Googlen PageRankin matematiikka

Juulia Hirvelä

Matematiikan pro gradu

Jyväskylän yliopisto
Matematiikan ja tilastotieteen laitos
Kevät 2023

Tiivistelmä: Juulia Hirvelä, *Googlen PageRankin matematiikka* (engl. *Mathematics of Google's PageRank*), matematiikan pro gradu -tutkielma, 31 sivua., Jyväskylän yliopisto, Matematiikan ja tilastotieteen laitos, kevät 2023.

Tämän tutkielman tarkoituksena on esitellä matematiikkaa, johon Googlen hakutulosten järjestämiseen käyttämä PageRank-algoritmi perustuu. Tutkielmassa hyödynnetään lineaarialgebran, graafien ja Markovin ketjujen teorian yhteyksiä, joiden avulla algoritmin matemaattinen muoto saadaan esitettyä.

Hyvä hakukone tarjoaa hakutuloksissaan ensimmäisenä sellaisia sivuja, jotka ovat netinselaajan mielestä hyödyllisiä. Google-hakukone käyttää sivun tärkeyden selvittämiseen omaa PageRank-algoritmiaan, joka laskee jokaiselle nettisivulle tärkeysarvon eli PageRankin. Hakutulokset järjestetään tämän arvon perusteella suuruusjärjestykseen. Sivun PageRank määräytyy sivulle johtavien linkkien määrästä ja viittaavien sivujen tärkeydestä.

Algoritmi perustuu World Wide Webin linkkirakenteen esittämiseen suunnattuna graafina. Graafin informaatiosta muodostetaan matriisi, jonka itseisarvoltaan suurinta ominaisarvoa vastaava ominaisvektori on niin sanottu PageRank-vektori. Algoritmin tarkoituksena on saada selville tämä vektori, koska se pitää sisällään sivujen PageRankit. Ominaisarvon ja sitä vastaavaa vektorin laskeminen suoraan matriisin ominaisarvoyhtälöstä olisi kuitenkin liian työlästä, joten tätä varten työssä esitellään iteratiivinen prosessi nimeltä potenssimenetelmä, jossa mielivaltaisesti valittua aloitusvektoria kerrotaan toistuvasti edellä muodostetulla matriisilla.

Jotta menetelmää voidaan hyödyntää, täytyy ensin varmistua siitä, että sen avulla laskettava vektorijono suppenee. Tämä asettaa edellä muodostetulle matriisille tietyt vaatimukset. Tutkielmassa huomataan, että mikäli matriisi on muistittoman stokastisen prosessin primitiivinen siirtymämatriisi, potenssimenetelmällä muodostettu vektorijono suppenee kohti PageRank-vektoria, joka vastaa itse asiassa kyseisen stokastisen prosessin tasapainotilaa. Jonon suppenemisen ehtoja selvittäessä tutustutaan muun muassa redusoituviin matriiseihin, Perronin ja Frobeniuksen lauseeseen sekä Markovin ketjuihin.

Sisällys

Johdanto 1

Luku 1. Matriisien ja vektoreiden perusominaisuuksia 3
1.1. Merkintöjä 3
1.2. Ominaisarvot 3
1.3. Normit 6
1.3.1. Matriisinormi 7

Luku 2. Hakukoneen matematiikka 10
2.1. Lyhyesti verkkoteoriasta 10
2.2. Erityisiä matriiseja sekä Perronin ja Frobeniuksen lause 12
2.3. Potenssimenetelmä 17
2.4. Markovin ketjut 19

Luku 3. PageRank 23
3.1. Yhteenlaskukaava 23
3.2. Yhteenlaskukaavan matriisiesitys 23
3.3. Stokastisen ja primitiivisen matriisin muodostaminen 25

Liite A. Perronin ja Frobeniuksen lauseen todistus 27

Lähdeluettelo 31

Johdanto

Internetistä on muodostunut yksi tärkeimmistä tietolähteistä. Tietoa on tarjolla valtavasti ja se on jatkuvasti saatavilla. Tiedon löytämisessä hyödynnetään internetin hakukoneita, jotka poimivat hakusanaa vastaavia tuloksia nettisivujen kokoelmistaan. Tiedonjanon täyttymisessä keskeistä on se, kuinka pian hakutuloksia selatessaan netinselaaja löytää juuri sen itselleen oleellisimman sivun. Tämä oleellisimmaksi valittu sivu on usein myös muiden samasta aiheesta kiinnostuneiden mielestä se relevantein sivu. Tätä ajatusta myös hakukoneiden laatijat ovat hyödyntäneet suunnitellessaan tehokasta työkalua hakutulosten järjestämiseen. Hyvä hakukone tarjoaa hakutuloslistauksessaan ensimmäisenä kaikkein relevanteinta sivua. Jotta hakukone voisi täyttää tämän vaatimuksen, täytyy ensin määritellä, mitä sivun relevanssi tarkoittaa sekä keksiä, kuinka se pystytään määrittämään jokaiselle nettisivulle.

Suosituimman hakukoneen Googlen hakutulosten järjestys perustuu useisiin eri tekijöihin, joista yksi on hakukoneen oma algoritmi, PageRank. Algoritmin avulla kullekin sivulle lasketaan tärkeysarvo. Tulokset listataan tärkeysarvon perusteella suuruusjärjestykseen niin, että suurimman arvon saanut sivu on listan ensimmäisenä. Sivun tärkeysarvo eli PageRank määräytyy karkeasti siitä, kuinka monta linkkiä muilta nettisivuilta johtaa kyseiselle sivulle. Toisin sanoen jos sivua suositellaan paljon muilla sivuilla, se lasketaan tärkeäksi. Myös suositukset painotetaan niin, että tärkeän sivun antama suositus on arvokkaampi. Tässä työssä esitellään PageRank-algoritmin taustalla olevaa matematiikkaa ja lopulta muotoillaan algoritmin matemaattinen esitys.

Nettisurffailun mallintamiseen käytetään algoritmin kehittäjien keksimää ajatusta satunnaisesta netinselaajasta, joka liikkuu pitkin World Wide Webin linkkirakennetta valiten seuraavan klikattavan linkkinsä täysin sattumanvaraisesti. Tällöin jos nettisivulla on useampi linkki, netinselaajalla on sivulle päädyttyään yhtä suuri todennäköisyys valita mikä tahansa linkeistä. Mallinnus on muistiton, koska seuraavan linkin valinta ei riipu siitä, millä sivuilla netinselaaja on aiemmin vieraillut. PageRank-arvo muodostuu pitkällä aikavälillä siitä aikaosuudesta, jonka netinselaaja yhteensä viettää kullakin sivulla.

Algoritmin muotoilua varten nettisivut ja niiden väliset suhteet eli linkit täytyy muuttaa matemaattiseen muotoon. Tätä varten työssä käydään läpi graafiteoriaa, jonka avulla World Wide Web saadaan esitettyä suunnattuna graafina, jossa solmut edustavat nettisivuja ja nuolet sivujen välisiä linkkejä. Jos graafi täyttää tietyt edellytykset, satunnaisen netinselaajan liikkuminen pitkin tätä graafia on stokastinen prosessi, tarkemmin Markovin ketju. Prosessin tasapainotila vastaa netinselaajan sivulla viettämää aikaosuutta. PageRank saadaan siis laskettua selvittämällä tämä tilajakauma, mutta huomioitavaa on, että kaikki Markovin ketjut eivät kuitenkaan saavuta tasapainotilaa. Tämä asettaa graafille tai oikeammin sen matriisivastineelle erityisiä vaatimuksia, joita johdettaessa tutustutaan muun muassa redusoituviin ja primitiivisiin matriiseihin sekä Perronin ja Frobeniuksen lauseeseen.

Työssä käsiteltävä matematiikka on rajattu niin, että se riittää PageRank-algoritmin ymmärtämiseen. Joidenkin väittämien todistukset sivuutetaan, koska niissä vaadittava osaaminen ei ole algoritmin ymmärtämisen kannalta oleellista. Tällaisen väittämän yhteydessä on kuitenkin viite teokseen, josta todistus on löydettävissä. Luvussa 1 kerrataan lyhyesti lineaarialgebran perusteista työn kannalta tärkeitä asioita. Lisäksi luvussa tutustutaan matriisin vasemmanpuoleisiin ominaisvektoreihin ja normeihin sekä määritellään matriisin spektri ja spektraalisäde. Luvun 2 ensimmäisessä osiossa käydään lyhyesti läpi graafiteorian alkeita ja opetellaan liikkumaan graafin ja sen matriisiesityksen välillä. Seuraavaksi määritellään erilaisia matriiseja pitäen edelleen mielessä matriisin yhteys graafiin. Työn kannalta yksi tärkeimmistä tuloksista esitetään lauseessa 2.16, joka on nimetty kehittäjiensä mukaan Perronin ja Frobeniuksen lauseeksi. Luvun 2 lopussa esitellään potenssimenetelmä sekä Markovin ketjut. Luvun 2 viimeinen lause (2.24) muotoilee ehdot, joiden täyttyessä PageRank-arvot saadaan laskettua. Viimeisessä luvussa esitellään algoritmin matemaattinen muoto ja osoitetaan, että sen avulla saadaan todella selvitettyä jokaiselle sivulle PageRank-arvo.

Työ pohjautuu pääosin Langvillen ja Meyerin [**1**] teokseen, josta myös PageRank-algoritmin muotoilu on peräisin. Lineaarialgebran tuloksissa on lisäksi tukeuduttu kirjaan [**2**] ja graafiteorian täydentämisessä on auttanut [**3**].

LUKU 1

Matriisien ja vektoreiden perusominaisuuksia

1.1. Merkintöjä

Tässä työssä käsiteltävät matriisit ovat pääasiassa reaalisia ja usein epänegatiivisia, mutta näiden matriisien ominaisarvoissa saattaa esiintyä ei-reaalisia arvoja, joten pitää huomioida myös kompleksisten vektoreiden ja matriisien mahdollisuus. Kompleksilukujen tunteminen ei kuitenkaan ole välttämätöntä työssä tarkasteltavan matematiikan ymmärtämiseksi.

Matriiseja merkitään lihavoiduilla isoilla kirjaimilla. Matriisissa $\boldsymbol{A}_{m\times n}$ on m riviä ja n saraketta. Merkinnällä a_{ij} viitataan matriisin $\boldsymbol{A}$ i. rivin ja j. sarakkeen alkioon. Jos $m \times n$-matriisi sisältää ainoastaan reaalisia alkioita, merkitään $\boldsymbol{A} \in \mathbb{R}^{m\times n}$.

Vektori on matriisi, jossa on vain yksi rivi tai sarake. Yhden sarakkeen matriisi

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}$$

on pystyvektori, jolle voidaan käyttää myös merkintää $\boldsymbol{x} = (x_1, x_2, \ldots, x_m)$. Pystyvektorin transpoosi on $1 \times m$-matriisi $\boldsymbol{x}^\top = [x_1 \quad x_2 \quad \cdots \quad x_m]$, jota sanotaan vaakavektoriksi. Tässä työssä pysty- ja vaakavektoria merkitään lihavoidulla pienellä kirjaimella tai symbolilla, mutta vaakavektorin merkinnässä on lisäksi mukana matriisin transpoosia merkitsevä oikean yläindeksin $\top$.

Jos matriisin kaikki alkiot ovat nollia, matriisia sanotaan nollamatriisiksi ja merkitään $\mathbf{0}$. Jos matriisin $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ alkiot ovat positiivisia, merkitään $\boldsymbol{A} > \mathbf{0}$. Jos taas matriisin alkiot ovat epänegatiivisia, merkitään tällöin $\boldsymbol{A} \geq \mathbf{0}$. Vastaavia merkintöjä käytetään vektoreille.

1.2. Ominaisarvot

Ominaisvektorin kertominen matriisilla vastaa vektorin kertomista skalaarilla.

Määritelmä 1.1. Luku $\lambda \in \mathbb{C}$ on neliömatriisin $\boldsymbol{A}$ (oikeanpuoleinen) ominaisarvo, jos on olemassa vektori $\boldsymbol{x} \neq \mathbf{0}$, jolle

$$\boldsymbol{A}\boldsymbol{x} = \lambda \boldsymbol{x}.$$

Vektoria $\boldsymbol{x}$ kutsutaan ominaisarvoa λ vastaavaksi (oikeanpuoleiseksi) ominaisvektoriksi.

Matriisin $\boldsymbol{A}_{n\times n}$ ominaisarvot ovat karakteristisen yhtälön $\det(\boldsymbol{A} - \lambda \boldsymbol{I}) = 0$ ratkaisut, joita on ratkaisujen kertaluvut huomioiden n kappaletta.

Esimerkki 1.2. a) Olkoon
$$\boldsymbol{A} = \begin{bmatrix} 4 & 2 \\ 1 & 3 \end{bmatrix}.$$
Lasketaan matriisin $\boldsymbol{A}$ ominaisarvot.

Selvitetään karakteristisen yhtälön $\det(\boldsymbol{A} - \lambda \boldsymbol{I}) = 0$ ratkaisut.
$$\begin{aligned} \det(\boldsymbol{A} - \lambda \boldsymbol{I}) &= \det\left(\begin{bmatrix} 4-\lambda & 2 \\ 1 & 3-\lambda \end{bmatrix}\right) = (4-\lambda)(3-\lambda) - 2 \\ &= \lambda^2 - 7\lambda + 10 = (\lambda - 2)(\lambda - 5). \end{aligned}$$
Yhtälön ratkaisut ovat $\lambda = 2$ ja $\lambda = 5$, joten matriisin $\boldsymbol{A}$ ominaisarvot ovat 2 ja 5.

b) Olkoon
$$\boldsymbol{B} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}.$$
Lasketaan matriisin $\boldsymbol{B}$ ominaisarvot kuten a-kohdassa. Matriisin karakteristinen yhtälö on $\lambda^2 - 2\lambda + 2 = 0$. Toisen asteen yhtälön ratkaisukaavalla saadaan
$$\lambda = \frac{2 \pm \sqrt{-4}}{2} = \frac{2 \pm 2i}{2} = 1 \pm i.$$
Matriisin $\boldsymbol{B}$ ominaisarvot ovat siis kompleksiluvut $1 + i$ ja $1 - i$.

Edellisen määritelmän ominaisvektoria kutsutaan joissain tilanteissa oikeanpuoleiseksi ominaisvektoriksi, koska matriisilla on myös vasemmanpuoleinen ominaisvektori, joka on vaakavektori.

Määritelmä 1.3. Luku $\lambda \in \mathbb{C}$ on neliömatriisin $\boldsymbol{A}$ vasemmanpuoleinen ominaisarvo, jos on olemassa vektori $\boldsymbol{y}^\top \neq \boldsymbol{0}$, jolle
$$\boldsymbol{y}^\top \boldsymbol{A} = \lambda \boldsymbol{y}^\top.$$
Vektoria $\boldsymbol{x}$ kutsutaan ominaisarvoa λ vastaavaksi vasemmanpuoleiseksi ominaisvektoriksi.

Oikean- ja vasemmanpuoleisen ominaisvektorin ominaisarvot voidaan erotella vastaavalla tavalla, mutta seuraavaksi huomataan, että tätä jaottelua ei tarvita.

