Matrične i tenzorske metode u analizi podataka

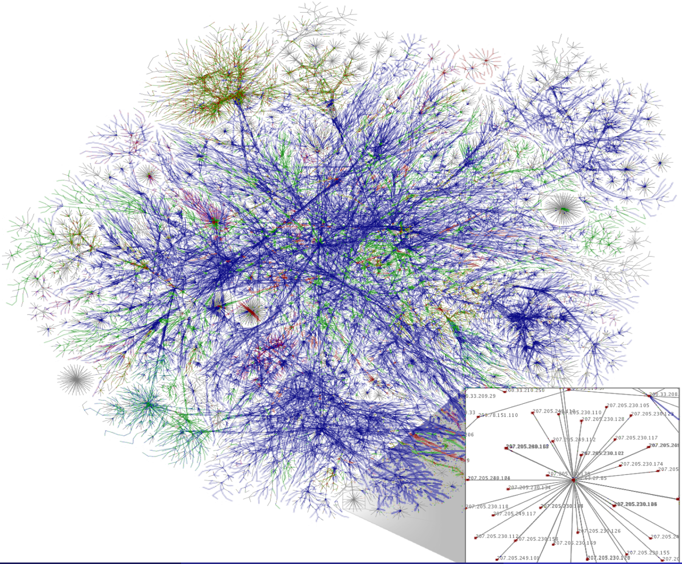
©Zlatko Drmač
PMF Matematički Odsjek, Zagreb

2024, Zagreb

Zagreb 2024

Sadržaj

Primjer: Internet tražilica Google i rangiranje mrežnih stranica

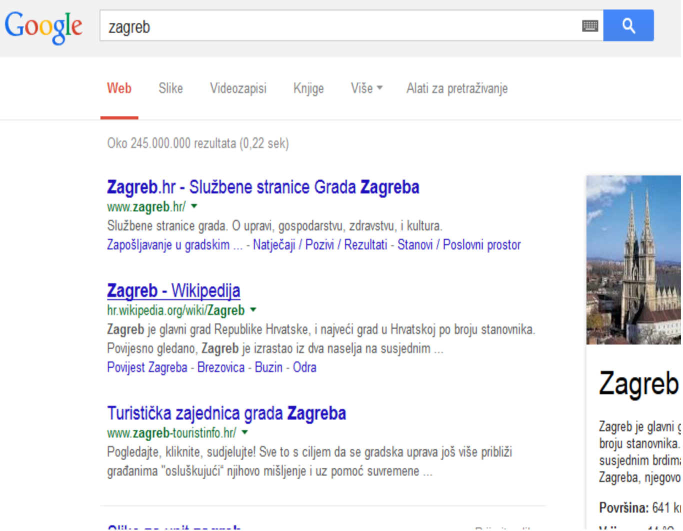


WWW je ogroman, googol=10^{100}

GB = Sorted on Google and Bing

BG = Sorted on Bing and Google

Svaka pretraga na WWW daje puuunoo rezultata



Nije prakično pregledavati milijune stranica nakon svake pretrage

- Nemoguće je pregledati milijune stranica da bi se došlo do kvalitetne informacije.
- Bilo bi idealno kada bi se stranice moglo poredati tako da prvo budu izlistane one koje su najbolje.
- Kako poredati 245000000 stranica jedne pretrage u 0.22 sekunde, počevši od najboljih? (Google obrađuje preko 70000 pretraga svake sekunde.)
- Ali, što u ovom kontekstu znači biti najbolji?
- Što god to značilo, kako to izvesti/implementirati da može efikasno funkcionirati?

Cijena? Sitnica.

THE $25,000,000,000* EIGENVECTOR THE LINEAR ALGEBRA BEHIND GOOGLE

KURT BRYAN† AND TANYA LEISE‡

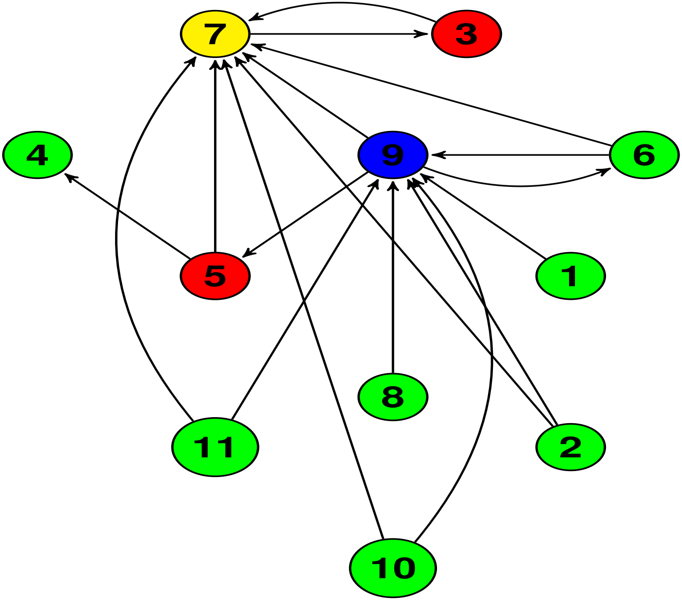
Abstract. Google's success derives in large part from its PageRank algorithm, which ranks the importance of webpages according to an eigenvector of a weighted link matrix. Analysis of the PageRank formula provides a wonderful applied topic for a linear algebra course. Instructors may assign this article as a project to more advanced students, or spend one or two lectures presenting the material with assigned homework from the exercises. This material also complements the discussion of Markov chains in matrix algebra. Maple and Mathematica files supporting this material can be found at www.rose-hulman.edu/~bryan.

Key words. linear algebra, PageRank, eigenvector, stochastic matrix

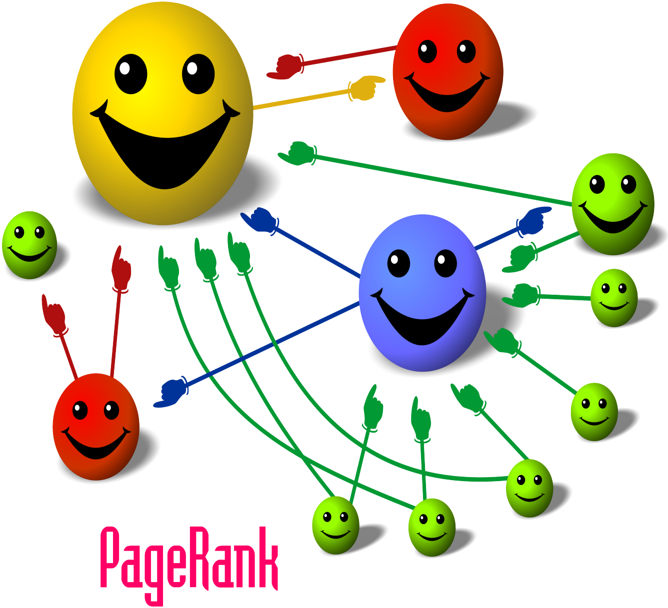
AMS subject classifications. 15-01, 15A18, 15A51

1. Introduction. When Google went online in the late 1990's, one thing that set it apart from other search engines was that its search result listings always seemed deliver the "good stuff" up front. With other search engines you often had to wade through screen after screen of links to irrelevant web pages that just happened to match the search text. Part of the magic behind Google is its PageRank algorithm, which quantitatively rates the importance of each page on the web, allowing Google to rank the pages and thereby present to the user the more important (and typically most relevant and helpful) pages first.

