

Ole Fredrik Brevig, Arne B. Sletsjøe

Kompendium, MAT 1060

Innhold

1 Matriseregning **3**
- 1.1 Vektorer . . . 3
- 1.2 Skalarprodukt . . . 7
- 1.3 Matriser . . . 8
- 1.4 Kvadratiske matriser . . . 13
- 1.5 Lineære likningssytemer . . . 18
- 1.6 Determinanter . . . 26
- 1.7 Spektralteori . . . 31
- 1.8 Oppgaver med løsning . . . 38

2 Dynamiske systemer **43**
- 2.1 Diskrete dynamiske systemer . . . 43
- 2.2 Lukkede dynamiske systemer . . . 48
- 2.3 Oppgaver med løsning . . . 52

3 Euklidske vektorrom **56**
- 3.1 Lineære underrom . . . 56
- 3.2 Lineære avbildninger . . . 60
- 3.3 Generelle vektorrom* . . . 71
- 3.4 Oppgaver med løsning . . . 74

4 Geometri i planet **77**
- 4.1 Geometrisk tolkning av lineære avbildninger . . . 77
- 4.2 Stive bevegelser i planet . . . 82
- 4.3 Symmetrier av plane figurer . . . 93
- 4.4 Plane krystaller . . . 94
- 4.5 Oppgaver med løsning . . . 101

5 Geometri i rommet **104**
- 5.1 Vektorer i $\mathbb{R}^3$ . . . 104
- 5.2 Ortogonale 3×3-matriser . . . 109
- 5.3 Isometrier i $\mathbb{R}^3$ . . . 113
- 5.4 Oppgaver med løsning . . . 116

6 Gruppeteori **118**
- 6.1 Symmetrigrupper . . . 118
- 6.2 Abstrakte grupper . . . 122
- 6.3 Den ortogonale gruppa $O(3)$ . . . 124
- 6.4 Symmetrier av molekyler . . . 128

6.5 Oppgaver med løsning 129

7 Indreprodukt **130**
7.1 Et generelt indreproduktbegrep 130
7.2 Eksempler på indreprodukt 133
7.3 Dekomponering av vektorer ved indreprodukt 136
7.4 Fourier-rekker 139
7.5 Oppgaver med løsning 146

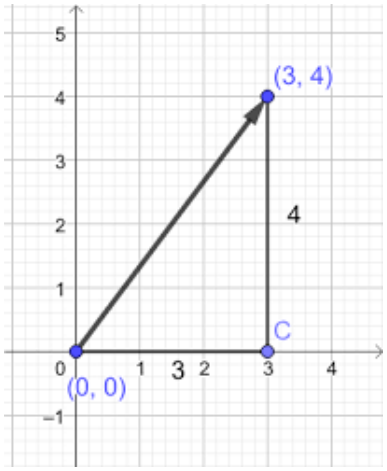
8 Oppgavesett **149**
8.1 Oppgavesett 149
8.2 Fasit til oppgavesett 159

Kapittel 1

Matriseregning

1.1 Vektorer

Et grunnleggende begrep i lineær algebra er vektorer. En vektor kan illustreres geometrisk som en pil, med gitt retning og lengde. Retningen på vektoren følger retningen på pila. To vektorer er like dersom de har samme lengde og retning. Vektorenes plassering i planet (eller rommet) har ingen betydning. To eksempler på vektorer, hentet fra fysikk, er hastighet og kraft. Både hastighet og kraft skjer i vektorens retning og lengden av vektoren illustrerer henholdsvis farten og styrken i krafta.



En annen måte å framstille vektorer er som såkalte tall-tupler, f.eks. $(3,4)$. Den geometriske tolkningen av vektoren $(3,4)$ er en pil som starter i origo og ender i punktet $(3,4)$. Dette gir oss en 1-1-sammenheng mellom punkter i planet, vektorer og tallpar. Det fine med å bruke tall-tupler er at denne notasjonen ikke begrenser seg til å knytte vektorer til geometri i planet eller rommet. 5-tuplet $(2,4,-1,3,0)$ har ikke noen klar geometrisk tolkning, men kan likevel være svært nyttig for oss i mange sammenhenger.

Det er vanlig å bruke fet skrift for å betegne en vektor, som i $\mathbf{v}$ og $\mathbf{w}$, eller ved en pil over vektornavnet, som i $\overrightarrow{v}$. For lengden av en vektor bruker vi notasjonen $\|\mathbf{v}\|$. Siden

lengden av en vektor $\mathbf{v} = (a_1, a_2)$ i planet,

$$\|\mathbf{v}\| = \|(a_1, a_2)\| = \sqrt{a_1^2 + a_2^2}$$

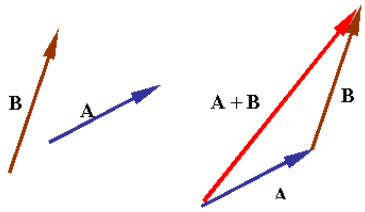
er gitt ved å bruke Pythagoras formel, etteraper vi dette også for mer generelle vektorer, $\mathbf{v} = (a_1, \ldots, a_n)$ hvor n er et vilkårlig positivt heltall, og sier at lengden av $\mathbf{v}$ er gitt ved

$$\|\mathbf{v}\| = \|(a_1, \ldots, a_n)\| = \sqrt{a_1^2 + \cdots + a_n^2}$$

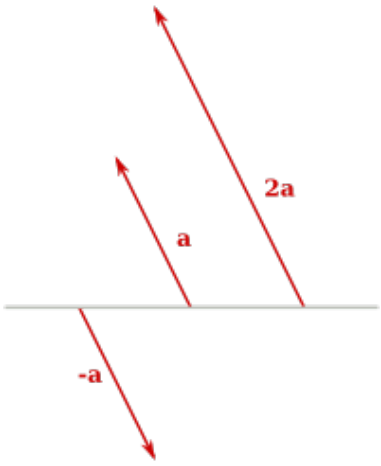
Eksempel 1.1.1.

$$\|(2,4,-1,3,0)\| = \sqrt{2^2 + 4^2 + (-1)^2 + 3^2 + 0^2} = \sqrt{30}$$

Vi kan legge sammen vektorer. Hvis de to vektorene er illustrert ved hver sin pil, finner vi summen av dem ved å legge pilene etter hverandre, Summen av vektornene er da gitt ved en pil som starter i startpunktet for den ene vektoren og ender i endepunktet for den andre.



Hvis vi multipliserer en vektor med et tall, så beholder vi retningen og endrer lengden i henhold til størrelsen på tallet. Vektoren $2\mathbf{v}$ peker i samme retning som $\mathbf{v}$, men er dobbelt så lang. Multiplikasjon med negative tall gir oss vektorer i motsatt retning, så vektoren $-\mathbf{v}$ er like lang som $\mathbf{v}$, men peker stikk motsatt vei. Spesielt har vi $1 \cdot \mathbf{v} = \mathbf{v}$.



Vi har tilsvarende operasjoner for n-tupler av reelle tall. La

$$\mathbf{v} = (a_1, a_2, \ldots, a_n) \in \mathbb{R}^n$$

Da har vi

$$2\mathbf{v} = (2a_1, 2a_2, \ldots, 2a_n), \quad -\mathbf{v} = (-a_1, -a_2, \ldots, -a_n)$$

Tallene som inngår i en vektor kalles **koordinater** og addisjon og multiplikasjon med reelle tall skjer **koordinatvis**;

$$(a_1,\ldots,a_n)+(b_1,\ldots,b_n)=(a_1+b_1,\ldots,a_n+b_n)$$
$$k\cdot(a_1,\ldots,a_n)=(ka_1,\ldots,ka_n)$$

for to vektorer $(a_1,\ldots,a_n)$ og $(b_1,\ldots,b_n)$ og et reelt tall $k\in\mathbb{R}$. Den eneste begrensningen er at ved addisjon av vektorer må **størrelsen** (dvs. antall koordinater) på de to vektorene være den samme.

Vi kaller multiplikasjon av en vektor med et tall for **skalarmultiplikasjon**. Skalarmultiplikasjon tilfredsstiller de samme regnereglene som vanlig multiplikasjon, dvs. $r(s\mathbf{v})=(rs)\mathbf{v}$ og $(r+s)\mathbf{v}=r\mathbf{v}+s\mathbf{v}$ for reelle tall r og s og en vektor $\mathbf{v}$.

Eksempel 1.1.2.

$$\begin{aligned}(2+3)(2,-1,3,0)&=2(2,-1,3,0)+3(2,-1,3,0)\\&=(4,-2,6,0)+(6,-3,9,0)\\&=(10,-5,15,0)\\5(2,-1,3,0)&=(10,-5,15,0)\end{aligned}$$

Vektoraddisjon er **kommutativ** ($\mathbf{a}+\mathbf{b}=\mathbf{b}+\mathbf{a}$), og **assosiativ** $\mathbf{a}+(\mathbf{b}+\mathbf{c})=(\mathbf{a}+\mathbf{b})+\mathbf{c}$), dvs. at rekkefølgen på vektorene og vektoraddisjonene ikke spiller noen rolle. Dette kommer av at vi har tilsvarende egenskaper for reelle tall. Subtraksjon av to vektorer definerer vi ved å kombinere multiplikasjon av en vektor med et negativt tall med vektoraddisjon. Vi definerer null-vektoren $\mathbf{0}$ som en vektor med bare 0-er.

$$\mathbf{0}=(0,0,\ldots,0)$$

Den har lengde 0 og har derfor ikke noen spesiell retning. Videre har vi $0\cdot\mathbf{a}=\mathbf{0}$ for alle vektorer $\mathbf{a}$.

Det viser seg at det er hensiktsmessig å bruke to ulike måter å angi en vektor. Vi kan skrive den som en **radvektor**;

$$(a_1,\ldots,a_n)$$

eller som en **søylevektor**;

$$\begin{pmatrix}a_1\\\vdots\\a_n\end{pmatrix}$$

De to skrivemåtene angir essensielt den samme vektoren, men vi skal bruke dem litt forskjellig. Bakgrunnen for denne forskjellen vil bli klarere når vi kommer til matriser. Overgangen fra den ene til den andre formen har også et eget navn; vi kaller det å **transponere**, og skriver

$$(a_1,\ldots,a_n)^T=\begin{pmatrix}a_1\\\vdots\\a_n\end{pmatrix}$$

og omvendt. Vektorene er de **transponerte** til hverandre.

Hvis vi har gitt en mengde $\mathscr{V}=\{\mathbf{v}_1,\ldots,\mathbf{v}_m\}$ av vektorer kan vi produsere enda flere vektorer ved å danne **lineærkombinasjoner** av vektorene.

Definisjon 1.1.1. *En **lineærkombinasjon** av vektorene $\mathbf{v}_1,\ldots,\mathbf{v}_m$ er en sum*

$$\mathbf{w} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_m\mathbf{v}_m$$

for reelle tall $a_1,\ldots,a_m \in \mathbb{R}$.

Dersom dette er den eneste måten vi kan skrive $\mathbf{w}$ som en lineærkombinasjon av $\mathbf{v}_1,\ldots,\mathbf{v}_m$ sier vi at mengden $\mathscr{V} = \{\mathbf{v}_1,\ldots,\mathbf{v}_m\}$ er **lineært uavhengig**. Dersom en mengde av vektorer er lineært uavhengig, vil denne egenskapen gjelde for alle valg av $\mathbf{w}$, også for nullvektoren $\mathbf{0}$. Den formelle definisjonen av lineær uavhengighet ser derfor slik ut:

Definisjon 1.1.2. *En mengde $\mathscr{V} = \{\mathbf{v}_1,\ldots,\mathbf{v}_m\}$ av vektorer sies å være **lineært uavhengig** dersom*

$$\sum_{i=1}^{m} a_i\mathbf{v}_i = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_m\mathbf{v}_m = \mathbf{0}$$

*medfører at alle $a_i = 0$, $i = 1,2,\ldots,m$. En mengde av vektorer som ikke er lineært uavhengig sies å være **lineært avhengig**.*

Hvis vi antar at mengden $\mathscr{V} = \{\mathbf{v}_1,\ldots,\mathbf{v}_m\}$ er lineært uavhengig og samtidig tenker oss at en vektor $\mathbf{w}$ kan skrives på to ulike måter som lineærkombinasjoner av vektorene i $\mathscr{V}$;

$$\mathbf{w} = a_1\mathbf{v}_1 + \cdots + a_m\mathbf{v}_m = b_1\mathbf{v}_1 + \cdots + b_m\mathbf{v}_m$$

så følger det at

$$\mathbf{w} - \mathbf{w} = (a_1 - b_1)\mathbf{v}_1 + \cdots + (a_m - b_m)\mathbf{v}_m = \mathbf{0}$$

Men i henhold til definisjonen av lineær uavhengighet betyr det at alle koeffiisientene $(a_j - b_j) = 0$ og de to skrivemåtene for $\mathbf{w}$ er de samme.

Eksempel 1.1.3. *Vektorene*

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 1 \\ 4 \end{pmatrix}$$

er lineært avhengige siden

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \\ 4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

mens vektorene

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

er lineært uavhengige siden

$$a_1\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + a_2\begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} + a_3\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

kun er mulig dersom $a_1 = a_2 = a_3 = 0$.

1.2 Skalarprodukt

Siden vi kan legge sammen vektorer av samme størrelse er det naturlig å spørre om vi også kan gange sammen vektorer. Svaret er at det kan vi. Produktet er imidlertid ikke en ny vektor, men derimot et reelt tall. Produktet kalles **skalarproduktet** eller **prikkproduktet** av de to vektorene. For å regne ut skalarproduktet av to vektorer må vektorene ha samme størrelse.

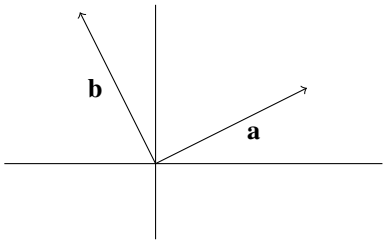
Definisjon 1.2.1. *La* $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ *og* $\mathbf{w} = (w_1, w_2, \ldots, w_n)$ *være to vektorer av samme størrelse.* ***Skalarproduktet*** *av vektorene* $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ *og* $\mathbf{w} = (w_1, w_2, \ldots, w_n)$ *er gitt ved*

$$\begin{aligned}\mathbf{v} \cdot \mathbf{w} &= (v_1, v_2, \ldots, v_n) \cdot (w_1, w_2, \ldots, w_n) \\ &= v_1 w_1 + v_2 w_2 + \cdots + v_n w_n\end{aligned}$$

Eksempel 1.2.1. *De to vektorene* $\mathbf{a} = (2,1)$ *og* $\mathbf{b} = (-1,2)$ *har skalarprodukt*

$$(2,1) \cdot (-1,2) = 0$$

Figuren viser en geometrisk framstilling av $\mathbf{a}$ *og* $\mathbf{b}$ *som piler.*



Vi observerer at vektorene står normalt på hverandre. Det er et generelt faktum for alle 2- og 3-vektorer at skalarprodukt lik 0 betyr at vektorene står normalt på hverandre.

Eksempel 1.2.2.

$$(2,1,3,-1) \cdot (1,0,0,-2) = 2 + 0 + 0 + 2 = 4$$

$$(2,1,3,-1) \cdot (2,1,3,-1) = \|(2,1,3,-1)\|^2 = 2^2 + 1^2 + 3^2 + (-1)^2 = 15$$

Hvis vi tar skalarproduktet av vektoren $\mathbf{a} = (2,1)$ med seg selv, gjør vi en annen observasjon;

$$\mathbf{a} \cdot \mathbf{a} = (2,1) \cdot (2,1) = 2^2 + 1^2$$

som viser at $\mathbf{a} \cdot \mathbf{a} = \|\mathbf{a}\|^2$ eller $\|\mathbf{a}\| = \sqrt{\mathbf{a} \cdot \mathbf{a}}$.

Dette gjelder ikke kun for 2-vektorer, generelt har vi

$$(v_1, \ldots, v_n) \cdot (v_1, \ldots, v_n) = v_1^2 + \cdots + v_n^2 = \|(v_1, \ldots, v_n)\|^2$$

eller

$$\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}}$$

For to vektorer som peker i samme retning, $\mathbf{a} = (a_1, a_2)$ og $\mathbf{b} = k\mathbf{a} = (ka_1, ka_2)$ har vi

$$\mathbf{a} \cdot \mathbf{b} = (a_1, a_2) \cdot (ka_1, ka_2) = k(a_1^2 + a_2^2) = \sqrt{a_1^2 + a_2^2} \cdot \sqrt{k^2 a_1^2 + k^2 a_2^2} = \|\mathbf{a}\| \cdot \|\mathbf{b}\|$$

Så skalarproduktet av to parallelle vektorer er gitt ved produktet av lengdene av vektorene.

En intuitiv tolkning av skalarproduktet baserer seg på de to ytterlighetene vi har beskrevet over, skalarproduktet av to vektorer måler i hvilken grad to vektorer peker i samme retning.

1.3 Matriser

Det neste begrepet vi skal se på er **matriser**.

Definisjon 1.3.1. *En reell $n \times m$-**matrise** A er en tabell med n **rader** og m **søyler**.*

$$A = \begin{pmatrix} a_{11} & \dots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nm} \end{pmatrix},$$

*der **elementene** a_{ij}, $i = 1,\dots,n$, $j = 1,\dots,m$, i matrisen er reelle tall. Vi bruker ofte den kortfattede skrivemåten*

$$A = (a_{ij})$$

De to tallene m og n kan være hvilke som helst positive heltall. Siden det ene tallet angir antall rader og det andre antall søyler, er rekkefølgen helt vesentlig. Matrisen

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

er en 3×2-matrise, mens

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

er en 2×3-matrise. Generelt har vi $m \times n =$ (antall rader) $\times$ (antall søyler).

Sammenhengen mellom de to matrisene over svarer til sammenhengen mellom rad- og søylevektorer som transponerte av hverandre. Vi kan formalisere dette i en generell definisjon:

Definisjon 1.3.2. *La A være $n \times m$-matrisen*

$$A = \begin{pmatrix} a_{11} & \dots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nm} \end{pmatrix},$$

Da sier vi at $m \times n$-matrisen

$$A^T = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix},$$

*er den **transponerte** av A. Transponering innebærer å bytte om på rader og søyler.*

Eksempel 1.3.1.

$$\begin{pmatrix} 1 & -1 \\ -3 & 0 \\ 1 & 2 \end{pmatrix}^T = \begin{pmatrix} 1 & -3 & 1 \\ -1 & 0 & 2 \end{pmatrix}$$

Siden to ombytter bringer oss tilbake til utgangspunktet vil vi generelt ha at $(A^T)^T = A$.

Vi trenger ikke å begrense oss til reelle elementer i matrisen, vi kan gjerne studere komplekse matriser, hvor elementene i matrisen er komplekse tall. Men vi velger å holde oss til reelle tall dersom ikke annet er angitt.

Vi kan også betrakte vektorer som matriser. En **radvektor** $(a_1, \ldots, a_m)$ kan vi oppfatte som en $1 \times m$-matrise, og en **søylevektor**

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

som en $n \times 1$-matrise.

Hvis vi har to matriser av samme størrelse, kan vi legge dem sammen, omtrent som vi gjør med vanlige tall.

Definisjon 1.3.3. *La A og B være to reelle $n \times m$-matriser, $A = (a_{ij})$ og $B = (b_{ij})$. Da er summen av dem gitt ved $A + B = (a_{ij} + b_{ij})$, som igjen er en reell $n \times m$-matrise. Vi kan også gange en matrise med et reelt tall c. Det gjør vi ved å gange alle elementene i matrisen med det samme tallet, $c \cdot (a_{ij}) = (ca_{ij})$. Vi kaller denne operasjonen for* ***skalar multiplikasjon****.*

Eksempel 1.3.2. *La*

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 1 & -1 & 2 \\ 2 & 0 & -1 \end{pmatrix} \text{ og } B = \begin{pmatrix} -1 & 0 & 1 \\ 1 & 1 & -1 \\ 2 & 1 & 0 \end{pmatrix}$$

Da har vi

$$A + B = \begin{pmatrix} 1-1 & 2+0 & 0+1 \\ 1+1 & -1+1 & 2-1 \\ 2+2 & 0+1 & -1+0 \end{pmatrix} = \begin{pmatrix} 0 & 2 & 1 \\ 2 & 0 & 1 \\ 4 & 1 & -1 \end{pmatrix}$$

Eksempel 1.3.3. *La A være som i forrige eksempel. Da har vi*

$$2A = 2\begin{pmatrix} 1 & 2 & 0 \\ 1 & -1 & 2 \\ 2 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 2 & 4 & 0 \\ 2 & -2 & 4 \\ 4 & 0 & -2 \end{pmatrix}$$

Vi kan også multiplisere en matrise med det reelle tallet 0. Det gir oss $\mathbf{0}$-matrisen (her av størrelse 2×3).

$$\mathbf{0} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

Eksempel 1.3.4. *Vi har*

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

og

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

Den neste proposisjonen gir en oppsummering av de formelle egenskapene for matriser vi har studert:

Proposisjon 1.3.4. *La A, B og C være reelle $n \times m$-matriser. Da har vi at:*

i) Addisjon av matriser er assosiativ, dvs,

$$(A+B)+C = A+(B+C)$$

ii) Addisjon av matriser er kommutativ, dvs,

$$A+B = B+A$$

iii) Null-matrisen er et 0-element, det vil si $A+\mathbf{0} = A$

iv) Enhver matrise har en additiv invers, det vil si en matrise $-A = (-1)A$ som oppfyller $A+(-A) = \mathbf{0}$.

v) Skalar multiplikasjon er assosiativ, dvs,

$$c(dA) = (cd)A$$

vi) Skalar multiplikasjon er distributiv over addisjon av matriser, dvs,

$$(c+d)A = cA+dA$$

og

$$c(A+B) = cA+cB$$

vii) $1 \cdot A = A$.

Bevisene for alle disse påstandene er mer eller mindre trivielle og vi lar det være opp til leseren å skrive dem ut.

Under forutsetning av at størrelsene passer sammen kan matriser multipliseres med hverandre.

Definisjon 1.3.5. *La A være en $n \times p$-matrise og B en $p \times m$-matrise, $A = (a_{ij})$ og $B = (b_{ij})$. Da er produktet av dem gitt ved $A \cdot B = (c_{ij})$ hvor*

$$c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj}$$

og produktet er en $n \times m$-matrise.

Det er viktig at antallet søyler i den første matrisen er det samme som antall rader i den andre. Denne definisjonen kan være litt vanskelig å få tak på, men det hjelper kanskje med et eksempel.

Eksempel 1.3.5. *La*

$$A = \begin{pmatrix} 0 & 2 & 1 \\ 2 & 0 & 1 \\ 3 & 1 & -1 \end{pmatrix}, \quad B = \begin{pmatrix} -1 & 1 \\ 3 & 0 \\ -1 & -2 \end{pmatrix}$$

være en 3×3*-, henholdsvis* 3×2*-matrise. Da passer størrelsene slik at vi kan definere produktet* $A \cdot B$*.*

$$(3 \times 3) \cdot (3 \times 2) = (3 \times 2)$$

Det blir en ny 3×2*-matrise*

$$A \cdot B = \begin{pmatrix} 5 & -2 \\ -3 & 0 \\ 1 & 5 \end{pmatrix}$$

hvor elementene framkommer slik;

$$\begin{aligned} 5 &= (0,2,1) \cdot (-1,3,-1) = 0 \cdot (-1) + 2 \cdot 3 + 1 \cdot (-1) \\ -2 &= (0,2,1) \cdot (1,0,-2) = 0 \cdot 1 + 2 \cdot 0 + 1 \cdot (-2) \\ -3 &= (2,0,1) \cdot (-1,3,-1) = 2 \cdot (-1) + 0 \cdot 3 + 1 \cdot (-1) \\ &\textit{osv.} \end{aligned}$$

Elementene i matriseproduktet framkommer ved å ta skalarproduktene mellom rader i den høyre matrisen med søyler i den venstre. Det betyr at for to sølevektorer

$$\mathbf{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

så kan vi skrive skalarproduktet

$$a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$$

som et matriseprodukt

$$\mathbf{a}^T \mathbf{b} = (a_1, \ldots, a_n) \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = (a_1 b_1 + a_2 b_2 + \cdots + a_n b_n)$$

Dette produktet vil være en 1×1-matrise, mens skalarproduktet er et tall. Siden det er en opplagt 1-1-sammenheng mellom reelle tall og reelle 1×1-matriser; $a \mapsto (a)$, kommer vi til å bruke disse notasjonene litt om hverandre.

De formelle egenskapene ved matrise-multiplikasjon er som følger:

Proposisjon 1.3.6. *La A, B og C være reelle matriser av rett størrelse for de aktuelle formålene i denne proposisjonen. Da har vi at:*

i) Matrise-multiplikasjon er assosiativ, dvs,

$$(A \cdot B) \cdot C = A \cdot (B \cdot C)$$

ii) Matrise-multiplikasjon er distributiv over addisjon, dvs,

$$(A+B)\cdot C = A\cdot C + B\cdot C$$

og

$$A\cdot(B+C) = A\cdot B + A\cdot C$$

Proposisjonen bevises ved å bruke definisjonene direkte og er utelatt.

Eksempel 1.3.6. *La matrisene A og B være som i forrige eksempel og la*

$$C = \begin{pmatrix} 1 & -1 \\ 2 & 0 \end{pmatrix}$$

Da har vi

$$(A\cdot B)\cdot C = \begin{pmatrix} 5 & -2 \\ -3 & 0 \\ 1 & 5 \end{pmatrix} \cdot \begin{pmatrix} 1 & -1 \\ 2 & 0 \end{pmatrix} = \begin{pmatrix} 1 & -5 \\ -3 & 3 \\ 11 & -1 \end{pmatrix}$$

På den annen side har vi

$$B\cdot C = \begin{pmatrix} -1 & 1 \\ 3 & 0 \\ -1 & -2 \end{pmatrix} \cdot \begin{pmatrix} 1 & -1 \\ 2 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 3 & -3 \\ -5 & 1 \end{pmatrix}$$

og dermed

$$A\cdot(B\cdot C) = \begin{pmatrix} 0 & 2 & 1 \\ 2 & 0 & 1 \\ 3 & 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 3 & -3 \\ -5 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -5 \\ -3 & 3 \\ 11 & -1 \end{pmatrix}$$

som bekrefter punkt i) i proposisjonen.

Vær oppmerksom at rekkefølgen på faktorene i en matrisemultiplikasjon er helt vesentlig. Normalt har vi at $AB = \neq BA$. Det er til og med sånn at selv om produktet den ene veien (AB) er veldefinert, trenger ikke produktet den andre veien (BA) å være definert i det hele tatt. F.eks. er produktet AB av en 2×3-matrise A og en 3×4-matrise B helt ok, mens den motsatte rekkefølgen, BA ville være et produkt av en 3×4-matrise med en 2×3-matrise som jo ikke lar seg gjøre.

Dette har betydning for transponering av et produkt. Regelen er da at $(AB)^T = B^T A^T$, vi må altså bytte om rekkefølgen på faktorene når vi transponerer.

I det neste eksemplet skal vi se på en praktisk anvendelse av matrisemultiplikasjon. Vi tar utgangspunkt i næringsmiddeltabellen for ingrediensene som inngår i vafler og pannekaker. Dette skal vi kombinere med oppskrifter på vaffel- og pannekakerøre for å beregne hvor mye protein, fett og karbohydrat det er i vafler og pannekaker.

Ingredienser	Hvetemel	Sukker	Melk	Egg	Smør
Vaffelrøre (1000g)	240	90	400	171	100
Pannekakerøre (1000g)	195	0	542	247	16

Næringsinnhold pr. 100 gram vare	Protein	Fett	Karbohydrat
Hvetemel	11,5	1,5	71,0
Sukker	0,0	0,0	100,0
Melk	3,4	3,9	4,8
Egg	12,5	11,3	0,7
Smør	0,5	82,0	0,2
Energiinnhold i 1 gram	17kJ	37kJ	17kJ

Vi betrakter de to tabellene som matriser (elementene i R angis pr gram);

$$N = \begin{pmatrix} 11,5 & 1,5 & 71,0 \\ 0,0 & 0,0 & 100,0 \\ 3,4 & 3,9 & 4,8 \\ 12,5 & 11,3 & 0,7 \\ 0,5 & 82,0 & 0,2 \end{pmatrix} \quad R = \begin{pmatrix} 0,24 & 0,09 & 0,4 & 0,17 & 0,1 \\ 0,20 & 0 & 0,54 & 0,25 & 0,02 \end{pmatrix}$$

og i tillegg ser vi på energiinnholdet som en vektor

$$E = \begin{pmatrix} 17 \\ 37 \\ 17 \end{pmatrix}$$

Matrisen N er en 5×3-matrise siden vi har 5 ingredienser og 3 næringsstoffer. Matrisen R er en 2×5-matrise siden vi har 2 sluttprodukter og 5 ingredienser. Når vi tar produktet baker vi på sett og vis ingriediensene inn mellom sluttproduktet og næ?ngsstoffene. Produktet RN vil gi oss en 2×3-matrise hvor radene svarer til vaffel- og pannekakerøre og søylene gir innholdet av protein, fett og karbohydrat i 100g med røre.

$$RN = \begin{pmatrix} 6,3 & 12,1 & 28,1 \\ 7,2 & 6,5 & 16,6 \end{pmatrix}$$

Energiinnholdet i vafler og pannekaker finner vi å regne ut produktet RNE;

$$RNE = \begin{pmatrix} 6,3 & 12,1 & 28,1 \\ 7,2 & 6,5 & 16,6 \end{pmatrix} \begin{pmatrix} 17 \\ 37 \\ 17 \end{pmatrix} = \begin{pmatrix} 1031 \\ 646 \end{pmatrix}$$

Benevnelsen i denne vektoren er kJ. Hvis vi skal omregne til kcal må vi gange med 0,239, som gir henholdsvis 246 og 154 kcal for vafler og pannekaker.

1.4 Kvadratiske matriser

Kvadratiske matriser er matriser med like mange rader som søyler, dvs. matriser av størrelse $n \times n$ for et positivt heltall n. Slik matriser har mange fine egenskaper. Blant annet finnes det en **identitetsmatrise**, I (eller I_n dersom vi har behov for a angi størrelsen) gitt ved $n \times n$-matrisen

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

dvs. en **diagonal** matrise med 1-ere på **hoveddiagonalen** (elementer med samme rad- og søyle-nummer), og 0-er ellers. Denne matrisen har mange av de samme egenskapene som tallet 1 og kalles derfor også for **enhetsmatrisen**. Spesielt har vi for en kvadratisk $n \times n$-matrise A at $A \cdot I_n = I_n \cdot A = A$.

Kvadratiske matriser har også den nyttige egenskapen at produktet av to $n \times n$-matriser er en ny $n \times n$-matrise.

Alle tall bortsett fra 0 har multiplikative inverser, dvs. det finnes et annet reelt tall som multiplisert med tallet gir oss 1. For eksempel er $\frac{1}{2}$ den multiplikative inversen til 2 fordi $2 \cdot \frac{1}{2} = 1$. På samme måte kan kvadratiske matriser ha inverser. I likhet med tallene har ikke $\mathbf{0}$-matrisen noen invers, men ulikt tallene finnes det mange andre kvadratiske matriser som heller ikke har noen invers.

Definisjon 1.4.1. *La A være en kvadratisk $n \times n$-matrise. Anta at det finnes en annen $n \times n$-matrise B slik at $A \cdot B = B \cdot A = I_n$. Da sier vi at B er den inverse matrisen til A, og vi skriver $B = A^{-1}$. Dersom det finnes en invers til A sier vi at matrisen A er **invertibel**.*

Man kan vise at dersom det finnes en B slik at $A \cdot B = I$, så vil også $B \cdot A = I$, men vi går ikke inn på beviset her.

Eksempel 1.4.1. *Matrisen*

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

har invers

$$A^{-1} = \begin{pmatrix} -1 & -2 \\ 1 & 1 \end{pmatrix}$$

fordi

$$A \cdot A^{-1} = \begin{pmatrix} 1 & 2 \\ -1 & -1 \end{pmatrix} \cdot \begin{pmatrix} -1 & -2 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Vi har også

$$A^{-1} \cdot A = \begin{pmatrix} -1 & -2 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ -1 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

De formelle egenskapene for inverse matriser er gitt i den neste proposisjonen.

Proposisjon 1.4.2. *La A være en invertibel $n \times n$-matrise. Da har vi at*

i) Den inverse matrisen er entydig bestemt av A,

ii) $(A^{-1})^{-1} = A$

ii) $(AB)^{-1} = B^{-1}A^{-1}$

Bevis. For å bevise i) antar vi at det finnes to inverser B_1 og B_2 til A. Bruker vi den definerende egenskapen til den inverse matrisen, sammen med assosiativitet av matrise-multiplikasjon, får vi

$$B_1 = B_1 \cdot I_n = B_1 \cdot (A \cdot B_2) = (B_1 \cdot A) \cdot B_2 = I_n \cdot B_2 = B_2$$

Med andre ord, de to matrisene vi startet med kan ikke være forskjellige. De to andre punktene følger nå fra entydigheten i i) og ved definisjonen av den inverse. F.eks. er A^{-1} definert ved at

$$(A^{-1}) \cdot (A^{-1})^{-1} = I$$

Men vi har også at $(A^{-1}) \cdot A = I$. Entydigheten av inversen til A^{-1} gir da at $(A^{-1})^{-1} = A$. □

Eksempel 1.4.2. *Det er ikke alltid enkelt å regne ut inversen til en matrise. Senere skal vi se på noen forskjellige teknikker som vi kan benytte oss av. Men i noen tilfeller er det relativt enkelt å beregne inverser. Spesielt gjelder det for diagonale matriser. I det tilfellet kan vi bare invertere elementene på diagonalen (under forutsetning av at disse er* $\neq 0$*):*

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

Grunnen er at produktet av to diagonale matriser er en diagonal matrise bestående av *samplasserte* produkter på diagonalen.

$$\begin{pmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{pmatrix} \cdot \begin{pmatrix} b_{11} & 0 & 0 \\ 0 & b_{22} & 0 \\ 0 & 0 & b_{33} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} & 0 & 0 \\ 0 & a_{22}b_{22} & 0 \\ 0 & 0 & a_{33}b_{33} \end{pmatrix}$$

En spesiell egenskap for kvadratiske matriser er at de har samme størrelse som den transponerte matrisen. Den transponerte av en $n \times m$-matrise er en $m \times n$-matrise, så hvis $n = m$ vil begge matrisene være $n \times n$-matriser.

Definisjon 1.4.3. *En kvadratisk matrise A som er lik sin egen transponerte, kalles en* ***symmetrisk*** *matrise. Vi skriver denne betingelsen som* $A^T = A$.

Eksempel 1.4.3. *Symmetriske matriser kan vi speile om hoveddiagonalen uten at de endrer seg. Et eksempel på en symmetrisk matrise er*

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

Eksempel 1.4.4. *La A være en* $n \times m$*-matrise. Da er* A^T *en* $m \times n$*-matrise og produktene* AA^T *og* A^TA *er kvadratiske matriser, henholdsvis av størrelse* $n \times n$ *og* $m \times m$*. Siden vi har*

$$(AA^T)^T = (A^T)^T A^T = AA^T$$

og tilsvarende for A^TA*, vil begge produktene være symmetriske matriser.*

Den transponerte av en matrise har noen egenskaper som minner om egenskapene til den inverse. På samme måte som at $(A^T)^T = A$ har vi $(A^{-1})^{-1} = A$ og produktformelen $(AB)^T = B^TA^T$ stemmer også overens med produktformelen for inverser: $(AB)^{-1} = B^{-1}A^{-1}$. Det leder oss til den neste definisjonen

Definisjon 1.4.4. *En invertibel kvadratisk matrise A som oppfyller* $A^{-1} = A^T$ *kalles en* ***ortogonal*** *matrise.*

Eksempel 1.4.5. *Matrisen*

$$A = \begin{pmatrix} 0 & -0.8 & -0.6 \\ 0.8 & -0.36 & 0.48 \\ 0.6 & 0.48 & -0.64 \end{pmatrix}$$

er ortogonal. Spesielt har vi

$$\begin{pmatrix} 0 & -0.8 & -0.6 \\ 0.8 & -0.36 & 0.48 \\ 0.6 & 0.48 & -0.64 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0.8 & 0.6 \\ -0.8 & -0.36 & 0.48 \\ -0.6 & 0.48 & -0.64 \end{pmatrix} = I$$

Grunnen til at slike matriser kalles ortogonale er at de har følgende viktige egenskap;

Proposisjon 1.4.5. *Skalarproduktet av to forskjellige søyler i en ortogonal matrise er 0, mens skalarproduktet av en søyle med seg selv er 1.*

Bevis. Siden den inverse er den transponerte, så har vi $A \cdot A^T = I_n$. Hvis vi kaller søylene i A for $\mathbf{a}_1, \ldots, \mathbf{a}_n$, så sier denne betingelsen at $\mathbf{a}_i \cdot \mathbf{a}_j = 0$ for $i \neq j$, og $\mathbf{a}_i \cdot \mathbf{a}_i = 1$. □
Så en kvadratisk matrise kalles ortogonal fordi søylene i matrisen er ortogonale, dvs. står normalt på hverandre. I tillegg har de alle lengde 1.

Proposisjon 1.4.6. *Produktet av to ortogonale matriser er en ortogonal matrise.*

Bevis. La A og B være to ortogonale matriser. Da har vi

$$(AB)^{-1} = B^{-1}A^{-1} = B^T A^T = (AB)^T$$

□

Mange ulike typer av kvadratiske matriser har fått egne navn, siden de spiller sentrale roller i forskjellige sammenhenger.

Definisjon 1.4.7. *En kvadratisk matrise A sies å være* ***øvre triangulær*** *dersom alle elementene under hoveddiagonalen er 0. Matrisen er* ***nedre triangulær*** *dersom alle elementene over hoveddiagonalen er 0. En kvadratisk matrise som er både øvre og nedre triangulær (dvs. alle elementene bortsatt fra de på diagonalen er 0) kalles en* ***diagonal*** *matrise.*

Eksempel 1.4.6. *Matrisen*

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

er et eksempel på en øvre triangulær matrise, mens

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

er et eksempel på en nedre triangulær matrise.

Et eksempel på en diagonal matrise er

$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

Det finnes metoder og formler for å regne ut den inverse til en matrise. Den enkleste formelen er naturlig nok for en 2×2-matrise.

Eksempel 1.4.7. *For en* 2×2*-matrise*

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

har vi den inverse matrisen

$$A^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$$

under forutsetning av at $a_{11}a_{22} - a_{12}a_{21} \neq 0$*. Hvis* $a_{11}a_{22} - a_{12}a_{21} = 0$ *har ikke matrisen noen invers.*

Eksempel 1.4.8. *Vi skal regne ut inversen til* 2×2*-matrisen*

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

Vi bruker formelen over og får

$$A^{-1} = \frac{1}{1 \cdot 1 - 2(-2)} \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix} = \frac{1}{5} \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}$$

Vi kan kontrollere svaret ved å regne ut

$$\begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} \cdot \frac{1}{5} \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Ved å bruke matriseformen for prikkprodukt kan vi definere en stor klasse av ortogonale matriser, omtalt som Householder-transformasjoner. Householder-transformasjoner er oppkalt etter Alston Scott Householder, som introduserte denne konstruksjonen i 1958;

Definisjon 1.4.8. *La* $\mathbf{v}$ *være en ikke-triviell vektor (Ikke-triviell betyr at den er forskjellig fra 0-vektoren). Vi kaller matrisen*

$$Q_{\mathbf{v}} = I - 2\frac{\mathbf{v}\mathbf{v}^T}{\mathbf{v}^T\mathbf{v}}$$

for en ***Householder****-transformasjon. Merk at telleren i brøken er en* $n \times n$*-matrise, mens nevneren er en* 1×1*- matrise (som vi identifiserer med et tall).*

Uttrykket i denne definisjonen kan se komplisert ut, men det er ikke så vanskelig å regne ut.

Eksempel 1.4.9. *La*

$$\mathbf{v} = \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix}$$

Da har vi

$$\mathbf{v}^T\mathbf{v} = (2^2 + 1^2 + (-2)^2) = 9$$

og

$$Q_{\mathbf{v}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \frac{2}{9} \begin{pmatrix} 2 \cdot 2 & 2 \cdot 1 & 2 \cdot (-2) \\ 2 & 1 & -2 \\ -2 \cdot 2 & -2 \cdot 1 & -2 \cdot (-2) \end{pmatrix} = \frac{1}{9} \begin{pmatrix} 1 & -4 & 8 \\ -4 & 7 & 4 \\ 8 & 4 & 1 \end{pmatrix}$$

Proposisjon 1.4.9. *Householder-transformasjonen $Q_{\mathbf{v}}$ er en ortogonal matrise.*

Bevis. Vi har

$$Q_{\mathbf{v}}^T = I^T - 2(\frac{\mathbf{v}\mathbf{v}^T}{\mathbf{v}^T\mathbf{v}})^T = I - 2\frac{\mathbf{v}\mathbf{v}^T}{\mathbf{v}^T\mathbf{v}} = Q_{\mathbf{v}}$$

Samtidig har vi

$$Q_{\mathbf{v}}^2 = (I - 2\frac{\mathbf{v}\mathbf{v}^T}{\mathbf{v}^T\mathbf{v}})^2 = I - 4\frac{\mathbf{v}\mathbf{v}^T}{\mathbf{v}^T\mathbf{v}} + 4(\frac{\mathbf{v}\mathbf{v}^T}{\mathbf{v}^T\mathbf{v}})^2 = I$$

siden vi har

$$(\frac{\mathbf{v}\mathbf{v}^T}{\mathbf{v}^T\mathbf{v}})^2 = \frac{\mathbf{v}\mathbf{v}^T\mathbf{v}\mathbf{v}^T}{\mathbf{v}^T\mathbf{v}\mathbf{v}^T\mathbf{v}} = \frac{\mathbf{v}^T\mathbf{v}}{\mathbf{v}^T\mathbf{v}} \cdot \frac{\mathbf{v}\mathbf{v}^T}{\mathbf{v}^T\mathbf{v}} = \frac{\mathbf{v}\mathbf{v}^T}{\mathbf{v}^T\mathbf{v}}$$

□

Vi ser av beviset at Householder-transformasjonen ikke bare er ortogonal, men den oppfyller i tillegg at $Q_{\mathbf{v}}^2 = I$. Householder-transformasjoner har en spesiell geometrisk betydning. Dette kommer vi tilbake til i et senere kapittel.

1.5 Lineære likningssytemer

Når man brenner metangass sammen med oksygen, vil sluttproduktene være karbondioksyd og vann, i tillegg til frigjort energi. Reaksjonslikningen for denne reaksjonen er

$$CH_4 + O_2 \to CO_2 + H_2O$$

Denne likningen er ikke balansert, dvs. den tar ikke hensyn til at antall atomer må være konstant i reaksjonen. Vår oppgave er å balansere likningen. La nå x_1, x_2, x_3 og x_4 være antall molekyler som inngår i reaksjonslikningen;

$$x_1CH_4 + x_2O_2 \to x_3CO_2 + x_4H_2O$$

Vi teller opp antallene av de tre atomene, C, O og H som inngår i likningen og kravet om at antallene skal være bevart i reaksjonen gir oss likningssystemet

$$\begin{aligned} x_1 &= x_3 \\ 2x_2 &= 2x_3 + x_4 \\ 4x_1 &= 2x_4 \end{aligned}$$

Vi kan rydde dette systemet for å få det på en mer strukturert form;

$$\begin{aligned} x_1 \qquad - x_3 \qquad &= 0 \\ 2x_2 - 2x_3 - x_4 &= 0 \\ 4x_1 \qquad - 2x_4 &= 0 \end{aligned}$$

Dette er et homogent likningssystem med 3 likninger og 4 ukjente. Systemet er underbestemt og har mange løsninger. En av løsningene av systemet kan være $x_1 = x_3 = 1$, $x_2 = x_4 = 2$, som gir reaksjonslikningen

$$CH_4 + 2O_2 \to CO_2 + 2H_2O$$

En annen løsning vil være

$$3CH_4 + 6O_2 \rightarrow 3CO_2 + 6H_2O$$

eller et hvilket som helst multiplum av den enkleste reaksjonslikningen. Selv om dette gir oss en frihetsgrad i løsningen er vi ikke så interessert i de andre løsningene i dette tilfellet, den enkle likningen gir oss alt vi trenger. I et matematisk språk vil vi si at løsningen utgjør en **basis** for **løsningsrommet**.

I dette kapitlet skal vi studere lineære likningssystemer og se på metoder for å finne løsningene.

Definisjon 1.5.1. *Et sett av lineære likninger*

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

*kalles et **lineært likningssystem** i n variable. Dersom alle $b_i = 0$ sier vi at likningssystemet er **homogent**. I motsatt fall er det **inhomogent**.*

Vi kan skrive et lineært likningssystem på matriseform;

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

I hvilken grad vi kan finne løsninger til et lineært likningssystem avhenger av forholdet mellom antall ukjente og antall likninger. Vi tar utgagnspunkt i at løsningene $\mathbf{x} = (x_1, \dots, x_n)$ har n variable som vi kan velge fritt. Dette omtales ofte som at vi har n **frihetsgrader** eller n **ukjente**. Hver likning gir oss en begrensning på hvilke $\mathbf{x}$ vi kan velge, og reduserer dermed antall frihetsgrader med 1. Det betyr at med m likninger, så står vi igjen med $n - m$ frihetsgrader, under forutsetning av at dette tallet er større enn eller lik 0. Et system som har flere likninger enn frihetsgrader sies å være **overbestemt**. Det systemet vil normalt ikke ha noen løsninger. Det motsatte tilfellet kalles et **underbestemt** system. Da har vi flere frihetsgrader enn likninger, og vi vil normalt ha mange løsninger. Men et underbestemt system trenger ikke å ha noen løsninger. F.eks. vil ikke systemet

$$\begin{aligned} x + y + z &= 2 \\ x + y + z &= 3 \end{aligned}$$

ha noen løsninger, selv om vi har 3 ukjente og kun 2 likninger. På den annen side kan et tilsynelatende overbestemt system likevel ha løsninger hvis flere av likningene er like eller eventuelt summer av hverandre. F.eks. er ikke systemet

$$\begin{aligned} x + y &= 2 \\ x - y &= 0 \\ 3x + y &= 4 \end{aligned}$$

overbestemt, på tross av at vi har flere likninger enn ukjente. Den siste likningen $3x+y=4$ kan skrives som 2 ganger den første likningen pluss den andre. Det betyr at den siste likningen ikke gir oss noe mer informasjon om løsningen enn vi allerede har fra de to første likningene, og derfor heller ikke noen flere begrensniger på hvilke løsninger som er mulig. Likningssystemet er derfor **ekvivalent** med et system med 2 ukjente og 2 likninger.

