MAT 1120

Obligatorisk oppgave 1

Jonas Semprini Næss

16. september 2020

Oppgave 1:

Skriv opp link-matrisen A for weben gitt ved Figur 2. Bestem deretter en basis for nullrommet til matrisen $A - I$, og finn den unike score-vektoren (husk at summen av elementene i score-vektoren skal være 1). Angi den tilhørende rangeringen av dokumentene.

Løsning:

Vi har fra Figur (1) følgende relasjon mellom de fire dokumentene skrevet på vektorform.

$$n_1 = \left(0, \frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$$

$$n_2 = \left(0, 0, \frac{1}{2}, \frac{1}{2}\right)$$

$$n_3 = \left(\frac{1}{2}, 0, 0, \frac{1}{2}\right)$$

$$n_4 = (1, 0, 0, 0)$$

Hvor $n_1, \ldots, n_4$ svarer til dokumentene i weben og entriene i vektoren til linken mellom tilhørende dokument. Setter vi så kolonnevektorene sammen til en matrise får vi linkmatrisen

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

For å finne en basis for nullrommet til $C = A - I$ radreduserer vi matrisen til vi når radredusert form. Det gir

$$\mathrm{rref}(A - I \mid \mathbf{0}) \sim \cdots \sim \begin{pmatrix} 1 & 0 & 0 & -\frac{4}{3} & 0 \\ 0 & 1 & 0 & -\frac{4}{9} & 0 \\ 0 & 0 & 1 & -\frac{2}{3} & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Vi ser at det kun eksiterer én fri variabel (x_4) hvilket betyr at nullrommet til matrisen $C = A - I$ er utspent av vektoren

$$\underline{\underline{B = \left\{ \frac{1}{3} \begin{pmatrix} 4 \\ \frac{4}{3} \\ 2 \\ 3 \end{pmatrix} \right\}}}.$$

For å beregne den unike score-vektoren tar vi summen av entriene $(4 + \frac{4}{3} + 2 + 3 = \frac{31}{3})$, og deler følgelig hver entri på denne summen. Det gir dermed at score-vektoren S blir

$$\underline{\underline{S = \begin{pmatrix} \frac{12}{31} \\ \frac{4}{31} \\ \frac{6}{31} \\ \frac{9}{31} \end{pmatrix}}}.$$

Nå som score-vektoren er funnet er rangeringen enkel, der rangeringsvektoren ser slik ut

$$\underline{\underline{\mathrm{Rangering}(S) = \begin{pmatrix} \frac{12}{31} \\ \frac{9}{31} \\ \frac{6}{31} \\ \frac{4}{31} \end{pmatrix} = \begin{pmatrix} P(n_1) \\ P(n_4) \\ P(n_3) \\ P(n_2) \end{pmatrix}}}$$

Oppgave 2:

Figur 3 viser det vi kaller en usammenhengende web, det vil si at weben kan kan deles opp i områder som ikke refererer til hverandre. Skriv opp link-matrisen A, og bestem en basis for nullrommet til matrisen A - I. Finnes det en unik score-vektor?

Løsning:

Bruker vi samme ressonement som i oppgave 1.) har vi at linken mellom hvert området kan skrives som vektorene

$$n_1 = (0, 0, 0, 0, 1)$$

$$n_2 = (0, 0, 0, 1, 0)$$

$$n_3 = \left(0, \frac{1}{2}, 0, \frac{1}{2}, 0\right)$$

$$n_4 = (1, 0, 0, 0, 0)$$

og link-matrisen A blir følgelig

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

og for basisen til $\mathrm{Nul}(A - I)$ får vi

$$\mathrm{rref}(A - I \mid \mathbf{0}) \sim \cdots \sim \begin{pmatrix} 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

der vi observerer at vi har to frie variable (henholdsvis x_4, x_5) hvilket betyr at vi kan skrive basisvariablene på formen

$$\text{I. } x_1 = x_5$$

$$\text{II. } x_2 = x_4$$

$$\text{III. } x_3 = 0$$

der resulterende vektorer(basis) fra likningssytemet blir

$$B = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right\}.$$

Vi kan nå ut i fra basisen konkludere med at det ikke finnes én unik score-vektor ettersom basisen utgjør to lineært uavhengige vektorer. Dette ville gitt en selvmotsigelse hvis vi skulle funnet en score-vektor siden det finnes uendelig mange kombinasjoner av x_4, x_5 og ergo ikke en entydig løsning.

