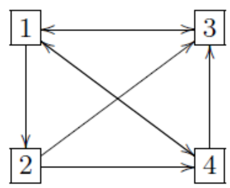
Oppgave 1



Figur 2: Web for oppgave 1

Fra figur 2, kan vi se at:

$$n_1 = 3, \quad n_2 = 2, \quad n_3 = 1, \quad n_4 = 2$$

Da må vi bygge opp link-matrisen $A = [a_{ij}]$ slik at $a_{ij} = 1/n_j$ hvis dokument j har en hyperlenke til dokument i, og $a_{ij} = 0$ ellers. Matrisen A er da:

$$A = \begin{bmatrix} 0 & 0 & 1 & 1/2 \\ 1/3 & 0 & 0 & 0 \\ 1/3 & 1/2 & 0 & 1/2 \\ 1/3 & 1/2 & 0 & 0 \end{bmatrix} \Leftrightarrow A - I = \begin{bmatrix} -1 & 0 & 1 & 1/2 \\ 1/3 & -1 & 0 & 0 \\ 1/3 & 1/2 & -1 & 1/2 \\ 1/3 & 1/2 & 0 & -1 \end{bmatrix}$$

For å finne score-vektoren for A, trenger vi å finne nullrommet for matrisen $A - I$. Dette kan vi regne ut ved hjelp av rad-reduksjon algoritmen på denne matrisen:

$$\left[\begin{array}{cccc|c} -1 & 0 & 1 & 1/2 & 0 \\ 1/3 & -1 & 0 & 0 & 0 \\ 1/3 & 1/2 & -1 & 1/2 & 0 \\ 1/3 & 1/2 & 0 & -1 & 0 \end{array}\right] \sim \left[\begin{array}{cccc|c} -2 & 0 & 2 & 1 & 0 \\ 1 & -3 & 0 & 0 & 0 \\ 2 & 3 & -6 & 3 & 0 \\ 2 & 3 & 0 & -6 & 0 \end{array}\right]$$

$$\sim \left[\begin{array}{cccc|c} 1 & 0 & 0 & -2 & 0 \\ 0 & 3 & 0 & -2 & 0 \\ 0 & 0 & 2 & -3 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] \sim \ldots \sim \left[\begin{array}{cccc|c} 1 & 0 & 0 & -2 & 0 \\ 0 & 1 & 0 & -2/3 & 0 \\ 0 & 0 & 1 & -3/2 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

La x være en fri variabel, kan vi skrive nullrommet slik:

$$\mathrm{Nul}(A - I) = \begin{pmatrix} 2x \\ 2x/3 \\ 3x/2 \\ x \end{pmatrix} \tag{1}$$

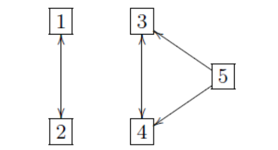
For å få en score-vektor, trenger vi å finne en riktig verdi for x slik at summen av elementene i (1) blir 1:

$$2x + \frac{2x}{3} + \frac{3x}{2} + x = x\left(2 + \frac{2}{3} + \frac{3}{2} + 1\right) = 1 \Leftrightarrow x = \frac{6}{31}$$

Score-vektor $\mathbf{x}$ for link-matrise A er da:

$$\mathbf{x} = \frac{6}{31}\begin{pmatrix} 2 \\ 2/3 \\ 3/2 \\ 1 \end{pmatrix} = \begin{pmatrix} 12/31 \\ 4/31 \\ 9/31 \\ 6/31 \end{pmatrix}$$

Oppgave 2



Figur 3: Web for oppgave 2

Link-matrise A for figur 3 er:

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1/2 \\ 0 & 0 & 1 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \Leftrightarrow A - I = \begin{bmatrix} -1 & 1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 1/2 \\ 0 & 0 & 1 & -1 & 1/2 \\ 0 & 0 & 0 & 0 & -1 \end{bmatrix}$$