Lause 1.4. *Neliömatriisin oikean- ja vasemmanpuoleiset ominaisarvot ovat samat.*

Todistus. Olkoon $\boldsymbol{y}^\top$ matriisin $\boldsymbol{A}$ vasemmanpuoleista ominaisarvoa λ vastaava vasemmanpuoleinen ominaisvektori. Määritelmän mukaan
$$\begin{aligned} & \boldsymbol{y}^\top \boldsymbol{A} = \lambda \boldsymbol{y}^\top \\ \Leftrightarrow \quad & \boldsymbol{y}^\top \boldsymbol{A} - \boldsymbol{y}^\top \lambda \boldsymbol{I} = \boldsymbol{0} \\ \Leftrightarrow \quad & \boldsymbol{y}^\top (\boldsymbol{A} - \lambda \boldsymbol{I}) = \boldsymbol{0}. \end{aligned}$$
Otetaan yhtälöstä puolittain transpoosi, jolloin ominaisuudesta $(\boldsymbol{A}\boldsymbol{B})^\top = \boldsymbol{B}^\top \boldsymbol{A}^\top$ seuraa
$$(\boldsymbol{A} - \lambda \boldsymbol{I})^\top \boldsymbol{y} = \boldsymbol{0}.$$

Ominaisarvo λ on karakteristisen yhtälön $\det((\boldsymbol{A}-\lambda\boldsymbol{I})^\top)=0$ ratkaisu. Determinantille pätee $\det(\boldsymbol{B})=\det(\boldsymbol{B}^\top)$, joten $\det((\boldsymbol{A}-\lambda\boldsymbol{I})^\top)=\det(\boldsymbol{A}-\lambda\boldsymbol{I})$. Vasemmanpuoleiset ominaisarvot ovat siten saman yhtälön $\det(\boldsymbol{A}-\lambda\boldsymbol{I})=0$ ratkaisuja kuin oikeanpuoleiset ominaisarvotkin. $\square$

Huomautus 1.5. Edellisessä todistuksessa huomattiin, että yhtälöt

$$\boldsymbol{y}^\top(\boldsymbol{A}-\lambda\boldsymbol{I})=\boldsymbol{0}$$

ja

$$(\boldsymbol{A}-\lambda\boldsymbol{I})^\top\boldsymbol{y}=\boldsymbol{0}$$

ovat yhtäpitäviä. Koska $(\boldsymbol{A}-\lambda\boldsymbol{I})^\top=\boldsymbol{A}^\top-\lambda\boldsymbol{I}$, niin ominaisarvoa λ vastaava vasemmanpuoleinen ominaisvektori $\boldsymbol{y}^\top$ voidaan ratkaista myös yhtälön

$$(\boldsymbol{A}^\top-\lambda\boldsymbol{I})\boldsymbol{y}=\boldsymbol{0}$$

avulla. Vektori $\boldsymbol{y}^\top$ on yhtälön pystyvektorin $\boldsymbol{y}$ transpoosi. Lisäksi yhtälöstä havaitaan, että vektori $\boldsymbol{y}$ on matriisin $\boldsymbol{A}$ transpoosin $\boldsymbol{A}^\top$ ominaisarvoa λ vastaava oikeanpuoleinen ominaisvektori.

Jos matriisi on symmetrinen, yhtä ominaisarvoa vastaavat oikean- ja vasemmanpuoleinen ominaisvektori ovat toistensa transpooseja, koska tällöin $\boldsymbol{A}=\boldsymbol{A}^\top$. Muussa tapauksessa näin ei välttämättä ole, kuten huomataan seuraavassa esimerkissä.

Esimerkki 1.6. Lasketaan matriisin

$$\boldsymbol{A}=\begin{bmatrix}4 & 2\\ 1 & 3\end{bmatrix}$$

ominaisarvoa $\lambda=2$ vastaava oikean- ja vasemmanpuoleinen ominaisvektori.

Oikeanpuoleinen ominaisvektori $\boldsymbol{x}$ toteuttaa yhtälön $\boldsymbol{A}\boldsymbol{x}=2\boldsymbol{x}$ yhtälön eli

$$(\boldsymbol{A}-2\boldsymbol{I})\boldsymbol{x}=\boldsymbol{0}.$$

Nyt

$$(\boldsymbol{A}-2\boldsymbol{I})\boldsymbol{x}=\begin{bmatrix}2 & 2\\ 1 & 1\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}=\begin{bmatrix}0\\ 0\end{bmatrix}\Leftrightarrow\begin{cases}2x_1+2x_2 & =0\\ x_1+x_2 & =0.\end{cases}$$

Yhtälö toteutuu, kun $x_2=-x_1, x_1\in\mathbb{R}$, joten ominaisarvoa $\lambda=2$ vastaavat oikeanpuoleiset ominaisvektorit ovat vektorin $\boldsymbol{x}=(1,-1)$ monikerrat.

Vasemmanpuoleinen ominaisvektori $\boldsymbol{y}^\top$ toteuttaa yhtälön $\boldsymbol{y}^\top\boldsymbol{A}=2\boldsymbol{y}^\top$ eli

$$\boldsymbol{y}^\top(\boldsymbol{A}-2\boldsymbol{I})=\boldsymbol{0},$$

joka saadaan transponoimalla muotoon

$$\begin{aligned}&(\boldsymbol{A}-2\boldsymbol{I})^\top\boldsymbol{y}=\boldsymbol{0}\\ \Leftrightarrow\quad&\begin{bmatrix}2 & 1\\ 2 & 1\end{bmatrix}\begin{bmatrix}y_1\\ y_2\end{bmatrix}=\begin{bmatrix}0\\ 0\end{bmatrix}.\end{aligned}$$

Yhtälö totetuu, kun $y_2=-2y_1, y_1\in\mathbb{R}$, joten ominaisarvoa $\lambda=2$ vastaavat vasemmanpuoleiset ominaisvektorit ovat vektorin $\boldsymbol{y}^\top=[1\quad -2]$ monikerrat.

Määritelmä 1.7. Neliömatrisiin $\boldsymbol{A}$ ominaisarvojen joukkoa merkitään $\sigma(\boldsymbol{A})$ ja kutsutaan matriisin $\boldsymbol{A}$ *spektriksi*.

MÄÄRITELMÄ 1.8. Neliömatriisin $\boldsymbol{A}$ spektraalisäde on

$$\rho(\boldsymbol{A}) = \max_{\lambda \in \sigma(\boldsymbol{A})} |\lambda|.$$

1.3. Normit

Yleisin tapa mitata vektorin suuruutta on sen Euklidisen normin eli pituuden määrittäminen. Vektorin $\boldsymbol{x} \in \mathbb{R}^n$ pituus on

$$\|\boldsymbol{x}\|_2 = \sqrt{\sum_{i=1}^{n} x_i^2}.$$

Pituus ei kuitenkaan sovellu vektorin suuruuden mitaksi kaikissa tapauksissa, joten määritellään tätä varten vektorin p-normi.

MÄÄRITELMÄ 1.9. Vektorin $\boldsymbol{x} \in \mathbb{R}^n$ p-normi määritellään kaavalla

$$\|\boldsymbol{x}\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{\frac{1}{p}},$$

missä $p \geq 1$.

Euklidinen normi on p-normin erityistapaus, kun $p = 2$. Kaksi muuta tässä työssä hyödyllistä erityistapausta ovat 1-normi ja maksiminormi. Vektorin 1-normi saadaan laskemalla vektorin $\boldsymbol{x} \in \mathbb{R}^n$ alkioiden itseisarvojen summa:

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

Maksiminormi saadaan, kun p-normin määritelmässä $p \to \infty$, jolloin kaava supistuu muotoon

$$\|\boldsymbol{x}\|_\infty = \max_{1 \leq i \leq n} |x_i|.$$

Vektorin p-normi täyttää yleisen normin määritelmässä esitetyt ehdot, jotka on muotoiltu seuraavaan lauseeseen.

LAUSE 1.10. *Olkoot $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^n$ ja olkoon $\alpha \in \mathbb{C}$. Tällöin*

(i) $\|\boldsymbol{x}\| \geq 0$,
(ii) $\|\boldsymbol{x}\| = 0 \Leftrightarrow \boldsymbol{x} = \boldsymbol{0}$,
(iii) $\|\alpha \boldsymbol{x}\| = |\alpha| \|\boldsymbol{x}\|$,
(iv) $\|\boldsymbol{x} + \boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$.

Vektorin p-normin määritelmä ja ominaisuudet pätevät sekä pysty- että vaakavektoreille. Lisäksi $\|\boldsymbol{x}\| = \|\boldsymbol{x}^\top\|$, kun $\boldsymbol{x}$ on pystyvektori.

1.3.1. Matriisinormi. Myös matriisien suuruutta voidaan mitata matriiseille määritellyllä normilla. Hyödynnetään matriisinormin määrittelyssä vektorin p-normia, jolloin saadaan vektorinormin indusoima matriisinormi.

MÄÄRITELMÄ 1.11. Matriisin $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ vektorinormin indusoima matriisinormi on

$$\|\boldsymbol{A}\|_p = \max_{\|\boldsymbol{x}\|_p=1} \|\boldsymbol{A}\boldsymbol{x}\|_p,$$

missä $p \geq 1$ ja $\boldsymbol{x} \in \mathbb{R}^{n\times 1}$.

Aiemmin esiteltyjen vektorinormin erityistapausten indusoimat matriisinormit pystytään johtamaan yksinkertaisempaan muotoon.

LAUSE 1.12. *Vektorin 1-normin indusoima matriisinormi on matriisin alkioiden itseisarvojen suurin sarakesumma. Matriisille* $\boldsymbol{A} \in \mathbb{R}^{m\times n}$

$$\|\boldsymbol{A}\|_1 = \max_{1\leq j\leq n} \sum_{i=1}^{m} |a_{ij}|.$$

Maksiminormin indusoima matriisinormi on matriisin alkioiden itseisarvojen suurin rivisumma.

$$\|\boldsymbol{A}\|_\infty = \max_{1\leq i\leq m} \sum_{j=1}^{n} |a_{ij}|.$$

TODISTUS. Osoitetaan ensin 1-normin indusoiman matriisinormin kaava. Olkoon $\boldsymbol{x} \in \mathbb{R}^{n\times 1}$ siten, että $\|\boldsymbol{x}\|_1 = 1$. Tällöin

$$\|\boldsymbol{A}\boldsymbol{x}\|_1 = \sum_{i=1}^{m} \left| \sum_{j=1}^{n} a_{ij}x_j \right| \leq \sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}x_j| = \sum_{j=1}^{n} \left(|x_j| \sum_{i=1}^{m} |a_{ij}| \right).$$

Valitaan $\tilde{j}$ siten, että

$$\max_{1\leq j\leq n} \sum_{i=1}^{m} |a_{ij}| = \sum_{i=1}^{m} |a_{i\tilde{j}}|.$$

Nyt

$$\sum_{j=1}^{n} \left(|x_j| \sum_{i=1}^{m} |a_{ij}| \right) \leq \sum_{j=1}^{n} |x_j| \sum_{i=1}^{m} |a_{i\tilde{j}}| = \|\boldsymbol{x}\|_1 \sum_{i=1}^{m} |a_{i\tilde{j}}| = \sum_{i=1}^{m} |a_{i\tilde{j}}|.$$

Niinpä $\max_{\|\boldsymbol{x}\|_1=1} \|\boldsymbol{A}\boldsymbol{x}\|_1 \leq \max_{1\leq j\leq n} \sum_{i=1}^{m} |a_{ij}|$. Lisäksi

$$\sum_{i=1}^{m} |a_{i\tilde{j}}| = \|\boldsymbol{A}\boldsymbol{e}_{\tilde{j}}\|_1 \leq \max_{\|\boldsymbol{x}\|_1=1} \|\boldsymbol{A}\boldsymbol{x}\|_1,$$

missä $\boldsymbol{e}_{\tilde{j}}$ on $\tilde{j}$. standardikannan vektori. Kun yhdistetään tulokset, saadaan

$$\max_{\|\boldsymbol{x}\|_1=1} \|\boldsymbol{A}\boldsymbol{x}\|_1 \leq \max_{1\leq j\leq n} \sum_{i=1}^{m} |a_{ij}| \leq \max_{\|\boldsymbol{x}\|_1=1} \|\boldsymbol{A}\boldsymbol{x}\|_1,$$

joten

$$\|\boldsymbol{A}\|_1 = \max_{\|\boldsymbol{x}\|_1=1} \|\boldsymbol{A}\boldsymbol{x}\|_1 = \max_{1\leq j\leq n} \sum_{i=1}^{m} |a_{ij}|.$$

Osoitetaan sitten maksiminormin indusoiman matriisinormin kaava. Olkoon nyt $\boldsymbol{x} \in \mathbb{R}^{n\times 1}$ niin, että $\|\boldsymbol{x}\|_\infty = 1$. Nyt

$$\begin{aligned}\|\boldsymbol{Ax}\|_\infty &= \max_{1\le i\le m}\left|\sum_{j=1}^n a_{ij}x_j\right| \le \max_{1\le i\le m}\sum_{j=1}^n |a_{ij}x_j| = \max_{1\le i\le m}\sum_{j=1}^n |a_{ij}||x_j| \\ &\le \max_{1\le i\le m}\sum_{j=1}^n |a_{ij}| \max_{1\le k\le n}|x_k| = \max_{1\le i\le m}\sum_{j=1}^n |a_{ij}|\|\boldsymbol{x}\|_\infty = \max_{1\le i\le m}\sum_{j=1}^n |a_{ij}|.\end{aligned}$$

Valitaan $\tilde{i}$ siten, että

$$\max_{1\le i\le m}\sum_{j=1}^n |a_{ij}| = \sum_{j=1}^n |a_{\tilde{i}j}|.$$

Olkoon $\boldsymbol{y} \in \mathbb{R}^n$ niin, että

$$y_j = \begin{cases} 1, & \text{jos } a_{\tilde{i}j} \ge 0 \\ -1, & \text{jos } a_{\tilde{i}j} < 0, \end{cases}$$

jolloin $\|\boldsymbol{y}\|_\infty = 1$. Tällöin

$$\sum_{j=1}^n |a_{\tilde{i}j}| = \sum_{j=1}^n a_{\tilde{i}j}y_j \le \max_{\|\boldsymbol{x}\|_\infty=1}\|\boldsymbol{Ax}\|_\infty.$$

Niinpä

$$\max_{\|\boldsymbol{x}\|_\infty=1}\|\boldsymbol{Ax}\|_\infty \le \max_{1\le i\le m}\sum_{j=1}^n |a_{ij}| \le \max_{\|\boldsymbol{x}\|_\infty=1}\|\boldsymbol{Ax}\|_\infty,$$

joten

$$\|\boldsymbol{A}\|_\infty = \max_{\|\boldsymbol{x}\|_\infty=1}\|\boldsymbol{Ax}\|_\infty = \max_{1\le i\le m}\sum_{j=1}^n |a_{ij}|.$$

□

Jos $\boldsymbol{x} \ne \boldsymbol{0}$, niin matriisinormin määritelmän avulla saadaan

$$\|\boldsymbol{Ax}\| = \left\|\boldsymbol{A}\frac{\boldsymbol{x}}{\|\boldsymbol{x}\|}\|\boldsymbol{x}\|\right\| = \left\|\boldsymbol{A}\frac{\boldsymbol{x}}{\|\boldsymbol{x}\|}\right\|\|\boldsymbol{x}\| \le \max_{\|\boldsymbol{x}\|=1}\|\boldsymbol{Ax}\|\|\boldsymbol{x}\| = \|\boldsymbol{A}\|\|\boldsymbol{x}\|.$$

Epäyhtälö $\|\boldsymbol{Ax}\| \le \|\boldsymbol{A}\|\|\boldsymbol{x}\|$ toteutuu myös silloin, kun $\boldsymbol{x} = \boldsymbol{0}$. Tällöin sanotaan, että matriisinormi on yhteensopiva sen indusoineen vektorinormin kanssa. Matriisinormi siis kertoo, kuinka paljon vektori enintään venyy, kun sitä kerrotaan matriisilla.

