_i \geq \rho'(P)h_i, \text{ ha } \ i \neq k. \end{aligned} \tag{17}$$

Definiáljunk egy $\mathbf{h}' = \mathbf{h} + \epsilon\mathbf{e_k}$ vektort, ahol $\epsilon > 0$ szám, és az $\mathbf{e_k}$ vektor k-adik komponense 1, a többi pedig 0. Ezt helyettesítsük be a (16) formulába. Ekkor $(P\mathbf{h}')_k > (P\mathbf{h})_k > \rho'(P)h_k$ miatt, ha $\epsilon > 0$ elég kicsi, még az is igaz lesz, hogy

$$\begin{aligned} &(P\mathbf{h})_k > \rho'(P)h_k + \rho'(P)\epsilon = \rho'(P)(\mathbf{h}')_k, \text{ így} \\ &(P\mathbf{h}')_k > \rho'(P)h'_k. \end{aligned} \tag{18}$$

Az i-edik komponensben pedig, $i \neq k$-ra

$$(P\mathbf{h}')_i > (P\mathbf{h})_i \geq (\rho'(P)\mathbf{h})_i = (\rho'(P)\mathbf{h}')_i. \tag{19}$$

Tehát elég kis ϵ esetén

$$P\mathbf{h}' > \rho'(P)\mathbf{h}'. \tag{20}$$

Mivel itt szigorú egyenlőtlenség teljesül, ezért a $\rho'(P)$ helyére $(\rho'(P) + \delta)$-t írhatunk, elég kis δ esetén. Ezzel pedig ellentmondásra jutottunk, mivel $\rho'(P)$ maximális volt, így beláttuk, hogy $\rho'(P)$ a P mátrix sajátértéke, és $\mathbf{h}$ pedig a hozzá tartozó nemnegatív sajátvektor. A továbbiakban nevezzük a $\rho'(P)$ értéket λ_{max}-nak, mert ez jobban tükrözi azt, hogy ez a maximális abszolútértékű sajátérték.

A (i) rész bizonyításához szükséges még belátni, hogy $\mathbf{h}$ pozitív. Mivel $P > 0$, és $\mathbf{h} \geq 0$, valamint $h \neq 0$ miatt $P\mathbf{h} > 0$, és így $P\mathbf{h} = \lambda_{max}\mathbf{h}$-ból azt kapjuk, hogy $\mathbf{h} > 0$. Ezzel a tétel (i) részét beláttuk.

Most mutassuk meg, hogy λ_{max} geometriai multiplicitása 1, azaz a hozzá tartozó sajátaltér dimenziója 1. Legyen $\mathbf{h}$ egy pozitív sajátvektor a λ_{max} sajátértékkel. Indirekt módon tegyük fel, hogy létezik $\mathbf{y} \in S_{\lambda_{max}}$, $\mathbf{y} \neq 0$, és $\mathbf{y}$ független $\mathbf{h}$-tól. Tehát $\mathbf{h} \pm c\mathbf{y} \neq 0$ tetszőleges $c \in \mathbb{R}$ szám esetén, és elég kis c választással $\mathbf{h} \pm c\mathbf{y} > 0$. Viszont ha növeljük a c értéket, akkor vagy $\mathbf{h} + c\mathbf{y}$ vagy $\mathbf{h} - c\mathbf{y}$ egyik komponense 0-vá fog válni úgy, hogy a többi komponens is nemnegatív lesz. Ez pedig ellentmond annak, amit az előbb beláttunk, hogy a λ_{max}-hoz tartozó nemnegatív sajátvektor pozitív.

A (ii) rész bizonyításához még be kell látnunk, hogy λ_{max} algebrai multiplicitása is 1. Tehát a karakterisztikus polinomnak nem osztója az $(\mathbf{x} - \lambda_{max})^2$. Ehhez ismét egy lemmát fogunk használni.

3.4. Lemma. *Ha $(x-\lambda)^2 \mid k_A(x)$ valamely $A \in T^{n\times n}$ tetszőleges test feletti mátrixra és $\mathbf{v}$ egy λ-hoz tartozó sajátvektor, akkor létezik egy $\mathbf{u} \neq 0$ vektor, ami független $\mathbf{v}$-től és $A\mathbf{u} = \lambda\mathbf{u} + c\mathbf{v}$ teljesül egy c konstansra.*

Bizonyítás. Írjuk föl az A-hoz tartozó lineáris transzformációt egy olyan bázisban, melynek első eleme $\mathbf{v}$. Azt kapjuk, hogy $A \sim \begin{pmatrix} \lambda & \star \\ \mathbf{0} & B \end{pmatrix}$ és a kifejtési tételből $k_A(x) = (\lambda - x)\cdot k_B(x)$. Így λ sajátértéke B-nek, tehát

$$A \sim \begin{pmatrix} \lambda & \star & \star \\ \mathbf{0} & \lambda & \star \\ \mathbf{0} & \mathbf{0} & C \end{pmatrix}. \tag{21}$$

Ez pedig azt jelenti, hogy A hasonlít egy ilyen alakú mátrixhoz, vagyis létezik olyan bázis, aminek eleme $\mathbf{v}$ és valamilyen $\mathbf{u}$, amelyekre $A\mathbf{v} = \lambda\mathbf{v}$ és $A\mathbf{u} = \lambda\mathbf{u} + c\mathbf{v}$. $\square$

Most azt fogjuk belátni, hogy a λ_{max} sajátértékhez nem létezik ilyen $\mathbf{u}$ vektor. Ehhez tegyük föl, hogy $\mathbf{u}$ mégis ilyen. Ekkor $P\mathbf{u} = \lambda_{max}\mathbf{u} + c\mathbf{h}$, ahol $\mathbf{h}$ továbbra is a λ_{max}-hoz tartozó pozitív sajátvektor, $\mathbf{u}$ és $\mathbf{v}$ pedig lineárisan függetlenek. Azt is tudjuk, hogy $c \neq 0$, mert $\dim S_{\lambda(P)} = 1$, ezért az $\mathbf{u}$ nem sajátvektor. Nyilván $\mathbf{u}$ megválasztható úgy is, hogy $c > 0$ teljesüljön. Ez után $\mathbf{u}$ helyére írjuk be az $\mathbf{u}' = \mathbf{u} + b\mathbf{h}$ vektort.

$$P\mathbf{u}' = P\mathbf{u} + Pb\mathbf{h} = \lambda_{max}\mathbf{u} + c\mathbf{h} + \lambda_{max}b\mathbf{h} = \lambda_{max}\mathbf{u}' + c'\mathbf{h} \tag{22}$$

Ha $c, b > 0$, akkor $c' > 0$ is teljesül, továbbá elég nagy $b > 0$ választással $\mathbf{u}' > 0$ is elérhető. De mivel, ha $\mathbf{u}' > 0$, és $P\mathbf{u}' = \lambda_{max}\mathbf{u}' + c'\mathbf{h}$ esetén $P\mathbf{u}' > \lambda_{max}\mathbf{u}'$, így $\exists\delta > 0$, melyre $P\mathbf{u}' \geq (\lambda_{max} + \delta)\mathbf{u}'$. Ez viszont ellentmond annak, hogy λ_{max} a $p(P)$ maximális eleme. Így beláttuk, hogy a domináns sajátérték egyszeres sajátértéke P-nek, ami a (ii) állítás volt.