Oppgave 3:

Er link-matrisene fra Oppgave 1 og 2 stokastiske? Er de regulære? Kan du forklare med enkle ord hvordan en web kan gi en link-matrise som ikke er stokastisk?

Løsning:

Følger vi definisjonen fra kapittel 4.9 i "Linear Algebra and its Applications" er matrisen stokastisk (eller mer presist venstre stokastisk) hvis og bare hvis summen av entriene i hver kolonne summerer til 1, da gitt at hver entri er et ikke-negativt tall større eller lik null, men mindre eller lik 1.

Matematisk forklart kan vi anta at vi har en matrise K med n rader og n

kolonner slik at

$$K^{n\times n} = \begin{pmatrix} K_{1,1} & \cdots & K_{1,j} & \cdots & K_{1,n} \\ K_{2,1} & \cdots & K_{2,j} & \cdots & K_{2,n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ K_{i,1} & \cdots & K_{i,j} & \cdots & K_{i,n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ K_{n,1} & \cdots & K_{n,j} & \cdots & K_{n,n} \end{pmatrix}$$

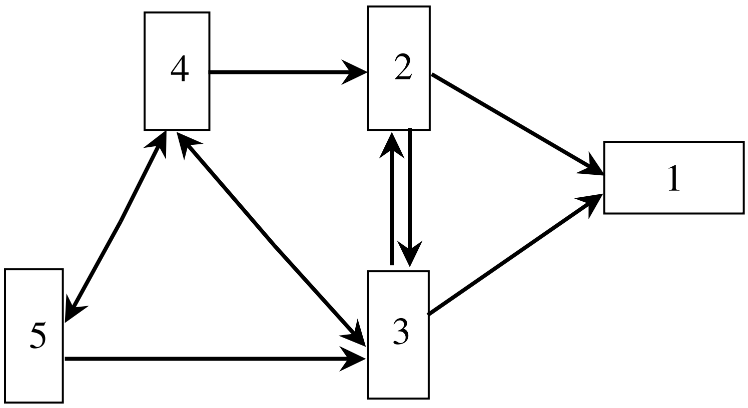
hvor summen av entriene i hver kolonne må summere til 1.

$$\boxed{\sum_{j=1}^{n} K_{i,j} = 1}$$

Ser vi da på matrisene i begge eksemplene ser vi raskt at begge oppfyller dette kravet, hvilket betyr at begge er stokastiske matriser.

For å sjekke om matrisene er regulære kan vi benytte oss av Perron-Frobenius teoremet som sier "Enhver stokastisk matrise K har en likevektsvektor $\boldsymbol{x}$. Hvis K er regulær så er $\boldsymbol{x}$ entydig bestemt." Her er likevektsvektoren relatert til score-vektoren, og fra hva vi fant ut i oppgave 1 og 2.) var det kun link-matrisen i oppgave 1.) som hadde en unik/entydig bestemt score-vektor. Derfor er link-matrisen i oppgave 1.) regulær og ikke link-matrisen i oppgave 2.).

Vi har vist hvordan en matrise kan være stokastisk og regulær, men en link-matrise i en web kan også gi en ikke-stokatisk matrise. Dette kan visualiseres ved skissen under



Figur 1: Web som vil gi en ikke-stokastisk matrise

Her vil link-matrisen til følgende web være

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

hvor summen av entriene i kolonne 1 er lik null som bryter med kravet om at matrisen skal være stokastisk. Mer generelt kan vi definere tilfellet slik

Anta at vi har en web W med n dokumenter hvor linkene mellom hvert dokument kan skrives på formen $W_n = \{1, 2 \ldots, n-1\}$. Da kan det eksistere et dokument W_i for $i \in W$ slik at $W_i = \{\emptyset\}$, og dermed gi en kolonnevektor med kun nuller som entrier.