Neliömatriisin vektorinormilla indusoitua matriisinormia rajoittaa alhaalta matriisin spektraalisäde.

Lause 1.13. *Olkoon $\boldsymbol{A} \in \mathbb{R}^{n\times n}$. Tällöin p-normin indusoimilla matriisinormeilla*

$$\rho(\boldsymbol{A}) \le \|\boldsymbol{A}\|.$$

Todistus. Olkoon λ matriisin $\boldsymbol{A}$ ominaisarvo, jolla $|\lambda| = \rho(\boldsymbol{A})$, ja olkoon $\boldsymbol{x} \in \mathbb{R}^{n\times 1}$ sitä vastaava ominaisvektori. Ominaisarvoyhtälöstä $\boldsymbol{Ax} = \lambda\boldsymbol{x}$ ja matriisinormin ja sen indusoineen vektorinormin yhteensopivuudesta seuraa

$$\begin{aligned} & |\lambda|\|\boldsymbol{x}\| = \|\lambda\boldsymbol{x}\| = \|\boldsymbol{Ax}\| \leq \|\boldsymbol{A}\|\|\boldsymbol{x}\| \\ \Leftrightarrow \quad & |\lambda|\|\boldsymbol{x}\| \leq \|\boldsymbol{A}\|\|\boldsymbol{x}\| \\ \Leftrightarrow \quad & |\lambda| \leq \|\boldsymbol{A}\|. \end{aligned}$$

Koska $|\lambda| = \rho(\boldsymbol{A})$, niin $\rho(\boldsymbol{A}) \leq \|\boldsymbol{A}\|$. $\square$

LUKU 2

Hakukoneen matematiikka

2.1. Lyhyesti verkkoteoriasta

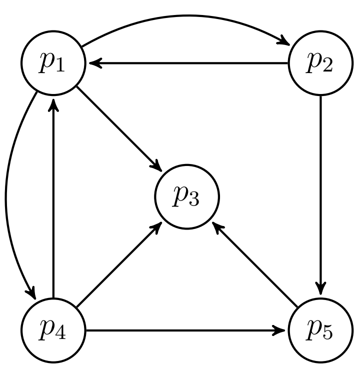
MÄÄRITELMÄ 2.1. Graafi G on järjestetty pari (V, E), missä $V = \{v_1, v_2, \ldots, v_n\}$ on joukko solmuja ja E on joukko solmupareja $\{v_i, v_j\}, i \neq j$, joita kutsutaan kaariksi.

Graafi voidaan esittää kuviona, jossa solmut kuvataan pisteinä ja kaaret solmuja yhdistävinä viivoina. Jos kaaret ovat yksisuuntaisia, puhutaan suunnatusta graafista.

MÄÄRITELMÄ 2.2. Graafi G on suunnattu graafi, jos sen kaaret ovat järjestettyjä solmupareja. Suunnatun graafin kaaria kutsutaan nuoliksi.

Parissa (v_i, v_j) solmu v_i on nuolen alkupiste ja solmu v_j loppupiste. Tällöin $(v_i, v_j) \neq (v_j, v_i)$. Kahden solmun välillä voi olla korkeintaan kaksi nuolta, yksi kumpaankin suuntaan. Suunnatussa graafissa voi olla myös luuppeja, jossa nuolen alkupisteenä ja loppupisteenä on sama solmu.

ESIMERKKI 2.3. World Wide Webin rakenne voidaan esittää suunnattuna graafina, jossa nettisivuja kuvataan solmuilla ja sivujen välisiä linkkejä nuolilla. Jos linkki vie sivulta p_i sivulle p_j, nuoli osoittaa solmusta p_i solmuun p_j.



KUVA 2.1. Graafi viisisivuisesta webistä

Polku on solmuista ja kaarista muodostettu äärellinen jono

$$P = v_{i0}, (v_{i0}, v_{i1}), v_{i1}, \ldots, (v_{i(k-1)}, v_{ik}), v_{ik},$$

jossa solmut ja niitä yhdistävät kaaret vuorottelevat. Polun P alkupiste on solmu v_{i0} ja päätepiste v_{ik}. Polun alkupiste voi olla sama kuin loppupiste, mutta muut graafin solmut esiintyvät polussa korkeintaan kerran. Suunnatussa graafissa nuolia pitkin voidaan kulkea vain nuolen osoittamaan suuntaan.

Määritelmä 2.4. Suunnattu graafi on *vahvasti yhtenäinen*, jos sen jokaisesta solmusta on polku graafin kaikkiin muihin solmuihin.

Kuvan 2.1 graafi ei ole vahvasti yhtenäinen, koska esimerkiksi solmusta p_3 ei ole polkua solmuun p_1.

Graafi voidaan esittää kuvion lisäksi matriisin avulla.

Määritelmä 2.5. Olkoon $G = (V, E)$ suunnattu graafi, missä $V = \{v_1, \ldots, v_n\}$. Graafin vierusmatriisi $\boldsymbol{A}$ on $n \times n$-matriisi, jossa

$$a_{ij} = \begin{cases} 1, & \text{jos } (v_i, v_j) \in E \\ 0, & \text{jos } (v_i, v_j) \notin E. \end{cases}$$

Matriisin ja graafin välillä voidaan liikkua myös toiseen suuntaan, jolloin neliömatriisi tulkitaan graafina. Tällöin $(v_i, v_j) \in E$ jos ja vain jos $a_{ij} \neq 0$.

Määritelmä 2.6. Suunnatussa graafissa solmun v lähtöaste $d(v)$ on niiden nuolien lukumäärä, joiden alkupiste solmu v on.

Esimerkki 2.7. Kuvan 2.1 graafin vierusmatriisi on

$$\begin{bmatrix} 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}.$$

Matriisin toisen rivin ensimmäinen ja viimeinen alkio ovat ykkösiä ja muut ovat nollia, joten solmusta p_2 on nuoli solmuihin p_1 ja p_5. Neljännen sarakkeen ensimmäinen alkio on yksi, joten solmuun p_4 osoittaa vain yksi nuoli.

Solmun p_1 lähtöaste on $d(p_1) = 3$. Sivulla p_1 on siis kolme linkkiä.

Jatkossa tullaan tarvitsemaan rivinormalisoitua vierusmatriisia, jossa jokainen nollasta poikkeava vierusmatriisin $\boldsymbol{A}$ alkio korvataan luvulla $\frac{1}{d(v_i)}$. Käytetään tällaisesta matriisista jatkossa nimitystä linkkimatriisi.

Esimerkki 2.8. Kuvan 2.1 graafin linkkimatriisi $\boldsymbol{H}$ on

$$\boldsymbol{H} = \begin{bmatrix} 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 \\ \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 \\ \frac{1}{3} & 0 & \frac{1}{3} & 0 & \frac{1}{3} \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}.$$

Linkkimatriisissa nollien sijainti on sama kuin graafin vierusmatriisissa. Linkkimatriisin nollasta eroavat luvut voidaan tulkita todennäköisyyksinä, kun oletetaan, että yhdestä solmusta lähtevien nuolten loppupisteisiin siirrytään samalla todennäköisyydellä. Alkio a_{ij} kertoo todennäköisyyden siirtyä solmusta v_i solmuun v_j. Esimerkissä 2.8 sivulta p_1 valitaan linkki sivulle p_4 todennäköisyydellä $\frac{1}{3}$ eli $\frac{1}{3}$ klikkauksista sivulla p_1 johtaa sivulle p_4. Sivulta p_5 taas päädytään aina sivulle p_3, koska $h_{53} = 1$.

2.2. Erityisiä matriiseja sekä Perronin ja Frobeniuksen lause

MÄÄRITELMÄ 2.9. Permutaatiomatriisi $\boldsymbol{P}$ on neliömatriisi, joka muodostetaan vaihtamalla identiteettimatriisin $\boldsymbol{I}$ rivien järjestystä.

Identiteettimatriisin $\boldsymbol{I}_{n\times n}$ rivit voidaan järjestää $n!$ eri järjestykseen, joten $n \times n$-kokoisia permutaatiomatriiseja on $n!$ kappaletta.

LAUSE 2.10. *Matriisin $\boldsymbol{A}$ kertominen permutaatiomatriisilla $\boldsymbol{P}$ vaihtaa matriisin $\boldsymbol{A}$ rivien ja sarakkeiden järjestystä seuraavasti:*

- *Vasemmalta kertominen muuttaa matriisin $\boldsymbol{A}$ rivien järjestyksen vastaamaan permutaatiomatriisin rivien järjestystä.*
- *Oikealta kertominen muuttaa matriisin $\boldsymbol{A}$ sarakkeiden järjestyksen vastaamaan permutaatiomatriisin sarakkeiden järjestystä.*

Lauseen todistus on löydettävissä teoksesta [**5**].

ESIMERKKI 2.11. Lasketaan matriisitulo $\boldsymbol{AP}$, missä

$$\boldsymbol{A} = \begin{bmatrix} 0 & 1 & 2 \\ 0 & 3 & 0 \\ 4 & 5 & 6 \end{bmatrix} \text{ ja } \boldsymbol{P} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Kun permutaatiomatriisin $\boldsymbol{P}$ sarakkeiden järjestystä verrataan identiteettimatriisiin, huomataan, että niiden kolmas sarake on sama, mutta matriisissa $\boldsymbol{P}$ ensimmäisen ja toisen sarakkeen paikkaa on vaihdettu. Merkitään matriisin $\boldsymbol{A}$ j:nnettä saraketta merkinnällä $\boldsymbol{a}_j$, ja merkitään permutaatiomatriisin $\boldsymbol{P}$ sarakkeita standardikannan vektorien merkinnällä $\boldsymbol{e}_j$. Tällöin

$$\boldsymbol{AP} = \begin{array}{c} \begin{array}{ccc} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \boldsymbol{a}_3 \end{array} \\ \begin{bmatrix} 0 & 1 & 2 \\ 0 & 3 & 0 \\ 4 & 5 & 6 \end{bmatrix} \end{array} \begin{array}{c} \begin{array}{ccc} \boldsymbol{e}_2 & \boldsymbol{e}_1 & \boldsymbol{e}_3 \end{array} \\ \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{array} = \begin{array}{c} \begin{array}{ccc} \boldsymbol{a}_2 & \boldsymbol{a}_1 & \boldsymbol{a}_3 \end{array} \\ \begin{bmatrix} 1 & 0 & 2 \\ 3 & 0 & 0 \\ 5 & 4 & 6 \end{bmatrix} \end{array}.$$

Tulomatriisi koostuu siis samoista sarakkeista kuin matriisi $\boldsymbol{A}$, mutta ensimmäinen ja toinen sarake ovat vaihtaneet paikkaa. Kolmas sarake on pysynyt entisellään. Tulomatriisin sarakkeet vastaavat siten permutaatiomatriisin $\boldsymbol{P}$ sarakkeiden järjestystä suhteessa identiteettimatriisiin. Sama tulos saadaan, kun hyödynnetään lausetta 2.10.

MÄÄRITELMÄ 2.12. Neliömatriisin $\boldsymbol{A}$ on sanotaan olevan redusoituva, jos on olemassa permutaatiomatriisi $\boldsymbol{P}$ siten, että

$$\boldsymbol{P}^\top \boldsymbol{AP} = \begin{bmatrix} \boldsymbol{X} & \boldsymbol{Y} \\ \boldsymbol{0} & \boldsymbol{Z} \end{bmatrix},$$

missä $\boldsymbol{X}$ ja $\boldsymbol{Z}$ ovat neliömatriiseja. Matriisi $\boldsymbol{A}$ on redusoitumaton, jos se ei ole redusoituva.

ESIMERKKI 2.13. Olkoon matriisi $\boldsymbol{A}$ kuten esimerkissä 2.11. Selvitetään, onko $\boldsymbol{A}$ redusoituva.

Lasketaan määritelmän 2.12 mukainen matriisitulo esimerkin 2.11 permutaatiomatriisilla

$$\boldsymbol{P}_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Huomataan ensin, että $\boldsymbol{P}_1^\top = \boldsymbol{P}_1$. Tulomatriisi $\boldsymbol{AP}_1$ laskettiin jo esimerkissä 2.11. Matriisilla $\boldsymbol{P}_1^\top$ kertominen vasemmalta vaihtaa matriisin $\boldsymbol{AP}_1$ ensimmäisen ja toisen rivin paikkaa. Tällöin

$$\boldsymbol{P}_1^\top \boldsymbol{AP}_1 = \boldsymbol{P}_1^\top(\boldsymbol{AP}_1) = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 2 \\ 3 & 0 & 0 \\ 5 & 4 & 6 \end{bmatrix} = \begin{bmatrix} 3 & 0 & 0 \\ 1 & 0 & 2 \\ 5 & 4 & 6 \end{bmatrix}.$$

Saatu matriisi ei ole määritelmän 2.12 mukaisessa muodossa, mutta se ei vielä riitä kertomaan, onko $\boldsymbol{A}$ redusoituva. Valitsemalla permutaatiomatriisi

$$\boldsymbol{P}_2 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix},$$

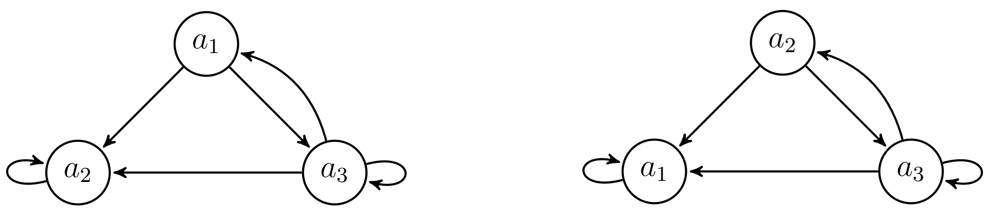
saadaan

$$\boldsymbol{P}_2^\top \boldsymbol{AP}_2 = \begin{bmatrix} 6 & 4 & 5 \\ 2 & 0 & 1 \\ 0 & 0 & 3 \end{bmatrix} = \left[\begin{array}{cc|c} 6 & 4 & 5 \\ 2 & 0 & 1 \\ \hline 0 & 0 & 3 \end{array}\right]$$

Koska $\boldsymbol{X} = \begin{bmatrix} 6 & 4 \\ 2 & 0 \end{bmatrix}$ ja $\boldsymbol{Z} = \begin{bmatrix} 3 \end{bmatrix}$ ovat neliömatriiseja ja matriisin vasen alalohko sisältää pelkästään nollia, määritelmän 2.12 nojalla matriisi $\boldsymbol{A}$ on redusoituva.