Most igazoljuk a tétel (iii) állítását. Tegyük föl, hogy $\kappa \neq \lambda(P)$ sajátértéke a P mátrixnak, és $\mathbf{y}$ egy hozzá tatozó sajátvektor. Ekkor $P\mathbf{y} = \kappa\mathbf{y}$, tehát

$$\sum_{j=1}^{n} p_{ij}y_j = \kappa y_i, \forall 1 \leq i \leq n. \tag{23}$$

Alkalmazva a háromszög-egyenlőtlenséget azt kapjuk, hogy

$$\sum_{j=1}^{n} p_{ij}|y_j| \geq |\sum_{j=1}^{n} p_{ij}y_j| = |\kappa|\cdot|y_i|, \forall 1 \leq i \leq n. \tag{24}$$

Ezért a $|\kappa| \in p(P)$, tehát $|\kappa| \leq \lambda_{max}$. Ha egyenlőség áll fenn köztük, akkor a korábbi érvelések egyike miatt az az $\mathbf{y}'$ vektor, melyre $y'_i = |y_i|$ a λ_{max}-hoz tartozó sajátvektor, ezért $\dim S_{\lambda max} = 1$ miatt $\mathbf{y}'$ a $\mathbf{h}$ többszöröse, tehát $|y_i| = ch_i$. Mivel ekkor a háromszög-egyenlőtlenség egyenlőséggel teljesül, ez azt jelenti, hogy az $\mathbf{y}$ vektor komponensei azonos állásúak, azaz

$$\operatorname{Arg} y_i = \operatorname{Arg} y_j, \forall i, j \tag{25}$$

tehát $\mathbf{y} = \epsilon \mathbf{y}' = \epsilon c \mathbf{h}$, ahol $c > 0$ és $\epsilon \in \mathbb{C}$, $|\epsilon| = 1$. Ez viszont azt jelenti, hogy $\kappa = \lambda(P)$, ami pontosan (iii) állítást igazolja.

Végezetül megmutatjuk, hogy a domináns sajátértéken kívül nincs olyan sajátérték, aminek van nemnegatív sajátvektora. Tegyük föl, hogy $\mathbf{f} \geq 0$ a P mátrix sajátvektora, tehát $P\mathbf{f} = \mu\mathbf{f}$. Azt kell megmutatnunk, hogy $\mu = \lambda_{max}$. Legyen $\mathbf{g} > 0$ vektor a P^T domináns sajátértékéhez tartozó pozitív sajátvektora, tehát a transzponáltjára igaz, hogy $\mathbf{g}^T P = \lambda_{max}\mathbf{g}^T$. Ekkor

$$(\mathbf{g}, P\mathbf{f}) = (\mathbf{g}, \mu\mathbf{f}) = \mu(\mathbf{g}, \mathbf{f}) \tag{26}$$

Viszont ha másképp alakítjuk át a szorzatot, akkor azt kapjuk, hogy

$$(\mathbf{g}, P\mathbf{f}) = (P^T\mathbf{g}, \mathbf{f}) = (\lambda_{max}\mathbf{g}, \mathbf{f}) = \lambda_{max}(\mathbf{g}, \mathbf{f}). \tag{27}$$

Tehát $\mu(\mathbf{g}, \mathbf{f}) = \lambda(P)(\mathbf{g}, \mathbf{f})$, amiből az következik, hogy

$$\mu = \lambda(P), \tag{28}$$

mert ha $\mathbf{g} > 0$ és $\mathbf{f} \geq 0$, valamint $\mathbf{f} \neq 0$, akkor $(\mathbf{g}, \mathbf{f}) \neq 0$. Ezzel beláttuk a (iv) állítást is. $\square$

3.5. Definíció. (Domináns sajátérték) A fenti tulajdonságokkal rendelkező sajátértéket *domináns sajátértéknek* nevezzük. A $\mathbf{h}$ pedig a hozzá tartozó *domináns sajátvektor*, amire igaz, hogy $\sum h_i = 1$.

3.6. Megjegyzés. Ez megegyezik azzal a speciális λ_1 sajátértékkel, amiről a hatványmódszer kapcsán esett szó.

3.2. A Perron-tétel alkalmazása

3.7. Definíció. (Sztochasztikus mátrix) Egy $A \in \mathbb{R}^{n \times n}$ mátrix *sztochasztikus*, ha minden eleme nemnegatív, és igaz rá az, hogy

$$\sum_{i=0}^{n} a_{ij} = 1, \text{ ahol } j = 1, \ldots, n, \tag{29}$$

tehát az egy oszlopban álló elemeinek összege 1.

3.8. Tétel. *Ha az S egy pozitív sztochasztikus mátrix, akkor teljesülnek rá az alábbiak.*

(i) Az S domináns sajátértéke 1, tehát $\lambda(S) = 1$.

(ii) Ha $\mathbf{x} \geq 0$ tetszőleges vektor, akkor

$$\lim_{k\to\infty} S^k \mathbf{x} = c\mathbf{h}, \tag{30}$$

ahol $\mathbf{h}$ a domináns sajátvektor és c pedig egy $\mathbf{x}$-től függő pozitív konstans.

Bizonyítás. Tudjuk, hogy ha S pozitív, akkor S^T is szintén pozitív. Egy korábbi tételből azt is tudjuk, hogy ha λ sajátértéke az S^T mátrixnak, akkor az S-nek is sajátértéke. Könnyen látható, hogy a csupa egyesből álló oszlopvektor az S^T sajátvektora 1 sajátértékkel és mivel S^T pozitív, ezért a csupa egyes vektor a domináns sajátvektora, és így az 1 domináns sajátértéke. Ezzel be is láttuk a tétel (i) állítását.

Az (ii) bizonyításához legyen az S mátrix hasonló J-hez, ami egy Jordan-alakú mátrix. Tudjuk, hogy az 1 domináns sajátérték, tehát a J úgy néz ki hogy,

$$J = \begin{pmatrix} 1 & 0 \\ 0 & J' \end{pmatrix}. \tag{31}$$

A J' szintén Jordan-alakú és sajátértékei abszolútértékben biztosan kisebbek mint 1, a Perron-tétel miatt. Ezért ha hatványozzuk a J mátrixot, akkor azt kapjuk, hogy

$$\lim_{k\to\infty} J^k = \lim_{k\to\infty} \begin{pmatrix} 1 & 0 \\ 0 & (J')^k \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & \mathbf{0} \end{pmatrix}, \tag{32}$$

valamint, ha jobbról szorozzuk az $\mathbf{x}$ vektorral, akkor

$$\lim_{k\to\infty} J^k \mathbf{x} = \begin{pmatrix} x_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{33}$$

A hasonlóság azt jelenti, hogy $J = T^{-1}ST$, ahol $T = (\mathbf{h}, \ldots)$ mátrix. Ezt átalakítva kapjuk, hogy

$$\begin{aligned} S &= TJT^{-1}, \text{ és} \\ S^k\mathbf{x} &= TJ^kT^{-1}\mathbf{x}. \end{aligned} \tag{34}$$

Legyen $T^{-1}\mathbf{x} = \begin{pmatrix} c \\ \star \\ \vdots \\ \star \end{pmatrix}$. Ekkor az előzőekből következik, hogy

$$\lim_{k\to\infty} J^k T^{-1}\mathbf{x} = \begin{pmatrix} c \\ 0 \\ \vdots \\ 0 \end{pmatrix} \tag{35}$$

