UNIVERSITETET
I OSLO

Masteroppgave

Strukturen til Dobbelt Stokastiske Matriser

Lars Hurlen

Matematikk, Lektorprogrammet
30 studiepoeng

Matematisk institutt
Det matematisk-naturvitenskapelige fakultet

Våren 2023

Lars Hurlen

Strukturen til Dobbelt Stokastiske Matriser

Veileder:
Geir Dahl

Sammendrag

Målet med denne oppgaven er å utforske strukturen til dobbelt stokastiske matriser. Hovedfokuset er å utforske verdien og posisjonen til elementene i dobbelt stokastiske matriser. Spesielt ser vi på strukturen i matrisene når enkelte av elementene er null. Vi viser at antall nuller og ikke-nuller, samt deres posisjon, er en begrensenede faktor for hvorvidt matrisen er dobbelt stokastisk. Dette framstilles ved konstruksjon av matriser, og ved hjelp av Birkhoffs teorem, hovedteoremet i oppgaven.

Innhold

Tabeller

Figurer

Anerkjennelser

Jeg ønsker å takke min veileder Geir Dahl for veiledning og motiverende innspill. Han oppfordret meg og mine medstudenter til å arbeide utforskende og selvstendig med oppgaven, noe som viste seg å være svært interessant og lærerikt. Det utforskende arbeidet har tillatt meg å kunne grave i teori, men også anvende teorien på sammenhenger jeg selv oppdaget underveis i arbeidet. Jeg vil også takke mine medstudenter på Lektorprogrammet for fem flotte år på Blindern, og takk til de på lesesalen denne våren på Niels Henrik Abels hus. Jeg har verdsatt det sosiale vel så mye som arbeidet med oppgaven, og den trivelige stemningen har virket motiverende på arbeidet mitt.

Lars Hurlen

Lars Hurlen
Oslo, Mai 2023

Kapittel 1

Introduksjon

Denne oppgaven kan ses på som todelt. I første del av oppgaven utledes teori om konveksitet og grafer, hvor vi blant annet utleder Königs teorem. Dette danner bakgrunnen for hoveddelen av oppgaven som omhandler Birkhoffs teorem og dobbelt stokastiske matriser. Denne delen inneholder utforskende arbeid hvor vi ser på strukturen til slike dobbelt stokastiske matriser.

Hoveddelen av oppgaven er en utforskende del som er gjort på bakgrunn av teori hentet fra litteratur oppgitt i referanselisten. Ettersom bevisene kan finnes i de bøkene, utelates de fleste fra oppgaveteksten. Eksempler til teoremene, figurer og utregninger er egne bidrag. I den utforskende delen av oppgaven i kapittel 6 og 7 er store deler av arbeidet mitt eget, og resultatene er funnet gjennom det utforskende arbeidet gjort i oppgaven. Det utforskende arbeidet finner man i delkapitlene 5.3, 6.2, 6.3, 6.4 og kapittel 7. Der anvender vi teorien fra de tidligere kapitlene. Kapittel 5.3 viser en sammenheng mellom nettverkstrøm, graf og dobbelt stokastiske matriser. Kapittel 6.2 - 6.4 diskuterer strukturen til dobbelt stokastiske matriser, og kapittel 7 diskuterer matriser som er mindre begrenset enn dobbelt stokastiske matriser.
Egne bidrar inkluderer også seks bevis jeg har gjort på egenhånd:

- Korollar 5.2.3
- Teorem 5.2.5
- Teorem 6.1.5
- Korollar 6.3.1 (s. 38-39 og s. 43)
- Korollar 7.1.3
- Korollar 7.2.4

Kapittel 2

Konveksitet

Dette kapitlet danner teorigrunnlaget for det oppgaven skal omhandle. Vi vil her introdusere begreper og teoremer som gir en bakgrunn for hoveddelen i oppgaven, Birkhoff-polytopen. Flere av teoremene presenteres uten bevis, men bevisene kan finnes i heftet til Christophersen og Ranestad [CR22].

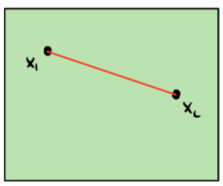
2.1 Konvekse mengder

I dette kapitlet skal vi gi noen viktige definisjoner relatert til konveksitet og konvekse mengder som vil være sentralt i hoveddelen av oppgaven. Eksempler på konvekse mengder er trekanter og firkanter i $\mathbb{R}^2$. Definisjonene til dette kapitlet er hentet fra et kompendium i geometri til faget MAT2500 (geometri) [CR22].

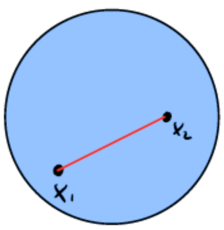
Definisjon 2.1.1. En mengde $C \in \mathbb{R}^n$ er *konveks* hvis

$$(1 - \lambda)x_1 + \lambda x_2 \in C$$

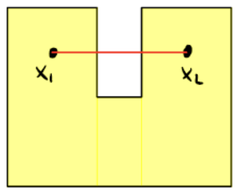
for alle $x_1, x_2 \in C$ og alle $0 \leq \lambda \leq 1$.
For å få en forståelse for dette rent geometrisk, betyr dette at i planet $\mathbb{R}^2$ og i rommet $\mathbb{R}^3$ at C inneholder linjesegmentet mellom hvert par av punkter i C.



(a) Konveks mengde i $\mathbb{R}^2$



(b) Konveks mengde i $\mathbb{R}^2$

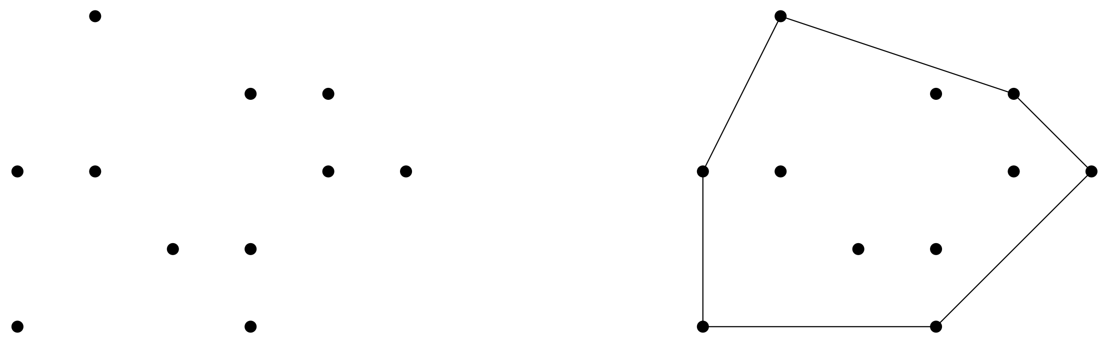


(c) Ikke konveks mengde i $\mathbb{R}^2$

Figur 2.1: Konvekse og ikke konvekse mengder i $\mathbb{R}^2$

Definisjon 2.1.2 (Konveks innhylling). Den *konvekse innhyllingen* til en mengde $P \in \mathbb{R}^n$, er den minste konvekse mengden C som inneholder alle punktene til P. Man kan også definere konveks innhylling som den konvekse kombinasjonen av alle punkter i P.

Når man finner den konvekse innhyllingen av en mengde, får man en ny mengde som er konveks. Alle konvekse innhyllinger er dermed konvekse, noe som kommer til anvendelse senere i oppgaven.



(a) Mengde C

(b) Konveks innhylling av C

Figur 2.2: Konveks innhylling

Definisjon 2.1.3 (Ekstrempunkter). Gitt en konveks mengde $C \in \mathbb{R}^n$, kalles et punkt $z \in C$ et *ekstrempunkt* til C hvis det *ikke* eksiterer en $\lambda \in (0,1)$ slik at $z = (1-\lambda)x_1 + \lambda x_2$ for alle $x_1, x_2 \in C$ hvor $x_1 \neq x_2$.

Et punkt i en mengde er et ekstrempunkt hvis det ikke kan skrives som en konveks kombinasjon av noen andre punkter i mengden. Både denne definisjonen, og definisjonen om konveksitet, vil være sentrale senere i oppgaven. En konveks kombinasjon minner mye om en lineær kombinasjon, men konstantene i en konveks kombinasjon skal summere til 1.

Definisjon 2.1.4. Et punkt z kalles en *lineær kombinasjon* av punktene $x_1, x_2, \ldots, x_n$ dersom det finnes konstanter $c_1, c_2, \ldots, c_n$ slik at

$$z = c_1x_1 + c_2x_2 + \cdots + c_nx_n$$

Definisjon 2.1.5. En *konveks kombinasjon* w er en lineær kombinasjon

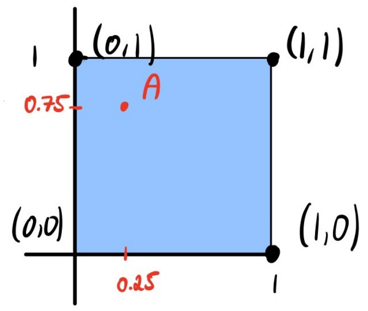
$$w = c_1x_1 + c_2x_2 + \cdots + c_nx_n$$

hvor konstantene $c_1 + c_2 + \cdots + c_n = 1$ og $c_i \geq 0$.

Teorem 2.1.6. *I en konveks mengde kan alle punktene i mengden skrives som en konveks kombinasjon av ekstrempunktene.*

Eksempel 2.1.7. Gitt et rektangel i $\mathbb{R}^2$ med ekstrempunktene $x_1 = (0,0), x_2 = (1,0), x_3 = (1,1), x_4 = (0,1)$ og et punkt $A = (0.25, 0.75)$ som ligger innhyllet i rektangelet, kan vi skrive A som en konveks kombinasjon av to av ekstrempunktene. Vi bruker punktene x_2, x_4 og $\lambda = 0.25$

$$A = (1-0.25)\cdot x_4 + 0.25\cdot x_2 = 0.75\cdot(0,1) + 0.25\cdot(1,0) = (0,0.75) + (0.25,0) = (0.25,0.75)$$



Figur 2.3: $A = 0.75 \cdot x_4 + 0.25 \cdot x_2$

□

2.2 Polytop

Definisjon 2.2.1. En mengde $P \subset \mathbb{R}^n$ kalles en *polytop* hvis den er den konvekse innhyllingen av en endelig mengde punkter i $\mathbb{R}^n$.

Et (konvekst) *polygon* er den konvekse innhyllingen av en endelig mengde punkter i $\mathbb{R}^2$. Et (konvekst) *polyeder* er den konvekse innhyllingen av et endelig antall punkter i $\mathbb{R}^3$. For å slippe å navngi uendelig mange polytoper av forskjellig dimensjoner, bruker vi *n-polytop* hvor n er størrelsen til dimensjonen til polytopen.

(a) Polygon

(b) Polyeder

Figur 2.4: Polytoper

Definisjon 2.2.2 (Hjørne (node)). I polytoper møtes flere av kantene og sideflatene. Snittet av kanter og sideflater kalles *hjørner*. Vi navngir disse punktene som enten hjørner eller *noder*.

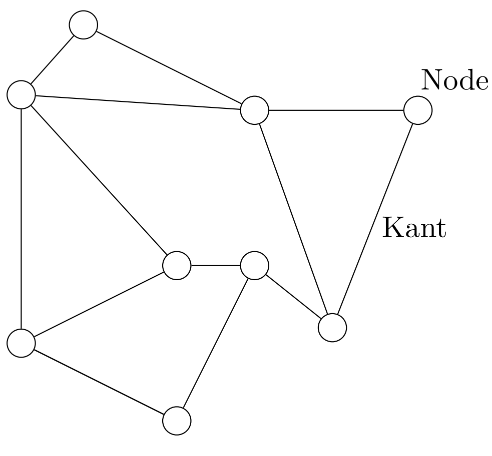
Kapittel 3

Graf

3.1 Graf

For å kunne bevise et av hovedteoremene i oppgaven trenger vi en grundig innføring i *grafteori*. For å få en intuitiv forståelse av hva en *graf* er, kan vi bruke eksempler fra hverdagen. Vi har flere eksempler på mengder punkter med linjer mellom par av disse punktene. Punktene kan representere mennesker, hvor linjene er et vennskapsbånd mellom to venner [BM76]. Geir Dahl [Dah01] bruker veikryss og gater som en illustrasjon som eksempel på en tilknytting av punkter og tilknytninger mellom punktene. Hver gate forbinder to gatekryss.

Vi har visse gatekryss, som vi kaller *noder* i grafen og vi har visse gater, som vi kaller *kanter* i grafen. En kant er en forbindelse mellom to noder [Dah01].



Figur 3.1: En graf som illustrerer et veinettverk.

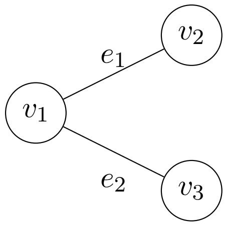
For å definere en graf vil vi bruke definisjonen til Gross m.fl. fra boken *Handbook of Graph Theory* [GYZ13]

Definisjon 3.1.1. En *graf* $G = (V, E)$ består av to mengder V og E.

- Elementene i V kalles *noder* (engelsk: vertices)
- Elementene i E kalles kanter (engelsk: edges)
- Hver kant er assosiert med én eller to noder som kalles *endepunktene*. En kant kobler sammen to endepunkter.