Edellisessä esimerkissä matriisin redusoituvuus saatiin osoitettua melko helposti, koska permutaatiomatriisi oli sattumalta sopiva jo toisella yrittämällä. Matriisin koon kasvaessa määritelmän 2.12 käyttö redusoituvuuden selvittämiseksi muuttuu kuitenkin nopeasti erittäin työlääksi. Jos halutaan varmistua, että esimerkiksi 4×4-matriisi on redusoitumaton, täytyy matriisitulo laskea jo 24 eri permutaatiomatriisin kanssa.

Kuvassa 2.2 esitetään esimerkin 2.13 matriiseja $\boldsymbol{A}$ ja $\boldsymbol{P}_1^\top \boldsymbol{AP}_1$ vastaavat graafit. Graafit ovat muuten yhtenäiset, mutta solmut a_1 ja a_2 ovat vaihtuneet keskenään. Määritelmän 2.12 mukainen operaatio nimeää matriisia vastaavan graafin solmut uudelleen, mutta säilyttää nuolet ennallaan. Tätä huomiota hyödynnetään myös seuraavassa lauseessa, joka yksinkertaistaa matriisin redusoituvuuden tutkimista huomattavasti.



(A) $\boldsymbol{A}$ (B) $\boldsymbol{P}_1^\top \boldsymbol{AP}_1$

KUVA 2.2. Esimerkin 2.13 matriiseja vastaavat graafit

LAUSE 2.14. *Neliömatriisi $\boldsymbol{A}$ on redusoitumaton jos ja vain jos sitä vastaava suunnattu graafi on vahvasti yhtenäinen.*

TODISTUS. Oletetaan, että matriisi $\boldsymbol{A}_{n\times n}$ on redusoituva, jolloin sopivalla permutaatiomatriisilla $\boldsymbol{P}$ saadaan

$$\boldsymbol{B} = \boldsymbol{P}^\top \boldsymbol{A}\boldsymbol{P} = \begin{bmatrix} \boldsymbol{B}_{11} & \boldsymbol{B}_{12} \\ \boldsymbol{0} & \boldsymbol{B}_{22} \end{bmatrix} = \begin{bmatrix} b_{11} & \cdots & b_{1k} & b_{1(k+1)} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ b_{k1} & \cdots & b_{kk} & b_{k(k+1)} & \cdots & b_{kn} \\ 0 & \cdots & 0 & b_{(k+1)(k+1)} & \cdots & b_{(k+1)n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & b_{n(k+1)} & \cdots & b_{nn} \end{bmatrix},$$

missä $1 \leq k < n$. Osoitetaan, että matriisia $\boldsymbol{B}$ vastaava suunnattu graafi ei ole vahvasti yhtenäinen. Matriisin $\boldsymbol{B}$ rivin $i > k$ ensimmäiset k alkiota ovat nollia. Niinpä matriisia vastaavan suunnatun graafin solmusta b_i ei ole nuolta solmuihin $b_1, b_2, \ldots, b_k$. Toisin sanoen solmujen joukosta $\{b_{k+1}, b_{k+2}, \ldots, b_n\}$ ei ole nuolta joukkoon $\{b_1, b_2, \ldots, b_k\}$. Jotta solmusta b_{k+1} olisi polku solmuun b_1, täytyisi joukkojen välillä olla vähintään yksi nuoli. Näin ei ole, joten graafi ei ole vahvasti yhtenäinen. Matriisin $\boldsymbol{A}$ graafi saadaan matriisin $\boldsymbol{B}$ graafista nimeämällä solmut uudestaan, joten matriisia $\boldsymbol{A}$ vastaava graafi ei myöskään ole vahvasti yhtenäinen.

Oletetaan nyt, että n solmuinen suunnattu graafi ei ole vahvasti yhtenäinen. Tällöin jostain graafin solmusta a_i ei ole polkua johonkin graafin solmuun a_j. Olkoon V niiden solmujen joukko, joihin ei ole polkua solmusta a_i. Voidaan olettaa, että $V = \{a_1, a_2, \ldots a_k\}$, koska jos näin ei olisikaan, voitaisiin solmut numeroida uudestaan vastaamaan oletusta. Oletuksen perusteella $V \neq \emptyset$. Olkoon W niiden solmujen joukko, joihin solmusta a_i on polku. Jos $W = \emptyset$, niin solmusta a_i ei ole nuolta yhteenkään solmuun ja valitsemalla $a_i = a_n$ graafin vierusmatriisin $\boldsymbol{A}$ rivi n sisältää pelkästään nollia, jolloin $\boldsymbol{A}$ on redusoituva. Jos $W \neq \emptyset$, niin $W = \{a_{k+1}, a_{k+2}, \ldots, a_n\}$. Joukon W solmuista ei voi olla nuolta joukon V solmuihin, sillä muuten myös solmusta a_i olisi polku joukkoon V, mikä olisi ristiriidassa oletuksen kanssa. Niinpä matriisin $\boldsymbol{A}$ rivien $k+1, \ldots, n$ ensimmäiset k alkiota ovat nollia, jolloin

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & \cdots & a_{1k} & a_{1(k+1)} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{kk} & a_{k(k+1)} & \cdots & a_{kn} \\ 0 & \cdots & 0 & a_{(k+1)(k+1)} & \cdots & a_{(k+1)n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{n(k+1)} & \cdots & a_{nn} \end{bmatrix}.$$

Tällöin matriisi $\boldsymbol{A}$ on redusoituva.

Niinpä matriisi on redusoituva jos ja vain jos sen suunnattu graafi ei ole vahvasti yhtenäinen. □

ESIMERKKI 2.15. Kuvan 2.2a graafista nähdään helposti, että solmusta a_2 ei ole polkua solmuun a_3, joten esimerkin 2.13 matriisin $\boldsymbol{A}$ graafi ei ole vahvasti yhtenäinen. Niinpä matriisi on lauseen 2.14 mukaan redusoituva, kuten myös esimerkissä 2.13 todettiin.

Kun suunnatun graafin vierusmatriisista otetaan transpoosi, jokaisen graafin nuolen suunta vaihtuu. Jos matriisin $\boldsymbol{A}$ alkio $a_{ij} \neq 0$, niin solmusta a_i on nuoli solmuun a_j. Koska transpoosin $\boldsymbol{A}^\top$ alkio $a_{ji}^{(t)} = a_{ij}$, niin transpoosin graafissa solmusta $a_j^{(t)}$ on nuoli solmuun $a_i^{(t)}$. Siten myös kaikkien polkujen suunta vaihtuu. Tästä seuraa, että jos graafi on vahvasti yhtenäinen, niin sen vierusmatriisin transpoosia vastaava graafi on myös vahvasti yhtenäinen. Vastaavasti jos jostain solmusta ei ole polkua toiseen solmuun, niin vierusmatriisin transpoosin graafista puuttuu vastakkaiseen suuntaan kulkeva polku, eikä siten kumpikaan graafi ole vahvasti yhtenäinen. Tällöin $\boldsymbol{A}$ on redusoituva täsmälleen silloin, kun $\boldsymbol{A}^\top$ on redusoituva.

Redusoitumattoman epänegatiivisen matriisin ominaisarvoille ja -vektoreille pätee erityisiä ominaisuuksia, joista tärkeimmät esitellään seuraavassa lauseessa. Merkittävän tuloksen takana ovat saksalaiset matemaatikot Oskar Perron ja Georg Frobenius. Lauseen todistus löytyy liitteestä A.

LAUSE 2.16 (Perronin ja Frobeniuksen lause). *Olkoon $\boldsymbol{A}_{n\times n} \geq \boldsymbol{0}$ redusoitumaton. Tällöin seuraavat ovat totta:*

(i) *Matriisilla $\boldsymbol{A}$ on reaalinen ominaisarvo $r > 0$ siten, että $r = \rho(\boldsymbol{A})$.*

(ii) *Ominaisarvoa r vastaa oikeanpuoleinen ominaisvektori $\boldsymbol{x}$, jolle*

$$\boldsymbol{x} > \boldsymbol{0}, \|\boldsymbol{x}\|_1 = 1.$$

Vektoria $\boldsymbol{x}$ sanotaan matriisin $\boldsymbol{A}$ oikeanpuoleiseksi Perronin vektoriksi. Vasemmanpuoleisella Perronin vektorilla $\boldsymbol{y}^\top$ on vastaavat ominaisuudet.

(iii) *Ominaisarvoon r ei liity Perronin vektoreiden lisäksi muita lineaarisesti riippumattomia ominaisvektoreita.*

(iv) *Kaikki matriisin $\boldsymbol{A}$ positiiviset ominaisvektorit ovat Perronin vektoreiden monikertoja.*

Jos epänegatiivisella redusoitumattomalla matriisilla on ainoastaan yksi ominaisarvo spektraalikehällä, matriisin sanotaan olevan *primitiivinen*. Primitiivisen matriisin skaalatuille potensseille pätee seuraava tulos, joka osoittautuu PageRank-algoritmin kannalta erittäin tärkeäksi.

LAUSE 2.17. *Redusoitumaton epänegatiivinen matriisi $\boldsymbol{A}$, jolle $r = \rho(\boldsymbol{A})$, on primitiivinen jos ja vain jos raja-arvo $\lim_{k\to\infty} \left(\frac{\boldsymbol{A}}{r}\right)^k$ on olemassa, jolloin*

$$\lim_{k\to\infty} \left(\frac{\boldsymbol{A}}{r}\right)^k = \frac{\boldsymbol{x}\boldsymbol{y}^\top}{\boldsymbol{y}^\top\boldsymbol{x}} > \boldsymbol{0},$$

missä $\boldsymbol{x}$ ja $\boldsymbol{y}^\top$ ovat matriisin $\boldsymbol{A}$ oikean- ja vasemmanpuoleinen Perronin vektori.

Lauseen todistus on löydettävissä teoksesta [**2**]. Matriisin primitiivisyys voidaan selvittää muillakin tavoilla kuin tutkimalla ominaisarvoja tai raja-arvon olemassaoloa.

LAUSE 2.18. *Epänegatiiviselle matriisille $\boldsymbol{A}$ pätevät seuraavat kaksi asiaa:*

- *$\boldsymbol{A}$ on primitiivinen, jos $\boldsymbol{A}$ on redusoitumaton ja sillä on ainakin yksi positiivinen diagonaalialkio.*
- *$\boldsymbol{A}$ on primitiivinen jos ja vain jos $\boldsymbol{A}^m > \boldsymbol{0}$ jollain $m > 0$.*

TODISTUS. Ensimmäisen kohdan todistus löytyy kokonaisuudessaan teoksesta [**2**]. Todistus perustuu tietoon siitä, että epänegatiivisen ja epäprimitiivisen matriisin jälki on nolla. Todistuksessa huomataan, että jos matriisi $\boldsymbol{A}$ oletetaan epäprimitiiviseksi, niin tällöin sen ominaisarvojen summa on nolla, jolloin myös matriisin jäljen täytyisi olla nolla. Tämä on ristiriita, sillä matriisilla $\boldsymbol{A}$ oletettiin olevan ainakin yksi positiivinen diagonaalialkio. Tällöin matriisin täytyy olla primitiivinen.

Osoitetaan jälkimmäinen kohta kahdessa osassa. Oletetaan ensin, että matriisi $\boldsymbol{A}$ on primitiivinen. Nyt lauseen 2.17 mukaan matriisin $\left(\frac{\boldsymbol{A}}{r}\right)^k$ raja-arvo on positiivinen matriisi, kun $k \to \infty$. Raja-arvon ominaisuuksien perusteella on olemassa $M \in \mathbb{N}$ siten, että kaikilla $k > M$ matriisi $\left(\frac{\boldsymbol{A}}{r}\right)^k$ on positiivinen. Perronin Frobeniuksen lauseen mukaan $r > 0$, joten $\boldsymbol{A}^k > \boldsymbol{0}$, kun $k > M$, mikä haluttiin osoittaa.

Oletetaan nyt, että jollain luvulla $m > 0$ matriisi $\boldsymbol{A}^m$ on positiivinen. Matriisin $\boldsymbol{A}^m$ yksikään alkio ei ole nolla, joten se on redusoitumaton. Oletetaan, että matriisi $\boldsymbol{A}$ on redusoituva, jolloin jollain permutaatiomatriisilla $\boldsymbol{P}$

$$\boldsymbol{A} = \boldsymbol{P}^\top \begin{bmatrix} \boldsymbol{B}_{11} & \boldsymbol{B}_{12} \\ \boldsymbol{0} & \boldsymbol{B}_{22} \end{bmatrix} \boldsymbol{P},$$

missä $\boldsymbol{B}_{11}$ ja $\boldsymbol{B}_{22}$ ovat neliömatriiseja. Tällöin

$$\boldsymbol{A}^m = (\boldsymbol{P}^\top)^m \begin{bmatrix} {\boldsymbol{B}_{11}}^m & \boldsymbol{C} \\ \boldsymbol{0} & {\boldsymbol{B}_{22}}^m \end{bmatrix} \boldsymbol{P}^m,$$

missä $\boldsymbol{B}_{11}^m$ ja $\boldsymbol{B}_{22}^m$ ovat edelleen neliömatriiseja. Permutaatiomatriisin kertominen itsellään muuttaa vain rivien järjestystä, joten matriisi $\boldsymbol{P}^m$ on edelleen permutaatiomatriisi. Lisäksi transpoosille pätee $(\boldsymbol{P}^\top)^m = \boldsymbol{P}^\top\boldsymbol{P}^\top \cdots \boldsymbol{P}^\top = (\boldsymbol{P}\boldsymbol{P}\cdots\boldsymbol{P})^\top = (\boldsymbol{P}^m)^\top$, joten tällöin myös $\boldsymbol{A}^m$ olisi redusoituva matriisi, mikä ei voi pitää paikkansa, koska aiemmin todettiin, että se on redusoitumaton matriisi. Siispä $\boldsymbol{A}$ on redusoitumaton.

Perronin ja Frobeniuksen lauseen perusteella voidaan olettaa, että matriisilla $\boldsymbol{A}$ on olemassa ominaisarvo $r = \rho(\boldsymbol{A})$ siten, että $r \geq |\lambda|$, kun $\lambda \in \sigma(\boldsymbol{A})$. Koska

$$\boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x} \Leftrightarrow \boldsymbol{A}^m\boldsymbol{x} = \lambda^m\boldsymbol{x},$$

niin tällöin matriisin $\boldsymbol{A}^m$ ominaisarvoille pätee $r^m \geq |\lambda^m|$, missä $r^m = \rho(\boldsymbol{A}^m)$. Lauseen 2.18 ensimmäisen kohdan perusteella matriisi $\boldsymbol{A}^m$ on primitiivinen, joten r^m on ainoa ominaisarvo spektraalikehällä eli $r^m > |\lambda^m| = |\lambda|^m$. Tästä seuraa, että $r > |\lambda|$, joten myös matriisilla $\boldsymbol{A}$ on vain yksi ominaisarvo spektraalikehällä, joten myös matriisi $\boldsymbol{A}$ on primitiivinen. □

Huomautus 2.19. Lauseen 2.17 toinen suunta saadaan osoitettua edellisen lauseen avulla, sillä tässä todistuksessa käytettävien lauseen tulosten osoittamiseen ei tarvita lausetta 2.17.