For å få bedre oversikt over mulige løsninger av et lineært likningssystem finnes det en metode for å forenkle likningssystemet. Med utgangspunkt i likningssystemet

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

kan vi skrive opp den **utvidede koeffisientmatrisen**

$$\begin{pmatrix} a_{11} & \dots & a_{1n} & b_n \\ \vdots & & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} & b_m \end{pmatrix}$$

som består av alle koeffisientene i likningssystemet med en ekstra søyle helt til høyre. Denne søylen inneholder alle tallene til høyre for likhetstegnene. Verktøyet vi skal bruke i denne metoden kalles **elementære radoperasjoner**.

Definisjon 1.5.2. *Det finnes 3 ulike typer elementære radoperasjoner som kan utføres på et likningssystem eller systemets utvidede koeffisientmatrise:*

i) Multiplisere en likning (rad) med et reelt tall.

ii) Bytte om to likninger (rader).

iii) Legge et multiplum av en likning (rad) til en annen likning (rad).

Proposisjon 1.5.3. *Løsningene til et lineært likningssystem endres ikke av elementære radoperasjoner.*

Bevis. At de to første ikke endrer på løsningene er nokså opplagt. Det er klart at dersom $(x_1,\dots,x_n)$ er en løsning av det opprinnelige likningssystemet, så vil $(x_1,\dots,x_n)$ også vaere en løsning etter at vi har lagt et multiplum av en likning til en annen. Anta nå at $(y_1,\dots,y_n)$ er en løsning av det nye systemet. Nå kan vi komme tilbake til det opprinnelige systemet ved å gjøre den samme radoperasjonen, med den lille forskjellen at vi trekker fra i stedet for å legge til. Dermed ender vi opp med de samme løsningene for de to systemene. □

Eksempel 1.5.1. *Vi har gitt et likningssystem*

$$\begin{aligned} x_1 + 2x_2 - x_3 &= 1 \\ -x_1 + 3x_2 + x_3 &= 4 \\ 2x_1 - x_2 + x_3 &= 3 \end{aligned}$$

Vi starter med å legge -2 ganger den første likningen til den tredje og får

$$\begin{aligned} x_1 + 2x_2 - x_3 &= 1 \\ -x_1 + 3x_2 + x_3 &= 4 \\ -5x_2 + 3x_3 &= 1 \end{aligned}$$

(merk at vi beholder den første likingen intakt). Dernest legger vi den første likningen til den andre og får

$$\begin{aligned} x_1 + 2x_2 - x_3 &= 1 \\ 5x_2 &= 5 \\ -5x_2 + 3x_3 &= 1 \end{aligned}$$

Nå har vi flere muligheter, men vi velger å legge den andre likningen til den tredje og likningssystemet blir da

$$\begin{aligned} x_1 + 2x_2 - x_3 &= 1 \\ 5x_2 &= 5 \\ 3x_3 &= 6 \end{aligned}$$

Nå er det naturlig å bruke den andre elementære operasjonen, første multiplisere andre rad med $\frac{1}{5}$*:*

$$\begin{aligned} x_1 + 2x_2 - x_3 &= 1 \\ x_2 &= 1 \\ 3x_3 &= 6 \end{aligned}$$

og følge raskt opp med å multiplisere den tredje raden med $\frac{1}{3}$*:*

$$\begin{aligned} x_1 + 2x_2 - x_3 &= 1 \\ x_2 &= 1 \\ x_3 &= 2 \end{aligned}$$

Herfra er det enkelt å se at løsningen av likningssystemet blir $x_1 = 1$*,* $x_2 = 1$ *og* $x_3 = 2$*. Hvis vi skal holde oss strengt til radoperasjoner vil de to siste skrittene være å trekke 2 ganger andre likning fra den første og deretter legge den siste likningen til den første. Det ville gitt oss systemet*

$$\begin{aligned} x_1 &= 1 \\ x_2 &= 1 \\ x_3 &= 2 \end{aligned}$$

Vi kan sette løsningen inn i det opprinnelige systemet og se at det stemmer.

Vi kan gjenta denne prosessen utført på den utvidede koeffisientmatrisen i stedet for likningssystemet:

$$\begin{pmatrix} 1 & 2 & -1 & 1 \\ -1 & 3 & 1 & 4 \\ 2 & -1 & 1 & 3 \end{pmatrix} \overset{III-2\cdot I}{\rightsquigarrow} \begin{pmatrix} 1 & 2 & -1 & 1 \\ -1 & 3 & 1 & 4 \\ 0 & -5 & 3 & 1 \end{pmatrix} \overset{II+1\cdot I}{\rightsquigarrow} \begin{pmatrix} 1 & 2 & -1 & 1 \\ 0 & 5 & 0 & 5 \\ 0 & -5 & 3 & 1 \end{pmatrix}$$

$$\overset{III+1\cdot II}{\rightsquigarrow}\begin{pmatrix}1&2&-1&1\\0&5&0&5\\0&0&3&6\end{pmatrix}\overset{\frac{1}{5}II}{\rightsquigarrow}\begin{pmatrix}1&2&-1&1\\0&1&0&1\\0&0&3&6\end{pmatrix}\overset{\frac{1}{3}III}{\rightsquigarrow}\begin{pmatrix}1&2&-1&1\\0&1&0&1\\0&0&1&2\end{pmatrix}$$

$$\overset{I-2\cdot II}{\rightsquigarrow}\begin{pmatrix}1&0&-1&-1\\0&1&0&1\\0&0&1&2\end{pmatrix}\overset{I+1\cdot III}{\rightsquigarrow}\begin{pmatrix}1&0&0&1\\0&1&0&1\\0&0&1&2\end{pmatrix}$$

hvor vi med $\overset{I+1\cdot III}{\rightsquigarrow}$ mener at vi her har utført radoperasjonen som legger 1 ganger den tredje raden til den første raden. Ombytte av rader skriver vi som $I \rightsquigarrow II$ og multiplikasjon av en rad som $II \rightsquigarrow 5 \cdot II$.

Definisjon 1.5.4. *Vi sier at et lineært likningssystem på formen*

$$\begin{aligned}a_{11}x_1+a_{12}x_2+a_{13}x_3+\cdots+a_{1n}x_n&=b_1\\a_{22}x_2+a_{23}x_3+\cdots+a_{2n}x_n&=b_2\\a_{23}x_3+\cdots+a_{2n}x_n&=b_2\\&\vdots\\a_{mn}x_n&=b_m\end{aligned}$$

*er på **trappeform**, og at systemet*

$$\begin{aligned}a_{11}x_1&=b_1\\a_{22}x_2&=b_2\\a_{23}x_3&=b_2\\&\vdots\\a_{mn}x_n&=b_m\end{aligned}$$

*er på **redusert trappeform**.*

Ved å bringe systemet over på trappeform er det enkelt å lese av om systemet er over- eller underbestemt, eller om det har en entydig løsning. Og som vi har sett, også å finne løsningen. Trappeformen til systemet

$$\begin{aligned}x+y+z&=2\\x+y+z&=3\end{aligned}$$

vil være

$$\begin{aligned}x+y+z&=2\\0&=3\end{aligned}$$

som åpenbart ikke har noen løsning, mens systemet

$$\begin{aligned}x+y&=2\\x-y&=0\\3x+y&=4\end{aligned}$$

som på trappeform ser ut som

$$\begin{aligned} x+y &= 2 \\ -2y &= -2 \\ -2y &= -2 \end{aligned}$$

opplagt har en løsning, selv om det i utgangspunktet var 3 likninger med 2 ukjente. Et annet eksempel er likningssystemet

$$\begin{aligned} x+2y-z &= 2 \\ -x+3y-2z &= -1 \\ 2x-y+3z &= 3 \end{aligned}$$

som har trappeform

$$\begin{aligned} x+2y-z &= 2 \\ 5y-3z &= 1 \\ 0 &= 0 \end{aligned}$$

Her vil vi ha mange løsninger, f.eks. kan vi velge z fritt og få $x = \frac{8-z}{5}$ og $y = \frac{3z+1}{5}$.

Likningssystemer med like mange frihetsgrader som likninger kaller vi for **kvadratiske** likningssystemer. Kvadratiske likningssystemer kan være underbestemt, **entydig bestemt** eller overbestemt, avhengig av hvordan likningene ser ut. Entydig bestemt betyr at likningssystemet har en entydig løsning.

Likningssystemet

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

har **koeffisientmatrise**

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$$

som er en kvadratisk matrise. Vektorene som inngår kan vi skrive

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{og} \mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

Det gir oss skriveformen $A\mathbf{x} = \mathbf{b}$ for likningssytemet.

Proposisjon 1.5.5. *La*

$$A\mathbf{x} = \mathbf{b}$$

være et likningsystem med n likninger og n ukjente, dvs. at A er en kvadratisk matrise. Dersom A er invertibel, så har likningssystemet en entydig løsning gitt ved

$$\mathbf{x} = A^{-1}\mathbf{b}$$

Bevis. Vi multipliserer likningen $A\mathbf{x} = \mathbf{b}$ med A^{-1} fra venstre. Det gir for venstresiden

$$A^{-1}(A\mathbf{x}) = (A^{-1}A)\mathbf{x} = I \cdot \mathbf{x} = \mathbf{x}$$

og for høyresiden $A^{-1}\mathbf{b}$. □

For å kunne utnytte resultatet i forrige proposisjon trenger vi å kunne regne ut inversen til en invertibel kvadratisk matrise. Igjen skal vi nyttiggjøre oss av de elementære radoperasjonene, og vi skal vise framgangsmåten på et eksempel. En 3×3-matrise A er gitt ved

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

Vi skal finne A^{-1}. Ideen er å utvide 3×3-matrisen til en 3×6-matrise hvor den høyre halvparten er identitetsmatrisen (generelt utvider vi en $n \times n$-matrise til en $n \times 2n$-matrise.):

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

Så gjennomfører vi elementære radoperasjoner på denne matrisen helt til den venstre 3×3-delen er blitt identitetsmatrisen (hvis det er mulig). Den høyre 3×3-delen vil da være den inverse matrisen A^{-1}:

$$\begin{pmatrix} 1 & 2 & -1 & 1 & 0 & 0 \\ -1 & 3 & 1 & 0 & 1 & 0 \\ 2 & -1 & 1 & 0 & 0 & 1 \end{pmatrix} \overset{III-2\cdot I}{\rightsquigarrow} \begin{pmatrix} 1 & 2 & -1 & 1 & 0 & 0 \\ -1 & 3 & 1 & 0 & 1 & 0 \\ 0 & -5 & 3 & -2 & 0 & 1 \end{pmatrix}$$

$$\overset{II+1\cdot I}{\rightsquigarrow} \begin{pmatrix} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & 5 & 0 & 1 & 1 & 0 \\ 0 & -5 & 3 & -2 & 0 & 1 \end{pmatrix} \overset{III+1\cdot II}{\rightsquigarrow} \begin{pmatrix} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & 5 & 0 & 1 & 1 & 0 \\ 0 & 0 & 3 & -1 & 1 & 1 \end{pmatrix}$$

$$\overset{\frac{1}{5}II}{\rightsquigarrow} \begin{pmatrix} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & 1 & 0 & \frac{1}{5} & \frac{1}{5} & 0 \\ 0 & -5 & 3 & -2 & 0 & 1 \end{pmatrix} \overset{\frac{1}{3}III}{\rightsquigarrow} \begin{pmatrix} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & 1 & 0 & \frac{1}{5} & \frac{1}{5} & 0 \\ 0 & 0 & 1 & -\frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$$

$$\overset{I-2\cdot II}{\rightsquigarrow} \begin{pmatrix} 1 & 0 & -1 & \frac{3}{5} & -\frac{2}{5} & 0 \\ 0 & 1 & 0 & \frac{1}{5} & \frac{1}{5} & 0 \\ 0 & 0 & 1 & -\frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix} \overset{I+1\cdot III}{\rightsquigarrow} \begin{pmatrix} 1 & 0 & 0 & \frac{4}{15} & -\frac{1}{15} & \frac{1}{3} \\ 0 & 1 & 0 & \frac{1}{5} & \frac{1}{5} & 0 \\ 0 & 0 & 1 & -\frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$$

som betyr at

$$A^{-1} = \begin{pmatrix} \frac{4}{15} & -\frac{1}{15} & \frac{1}{3} \\ \frac{1}{5} & \frac{1}{5} & 0 \\ -\frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix} = \frac{1}{15} \begin{pmatrix} 4 & -1 & 5 \\ 3 & 3 & 0 \\ -5 & 5 & 5 \end{pmatrix}$$

En alternativ skrivemåte for elementære radoperasjoner er som multiplikasjon fra venstre med elementære matriser. Det finnes tre typer av elementære matriser, svarende til de tre typene av elementære radoperasjoner. Elementære matriser er kvadratiske, invertible og de kan ha alle størrelser. I stedet for å skrive opp en geenrell definisjon skal vi gi eksempler på alle tre typer for 3×3-matriser. Alle matrisene tar utgangspunkt i identitetsmatrisen

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

Den første operasjonen er å bytte om på to rader. Dette svarer til multiplikasjon med en matrise der vi har byttet om på de tilsvrende radene i identitetsmatrisen. F.eks. vil ombytting av rad 1 og 2 svare til multiplikasjon fra venstre med matrisen

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

Den neste operasjonen er å multiplisere en rad med et fast tall. Dette erstatter vi med multiplikasjon med matrisen

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

hvis oppdraget var å multiplisere tredje rad med tallet k. Til slutt har vi den tredje operasjonen, hvor vi legger til et multiplum av en rad til en annen. Vi kan f.eks. se på $\overset{III-2\cdot I}{\rightsquigarrow}$, dvs. at 2 ganger første rad trekkes fra tredje rad. Den elementære matrisen som uttrykker dette finner vi ved å gjennomføre radoperasjonen på identitetsmatrisen;

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \overset{III-2\cdot I}{\rightsquigarrow} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix}$$

Radoperasjonen

$$\begin{pmatrix} 1 & 2 & -1 & 1 & 0 & 0 \\ -1 & 3 & 1 & 0 & 1 & 0 \\ 2 & -1 & 1 & 0 & 0 & 1 \end{pmatrix} \overset{III-2\cdot I}{\rightsquigarrow} \begin{pmatrix} 1 & 2 & -1 & 1 & 0 & 0 \\ -1 & 3 & 1 & 0 & 1 & 0 \\ 0 & -5 & 3 & -2 & 0 & 1 \end{pmatrix}$$

vil da erstattes av

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & -1 & 1 & 0 & 0 \\ -1 & 3 & 1 & 0 & 1 & 0 \\ 2 & -1 & 1 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & -1 & 1 & 0 & 0 \\ -1 & 3 & 1 & 0 & 1 & 0 \\ 0 & -5 & 3 & -2 & 0 & 1 \end{pmatrix}$$

Vi kan bruke denne teknikken til å vise at vår konstruksjon av inverse matriser gir rett resultat. Vi skriver for enkelthets skyld

$$\begin{pmatrix} 1 & 2 & -1 & 1 & 0 & 0 \\ -1 & 3 & 1 & 0 & 1 & 0 \\ 2 & -1 & 1 & 0 & 0 & 1 \end{pmatrix} = (A;I)$$

og det er enekelt å vise at multiplikasjon fra venstre er gitt ved $M(A;B) = (MA;MB)$. Anta nå at

$$E_r E_{r-1} \dots E_2 E_1 A = I$$

hvor E_1, E_2, …, E_{r-1}, E_r svarer til de elementære radoperasjonene som skal til for å omdanne A til identitetsmatrisen. Det gir

$$E_r E_{r-1} \dots E_2 E_1 (A;I) = (I; E_r E_{r-1} \dots E_2 E_1)$$

Men $E_r E_{r-1} \dots E_2 E_1 A = I$ betyr også at $E_r E_{r-1} \dots E_2 E_1 = A^{-1}$.

1.6 Determinanter

Dersom vi anvender metodene fra forrige kapittel på det lineære likningssystemet

$$a_{11}x + a_{12}y = b_1$$
$$a_{21}x + a_{22}y = b_2$$

finner vi at systemet har løsning gitt ved

$$(a_{11}a_{22} - a_{12}a_{21})x = a_{22}b_1 - a_{12}b_2$$
$$(a_{11}a_{22} - a_{12}a_{21})y = -a_{21}b_1 + a_{11}b_2$$

under forutsetning av at $a_{11}a_{22} - a_{12}a_{21} \neq 0$. Det samme uttrykket dukket opp i formelen for den inverse til 2×2-matrisen

$$A = \begin{pmatrix} a_{11} & a_{11} \\ a_{21} & a_{22} \end{pmatrix}$$

gitt ved

$$A^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$$

hvor det også er en forutsetning at $a_{11}a_{22} - a_{12}a_{21} \neq 0$ for at vi har en invers matrise. Uttrykket er med andre ord en viktig størrelse og har derfor fått en egen betegnelse, **determinanten** til matrisen A.

Definisjon 1.6.1. *La*

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

være en 2×2*-matrise. Da er* ***determinanten*** $\det(A)$ *til A gitt ved*

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

Eksempel 1.6.1.

$$\det \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} = 2 \cdot 2 - 1 \cdot 1 = 3$$

Med denne definisjonen kan vi skrive opp den inverse til en 2×2-matrise $A = (a_{ij})$ på en alternativ form;

Proposisjon 1.6.2. *La A være en invertibel* 2×2*-matrise. Da har vi*

$$A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$$

Det er ikke bare 2×2-matriser som har determinanter, alle kvadratiske matriser har determinanter. For n større enn 2 er det ikke like enkelt å regne ut determinanten til en matrise, men det finnes en grei rekursjonsformel. Først tar vi den formelle definisjonen (eller Leibniz' formel, som den også kalles).

Definisjon 1.6.3. *La*

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$$

være en $n \times n$-matrise. Da er determinanten til A gitt ved

$$\det(A) = \sum_{\sigma \in S_n} (-1)^{sgn(\sigma)} \prod_{i=1}^{n} a_{i,\sigma(i)}$$

Her betyr S_n mengden av alle ombyttinger (eller permutasjoner) av mengden $\{1,2,\dots,n\}$ og $sgn(\sigma)$ teller hvor mange ombyttinger av to elementer som trengs for å få til ombyttingen σ. Definisjonen sier altså at for å beregne determinanten til en matrise skal vi ta alle mulige kombinasjoner av å plukke ut ett element fra hver rad, og samtidig fra hver søyle, gange dem sammen, sette pluss eller minus foran avhengig av hvilken rekkefølge elementene er valgt ut, for så å legge alle disse produktene sammen.

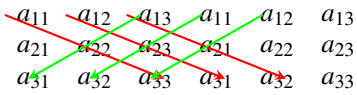
Eksempel 1.6.2. *For en 3×3-matrise*

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

har vi

$$\det(A) = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{32}a_{21} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}$$

Det er en enkel huskeregel for denne formelen for 3×3-determinanter. Vi skriver to kopier av matrisen ved siden av hverandre (vi trenger strengt tatt ikke den tredje søylen). Vi ganger sammen tallene langs de 6 pilene og summerer med positivt fortegn langs de røde og negativt fortegn langs de grønne.



Eksempel 1.6.3.

$$\det \begin{pmatrix} 2 & 0 & -1 \\ 1 & 1 & 0 \\ -2 & -1 & 1 \end{pmatrix} = \begin{array}{ccccc} 2 & 0 & -1 & 2 & 0 \\ 1 & 1 & 0 & 1 & 1 \\ -2 & -1 & 1 & -2 & -1 \end{array}$$

$$\begin{aligned} &= 2 \cdot 1 \cdot 1 + 0 \cdot 0 \cdot (-2) + (-1) \cdot 1 \cdot (-1) \\ &\quad - (-1) \cdot 1 \cdot (-2) - 2 \cdot 0 \cdot (-1) - 0 \cdot 1 \cdot 1 \\ &= 2 + 0 + 1 - 2 - 0 - 0 = 1 \end{aligned}$$

For store matriser er det tungvint å regne ut determinanten på denne måten. Vi har derfor en alternativ definisjon (Laplaces formel) for determinanten, gitt ved en **rekursiv formel**:

Proposisjon 1.6.4. *La*

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$$

være en $n \times n$-matrise. Da er determinanten til A gitt ved

$$\det(A) = \sum_{j=1}^{n} (-1)^{j-1} a_{j1} \det(A_j)$$

der A_j er $(n-1) \times (n-1)$-matrisen som framkommer ved stryke første søyle og j-te rad i matrisen A, dvs.

$$A_j = \begin{pmatrix} a_{12} & a_{13} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{j-1,2} & a_{j-1,3} & \dots & a_{j-1,n} \\ a_{j+1,2} & a_{j+1,3} & \dots & a_{j+1,n} \\ \vdots & \vdots & & \vdots \\ a_{n2} & a_{n3} & \dots & a_{nn} \end{pmatrix}$$

Siden vi har en enkel formel for determinanten til en 2×2-matrise, kan vi bruke denne formelen rekursivt til å beregne determinanten til større matriser. Merk at vi også har en alternativ notasjon for deteminanter. For

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$$

bruker vi de to ekvivalente skriveformene;

$$\det(A) = \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix}$$

De buede parantesene i matrisenotasjonen byttes ut med rette streker for den tilhørende determinanten.

Eksempel 1.6.4. *Vi skal regne ut determinanten til 3×3-matrisen*

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

ved hjelp av rekursjonsformelen. Vi har

$$\begin{aligned} \det(A) &= 2\begin{vmatrix} 1 & 0 \\ -1 & 1 \end{vmatrix} - 1\begin{vmatrix} 0 & -1 \\ -1 & 1 \end{vmatrix} - 2\begin{vmatrix} 0 & -1 \\ 1 & 0 \end{vmatrix} \\ &= 2(1-0) - (0-1) - 2(0+1) = 1 \end{aligned}$$

I determinanten $\begin{vmatrix} 1 & 0 \\ -1 & 1 \end{vmatrix}$ har vi fjernet første søyle og første rad i matrisen A, i $\begin{vmatrix} 0 & -1 \\ -1 & 1 \end{vmatrix}$ har vi fjernet første søyle og andre rad og i $\begin{vmatrix} 0 & -1 \\ 1 & 0 \end{vmatrix}$ første søyle og tredje rad.

Det er noen nyttige regneregler knyttet til determinanter. Den kanskje aller viktigste er produktformelen.

Teorem 1.6.5. *La A og B være to kvadratiske matriser av samme størrelse. Da har vi*

$$\det(A \cdot B) = \det(A) \cdot \det(B)$$

Vi skal ikke vise det her, det involverer en god del begreper som vi ikke vil ha bruk for til annet enn å bevise formelen. Men vi kan skrive ut formelen i 2×2-tilfellet;

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$

Det gir at

$$A \cdot B = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix}$$

og

$$\begin{aligned} \det(A \cdot B) &= (a_{11}b_{11} + a_{12}b_{21})(a_{21}b_{12} + a_{22}b_{22}) - (a_{11}b_{12} + a_{12}b_{22})(a_{21}b_{11} + a_{22}b_{21}) \\ &= a_{11}b_{11}a_{21}b_{12} + a_{11}b_{11}a_{22}b_{22} + a_{12}b_{21}a_{21}b_{12} + a_{12}b_{21}a_{22}b_{22} \\ &\quad - a_{11}b_{12}a_{21}b_{11} - a_{12}b_{22}a_{22}b_{21} - a_{11}b_{12}a_{22}b_{21} - a_{12}b_{22}a_{21}b_{11} \\ &= a_{11}a_{22}b_{11}b_{22} - a_{11}a_{22}b_{12}b_{21} - a_{12}a_{21}b_{22}b_{11} + a_{12}a_{21}b_{21}b_{12} \\ &= (a_{11}a_{22} - a_{12}a_{21})(b_{11}b_{22} - b_{12}b_{21}) \\ &= \det(A) \cdot \det(B) \end{aligned}$$

Proposisjon 1.6.6. *Determinanten til en matrise med en rad eller søyle av 0-er er 0.*

Bevis. Siden alle leddene som inngår i determinanten inneholder en faktor fra hver rad og hver søyle, vil i dette tilfelle alle leddene inneholde en faktor 0, og determinanten er derfor 0. □

Eksempel 1.6.5.

$$\det \begin{pmatrix} 2 & 1 & 0 \\ 1 & -1 & 2 \\ 0 & 0 & 0 \end{pmatrix} = 0$$

Proposisjon 1.6.7. *Determinanten til en triangulær matrise er produktet av elementene på diagonalen.*

Bevis. Samtlige ledd som inngår i formelen for determinanten inneholder faktorer både fra oversiden og fra undersiden av diagonalen, bortsett fra ett. Det ene er produktet av elementene på diagonalen. I formelen inngår dette leddet alltid med positivt fortegn. □
Det følger at determinanten til identitetsmatrisen er 1, og determinanten til $k \cdot I$ er k^n.

Eksempel 1.6.6.

$$\det \begin{pmatrix} 2 & 1 & -3 \\ 0 & 3 & -2 \\ 0 & 0 & 1 \end{pmatrix} = 2 \cdot 3 \cdot 1 = 6$$

Proposisjon 1.6.8. *La A være en $n \times n$-matrise. Da har vi*

$$\det(k \cdot A) = k^n \cdot \det(A)$$

Bevis. Vi kan skrive

$$k \cdot A = (kI) \cdot A$$

Dermed følger resultatet fra kommentaren etter proposisjonen over. □

Proposisjon 1.6.9. *Determinanten til en matrise og dens transponerte er like.*

Bevis. Siden determinanten er satt sammen av alle produkter av ett element fra hver rad og hver søyle, kan vi bytte om rader og søyler og få de samme leddene. Fortegnene vi introduserer er laget slik at de ikke endrer seg om vi bytter om rader og søyler. □

Eksempel 1.6.7. *Vi har*

$$\det \begin{pmatrix} 2 & 0 & -1 \\ 1 & 1 & 0 \\ -2 & -1 & 1 \end{pmatrix} = 1$$

og for den transponerte

$$\det \begin{pmatrix} 2 & 1 & -2 \\ 0 & 1 & -1 \\ -1 & 0 & 1 \end{pmatrix} = 1$$

Proposisjon 1.6.10. *Bytter vi om to rader eller to søyler i en matrise skifter determinanten tegn.*

Bevis. Bytter vi om to rader (eller to søyler) vil alle leddene i determinanten bli de samme, men alle tegnene vil endre seg. Igjen bruker vi det at fortegnene er satt slik at determinanten ved denne prosessen vil skifte tegn. □

Eksempel 1.6.8. *Vi har*

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

mens

$$\begin{vmatrix} a_{21} & a_{22} \\ a_{11} & a_{12} \end{vmatrix} = a_{21}a_{12} - a_{11}a_{22} = -(a_{11}a_{22} - a_{12}a_{21})$$

Eksempel 1.6.9. *Vi har*

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

mens

$$\det \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} = -3$$

Proposisjon 1.6.11. *Determinanten til den inverse matrisen er gitt ved*

$$\det(A^{-1}) = \frac{1}{\det(A)}$$

Bevis. Siden vi har $A \cdot A^{-1} = I$ følger det at

$$\det(A) \cdot \det(A^{-1}) = \det(I) = 1$$

□

Eksempel 1.6.10. *La*

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

Da har vi

$$\det(A) = 5$$

mens determinanten til den inverse matrisen er

$$\det(A^{-1}) = \det\begin{pmatrix} \frac{1}{5} & \frac{-2}{5} \\ \frac{2}{5} & \frac{1}{5} \end{pmatrix} = \frac{1}{5}^2 - \frac{2}{5}(-\frac{2}{5}) = \frac{1}{5}$$

Proposisjon 1.6.12. *Determinanten til en ortogonal matrise er* ± 1.

Bevis. En ortogonal matrise er definert ved at $A^{-1} = A^T$. Vi har vist at $\det(A^T) = \det(A)$ og siden vi samtidig har at $\det(A) \cdot \det(A^{-1}) = 1$ følger det at $\det(A)^2 = 1$. □

Eksempel 1.6.11.

$$\begin{vmatrix} 0 & -0.8 & -0.6 \\ 0.8 & -0.36 & 0.48 \\ 0.6 & 0.48 & -0.64 \end{vmatrix}$$
$$= -0.8 \cdot 0.48 \cdot 0.6 - 0.6 \cdot 0.8 \cdot 0.48 - 0.8 \cdot 0.8 \cdot 0.64 - 0.6 \cdot 0.36 \cdot 0.6$$
$$= -0.2304 - 0.2304 - 0.4096 - 0.1296 = -1$$

Ved å bruke produktformelen for determinanter følger det at en invertibel matrise må ha determinant ulik 0.

$$\det(A) \cdot \det(A^{-1}) = \det(I) = 1$$

Når produktet av to tall er 1, eller mer generelt ulik 0, er det ikke mulig at noen av tallene er lik 0. I så fall ville jo produktet av dem være 0. Faktisk er det motsatte også sant; dersom determinanten til en kvadratisk matrise er ulik 0, så er matrisen inverterbar. Vi skal ikke gå inn på beviset for dette resultatet her.

Proposisjon 1.6.13. *En kvadratisk matrise A er invertibel hvis og bare hvis determinanten* $\det(A) \neq 0$.

1.7 Spektralteori

Til en kvadratisk matrise finnes det alltid noen vektorer som er spesielt knyttet til matrisen. Disse vektorene har den nyttige egenskapen at når vi ganger dem med matrisen så endrer det ikke på forholdet mellom koordinatene til vektoren. Slike vektorer kalles **egenvektorer** for matrisen og det finnes en svært omfattende teori som tar for seg disse. I dette kapitlet skal vi se på den grunnleggende teorien for disse vektorene, en teori som ofte omtales med det klingende navnet **spektralteori**.

Definisjon 1.7.1. *Gitt en kvadratisk* $n \times n$*-matrise A. En vektor* $\mathbf{v} \neq \mathbf{0}$ *kalles en* ***egenvektor*** *for A dersom*

$$A \cdot \mathbf{v} = \lambda \cdot \mathbf{v}$$

for et reelt tall λ*. Tallet* λ *kalles for den tilhørende* ***egenverdien***.

For en egenvektor $\mathbf{v}$ vil også alle vektorer på formen $a\mathbf{v}$ for et reelt tall $a \in \mathbb{R}$, $a \neq 0$, være en egenvektor med samme egenverdi λ:

$$A(a\mathbf{v}) = aA\mathbf{v} = a\lambda\mathbf{v} = \lambda(a\mathbf{v})$$

Mengden av egenvektorer med samme egenverdi λ kalles **egenrommet** til egenverdien λ. (Rent formelt er ikke 0-vektoren en egenvektor, men det likevel hensiktsmessig å inkludere $\mathbf{0}$ i egenrommet til alle egenverdier). Det finnes åpenbart vektorer i egenrommet som er lineært avhengige, slik som vi har sett ovenfor (vektorene $\mathbf{v}$ og $a\mathbf{v}$ er lineært avhengige for alle valg av $a \neq 0$). Men det kan også være at egenrommet inneholder lineært uavhengige vektorer.

Eksempel 1.7.1. *En hvilken som helst vektor vil være en egenvektor for identitetsmatrisen I_n med egenverdi 1 siden*

$$I_n\mathbf{v} = 1 \cdot \mathbf{v}$$

Det betyr at egenrommet til I_n for egenverdien 1 er hele $\mathbb{R}^n$.

Anta at $\mathbf{v}$ er en egenvektor for en matrise A, med tilhørende egenverdi λ. Da har vi

$$A\mathbf{v} = \lambda\mathbf{v}$$

Alternativt kan vi skrive denne likningen som

$$A\mathbf{v} - \lambda\mathbf{v} = (A - \lambda I)\mathbf{v} = \mathbf{0}$$

Egenvektorer er alltid antatt å være forskjellige fra $\mathbf{0}$-vektoren. Det betyr at likningssytemet $(A - \lambda I)\mathbf{v} = \mathbf{0}$ skal ha en ikke-triviell løsning. Siden dette er et homogent likningsystem er det ekvivalent med at determinanten til koeffisientmatrisen er lik 0. Vi har i dette resonnementet kun antatt at λ er en egenverdi, uten å kjenne den noe nærmere. Vi betrakter derfor foreløpig λ som en ukjent og prøver å finne en løsning av likningen

$$\det(A - \lambda I) = \mathbf{0}$$

Dersom vi skriver dette ut ser vi at dette gir oss en polynomial likning i λ.

Definisjon 1.7.2. *La A være en kvadratisk matrise, og la λ være en ukjent størrelse. Polynomet*

$$\chi_A(\lambda) = \det(A - \lambda I)$$

*kalles det **karakteristiske polynomet** til matrisen A, og røttene til polynomet er egenverdiene til matrisen.*

Eksempel 1.7.2. *Det karakteristiske polynomet til matrisen*

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

er gitt ved

$$\begin{aligned}
\chi_A(\lambda) &= \det\left(\begin{pmatrix} 2 & 3 \\ -1 & 1 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right) \\
&= \det\begin{pmatrix} 2-\lambda & 3 \\ -1 & 1-\lambda \end{pmatrix} \\
&= (2-\lambda)(1-\lambda) - 3 \cdot (-1) \\
&= \lambda^2 - 3\lambda + 5
\end{aligned}$$

Det karakteristiske polynomet til en $n \times n$-matrise vil være et polynom av grad n i λ. Matrisen kan derfor ha inntil n egenverdier, de n røttene til det karakteristiske polylynomet.

Vi skal regne ut denne determinanten, og da får vi bruk for et begrep til.

Definisjon 1.7.3. *La A være en kvadratisk matrise;*

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$$

Da er ***trasen*** *(også kalt* ***sporet****) til A gitt som summen av diagonal-elementene*

$$\operatorname{tr}(A) = a_{11} + a_{22} + \cdots + a_{nn}$$

For en 2×2-matrise er det enkelt å se at det karakteristiske polynomet kan skrives

$$\chi_A(\lambda) = \lambda^2 - \operatorname{tr}(A)\lambda + \det(A)$$

Det faller direkte ut dersom vi skriver ut det karakteristiske polynomet på elementform. Vi har

$$\begin{aligned} \det(A - \lambda I) &= \begin{vmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{vmatrix} \\ &= \lambda^2 - (a_{11} + a_{22})\lambda + a_{11}a_{22} - a_{12}a_{21} \\ &= \lambda^2 - \operatorname{tr}(A)\lambda + \det(A) \end{aligned}$$

Eksempel 1.7.3. *Vi skal beregne egenverdiene til matrisen*

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

Vi har $\operatorname{tr}(A) = 2$ *og* $\det(A) = -3$. *Det gir karakteristisk polynom*

$$\chi_A(\lambda) = \lambda^2 - 2\lambda - 3$$

For å finne røttene bruker vi abc-formelen;

$$\lambda = \frac{2 \pm \sqrt{2^2 - 4(-3)}}{2} = 1 \pm 2$$

Egenverdiene blir $\lambda_1 = 3$ *og* $\lambda_2 = -1$.

Hvis vi bruker abc-formelen direkte på det generelle uttrykket $\lambda^2 - \operatorname{tr}(A)\lambda + \det(A) = 0$, får vi egenverdiene gitt ved

$$\lambda = \frac{\operatorname{tr}(A) \pm \sqrt{\operatorname{tr}(A)^2 - 4\det(A)}}{2}$$

For større matriser medfører det litt mer regning å finne egenverdiene. For å gi en formel i 3×3-tilfellet trenger vi en mellomregning.

La A være en 3×3-matrise. Utregning gir

$$\begin{aligned} \frac{1}{2}(\operatorname{tr}(A)^2 - \operatorname{tr}(A^2)) = a_{11}a_{22} + a_{11}a_{33} + a_{22}a_{33} \\ - a_{12}a_{21} - a_{13}a_{31} - a_{23}a_{32} \end{aligned}$$

Proposisjon 1.7.4. *La A være en 3×3-matrise. Da har vi*

$$\chi_A(\lambda) = -\lambda^3 + \operatorname{tr}(A)\lambda^2 - \frac{1}{2}((\operatorname{tr}(A))^2 - \operatorname{tr}(A^2))\lambda + \det(A)$$

Bevis. Beviset er en ren utregning;

$$\begin{aligned}
\chi_A(\lambda) = \det(A-\lambda I) &= \begin{vmatrix} a_{11}-\lambda & a_{12} & a_{13} \\ a_{21} & a_{22}-\lambda & a_{23} \\ a_{31} & a_{32} & a_{33}-\lambda \end{vmatrix} \\
&= (a_{11}-\lambda)(a_{22}-\lambda)(a_{33}-\lambda) + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\
&\quad - (a_{11}-\lambda)a_{23}a_{32} - (a_{22}-\lambda)a_{13}a_{31} - (a_{33}-\lambda)a_{12}a_{21} \\
&= -\lambda^3 + (a_{11}+a_{22}+a_{33})\lambda^2 - (a_{11}a_{22}+a_{11}a_{33}+a_{22}a_{33}-a_{12}a_{21}-a_{13}a_{31}-a_{23}a_{32})\lambda \\
&\quad + (a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{23}a_{32}-a_{11}a_{23}a_{32}-a_{22}a_{13}a_{31}-a_{33}a_{12}a_{21})
\end{aligned}$$

som ved innsetting gir formelen i proposisjonen. □

Eksempel 1.7.4. *Vi skal regne ut det karakteristiske polynomet til*

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

Vi har

$$A^2 = \begin{pmatrix} -1 & -2 & -2 \\ -2 & -1 & -2 \\ 1 & 1 & -4 \end{pmatrix}$$

som gir $\operatorname{tr}(A) = 2$, $\operatorname{tr}(A^2) = -6$ *og* $\det(A) = 4$. *Det gir*

$$\begin{aligned}
\chi_A(\lambda) &= -\lambda^3 + 2\lambda^2 - \frac{1}{2}(2^2 - (-6))\lambda + 4 \\
&= -\lambda^3 + 2\lambda^2 - 5\lambda + 4
\end{aligned}$$

Vi ser at $\lambda = 1$ *er en løsning av likningen* $\chi_A(\lambda) = 0$. *Polynomdivisjon gir*

$$\chi_A(\lambda) = -(\lambda-1)(\lambda^2 - \lambda + 4)$$

og abc-formelen gir at vi i dette tilfellet ikke får noen flere reelle egenverdier.

Vi kan bruke det karakteristiske polynomet til å studere egenskaper ved ulike matriser.

Proposisjon 1.7.5. *La A være en invertibel $n\times n$-matrise, og $\mathbf{v}$ en egenvektor med tilhørende egenverdi λ. Da er $\mathbf{v}$ også egenvektor for A^{-1} med egenverdi $\frac{1}{\lambda}$.*

Bevis. Vi har $A\mathbf{v} = \lambda\mathbf{v}$. Siden A er invertibel kan vi multiplisere med A^{-1} på begge sider av likhetstegnet, og vi får $\mathbf{v} = A^{-1}\lambda\mathbf{v}$. Deler vi begge sider med λ får vi

$$A^{-1}\mathbf{v} = \frac{1}{\lambda}\mathbf{v}$$

□

Proposisjon 1.7.6. *En kvadratisk matrise og dens transponerte har samme karakteristiske polynom og derfor de samme egenverdiene.*

Merk at de to matrisene har samme egenverdier, vi sier ikke noe om egenvektorene.
Bevis. Siden en matrise og dens transponerte har samme determinant følger det at

$$\chi_{A^T}(\lambda) = \det(A^T - \lambda I) = \det(A^T - \lambda I)^T = \det((A^T)^T - \lambda I^T) = \det(A - \lambda I) = \chi_A(\lambda)$$

□

Proposisjon 1.7.7. *Alle (de reelle) egenverdiene til en ortogonal $n \times n$-matrise har absoluttverdi 1.*

Proposisjonen er riktig også for de komplekse egenverdiene, men det krever et litt annet argument.
Bevis. La $\mathbf{v}$ være en egenvektor for A med egenverdi λ. Da har vi på den ene side

$$(A\mathbf{v})^T \cdot A\mathbf{v} = \mathbf{v}^T A^T A\mathbf{v} = \mathbf{v}^T\mathbf{v}$$

og på den andre side

$$(\lambda\mathbf{v})^T \cdot \lambda\mathbf{v} = \lambda^2 \mathbf{v}^T\mathbf{v}$$

Det betyr at

$$\mathbf{v}^T\mathbf{v} = (A\mathbf{v})^T \cdot A\mathbf{v} = (\lambda\mathbf{v})^T \cdot \lambda\mathbf{v} = \lambda^2\mathbf{v}^T\mathbf{v}$$

Vi har $A\mathbf{v} = \lambda\mathbf{v}$ og siden $\mathbf{v} \neq \mathbf{0}$, så er $\mathbf{v}^T\mathbf{v} = v_1^2 + v_2^2 + \cdots + v_n^2 > 0$. Det følger at $\lambda^2 = 1$. □

Av alle mulige egenverdier skiller $\lambda = 1$ seg ut. En egenvektor med 1 som egenverdi forandres ikke i det hele tatt når vi multipliserer matrisen med den. En slik egenvektor kaller vi en **likevektstilstand**. Navnet er velvalgt fordi, som vi skal se, dersom matrisen modellerer et **dynamisk system** (mer om det i neste kapittel) så beskriver vektorer **tilstander** for systemet. Når tilstanden ikke endres at systemet har vi oppnådd det vi kaller en likevekt, derav navnet **likevektstilstand**.

Proposisjon 1.7.8. *La A være en ortogonal $n \times n$-matrise med $\det(A) = 1$ for et oddetall n. Da tillater A en likevektstilstand, dvs. A har en egenvektor $\mathbf{v}$ med tilhørende egenverdi $\lambda = 1$.*

Bevis. Vi har

$$\begin{aligned}\chi_A(1) &= \det(A - I) = \det(A - AA^T) = \det(A) \cdot \det(I - A^T) \\ &= (-1)^n \det(A) \cdot \det(A - I) = -\det(A) \cdot \chi_A(1)\end{aligned}$$

Siden vi har antatt at $\det(A) = 1$ følger det at $\chi_A(1) = 0$, som er ekvivalent med at $\lambda = 1$ er en egenverdi. □

Vi skal til slutt ta med et viktig resultat, det såkalte Cayley-Hamilton-teoremet. Resultatet sier at en matrise passer i sitt egen karakteristiske polynom. Dette gir mening siden kvadratiske matriser kan legges sammen og multipliseres med hverandre. Dermed kan vi også regne ut verdien av et polynom når vi plugger inn en matrise i stedet for et tall.

Teorem 1.7.9 (Cayley-Hamilton)**.** *La A være en $n \times n$-matrise med karakteristisk polynom χ_A. Da har vi*

$$\chi_A(A) = 0$$

Vi skal ikke bevise dette resultatet i sin fulle generalitet, vi nøyer oss med å se på et argument for 2×2-matriser.

Bevis. La

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

være en 2×2-matrise. Da har vi

$$\begin{aligned} \chi_A(A) &= A^2 - \operatorname{tr}(A) \cdot A - \det(A) \\ &= \begin{pmatrix} a_{11}^2 + a_{12}a_{21} & a_{12}(a_{11} + a_{22}) \\ a_{21}(a_{11} + a_{22}) & a_{12}a_{21} + a_{22}^2 \end{pmatrix} - (a_{11} + a_{22}) \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \\ &\quad + (a_{11}a_{22} - a_{12}a_{21}) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \mathbf{0} \end{aligned}$$

□

Vi kan bruke Cayley-Hamilton-teoremet til å regne ut den inverse til en 3×3-matrise. Vi har tidligere sett at vi har

$$\chi_A(\lambda) = -\lambda^3 + \operatorname{tr}(A)\lambda^2 - \frac{1}{2}((\operatorname{tr}(A))^2 - \operatorname{tr}(A^2))\lambda + \det(A)$$

Cayley-Hamilton gir at

$$-A^3 + \operatorname{tr}(A)A^2 - \frac{1}{2}((\operatorname{tr}(A))^2 - \operatorname{tr}(A^2))A + \det(A) \cdot I = 0$$

Siden A er invertibel kan vi multiplisere likningen med A^{-1}. Det gir

$$-A^2 + \operatorname{tr}(A)A - \frac{1}{2}((\operatorname{tr}(A))^2 - \operatorname{tr}(A^2)) \cdot I + \det(A)A^{-1} = 0$$

Og med litt enkel manipulasjon kommer vi fram til formelen

$$A^{-1} = \frac{1}{\det(A)}(A^2 - \operatorname{tr}(A)A + \frac{1}{2}((\operatorname{tr}(A))^2 - \operatorname{tr}(A^2)) \cdot I)$$

Dette er ikke noen enkel formel å regne med, vi tar den med mer for moro skyld. Men på tross av det ser vi på et eksempel hvor vi bruker formelen til å regne ut den inverse til en 3×3-matrise.

Eksempel 1.7.5. *Vi skal regne ut den inverse matrisen til*

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

Vi har tidligere sett at $\det(A) = 1$. *Dessuten ser vi lett at* $\operatorname{tr}(A) = 2 + 1 + 1 = 4$. *Videre har vi*

$$A^2 = \begin{pmatrix} 6 & 1 & -3 \\ 3 & 1 & -1 \\ -7 & -2 & 3 \end{pmatrix}$$

og $\operatorname{tr}(A^2) = 6+1+3 = 10$. *Dermed har vi at*

$$\begin{aligned} A^{-1} &= \begin{pmatrix} 6 & 1 & -3 \\ 3 & 1 & -1 \\ -7 & -2 & 3 \end{pmatrix} - 4 \begin{pmatrix} 2 & 0 & -1 \\ 1 & 1 & 0 \\ -2 & -1 & 1 \end{pmatrix} + \frac{1}{2}((16-10) \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 1 & 1 \\ -1 & 0 & -1 \\ 1 & 2 & 2 \end{pmatrix} \end{aligned}$$

1.8 Oppgaver med løsning

Eksempel 1.8.1. *La*

$$A = \begin{pmatrix} 3 & 2 \\ 4 & 3 \end{pmatrix}$$

Vis at A er invertibel og finn A^{-1}. Regn ut det karakteristiske polynomet og finn egenverdiene til A.

Løsning.

1. **Metodevalg:** For å avgjøre om A er invertibel regner vi ut determinanten etter formelen

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

Å vise at matrisen er invertibel er ekvivalent med å vise at determinanten er $\neq 0$. Når vi har gjort det kan vi bruke formelen for den inverse til en 2×2-matrise;

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}^{-1} = \frac{1}{\det(A)} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$$

Vi har at $\operatorname{tr}(A) = a_{11} + a_{22}$ og dermed kan vi regne ut det karakteristiske polynomet etter formelen

$$\chi_A(\lambda) = \lambda^2 - \operatorname{tr}(A)\lambda + \det(A)$$

For å finne egenverdiene setter vi $\chi_A(\lambda) = 0$ og løser med hensyn på λ ved å bruke *abc*-formelen.