és így azt kapjuk, hogy

$$\lim_{k\to\infty} TJ^kT^{-1}\mathbf{x} = \lim_{k\to\infty} S^k\mathbf{x} = T\begin{pmatrix} c \\ 0 \\ \vdots \\ 0 \end{pmatrix} = c\mathbf{h}. \tag{36}$$

Mivel $\mathbf{x} \geq 0$ és $S > 0$, ezért $c \geq 0$. Jelölje ismét $\xi = \begin{pmatrix} 1 & \dots & 1 \end{pmatrix}$ vektort. Ekkor $\xi S = \xi$, mivel S sztochasztikus mátrix, és így

$$\xi S^N\mathbf{x} = \xi\mathbf{x} = \sum_{i=1}^{n} x_i. \tag{37}$$

A szorzás folytonossága miatt pedig

$$\lim_{n\to\infty} \xi S^N\mathbf{x} = \lim_{n\to\infty} \sum_{i=1}^{n} x_i = \xi(c\mathbf{h}) = c. \tag{38}$$

Tehát $\mathbf{x} \neq 0$ és $\mathbf{x} \geq 0$ esetén teljesül, hogy a c egy pozitív konstans. $\square$

3.3. Általánosítás nemnegatív mátrixokra

A tételkör alkalmazásának szempontjából fontos megvizsgálni, hogy a Perron-tétel állításaiból mi tartható meg nemnegatív mátrix esetén. Könnyen észrevehető, hogy egyes nemnegatív mátrixoknál nem teljesül minden állítás. Ilyen például az egységmátrix, aminek az 1 lehetne domináns sajátértéke, de nem igaz, hogy egyszeres sajátérték, mert a karakterisztikus egyenletnek

kétszeres gyöke. Másik példa az $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ mátrix. Ennek a -1 is, és a $+1$ is sajátértéke, továbbá nincs pozitív elemű sajátvektora.

Most tehát lássuk, hogy milyen feltételeknek kell teljesülnie egy adott mátrixra, hogy létezzen a Perron-tételben definiált sajátértéke, illetve a tétel melyik részei nem igazak, ha létezik 0 eleme is a mátrixnak.

3.9. Definíció. (Irreducibilis és reducibilis mátrix) Azt mondjuk, hogy az $A \in \mathbb{R}^{n\times n}$, $A \geq 0$ mátrix *irreducibilis*, ha $\nexists P$ permutáció mátrix, amire

$$P^{-1}AP = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}, \tag{39}$$

ahol A_{11}, A_{22} négyzetes mátrixok. Egyébként A *reducibilis*.

3.10. Definíció. Legyen $A \in \mathbb{R}^{n\times n}$, $A \geq 0$ mátrix *gráfja* G_A, ami egy n csúcsú irányított gráf, és ha

$$a_{ij} = \begin{cases} > 0 & \text{, akkor } \exists i \to j\text{, és} \\ 0 & \text{, akkor } \nexists i \to j. \end{cases} \tag{40}$$

3.11. Definíció. (Erősen összefüggő gráf) Egy irányított gráf *erősen összefüggő*, ha minden v_i és v_j csúcshoz létezik irányított út v_i-ből v_j-be.

3.12. Definíció. (Primitív mátrix) Az $A \in \mathbb{R}^{n\times n}$ mátrix *primitív*, ha van olyan m szám, amire $A^m > 0$.

3.13. Tétel. *A nemnegatív $A \in \mathbb{R}^{n\times n}$ mátrixra ekvivalensek az alábbi állítások:*

(i) Az A mátrix irreducibilis.

(ii) G_A erősen összefüggő gráf.

(iii) $(A+I)^n > 0$, tehát $(A+I)$ primitív.

3.14. Megjegyzés. A tétel elemi kombinatorikai eszközökkel könnyen belátható. A bizonyítás alapgondolata, hogy az (iii) állítás jelentése gráfos megközelítésben az, hogy a $G_{(A+I)}$ gráf minden csúcspárja között van n hosszú út. Az $(A+I)$-hez tartozó gráf pedig annyiban különbözik az eredeti G_A-tól, hogy minden csúcshoz kerül egy hurok él. Ezért ha a G_A erősen összefüggő,

akkor G_{A+I}-ben bármely két csúcs közt van pontosan n hosszúságú út. Megfordítva, ha G_{A+I}-ben bármely két csúcs közt van n hosszúságú út, akkor G_{A+I} és így nyilván G_A is erősen összefüggő.

Ezt a tételt a későbbiekben fogjuk alkalmazni.

3.15. Tétel. *Ha a P mátrix primitív, akkor P-nek van domináns sajátértéke.*

Bizonyítás. Ha P^m pozitív, akkor a Perron-tétel miatt létezik domináns sajátértéke. Ezt jelölje μ és $\mathbf{v}$ legyen egy hozzá tartozó pozitív sajátvektor. Tehát

$$P^m\mathbf{v} = \mu\mathbf{v}. \tag{41}$$

Ekkor

$$P^{m+1}\mathbf{v} = P^m P\mathbf{v} = \mu P\mathbf{v}, \tag{42}$$

amiből az következik, hogy $P\mathbf{v}$ is a P^m mátrix μ sajátértékhez tartozó sajátvektora, ezért $P\mathbf{v} = c\mathbf{v}$, ugyanis $\dim S_\mu = 1$. Ha P^m hatványaira ismételten alkalmazzuk a fenti számítást, akkor azt kapjuk, hogy $P^m\mathbf{v} = c^m\mathbf{v}$, azaz $c^m = \mu$. Mivel $P\mathbf{v} = c\mathbf{v}$, tudjuk, hogy c egy pozitív, valós szám, ezért a $\mu^{\frac{1}{n}}$ valós. Így azt kapjuk, hogy a $c = \mu^{\frac{1}{n}}$ a P mátrix domináns sajátértéke. $\square$

3.16. Tétel. (Frobenius) *Ha $A \neq 0$ nemnegatív irreducibilis mátrix, akkor a $\rho(A)$ spektrálsugár az A mátrix egy sajátértéke és igazak rá az alábbi állítások.*

(i) *$\rho(A)$ pozitív és, van olyan $\rho(A)$-hoz tartozó $\mathbf{v}$ sajátvektor, melynek komponensei pozitívak*

$$A\mathbf{v} = \rho(A)\mathbf{v}, \ \textit{ahol} \ \ \mathbf{v} \geq 0. \tag{43}$$

(ii) *$\rho(A)$ egyszeres sajátérték.*

(iii) *Az A mátrix bármely κ sajátértékére igaz, hogy $|\kappa| \leq \rho(A)$, és ha k darab olyan sajátérték van, amire $|\kappa| = \rho(A)$, akkor ezek a sajátértékek:*

$$\kappa_j = (\cos j\frac{2\pi}{k} + i\sin j\frac{2\pi}{k})\rho(A), \ \textit{ahol} \ \ j = 0, \ldots, k-1. \tag{44}$$

Bizonyítás. A bizonyításhoz használjuk a Perron-tétel bizonyításánál alkalmazott halmazt, tehát $p(A)$ legyen azon $\lambda \in \mathbb{R}$, $\lambda \geq 0$ számok halmaza, amelyek esetén létezik olyan $\mathbf{x} \geq 0$, $\mathbf{x} \neq \mathbf{0}$ vektor, amelyre

$$A\mathbf{x} \geq \lambda \mathbf{x}. \tag{45}$$

A továbbiakban használjuk az alábbi jelöléseket.

$$\rho'(A) = \sup p(A), \tag{46}$$

és $\mathbf{v} \geq 0$-ra

$$\rho'_{\mathbf{v}}(A) = \sup\{r \geq 0 \mid A\mathbf{v} \geq r\mathbf{v}\}. \tag{47}$$

Vegyük észre, hogy $\rho'(A) = \sup\{\rho'_{\mathbf{v}}(A) : \mathbf{v} \geq 0\}$.