Matriisi $\frac{\boldsymbol{A}}{r} \geq \boldsymbol{0}$, koska Perronin ja Frobeniuksen lauseen nojalla $r > 0$. Lisäksi $\frac{\boldsymbol{A}}{r}$ on redusoitumaton, koska matriisin nolla-alkioiden sijainti on sama kuin matriisissa $\boldsymbol{A}$.

Oletetaan, että

$$\lim_{k\to\infty} \left(\frac{\boldsymbol{A}}{r}\right)^k = \frac{\boldsymbol{x}\boldsymbol{y}^\top}{\boldsymbol{y}^\top\boldsymbol{x}} > \boldsymbol{0}.$$

Jos matriisi $\boldsymbol{A}$ ei olisi primitiivinen, niin lauseen 2.18 toisen kohdan mukaan jokaisella $k > 0$ vähintään yksi matriisin $\boldsymbol{A}^k$ alkio olisi nolla. Tällöin myös matriisissa $\left(\frac{\boldsymbol{A}}{r}\right)^k$ olisi vähintään yksi nolla-alkio. Tämä on ristiriidassa oletuksen kanssa, jossa todettiin, että $\lim_{k\to\infty}\left(\frac{\boldsymbol{A}}{r}\right)^k$ on positiivinen matriisi. Niinpä matriisin $\boldsymbol{A}$ täytyy olla primitiivinen.

Lauseen todistuksen toinen puoli vaatii osaamista sellaisilta aihealueilta, jotka eivät ole työssä käsiteltävän sovelluksen kannalta oleellisia. Jos tunnetaan esimerkiksi Jordanin lohkomatriisille pätevät ominaisuudet, ja tiedetään ominaisarvon r algebrallisen kertaluvun olevan yksi, pystytään todistamaan lauseen raja-arvon olemassaolo [**4**].

2.3. Potenssimenetelmä

Luvussa 1 kerrottiin, kuinka matriisin ominaisarvot ja -vektorit saadaan määritettyä suoraan ominaisarvoyhtälöstä. Jos matriisi on suuri, tuolla tavalla laskeminen on työlästä, erityisesti jos ollaan kiinnostuneita vain itseisarvoltaan suurimmasta ominaisarvosta ja sitä vastaavasta ominaisvektorista. Tämän parin löytämiseen voidaan käyttää potenssimenetelmää. Menetelmä on iteroiva ja sen avulla muodostettu jono suppenee kohti itseisarvoltaan aidosti suurinta ominaisarvoa vastaavaa ominaisvektoria. Menetelmä on tyypillisesti muotoiltu niin, että sillä löydetään matriisin oikeanpuoleinen ominaisvektori. Lisäksi menetelmän avulla saadaan laskettua myös vektoria vastaava ominaisarvo. Myöhemmin osoittautuu, että työssä käsiteltävän sovelluksen kannalta on oleellisempaa löytää itseisarvoltaan suurinta ominaisarvoa vastaava vasemmanpuoleinen ominaisvektori, eikä ole tarpeen määrittää vektoria vastaavaa ominaisarvoa. Niinpä tässä työssä esiteltävä potenssimenetelmä poikkeaa hieman menetelmän tyypillisestä muotoilusta, joka on löydettävissä teoksesta [**5**]. Tästä johtuen menetelmän johdattelussa käytetään matriisin transpoosia.

Olkoon $\boldsymbol{A}^\top$ $n \times n$-matriisi, jolla on n lineaarisesti riippumatonta ominaisvektoria $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ ja niitä vastaavat ominaisarvot, joille pätee

$$|\lambda_1| > |\lambda_2| \geq \cdots \geq |\lambda_n|.$$

Huomautuksen 1.5 perusteella näiden vektoreiden tranpoosit $\boldsymbol{v}_1^\top, \ldots, \boldsymbol{v}_n^\top$ ovat matriisin $\boldsymbol{A}$ vasemmanpuoleisia ominaisvektoreita vastaten samoja ominaisarvoja $\lambda_1, \ldots, \lambda_n$.

Matriisin $\boldsymbol{A}^\top$ ominaisvektorit muodostavat avaruuden $\mathbb{R}^n$ kannan, joten mielivaltainen vektori $\boldsymbol{y}_0 \in \mathbb{R}^n$ voidaan ilmaista ominaisvektoreiden lineaarikombinaationa

$$\boldsymbol{y}_0 = c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_n\boldsymbol{v}_n,$$

missä $c_i \in \mathbb{R}, i = 1, \ldots, n$. Kun yhtälö kerrotaan puolittain matriisilla $\boldsymbol{A}^\top$, saadaan

$$\boldsymbol{A}^\top\boldsymbol{y}_0 = c_1\boldsymbol{A}^\top\boldsymbol{v}_1 + c_2\boldsymbol{A}^\top\boldsymbol{v}_2 + \cdots + c_n\boldsymbol{A}^\top\boldsymbol{v}_n,$$

ja koska $\boldsymbol{A}^\top\boldsymbol{v}_i = \lambda_i\boldsymbol{v}_i$, niin

$$\begin{aligned} \boldsymbol{A}^\top\boldsymbol{y}_0 &= c_1\lambda_1\boldsymbol{v}_1 + c_2\lambda_2\boldsymbol{v}_2 + \cdots + c_n\lambda_n\boldsymbol{v}_n \\ \Leftrightarrow \quad \left(\boldsymbol{A}^\top\right)^k\boldsymbol{y}_0 &= c_1\lambda_1^k\boldsymbol{v}_1 + c_2\lambda_2^k\boldsymbol{v}_2 + \cdots + c_n\lambda_n^k\boldsymbol{v}_n. \end{aligned}$$

Jaetaan nyt yhtälö ensin puolittain luvulla λ_1^k ja otetaan sitten transpoosi, jolloin

$$\begin{aligned} \left(\frac{\boldsymbol{A}^\top}{\lambda_1}\right)^k\boldsymbol{y}_0 &= c_1\boldsymbol{v}_1 + c_2\left(\frac{\lambda_2}{\lambda_1}\right)^k\boldsymbol{v}_2 + \cdots + c_n\left(\frac{\lambda_n}{\lambda_1}\right)^k\boldsymbol{v}_n \\ \Leftrightarrow \quad \boldsymbol{y}_0^\top\left(\frac{\boldsymbol{A}}{\lambda_1}\right)^k &= c_1\boldsymbol{v}_1^\top + c_2\left(\frac{\lambda_2}{\lambda_1}\right)^k\boldsymbol{v}_2^\top + \cdots + c_n\left(\frac{\lambda_n}{\lambda_1}\right)^k\boldsymbol{v}_n^\top. \end{aligned}$$

Koska λ_1 on itseisarvoltaan aidosti suurin ominaisarvo, niin

$$\left|\frac{\lambda_i}{\lambda_1}\right| < 1 \quad \text{kaikilla } i = 2, 3, \ldots, n,$$

joten

$$\lim_{k\to\infty} \left(\frac{\lambda_i}{\lambda_1}\right)^k = 0.$$

Tällöin

$$\lim_{k\to\infty} \boldsymbol{y}_0^\top \left(\frac{\boldsymbol{A}}{\lambda_1}\right)^k = c_1 \boldsymbol{v}_1^\top,$$

kunhan $c_1 \neq 0$. Näin ollen vektorijono $\left(\boldsymbol{y}_0^\top \left(\frac{\boldsymbol{A}}{\lambda_1}\right)^k\right)$ suppenee kohti vektoria $c_1 \boldsymbol{v}_1^\top$. Vektori $c_1 \boldsymbol{v}_1^\top$ on matriisin $\boldsymbol{A}$ skaalattu vasemmanpuoleinen ominaisvektori, joka vastaa ominaisarvoa λ_1.

Tässä menetelmässä on kuitenkin ongelma, sillä vektorin $\boldsymbol{y}_0^\top \boldsymbol{A}^k$ pituutta rajoitetaan luvulla λ_1^k, mutta lukua λ_1 ei tunneta. Niinpä vektoria $\boldsymbol{y}_0^\top \left(\frac{\boldsymbol{A}}{\lambda_1}\right)^k$ ei pystytä laskemaan. Skaalaukseen voidaan käyttää myös vektorin 1-normia, jolloin potenssimenetelmä voidaan esittää seuraavana iteraationa:

Valitaan aloitusvektori $\boldsymbol{y}_0^\top \neq \boldsymbol{0}$. Asetetaan

$$\boldsymbol{y}_{k+1}^\top = \frac{\boldsymbol{y}_k^\top \boldsymbol{A}}{\|\boldsymbol{y}_k^\top \boldsymbol{A}\|_1}.$$

Menetelmällä saadaan vektorijono $\left(\boldsymbol{y}_k^\top\right)$, jonka jokaisen aloitusvektorin jälkeisen vektorin 1-normi on yksi. Termit voidaan esittää aloitusvektorin $\boldsymbol{y}_0^\top$ avulla:

$$\begin{aligned}
\boldsymbol{y}_1^\top &= \frac{\boldsymbol{y}_0^\top \boldsymbol{A}}{\|\boldsymbol{y}_0^\top \boldsymbol{A}\|_1}, \\
\boldsymbol{y}_2^\top &= \frac{\boldsymbol{y}_0^\top \boldsymbol{A}^2}{\|\boldsymbol{y}_0^\top \boldsymbol{A}^2\|_1}, \\
&\vdots \\
\boldsymbol{y}_k^\top &= \frac{\boldsymbol{y}_0^\top \boldsymbol{A}^k}{\|\boldsymbol{y}_0^\top \boldsymbol{A}^k\|_1}.
\end{aligned}$$

Edellä havaittiin, että jono $\left(\boldsymbol{y}_0^\top \boldsymbol{A}^k\right)$ suppenee matriisin $\boldsymbol{A}$ ominaisarvoa λ_1 vastaavan vasemmanpuoleisen ominaisvektorin $\boldsymbol{y}^\top$ suuntaan, joten myös potenssimenetelmän muodostaman jonon $\left(\boldsymbol{y}_k^\top\right)$ raja-arvo on vektorin $\boldsymbol{y}^\top$ monikerta.

Seuraavassa osiossa tullaan huomaamaan, että jos matriisi $\boldsymbol{A}$ on epänegatiivinen ja sen jokainen rivisumma on yksi, niin myös $\lambda_1 = 1$. Tällöin iteraatio supistuu muotoon $\boldsymbol{y}_{k+1}^\top = \boldsymbol{y}_k^\top \boldsymbol{A}$. Havaitaan myös, että jos matriisi $\boldsymbol{A}$ on lisäksi primitiivinen sekä aloitusvektori on epänegatiivinen ja sen komponenttien summa on yksi, niin menetelmällä saatu vektorijono suppenee matriisin vasemmanpuoleiseen Perronin vektoriin.

2.4. Markovin ketjut

Stokastinen prosessi on joukko satunnaismuuttujia $\{X_t\}_{t=0}^{\infty}$, joilla on yhteinen arvojoukko S. Arvojoukkoa kutsutaan prosessin tila-avaruudeksi. Parametri t tulkitaan tyypillisesti aikana, jolloin X_t on prosessin tila hetkellä t. Esimerkiksi netinselaus voidaan ajatella stokastinen prosessina, jonka tila-avaruuden muodostaa kaikkien nettisivujen kokoelma ja satunnaismuuttuja X_t on se sivu, jota hetkellä t tarkastellaan.

Aika voi olla diskreetti tai jatkuva, ja tila-avaruus äärellinen tai ääretön. Seuraavaksi esiteltävässä prosessissa tila-avaruus on äärellinen ja aika diskreetti, joten aikaa merkitään kirjaimella n. Palautetaan prosessin määritelmää varten mieleen ehdollinen todennäköisyys: Jos $\mathbb{P}(B) > 0$, niin tapahtuman A ehdollinen todennäköisyys ehdolla B on

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

MÄÄRITELMÄ 2.20. Stokastinen prosessi $X = \{X_0, X_1, \dots\}$ tila-avaruudessa S on Markovin ketju, jos

$$\mathbb{P}(X_n = s_n | X_{n-1} = s_{n-1}, X_{n-2} = s_{n-2} \dots, X_0 = s_0) = \mathbb{P}(X_n = s_n | X_{n-1} = s_{n-1})$$

kaikilla $n \geq 1$ ja jokaisella tilalla $s_0, \dots, s_n \in S$.

Määritelmässä esitetty ehto Markovin ketjulle on ns. *Markovin ominaisuus*. Markovin ominaisuuden täyttävän prosessin seuraava tila riippuu vain sitä edeltävästä tilasta. Markovin ketju on siten muistiton.

Siirtymätodennäköisyys on todennäköisyys sille, että prosessi on tilassa s_j hetkellä n, kun edeltävällä hetkellä prosessi oli tilassa s_i. Merkitään

$$p_{ij}(n) = \mathbb{P}(X_n = s_j | X_{n-1} = s_i).$$

Siirtymätodennäköisyyden voidaan yksinkertaisemmin ajatella olevan todennäköisyys siirtyä tilasta s_i tilaan s_j hetkellä n. Jatkossa oletetaan, että siirtymätodennäköisyys Markovin ketjussa ei riipu hetkestä n, jolloin $p_{ij}(n)$ on vakio kaikilla n. Markovin ketjun siirtymätodennäköisyydet voidaan esittää siirtymämatriisina

$$\boldsymbol{P}_{k \times k} = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1k} \\ p_{21} & p_{22} & \cdots & p_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ p_{k1} & p_{k2} & \cdots & p_{kk} \end{bmatrix},$$

kun prosessin tila-avaruus on $S = \{s_1, \dots, s_k\}$. Siirtymämatriisin alkiot ovat todennäköisyyksiä, joten matriisi on epänegatiivinen. Ketju saa jokaisena hetkenä jonkin arvon tila-avaruudestaan, joten siirtymämatriisin jokaisen rivin $i = 1, \dots, k$ alkioille pätee $\sum_{j=1}^{k} p_{ij} = 1$. Jos matriisilla on nämä kaksi ominaisuutta, sen sanotaan olevan stokastinen matriisi.

HUOMAUTUS 2.21. Stokastisen matriisin $\boldsymbol{P}_{k \times k}$ kunkin rivin alkioiden summa on yksi, jolloin matriisille pätee

$$\boldsymbol{P}\boldsymbol{e} = \boldsymbol{e},$$

missä $\boldsymbol{e} = (1, \dots, 1)$. Niinpä $\boldsymbol{e}$ on stokastisen matriisin ominaisvektori ja sitä vastaava ominaisarvo on yksi.

Tästä saadaan johdettua stokastisen matriisin tärkeä ominaisuus.

LAUSE 2.22. *Stokastiselle matriisille* $\rho(\boldsymbol{P}) = 1$.