2. **Regning:**

$$\det(A) = 3 \cdot 3 - 2 \cdot 4 = 1 \neq 0$$

og matrisen er derfor invertibel. Videre har vi

$$A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} 3 & -2 \\ -4 & 3 \end{pmatrix} = \begin{pmatrix} 3 & -2 \\ -4 & 3 \end{pmatrix}$$

Trasen til A er gitt ved $\operatorname{tr}(A) = 3 + 3 = 6$ og vi har karakteristisk polynom

$$\chi_A(\lambda) = \lambda^2 - 6\lambda + 1$$

Innsetting i *abc*-fomelen gir

$$\lambda = \frac{-(-6) \pm \sqrt{(-6)^2 - 4}}{2} = 3 \pm 2\sqrt{2}$$

siden $\sqrt{32} = \sqrt{2^4 \cdot 2} = 2^2\sqrt{2}$.

Eksempel 1.8.2. *La*

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

Vis at A er invertibel og finn A^{-1}. Regn ut det karakteristiske polynomet og finn egenverdiene til A.

Løsning.

1. **Metodevalg:** For å avgjøre om A er invertibel regner vi ut determinanten etter formelen

$$\begin{aligned}\det\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \\ = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{32}a_{21} \\ - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}\end{aligned}$$

Å vise at matrisen er invertibel er ekvivalent med å vise at determinanten er $\neq 0$. For åregne ut den inverse matrisen kan vi bruke formelen

$$A^{-1} = \frac{1}{\det(A)}(A^2 - (\operatorname{tr}(A))A + \frac{1}{2}((\operatorname{tr}(A))^2 - \operatorname{tr}(A^2)) \cdot I)$$

Det betyr at vi trenger $\operatorname{tr}(A)$ og $\operatorname{tr}(A^2)$. Disse må vi regne ut. Trasen er gitt ved summen av diagonalelementene. Den kan vi regne ut for A^2 etter at vi har kvadrert matrisen.

For å regne ut det karakteristiske polynomet til A bruker vi formelen

$$\chi_A(\lambda) = \det(A - \lambda I)$$

eller alternativt

$$\chi_A(\lambda) = -\lambda^3 + (\operatorname{tr}(A))\lambda^2 - \frac{1}{2}((\operatorname{tr}(A))^2 - \operatorname{tr}(A^2))\lambda + \det(A)$$

For å finne egenverdiene setter vi $\chi_A(\lambda) = 0$ og løser med hensyn på λ. Siden dette blir et 3.gradspolynom har vi ikke noen formel slik vi har for 2.gradspolynomer. Men hvis vi via andre metoder kan finne en av egenverdiene, så kan vi ved hjelp av polynomdivisjon redusere problemet til å løse en 2. gradslikning.

2. **Regning:**

$$\det(A) = 0 - 2 - 2 - (-4) - 0 - (-1) = 1 \neq 0$$

og matrisen er derfor invertibel. Videre har vi

$$A^2 = \begin{pmatrix} 1-2+2 & 1-2+1 & 1-1+0 \\ -2+4-2 & -2+4-1 & -2+2+0 \\ 2-2+0 & 2-2+0 & 2-1+0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I$$

Det betyr at $\operatorname{tr}(A^2) = 1+1+1 = 3$, mens $\operatorname{tr}(A) = 1-2+0 = -1$. Det gir

$$A^{-1} = I - (-A + \frac{1}{2}(1-3) \cdot I) = I + A - I = A$$

Dette kunne vi alternativt sett da vi regnet ut A^2 og fant at denne var identitetsmatrisen. Pr. definisjon av den inverse sier dette at A er sin egen invers.

Det karakteristiske polynomet er gitt ved

$$\begin{aligned}\chi_A(\lambda) &= -\lambda^3 - \lambda^2 - \frac{1}{2}(1-3)\lambda + 1 \\ &= -\lambda^3 - \lambda^2 + \lambda + 1\end{aligned}$$

For å finne røttene til dette 3.gradspolynomet observerer vi at siden $A^2 = I$, så vil vi for en egenvektor $\mathbf{v}$ med egenverdi λ ha

$$\mathbf{v} = I\mathbf{v} = A \cdot A\mathbf{v} = A\lambda\mathbf{v} = \lambda^2\mathbf{v}$$

som betyr at $\lambda^2 = 1$, det vil si at $\lambda = \pm 1$. Vi regner ut og finner at $\chi_A(1) = 0$ og $\chi_A(-1) = 0$. Det betyr at $\lambda - 1$ og $\lambda + 1$ deler $\chi_A(\lambda)$. Polynomdivisjon gir da at

$$\frac{\chi_A(\lambda)}{(\lambda-1)(\lambda+1)} = \lambda + 1$$

og vi får

$$\chi_A(\lambda) = (\lambda-1)(\lambda+1)^2$$

Det betyr at vi har to egenverdier, $\lambda_1 = 1$ og $\lambda_2 = -1$.

Eksempel 1.8.3. *La*

$$\mathbf{v} = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$$

Regn ut Householder-transformasjonen $Q_\mathbf{v}$ tilordnet denne vektoren. Vis at $Q_\mathbf{v}(\mathbf{v}) = -\mathbf{v}$. La så

$$\mathbf{w} = \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}$$

Vis at $Q_\mathbf{v}(\mathbf{w}) = \mathbf{w}$.

Vi skal se at disse resultatene ikke svarer til noen sære tilfeller. La derfor $\mathbf{v}$ være en vilkårlig vektor, og $\mathbf{w}$ en vektor som oppfyller $\mathbf{v}^T\mathbf{w} = 0$. Det betyr at de to vektorene stå normalt på hverandre. Vis at

$$Q_\mathbf{v}(\mathbf{v}) = -\mathbf{v}$$

og at

$$Q_\mathbf{v}(\mathbf{w}) = \mathbf{w}$$

Løsning.

1. **Metodevalg:** Vi bruker den generelle formen for Householder-transformasjonen

$$Q_\mathbf{v} = I - 2\frac{\mathbf{v}\mathbf{v}^T}{\mathbf{v}^T\mathbf{v}}$$

og regner ut. I det første tilfellet med innsatte tall, i det andre tilfellet med symboler.

2. **Regning:** Vi har

$$\mathbf{v}^T\mathbf{v} = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} \cdot (2\ 1\ 1) = 6$$

og

$$\mathbf{v}\mathbf{v}^T = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 1 & 1 \\ 2 & 1 & 1 \end{pmatrix}$$

Det gir

$$Q_{\mathbf{v}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \frac{1}{3}\begin{pmatrix} 4 & 2 & 2 \\ 2 & 1 & 1 \\ 2 & 1 & 1 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} -1 & -2 & -2 \\ -2 & 2 & -1 \\ -2 & -1 & 2 \end{pmatrix}$$

og dermed

$$Q_{\mathbf{v}}(\mathbf{v}) = \frac{1}{3}\begin{pmatrix} -1 & -2 & -2 \\ -2 & 2 & -1 \\ -2 & -1 & 2 \end{pmatrix}\begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 \\ -1 \\ -1 \end{pmatrix} = -\mathbf{v}$$

Videre har vi

$$Q_{\mathbf{v}}(\mathbf{w}) = \frac{1}{3}\begin{pmatrix} -1 & -2 & -2 \\ -2 & 2 & -1 \\ -2 & -1 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix} = \mathbf{w}$$

Mer generelt har vi

$$Q_{\mathbf{v}}(\mathbf{v}) = \mathbf{v} - 2\frac{(\mathbf{v}\mathbf{v}^T)\mathbf{v}}{\mathbf{v}^T\mathbf{v}} = \mathbf{v} - 2\frac{\mathbf{v}(\mathbf{v}^T\mathbf{v})}{\mathbf{v}^T\mathbf{v}} = \mathbf{v} - 2\mathbf{v} = -\mathbf{v}$$

og

$$Q_{\mathbf{v}}(\mathbf{w}) = \mathbf{w} - 2\frac{(\mathbf{v}\mathbf{v}^T)\mathbf{w}}{\mathbf{v}^T\mathbf{v}} = \mathbf{w} - 2\frac{\mathbf{v}(\mathbf{v}^T\mathbf{w})}{\mathbf{v}^T\mathbf{v}} = \mathbf{w} - 0 = \mathbf{w}$$

Eksempel 1.8.4. *Vi har gitt et likningssystem*

$$\begin{aligned} 2x + 4y &= 2 \\ -x + 2y &= -5 \\ x - y &= 4 \end{aligned}$$

Er systemet overbestemt, underbestemt eller bestemt? Dersom det er bestemt, finn et minimalt, ekvivalent likningssystem og regn ut determinanten til den tilhørende koeffisientmatrisen A. Regn også ut A^{-1}, og bruk dette til å løse likningsystemet.

Løsning.

1. **Metodevalg:** Et likningsystem med 3 likninger og 2 ukjente er ved første øyekast et overbestemt system, men dersom det finnes en eller flere overflødige likninger kan systemet gjerne være bestemt og til og med underbestemt. Dette må vi sjekke opp. Dersom systemet er bestemt vil det reduserte likningsystemet ha en invertibel koeffisientmatrise, og vi regner ut determinanten med den vanlige formelen for determinanten av en 2×2-matrise,

$$\det(A) = a_{11}a_{22} - a_{12}a_{21}$$

Når det er på plass kan vi regne ut den inverse matrisen ved formelen

$$A^{-1} = \frac{1}{\det(A)}\begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$$

og til slutt løse likningsystemet ved å regne ut

$$\mathbf{x} = A^{-1}\mathbf{b}$$

2. **Regning:** Først observerer vi at den første likningen er det samme som 6 ganger den andre pluss 8 ganger den tredje. Det betyr at den første likningen er overflødig. Likningsystemet reduserer seg til

$$\begin{aligned} -x+2y &= -5 \\ x-y &= 4 \end{aligned}$$

som vi ikke kan redusere ytterligere. Koeffisientmatrisen tildette systemet er

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

som har determinant

$$\det(A) = \begin{vmatrix} -1 & 2 \\ 1 & -1 \end{vmatrix} = (-1)^2 - 2 = -1 \neq 0$$

Høyresiden i likningen er gitt ved

$$\mathbf{b} = \begin{pmatrix} -5 \\ 4 \end{pmatrix}$$

Dermed får vi den inverse matrisen gitt ved

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

som gir oss løsningen

$$\mathbf{x} = A^{-1}\mathbf{b} = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} -5 \\ 4 \end{pmatrix} = \begin{pmatrix} 3 \\ -1 \end{pmatrix}$$

det vil si $x = 3$ og $y = -1$.

Kapittel 2

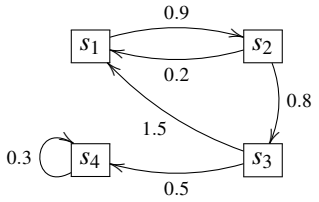
Dynamiske systemer

Mange fenomener innen naturvitenskap kan beskrives med det som i matematisk teori kalles et **dynamisk system**. Felles for alle dynamiske systemer er at vi betrakter et sett av **tilstander** og en regel for hvordan systemet utvikler seg over tid, det vi kaller **dynamikken** til systemet. Vi skiller mellom to typer dynamiske systemer, gjennom hvordan de utvikler seg. **Kontinuerlige systemer** har en fast, kontinuerlig utvikling, f.eks. slik som verdens befolkning, nedbryting av et stoff i kroppen eller radioaktiv utstråling fra et radioaktivt materiale. **Diskrete systemer** utvikler seg skrittvis, f.eks. slik som størrelsen på en populasjon av dyr som stort sett føder barn på en bestemt tid på året. Da måler vi tilstanden, dvs. antall dyr i ulike årsklasser, ved et bestemt tidspunkt på året, og dynamikken forteller oss noe om endringene fra år til år.

2.1 Diskrete dynamiske systemer

Vi kan illustrere et diskret dynamisk system med en figur: Størrelsene som inngår i modellen får hver sin boks, mens dynamikken beskrives med piler mellom boksene. Et tall over pila indikerer hvor stor endringen fra en tilstand til en annen er.

Eksempel 2.1.1. *Vi betrakter et system bestående av 4 størrelser, s_1, s_2, s_3, s_4:*



Systemet kan f.eks. beskrive en populasjon som vi deler inn i 4 klasser. Størrelsen s_1 gir antall individer under 1 år, størrelsen s_2 er antall individer mellom 1 og 2 år, s_3 antall mellom 2 og 3 år, og s_4 er eldre individer. Dynamikken beskriver hva som skjer med antallene ett år fram i tid. Pila fra s_1 til s_2 med tallet 0.9 over beskriver at 90% av alle i den laveste årsklassen overlever til neste år. Tilsvarende for pilene mellom s_2 og s_3, og mellom s_3 og s_4. Siden s_4 favner alle eldre årsklasser har vi en pil fra denne boksen inn i seg selv, hvor tallet forteller hvor stor andel som overlever til neste år.

Pilene fra de tre størrelsene s_2, s_3 og s_4 til s_1 modellerer reproduksjon. Individene i klassen s_2 føder barn og bidrar dermed at antallet i s_1 øker. Reproduksjonsraten er her

satt til 0.2. Individene i klassen s_3 har som vi ser mye høyere reproduksjonsrate, mens "de gamle" i s_4 ikke har noen reproduksjon (0.0). De yngste i klassen s_1 har heller ikke noen reproduksjon, og det går derfor ikke noen pil fra s_1 til seg selv.

Ved et bestemt tidspunkt, som vi setter til år $n = 0$, har vi en fordeling av individer gitt ved $s_1 = 1000$, $s_2 = 700$, $s_3 = 500$, og $s_4 = 200$. Dette setter vi inn i en **tilstandsvektor** $\mathbf{v}_0$;

$$\mathbf{v}_0 = \begin{pmatrix} 1000 \\ 700 \\ 500 \\ 200 \end{pmatrix}$$

Ved å bruke tallene i diagrammet kan vi nå regne ut antallene i hver klasse i år $n = 1$, gitt ved tilstandsvektoren $\mathbf{v}_1$.

$$\mathbf{v}_1 = \begin{pmatrix} 890 \\ 900 \\ 560 \\ 310 \end{pmatrix}$$

Vi kan fortsette å regne ut $\mathbf{v}_2$, $\mathbf{v}_3$, osv., men det er mer hensiktsmessig å forenkle denne prosessen ved å bruke matriseregning.

Matrisen beskriver det som skjer i det dynamiske systemet. Siden matrisen skal beskrive hva som skjer fra tilstand til den neste, kaller vi matrisen for en **overgangsmatrise**. Overgangsmatrisen tilordnet eksemplet over er gitt ved en 4×4-matrise (fordi vi har 4 forskjellige størrelser som inngår i modellen)

$$A = \begin{pmatrix} 0 & 0.2 & 1.5 & 0.0 \\ 0.9 & 0 & 0 & 0 \\ 0 & 0.8 & 0 & 0 \\ 0 & 0 & 0.5 & 0.3 \end{pmatrix}$$

Vi kjenner igjen tallene fra diagrammet, men hvorfor står de akkurat sånn plassert i tabellen? Tallet som står i 2. rad og 1. søyle er det samme som står over pila i diagrammet fra klassen s_2 til klassen s_1. Tilsvarende for de andre tallene. I de tilfellene vi ikke har en pil i diagrammet, setter vi inn en 0 i matrisen.

Tilstandsvektorene etter $n = 1, 2, \ldots$ år er gitt ved å multiplisere forrige tilstandsvektor med matrisen A.

$$\mathbf{v}_1 = A \cdot \mathbf{v}_0, \quad \mathbf{v}_2 = A \cdot \mathbf{v}_1, \quad \text{osv.}$$

Det gir i vårt eksempel

$$\begin{aligned}
\mathbf{v}_1 &= A \cdot \mathbf{v}_0 \\
&= \begin{pmatrix} 0 & 0.2 & 1.5 & 0.0 \\ 0.9 & 0 & 0 & 0 \\ 0 & 0.8 & 0 & 0 \\ 0 & 0 & 0.5 & 0.3 \end{pmatrix} \begin{pmatrix} 1000 \\ 700 \\ 500 \\ 200 \end{pmatrix} \\
&= \begin{pmatrix} 0 \cdot 1000 + 0.2 \cdot 700 + 1.5 \cdot 500 + 0 \cdot 200 \\ 0.9 \cdot 1000 + 0 \cdot 700 + 0 \cdot 500 + 0 \cdot 200 \\ 0 \cdot 1000 + 0.8 \cdot 700 + 0 \cdot 500 + 0 \cdot 200 \\ 0 \cdot 1000 + 0 \cdot 700 + 0.5 \cdot 500 + 0.3 \cdot 200 \end{pmatrix} \\
&= \begin{pmatrix} 890 \\ 900 \\ 560 \\ 310 \end{pmatrix}
\end{aligned}$$

og

$$\mathbf{v}_2 = A \cdot \mathbf{v}_1 = \begin{pmatrix} 0 & 0.2 & 1.5 & 0.0 \\ 0.9 & 0 & 0 & 0 \\ 0 & 0.8 & 0 & 0 \\ 0 & 0 & 0.5 & 0.3 \end{pmatrix} \begin{pmatrix} 890 \\ 900 \\ 560 \\ 310 \end{pmatrix} = \begin{pmatrix} 1020 \\ 801 \\ 720 \\ 373 \end{pmatrix}$$

Det siste uttrykket kan vi skrive litt om;

$$\mathbf{v}_2 = A \cdot \mathbf{v}_1 = A \cdot (A \cdot \mathbf{v}_0) = (A \cdot A) \cdot \mathbf{v}_0 = A^2 \cdot \mathbf{v}_0$$

Eksempel 2.1.2. *Vi kan se på potenser A^k av A for ulike verdier av k gitt i det første eksempelet. I dette tilfellet har vi fått hjelp av en datamaskin, og vi finner (med to riktige desimaler) at*

$$A^2 = \begin{pmatrix} 0.18 & 1.2 & 0 & 0 \\ 0 & 0.18 & 1.35 & 0 \\ 0.72 & 0 & 0 & 0 \\ 0 & 0.4 & 0.15 & 0.09 \end{pmatrix}$$

$$A^{10} = \begin{pmatrix} 0.23 & 1.01 & 1.93 & 0 \\ 1.16 & 0.23 & 0.85 & 0 \\ 0.45 & 1.03 & 0.11 & 0 \\ 0.18 & 0.31 & 0.71 & 0 \end{pmatrix}$$

Potensen A^{10} gir overgangsmatrisen for en 10-års-periode. Vi kan bruke denne på aldersfordelingen $\mathbf{v}_0$ vi opprinnelig hadde. Det gir

$$\mathbf{v}_{10} = A^{10} \cdot \begin{pmatrix} 1000 \\ 700 \\ 500 \\ 200 \end{pmatrix} = \begin{pmatrix} 1901 \\ 1741 \\ 1231 \\ 757 \end{pmatrix}$$

Som vi ser har det totale antallet individer økt, fra 2400 til 5630, og fordelingen mellom aldersgruppene har endret seg. Hvis vi regner ut hva som skjer etter enda ett år, så

kommer vi til å se konturene av noe interessant.

$$\mathbf{v}_{11} = A^{11} \cdot \begin{pmatrix} 1000 \\ 700 \\ 500 \\ 200 \end{pmatrix} = \begin{pmatrix} 2194 \\ 1710 \\ 1393 \\ 842 \end{pmatrix}$$

Tallene fortsetter å øke, fra totalt 5630 etter 10 år, til 6139 etter 11 år, dvs. en økning på 9%. Dersom vi gjør den samme beregningen mellom A^{100} og A^{101}, finner vi en økning på 8,4%. Forskjellen i økningen er ikke så stor, og dersom vi ser på enda høyere potenser av A, vil vi se at økningen stabiliserer seg på 8,4%. Men vi finner ettert hvert en enda dypere sammenheng. Antallet individer i hver aldersgruppe øker også med 8,4%. Hvis vi hadde gjort det samme med A^{1000} og A^{1001}, så ville vi ha fått tilnærmet samme resultat. Det betyr at vi har funnet en tilstandsvektor $\mathbf{v}$ som oppfyller

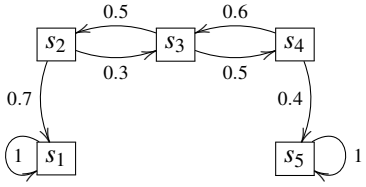
$$A \cdot \mathbf{v} = 1,084 \cdot \mathbf{v}$$

(For ordens skyld, så har vi at $\mathbf{v}_{100} = (2943, 2443, 1802, 1149)$ i hele 1000.) Tilstandsvektoren $\mathbf{v}$ er derfor en egenvektor for matrisen.

En egenvektor omtales ofte som en **stabil** fordeling for systemet. Komponentene i egenvektoren endrer seg når vi multipliserer med matrisen, men forholdet mellom dem bevares. En likevektstilstand er dermed et spesialtilfelle av en stabil fordeling.

Vi skal se på et eksempel til.

Eksempel 2.1.3. *I dette eksemplet lar vi det dynamiske systemet være beskrevet av følgende figur:*



Den tilhørende overgangsmatrisen blir

$$P = \begin{bmatrix} 1 & 0.7 & 0 & 0 & 0 \\ 0 & 0 & 0.5 & 0 & 0 \\ 0 & 0.3 & 0 & 0.6 & 0 \\ 0 & 0 & 0.5 & 0 & 0 \\ 0 & 0 & 0 & 0.4 & 1 \end{bmatrix}$$

Denne overgangsmatrisen har en spesiell egenskap, nemlig at summen av tallene i hver kolonne er 1.

Definisjon 2.1.1. *En matrise med ikke-negative elementer og der summen av tallene i hver kolonne er 1 kalles en* ***stokastisk matrise****.*

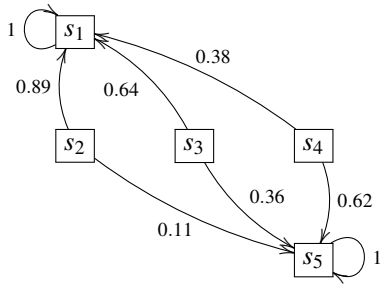
Ser vi på modellen så gjenkjenner vi dette ved at summen av "ut-verdiene" er 1 i alle boksene. Med andre ord, ingen ting kommer til eller fjernes fra systemet. Vi kan tenke på tallene i matrisen som sannsynligheter for at man beveger seg mellom de ulike boksene. Det er grunnen til at vi i dette tilfellet har kalt overgangsmatrisen for P, P for probability. Dette passer også godt med at alle tallene i matrisen skal være positive eller 0, og nødvendigvis mindre enn eller lik 1, siden summen er lik 1.

Ordet **stokastisk** kommer av det greske ordet $\sigma\tau o\chi o\varsigma$ [stochos] som betyr *gjetning* og det henspeiler på at matrisen oppfattes som en tabell av sannsynligheter.

Vi kan bruke en datamaskin til å regne ut høye potenser av P, og da ser vi noe interessant:

$$P^{100} = \begin{bmatrix} 1 & 0.89 & 0.64 & 0.38 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0.11 & 0.36 & 0.62 & 1 \end{bmatrix}$$

Alt i midten av matrisen blir 0. Vi kan illustrere P^{100} med en figur:



Her har vi gjort det motsatte av det vi startet med å gjøre. Med utgangspunkt i en overgangsmatrise har vi laget en figur som gir oss nøyaktig samme informasjon. Hvis vi starter med en tilstandsvektor

$$\mathbf{v}_0 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

så vil vi etter 100 skritt ende med

$$\mathbf{v}_{100} = P^{100} \cdot \mathbf{v}_0 = \begin{pmatrix} 0.89 \\ 0 \\ 0 \\ 0 \\ 0.11 \end{pmatrix}$$

Dette kan vi tolke som at dersom vi starter i klassen s_2, så vil vi i 89% av tilfellene ende opp i klassen s_1 etter 100 skritt, og resten i klassen s_5. Vi ender aldri opp i de tre midterste klassene.

Vi har sett at det er en 1-1 sammenheng mellom figurene med bokser og piler og den tilsvarende overgangsmatrisen.

$$\mathbf{v}_{n+1} = P \cdot \mathbf{v}_n, \quad n \geq 0$$

Dersom vi skriver ut dette i detalj får vi en tredje måte å beskrive systemet. Vi bruker notasjonen $\mathbf{v}_{n,i}$ for i-te koordinat i vektoren $\mathbf{v}_n$. Uttrykket $\mathbf{v}_{n+1} = P \cdot \mathbf{v}_n$ skriver vi som (med tall fra *Eksempel* 2.1.1)

$$\begin{pmatrix} \mathbf{v}_{n+1,1} \\ \mathbf{v}_{n+1,2} \\ \mathbf{v}_{n+1,3} \\ \mathbf{v}_{n+1,4} \end{pmatrix} = \begin{pmatrix} 0 & 0.2 & 1.5 & 0.0 \\ 0.9 & 0 & 0 & 0 \\ 0 & 0.8 & 0 & 0 \\ 0 & 0 & 0.5 & 0.3 \end{pmatrix} \begin{pmatrix} \mathbf{v}_{n,1} \\ \mathbf{v}_{n,2} \\ \mathbf{v}_{n,3} \\ \mathbf{v}_{n,4} \end{pmatrix} = \begin{pmatrix} 0.2\mathbf{v}_{n,2} + 1.5\mathbf{v}_{n,3} \\ 0.9\mathbf{v}_{n,1} \\ 0.8\mathbf{v}_{n,2} \\ 0.5\mathbf{v}_{n,3} + 0.3\mathbf{v}_{n,4} \end{pmatrix}$$

som gir oss likningsettet

$$\begin{aligned}
\mathbf{v}_{n+1,1} &= 0.2\mathbf{v}_{n,2} + 1.5\mathbf{v}_{n,3} \\
\mathbf{v}_{n+1,2} &= 0.9\mathbf{v}_{n,1} \\
\mathbf{v}_{n+1,3} &= 0.8\mathbf{v}_{n,2} \\
\mathbf{v}_{n+1,4} &= 0.5\mathbf{v}_{n,3} + 0.3\mathbf{v}_{n,4}
\end{aligned}$$

Dermed har vi tre ulike måter å presentere modellen. Alle måtene gir oss samme informasjon. Hvilken vi velger avhenger av hva vi skal bruke modellen til og hvordan vi skal bruke den.

2.2 Lukkede dynamiske systemer

Vi skal se litt nøyere på en bestemt klasse av dynamiske systemer, som vi kaller **lukkede systemer**. Lukkede systemer modellerer prosesser som er isolert fra omgivelsene, i den forstand at ingenting kommer ut eller føres inn i systemet. Et eksempel kan være gassmolekylene i en lukket og isolert beholder.

For en vektor $\mathbf{v} = (v_1, \ldots, v_n)$ lar vi

$$\sigma(\mathbf{v}) = \sum_{i=1}^{n} v_i$$

betegne **komponentsummen** til vektoren.

Definisjon 2.2.1. *Vi sier at et dynamisk system er **lukket** dersom elementene i overgangsmatrisen A er ikke-negative og komponentsummen til en tilstandsvektor ikke endres av dynamikken i systemet, dvs. dersom* $\mathbf{v}$ *er en tilstandsvektor, så er komponentsummene av* $\mathbf{v}$ *og* $A\mathbf{v}$ *like;*

$$\sigma(\mathbf{v}) = \sigma(A\mathbf{v})$$

Proposisjon 2.2.2. *Dersom et dynamisk system er lukket vil den tilhørende overgangsmatrisen være en stokastisk matrise. Det omvendte er også tilfellet, det tilhørende dynamiske systemet til en stokastisk matrise er lukket.*

Bevis. La A være overgangsmatrisen til det dynamiske systemet, og la $\mathbf{e}_j$ være vektoren med en 1-er i koordinat j og 0-er ellers. Da vil produktet $A\mathbf{e}_j$ være den j-te søylen i matrisen A. Siden komponentsummen til $\mathbf{e}_j$ er $\sigma(\mathbf{e}_i) = 1$, må komponentsummen til hver søyle i A være 1, og matrisen er stokastisk.

La A være en stokastisk matrise gitt ved

$$A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix},$$

der

$$a_{1,j} + a_{2,j} + \cdots + a_{n,j} = 1 \quad \text{for alle } j = 1, \ldots, n$$

og la

$$\mathbf{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix},$$

være en tilstandsvektor. Da har vi

$$A\mathbf{v} = \begin{pmatrix} a_{1,1}v_1 + \cdots + a_{1,n}v_n \\ \vdots \\ a_{n,1}v_1 + \cdots + a_{n,n}v_n \end{pmatrix},$$

og

$$\begin{aligned} \sigma(A\mathbf{v}) &= \sum_{j=1}^{n} (a_{j,1}v_1 + \cdots + a_{j,n}v_n) \\ &= (\sum_{j=1}^{n} a_{j,1})v_1 + \cdots + (\sum_{j=1}^{n} a_{j,n})v_n \\ &= v_1 + \cdots + v_n = \sigma(\mathbf{v}) \end{aligned}$$

og systemet er lukket. □

Proposisjon 2.2.3. *Et lukket system har en likevektstilstand.*

Bevis. La $A = (a_{ij})$ være den stokastiske overgangsmatrisen til det lukkede systemet. En likevektstilstand $\mathbf{v}$ må tilfredsstille et likningssystem

$$\begin{aligned} a_{11}v_1 + a_{12}v_2 + \cdots + a_{1n}v_n &= v_1 \\ a_{21}v_1 + a_{22}v_2 + \cdots + a_{2n}v_n &= v_2 \\ &\vdots \\ a_{n1}v_1 + a_{n2}v_2 + \cdots + a_{nn}v_n &= v_n \end{aligned}$$

Den siste likningen kan vi skrive som

$$\begin{aligned} v_n &= a_{n1}v_1 + a_{n2}v_2 + \cdots + a_{nn}v_n \\ &= (1 - \sum_{i=1}^{n-1} a_{i1})v_1 + \cdots + (1 - \sum_{i=1}^{n-1} a_{in})v_n \\ &= (v_1 + \cdots + v_n) - (v_1 + \cdots + v_{n-1}) = v_n \end{aligned}$$

som betyr at den er overflødig. Dermed har vi $n-1$ likninger med n ukjente, og ved å flytte høyresiden over til venstresiden får vi et homogent likningssystem. Vi skal siden se at slike systemer har ikke-trivielle løsninger. Det betyr at det finnes en egenvektor med egenverdi $\lambda = 1$. □

Proposisjon 2.2.4. *La λ være en egenverdi for en stokastisk matrise A. Da er absoluttverdien til λ er mindre enn eller lik 1, dvs. $|\lambda| \leq 1$.*

Bevis. Vi så i Prop. 1.6.6 at en matrise og dens transponerte har de samme egenverdiene. Det betyr at det er like bra å vise resultatet for A^T som for A. Grunnen til at vi vil gjøre akkurat det er at det for dette beviset er mer hensiktsmessig at vi har radsummer lik 1 (og ikke søylesummer). Premisset i proposisjonen er dermed at A^T har radsummer lik 1.

La $\mathbf{v} = (v_1, \ldots, v_n)$ være en egenvektor for A^T med egenverdi λ. Anta at v_k er komponenten med størst absoluttverdi. Siden $\mathbf{v} \neq \mathbf{0}$ har vi nødvendigvis at $v_k > 0$. Hvis vi skriver ut $A\mathbf{v} = \lambda\mathbf{v}$ på koordinatform, får vi

$$a_{j1}v_1 + a_{j2}v_2 + \cdots + a_{jn}v_n = \lambda v_j$$

for alle $j = 1, 2, \ldots, n$. Spesielt vil dette også gjelde for $j = k$.

Det følger nå at (ved å bruke trekantulikheten og at elementene i A er positive):

$$\begin{aligned}
|\lambda| \cdot |v_k| &= |a_{k1}v_1 + a_{k2}v_2 + \cdots + a_{kn}v_n| \\
&\leq a_{k1}|v_1| + a_{k2}|v_2| + \cdots + a_{kn}|v_n| \\
&\leq a_{k1}|v_k| + a_{k2}|v_k| + \cdots + a_{kn}|v_k| \\
&= (a_{k1} + a_{k2} + \cdots + a_{kn})|v_k| \\
&= |v_k|
\end{aligned}$$

Siden $|v_k| > 0$ følger det at $|\lambda| \leq 1$.

□

Siden radsummene i A^T er 1 følger det at $\mathbf{1} = (1, \ldots, 1)^T$ er en egenvektor med egenverdi 1. Det betyr at for en egenvektor for en annen egenverdi enn 1 vil ikke alle koordinatene være like. Ulikheten i beviset for proposisjonen over vil det tilfellet være en ekte ulikhet dersom alle elementene i A er positive (og ikke 0), noe som vi skriver $A > 0$. Dette er essensen i beviset for Markovs teorem. Markovs teorem sier:

Proposisjon 2.2.5. *La A være en stokastisk matrise hvor alle elementene er positive. Da vil den eneste egenverdien med absoluttverdi 1 være 1.*

Faktisk er resultatet litt mer generelt. Man kan vise at det holder at positiviteten er oppfyllt for en potens av A, dvs. at det eksisterer et positivt heltall s slik at $A^s > 0$.

En konsekvens av Markovs teorem er at dersom vi starter med en tilstandsvektor $\mathbf{w} = \mathbf{v}_0 + \sum_{i=1}^{m} b_i \mathbf{v}_i$ der $\mathbf{v}_0$ er en likevektstilstand og $\mathbf{v}_i$ er egenvektorer med egenverdi ulik 1, som betyr at absoluttverdien er ekte mindre enn 1, så vil

$$A^n \mathbf{w} = 1^n \mathbf{v}_0 + \sum_{i=1}^{m} b_i \lambda_i^n \mathbf{v}_i$$

og siden

$$\lim_{n \to \infty} \lambda^n = 0$$

for $|\lambda| < 1$, så følger det at

$$\lim_{n \to \infty} A^n \mathbf{w} = \mathbf{v}_0$$

dvs. at vi ender opp i likevektstilstanden.

Vi går tilbake til *Eksempel* 2.1.3, og 5×5-matrisen P gitt ved

$$P = \begin{bmatrix} 1 & 0.7 & 0 & 0 & 0 \\ 0 & 0 & 0.5 & 0 & 0 \\ 0 & 0.3 & 0 & 0.6 & 0 \\ 0 & 0 & 0.5 & 0 & 0 \\ 0 & 0 & 0 & 0.4 & 1 \end{bmatrix}$$

Matrisen P er stokastisk og vi vet da at en av egenverdiene er 1. De andre egenverdiene er i avtagende rekkefølge

$$1 \geq 1 \geq \frac{3\sqrt{5}}{10} \geq 0 \geq -\frac{3\sqrt{5}}{10}$$

Egenrommet til $\lambda_1 = 1$ er gitt ved alle vektorer på formen

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

hvor x og y er vilkårlige reelle tall. Vi har altså to lineært uavhengige egenvektorer for egenverdien $\lambda = 1$. Slike tilstander vil derfor være likevektstilstander for systemet. For vektorer av typen

$$\mathbf{v} = \begin{pmatrix} 0 \\ a \\ b \\ c \\ 0 \end{pmatrix}$$

hvor a, b og c er ikke-negative tall har vi

$$P^2\mathbf{v} = (0,7a+0,35b)\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + (0,2b+0,4c)\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 0 \\ 0,15a+0,3c \\ 0,45b \\ 0,15a+0,3c \\ 0 \end{pmatrix}$$

Det betyr at

$$\sigma\begin{pmatrix} 0,15a+0,3c \\ 0,45b \\ 0,15a+0,3c \end{pmatrix} = 0.3a+0.45b+0.6c < \mu(a+b+c) = \sigma\begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

for alle μ slik at $0.6 < \mu < 1$ (dette finner vi ved litt regning). Dermed har vi

$$\lim_{n\to\infty} P^n\mathbf{v} = \begin{pmatrix} x \\ 0 \\ 0 \\ 0 \\ y \end{pmatrix}$$

for to tall x og y med $x+y=1$. Fra *Eksempel* 2.1.3 vet vi at $x \approx 0,89a+0,64b+0,38c$ og $y \approx 0,11a+0,36b+0,62c$. Modellen beskriver et system der vi starter med 5 størrelser, og hvor etter hvert alt som inngår i modellen flytter seg fra de tre midterste boksene og ut i de to ytterste.

2.3 Oppgaver med løsning

Eksempel 2.3.1. *I en populasjon forekommer to forskjellige genvarianter A og B. Som et resultat av mutasjoner vil det skje endringer i antallet av de to genvariantene, gitt ved den dynamiske modellen*

$$\mathbf{y}_{t+1} = \begin{pmatrix} 0.9 & 0 \\ 0.1 & 1 \end{pmatrix} \mathbf{y}_t$$

a) Anta at på tidspunktet t så vil fordelingen mellom de to genvariantene være

$$\begin{pmatrix} \mathbf{y}_{t,A} \\ \mathbf{y}_{t,B} \end{pmatrix} = \begin{pmatrix} 600 \\ 400 \end{pmatrix}$$

Hva vil fordelingen være ved tiden $t+1$?

b) Overgangen fra tilstanden ved tiden t til tilstanden ved tiden $t+2$ kan beskrives ved matrisemultiplikasjon med

$$\begin{pmatrix} 0.9 & 0 \\ 0.1 & 1 \end{pmatrix}^2$$

Regn ut dette produktet.

c) En likevektstilstand for denne modellen finner vi ved å sette $\mathbf{y}_{t+1} = \mathbf{y}_t$. Finn en likevektstilstand

$$\mathbf{y} = \begin{pmatrix} u \\ v \end{pmatrix}$$

slik at

$$\mathbf{y} = \begin{pmatrix} 0.9 & 0 \\ 0.1 & 1 \end{pmatrix} \mathbf{y}$$

og der $u+v = 1000$.

Løsning.

1. **Metodevalg:** Vi skal multiplisere sammen matriser, eventuelt matriser og søylevektorer. De aktuelle formelene skrevet ut for 2×2-matriser er

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} a_{11}v_1 + a_{12}v_2 \\ a_{21}v_1 + a_{22}v_2 \end{pmatrix}$$

Og for to matriser:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \cdot \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix}$$

2. **Regning:** Vi får

$$\begin{pmatrix} 0.9 & 0 \\ 0.1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 600 \\ 400 \end{pmatrix} = \begin{pmatrix} 0.9 \cdot 600 + 0 \cdot 400 \\ 0.1 \cdot 600 + 1 \cdot 400 \end{pmatrix} = \begin{pmatrix} 540 \\ 460 \end{pmatrix}$$

Videre har vi

$$\begin{pmatrix} 0.9 & 0 \\ 0.1 & 1 \end{pmatrix}^2 = \begin{pmatrix} 0.9\cdot 0.9+0\cdot 0.1 & 0.9\cdot 0+0\cdot 1 \\ 0.1\cdot 0.9+1\cdot 0.1 & 0\cdot 0+1\cdot 1 \end{pmatrix}$$
$$= \begin{pmatrix} 0.81 & 0 \\ 0.19 & 1 \end{pmatrix}$$

En ekvivalent formulering av matriseproduktet

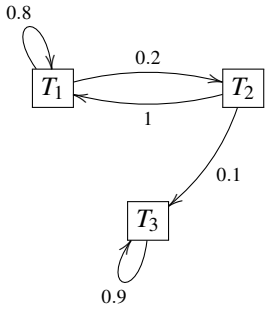
$$\mathbf{y} = \begin{pmatrix} 0.9 & 0 \\ 0.1 & 1 \end{pmatrix} \mathbf{y}$$

er som likningssytemet

$$u = 0.9u$$
$$v = 0.1u + v$$

og samtidig $u+v = 1000$. Dette gir at $u = 0$ og $v = 1000$.

Eksempel 2.3.2. *En modell for en populasjonen består av tre klasser. Dynamikken i modellen er gitt som i illustrasjonen.*



Still opp den tilhørende overgangsmatrisen. Kan du tenke deg til hva som skjer i modellen i det lange løp? Vis at dersom utgangspunktet er en tilstandsvektor

$$\mathbf{v}_0 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

så vil populasjonen dø ut. Vis også at dersom utgangspunktet er

$$\mathbf{v}_0 = \begin{pmatrix} 5 \\ 1 \\ 1 \end{pmatrix}$$

så vil populasjonen holde seg konstant. Faktisk er det slik at dersom vi starter med tilstandsvektoren

$$\mathbf{v}_0 = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

der a og b ikke begge er 0, så vil vi i det lange løp ende opp med en tilstandsvektor gitt ved

$$\mathbf{v}_\infty = \frac{a+b}{6}\begin{pmatrix} 5 \\ 1 \\ 1 \end{pmatrix}$$

Forklar hvorfor det er slik.

Løsning.

1. **Metodevalg:** Vi oversetter en boks-illustrasjon av et dynamisk system til en matrise M ved følgende oppskrift: Siden systemet har 3 klasser, vil overgangsmatrisen være en 3×3-matrise. På plass (i, j), dvs. i-te rad og j-te søyle, setter vi inn tallet som står ved pila som går fra tilstand j til tilstand i. Vi regner ut tilstandsvektorer med utgangspunkt i $\mathbf{v}_0$ ved å multiplisere med M;

$$\mathbf{v}_1 = M \cdot \mathbf{v}_0$$

osv. Dersom vi har en tilstandsvektor $\mathbf{v}$ med en stabil fordeling, dvs. at

$$\mathbf{v}_{n+1} = M \cdot \mathbf{v}_n = \lambda \mathbf{v}_n$$

for et reelt tall λ, så kan vi si noe om utviklingen av systemet. Dersom $|\lambda| < 1$, så vil $\mathbf{v}_n \to 0$ i dette tilfellet. Dersom $\lambda = 1$ vil tilstanden holde seg konstant.

2. **Regning:** Vi bruker oppskriften over og finner at

$$M = \begin{pmatrix} 0.8 & 1 & 0 \\ 0.2 & 0 & 0 \\ 0 & 0.1 & 0.9 \end{pmatrix}$$

Vi regner ut

$$M = \begin{pmatrix} 0.8 & 1 & 0 \\ 0.2 & 0 & 0 \\ 0 & 0.1 & 0.9 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0.9 \end{pmatrix}$$

Det betyr at vi har en stabil fordeling med $\lambda = 0.9$, som betyr at populasjonen vil dø ut. Det neste eksempelet på en $\mathbf{v}_0$ gir oss

$$M = \begin{pmatrix} 0.8 & 1 & 0 \\ 0.2 & 0 & 0 \\ 0 & 0.1 & 0.9 \end{pmatrix} \cdot \begin{pmatrix} 5 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 5 \\ 1 \\ 1 \end{pmatrix}$$

som betyr at i dette tilfellet vil populasjonen holde seg konstant. I det siste eksemplet får vi

$$\begin{pmatrix} 0.8 & 1 & 0 \\ 0.2 & 0 & 0 \\ 0 & 0.1 & 0.9 \end{pmatrix} \cdot \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0.8a+b \\ 0.2a \\ 0.1b+0.9c \end{pmatrix}$$

I dette tilfellet ser vi at summen av de to første komponentene er konstant lik $a+b$ og hvis vi skal ha

$$0.8a+b = a \quad \text{og} \quad 0.2a = b$$

så må vi ha $a = 5b$. Dersom summen skal være konstant betyr det at $(a,b) = \frac{a+b}{6}(5,1)$. Setter vi dette inn i utregningen får vi

$$M \begin{pmatrix} \frac{5(a+b)}{6} \\ \frac{a+b}{6} \\ c \end{pmatrix} = \begin{pmatrix} \frac{5(a+b)}{6} \\ \frac{a+b}{6} \\ \frac{a+b}{60} + 0.9c \end{pmatrix}$$

og den tredje koordinaten vil ha utviklingen

$$\begin{aligned} c, \quad & \frac{a+b}{60} + 0.9c, \quad \frac{a+b}{60} + 0.9\frac{a+b}{60} + 0.9^2 c, \\ & \frac{a+b}{60} + 0.9\frac{a+b}{60} + 0.9^2\frac{a+b}{60} + 0.9^3 c, \quad \ldots \end{aligned}$$

Bruker vi formelen for summen av en geometrisk rekke med første ledd $\frac{a+b}{60}$ og faktor 0.9 får vi at etter hvert så vil den siste komponenten nærme seg mot

$$\lim_{n\to\infty} 0.9^n c + \frac{\frac{a+b}{60}}{1-0.9} = 0 + \frac{\frac{a+b}{60}}{\frac{1}{10}} = \frac{a+b}{6}$$

Kapittel 3

Euklidske vektorrom

3.1 Lineære underrom

Vi starter med litt repetisjon fra Kapittel 1. La $V = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m\}$ være en mengde av vektorer i $\mathbb{R}^n$. En **lineærkombinasjon** av vektorene er en sum

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_m\mathbf{v}_m$$

hvor $x_1, x_2, \ldots, x_m$ er reelle tall. Mengden $\mathscr{V} = \{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ sies å være **lineært uavhengig** dersom den eneste muligheten for at

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_m\mathbf{v}_m = \mathbf{0}$$

er at alle $x_i = 0$, $i = 1, 2, \ldots, m$. En mengde av vektorer som ikke er lineært uavhengig sies å være **lineært avhengig**.

En alternativ formulering av denne definisjonen er at en mengde er lineært avhengig dersom minst en av vektorene i mengden kan skrives som en lineær kombinasjon av de andre.

Definisjon 3.1.1. *La V være en ikke-tom mengde av vektorer i $\mathbb{R}^n$. Vi sier at V er **lukket under lineære kombinasjoner** dersom*

$$a\mathbf{v} + b\mathbf{w} \in V$$

for alle vektorer $\mathbf{v}, \mathbf{w} \in V$ og reelle tall $a, b \in \mathbb{R}$.

Definisjon 3.1.2. *En ikke-tom delmengde av $\mathbb{R}^n$ kalles et **lineært underrom** av $\mathbb{R}^n$ dersom det er lukket under lineære kombinasjoner.*

Med denne definisjonen vil $\mathbb{R}^n$ være et lineært underrom av seg selv og selvfølgelig det største mulige underrommet. Ved å sette a og b til å være 0 i Definisjon 3.1.1, følger det at $\mathbf{0} \in V$ for ethvert lineært underrom $V \subset \mathbb{R}^n$. I tillegg er også $\{\mathbf{0}\}$ et lineært underrom i seg selv, siden $a\mathbf{0} + b\mathbf{0} = \mathbf{0}$ for alle reelle tall a og b. Dette er nødvendigvis det minste lineære underrommet av $\mathbb{R}^n$. Det er vanlig å omtale $V = \{\mathbf{0}\}$ som det **trivielle** underrommet i $\mathbb{R}^n$.

Eksempel 3.1.1. *La $x \in \mathbb{R}$ være et rellt tall forskjellig fra 0 og y et vilkårlig reelt tall. Da har vi*

$$y = \frac{y}{x} \cdot x$$

som betyr at ethvert reelt tall er en lineærkombinasjon av et vilkårlig tall x, forskjellig fra 0. Følgelig vil de eneste mulige lineære underrommene av $\mathbb{R}$ være $\mathbb{R}$ selv og $\{\mathbf{0}\}$.

Formuleringen $y = \frac{x}{y} \cdot x$ er en lineærkombinasjon av x kan virke litt rar, siden det åpenbart ikke er noe å kombinere her. Men dersom vi i stedet skriver

$$y = \frac{x}{y} \cdot x + b \cdot 0$$

så passer det bedre med oppsettet i definisjonen og også rent språklig.