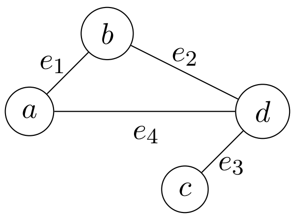
Kantmengden til en graf $G = (V, E)$ består, naturlig nok, av kanter. Vi skriver kantene $e \in E$ som en mengde av nodene, $e = (u, v)$. Der u og v er noder, $u, v \in V$.
Enhver graf kan representeres ved en tegning i planet ved å tegne nodene som punkter og kantene som linjer mellom nodene. For en graf vil ikke en slik tegning være unik, men en nyttig fremstilling for å forstå sammenhengen mellom noder og kanter. Alle grafene i denne oppgaven er *enkle grafer*, og vil ikke ha kanter hvor begge endepunktene er samme node, og to kanter vil ikke dele de samme to endepunktene.

Eksempel 3.1.2. Gitt en graf $G = (V, E)$ der $V = \{v_1, v_2, v_3\}$ og $E = \{e_1, e_2\}$ der $e_1 = \{v_1, v_2\}, e_2 = \{v_1, v_3\}$.



□

Eksempel 3.1.3. Gitt en graf $F = (V, E)$ der $V = \{a, b, c, d\}$ og $E = \{e_1, e_2, e_3, e_4\}$ der $e_1 = \{a, b\}, e_2 = \{b, d\}, e_3 = \{c, d\}, e_4 = \{d, a\}$.



Det spiller ingen rolle om vi skriver $e_4 = \{d, a\}$ eller $e_4 = \{a, d\}$ da retning foreløpig ikke har noe å si for kanten. □

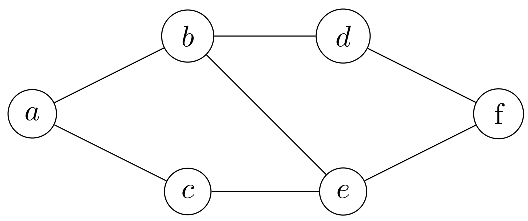
3.1.1 Bipartitt graf

I store deler av denne oppgaven skal vi se på et mer spesielt tilfelle av grafer, nemlig bipartitte grafer. En bipartitt graf er en graf hvor nodene til grafen kan deles inn i to delmengder og ingen av nodene i hver delmengde deler kant med noen andre noder i den respektive delmengden [BM76].

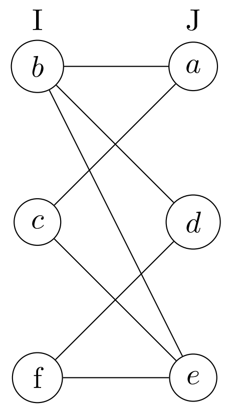
Definisjon 3.1.4 (Bipartitt graf). En graf $G = (V, E)$ er en *bipartitt graf* hvis $V = I \cup J$ og I og J er disjunkte. Hver kant $e \in E$ er på formen $e = (i, j)$ der $i \in I, j \in J$.

En vanlig måte å representere en bipartitt graf, er med en figur hvor vi deler nodemengdene I og J i to kolonner. Det vil ikke være nødvendig å skrive opp en liste av alle elementene i V og E, det er tilstrekkelig å tegne nodene og kantene.

Eksempel 3.1.5. $G = (V, E)$ er en bipartitt graf.



Det er ikke nødvendigvis åpenbart hvordan man kan dele nodene inn i to delmengder, I og J, men ved å dele nodene opp i to kolonner er det enklere å se at nodene kan deles i to delmengder.



G er en bipartitt graf som kan deles inn i $I = \{b, c, f\}$ og $J = \{a, d, e\}$, og $V = I \cup J$. $\square$

Videre vil vi plassere nodene til bipartitte grafer i to kolonner, I og J. Hver node i I vil være navngitt ved $i_1, i_2, \ldots, i_n$ og hver node i J vil være navngitt ved $j_1, j_2, \ldots, j_m$. Der n og m er antall noder i I og J.

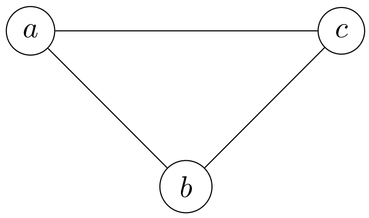
Definisjon 3.1.6 (Insident). En node u er *insident* til en kant e hvis u er et endepunkt(node) til e, altså at $e = (u, v)$ for $v \in V$. Det er også vanlig å si at en node er *inntil* en kant.

Definisjon 3.1.7 (Nodedekning). Gitt en graf $G = (V, E)$ kalles en delmengde $S \subset V$ en *nodedekning* hvis nodene i S er inntil hver kant i G.

3.2 Graf til matrise

Til nå har vi sett på hvordan en graf kan representeres som en figur, og det vil i mange sammenhenger være den mest intuitive måten å framstille grafer på, men det er ikke den eneste. Hoveddelen av denne oppgaven vil omhandle matriser, og en graf kan også representeres ved en matrise.

Definisjon 3.2.1. Gitt en graf $G = (V, E)$, kan vi konstruere en $n \times n$ matrise, A, som representerer G. En slik matrise kalles en *nabomatrise*. Hver rad $i \in 0, 1, \ldots, n$ og kolonne $j \in 0, 1, \ldots, n$ er hver node $v_1, v_2, \ldots v_n$ og verdien til hvert element $a_{ij} \in A$ er antall kanter $e \in E$ mellom nodene v_i og v_j.



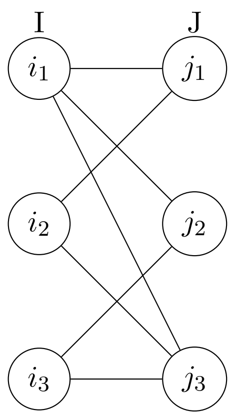
Figur 3.2: Eksempel

Grafen i figur 3.2 har tre noder $V = \{a, b, c\}$ og tre kanter $E = \{(a, b), (b, c), (a, c)\}$. For å konstruere en nabomatrise, setter vi nodene $a, b, c \in V$ på hver sin rad i og kolonne j. Vi får nabomatrisen

$$A = \begin{array}{c} \\ \text{a} \\ \text{b} \\ \text{c} \end{array} \begin{array}{c} \begin{array}{ccc} \text{a} & \text{b} & \text{c} \end{array} \\ \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \end{array}$$

hvor hvert element $a_{ij} \in A$ er antall kanter mellom nodene. I posisjon $i = 2, j = 1$ er elementet 1, som forteller oss at det er én kant mellom node b og a. Størrelsen til en nabomatrise er lik antall noder, n, og vil alltid være kvadratisk.

Eksempel 3.2.2. Gitt en bipartitt graf $G = (V, E)$ hvor $V = I \cup J$



vil nabomatrisen til G være

$$\begin{array}{c} \\ \text{i1} \\ \text{i2} \\ \text{i3} \\ \text{j1} \\ \text{j2} \\ \text{j3} \end{array} \begin{array}{c} \begin{array}{cccccc} \text{i1} & \text{i2} & \text{i3} & \text{j1} & \text{j2} & \text{j3} \end{array} \\ \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

Ettersom det ikke går kanter mellom nodene i I og det ikke går kanter mellom nodene i J, består nabomatrisen av to deler med 3×3 $\square$

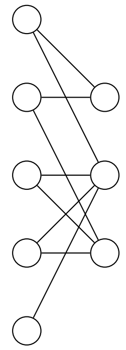
For bipartitte grafer er det vanlig å representere en graf med en *redusert nabomatrise*. Hver rad representerer hver $i \in I$, og hver kolonne representerer de forskjellige $j \in J$. Dette gjør vi ettersom vi vet at ingen av nodene i I deler kant med hverandre og ingen noder i J deler kant med hverandre.

Teorem 3.2.3. *For en bipartitt graf, $G = (V, E)$ med $n \times m$ noder kan enhver nabomatrise, N, skrives som en $n \times m$ matrise*

$$N = \begin{pmatrix} 0_n & B \\ B^T & 0_m \end{pmatrix}$$

*Der 0_n er en $n \times n$- og 0_m er en $m \times m$ matrise hvor hvert element er 0. B kalles den **reduserte nabomatrisen** til G, og B^T er B transponert. B er en $n \times m$ matrise.*

Eksempel 3.2.4. Gitt en bipartitt graf $G = (V, E)$ med 5 noder i I og 3 noder i J, skal vi konstruere den reduserte nabomatrisen B.



$$B = \begin{array}{cc} & \begin{array}{ccc} j_1 & j_2 & j_3 \end{array} \\ \begin{array}{c} i_1 \\ i_2 \\ i_3 \\ i_4 \\ i_5 \end{array} & \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix} \end{array}$$

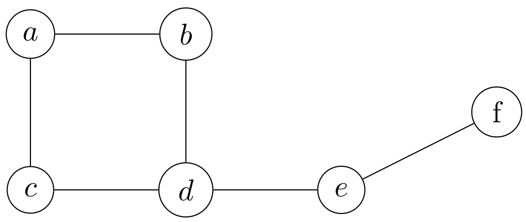
□

3.3 Matching

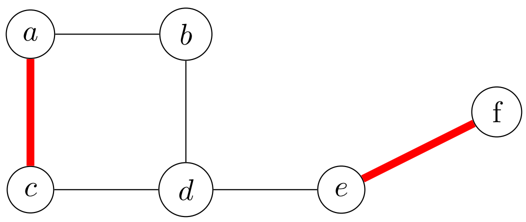
I grafteori er det vanlig å snakke om delmengder av grafer. En graf består som kjent av noder og kanter, men også innenfor disse mengdene er det vanlig å se på delmengder. I en bipartitt graf separerer vi nodene i to delmengder, I og J, men også kantene kan deles opp i delmengder. En matching er en delmengde av kantene hvor ingen av kantene i delmengden har en node som felles endepunkt. Teori om matching [GYZ13], [AMO93] er sentralt i Königs teorem som vi senere skal utlede.

Definisjon 3.3.1. Gitt en graf $G = (V, E)$, vil en kantmengde $M \in E$ være en *matching* når ingen av kantene i M er inntil samme node. Nodene i M sies å være *matchet*, og nodene som ikke er i M sies å være *ikke matchet*. Når vi snakker om størrelsen til en matching, altså antall kanter r med i M, er det vanlig å si at M har kardinalitet r.

Eksempel 3.3.2. Gitt grafen $G = (V, E)$



har vi en matching $M_1 = \{(a, c), (e, f)\}$ som dekker kantene fra a til c og fra e til f, så nodene a, b, e, f er matchet. b og d er ikke matchet.



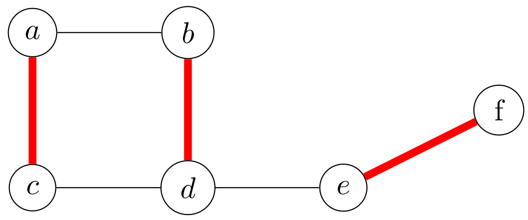
Dette eksemplet illustrerer en matching, men den viser ikke den maksimale matchingen.

□

Definisjon 3.3.3. Den *maksimale matchingen* til en graf $G = (V, E)$ er en matching hvor det ikke kan legges til flere kanter i matchingen, og er med andre ord ikke en delmengde av en annen matching.

Definisjon 3.3.4. En matching M i grafen $G = (V, E)$ kalles en *maksimum matching* når det ikke finnes matchinger med større kardinalitet enn M. Da har vi dekket maksimum antall kanter i G. Det kan være flere forskjellige maksimum matchinger.

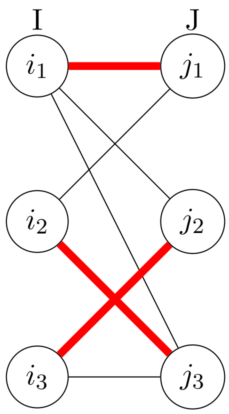
Eksempel 3.3.5. Gitt samme grafen G som i eksempel 3.3.7 er $M_2 = \{(a, c), (e, f), (b, d)\}$ en matching.



M_2 er en maksimal matching, men den er ikke unik. $M_3 = \{(a, b), (c, d), (e, f)\}$, er også en maksimal matching. En maksimal matching trenger dermed ikke være unik. M_2 er også maksimum matching, i tillegg til at den også er en perfekt matching. □

Definisjon 3.3.6. Gitt en graf $G = (V, E)$ er en *perfekt matching* er en matching som dekker alle nodene i G.

Eksempel 3.3.7. Vi ser på samme graf, $G = (V, E)$, som i eksempel 3.1.5, men nå med nye navn på nodene.



$M_1 = \{(i_1, j_1), (i_2, j_3), (i_3, j_2)\}$ er en perfekt matching da alle nodene er insident til kanter som er med i matchingen. En perfekt matching vil også være en maksimum matching og en maksimal matching. □

En graf kan representeres både ved en figur og ved en matrise. En matching kan også illustreres ved en figur og ved en matrise. Figuren kan omskrives til en nabomatrise, og en matching kan dermed illustreres i matrisen. Grafen i forrige eksempel er en bipartitt

graf, og ved bipartitte grafer er det vanlig å bruke en *redusert nabomatrise* der radene svarer til I og kolonnene svarer til J.
Fra eksempel 3.3.7 kan grafen omskrives til den reduserte nabomatrisen

$$A = \begin{array}{c c} & \begin{array}{ccc} j_1 & j_2 & j_3 \end{array} \\ \begin{array}{c} i_1 \\ i_2 \\ i_3 \end{array} & \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \end{array}$$

Fra eksempel 3.3.7 vet vi at grafen G har en perfekt matching. En perfekt matching er en delmengde av kantene som dekker alle nodene, men ingen av kantene er inntil samme node. På matriseform vil dette være å velge ikke-nuller, altså kanter mellom nodene, slik at ingen noder blir valgt to ganger. En perfekt matching vil altså være å velge én ikke-null på hver rad og kolonne.