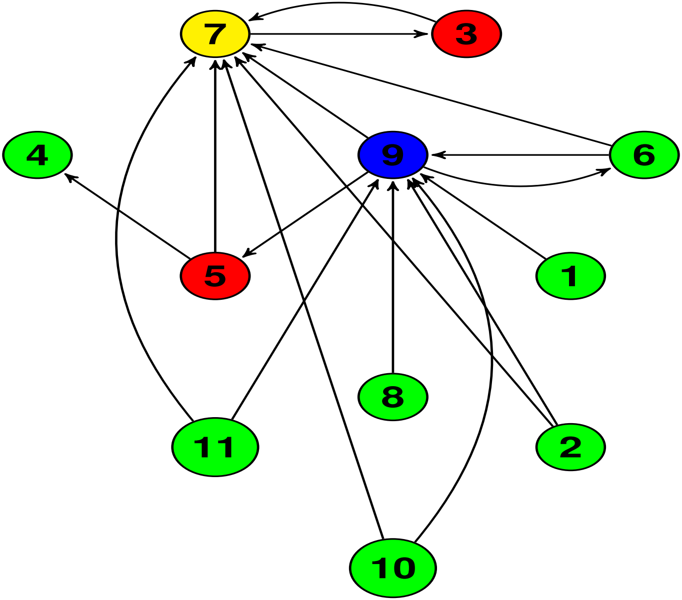
WWW = usmjereni graf



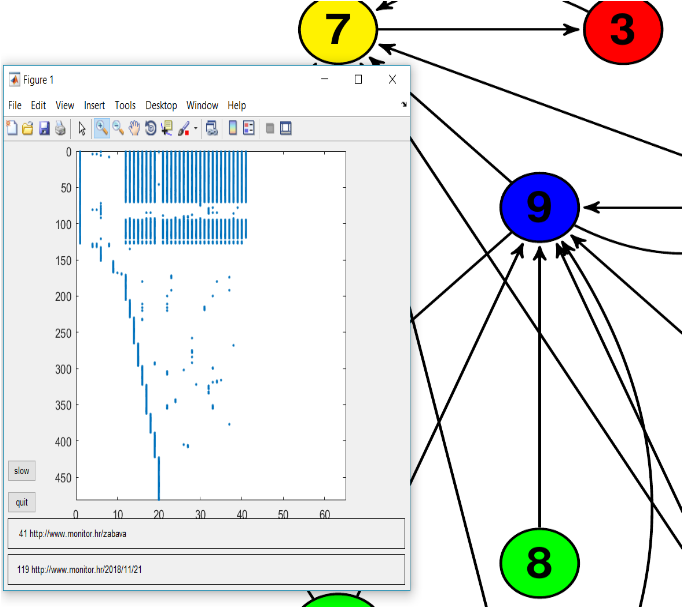
Google Page Rank



WWW = usmjereni graf



WWW crawler/spyder



usmjereni graf = matrica

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

$\mathbf{G_{ij} = 1}$ **ako vrh j ima link na vrh i**

Težine (važnosti) vrhova

Važnost i-tog vrha neka je $x_i \geq 0$.
Vrh i je $\boxed{\text{važniji od}}$ vrha j ako je $x_i \boxed{>} \mathbf{x_j}$.

- **Vrh je važan ako na njega pokazuju važni vrhovi.**
- **Ako neki vrh pokazuje na puno drugih vrhova, važnost takvih linkova je proporcionalno manja.**

Neka je n_j ukupan broj linkova iz vrha j, tj. (z)broj jedinica u j–tom stupcu matrice G.
Neka je L_i skup vrhova koji imaju link na i.

Sve sažeto u jednu formulu glasi:

$$x_i = \sum_{j \in L_i} \frac{1}{n_j} x_j$$

$Ax = 1 \cdot x$

$$x_i = \sum_{j \in L_i} \frac{1}{n_j} x_j, \quad i = 1, \ldots, n. \tag{1}$$

Ako je $n_j = 0$, onda je j-ti stupac nula, tj. j-ti vrh nema izlaznih linkova. U tom slučaju taj stupac od G npr. napunimo elementima 1. Nakon te korekcije, elemente stupca podijelimo sumom njegovih elemenata. Tako korigiranu G označimo s A. Stavimo

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad A = \begin{pmatrix} A_{11} & A_{12} & \ldots & A_{1n} \\ A_{21} & A_{22} & \ldots & A_{2n} \\ \vdots & \vdots & \ldots & \vdots \\ A_{n1} & A_{n2} & \ldots & A_{nn} \end{pmatrix}, \quad A_{ij} = \begin{cases} \frac{1}{n_j} & \text{ako } j \longrightarrow i \\ 0 & \text{inače} \end{cases}$$

Tada (1) glasi

$$x = Ax \ \text{ tj. } \ \mathbf{Ax = 1 \cdot x}$$

Stohastičke matrice

Definicija

Realna kvadratna matrica A je stohastička po stupcima (retcima) ako su joj svi elementi nenegativni i suma elemenata u svakom stupcu (retku) je jedan. Kažemo da je A dvostruko stohastička ako je stohastička i po stupcima i po retcima.

Propozicija

Svaka stupčano (retčano) stohastička matrica ima jedinicu za svojstvenu vrijednost.

Dokaz: Ako je A stupčano stohastička, onda je A^T retčano stohastička, i jedinica je njena svojstvena vrijednost jer vrijedi

$$\begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \vdots & \vdots & \dots & \vdots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^n A_{j1} \\ \sum_{j=1}^n A_{j2} \\ \vdots \\ \sum_{j=1}^n A_{jn} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = 1 \cdot \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = 1 \cdot e$$

Zbog $\det(\lambda I - A) = \det(\lambda I - A^T)$, A i A^T imaju iste svojstvene vrijednosti.

Pitanja, problemi

Dakle, u ovom modelu, rangiranje ovisi o komponentama svojstvenog vektora x matrice A, tj. rješenja problema $Ax = x$, $x \neq \mathbf{0}$; zapravo trebamo $x_i \geq 0$. Zašto bi A imala svojstveni vektor s nenegativnim elementima? Što ako A ima nekoliko lineanro nezavisnih svojstvenih vektora za $\lambda = 1$? Na primjer, sljededeća matrica A ima svojstvene vektore x, y za $\lambda = 1$:

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1/2 \\ 0 & 0 & 1 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad x = \begin{pmatrix} 1/2 \\ 1/2 \\ 0 \\ 0 \\ 0 \end{pmatrix}, y = \begin{pmatrix} 0 \\ 0 \\ 1/2 \\ 1/2 \\ 0 \end{pmatrix}$$

Onda je i $z = (3/4)x + (1/4)y = (3/8, 3/8, 1/8, 1/8, 0)^T$ svojstveni vektor, $Az = z$.
Dakle, naš ključni objekt u modelu nije jedinstveno određen i njegova svojstva ključna za rangiranje se mijenjaju na način koji oduzima smisao našem modelu. S druge strane, promjena x u αx s $\alpha > 0$ i $x_i \geq 0$ nije problematična ($x_i \geq x_j \Leftrightarrow \alpha x_i \geq \alpha x_j$)

Primjer

Višestrukost svojstvene vrijednosti (i problem za naš model) slijedi npr. iz blok dijagonalne strukture: Ako je A oblika

$$A = \begin{pmatrix} A_1 & 0 & 0 & 0 \\ 0 & A_2 & 0 & 0 \\ & & \ddots & \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & A_k \end{pmatrix}, \quad A_i \text{ stupčano stohastička; } A_i v^{(i)} = v^{(i)}$$

$$w^{(1)} = \begin{pmatrix} v^{(1)} \\ 0 \\ \vdots \\ 0 \end{pmatrix}, w^{(2)} = \begin{pmatrix} 0 \\ v^{(2)} \\ \vdots \\ 0 \end{pmatrix}, w^{(k)} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ v^{(k)} \end{pmatrix}$$

$$Aw^{(i)} = w^{(i)}, \;\; i = 1, \ldots k; \quad A(\sum_{j=1}^{k} \alpha_j w^{(j)}) = (\sum_{j=1}^{k} \alpha_j w^{(j)})$$

Tu je i problem vrhova iz kojih ne izlaze bridovi (linkovi) - tzv. *dangling nodes* kojima u matrici A odgovara nul stupac. Jednostavan *quick-fix* je takve stupce popuniti slučajnim nenegativnim vektorom sa sumom komponenti jedan. Napravimo to.

Modifikacija matrice A

Stavimo

$$S = \left(\frac{1}{n}\right)_{i,j=1}^{n} = ff^T, \ \ f = (\frac{1}{\sqrt{n}}, \ldots, \frac{1}{\sqrt{n}})^T$$

S je simetrična, ranga jedan, $Se = e$, $e = (1, \ldots, 1)^T$. Definiramo

$$M = M_\alpha = (1-\alpha)A + \alpha S, \ \ \alpha \in [0,1].$$

M je stupčano stohastička:

$$e^T M = (1-\alpha)e^T A + \alpha e^T S = (1-\alpha)e^T + \alpha e^T = e^T$$

$M_0 = A$; $M_1 = S$. Diskusija.
Za $\alpha > 0$ je $M_{ij} > 0$ za sve i, j (kažemo da je M pozitivna).
Pozitivnost matrice po elementima je svojstvo koje ima netrivijalne implikacije na strukturu svojstvenih vrijednosti i vektora.

Pozitivne matrice

Propozicija

Ako je M pozitivna (po elementima) i stupčano stohastička, onda svaki (realni) svojstveni vektor koji pripda svojstvenoj vrijednosti $\lambda = 1$ ima samo pozitivne ili samo negative komponente.

DOKAZ: Dakle imamo $x \neq 0$ i

$$x = Mx, \quad x_i = \sum_{j=1}^{n} M_{ij} x_j, \quad M_{ij} > 0.$$

Pretpostavimo da komponente od x nisu sve istog predznaka. Tada je

$$|x_i| = \left| \sum_{j=1}^{n} \overbrace{M_{ij}}^{>0} \underbrace{x_j}_{\pm} \right| < \sum_{j=1}^{n} M_{ij} |x_j|$$

(nejednakost je stroga zbog miješanih predznaka)

Nastavak dokaza Propozicije

Dakle, $|x_i| < \sum_{j=1}^n M_{ij}|x_j|$. Prosumiramo po svim i:

$$\sum_{i=1}^n |x_i| < \sum_{i=1}^n \sum_{j=1}^n M_{ij}|x_j| = \sum_{j=1}^n \underbrace{\sum_{i=1}^n M_{ij}}_{1} |x_j| = \sum_{j=1}^n |x_j|$$

Dakle, x ne može imati i strogo pozitivne i strogo negativne komponente. Neka su npr. svi $x_i \geq 0$. Iz $x \neq \mathbf{0}$ i

$$x_i = \sum_{j=1}^n M_{ij}x_j \ \ (\text{za barem jedan } j \text{ je } x_j > 0 \text{ pa i } M_{ij}x_j > 0)$$

slijedi da je $x_i > 0$. Analogno dokazujemo za sve $x_i \leq 0$.