Eksempel 3.1.2. *Det er tre ulike typer av lineære underrom av $\mathbb{R}^2$;*

- *Det trivielle underrommet $\{\mathbf{0}\}$,*
- *Mengden av vektorer på samme rette linje gjennom origo,*
- *Hele rommet $\mathbb{R}^2$.*

Eksempel 3.1.3. *Det er fire ulike typer av lineære underrom av $\mathbb{R}^3$;*

- *Det trivielle underrommet $\{\mathbf{0}\}$,*
- *Mengden av vektorer på samme rette linje gjennom origo,*
- *Mengden av vektorer i samme plan gjennom origo,*
- *Hele rommet $\mathbb{R}^3$.*

Vårt neste mål er å forsøke å si noe om størrelsen på et lineært underrom. Det er ikke vanskelig å overbevise seg om at et ikke-trivielt lineært underrom, vil inneholde uendelig mange vektorer. Dermed vil det gi lite mening i å måle størrelsen ved å telle anall elementer. Et bedre mål er å klassifisere underrommene gjennom **dimensjon**. I Eksempel 3.1.3 tenker vi kanskje intuitivt at rommet har dimensjon 3, at et plan har dimensjon 2 og at en rett linje har dimensjon 1. For at dimensjonsbegrepet skal være troverdig bør dette faktisk stemme.

Før vi kan gi en presis matematisk definisjon av dimensjon, trenger vi litt mer maskineri. Først skal vi se på en måte å lage lineære underrom.

La $\mathscr{V}$ være en mengde vektorer i $\mathbf{R}^n$. Da er $V = \text{span}\mathscr{V}$ mengden av alle vektorer som kan skrives som en linæerkombinasjon av vektorer fra $\mathscr{V}$. Det følger nesten umiddelbart fra definisjonen at V er et lineært underrom av $\mathbb{R}^n$. Vi sier at mengden $\mathscr{V}$ **utspenner** underrommet V .

Eksempel 3.1.4. *Mengden*

$$\mathscr{V} = \left\{ \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \right\}$$

utspenner det lineære underrommet

$$V = span\mathscr{V} = \left\{ a \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} + b \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \,\middle|\, a, b \in \mathbb{R} \right\} = \left\{ \begin{pmatrix} a+b \\ a-b \\ 2a \end{pmatrix} \,\middle|\, a, b \in \mathbb{R} \right\}$$

Hvis vi setter $x = a+b$, $y = a-b$ og $z = 2a$, så får vi

$$x + y - z = (a+b) + (a-b) - 2a = 0$$

Underrommet V består derfor av alle punkter $(x,y,z) \in \mathbb{R}^3$ som oppfyller $x+y-z = 0$.

Ulike mengder kan utspenne samme lineære underrom. I eksemplet over vil f.eks. mengdene

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\} \quad \text{og} \quad \left\{ \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\}$$

begge utspenne V. I den siste mengden har vi en lineær relasjon

$$\begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} - \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} - 2 \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \mathbf{0}$$

som betyr at den er lineært avhengig. Hvis vi dropper en av de tre vektorene så vil de gjenværende fortsatt spenne ut det samme underrommet.

Definisjon 3.1.3. *La V være et lineært underrom av $\mathbb{R}^n$. En mengde $\mathscr{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_r\} \subset V$ kalles en **basis** for V dersom mengden er lineært uavhengig og **utspenner** hele V, dvs. at for alle $\mathbf{v} \in V$, så finnes en og bare en mengde av reelle tall $a_1, \ldots a_r \in \mathbb{R}$ slik at*

$$\mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_r\mathbf{v}_r$$

Grunnen til at vi krever at en basis skal være lineært uavhengig er at vi vil at rommet skal utspenne mengden med så få elementer som mulig. En basis er en minimal generatormengde i den forstand at hvis vi fjerner et element fra basisen, så vil underrommet som basisen utspenner bli ekte redusert. For en lineært avhengig mengde er det alltid mulig å fjerne et element uten at mengden som utspennes endres.

Eksempel 3.1.5. *Mengden*

$$\mathscr{V} = \{ \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \end{pmatrix} \}$$

utspenner hele $\mathbb{R}^2$; for et vilkårlig element

$$\begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^2$$

kan vi f.eks. skrive

$$\begin{pmatrix} a \\ b \end{pmatrix} = b \begin{pmatrix} 2 \\ 1 \end{pmatrix} + (a - 2b) \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Mengden er imidlertid ikke lineært uavhengig (sett $a = b = 2$) og danner derfor ikke en basis. Fjerner vi det siste elementet $\begin{pmatrix} 2 \\ 2 \end{pmatrix}$ fra mengden har vi fortsatt nok vektorer til å spenne ut hele $\mathbb{R}^2$. De to elementene

$$\begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

er lineært uavhengige og siden de samtidig utspenner hele $\mathbb{R}^2$ utgjør de en basis for $\mathbb{R}^2$.

Eksempel 3.1.6. *La*

$$\mathscr{B} = \{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \}$$

være tre elementer i $\mathbb{R}^3$. $\mathscr{B}$ *er lineært uavhengig siden*

$$a\begin{pmatrix}1\\0\\0\end{pmatrix}+b\begin{pmatrix}0\\1\\0\end{pmatrix}+c\begin{pmatrix}0\\0\\1\end{pmatrix}=\begin{pmatrix}a\\b\\c\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}$$

kun er mulig dersom $a=b=c=0$. *Vi ser også at* $\mathscr{B}$ *utspenner hele* $\mathbb{R}^3$ *siden et vilkårlig element*

$$\begin{pmatrix}a\\b\\c\end{pmatrix}=a\begin{pmatrix}1\\0\\0\end{pmatrix}+b\begin{pmatrix}0\\1\\0\end{pmatrix}+c\begin{pmatrix}0\\0\\1\end{pmatrix}$$

Basisen $\mathscr{B}$ *kalles vanligvis for* ***standardbasisen*** *for* $\mathbb{R}^3$.

Eksempel 3.1.7. *La*

$$\mathscr{B}=\{\begin{pmatrix}1\\1\\1\end{pmatrix},\begin{pmatrix}-1\\1\\0\end{pmatrix},\begin{pmatrix}2\\-1\\1\end{pmatrix}\}$$

være en mengde av tre elementer i $\mathbb{R}^3$. *For å avgjøre om* $\mathscr{B}$ *er lineært uavhengig setter vi*

$$a\begin{pmatrix}1\\1\\1\end{pmatrix}+b\begin{pmatrix}-1\\1\\0\end{pmatrix}+c\begin{pmatrix}2\\-1\\1\end{pmatrix}=\begin{pmatrix}a-b+2c\\a+b-c\\a+c\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}$$

Det gir oss likningssettet

$$\begin{aligned}a-b+2c&=0\\a+b-c&=0\\a+c&=0\end{aligned}$$

som har som eneste løsning $a=b=c=0$ *og mengden er derfor lineært uavhengig. Vi ser også at* $\mathscr{B}$ *utspenner hele* $\mathbb{R}^3$ *siden et vilkårlig element kan skrives som en lineær kombinasjon av de tre vektorene;*

$$\begin{pmatrix}x\\y\\z\end{pmatrix}=(x+y-z)\begin{pmatrix}1\\1\\1\end{pmatrix}+(-2x-y+3z)\begin{pmatrix}-1\\1\\0\end{pmatrix}+(-x-y+2z)\begin{pmatrix}2\\-1\\1\end{pmatrix}$$

La $\mathscr{B}$ være en basis for et lineært underrom V. Vi skal se at alle basiser for V har like mange elementer som $\mathscr{B}$.

Lemma 3.1.4. *La* $\mathscr{B}=\{\mathbf{v}_1,\ldots,\mathbf{v}_r\}$ *være en basis for et lineært underrom* V, *og la* $\mathscr{W}=\{\mathbf{w}_1,\ldots,\mathbf{w}_s\}\subset V$, $\mathbf{w}_j\neq 0$, *være en mengde med flere elementer enn det er i basisen* $\mathscr{B}$. *Da er mengden* $\mathscr{W}$ *lineært avhengig.*

I en litt annen versjon kalles dette lemmaet vanligvis for **Steinitz' erstatningsresultat**.

Bevis. Anta at mengden $\mathscr{W}$ er lineært uavhengig. Siden $\mathscr{B}$ er en basis kan vi skrive $\mathbf{w}_1=b_1\mathbf{v}_1+\cdots+b_r\mathbf{v}_r$ hvor ikke alle $b_j=0$. Ved å bytte om på indeksene, kan vi anta at anta at vi har ordnet basiselementene slik at $b_1\neq 0$. Da har vi

$$\mathbf{v}_1=\frac{1}{b_1}\mathbf{w}_1-\frac{b_2}{b_1}\mathbf{v}_2-\cdots-\frac{b_r}{b_1}\mathbf{v}_r$$

Det betyr at vi kan erstatte $\mathbf{v}_1$ med $\mathbf{w}_1$ i basisen $\mathscr{B}$ og det vil fortsatt være en basis. Vi gjentar denne prosedyren til vi har byttet ut alle $\mathbf{v}_j$ med $\mathbf{w}_j$. (Dette er mulig siden begge mengdene er lineært uavhengige). Vi ender opp med at $\{\mathbf{w}_1,\ldots,\mathbf{w}_r\}$ danner en basis for V. Men det betyr at $\mathbf{w}_{r+1}$ kan skrives som en lineær kombinasjon av $\mathbf{w}_1,\ldots,\mathbf{w}_r$, og som strider mot antakelsen om at $\mathscr{W}$ er lineært uavhengig. Altså var antakelsen feil. □

Det følgende resultatet er en direkte konsekvens av Steinitz' erstatnings-resultat:

Teorem 3.1.5. *Gitt et lineært underrom V og anta at $\mathscr{B}_1 = \{\mathbf{v}_1,\ldots,\mathbf{v}_r\}$ og $\mathscr{B}_2 = \{\mathbf{w}_1,\ldots,\mathbf{w}_s\}$ er to forskjellige basiser for V. Da har de to basisene like mange elementer, dvs. $r = s$.*

Definisjon 3.1.6. *La V være et lineært underrom, og anta at $\mathscr{B} = \{\mathbf{v}_1,\ldots,\mathbf{v}_r\}$ er en basis for V. Vi definerer **dimensjonen** til underrommet V til å være antall elementer r i basisen $\mathscr{B}$. Siden alle basiser i henhold til Steinitz' erstatnings-resultat må ha like mange elementer, er dimensjonen til V uavhengig av valg av basis $\mathscr{B}$.*

Eksempel 3.1.8. *Dimensjonen til $\mathbb{R}^n$ er n. Dette kan vi se siden alle elementer i $\mathbb{R}^n$ er på formen $(a_1,\ldots,a_n)$ og*

$$(a_1,\ldots,a_n) = \sum_{i=1}^{n} a_i\mathbf{e}_i$$

*der $\mathbf{e}_i = (0,0,\ldots,0,1,0,\ldots,0)$ er et n-tuppel med en 1-er på plass i og resten 0-er. Vi kaller $\mathscr{S} = \{\mathbf{e}_1,\ldots,\mathbf{e}_n\}$ for **standardbasisen** for $\mathbb{R}^n$. At mengden er lineært uavhengig følger av at dersom*

$$(a_1,\ldots,a_n) = \sum_{i=1}^{n} a_i\mathbf{e}_i = \mathbf{0}$$

så må alle $a_i = 0$.

Det siste eksemplet viser at dimensjons-begrepet passer med vår intuitive oppfatning av dimensjon, planet $\mathbb{R}^2$ har dimensjon 2 og rommet $\mathbb{R}^3$ har dimensjon 3. Dimensjonen til det trivielle underrommet $V = \{0\}$ settes lik 0.

3.2 Lineære avbildninger

Etter å ha definert lineære underrom av $\mathbb{R}^n$ er det naturlig å se på avbildninger (funksjoner) $T : \mathbb{R}^n \to \mathbb{R}^m$ som respekterer den lineære strukturen.

Definisjon 3.2.1. *En avbildning $T : \mathbb{R}^n \to \mathbb{R}^m$ kalles en **lineæravbildning** (eller en **lineær avbildning**) dersom*

$$T(a\mathbf{x}+b\mathbf{y}) = aT(\mathbf{x})+bT(\mathbf{y})$$

*for alle $\mathbf{x},\mathbf{y} \in \mathbb{R}^n$ og $a,b \in \mathbb{R}$. Dersom $m = n$, kaller vi avbildningen $T : \mathbb{R}^n \to \mathbb{R}^n$ en **lineær operator**.*

Proposisjon 3.2.2. *La $T : \mathbb{R}^n \to \mathbb{R}^m$ være en lineæravbildning og la $\mathbf{v} \in \mathbb{R}^n$. Da har vi at $T(\mathbf{0}) = \mathbf{0}$ og $T(-\mathbf{v}) = -T(\mathbf{v})$.*

Bevis. Vi har

$$T(\mathbf{0}) = T(0 \cdot \mathbf{v}) = 0 \cdot T(\mathbf{v}) = \mathbf{0}$$

og

$$T(-\mathbf{v}) + T(\mathbf{v}) = T(-\mathbf{v} + \mathbf{v}) = T(\mathbf{0}) = \mathbf{0}$$

Det følger at $T(-\mathbf{v}) = -T(\mathbf{v})$. □

Eksempel 3.2.1. *En lineær avbildning $T : \mathbb{R} \to \mathbb{R}$ må være på formen $T(x) = ax$ for et vilkårlig reellt tall $a \in \mathbb{R}$. Dette følger av at $T(x) = xT(1)$, og at vi setter $T(1) = a$.*

Eksempel 3.2.2. *Avbildningen gitt ved*

$$T(x_1, x_2) = (x_1 + x_2, x_2 - 2x_1)$$

er en lineær operator på $\mathbb{R}^2$.

La A være en $m \times n$-matrise og definer avbildningen $T : \mathbb{R}^n \to \mathbb{R}^m$ ved

$$T(\mathbf{v}) = A\mathbf{v}$$

for en vektor $\mathbf{v} \in \mathbb{R}^n$. Da har vi for vektorer $\mathbf{v}$ og $\mathbf{w}$ og reelle tall $a, b \in \mathbf{R}$

$$\begin{aligned} T(a\mathbf{v} + b\mathbf{w}) &= A(a\mathbf{v} + b\mathbf{w}) \\ &= aA\mathbf{v} + bA\mathbf{w} \\ &= aT(\mathbf{v}) + bT(\mathbf{w}) \end{aligned}$$

som viser at multiplikasjon med en matrise gir oss en lineæravbildning.

Dette resultatet gjelder også den andre veien, i den forstand at enhver lineæravbildning kan skrives som et produkt med en matrise. I stedet for å skrive opp et teorem gir vi her en oppskrift på hvordan matrisen A kan finnes fra T. Vi starter med standardbasisen $\mathscr{S} = \{e_1, e_2, \ldots, e_n\}$ for $\mathbb{R}^n$. Enhver vektor $\mathbf{x}$ i $\mathbb{R}^n$ har da en unik representasjon

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n$$

Siden T er en lineæravbildning, får vi

$$T(\mathbf{x}) = x_1T(\mathbf{e}_1) + x_2T(\mathbf{e}_2) + \cdots + x_nT(\mathbf{e}_n)$$

Lar vi A være matrisen med $T(\mathbf{e}_i)$ som kolonner;

$$A = (T(\mathbf{e}_1) \mid T(\mathbf{e}_2) \mid \ldots \mid T(\mathbf{e}_n))$$

så får vi en $m \times n$-matrise som oppfyller

$$A\mathbf{x} = x_1T(\mathbf{e}_1) + x_2T(\mathbf{e}_2) + \ldots x_nT(\mathbf{e}_n) = T(\mathbf{x})$$

Vi sier at A er matriseformen til lineæravbildningen T relativt til standardbasisen og skriver $A = [T]_{\mathscr{S}}$.

Eksempel 3.2.3. *La $T : \mathbb{R}^2 \to \mathbb{R}^3$ være den lineære avbildningen*

$$T(x, y) = \begin{pmatrix} x + y \\ x - y \\ 2x \end{pmatrix}$$

Standardbasisen for $\mathbb{R}^2$ er gitt ved $\mathscr{S} = \{\mathbf{e}_1, \mathbf{e}_2\}$ hvor

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$T(\mathbf{e}_1) = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, \qquad T(\mathbf{e}_2) = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$$

slik at

$$[T]_{\mathscr{S}} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 2 & 0 \end{pmatrix}$$

Vi kjenner igjen $\mathscr{B} = \{T(\mathbf{e}_1), T(\mathbf{e}_2)\}$ som basisen til underrommet fra Eksempel 3.1.4. Altså avbilder T (det vanlige) planet $\mathbb{R}^2$ til underrommet av $\mathbb{R}^3$ beskrevet ved ligningen x+y-z = 0, som jo definerer et plan.

Eksempel 3.2.4. *Vi betrakter lineæravbildningen $T : \mathbb{R}^2 \to \mathbb{R}^2$ gitt ved*

$$T(x_1, x_2) = \begin{pmatrix} x_1 + x_2 \\ x_2 - 2x_1 \end{pmatrix}$$

La $\mathscr{S}$ være standardbasisen for $\mathbb{R}^2$;

$$\mathscr{S} = \{\mathbf{e}_1, \mathbf{e}_2\}$$

Da har vi at

$$T(\mathbf{e}_1) = T(1,0) = \begin{pmatrix} 1 \\ -2 \end{pmatrix} \qquad T(\mathbf{e}_2) = T(0,1) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

og

$$[T]_{\mathscr{S}} = \begin{pmatrix} 1 & 1 \\ -2 & 1 \end{pmatrix}$$

Lineæravbildningen blir nå gitt ved matrisemultiplikasjon;

$$[T]_{\mathscr{S}} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ x_2 - 2x_1 \end{pmatrix}$$

Et litt forvirrende faktum er at ulike matriser kan beskrive samme lineæravbildning. Dette skyldes at matrisen som beskriver en lineæravbildning ikke bare er bestemt av avbildningen, men også av hvilken basis vi opererer med. Hvis vi skifter til en annen basis, får vi en annen matrise, selv om lineæravbilnigen er den samme.

For å studere lineære avbildninger nærmere skal vi introdusere noen lineære underrom som vil hjelpe oss i kartleggingen. Vi begynner med å definere **søylerommet** til en matrise.

Definisjon 3.2.3. ***Søylerommet*** *til $m \times n$-matrisen*

$$A = (\mathbf{a}_1 \mid \mathbf{a}_2 \mid \ldots \mid \mathbf{a}_n)$$

*er underrommet col $A = span\{\mathbf{a}_1, \mathbf{a}_2, \ldots . \mathbf{a}_n a) \subset \mathbb{R}^m$ Dimensjonen til col A kalles **rangen** til A og vi skriver rg $A = dim(col\, A)$.*

Hvis $\mathbf{y}$ er en vektor i col A, så finnes tall $x_1, x_2, \ldots, x_n$ slik at

$$\mathbf{y} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n = A\mathbf{x}$$

der $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T$. Vi ser fra dette at col A består av vektorene $\mathbf{y}$ i $\mathbb{R}^m$ slik at det finnes en vektor $\mathbf{x} \in \mathbb{R}^n$ som oppfyller $A\mathbf{x} = \mathbf{y}$. Hvis vi tenker på A som en lineær avbildning fra $\mathbb{R}^n$ til $\mathbb{R}^m$ vil kolonnerommet være nøyaktig det samme som bildet av avbildningen.

Eksempel 3.2.5. *La oss betrakte matrisene*

$$A_1 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 2 & 0 \end{pmatrix} \quad og \quad A_2 = \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 1 \\ 2 & 0 & 1 \end{pmatrix}$$

der A_1 er en lineær avbildning fra $\mathbb{R}^2$ til $\mathbb{R}^3$ og A_2 en lineær avbildning fra $\mathbb{R}^3$ til $\mathbb{R}^3$. Fra Eksempel 3.1.4 og den påfølgende diskusjonen ser vi at

$$col\, A_1 = col\, A_2 = span\left\{ \begin{pmatrix} a+b \\ a-b \\ 2a \end{pmatrix} \,\middle|\, a, b \in \mathbb{R} \right\}$$

Det betyr at rg A_1 = rg $A_2 = 2$

Siden col A er et underrom av $\mathbb{R}^m$, så kan ikke rg A bli større enn m. Hvis $n > m$ betyr dette at A sender det n-dimensjonale $\mathbb{R}^n$ til et underrom av $\mathbb{R}^m$ av mindre dimensjon. Altså er det naturlig å tenke at noe går tapt i den lineær avbildningen. Følgende definisjon fanger opp dette.

Definisjon 3.2.4. *La A være en $m \times n$-matrise. **Nullrommet** til A er underrommet av $\mathbb{R}^n$ som består av de vektorene $\mathbf{x}$ som oppfyller*

$$A\mathbf{x} = \mathbf{0}$$

*Vi skriver $\mathcal{N}(A)$ for nullrommet til A. Dimensjonen til $\mathcal{N}(A)$ kalles **nulliteten** til A og vi skriver nullity $A = dim(\mathcal{N}(A))$.*

Siden $\mathcal{N}(A)$ er et underrom av $\mathbb{R}^n$ vet vi at nullity A ikke kan bli større enn n. Definisjonen av nullrom har sin opplagte parallell for lineæravbildninger.

Definisjon 3.2.5. *La $T : V \to W$ være en lineæravbildning. Mengden*

$$\mathcal{N}(T) = \{\mathbf{v} \in V \,|\, T(\mathbf{v}) = \mathbf{0}\}$$

*kalles **nullrommet** til T.*

Det er kanskje greit å dobbeltsjekke at $\mathcal{N}(A)$ faktisk er et linært underrom.

Proposisjon 3.2.6. *Nullrommet $\mathcal{N}(A)$ til A er et lineært underrom av V.*

Bevis. La $\mathbf{v}, \mathbf{w} \in \mathcal{N}(A)$, og $a, b \in \mathbb{R}$. Da har vi

$$A(a\mathbf{v} + b\mathbf{w}) = aA\mathbf{v} + bA\mathbf{w} = a \cdot \mathbf{0} + b \cdot \mathbf{0} = \mathbf{0}$$

som viser at nullrommet er et lineært underrom. □

Eksempel 3.2.6. *Vi betrakter en lineæravbildning $T : \mathbb{R}^3 \to \mathbb{R}^3$ gitt ved*

$$T(x_1,x_2,x_3) = (x_1+x_2+x_3, x_1-x_2+2x_3, 2x_1+3x_3)$$

Vi er interessert i å finne tripler (x_1,x_2,x_3) slik at $T(x_1,x_2,x_3) = (0,0,0)$. Det betyr at vi må løse likningssystemet

$$\begin{aligned} x_1+x_2+x_3 &= 0 \\ x_1-x_2+2x_3 &= 0 \\ 2x_1+3x_3 &= 0 \end{aligned}$$

Her observerer vi at den tredje likningen faktisk er summen av de to første, og derfor er overflødig. På den annen side gir den tredje likningen at

$$x_3 = -\frac{2}{3}x_1$$

og setter vi det inn i den første likningen, får vi

$$x_2 = -x_1 - x_3 = -x_1 + \frac{2}{3}x_1 = -\frac{1}{3}x_1$$

Samler vi disse opplysningene ser vi at

$$T(x_1, -\frac{1}{3}x_1, -\frac{2}{3}x_1) = (0,0,0)$$

Alle vektorer på formen

$$(x_1, -\frac{1}{3}x_1, -\frac{2}{3}x_1), \quad x_1 \in \mathbb{R}$$

utgjør et vektorrom, det såkalte nullrommet til avbildningen T.

Eksempel 3.2.7. *La $T : \mathbb{R}^2 \to \mathbb{R}^2$ være gitt ved*

$$T(\mathbf{v}) = A\mathbf{v}$$

der A er matrisen

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

Vi har

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

og $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ ligger derfor i nullrommet til T. Faktisk vil denne vektoren alene danne en basis for nullrommet til T.

Eksempel 3.2.8. *La $T : \mathbb{R}^3 \to \mathbb{R}^3$ være gitt ved*

$$T(\mathbf{v}) = A\mathbf{v}$$

der A er matrisen

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

Dersom en vektor

$$\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$$

ligger i nullrommet til T, betyr det at

$$v_1 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + v_2 \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + v_3 \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} = \mathbf{0}$$

Men vi har i et tidligere eksempel sett at søylevektorene til A er lineært uavhengig, som betyr at

$$\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \mathbf{0}$$

og nullrommet bestå kun av nullvektoren.

Proposisjon 3.2.7. *La $T : V \to V$ være en lineær avbildning gitt ved en kvadratisk matrise A;*

$$T(\mathbf{v}) = A\mathbf{v}$$

Da har vi at $\mathscr{N}(T) = \mathbf{0}$ hvis og bare hvis A er invertibel.

Bevis. Anta A er invertibel og la $T(\mathbf{v}) = A\mathbf{v} = \mathbf{0}$. Det betyr at $\mathbf{v} = A^{-1}\mathbf{0} = \mathbf{0}$. Anta så at $\mathscr{N}(T) = \mathbf{0}$. La $\mathbf{v} = (v_1, \ldots, v_n)$ være en lineær avhengighet mellom søylene i A. Da har vi at $A\mathbf{v} = \mathbf{0}$. Siden nullrommet er trivielt betyr det at $\mathbf{v} = \mathbf{0}$ som igjen betyr at søylene i mtraisen er lineært uavhengig. Det følger at A er invertibel. □

La oss ta et steg tilbake og tenke på en $m \times n$-matrise A som en lineær avbildning. Vi starter her i $\mathbb{R}^n$ hvor vi har n dimensjoner å boltre oss i. Hvis nullity $A = k$ for et tall $0 \leq k \leq n$ kan vi se for oss at vi sender k dimensjoner til 0. Da er det ikke fullstendig unaturlig å se for seg at col A består av de resterende $n - k$ dimensjonene. Det viser ser også å være tilfelle.

Det er en nær sammenheng mellom rangen til en lineær avbildning og dimensjonen til nullrommet. Vi har sett at vi kan uttrykke lineære sammenhenger mellom søylene i en matrise ved n-tupler, slik som i et tidligere eksempel: Vektorene

$$\begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \end{pmatrix}$$

er lineært avhengig og vi har

$$2\begin{pmatrix} 2 \\ 1 \end{pmatrix} - 2\begin{pmatrix} 1 \\ 0 \end{pmatrix} - \begin{pmatrix} 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Setter vi de tre vektorene sammen i en 2×3-matrise

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

så vil den lineære avhengigheten gi oss et en vektor i nullrommet;

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

Dette gir oss en korrespondanse mellom nullvektorer og lineære avhengigheter mellom søylene i matrisen. Vi kan formulere dette i et generelt resultat:

Teorem 3.2.8. *En lineæravbildning $T : \mathbb{R}^n \to \mathbb{R}^m$ er gitt ved $T(\mathbf{v}) = A\mathbf{v}$ for en $m \times n$-matrise A. Da har vi*

$$\operatorname{rg} T + \dim \mathscr{N}(T) = n$$

Bevis. La $\mathscr{B}_0 = \{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ være en basis for nullrommet til A. Vi kan utvide denne basisen til en basis for hele $\mathbb{R}^n$, $\mathscr{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_r, \mathbf{w}_{r+1}, \ldots, \mathbf{w}_n\}$. Vi skal vise at $\operatorname{rg} T = n - r$. La $\mathbf{u} = T(\mathbf{v})$. Siden $\mathscr{B}$ er en basis for $\mathbb{R}^n$ kan vi skrive

$$\mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_r\mathbf{v}_r + a_{r+1}\mathbf{w}_{r+1} + \cdots + a_n\mathbf{w}_n$$

og

$$\begin{aligned}\mathbf{u} = T(\mathbf{v}) &= a_1T(\mathbf{v}_1) + \cdots + a_rT(\mathbf{v}_r) + a_{r+1}T(\mathbf{w}_{r+1}) + \cdots + a_nT(\mathbf{w}_n) \\ &= a_{r+1}T(\mathbf{w}_{r+1}) + \cdots + a_nT(\mathbf{w}_n)\end{aligned}$$

Det betyr at $T(\mathbf{w}_{r+1}), \ldots, T(\mathbf{w}_n)$ spenner ut $T(\mathbb{R}^n)$. Samtidig har vi at dersom

$$\begin{aligned}&a_{r+1}T(\mathbf{w}_{r+1}) + \cdots + a_nT(\mathbf{w}_n) \\ &\quad = T(a_{r+1}\mathbf{w}_{r+1} + \cdots + a_n\mathbf{w}_n) = \mathbf{0}\end{aligned}$$

så ligger $a_{r+1}\mathbf{w}_{r+1} + \cdots + a_n\mathbf{w}_n$ i nullrommet til A, dvs.

$$a_{r+1}\mathbf{w}_{r+1} + \cdots + a_n\mathbf{w}_n = a_1\mathbf{v}_1 + \cdots + a_r\mathbf{v}_r$$

Dermed kan vi skrive

$$a_1\mathbf{v}_1 + \cdots + a_r\mathbf{v}_r - a_{r+1}\mathbf{w}_{r+1} - \cdots - a_n\mathbf{w}_n = \mathbf{0}$$

og siden $\mathscr{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_r, \mathbf{w}_{r+1}, \ldots, \mathbf{w}_n\}$ er en basis for $\mathbb{R}^n$ og derfor danner en lineært uavhengig mengde, så må vi ha $a_{r+1} = \cdots = a_n = 0$ og mengden $\{T(\mathbf{w}_{r+1}), \ldots, T(\mathbf{w}_n)\}$ er lineært uavhengig. Det følger at dimensjonen til $T(\mathbb{R}^n)$ som er det samme som rangen til T, er $n - r$. □

Eksempel 3.2.9. *La $T : \mathbb{R}^2 \to \mathbb{R}^2$ være gitt ved*

$$T(\mathbf{v}) = A\mathbf{v}$$

der A er matrisen

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

Søylerommet til A er generert av

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

og nullrommet av

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

Dermed har vi

$$\operatorname{rg} T + \dim \mathscr{N}(T) = 1 + 1 = 2$$

Siden det er n søyler i en $n \times n$-matrise vil dimensjonen til søylerommet alltid være mindre enn eller lik n. Dersom søylene utgjør en lineært uavhengig mengde vil matrisen ha **maksimal rang**, det vil si n for en $n \times n$-matrise. Maksimal rang betyr at søylene utspenner hele $\mathbb{R}^n$.

Proposisjon 3.2.9. *En $n \times n$-matrise A har maksimal rang hvis og bare hvis matrisen er invertibel.*

Bevis. Observer først at dersom en vektor $\mathbf{w}$ ligger i spennet av søylevektorene til A, så er det ekvivalent med at det finnes en annen vektor $\mathbf{v}$ slik at $A\mathbf{v} = \mathbf{w}$.

Anta at A er invertibel og $\mathbf{w}$ en vektor i $\mathbb{R}^n$. Sett $\mathbf{v} = A^{-1}\mathbf{w}$. Da har vi

$$A\mathbf{v} = AA^{-1}\mathbf{w} = \mathbf{w}$$

og $\mathbf{w}$ ligger i søylerommet. Anta så at matrisen har maksimal rang. Det betyr at for en hver vektor $\mathbf{w}$ i søylerommet til A, så finnes det en vektor $\mathbf{v}$ slik at $A\mathbf{v} = \mathbf{w}$. Siden A har maksimal rang, vil standardvektorene $\mathbf{e}_i$ ligge i søylerommet til A. La $\mathbf{v}_i$ oppfylle $A\mathbf{v}_i = \mathbf{e}_i$, og la B være $n \times n$-matrisen med $\mathbf{v}_i$ som sine søyler. Da vil

$$AB = I$$

som viser at $B = A^{-1}$ og A er invertibel. □

Teorem 3.2.10. *La $A = (\mathbf{a}_1 \mid \mathbf{a}_2 \mid \ldots \mid \mathbf{a}_n)$ være en $n \times n$-matrise. Følgende utsagn er ekvivalente:*

(a) A er invertibel.

(b) $\det A \neq 0$.

(c) Mengden $\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n\}$ er lineært uavhengig.

(d) $\mathrm{rg}A = n$.

(e) $\mathrm{nullity}A = 0$.

Bevis. Vi har allerede sett at (a), (b) og (c) er ekvivalente. Det følger av (c) at $\mathrm{col}A = \mathbb{R}^n$, så $\mathrm{rg}A = \dim A = 0$. Fra dimensjonsteoremet får vi at a $\mathrm{nullity}A = n - \mathrm{rg}A = 0$, så (d) impliserer (e). Hvis A ikke er invertibel, så finnes en $\mathbf{x} \neq \mathbf{0}$ slik at $A\mathbf{x} = \mathbf{0}$. Dette viser at (e) impliserer (a). □

I et tidligere kapittel definerte vi egenvektorer og egenverdier for en matrise. Vi skal se at egenvektorer som svarer til forskjellige egenverdier danner lineært uavhengige mengder.

Proposisjon 3.2.11. *La $T : \mathbb{R}^n \to \mathbb{R}^n$ være en lineæravbildning gitt ved $T(\mathbf{v}) = A\mathbf{v}$ for en $n \times n$-matrise A. La videre $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ være egenvektorer for A svarende til ulike egenverdier $\lambda_1, \lambda_2, \ldots, \lambda_r$. Da er mengden*

$$\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$$

lineært uavhengig.

Bevis. La først $r = 2$, og anta at $a_1\mathbf{v}_1 + a_2\mathbf{v}_2 = \mathbf{0}$. Anta at det finnes en lineær avhengighet mellom de to vektorene, anta f.eks at $a_1 \neq 0$. Da kan vi skrive

$$\mathbf{v}_1 = -\frac{a_2}{a_1}\mathbf{v}_2$$

og siden $\mathbf{v}_1$ og $\mathbf{v}_2$ begge er egenvektorer får vi

$$T(\mathbf{v}_1) = \lambda_1\mathbf{v}_1 = -\frac{a_2}{a_1}\lambda_1\mathbf{v}_2$$

og samtidig

$$T(\mathbf{v}_1) = T(-\frac{a_2}{a_1}\mathbf{v}_2) = -\frac{a_2}{a_1}T(\mathbf{v}_2) = -\frac{a_2}{a_1}\lambda_2\mathbf{v}_2$$

Det følger at

$$-\frac{a_2}{a_1}\lambda_1 = -\frac{a_2}{a_1}\lambda_2$$

og siden $\lambda_1 \neq \lambda_2$ så må $a_2 = 0$. Siden $\mathbf{v}_1 \neq \mathbf{0}$ følger det at også $a_1 = 0$, som gir oss en motsetning.

Anta så at $r = 3$ og at vi har en lineær avhengighet

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + a_3\mathbf{v}_3 = \mathbf{0}$$

Vi kan uten videre anta at $a_1 \neq 0$ og ved å dele likningen på a_1 kan vi like godt sette a_1 til å være 1. Det gir

$$\mathbf{v}_1 = -a_2\mathbf{v}_2 - a_3\mathbf{v}_3$$

Bruker vi nå T på dette uttrykket får vi

$$\lambda_1\mathbf{v}_1 = -a_2\lambda_2\mathbf{v}_2 - a_3\lambda_3\mathbf{v}_3$$

Vi multipliserer uttrykket over med λ_1 og trekker de to likningene fra hverandre. Det gir

$$\mathbf{0} = -a_2(\lambda_2 - \lambda_1)\mathbf{v}_2 - a_3(\lambda_3 - \lambda_1)\mathbf{v}_3$$

Dermed er vi tilbake i situasjonen med to egenvektorer og siden $\lambda_2 - \lambda_1, \lambda_3 - \lambda_1 \neq 0$ følger det at koeffisientene $a_2 = a_3 = 0$, og vektorene er lineært uavhengige. For $r > 3$ følger vi nå nøyaktig samme prosedyre. □

Vi har sett at en $n \times n$-matrise kan ha maksimalt n egenverdier. Anta nå at A har nøyaktig n forskjellige egenverdier. Det betyr at vi kan finne n egenvektorer som utgjør en lineært uavhengig mengde. Siden vektorrommet vi opererer på i dette tilfellet er av dimensjon n betyr det at vi kan finne en basis av egenvektorer. Vi formulerer dette i en proposisjon.

Proposisjon 3.2.12. *La V være et n-dimensjonalt vektorrom. La $T : V \to V$ være en lineæravbildning gitt ved $T(\mathbf{v}) = A\mathbf{v}$ for en $n \times n$-matrise A. Anta at A har n forskjellige egenverdier. Da finnes en basis for V av egenvektorer for A.*

Eksempel 3.2.10. *La A være 2×2-matrisen*

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

Det karakteristiske polynomet til A er gitt ved

$$\chi_A(\lambda) = \lambda^2 - 3\lambda + 2$$

som gir egenverdier $\lambda_1 = 1$, $\lambda_2 = 2$, *med tilhørende egenvektorer*

$$\mathbf{v}_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$$

Vi ser at mengden

$$\{\mathbf{v}_1, \mathbf{v}_2\}$$

er lineært uavhengig og danner en basis for $\mathbb{R}^2$.

For stokastiske matriser, dvs. matriser med søylesum 1, kan vi si noe generelt om egenvektorene.

La P være en stokastisk matrise, med egenverdier $\lambda_1 = 1 > \lambda_2 \geq \cdots \geq \lambda_n$ og tilhørende egenvektorer $\mathbf{v}_i$ for $i = 1, 2, \ldots$.

Definer $\sigma(\mathbf{v}) = v_1 + \cdots + v_m$ til å være komponentsummen til vektoren $\mathbf{v}$.

Proposisjon 3.2.13. *Alle egenvektorene til en stokastisk matris, bortsett fra de med egenverdi 1, har komponentsum lik 0, dvs.*

$$\sigma(\mathbf{v}_i) = 0 \quad i = 2, 3, \ldots, n$$

Bevis. Vi har at $P(\mathbf{v}_i) = \lambda_i \mathbf{v}_i$ og siden P er stokastisk, så er $\sum_j p_{ij} = 1$ for alle i. La $\mathbf{v}_i = (v_1, \ldots, v_m)$. Da har vi

$$\sigma(P(\mathbf{v}_i)) = \sum_{j=1}^{m} \sum_{k=1}^{m} p_{jk} v_k = \sum_{j=1}^{m} v_k = \sigma(\mathbf{v}_i)$$

og samtidig

$$\sigma(P(\mathbf{v}_i)) = \sigma(\lambda \mathbf{v}_i) = \lambda \sigma(\mathbf{v}_i)$$

Det gir at

$$\sigma(\mathbf{v}_i) = \lambda \sigma(\mathbf{v}_i)$$

og derfor $\sigma(\mathbf{v}_i) = 0$ siden $\lambda \neq 1$. □

En størrelse som er nært knyttet til komponentsummen er den såkalte **Taxi-lengden** til en vektor. Taxi-lengden regnes ut som komponentsummen til vektoren vi får ved å ta absoluttverdien av alle komponentene,

Definisjon 3.2.14. *La* $\mathbf{v} = (v_1, \ldots, v_n)$ *være en vektor i* $\mathbb{R}^n$. *Vi definerer* ***Taxi-lengden*** $|\mathbf{v}|_1$ *til* $\mathbf{v}$ *ved*

$$|\mathbf{v}|_1 = \sum_{k=1}^{n} |v_k|$$

Anta nå at vi har et lukket dynamisk system, gitt ved en stokastisk matrise P. For en tilstand for systemet, dvs. for en vektor $\mathbf{w}$, lar vi $E(\mathbf{w}) = |\mathbf{v}_1|_1$ uttrykke **energien** i tilstanden, mens $S(\mathbf{w}) = |\mathbf{v}_2|_1$, uttrykker **entropien**. Begrepene er hentet fra termodynamikk, ikke uten grunn. I analogien med termodynamikk er det naturlig å knytte den nest største egenverdien λ_2 til tilstandens temperatur gjennom formelen

$$T(P) = \log \frac{1}{|\lambda_2|}$$

Tilstanden $\mathbf{w}$ kan skrives som en lineærkombinasjon av egenvektorer, som vi skriver om og får

$$\begin{aligned}\mathbf{w} &= \mathbf{v}_1 + \mathbf{v}_2 + \cdots + \mathbf{v}_m \\ &= E(\mathbf{w})\frac{\mathbf{v}_1}{|\mathbf{v}_1|_1} + S(\mathbf{w})\frac{\mathbf{v}_2}{|\mathbf{v}_2|_1} + \cdots + \mathbf{v}_m\end{aligned}$$

Multiplikasjon med den stokastiske matrisen P gir

$$\begin{aligned}P\mathbf{w} &= P\mathbf{v}_1 + P\mathbf{v}_2 + \cdots + P\mathbf{v}_m \\ &= E(\mathbf{w})\frac{\mathbf{v}_1}{|\mathbf{v}_1|_1} + \lambda_2 S(\mathbf{w})\frac{\mathbf{v}_2}{|\mathbf{v}_2|_1} + \cdots + \lambda_m\mathbf{v}_m\end{aligned}$$

og dermed

$$S(P\mathbf{w}) = \lambda_2 S(\mathbf{w})$$

eller

$$T(P) = \log\frac{S(\mathbf{w})}{S(P\mathbf{w})}$$

som gir

$$\frac{S(P\mathbf{w})}{S(\mathbf{w})} = e^{-T(P)}$$

eller

$$\frac{S(P\mathbf{w}) - S(\mathbf{w})}{S(\mathbf{w})} = e^{-T(P)} - 1$$

Dette kan vi tolke som at jo lavere temperatur, dvs. $T(p) \to 0$, så vil entropi-endringen også nærme seg 0. Det betyr at det ikke skjer noen endring i systemet. For høye temperaturer vil vi observere en stor endring i entropien, med andre ord systemet vil rask gå mot likevektstilstanden, hvor vi har den laveste entropien (dvs. størst mulig uorden).

Eksempel 3.2.11. *Vi har gitt en stokastisk 2×2-matrise*

$$P = \begin{pmatrix} p & q \\ 1-p & 1-q \end{pmatrix}$$

Vi kan regne ut egenverdiene ved å finne nullpunktene til det karakteristiske polynomet

$$\det(P - \lambda I) = \lambda^2 - (p+1-q)\lambda + (p-q) = 0$$

som gir $\lambda_1 = 1$ og $\lambda_2 = p-q$. De tilhørende egenvektorene er (normert til taxi-lengde 1)

$$\mathbf{v}_1 = \frac{1}{p+1-q}(q, 1-p), \quad \mathbf{v}_2 = \frac{1}{2}(1,-1)$$

For en tilstandsvektor har vi

$$\mathbf{w} = \begin{pmatrix} x \\ y \end{pmatrix} = (x+y)\mathbf{v}_1 + 2\frac{(1-p)x - qy}{1+q-p}\mathbf{v}_2$$

som gir at

$$\begin{aligned}E(\mathbf{w}) &= x+y \\ S(\mathbf{w}) &= 2\frac{(1-p)x - qy}{1+q-p}\end{aligned}$$

og temperatur

$$T(P) = \log\frac{1}{|p-q|}$$

3.3 Generelle vektorrom*

I dette avsnittet skal vi se hvordan vi kan generalisere teorien for lineære underrom til en mer abstrakt setting hvor V ikke lenger er en delmengde av $\mathbb{R}^n$. Det generaliserte begrepet kalles et (abstrakt) **vektorrom**.

Definisjon 3.3.1. *Et abstrakt **vektorrom** V (også kalt et **lineært rom**) er en mengde med to operasjoner, **addisjon**, dvs. at for $\mathbf{x},\mathbf{y} \in V$, så er summen $\mathbf{x}+\mathbf{y} \in V$, og **skalar multiplikasjon**, dvs. hvis $\mathbf{x} \in V$ og $a \in \mathbb{R}$ er et reelt tall, så er $a\mathbf{x} \in V$. De to operasjonene må oppfylle betingelsene:*

*i) Addisjon er **assosiativ**, dvs. $(\mathbf{x}+\mathbf{y})+\mathbf{z} = \mathbf{x}+(\mathbf{y}+\mathbf{z})$ for alle $\mathbf{x},\mathbf{y},\mathbf{z} \in V$.*

*ii) Addisjon er **kommutativ**, dvs. $\mathbf{x}+\mathbf{y} = \mathbf{y}+\mathbf{x}$ for alle $\mathbf{x},\mathbf{y} \in V$.*

*iii) Det finnes et **nøytralt addidtivt element** eller $\mathbf{0}$, slik at $\mathbf{x}+\mathbf{0} = \mathbf{x}$ for alle $\mathbf{x} \in V$.*

*iv) Til alle elementer $\mathbf{x} \in V$ så finnes det et element $-\mathbf{x}$ slik at $\mathbf{x}+(-\mathbf{x}) = 0$. Vi kaller elementet den **additive inversen** til $\mathbf{x}$.*

*v) Skalar multiplikasjon er **assosiativ**, dvs. for alle reelle tall $a,b \in \mathbb{R}$ og $\mathbf{x} \in V$, så er $a(b\mathbf{x}) = (ab)\mathbf{x}$.*

vi) Vi har $1 \cdot \mathbf{x} = \mathbf{x}$ for alle $\mathbf{x} \in V$.

*vii) De to operasjonene er **distributive over hverandre**, dvs. $(a+b)\mathbf{x} = a\mathbf{x}+b\mathbf{x}$ og $a(\mathbf{x}+\mathbf{y}) = a\mathbf{x}+a\mathbf{y}$ for alle $a,b \in \mathbb{R}$ og $\mathbf{x},\mathbf{y} \in V$.*

*Elementene i et vektorrom kalles ofte for **vektorer** (selv om de ikke ser ut som det vi vanligvis forbinder med vektorer), og de reelle tallene vi ganger med kalles **skalarer**.*

De lineære underrommene vi har sett på til nå er spesialtilfeller av vektorrom. De oppfyller alle betingelsene gitt i Definisjon 3.3.1, i tillegg til at de er delmengder av $\mathbb{R}^n$ for et positivt heltall n. I de neste eksemplene ser vi på en type vektorrom som har uendelig dimensjon og derfor ikke kan legges inn i noen $\mathbb{R}^n$.

Eksempel 3.3.1. *La $\mathscr{F}(I)$ være mengden av funksjoner definert på et intervall $I \subset \mathbb{R}$. Addisjon og skalar multiplikasjon er gitt som ordinær addisjon og multiplikasjon av funksjoner;*

$$(f+g)(x) = f(x)+g(x)$$
$$(cf)(x) = cf(x)$$

for alle $f,g \in \mathscr{F}(I)$ og $c \in \mathbb{R}$ og 0-elementet er konstantfunksjonen $f(x) = 0$. Mengden $\mathscr{F}(I)$ danner da et vektorrom av uendelig dimensjon.