Legyen $\epsilon > 0$ esetén $A_\epsilon = A + \epsilon J$, ahol $J = \begin{pmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{pmatrix}$. Ekkor az A_ϵ pozitív mátrix, ezért tudjuk, hogy $\rho(A_\epsilon) = \rho'(A_\epsilon)$, és $\exists \mathbf{v}_\epsilon > 0$, amire $A_\epsilon \mathbf{v}_\epsilon = \rho(A_\epsilon)\mathbf{v}_\epsilon$.

Ha $\epsilon \leq 1$, akkor a fenti jelöléseket használva azt kapjuk, hogy

$$\rho'_{\mathbf{v}}(A) \leq \rho'_{\mathbf{v}}(A_\epsilon) \leq \rho'_{\mathbf{v}}(A_1), \tag{48}$$

amiből az következik, hogy

$$\sup\{\rho'_{\mathbf{v}}(A) : \mathbf{v} \geq 0\} \leq \sup\{\rho'_{\mathbf{v}}(A_\epsilon) : \mathbf{v} \geq 0\} \leq \sup\{\rho'_{\mathbf{v}}(A_1) : \mathbf{v} \geq 0\}, \tag{49}$$

azaz,

$$\rho'(A) \leq \rho'(A_\epsilon) \leq \rho'(A_1), \tag{50}$$

és itt az is igaz, hogy $\rho'(A_\epsilon) = \rho(A_\epsilon)$ és $\rho'(A_1) = \rho(A_1)$.

Válasszunk most egy 0-hoz tartó sorozatot, például az $\frac{1}{n}$ sorozatot. Ekkor $A_{\frac{1}{n}}$ korlátos, és $\rho(A_{\frac{1}{n}})$ az előbbiek szerint szintén korlátos. Tudjuk, hogy minden n-hez választhatunk egy olyan $\mathbf{v}_{\frac{1}{\mathbf{n}}} > 0$ vektort, hogy $|\mathbf{v}_{\frac{1}{n}}| = 1$, és $A_{\frac{1}{n}}\mathbf{v}_{\frac{1}{\mathbf{n}}} = \rho(A_{\frac{1}{n}})\mathbf{v}_{\frac{1}{\mathbf{n}}}$. Ez a $\mathbf{v}_{\frac{1}{\mathbf{n}}}$ szintén korlátos, ezért $\exists n_i$ részsorozat, amire

$$\begin{aligned} \rho(A_{\frac{1}{n_i}}) &\to \rho, \text{ ahol } \rho \neq 0, \\ \mathbf{v}_{\frac{1}{\mathbf{n_i}}} &\to \mathbf{v}, \text{ ahol } \mathbf{v} \geq 0, |\mathbf{v}| = 1 \\ A_{\frac{1}{n_i}} &\to A. \end{aligned} \tag{51}$$

Tehát az

$$A_{\frac{1}{n_i}}\mathbf{v}_{\frac{1}{\mathbf{n_i}}} = \rho(A_{\frac{1}{n_i}})\mathbf{v}_{\frac{1}{\mathbf{n_i}}} \tag{52}$$

egyenlőség miatt $A\mathbf{v} = \rho\mathbf{v}$, azaz ρ sajátértéke A-nak. A ρ értékről pedig tudjuk, hogy

$$\rho \leq \rho(A) \leq \rho'(A) \leq \rho'(A_{\frac{1}{n_i}}), \tag{53}$$

és mivel $\rho'(A_{\frac{1}{n_i}}) \to \rho$, ezért $\rho = \rho(A) = \rho'(A)$. Ezzel pedig beláttuk a tétel (i) állítását.

Az (ii) rész bizonyításához használjuk a (3.13) tételt. Az irreducibiliásból következik, hogy $(A + I)$ primitív. Egy korábbi tételből pedig tudjuk, hogy létezik domináns sajátértéke pozitív sajátvektorral.

Ha $A\mathbf{v} = \lambda\mathbf{v}$, akkor $(A+I)\mathbf{v} = (\lambda+1)\mathbf{v}$, tehát ha $\mathbf{v}$ sajátvektora A-nak, akkor sajátvektora az $(A + I)$ mátrixnak, és ha λ sajátértéke A-nak, akkor $(\lambda + 1)$ sajátértéke $(A + I)$-nek.

Ha a $(\lambda + 1)$ az $(A + I)$ domináns sajátértéke, ekkor a $(\lambda + 1)$ egyszeres sajátérték, tehát a λ is egyszeres sajátértéke A-nak. Ez a gondolatmenet azt is mutatja, hogy ha A irreducibilis, akkor a domináns sajátértékhez, $\rho(A)$-hoz tartozó nemnegatív sajátvektor határozottan pozitív. Ezzel erősíthetjük az (i) állítást.

Most lássuk be a (iii) állítást. Legyen $|\kappa| = \rho(A)$ és $\mathbf{v}$ egy hozzá tartozó sajátvektor. Ekkor úgy is megfogalmazhatjuk az állítást, hogy $\kappa = \rho(A)c$, ahol $c \in \mathbb{C}$, $|c| = 1$. Megmutatjuk, hogy ekkor A hasonló cA-hoz.

Jelölje $\mathbf{w} = \begin{pmatrix} |v_1| \\ \vdots \\ |v_n|, \end{pmatrix}$ ekkor $v_i = d_i|v_i|$, azaz $v_i = d_i w_i$, ahol $d_i \in \mathbb{C}$, $|d_i| = 1$.

A fenti jelöléseket használva kapjuk, hogy

$$\rho(A)cd_iw_i = \sum_j a_{ij}(d_jw_j), \tag{54}$$

valamint

$$\rho'(A)w_i = \rho(A)w_i = |\rho(A)cd_iw_i| = |\sum_j a_{ij}d_jw_j| \leq a_{ij}w_j, \forall 1 \leq i \leq n. \tag{55}$$

A korábbiakhoz hasonlóan azt kapjuk, hogy $\mathbf{w}$ sajátvektor $\rho(A)$-hoz, és ha A irreducibilis, akkor a (ii) bizonyítása utáni megjegyzés értelmében $\mathbf{w} > 0$.

Így egyenlőség van a háromszög-egyenlőtlenségben, ami azt jelenti, hogy $a_{ij}d_j$ egyirányú $\forall i$-re. Ugyanakkor $\sum_j a_{ij}d_jw_j$ egyirányú cd_i-vel. Ezért $d_i^{-1}a_{ij}d_j$ egyirányú c-vel $\forall ij$-re, és $|d_i^{-1}a_{ij}d_j| = a_{ij}$, így $d_i^{-1}a_{ij}d_j = ca_{ij}, \forall i, j$.

