

MAT1120 Oblig 1

Daniel Heinesen, daniehei

18.09.2016

Oppgave 1:

Linkmatrisen er en matrise som viser linkene mellom dokumenter. Setter sier at kolonnene representerer dokumentene som linker, mens radene er hvor disse linkes. Vi begynner med å sette en 1 som element a_{ji} om dokument i linkes til dokument j . Vi ser vi at det går en link fra dokument 1 til 2, 3 og 4. Vi får da en ener i rad 2,3 og 4 i kolonne. Linken fra dokument 2 til 3 og 4 utgjør på samme måte kolonne 2. Samme for kolonne 3 og 4. Vi får da matrisen:

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

Men for at dette skal være en stokastisk matrise må summen av elementene i kolonnene være lik 1. Vi tar derfor og deler elementene i kolonnen på summen av elementene. Vi får da den stokastiske linkmatrisen:

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

For å finne den unike scorevektoren vet vi at vi må løse

$$\begin{aligned} \mathbf{Ax} &= \mathbf{x} \\ \Rightarrow (\mathbf{A} - \mathbf{I})\mathbf{x} &= 0 \end{aligned}$$

Vi må altså finne nullrommet til $\mathbf{A} - \mathbf{I}$. Vi gjør dette ved å radredusere $\mathbf{A} - \mathbf{I}$:

$$\begin{aligned} (\mathbf{A} - \mathbf{I}) &= \begin{pmatrix} -1 & 0 & 1 & 1/2 \\ 1/3 & -1 & 0 & 0 \\ 1/3 & 1/2 & -1 & 1/2 \\ 1/3 & 1/2 & 0 & -1 \end{pmatrix} \\ &\sim \begin{pmatrix} 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & -2/3 \\ 0 & 0 & 1 & -3/2 \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

Dette gir oss at:

$$\begin{aligned} x_1 &= 2x_4 \\ x_2 &= \frac{2}{3}x_4 \end{aligned}$$

$$x_3 = \frac{3}{2}x_4$$

Og siden det bare er 0'er i siste rad, så har vi én fri variabel: x_4 . Setter vi $x_4 = 6$ for å få heltall i vektoren:

$$\text{Nul}(\mathbf{A} - \mathbf{I}) = \text{span}\left\{\begin{bmatrix} 12 \\ 4 \\ 9 \\ 6 \end{bmatrix}\right\}$$

Alt vi trenger å gjøre nå er å normalisere basisen til nullrommet, og vi har scorevektoren. Summen av elementene i vektoren er 31. Så den unike scorevektoren er:

$$\frac{1}{31}\begin{bmatrix} 12 \\ 4 \\ 9 \\ 6 \end{bmatrix} = \begin{bmatrix} 0.38710 \\ 0.12903 \\ 0.29032 \\ 0.19355 \end{bmatrix}$$

Rankingen blir da:

Ranking	Document
1	1
2	3
3	4
4	2

Oppgave 2:

Bruker vi samme metode som over finner vi at linkmatrisen for system 2 er:

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

Vi kan allerede her se at vi kommer til å få problemer, siden 1 og 2, og 3 og 4 bare sender frem og tilbake til hverandre noe som betyr at vi ikke kommer til å oppnå noe likevekt mellom dem. Men la oss bevise det matematisk. Vi regner derfor ut $Nul(\mathbf{A} - \mathbf{I})$:

$$\begin{aligned} (\mathbf{A} - \mathbf{I}) &= \begin{pmatrix} -1 & 1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 1/2 \\ 0 & 0 & 1 & -1 & 1/2 \\ 0 & 0 & 0 & 0 & -1 \end{pmatrix} \\ &\sim \begin{pmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

Vi setter så opp variablene:

$$x_1 = x_2$$

$$x_3 = x_4$$

$$x_5 = 0$$

Vi har er 2 rader med bare 0'er, som betyr at vi har 2 frie variabler, x_2 og x_4 . Setter vi at disse er lik 1, får vi:

$$Nul(\mathbf{A} - \mathbf{I}) = span\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \right\}$$

Siden vi har 2 basiser er det ikke mulig å finne en unik scorevektor, akkurat som vi kunne se ut i fra linkmatrisen/diagrammet. Siden det ikke finnes

en unik scorevektor kan vi heller ikke finne en rank for disse dokumentene.