Pomoćni rezultat

Propozicija

Neka su v, w linearno nezavisni vektori u $\mathbb{R}^m$, $m \geq 2$. Tada postoje $s, t \in \mathbb{R}$, $|s| + |t| > 0$, tako da vektor $x = sv + tw$ ima i pozitivne i negativne komponente.

DOKAZ:Po pretpostavci je $v \neq \mathbf{0}$, $w \neq \mathbf{0}$. Stavimo $d = \sum_{i=1}^m v_i$. Ako je $d = 0$, onda v ima miješane predznake pa uzimamo $s = 1$, $t = 0$.
Ako je $d \neq 0$, stavimo

$$s = -\frac{\sum_{i=1}^n w_i}{d}, \;\; t = 1; \;\; x = sv + tw.$$

Zbog linearne nezavisnosti je $x \neq \mathbf{0}$ i

$$\sum_{i=1}^n x_i = s\sum_{i=1}^n v_i + t\sum_{i=1}^n w_i = -\frac{\sum_{i=1}^n w_i}{d}\sum_{i=1}^n v_i + \sum_{i=1}^n w_i = 0.$$

Dakle, x ima miješane predznake.

Lema

Lema

Neka je M pozitivna (po elementima) i stupčano stohastička matrica. Tada je svojstveni potprostor svojstvene vrijednosti $\lambda = 1$ jednodimenzionalan.

Dokaz: Prepostavimo suprotno. Neka postoje dva linearno nezavisna svojstvena vektora, v i w: $Mv = v$, $Mw = w$. Neka su s, t skalari takvi da je $|s| + |t| > 0$ (nisu oba jednaka nuli). Tada je $z = sv + tw$ također svojstveni vektor i kao takav mora imati komponente istog predznaka (sve > 0 ili sve < 0). Ali upravo smo pokazali da možemo odabrati skalare s i t tako da z ima komponente oba predznaka. Kontradikcija!

Dakle, imamo jednodimenzionalni (realni) svojstveni potprostor u kojem su svi netrivijalni vektori s komponentama istog predznaka (sve > 0 ili sve < 0). Normiranjem s $x_i > 0$ i $\sum_{i=1}^{n} x_i = 1$ smo izdvojili jedinstveno specificiran svojstveni vektor x.

Računanje vektora x

Dakle, imamo matricu M koja je stupčano stohastička, pozitivna po elementima i znamo da je $\lambda = 1$ njena svojstvena vrijednost geometrijske kratnosti jedan. Tražimo pripadni svojstveni vektor x, normiran tako da je $x_i > 0$, $\sum_{i=1}^n x_i = 1$.

Primijetimo da je

$$\|M\|_1 = \max_{j=1:n} \sum_{i=1}^n |M_{ij}| = \max_{j=1:n} \sum_{i=1}^n M_{ij} = 1$$

Ako je λ_i proizvoljna svojstvena vrijednost od M s pripadnim svojstvenim vektorom z, $Mz = \lambda_i z$ onda je

$$|\lambda_i| \|z\|_1 = \|\lambda_i z\|_1 = \|Mz\|_1 \leq \|M\|_1 \|z\|_1 \Longrightarrow |\lambda_i| \leq \|M\|_1 = 1$$

što znači da je $\lambda_1 = 1$ dominantna svojstvena vrijednost,

$$1 = |\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_n|$$

Metoda potencija za računanje x

Za računanje svojstvenog vektora dominantne svojstvene vrijednosti se može koristiti metoda potencija. S nekim zadanim početnim $x^{(0)}$, generiramo niz vektora

$$x^{(k+1)} = \frac{Mx^{(k)}}{\|Mx^{(k)}\|_1}, \quad k = 0, 1, 2, \ldots$$

i očekujemo konvergenciju k traženom vektoru kada $k \to \infty$. Slijedi analiza.

Propozicija

Neka je M pozitivna stupčano stohastička $n \times n$ matrica i

$$V = \{v \in \mathbb{R}^n \; : \; \sum_{i=1}^{n} v_i = 0\} = e^{\perp}$$

Tada je $MV \subseteq V$ ($\forall v \in V$, $Mv \in V$) i za svaki $v \in V$ je $\|Mv\|_1 \leq c\|v\|_1$, s

$$c = \max_{j=1:n} |1 - 2 \min_{i=1:n} M_{ij}| < 1.$$

Dokaz propozicije

DOKAZ: Uzmimo $v \in V$ i izračunajmo $w = Mv$. Provjerimo da li je $w \in V$:

$$\sum_{i=1}^{n} w_i = \sum_{i=1}^{n} \sum_{j=1}^{n} M_{ij} v_j = \sum_{j=1}^{n} \sum_{i=1}^{n} M_{ij} v_j = \sum_{j=1}^{n} v_j = 0 \Rightarrow w \in V$$

Ostaje ocijeniti $\|w\|_1 = \sum_{i=1}^{n} |w_i|$. Neka je $\epsilon_i = \mathrm{predznak}(w_i)$.

$$\|w\|_1 = \sum_{i=1}^{n} \epsilon_i w_i = \sum_{i=1}^{n} \epsilon_i \sum_{j=1}^{n} M_{ij} v_j .$$

Ako je $w = 0$, onda je trivijalno $\|Mv\|_1 = 0 \leq c\|v\|_1$. Neka je $w \neq \mathbf{0}$. Jer je $w \in V$, u nizu $\epsilon_1, \ldots, \epsilon_n$ su zastupljena oba predznaka.

$$\|w\|_1 = \sum_{j=1}^{n} v_j \sum_{i=1}^{n} M_{ij} \epsilon_i = \sum_{j=1}^{n} v_j a_j . \quad \text{Što možemo reći o } a_j?$$

Nastavak dokaza Propozicije

Ocjenjujemo $a_j = \sum_{i=1}^n M_{ij}\epsilon_i$. Koristeći

$$M_{ij} > 0, \;\; M_{ij} < 1, \;\; \sum_{i=1}^n M_{ij} = 1,$$

imamo

$$a_j = \sum_{i=1}^n M_{ij} \underbrace{\epsilon_i}_{\pm} \leq 1 - 2\min_i M_{ij} < 1.$$

Isto vrijedi i za $-a_j$ pa je $a_j \geq -1 + 2\min_i M_{ij} > -1$ pa je

$$|a_j| \leq |1 - 2\min_i M_{ij}| < 1.$$

Dakle

$$\|Mv\|_1 = \|w\|_1 = \sum_{j=1}^n v_j a_j \leq \sum_{j=1}^n |a_j||v_j| \leq \max_j |a_j| \|v\|_1.$$

Konvergencija

Teorem o konvergenciji

Svaka pozitivna stupčano stohastička matrica M ima jedinstven vektor x s pozitivnim elementima i $\|x\|_1 = 1$,za kojeg je $Mx = x$. Vektor x se može dobiti kao

$$x = \lim_{k\to\infty} M^k x^{(0)}$$

gdje je $x^{(0)}$ proizvoljan pozitivan vektor s $\|x^{(0)}\|_1 = 1$.

DOKAZ: Neka je $x^{(0)} > 0$ (po komponentama) i $\|x^{(0)}\|_1 = 1$.
Stavimo $x^{(0)} = x + v$. (Ovim definiramo $v = x^{(0)} - x$.) Odmah vidimo da je $\sum_i v_i = 0$, tj. $v \in V$. Sjetimo se, $x^{(k+1)} = Mx^{(k)}/\|Mx^{(k)}\|_1$, gdje je stalno $\|Mx^{(k)}\|_1 = 1$. Dakle

$$x^{(k)} = M^k x^{(0)} = \underbrace{M^k x}_{x} + \underbrace{M^k v}_{\in V}, \quad \|M^k v\|_1 \le c^k \|v\|_1$$

Dakle, $\|x^{(k)} - x\|_1 \to 0$ kada $k \to \infty$.

Detalji

Sjetimo se, naša matrica M je oblika $M = (1-\alpha)A + \alpha S$, gdje su obe A i S stupčano stohastičke. Umjesto ranije spomenute $S = ff^T$ možemo koristiti $S = se^T$, gdje je s stohastički vektor, $s_i \geq 0$, $e^T s = 1$. Iteraciju $x^{(k+1)} = Mx^{(k)}$ implementiramo kao

$$x^{(k+1)} = (1-\alpha)Ax^{(k)} + \alpha se^T x^{(k)} = (1-\alpha)Ax^{(k)} + \alpha s$$

tj. matricu S ne formiramo (velike je dimenzije!) i efikasno koristimo rijetku popunjenost matrice A (za računanje $Ax^{(k)}$).