Ebből pedig az következik, hogy $D = \begin{pmatrix} d_i & & \\ & \ddots & \\ & & d_n \end{pmatrix}$ diagonális mátrixra $D^{-1}AD = cA$, azaz A hasonló a cA mátrixhoz.

Tehát ha $c\rho(A)$ sajátérték, akkor a $c^2\rho(A)$ is sajátérték, amiből következik, hogy $c^3\rho(A)$ is az, és így tovább. De tudjuk, hogy véges számú sajátérték van, ezért $\exists k \neq l$, hogy $c^k\rho(A) = c^l\rho(A)$, azaz $c^k = c^l$, tehát a s egységgyök.

Legyen most $\epsilon = \frac{\kappa}{\rho(A)}$ olyan, hogy ϵ szöge minimális a spektrálkörön lévő κ sajátértékek között. Az előbb beláttuk, hogy a sajátértékek halmaza invariáns az ϵ egységgyökökkel való szorzásra, ezért

$$\epsilon = \cos\frac{2\pi}{k} + i\sin\frac{2\pi}{k} \tag{56}$$

alakú, és a spektrálkörön lévő sajátértékek halmaza megegyezik az

$$\{\epsilon^j \cdot \rho(A) \mid 0 \leq j \leq k-1\} \tag{57}$$

halmazzal. Ezzel pedig beláttuk az (iii) állítást is. □

3.17. Megjegyzés. Láthatjuk, hogy az (i) állítás bizonyításakor nem használtuk azt a feltételt, hogy a mátrix irreducibilis. Ez az állítás tehát igaz minden nemnegatív mátrixra.

A tétel (iii) részéről is belátható, hogy teljesül minden nemnegatív mátrix esetén, viszont bonyolultabbá teszi a bizonyítást, ha az irreducibilitásról szóló feltételt elhagyjuk.

4. A Page Rank algoritmus

A korábban tárgyalt Perron-Frobenius tételkörnek van egy olyan alkalmazása, amit valószínűleg mindannyian ismerünk, és napi szinten használunk is. Ez a Google keresőmotor egyik legfontosabb eleme, a Page Rank algoritmus, ami egy adott kulcsszóra érkező találatokat rendezi sorba úgy, hogy egy értéket rendel minden oldalhoz. Ezt a számot is Page Ranknek nevezik, és megadja, hogy az adott oldal milyen szerepet tölt be az internet hivatkozási gráfjában.

4.1. Keresőmotorok a Page Rank előtt

A 90-es évek elején a World Wide Web Worm nevű első keresőmotor körülbelül 110 ezer oldal között tudott keresni és ekkor még a felhasználók köre is jóval kisebb volt. Naponta átlagosan 1500 keresést hajtott végre. Ezek a számok viszont az évek során rohamosan nőttek. Pár év alatt az oldalak száma 100 millióra nőtt, és naponta körülbelül 20 millió lekérdezés érkezett be. Ez a növekedés azt eredményezte, hogy a találatok között sok olyan oldal is szerepelt, ami kevésbé volt releváns a felhasználók által keresett témában. Ezért megnőtt az igény egy olyan keresőre, ami képes felvenni a lépést a világháló folyamatos fejlődésével és növekedésével.

4.2. Motiváció

Larry Page és Sergey Brin 1998-ban fejlesztették ki a Google keresőmotort. Munkájuk során számos elvárásnak kellett eleget tenniük ahhoz, hogy keresőjük a jövőben is képes legyen kiszolgálni a felhasználókat. Többek között figyelniük kellett arra, hogy a folyamatosan növekedő hálózat adatai mindig frissen rendelkezésre álljanak, és hogy ezeket hatékonyan tárolják. Szerették volna, hogy egy olyan kereső készüljön, ami a későbbiekben is fejleszthető, és amibe az új technológiákat is könnyen be lehet építeni. A dolgozatom szempontjából legfontosabb a találatok minőségének javítása volt. Ugyanis egy átlagos felhasználó körülbelül 10-20 oldalt hajlandó átböngészni a keresése során, ezért fontos, hogy a több millió találat közül a legfontosabbak kerüljenek előre a találatok listázása során. Felmerülhet, hogy a látogatások számának segítségével állapítsuk meg ezt a rangsort, de sok olyan oldal van, ahol nem rögzítik ezt az adatot.

4.3. A Page Rank alapötlete

A fejlesztők úgy vélték, hogy jól tükrözi a felhasználók szerinti fontossági sorrendet, ha az oldalak az egymás közti linkstruktúra segítségével vannak sorba állítva. Emögött az az elképzelés áll, hogy ha valaki hivatkozást helyez el a saját oldalán, akkor ezt azért teszi, mert szerinte hasznos információt tartalmaz a másik oldal. Ezért ezt úgy kell elképzelni, mintha minden oldal a hivatkozásaival egy szavazatot adna le, viszont ezek nem egyenértékűek: a fontos oldalak szavazatai többet érnek. Az pedig, hogy egy oldal mikor fontos, az azon múlik, hogy hány, és mennyire fontos oldal mutat rá.

Larry Page és Sergey Brin megjelentettek egy cikket a keresőmotorról az 1998-as induláskor. Ebben írnak arról, hogy az általuk megalkotott algoritmus működésének alapja a következő. Képzeljünk el egy az interneten keresgélő („szörföző") személyt, aki véletlenszerűen választ az oldalakon található hivatkozások közül és így halad egyik oldalról a másikra, viszont amikor ezt megunja, akkor véletlenszerűen átugrik egy másik oldalra. Az algoritmus ezt a működést követve rendez sorba, és a Page Rank érték pont azt adja meg, hogy mennyi a valószínűsége annak, hogy a véletlen szörföző egy adott oldalra érkezik.

A cikkben a következőket javasolják. Tegyük föl, hogy az A-val jelölt weboldalra k darab másik oldal mutat. Jelölje ezeket $T_1, \ldots, T_k$, és az oldalakon lévő hivatkozások számát jelölje $C(T_i)$. Ha $PR(A)$ jelöli az A oldal Page Rank értékét, akkor első közelítésben $PR(A)$ a $PR(T_i)$ értékekből tevődik össze, természetesen mindegyik $PR(T_i)$ értéket le kell osztani $C(T_i)$-vel. Technikai okok miatt, (például mert a szörföző úgy is továbbhaladhat, hogy nem egy hivatkozást követ, hanem bekér egy konkrét oldalt) ezt a képletet célszerű egy úgynevezett csillapító tényezővel módosítani (jelölje ezt d). Így a következő képletet kapjuk:

$$PR(A) = \frac{(1-d)}{n} + d\left(\frac{PR(T_1)}{C(T_1)} + \ldots + \frac{PR(T_k)}{C(T_k)}\right). \tag{58}$$

4.4. Az egyszerű Page Rank algoritmus működése

Ha a weblapok hálózatát úgy képzeljük el, mint egy irányított gráfot, ahol a csúcsok az oldalak, és az élek a hivatkozások két oldal közt, akkor fel tudunk írni egy olyan mátrixot, ami jól reprezentálja az oldalak közti függőségeket.