TODISTUS. Lauseen 1.13 mukaan jokaiselle matriisinormille pätee $\rho(\boldsymbol{P}) \leq \|\boldsymbol{P}\|$. Stokastiselle matriisille

$$\|\boldsymbol{P}\|_\infty = \max_{1\leq i\leq k} \sum_{j=0}^{k} p_{ij} = 1.$$

Matriisin eräs ominaisarvo on yksi, joten

$$1 \leq \rho(\boldsymbol{P}) \leq \|\boldsymbol{P}\|_\infty = 1.$$

□

Markovin ketjun tilajakauma hetkellä n on vaakavektori

$$\boldsymbol{p}^\top(n) = \begin{bmatrix} p_1(n) & p_2(n) & \cdots & p_k(n) \end{bmatrix},$$

missä $p_j(n) = \mathbb{P}(X_n = s_j)$, eli $p_j(n)$ on todennäköisyys sille, että ketju on hetkellä n tilassa s_j. Niinpä ketjun aloitusjakauman $\boldsymbol{p}^\top(0)$ alkio $p_j(0)$ on todennäköisyys sille, että ketju alkaa tilasta s_j. Tilajakauma on stokastinen.

ESIMERKKI 2.23. Kuvassa 2.3 on esitetty kolmesta sivusta ja sivujen välisistä linkeistä koostuva web suunnattuna graafina. Graafin linkkimatriisi on

$$\boldsymbol{P} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 1 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}.$$

Graafin linkkimatriisi on stokastinen, koska jokaisen rivin alkioiden summa on yksi. Kuten esimerkissä 2.8 todettiin, linkkimatriisin i. rivin j. alkio voidaan tulkita todennäköisyytenä siirtyä sivulta a_i sivulle a_j, jos seuraava linkki valitaan satunnaisesti. Tähän todennäköisyyteen ei vaikuta se, kuinka sivulle a_i on päädytty. Tällainen netinselaus täyttää siten Markovin ominaisuuden ja matriisi $\boldsymbol{P}$ on Markovin ketjun siirtymämatriisi. Jos jollain webin sivulla ei olisi ainuttakaan linkkiä, netinselaus ei olisi Markovin ketju, sillä silloin sivua edustava solmu ei olisi yhdenkään nuolen alkupiste, mikä näkyisi linkkimatriisissa nollarivinä. Tällöin linkkimatriisi ei olisi stokastinen.

Jos aloitussivu valitaan täysin satunnaisesti, prosessin aloitusjakauma on

$$\boldsymbol{p}^\top(0) = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}.$$

Tällöin

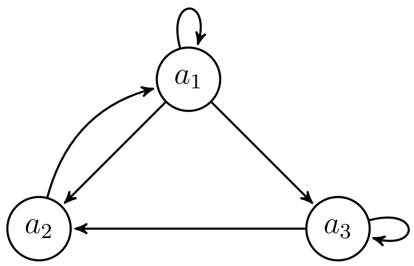
$$\boldsymbol{p}^\top(1) = \boldsymbol{p}^\top(0)\boldsymbol{P} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 1 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} \frac{4}{9} & \frac{5}{18} & \frac{5}{18} \end{bmatrix}.$$

Koska $\boldsymbol{p}^\top(1) > \boldsymbol{0}$, netinselaaja voi hetkellä $n = 1$ olla millä tahansa kolmesta sivusta.

Jos netinselaus aloitetaan sivulta a_3, prosessin aloitusjakauma on

$$\boldsymbol{p}^\top(0) = [0 \quad 0 \quad 1].$$

Tällöin $\boldsymbol{p}^\top(1) = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$, joten hetkellä $n = 1$ netinselaaja ei voi olla sivulla a_1. Tämä voidaan havaita myös graafista, jossa solmusta a_3 ei lähde nuolta solmuun a_1. Sivulta a_3 ei siten pääse sivulle a_1 yhden linkin kautta.



KUVA 2.3. Graafi kolmesivuisesta webistä

Aloitusjakauman $\boldsymbol{p}^\top(0)$ avulla pystytään laskemaan hetken $n = 1$ tilajakauma tila-avaruudessa $S = s_0, s_1, \dots, s_k$. Jokaiselle j

$$\begin{aligned}
p_j(1) &= \mathbb{P}(X_1 = s_j) = \mathbb{P}\big(X_1 = s_j \cap (X_0 = s_0 \cup X_0 = s_1 \cup \cdots \cup X_0 = s_k)\big) \\
&= \mathbb{P}\big((X_1 = s_j \cap X_0 = s_0) \cup (X_1 = s_j \cap X_0 = s_1) \cup \cdots \cup (X_1 = s_j \cap X_0 = s_k)\big) \\
&= \mathbb{P}(X_1 = s_j \cap X_0 = s_0) + \mathbb{P}(X_1 = s_j \cap X_0 = s_1) + \cdots + \mathbb{P}(X_1 = s_j \cap X_0 = s_k) \\
&= \sum_{i=0}^{k} \mathbb{P}(X_1 = s_j \cap X_0 = s_i).
\end{aligned}$$

Kokonaistodennäköisyyden kaavalla saadaan

$$\begin{aligned}
\sum_{i=0}^{k} \mathbb{P}(X_1 = s_j \cap X_0 = s_i) &= \sum_{i=0}^{k} \mathbb{P}(X_0 = s_i)\mathbb{P}(X_0 = s_j | X_1 = s_i) \\
&= \sum_{i=0}^{k} p_i(0) p_{ij}.
\end{aligned}$$

Summa $\sum_{i=0}^{k} p_i(0)p_{ij}$ vastaa vaakavektorin $\boldsymbol{p}^\top(0)$ ja siirtymämatriisin $\boldsymbol{P}$ j:nnen sarakkeen tuloa, joten $\boldsymbol{p}^\top(1) = \boldsymbol{p}^\top(0)\boldsymbol{P}$. Minkä tahansa hetken tilajakauma saadaan laskettua edellisen hetken tilajakauman ja siirtymämatriisin avulla. Tällöin

$$\boldsymbol{p}^\top(n) = \boldsymbol{p}^\top(n-1)\boldsymbol{P} = \boldsymbol{p}^\top(n-2)\boldsymbol{P}^2 = \boldsymbol{p}^\top(n-3)\boldsymbol{P}^3 = \cdots = \boldsymbol{p}^\top(0)\boldsymbol{P}^n.$$

Niinpä aloitusjakaumasta $\boldsymbol{p}^\top(0)$ pystytään laskemaan minkä tahansa hetken n tilajakauma kaavalla

$$\boldsymbol{p}^\top(n) = \boldsymbol{p}^\top(0)\boldsymbol{P}^n,$$

joka on potenssimenetelmä skaalaustekijällä yksi. Matriisin $\boldsymbol{P}^n$ rivin i ja sarakkeen j alkio kuvaa todennäköisyyttä sille, että tilalta s_i päädytään tilalle s_j täsmälleen n hetken kuluttua.

Markovin ketjun tasapainotila on tilajakauma $\boldsymbol{\pi}^\top$, joka toteuttaa ehdon

$$\boldsymbol{\pi}^\top \boldsymbol{P} = \boldsymbol{\pi}^\top,$$

missä $\boldsymbol{P}$ on ketjun siirtymämatriisi. Tasapainotila on siis matriisin $\boldsymbol{P}$ ominaisarvoa yksi vastaava vasemmanpuoleinen ominaisvektori. Lauseen 2.22 sekä Perronin ja Frobeniuksen lauseen perusteella tällainen tilajakauma $\boldsymbol{\pi}^\top$ on olemassa, jos siirtymämatriisi $\boldsymbol{P}$ on redusoitumaton. Jos Markovin ketjun siirtymämatriisi on redusoitumaton, miltä tahansa tilalta pystytään saavuttamaan jokainen muu tila, mutta ei välttämättä yhdellä siirtymällä. Jos siirtymämatriisi on primitiivinen, jollain hetkellä n matriisi

$\boldsymbol{P}^n$ on positiivinen, mikä tarkoittaa sitä, että jokaiselta tilalta on mahdollista siirtyä mille tahansa tilalle n siirtymällä. Lisäksi tällöin tilajakauma $\boldsymbol{p}^\top(n)$ suppenee kohti tasapainotilaa.

LAUSE 2.24. *Olkoon $\boldsymbol{P}$ Markovin ketjun siirtymämatriisi ja $\boldsymbol{p}^\top(n)$ hetken n tilajakauma. Jos $\boldsymbol{P}$ on primitiivinen, niin $\lim_{n\to\infty} \boldsymbol{p}^\top(n)$ on olemassa, ja*

$$\lim_{n\to\infty} \boldsymbol{p}^\top(n) = \boldsymbol{\pi}^\top,$$

missä $\boldsymbol{\pi}^\top$ on ketjun tasapainotila.

TODISTUS. Olkoon ketjun tila-avaruudessa k tilaa. Matriisi $\boldsymbol{P}$ on stokastinen, joten $\boldsymbol{P} \geq \boldsymbol{0}$ ja $\rho(\boldsymbol{P}) = 1$. Huomautuksen 2.21 nojalla vektori $\boldsymbol{e} = (1, \ldots, 1)$ on matriisin $\boldsymbol{P}$ ominaisarvoa yksi vastaava oikeanpuoleinen ominaisvektori. Tällöin vektori $\frac{\boldsymbol{e}}{k} > \boldsymbol{0}$,

$$\boldsymbol{P}\frac{\boldsymbol{e}}{k} = \frac{\boldsymbol{e}}{k}$$

ja $\|\frac{\boldsymbol{e}}{k}\|_1 = 1$, joten vektori $\frac{\boldsymbol{e}}{k}$ on matriisin oikeanpuoleinen Perronin vektori. Tasapainotila $\boldsymbol{\pi}^\top > \boldsymbol{0}$ on matriisin vasemmanpuoleinen Perronin vektori, sillä määritelmän mukaan $\|\boldsymbol{\pi}^\top\|_1 = 1$ ja $\boldsymbol{\pi}^\top \boldsymbol{P} = \boldsymbol{\pi}^\top$. Nyt lauseen 2.17 mukaan

$$\lim_{n\to\infty} \boldsymbol{P}^n = \frac{\frac{\boldsymbol{e}}{k}\boldsymbol{\pi}^\top}{\boldsymbol{\pi}^\top \frac{\boldsymbol{e}}{k}} = \frac{\boldsymbol{e}\boldsymbol{\pi}^\top}{\boldsymbol{\pi}^\top \boldsymbol{e}} = \boldsymbol{e}\boldsymbol{\pi}^\top = \begin{bmatrix} \pi_1 & \pi_2 & \cdots & \pi_k \\ \pi_1 & \pi_2 & \cdots & \pi_k \\ \vdots & \vdots & \ddots & \vdots \\ \pi_1 & \pi_2 & \cdots & \pi_k \end{bmatrix} > \boldsymbol{0}.$$

Tällöin

$$\lim_{n\to\infty} \boldsymbol{p}^\top(n) = \lim_{n\to\infty} \boldsymbol{p}^\top(0)\boldsymbol{P}^n = \boldsymbol{p}^\top(0)\boldsymbol{e}\boldsymbol{\pi}^\top.$$

Aloitusjakauma $\boldsymbol{p}^\top(0)$ on stokastinen vektori, joten

$$\boldsymbol{p}^\top(0)\boldsymbol{e} = \sum_{i=1}^{k} p_i e_i = \sum_{i=1}^{k} p_i \cdot 1 = 1 \tag{2.1}$$

Niinpä

$$\lim_{n\to\infty} \boldsymbol{p}^\top(n) = \boldsymbol{\pi}^\top.$$

□

Edellinen lause takaa sen, että jos potenssimenetelmää sovelletaan primitiiviseen ja stokastiseen matriisiin, niin menetelmällä saatu vektorijono suppenee. Lisäksi jos aloitusvektoriksi valitaan stokastinen vektori, jono suppenee matriisin vasemmanpuoleiseen Perronin vektoriin. Edellisen todistuksen yhtälön 2.1 perusteella aloitusvektoriksi voidaan valita mikä tahansa stokastinen vektori.

LUKU 3

PageRank

3.1. Yhteenlaskukaava

PageRank laskettiin alunperin toistamalla yksinkertaista yhteenlaskua. Olkoon $\{p_1, \ldots, p_n\}$ webin sivujen joukko. Olkoon J_i niiden indeksien joukko, joita vastaavat sivut sisältävät linkin sivulle p_i. Sivun p_i PageRank $r(p_i)$ saadaan selville laskemalla yhteen joukkoa J_i vastaavien sivujen painotettujen PageRankien summa:

$$r(p_i) = \sum_{j \in J_i} \frac{r(p_j)}{d(p_j)}.$$

Painokertoimessa $\frac{1}{d(p_j)}$ luku $d(p_j)$ on sivulta p_j löytyvien linkkien määrä. Kaavaa ei kuitenkaan pystytä käyttämään, koska yhteenlaskettavia PageRankeja ei tiedetä. Ongelma voidaan sivuuttaa iteroinnilla, jossa lähdetään liikkeelle oletuksesta, että jokaisen sivun PageRank on alussa $\frac{1}{n}$. Merkitään sivun p_i PageRankia $k+1$. iteraatiolla merkinnällä $r_{k+1}(p_i)$. Nyt $r_0(p_i) = \frac{1}{n}$ kaikille $i = 1, \ldots, n$ ja

$$r_{k+1}(p_i) = \sum_{j \in J_i} \frac{r_k(p_j)}{d(p_j)}.$$

Seuraavassa esimerkissä lasketaan aiemmin esitellyn webin joitakin ensimmäisiä PageRankeja yhteenlaskukaavalla.

Esimerkki 3.1. Kuvan 2.1 webin kaikkien sivujen PageRank on alussa $\frac{1}{5}$. Lasketaan sivun p_1 kaksi seuraavaa PageRankia. Sivut p_2 ja p_4 sisältävät linkin sivulle p_1, joten summaan tarvitaan niiden edellisen iteraation PageRankeja.

$$\begin{aligned}
r_1(p_1) &= \frac{r_0(p_2)}{d(p_2)} + \frac{r_0(p_4)}{d(p_4)} = \frac{\frac{1}{5}}{2} + \frac{\frac{1}{5}}{3} = \frac{1}{6} \\
r_1(p_2) &= \frac{1}{15} = r_1(p_4) \\
r_2(p_1) &= \frac{r_1(p_2)}{d(p_2)} + \frac{r_1(p_4)}{d(p_4)} = \frac{\frac{1}{15}}{2} + \frac{\frac{1}{15}}{3} = \frac{1}{18}
\end{aligned}$$

Kaikkien sivujen ensimmäisiä PageRankeja esitetään taulukossa 3.1. Kolmannen iteraation jälkeen sivun p_3 PageRank on suurin, joten se näytettäisiin hakutuloksissa ensimmäisenä.

3.2. Yhteenlaskukaavan matriisiesitys

Yhteenlaskukaavassa jokaisen iteraation PageRank lasketaan yhdelle nettisivulle kerrallaan. Kaavan painokertoimet pysyvät jokaisella iteraatiolla muuttumattomina.