Eksempel 3.3.2. *Mengden $C(I)$ av kontinuerlige funksjoner på et intervall I danner et undervektorrom av vektorrommet $\mathscr{F}(I)$ i Eksempel 3.3.1.*

Vi kan også restriktere oss til helt spesielle funksjoner, f.eks. vil mengden $\{ae^x + be^{-x} \mid a,b \in \mathbb{R}\}$ være lukket under addisjon og skalar multiplikasjon og derfor danne et vektorrom. Dette vektorrommet er imidlertid endelig-dimensjonalt av dimensjon 2 og avbildningen

$$\{ae^x + be^{-x} \mid a,b \in \mathbb{R}\} \to \mathbb{R}^2$$

der $ae^x + be^{-x} \mapsto (a,b)$ identifiserer vektorrommet med $\mathbb{R}^2$.

Som eksemplene viser trenger ikke et vektorrom å ha endelig dimensjon. Vi kan likevel skrive opp en basis, selv om basisen ikke er endelig.

Proposisjon 3.3.2. *La $\mathscr{P}$ være vektorrommet av polynomer med reelle koeffisienter. Mengden*

$$\mathscr{B} = \{1, x, x^2, x^3, \dots, \}$$

danner en basis for $\mathscr{P}$.

Bevis. Elementene i $\mathscr{P}$ er på formen

$$a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$$

for et ikke-negativt tall n og $a_i \in \mathbb{R}$. Det betyr at $\mathscr{B}$ spenner ut hele $\mathscr{P}$. Vi må også vise at $\mathscr{B}$ er lineært uavhengig. Anta det motsatte, det vil si at det finnes reelle tall $a_0, a_1, \dots, a_n$ slik at

$$a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n = 0$$

Ingen ikke-trivielle polynomer av grad n kan ha mer enn n røtter. La x være et reelt tall som ikke er blant disse røttene. Det betyr at

$$a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n \neq 0$$

og vår antagelse måtte i utgangspunktet ha vært feil. Dermed har vi vist at $\mathscr{B}$ er en basis for $\mathscr{P}$. □

En delmengde av et vektorrom som selv er et vektorrom kalles et **underrom**. For å vise at en delmengde er et underrom holder det å vise at mengden er lukket under addisjon og skalar multiplikasjon, dvs. at dersom $\mathbf{x}, \mathbf{y} \in V$ og $a, b \in \mathbb{R}$ er reelle tall, så er også $a\mathbf{x} + b\mathbf{y} \in V$. Merk at dette ene kravet inneholder både lukkethet under addisjon (sett $a = b = 1$) og skalar multiplikasjon (sett $b = 0$).

Definisjon 3.3.3. *La V være et vektorrom, og la $\mathscr{V} = \{\mathbf{v}_1, \mathbf{v}_2, \dots\}$ være en endelig eller uendelig delmengde av V. La W være det minste vektorrommet som inneholder $\mathscr{V}$. Vi sier at mengden $\mathscr{V}$ genererer W, og vi skriver*

$$W = span(\mathbf{v}_1, \mathbf{v}_2, \dots)$$

I dette tilfellet vil W utgjøre et underrom av V.
Et monom er et polynom med kun ett ledd, mono=en, poly=mange.

Eksempel 3.3.3. *Mengden av polynomer av grad mindre enn eller lik n, som vi skriver $\mathscr{P}_n$, utgjør et underrom av vektorrommet av alle polynomer. Rommet er generert av **monomene** av grad mindre enn eller lik n;*

$$\mathscr{P}_n = \langle 1, x, x^2, x^3, \dots, x^n \rangle$$

Eksempel 3.3.4. *Vi lar V være vektorrommet av alle deriverbare funksjoner på hele $\mathbb{R}$. Derivasjon er en lineær operator på dette rommet, dvs. hvis vi setter $D(f(x)) = f'(x)$ så har vi*

$$\begin{aligned} D(af(x) + bg(x)) &= \frac{d}{dx}(af(x) + bg(x)) \\ &= af'(x) + bg'(x) \\ &= aD(f(x)) + bD(g(x)) \end{aligned}$$

Eksempel 3.3.5. *La $\mathscr{T} = \langle \cos t, \sin t \rangle$. Da virker derivasjonsoperatoren D som en lineær operator på $\mathscr{T}$ fordi vi har*

$$D(\cos t) = -\sin t$$
$$D(\sin t) = \cos t$$

Som vi har sett av eksemplene kan elementene i et vektorrom gjerne være funksjoner. I slike tilfeller er det naturlig å bruke betegnelsen **egenfunksjoner** i stedet for egenvektorer, men definisjonen er den samme.

Definisjon 3.3.4. *La $V = \mathscr{F}(I)$ være et vektorrom av funksjoner, og $T : V \to V$ en lineær operator på dette rommet. En ikke-triviell funksjon f som oppfyller $T(f) = \lambda f$ for et reelt tall λ kalles en **egenfunksjon** for operatoren.*

Eksempel 3.3.6. *La $V = \{ae^x + be^{-x} \mid a, b \in \mathbb{R}\}$ og la $D : V \to V$ være derivasjonsoperatoren. Da er $f_1(x) = e^x$ en egenfunksjon for D med egenverdi $\lambda_1 = 1$ og $f_2(x) = e^{-x}$ en egenfunksjon med egenverdi $\lambda_2 = -1$.*

Eksempel 3.3.7. *La $\mathscr{T} = \{\sin t, \cos t\}$ og $\Delta = D^2 : \mathscr{T} \to \mathscr{T}$ være den 1-dimensjonale **Laplaceoperatoren**. Da er alle lineære kombinasjoner $a \sin x + b \cos x$ egenfunksjoner for Δ med egenverdi $\lambda = -1$;*

$$\begin{aligned} D^2(a\sin x + b\cos x) &= D(a\cos x - b\sin x) \\ &= -a\sin x - b\cos x \qquad\qquad = -(a\sin x + b\cos x) \end{aligned}$$

3.4 Oppgaver med løsning

Eksempel 3.4.1. *La $T : \mathbb{R}^3 \to \mathbb{R}^3$ være gitt ved $T(x_1,x_2,x_3) = (x_1 - 2x_2 + x_3, x_2 - x_3, 2x_1 - 2x_2)$.*

a) Vis at T er en lineæravbildning.

b) La $\mathscr{S} = \{\mathbf{e}_1 = \begin{pmatrix}1\\0\\0\end{pmatrix}, \mathbf{e}_2 = \begin{pmatrix}0\\1\\0\end{pmatrix}, \mathbf{e}_3 = \begin{pmatrix}0\\0\\1\end{pmatrix}\}$ være standardbasisen i $\mathbb{R}^3$. Skriv opp en 3×3-matrise $[T]$ der søylene er gitt ved $T(\mathbf{e}_i)$ for $i = 1,2,3$.

c) Finn dimensjonen til nullrommet til T.

d) Regn ut rangen til $[T]$ ved å bruke dimensjonsformelen.

Løsning.

1. **Metodevalg:** For å vise at vi har en lineæravbildning må vi regne ut $T(a\mathbf{x} + b\mathbf{y})$ og vise at dett er det samme som $aT(\mathbf{x}) + bT(\mathbf{y})$. Dimensjonen til nullrommet finner vi ved å telle antall frihetsgrader i likningssettet $T(\mathbf{x}) = \mathbf{0}$.

2. **Regning:** a) I første komponent har vi at

$$\begin{aligned}(ax_1 + by_1) - 2(ax_2 + by_2) + (ax_3 + by_3)\\ = a(x_1 - 2x_2 + x_3) + b(y_1 - 2y_2 + y_3)\end{aligned}$$

og tilsvarende for de andre komponentene.

$$\begin{aligned}[T] &= (T(\mathbf{e}_1)\ T(\mathbf{e}_2)\ T(\mathbf{e}_3))\\ &= \begin{pmatrix}1 & -2 & 1\\ 0 & 1 & -1\\ 2 & -2 & 0\end{pmatrix}\end{aligned}$$

Vi skal løse

$$\begin{aligned}T(x_1,x_2,x_3) &= (x_1 - 2x_2 + x_3, x_2 - x_3, 2x_1 - 2x_2)\\ &= (0,0,0)\end{aligned}$$

dvs. likningssytemet

$$\begin{aligned}x_1 - 2x_2 + x_3 &= 0\\ x_2 - x_3 &= 0\\ 2x_1 - 2x_2 &= 0\end{aligned}$$

Vi ser av de to siste likningene at $x_1 = x_2$ og $x_2 = x_3$. Dette er oppfylt i den første likningen og vi har en frihetsgrad i nullrommet, det vil si at dimensjonen er 1. Rangen blir da i henhold til dimensjonsformelen

$$\operatorname{rg} T = 3 - 1 = 2$$

Eksempel 3.4.2. *La A være 3×3-matrisen*

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

a) Finn rangen til A.

b) Bruk dimensjonsformelen til å regne ut dimensjonen til nullrommet til A.

c) Forklar hvorfor 0 er en egenverdi for A og finn en egenvektor med egenverdi 0.

Løsning.

1. **Metodevalg:** Rangen til A finner vi ved å lete etter lineært uavhengige søylevektorer. Det kan vi så sette inn i dimensjonsformelen for å finne dimensjonen til nullrommet.

Generelt har vi at elementer i nullrommet er det samme som egenvektorer med egenverdi 0.

2. **Regning:** Vi ser at de to første søylene er lineært uavhengig (ved inspeksjon av den tredje komponenten som er 0 i den ene vektoren og ikke i den andre). Det betyr at rangen er minst 2. Vi kan sjekke om vi har maksimal rang ved å regne ut determinanten. Den blir

$$\begin{vmatrix} 1 & 1 & 1 \\ 2 & -1 & 0 \\ 3 & 0 & 1 \end{vmatrix} = -1+0+0+3-2+0 = 0$$

som betyr at rangen er 2. Dimensjonsformelen gir da at dimensjonen til nullrommet er $3-2=1$. Det betyr at det er en ikke-triviell vektor i nullrommet, og denne er også en egenvektor med egenverdi 0.

Vi løser

$$A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

og finner at

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = x \cdot \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix}$$

Eksempel 3.4.3. *La V være vektorrommet generert av funksjonene*

$$\cos x, \sin x, \sin^2 x, \sin 2x, 1$$

a) Vis at funksjonene danner en lineært uavhengig mengde av funksjoner.

b) Hva er dimensjonen til V?

c) La D være derivasjonsoperatoren; $D(f(x)) = f'(x)$. Vis at D er en lineær operator på V, dvs. at D brukt på hver enkelt generator ligger i V.

Løsning.

1. **Metodevalg:** Vi kan prøve å sette opp en lineær avhengighet

$$a\cos x + b\sin x + c\sin^2 x + d\sin 2x + e\cdot 1 = 0$$

og så sette inn ulike verdier for x for langsom å tvinge alle koeffisientene til å bli 0. Når vi har gjort det kjenner vi antall lineære avhengigheter og vi kan da beregne dimensjonen.

For å vise at operatoren er en lineær operator må vi vise at den passer inn i formelen i definisjonen og at den deriverer alle generatorene på elementer i vektorrommet.

2. **Regning:** Vi tar utgangspunkt i

$$a\cos x + b\sin x + c\sin^2 x + d\sin 2x + e\cdot 1 = 0$$

Setter vi inn $x = 0$ får vi $a + e = 0$. Setter vi inn $x = \frac{\pi}{2}$ får vi $b + c + e = 0$. Setter vi inn $x = \frac{\pi}{3}$ får vi $\frac{a}{2} + \frac{b\sqrt{3}}{2} + \frac{3c}{4} + \frac{d\sqrt{3}}{2} + e = 0$. Slik kan vi fortsette med ulike valg og vi ser etter hvert at eneste løsning er at alle koeffisientene er 0. Dimensjonen til rommet blir da 5.

Derivasjon er en lineær operator siden

$$\frac{\partial}{\partial x}(af(x) + bg(x)) = a\frac{\partial f}{\partial x} + b\frac{\partial g}{\partial x}$$

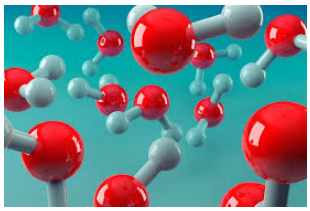
For elementene har vi

$$\begin{aligned} D(\cos x) &= -\sin x \\ D(\sin x) &= \cos x \\ D(\sin^2 x) &= 2\sin x\cos x = \sin 2x \\ D(\sin 2x) &= 2\cos 2x = 2(1 - 2\sin^2 x) \\ D(1) &= 0 \end{aligned}$$

Kapittel 4

Geometri i planet

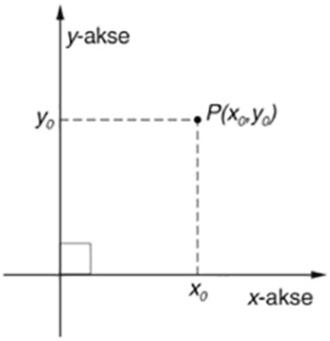
I dette og det neste kapitlet skal vi studere vektorrom i 2 og 3 dimensjoner, dvs. $\mathbb{R}^2$ og $\mathbb{R}^3$. Vi har valgt å kalle kapitlene geometri i planet eller rommet fordi vi i utgangspanktet skal bruke den generelle teorien for vektorrom til å forstå geometrien i planet og rommet. Mye av det vi gjør her vil være sentralt grunnlagsstoff når vi siden skal studere symmetrier av geometriske objekter, f.eks. molekyler eller krystaller.



4.1 Geometrisk tolkning av lineære avbildninger

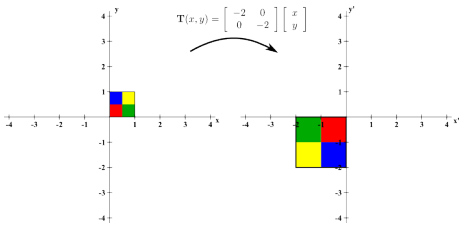
Vi starter med å studere vektorer i planet. Vi vil etter hvert ha mer bruk for å ha god kunnskap om romlige vektorer, men det er enklere å gi en innføring i teorien for vektorer i planet, for så å generalisere til vektorer i rommet.

Som vi har sett tidligere bruker vi alltid notasjonen $\mathbb{R}^2$ for det reelle planet. Normalt vil vi bruke samme notasjon for et punkt og en vektor. Når vi har bruk for begge deler samtidig, vil det framgå av sammenhengen hva som er hva. Punkter i planet vil vi da betegne med en stor bokstav, f.eks. $P \in \mathbb{R}^2$, mens en vektor som oftes uttrykkes med en liten, boldface bokstav, $\mathbf{v} \in \mathbb{R}^2$. Begge deler spesifiseres med et par av reelle tall (a,b). Tallene a og b er punktets eller vektorens koordinater. **Standardbasisen** $\{\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}\}$ danner et rettvinklet koordinatsystem. Punktet $P = P(x_0, y_0)$ på figuren har x-koordinat x_0 og y-koordinat y_0.



Figur 4.1. *Et rettvinklet koordianatsystem. Standardbasises svarer til verdien 1 på de to aksene.*

Vi skal se på hvordan vi kan tolke lineæravbildninger geometrisk. Relativt til standardbasisen vil en lineær operator $T : \mathbb{R}^2 \to \mathbb{R}^2$ være gitt ved en 2×2-matrise der de to søylene utgjør bildene av de to basisvektorene $\mathbf{e}_1$ og $\mathbf{e}_2$. Avbildningen er bestemt av disse to vektorene, så vi kan illustrerer avbildningen T ved å se på firkanten som er bildet av enhetskvadratet med hjørner $(0,0)$, $(1,0)$, $(0,1)$ og $(1,1)$. (se figur)



Figur 4.2. *Rotasjon og skalering*

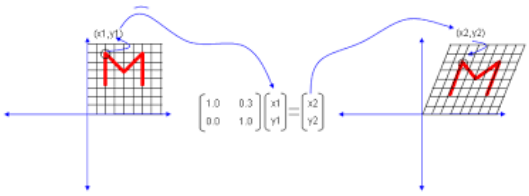
Alle invertible lineære operatorer i planet kan skrives som en sammensetning av **elementære** operasjoner, gitt ved **elementære matriser**. Dette er de samme elementære matrisene vi så på i kapittel 1.

Definisjon 4.1.1. *Det finnes tre typer elementære operasjoner i planet;*

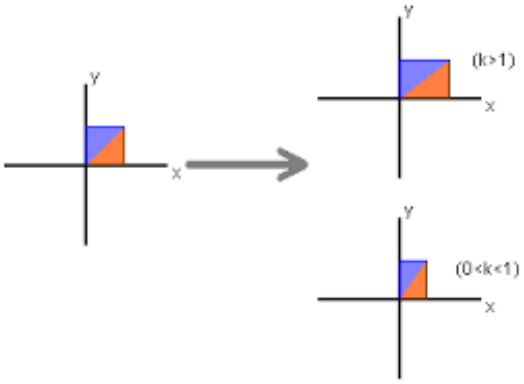
i) ***Skalering****, med matrise* $X_q = \begin{pmatrix} q & 0 \\ 0 & 1 \end{pmatrix}$ *(eller* $Y_q = \begin{pmatrix} 1 & 0 \\ 0 & q \end{pmatrix}$ *eller* $S_q = \begin{pmatrix} q & 0 \\ 0 & q \end{pmatrix}$*) for et reelt tall* $q \neq 0$ *og som forstørrer eller forminsker kvadratet i en eller to retninger.*

ii) ***Permutering****, med matrise* $\sigma = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ *som speiler kvadratet om linja* $y = x$*.*

i) ***Forskyvning****, med matrise* $F_k = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$ *for et reelt tall* $k \neq 0$ *som skyver øvre del av kvadratet, men beholder grunnlinjen.*



Figur 4.3. *Forskyvning*



Figur 4.4. *Horisontal skalering*

Rotasjoner er ikke blant de elementære operasjonene i planet. Men som et svært nyttig begrep kan det være greit å se hvordan en vilkårlig rotasjon kan settes sammen av elementære operasjoner. En rotasjon i planet med sentrum i origo og vinkel θ er gitt ved 2×2-matrisen

$$\rho_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

Denne matrisen avbilder

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}$$

og vi kan skrive den som et produkt av elementære matriser (for $\cos\theta, \sin\theta \neq 0$).

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -\frac{\cos\theta}{\sin\theta} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{\cos\theta} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -\sin\theta \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \frac{\sin\theta}{\cos\theta} & 1 \end{pmatrix}$$

Sjekk dette, og bestem også hvordan dette kan gjøres dersom $\cos\theta = 0$ eller $\sin\theta = 0$.

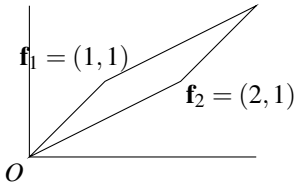
Proposisjon 4.1.2. *En hver invertibel 2×2-matrise A kan skrives som et produkt av elementære matriser.*

Bevis. Vi skal gi et geometrisk bevis for denne påstanden. Vi har allerede sett at enhetskvadratet avbildes på en firkant når vi multipliserer med A. Origo avbildes på origo, og bildet av de to enhetsvektorene $\mathbf{e}_1$ og $\mathbf{e}_2$ kaller vi $\mathbf{f}_1$ og $\mathbf{f}_2$.

Først må vi avklare om $(\mathbf{f}_1, \mathbf{f}_2)$ er orientert på samme måte som $(\mathbf{e}_1, \mathbf{e}_2)$. Hvis det ikke er tilfelle må vi multiplisere med permutasjonsmatrisen for å endre rekkefølgen på de to sidekantene. Deretter roterer vi firkanten tilbake slik at $\mathbf{f}_1$ ligger på x-aksen. Samtidig skalerer vi slik at $\mathbf{f}_1$ får lengde 1.

Nå har vi kommet fram til en firkant hvor grunnlinjen er enhetsvektoren langs x-aksen. Nå gjør vi en forskyvning slik at $\mathbf{f}_2$ plasseres på y-aksen, og til slutt gjør vi en skalering i vertikal retning for å skikre at høyden på firkanten er 1. □

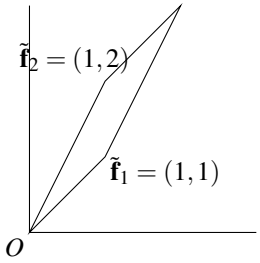
Vi skal gjennomføre denne prosessen på et eksempel. Figuren illustrerer bildet av enhetskvadratet etter multiplikasjon med $A = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}$



Det første vi ser er at $\mathbf{f}_1$ og $\mathbf{f}_2$ står i motsatt rekkefølge av $\mathbf{e}_1$ og $\mathbf{e}_2$. Det betyr at vi må bytte dem rundt, som igjen betyr at vi må multiplisere med permutasjonsmatrisen $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Det gir oss matrisen

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

og en figur hvor vi har speilet firkanten om $y = x$;



Vi lar hele tiden $\tilde{\mathbf{f}}_i$ betegne bildet av den opprinnelige $\mathbf{f}_i$.

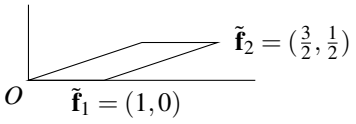
Neste step er å rotere firkanten slik at den høyre sidekanten (sett fra origo) flyttes til x-aksen. Vinkelen mellom x-aksen og sidekanten er $\frac{\pi}{4}$, så vi må rotere med den motsatte vinkelen $\theta = -\frac{\pi}{4}$. Det svarer til matrisen $\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$ og vi får

$$\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 2 & 3 \\ 0 & 1 \end{pmatrix}$$

Vi vil gjerne at høyre sidekant skal ha lengde 1, så vi skalerer med en faktor $\frac{1}{\sqrt{2}}$. Det gir oss

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

Figuren ser nå ut som



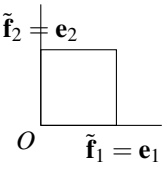
Vi følger på med å forskyve horisontalt slik at venstre sidekant faller sammen med y-aksen. Det svarer til å multiplisere med matrisen $\begin{pmatrix} 1 & -3 \\ 0 & 1 \end{pmatrix}$, og vi får

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

Til slutt skalerer vi med en faktor 2 i vertikal retning;

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

og vi ender opp med enhetskvadratet



Den sammensatte operasjonen for å bringe firkanten tilbake til enhetskvadratet er dermed gitt ved et produkt av elementære matriser;

$$\begin{aligned} B &= S_2 \cdot F_{-3} \cdot S_{\frac{1}{\sqrt{2}}} \cdot \rho_{-\frac{\pi}{4}} \cdot \sigma \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & -3 \\ 0 & 1 \end{pmatrix} \frac{1}{\sqrt{2}} I \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix} \end{aligned}$$

Siden A bringer enhetskuben over i den skjeive firkanten og B bringer firkanten tilbake til enhetskuben burde det ikke være noen stor overraskelse at $B = A^{-1}$. Utregningen

$$A \cdot B = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

bekrefter denne påstanden.

Vi har ved et geometrisk resonnement vist hvordan vi kan skrive en hver invertibel matrise som et produkt av elementære matriser. Hvis vi i tillegg skal ta med ikke-invertible matriser må vi føye en ny matrise til listen over elementære matriser. For å kunne redusere rangen til en matrise trenger vi en **projeksjon**, f.eks. gitt ved

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

Denne matrisen projiserer vertikalt ned på x-aksen.

$$P\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a \\ 0 \end{pmatrix}$$

Mer generelle projeksjoner vil kunne skrives som en kombinasjon av denne projeksjonen og passende rotasjoner og translasjoner.

Eksempel 4.1.1. *Lineæravbildningen gitt ved*

$$\mathbf{v} \to \frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\mathbf{v}$$

projiserer en vilkårlig vektor normalt inn i underrommet $\{(t,-t)\,|\,t \in \mathbb{R}\}$. *Denne linja kan vi dreie til x-aksen ved en rotasjon med vinkel* $\frac{\pi}{4}$. *Deretter projiserer vi vertikalt på x-aksen, for så å dreie tilbake. Det gir oss den sammensatte matrisen*

$$\begin{aligned} &\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \\ &\qquad = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \end{aligned}$$

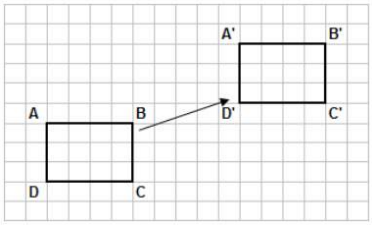
4.2 Stive bevegelser i planet

I dette kapitlet skal vi se på en bestemt type avbildninger av planet inn i seg selv, kalt **stive bevegelser**. Noen slike avbildninger er lineære og derfor gitt ved en 2×2-matrise, men som vi snart skal se, har vi også bruk for en annen type avbildninger, det som kalles **translasjoner**.

Definisjon 4.2.1. *En **translasjon** i planet er en avbildning* $T_{\mathbf{a}} : \mathbb{R}^2 \to \mathbb{R}^2$ *gitt ved*

$$T_{\mathbf{a}}(\mathbf{x}) = \mathbf{x} + \mathbf{a}$$

der $\mathbf{a} \in \mathbb{R}^2$.



Ikke-trivielle translasjoner (dvs. der $\mathbf{a} \neq \mathbf{0}$) er ikke lineære siden vi har

$$T_{\mathbf{a}}(\mathbf{x}+\mathbf{y}) = \mathbf{x}+\mathbf{y}+\mathbf{a}$$

mens

$$T_{\mathbf{a}}(\mathbf{x}) + T_{\mathbf{a}}(\mathbf{y}) = \mathbf{x}+\mathbf{a}+\mathbf{y}+\mathbf{a} = \mathbf{x}+\mathbf{y}+2\mathbf{a}$$

og dermed

$$T_{\mathbf{a}}(\mathbf{x}) + T_{\mathbf{a}}(\mathbf{y}) \neq T_{\mathbf{a}}(\mathbf{x}+\mathbf{y})$$

Avstander i planet eller lengder av vektorer finner vi ved å bruke Pythagoras setning.

Definisjon 4.2.2. *La (a_1,b_1) og (a_2,b_2) være to punkter i planet. Avstanden mellom punktene er gitt ved*

$$\|(a_1,b_1)-(a_2,b_2)\| = \sqrt{(a_1-a_2)^2+(b_1-b_2)^2}$$

Lengden av en vektor svarer til avstanden fra vektorens koordinater til origo, dvs. vi setter $(a_2,b_2)=(0,0)$, og lengden av $\mathbf{v}=\begin{pmatrix}a_1\\b_1\end{pmatrix}$ blir da

$$\|\mathbf{v}\| = \|(a_1,b_1)\| = \|(a_1,b_1)-(0,0)\| = \sqrt{a_1^2+b_1^2}$$

Dette kan vi også uttrykke ved matrisemultiplikasjon.

$$\|\mathbf{v}\|^2 = \mathbf{v}^T\cdot\mathbf{v}$$

Eksempel 4.2.1.

$$\|(3,4)\| = \sqrt{(3\ 4)\begin{pmatrix}3\\4\end{pmatrix}} = \sqrt{3^2+4^2} = \sqrt{25} = 5$$

Får å studere figurer i planet eller legemer i rommet, ser vi på avbildninger som ikke forandrer formen på figuren eller legemet. Slike avbildninger tilhører en klasse av avbildninger som kalles **stive bevegelser**.

Definisjon 4.2.3. *En avbildning $m:\mathbb{R}^2\to\mathbb{R}^2$ kalles en **stiv bevegelse** eller en **isometri** hvis den bevarer avstander, dvs. for vilkårlig valgte punkter $P,Q\in\mathbb{R}^2$ så er $\|P-Q\| = \|m(P)-m(Q)\|$.*

Proposisjon 4.2.4. *a) Identitetsavbildningen I er en stiv bevegelse.*

b) Dersom $m_1,m_2:\mathbb{R}^2\to\mathbb{R}^2$ er to stive bevegelser, da er også sammensetningen av de to avbildningene en stiv bevegelse.

c) Dersom m er en stiv bevegelse, så er også den inverse avbildningen m^{-1} en stiv bevegelse.

Bevis. Følger mer eller mindre direkte fra definisjonen av en stiv bevegelse. □

Eksempler på isometrier er rene forflytninger og rotasjoner.

Eksempel 4.2.2. *La $T_{\mathbf{e}_1}$ være translasjonen gitt ved*

$$T_{\mathbf{e}_1}(\mathbf{x}) = \mathbf{x}+(1,0)$$

Da vil $T_{\mathbf{e}_1}$ avbilde punktet $(a,0)$ på x-aksen på $(a+1,0)$ som også ligger på x-aksen.

Siden vi har

$$T_{\mathbf{a}}; P\mapsto P+\mathbf{a}$$

får vi

$$T_{\mathbf{b}}(T_{\mathbf{a}}(P)) = T_{\mathbf{a}}(P)+\mathbf{b} = (P+\mathbf{a})+\mathbf{b} = P+(\mathbf{a}+\mathbf{b})$$

og det følger at

$$T_{\mathbf{b}}\circ T_{\mathbf{a}} = T_{\mathbf{a}+\mathbf{b}}$$

Summen av to translasjoner er en ny translasjon.

I tillegg til translasjoner skal vi fokusere på to andre typer av isometrier i planet, og etter hvert skal vi se at alle isometrier kan skrives som sammensetninger av translasjoner og disse to.

Definisjon 4.2.5. *(i)* ***Rotasjon****: Rotasjon med en vinkel $\theta \neq 0$ om et fast punkt P.*

(ii) ***Refleksjon****: Refleksjon om en linje ℓ.*

Lemma 4.2.6. *La $m : \mathbb{R}^2 \to \mathbb{R}^2$ være en avbildning av planet på seg selv. Da er følgende utsagn om m ekvivalente:*

(a) Avbildningen m er en isometri som fikserer origo.

(b) Avbildningen m bevarer skalarprodukt, dvs. for alle $P, Q \in \mathbb{R}^2$ har vi

$$m(P) \cdot m(Q) = P \cdot Q$$

(c) Avbildningen m er venstre-multiplikasjon med en ortogonal matrise og derfor lineær.

Bevis. En isometri tilfredsstiller

$$(m(P) - m(Q)) \cdot (m(P) - m(Q)) = (P - Q) \cdot (P - Q)$$

for alle vektorer P og Q. Setter vi $Q = \mathbf{O}$ og antar at m fikserer origo, dvs. $m(Q) = \mathbf{O}$, så følger det at

$$m(P) \cdot m(P) = P \cdot P$$

for alle P. Kombinasjon av disse to likhetene gir at (a) $\implies$ (b).

Vi kan også vise at (b) $\implies$ (c). Anta at m bevarer skalarproduktet og la $\mathscr{S} = \{\mathbf{e}_1, \mathbf{e}_2\}$ være standardbasisen for $\mathbb{R}^2$. La $A = (m(\mathbf{e}_1), m(\mathbf{e}_2))$ være matrisen til m relativt til $\mathscr{S}$. Egenskapen i (b) gir at A er en ortogonal matrise, og det samme vil være tilfelle for A^{-1}. Sammensetningen $A^{-1}m$ vil bevare skalarproduktet, og i tillegg fiksere hvert basiselement. Det følger at $A^{-1}m$ er identiteten og derfor har vi $A^{-1}m(P) = P$, eller $m(P) = AP$.

Til slutt antar vi at (c) er sann og viser (a). Vi har

$$\begin{aligned} \|m(P) - m(Q)\|^2 &= (m(P) - m(Q))^T (m(P) - m(Q)) \\ &= (AP - AQ)^T (AP - AQ) \\ &= (A(P - Q))^T (A(P - Q)) \\ &= (P - Q)^T A^T A (P - Q) \\ &= (P - Q)^T (P - Q) = \|P - Q\|^2 \end{aligned}$$

siden ortogonale matriser oppfyller likheten $A^T A = Id$. I tillegg har vi at $m(\mathbf{O}) = A \cdot \mathbf{O} = \mathbf{O}$. □

Ortogonale 2×2-matriser $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ er definert gjennom relasjonen $A^T A = Id$, det vil si

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} a & c \\ b & d \end{pmatrix} = \begin{pmatrix} a^2 + b^2 & ac + bd \\ ac + bd & c^2 + d^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

eller $a^2 + b^2 = c^2 + d^2 = 1$, $ac + bd = 0$. Det gir at $a^2 + b^2 = 1$ og $(c, d) = \pm(-b, a)$. Det betyr at alle ortogonale 2×2-matriser er på formen

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix}, \quad \text{eller} \quad \begin{pmatrix} a & b \\ b & -a \end{pmatrix}$$

Eksempel 4.2.3. *La*

$$A = \begin{pmatrix} 0.6 & -0.8 \\ 0.8 & 0.6 \end{pmatrix}$$

Da har vi $0.6^2 + 0.8^2 = 0.36 + 0.64 = 1$ *og søylene står normalt på hverandre. A er derfor en ortogonal matrise.*

Proposisjon 4.2.7. *Enhver isometri kan skrives som en sum av en ortogonal lineær operator og en translasjon, dvs.*

$$m(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$$

Bevis. La $\mathbf{b} = m(\mathbf{0})$. Da har vi at $T_{-\mathbf{b}}(\mathbf{b}) = \mathbf{0}$ og $T_{-\mathbf{b}}m$ er en stiv bevegelse som holder origo i ro; $T_{-\mathbf{b}}m(\mathbf{0}) = \mathbf{0}$. Fra Lemmaet får vi at

$$T_{-\mathbf{b}}m(\mathbf{x}) = A\mathbf{x}$$

eller at $m(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$. □

Vi deler isometrier i planet inn i to hovedgrupper, **orienterings-bevarende** og **orienterings-reverserende**.

Definisjon 4.2.8. *En stiv bevegelse* $m(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$ *i planet er* ***orienterings-bevarende*** *dersom* $\det(A) = 1$ *og den er* ***orienterings-reverserende*** *dersom* $\det(A) = -1$.

I praksis betyr ordningsbevarende at avbildningen beholder rekkefølgen på basisvektorene når vi betrakter dem fra origo og følger en bestemt omløpsordning, f.ek.s mot klokka.

For en translasjon har vi at A er identitetsmatrisen, og den er derfor orienteringsbevarende.

En rotasjon i planet kan uttrykkes ved matrisen

$$\rho_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

for en rotasjonsvinkel $\theta \in [0, 2\pi)$ og $\det(\rho_\theta) = 1$. Vi kan f.eks. sette $\theta = \frac{\pi}{2}$. Det gir

$$\rho_{\frac{\pi}{2}} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

som tar x-aksen, gitt ved $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ på $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$, og y-aksen, gitt ved $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ på $\begin{pmatrix} -1 \\ 0 \end{pmatrix}$.

Hvis vi setter sammen to rotasjoner med rotasjonsvinkler θ_1 og θ_2 vil vi gjerne at svaret skal bli en rotasjon med vinkel $\theta_1 + \theta_2$. Det stemmer også analytisk ved å bruke formelene for cosinus og sinus til vinkelsummer;

$$\begin{aligned} \rho_{\theta_1}\rho_{\theta_2} &= \begin{pmatrix} \cos\theta_1 & -\sin\theta_1 \\ \sin\theta_1 & \cos\theta_1 \end{pmatrix} \begin{pmatrix} \cos\theta_2 & -\sin\theta_2 \\ \sin\theta_2 & \cos\theta_2 \end{pmatrix} \\ &= \begin{pmatrix} \cos(\theta_1+\theta_2) & -\sin(\theta_1+\theta_2) \\ \sin(\theta_1+\theta_2) & \cos(\theta_1+\theta_2) \end{pmatrix} \\ &= \rho_{\theta_1+\theta_2} \end{aligned}$$

hvor vi har brukt at

$$\cos(\theta_1+\theta_2)=\cos\theta_1\cos\theta_2-\sin\theta_1\sin\theta_2$$

og

$$\sin(\theta_1+\theta_2)=\sin\theta_1\cos\theta_2+\cos\theta_1\sin\theta_2$$

Eksempel 4.2.4. *Rotasjon med en vinkel $\frac{\pi}{3}$ er gitt ved matrisen*

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

Det gir f.eks. for A^3;

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

som svarer til rotasjon med en vinkel π.

Refleksjoner vil alltid oppfylle relasjonen $r^2=I$. I et passende valgt koordinatsystem kan vi skrive en refleksjon på formen

$$r=\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$$

Denne oppfyller $r^2=I$, og den har determinant lik -1.

Refleksjonen

$$r=\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$$

avbilder $\begin{pmatrix}a\\ b\end{pmatrix}$ på $\begin{pmatrix}a\\ -b\end{pmatrix}$, som betyr at den reflekterer om x-aksen. Siden den er gitt ved matrise-multiplikasjon er den også lineær. Vi har også refleksjoner om vilkårlige linjer gjennom origo.

La $y=ax$ være en rett linje gjennom origo. Vi plukker ut et punkt (x,y) på linja med $\|(x,y)\|=1$;

Lemma 4.2.9. *Punktet*

$$P=(\frac{1}{\sqrt{1+a^2}},\frac{a}{\sqrt{1+a^2}})$$

ligger på linja $y=ax$ og $\|P\|=1$.

Bevis. Vi har

$$y=\frac{a}{\sqrt{1+a^2}}=a\cdot\frac{1}{\sqrt{1+a^2}}=ax$$

og

$$\|(\frac{1}{\sqrt{1+a^2}},\frac{a}{\sqrt{1+a^2}})\|=\sqrt{(\frac{1}{\sqrt{1+a^2}})^2+(\frac{a}{\sqrt{1+a^2}})^2}=1$$

□

Siden punktet

$$P = (\frac{1}{\sqrt{1+a^2}}, \frac{a}{\sqrt{1+a^2}})$$

ligger på enhetssirkelen finnes det en entydig vinkel θ mellom 0 og 2π (0 inkludert, men ikke 2π) slik at

$$P = (\cos\theta, \sin\theta)$$

Dermed kan vi gi en refleksjon om linja $y = ax$ ved først å rotere med en vinkel $-\theta$ (som bringer linja til x-aksen), så speile om x-aksen, og så rotere tilbake med vinkelen θ. På matriseform får vi

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$
$$= \begin{pmatrix} \cos^2\theta - \sin^2\theta & 2\sin\theta\cos\theta \\ 2\sin\theta\cos\theta & \sin^2\theta - \cos^2\theta \end{pmatrix}$$

Bruker vi nå at

$$\cos\theta = \frac{1}{\sqrt{1+a^2}}, \quad \sin\theta = \frac{a}{\sqrt{1+a^2}}$$

og dermed

$$\cos^2\theta - \sin^2\theta = \frac{1-a^2}{1+a^2}, \quad 2\sin\theta\cos\theta = \frac{2a}{1+a^2}$$

får vi at refleksjonen om linja $y = ax$ er gitt ved

$$\begin{pmatrix} \frac{1-a^2}{1+a^2} & \frac{2a}{1+a^2} \\ \frac{2a}{1+a^2} & \frac{a^2-1}{1+a^2} \end{pmatrix}$$

Retningen på linja $y = ax$ er gitt ved en vektor $(1, a)$. Det betyr at vektoren $\mathbf{v} = \begin{pmatrix} a \\ -1 \end{pmatrix}$ står normalt på linja. Vi kan regne ut Householder-matrisen til denne vektoren.

$$\begin{aligned} \mathscr{Q}_{\mathbf{v}} &= I - 2\frac{\mathbf{v}\mathbf{v}^T}{\mathbf{v}^T\mathbf{v}} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \frac{2}{1+a^2}\begin{pmatrix} a^2 & -a \\ -a & 1 \end{pmatrix} \\ &= \begin{pmatrix} \frac{1-a^2}{1+a^2} & \frac{2a}{1+a^2} \\ \frac{2a}{1+a^2} & \frac{a^2-1}{1+a^2} \end{pmatrix} \end{aligned}$$

Vi ser at de to matrisene blir det samme, men vi kommer fram til svaret på to forskjellige måter.

Hvis vi setter sammen to refleksjoner, om forskjellige linjer, så vil determinanten til den sammensatte avbildningen være 1. Det følger av produktformelen for determinanter og at $(-1)(-1) = 1$. Sammensetningen er fortsatt en stiv bevegelse og den er en lineær avbildning, altså er den en rotasjon.

Proposisjon 4.2.10. *La ℓ_1 og ℓ_2 være to linjer gjennom origo, og la θ være vinkelen mellom dem. Da vil sammensetningen av refleksjoner over de to linjene svare til en rotasjon med vinkel 2θ.*

Bevis. For enkelthets skyld lar vi ℓ_2 være x-aksen, og ℓ_1 er gitt ved $y = ax$. Da vet vi fra det vi har gjort tidligere at speiling om linja ℓ_1 er gitt ved matrisen

$$\begin{pmatrix} \frac{1-a^2}{1+a^2} & \frac{2a}{1+a^2} \\ \frac{2a}{1+a^2} & \frac{a^2-1}{1+a^2} \end{pmatrix} = \frac{1}{1+a^2}\begin{pmatrix} 1-a^2 & 2a \\ 2a & a^2-1 \end{pmatrix}$$

Vi vet også at speiling om x-aksen er gitt ved matrisen

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

Setter vi disse to sammen, først speiling om x-aksen, dernes om $y = ax$, får vi på matriseform

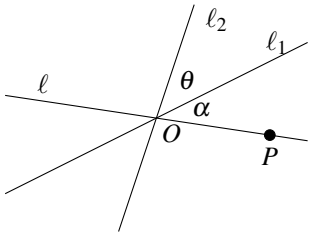
$$\begin{aligned} &\frac{1}{1+a^2}\begin{pmatrix} 1-a^2 & 2a \\ 2a & a^2-1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \\ &\qquad = \frac{1}{1+a^2}\begin{pmatrix} 1-a^2 & -2a \\ 2a & 1-a^2 \end{pmatrix} \end{aligned}$$

Fra argumentene over har vi at

$$\begin{aligned} &\frac{1}{1+a^2}\begin{pmatrix} 1-a^2 & -2a \\ 2a & 1-a^2 \end{pmatrix} \\ &\qquad = \begin{pmatrix} \cos^2\theta - \sin^2\theta & -2\sin\theta\cos\theta \\ 2\sin\theta\cos\theta & \cos^2\theta - \sin^2\theta \end{pmatrix} \\ &\qquad = \begin{pmatrix} \cos 2\theta & -\sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{pmatrix} = \rho_{2\theta} \end{aligned}$$

Denne matrisen gjenkjenner vi altså som en rotasjon med en vinkel 2θ. □

Vi kan føre et mer geometrisk bevis for denne proposisjonen.



Bevis. [Alternativt] La P være et vilkårlig punkt i planet og la ℓ være linja gjennom origo og P (se figuren over). Vi kaller vinkelen mellom ℓ og ℓ_1 for α. Det betyr at vinkelen mellom ℓ og ℓ_2 er $\theta + \alpha$. Refleksjon om ℓ_1 bringer linja ℓ til en linje $\tilde{\ell}$ som har vinkel $-\alpha$ med ℓ_1 og $\theta - \alpha$ med linja ℓ_2. Refleksjon om ℓ_2 bringer linja $\tilde{\ell}$ til en linje som har vinkel $-(\theta - \alpha) = \alpha - \theta$ med linja ℓ_2. Det betyr at sammensetningen tar en linje med vinkel $\theta + \alpha$ om ℓ_2 til en linje med vinkel $\alpha - \theta$ om ℓ_2. Sammensetningen roterer derfor linja ℓ med en vinkel $\theta + \alpha - (\alpha - \theta) = 2\theta$. □

Rotasjoner som matrise-multiplikasjon svarer til at vi roterer om origo. Vi er interessert i å rotere om et vilkårlig punkt i planet, eller reflektere om en vilkårlig linje. Måten vi fikser dette på er ved translasjon å flytte henholdsvis rotasjonssenter og speilings-linje til origo. Så gjennomfører vi rotasjonen og/eller refleksjonen, for så å flytte rotasjonssenteret tilbake til origo.

Eksempel 4.2.5. *Vi skal se på avbildningen i planet som roterer med vinkel $\theta = \frac{\pi}{2}$ om punktet $Q = (1,1)$ (mot klokka). Vi flytter først rotasjonssenteret til origo med translasjonsavbildningen $T_{-(1,1)}$. Deretter gjennomfører vi rotasjonen $\rho_{\frac{\pi}{2}}$, for så å flytte tilbake igjen med $T_{(1,1)}$. Dette gir*

$$\begin{aligned} T_{(1,1)} \circ \rho_{\frac{\pi}{2}} \circ T_{-(1,1)}(\mathbf{x}) &= \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} (\mathbf{x} - \begin{pmatrix} 1 \\ 1 \end{pmatrix}) + \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \mathbf{x} - \begin{pmatrix} -1 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \mathbf{x} + \begin{pmatrix} 2 \\ 0 \end{pmatrix} \end{aligned}$$

Vi kan formulere dette eksempelet mer generelt. Anta at vi skal rotere om et punkt Q, og rotasjonen er gitt ved multiplikasjon med en matrise A. Vi starter med et vilkårlig punkt $\mathbf{x}$. Vi flytter rotasjonssenteret til origo, og punktet $\mathbf{x}$ følger etter, beskrevet ved $\mathbf{v} - Q$. Nå er rotasjonssenteret origo, og rotasjon gir oss

$$A(\mathbf{v} - Q) = A\mathbf{v} - AQ$$

Til slutt flytter vi origo tilbake til Q, og vi ender opp med

$$\mathbf{v} \to A\mathbf{v} - AQ + Q$$

Hvis vi setter sammen to rotasjoner med forskjellig rotasjonssenter viser det seg at vi ender opp med en rotasjon med et tredje rotasjonssenter. Vi lar de to rotasjonene være gitt ved

$$\mathbf{v} \to A_1\mathbf{v} - A_1Q_1 + Q_1$$

og

$$\mathbf{v} \to A_2\mathbf{v} - A_2Q_2 + Q_2$$

Setter vi dem sammen, får vi

$$\begin{aligned} \mathbf{v} \to & A_2(A_1\mathbf{v} - A_1Q_1 + Q_1) - A_2Q_2 + Q_2 \\ & = A_2A_1\mathbf{v} - A_2A_1Q_1 + A_2Q_1 - A_2Q_2 + Q_2 \end{aligned}$$

Vi kan finne rotasjonssenteret til den sammensatte avbildningen ved å løse

$$A_2A_1\mathbf{v} - A_2A_1Q_1 + A_2Q_1 - A_2Q_2 + Q_2 = \mathbf{v}$$

Det gir rotasjonssenter

$$\mathbf{v} = (A_2A_1 - I)^{-1}(A_2A_1Q_1 - A_2Q_1 + A_2Q_2 - Q_2)$$

og rotasjonen er gitt ved matrisen A_2A_1. Hvis vi lar $Q_1 = Q_2 = Q$ reduserer dette uttrykket seg til $\mathbf{v} = Q$. Hvis vi lar $A = A_2 = A_1^{-1}$ får vi

$$\begin{aligned} \mathbf{v} \to A_2A_1\mathbf{v} & - A_2A_1Q_1 + A_2Q_1 - A_2Q_2 + Q_2 \\ & = \mathbf{v} - Q_1 + AQ_1 - AQ_2 + Q_2 \\ & = \mathbf{v} + (I - A)(Q_2 - Q_1) \end{aligned}$$

som betyr at i det tilfellet er den sammensatte bevegelsen en translasjon.