Oppgave 3):

Begge matrisene i oppgavene over er stokastiske, siden vi sikret oss om at summen av elementene i kolonnene var 1, og de er begge kvadratiske.

For å være regulære må matrisene ha en potens k hvor alle elementene i A^k er positive, da vil, for alle potenser A^m , $m > k$ også bare ha positive elementer. Matrisen vil da ha en unik scorevektor (teorem 18).

Matrisen i oppgave 1 har 0-elementer, men sjekker man på matlab finner man at A^6 har bare positive elementer, og det finnes en unik scorevektor, hvilket vi vet siden vi allerede har funnet den. Denne er derfor regulær. Vi vet også at denne matrisen har en unik scorevektor, som sier oss at den er regulær.

For matrisen i oppgave 2 vet vi allede at det ikke finnes en unik scorevektor, så den kan per teorem 18 ikke være regulær.

En enkel måte å få en web med ikke-stokastisk linkmatrise er å ha et dokument som ikke har noen linker. Vi vil da ha en kolonne med bare 0'ere. Summen av denne kolonnen vil da ikke være lik 1, og matrisen er derfor ikke stokastisk.

Oppgave 4):

$$\mathbf{M} = (1 - m)\mathbf{A} - m\mathbf{S}$$

For å finne ut om M er stokastisk kan vi summere en virkårlig kolonne, og ser om det blir 1. Kolonnen er:

$$\mathbf{m}_i = \begin{bmatrix} (1 - m)a_{1i} + m/n \\ \vdots \\ (1 - m)a_{ji} + m/n \end{bmatrix}$$

Summen av elementene blir:

$$\begin{aligned} \sum_j m_{ji} &= (1 - m) \underbrace{(a_{1i} + \dots + a_{ji})}_{=1} + n \frac{m}{n} \\ &= (1 - m) + m = 1 \end{aligned}$$

Vi vet A er kvadratisk, og S er per definisjon kvadratisk. M er derfor stokastisk så lenge A er stokastisk.

For at \mathbf{M} skal være regulær må elementene være strengt positive. Vi vet at alle elementene i \mathbf{A} er ≥ 0 , derfor vil også $(1 - m)\mathbf{A}$ også være det. $m\mathbf{S} = m/n > 0$. Dette betyr at $(1 - m)\mathbf{A} + m\mathbf{S}$ bare har positive elementer. Siden matriseprodukten er dotproduktet av rader og kolonner, hvor alle elementene er positive tall, vil også produktet bare ha positive elementer. Derfor vil \mathbf{M}^k for $k > 0$ bare ha strengt positive elementer, og \mathbf{M} er regulær.

Oppgave 5):

For linkmatrisen \mathbf{A} i oppgave 1, får vi, med $m = 0.15$:

$$M = \begin{pmatrix} 0.03 & 0.88 & 0.03 & 0.03 & 0.03 \\ 0.88 & 0.03 & 0.03 & 0.03 & 0.03 \\ 0.03 & 0.03 & 0.03 & 0.88 & 0.455 \\ 0.03 & 0.03 & 0.88 & 0.03 & 0.455 \\ 0.03 & 0.03 & 0.03 & 0.03 & 0.03 \end{pmatrix}$$

Ved bruk av matlab finner vi at:

$$Nul(\mathbf{M} - \mathbf{I}) = span\left\{ \begin{bmatrix} -0.405429 \\ -0.405429 \\ -0.577736 \\ -0.577736 \\ -0.060814 \end{bmatrix} \right\}$$

Normaliserer vi basisen finner vi scorevektoren for \mathbf{M}

$$\begin{bmatrix} 0.2 \\ 0.2 \\ 0.285 \\ 0.285 \\ 0.03 \end{bmatrix}$$

Oppgave 6):