TAULUKKO 3.1. Webin 2.1 sivujen PageRankeja

	r_0	r_1	r_2	r_3
p_1	$\frac{1}{5}$	$\frac{1}{6}$	$\frac{1}{18}$	$\frac{5}{108}$
p_2	$\frac{1}{5}$	$\frac{1}{15}$	$\frac{1}{18}$	$\frac{1}{54}$
p_3	$\frac{1}{5}$	$\frac{1}{3}$	$\frac{11}{45}$	$\frac{5}{54}$
p_4	$\frac{1}{5}$	$\frac{1}{15}$	$\frac{1}{18}$	$\frac{1}{54}$
p_5	$\frac{1}{5}$	$\frac{1}{6}$	$\frac{1}{18}$	$\frac{5}{108}$

Muokataan yhteenlaskukaavaa niin, että yhteenlaskuun saadaan näkyville PageRankit myös niiltä sivuilta, jotka eivät sisällä linkkiä sivulle p_i. Olkoon I_i näiden sivujen indeksien joukko. Jotta summa pysyy samana, käytetään näiden sivujen painokertoimena nollaa. Kaava saadaan muotoon

$$r_{k+1}(p_i) = \sum_{j \in J_i} \frac{r_k(p_j)}{d(p_j)} + \sum_{j \in I_i} (0 \cdot r_k(p_j)) \,.$$

Jokaisen sivun PageRankia laskettaessa yhteenlaskussa esiintyy nyt kaikkien sivujen edellisen iteraation PageRank. Painokertoimet pysyvät jokaisella iteraatiolla samana. Jos PageRankin $r_k(p_i)$ kaavan yhteenlaskettavat järjestetään indeksien mukaiseen järjestykseen, huomataan, että painokertoimet vastaavat webin muodostavan graafin linkkimatriisin i. riviä. Jos lisäksi muodostetaan vektori

$$\boldsymbol{\pi}_k^\top = \begin{bmatrix} r_k(p_1) & r_k(p_2) & \dots & r_k(p_n) \end{bmatrix},$$

yhteenlaskukaava saadaan esitettyä lyhyesti kahden matriisin kertolaskuna:

$$\boldsymbol{\pi}_{k+1}^\top = \boldsymbol{\pi}_k^\top \boldsymbol{H}, \tag{3.1}$$

missä $\boldsymbol{H}$ on linkkimatriisi. Iteraatio on matriisiin $\boldsymbol{H}$ sovellettu potenssimenetelmä. Edellisen luvun perusteella potenssimenetelmällä muodostettu vektorijono suppenee, jos matriisi täyttää lauseen 2.24 ehdot. Vaatimusten täyttyessä algoritmilla saadaan laskettua vektorijonon raja-arvo, joka on jokaisen nettisivun PageRank-arvoista koostuva PageRank-vektori $\boldsymbol{\pi}^\top$.

ESIMERKKI 3.2. Edellisessä esimerkissä 3.1 laskettiin kuvan 2.1 webin kolme ensimmäistä PageRankia. Lasketaan esimerkissä 2.8 esitetyn linkkimatriisin $\boldsymbol{H}$ avulla seuraava PageRank. Taulukon 3.1 r_3-sarake muodostaa PageRank-vektorin $\boldsymbol{\pi}_3$.

$$\begin{aligned}\boldsymbol{\pi}_4^\top = \boldsymbol{\pi}_3^\top \boldsymbol{H} &= \begin{bmatrix} \frac{5}{108} & \frac{1}{54} & \frac{5}{54} & \frac{1}{54} & \frac{5}{108} \end{bmatrix} \begin{bmatrix} 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 \\ \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 \\ \frac{1}{3} & 0 & \frac{1}{3} & 0 & \frac{1}{3} \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} \frac{5}{324} & \frac{5}{324} & \frac{11}{162} & \frac{5}{324} & \frac{5}{324} \end{bmatrix}.\end{aligned}$$

Sivujen p_1, p_2, p_4 ja p_5 neljäs PageRank on $\frac{5}{324}$ ja sivun p_3 on $\frac{11}{162}$.

3.3. Stokastisen ja primitiivisen matriisin muodostaminen

Edellä todettiin, että PageRank-vektori $\boldsymbol{\pi}^\top$ pyritään laskemaan yhtälön 3.1 mukaisella potenssimenetelmällä. Linkkimatriisi $\boldsymbol{H}_{n\times n}$ muistuttaa Markovin ketjun siirtymämatriisia, mutta matriisin mahdolliset nollarivit aiheuttavat sen, ettei linkkimatriisi ole stokastinen. Linkkimatriisin nollarivit havaitaan graafissa solmuina, joista ei lähde yhtäkään nuolta muihin solmuihin tai edes solmuun itseensä. Tällaiselta nettisivulta ei siten pääse siirtymään muille sivuille, joten netinselaaja jää jumiin.

Jos matriisi ei ole stokastinen, siihen käytetty potenssimenetelmä ei välttämättä tuota suppenevaa jonoa. Matriisi voidaan kuitenkin muuttaa stokastiseksi korvaamalla jokainen nollarivi vektorilla $\frac{1}{n}\boldsymbol{e}^\top$, joka vaihtaa jokaisen nollarivin alkion tilalle luvun $\frac{1}{n}$. Korvataan samalla vektorilla myös sellaiset rivit, joiden ainoa positiivinen alkio on matriisin diagonaalialkio, jotta päästään eroon sellaisista solmuista, joiden ainoa nuoli osoittaa solmuun itseensä. Stokastinen matriisi $\boldsymbol{S}$ saadaan nyt kaavalla

$$\boldsymbol{S} = \boldsymbol{H} + \boldsymbol{a}\frac{1}{n}\boldsymbol{e}^\top,$$

missä vektorin $\boldsymbol{a}$ komponentti $a_i = 1$, jos matriisin $\boldsymbol{H}$ i. rivi on nollarivi tai $h_{ii} = 1$, muuten $a_i = 0$. Tämän muutoksen ansiosta jokaisesta matriisin määräämästä graafin solmusta lähtee ainakin yksi nuoli toiseen solmuun. Nyt netinselaaja ei voi jäädä jumiin, kun jokaiselta nettisivulta pääsee siirtymään jollekin toiselle sivulle.

Nyt matriisi $\boldsymbol{S}$ on Markovin ketjun siirtymämatriisi. Jotta potenssimenetelmällä saatu vektorijono suppenee, täytyy siirtymämatriisin olla lauseen 2.24 mukaan primitiivinen. Muokataan matriisia $\boldsymbol{S}$ niin, että sen jokainen nolla-alkio muutetaan positiiviseksi luvuksi. Muodostetaan uusi matriisi $\boldsymbol{G}$ seuraavalla kaavalla:

$$\boldsymbol{G} = \alpha\boldsymbol{S} + (1-\alpha)\frac{1}{n}\boldsymbol{e}\boldsymbol{e}^\top, \tag{3.2}$$

missä $\alpha \in\,]0,1[$ ja $\frac{1}{n}\boldsymbol{e}\boldsymbol{e}^\top$ on $n \times n$-matriisi, jonka jokainen alkio on luku $\frac{1}{n}$. Kerroin α määrää sen, että minkä osuuden ajasta netinselaaja klikkailee sivuilta löytyviä linkkejä seuraten matriisia $\boldsymbol{S}$ vastaavaa graafia. Loput ajasta netinselaaja siirtyy satunnaisesti jollekin webin sivulle. Tämän ajan netinselaaja seuraa siis matriisia $\frac{1}{n}\boldsymbol{e}\boldsymbol{e}^\top$ vastaavan graafin rakennetta, jolloin jokaiselta sivulta on yhtä suuri todennäköisyys siirtyä mille tahansa webin sivulle.

Kaavalla 3.2 saatu matriisi on niin sanottu Google-matriisi. Tähän matriisiin käytetyn potenssimenetelmän muodostama jono suppenee kohti PageRank-vektoria.

LAUSE 3.3. *Olkoon $\boldsymbol{G}$ Google-matriisi ja $\boldsymbol{\pi}_0^\top$ stokastinen vektori. Jos*

$$\boldsymbol{\pi}_{k+1}^\top = \boldsymbol{\pi}_k^\top\boldsymbol{G},$$

niin vektorijono $\boldsymbol{\pi}_k^\top$ suppenee kohti vektoria $\boldsymbol{\pi}^\top$, jolle $\boldsymbol{\pi}^\top\boldsymbol{G} = \boldsymbol{\pi}^\top$.

TODISTUS. Matriisin $\boldsymbol{G}$ jokaiselle riville i pätee

$$\begin{aligned}\sum_{j=1}^n g_{ij} &= \sum_{j=1}^n \left(\alpha s_{ij} + (1-\alpha)\frac{1}{n}\right) = n(1-\alpha)\frac{1}{n} + \sum_{j=1}^n \alpha s_{ij}\\ &= 1-\alpha+\alpha\sum_{j=1}^n s_{ij}.\end{aligned}$$

Stokastisen matriisin $\boldsymbol{S}$ jokaiselle riville i $\sum_{j=1}^{n} s_{ij} = 1$. Niinpä myös $\sum_{j=1}^{n} g_{ij} = 1$, joten matriisi $\boldsymbol{G}$ on stokastinen. Lisäksi $\boldsymbol{G} > \mathbf{0}$, koska $(1-\alpha)\frac{1}{n}\boldsymbol{e}\boldsymbol{e}^\top > \mathbf{0}$. Siis $\boldsymbol{G}^1 > \mathbf{0}$, joten lauseen 2.18 perusteella matriisi $\boldsymbol{G}$ on primitiivinen. Nyt lauseen 2.24 nojalla iteraatiolla

$$\boldsymbol{\pi}_{k+1}^\top = \boldsymbol{\pi}_k^\top \boldsymbol{G}$$

saatu vektorijono suppenee kohti PageRank-vektoria $\boldsymbol{\pi}^\top$, jolle $\boldsymbol{\pi}^\top \boldsymbol{G} = \boldsymbol{\pi}^\top$. $\square$

LIITE A

Perronin ja Frobeniuksen lauseen todistus

Todistuksessa mukaillaan teoksessa [**6**] esitettyä todistusta.

Käytetään todistuksessa Brouwerin kiintopistelausetta. Kiintopistelauseen todistus on löydettävissä esimerkiksi lähteestä [**7**].

LAUSE A.1. *Olkoon $S \subset \mathbb{R}^n$ kompakti ja konveksi joukko. Jos kuvaus $T\colon S \to S$ on jatkuva, on olemassa $\boldsymbol{x} \in S$ siten, että $T(\boldsymbol{x}) = \boldsymbol{x}$.*

Joukko S on konveksi, jos jokaiselle $x, y \in S$ pätee

$$(1-\delta)x + \delta y \in S$$

kaikilla $\delta \in [0,1]$.

TODISTUS. [Perronin ja Frobeniuksen lause] Olkoon $\boldsymbol{A}_{n\times n} \geq \boldsymbol{0}$ redusoitumaton. Osoitetaan ensin, että matriisilla $\boldsymbol{A}$ on positiivinen ominaisarvo.

Olkoon $S = \{\boldsymbol{x} \in \mathbb{R}^n \colon \boldsymbol{x} \geq 0 \text{ ja } \|\boldsymbol{x}\|_1 = 1\}$. Joukko S on rajoitettu, sillä jos $\boldsymbol{x} \in S$, niin $0 \leq x_i \leq 1$ kaikilla $i = 1, \ldots, n$ eli $S \subset [0,1]^n$. Joukko S on joukkojen $\{\boldsymbol{x} \in \mathbb{R}^n \colon \boldsymbol{x} \geq \boldsymbol{0}\}$ ja $\{\boldsymbol{x} \in \mathbb{R}^n \colon \|\boldsymbol{x}\| = 1\}$ leikkaus. Joukko $\{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{x} \geq \boldsymbol{0}\}$ on suljettu, koska sen komplementti on avoin. Myös joukko $\{\boldsymbol{x} \in \mathbb{R}^n \colon \|\boldsymbol{x}\| = 1\}$ on suljettu: Funktio $f\colon \mathbb{R}^n \to \mathbb{R}, f(\boldsymbol{x}) = \|\boldsymbol{x}\|_1$ on normifunktiona jatkuva, joten suljetun joukon $\{1\} \subset \mathbb{R}$ alkukuva $f^{-1}(1)$ on suljettu. Tällöin S on kahden suljetun joukon leikkauksena suljettu. Koska joukko S on sekä suljettu että rajoitettu, niin S on kompakti. Lisäksi jos $\boldsymbol{x}, \boldsymbol{y} \in S$, niin kaikilla $\delta \in [0,1]$ pätee

$$(1-\delta)\boldsymbol{x} + \delta\boldsymbol{y} \geq \boldsymbol{0}$$

sekä

$$\|(1-\delta)\boldsymbol{x} + \delta\boldsymbol{y}\|_1 = (1-\delta)\|\boldsymbol{x}\|_1 + \delta\|\boldsymbol{y}\|_1. \tag{A.1}$$

Koska $\boldsymbol{x}, \boldsymbol{y} \in S$, niin $\|\boldsymbol{x}\|_1 = 1 = \|\boldsymbol{y}\|_1$ ja siten edellisestä yhtälöstä saadaan

$$1 - \delta + \delta = 1.$$

Siis $(1-\delta)\boldsymbol{x} + \delta\boldsymbol{y} \in S$, joten S on konveksi. Yhtälö A.1 perustelee sen, miksi joukon S määrittelyssä käytetään juuri 1-normia. Esimerkiksi euklidiselle normille yhtälö ei päde, sillä

$$\sqrt{\sum_{i=1}^{n} ((1-\delta)x_i + \delta y_i)^2} \neq (1-\delta)\sqrt{\sum_{i=1}^{n} x_i^2} + \delta\sqrt{\sum_{i=1}^{n} y_i^2}.$$

Jokaiselle $\boldsymbol{x} \in S$ pätee $\boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{0}$, koska $\boldsymbol{x} \geq \boldsymbol{0}$ ja $\boldsymbol{A} \geq \boldsymbol{0}$. Lisäksi $\boldsymbol{A}\boldsymbol{x} \neq \boldsymbol{0}$, sillä jos olisi $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$, niin vektorin $\boldsymbol{A}\boldsymbol{x}$ jokaiselle komponentille $(Ax)_i$ pätisi $\sum_{j=0}^{n} a_{ij}x_j = 0$. Summan termit ovat epänegatiivisia, joten kaikilla j joko $a_{ij} = 0$ tai $x_j = 0$. Tiedetään, että $x_j > 0$ jollain j, joten tätä indeksiä j vastaava $a_{ij} = 0$. Tämä tarkoittaisi sitä, että matriisin $\boldsymbol{A}$ jokaisen rivin i sarakkeen j alkio olisi nolla eli matriisin j. sarake olisi nollasarake. Tällöin matriisi $\boldsymbol{A}$ olisi redusoituva, mikä oletuksen perusteella ei voi pitää paikkansa.

Määritellään nyt kuvaus $T(\boldsymbol{x}) = \boldsymbol{Ax}/\|\boldsymbol{Ax}\|_1$, joka on kuvaus joukolta S samaan joukkoon S, sillä $T(\boldsymbol{x}) \geq \boldsymbol{0}$ ja

$$\left\| \frac{\boldsymbol{Ax}}{\|\boldsymbol{Ax}\|_1} \right\|_1 = \frac{\|\boldsymbol{Ax}\|_1}{\|\boldsymbol{Ax}\|_1} = 1$$

kaikilla $\boldsymbol{x} \in S$.