Teorem

Ako su $\lambda_1 = 1, \lambda_2(A), \ldots, \lambda_n(A)$ svojstvene vrijednosti matrice A, onda su svojstvene vrijednosti matrice M, redom, 1 i $\lambda_i = (1-\alpha)\lambda_i(A)$:

$$1, (1-\alpha)\lambda_2(A), (1-\alpha)\lambda_3(A), \ldots, (1-\alpha)\lambda_n(A)$$

Dokaz Teorema

Stavimo $\hat{e} = e/\sqrt{n}$ i $\hat{e}$ dopunimo do ortogonalne $n \times n$ matrice $U = (\hat{e}, U_1)$. Primijetimo da je $\hat{e}^T A = \hat{e}^T$. Računamo sličnost $U^T A U$ (jer je $U^T = U^{-1}$) i zatim $U^T M U$:

$$
\begin{aligned}
U^T A U &= \begin{pmatrix} \hat{e}^T \\ U_1^T \end{pmatrix} A(\hat{e}, U_1) = \begin{pmatrix} \hat{e}^T A \\ U_1^T A \end{pmatrix} (\hat{e}, U_1) = \begin{pmatrix} \hat{e}^T \\ U_1^T A \end{pmatrix} (\hat{e}, U_1) \\
&= \begin{pmatrix} \hat{e}^T \hat{e} & \hat{e}^T U_1 \\ U_1^T A \hat{e} & U_1^T A U_1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ w & T \end{pmatrix}, \ \mathrm{eig}(T) = \{\lambda_2(A), \ldots, \lambda_n(A)\} \\
U^T M U &= U^T((1-\alpha)A + \alpha s e^T)U \\
&= (1-\alpha) \begin{pmatrix} 1 & 0 \\ w & T \end{pmatrix} + \alpha \begin{pmatrix} 1/\sqrt{n} \\ U_1^T s \end{pmatrix} \begin{pmatrix} \sqrt{n} & 0 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 \\ \star & (1-\alpha)T \end{pmatrix} \\
&\Longrightarrow \mathrm{eig}(M) = \{1, (1-\alpha)\lambda_2(A), \ldots, (1-\alpha)\lambda_n(A)\}
\end{aligned}
$$

O konvergenciji

Pogledajmo o čemu ovisi brzina konvergencije. Neka je M dijagonalizabilna matrica, tj. neka ima n linearno nezavisnih svojstvenih vektora $\mathsf{v}_1, \ldots, \mathsf{v}_n$ i neka su pripadne svojstvene vrijednosti indeksirane tako da je $1 = \lambda_1 \geq |\lambda_2| \geq \cdots \geq |\lambda_n|$. Neka je također $|\lambda_1| = 1 > |\lambda_2|$. Koristeći relacije $M\mathsf{v}_j = \lambda_j \mathsf{v}_j$ i reprezentaciju $x^{(0)} = \beta_1 \mathsf{v}_1 + \beta_2 \mathsf{v}_2 + \cdots + \beta_n \mathsf{v}_n$ dobijemo

$$\begin{aligned} M^k x^{(0)} &= \beta_1 \lambda_1^k \mathsf{v}_1 + \beta_2 \lambda_2^k \mathsf{v}_2 + \cdots + \beta_n \lambda_n^k \mathsf{v}_n \\ &= \underbrace{\lambda_1^k}_{1} \left(\beta_1 \mathsf{v}_1 + \beta_2 \underbrace{\left(\frac{\lambda_2}{\lambda_1}\right)^k}_{\to 0} \mathsf{v}_2 + \cdots + \beta_n \underbrace{\left(\frac{\lambda_n}{\lambda_1}\right)^k}_{\to 0} \mathsf{v}_n \right) \end{aligned}$$

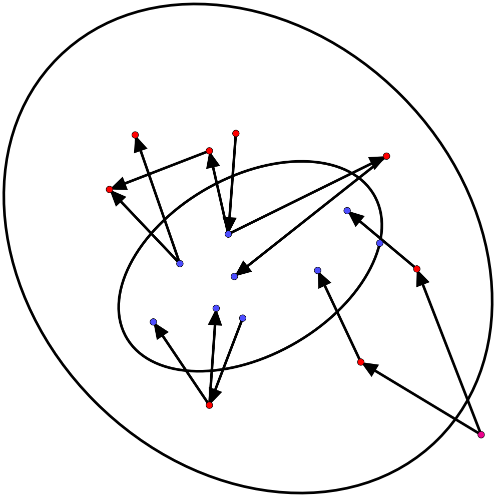
Dakle, bitno je kako brzo $|\lambda_2|^k \to 0$. Kako je $\lambda_2 = (1-\alpha)\lambda_2(A)$ s $|\lambda_2(A)| \leq 1$ (nadamo se i $|\lambda_2| < 1$) i $\alpha \in [0,1)$, vidimo da odabirom $\alpha \in (0,1)$ možemo utjecati na brzinu konvergencije. Diskusija.

Literatura

Literatura

- Kurt Bryan, Tanya Leise: The \$ 25,000,000,000 Eigenvector: The Linear Algebra behind Google. SIAM Review Vol. 48, No. 3, pp. 569-581, 2006.
- Lars Eldén: A Note on the Eigenvalues of the Google Matrix (arXiv:math/0401177)
- Roger Nussbaum: Notes on the Second Eigenvalue of the Google Matrix (arXiv:math/0307056)

Hyperlink-Induced Topic Search (HITS)



$\wp$=upit; $O_\wp$=pronađeni relevantni dokumenti, koje rangiramo na neki način (npr. PageRank) i uzmemo podskup od k "najboljih"

$R_\wp = \{\bullet\} \subset O_\wp$. Ako $\bullet \to \bullet$ ili $\bullet \to \bullet$ (ili $\bullet \to \bullet \to \bullet$) onda promatramo skup $S_\wp = R_\wp \cup \{\bullet, \bullet\}$. Neka je $G[S_\wp]$ inducirani graf. S odgovarajućom strukturom podataka su ovo jednostavne operacije. Sada rangiramo vrhove u $G[S_\wp]$.

Rangiranje vrhova u $G[S_\wp]$

Osim rangiranja tipa PageRank, vrhove u grafu možemo rangirati s obzirom na različita svojstva koja su od interesa u konkretnoj primjeni.

Hubovi (čvorišta) i autoriteti

Kažemo da je vrh i

- dobar izvor informacije o lokaciji kvalitetnog sadržaja ili dobro čvorište (*hub*) ako sadrži linkove na vrhove koji su dobri autoriteti u smislu sadržavanja kvalitetne informacije
- dobar autoritet (sadrži kvalitetnu informaciju o nekoj temi) ako na njega pokazuju dobri hubovi.

Svakom vrhu pridružujemo uređen par (a_i, h_i) nenegativnih brojeva kao mjeru za "biti dobar autoritet" i "biti dobar hub".

$$\overbrace{\bullet}^{hub} \rightarrow \underbrace{\bullet}_{autoritet} \leftarrow \overbrace{\bullet}^{hub}; \quad \overbrace{\bullet}^{autoritet} \leftarrow \underbrace{\bullet}_{hub} \rightarrow \overbrace{\bullet}^{autoritet}$$

Autoriteti i hubovi

Težine $a_i \geq 0$, $h_i \geq 0$ definiramo rekurzivno s

$a_i = \sum_{j \to i} h_j$ i je dobar autoritet koliko su dobri hubovi koji pokazuju na njega

$h_i = \sum_{i \to j} a_j$ i je dobar hub koliko su dobri autoriteti na koje pokazuje

Kako odrediti a_i, h_i? Da li ima smisla h_i skalirati kao za PageRank? Diskusija. Stavimo $a = (a_1, \ldots, a_n)^T$, $h = (h_1, \ldots, h_n)^T$ i

$$L = (L_{ij})_{i,j=1}^n, \quad L_{ij} = \begin{cases} 1 & \text{ako } i \to j \\ 0 & \text{inače} \end{cases}$$

Ako gronji princip iteriramo i generiramo nizove $a^{(t)}$ i $h^{(t)} = La^{(t)}$ dobijemo $a^{(t+1)} = L^T h^{(t)} = L^T L a^{(t)}$, $h^{(t+1)} = La^{(t+1)} = LL^T h^{(t)}$. Te relacije korigiramo normiranjem u

$$a^{(t+1)} = \frac{L^T L a^{(t)}}{\|L^T L a^{(t)}\|_2}, \quad h^{(t+1)} = \frac{LL^T h^{(t)}}{\|LL^T h^{(t)}\|}.$$

Autoriteti i hubovi

Dakle dobili smo dva niza generirana metodom potencija za $L^T L$ i LL^T.

$$a^{(t+1)} = \frac{L^T L a^{(t)}}{\|L^T L a^{(t)}\|_2}, \quad h^{(t+1)} = \frac{LL^T h^{(t)}}{\|LL^T h^{(t)}\|}.$$

Matrice $L^T L$ i LL^T imaju iste svojstvene vrijednosti, pa ako je λ_1 dominantna svojstvena vrijednost ($|\lambda_1| > \max_{i>1} |\lambda_i|$) onda imamo konvergenciju: postoje a i h, $\|a\|_2 = \|h\|_2 = 1$ tako da je

$$L^T L a = \lambda_1 a, \quad LL^T h = \lambda_1 h$$

Veza sa SVD od L

Neka je $L = U\Sigma V^T$ SVD od L; $\Sigma = \mathrm{diag}(\sigma_i)_{i=1}^n$, $\sigma_1 \geq \cdots \geq \sigma_n \geq 0$; $U = (u_1, \ldots, u_n)$ i $V = (v_1, \ldots, v_n)$ ortogonalne.

$$(LL^T)u_1 = \sigma_1^2 u_1, \;\; (L^T L)v_1 = \sigma_1^2 v_1, \;\; Lv_1 = \sigma_1 u_1, \;\; L^T u_1 = \sigma_1 v_1.$$

Kao i prije, postavlja se pitanje jedinstvenosti, efikasnog računanja itd.