Oppgave 4

Anta at A er stokastisk. Begrunn at matrisen M gitt ved (3) er stokastisk og regulær. (M vil dermed ha en unik score-vektor).

Løsning:

Vi har at $M = (1-m)A + mS$ for $m \in (0,1)$, $S^{n\times n} = \begin{pmatrix} \frac{1}{n} & \cdots & \frac{1}{n} \\ \vdots & \ddots & \vdots \\ \frac{1}{n} & \cdots & \frac{1}{n} \end{pmatrix}$ og

$A^{n\times n} = \begin{pmatrix} A_{1,1} & \cdots & A_{1,j} & \cdots & A_{1,n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ A_{n,1} & \cdots & A_{n,j} & \cdots & A_{n,n} \end{pmatrix}$.

Summerer vi nå over (i´te kolonne) på venstre og høyre side av likningen over har vi

$$\sum_{j=1}^{n} M_{i,j} = (1-m)\sum_{j=1}^{n} A_{i,j} + m\sum_{j=1}^{n} S_{i,j}$$
$$1 = 1 - m + m$$
$$1 = 1$$

og dermed har vi vist at M er stokastisk.

Videre ønsker vi å vise at M er regulær, hvilket betyr at alle entriene i M må være større enn null for minst en potens av M. Tar vi utgangspunktet i utrykket for M har vi

$$\begin{aligned}
M &= (1-m)A + mS \\
&= (1-m)\begin{pmatrix} A_{1,1} & \cdots & A_{1,j} & \cdots & A_{1,n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ A_{n,1} & \cdots & A_{n,j} & \cdots & A_{n,n} \end{pmatrix} + m\begin{pmatrix} \frac{1}{n} & \cdots & \frac{1}{n} \\ \vdots & \ddots & \vdots \\ \frac{1}{n} & \cdots & \frac{1}{n} \end{pmatrix} \\
&= \begin{pmatrix} A_{1,1}(1-m) & \cdots & A_{1,j}(1-m) & \cdots & A_{1,n}(1-m) \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ A_{n,1}(1-m) & \cdots & A_{n,j}(1-m) & \cdots & A_{n,n}(1-m) \end{pmatrix} + \begin{pmatrix} \frac{m}{n} & \cdots & \frac{m}{n} \\ \vdots & \ddots & \vdots \\ \frac{m}{n} & \cdots & \frac{m}{n} \end{pmatrix} \\
&= \begin{pmatrix} A_{1,1}(1-m) + \frac{m}{n} & \cdots & A_{1,j}(1-m) + \frac{m}{n} & \cdots & A_{1,n}(1-m) + \frac{m}{n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ A_{n,1}(1-m) + \frac{m}{n} & \cdots & A_{n,j}(1-m) + \frac{m}{n} & \cdots & A_{n,n}(1-m) + \frac{m}{n} \end{pmatrix}
\end{aligned}$$

Studerer vi sluttmatrisen ser vi at leddene inneholdt i $(1-m)A$ kan være null ettersom A er stokastisk, men leddet $\frac{m}{n}$ vil alltid være strengt større enn null grunnet $m \in (0,1)$ og $n > 0$ hvilket betyr at $\frac{m}{n} > 0$, og M er dermed regulær fordi en vilkårlig potens av et tall større null vil igjen gi et tall større enn null.

Oppgave 5

Skriv opp Google-matrisen M når link-matrisen A er den du fant i Oppgave 2,og m velges lik 0.1. Bruk Matlab-kommandoen null til å bestemme nullrommet til M - I. Angi så den unike score-vektoren til M.