I grafen G er det relativt greit å finne en perfekt matching på matriseform.

$$\begin{array}{c c} & \begin{array}{ccc} j_1 & j_2 & j_3 \end{array} \\ \begin{array}{c} i_1 \\ i_2 \\ i_3 \end{array} & \begin{bmatrix} \boxed{1} & 1 & 1 \\ 1 & 0 & \boxed{1} \\ 0 & \boxed{1} & 1 \end{bmatrix} \end{array}$$

Dette representerer en perfekt matching, $M_2 = \{(i_1, j_1), (i_3, j_2), (i_2, j_3)\}$ som er den samme vi fant i eksempel 3.3.7. Det er mulig å velge en annen perfekt matching, og vi kunne valgt den samme som i eksempel 3.3.7.

Kapittel 4

Königs teorem

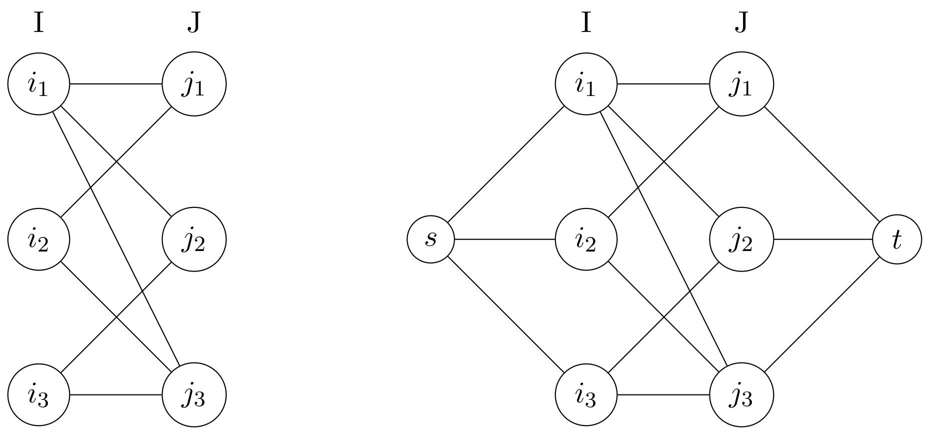
I dette kapitlet skal Königs teorem utledes for senere å kunne bevise Birkhoffs teorem. Königs teorem tar utgangspunkt i en bipartitt graf, og mer nøyaktig matching i en bipartitt graf. For å bevise Königs teorem brukes resultatet fra et annet teorem, maks strøm min kutt teoremet. Det er et kjent grafteoretisk teorem som omhandler nettverkstrøm. Beviset for maks strøm min kutt skal ikke utledes, men selve teoremet blir forklart samt teoremets sammenheng til Königs teorem vil også bli forklart. For å kunne gjøre dette trengs kunnskap om nettverkstrøm, noe som vil bli presentert i dette kapitlet før utledningen av Königs teorem. Königs teorem sier at verdien til en maksimal matching vil være lik verdien til den minste nodedekningen, og teoremet om maks strøm min kutt skal brukes til å vise at dette vil stemme for alle bipartitte grafer.

4.1 St-graf

Når man jobber med nettverkstrøm i en bipartitt graf $G = (V, E)$, er det vanlig å utvide grafen ved å legge til en startnode s og en sluttnode t. Vi kaller s for *kilden* og t for *sluket*. Man kan se for seg at det kommer strøm ut av kilden s som til slutt renner ut i sluket t. Nodene i G kan deles i to delmengder I og J, og alle nodene $i \in I$ deler kant med s, og alle nodene $j \in J$ deler kant med t. Både definisjonen av st-graf og de andre definisjonene i dette kapitlet er hentet fra Geir Dahls innføring i grafteori [Dah01].

Definisjon 4.1.1. Gitt en graf $G = (V, E)$ hvor $V = I \cup J$ hvor G er en del av nettverket $G' = (V', E')$ der $V' = V \cup (s, t)$ og $E' = E \cup \{(s, i) : i \in I\} \cup \{(j, t) : j \in J\}$. $G' = (V', E')$ kalles en *st-graf*.

En st-graf $G' = (V', E')$ kan illustreres ved figuren



(a) $G = (V, E)$ (b) $G' = (V', E')$

4.2 Nettverkstrøm

Definisjon 4.2.1 (Rettet graf). En *rettet graf* $G = (V, E)$ er en graf hvor hver kant har en bestemt retning. Retningen indikeres ofte med piler satt på hver kant. Et *rettet nettverk* er en rettet graf der det for hver kant går en strøm mellom to noder. Hvert par $e = (u, v)$ for $e \in E$ kalles en rettet kant (linje), der u er startnoden og v er sluttnoden.

Definisjon 4.2.2 (En vei). Gitt en graf $G = (V, E)$ er en *vei* en sekvens P av påfølgende kanter

$$P : (v_0, v_1), (v_1, v_2), \ldots, (v_{k-1}, v_k)$$

Definisjon 4.2.3 (Utgående- og inngående stjerne). Gitt en rettet graf $G = (V, E)$, finnes en *utgående stjerne* hvis alle kantene $e \in E$ ut fra en node u har samme retning til alle nodene $w \in V$.

$$\delta^+(v) = \{e \in E : e = (u, v) \text{ for } v \in V\}$$

Vi har en *inngående stjerne* når retningen er motsatt, og alle kantene $e \in E$ inn til u går med samme retning fra $v \in V$.

$$\delta^-(v) = \{e \in E : e = (v, u) \text{ for en } v \in V\}$$

Definisjon 4.2.4 (Kutt). Gitt en graf $G = (V, E)$ og en ikke-tom delmengde av nodene $S \subset V$ kalles kantmengden

$$\delta^+(S) = \{(u, v) \in E : u \in S, v \in V \setminus S\}$$

et *kutt*.

Definisjon 4.2.5 (Nettverk). Et *nettverk* er en rettet graf med kapasiteter knyttet til kantene. Vi har en rettet graf D med kanter E og noder V. Kapasiteten $c_{u,v} \geq 0$ for hver rettet kant (u, v) er øvre grense for hvor mye strøm som kan gå gjennom kanten. $c_{i,j}$ er *kapasiteten* til noden (i, j).

Kapasiteten til et kutt er summen av kapasitetene til kantene i kuttet.

Definisjon 4.2.6 (Strøm). For hver rettet kant $(u, v) \in E$ har vi en variabel $x_{i,j}$ som skal si hvor mye som strømmer i kanten (u, v) fra node u til node v.

$$x : E \to \mathbb{R}_+$$

Definisjon 4.2.7 (Strømbevaring). I en rettet graf $G = (V, E)$ har vi to noder s og t, og for alle $v \in V \setminus \{s, t\}$ krever vi at den totale strømmen inn i v er lik den totale strømmen ut av v.

$$\sum_{(u,v)\in\delta^-(v)} x_{u,v} = \sum_{(u,v)\in\delta^+(v)} x_{v,w}, \text{ for alle } v \in V \setminus \{s, t\} \tag{4.1}$$

Vi krever også at kapasiteten skal overholdes, altså at strømmen skal være mindre eller lik kapasiteten i hver kant.

$$0 \le x_{u,v} \le c_{u,v} \text{ for alle } (u, v) \in E \tag{4.2}$$

Hvis funksjonen x tilfredsstiller 4.1 og 4.2 kalles x en *st-strøm*. Hver st-strøm er tilknyttet en verdi

$$v(x) = \sum_{(s,v)\in E} x_{s,v}$$

$v(x)$ er lik mengden som strømmer ut fra noden s.

Lemma 4.2.8. *Gitt en rettet graf $G = (V, E)$ har vi en utgående stjerne s og en inngående stjerne t, og for alle nodene $v \in V \setminus \{s, t\}$ er strømmen bevart. Da er strømmen ut fra s lik strømmen inn til t.*

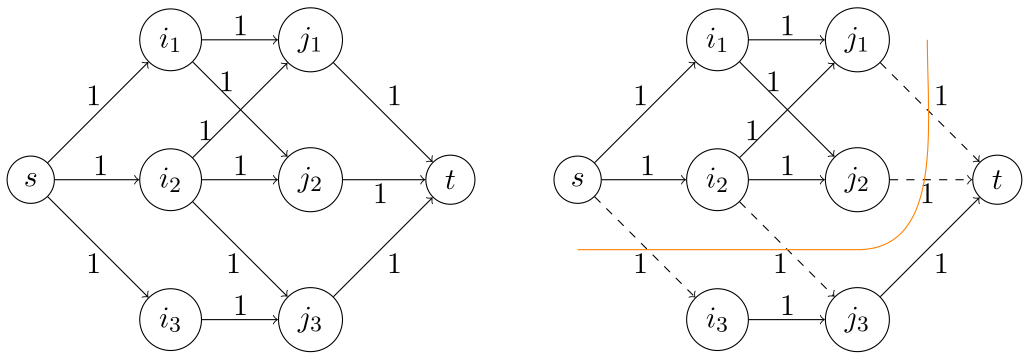
$$\sum_{e\in\delta^+(s)} x_{s,v} = \sum_{e\in\delta^-(t)} x_{v,t}, \textit{ for alle } v \in V \setminus \{s, t\}$$

Definisjon 4.2.9 (St-kutt). Gitt en graf G og nodene s og t. Hvis vi gjør et kutt og deler G inn i to mengder A og B, der $s \in A$ og $t \in B$ har vi et *st-kutt*.

4.3 Maks strøm min kutt

Maksimum strøm og minimum kutt-teoremet (maks strøm min kutt) er et teorem som baserer seg på to problemer i grafteori, maksimum strøm i en graf og minimum kutt i en graf. Teoremet sier at verdien til den største strømmen som kan gå i en st-graf fra s til t, er lik kapasiteten til det minste st-kuttet man kan gjøre i den samme grafen.

Definisjon 4.3.1 (Maks strøm-problemet). Gitt en rettet st-graf $G' = (V', E')$ og en ikke-negativ kapasitetsfunksjon c, finn en st-strøm x med størst mulig verdi $v(x)$.



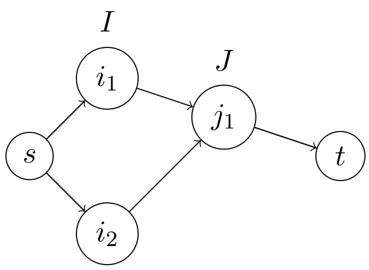
(a) St-strøm med kapasiteter på 1 til hver kant

(b) Et st-kutt med kapasitet på 4

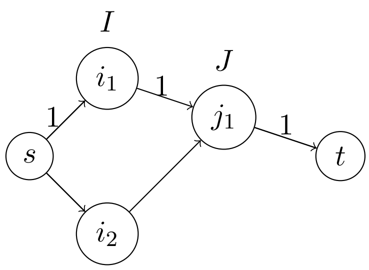
Definisjon 4.3.2 (Min kutt-problemet). Gitt en rettet st-graf $G' = (V', E')$ og en ikke-negativ kapasitetsfunksjon c , finn et st-kutt $\delta^+(S)$ med lavest mulig kapasitet, der $S \subset V$.

Teorem 4.3.3 (Max strøm min kutt). *For enhver rettet st-graf med kapasitetsfunksjon c på kantene, vil verdien av maksimum strøm være lik verdien til kapasiteten til minimum kutt.*

Eksempel 4.3.4. Gitt en st-graf $G' = (V', E')$ skal vi finne maks strøm fra s til t. Hver kant har kapasitet på $c = 1$.



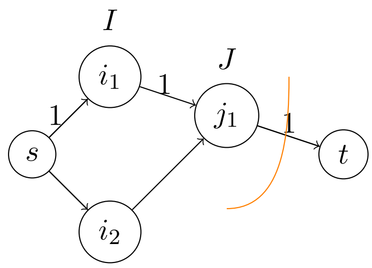
Fra s til t vil maks strøm gå langs én vei, og det sendes maksimalt med strøm ut fra s. Det går 1 ut av én av kantene fra s, og strømmen vil uansett gå gjennom j_1 før den går til t. Her er tallet 1 på figuren strømmen som går langs hver kant.



Maks strøm blir dermed 1 fra s til t.

$$\sum_{e \in \delta^+(s)} x(e) = 1$$

For å illustrere et kutt kan vi tegne en orange strek på grafen for å se hvor streken deler nodene i to mengder. Én mengde som inkluderer s og én mengde som inkluderer t. Målet er å finne flaskehalsen til strømmen, og den ser vi må være kanten fra den øverste noden i J til t.



Min kutt blir 1 ettersom det den totale kapasiteten av det minste st-kuttet er 1.

$$\sum c(e) = 1$$

.