Definicija

Kažemo da je $\mathsf{A} \in \mathbb{R}^{n\times n}$, $n \geq 2$, reducibilna ako postoji matrica permutacije P i $r \in \{1, \dots, n-1\}$ tako da je

$$\mathsf{P}^T \mathsf{A} \mathsf{P} = \begin{pmatrix} \tilde{\mathsf{A}}_{[11]} & \tilde{\mathsf{A}}_{[12]} \\ \mathbf{0} & \tilde{\mathsf{A}}_{[22]} \end{pmatrix}, \quad \tilde{\mathsf{A}}_{[11]} \in \mathbb{R}^{r\times r}. \tag{1}$$

Za $n = 1$ je A reducibilna ako je $\mathsf{A} = (0)$. Matrica A je ireducibilna ako nije reducibilna.

Ako je A simetrična, onda je i $\mathsf{P}^T\mathsf{A}\mathsf{P}$ simetrična, pa je $\tilde{\mathsf{A}}_{[12]} = \mathbf{0}$. Ako je neka od matrica $\tilde{\mathsf{A}}_{[11]}$, $\tilde{\mathsf{A}}_{[22]}$ reducibilna onda je na isti način možemo permutacijskom transformacijom sličnosti svesti na blok–trokutasti oblik.

Ireducibilne matrice

Postupak primjenjujemo rekurzivno sve dok na kraju ne dobijemo blok–trokutastu matricu

$$
\mathsf{P}^T \mathsf{A} \mathsf{P} = \begin{pmatrix} \mathsf{A}_{[11]} & \mathsf{A}_{[12]} & \cdots & \mathsf{A}_{[1k]} & \cdots & \mathsf{A}_{[1,\ell]} \\ & \mathsf{A}_{[22]} & \mathsf{A}_{[23]} & \cdots & \cdots & \mathsf{A}_{[2,\ell]} \\ & & \ddots & \ddots & \cdots & \vdots \\ & & & \mathsf{A}_{[kk]} & \cdots & \mathsf{A}_{[k,\ell]} \\ & & & & \ddots & \vdots \\ & & & & & \mathsf{A}_{[\ell\ell]} \end{pmatrix}. \tag{2}
$$

u kojoj su svi dijagonalni blokovi ireducibilne matrice ili 1×1 nul–matrice. Blok gornje trokutasta matrica sa ireducibilnim dijagonalnim blokovima u relaciji (1) se zove ireducibilna normalna forma matrice reducibilne A. Ako je A simetrična, onda je i $\mathsf{P}^T\mathsf{A}\mathsf{P}$ simetrična, pa je $\tilde{\mathsf{A}}_{[ij]} = \mathbf{0}$ za sve $i < j$.

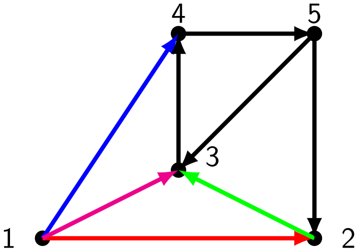
Grafovska interpretacija ireducibilnosti matrice

Ireducibilnost matrice je svojstvo kombinatorne prirode i kao takvoga ga možemo prirodnije karakterizirati preko jednog kombinatornog objekta izvedenog iz matrice.

Definicija

Za matricu A $\in \mathcal{M}_n$ vežemo usmjereni graf $\Gamma(\mathrm{A})$ koji ima n vrhova $\nu_1, \nu_2, \ldots, \nu_n$ i u kojem postoji usmjereni brid $\overrightarrow{\nu_i \nu_j}$ ako i samo ako je $a_{ij} \neq 0$. Put u Γ od vrha ν_k do vrha ν_ℓ je niz usmjerenih bridova $\overrightarrow{\nu_k \nu_{i_1}}, \overrightarrow{\nu_{i_1} \nu_{i_2}}, \ldots, \overrightarrow{\nu_{i_{j-1}} \nu_\ell}$, pri čemu je j duljina puta. Kažemo da je $\Gamma(\mathrm{A})$ jako povezan ako su svaka dva njegova vrha povezana nekim putem.

$$A = \begin{pmatrix} 0 & \star & \star & \star & 0 \\ 0 & 0 & \star & 0 & 0 \\ 0 & 0 & 0 & \star & 0 \\ 0 & 0 & 0 & 0 & \star \\ 0 & \star & \star & 0 & 0 \end{pmatrix}, \quad \star \neq 0$$



Ireducibilne matrice

Propozicija

Matrica $\mathsf{A} \in \mathbb{C}^{n\times n}$ je ireducibilna ako i samo ako je njen graf $\Gamma(\mathsf{A})$ jako povezan.

DOKAZ: Neka je A reducibilna i neka s nekom permutacijom P vrijedi

$$\tilde{\mathsf{A}} = \mathsf{P}^T\mathsf{A}\mathsf{P} = \begin{array}{c c} & \begin{array}{cc} \tilde{\mathcal{V}}_2 & \tilde{\mathcal{V}}_1 \end{array} \\ \begin{array}{c} \tilde{\mathcal{V}}_2 \\ \tilde{\mathcal{V}}_1 \end{array} & \begin{pmatrix} \blacksquare & \blacksquare \\ \mathbf{0} & \blacksquare \end{pmatrix} \end{array}$$

Grafovi $\Gamma(\mathsf{A})$ i $\Gamma(\tilde{\mathsf{A}})$ se razlikuju samo u oznakama čvorova, označimo ove druge sa $\tilde{\nu}_1, \ldots, \tilde{\nu}_n$. Ako bismo tražili put u $\Gamma(\tilde{\mathsf{A}})$ koji počinje u nekom čvoru iz $\tilde{\mathcal{V}}_1 = \{\tilde{\nu}_{r+1}, \ldots, \tilde{\nu}_n\}$ a završava u nekom od čvorova iz $\tilde{\mathcal{V}}_2 = \{\tilde{\nu}_1, \ldots, \tilde{\nu}_r\}$, onda bi na tom putu morao postojati brid $\overrightarrow{\tilde{\nu}_i\tilde{\nu}_j}$ sa $\tilde{\nu}_i \in \tilde{\mathcal{V}}_1$, $\tilde{\nu}_j \in \tilde{\mathcal{V}}_2$, tj. morao bi biti $\tilde{a}_{ij} \neq 0$ – ali to je nemoguće jer je taj dio matrice $\tilde{\mathsf{A}}$ jednak nuli. Dakle, $\Gamma(\tilde{\mathsf{A}})$ pa niti $\Gamma(\mathsf{A})$ nije jako povezan.

Ireducibilne matrice ... nastavak dokaza

Sada pretpostavimo da $\Gamma(\mathrm{A})$ nije jako povezan. To znači da postoje dva vrha ν_k i ν_ℓ iz skupa vrhova takva da ne postoji put od ν_k do ν_ℓ. Skup $\mathcal{V}$ svih vrhova particionirajmo na sljedeći način:

- u $\mathcal{V}_1$ neka sadrži ν_ℓ i sve vrhove koji su putovno povezani s ν_ℓ, $\mathcal{V}_1 = \{\nu_\ell, \ldots\}$
- $\mathcal{V}_2 = \mathcal{V} \setminus \mathcal{V}_1 = \{\nu_k, \ldots\}$.

Sada uvedemo renumeraciju vrhova tako da prvo izlistamo one iz $\mathcal{V}_1$, a zatim iz $\mathcal{V}_2$. Neka je $|\mathcal{V}_1| = r$. S odgovarajućom permutacijom P transformiramo A u

$$\tilde{\mathrm{A}} = \mathrm{P}^T \mathrm{A} \mathrm{P} = \begin{array}{c c} & \begin{array}{cc} \mathcal{V}_1 & \mathcal{V}_2 \end{array} \\ \begin{array}{c} \mathcal{V}_1 \\ \mathcal{V}_2 \end{array} & \begin{pmatrix} \blacksquare & \blacksquare \\ \mathbf{0} & \blacksquare \end{pmatrix} \end{array}$$

Ako bi za neki $i \in \{r+1, \ldots, n\}$ i $j \in \{1, \ldots, r\}$ bilo $\tilde{a}_{ij} \neq 0$, onda bi to značilo da je neki vrh iz $\mathcal{V}_2$ povezan bridom sa nekim vrhom iz $\mathcal{V}_1$, a time onda i putem sa ν_ℓ. No, to je kontradikcija s definicijom skupa $\mathcal{V}_2$ pa je $\tilde{\mathrm{A}}$ reducibilna. Time je dokaz završen. Diskusija.