Ezért definiálunk n darab laphoz egy $n \times n$-es mátrixot úgy, hogy ha az

i-edik oldalon m darab hivatkozás található, akkor

$$A_{ij} = \begin{cases} \frac{1}{m} & \text{ha } j\text{-ről mutat link } i\text{-re} \\ 0 & \text{egyébként.} \end{cases} \tag{59}$$

Az oldalak Page Rank értékét pedig tároljuk egy vektorban, amit jelöljünk $\mathbf{p}$-vel, és $\mathbf{p} = (x_1, \ldots, x_n)^T$. Az első oldal Page Rank értéke tehát x_1, a másodiké x_2 és így tovább. Az oldalak fontosságát, azaz a Page Rank értékét egy iterációs eljárással kapjuk meg. Kezdetben minden oldal rangját egyenlőnek vesszük, ezért

$$\mathbf{p_0} = \begin{pmatrix} \frac{1}{n} \\ \vdots \\ \frac{1}{n} \end{pmatrix}. \tag{60}$$

Rekurzívan pedig úgy adjuk meg a $\mathbf{p_k}$ értékét, hogy

$$\mathbf{p_k} = A\mathbf{p_{k-1}}. \tag{61}$$

azaz

$$\mathbf{p_k} = A^k\mathbf{p_0}. \tag{62}$$

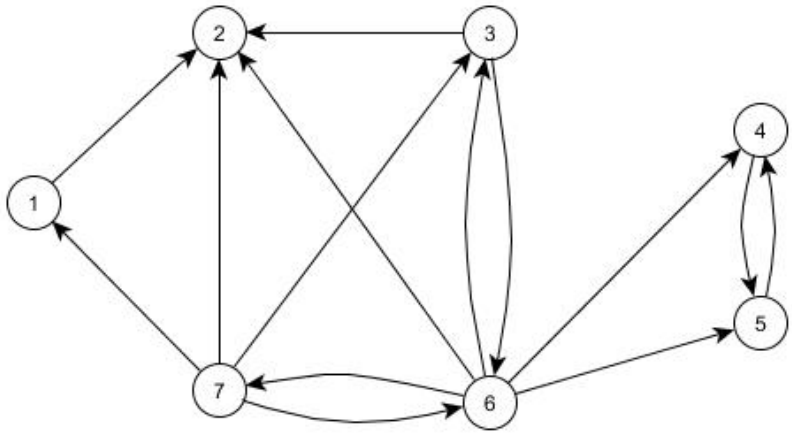
Az egyes oldalak Page Rank értékét így a $\mathbf{p_k}$ vektorsorozat határértéke fogja megadni.

Az algebrai bevezető részben felmerült, hogy a Jordan-blokkok hatványozása egyszerűbb, így az A mátrix hatványozása is hatékonyabb lehet, ha a mátrixot Jordan-normálalakra hozzuk.

4.5. Példa

Az alábbi példával szeretném szemléltetni az algoritmus működését. Természetesen ez a kis gráf a teljes hálózattal összemérhetetlen, mégis jól látható, hogy hogyan működik az algoritmus, és hogy milyen problémákba ütközhetünk a rangsorolás során.

Lássunk tehát hét weblapot, amik hivatkoznak egymásra. Ezeket a kapcsolatokat az alábbi gráf szemlélteti.



Az ehhez tartozó, korábban definiált mátrix a következő.

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{4} \\ 1 & 0 & \frac{1}{2} & 0 & 0 & \frac{1}{5} & \frac{1}{4} \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{5} & \frac{1}{4} \\ 0 & 0 & 0 & 0 & 1 & \frac{1}{5} & 0 \\ 0 & 0 & 0 & 1 & 0 & \frac{1}{5} & 0 \\ 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & \frac{1}{4} \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{5} & 0 \end{pmatrix} \tag{63}$$

A kiinduló vektor pedig $\mathbf{p_0} = \begin{pmatrix} \frac{1}{7} \\ \vdots \\ \frac{1}{7} \end{pmatrix}$.

A korábbi képletbe való behelyettesítés után a következő értékeket kapjuk:

$$\begin{aligned}
\mathbf{p_1}^T &= \begin{pmatrix} 0{,}04 & 0{,}28 & 0{,}06 & 0{,}17 & 0{,}17 & 0{,}11 & 0{,}03 \end{pmatrix}, \\
\mathbf{p_2}^T &= \begin{pmatrix} 0{,}01 & 0{,}10 & 0{,}03 & 0{,}19 & 0{,}19 & 0{,}04 & 0{,}02 \end{pmatrix}, \\
\mathbf{p_3}^T &= \begin{pmatrix} 0{,}00 & 0{,}03 & 0{,}01 & 0{,}20 & 0{,}20 & 0{,}02 & 0{,}01 \end{pmatrix}, \\
\mathbf{p_4}^T &= \begin{pmatrix} 0{,}00 & 0{,}02 & 0{,}01 & 0{,}20 & 0{,}20 & 0{,}01 & 0{,}00 \end{pmatrix}, \\
\mathbf{p_5}^T &= \begin{pmatrix} 0{,}00 & 0{,}01 & 0{,}00 & 0{,}21 & 0{,}21 & 0{,}00 & 0{,}00 \end{pmatrix}.
\end{aligned}$$

Láthatjuk, hogy néhány iteráció után a 4. és 5. oldal begyűjti az összes szavazatot. Ezen oldalak Page Rank értéke 0,21, a többié viszont már az 5.

iteráció után nagyon kicsi. A második oldal értéke is, pedig a gráfot tekintve azt várnánk, hogy begyűjti a szavazatokat, mivel nincs rajta hivatkozás. Ennek az az oka, hogy a 2. oszlopban csupa 0 szerepel, és ezért a mátrix nem sztochasztikus. Ezt tudjuk úgy orvosolni, hogy a 2. sor 2. elemét átírjuk 1-re, ami azt jelenti, hogy ha a 2. lapra érünk akkor 1 valószínűséggel ott is maradunk.

Ezt a mátrixot behelyettesítve ténylegesen azt kapjuk amit várunk, hogy a 2. weblap gyűjti be a legtöbb szavazatot.

$$\begin{aligned}
\mathbf{p_1}^T &= \begin{pmatrix} 0{,}04 & 0{,}42 & 0{,}06 & 0{,}17 & 0{,}17 & 0{,}11 & 0{,}03 \end{pmatrix}, \\
\mathbf{p_2}^T &= \begin{pmatrix} 0{,}01 & 0{,}52 & 0{,}03 & 0{,}19 & 0{,}19 & 0{,}04 & 0{,}02 \end{pmatrix}, \\
\mathbf{p_3}^T &= \begin{pmatrix} 0{,}00 & 0{,}55 & 0{,}01 & 0{,}20 & 0{,}20 & 0{,}02 & 0{,}01 \end{pmatrix}, \\
\mathbf{p_4}^T &= \begin{pmatrix} 0{,}00 & 0{,}57 & 0{,}01 & 0{,}20 & 0{,}20 & 0{,}01 & 0{,}00 \end{pmatrix}, \\
\mathbf{p_5}^T &= \begin{pmatrix} 0{,}00 & 0{,}58 & 0{,}00 & 0{,}21 & 0{,}21 & 0{,}00 & 0{,}00 \end{pmatrix}.
\end{aligned}$$

4.6. Problémák az egyszerű működéssel

A fenti algoritmus egyelőre még csak azt foglalja magában, hogy a szörföző mindig választ véletlenszerűen egy oldalt és így halad oldalról oldalra. Ezzel viszont akad némi probléma.