Maks strøm = min kutt = 1. □

4.4 Königs teorem

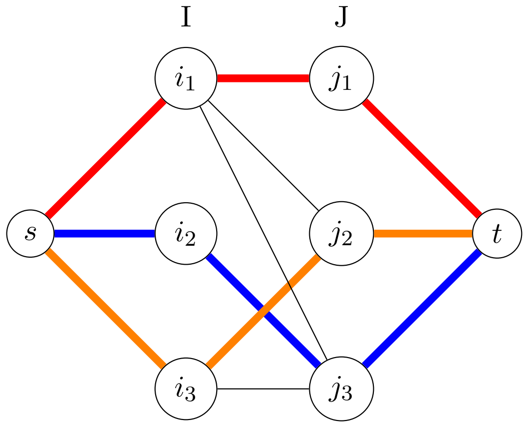
Teorem 4.4.1 (Fakta). *La $G = (V, E)$ være en bipartitt graf og la $G' = (V', E')$ være nettverksstrømmen gitt av G slik at $c(e) = 1$ for alle $e \in E'$ Da er verdien til maks strøm lik størrelsen til den maksimume matchingen i E.*

Teorem 4.4.2 (Königs teorem). *Gitt en bipartitt graf $G = (V, E)$, vil maksimum matching i G, $\rho(G)$, være lik minimum nodedekning av G, $\tau(G)$. Königs teorem sier altså at i en bipartitt graf G vil størrelsen $\rho(G)$ være lik størrelsen $\tau(G)$, $\rho(G) = \tau(G)$.*

Bevis. Vi utvider G til en rettet st-strøm, G', og vil vise at i dette nye nettverket får vi at maks strøm er lik maksimum matching og minimum kutt blir det samme som minimum nodedekning. Vi kan deretter bruke maks strøm min kutt til å konkludere at maksimal matching må være lik minimum nodedekning. For å illustrere beviset, vil vi bruke en eksempelfigur underveis.

$G = (V, E)$ og $G' = (V', E')$ der $V' = V \cup \{s, t\}$, der s er kilden og t er sluket til den rettede grafen G'. E' består av E og alle de nye kantene fra s til nodene i I og kantene fra nodene i J til t. Vi tilegner hver kant kapasiteter, og kantene som allerede var i E får kapasitet på uendelig. De nye kantene fra s og de nye kantene inn til t, får kapasitet på 1. Da er premisset satt.

Gitt en matching med kardinalitet r, er det enkelt å finne en strøm med størrelse r. Vi sender en strøm på 1 langs hver kant ut fra s og fortsetter langs en kant fra I til J som er en del av matchingen. Da vil alle nodene i J ha inngående strøm på 1, og det vil gå 1 strøm langs hver kant inn til t. Dermed vil maks strøm i G' korrespondere til matchingen i G sin maksimale kardinalitet.



Figur 4.3: Maksimal strøm = 3

La W være en nodedekning i G med k noder, og la $W(I) = W \cap I$ og $W(J) = W \cap J$ være delmengder av W. Videre, la $I' = I - W(I)$ og $Y' = Y - W(Y)$. I' og J' er nodene som ikke er med i nodedekningen W. La $S = s \cup W(J) \cup Y'$ og $T = t \cup W(I) \cup I$. Kuttet (S, T) i G' har kardinalitet k, samme kardinalitet som W. Dermed vil en nodedekning definere et kutt med lik kardinalitet.

Vi har et kutt (S, T) i G' med verdi k. Kantene fra I til J har uendelig kapasitet, så hver kant fra S til T må enten gå fra s til I eller fra J til t, og de kantene har kapasitet 1. Kuttet har verdi k, så (S, T) er lik k.

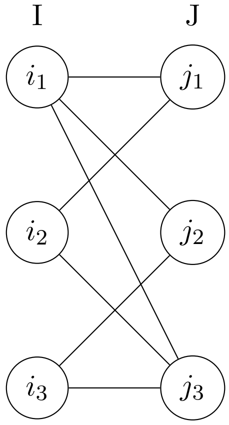
La W være unionen av kantene fra s til I slik at kantene er i (S, T), og la kantene fra J til t også være i (S, T). Da ha r W k kanter. For enhver nodedekning i G er det er korresponderende kutt i G'. Dermed vil et minimum kutt korrespondere til en minimum nodedekning i G.

Til nå har vi sett at maksimal strøm har lik verdi som en kardinalitet til en maksimal matching, og verdien til et minimum kutt tilsvarer verdien til den minste nodedekningen. Det følger fra maks-strøm min-kutt-teoremet at minimum kutt har lik verdi som maksimal strøm. Da kan vi konkludere at kardinaliteten til en maksimal matching vil ha lik verdi som den minste nodedekningen.

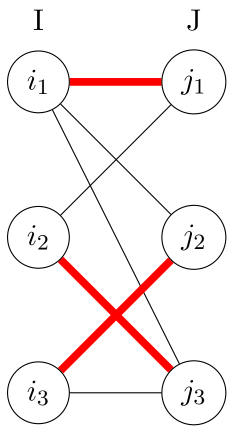
$\blacksquare$

Königs teorem sier at i en graf vil verdien av den maksimume matchingen vil være lik verdien til den minste nodedekningen.

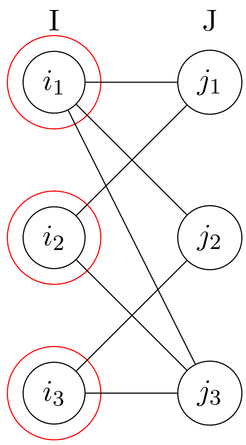
Eksempel 4.4.3. Gitt en bipartitt graf A, hvor vi for enkelhets skyld kaller nodene i I og J for i_n og j_n for $n = 1, 2, 3$.



For å finne $\rho(A)$ skal vi finne maksimum matching, og fra et tidligere eksempel vet vi at vi i dette tilfelle kan maksium matching se slik ut Her blir $\rho(A) = 3$. Vi skal nå finne den minste nodedekningen som gjør at hver kant er insident med en av nodene i nodedekningen. Vi må velge minst tre noder for å at alle kantene skal med. $\tau(A) = 3$. Dermed får vi $\rho(A) = \tau(A) = 3$, som vi kan se illustrert i figur 4.4.



(a) Maksimum matching



(b) Minste nodedekning

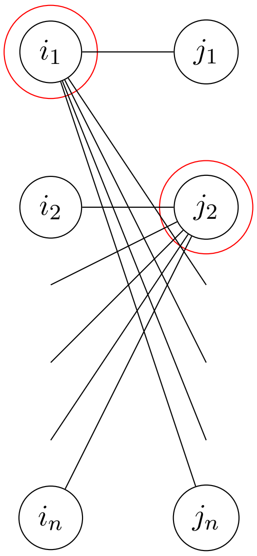
Figur 4.4: Eksempel på Königs teorem

□

Det skal sies at i en bipartitt graf er det alltid tilstrekkelig å velge alle nodene i I for å få en nodedekning, for da er alle kantene med i dekningen. Det betyr likevel ikke nødvendigvis at vi har den minste nodedekningen.

Eksempel 4.4.4. Gitt en biparitt graf $B = (V, E)$ med $2n$ noder og nodemengdene $I \in E$ og $J \in E$ med kantene $\{i_1 j_k\}$ for $k = \{1, 3, 4 \ldots, n\}$ og $\{j_2 i_m\}$ for $m = \{2, 3, 4 \ldots, n\}$

der $i \in I$ og $j \in J$.



Det holder å inkludere i_1 og j_2 for å få en nodedekning med alle kantene. $\square$

4.4.1 Matriseform

Hvordan ser Königs teorem ut på matriseform? Gitt en graf $G = (V, E)$ har vi en redusert nabomatrise A med kun 0 og 1 som elementer i matrisen. $\rho(A)$ er maksimum antall ikke-nuller i A, der ingen par er på samme rad eller kolonne. $\tau(A)$ er minimum antall linjer i A som dekker alle ikke-nuller. Der en *linje* er enten en hel rad eller kolonne. Da er $\rho(A) = \tau(A)$.

Eksempel 4.4.5.

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

Ved å velge en ikke-null i A, har vi valgt et element som ikke deler rad eller kolonne med noen andre ikke-nuller, da vi bare har ett element i mengden vår. Vi ser på to elementer, $a_{13} = 1$ og $a_{32} = 1$, og vi kan enkelt se at dette er maks antall ikke-nuller vi kan velge som ikke deler rad og kolonne. Dermed er $\rho(A) = 2$.

Hvis vi velger alle tre kolonnene eller alle tre radene, vil vi få dekket alle ikke-nuller i A. Vi er imidlertid ute etter min. antall linjer, så det vil være tilstrekkelig i dette tilfelle å velge kolonne 1 og kolonne 2 for å dekke alle ikke-nuller i A. Dermed er $\tau(A) = 2$.

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

$\square$

Eksempel 4.4.6. Vi ser på enda et eksempel med matrisen

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

Vi kan velge oss flere utvalg av ikke-nuller som ikke er på samme rad eller kolonne, men det maksimale antallet blir fortsatt 2. Dermed er $\rho(B) = 2$. Minimum antall linjer er 2 hvor vi får dekket alle ikke-nuller i B. Dermed er $\tau(B) = 2$.

$$B = \begin{pmatrix} 1 & 1 & \boxed{1} \\ 1 & 0 & 0 \\ \boxed{1} & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

$\square$

I begge de foregående eksemplene har vi at $\rho(A) = \tau(A)$ og $\rho(B) = \tau(B)$.

Kapittel 5

Dobbelt stokastiske matriser

I dette kapitlet skal vi presentere det mest sentrale for oppgaven: Dobbelt stokastiske matriser, heretter forkortet d.s. For en d.s. matrise er en $n \times n$ matrise er hver radsum og kolonnesum 1. Teorien og definisjonene vil være hentet fra boken *Matrix Analysis* av Rajendra Bhatia [Bha97].

5.1 Teori

Definisjon 5.1.1 (Radsum). Gitt en $m \times n$ matrise A kalles summen av alle elementene i rad i i A, r_i, for *radsum* i A.

$$A_{mn} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

Radsummene til A er, for en gitt rad, i,

$$r_i = \sum_{j=1}^{n} a_{ij}$$

Definisjon 5.1.2 (Kolonnesum). Gitt en $m \times n$ matrise A kalles summen av alle elementene i kolonne j i A, k_j, for *kolonnesum* i A.

$$A_{m,n} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

Kolonnesummene til A er, for en gitt kolonne, j,

$$k_j = \sum_{i=1}^{m} a_{ij}$$

Definisjon 5.1.3 (Dobbelt stokastisk matrise). En *dobbelt stokastisk matrise* A er en $n \times n$ matrise hvor alle radsummene og kolonnesummene i A er 1, og elementene i matrisen er større eller lik 0.

- $\sum_{j=1}^{n} a_{ij} = 1, \text{ for } i = 1, ..., n$
- $\sum_{i=1}^{m} a_{ij} = 1, \text{ for } j = 1, ..., n$
- $a_{i,j} \geq 0, \text{ for } i, j = 1, ..., n$

Eksempel 5.1.4. Den mest trivielle d.s. matrisen er 1×1 matrisen

$$A = \begin{pmatrix} 1 \end{pmatrix}$$

som oppfyller kravene til at radsummen og kolonnesummen er 1. □

Eksempel 5.1.5. Gitt en 2×2 matrise

$$B = \begin{pmatrix} 0.2 & 0.8 \\ 0.8 & 0.2 \end{pmatrix}$$

som oppfyller kravene til at radsummen og kolonnesummen er 1, er B d.s. Elementene i B er enten 0.2 eller 0.8, og generelt vil en 2×2 matrise være svært begrenset med tanke på hvilke verdier elementene kan ha. Vi skal senere se på hvordan den generelle formen til en 2×2 ser ut. □

Eksempel 5.1.6. Vi har 3×3-matrisen B

$$C = \begin{pmatrix} 0.2 & 0.3 & 0.5 \\ 0.6 & 0.1 & 0.3 \\ 0.2 & 0.6 & 0.2 \end{pmatrix}$$

Her ser vi radsummene r_1, r_2, r_3 og kolonnesummene k_1, k_2, k_3 bli

$$\begin{aligned} r_1 &= 0.2 + 0.3 + 0.5 = 1 \\ r_2 &= 0.6 + 0.1 + 0.3 = 1 \\ r_3 &= 0.2 + 0.6 + 0.2 = 1 \\ k_1 &= 0.2 + 0.6 + 0.2 = 1 \\ k_2 &= 0.3 + 0.1 + 0.6 = 1 \\ k_3 &= 0.5 + 0.3 + 0.2 = 1 \end{aligned}$$

C er en 3×3-matrise hvor linjesummene er 1, altså er C d.s. □

Eksempel 5.1.7. En 2×2-matrise

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

har radsummene r_1, r_2 og kolonnesummene k_1, k_2

$$r_1 = 1 + 1 = 2$$
$$r_2 = 0 + 0 = 0$$
$$k_1 = 1 + 0 = 1$$
$$k_2 = 1 + 0 = 1$$

Kolonnesummene er 1, men radsummene er 2 og 0, så D er ikke d.s. □

Spesialtilfeller av d.s. matriser er *permutasjonsmatriser*. Permutasjonsmatrisene er sentrale da alle d.s. er konvekse kombinasjoner av permutasjonsmatrisene. Dette skal vi ser mer inngående på i neste kapittel.

Definisjon 5.1.8. En *permutasjonsmatrise* er en $n \times n$ matrise som oppfyller disse kravene

- elementene er 0 eller 1
- det er bare én 1 per rad
- det er bare én 1 per kolonne

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

Figur 5.1: Eksempler på permutasjonsmatriser

Ettersom permutasjonsmatriser kun har én 1 per rad og kolonne, og resten av elementene er 0, vil rad- og kolonnesumeme være 1. Permutasjonsmatrisene er dermed d.s.

Definisjon 5.1.9 (Diagonal). Gitt en $n \times n$ matrise $A = (a_{ij})$, og $\{\sigma_1, \sigma_2, \dots, \sigma_n\}$ er en permutasjon på n elementer, så er mengden $\{a_{1\sigma_1}, a_{2\sigma_2}, \dots, a_{n\sigma_n}\}$ en *diagonal* til A. Hver diagonal inneholder nøyaktig ett element fra hver rad og kolonne.
En *positiv diagonal* er en diagonal i en matrise hvor alle elementene er større enn null.