En spesielt interessante type isometrier er de som oppfyller $m^k = I$ for et helt tall $k \geq 2$. Det betyr at

$$m^k(\mathbf{x}) = A^k\mathbf{x} + (\sum_{j=0}^{k-1} A^j)\mathbf{b} = \mathbf{x}$$

eller

$$(A^k - I)\mathbf{x} + (\sum_{j=0}^{k-1} A^j)\mathbf{b} = \mathbf{0}$$

Dette skal gjelde for alle $\mathbf{x}$. Spesielt skal det gjelde for $\mathbf{x} = \mathbf{0}$, som betyr at

$$(\sum_{j=0}^{k-1} A^j)\mathbf{b} = \mathbf{0}$$

og derfor at også

$$(A^k - I)\mathbf{x} = \mathbf{0}$$

for alle $\mathbf{x}$. Men det betyr at

$$A^k = I$$

Vi sier at matrisen A har **endelig orden**, i dette tilfellet **orden** k.

Eksempel 4.2.6. *I det forrige eksemplet så vi på rotasjoner om* $(1,1)$ *med rotasjonsvinkel* $\frac{\pi}{2}$*. Vi kaller denne avbildningen S. Hvis vi skriver* $\mathbf{x} = (x_1, x_2)$*, får vi*

$$S(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 2 \\ 0 \end{pmatrix} = \begin{pmatrix} -x_2 + 2 \\ x_1 \end{pmatrix}$$

Det gir

$$S(S(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix})) = S(\begin{pmatrix} -x_2 + 2 \\ x_1 \end{pmatrix}) = \begin{pmatrix} -x_1 + 2 \\ -x_2 + 2 \end{pmatrix}$$

og

$$S(S(S(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}))) = \begin{pmatrix} x_2 - 2 + 2 \\ -x_1 + 2 \end{pmatrix} = \begin{pmatrix} x_2 \\ -x_1 + 2 \end{pmatrix}$$

og endelig

$$S(S(S(S(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix})))) = \begin{pmatrix} x_1 - 2 + 2 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

som betyr at $S^4 = I$*, og rotasjon med vinkelen* $\frac{\pi}{2}$ *har orden 4.*

Eksempel 4.2.7. *Vi skal se på en speiling R om linja* $x + y = 1$*. Vi plukker et punkt på linja, f.eks.* $(1,0)$ *(her kan vi faktisk velge et vilkårlig punkt på linja, svaret vil bli det samme uansett). Translasjon med* $-(1,0)$ *flytter linja til en linje gjennom origo. Deretter roterer vi linja med en vinkel* $\theta = \frac{\pi}{4}$ *og legger linja langs med x-aksen. Så reflekterer vi om x-aksen, og avslutter med å gjøre rotasjonen og translasjonen i motsatt retning av hva vi startet med. Det gir*

$$R = T_{(1,0)}\rho_{-\frac{\pi}{4}}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\rho_{\frac{\pi}{4}}T_{-(1,0)}$$

Nå har vi

$$\rho_{\pm\frac{\pi}{4}} = \begin{pmatrix} \frac{\sqrt{2}}{2} & \mp\frac{\sqrt{2}}{2} \\ \pm\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix} = \frac{\sqrt{2}}{2}\begin{pmatrix} 1 & \mp 1 \\ \pm 1 & 1 \end{pmatrix}$$

og derfor

$$\begin{aligned}
&\rho_{-\frac{\pi}{4}} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \rho_{\frac{\pi}{4}} \\
&= \frac{\sqrt{2}}{2} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \frac{\sqrt{2}}{2} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \\
&= \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}
\end{aligned}$$

Men det gir

$$\begin{aligned}
R(\mathbf{x}) &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} (\mathbf{x} - \begin{pmatrix} 1 \\ 0 \end{pmatrix}) \\
&= \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \mathbf{x} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\
&= \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \mathbf{x} + \begin{pmatrix} 1 \\ 1 \end{pmatrix}
\end{aligned}$$

Denne avbildningen oppfyller

$$\begin{aligned}
R^2(\mathbf{x}) &= \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} (\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \mathbf{x} + \begin{pmatrix} 1 \\ 1 \end{pmatrix}) + \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\
&= \mathbf{x} + \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \mathbf{x}
\end{aligned}$$

I tillegg ser vi at for punkter på speilingslinja, $(x, 1-x)$ *så får vi*

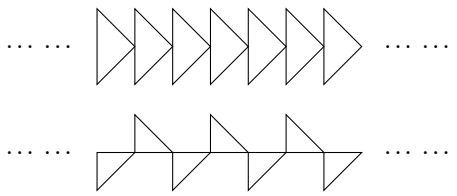
$$\begin{aligned}
R(\begin{pmatrix} x \\ 1-x \end{pmatrix}) &= \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} x \\ 1-x \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\
&= \begin{pmatrix} x-1 \\ -x \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} x \\ 1-x \end{pmatrix}
\end{aligned}$$

Refleksjoner har altså orden 2.

I tillegg til translasjoner, rotasjoner og refleksjoner setter vi navn på enda en type av stive bevegelser. En **glide-refleksjon** er en spesiell sammensetning av en refleksjon og en translasjon.

Definisjon 4.2.11. *La r være en refleksjon og T en translasjon langs refleksjonsaksen til r. Sammensetningen* $T \circ r$ *kalles en* ***glide-refleksjon****.*

Glide-refleksjoner illustreres ofte med en **bord**, dvs. en figur som gjentar seg langs en rett linje. Betrakt følgende to border:



Begge bordene kan reflekteres om midtaksen. Forskjellen er at i den nederste må vi flytte oss ett hakk til siden etter at vi har reflektert. Det er typisk for en glide-refleksjon.

Felles for rotasjoner og refleksjoner er at noen av punktene i figurene avbildet på seg selv; en rotasjon vil alltid avbilde rotasjonssenteret på seg selv, mens en refleksjon avbilder hele refleksjonsaksen på seg selv. Dette er en viktig egenskap og har derfor fått sitt eget navn.

Definisjon 4.2.12. *La $m : \mathbb{R}^2 \to \mathbb{R}^2$ være en stiv bevegelse i planet. Et punkt i planet som ved m avbildes på seg selv, $m(P) = P$, kalles et **fikspunkt** for m. Mengden av alle fikspunkter;*

$$\mathscr{F}_m = \{P \in \mathbb{R}^2 \,|\, m(P) = P\}$$

*kalles **fikspunktsmengden** til m.*

Fikspunktsmengden til en speiling S_ℓ om en linje ℓ er linja ℓ; $\mathscr{F}_{S_\ell} = \ell$. Fikspunktsmengden til en rotasjon derimot er kun ett punkt, rotasjonssenteret. Translasjoner har ingen fikspunkter, siden

$$T_{\mathbf{a}}(\mathbf{x}) = \mathbf{x} + \mathbf{a} = \mathbf{x}$$

betyr at $\mathbf{a} = \mathbf{0}$. Avbildningen $T_{\mathbf{0}}$ kaller vi normalt ikke en translasjon siden det faktisk er identitetsavbildningen, og denne har hele planet som sin fikspunktsmengde. For å fange opp glide-refleksjoner kan vi utvide definisjonen av fikspunkter til større mengder enn bare et punkt:

Definisjon 4.2.13. *La $m : \mathbb{R}^2 \to \mathbb{R}^2$ være en stiv bevegelse i planet. Vi sier at en delmengde $F \subset \mathbb{R}^2$ av planet som m avbilder på seg selv **fikseres** av m.*

Merk at en mengde som fikseres av m ikke trenger å være en fikspunktmengde. I en fikspunktmengde skal alle punktene avbildes på seg selv, noe som ikke trenger å være tilfelle for en delmengde som fikseres av m. I det siste tilfellet kan det være at punktene i mengden avbildes på andre punkter i mengden, det vesentlige er ikke om punktene fikseres, men at de ikke sendes ut av delmengden.

Siden fikspunktsmengdene er vesensforskjellige for de forskjellige typene av stive bevegelser, kan vi bruke denne egenskapen til å karakterisere dem. Det gir oss følgende resultat:

Proposisjon 4.2.14. *La $m : \mathbb{R}^2 \to \mathbb{R}^2$ være en stiv bevegelse i planet med fikspunktsmengde $\mathscr{F}_m$. Da har vi følgende karakterisering:*

i) Hvis $\mathscr{F}_m = \mathbb{R}^2$, så er m identitetsavbildningen.

ii) Hvis $\mathscr{F}_m$ er en linje, så er m en speiling.

iii) Hvis $\mathscr{F}_m$ er et punkt, så er m en rotasjon.

iv) Hvis $\mathscr{F}_m = \emptyset$ (den tomme mengden), så er m en translasjon .

v) Hvis $\mathscr{F}_m = \emptyset$, men m fikserer en linje, så er m en glide-refleksjon.

Bevis. Følger fra argumentet over og klassifikasjonen av stive bevegelser i planet. □

Eksempel 4.2.8. *La* $m : \mathbb{R}^2 \to \mathbb{R}^2$ *være gitt ved*

$$m(\mathbf{x}) = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \mathbf{x} + \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Vi finner fikspunktene til m ved å løse likningen

$$\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \mathbf{x} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \mathbf{x}$$

Vi omformer den til

$$(\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix})\mathbf{x} = \begin{pmatrix} -1 & -1 \\ -1 & -1 \end{pmatrix} \mathbf{x} = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$$

som har løsning $\{(x_1, x_2) \in \mathbb{R}^2 \,|\, x_1 + x_2 = 1\}$. *Dette er en hel linje, så den stive bevegelsen er en speiling.*

4.3 Symmetrier av plane figurer

Vi skal se litt nærmere på begrepet symmetri av en mengde F i planet, og vi begynner med å repetere definisjonen.

Definisjon 4.3.1. *La* $F \subset \mathbb{R}^2$ *være en mengde (endelig eller uendelig) av punkter i planet. En stiv bevegelse* $m : \mathbb{R}^2 \to \mathbb{R}^2$ *kalles en* ***symmetri*** *av* F *dersom* $m(F) = F$. *Mengden av symmetrier for* F *skriver vi som* $\mathrm{Sym}(F)$, *og vi bruker navnet* ***symmetrigruppe*** *om denne mengden.*

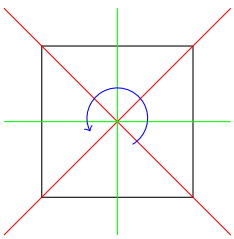
Eksempel 4.3.1. *La* ℓ *være en linje i planet. Da vil translasjoner langs linja være symmetrier for linja. Det samme vil være tilfelle for refleksjoner og glide-refleksjoner langs linja.*

Eksempel 4.3.2. *La* C *være enhetskvadratet med hjørner* $(0,0)$, $(1,0)$, $(0,1)$ *og* $(1,1)$. *Symmetriene* $\mathrm{Sym}(C)$ *til* C *er gitt ved*

(i) Identitetsavbildningen (alltid en symmetri),

(ii) Rotasjoner om midtpunktet $(\frac{1}{2}, \frac{1}{2})$ *med en vinkel som er et helt multippel av* $\frac{\pi}{2}$,

(iii) Speilinger om midtlinjene $x = \frac{1}{2}$ *og* $y = \frac{1}{2}$ *(grønne linjer), dessuten diagonalene* $y = x$ *og* $y = 1 - x$ *(røde linjer).*



Eksempel 4.3.3. *Symmetriene til en sirkel er alle rotasjoner om sentrum av sirkelen, i tillegg til alle speilinger om en vilkårlig valgt diameter.*

Symmetri-grupper har noen viktige, felles egenskaper.

Proposisjon 4.3.2. *La $F \subset \mathbb{R}^2$ være en mengde av punkter i planet med tilhørende symmetri-gruppe* $\mathrm{Sym}(F)$*. Da har vi*

a) Identitetsavbildningen er med i $\mathrm{Sym}(F)$*.*

b) Dersom $m_1, m_2 : \mathbb{R}^2 \to \mathbb{R}$ er to symmetrier av F, så er også sammensetningen av de to avbildningene en symmetri.

c) Dersom en avbildning m er med i $\mathrm{Sym}(F)$*, så er også den inverse avbildningen en symmetri.*

Bevis. Resultatet følger direkte fra et tilsvarende resultat for stive bevegelse. □

Når vi studerer ulike fysiske objekter, f.eks. molekyler, vil mange egenskaper for objektene være knyttet til symmetrier for objektet. Et eksempel på fysiske objekter der symmetri-strukturen har avgjørende betydning for egenskaper ved objektene er krystaller, som jo finnes i mange varianter i naturen. I det neste kapittelet skal vi studere krystaller gjennom å klassifisere deres symmetriegenskaper.

4.4 Plane krystaller

I krystaller danner molekylene et regelmessig mønster som gjentar seg i alle retninger. Hele strukturen er bestemt av et lite utsnitt, som vi kaller et **fundamentalområde** eller en **enhetscelle** for krystallen. I dette kapitlet skal vi se på krystaller i planet. Det gir ikke helt mening, siden krystaller jo er romlige. Men det er enklere å studere plane krystaller enn romlige, og prinsipielt er det ikke så stor forskjell, selv om mulighetene er mye større i rommet enn i planet. I rommet finnes det 230 forskjellige strukturer, mens det i planet bare er 17 forskjellige i. Plane krystaller kalles også for **tapetmønstre** siden tapeter på samme måte som de plane krystallstrukturene er bestemt av et lite utsnitt som gjentar seg i to retninger.

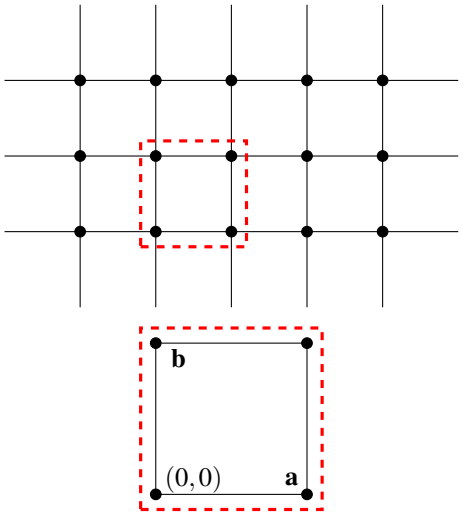
Definisjon 4.4.1. *En **plan krystallstruktur** $\mathscr{M}$ er en mengde av punkter i planet slik at det finnes en endelig mengde M og to lineært uavhengige vektorer* $\mathbf{a}$ *og* $\mathbf{b}$ *som genererer $\mathscr{M}$;*

$$\mathscr{M} = \{T_{n\mathbf{a}+m\mathbf{b}}(M) \,|\, n, m \in \mathbb{Z}\}$$

*En **enhetscelle** er en minste del av krystallstrukturen som viser hvordan strukturen fyller hele planet, og som med unntak av translasjonene bestemmer alle symmetriene til strukturen.*

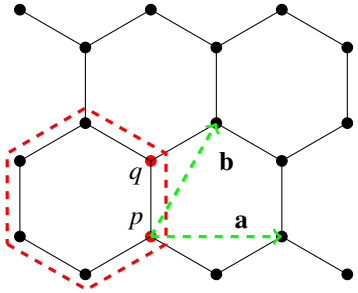
En krystallstruktur kan sees på som at enhetcellen gjentas igjen og igjen i to retninger. Vi kan alltid la M i definisjonen over være en enhetscelle, men det vil normalt finnes mindre mengder som gjør samme jobben, slik vi skal se i de neste eksemplene.

Eksempel 4.4.1.



Et rektangulært rutenett kan betraktes som en plan krystallstruktur gitt av ett punkt, $M = \{(0,0)\}$*. Den markerte enhetscellen består av de fire hjørnene* $(0,0)$, $\mathbf{a}$, $\mathbf{b}$ *og* $\mathbf{a}+\mathbf{b}$*. Krystallstrukturen framkommer ved å kopiere denne enhetscellen i to ortogonale retninger*

Eksempel 4.4.2.



En ***hexagonalt*** *struktur med en enhetscelle markert. Krystallstrukturen er definert av mengden* $M = \{p,q\}$ *bestående av de to røde punktene og med* $\mathbf{a}$ *og* $\mathbf{b}$ *som markert med grønt på figuren. Enhetscellen består av de 6 punktene* p, q, $p-\mathbf{a}$, $q-\mathbf{b}$, $q-\mathbf{a}$ *og* $p+\mathbf{b}-\mathbf{a}$*.*

Krystallstrukturer danner **diskrete** konfigurasjoner. Det betyr at det finnes en minste avstand mellom to punkter i strukturen.

Proposisjon 4.4.2. *La* $\mathscr{M}$ *være en plan krystallstruktur. Da finnes det et positivt tall* $d = d(\mathscr{M})$ *slik at alle avstander mellom punktene i* M *er større enn eller lik med* d*.*

Bevis. Vi lar d_0 være den minste avstanden mellom et punkt i enhetscellen og punktene i krystallstrukturen. Siden enhetscellen betsår av endelig mange punkter og de nærmeste punktene nødvendigvis må ligge innenfor en endelig mengde av andre punkter i krystallstrukturen finnes det en slik minste avstand. Siden krystallstrukturen bare gjentar seg utover planet vil det ikke være noen avstander andre steder som er mindre. □

Pr. definisjon vil en krystallstruktur ha to translasjonssymmetrier, $T_{\mathbf{a}}$ og $T_{\mathbf{b}}$. Strukturen kan også ha rotasjonssymmetrier, slik som i eksemplet over. Anta at vi har en slik rotasjonssymmetri ρ med **rotasjonssenter** O . Vi lar O være origo i et koordinatsystem. Rotasjon om O vil da være en lineær avbildning. La $\mathbf{c}$ være det minste translatet i $\mathscr{M}$. Med det mener vi at $\|\mathbf{c}\|$ er minimum blant alle $\mathbf{c}$ slik at $\mathscr{M}+\mathbf{c}=\mathscr{M}$. La $P\in\mathscr{M}$. Da har vi $P+\mathbf{c}\in\mathscr{M}$, og også $\rho^{-1}(P+\mathbf{c})\in\mathscr{M}$. Videre har vi at $\rho^{-1}(P+\mathbf{c})-\mathbf{c}\in\mathscr{M}$, og endelig

$$\rho(\rho^{-1}(P+\mathbf{c})-\mathbf{c})=P+\mathbf{c}-\rho(\mathbf{c})\in\mathscr{M}$$

siden ρ er lineær. Vi kan bruke cosinus-setningen på en trekant med hjørner $(0,0)$, $\mathbf{c}$ og $\rho(\mathbf{c})$. Hvis vi setter $\|\mathbf{c}\|=\|\mathbf{c}-(0,0)\|=\|\rho(\mathbf{c})-(0,0)\|=1$, så vil

$$\|\rho(\mathbf{c})-\mathbf{c}\|^2=1^2+1^2-2\cos\theta=2(1-\cos\theta)$$

hvor θ er rotasjonsvinkelen. Siden $\rho(\mathbf{c})-\mathbf{c}$ er et translat må vi ha

$$\|\rho(\mathbf{c})-\mathbf{c}\|\geq 1$$

og det følger at $1-\cos\theta\geq\frac{1}{2}$ eller $\cos\theta\leq\frac{1}{2}$. Det betyr at rotasjonsvinkelen $\theta\geq\frac{\pi}{3}$.

En annen konsekvens av at vi har en minste avstand i krystallstrukturen er at rotasjonsymmetrien må være endelig, det vil si at $\rho^k=I$ for et naturlig tall k. Hvis det ikke var tilfelle, ville vi hatt uendelig mange krystallpunkter rundt rotasjonssenteret, i samme avstand fra senteret, noe som ikke er forenlig med en minste avstand mellom punktene.

Proposisjon 4.4.3. *En rotasjonssymmetri ρ for en plan krystallstruktur må ha orden 1, 2, 3, 4 eller 6.*

Bevis. Det følger av argumentet over at rotasjonsvinkelen θ må være større enn eller lik $\frac{\pi}{3}$. Hvis $r\cdot\theta=2\pi$, følger det at $r\leq 6$. Tilfellet $r=5$ trenger litt særbehandling. Igjen lar vi rotasjonssenteret være origo i koordinatsystemet og vi lar $\mathbf{c}$ være det minste translatet i $\mathscr{M}$. La $P\in\mathscr{M}$ et vilkårlig punkt i krystallstrukturen Da har vi

$$\rho^2(\rho^{-2}(P+\mathbf{c})+\mathbf{c})=P+\mathbf{c}+\rho^2(\mathbf{c})\in\mathscr{M}$$

for alle $P\in\mathscr{M}$. Hvis vi lar ρ være en rotasjon av orden 5 får vi

$$\begin{aligned}I+\rho^2&=\begin{pmatrix}1&0\\0&1\end{pmatrix}+\begin{pmatrix}\cos\frac{4\pi}{5}&-\sin\frac{4\pi}{5}\\\sin\frac{4\pi}{5}&\cos\frac{4\pi}{5}\end{pmatrix}\\&=\begin{pmatrix}\cos\frac{4\pi}{5}+1&-\sin\frac{4\pi}{5}\\\sin\frac{4\pi}{5}&\cos\frac{4\pi}{5}+1\end{pmatrix}\end{aligned}$$

Hvis vi bruker denne operatoren på en vektor av lengde 1, f.eks. $(1,0)$, får vi

$$(I+\rho^2)(\begin{pmatrix}1\\0\end{pmatrix})=\begin{pmatrix}\cos\frac{4\pi}{5}+1\\\sin\frac{4\pi}{5}\end{pmatrix}$$

med (kvadratet av) absoluttverdien;

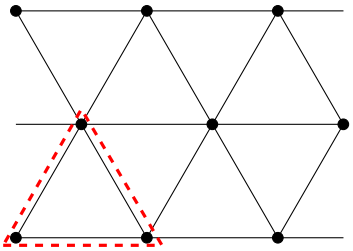
$$\begin{aligned}(\cos\frac{4\pi}{5}+1)^2+(\sin\frac{4\pi}{5})^2&=2+2\cos\frac{4\pi}{5}\\&=2+2\cdot\frac{1}{4}(-1-\sqrt{5})\\&=\frac{3}{2}-\frac{\sqrt{5}}{2}<1\end{aligned}$$

Det betyr at

$$\|\mathbf{c}+\rho^2(\mathbf{c})\| < \|\mathbf{c}\|$$

som gir oss en motsigelse siden $\mathbf{c}$ var antatt å gi det minste translatet. □

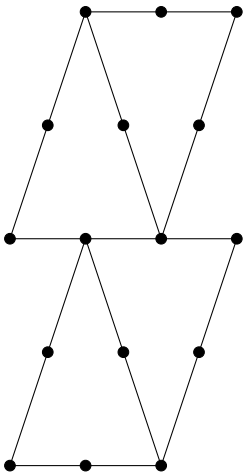
Eksempel 4.4.3.



En plan krystallstruktur med rotasjonssymmetri av orden 6.

Krystallstrukturen kan også ha refleksjonssymmetrier. Det innebærer at vi kan finne en linje ℓ slik at speiling om linja er en symmetri av $\mathscr{M}$.

Eksempel 4.4.4.



En plan krystallstruktur med flere refleksjonssymmetrier

I tillegg til de rene translasjonen, rotasjonene og refleksjonene kan vi sette sammen symmetrier. Det gir oss flere muligheter:

i) Sammensetting av to translasjoner $T_{\mathbf{a}}$ og $T_{\mathbf{b}}$ gir oss en ny translasjon $T_{\mathbf{a}+\mathbf{b}}$.

ii) Sammensetting av to rotasjoner gir oss en ny rotasjon om et tredje rotasjonssenter.

iii) Sammensetting av to refleksjoner gir oss en ny rotasjon med skjæringspunktet mellom de to refleksjonslinjene som rotasjonssenter.

iv) Sammensetting av en translasjon $T_{\mathbf{a}}$ og en rotasjon om punktet P gir oss en tilsvarende rotasjon om punktet $P-(A-I)^{-1}A\mathbf{a}$, hvor rotasjonen er gitt ved matrisen A. Gjør vi rotasjonen først, vil sammensetningen være en tilsvarende rotasjon om punktet $P+(A-I)^{-1}\mathbf{a}$.

v) Translasjon og refleksjon gir oss en glide-refleksjon, tilsvarende for motsatt rekkefølge.

vi) En rotasjon satt sammen med en refleksjon gir oss en ny refleksjon, hvor refleksjonslinja er både flyttet og dreid.

En liten kommentar til punkt v). En glide-refleksjon er definert ved en refleksjon etterfulgt av en translasjon langs refleksjonsaksen. Vi trenger ikke presiseringen om at translasjonen skal være langs refleksjonsaksen. Enhver sammensetning av en refleksjon og en translasjon vil gi oss en glide-refleksjon, slik det neste eksempelet illustrerer:

Eksempel 4.4.5. *La r være refleksjon om x-aksen, gitt ved*

$$r = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

og $T_{\mathbf{a}}$ en vilkårlig translasjon med translat $\mathbf{a} = (a,b)$. Da vil sammensettingen $T_{\mathbf{a}} \circ r$ gi oss avbildningen

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} x+a \\ -y+b \end{pmatrix}$$

La ℓ være linja gitt ved $y = \frac{b}{2}$. Refleksjon om linja ℓ finner vi ved først å flytte linja ℓ til x-aksen, så reflektere om x-aksen og tilslutt x-aksen tilbake til ℓ. Denne avbildningen er gitt ved

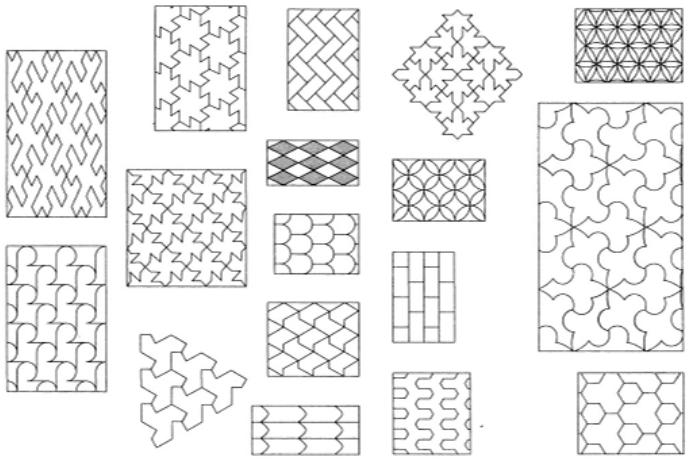
$$T_{(0,\frac{b}{2})} \circ r \circ T_{(0,-\frac{b}{2})} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ -y+b \end{pmatrix}$$

En translasjon langs ℓ er gitt ved $T_{(a,0)}$ og følger vi nå på med denne ender vi opp med

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x+a \\ -y+b \end{pmatrix}$$

som var det vi ville. Tilsvarende kan vi gjøre for den motsatte sammensettingen $r \circ T_{\mathbf{a}}$. I det tilfellet vil den nye refleksjonsaksen være gitt ved $y = -\frac{b}{2}$.

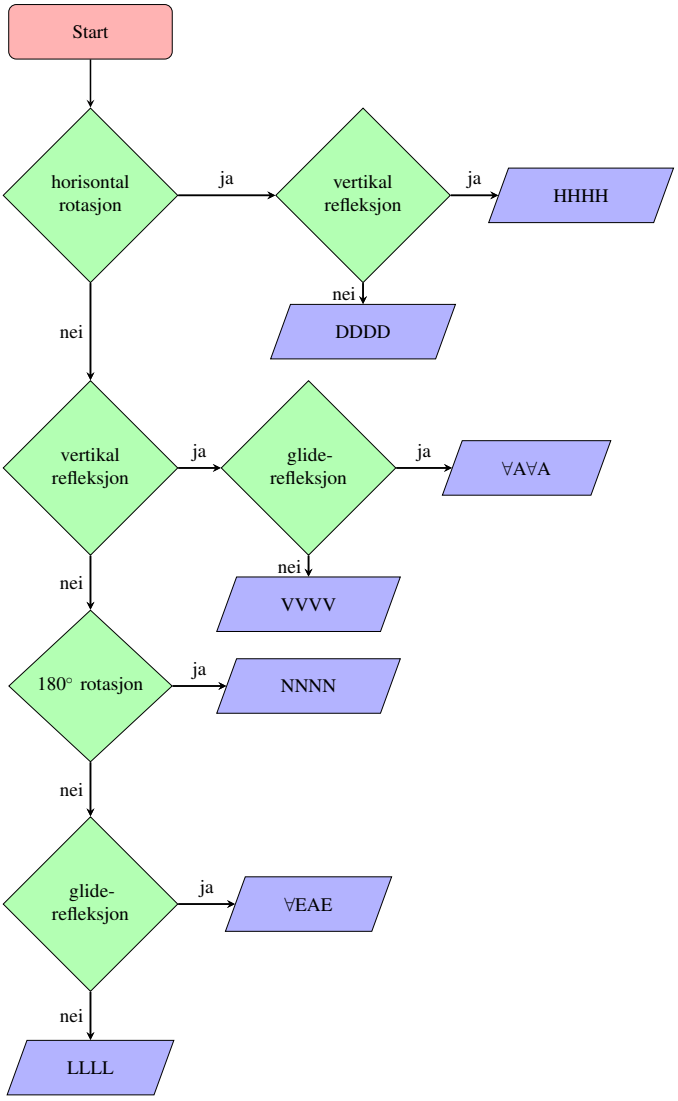
Til sammen finnes det 17 ulike symmetri-konfigurasjoner i planet. Vi skal ikke gi noen full klassifikasjon her, men vi gir noen eksempler. I den generelle klassifikasjonen vil man naturlig stille opp noen spørsmål i en standard rekkefølge. Det første er å avklare ordenen til rotasjons-symmetriene, 1, 2, 3, 4 eller 6. Dernest sjekker vi om det finnes refleksjons-symmetrier. Når det er avklart gjelder det å få oversikt over sammenhengen mellom rotasjoner refleksjoner og translasjoner. Her dukker det opp fenomener som: rotasjonssenter på eller utenfor refleksjons-linjer, glide-refleksjoner eller ordinære refleksjoner og glide-akser sammenfallende med refleksjons-linjer. Figuren gir en skjematisk framstilling av de 17 mønstrene



Den tilsvarende analysen i det 1-dimensjonale tilfellet er betydelig enklere. Det finnes 7 ulike border (symmetri-strukturer). Disse er listet opp i tabellen under:

Bord	Translasjon	180° Rotasjon	Horisontal refleksjon	Vertikal refleksjon	Ikke-triviell glide-refleksjon
LLLLLLL	ja	nei	nei	nei	nei
NNNNNN	ja	ja	nei	nei	nei
DDDDDD	ja	nei	ja	nei	nei
VVVVVV	ja	nei	nei	ja	nei
HHHHHH	ja	ja	ja	ja	nei
∀A∀A∀A∀A	ja	ja	nei	ja	ja
∀EAE∀EAE	ja	nei	nei	nei	ja

En alternativ framstilling av den samme klassifikasjonen er som et flytdiagram:



4.5 Oppgaver med løsning

Eksempel 4.5.1. *En lineær avbildning $T : \mathbb{R}^2 \to \mathbb{R}^2$ er gitt ved*

$$T(x,y) = (\sqrt{3}x + y, x + \sqrt{3}y)$$

a) Skriv opp avbildningen på matriseform,

$$T(\begin{pmatrix} x \\ y \end{pmatrix}) = A \begin{pmatrix} x \\ y \end{pmatrix}$$

b) Forklar hvorfor T avbilder enhetskvadratet på firkanten med hjørner $(0,0)$, $(\sqrt{3},1)$, $(1,\sqrt{3})$ og $(1+\sqrt{3},1+\sqrt{3})$.

c) Finn et produkt av elementære matriser som avbilder firkanten i b) på enhetskvadratet.

d) Regn ut A^{-1} ved å gange sammen produktet i c).

Løsning.

1. **Metodevalg:** Vi skal dekomponere den inverse til A ved en sekvens av elementære matriser, først en permutasjon hvis nødvendig, dernest en rotasjon og eventuelt en horisontal skalering. Så en forskyvning og eventuelt en vertikal skalering.

2. **Regning:** Matrisen A er gitt ved

$$(T(\begin{pmatrix} 1 \\ 0 \end{pmatrix})\ T(\begin{pmatrix} 0 \\ 1 \end{pmatrix})) = \begin{pmatrix} \sqrt{3} & 1 \\ 1 & \sqrt{3} \end{pmatrix}$$

Hjørnene i firkanten er rett orientert, så vi trenger ikke noen permutasjon. Vinkelen mellom x-aksen og $(\sqrt{3},1)$ er $\frac{\pi}{6}$ siden lengden av $(\sqrt{3},1)$ er $\sqrt{\sqrt{3}^2+1^2} = \sqrt{3+1} = 2$ og $\cos\frac{\pi}{6} = \frac{\sqrt{3}}{2}$. Rotasjon med en vinkel $-\frac{\pi}{6}$ er gitt ved

$$\begin{pmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix}$$

Denne avbilder $(\sqrt{3},1)$ på

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

og $(1,\sqrt{3})$ på

$$\begin{pmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix} \begin{pmatrix} 1 \\ \sqrt{3} \end{pmatrix} = \begin{pmatrix} \sqrt{3} \\ 1 \end{pmatrix}$$

Horisontal skalering med

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

sender $\begin{pmatrix}2\\0\end{pmatrix}$ på $\begin{pmatrix}1\\0\end{pmatrix}$ og $\begin{pmatrix}\sqrt{3}\\1\end{pmatrix}$ på $\begin{pmatrix}\frac{\sqrt{3}}{2}\\1\end{pmatrix}$.

Neste skritt er en forskyvning som tar $\begin{pmatrix}\frac{\sqrt{3}}{2}\\1\end{pmatrix}$ inn på y-aksen. Den er gitt ved

$$\begin{pmatrix}1 & -\frac{\sqrt{3}}{2}\\0 & 1\end{pmatrix}$$

og vi får

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

Det gir oss sammensetting (som gir A^{-1});

$$\begin{aligned}A^{-1}&=\begin{pmatrix}1 & -\frac{\sqrt{3}}{2}\\0 & 1\end{pmatrix}\cdot\begin{pmatrix}\frac{1}{2} & 0\\0 & 1\end{pmatrix}\cdot\begin{pmatrix}\frac{\sqrt{3}}{2} & \frac{1}{2}\\-\frac{1}{2} & \frac{\sqrt{3}}{2}\end{pmatrix}\\&=\begin{pmatrix}\frac{\sqrt{3}}{2} & -\frac{1}{2}\\-\frac{1}{2} & \frac{\sqrt{3}}{2}\end{pmatrix}\end{aligned}$$

Vi kan sjekke at svaret er riktig ved å regne ut

$$\begin{pmatrix}\frac{\sqrt{3}}{2} & -\frac{1}{2}\\-\frac{1}{2} & \frac{\sqrt{3}}{2}\end{pmatrix}\cdot\begin{pmatrix}\sqrt{3} & 1\\1 & \sqrt{3}\end{pmatrix}=\begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}$$

Eksempel 4.5.2. *Vi skal sette sammen to stive bevegelser. Den ene, T_1, er en rotasjon med en vinkel $\frac{\pi}{4}$ om punktet $(1,1)$, og den andre, T_2, er en refleksjon om linja $y=x$.*

a) Skriv opp formler på formen $T_i(\mathbf{v})=A_i\mathbf{v}+\mathbf{b}_i$ for de to avbildningene.

b) Regn ut komposisjonen $T_1\circ T_2$, og finn fikspunktene. Bruk det til å avgjøre hva slags stiv bevegelse dette er.

Løsning.

1. **Metodevalg:** Vi bruker formelene for rotasjon og refleksjon gitt i dette kapitlet i kompendiet.

2. **Regning:** For rotasjonen har vi

$$\begin{aligned}T_1(\mathbf{x})&=\begin{pmatrix}\cos\frac{\pi}{4} & -\sin\frac{\pi}{4}\\\sin\frac{\pi}{4} & \cos\frac{\pi}{4}\end{pmatrix}(\mathbf{x}-\begin{pmatrix}1\\1\end{pmatrix})+\begin{pmatrix}1\\1\end{pmatrix}\\&=\begin{pmatrix}\frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2}\\\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2}\end{pmatrix}\mathbf{x}+\begin{pmatrix}1\\1-\sqrt{2}\end{pmatrix}\end{aligned}$$

Refleksjon om linja $y=x$ er gitt ved

$$T_2(\mathbf{x})=\begin{pmatrix}0 & 1\\1 & 0\end{pmatrix}\mathbf{x}$$

Sammensettingen blir da

$$\begin{aligned} T_1(T_2(\mathbf{x})) &= \begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \mathbf{x} + \begin{pmatrix} 1 \\ 1-\sqrt{2} \end{pmatrix} \\ &= \begin{pmatrix} -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix} \mathbf{x} + \begin{pmatrix} 1 \\ 1-\sqrt{2} \end{pmatrix} \end{aligned}$$

Fikspunktene til denne funksjonen finner vi ved å sette $T_1(T_2(\mathbf{x})) = \mathbf{x}$. Det gir

$$\begin{pmatrix} -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix} \mathbf{x} + \begin{pmatrix} 1 \\ 1-\sqrt{2} \end{pmatrix} = \mathbf{x}$$

eller

$$\begin{pmatrix} -\frac{\sqrt{2}}{2}-1 & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2}-1 \end{pmatrix} \mathbf{x} = \begin{pmatrix} -1 \\ -1+\sqrt{2} \end{pmatrix}$$

Setter vi $\mathbf{x} = (x_1, x_2)$ så må vi løse likningssettet

$$\begin{aligned} (-\sqrt{2}-2)x_1 + \sqrt{2}x_2 &= -2 \\ \sqrt{2}x_1 + (\sqrt{2}-2)x_2 &= 2\sqrt{2}-2 \end{aligned}$$

Vi kan gange den nederste likningen med $-1-\sqrt{2}$. Det gir likningen

$$(-\sqrt{2}-2)x_1 + \sqrt{2}x_2 = -2$$

som er samme likning som den første. Det betyr at vi har en hel fiks-linje, gitt ved

$$(\sqrt{2}+2)x_1 - \sqrt{2}x_2 = 2$$

Altså er komposisjonen en refleksjon om denne linja.

Kapittel 5

Geometri i rommet

I dette kapitlet skal vi konsentrere oss om isometrier i $\mathbb{R}^3$. Det er stort sammenfall mellom teoriene i 2 og 3 dimensjoner, og mange av resultatene fra forrige kapittel er gyldig også i $\mathbb{R}^3$. Vi tar med resultatene her, men hopper over bevisene der det ikke er noe forskjell.

Vi begynner imidlertid med noen spesielle regneregler for vektorer i rommet, og som ikke nødvendigvis har noen paralleller i planet.

5.1 Vektorer i $\mathbb{R}^3$

I mange tilfeller kan det være vanskelig å skille mellom punkter og vektorer i rommet, begge deler angis med 3-tupler av reelle tall. Imidlertid er det en forskjell på dem. Et punkt P i rommet eksisterer uavhengig av valg av koordinatsystem, vi trenger faktisk ikke bestemme oss for noe koordinatsystem i det hele tatt. Beskrivelsen som et 3-tuppel derimot baserer seg på et valg av koordinatsystem, og beskriver punktets plassering i forhold til dette koordinatsystemet. En vektor $\mathbf{v}$ kan vi tenke på som en pil i rommet. En pil har en retning og en lengde, og begge disse størrelsene er uavhengig av valg av koordinatsystem. Når vi har bestemt oss for et koordinatsystem kan vi plassere en vektor (pil) med sin start-ende i origo. Den andre enden av pila peker da på et punkt i rommet. Vektoren vil da bli gitt ved koordinatene til dette punktet, relativt til det valgte koordinatsystemet.

Noen resultater og størrelser vil være avhengig av valg av koordinatsystem, mens andre vil ikke være det. En vanlig størrelse som er avhengig av valg av koordinatsystem, eller mer spesifikt valg av enhet på aksene, er lengden av en vektor $\mathbf{v}$. Når vi har valgt et koordinatsystem kan vi angi en vektor ved et 3-tuppel $\mathbf{v} = (a,b,c)$. Lengder av vektoren finner vi ved å bruke Pythagoras setning.

Definisjon 5.1.1. *La $\mathbf{v} = (a,b,c)$ være en vektor i $\mathbb{R}^3$. **Lengden** av $\mathbf{v}$ er gitt ved*

$$\|\mathbf{v}\| = (\mathbf{v}^T\mathbf{v})^{\frac{1}{2}} = \sqrt{a^2+b^2+c^2}$$

Vi skal hovedsakelig være interessert i **ortonormale** koordinatsystem. Ortonormale koordinatsystem henger sammen med ortonormale basiser for $\mathbb{R}^3$. I en ortonormal basis har basisvektorene lengde 1 og de står normalt på hverandre. **Standardbasisen**

$\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ der

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

er et eksempel på en ortonormal basis. Fortrinnet til ortonormale basiser er at det er enkelt å beregne koordinatene til et punkt eller en vektor. En vektor $\mathbf{v} \in \mathbb{R}^3$ har koordinater relativt til standardbasisen gitt ved skalarproduktet

$$v_i = \mathbf{v} \cdot \mathbf{e}_i \quad i = 1, 2, 3$$

og tilsvarende for en vilkårlig annen ortonormal basis.

Merk at det er vanlig å bruke notasjonen

$$\mathbf{i} = \mathbf{e}_1 \quad \mathbf{j} = \mathbf{e}_2 \quad \mathbf{k} = \mathbf{e}_3$$

for standardbasisen i $\mathbb{R}^3$.

Eksempel 5.1.1. *Mengden $\mathscr{B} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ der*

$$\mathbf{v}_1 = \frac{1}{3}\begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}, \ \mathbf{v}_2 = \frac{1}{3\sqrt{5}}\begin{pmatrix} 4 \\ 2 \\ -5 \end{pmatrix}, \ \mathbf{v}_3 = \frac{1}{\sqrt{5}}\begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix}$$

danner en ortonormal basis for $\mathbb{R}^3$.

I mange tilfeller vil vi ha gitt en vektor (av lengde 1), og vi ønsker å finne en ortonormal basis som involverer denne vektoren. I planet har vi en enkel formel for å finne en vektor som står normalt på en annen vektor. I 3-rommet har vi også formler for dette.

Lemma 5.1.2. *La*

$$\mathbf{u} = \begin{pmatrix} a \\ b \\ c \end{pmatrix} \in \mathbb{R}^3$$

være en vektor med lengde 1, dvs. $a^2 + b^2 + c^2 = 1$. Dersom $(a,b) \neq (0,0)$ vil vektorene

$$\mathbf{v} = \frac{1}{\sqrt{a^2+b^2}}\begin{pmatrix} -b \\ a \\ 0 \end{pmatrix}$$

og

$$\mathbf{w} = \frac{1}{\sqrt{a^2+b^2}}\begin{pmatrix} ac \\ bc \\ -(a^2+b^2) \end{pmatrix}$$

ha lengde 1 og stå normalt på hverandre og på $\mathbf{u}$*, det vil si*

$$\mathbf{v} \cdot \mathbf{v} = \mathbf{w} \cdot \mathbf{w} = 1$$
$$\mathbf{u} \cdot \mathbf{v} = \mathbf{u} \cdot \mathbf{w} = \mathbf{v} \cdot \mathbf{w} = 0$$

Bevis. Ren utregning. □

Eksempel 5.1.2. *Hvis vi bruker formelen i lemmaet på*

$$\mathbf{u} = \frac{1}{3}\begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}$$

får vi

$$\mathbf{v} = \frac{1}{\sqrt{5}}\begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix}, \quad \mathbf{w} = \frac{1}{3\sqrt{5}}\begin{pmatrix} 4 \\ 2 \\ -5 \end{pmatrix}$$

Det finnes en vektor-konstruksjon som er helt spesiell for $\mathbb{R}^3$, det såkalte **krysspro-duktet** av to vektorer. Den viktigste egenskapen til dette produktet er at den konstruerte vektoren står normalt på de to vi starter med. Dette kan vi selvfølgelig utnytte når vi skal konstruere ortonormale basiser.

Definisjon 5.1.3. *La $\mathbf{v}, \mathbf{w} \in \mathbb{R}^3$ være to vektorer. Vi definerer **kryssproduktet** $\mathbf{v} \times \mathbf{w}$ av de to vektorene ved formelen*

$$\begin{aligned} \mathbf{v} \times \mathbf{w} &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix} \\ &= (v_2 w_3 - v_3 w_2)\mathbf{i} + (v_3 w_1 - v_1 w_3)\mathbf{j} + (v_1 w_2 - v_2 w_1)\mathbf{k} \\ &= \begin{pmatrix} v_2 w_3 - v_3 w_2 \\ v_3 w_1 - v_1 w_3 \\ v_1 w_2 - v_2 w_1 \end{pmatrix} \end{aligned}$$

Eksempel 5.1.3. *Vi skal regne ut kryssproduktet $\mathbf{v} \times \mathbf{w}$ for vektorene $\mathbf{v} = (2,1,0)$ og $\mathbf{w} = (-1,2,1)$. Vi har*

$$\mathbf{v} \times \mathbf{w} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 2 & 1 & 0 \\ -1 & 2 & 1 \end{vmatrix} = \mathbf{i} - 2\mathbf{j} + 5\mathbf{k}$$

Proposisjon 5.1.4. *Kryssproduktet har følgende egenskaper ($\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^3$, $k \in \mathbb{R}$);*

i) Lineært i begge faktorer;

$$\begin{aligned} (\mathbf{u} + \mathbf{v}) \times \mathbf{w} &= \mathbf{u} \times \mathbf{w} + \mathbf{v} \times \mathbf{w} \\ \mathbf{u} \times (\mathbf{v} + \mathbf{w}) &= \mathbf{u} \times \mathbf{v} + \mathbf{u} \times \mathbf{w} \\ (k\mathbf{u}) \times \mathbf{v} &= \mathbf{u} \times (k\mathbf{v}) = k(\mathbf{u} \times \mathbf{v}) \end{aligned}$$

ii) Anti-kommutativt;

$$\mathbf{u} \times \mathbf{v} = -\mathbf{v} \times \mathbf{u}$$

iii) Ortogonalitet;

$$\mathbf{u} \cdot (\mathbf{u} \times \mathbf{v}) = \mathbf{v} \cdot (\mathbf{u} \times \mathbf{v}) = 0$$

iv) Selv-utslettende;

$$\mathbf{u} \times \mathbf{u} = 0$$

v) Tilfredsstiller Jacobi-identiteten;

$$\mathbf{u}\times(\mathbf{v}\times\mathbf{w})+\mathbf{v}\times(\mathbf{w}\times\mathbf{u})+\mathbf{w}\times(\mathbf{u}\times\mathbf{v})=0$$

Egenskap iii) sier at kryssproduktet står normalt på de to vektorene som inngår i produktet.
Bevis. Følger direkte fra definisjonen. □

Eksempel 5.1.4. *La* $\mathbf{u}_1=(1,-1,0)$, $\mathbf{u}_2=(1,1,-1)$. *Vi kan bruke kryssproduktet til å finne en tredje vektor* $\mathbf{u}_3$ *slik at de tre vektorene danner en lineært uavhengig mengde,. Utregning gir*

$$\mathbf{u}_1\times\mathbf{u}_2=\begin{vmatrix}\mathbf{i}&\mathbf{j}&\mathbf{k}\\1&-1&0\\1&1&-1\end{vmatrix}=(1,1,2)$$

Siden kryssproduktet er forskjellig fra 0 vet vi at de to vektorene $\mathbf{u}_1$, $\mathbf{u}_2$ *er lineært uavhengig. I tillegg vet vi at kryssproduktet står normalt på de to vektorene. Mengden* $\{\mathbf{u}_1,\mathbf{u}_2,\mathbf{u}_1\times\mathbf{u}_2\}$ *er derfor lineært uavhengig.*

Eksempel 5.1.5. *Vi skal regne ut Jacobi-identiteten for vektorene* $\mathbf{u}_1=(1,0,0)$, $\mathbf{u}_2=(0,1,0)$ *og* $\mathbf{u}_3=(a,b,c)$. *Vi har*

$$\begin{aligned}\mathbf{u}_1\times\mathbf{u}_2&=(0,0,1)\\\mathbf{u}_3\times\mathbf{u}_1&=(0,c,-b)\\\mathbf{u}_2\times\mathbf{u}_3&=(c,0,-a)\end{aligned}$$

som gir

$$\begin{aligned}(\mathbf{u}_1\times\mathbf{u}_2)\times\mathbf{u}_3&=(-b,a,0)\\(\mathbf{u}_3\times\mathbf{u}_1)\times\mathbf{u}_2&=(b,0,0)\\(\mathbf{u}_2\times\mathbf{u}_3)\times\mathbf{u}_1&=(0,-a,0)\end{aligned}$$

Vi ser at summen av de tre vektorene er 0.