Pozitivne matrice

Specijalni slučaj ireducibilnih su pozitivne matrice. Posebno nas zanima spektralni radijus $\mathrm{spr}(\mathsf{A}) = \max\{|\lambda| \,:\, \lambda \text{ svojstvena vrijednost od } A\}$.

Perronov teorem za pozitivne matrice

Neka je $\mathsf{A} \in \mathbb{R}^{n\times n}$ pozitivna matrica. Tada vrijedi:

- $\mathrm{spr}(\mathsf{A}) > 0$ i $\mathrm{spr}(\mathsf{A})$ je svojstvena vrijednosti od A algebarske kratnosti jedan, pri čemu se pripadni svojstveni vektor v može odabrati sa realnim pozitivnim komponentama, $\mathsf{Av} = \mathrm{spr}(\mathsf{A})\mathsf{v}$, $\mathsf{v} > \mathbf{0}$.
- $\mathrm{spr}(\mathsf{A})$ je jedina svojstvena vrijednost koja postiže maksimalnu apsolutnu vrijednost, tj. $(\forall\lambda \in \mathrm{Spektar}(\mathsf{A}))\ (\lambda \neq \mathrm{spr}(\mathsf{A}) \Longrightarrow |\lambda| < \mathrm{spr}(\mathsf{A}))$.
- Neka su $\mathsf{u} > \mathbf{0}$, $\mathsf{v} > \mathbf{0}$ lijevi i desni svojstveni vektor koji pripadaju $\mathrm{spr}(\mathsf{A})$. Tada je $\lim_{m\to\infty}(\mathsf{A}/\mathrm{spr}(\mathsf{A}))^m = \mathsf{vu}^T$.

Nenegativnu matricu možemo uvijek napisati kao limes pozitivnih matrica i to otvara prostor da neke rezultate proširimo na nenegativne matrice. Ipak valja biti oprezan jer neka važna svojstva matrice kao npr. rang ili stroga pozitivnost spektralnog radijusa nisu nužno sačuvana u limesu.

Nenegativne matrice

Teorem

Neka je $\mathsf{A} \in \mathbb{R}^{n\times n}$ nenegativna matrica. Tada je $\operatorname{spr}(\mathsf{A})$ svojstvena vrijednost od A i pripadni svojstveni vektor v možemo odabrati da bude nenegativan, $\mathsf{Av} = \operatorname{spr}(\mathsf{A})\mathsf{v}$, $\mathsf{v} \geq \mathbf{0}$, $\mathsf{v} \neq \mathbf{0}$.

Nenegativne matrice tek uz dodatan uvjet ireducibilnosti uživaju neka svojstva pozitivnih matrica.
Diskusija: što nam znači ireducibilnost u primjeni na analizu vrhova neke mreže? Sjetimo se kako smo u algoritmu PageRank nenegativnoj matrici dodali strogo pozitivnu i tako osigurali neka dobra svojstva algoritma.

Nenegativne ireducibilne matrice

Perronov teorem za nenegativne ireducibilne matrice

Neka je $\mathsf{A} \in \mathbb{R}^{n\times n}$ nenegativna i ireducibilna matrica. Tada vrijedi:

- $\operatorname{spr}(\mathsf{A}) > 0$ i $\operatorname{spr}(\mathsf{A})$ je svojstvena vrijednosti od A algebarske kratnosti jedan, pri čemu se pripadni svojstveni vektor v može odabrati sa realnim pozitivnim komponentama, $\mathsf{Av} = \operatorname{spr}(\mathsf{A})\mathsf{v}$, $\mathsf{v} > \mathbf{0}$.
- Spektralni radijus je strogo rastuća funkcija: Ako povećamo bilo koji element od A i dobijemo $\tilde{\mathsf{A}} \geq \mathsf{A}$ i $\tilde{\mathsf{A}} \neq \mathsf{A}$, onda je $\operatorname{spr}(\tilde{\mathsf{A}}) > \operatorname{spr}(\mathsf{A})$.
- Ako u spektru od A točno ℓ svojstvenih vrijednosti ima modul jednak $\operatorname{spr}(\mathsf{A})$, onda su one jednake

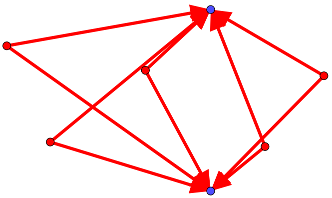
$$\operatorname{spr}(\mathsf{A})\mathbf{e}^{2\pi \mathrm{i} j/\ell},\ j = 0, \ldots, \ell - 1.$$

$L^T L$ i LL^T – ko-citiranost

Dakle, naš model za rangiranje u smislu autoriteta i hubova ovisi o strukturi matrica $L^T L$ i LL^T.
Koju informaciju zapravo sadrže elementi tih matrica?

$$\text{za } i \neq j \text{ stavimo } \; C_{ij} = (L^T L)_{ij} = \sum_{k=1}^{n} \underbrace{L_{ki}}_{k \to i} \underbrace{L_{kj}}_{k \to j}$$

C_{ij} broji koliko vrhova istovremeno pokazuje na ("citira") i i j; C_{ij} zovemo ko-citiranost. Stavimo po definiciji $C_{ii} = 0$.



$$d_i = \sum_{k=1}^{n} L_{ki} L_{ki} = (L^T L)_{ii}; \quad D = \mathrm{diag}(d_i)$$

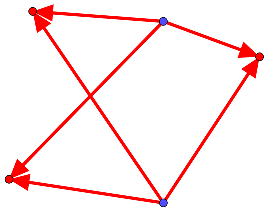
$$\max(0, d_i + d_j - n) \leq C_{ij} \leq \min(d_i, d_j)$$

$$L^T L = D + C$$

$L^T L$ i LL^T – ko-referenciranje

za $i \neq j$ stavimo $R_{ij} = (LL^T)_{ij} = \sum_{k=1}^{n} \underbrace{L_{ik}}_{i \to k} \underbrace{L_{jk}}_{j \to k}$

R_{ij} broji na koliko vrhova istovremeno pokazuju ("referenciraju ih") i i j;
R_{ij} zovemo ko-referenciranje. Stavimo po definiciji $R_{ii} = 0$.



$$o_i = \sum_{k=1}^{n} L_{ki} L_{ki} = (L^T L)_{ii}; \quad O = \mathrm{diag}(o_i)$$

$$LL^T = O + R$$

Implementacija i primjene. Diskusija.
Više detalja u

- Chris H. Q. Ding, Hongyuan Zha, Xiaofeng He, Horst D. Simon: *Link Analysis: Hubs and Authorities on the World Wide Web*, SIAM Review 46(2), 2002.

Poopćenje - složenija svojstva vrhova

$$\overbrace{\bullet}^{hub} \rightarrow \underbrace{\bullet}_{autoritet} \leftarrow \overbrace{\bullet}^{hub}; \qquad \overbrace{\bullet}^{autoritet} \leftarrow \underbrace{\bullet}_{hub} \rightarrow \overbrace{\bullet}^{autoritet}$$

$$h_j = \sum_{i,j\to i} a_i; \quad a_j = \sum_{i,i\to j} h_i; \quad h = (h_j), \quad a = (a_j) \in \mathbb{R}^n_+$$

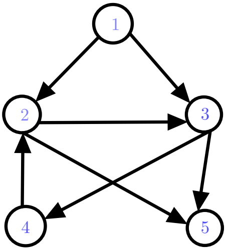
Kompaktan matrični zapis za $a^{(k+1)} = L^T h^{(k)}$, $h^{(k+1)} = La^{(k)}$.:

$$z_{k+1} = \begin{pmatrix} h \\ a \end{pmatrix}_{(k+1)} = \underbrace{\begin{pmatrix} \mathbf{0} & L \\ L^T & \mathbf{0} \end{pmatrix}}_{M} \begin{pmatrix} h \\ a \end{pmatrix}_{(k)} = Mz_k, \quad M = M^T$$

pa možemo promatrati metodu potencija

$$z_0 = \text{početna iteracija} > \mathbf{0}\ ; \quad z_{k+1} = \frac{Mz_k}{\|Mz_k\|}, \quad k = 0, 1, 2, \ldots$$

Poopćenje - složenija svojstva vrhova



Proučavamo svojstva vrhova grafa $\mathcal{G}_B$ uspoređujući s model-grafom $\mathcal{G}_A$.

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

Svakom vrhu i u $\mathcal{G}_B$ pridružujemo tri ocjene x_{i1}, x_{i2}, x_{i3} prema sličnosti pozicije s vrhovima u modelu.

$$x_{i1} = \sum_{j:(i \to j)(1 \to 2)} x_{j2} \quad \text{(1 pokazuje na 2)} \qquad x_1 = (x_{i1})_{i=1}^{n}$$

$$x_{i2} = \sum_{j:j \to i(1 \to 2)} x_{j1} + \sum_{j:i \to j(2 \to 3)} x_{j3} \qquad x_2 = (x_{i2})_{i=1}^{n}$$

$$x_{i3} = \sum_{j:j \to i(2 \to 3)} x_{j2} \qquad x_3 = (x_{i3})_{i=1}^{n}$$

Da li te brojeve možemo dobiti iteracijama?