Ved å bruke rad-operasjonene på matrisen $A - I$, får vi den følgende redusert trappeform:

$$\begin{bmatrix} -1 & 1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 1/2 \\ 0 & 0 & 1 & -1 & 1/2 \\ 0 & 0 & 0 & 0 & -1 \end{bmatrix} \sim \dots \sim \begin{bmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Hvis x_1 og x_2 er to frie variabler, så er:

$$\text{Nul}(A - I) = \begin{pmatrix} x_1 \\ x_1 \\ x_2 \\ x_2 \\ 0 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \text{ hvor vektorene } \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \text{ og } \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} \text{ danner en basis for } \text{Nul}(A - I).$$

På samme måte som i oppgave 1, kan vi enkelt regne ut x_1 og x_2 slik at vi ender opp med en score-vektor:

$$\begin{aligned} x_1 + x_1 + x_2 + x_2 + 0 = 2(x_1 + x_2) = 1 \\ \Leftrightarrow x_1 + x_2 = \frac{1}{2} \end{aligned} \tag{2}$$

Men siden likningen (2) indikerer at det er uendlig mange muligheter for hva x_1 og x_2 kan bli, så finnes det ikke en unik score-vektor for weben i figur 3.

Oppgave 3

Begge de to link-matrisene i oppgave 1 og 2 er stokastiske. Det er fordi alle koeffisientene er større eller lik 0, og summen av alle elementene i en kolonne er lik 1.

Link-matrise A i oppgave 1 er regulær, og dette kan bevises med eksistensbevis. En matrise er stokastisk regulær hvis det er minst et tall for $k \in \mathbb{N}$, slik at A^k er en matrise hvor ingen av koeffisientene er mindre eller lik 0. Denne kondisjonen tilfredstilles for link-matrise A i oppgave 1 hvis k er lik 4:

$$A^4 = \begin{bmatrix} 0 & 0 & 1 & 1/2 \\ 1/3 & 0 & 0 & 0 \\ 1/3 & 1/2 & 0 & 1/2 \\ 1/3 & 1/2 & 0 & 0 \end{bmatrix}^4 = \begin{bmatrix} 1/3 & 3/8 & 5/12 & 11/24 \\ 5/36 & 1/12 & 1/6 & 1/12 \\ 11/36 & 1/3 & 1/4 & 7/24 \\ 2/9 & 5/24 & 1/6 & 1/6 \end{bmatrix}$$

For link-matrisen A i oppgave 2, kan vi bruke kontrapositivt bevis til å vise at A ***ikke*** er regulær. Teorem 18 i læreboka (avsnitt 4.9) sier at: «Hvis A er en regulær stokastisk matrise, så finnes det en unik score-vektor for A». Kontrapositive påstanden for denne teoremen er da: «Hvis matrise A *ikke* har en unik score-vektor, så er A *ikke* en regulær stokastisk matrise». Siden vi vet allerede fra besvarelsen i oppgave 2 at link-matrisen A (som er stokastisk) ***ikke*** har en unik score-vektor, så følger det at A ***ikke*** er regulær.

En web kan gi en link-matrise M som ikke er stokastisk hvis det er minst et dokument j som ikke har noen hyper-lenker til et annet dokument (dvs. $n_j = 0$). Dette er fordi hvis $n_j = 0$, så er kolonne j lik null-vektoren ($\mathbf{0}$); og det betyr at summen av elementene i kolonne j er lik 0, ikke lik 1. Derfor er det umulig for M å være stokastisk i dette tilfellet.

Oppgave 4

Her antar vi at A er stokastisk, og dette betyr at:

$$a_{ij} \geq 0 \tag{4a}$$

$$\sum_{i=1}^{n} a_{ij} = 1 \tag{4b}$$

La $M = [\mathbf{m}_1 \quad \cdots \quad \mathbf{m}_n]$ være Google-matrisen og $A = [\mathbf{a}_1 \quad \cdots \quad \mathbf{a}_n]$ være den assosierte link-matrisen. Fra definisjonen for M, kan vi skrive matrise M slik:

- $\mathbf{m}_j = (1-m)\mathbf{a}_j + I_n \begin{pmatrix} m/n \\ \cdots \\ m/n \end{pmatrix}$, hvor $j \in \{1, \ldots, n\}$ og I_n er den $n \times n$ identitet matrisen.