Eksempel 5.1.10. Enerne i A ligger på samme diagonal, og enerne i B ligger på samme diagonal.

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

□

5.2 Egne bevis

I dette delkapitlet skal vi se på noen egenskaper ved d.s. matriser hvor jeg selv har utledet bevisene.

Definisjon 5.2.1. La A være en $n \times n$ matrise. Matrisen B kalles den inverse til matrisen A hvis

$$AB = BA = I$$

der I er identitetsmatrisen. Hvis en slik B finnes, sier vi at A er *invertibel.*

Definisjon 5.2.2. Gitt en d.s. matrise A, er

$$Ae = e$$
$$Ae^T = e^T$$

hvor e er en $n \times 1$-matrise

$$e = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

Dette uttrykker at rad- og kolonnesummene er 1.

Lemma 5.2.3. *Gitt at en $n \times n$ matrise A har en invers B, og alle radsummene til A er 1, vil alle radsummene til B også være* 1.

Bevis. A er en $n \times n$ matrise

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{pmatrix}$$

For å vise at radsummene til B er 1, holder det å vise $Be = e$.
Vi har en $d.s.$-matrise A og den inverse B.

$$BA = I$$
$$BAe = Ie$$

Vi vet at radsummen til A er 1, så $Ae = e$ og at $Ie = e$ da I er identiteten.

$$Be = e$$

$Be = e$, noe som betyr at radsummene til B er 1.
Det samme kan vises for kolonnesummene.

$$BA = I$$
$$BAe^T = Ie^T$$

$Be^T = e^T$, dermed er kolonnesummene til B 1.

■

Eksempel 5.2.4. Gitt en d.s. matrise

$$A = \begin{pmatrix} 0.2 & 0.3 & 0.5 \\ 0.3 & 0.6 & 0.1 \\ 0.5 & 0.1 & 0.4 \end{pmatrix}$$

kan vi finne den inverse matrisen til A, gitt at determinanten ikke er null.

$$B = \begin{pmatrix} -23/11 & 7/11 & 27/11 \\ 7/11 & 17/11 & -13/11 \\ 27/11 & -13/11 & -3/11 \end{pmatrix}$$

Alle rad- og kolonnesummene er 1, men ikke alle elementene er ikke-negative. Den inverse matrisen B er dermed ikke d.s., men har rad- og kolonnesummer lik 1. □

Teorem 5.2.5. *Gitt to d.s. matriser, P_1 og P_2, vil produktet $P = P_1 \cdot P_2$ også være d.s.*

Bevis. En d.s. matrise A oppfyller likningen $Ae = e$

$$P = P_1 P_2$$
$$Pe = P_1 P_2 e$$
$$Pe = P_1 e$$
$$Pe = e$$

Det samme gjelder for kolonnesummene. A må oppfylle $Ae^T = e^T$

$$P = P_1 P_2$$
$$Pe^T = P_1 P_2 e^T$$
$$Pe^T = P_1 e^T$$
$$Pe^T = e^T$$

■

Eksempel 5.2.6. Vi kan se på et tilfelle med to d.s. 3×3 matriser

$$P_1 = \begin{pmatrix} 0.2 & 0.1 & 0.7 \\ 0.7 & 0.3 & 0 \\ 0.1 & 0.6 & 0.3 \end{pmatrix}, P_2 = \begin{pmatrix} 0.5 & 0.5 & 0 \\ 0.1 & 0.2 & 0.7 \\ 0.4 & 0.3 & 0.3 \end{pmatrix}$$

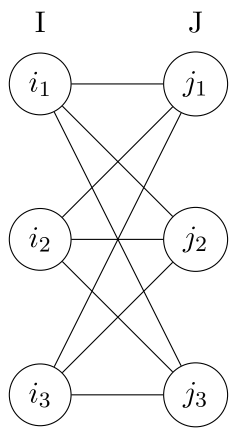
$$P = P_1 \cdot P_2 = \begin{pmatrix} 0.39 & 0.33 & 0.28 \\ 0.38 & 0.41 & 0.21 \\ 0.23 & 0.26 & 0.51 \end{pmatrix}$$

Rad- og kolonnesummene til P er 1, og hvert element er ikke-negativt. Dermed er P d.s.

$\square$

5.3 Fra nettverkstrøm til d.s. matrise

Fra kapittel 3 så vi hvordan vi kunne representere en graf ved en matrise. I tillegg så vi på hvordan vi kunne konstruere en redusert nabomatrise for å representere en bipartitt graf. I kapittel 4 så vi på hvordan teori om nettverkstrøm kan bevise Königs teorem. Videre skal vi ta innhold fra begge kapitlene og se på en sammenheng mellom nettverkstrøm og hvordan vi kan konstruere en redusert nabomatrise som vil være d.s.
For enkelhets skyld skal vi se på en bipartitt graf $G = (V, E)$ hvor $|V| = 6$, der $|I| = 3$ og $|J| = 3$. Hver node i I deler kant med hver node i J.

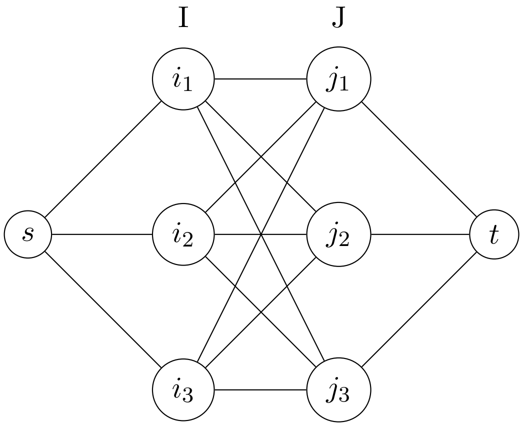


Figur 5.2: $G = (V, E)$

Denne grafen har den reduserte nabomatrisen

$$B = \begin{array}{c c} & \begin{array}{ccc} j_1 & j_2 & j_3 \end{array} \\ \begin{array}{c} i_1 \\ i_2 \\ i_3 \end{array} & \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \end{array}$$

Med utgangspunkt i $G = (V, E)$ introduserer vi den rettede st-grafen $G' = (V', E')$ hvor $V' = I' \cup J'$ og hver node i I deler kant med hver node i J.



Figur 5.3: $G' = (V', E')$

Fra tidligere vet vi at all strøm ut fra kilden s vil være lik all strøm inn til sluket t, og all strøm inn i hver node i I og J vil være bevart. Langs kantene (s, i_k) er strømmen $x_k = 1$, der $k = 1, 2, 3$. Det betyr at for alle $i \in I$ er strømmen inn 1, og ettersom strømmen er bevart i hver node, vil strømmen ut også være 1. For nodene $j \in J$ er strømmen også bevart. Den totale mengden strøm inn i t er lik strømmen ut av s, altså 3. Derfor må strømmen ut av j_1, j_2, j_3 være totalt 3.

Grafen G kan sies å være innbakt i G'. Tanken er å konstruere den reduserte nabomatrisen til G med utgangspunkt i strømmen som går i G'. Tidligere har vi skrevet en ener i posisjonen i matrisen hvor to noder deler en kant, men nå skal vi heller skalere hvert element med hvor mye strøm som faktisk går i den spesifikke kanten. Går det 0.2 strøm mellom i_1 og j_1 vil posisjon $b_{i_1 j_1} = 0.2$ i den reduserte nabomatrisen B. Dette gjør vi for all strøm i alle kanter. Ettersom målet vårt er å få B til å bli d.s. vet vi at ingen av kantene kan ha over 1 i strøm. I tillegg kan ingen av nodene ha mer enn 1 i strøm inn, for da vil vi få over 1 på raden og matrisen blir ikke d.s. Faktisk må hver node j_1, j_2, j_3 ha inngående strøm på nøyaktig 1. Reglene er satt:

- kapasitetskrav på 1.
- ingen strøm inn til hver node over 1
- strøm inn til j_1, j_2, j_3 er 1.

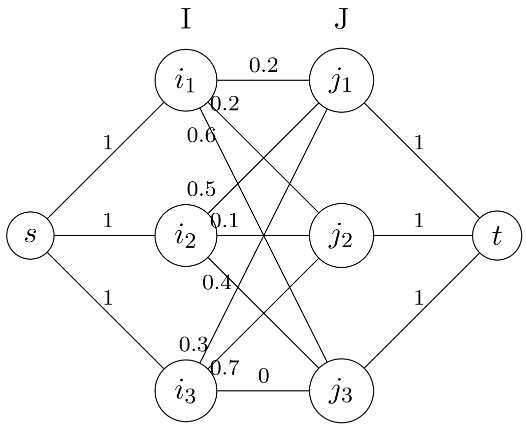
Ser vi på node i_1 har den strøm inn på 1 fra kanten (s, i_1) og strøm ut på 1 fordelt på tre kanter (i_1, j_k) der $k = 1, 2, 3$. Dermed er $x_{i_1 j_1} + x_{i_1 j_2} + x_{i_1 j_3} = 1$. Ser vi på strømmen inn i j_1, vet vi at den må være lik 1. Da får vi at $x_{i_1 j_1} + x_{i_2 j_1} + x_{i_3 j_1} = 1$. Dette kan vi gjøre for alle nodene som er felles mellom G og G', altså nodene i I og J. Nå skal vi konstruere den reduserte nabomatrisen B. Da vil hvert element i matrisen være strømmen mellom de to nodene som hver rad og kolonne er assosiert med.

$$B = \begin{array}{c} \\ i_1 \\ i_2 \\ i_3 \end{array} \begin{array}{c} \begin{array}{ccc} j_1 & j_2 & j_3 \end{array} \\ \begin{bmatrix} x_{i_1j_1} & x_{i_1j_2} & x_{i_1j_3} \\ x_{i_2j_1} & x_{i_2j_2} & x_{i_2j_3} \\ x_{i_3j_1} & x_{i_3j_2} & x_{i_3j_3} \end{bmatrix} \end{array}$$

Hver radsum er total strøm ut fra hver node $i \in I$, og vi vet at den er 1. Derfor er alle radsummene lik 1. Hver kolonnesum er total strøm inn i hver node $j \in J$, som vi har satt til å være 1. Derfor er kolonnesummene også lik 1. Dermed er den reduserte nabomatrisen B d.s.

Vi har nå sett på en sammenheng mellom nettverkstrøm og d.s. matriser. Det krevde noen begrensninger for å konstruere den reduserte nabomatrisen, men sammenhengen mellom nettverkstrøm og d.s. matriser er tydelig. For å konstruere en 3×3 d.s. matrise trengte vi en graf G med 3 noder i I og J. Skal vi konstruere en 4×4 d.s. matrise trenger vi en graf med 4 noder i I og J. Skal vi konstruere en $n \times n$ matrise trenger vi en graf med n noder i I og J.

Eksempel 5.3.1. Gitt en rettet st-strøm $G' = (V', E')$ skal vi konstruere en redusert nabomatrise hvor hver posisjon i nabomatrisen er strømmen langs kantene mellom I og J. Verdiene langs kantene i figuren indikerer strømmen langs kantene.



Den reduserte nabomatrisen til G blir

$$B = \begin{array}{c} \\ i_1 \\ i_2 \\ i_3 \end{array} \begin{array}{c} \begin{array}{ccc} j_1 & j_2 & j_3 \end{array} \\ \begin{bmatrix} 0.2 & 0.2 & 0.6 \\ 0.5 & 0.1 & 0.4 \\ 0.3 & 0.7 & 0 \end{bmatrix} \end{array}$$

□

Kapittel 6

Birkhoff-polytopen

Til nå har vi sett på grafteori og nettverkstrøm. Videre skal vi bevege oss inn i matriseverden. Bakgrunnen for matrisene vi skal se på har rot i grafteori og nettverksstrøm. Dette kapitlet er hoveddelen av oppgaven, nemlig mengden av d.s. matriser, Birkhoff-polytopen. Vi vil presentere den før vi skal gå inn i en mer utforskende fase om strukturen til d.s. matriser, altså elementene i Birkhoff-polytopen.

6.1 Birkhoff polytopen

I 1946 [MOA11] presenterte George David Birkhoff sitt teoreom om d.s. matriser. Vi kaller mengden av alle d.s. matriser for Birkhoff-polytopen. I dette kapitlet skal vi se på elementene i denne mengden og strukturen til de d.s. matrisen kan se ut. D.s. matriser har begrensninger når det gjelder hvilke verdier hvert element i matrisene kan ha, og ved å innføre enda flere begrensninger enn de som gjelder for definisjonen av d.s. matriser kan vi se at det er distinkte mønstre matrisene følger. Vi skal innlede kapitlet med å presentere Birkhoffs teorem, og så utforske hvordan matriser i Birkhoff-polytopen kan se ut.

Definisjon 6.1.1 (Birkoff-polytopen). Birkoff-polytopen, Ω_n, er mengden av alle $n \times n$ d.s. matriser.

$$\Omega_n = \{A \in \mathbb{R}_{\geq 0}^{n \times n} | \text{for alle } j = \{1, ..., n\}, \sum_i a_{i,j} = \sum_i a_{j,i} = 1\}$$

Vi bruker Ω_n for Birkhoff-polytopen hvor n er dimensjonen til en $n \times n$ matrise.

Teorem 6.1.2 (Birkhoffs teorem). *Birkhoff-polytopen, alstå mengden av alle $n \times n$ d.s. matriser, er en konveks mengde der ekstrempunktene er permutasjonsmatrisene.*

Birkhoffs teorem sier at Ω_n er den konvekse innhyllingen av mengden $n \times n$ permutasjonsmatriser og at permutasjonsmatrisene er hjørnene til Ω_n. Beviset er utledet i *Inequalities: Theory of Majorization and Its Applications* [MOA11].