Kompaktan zapis i poopćenje

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_{(k+1)} = \begin{pmatrix} \mathbf{0} & B & \mathbf{0} \\ B^T & \mathbf{0} & B \\ \mathbf{0} & B^T & \mathbf{0} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_{(k)}$$

Kako ovo izgleda ako $\mathcal{G}_B$ ima n_B a $\mathcal{G}_A$ n_A vrhova? Neka je x_{ij} sličnost $i$tog vrha iz $\mathcal{G}_B$ s $j$tim vrhom iz $\mathcal{G}_A$; $i = 1, \ldots, n_B$, $j = 1, \ldots, n_A$. Kako svaki vrh ima ulazne i izlazne bridove prethodni primjer direktno poopćavamo sljedećim iteracijama

$$x_{ij}^{(k+1)} = \sum_{\substack{r\,:\,(r \to i)_B \\ s\,:\,(s \to j)_A}} x_{rs}^{(k)} + \sum_{\substack{r\,:\,(i \to r)_B \\ s\,:\,(j \to s)_A}} x_{rs}^{(k)}$$

Ako definiramo $n_B \times n_A$ matrice $X_k = (x_{ij}^{(k)})$, $k \geq 0$, onda nakon neke startne X_0 imamo iteracije

$$X_{k+1} = BX_kA^T + B^TX_kA \equiv \mathcal{L}(X_k)$$

$\mathcal{L}$ je linearan operator; ako dodamo normiranje imamo metodu potencija. Konvergencija?

$C = A \otimes B$

Definicija: Kroneckerov produkt

Neka su $A \in \mathbb{C}^{m\times n}$, $B \in \mathbb{C}^{p\times q}$ proizvoljne. Kroneckerov produkt $C = A \otimes B$ je definiran s

$$C = \begin{pmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{pmatrix} \in \mathbb{C}^{m\cdot p\times n\cdot q}$$

Primjer:

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}, C = A\otimes B = \begin{pmatrix} a_{11}b_{11} & a_{11}b_{12} & a_{12}b_{11} & a_{12}b_{12} \\ a_{11}b_{21} & a_{11}b_{22} & a_{12}b_{21} & a_{12}b_{22} \\ a_{21}b_{11} & a_{21}b_{12} & a_{22}b_{11} & a_{22}b_{12} \\ a_{21}b_{21} & a_{21}b_{22} & a_{22}b_{21} & a_{22}b_{22} \end{pmatrix}$$

Ako je $X = (x_1, \dots, x_n)$ stupčana particija $m \times n$ matrice, onda je operator $\mathrm{vec} : \mathbb{C}^{m\times n} \to \mathbb{C}^{m\times n}$ definiran s

$$\mathrm{vec}(X) = (x_1^T, x_2^T, \dots, x_{n-1}^T, x_n^T)^T = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

izomorfizam vektorskih prostora.

Svojstva množenja $C = A \otimes B$

- For $\alpha \in \mathbb{C}$, $A \otimes (\alpha B) = \alpha(A \otimes B)$
- $(A + B) \otimes C = A \otimes C + B \otimes C$
- $A \otimes (B + C) = A \otimes B + A \otimes C$
- $(A \otimes B) \otimes C = A \otimes (B \otimes C)$
- $(A \otimes B)^T = A^T \otimes B^T$; $(A \otimes B)^* = A^* \otimes B^*$
- Nul matrica: $\mathbf{0}_{mn \times mn} = \mathbf{0}_{m \times m} \otimes \mathbf{0}_{n \times n}$
- Jedinična matrica: $\mathbf{I}_{mn \times mn} = \mathbf{I}_{m \times m} \otimes \mathbf{I}_{n \times n}$
- $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$
- $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$ (ako postoje A^{-1}, B^{-1})
- $\mathrm{vec}(AX) = (I \otimes A)\mathrm{vec}(X)$
- $\mathrm{vec}(XB) = (B^T \otimes I)\mathrm{vec}(X)$
- $\mathrm{vec}(CXD) = (D^T \otimes C)\mathrm{vec}(X)$
- ...

Iteracije

Ako na relaciju

$$X_{k+1} = BX_kA^T + B^TX_kA$$

primijenimo operator vec dobijemo

$$\operatorname{vec}(X_{k+1}) = \underbrace{(A \otimes B + A^T \otimes B^T)}_{M} \operatorname{vec}(X_k)$$

gdje je $M \geq 0$ (po elementima) i $M = M^T$ jer je $(A \otimes B)^T = A^T \otimes B^T$.

U našem motivacijskom primjeru smo imali matrice

$$B = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 \end{pmatrix}, \quad A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad M = \begin{pmatrix} \mathbf{0} & B & \mathbf{0} \\ B^T & \mathbf{0} & B \\ \mathbf{0} & B^T & \mathbf{0} \end{pmatrix}.$$

Primijetimo

$$A \otimes B = \begin{pmatrix} 0 & B & 0 \\ 0 & 0 & B \\ 0 & 0 & 0 \end{pmatrix}, \quad A^T \otimes B^T = \begin{pmatrix} 0 & 0 & 0 \\ B^T & 0 & 0 \\ 0 & B^T & 0 \end{pmatrix}$$

Pomoćni rezultat

Teorem

Neka je M nenegativna simetrična matrica spektralnog radijusa ρ. tada su algebarska i geometrijska kratnost Perronovog korjena ρ jednake. Nadalje, postoji nenegativna matrica $X \geq 0$ (nejednakost po svim elementima) čiji stupci razapinju invarijantan (tj. spektralni) potprostor $\mathcal{X}$ od ρ, i elementi pripadnog ortogonalnog projektora (matrice) Π na $\mathcal{X}$ su nenegativni.

DOKAZ: Jer je M nenegativna, ρ je svojstvena vrijednost od M. Jer je M simetrična (dijagonalizabilna s orthogonalnom transformacijom sličnosti) algebarska i geometrijska kratnost bilo koje svojstvene vrijednosti su jednake. $M = \mathbf{0}$ je trivijalno; neka je $M \neq 0$. Postoji permutacija P tako da je

$$P^T M P = \begin{pmatrix} M_1 & 0 & \dots & 0 \\ 0 & M_2 & 0 & 0 \\ & & \ddots & \\ 0 & 0 & & 0 \\ 0 & 0 & 0 & M_k \end{pmatrix}, \quad M_i \text{ ireducibilne ili } \mathbf{0}$$

$\rho(M) = \max_i \rho(M_i) = \rho(M_{i_j})$, M_{i_k} ireducibilne, $j = 1, \ldots, \ell$.

Nastavak dokaza teorema

Jer je za $j = 1, \ldots, \ell$ M_{i_j} ireducibilna, $\rho(M_{ij})$ je jednostruka i pripadni svojstveni vektor y_{i_j} $(M_{i_j} y_{i_j} = \rho(M_{i_j}) y_{i_j})$ se može odabrati pozitivan po komponentama i $\|y_{i_j}\|_2 = 1$. Definirajmo

$$x_{i_j} = \begin{pmatrix} 0 \\ \vdots \\ y_{i_j} \\ \vdots \\ 0 \end{pmatrix}$$

gdje je pozicija y_{i_j} u x_{i_j} analogna poziciji M_{i_j} u M. Stavimo

$$X = (x_{i_1}, x_{i_2}, \ldots, x_{i_\ell})$$

i odmah uočimo $X^T X = I_\ell$.

Dakle, $X \geq 0$, $X^T X = I_\ell$ i $\Pi = X X^T \geq 0$ je ortogonalni projektor na $\mathcal{X} = \mathrm{Slika}(X)$.

Ostaje proučiti konvergenciju niza iteracija

$$z_{k+1} = \frac{M z_k}{\|M z_k\|_2}$$

Konvergencija

Teorem

Neka je M nenegativna simetrična matrica spektralnog radijusa $\rho > 0$ i neka je $z_0 > \mathbf{0}$ pozitivan vektor s kojim generiramo

$$z_{k+1} = \frac{Mz_k}{\|Mz_k\|_2}, \quad k = 0, 1, 2, \dots$$

- Ako $-\rho$ nije svojstvena vrijednost od M tada

$$z_k \longrightarrow \frac{\Pi z_0}{\|\Pi z_0\|_2} \; (k \to \infty)$$

  gdje je Π ortogonalni projektor na invarijantni potprostor pridružen Perronovom korjenu ρ.