Az egyik probléma akkor merül fel, ha van olyan oldal, amin nincs elhelyezve hivatkozás. (Az előbbi példában ilyen volt a 2. csúcs.) Az ilyen zsákutca jellegű oldalnak az A-ban csupa nullából álló oszlopa van, tehát az iterációban nem osztja szét a saját szavazatát, és mint az előbb láttuk, torzulnak az eloszlás értékei. Ha pedig azt föltételezzük, hogy a zsákutcából nem megyünk tovább, és mint a példánál tettük, a 0-s oszlopot módosítjuk azzal, hogy a főátló elemét 1-esnek választjuk, akkor az ilyen oldalak jogtalanul fognak előre kerülni a rangsorolásban.

Hasonló probléma lép fel, ha a gráf csúcsainak a halmaza erősen összefüggő, de egy valódi részhalmazból nem vezet ki él. Ez azt fogja eredményezni, hogy ez a részhalmaz szintén be fogja gyűjteni a szavazatokat, és a többi csúcs rangja 0 lesz egy idő után. Ezért van szükség az iteráció egy kibővített verziójára.

4.7. Kibővített Page Rank algoritmus

A előbb említett problémák miatt az algoritmusnak nem ez az egyszerű verziója került be a Google keresőmotorjába, hanem egy kibővített formája.

A zsákutca problémára az is lehet egy megoldás, hogy az A mátrix csupa nullából álló oszlopainak elemeit kicseréljük $\frac{1}{n}$-re. Így ez azt jelenti, hogy ha a szörföző egy olyan oldalra ér, amin nincs hivatkozás, akkor egyenlő valószínűséggel választ egy másik oldalt. Ebben az esetben a következő eredményeket kapjuk:

$$\begin{aligned}
\mathbf{p_1}^T &= \begin{pmatrix} 0{,}06 & 0{,}30 & 0{,}08 & 0{,}19 & 0{,}19 & 0{,}13 & 0{,}05 \end{pmatrix}, \\
\mathbf{p_2}^T &= \begin{pmatrix} 0{,}05 & 0{,}18 & 0{,}08 & 0{,}26 & 0{,}26 & 0{,}10 & 0{,}07 \end{pmatrix}, \\
\mathbf{p_3}^T &= \begin{pmatrix} 0{,}04 & 0{,}16 & 0{,}06 & 0{,}30 & 0{,}30 & 0{,}08 & 0{,}04 \end{pmatrix}, \\
\mathbf{p_4}^T &= \begin{pmatrix} 0{,}03 & 0{,}12 & 0{,}05 & 0{,}34 & 0{,}34 & 0{,}06 & 0{,}04 \end{pmatrix}, \\
\mathbf{p_5}^T &= \begin{pmatrix} 0{,}03 & 0{,}10 & 0{,}04 & 0{,}37 & 0{,}37 & 0{,}05 & 0{,}03 \end{pmatrix}.
\end{aligned}$$

Látjuk, hogy a 2. oldal szétosztotta a szavazatait, és ismét a 4. és 5. oldal lett a legmagasabb rangú.

A fenti ötletet átgondolva, feltételezhetjük, hogy a szörföző nem csak akkor ugrik át véletlenszerűen egy másik oldalra, amikor egy zsákutca oldalra ér, hanem bármikor megunhatja azt, hogy hivatkozások mentén böngészik, és véletlenül választhat egy másik weblapot. Az algoritmus bővítésével ezt a jelenséget is beépítették a fejlesztők, ezért vezették be a $0 < d < 1$ paramétert. Ilyenkor azt feltételezzük, hogy a szörföző d valószínűséggel követi a linkeket, $(1-d)$ valószínűséggel viszont ugrik egy véletlenszerűen választott oldalra. A Google alapításakor megjelent cikk szerint általában $d = 0{,}85$.

Ezekkel pedig úgy módosul a korábbi A mátrix, hogy

$$A' = (1-d)\frac{1}{n}E + dA, \tag{64}$$

ahol $E = \begin{pmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{pmatrix}$.

Így megoldottuk az irreducibilitás problémáját, azaz nem lesznek olyan halmazok, amiből nem vezet ki él.

Most térjünk vissza a korábbi példához. Ha a d értéke 0,85, akkor az

$$A' = 0{,}15\frac{1}{7}E + 0{,}85A. \tag{65}$$

Ha ezt a módosított A' mátrixot használva számolunk az iterációs lépésekben, akkor azt kapjuk hogy,

$$\begin{aligned}
\mathbf{p_1}^T &= \begin{pmatrix}0{,}05 & 0{,}26 & 0{,}08 & 0{,}17 & 0{,}17 & 0{,}11 & 0{,}05\end{pmatrix}, \\
\mathbf{p_2}^T &= \begin{pmatrix}0{,}03 & 0{,}12 & 0{,}05 & 0{,}18 & 0{,}18 & 0{,}06 & 0{,}04\end{pmatrix}, \\
\mathbf{p_3}^T &= \begin{pmatrix}0{,}02 & 0{,}08 & 0{,}03 & 0{,}18 & 0{,}18 & 0{,}04 & 0{,}02\end{pmatrix}, \\
\mathbf{p_4}^T &= \begin{pmatrix}0{,}02 & 0{,}06 & 0{,}02 & 0{,}17 & 0{,}17 & 0{,}03 & 0{,}02\end{pmatrix}, \\
\mathbf{p_5}^T &= \begin{pmatrix}0{,}01 & 0{,}04 & 0{,}02 & 0{,}16 & 0{,}16 & 0{,}02 & 0{,}02\end{pmatrix}.
\end{aligned}$$

4.8. Markov-láncok és a Page Rank

Most a problémát egy teljesen más szemszögből fogom vizsgálni, a valószínűségszámítás eszközeit használva. A fejlesztők is így fogalmazták meg cikkükben a Page Rank működését, mint egy valószínűségi eloszlást, ami megadja, hogy a szörföző milyen eséllyel látogat el egy adott oldalra.

4.1. Definíció. (Markov-tulajdonság, Markov-lánc) Ha az X_n, $n \in \mathbb{N}$ folyamat, ahol az X_n *valószínűségi változók* I értékkészlete véges halmaz. Azt mondjuk X_n-re, hogy *Markov-tulajdonságú*, ha igaz rá, hogy

$$P(X_{n+1} = x_{n+1} | X_n = x_n, \ldots, X_0 = x_0) = P(X_{n+1} = x_{n+1} | X_n = x_n) \quad (66)$$

bármely i esetén. Tehát ha az $(i+1)$-edik időpillanatbeli állapot csak az előző, i-edik időpontbeli állapottól függ.

A Markov-tulajdonságú folyamatokat *Markov-folyamatoknak* hívjuk. Az olyan véges Markov-folyamatot pedig, aminek állapottere diszkrét, *Markov-láncoknak* nevezzük.

Ez azt jelenti, hogy a folyamat egy bizonyos állapota teljesen magába foglalja az összes olyan információt, ami befolyásolhatja a jövőbeli változást.

A Page Rank is ilyen tulajdonságú, ugyanis az algoritmus szerint a szörfőzőt nem befolyásolja, hogy korábban melyik oldalakon járt, csak az, hogy az adott lépésben mely hivatkozások mentén tud továbbhaladni.