Det er verdt å merke seg at lengden av kryssproduktet er lik arealet av parallellogrammet utspent av de to vektorene som inngår. Dett er det lett å se dersom vi antar at den ene vektoren er $\mathbf{u}=(1,0,0)$. Hvis den andre vektoren er gitt ved $\mathbf{v}=(v_1,v_2,v_3)$, så vil arealet av parallellogrammet være gitt ved $\sqrt{v_2^2+v_3^2}$. Samtidig vil kryssproduktet av $\mathbf{u}$ og $\mathbf{v}$ være $\mathbf{u}\times\mathbf{v}=(0,-v_3,v_2)$, og lengdene stemmer overens.

Formelen for den ortonormale basisen gitt i lemmaet over baserer seg på kryssproduktet. Vi har

$$\mathbf{u}=\begin{pmatrix}a\\b\\c\end{pmatrix},\quad \mathbf{v}=\frac{1}{\sqrt{a^2+b^2}}\begin{pmatrix}-b\\a\\0\end{pmatrix}$$

$$\mathbf{w}=\mathbf{v}\times\mathbf{u}=\frac{1}{\sqrt{a^2+b^2}}\begin{pmatrix}ac\\bc\\-(a^2+b^2)\end{pmatrix}$$

For standardbasisen har vi de svært vakre sammenhengene

$$\mathbf{i}\times\mathbf{j}=\mathbf{k},\quad \mathbf{j}\times\mathbf{k}=\mathbf{i},\quad \mathbf{k}\times\mathbf{i}=\mathbf{j}$$

Eksempel 5.1.6. *Vi kan bruke disse formelene sammen med linearitetsegenskapene til vektorproduktet til å forenkle utregninger: La* $\mathbf{u}_1 = \mathbf{i} - \mathbf{j}$, $\mathbf{u}_2 = \mathbf{i} + \mathbf{j} - \mathbf{k}$. *Da har vi*

$$\begin{aligned}\mathbf{u}_1 \times \mathbf{u}_2 &= (\mathbf{i} - \mathbf{j}) \times (\mathbf{i} + \mathbf{j} - \mathbf{k}) \\ &= \mathbf{i} \times \mathbf{i} + \mathbf{i} \times \mathbf{j} - \mathbf{i} \times \mathbf{k} - \mathbf{j} \times \mathbf{i} - \mathbf{j} \times \mathbf{j} + \mathbf{j} \times \mathbf{k} \\ &= 0 + \mathbf{k} + \mathbf{j} + \mathbf{k} - 0 + \mathbf{i} \\ &= \mathbf{i} + \mathbf{j} + 2\mathbf{k}\end{aligned}$$

Vi kan kombinere kryssprodukt og skalarprodukt i det som kalles trippelproduktet.

Definisjon 5.1.5. ***Trippelproduktet*** *av tre vektorer* $\mathbf{u}, \mathbf{v}, \mathbf{w}$ *er definert ved*

$$[\mathbf{u}, \mathbf{v}, \mathbf{w}] = (\mathbf{u} \times \mathbf{v}) \cdot \mathbf{w}$$

Ved å bruke definisjonen av kryssprodukt kan vi gi en fin formel for trippelproduktet;

Proposisjon 5.1.6. *La* $\mathbf{u} = (u_1, u_2, u_3)$, $\mathbf{v} = (v_1, v_2, v_3)$ *og* $\mathbf{w} = (w_1, w_2, w_3)$ *være tre vektorer i rommet. Da har vi at trippelproduktet*

$$[\mathbf{u}, \mathbf{v}, \mathbf{w}] = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix} = \det(A)$$

der A er matrisen med søyler (eller rader) $\mathbf{u}, \mathbf{v}, \mathbf{w}$.

Bevis. Følger direkte fra formelene for kryss- og skalar-produkt. □

Eksempel 5.1.7. *Vi skal regne ut trippelproduktet av vektorene* $\mathbf{u} = (1,1,0)$, $\mathbf{v} = (0,1,1)$ *og* $\mathbf{w} = (1,0,1)$. *Vi bruker formelen*

$$[\mathbf{u}, \mathbf{v}, \mathbf{w}] = \begin{vmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{vmatrix} = 2$$

Lemma 5.1.7. *Vi har*

$$(\mathbf{u} \times \mathbf{v}) \cdot \mathbf{w} = (\mathbf{v} \times \mathbf{w}) \cdot \mathbf{u} = -(\mathbf{w} \times \mathbf{v}) \cdot \mathbf{u}$$

Bevis. Følger fra proposisjonen over siden to ombyttinger av rader i en determinant ikke forandrer verdien av determinanten. □

Det siste resultatet gir at vi kan bytte om rekkefølgen i trippelproduktet uten å forandre verdien. Det eneste vi må passe på er fortegnet. Reglen er at så lenge vi beholder rekkefølgen på faktorene, så endres ikke tegnet. Med rekkefølge mener vi rekkefølge i **syklisk** forstand, det vil si at 123, 231 og 312 gir samme rekkefølge, mens 132, 321, 213 gir den motsatte rekkefølgen.

Trippelproduktet oppfører seg ganske pent i forhold til å gange vektorene med en matrise.

Proposisjon 5.1.8. *La A være en* 3×3*-matrise. Da har vi*

$$[A\mathbf{u}, A\mathbf{v}, A\mathbf{w}] = \det(A)[\mathbf{u}, \mathbf{v}, \mathbf{w}]$$

Bevis. Trippelproduktet $[A\mathbf{u}, A\mathbf{v}, A\mathbf{w}]$ er gitt ved determinanten til en matrise med de tre vektorene som sine søylevektorer. Men det er det samme som determinanten til matriseproduktet

$$A \cdot (\mathbf{u}\ \mathbf{v}\ \mathbf{w})$$

□

Eksempel 5.1.8. *La A være matrisen gitt ved*

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

og

$$\mathbf{u} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix},\ \mathbf{v} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix},\ \mathbf{w} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

Vi skal regne ut trippelproduktet

$$[A\mathbf{u}, A\mathbf{v}, A\mathbf{w}]$$

på to måter. Først beregner vi

$$A\mathbf{u} = \begin{pmatrix} 1 \\ -3 \\ 0 \end{pmatrix},\ A\mathbf{v} = \begin{pmatrix} -1 \\ -1 \\ -2 \end{pmatrix},\ A\mathbf{w} = \begin{pmatrix} 0 \\ -2 \\ 0 \end{pmatrix}$$

Det gir trippelprodukt

$$[A\mathbf{u}, A\mathbf{v}, A\mathbf{w}] = \begin{vmatrix} 1 & -3 & 0 \\ -1 & -1 & -2 \\ 0 & -2 & 0 \end{vmatrix} = -4$$

Så regner vi ut

$$\det(A) = \begin{vmatrix} 1 & 0 & -1 \\ -2 & -1 & 0 \\ 1 & -1 & -1 \end{vmatrix} = -2$$

og

$$[\mathbf{u}, \mathbf{v}, \mathbf{w}] = \begin{vmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{vmatrix} = 2$$

og produktet blir igjen -4.

5.2 Ortogonale 3×3-matriser

Ortogonale matriser A er definert gjennom relasjonen $A^T A = Id$. Alle søylene i en ortogonal matrise har derfor lengde 1 og de står normalt på hverandre, det vil si at skalar-produktet mellom to forskjellige søyler er 0. Det er historiske grunner som ligger bak navnet ortogonal matrise. Siden søylene har lengde 1 skulle man kanskje tenke seg at matrisene burde omtales som **ortonormale**. Men det gjør vi altså ikke.

Vi skal liste opp noen egenskaper ved ortogonale matriser som vi tidligere har gjennomgått, i tillegg til at det etter hvert kommer noen nye:

- Produktet av to ortogonale matriser er en ortogonal matrise.
- Determinanten til en ortogonal matrise er ± 1.
- Alle (de reelle) egenverdiene til en ortogonal $n \times n$-matrise har absoluttverdi 1.
- La A være en ortogonal 3×3-matrise med determinant $\det(A) = 1$. Da har A en egenvektor $\mathbf{v}$ med tilhørende egenverdi $\lambda = 1$.

Proposisjon 5.2.1. *La A være en ortogonal 3×3-matrise med determinant $\det(A) = -1$. Da har A en egenvektor $\mathbf{v}$ med tilhørende egenverdi $\lambda = -1$.*

Bevis. Vi har

$$\begin{aligned}\chi_A(-1) &= \det(A+I) = \det(A+AA^T) \\ &= \det(A) \cdot \det(I+A^T) = \det(A) \cdot \det(A+I) \\ &= \det(A) \cdot \chi_A(-1)\end{aligned}$$

Siden vi har antatt at $\det(A) = -1$ følger det at $\chi_A(-1) = 0$, som er ekvivalent med at $\lambda = -1$ er en egenverdi. □

Eksempel 5.2.1. *La A være den ortogonale matrisen*

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

Da er

$$\mathbf{v} = \begin{pmatrix} 1 \\ -1 \\ \sqrt{2}-1) \end{pmatrix}$$

en egenvektor med egenverdi -1.

Proposisjon 5.2.2. *La A være en ortogonal matrise, og la $\mathbf{v}$ være en egenvektor med egenverdi ± 1. La $W = \mathbf{v}^\perp$ være det **ortogonale komplementet** til $\mathbf{v}$ det vil si alle vektorer i rommet som stå normalt på $\mathbf{v}$. Da har vi $A \cdot W = W$.*

Bevis. La $\mathbf{w}$ oppfylle $\mathbf{w}^T \cdot \mathbf{v} = 0$. Da har vi

$$(A\mathbf{w})^T\mathbf{v} = \mathbf{w}^T A^T \mathbf{v} = \mathbf{w}^T A^{-1}\mathbf{v} = \pm \mathbf{w}^T \mathbf{v} = 0$$

og det følger at $A\mathbf{v}$ står normalt på $\mathbf{v}$. □

Proposisjon 5.2.3. *La A være en ortogonal 3×3-matrise, og la $\mathbf{v}, \mathbf{w} \in \mathbb{R}^3$. Da har vi*

$$A\mathbf{v} \cdot A\mathbf{w} = \mathbf{v} \cdot \mathbf{w}$$

Bevis. Vi har

$$A\mathbf{v} \cdot A\mathbf{w} = (A\mathbf{v})^T A\mathbf{w} = \mathbf{v}^T A^T A\mathbf{w} = \mathbf{v}^T\mathbf{w} = \mathbf{v} \cdot \mathbf{w}$$

□

Eksempel 5.2.2. *Vi betrakter den ortogonale* 3×3*-matrisen*

$$A = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & -\frac{2}{\sqrt{6}} \end{pmatrix}$$

og de to vektorene

$$\mathbf{v} = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}, \quad \mathbf{w} = \begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix}$$

Da har vi

$$A\mathbf{v} = \begin{pmatrix} \frac{\sqrt{2}+2}{\sqrt{6}} \\ -\frac{\sqrt{2}+2}{\sqrt{6}} \\ \frac{\sqrt{2}-4}{\sqrt{6}} \end{pmatrix}, \quad A\mathbf{w} = \begin{pmatrix} \frac{\sqrt{2}+3\sqrt{3}}{\sqrt{6}} \\ \frac{-\sqrt{2}+3\sqrt{3}}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} \end{pmatrix}$$

Videre har vi

$$\mathbf{v} \cdot \mathbf{w} = 1$$

og med litt mer regning,

$$A\mathbf{v} \cdot A\mathbf{w} = 1$$

som stemmer med resultatet over.

Resultatene og eksemplene ovenfor illustrerer hver for seg noen av de spesielle egenskapene ortogonale matriser har. Siden skal vi se at ortogonale matriser er nært knyttet til stive bevegelser i rommet. Det betyr at egenskapene til de ortogonale matrisene kan tolkes som egenskaper ved stive bevegelser. Det vil gi disse egenskapene en mer intuitiv geometrisk tolkning.

Vi har tidligere definert såkalte Householder-transformasjoner, like gjerne kalt Householder-matriser. Vi skal se at Householder-matriser faktisk er ortogonale med determinant -1.

Proposisjon 5.2.4. *La* $\mathbf{v} \neq \mathbf{0}$ *være en vektor i* $\mathbb{R}^3$*. Da er Householder-matrisen*

$$Q_{\mathbf{v}} = I - 2\frac{\mathbf{v}\mathbf{v}^T}{\mathbf{v}^T\mathbf{v}}$$

ortogonal og har determinant -1.

Bevis. Ortogonaliteten har vi bevist tidligere, så det holder å vise at determinanten er -1. Vi kan for enkelthet skyld anta at $\mathbf{v} = \mathbf{e}_1$. Siden vi i utgangspunktet kan velge å legge koordinatsystemet slik vi ønsker, mister vi ingenting ved å gjøre det på denne måten. Det gir for Householder-matrisen

$$Q_{\mathbf{e}_1} = I - 2\frac{\mathbf{e}_1\mathbf{e}_1^T}{\mathbf{e}_1^T\mathbf{e}_1} = I - 2\mathbf{e}_{11} = -\mathbf{e}_{11} + \sum_{i=2}^{n} \mathbf{e}_{ii}$$

som er en diagonal matrise med en -1-er og resten 1-ere. det betyr determinant -1. □
Vi vet at $Q_{\mathbf{v}}\mathbf{v} = -\mathbf{v}$ og at $Q_{\mathbf{v}}(\mathbf{w}) = \mathbf{w}$ for alle $\mathbf{w} \in \mathbf{v}^{\perp}$. Det betyr at Householder-matrisen gir oss en refleksjon gjennom planet $\mathbf{v}^{\perp}$, slik det neste eksemplet illustrerer.

Eksempel 5.2.3. *Vi regner ut Householder-matrisen for* $\mathbf{v} = (1,0,1)$. *Det gir*

$$Q_{\mathbf{v}} = I - 2\frac{\begin{pmatrix}1\\0\\1\end{pmatrix}(1\,0\,1)}{1^2+0^2+1^1} = \begin{pmatrix}0 & 0 & -1\\ 0 & 1 & 0\\ -1 & 0 & 0\end{pmatrix}$$

Et vilkårlig vektor i $\mathbf{v}^\perp$ *er gitt ved*

$$\begin{pmatrix}a\\b\\-a\end{pmatrix}$$

Det gir

$$Q_{\mathbf{v}}\begin{pmatrix}a\\b\\-a\end{pmatrix} = \begin{pmatrix}0 & 0 & -1\\ 0 & 1 & 0\\ -1 & 0 & 0\end{pmatrix}\begin{pmatrix}a\\b\\-a\end{pmatrix} = \begin{pmatrix}a\\b\\-a\end{pmatrix}$$

Siden standardbasisen er en ortonormal basis kan vi skrive

$$\mathbf{v} = (\mathbf{v}^T\mathbf{e}_1)\mathbf{e}_1 + (\mathbf{v}^T\mathbf{e}_2)\mathbf{e}_2 + (\mathbf{v}^T\mathbf{e}_3)\mathbf{e}_3$$

Det gir oss et direkte argument for følgende resultat:

Lemma 5.2.5. *La* $\mathbf{v},\mathbf{w} \in \mathbb{R}^3$. *Da har vi at* $\mathbf{v} = \mathbf{w}$ *hvis og bare hvis* $\mathbf{v}^T\mathbf{u} = \mathbf{w}^T\mathbf{u}$ *for alle* $\mathbf{u} \in \mathbb{R}^3$.

Proposisjon 5.2.6. *La* A *være en ortogonal* 3×3*-matrise, og la* $\mathbf{v},\mathbf{w} \in \mathbb{R}^3$. *Da har vi*

$$A\mathbf{v} \times A\mathbf{w} = (\det A)\cdot A(\mathbf{v}\times\mathbf{w})$$

Bevis. Det holder å vise at

$$[A\mathbf{v}, A\mathbf{w}, \mathbf{x}] = (\det A)A(\mathbf{v}\times\mathbf{w})\cdot\mathbf{x}$$

for alle $\mathbf{x} \in \mathbb{R}^3$. Vi har

$$\begin{aligned}[A\mathbf{v}, A\mathbf{w}, \mathbf{x}] &= [A\mathbf{v}, A\mathbf{w}, AA^{-1}\mathbf{x}]\\ &= \det(A)[\mathbf{v},\mathbf{w},A^{-1}\mathbf{x}]\\ &= \det(A)(\mathbf{v}\times\mathbf{w})\cdot A^{-1}\mathbf{x}\\ &= \det(A)A(\mathbf{v}\times\mathbf{w})\cdot AA^{-1}\mathbf{x}\\ &= \det(A)A(\mathbf{v}\times\mathbf{w})\cdot\mathbf{x}\end{aligned}$$

□

Eksempel 5.2.4. *Vi betrakter igjen den ortogonale* 3×3*-matrisen*

$$A = \begin{pmatrix}\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}}\\ -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}}\\ \frac{1}{\sqrt{3}} & 0 & -\frac{2}{\sqrt{6}}\end{pmatrix}$$

og de to vektorene

$$\mathbf{v} = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}, \quad \mathbf{w} = \begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix}$$

Vi har

$$A\mathbf{v} = \frac{1}{\sqrt{6}} \begin{pmatrix} \sqrt{2}+2 \\ -\sqrt{2}-2 \\ \sqrt{2}-4 \end{pmatrix}, \quad A\mathbf{w} = \frac{1}{\sqrt{6}} \begin{pmatrix} \sqrt{2}+3\sqrt{3} \\ -\sqrt{2}+3\sqrt{3} \\ \sqrt{2} \end{pmatrix}$$

og det følger med litt regning at

$$A\mathbf{v} \times A\mathbf{w} = \begin{pmatrix} -\sqrt{2}+2\sqrt{3}-\frac{1}{2}\sqrt{6} \\ -\sqrt{2}-2\sqrt{3}+\frac{1}{2}\sqrt{6} \\ \sqrt{6}+2\sqrt{3} \end{pmatrix}$$

Vi har også

$$\mathbf{v} \times \mathbf{w} = \begin{pmatrix} -6 \\ 2 \\ 3 \end{pmatrix}$$

og derfor

$$A(\mathbf{v} \times \mathbf{w}) = \begin{pmatrix} -2\sqrt{3}+\sqrt{2}+\frac{1}{2}\sqrt{6} \\ 2\sqrt{3}+\sqrt{2}-\frac{1}{2}\sqrt{6} \\ -2\sqrt{3}-\sqrt{6} \end{pmatrix}$$

Bruker vi at

$$\det(A) = \begin{vmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & -\frac{2}{\sqrt{6}} \end{vmatrix} = -1$$

får vi det aktuelle resultatet.

5.3 Isometrier i $\mathbb{R}^3$

I dette avsnittet skal vi se på isometrier i $\mathbb{R}^3$. Slik som tilfellet var med de ortogonale matrisene i forrige avsnitt, vil også mange av definisjonene og resultatene i dette avsnittet ha klare analogier med tilsvarende definisjoner og resultater i $\mathbb{R}^2$. I så fall gjengir vi kun resultatene, uten å ta med bevisene.

Definisjon 5.3.1. *En avbildning $m : \mathbb{R}^3 \to \mathbb{R}^3$ kalles en **isometri** (stiv bevegelse) hvis den bevarer avstander, dvs. for vilkårlig valgte punkter $P, Q \in \mathbb{R}^3$ så er $\|P-Q\| = \|m(P)-m(Q)\|$. En isometri som avbilder en delmengde $F \subset \mathbb{R}^3$ på seg selv kalles en **symmetri** av F.*

Proposisjon 5.3.2. *a) Identitetsavbildningen er en isometri.*

b) Dersom $m_1, m_2 : \mathbb{R}^3 \to \mathbb{R}^3$ er to isometrier, da er også sammensettingen av de to avbildningene en isometri.

c) Dersom m er en isometri, så er også den inverse avbildningen m^{-1} en isometri.

Mengder som oppfyller de tre punktene i proposisjonen kalles i matematikken for en **gruppe**.

Lemma 5.3.3. *La $m : \mathbb{R}^3 \to \mathbb{R}^3$ være en avbildning av planet på seg selv. Da er følgende utsagn om m ekvivalente:*

(a) Avbildningen m er en isometri som fikserer origo.

(b) Avbildningen m er venstre-multiplikasjon med en ortogonal matrise og derfor lineær.

Siden en isometri avbilder $\mathbb{R}^3$ på seg selv, vil den pr. definisjon være en symmetri av $\mathbb{R}^3$. Symmetrier som bevarer origo, og som i henhold til resultatet over er gitt ved multiplikasjon med en ortogonal matrise, kalles **punkt-symmetrier**. Vi skal i resten av kapitlet konsentrere oss om punkt-symmetrier, eller ekvivalent; ortogonale matriser. Det betyr at vi i hovedsak skal studere isometrier på formen

$$m(\mathbf{x}) = A\mathbf{x}$$

der A er en ortogonal 3×3-matrise.

Vi deler isometriene (eller punkt-symmetriene) inn i to hovedgrupper, **orienterings-bevarende**, med determinant 1 og **orienterings-reverserende** med determinant -1.

Definisjon 5.3.4. *Mengden av ortgonale 3×3-matriser kalles den **ortogonale gruppa** og skrives $O(3)$. Inneholdt i $O(3)$ ligger $SO(3)$, som er de ortogonale 3×3-matrisene med determinant 1.*

Teorem 5.3.5. *La*

$$m(\mathbf{x}) = A\mathbf{x}$$

være en punkt-symmetri definert ved en ortogonal 3×3-matrise A. Da kan A skrives som et produkt av en rotasjon og en refleksjon.

Bevis. Vi vet at en ortogonal matrise har determinant $\det(A) = \pm 1$. Hvis $\det(A) = 1$, så har vi tidligere vist at A har en egenvektor $\mathbf{v}$ med egenverdi 1, og at A tar det ortogonale komplementet $W = \mathbf{v}^\perp$ til denne vektoren på seg selv. Det betyr at restriktert til W så svarer A til en orienteringsbevarende isometri av et plan, det vil si at A svarer til en rotasjon av W.

Dersom $\det(A) = -1$, så har A en egenvektor $\mathbf{v}$ med tilhørende egenverdi -1. Matrisen A avbilder det ortogonale komplementet $\mathbf{v}^\perp$ til $\mathbf{v}$ på seg selv. Det gjør også Householder-matrisen $Q_\mathbf{v}$ og dermed også produktet av dem

$$A \cdot Q_\mathbf{v}$$

Dette produktet har determinant $(-1)^2 = 1$ og svarer til en rotasjon. Det betyr at A kan skrives som et produkt av denne rotasjonen og refleksjonen gitt ved $Q_\mathbf{v}$. □

Proposisjon 5.3.6. *To refleksjoner gir en rotasjon om skjæringslinja mellom de to refleksjonsplanene.*

Bevis. Siden refleksjoner er ortogonale matriser med determinant $\det(A) = -1$, vil produktet være en ortogonal matrise med determinant 1, altså en rotasjon. I tillegg vil denne avbildningen opplagt fiksere snitett av de to refleksjonsplanene, som dermed gir oss rotasjonsaksen. □

En ortogonal matrise som oppfyller

$$A^k = Id$$

sies å ha **endelig orden**, i dette tilfellet **orden** k. F.eks. vil en refleksjon ha orden 2.

Definisjon 5.3.7. *La $m : \mathbb{R}^3 \to \mathbb{R}^3$ være en punkt-symmetri. Et punkt $P \in \mathbb{R}^3$ som ved m avbildes på seg selv; $m(P) = P$, kalles et **fikspunkt** for m. Mengden av alle fikspunkter;*

$$\mathscr{F}_m = \{P \in \mathbb{R}^3 \,|\, m(P) = P\}$$

*kalles **fikspunktsmengden** til m.*

Klassifikasjon av punkt-symmetrier i rommet blir litt forskjellig fra det plane tilfellet;

Proposisjon 5.3.8. *La $m : \mathbb{R}^3 \to \mathbb{R}^3$ være en punkt-symmetri med fikspunktsmengde $\mathscr{F}_m$. Da har vi følgende karakterisering:*

i) Hvis $\mathscr{F}_m = \mathbb{R}^3$, så er m identitetsavbildningen.

ii) Hvis $\mathscr{F}_m$ er et plan, så er m en refleksjon.

iii) Hvis $\mathscr{F}_m$ er en linje, så er m en rotasjon.

iv) Hvis $\mathscr{F}_m$ er et punkt, så er m en sammensetting av en refleksjon og en rotasjon.

Bevis. Resultatet er en oppsummering av ting vi har vist tidligere. □

Proposisjon 5.3.9. *Enhver punkt-symmetri kan skrives som et produkt av 2 eller 3 refleksjoner.*

Bevis. På bakgrunn av hva vi har vist tidligere, holder det å vise at en vilkårlig rotasjon kan skrives som et produkt av to refleksjoner. Vi betrakter en rotasjon med rotasjonsvinkel θ. La $\mathbf{v}$ være rotasjonsaksen til rotasjonen, og velg en vektor $\mathbf{w}_1$ som står normalt på denne. La $Q_{\mathbf{w}_1}$ være den ene refleksjonen. Vi vet at denne refleksjonen fikserer normalplanet til $\mathbf{w}_1$ og derfor også $\mathbf{v}$. La videre $\mathbf{w}_2$ være en vektor som står normalt på $\mathbf{v}$ og som danner en vinkel $\frac{\theta}{2}$ med $\mathbf{w}_1$, og la $Q_{\mathbf{w}_2}$ være den andre refleksjonen. Vektoren $\mathbf{v}$ fikseres også av denne refleksjonen og $\mathbf{v}$ fikseres spesielt av begge de to refleksjonene. Nå kan vi bruke et tidligere resultat til å fastslå at rotasjonen om $\mathbf{v}$ er gitt ved $2\frac{\theta}{2} = \theta$ som var det vi ville ha. □

5.4 Oppgaver med løsning

Eksempel 5.4.1. *I denne oppgaven har vi oppgitt en vektor,*

$$\mathbf{u} = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$$

Finn to andre vektorer $\mathbf{v}$ *og* $\mathbf{w}$ *slik at de tre vektorene til sammen gir en basis for* $\mathbb{R}^3$.

Løsning.

1. **Metodevalg:** For å finne $\mathbf{v}$ skal vi bruke formelen i Lemma 5.1.2, mens $\mathbf{w}$ skal vi finne som kryssproduktet av $\mathbf{u}$ og $\mathbf{v}$.

2. **Regning:** Vi har $a = \frac{1}{\sqrt{6}}$ og $b = -\frac{1}{\sqrt{6}}$. Det gir $\sqrt{a^2+b^2} = \sqrt{\frac{1}{6}+\frac{1}{6}} = \frac{1}{\sqrt{3}}$ og dermed

$$\mathbf{v} = \frac{1}{\frac{1}{\sqrt{3}}} \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

Vektoren $\mathbf{w}$ finner vi ved

$$\begin{aligned} \mathbf{w} &= \frac{1}{\sqrt{6}} \frac{1}{\sqrt{2}} (\mathbf{i} - \mathbf{j} + 2\mathbf{k}) \times (\mathbf{i} + \mathbf{j}) \\ &= \frac{\sqrt{3}}{6} (\mathbf{i} \times \mathbf{i} + \mathbf{i} \times \mathbf{j} - \mathbf{j} \times \mathbf{i} - \mathbf{j} \times \mathbf{j} + 2\mathbf{k} \times \mathbf{i} + 2\mathbf{k} \times \mathbf{j} \\ &= \frac{\sqrt{3}}{6} (\mathbf{k} + \mathbf{k} + 2\mathbf{j} - 2\mathbf{i}) \\ &= \frac{\sqrt{3}}{6} (-2\mathbf{i} + 2\mathbf{j} + 2\mathbf{k}) = \frac{\sqrt{3}}{3} \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} \end{aligned}$$

Eksempel 5.4.2. *Vi har gitt to refleksjoner. Den ene tar* $\mathbf{v} = (0,2,1)$ *på* $-\mathbf{v}$*, den andre tar* $\mathbf{w} = (1,1,0)$ *på* $-\mathbf{w}$*. Finn skjæringslinja mellom de to refleksjonsplanene.*

Løsning.

1. **Metodevalg:** Refleksjonsplanene står normalt på de oppgitte vektorene. Skjæringslinja mellom planene vil derfor være definert av en vektor som står normalt på begge vektorene. Denne kan vi finne ved å ta kryssproduktet av de to vektorene.

2. **Regning:** Kryssproduktet er gitt ved

$$\begin{aligned} \mathbf{v} \times \mathbf{w} &= (2\mathbf{j} + \mathbf{k}) \times (\mathbf{i} + \mathbf{j}) \\ &= 2\mathbf{j} \times \mathbf{i} + 2\mathbf{j} \times \mathbf{j} + \mathbf{k} \times \mathbf{i} + \mathbf{k} \times \mathbf{j} \\ &= -2\mathbf{k} + 0 + \mathbf{j} - \mathbf{i} = (-1,1,-2) \end{aligned}$$

Skjæringslinja er linja langs vektoren $(-1,1,-2)$.

Eksempel 5.4.3. *La A være matrisen gitt ved*

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

Finn fikspunktene til matrisen og bestem hva slags isometri den representerer.

Løsning.

1. **Metodevalg:** For å finne fikspunktene til en matrise A, må vi løse $A\mathbf{v} = \mathbf{v}$. Vi lar $\mathbf{v} = (x, y, z)$ og regner ut.

2. **Regning:** Vi skal løse likningssystemet

$$\frac{1}{3}\begin{pmatrix} 2 & 2 & -1 \\ -2 & 1 & -2 \\ 1 & -2 & -2 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

Fikspunktmengden til A er det samme som nullrommet til $A - I$, som gir

$$(\frac{1}{3}\begin{pmatrix} 2 & 2 & -1 \\ -2 & 1 & -2 \\ 1 & -2 & -2 \end{pmatrix} - \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix})\begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0$$

eller

$$\frac{1}{3}\begin{pmatrix} -1 & 2 & -1 \\ -2 & -2 & -2 \\ 1 & -2 & -5 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0$$

Det betyr at hvis vi finner rangen til denne matrisen, så kan vi finne dimensjonen til nullrommet. Vi regner ut determinanten til matrisen

$$\begin{aligned} \frac{1}{3}\begin{vmatrix} -1 & 2 & -1 \\ -2 & -2 & -2 \\ 1 & -2 & -5 \end{vmatrix} &= (\frac{1}{3})^3(-10-4-4-2-20+4) \\ &= -\frac{4}{3} \neq 0 \end{aligned}$$

som betyr at rangen er 3, og nullrommet derfor har dimensjon 0. Det betyr at A uttrykker en sammensetting av en rotasjon og en refleksjon.

Kapittel 6

Gruppeteori

Gruppeteori handler om å studere **grupper**, det vil si mengder med en veldig spesifikk, men likevel enkel, struktur. Den mest sentrale delen av definisjonen av en gruppe er en binær operasjon. En binær operasjon er regel som til to elementer i en gruppe tilordner et tredje element. Eksempler på slike operasjoner er vanlig addisjon og multiplikasjon. Gitt to tall, a og b, så kan finnes det et nytt tall $a+b$ som er laget av de to andre. Tilsvarende vil også multiplikasjon gi oss et tredje tall, ab. Dette er to eksempler på **gruppeoperasjoner** på mengden $\mathbb{Z}$ av hele tall. Men grupper kan være så mye mer. Mengden av isometrier i planet utgjør også en gruppe, i det tilfellet er den binære gruppeoperasjonen sammensetting av isometrier. Felles for disse operasjonene er at de tilfredsstiller visse betingelser, som assosiativitet og eksistens av inverse elementer. Når vi skal lage en abstrakt versjon av disse konkrete eksemplene trekker vi ut et sett av vesentlige egenskaper som vi ønsker at alle grupper skal ha.

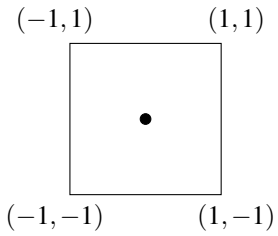
Grupper kan være endelige eller uendelige mengder og de kan ha en mer eller mindre kompleks struktur. Et eksempel på en gruppe med en enkel struktur er de hele tallene med addisjon som operasjon. En mye mer komplisert gruppestruktur er mengden av alle operasjoner på en Rubiks kube. Det finnes 43,252,003,274,489,856,000 slike og det sier seg selv at det ikke er lett å få oversikt over strukturen av denne gruppa. Selv ikke de som er gode til å løse Rubiks kube har oversikt over denne strukturen, det holder å kjenne til noen sentrale elementer og hvordan de er satt sammen av **generatorer** for gruppa for å løse kuben.

Før vi går til den abstrakte definisjonen av grupper skal vi se nærmere på strukturen til noen grupper vi allerede har vært bort i, nemlig symmetrigruppene til plane og romlige legemer.

6.1 Symmetrigrupper

Vi begynner med å se på et par eksempler på symmetrigrupper.

Eksempel 6.1.1. *Vi skal se på symmetriene til et kvadrat K med hjørner* $(\pm 1, \pm 1)$.



K har opplagt en rotasjonssymmetri med vinkel $\frac{\pi}{2}$ *gitt ved matrisen*

$$\rho = \rho_{\frac{\pi}{2}} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

og dermed er også ρ^2 *og* ρ^3 *rotasjonssymmetrier, mens* $\rho^4 = Id$. *I tillegg kan vi finne fire refleksjoner; vi kan reflektere om koordinataksene såvel som om de to diagonalene. La*

$$\mu = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

være en refleksjon om x-aksen. Da vil

$$\rho\mu = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

være en refleksjon om diagonalen $y = x$, *mens*

$$\rho^2\mu = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

er refleksjon om y-aksen og

$$\rho^3\mu = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$$

beskriver en refleksjon om diagonalen $y = -x$. *Det betyr at vi kan skrive de 8 symmetriene som*

$$Id, \rho, \rho^2, \rho^3, \mu, \rho\mu, \rho^2\mu, \rho^3\mu$$

Flere symmetrier finnes det faktisk ikke. Denne mengden av symmetrier kalles ofte for den ***dihedrale gruppa*** D_4. *Vi sier gjerne at denne gruppa er* ***generert*** *av rotasjonen* ρ *og refleksjonen* μ *siden alle elementene kan skrives som produkter av disse to. Merk at rekkefølgen på* ρ *og* μ *er vesentlig, vi har nemlig at*

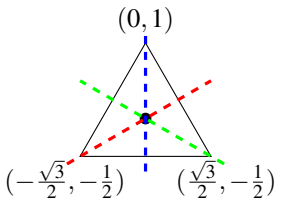
$$\mu\rho = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} = \rho^3\mu$$

så $\mu\rho \neq \rho\mu$. *Hvis vi samler alle disse opplysningene sammen kan vi skrive symmetrigruppa til kvadratet som*

$$D_4 = \langle \rho, \mu \,|\, \rho^4 = \mu^2 = Id, \mu\rho = \rho^3\mu \rangle$$

Den siste relasjonen betyr at dersom vi først roterer med en vinkel $\frac{\pi}{2}$ *og deretter speiler om x-aksen, er det det samme som å første speile om x-aksen og deretter rotere med en vinkel* $\frac{3\pi}{2} = -\frac{\pi}{2}$, *altså en rotasjon motsatt vei.*

Eksempel 6.1.2. *Vi kan gjøre tilsvarende som vi gjorde i ekempelet med kvadratet med en likesidet trekant T.*



T har en rotasjonssymmetri med vinkel $\frac{2\pi}{3}$ gitt ved matrisen

$$\rho = \rho_{\frac{2\pi}{3}} = \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}$$

Siden $3 \cdot \frac{\pi}{3} = 2\pi$ vil vi ha $\rho^3 = Id$, mens ρ^2 gir en rotasjon med en vinkel $\frac{4\pi}{3}$. Vi lar μ være refleksjon om y-aksen (den blå aksen), gitt ved

$$\mu = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

Da vil

$$\rho\mu = \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}$$

fiksere punktet $(\frac{\sqrt{3}}{2}, -\frac{1}{2})$ fordi

$$\rho\mu \begin{pmatrix} \frac{\sqrt{3}}{2} \\ -\frac{1}{2} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} \frac{\sqrt{3}}{2} \\ -\frac{1}{2} \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{3}}{2} \\ -\frac{1}{2} \end{pmatrix}$$

Siden alle lineæravbildninger fikserer origo vil $\rho\mu$ fiksere linja gjennom $(\frac{\sqrt{3}}{2}, -\frac{1}{2})$ og origo (den grønne linja). Tilsvarende vil

$$\rho^2\mu = \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}^2 \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}$$

fiksere den røde linja. Det betyr at trekanten har symmetrier

$$Id, \rho, \rho^2, \mu, \rho\mu, \rho^2\mu$$

*Denne mengden av symmetrier kalles ofte for den **symmetriske gruppa** S_3. Den er **generert** av rotasjonen ρ og refleksjonen μ siden alle elementene kan skrives som produkter av disse to. I tillegg har vi*

$$\begin{aligned} \mu\rho &= \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} = \rho^2\mu \end{aligned}$$

Dermed kan vi skrive

$$S_3 = \langle \rho, \mu \,|\, \rho^3 = \mu^2 = Id, \mu\rho = \rho^2\mu \rangle$$

Notasjonen for grupper som vi har brukt i eksemplet er en standard måte å presentere grupper. Foran den vertikale midtstreken ramser vi opp **generatorene** til gruppa, mens til høyre for midtstreken setter vi inn et minimalt sett av **relasjoner**. Det betyr at elementene i S_3 kan skrives som en eller annen sammensetting av generatorene, men hvor vi bruker relasjonene til å forenkle uttrykkene. Se f.eks. på uttrykket

$$\mu^2\rho^5\mu^3\rho\mu^4$$

Dette vil gi oss en symmetri av trekanten, men det finnes en mye enklere måte å skrive den på. For det første vet vi at $\rho^3 = \mu^2 = Id$. Det kan vi utnytte til forenklingen:

$$\mu^2\rho^5\mu^3\rho\mu^4 = (\mu^2)(\rho^3)\rho^2(\mu^2)\mu\rho(\mu^2)^2 = \rho^2\mu\rho$$

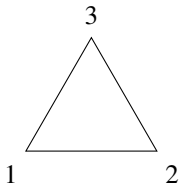
Disse to symmetriene har samme resultat siden f.eks. det å rotere en firkant fem ganger er det samme som å rotere to ganger. Ofte ønsker vi å samle generatorene sammen, og det kan vi også gjøre i dette tilfellet. Vi bruker relasjonen $\mu\rho = \rho^2\mu$ og får

$$\rho^2\mu\rho = \rho^2\rho^2\mu = \rho\mu$$

Det betyr at vi har

$$\mu^2\rho^5\mu^3\rho\mu^4 = \rho\mu$$

Vi kan se på dette på en annen måte også. Hvis vi nummerer hjørnene på trekanten som 1, 2 og 3



så kan vi uttrykke rotasjonen ρ ved $1 \to 2 \to 3 \to 1$ og μ ved $1 \to 2 \to 1$, og hvor 3 holdes i ro. Ser vi på uttrykket

$$\mu^2\rho^5\mu^3\rho\mu^4$$

så kan vi starte med 1 og se hvor vi ender. Vi begynner fra høyre med 4 ganger μ, deretter en ρ, 3 ganger μ, osv. Det gir oss

$$\begin{aligned}\mathbf{1} \to 2 \to &1 \to 2 \to 1 \to 2 \to 1 \to 2 \to 1 \to 2\\ &\to 3 \to 1 \to 2 \to 3 \to 3 \to \mathbf{3}\end{aligned}$$

og for hjørne 2;

$$\begin{aligned}\mathbf{2} \to 1 \to &2 \to 1 \to 2 \to 3 \to 3 \to 3 \to 3 \to 1\\ &\to 2 \to 3 \to 1 \to 2 \to 1 \to \mathbf{2}\end{aligned}$$

Siden to hjørner ikke kan avbildes på samme hjørne betyr det at $3 \to 1$. Vi ender altså opp ned $1 \to 3 \to 1$, og 2 holdes i ro. Dette er det samme som det mye enklere uttrykket $\rho\mu$, som tar $1 \to 2 \to 3$, $2 \to 1 \to 2$ og $3 \to 3 \to 1$.

Symmetriene til kvadratet og den likesidede trekanten består av rotasjoner og refleksjoner. Vi kan skille dem fra hverandre ved å se på determinanten til matrisene. I eksemplet med den likesidede trekanten har vi;

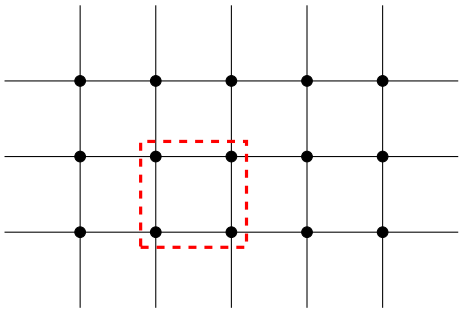
$$\det\rho = \begin{vmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{vmatrix} = (-\frac{1}{2})(-\frac{1}{2}) - (-\frac{\sqrt{3}}{2})\frac{\sqrt{3}}{2} = 1$$

mens

$$\det \mu = \begin{vmatrix} -1 & 0 \\ 0 & 1 \end{vmatrix} = (-1)\cdot 1 - 0\cdot 0 = -1$$

Produktregelen for determinanter gir da at determinanten blir 1 for alle symmetrier som inneholder en refleksjon, mens den er 1 for de rene rotasjonene.

Eksempel 6.1.3.



Et krystallmønster generert av de fire hjørnene i et kvadrat med en enhetscelle markert. Vi har allerede sett at enhetscellen har 8 symmetrier, 4 rotasjoner og 4 speilinger. I tillegg vil dette krystallmønstret ha ***translasjonssymmetrier****. Dersom hjørnene i enhetscellen er gitt ved* $(\pm 1, \pm 1)$ *vil translasjonene* $T_{(2,0)}$ *og* $T_{(0,2)}$ *være symmetrier. Sammen vil de generere alle translasjonssymmetriene, gitt ved* $T_{(2n,2m)}$ *for vilkårlige hele tall m og n. Det gir oss symmetrier*

$$D_4 \times \mathbb{Z}^2$$

Symmetriene av enhetscellen kalles i denne sammenheng for ***punktsymmetrier****.*

I det neste kapitlet skal vi lage et abstrakt rammeverk som beskriver det vi har sett på for symmetrigrupper.

6.2 Abstrakte grupper

Vi går rett på og definerer en (abstrakt) gruppe.

Definisjon 6.2.1. *En* ***gruppe*** G *er en mengde utstyrt med en binær operasjon* $\mu : G \times G \to G$ *(for enkelthet skyld skriver vi* $\mu(g,h) = gh$*) som oppfyller følgende betingelser:*

i) Assossiativitet;

$$(g_1 g_2)g_3 = g_1(g_2 g_3), \quad \text{for alle} g_1, g_2, g_3 \in G$$

ii) Eksistens av et enhetselement $e \in G$*;*

$$eg = ge = g, \quad \text{for alle} g \in G$$

iii) Eksistens av inverser; for alle $g \in G$ *så finnes et element* $g^{-1} \in G$ *slik at* $gg^{-1} = g^{-1}g = e$*.*

I den generelle definisjonen har vi skrevet gruppeoperasjonen som en multiplikasjon. I noen tilfeller er det mer naturlig å skrive operasjonen som en addisjon.

Definisjon 6.2.2. *En gruppe G sies å være **abelsk** (oppkalt etter Niels Henrik Abel) dersom for alle elementer $g,h \in G$ så er $gh = hg$. For abelske grupper skriver vi gjerne gruppeoperasjonen som $g+h$.*

Denmest berømte abelske gruppa er $\mathbb{Z}$, de hele tallene med $+$-operasjon. Vanlig addisjon er assosiativ, enhetselementet er 0 og inversen til et tall a er $-a$. Men det finnes også andre enkle abelske grupper, de såkalte **modulogruppene** $\mathbb{Z}_n$ med "klokkeaddisjon". F.eks. har vi

$$\mathbb{Z}_{12} = \{0,1,2,3,4,5,6,7,8,9,10,11\}$$

hvor vi legger sammen slik vi legger sammen klokkeslett, hver gang vi kommer til 12 begynner vi på nytt. $3+4$ er fortsatt 7, mens $8+9$ ikke er 17, men 5 $(= 17-12)$. Disse gruppene kalles også **sykliske** siden de er generert av ett element, f.eks. tallet 1; $2 = 1+1$, $3 = 1+1+1$, $4 = 1+1+1+1$ osv.

Definisjon 6.2.3. *La $H \subset G$ være en delmengde av en gruppe G som selv er en gruppe med den samme binære operasjonen som i G. Da sier vi at H er en **undergruppe** av G.*

Eksempel 6.2.1. *Vi kan skrive opp noen eksempler på undergrupper av $\mathbb{Z}_{12}$;*

$$\{0,2,4,6,8,10\} \quad \{0,3,6,9\}$$

er to undergrupper med henhodsvis 6 og 4 elementer.

Definisjon 6.2.4. *Dersom gruppa G er en endelig mengde kaller vi antall elementer i G for **ordenen** til gruppa, og vi bruker notasjonen $|G|$ for dette antallet.*

De sykliske gruppene $\mathbb{Z}_n$ har orden n, mens $\mathbb{Z}$ har uendelig orden.