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} 0.2 & 0.4 & 0.4 \\ 0.6 & 0.1 & 0.3 \\ 0.2 & 0.5 & 0.3 \end{pmatrix} \qquad \frac{1}{3}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

Figur 6.1: Eksempler på matriser i Birkoff-polytopen Ω_3

Teorem 6.1.3. *En ikke-negativ reell $n \times n$-matrise A er dobbelt stokastisk hvis, og bare hvis, det finnes permutasjonsmatriser $P_1, P_2, ..., P_t$ og positive reelle tall $c_1, c_2, ..., c_t$ slik at*

$$A = c_1P_1 + c_2P_2 + ... + c_tP_t \tag{6.1}$$

og

$$c_1 + c_2 + ... + c_t = 1 \tag{6.2}$$

Beviset for dette relativt simple teoremet bygger på Königs teorem som igjen ble bevist av maks strøm min kutt. Beviset er utledet i boken til Brualdi og Ryser [BR91]. Ettersom det er relativt kort vil vi presentere det her også.

Bevis. Hvis en matrise A oppfyller kravene til 6.1 og 6.2 er

$$AI = IA = I$$

og A er d.s.
Vi skal nå bevise motsatt implikasjon. En d.s. matrise A har n positive elementer hvor ingen av elementene er på samme rad eller kolonne (linje). Hvis dette ikke var tilfelle, følger det av Königs teorem 4.4.2 at vi kan få en dekning av alle positive elementer i A med e kanter og f kolonner, hvor $e + f < n$. Ettersom alle linjesummene til A er lik 1, følger det at $n \leq e + f < n$, og vi har en motstridelse. La P_1 være en permutasjonsmatrise av grad n med enere i samme posisjonene som de n positive elementene i A. La c_1 være den minste av disse n positive elementene. Da er $A - c_1P_1$ en skalert d.s. matrise, og minst én 0 mer eksisterer i $A - c_1P_1$ enn i A. Fortsetter vi denne logikken og bryter ned A i mindre og mindre biter, vil vi ende opp med den ønskede dekomposisjonen i 6.1. Multipliserer vi 6.1 med identitetsmatrisen I og dette impliserer 6.2. ■

Dette teoremet er sentralt for d.s. matriser. Det konstaterer at enhver d.s. matrise kan skrives som en konveks kombinasjon av permutasjonsmatriser. Vi vet fra tidligere at elementer i konvekse mengder kan skrives som en konveks kombinasjon av ekstrempunktene, og ettersom alle elementene i Ω_n kan skrives som en konveks kombinasjon av permutasjonsmatrisene, er permutasjonsmatrisene ekstrempunktene i Ω_n.

Eksempel 6.1.4. Vi har en 3×3 d.s. matrise

$$B = \begin{pmatrix} 0.5 & 0.5 & 0 \\ 0.5 & 0.2 & 0.3 \\ 0 & 0.3 & 0.7 \end{pmatrix}$$

Vi ser at alle rad- og kolonnesummene blir 1 og kan vise at B oppfyller likning 6.1 og 6.2.

$$B = \begin{pmatrix} 0.5 & 0.5 & 0 \\ 0.5 & 0.2 & 0.3 \\ 0 & 0.3 & 0.7 \end{pmatrix}$$

$$B = 0.2 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + 0.3 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} + 0.5 \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

som oppfyller 6.1 og $c_1 = 0.2$, $c_2 = 0.3$, $c_3 = 0.5$ oppfyller 6.2

$$c_1 + c_2 + c_3 = 0.2 + 0.3 + 0.5 = 1$$

□

Teorem 6.1.5. *Birkhoff-polytopen er konveks.*

Bevis. Vi skal vise at hvis $A, B \in \Omega_n$ så er også $M = (1-\lambda)A + \lambda B \in \Omega_n$ for $\lambda \in [0,1]$
Vi ser på to generelle $n \times n$ matriser $A, B \in \Omega_n$, og på $M = (1-\lambda)A + \lambda B$.

$$\begin{aligned} M &= (1-\lambda)A + \lambda B \\ &= (1-\lambda) \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} + \lambda \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nn} \end{pmatrix} \end{aligned}$$

Vi har radsummene og kolonnesummene til M blir

$$r_i = (1-\lambda) \sum_j a_{ij} + \lambda \sum_j b_{ij}$$

$$k_i = (1-\lambda) \sum_i a_{ij} + \lambda \sum_i b_{ij}$$

der

$$\sum_j a_{ij} = 1, \sum_i a_{ij} = 1$$

$$\sum_j b_{ij} = 1, \sum_i b_{ij} = 1$$

Vi får da at radsummen til M blir

$$r_i = (1-\lambda) + \lambda = 1$$

$$k_i = (1-\lambda) + \lambda = 1$$

Dermed er $M \in \Omega_n$, og Birkhoff-polytopen er konveks. ∎

Eksempel 6.1.6. Vi har matrisene $A = \begin{pmatrix} 0.2 & 0.3 & 0.5 \\ 0.1 & 0.4 & 0.5 \\ 0.7 & 0.3 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0.4 & 0.2 & 0.4 \\ 0.1 & 0.6 & 0.3 \\ 0.5 & 0.2 & 0.3 \end{pmatrix}$ og $\lambda = 0.3$

Skal vise at $M = (1-\lambda)A + \lambda B \in \Omega_n$

$$\begin{aligned} M = (1-0.3) &\begin{pmatrix} 0.2 & 0.3 & 0.5 \\ 0.1 & 0.4 & 0.5 \\ 0.7 & 0.3 & 0 \end{pmatrix} + 0.3 \begin{pmatrix} 0.4 & 0.2 & 0.4 \\ 0.1 & 0.6 & 0.3 \\ 0.5 & 0.2 & 0.3 \end{pmatrix} \\ &= \begin{pmatrix} 0.14 & 0.21 & 0.35 \\ 0.07 & 0.28 & 0.35 \\ 0.49 & 0.21 & 0 \end{pmatrix} + \begin{pmatrix} 0.12 & 0.06 & 0.12 \\ 0.03 & 0.18 & 0.09 \\ 0.15 & 0.06 & 0.09 \end{pmatrix} \\ &= \begin{pmatrix} 0.26 & 0.27 & 0.47 \\ 0.10 & 0.46 & 0.44 \\ 0.64 & 0.27 & 0.09 \end{pmatrix} \end{aligned}$$

Vi ser at rad- og kolonnesummene blir 1 og $m_{ij} \geq 0$, så M er en d.s. matrise og med i Ω_3. □

6.2 Strukturen til 2x2 og 3x3 d.s. matriser

Elementer i Ω_n er som vi skjønner relativt begrenset da vi har kravene om rad- og kolonnesummer, i tillegg til at hvert element ikke kan være negativt. Vi skal utforske om det er noen mønstre for strukturen til d.s. matriser. Vi skal begynne med matriser i Ω_1, Ω_2 og Ω_3 før vi etterhvert skal trekke konklusjoner for alle matriser i Ω_n. Vi har allerede sett på den mest trivielle d.s. matrisen, en 1×1 matrise som kun kan inneholde $a_{11} = 1$.

$$A = \begin{pmatrix} 1 \end{pmatrix}$$

For d.s. 2×2 matriser vil alle fire elementene direkte påvirke hverandres verdier. Det finnes uendelig mange 2×2 matriser da det er uendelig mange kombinasjoner av to tall som summerer til 1, men verdiene til elementene er likevel begrenset. En d.s. 2×2 vil ha formen

$$\begin{pmatrix} \alpha & 1-\alpha \\ 1-\alpha & \alpha \end{pmatrix}$$

der $\alpha \leq 1$. Når vi vet hva a_{11} er, α må elementene a_{12} og a_{21} være $1-\alpha$. Elementet a_{22} vil også være α for at rad 2 og kolonne 2 skal summere til 1.

En 3×3-matrise vil ikke være like begrenset. La oss se på tilfellet hvor vi har en 3×3 $d.s.$ matrise, og setter $a_{11} = \alpha_1$ og utvider med α_2 osv.

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

Ettersom vi skal ha radsummene og kolonnesummene lik 1, får vi denne matrisen

$$\begin{pmatrix} \alpha_1 & \alpha_2 & 1-\alpha_1-\alpha_2 \\ \alpha_3 & \alpha_4 & 1-\alpha_3-\alpha_4 \\ 1-\alpha_1-\alpha_3 & 1-\alpha_2-\alpha_4 & \alpha_1+\alpha_2+\alpha_3+\alpha_4-1 \end{pmatrix}$$

Hvor $\alpha_1, \alpha_2, \alpha_3, \alpha_4 \leq 1$ i tillegg til

$$\begin{aligned} \alpha_1 + \alpha_3 &\leq 1 \\ \alpha_2 + \alpha_4 &\leq 1 \\ \alpha_1 + \alpha_2 &\leq 1 \\ \alpha_3 + \alpha_4 &\leq 1 \\ \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 &\geq 1 \end{aligned}$$

Her må fire elementer være gitt skal vi kunne konstruere hele matrisen, $\alpha_1, \alpha_2, \alpha_3, \alpha_4$.

6.3 Strukturen på 0-er

Ettersom 1×1- og 2×2 matriser er så begrenset vil vi videre utforske strukturen til 3×3 matriser, og så bruke resultatene til å videre trekke konklusjoner for alle $n \times n$ matriser. Vi vet at en d.s. 3×3 matrise kan inneholde kun positive elementer. Vi skal nå se på hvor få positive elementer en d.s. 3×3 matrise kan inneholde. En d.s. matrise trenger ikke kun inneholde positive elementer. Den kan også inneholde tallet 0. Det viser seg at det er ikke helt vilkårlig hvor mange nuller det er i en d.s. matrise, og ut fra hvor mange nuller vi har i matrisen vil posisjonene til nullene avhenge av det antallet.

La oss se på slike matriser i Ω_3, nærmere bestemt matriser som inneholder én eller flere nuller. Permutasjonsmatrisene er eksempler på slike

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

Vil skal se på hvor mange nuller som trengs i en 3×3 matrise, og strukturen til plasseringen av nuller, for at matrisen fortsatt skal være i Ω_3. Er matrisen fylt av nuller, kan vi enkelt se at matrisen ikke er i Ω_3

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

Det kan maks være seks nuller for at det kan være en d.s. matrise, da syv og åtte nuller heller ikke er nok ettersom at minst én av rad- eller kolonnesummene blir 0.

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

Men er det nok å si at så lenge vi har maksimalt seks nuller i en matrise og linjesummene er maksimalt 1, så vil vi ha en d.s. matrise? Vi kan lett finne et moteksempel

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0.5 & 0 \\ 0 & 0.5 & 0 \end{pmatrix}$$

Vi kan se at vi trenger elementer i alle rader og kolonner for at vi skal ha en d.s. matrise. Hvis ikke vil vi få minst én rad- eller kolonnesum som blir 0. Minstekravet er altså tre elementer, med ett element i hver rad og kolonne. Disse kravene, inkludert det generelle kravet om at hver rad- og kolonnesum er 1, er kravene som trengs for å danne en permutasjonsmatrise.

6.3.1 n+1 ikke-nuller

Kan en d.s. 3×3 matrise ha fem nuller? Altså at vi har fire ikke-nuller i matrisen. Vi kan se på et eksempel av en matrise som ikke er d.s.

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0.5 & 0 \\ 0 & 0.5 & 0.5 \end{pmatrix}$$

Denne matrisen fungerer ikke, da $r_2 = k_2 = 0.5$. Vi trenger enda et element for at alle linjesummene skal bli 1, men da inneholder matrisen *fem* ikke-nuller. Det viser seg at det ikke finnes en 3×3 matrise med kun fire positive elementer. Faktisk finnes det ingen d.s. $n \times n$ matriser med $n + 1$ positive elementer.

Korollar 6.3.1. *For en d.s.* $n \times n$ *matrise* A *kan ikke* A *ha nøyaktig* $n + 1$ *ikke-nuller.*

Bevis. Vi skal bevise dette ved å konstruere en $n \times n$ matrise, og forsøke å få den til å være d.s. Vi antar at A har nøyaktig $n + 1$ ikke-nuller. Vi vet at det må være minst ett element i hver rad og kolonne, og i én av radene og kolonnene må vi nå ha to elementer som til sammen summerer til 1 for at A fortsatt skal være d.s. Da har vi $n - 1$ rader og kolonner med nøyaktig ett element, og de elementene må være 1. Dette betyr at for disse linjene kan ikke de andre elementene på de respektive linjene være noe annet enn 0. Ettersom linjesummene skal være lik 1. Vi kan ta utgangspunkt i en permutasjonsmatrise som har n elementer.

$$\begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

Ved å erstatte en av nullene med et element $\alpha > 0$ må vi trekke i fra α fra den eneren på den raden og kolonnen som α står. Ettersom alle elementene skal være minst 0 og linjesummen skal bli lik 1 tar vi $1 - \alpha$ der 1 stod på den linja.

$$\begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & \alpha \\ 0 & 0 & 0 & \dots & 0 & 1-\alpha \end{pmatrix}$$

Radsummen til den nest nederste raden vil bli $1 + \alpha$, så vi må også trekke i fra α på den eneren også.

$$\begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1-\alpha & \alpha \\ 0 & 0 & 0 & \dots & 0 & 1-\alpha \end{pmatrix}$$

Da oppstår et nytt problem ettersom kolonnesummen til kolonne $n - 1$ nå blir $1 - \alpha$. Ettersom $\alpha > 0$ må vi erstatte en null med α et eller annet sted på kolonnen, og vi har da erstattet to nuller med to ikke-negative elementer. Hvis vi erstatter en null med α på rad n og kolonne $n - 1$ vil vi ha en d.s. matrise.

$$\begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1-\alpha & \alpha \\ 0 & 0 & 0 & \dots & \alpha & 1-\alpha \end{pmatrix}$$

Matrisen har ikke klart å oppfylle kravet om å inneholde nøyaktig $n+1$ ikke-nuller. En d.s. matrise kan altså ikke ha $n+1$ positive elementer, men den kan ha $n+2$. ■

Det kan ikke eksistere $n \times n$ matriser i Ω_n med $n+1$ ikke-nuller. Da har vi fastslått at hvis vi skal ha en d.s. matrise som ikke er en permutasjonsmatrise trenger vi minst $n+2$ ikke-nuller og at det er en ikke-null i hver rad og kolonne. Finnes det nå flere matriser som ikke er i Ω_n? La oss se på denne 3×3 matrisen. Her representerer x ikke-nuller.