Det denne funksjonen gjør er at den lager en matrise med 0'ere og 1'ere spredd tilfeldig i matrisen, deretter sjekker den om det er noen kolonner som bare inneholder 0'ere, og om det er tilfelle blir en 1'er lagt til n-te rad. Elementene i diagonalen blir satt til 0. Siste kolonne blir sjekket separat, siden vi ikke kan legge til en 1 i n-te rad her, siden dette er på diagonalen. Derfor settes det heller en 1'er i første rad skulle det ikke være noen andre enere i kolonnen. Tilslutt normaliseres kolonnene.

Siden alle kolonnene nå bare har positive elementer og summen av elementene er 1, så har vi en stokastisk matrise. Men vi ville ikke at noen av

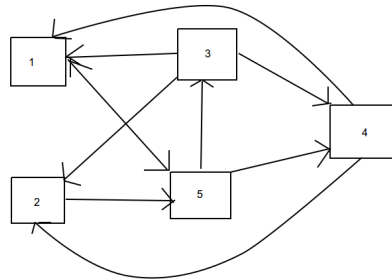
dokumentene skal linke til seg selv, men siden diagonalen bare inneholder 0'er, vil dette være tilfellet. Dette betyr at alle kravene for en linkmatrise er oppfylt.

Oppgave 7):

Jeg fikk den tilfeldige linkmatrisen:

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

Linkdiagrammet blir:



Figur 1: Et litt forvirrende diagram

Oppgave 8):

Dette er kodet i Octave og ikke matlab, men jeg tror det skal kunne kjøres i begge programmene:

```
1 function x=ranking(A,m = 0.15)
2     n = size(A,2);
3     rows = size(A,1);
4
5     if (rows != n) %Checks if the matrix is square
6         disp("This matrix is not nxn")
7         return
8     end
9     for k=1:(n) %Checks if the matrix is stocatic
10         if (sum(A(:,k)) != 1)
11             disp("This is not a stocatisk matrix");
12             return
13         end;
14     end;
15     S = ones(n,n);
16     M = (1-m).*A + (m/n).*S; %Makes the google matrix
17
18     x = null(M - eye(n)); %Finds the basis of the null
19         space
20
21     s = sum(x);
22
23     x = (1./s).*x; %Makes the sum of the elements = 1
```


Oppgave 9):

Som over, er dette kodet i Octave:

```
1 function x=rankingapprox(A,delta,m=0.15)
2     n = size(A,2); %Finds the number of columns
3     S = ones(n,n);
4     M = (1-m).*A + (m/n).*S;
5
6     x = (1/n).*ones(n,1); %Begins with a guess of x0
7     x = (1/(sum(x))).*x; %Makes the sum of elements = 1
8
9     x_old = x;
10    x = M*x;
11
12    while(max(abs(x-x_old)) > delta)
13        x_old = x; %Saves the old x
14        x = M*x; %Updates x
15    end;
```

Oppgave 10):

Som i oppgave 1 er

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

Vi kan få scorevektoren til $\mathbf{M}(\mathbf{A})$ ved både den nøyaktige utregningen og tilnærmingen. Vi kan forvente at for små δ skal disse være ganske like.

Ved bruk av *ranking*-funksjonen får jeg at scorevektoren for M er

$$\mathbf{x} = \begin{bmatrix} 0.36815 \\ 0.14181 \\ 0.28796 \\ 0.20208 \end{bmatrix}$$

Ved bruk av *rankingapprox* med $\delta = 0.005$ får vi

$$\mathbf{x} = \begin{bmatrix} 0.36639 \\ 0.14224 \\ 0.28840 \\ 0.20297 \end{bmatrix}$$

Vi kan se at disse 2 vektorene nesten er like, som vi forventer av funksjonene. Senker vi δ litt vil de komme nærmere hverandre, og ved enda mindre δ konvergerer de raskt mot hverandre.