4.2. Definíció. (Átmenet-valószínűség) Egy Markov-lánc *átmenetvalószínűsége* az n-edik lépésben

$$p_{ij}(n) = P(X_n = j | X_{n-1} = i), \text{ ahol } i, j \in I. \quad (67)$$

A négyzetes P mátrixot, aminek az i. sor j. eleme a p_{ij} átmenet-valószínűség, *átmenet-valószínűség mátrixnak* nevezzük.

4.3. Megjegyzés. Ez egy sztochasztikus mátrix, hiszen 1 a valószínűsége annak, hogy a folyamat a következő lépésben az értékkészlet elemei közül az egyikben lesz.

4.4. Definíció. Egy Markov-lánc *kezdeti eloszlása* az a vektor, amely megadja, hogy melyik állapotból milyen valószínűséggel indul ki a folyamat. Jelölje ezt a vektort $\mathbf{v_0}$.

4.5. Tétel. *Egy véges Markov-láncnak az eloszlása legyen* $\mathbf{v_n}$*, a kezdeti eloszlása* $\mathbf{v_0}$*, és az átmenet-valószínűség mátrixa pedig* P*. Ekkor*

$$\mathbf{v_n} = P^n \mathbf{v_0}. \tag{68}$$

Erre pedig tudjuk alkalmazni a pozitív sztochasztikus mátrixokra vonatkozó (3.7) tételt. Ugyanis, mint említettem, a P mátrix sztochasztikus. Igaz, hogy a tétel pozitív mátrixokra vonatkozik, de a Frobenius-tétel segítségével általánosítható, hogy nemnegatív mátrixok esetén is teljesüljön.

5. Egyéb rangsoroló algoritmusok

Ha már hálózati kereséssel és linkstruktúrát alkalmazó rangsorolásokkal foglalkozunk, akkor érdemes megemlíteni a kérdésfüggő algoritmusokat. Ezek abban különböznek a korábbitól, hogy amíg a Page Rank az összes oldalra meghatároz egy rangsort, ezek az algoritmusok csak a kérdéstől függően, az oldalak egy részére futnak le. Így persze kevesebb futási költséget igényelnek, viszont minden keresés alkalmával újraszámítódnak az értékek. Ezzel ellentétben a Page Rank csak egyszer fut le és a következő futásig, használhatóak a kapott értékek.

Két algoritmus ismeretes, a HITS és a SALSA. Ezekben két közös van, az egyik, a kéréstől függő részhalmaz előállításának módja, a másik pedig, hogy az oldalakat központi lapokra és linkgyűjteményekre osztják.

Egy kérdéshez tartozó oldalak halmazának meghatározása úgy történik, hogy veszünk egy alaphalmazt, ami egy hagyományos szöveges keresés első néhány találatából áll. Majd ezt bővítjük azokkal az oldalakkal, amik ezekre mutatnak, illetve amikre hivatkoznak az alaphalmazbeli oldalak. Így megkapjuk azt a részgráfot, aminek a csúcsait az algoritmusok rangsorolni fogják.

Az oldalak közti hivatkozásokat most is, a Page Rankhez hasonlóan, felfoghatjuk úgy mint szavazatokat, csak ez esetben kétféle szavazatot különböztetünk meg. Az egyik az $\mathbf{a}$, mint *authority*, ami az oldalra mutató élek, tehát a jó központi oldal mértéke, a másik pedig a $\mathbf{h}$, mint *hub*, az oldalról mutató élek, így a jó linkgyűjtemény mértéke.

A következő alfejezetben vázlatosan bemutatom a HITS algoritmus működését.

5.1. A HITS algoritmus

A HITS algoritmust 1998-ban Jon Kleinberg publikálta. Az linkstruktúrát leíró gráf szomszédsági mátrixa legyen A. Az $\mathbf{a}$, illetve $\mathbf{h}$ vektorok meghatározásához vezessünk be két jelölést. A $B(i) = \{k\colon k \to i\}$ az i. oldalra mutató lapokat jelenti, és a $C(i) = k\colon i \to k$ pedig azokat a lapokat jelöli, amikre az i. oldal mutat. Így minden i oldalra:

$$\begin{aligned} \mathbf{a}(i) &= \sum_{j \in B(i)} \mathbf{h}(j), \\ \mathbf{h}(i) &= \sum_{j \in C(i)} \mathbf{a}(j) \end{aligned} \tag{69}$$

tehát

$$\begin{aligned} \mathbf{a} &= A^T\mathbf{h}, \\ \mathbf{h} &= A\mathbf{a}, \end{aligned} \tag{70}$$

azaz

$$\begin{aligned} \mathbf{a} &= A^T A\mathbf{a}, \\ \mathbf{h} &= AA^T\mathbf{h}. \end{aligned} \tag{71}$$

Ez a modell leírja azt a működést, hogyha egy oldalra sok jó gyűjtő oldal mutat, akkor az valószínűleg egy hasznos központi oldal, és fordítva, egy oldal akkor lesz jó gyűjtő, ha az sok jó központi oldalra mutat.

Ezen vektorokat iteratívan a következőképpen számolhatjuk. Kezdetben

$$\mathbf{a_0} = \mathbf{h_0} = \begin{pmatrix} \frac{1}{n} \\ \vdots \\ \frac{1}{n} \end{pmatrix}, \tag{72}$$

majd minden lépésben

$$\begin{aligned} (a_i)_k &= \sum_{j \in B(i)} (h_j)_{k-1}, \\ (h_i)_k &= \sum_{j \in C(i)} (a_j)_{k-1}, \end{aligned} \tag{73}$$

a vektorok minden i. eleme esetén, azaz

$$\begin{aligned} \mathbf{a}_k &= A^T\mathbf{h}_{k-1}, \\ \mathbf{h}_k &= A\mathbf{h}_{k-1}. \end{aligned} \tag{74}$$

Az iteráció minden lépésénél normálni kell az így kapott vektorokat.

Kleinberg bizonyította, hogy az $\mathbf{a}$, és $\mathbf{h}$ vektorok iránya az $A^T A$, illetve a AA^T mátrixok domináns sajátértékhez tartozó, pozitív sajátvektor irányához tart.

Hivatkozások

[1] Alberto Peretti: *The importance of Perron-Frobenius Theorem in ranking problems*, 2014
http://dse.univr.it/home/workingpapers/wp2014n26.pdf

[2] Dr. Bodon Ferenc: *Adatbányászati algoritmusok*, 2010, 9. fejezet
http://www.cs.bme.hu/~bodon/magyar/adatbanyaszat/tanulmany/adatbanyaszat.pdf

[3] Faragó István és Horváth Róbert: *Numerikus Módszerek*, 2013
http://tankonyvtar.ttk.bme.hu/pdf/30.pdf

[4] Ivanyos Gábor előadásjegyzete,
http://math.bme.hu/~ig/alkalg/alkalg.pdf

[5] Jakabfy Tamás: *Gráfos rangsoroló algoritmusok a World Wide Web keresőiben*, Diplomamunka, 2002
http://web.cs.elte.hu/alkmatdiploma/jaksy.pdf

[6] Peter D. Lax: *Lineáris Algebra és alkalmazásai* Akadémiai Kiadó 2008, 16. fejezet

[7] Sergey Brin and Lawrence Page: *The Anatomy of a Large-Scale Hypertextual Web Search Engine*, 1998, 1–2. http://infolab.stanford.edu/~backrub/google.html

[8] Wettl Ferenc: *Lineáris algebra*, 2011
http://tankonyvtar.ttk.bme.hu/pdf/14.pdf