Løsning:

Løser oppgaven ved følgende Matlab-script

```
A = [0 0 0 0 1; 0 0 1/2 1 0; 0 0 0 0 0; 0 1 1/2 0 0; 1 0 0 0 0]

S = (1/5)*(ones(5))
I = eye(5)
m = 0.1
M = (1 - m)*A + m*S

C = null(M - I, 'r')
Su = sum(C)
Final = C/Su
```

ved kjøring av programmet får vi at den unike score-vektoren er $\underline{\underline{S = \begin{pmatrix} \frac{1}{5} \\ \frac{29}{100} \\ \frac{1}{50} \\ \frac{29}{100} \\ \frac{1}{5} \end{pmatrix}}}$

Oppgave 6

I denne oppgaven skal du studere Matlab-koden under, som returnerer en $n \times n$ matrise A. Her gir kommandoen ***round(rand(n,n))*** *en tilfeldig $n \times n$ matrise med bare nuller og enere.*

```
function A = randlinkmatrix(n)
    A = round(rand(n,n));
    for k = 1: (n-1)
        A(k,k) = 0;
        if (A(:,k) == 0)
            A(n,k) = 1;
        end
```

```
        s = sum(A(:,k));
        A(:,k) = (1/s) * A(:,k);
    end
    A(n,n) = 0;
    if(A(:,n) == 0)
        A(1,n) = 1;
    end
    s = sum(A(:,n));
    A(:,n) = (1/s) * A(:,n);
end
```

Løsning:

Koden starter med å initiere en tifeldig $n \times n$ matrise med nuller og enere. Deretter benyttes en for-løkke som løper over første kolonne til nest siste. Herfra endres diagonal-entriene i matrisen til null ettersom en stokastisk link-matrise har nuller langs diagonalen. Så kommer en if-blokk som sjekker om entriene i den gitte kolonnen er lik null. Stemmer dette settes siste entri i den gitte kolonnen til 1. Videre regnes summen av den gitte kolonnen, hvor så entriene blir delt på denne summen.

Etter for-løkke prosedyren sjekkes siste element i den siste raden og kolonnen for å oppnå at diagonalen blir 0. Likeledes som i for-løkken sjekkes entriene i siste kolonne om kun består av nullere og hvis så settes elementet i rad en, siste kolonne til 1. Sluttvis regnes summen av siste kolonnen, og blir som over delt på denne summen for å passe på at summen av kolonnen blir 1.

Oppgave 7

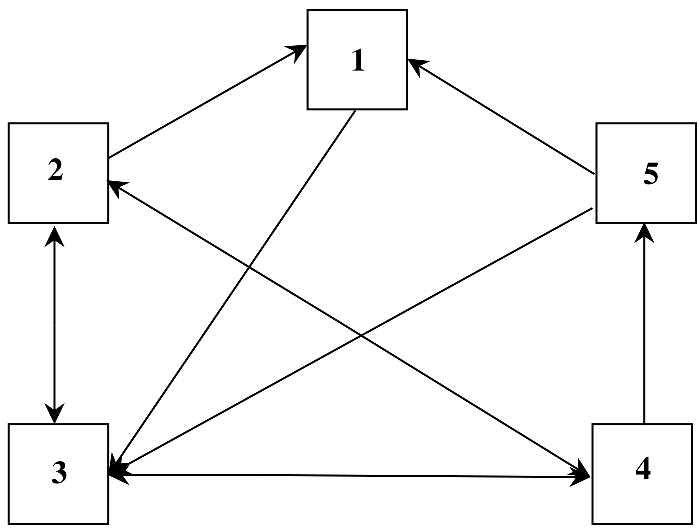
Kjør Matlab-funksjonen fra Oppgave 6 med $n = 5$, og tegn den tilhørende weben for link-matrisen som funksjonen returnerer.

Løsning:

Kjører vi koden over får vi tilhørende $A^{5\times 5}$ matrise.

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

Da har vi følgende skisse av weben til A



Figur 2: Web tilhørende matrise A i oppgave 7.)