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

Matrisen inneholder fem ikke-nuller som oppfyller kravet om minst $n+2$ ikke-nuller. Matrisen har en ikke-null i hver rad- og kolonne, men på rad 1 må disse to elementene summere til 1. Det vil si at $a_{11} + a_{13} = 1$ og både $a_{11}, a_{13} < 1$. På kolonne 3 står a_{13} alene og ettersom $a_{13} < 1$ vil kolonnesummen $k_3 < 1$ og vi har ikke oppfylt kravene til en d.s. matrise. Samme problemet får vi i rad 2. Hvis en ikke-null er *alene* på en rad eller kolonne, må det også være alene på sin kolonne eller rad. Når $n = 3$ og vi har $n+2 = 5$ ikke-nuller får vi et ganske begrenset utvalg av matriser. Vi får fire elementer som påvirker hverandre direkte, og ett element som er alene på en rad og kolonne.

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1-\alpha & \alpha \\ 0 & \alpha & 1-\alpha \end{pmatrix}, \quad \begin{pmatrix} 1-\alpha & 0 & \alpha \\ 0 & 1 & 0 \\ \alpha & 0 & 1-\alpha \end{pmatrix}, \quad \begin{pmatrix} \alpha & 1-\alpha & 0 \\ 0 & 0 & 1 \\ 1-\alpha & \alpha & 0 \end{pmatrix}$$

Figur 6.2: Noen eksempler

Forslag 6.3.2. *Vi har disse kravene for en d.s.* 3×3 *matrise med fem ikke-nuller:*

- *det må være minst én ikke-null i hver rad og kolonne.*
- *hvis et element står alene på en rad, må det også stå alene i kolonnen.*
- *hvis et element står alene på en kolonne, må det også stå alene i raden.*

Ut ifra disse kravene ser vi at et element alltid må stå alene på en kolonne og en rad når matrisen inneholder fem ikke-nuller, og det elementet vil være en 1. De fire siste danner en lignende struktur som for en 2×2 matrise.

$$\begin{pmatrix} \alpha & 1-\alpha \\ 1-\alpha & \alpha \end{pmatrix}$$

En 3×3 matrise med fem ikke nuller, er svært lik en 2×2 matrise, men har en kolonne og en rad ekstra med nuller og én 1.

6.3.2 For alle $n + 2$ ikke-nuller

Gitt en d.s. $n \times n$ matrise A med $n + 2$ ikke-nuller, vil A inneholde $n - 2$ enere og en form for 2×2 matrise med elementene $1 - \alpha, 1 - \alpha, \alpha, \alpha$, hvor $1 - \alpha$ og α er på samme rad eller kolonne.

$$A = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1-\alpha & \alpha \\ 0 & 0 & 0 & \dots & \alpha & 1-\alpha \end{pmatrix}$$

A vil være på denne formen, eller en permutasjon eller speiling hvor disse kravene er oppfylt:

- det må være minst én ikke-null i hver rad og kolonne.
- hvis et element står alene på en rad, må det også stå alene i kolonnen.
- hvis et element står alene på en kolonne, må det også stå alene i raden.

6.3.3 Flere ikke-nuller

Ser vi på kravene i forslag 6.3.2 kan vi også la dem gjelde for d.s. matriser med $n + 3$ ikke-nuller også.
La oss se på et par eksempler på 3×3 d.s. matriser der x er ikke-nuller.

$$\begin{pmatrix} x & 0 & x \\ 0 & x & x \\ x & x & 0 \end{pmatrix}, \begin{pmatrix} x & x & 0 \\ x & 0 & x \\ 0 & x & x \end{pmatrix}$$

Når en matrise inneholder $n + 3 = 6$ ikke-nuller, vil det si at matrisen har tre nuller. Hvis to av dem står på samme rad eller kolonne, vil matrisen ha et element som står alene på den raden og kolonnen, men ikke alene på kolonnen eller raden. Fra tidligere vet vi at dette ikke vil oppfylle kravene til en d.s. matrise. Det betyr at alle nullene må stå på forskjellige rader og kolonner, og strukturen til nullene ligger på en diagonal. For $n+4=7$ ikke-nuller har vi bare to nuller igjen å plassere. Så lenge de ikke står på samme rad eller kolonne, vil vi kunne konstruere en d.s. matrise. For $n+5=8$ ikke-nuller spiller det ingen rolle hvor den siste nullen blir plassert.

6.4 Birkhoff om strukturen til d.s. matriser

Birkhoffs teorem sier at enhver d.s. matrise kan skrives som en konveks kombinasjon av permutasjonsmatrisene

$$A = c_1P_1 + c_2P_2 + ... + c_tP_t$$

der $P_1, P_2, \ldots P_t$ er permutasjonsmatriser og $c_1, c_2, \ldots, c_t$ er konstatner som summerer til 1. Vi kan velge å se bort fra leddene hvor konstantene er 0, og vi får bare positive ledd da konstantene $c_1, c_2, \ldots, c_t$ er positive reelle tall og permutasjonsmatrisene kun har enere og nuller som elementer i matrisen. Den enkleste formen for en d.s. matrise er en permutasjonsmatrise.

Eksempel 6.4.1. Gitt en $n \times n$ d.s. matrise

$$P = \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 & 0 \\ 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1 & 0 \\ 0 & 0 & 0 & \ldots & 0 & 1 \end{pmatrix}$$

P kan skrives som en konveks kombinasjon av permutasjonsmatriser på denne måten

$$P = \lambda_1 \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 & 0 \\ 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1 & 0 \\ 0 & 0 & 0 & \ldots & 0 & 1 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 & 0 \\ 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1 & 0 \\ 0 & 0 & 0 & \ldots & 0 & 1 \end{pmatrix} + \cdots + \lambda_k \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 & 0 \\ 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1 & 0 \\ 0 & 0 & 0 & \ldots & 0 & 1 \end{pmatrix}$$

da må $\lambda_1 + \lambda_2 + \cdots + \lambda_k = 1$. Ettersom vi kun legger til den samme matrisen k ganger, trenger vi ikke å skrive opp den samme matrisen flere ganger med forskjellig λ, men heller legge sammen de forskjellige λ og kun legge til én unik matrise multiplisert med én unik λ. I dette tilfelle får vi $\lambda = 1$.

$$P = 1 \cdot \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 \end{pmatrix}$$

□

Når vi skriver en konveks kombinasjon av permutasjonsmatriser, kan vi legge sammen de matrisene som er like og heller skrive en kombinasjon av *unike* permutasjonsmatriser. Vi kan nå bruke Birkhoffs teorem til å bevise 6.3.1 som sa at en d.s. $n \times n$ matrise A ikke kan ha nøyaktig $n + 1$ ikke-nuller.

Bevis. Vi vet fra Birkhoff at en $n \times n$ d.s matrise A kan skrives som en kombinasjon av unike permutasjonsmatriser

$$A = c_1 P_1 + c_2 P_2 + ... + c_t P_t$$

der $P_1, P_2, \dots P_t$ er permutasjonsmatriser og $c_1, c_2, \dots, c_t$ er konstanter som summerer til 1. En hvilken som helst d.s. matrise må da ha minst n ikke-nuller, og vi ser fra eksempel 6.4.1 at vi ender opp med permutasjonsmatrisene.

Unike permutasjonsmatriser har enere langs en unik diagonal. Gitt at vi har en d.s. matrise som *ikke* er en permutasjonsmatrise, må vi legge sammen minst to unike permutasjonsmatriser (multiplisert med hver sin konstant).

$$A = c_1 \cdot \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 \end{pmatrix} + c_2 \cdot \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & \dots & 1 & 0 \end{pmatrix}$$

For at to permutasjonsmatriser skal være forskjellige, må minst to rader byttes om. For enkelhets skyld har vi byttet de to nederste radene, r_n og r_{n-1}, men resultatet vil gjelde for alle rader. Vi vet fra eksempel 6.4.1 at det minste antallet ikke-nuller er n, og ved å bytte om på to rader legger vi nå til to nye elementer i matrisen. Vi kan regne ut hva A må bli

$$A = \begin{pmatrix} c_1+c_2 & 0 & 0 & \dots & 0 & 0 \\ 0 & c_1+c_2 & 0 & \dots & 0 & 0 \\ 0 & 0 & c_1+c_2 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & c_1 & c_2 \\ 0 & 0 & 0 & \dots & c_2 & c_1 \end{pmatrix}$$

A må minst inneholde $n+2$ elementer, og kan dermed ikke inneholde $n+1$ elementer. ■

Ettersom en d.s. matrise er en sum av minst én permutasjonsmatrise, vil en d.s. matrise ha ikke-nuller på en positiv diagonal fra én eller flere permutasjonsmatriser. Enhver ikke-null i en d.s. matrise har et element som ligger i samme posisjon som en av enerne til permutasjonsmatrisene brukt i den konvekse kombinasjonen for å summere til den d.s. matrisen.

Eksempel 6.4.2. Vi kan se på et tidligere eksempel i lys av Birkhoff-teoremet. Vi har matrisen

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

B er d.s. hvis det finnes en konveks kombinasjon av unike permutasjonsmatriser. Med andre ord må enhver ikke-null i matrisen være på en diagonal med andre ikke-nuller. Vi kan enkelt vise at B ikke er d.s.

$$B = \begin{pmatrix} 1 & \boxed{1} & 1 \\ 1 & 0 & \boxed{0} \\ \boxed{1} & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 & \boxed{1} \\ 1 & \boxed{0} & 0 \\ \boxed{1} & 0 & 0 \end{pmatrix}$$

Disse tre er på en diagonal, og vi kan se at elementet $a_{31} = 1$ ikke er på en positiv diagonal, dermed er ikke B d.s.

Det kan imidlertid legges til at den enkleste måten å sjekke om B er d.s. er å bruke definisjonen om radsum og kolonnesum. □

Kapittel 7

Magisk Kvadrat

Etter nå å ha sett på forskjellige variasjoner av d.s. matriser med nuller, skal vi se på $n \times n$ matriser med $n \times n$ ikke-nuller. Det vil si at hvert element i en matrise er et positivt tall. I begynnelsen skal vi se på matriser hvor hvert av elementene også er et heltall. Med disse to betingelsene, at vi ikke skal ha nuller og at hver element skal være et heltall, beveger vi oss bort fra definisjonen av d.s. matriser. Det er nemlig kun én d.s. matrise som oppfyller disse kravene, 1×1 matrisen

$$A = \begin{pmatrix} 1 \end{pmatrix}$$

Vi skal likevel ikke bevege oss særlig langt vekk fra definisjonen av d.s. matriser. Vi skal kun utvide definisjonen, og der vi tidligere hadde krav om at linjesummene skulle være lik 1, skal vi nå kreve at linjesummene skal være lik det samme tallet l. Hvis vi også krever at diagonalene skal være lik l, i tillegg til at hvert element er unikt, får vi en matrise på formen av det som kalles et *magisk kvadrat*.

7.1 Magisk Kvadrat

Et *magisk kvadrat* A er en $n \times n$-matrise der hvert element er et unikt tall $1, 2, ..., n^2$, og hver rad, kolonne, diagonal har lik sum. Diagonalsummene kaller vi d_1 og d_2, der $d_1 = a_{11} + a_{22} + \cdots + a_{nn}$ og $d_1 = a_{n1} + a_{(n-1)2} + \cdots + a_{1n}$. En diagonalsum i denne sammenhengen er *ikke* en diagonal slik som en diagonal i permutasjonsmatriser. I dette tilfellet er en diagonal en mengde av elementene i A som går på skrå fra element a_{11} til a_{nn} og en annen diagonal fra a_{1n} til a_{n1}.