Kuvaus T on lineaarikuvauksen ja vektorinormin osamäärä, jotka kumpikin ovat jatkuvia kuvauksia joukossa S. Siten T on jatkuva ja kiintopistelauseen perusteella on olemassa $\boldsymbol{x}_0 \in S$, jolle $T(\boldsymbol{x}_0) = \boldsymbol{x}_0$. Tällöin

$$\begin{aligned} & \frac{\boldsymbol{Ax}_0}{\|\boldsymbol{Ax}_0\|_1} = \boldsymbol{x}_0 \\ \Leftrightarrow \qquad & \boldsymbol{Ax}_0 = \|\boldsymbol{Ax}_0\|_1 \boldsymbol{x}_0. \end{aligned}$$

Niinpä $\|\boldsymbol{Ax}_0\|_1 \in \sigma(\boldsymbol{A})$ ja $\boldsymbol{x}_0$ on sitä vastaava ominaisvektori. Koska $\boldsymbol{Ax}_0 \neq \boldsymbol{0}$, niin $\|\boldsymbol{Ax}_0\|_1 > 0$. Merkitään $r = \|\boldsymbol{Ax}_0\|_1$ ja $\boldsymbol{x} = \boldsymbol{x}_0$.

Osoitetaan, että $\boldsymbol{x} > \boldsymbol{0}$. Koska $\boldsymbol{x} \in S$, niin $\boldsymbol{x} \geq \boldsymbol{0}$, joten riittää osoittaa, että $x_i > 0$ kaikilla $i = 1, \ldots, n$. Tehdään vastaoletus, että $x_i = 0$ jollain i.

Olkoon $\boldsymbol{P}_{n\times n}$ permutaatiomatriisi siten, että $\boldsymbol{Px} = (\tilde{\boldsymbol{x}}, \boldsymbol{0})$, missä vektori $\tilde{\boldsymbol{x}}$ sisältää vektorin $\boldsymbol{x}$ kaikki positiiviset komponentit. Merkitään

$$\boldsymbol{PAP}^\top = \begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{bmatrix}.$$

Permutaatiomatriisille pätee $\boldsymbol{PP}^\top = \boldsymbol{I}$. Tällöin

$$\begin{aligned} & \boldsymbol{Ax} = \boldsymbol{AIx} = \boldsymbol{A}(\boldsymbol{P}^\top \boldsymbol{P})\boldsymbol{x} = r\boldsymbol{x} \\ \Leftrightarrow \qquad & \boldsymbol{PAP}^\top \boldsymbol{Px} = r\boldsymbol{Px} \\ \Leftrightarrow \qquad & \begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{bmatrix} \begin{bmatrix} \tilde{\boldsymbol{x}} \\ \boldsymbol{0} \end{bmatrix} = \begin{bmatrix} r\tilde{\boldsymbol{x}} \\ \boldsymbol{0} \end{bmatrix} \\ \Leftrightarrow \qquad & \begin{bmatrix} \boldsymbol{A}_{11}\tilde{\boldsymbol{x}} \\ \boldsymbol{A}_{21}\tilde{\boldsymbol{x}} \end{bmatrix} = \begin{bmatrix} r\tilde{\boldsymbol{x}} \\ \boldsymbol{0} \end{bmatrix}. \end{aligned}$$

Koska $\tilde{\boldsymbol{x}} > \boldsymbol{0}$, niin $\boldsymbol{A}_{21} = \boldsymbol{0}$, jotta $\boldsymbol{A}_{21}\tilde{\boldsymbol{x}} = \boldsymbol{0}$. Tällöin

$$\boldsymbol{PAP}^\top = \begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{0} & \boldsymbol{A}_{22} \end{bmatrix},$$

mikä ei voi olla totta, koska matriisi $\boldsymbol{A}$ on redusoitumaton. Niinpä ominaisarvoa r vastaava oikeanpuoleinen ominaisvektori $\boldsymbol{x} > \boldsymbol{0}$. Tällöin $\boldsymbol{x}$ on matriisin $\boldsymbol{A}$ oikeanpuoleinen Perronin vektori.

Osoitetaan, että ominaisarvoa r vastaava vasemmanpuoleinen ominaisvektori on positiivinen. Matriisi $\boldsymbol{A}^\top$ on epänegatiivinen ja redusoitumaton, joten todistuksen alkuosan perusteella $\boldsymbol{A}^\top \boldsymbol{y} = \lambda_1 \boldsymbol{y}$ jollain $\lambda_1 > 0$ ja $\boldsymbol{y} > \boldsymbol{0}$. Voidaan olettaa, että $\|\boldsymbol{y}\|_1 = 1$. Nyt

$$\lambda_1 \boldsymbol{x}^\top \boldsymbol{y} = \boldsymbol{x}^\top \lambda_1 \boldsymbol{y} = \boldsymbol{x}^\top \boldsymbol{A}^\top \boldsymbol{y} = r\boldsymbol{x}^\top \boldsymbol{y},$$

sillä $\boldsymbol{Ax} = r\boldsymbol{x} \Leftrightarrow \boldsymbol{x}^\top \boldsymbol{A}^\top = r\boldsymbol{x}^\top$. Koska $\boldsymbol{x}^\top \boldsymbol{y} > \boldsymbol{0}$ ja erityisesti $\boldsymbol{x}^\top \boldsymbol{y} \neq \boldsymbol{0}$, niin $r = \lambda_1$. Niinpä

$$\boldsymbol{A}^\top \boldsymbol{y} = r\boldsymbol{y} \Leftrightarrow \boldsymbol{y}^\top \boldsymbol{A} = r\boldsymbol{y}^\top, \tag{A.2}$$

joten vektori $\boldsymbol{y}^\top > \boldsymbol{0}$ on ominaisarvoa r vastaava vasemmanpuoleinen ominaisvektori. Koska $\|\boldsymbol{y}^\top\|_1 = 1$, niin $\boldsymbol{y}^\top$ on matriisin $\boldsymbol{A}$ vasemmanpuoleinen Perronin vektori.

Osoitetaan, että $r = \rho(\boldsymbol{A})$. Olkoon λ matriisin $\boldsymbol{A}$ ominaisarvo ja $\boldsymbol{z}$ ominaisarvoa λ vastaava ominaisvektori niin, että $\|\boldsymbol{z}\|_1 = 1$. Olkoon $\boldsymbol{z}_+ = (|z_1|, |z_2|, \ldots, |z_n|)$. Tällöin

$$\sum_{j=1}^{n} a_{ij}|z_j| \geq \left|\sum_{j=1}^{n} a_{ij} z_j\right| = |\lambda z_i| = |\lambda||z_i|,$$

joten $\boldsymbol{A}\boldsymbol{z}_+ \geq |\lambda|\boldsymbol{z}_+$. Kerrotaan epäyhtälö vasemmalta positiivisella vektorilla $\boldsymbol{y}^\top$, jolloin saadaan

$$\begin{aligned} & \boldsymbol{y}^\top \boldsymbol{A}\boldsymbol{z}_+ \geq |\lambda|\boldsymbol{y}^\top \boldsymbol{z}_+ \\ \Leftrightarrow \quad & r\boldsymbol{y}^\top \boldsymbol{z}_+ \geq |\lambda|\boldsymbol{y}^\top \boldsymbol{z}_+. \end{aligned}$$

Siis $r \geq |\lambda|$ eli $r = \rho(\boldsymbol{A})$.

Osoitetaan nyt, että ominaisarvoon r ei liity muita lineaarisesti riippumattomia ominaisvektoreita. Olkoon $\boldsymbol{v} \in \mathbb{R}^n, \boldsymbol{v} \neq \boldsymbol{0}$ ominaisarvoa r vastaava ominaisvektori. Oletetaan, että vektorit $\boldsymbol{v}$ ja $\boldsymbol{x}$ ovat lineaarisesti riippumattomat. Näytetään ensin, että on olemassa $c \in \mathbb{R}$ siten, että $\boldsymbol{x} - c\boldsymbol{v} \geq \boldsymbol{0}, \boldsymbol{x} - c\boldsymbol{v} \neq \boldsymbol{0}$ ja $x_i - cv_i = 0$ jollain $i = 1, \ldots, n$. Koska $\boldsymbol{v} \neq \boldsymbol{0}$, niin on olemassa epätyhjä joukko $I = \{i : v_i \neq 0\}$. Tällöin voidaan valita k niin, että $\left|\frac{x_k}{v_k}\right| = \min_{i\in I}\left|\frac{x_i}{v_i}\right|$. Asetetaan $c = \frac{x_k}{v_k}$. Nyt jos $v_i \neq 0$, niin

$$x_i - cv_i = x_i - \frac{x_k}{v_k}v_i \geq x_i - \frac{x_i}{v_i}v_i = 0,$$

ja jos $v_i = 0$, niin $x_i - cv_i = x_i > 0$. Tällöin siis $\boldsymbol{x} - \frac{x_k}{v_k}\boldsymbol{v} \geq \boldsymbol{0}$. Lisäksi $x_k - \frac{x_k}{v_k}v_k = 0$. Koska $\boldsymbol{x}$ ja $\boldsymbol{v}$ ovat lineaarisesti riippumattomat, niin $c_1\boldsymbol{x} + c_2\boldsymbol{v} = 0$ täsmälleen silloin, kun $c_1 = 0 = c_2$. Niinpä $\boldsymbol{x} - \frac{x_k}{v_k}\boldsymbol{v} \neq \boldsymbol{0}$ ja valittu c toteuttaa ehdot.

Ehdot täyttävän luvun c löytämiseksi on oleellista olettaa, että vektorit $\boldsymbol{x}$ ja $\boldsymbol{v}$ ovat lineaarisesti riippumattomat. Muuten vektori $\boldsymbol{x} - c\boldsymbol{v}$ olisi yhdensuuntainen vektorin $\boldsymbol{x}$ kanssa eli $\boldsymbol{x} - c\boldsymbol{v} = t\boldsymbol{x}$ jollain $t \in \mathbb{R}$. Koska $\boldsymbol{x} > \boldsymbol{0}$, niin vektorin $t\boldsymbol{x}$ komponentit olisivat kaikki joko positiivisia, negatiivisia tai nollia riippuen kertoimesta t. Tällöin ei siis voitaisi valita lukua c niin, että ehdot täyttyisivät.

Nyt siis $\boldsymbol{x}$ ja $\boldsymbol{v}$ ovat lineaarisesti riippumattomat ja luku c on valittu niin, että $\boldsymbol{x} - c\boldsymbol{v} \geq \boldsymbol{0}, \boldsymbol{x} - c\boldsymbol{v} \neq \boldsymbol{0}$ ja $x_i - cv_i = 0$ jollain $i = 1, \ldots, n$. Tällöin

$$\begin{aligned} & \boldsymbol{Ax} - \boldsymbol{A}c\boldsymbol{v} = r\boldsymbol{x} - rc\boldsymbol{v} \\ \Leftrightarrow \quad & \boldsymbol{A}(\boldsymbol{x} - c\boldsymbol{v}) = r(\boldsymbol{x} - c\boldsymbol{v}), \end{aligned}$$

joten vektori $\boldsymbol{x} - c\boldsymbol{v}$ on matriisin $\boldsymbol{A}$ ominaisvektori. Aiemmin havaittiin, että jos ominaisarvoa r vastaavalle ominaisvektorille $\boldsymbol{w}$ pätee $\boldsymbol{w} \geq \boldsymbol{0}$, niin täytyy olla $\boldsymbol{w} > \boldsymbol{0}$, koska matriisi $\boldsymbol{A}$ on redusoitumaton. Siten $\boldsymbol{x} - c\boldsymbol{v} > \boldsymbol{0}$. Tämä on ristiriidassa oletuksen kanssa, jossa oletettiin, että vähintään yksi vektorin $\boldsymbol{x} - c\boldsymbol{v}$ komponentti on nolla. Niinpä vektorit $\boldsymbol{v}$ ja $\boldsymbol{x}$ ovat lineaarisesti riippuvat.

Vastaavasti voidaan näyttää, että ominaisarvoon r ei liity muita lineaarisesti riippumattomia vasemmanpuoleisia ominaisvektoreita kuin vektori $\boldsymbol{y}^\top$. Tätä ennen tulee kuitenkin osoittaa, että myös vasemmanpuoleinen ominaisarvoon r liittyvä ominaisvektori $\boldsymbol{w}^\top > \mathbf{0}$, jos $\boldsymbol{w}^\top \geq \mathbf{0}$. Tämän todistus noudattaa samaa kaavaa kuin aiemmin esitetty perustelu sille, että $\boldsymbol{x} > \mathbf{0}$.

Näytetään vielä, että jokainen matriisin $\boldsymbol{A}$ positiivinen ominaisvektori on vektorin $\boldsymbol{x}$ tai $\boldsymbol{y}^\top$ monikerta. Olkoon $\boldsymbol{A}\boldsymbol{w} = \alpha\boldsymbol{w}$ niin, että $\boldsymbol{w} > \mathbf{0}$. Yhtälöstä A.2 seuraa

$$\alpha\boldsymbol{y}^\top\boldsymbol{w} = \boldsymbol{y}^\top\alpha\boldsymbol{w} = \boldsymbol{y}^\top\boldsymbol{A}\boldsymbol{w} = r\boldsymbol{y}^\top\boldsymbol{w}.$$

Koska $\boldsymbol{y}^\top\boldsymbol{w} > \mathbf{0}$, niin $\alpha = r$ ja siten $\boldsymbol{w}$ on vektorin $\boldsymbol{x}$ monikerta.

Jos $\boldsymbol{v}^\top\boldsymbol{A} = \beta\boldsymbol{v}^\top$ siten, että $\boldsymbol{v}^\top > \mathbf{0}$, niin

$$\beta\boldsymbol{u}^\top\boldsymbol{x} = \boldsymbol{u}^\top\beta\boldsymbol{x} = \boldsymbol{u}^\top\boldsymbol{A}\boldsymbol{x} = r\boldsymbol{u}^\top\boldsymbol{x}.$$

Siten $\beta = r$ ja $\boldsymbol{u}^\top$ on vektorin $\boldsymbol{y}^\top$ monikerta. $\square$

Lähdeluettelo

[1] Amy N. Langville ja Carl D. Meyer: *Google's PageRank and Beyond: The Science of Search Engine Rankings*. Princeton University Press, 2006.

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

[3] Pertti Koivisto ja Riitta Niemistö: *Graafiteoriaa*. 2. painos, Tampereen yliopisto, 2018.

[4] Gerard Debreu ja I. N. Herstein: Nonnegative Square Matrices, *Econometrica*, vol. 21 (1953), 597–607.

[5] Steven J. Leon: *Linear Algebra with Applications*. 8th Edition, Pearson, 2010.

[6] R. B. Bapat ja T. E. S. Raghavan: *Nonnegative Matrices and Applications*. Cambridge University Press, 1997.

[7] Hichem Ben-El-Mechaieh ja Yusef A. Mechaiekh: An elementary proof of the Brouwer's fixed point theorem, *Arab. J. Math.*, vol. 11 (2022), 179–188.