Oppgave 8

Skriv en Matlab-funksjon ranking(A) som returnerer (den unike) score-vektoren x for Google-matrisen M definert ved (3) med $m = 0.1$ når input-matrisen A er en stokastisk link-matrise. Hvis A ikke er stokastisk skal funksjonen returnere en feilmelding. Du kan gjerne anvende Matlab-kommandoen null for å bestemme nullrommet til $M - I$.

Løsning:

Vi løser oppgaven ved å lage følgende Matlab-script

```
function E = ranking(A, n)

    m = 0.1;
    S = (1/n)*(ones(n));
    I = eye(n);

    for i = 1:length(A)
        if (sum(A(:,i)) < 1 || sum(A(:,i)) > 1)
           disp('Matrisen er ikke stokastisk')
           break
        end
    M = (1-m)*A + m*S;
    C = null((M - I));
    D = sum(C);
    E = C/(D);
```

```

    end
end
```

Genererer vi en tilfeldig stokastisk matrise A ved hjelp av randlinkmatrix scriptet får den unike score-vektoren $S = \underline{\underline{\begin{pmatrix} \frac{1}{5} \\ \frac{22}{65} \\ \frac{1}{5} \\ \frac{59}{325} \\ \frac{2}{25} \end{pmatrix}}}$

(dette er kun et tilfeldig eksempel av en rekke mulige matriser man kan få ved bruk av randlinkmatrix scriptet.)

Oppgave 9

Skriv en Matlab-funksjon rankingapprox(A, delta) som beregner vektoren $\boldsymbol{x}_k$ *ved hjelp av (4) helt til den støter på et naturlig tall* k *som er slik at* $max_j|\boldsymbol{x}_{k(j)} - \boldsymbol{x}_{k(j)}| < delta$*; funksjonen skal da returnere vektoren* $\boldsymbol{x}_k$*. Her antas det at* A *er stokastisk og at delta er et positivt (lite) tall.*

Løsning:

Vi kan løse oppgaven ved å lage følgende Matlab-script

```
function p = rankingapprox(A, delta, n)

    S = (1/n)*(ones(n));
    m = 0.1;
    M = (1 - m)*A + m*S;
    x_inital = (1/n)*(ones(n,1))
    x1 = M * x_inital;
    xk = M * x1;

    while abs(max(xk) - max(x1)) >= delta
        x1 = xk;
        xk = M * x1;
    end
    p = x1;
end
```

Oppgave 10

Sammenlign vektorene som du får ut når du anvender funksjonene ranking(A) og rankingapprox(A, delta) på link-matrisen A fra Oppgave 1 og du velger delta = 0.01

Løsning:

Kjører vi ranking(A) med matrisen $A = \begin{pmatrix} 0 & 0 & \frac{1}{2} & 1 \\ \frac{1}{3} & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{2} & 0 & 0 \\ \frac{1}{3} & \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}$, får vi score-vektoren $S_1 \begin{pmatrix} 0.3746 \\ 0.1374 \\ 0.1992 \\ 0.2888 \end{pmatrix}$, og likeledes med rankingapprox(A, delta, n) får vi score-vektoren $S_2 = \begin{pmatrix} 0.3797 \\ 0.1307 \\ 0.1977 \\ 0.2918 \end{pmatrix}$. Den originale score-vektoren vi fikk i oppgave 1.) er $S_o = \begin{pmatrix} 0.3870 \\ 0.1290 \\ 0.1935 \\ 0.2903 \end{pmatrix}$ hvor vi dermed observerer at den approksimerte løsningen til score-vektoren i oppgave 1.) er den mest presise selv om differansen til ranking(A) - løsningen er minimal.

En viktig observasjon er at for delta = 0.01 så blir løsningen til den approkismerte vektoren mest presis, men for mindre og mindre deltaer så vil løsningen konvergere mot løsningen til Ranking(A), og således vike mer fra løsningen i oppgave 1.).