Da kan vi skrive den ij-koeffisienten til M på følgende måte:

$$m_{ij} = (1-m)a_{ij} + \frac{m}{n} \tag{4c}$$

Siden $a_{ij} \geq 0$ (fra $4a$), da er tallet 0 den minste verdien a_{ij} kan være i en hvilken som helst matrise A. Det betyr da at den minste verdien som m_{ij} kan være er m/n (se på likningen $4c$).

Men siden både m og n er positive tall, så er $m_{ij} > 0$. (4d)

Hvis vi bruker resultatet fra (4b) og (4c), kan vi skrive summen av elementene i en vilkårlig kolonne j i M slik:

$$\sum_{i=1}^{n} m_{ij} = \sum_{i=1}^{n} \left\{(1-m)a_{ij} + \frac{m}{n}\right\} = (1-m)\sum_{i=1}^{n}\left(a_{ij}\right) + m = (1-m) + m = 1 \tag{4e}$$

Fra (4d) og (4e), kan vi konkludere at M er en stokastisk regulær matrise.

Oppgave 5

Se Appendix A for MATLAB utskrift.

Google-matrise M for link-matrisen A i oppgave 2 er:

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

Basis-mengde for $\mathrm{Nul}(M - I)$ er $\left\{ \begin{pmatrix} 20/3 \\ 20/3 \\ 19/2 \\ 19/2 \\ 1 \end{pmatrix} \right\}$. Og den unike score vektoren til M er $\begin{pmatrix} 0.2 \\ 0.2 \\ 0.2850 \\ 0.2850 \\ 0.03 \end{pmatrix}$.

Oppgave 6

Denne funksjonen returnerer en link-matrise fordi:

(1) Linje 3 til 10 er for å håndtere de første $n - 1$ kolonner, mens linje 11 til 16 håndterer den siste kolonnen.

(2) Linje 4 og 11 nuller ut alle koeffisientene på diagonalen til matrisen. Dette her er viktig fordi vi antar at et dokument ikke kan ha en hyper-link til seg selv.

(3) Linje 5 til 6 sjekker om en kolonne er lik nullvektoren, og hvis dette er sann, så setter vi det *siste* element til 1 slik at summen av alle elementene i kolonnen er lik 1 (viktig fordi en link-matrise må være stokastisk). Denne sjekkningen gjør vi også på den siste kolonnen (linje 12 til 13). Men siden den siste element i den siste kolonnen ligger på diagonalen, derfor istedet for siste element, må vi sette den *første* element i siste kolonnen til 1.

```
1    function A=randlinkmatrix(n)
2      A = round(rand(n,n));
3      for k=1:(n-1)
4          A(k,k) = 0;
5          if (A(:,k) == 0)
6              A(n,k) = 1;
7          end
8          s = sum(A(:,k));
9          A(:,k) = (1/s) * A(:,k);
10     end
11     A(n,n) = 0;
12     if (A(:,n) == 0)
13         A(1,n) = 1;
14     end
15     s = sum(A(:,n));
16     A(:,n) = (1/s) * A(:,n);
```

(4) Hvis en kolonne ikke er lik nullvektoren, da må summen av elementene i kolonnen være større eller lik 1 (fordi den random matrisen som vi oppretter i linje 2 kan bare ha koeffisientene lik 0 eller 1). Da er det nødvendig å skalere hele kolonnen med en konstant (som er det "resiproke" tallet til den nåværende summen), slik vi har gjort i linje 8 til 9 og linje 15 til 16.