Eksempel 7.1.1. Vi har matrisen

$$A = \begin{pmatrix} 4 & 9 & 2 \\ 3 & 5 & 7 \\ 8 & 1 & 6 \end{pmatrix}$$

Vi kan regne ut radsummene, kolonnesummene og diagonalsummene

$$\begin{aligned}
r_1 &= 4+9+2 = 15 & r_2 &= 3+5+7 = 15 & r_3 &= 8+1+6 = 15 \\
k_1 &= 4+3+8 = 15 & k_2 &= 9+5+1 = 15 & k_3 &= 2+7+6 = 15 \\
d_1 &= 4+5+6 = 15 & d_2 &= 8+5+2 = 15 & &
\end{aligned}$$

Vi ser at alle radene, kolonnene og diagonale summerer til 15. Dermed er matrisen A et *magisk kvadrat.*

□

Definisjon 7.1.2. Rad-, kolonne- og diagonalsummen til et magisk kvadrat kalles ofte for den *magiske konstanten*, men vi velger å kalle den for *linjekonstanten l.*

Korollar 7.1.3. *Gitt et $n \times n$ magisk kvadrat A med linjesummen l vil $A' = \frac{1}{l} \cdot A$ være en d.s. matrise.*

Bevis. Vi har et $n \times n$ magisk kvadrat

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1(n-1)} & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2(n-1)} & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3(n-1)} & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{(n-1)1} & a_{(n-1)2} & a_{(n-1)3} & \cdots & a_{(n-1)(n-1)} & a_{(n-1)n} \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{n(n-1)} & a_{nn} \end{pmatrix}$$

Vi vet at linjesummene skal bli den samme konstanten, l, og finner $A' = \frac{1}{l} \cdot A$

$$A' = \frac{1}{l} \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1(n-1)} & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2(n-1)} & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3(n-1)} & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{(n-1)1} & a_{(n-1)2} & a_{(n-1)3} & \cdots & a_{(n-1)(n-1)} & a_{(n-1)n} \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{n(n-1)} & a_{nn} \end{pmatrix}$$

Hver radsum i A' er $r_i = \frac{1}{l}(a_{i1} + a_{i2} + \cdots + a_{in})$ for hver $i = 1, 2, \ldots, n$. Vi vet at hver radsum skal summeres til linjekonstanten l. Da blir $r_k = \frac{1}{l}(a_{i1} + a_{i2} + \cdots + a_{in}) = \frac{1}{l} \cdot l = 1$. Dermed er radsummene lik 1.

Hver kolonnesum i A' er $k_j = \frac{1}{l}(a_{1j} + a_{2j} + \cdots + a_{nj}$ for hver $j = 1, 2, \ldots, n$. Vi vet hver kolonnesum skal summeres til linjekonstanten l. Da blir $r_k = \frac{1}{l}(a_{1j} + a_{2j} + \cdots + a_{nj}) =$

$\frac{1}{l} \cdot l = 1$. Dermed er kolonnesummene lik 1. Diagonalsummene til A' vil også bli 1, men det stilles ikke krav til diagonalsummer i definisjonen av d.s. matriser. A' er d.s. ■

Ved å dividere alle elementene i et magisk kvadrat på linjekonstanten får vi en d.s. matrise. Magiske kvadrat skalert med $1/l$ er en delmengde av Ω_n og de har begrensninger som gjør at det ikke eksisterer uendelig mange slike matriser. I hvert fall ikke for små n. Ser vi bort fra rotasjon, speiling og permutasjoner av radene og kolonnene finnes det bare ett magisk kvadrat med størrelse 3×3. Den så vi på i eksempel 7.1.1.

7.2 Matriser med like linjesummer

Definisjon 7.2.1 (Matriser med lik linjekonstant). En *matrise med lik linjekonstant* B er en $n \times n$ matrise med radsummer og kolonnesummer lik samme linjekonstant l. Hvis B inneholder kun ikke-negative elementer, kaller vi B en *positiv* matrise med lik linjekonstant.

En matrise med lik linjekonstant er mindre begrenset enn et magisk kvadrat. Det stilles ikke krav om at alle tallene skal være forskjellige, og det stilles ikke krav til at elementene ikke kan være 0. I tillegg trenger ikke elementene å være heltall.

Eksempel 7.2.2. Vi har en 3×3-matrise, A,

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

Vi ser på radsummene og kolonnesummene

$$\begin{aligned} r_1 &= 3+0+0=3 & r_2 &= 0+1+2=3 & r_3 &= 0+2+1=3 \\ k_1 &= 3+0+0=3 & k_2 &= 0+1+2=3 & k_3 &= 0+2+1=3 \end{aligned}$$

Vi ser alle rad- og kolonnesummene er lik 3, og matrisen A er en matrise med lik linjekonstant. □

Som for magiske kvadrat, kan vi skalere ned elementene i matrisen med linjekonstanten, og vi får en d.s. matrise. Altså vil en matrise med lik linjekonstant $B' = \frac{1}{l} \cdot B$ være d.s.

Eksempel 7.2.3. Vi har en matrise med lik linjekonstant

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

Linjesummene, l, er 3, og vi finner $B' = \frac{1}{l} \cdot B$ er

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

Positive matriser med lik linjekonstant deler mange av de samme egenskapene og strukturene som d.s. matriser. Hele forrige kapittel om strukturer til d.s. gjelder også for positive matriser med lik linjekonstant. Den store forskjellen er at linjesummene til en positiv matrise med lik linjekonstant, kan være hva som helst av ikke-negative tall, mens for d.s. matriser skal den være 1.

For både magiske kvadrat og positive matriser med lik linjekonstant kan vi skrive matrisene som en vektet sum av permutasjonsmatriser. Magiske kvadrat er en delmengde av positive matriser med lik linjekonstant. Ved å vise at positive matriser med lik linjekonstant kan skrives som en vektet sum av permutasjonsmatriser, vil magiske kvadrat også kunne det.

Korollar 7.2.4. *En positiv matrise med lik linjekonstant A kan skrives som en vektet sum av permutasjonsmatriser.*

Bevis. Gitt en positiv matrise med lik linjekonstant A vet vi at vi har en d.s. matrise $A' = \frac{1}{l} \cdot A$ der l er linjesummene til A. Ettersom A' er d.s. vet vi fra Birkhoffs teorem at vi kan skrive A' som en vektet sum av permutasjonsmatriser.

$$A' = \sum_{i}^{t} c_i P_i$$

der $i = 1, 2, \ldots, t$. Vi har at $A' = \frac{1}{l} \cdot A$, dermed er $A = l \cdot A'$

$$\begin{aligned} A &= l \cdot (c_1 P_1 + c_2 P_2 + \ldots c_t P_t) \\ A &= lc_1 \cdot P_1 + lc_2 \cdot P_2 + \cdots + lc_t \cdot P_t \\ &= \sum_{i=1}^{t} lc_i \cdot P_i \end{aligned}$$

$c_1 + c_2 + \ldots c_t = 1$, og dermed vil de nye konstantene lc_t for $i = 1, 2, \ldots, t$ summeres til l fordi ser vi på summen av konstantene får vi $\sum_i^t lc_i \cdot = l \cdot \sum_i^n c_i = l \cdot 1$ da vi vet fra Birkhoff at $\sum_i^n c_i = 1$.
Altså kan alle positive matriser med lik linjekonstant A skrives som en vektet sum av permutasjonsmatrisene

$$\sum_{i=1}^{t} lc_i \cdot P_i$$

der konstantene

$$\sum_{i=1}^{t} lc_i = l$$

■

Eksempel 7.2.5. Vi kan se på et magisk kvadrat M av grad 3, og hvordan det kan summeres opp av permutasjonsmatriser

$$M = \begin{pmatrix} 4 & 9 & 2 \\ 3 & 5 & 7 \\ 8 & 1 & 6 \end{pmatrix}$$

$$\begin{aligned} M = 7 \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} + 4 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \\ + 2 \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \end{aligned}$$

For å konstruere det magiske kvadratet M, trenger vi summen av fem unike permutasjonsmatriser. Det finnes algoritmer og metoder for å bryte ned både magiske kvadrat og d.s. matriser ned i vektede summer av permutasjonsmatriser, men dette skal vi ikke gå inn på. Utregningen i forrige eksempel ble gjort for hånd.

7.3 Sudoku

Det finnes vel ikke en avis i verden som ikke dedikerer en del av en side til et par sudoku-brett. Sudoku er et velkjent spill med enkle krav. Som regel blir vi gitt et 9×9 brett hvor vi skal fylle ut alle tallene fra 1 til 9 langs hver rad og kolonne. Ingen tall skal gjentas på hver rad og kolonne. I tillegg er brettet delt i ni 3×3 bokser hvor alle tallene fra 1 til 9 også skal plasseres én gang. Noe av inspirasjonen til dette kapitlet er hentet fra Geir Dahls *Permutation matrices related to Sudoku* [Dah09].

Definisjon 7.3.1 (Latinsk kvadrat). Et latinsk kvadrat er en $n \times n$ matrise fylt med n elementer hvor hvert element befinner seg på hver rad og kolonne nøyaktig én gang.

Et Sudoku-brett er et 9×9 latinsk kvadrat med ni 3×3 kvadrater med tallene 1 til 9. De ni kvadratene er gitt på forhånd, og det er ikke vilkårlige 3×3 kvadrater på hele brettet. De ni kvadratene er markert med tykke streker i figur 7.1a.

2	5			3		9		1
	1				4			
4		7				2		8
		5	2					
				9	8	1		
	4				3			
			3	6			7	2
	7							3
9		3				6		4

(a) Et uløst Sudoku-brett

2	5	8	7	3	6	9	4	1
6	1	9	8	2	4	3	5	7
4	3	7	9	1	5	2	6	8
3	9	5	2	7	1	4	8	6
7	6	2	4	9	8	1	3	5
8	4	1	6	5	3	7	2	9
1	8	4	3	6	9	5	7	2
5	7	6	1	4	2	8	9	3
9	2	3	5	8	7	6	1	4

(b) Et ferdig løst Sudoku-brett

Figur 7.1: Sudoku-brett

Målet med sudoku er å ta et uløst sudoku-brett, det vil si et 9×9 brett hvor flere av tallene er gitte, og så fylle inn de manglende tallene. Hver rad, hver kolonne og hver 3×3 boks kun kan inneholde hvert tall én gang. Figur 7.1b viser et ferdig løst sudoku-brett. Posisjonene til hver verdi fra 1 til 9 ligger på hver sin diagonal. La oss se på posisjonene til enerne. Hver ener ligger alene på hver sin rad og kolonne.

								1
	1							
				1				
					1			
						1		
		1						
1								
			1					
							1	

Figur 7.2: Posisjonene til enerne

Ser man på dette brettet som en matrise, hvor de tomme posisjonene er nuller, danner dette brettet i figur 7.2 en permutasjonsmatrise. Faktisk ligger alle tallene på hver sin diagonal av en permutasjonsmatrise, og hvert eneste 9×9 sudokubrett S kan skrives som en lineærkombinasjon av 9 unike permutasjonsmatriser.

$$S = 1 \cdot P_1 + 2 \cdot P_2 + \cdots + 9 \cdot P_9$$

.

Bibliografi

[AMO93] Ahuja, R. K., Magnanti, T. L. og Orlin, J. B. *Network Flows: Theory, Algorithms, and Applications.* Prentice hall, 1993.

[Bha97] Bhatia, R. *Matrix Analysis.* Bd. 169. Springer, 1997.

[BM76] Bondy, J. A. og Murty, U. S. R. *Graph Theory with Applications.* New York: Elsevier, 1976.

[BR91] Brualdi, R. A. og Ryser, H. J. *Combinatorial Matrix Theory.* Cambridge [England]; New York: Cambridge University Press, 1991.

[CR22] Christophersen, J. A. og Ranestad, K. *Geometro.* Universitetet i Oslo, 2022.

[Dah01] Dahl, G. 2001, s. 1–42.

[Dah09] Dahl, G. «Permutation matrices related to Sudoku». I: *Linear Algebra and Its Applications - LINEAR ALGEBRA APPL* 430 (apr. 2009), s. 2457–2463. DOI: 10.1016/j.laa.2008.12.023.

[GYZ13] Gross, J. L., Yellen, J. og Zhang, P. *Handbook of Graph Theory, Second Edition.* 2nd. Chapman & Hall/CRC, 2013.

[MOA11] Marshall, A. W., Olkin, I. og Arnold, B. C. *Inequalities: Theory of Majorization and its Applications.* Second. Bd. 143. Springer, 2011. DOI: 10.1007/978-0-387-68276-1.

Tillegg A

Python-kode

A.1 Invers matrise

Her er koden brukt for å finne den inverse matrisen i eksempel 5.2.4. Elementene i matrisen ble gitt i desimaltall og ikke brøk, så vi måtte likevel regne ut hva determinanten var for å finne at hvert element var en brøk med 11 i nevneren.

```
import numpy as np

# Importerer linalg fra scipy
from scipy import linalg

# En 3x3 matrise
A = np.matrix([[2/10,3/10,5/10],
               [3/10, 6/10,1/10],
               [5/10,1/10,4/10]])

# 3x3-matrisen
print("3x3-matrisen A:")
print(A)

# Finner den inverse matrisen B
invers = linalg.inv(A)

# Printer resultatet av den inverse matrisen
print("Den inverse matrisen til A er:")
print(invers)
```

A.2 Matrisemultiplikasjon

Her er koden for å utføre matrisemultiplikasjonen i eksempel 5.2.6

```
    # Matrise P1
P1 = [[0.2, 0.1, 0.7],
      [0.7, 0.3, 0],
      [0.1, 0.6, 0.3]]

# Matrise P2
P2 = [[0.5, 0.5, 0],
      [0.1, 0.2, 0.7],
      [0.4, 0.3, 0.3]]

# Tom matrise
produkt = [[0, 0, 0],
           [0, 0, 0],
           [0, 0, 0]]

# Tre lokker for aa finne hvert element i produktet

for i in range(len(P1)):
    for j in range(len(P2[0])):
        for k in range(len(P2)):
            produkt[i][j] += P1[i][k] * P2[k][j]
for p in produkt:
    print(p)


# Resultat
[0.39, 0.33, 0.27999999999999997]
[0.38, 0.41, 0.21]
[0.22999999999999998, 0.26, 0.51]
```