- Ako $-\rho$ je svojstvena vrijednost od M, tada postoje

$$z_{par}(z_0) = \lim_{k\to\infty} z_{2k} = \frac{\Pi z_0}{\|\Pi z_0\|_2}, \quad z_{nepar}(z_0) = \lim_{k\to\infty} z_{2k+1} = \frac{\Pi M z_0}{\|\Pi M z_0\|_2}$$

Konvergencija

U oba slučaja je skup svih mogućih limesa jednak

$$\mathcal{Z} = \{z_{par}(z_0),\ z_{nepar}(z_0)\ :\ z_0 > \mathbf{0}\} = \{\frac{\Pi z}{\|\Pi z\|_2}\ :\ z > \mathbf{0}\}$$

i $z_{par}((1,\ldots,1)^T)$ je jedinstveni vektor maksimalne ℓ_1 norme u $\mathcal{Z}$.

Teorem

Neka su $\mathcal{G}_A$ i $\mathcal{G}_B$ grafovi s matricama susjedstva A, B. Neka je $X_0 > \mathbf{0}$ proizvoljna pozitivna (po elementima) matrica i neka je dan niz

$$X_{k+1} = \frac{BX_kA^T + B^TX_kA}{\|BX_kA^T + B^TX_kA\|_F},\quad k = 0, 1, 2, \ldots$$

Tada $X_{2k} \to X_{par}(X_0)$, $X_{2k+1} \to X_{nepar}(X_0)$ i u skupu $\mathbb{X} = \{X_{par}(X_0),\ X_{par}(X_0)\ :\ X_0 > \mathbf{0}\}$ je $X_{par}(\begin{pmatrix} 1 & \ldots & 1 \\ \vdots & \vdots & \vdots \\ 1 & \ldots & 1 \end{pmatrix})$ jedinstvena matrica maksimalne ℓ_1 (vektorske) norme.

Konvergencija

Definicija sličnosti vrhova

Matricu

$$X_{par}(\begin{pmatrix} 1 & \ldots & 1 \\ \vdots & \vdots & \vdots \\ 1 & \ldots & 1 \end{pmatrix}) \in \mathbb{R}_{+}^{n_B \times n_A}$$

koristimo za definiranje težina kao mjera sličnosti vrhova grafa $\mathcal{G}_B$ sa vrhovima grafa $\mathcal{G}_A$. Primijetimo da ovaj izbor ima smisla jer inicijalno nemamo preferencije (sve težine jednake)

Sada se vratimo na naše polazne, motivacijske primjere.

Primjer - centralnost i sinonimi

- Imamo rječnik nekog jezika.
- Pomoću rječnika konstruiramo usmjerni graf na sljedeći način: svaka riječ navedena u rječniku postaje vrh; vrhovi su numerirani s $1, \ldots, n$, prema redoslijedu kako su riječi navedene u rječniku. Iz riječi i postoji usmjereni brid u riječ j ako se j pojavljuje u definiciji od i.
- Za zadanu riječ r konstruiramo graf susjedstva $\mathcal{G}_r$ kojem su vrhovi svi oni na koje pokazuje r i koji pokazuju na r, sa svim njihovim međusobnim bridovima.
- Za vrhove grafa $\mathcal{G}_r$ izračunamo indekse sličnosti s vrhom 2 u grafu $1 \longrightarrow 2 \longrightarrow 3$ i sortiramo ih u padajući niz.
- Riječ s najvećim indeksom sličnosti je centralna i bliska po značenju/istoznačnica/sinonim riječi r.

Primjer - centralnost

- Imamo bazu podataka telefonskih poziva (ili nekih drugih vidova usmjerene komunikacije).
- Pomoću te baze konstruiramo usmjerni graf na sljedeći način: svaki korisnik/broj/osoba navedena u bazi postaje vrh; vrhovi su numerirani s $1, \ldots, n$, prema nekom redoslijedu. Iz vrha i postoji usmjereni brid u vrh j ako iz i postoji komunikacija/poziv k j.
- Za zadani vrh r konstruiramo graf susjedstva $\mathcal{G}_r$ kojem su vrhovi svi oni na koje pokazuje r i koji pokazuju na r, sa svim njihovim međusobnim bridovima.
- Za vrhove grafa $\mathcal{G}_r$ izračunamo indekse sličnosti s vrhom 2 u grafu $1 \longrightarrow 2 \longrightarrow 3$ i sortiramo ih u padajući niz.
- Vrh s najvećim indeksom sličnosti je centralni (s raznim interpretacijama) u grupi kontakata vrha r.

Primjer - centralnost

Ako je G_r matrica susjedstva grafa $\mathcal{G}_r$, i A matrica susjedstva modela $\textcircled{1} \longrightarrow \textcircled{2} \longrightarrow \textcircled{3}$, onda imamo

$$A = \left(\begin{smallmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{smallmatrix}\right), \quad M = \begin{pmatrix} \mathbf{0} & G_r & \mathbf{0} \\ G_r^T & \mathbf{0} & G_r \\ \mathbf{0} & G_r^T & \mathbf{0} \end{pmatrix}, \; M^2 = \begin{pmatrix} G_r G_r^T & \mathbf{0} & G_r G_r \\ \mathbf{0} & G_r^T G_r + G_r G_r^T & \mathbf{0} \\ G_r^T G_r^T & \mathbf{0} & G_r^T G_r \end{pmatrix}.$$

Teorem o indeksu centralnosti

Normalizirani indeksi centralnosti vrhova grafa $\mathcal{G}_r$ su komponente svojstvenog vektora matrice $G_r^T G_r + G_r G_r^T$ koji odgovara Perronovom vektoru, pod uvjetom da je on jednostruk. Inače su jednaki normaliziranoj projekciji vektora jedinica na dominantni invarijantni potprostor.

DOKAZ: Neka je P permutacija tako da je

$$M = P^T \left(\begin{smallmatrix} \mathbf{0} & E \\ E^T & \mathbf{0} \end{smallmatrix}\right) P, \;\; E = \begin{pmatrix} G_r \\ G_r^T \end{pmatrix}; \;\; PM^2P^T = \begin{pmatrix} EE^T & \mathbf{0} \\ \mathbf{0} & E^T E \end{pmatrix}.$$

Nastavak dokaza teorema o indeksu centralnosti

Neka je $E = U\Sigma V^T$ SVD matrice E i neka je σ dominantna singularna vrijednost kratnosti ℓ s dominatnim potprostorima razapetim s $U(:,1:\ell)$, $V(:,1:\ell)$. Vrijedi

$$EV(:,1:\ell) = \sigma U(:,1:\ell), \quad E^T U(:,1:\ell) = \sigma V(:,1:\ell),$$

$$E^T E V(:,1:\ell) = \sigma^2 V(:,1:\ell), \quad E E^T U(:,1:\ell) = \sigma^2 U(:,1:\ell).$$

Pripadni projektori (za $E^T E$ i EE^T) su $\Pi_v = VV^T$, $\Pi_u = UU^T$ a pripadni projektor Π za $M^2 = P^T(EE^T \oplus E^T E)P$ (pa i za M) je

$$\Pi = P^T \begin{pmatrix} \Pi_u & \mathbf{0} \\ \mathbf{0} & \Pi_v \end{pmatrix} P \quad \text{(ovdje je } P = \begin{pmatrix} I & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & I \\ \mathbf{0} & I & \mathbf{0} \end{pmatrix}\text{)}$$

pa je za vektor jedinica e (dimenzija od e iz konteksta!)

$$\Pi e = P^T \begin{pmatrix} \Pi_u & \mathbf{0} \\ \mathbf{0} & \Pi_v \end{pmatrix} Pe = P^T \begin{pmatrix} \Pi_u & \mathbf{0} \\ \mathbf{0} & \Pi_v \end{pmatrix} e = \begin{pmatrix} I & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & I \\ \mathbf{0} & I & \mathbf{0} \end{pmatrix} \begin{pmatrix} \Pi_u e \\ \Pi_v e \end{pmatrix}$$

Nastavak dokaza teorema o indeksu centralnosti

Dakle, imamo

$$\Pi e = \begin{pmatrix} I & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & I \\ \mathbf{0} & I & \mathbf{0} \end{pmatrix} \begin{pmatrix} \Pi_u e \\ \Pi_v e \end{pmatrix} = \begin{pmatrix} \star \\ \Pi_v e \\ * \end{pmatrix}.$$

Ako nas zanimaju samo indeksi centralnosti, oni su zapravo jednaki $\Pi_v e$, i mozemo ih dobiti koristeći dominantni svojstveni potprostor matrice $E^T E = G_r^T G_r + G_r G_r^T$, koja je puno (u ovom slučaju tri puta) manje dimenzije od dimenzije matrice M.

Analogno možemo analizirati i problem hubova i autoriteta, i to ostavljam za vježbu.

Literatura

Literatura

- Jon Kleinberg: *Authoritative sources in a hyperlinked environment*, Journal of the ACM. 46 (5): 604-632, 1999.
- Chris H. Q. Ding, Hongyuan Zha, Xiaofeng He, Horst D. Simon: *Link Analysis: Hubs and Authorities on the World Wide Web*, SIAM Review 46(2), 2002.
- Vincent D. Blondel, Anahí Gajardo, Maureen Heymans, Pierre Senellart, Paul Van Dooren: *A Measure of Similarity between Graph Vertices: Applications to Synonym Extraction and Web Searching*, SIAM Review 46(4), 647-666, 2004.
- `https://patents.google.com/patent/US6112202`