Oppgave 7

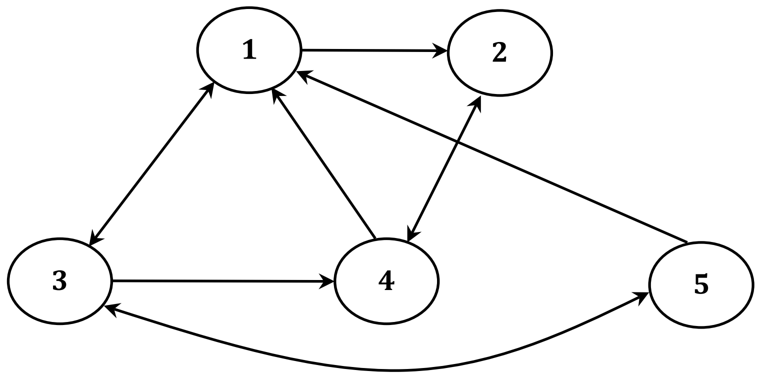
MATLAB kjøring

```
>> sym(randlinkmatrix(5))

ans =

[   0, 0, 1/3, 1/2, 1/2]
[ 1/2, 0,   0, 1/2,   0]
[ 1/2, 0,   0,   0, 1/2]
[   0, 1, 1/3,   0,   0]
[   0, 0, 1/3,   0,   0]
```

Tegning



Oppgave 8, 9, 10

Se Appendix B og C for MATLAB kode og utskrift.

Appendix A (for oppgave 5)

```
>> A = sym([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]);
>> m = sym(0.15);
>> M = (1-m)*A + (m/5)*ones(5);
>> eval(M)
ans =
    0.0300    0.8800    0.0300    0.0300    0.0300
    0.8800    0.0300    0.0300    0.0300    0.0300
    0.0300    0.0300    0.0300    0.8800    0.4550
    0.0300    0.0300    0.8800    0.0300    0.4550
    0.0300    0.0300    0.0300    0.0300    0.0300
>> basis = null(M-eye(5))
basis =
 20/3
 20/3
 19/2
 19/2
    1
>> score_vector = eval((1/sum(basis))*basis)
score_vector =
    0.2000
    0.2000
    0.2850
    0.2850
    0.0300
```

Appendix B (for oppgave 8 og 9)

```matlab
function ans = isStochastic(A)

% Check if A is a square matrix
[n,d] = size(A);
if (n ~= d)
    ans = false;
    return;
end

% Main loop to check if all columns satisfy the stochastic conditions
for k = 1:n
    col = A(:,k);
    if((sum(col) ~= 1) || (min(col) < 0))
        ans = false;
        return;
    end
end
ans = true; % If function reaches this point, then A must be stochastic
```

```matlab
function x = ranking(A)

if(~isStochastic(A))
    error('Cannot compute ranking for non-stochastic input matrix!');
    return;
end
m = sym(0.15);
n = size(A,1);
M = (1-m)*sym(A) + (m/n)*ones(n);
basis = null(M-eye(n));
x = eval((1/sum(basis))*basis); % eval is used here to get decimal numbers
```

```matlab
function curX = rankingapprox(A,delta)

m = sym(0.15);
n = size(A,1);
M = (1-m)*sym(A) + (m/n)*ones(n);
prevX = sym((1/n)*ones(n,1)); % previous X, initialized to x{0}
curX  = M*prevX;              % current X, initialized to x{1}

while (abs(max(curX-prevX)) > delta)
    prevX = curX;
    curX  = M*prevX;
end
curX = eval(curX); % eval is used here to get decimal numbers
```

Appendix C (for oppgave 10)

```
>> A = [0 0 1 1/2;
1/3 0 0 0;
1/3 1/2 0 1/2;
1/3 1/2 0 0];
>> ranking(A)

ans =

    0.3682
    0.1418
    0.2880
    0.2021

>> rankingapprox(A,0.005)

ans =

    0.3664
    0.1422
    0.2884
    0.2030
```

Fra MATLAB-utskriftet ovenfor, kan vi se at resultatene for $ranking(A)$ og $rankingapprox(A)$ er lik hverandre på 2 desimaler, og dette er fordi vi har brukt $\delta = 0.005$ (dvs. vi tillater ± 0.005 avvik fra den riktige score-vektoren)