Teorem 6.2.5 (Lagrange)**.** *La $H \subset G$ være en undergruppe av en endelig gruppe G. Da vil ordenen til H dele ordenen til G.*

Bevis. For hvert element $g \in G$ kan vi definere **restklassen** $gH = \{gh \,|\, h \in H\}$. Denne mengden har nøvendigvis like mange elementer som H. For to elementer $g_1, g_2 \in G$ har vi at enten så er $g_1H = g_2H$ eller $g_1H \cap g_2H = \emptyset$. Dette skyldes at dersom $g \in g_1H$, så kan vi skrive $g = g_1h$ eller $gh^{-1} = g_1$ for en $h \in H$. Da følger det at $g_1h' = gh^{-1}h_\prime \in gH$ siden H er en gruppe. Restklassene deler hele G opp i delmengder med like mange elemneter i hver som det er i H. Ordenen til H må derfor gå opp i ordenen til G. □

I eksemplet over så vi på to undergrupper av $\mathbb{Z}_{12}$. Den abelske gruppa $\mathbb{Z}_{12}$ har orden 12, og de to undergruppene hadde orden 6 og 4, som begge er tall som deler 12.

Definisjon 6.2.6. *La G være en gruppe og $g \in G$ et element. Det minste tallet d slik at $g^d = e$ kalles **ordenen** til elementet g, og vi bruker notasjonen $d = |g|$. Dersom det ikke finnes noen slik d sier vi at elementet har **uendelig orden**.*

I en abelsk gruppe hvor vi skriver operasjonen som +, vil orden til et element være det minste tallet d slik at

$$1+1+\cdots+1 = d \cdot 1 = d = 0$$

hvor vi tenker på 0 som enhetselementet i gruppa. F.eks. vil orden til elementene i $\mathbb{Z}_{12}$ være gitt ved

$$\begin{array}{lll} |0| = 1 & |1| = 12 & |2| = 6 \\ |3| = 4 & |4| = 3 & |5| = 12 \\ |6| = 2 & |7| = 12 & |8| = 3 \\ |9| = 4 & |10| = 6 & |11| = 12 \end{array}$$

Vi merker oss at alle elementer av orden 12 nødvendigvis vil generere hele gruppa. F.eks. har vi $5+5=10$, $5+5+5=3$, $5+5+5+5=8$, $5+5+5+5+5=1$. osv.

Proposisjon 6.2.7. *La $g \in G$ være et element i en endelig gruppe G. Da vil ordenen til g dele ordenen til G,*

Bevis. La $H = \{e, g, g^2, \ldots, g^{d-1}\}$ være undergruppa av G generert av g. Vi har $|g| = |H|$ og ved Lagranges teorem deler ordenen $|H|$ ordenen til gruppa. Siden ordenen til elementet g er det samme som omrdenen til H følger resultatet. □

Vi ser at for gruppa $\mathbf{Z}_{12}$ er dette resultatet oppfyllt, alle ordenene deler 12.

6.3 Den ortogonale gruppa $O(3)$

Vi har tidligere studert ortogonale 3×3-matriser, men vi repeterer for orden skyld definisjonen.

Definisjon 6.3.1. *En 3×3-matrise A kalles **ortogonal** dersom $A^T A = AA^T = Id$.*

Identitetsmatrisen er opplagt ortogonal og det følger fra definisjonen at dersom A er ortogonal, så er også $A^{-1} = A^T$ ortogonal:

$$(A^T)^T A^T = (AA^T)^T = I^T = I$$

Vi har også at komposisjonen av to ortogonale matriser er ortogonal:

$$\begin{aligned}(AB)^T(AB) &= (B^T A^T)(AB) = B^T(A^T A)B \\ &= B^T IB = B^T B = I\end{aligned}$$

Det betyr at de ortogonale matrisene tilfredsstiller alle kravene til å være en gruppe.

Definisjon 6.3.2. *Mengden av ortogonale 3×3-matriser danner en gruppe som vi kaller den **ortogonale gruppa**, $O(3)$.*

Siden det finnes uendelig mange ortogonale 3×3-matriser kan det være interessant å telle antall frihetsgrader. For første søyle kan vi velge to koordinater fritt, mens den tredje vil være bestemt (borsett fra at det vil være to muligheter) av at søylevektoren skal ha lengde 1. For den neste søylen har vi en frihetsgrad færre, siden denne også skal stå normal på den første. Når de to første søylene er bestemt, så vil den siste søylen være bestemt, igjen opp til et lite antall tilfeller der vi bytter ut noen plusser og minuser. Det betyr at at vi har 3 uendelig store frihetsgrader og en del tilfeller hvor vi kan velge fortegn.

De ortogonale matrisene har som vi har sett tidligere, determinant ± 1. De med determinant 1 har fått et eget navn.

Definisjon 6.3.3. *Delmengden av den ortogonale gruppa $O(3)$ som består av 3×3-matriser med determinant 1 kalles den **spesielle ortogonale gruppa**, med standard notasjon $SO(3)$.*

$SO(3)$ er også en gruppe, siden

- Produktet av to matriser med determinant 1 har determinant 1.

- Inversen til en matrise med determinant 1 har determinant 1.

- Identitetsmatrisen har determinant 1.

Ortogonale matriser med determinant 1 beskriver isometrier, og det samme gjelder de med determinant -1. Imidlertid vil de med determinant -1 "vrenge" et legeme. Vi kommer tilbake tildette i forbindelse med symmetrier av molekyler.

Eksempel 6.3.1. *Vi kan identifisere $SO(3)$ som symmetriene til en kuleflate. La S være kuleflaten gitt ved $x_1^2+x_2^2+x_3^2=1$, og la $(x_1,x_2,x_3) \in S$. La $A \in SO(3)$ være en ortogonal matrise med determinant 1. Da vil*

$$(y_1,y_2,y_3)A\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

også ha lengde 1, og derfor ligge på S, siden ortogonale matriser er bevarer lengder. Alle elementer i $SO(3)$ vil derfor være symmetrier.

Enhetsvektorene $\mathbf{i}$, $\mathbf{j}$ *og* $\mathbf{k}$ *ligger alle på kuleflaten og en symmetri $\phi : S \to S$ må derfor avbilde dem på vektorer av lengde 1. I tillegg står de tre enhetsvektorene normalt på hverandre og siden symmetrier må bevare vinkler, må også bildene av dem stå normalt på hverandre. Det betyr at symmetrien ϕ kan uttrykkes som multiplikasjon med en matrise der søylene står normalt på hverandre og har lengde 1. Når vi også tar med at en symmetri av kuleflaten må bevare orienteringen av enhetsvektorene, kommer vi fram til en hver symmetri kan representeres med et element i $SO(3)$.*

Kuleflaten er det romlige objektet som er mest symmetrisk. Andre legemer vil normalt ha mye færre symmetrier, men som for kuleflaten kan vi oppfatte symmetrier som ortogonale matriser. Siden identiteten alltid er en symmetri, inversen til en symmetri av et legeme er en symmetri av legemet og sammensetting av to symmetrier gir en ny symmetri, så vil mengden av symmetrier til et romlig legeme alltid danne en undergruppe av $SO(3)$.

Vi har tidligere vist at en ortogonal 3×3-matrise A med determinant 1 har en egenvektor $\mathbf{v}$ med egenverdi $\lambda = 1$, det vil si at det finnes en vektor som fikseres av A. A uttrykker dermed en rotasjon med rotasjonsakse $\mathbf{v}$. Vi skriver

$$A = \rho_{\mathbf{v},\theta}$$

hvor $\mathbf{v}$ er rotasjonsaksen og θ er rotasjonsvinkelen. La

$$Z_\alpha = \rho_{\mathbf{e}_3,\alpha}$$

være en rotasjon om z-aksen med vinkel α og

$$Y_\beta = \rho_{\mathbf{e}_2,\beta}$$

en rotasjon med en vinkel β om y-aksen. Vi skal se at vi kan skrive en vilkårlig rotasjon $A = \rho_{\mathbf{v},\theta}$ som et produkt av en Z_α og en Y_β.

Vi begynner med å rotere rundt z-aksen til $\mathbf{v}$ ligger i xz-planet. Det er intuitivt opplagt at vi alltid kan få til det, men kan også se det analytisk: xz-planet kjennetegnes ved at y-koordinaten er 0. Rotasjoner om z-aksen er gitt ved matriser

$$\begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Hvis vi skriver $\mathbf{v} = (v_1, v_2, v_3)$ får vi y-koordinaten til produktet $Z_\alpha \mathbf{v}$ gitt ved $\sin\alpha v_1 + \cos\alpha v_2$ som blir 0 dersom vi velger α slik at

$$\frac{\sin\alpha}{\cos\alpha} = -\frac{v_2}{v_1}$$

og $\cos\alpha = 0$ dersom $v_1 = 0$. Det betyr at Z_α har fraktet $\mathbf{v}$ til xz-planet og en rotasjon om y-aksen bringer resultatet videre til z-aksen. Det betyr at produktet

$$Y_\beta Z_\alpha$$

tar rotasjonsaksen $\mathbf{v}$ til z-aksen. Vi kan fortsette med en rotasjon om z-aksen, $Z_{\theta'}$ for så å rotere z-aksen tilbake til $\mathbf{v}$ med

$$Z_{-\alpha} Y_{-\beta}$$

Konklusjonen blir at en rotasjon med vinkel θ og akse $\mathbf{v}$ kan skrives som et produkt

$$Z_{-\alpha} Y_{-\beta} Z_{\theta'} Y_\beta Z_\alpha$$

og vi har vist at $SO(3)$ er generert av Z_α og Y_β.

Veldig mange legemer (f.eks. molekyler) kjennetegnes ved at de har kun endelig mange symmetrier. En kube (terning) har 24 symmetrier og et tetraeder har 12. Vi er derfor spesielt interessert i endelige undergrupper av $O(3)$.

For at en rotasjon skal ha endelig orden må vinkelen vi roterer med være et rasjonalt tall ganger 2π, dvs.

$$\theta = \frac{2n\pi}{m}, \quad m, n \in \mathbb{Z}$$

Hvis vi antar at $2n$ og m ikke har noen felles faktorer, så vil m være det minste hele tallet slik at $m\theta$ er et helt antall ganger 2π. Husk at vi ikke kan se forskjell på en symmetri som er en rotasjon med vinkelen 2π og identiteten.

Det viser seg, uten at vi skal gå inn på beviset her, at det kun finnes en begrenset liste av undergrupper av $SO(3)$.

Teorem 6.3.4. *Følgende liste gir en komplett beskrivelse av de endelige undergruppene av $SO(3)$:*

i) Endelige sykliske grupper C_n for $n \geq 1$,

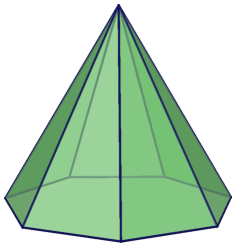
ii) Dihedrale grupper D_n for $n \geq 2$,

iii) Symmetrigruppa til et tetraeder; A_4,

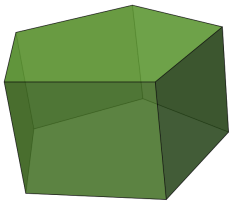
iv) Symmetrigruppa til en kube eller et oktaeder; S_4,

v) Symmetrigruppa til et dodekaeder eller et ikosaeder; A_5

For å illustrere teoremet kan vi se på figurer av de ulike legemene:



En 7-kantet pyramide har symmetrigruppe $C_7 = \langle \theta_{\frac{2\pi}{7}} \rangle$. De eneste symmetriene er rotasjoner om sentralaksen.

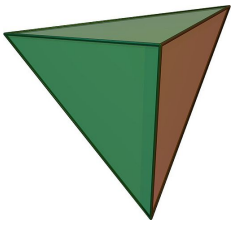


Et 5-kantet prisme har rotasjonssymmetrier av orden 5, samt en rotasjon med vinkel π om en akse midt mellom de to 5-kantene. De to symmetriene er relatert til hverandre ved

$$\theta_{\frac{2\pi}{5}}\mu = \mu\theta_{-\frac{2\pi}{5}}$$

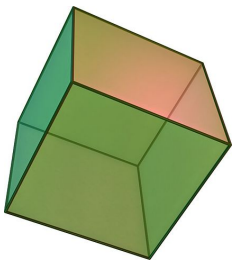
der $\theta = \theta_{\frac{2\pi}{5}}$ er rotasjonen av orden 5 og μ er rotasjonen av orden 2. Det betyr at symmetrigruppa er den dihedrale gruppa

$$D_5 = \langle \theta, \mu \,|\, \theta^5 = Id,\, \mu^2 = Id,\, \theta\mu = \mu\theta^4 \}$$

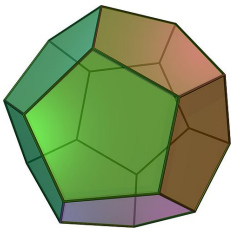


Tetraederet har 12 symmetrier generert av to rotasjoner av orden 3, men rundt hver sin akse. Hvis vi nummererer hjørnene i tetraederet med 1,2,3 og 4, kan vi beskrive den ene rotasjonen, ρ, ved $1 \to 2 \to 3 \to 1$, mens 4 holdes i ro, og den andre, μ, ved $1 \to 2 \to 4 \to 1$ og 3 holdes i ro. Både ρ og μ har orden 3, mens sammensettingen $\rho\mu$ er gitt ved $1 \to 3 \to 1$ og $2 \to 4 \to 2$ og har orden 2. Det samme gjelder for $\mu\rho$. Setter vi $\alpha = \rho\mu$ og $\beta = \mu\rho$ kan vi skrive symmetrigruppa i dette tilfellet som

$$\begin{aligned}\langle \rho, \alpha, \beta \,|\, \rho^3 = \alpha^2 = \beta^2 = Id,\\ \alpha\rho = \rho\beta, \beta\rho = \rho\alpha\beta, \alpha\beta = \beta\alpha \rangle\end{aligned}$$



En kube har 6 flater som hver er et kvadrat, noe som gir oss $4 \cdot 6 = 24$ symmetrier. Vi har 3 rotasjonsakser for rotasjoner av orden 4, samt 4 rotasjonsakser (diagonalene) for rotasjoner av orden 3.



Dodekaederet består av 12 5-kanter og har derfor $12 \cdot 5 = 60$ symmetrier.

6.4 Symmetrier av molekyler

Symmetrier av molekyler er stive isometrier, hvor vi inkluderer muligheten av å "vrenge" eller speile et molekyl, dvs. at vi aksepterer isometrier med determinant -1. Det betyr at vi holder oss i gruppa $O(3)$ og at vi er interessert i endelige undergrupper. Elementene i symmetrigruppene vil være av 3 ulike typer, eller eventuelt sammensettinger av dem. De tre typene er

i) Rotasjoner av endelig orden om en rotasjonsakse gjennom molekylet,

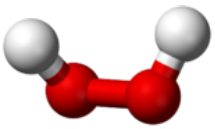
ii) Refleksjoner om et plan,

iii) "Vrenge" et molekyl, essensielt gitt ved $-Id$, som for 3×3-matriser har determinant -1.

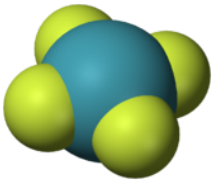
Her er noen eksempler på molekyler med ulike symmetrigrupper:



Sulfur-hexafluorid, SF_6, har symmetrigruppe som en kube eller et oktaeder.



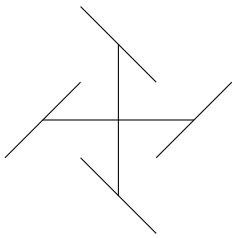
Hydrogenperoksyd, H_2O_2, har symmetrigruppe C_2.



Xenon-tetrafluorid, XeF_4, med symmetrigruppe D_4.

6.5 Oppgaver med løsning

Eksempel 6.5.1. *Beskriv symmetriene til figuren*



Løsning.

1. **Metodevalg:** Mulige symmetrier er rotasjoner og refleksjoner. Her er det bare å lete.

2. **Regning:** Figuren har en rotasjonssymmetri med vinkel $\frac{\pi}{2}$. Ingen refleksjoner.

Eksempel 6.5.2. *a) Finn alle generatorer til den sykliske gruppa* $\mathbb{Z}_{12}$.

b) Finn alle undergrupper til $\mathbb{Z}_{12}$.

Løsning.

1. **Metodevalg:** Generatorene til en abelsk gruppe G er et element a slik at $G = \{0, a, 2a, 3a, \ldots, (d-1)a\}$ hvor d er orden til gruppa. Undergrupper er delmengder av G som er lukket under addisjon.

2. **Regning:** Vi kan teste alle elementene i $\mathbb{Z}_{12}$ og se hva slags undergruppe de genererer. Vi skriver $\langle a \rangle$ for undergruppa generert av a:

$$\begin{aligned}
\langle 0 \rangle &= \{0\} \\
\langle 1 \rangle &= \mathbb{Z}_{12} \\
\langle 2 \rangle &= \{0,2,4,6,8,10\} \\
\langle 3 \rangle &= \{0,3,6,9\} \\
\langle 4 \rangle &= \{0,4,8\} \\
\langle 5 \rangle &= \{0,5,10,3,8,1,6,11,4,9,2,7\} = \mathbb{Z}_{12} \\
\langle 6 \rangle &= \{0,6\} \\
\langle 7 \rangle &= \{0,7,2,9,4,11,6,1,8,3,10,5\} = \mathbb{Z}_{12} \\
\langle 8 \rangle &= \{0,8,4\} \\
\langle 9 \rangle &= \{0,9,6,3\} \\
\langle 10 \rangle &= \{0,10,8,6,4,2\} \\
\langle 11 \rangle &= \}0,11,10,9,8,7,6,5,4,3,2,1\} = \mathbb{Z}_{12}
\end{aligned}$$

Vi ser at tallene 1, 5, 7 og 11 genererer hele gruppa, mens undergruppene er

$$\{0,6\} \quad \{0,4,8\} \quad \{0,3,6,9\} \quad \{0,2,4,6,8,10\}$$

Kapittel 7

Indreprodukt

Skalarproduktet av vektorer er et nyttig verktøy. Vi har sett at det kan brukes til å regne ut lengder av vektorer og å fastslå om vektorer står vinkelrett på hverandre. I tillegg har det en veldig enkel form og er lett å regne ut. Det er derfor nærliggende å spørre seg om det er mulig å generalisere skalarproduktet til andre og mer kompliserte vektorrom enn $\mathbb{R}^2$ og $\mathbb{R}^3$. I dette kapitlet skal vi se hvorden vi kan gjøre det. Måten vi skal gå fram på er å plukke ut de vesentlige egenskapene til skalarproduktet og introdusere et generelt begrep, et **indreprodukt**, som har akkurat disse egenskapene. Når et slikt generelt begrep er på plass kan vi prøve å finne ut hvilke egenskaper som følger av det generelle oppsettet, uten at vi baserer oss på egenskaper som er spsifikke for skalarproduktet.

7.1 Et generelt indreproduktbegrep

Hvilke egenskaper ved skalarproduktet skal vi plukke ut, og opphøye til generelle betingelser? Det mest grunnleggende ved skalarproduktet er at input er to vektorer av samme lengde, mens output er et reelt tall. Dette lar vi være grunnlaget når vi generaliserer begrepet; til et par av elementer i et generelt vektorrom tilordner vi et reelt tall. Vi skriver det som

$$\langle \mathbf{v}, \mathbf{w} \rangle \in \mathbb{R}$$

hvor $\mathbf{v}, \mathbf{w} \in V$, og V er et generelt vektorrom. Vektorrom er lineære rom der $a\mathbf{v} + b\mathbf{w} \in V$ når $\mathbf{v}, \mathbf{w} \in V$ og $a, b \in \mathbb{R}$. Dette er en egenskap vi gjerne vil bruke, og på samme måte som at skalarproduktet er linært i hver variabel, skal også indreproduktet ha denne egenskapen.

Siden skalarproduktet av en vektor med seg selv alltid er et ikke-negativt tall, og dette er en vesentlig side ved det at skalarproduktet gir oss lengder av vektorer, tar vi det med også. Den eneste vektoren som har lengde 0 er $\mathbf{0}$-vektoren og det er også en viktig egenskap. Rekkefølgen på faktorene i skalproduktet er uvesentlig, siden $\mathbf{v} \cdot \mathbf{w} = \mathbf{w} \cdot \mathbf{v}$ og det ønsker vi også å bruke. Dermed ender vi opp med følgende generelle definisjon:

Definisjon 7.1.1. *La V være et reelt vektorrom. Et **indreprodukt** på V er en regel som til et hvert par av elementer i V tilordner et reelt tall;*

$$\langle \mathbf{v}, \mathbf{w} \rangle \in \mathbb{R}$$

slik at følgende betingelser er oppfylt:

i) Indreproduktet er lineært i hver variabel, det vil si at

$$\langle a\mathbf{u}+b\mathbf{v},\mathbf{w}\rangle = a\langle \mathbf{u},\mathbf{w}\rangle + b\langle \mathbf{v},\mathbf{w}\rangle$$
$$\langle \mathbf{u},a\mathbf{v}+b\mathbf{w}\rangle = a\langle \mathbf{u},\mathbf{v}\rangle + b\langle \mathbf{u},\mathbf{w}\rangle$$

ii) Indreproduktet er symmetrisk, det vil si at

$$\langle \mathbf{u},\mathbf{v}\rangle = \langle \mathbf{v},\mathbf{u}\rangle$$

*iii) Indreproduktet er **positivt definit**, det vil si at*

$$\langle \mathbf{v},\mathbf{v}\rangle \geq 0$$

med likhet hvis og bare hvis $\mathbf{v}=0$.

Det vanlige skalarproduktet oppfyller alle disse betingelsene og danner dermed et indreprodukt på $\mathbb{R}^n$. Før vi begynner å lete etter andre indreprodukt skal vi se på en del egenskaper som vi kjenner igjen fra skalaproduktet, men som vi nå kan vise gjelder på bakgrunn av den generelle definisjonen av et indreprodukt.

Kravet om at indreproduktet skal være positivt definit, bygger på at skalarproduktet av en vektor med seg selv gir oss (kvadratet av) lengden av vektoren. Vi begynner derfor med å generalisere begrepet lengde, til det som kalles **normen** til en vektor:

Definisjon 7.1.2. *La* $\langle\ ,\ \rangle : V\times V\to\mathbb{R}$ *være et indreprodukt på et generelt vektorrom* V*. Da kaller vi*

$$||\mathbf{v}|| = \langle \mathbf{v},\mathbf{v}\rangle^{\frac{1}{2}}$$

*for **normen** til* $\mathbf{v}\in V$.

Hvis $\mathbf{v}\in\mathbb{R}^3$, så er normen til $\mathbf{v}$ det samme som vanlig euklidsk lengde, gitt ved $\sqrt{v_1^2+v_2^2+v_3^3}$. I alle de neste proposisjonene lar vi V være et generelt vektorrom og $\langle\ ,\ \rangle : V\times V\to\mathbb{R}$ et indreprodukt på V.

Proposisjon 7.1.3. *For alle* $\mathbf{u},\mathbf{v}\in V$ *gjelder **Cauchy-Schwartz** ulikhet;*

$$\langle \mathbf{u},\mathbf{v}\rangle \leq ||\mathbf{u}||\cdot||\mathbf{v}||$$

Bevis. Dersom $\mathbf{v}=\mathbf{0}$ er resultatet opplagt sant. Så vi antar $\mathbf{v}\neq\mathbf{0}$ og lar $\lambda = \frac{\langle \mathbf{u},\mathbf{v}\rangle}{||\mathbf{v}||^2}\in\mathbb{R}$. Da har vi

$$\begin{aligned}
0 &\leq ||\mathbf{u}-\lambda\mathbf{v}||^2\\
&= ||\mathbf{u}||^2 - 2\lambda\langle \mathbf{u},\mathbf{v}\rangle + \lambda^2||\mathbf{v}||^2\\
&= ||\mathbf{u}||^2 - 2\frac{\langle \mathbf{u},\mathbf{v}\rangle}{||\mathbf{v}||^2}\langle \mathbf{u},\mathbf{v}\rangle + \frac{\langle \mathbf{u},\mathbf{v}\rangle}{||\mathbf{v}||^4}||\mathbf{v}||^2\\
&= ||\mathbf{u}||^2 - \frac{\langle \mathbf{u},\mathbf{v}\rangle}{||\mathbf{v}||^2}\langle \mathbf{u},\mathbf{v}\rangle
\end{aligned}$$

og det følger at

$$||\mathbf{u}||^2\cdot||\mathbf{v}||^2 \geq \langle \mathbf{u},\mathbf{v}\rangle^2$$

□

Proposisjon 7.1.4. *For alle* $\mathbf{v}, \mathbf{w} \in V$ *gjelder* ***trekantulikheten****;*

$$||\mathbf{v}+\mathbf{w}|| \leq ||\mathbf{v}|| + ||\mathbf{w}||$$

Bevis. Vi har

$$\begin{aligned} ||\mathbf{v}+\mathbf{w}||^2 &= \langle \mathbf{v}+\mathbf{w}, \mathbf{v}+\mathbf{w} \rangle \\ &= ||\mathbf{v}||^2 + 2\langle \mathbf{v}, \mathbf{w} \rangle + ||\mathbf{w}||^2 \\ &\leq ||\mathbf{v}||^2 + 2||\mathbf{v}|| \cdot ||\mathbf{w}|| + ||\mathbf{w}||^2 \\ &= (||\mathbf{v}|| + ||\mathbf{w}||)^2 \end{aligned}$$

□

For skalarprodukt i $\mathbb{R}^2$ og $\mathbb{R}^3$ har vi formelen som sier at

$$\langle \mathbf{v}, \mathbf{w} \rangle = ||\mathbf{v}|| \cdot ||\mathbf{w}|| \cos\theta$$

der θ er vinklen mellom vektorene $\mathbf{v}$ og $\mathbf{w}$. Siden $\cos\theta$ alltid ligger mellom -1 og 1, kan vi bruke Cauchy-Schwartz ulikhet til å definere et abstrakt vinkelbegrep (som passer sammen med det vanlige begrepet i planet eller rommet):

Definisjon 7.1.5. *La* $\mathbf{u}, \mathbf{v} \in V$ *være to elementer i et generelt vektorrom* V*. Da definerer vi* ***vinkelen*** θ *mellom* $\mathbf{u}$ *og* $\mathbf{v}$ *ved at*

$$\cos\theta = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{||\mathbf{u}|| \cdot ||\mathbf{v}||}$$

I analogi med skalarproduktet sier vi at to vektorer $\mathbf{v}, \mathbf{w} \in V$ **står normalt på hverandre** eller er **ortogonale** dersom $\langle \mathbf{v}, \mathbf{w} \rangle = 0$. Det gir oss en **Pythagoras setning** for et hvert indreprodukt:

Proposisjon 7.1.6. *Dersom to elementer* $\mathbf{v}, \mathbf{w} \in V$ *i et vektorrom er ortogonale, så oppfyller de Pythagoras setning;*

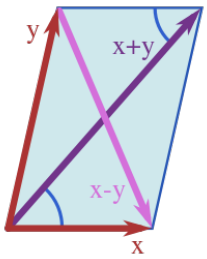
$$||\mathbf{v}||^2 + ||\mathbf{w}||^2 = ||\mathbf{v}+\mathbf{w}||^2$$

Bevis. Vi har

$$\begin{aligned} ||\mathbf{v}+\mathbf{w}||^2 &= \langle \mathbf{v}+\mathbf{w}, \mathbf{v}+\mathbf{w} \rangle \\ &= ||\mathbf{v}||^2 + 2\langle \mathbf{v}, \mathbf{w} \rangle + ||\mathbf{w}||^2 \\ &= ||\mathbf{v}||^2 + ||\mathbf{w}||^2 \end{aligned}$$

□

I tillegg til (u-)likhetene over har vi **parallellogram-loven**. Den sier at kvadratsummen av sidekantene i et parallellogram er lik kvadratsummen av diagonalene.



Proposisjon 7.1.7. *For alle* $\mathbf{v}, \mathbf{w} \in V$ *gjelder* ***parallellogram-loven****;*

$$||\mathbf{v}+\mathbf{w}||^2 + ||\mathbf{v}-\mathbf{w}||^2 = 2||\mathbf{v}||^2 + 2||\mathbf{w}||^2$$

Bevis. Vi har

$$\begin{aligned} ||\mathbf{v}+\mathbf{w}||^2 + ||\mathbf{v}-\mathbf{w}||^2 &= ||\mathbf{v}||^2 + 2\langle \mathbf{v}, \mathbf{w}\rangle + ||\mathbf{w}||^2 \\ &\quad + ||\mathbf{v}||^2 - 2\langle \mathbf{v}, \mathbf{w}\rangle + ||\mathbf{w}||^2 \\ &= 2||\mathbf{v}||^2 + 2||\mathbf{w}||^2 \end{aligned}$$

□

Et nært beslektet resultat til parallellogram-loven er den såkalte **polariserings-identiteten**;

$$\langle \mathbf{v}, \mathbf{w}\rangle = \frac{1}{4}\left(||\mathbf{v}+\mathbf{w}||^2 - ||\mathbf{v}-\mathbf{w}||^2\right)$$

7.2 Eksempler på indreprodukt

Det vanligste indreproduktet er selvfølgelig skalarproduktet på $\mathbb{R}^n$. Det er gitt ved

$$\mathbf{v}\cdot\mathbf{w} = v_1 w_1 + \cdots + v_n w_n$$

for to vektorer $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ og $\mathbf{w} = (w_1, w_2, \ldots, w_n)$. Med vår generelle definisjon av et indreprodukt kan vi imidlertid finne mange andre varianter.

For å bevise den neste proposisjonen trenger vi et hjelperesultat.

Lemma 7.2.1. *La* A *være en symmetrisk* 2×2*-matrise med to forskjellige egenverdier. La* $\mathbf{v}_1$ *og* $\mathbf{v}_2$ *være egenvektorer assosiert med hver sin egenverdi. Da har vi*

$$\mathbf{v}_1\cdot\mathbf{v}_2 = 0$$

Bevis. Siden $\mathbf{v}_1^T A\mathbf{v}_2$ er en 1×1-matrise er den lik sin egen transponerte, og siden $A^T = A$ følger det at

$$\mathbf{v}_1^T A\mathbf{v}_2 = \left(\mathbf{v}_1^T A\mathbf{v}_2\right)^T = \mathbf{v}_2^T A^T\mathbf{v}_1 = \mathbf{v}_2^T A\mathbf{v}_1$$

Det betyr at

$$0 = \mathbf{v}_1^T A\mathbf{v}_2 - \mathbf{v}_2^T A\mathbf{v}_1 = \lambda_2\mathbf{v}_1^T\mathbf{v}_2 - \lambda_1\mathbf{v}_2^T\mathbf{v}_1$$

Men nå har vi $\mathbf{v}_1^T\mathbf{v}_2 = \mathbf{v}_2^T\mathbf{v}_1$ og det følger at

$$0 = (\lambda_2 - \lambda_1)\mathbf{v}_1^T\mathbf{v}_2$$

Siden $\lambda_1 \neq \lambda_2$ følger resultatet. □

Proposisjon 7.2.2. *La* A *være en symmetrisk* 2×2*-matrise med to positive egenverdier. Da beskriver*

$$\langle \mathbf{x}, \mathbf{y}\rangle_A = \mathbf{x}^T A\mathbf{y}$$

et indreprodukt på $\mathbb{R}^2$*.*

Siden vi bruker notasjonen $\langle , \rangle$ som en fellesbetegnelse for indreprodukt, hekter vi på en indeks for å markere ulike former for indreprodukter. I dette tilfellet hekter vi på en A siden matrisen A er en vesentlig del av definisjonen.

Bevis. Linearitet og symmetri er nokså opplagt oppfylt, så det gjenstår å vise at indreproduktet er positivt definit. Et hvert element i $\mathbb{R}^2$ kan skrives som en lineærkombinasjon av egenvektorer for A, $\mathbf{x} = \mathbf{v}_1 + \mathbf{v}_2$. Det gir

$$\begin{aligned}\langle \mathbf{x}, \mathbf{x} \rangle_A &= \mathbf{x}^T A \mathbf{x} \\ &= \mathbf{v}_1^T A \mathbf{v}_1 + \mathbf{v}_1^T A \mathbf{v}_2 + \mathbf{v}_2^T A \mathbf{v}_1 + \mathbf{v}_2^T A \mathbf{v}_2 \\ &= \lambda_1 \mathbf{v}_1 \cdot \mathbf{v}_1 + \lambda_2 \mathbf{v}_1^T \mathbf{v}_2 + \lambda_1 \mathbf{v}_2^T \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 \cdot \mathbf{v}_2 \\ &= \lambda_1 ||\mathbf{v}_1||^2 + \lambda_2 ||\mathbf{v}_2||^2 \geq 0\end{aligned}$$

siden egenverdiene er positive. Det betyr også at produktet er 0 hvis og bare hvis $\mathbf{v}_1 = \mathbf{v}_2 = \mathbf{0}$. □

Eksempel 7.2.1. *La A være 2×2-matrisen*

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

Matrisen er symmetrisk med positive egenverdier $\lambda_1 = 1$ og $\lambda_2 = 3$. Det betyr at

$$\langle , \rangle_A$$

gir et indreprodukt på $\mathbb{R}^2$. Vi kan se at matrisen er positivt definit ved å betrakte utregningen

$$\begin{aligned}\langle \mathbf{v}, \mathbf{v} \rangle &= (v_1, v_2) \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \\ &= 2v_1^2 - 2v_1 v_2 + 2v_2^2 \\ &= v_1^2 + (v_1 - v_2)^2 + v_2^2 \geq 0\end{aligned}$$

siden dette er en sum av kvadrater, og de er alltid ikke-negative. Produktet er også 0 hvis og bare hvis vi har $v_1 = v_2 = 0$.

Selv om vi vanligvis betrakter matriser som lineæravbildninger, kan vi også se på rommet av matriser som et vektorrom. Vi lar $M_{n \times m}(\mathbb{R})$ betegne mengden av $n \times m$-matriser med reelle elementer. Summen av to matriser er en ny matrise, og vi kan multiplisere matriser med et reelt tall. Det finnes en $\mathbf{0}$-matrise og multiplikasjon med -1 gir oss additivt inverse matriser. Det betyr at denne mengden oppfyller alle kravene til å være et vektorrom.

Vi skal definere et indreprodukt på dette vektorrommet. For to $n \times m$-matriser A og B definerer vi

$$\langle A, B \rangle_{n \times n} = \operatorname{tr}(A^T B)$$

Vi minner om at **trasen** til en kvadratisk matrise $\operatorname{tr}(A)$ er summen av diagonalelementene til A;

$$\operatorname{tr}(A) = a_{11} + a_{22} + \cdots + a_{nn}$$

Merk at produktet $A^T B$ alltid vil være kvadratisk selv om A og B ikke er det.

Proposisjon 7.2.3. *La A og B være to $n \times m$-matriser. Da definerer*

$$\langle A, B \rangle_{n \times n} = \operatorname{tr}(A^T B)$$

et indreprodukt på vektorrommet av $n \times n$-matriser, kalt ***Frobenius-indreproduktet****.*

Bevis. Linearitet følger greit siden både matriseprodukt og trase er lineære operasjoner. La $A = (a_{ij})$ og $B = (b_{ij})$. Da har vi

$$(A^T B)_{kl} = a_{k1}b_{l1} + a_{k2}b_{l2} + \cdots + a_{kn}b_{ln}$$

og

$$\begin{aligned}\operatorname{tr}(A^T B) = &(a_{11}b_{11} + a_{12}b_{12} + \cdots + a_{1n}b_{1n}) \\ &+ \cdots + (a_{m1}b_{m1} + a_{m2}b_{m2} + \cdots + a_{mn}b_{mn})\end{aligned}$$

Produktet er åpenbart symmetrisk. Setter vi $A = B$ inn i formelen, får vi

$$\begin{aligned}\operatorname{tr}(A^T A) = &(a_{11}^2 + a_{12}^2 + \cdots + a_{1n}^2) \\ &+ \cdots + (a_{m1}^2 + a_{m2}^2 + \cdots + a_{mn}^2)\end{aligned}$$

Det er klart at denne summen av kvadrater alltid er positiv (eller 0), og at den er 0 hvis og bare hvis alle leddene er 0. Men det er det samme som å si at $A = 0$. □

Vi kan også definere en helt annen type indreprodukt. La $f(x)$ og $g(x)$ være to integrerbare funksjoner definert på et intervall $[a, b]$. La

$$\langle f(x), g(x) \rangle_{\int} = \int_a^b f(x)g(x)\,dx$$

Proposisjon 7.2.4. *Produktet $\langle\,,\,\rangle_{\int}$ definerer et indreprodukt.*

Bevis. Det er klart at produktet er lineært i hver variabel, og at det er symmetrisk. Vi har

$$\langle f(x), f(x) \rangle_{\int} = \int_a^b f(x)^2\,dx \geq 0$$

fordi integranden er ≥ 0 og av samme grunn så er produktet 0 hvis og bare hvis $f(x) = 0$, det vil si konstantfunksjonen lik 0. □

Eksempel 7.2.2. *Vi lar $V = \mathscr{P}_1$, dvs. vektorrommet av polynomer av grad mindre enn eller lik 1. Et generelt element i V er på formen*

$$f(x) = ax + b$$

og vi kan regne ut indreproduktet $\langle f(x), f(x) \rangle_{\int}$ over intervallet $[0, 1]$ av to vilkårlige førstegradspolynomer.

$$\begin{aligned}\langle f_1(x), f_2(x) \rangle_{\int} &= \int_0^1 (a_1x + b_1)(a_2x + b_2)\,dx \\ &= \int_0^1 a_1a_2x^2 + (a_1b_2 + a_2b_1)x + b_1b_2\,dx \\ &= [\frac{a_1a_2}{3}x^3 + \frac{a_1b_2 + a_2b_1}{2}x^2 + b_1b_2x]_0^1 \\ &= \frac{a_1a_2}{3} + \frac{a_1b_2 + a_2b_1}{2} + b_1b_2\end{aligned}$$

Vektorrommet $\mathscr{P}_1$ er 2-dimensjonalt og vi kan identifisere det med $\mathbb{R}^2$ gjennom korrespondansen

$$ax+b \leftrightarrow (a,b)$$

Denne sammenhengen kan vi trekke videre og betrakte indreproduktet på $\mathbb{R}^2$ gitt ved

$$\langle \mathbf{x}, \mathbf{y} \rangle_A = \mathbf{x}^T A \mathbf{y}$$

hvor A er matrisen

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

Dette gir oss det samme indreproduktet som i eksemplet ovenfor. Vi kan overbevise oss om det ved å regne ut indreproduktet

$$\begin{aligned} \langle (a_1,b_1),(a_2,b_2) \rangle_A &= (a_1\, b_1) \begin{pmatrix} \frac{1}{3} & \frac{1}{2} \\ \frac{1}{2} & 1 \end{pmatrix} \begin{pmatrix} a_2 \\ b_2 \end{pmatrix} \\ &= \frac{1}{3}a_1a_2 + \frac{1}{2}(a_1b_2 + a_2b_1) + b_1b_2 \end{aligned}$$

som er identisk likt det vi fant i eksemplet over.

Eksempel 7.2.3. *Vi holder oss til eksemplet gitt over, hvor $V = \mathscr{P}_1$ og indreproduktet gitt ved integralet over intervallet $[0,1]$. Polynomet $f(x) = 1$ har norm*

$$||1||^2 = \int_0^1 1^2\, dx = 1$$

og indreproduktet av 1 med et vilkårlig polynom $g(x) = ax+b$ er gitt ved

$$\langle 1, ax+b \rangle_{\int} = \int_0^1 ax+b\, dx = \frac{1}{2}a+b$$

Det betyr at hvis vi velger a og b slik at $\frac{1}{2}a+b=0$, så vil dette indreproduktet være 0. Vi har altså at polynomene $f(x)=1$ og $g(x)=2ax-a$ står normalt på hverandre. Vi kan også regne ut normen til $g(x)$;

$$\begin{aligned} ||2ax-a||^2 &= \int_0^1 (2ax-a)^2\, dx \\ &= \int_0^1 4a^2x^2 - 4a^2x + a^2\, dx \\ &= \frac{4}{3}a^2 - 2a^2 + a^2 = \frac{1}{3}a^2 \end{aligned}$$

Dersom vi setter $a = \sqrt{3}$ får vi $||2\sqrt{3}x - \sqrt{3}|| = 1$. Det betyr at $\{1, 2\sqrt{3}x - \sqrt{3}\}$ danner en ortonormal basis for $\mathscr{P}_1$ med dette indreproduktet.

7.3 Dekomponering av vektorer ved indreprodukt

Vi kan utnytte ortogonalitetsegenskapene til indreproduktet til å **dekomponere** vektorer relativt til en ortonormal basis. La $\mathbf{x} \in V$ være et element i et vektorrom med et indreprodukt $\langle\, , \rangle$ og la $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m\}$ være en ortonormal basis for dette vektorrommet. Da kan vi skrive $\mathbf{x}$ som en lineærkombinasjon av basisvektorene;

$$\mathbf{x} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_m\mathbf{v}_m$$

Siden basisen er ortonormal får vi

$$\langle \mathbf{x}, \mathbf{v}_j \rangle = a_j$$

og derfor

$$\mathbf{x} = \langle \mathbf{x}, \mathbf{v}_1 \rangle \mathbf{v}_1 + \langle \mathbf{x}, \mathbf{v}_2 \rangle \mathbf{v}_2 + \cdots + \langle \mathbf{x}, \mathbf{v}_m \rangle \mathbf{v}_m$$

Dette kalles en dekomponering av vektoren $\mathbf{x}$ og leddene $\langle \mathbf{x}, \mathbf{v}_j \rangle \mathbf{v}_j$ kalles komponentene til $\mathbf{x}$. Vi kan gjøre den samme konstruksjonen i tilfellet av at V er uendelig-dimensjonalt. Da kan vi finne komponentene på samme måte, men i det tilfellet vil summen over bli en uendelig rekke og vi kan ikke lenger være sikre på at et element i vektorrommet kan skrives som en sum av sine komponenter. Vi har et konvergensproblem. Vi skal komme litt tilbake til dette etter hvert.

Det er ikke sikkert at vi alltid har en ortonormal basis tilgjengelig, men vi kan fortsatt dekomponere en vektor.

Definisjon 7.3.1. *Gitt to vektorer* $\mathbf{v}$ *og* $\mathbf{w}$*. Vi definerer* ***ortogonalprojeksjonen*** *av* $\mathbf{v}$ *på* $\mathbf{w}$ *ved*

$$pr_{\mathbf{w}}(\mathbf{v}) = \frac{\mathbf{v} \cdot \mathbf{w}}{|\mathbf{w}|^2} \mathbf{w}$$

Eksempel 7.3.1. *Ortogonalprojeksjonen av* $\mathbf{v} = (2,1,0)$ *på* $\mathbf{w} = (1,1,1)$ *er gitt ved*

$$pr_{\mathbf{w}}(\mathbf{v}) = \frac{(2,1,0) \cdot (1,1,1)}{3} (1,1,1) = (1,1,1)$$

Eksempel 7.3.2. *La* M_2 *være vektorrommet av* 2×2*-matriser med indreproduktet gitt ved*

$$\langle A, B \rangle = \operatorname{tr}(A^T B)$$

Vi skal finne en ortonormal basis for dette vektorrommet. Vi tar utgangspunkt i identitetsmatrisen. Vi har

$$\langle I, I \rangle = \operatorname{tr}(I^T I) = \operatorname{tr}(I) = 2$$

For å få en basisvektor med norm 1, må vi derfor dele I med $\sqrt{2}$*. Det gir oss den første basisvektoren*

$$\mathbf{m}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Vi går videre og ser på

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

Denne står normalt på $\mathbf{m}_1$ *fordi vi har*

$$\langle \mathbf{m}_1, A \rangle = \operatorname{tr}\left(\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}^T \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right) = 0$$

og den har norm

$$\langle A, A \rangle^{\frac{1}{2}} = \operatorname{tr}\left(\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}^T \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right)^{\frac{1}{2}} = \sqrt{2}$$

Vi setter

$$\mathbf{m}_2 = \frac{1}{\sqrt{2}} A = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Da er vi ferdig med de diagonale matrisene, og vi tar for oss

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

Vi regner ut

$$\langle \mathbf{m}_1, B \rangle = \langle \mathbf{m}_2, B \rangle = 0$$

så ortogonaliteten er i orden, og

$$\langle B, B \rangle = \operatorname{tr}\left(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^T \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right) = \operatorname{tr}\left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right) = 2$$

Vi setter

$$\mathbf{m}_3 = \frac{1}{\sqrt{2}} B = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Til slutt lar vi

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

Vi regner ut

$$\langle \mathbf{m}_1, C \rangle = \langle \mathbf{m}_2, C \rangle = \langle \mathbf{m}_3, C \rangle = 0$$

og

$$\begin{aligned} \langle C, C \rangle &= \operatorname{tr}\left(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}^T \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \right) \\ &= \operatorname{tr}\left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right) = 2 \end{aligned}$$

så vi setter

$$\mathbf{m}_4 = \frac{1}{\sqrt{2}} C = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

Dermed har vi en ortonormal basis gitt ved

$$\{\mathbf{m}_1, \mathbf{m}_2, \mathbf{m}_3, \mathbf{m}_4\}$$

Hvis vi har gitt en vilkårlig matrise

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

så får vi

$$\begin{aligned} \langle \mathbf{m}_1, M \rangle &= \frac{1}{\sqrt{2}} \operatorname{tr}\left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \right) \\ &= \frac{1}{\sqrt{2}} (a + d) \end{aligned}$$

Vi regner ut tilsvarende for de andre og finner

$$\langle \mathbf{m}_2, M \rangle = \frac{1}{\sqrt{2}}(a-d)$$
$$\langle \mathbf{m}_3, M \rangle = \frac{1}{\sqrt{2}}(b+c)$$
$$\langle \mathbf{m}_4, M \rangle = \frac{1}{\sqrt{2}}(b-c)$$

Dermed kan vi skrive

$$M = \frac{1}{\sqrt{2}}((a+d)\mathbf{m}_1 + (a-d)\mathbf{m}_2 + (b+c)\mathbf{m}_3 + (b-c)\mathbf{m}_4)$$

Vi har tidligere sett at dekomponering av vektorer er spesielt enkelt i forhold til en ortogonal basis. Vi skal se på noen eksempler på hvordan dette kan se ut i praksis. Først gir vi det en formell definisjon.

Definisjon 7.3.2. *La V være et vektorrom (gjerne uendelig-dimensjonalt), og la $\mathscr{B}$ være en ortogonal basis for V. Da kaller vi $(V, \mathscr{B})$ et* ***ortogonalt system****.*

Bemerkning 7.3.3. *Vi definerer ortogonale system, og ikke ortonormale. Forskjellen er ubetydelig siden det er enkelt å gjøre et ortogonalt system til et ortonormalt system bare ved å dele alle basiselementer med sin lengde. Imidlertid viser det seg av det er mest hensiktsmessig å ta utgangspunkt i et ortogonalt system, slik vi skal se på de følgende eksemplene.*

7.4 Fourier-rekker

I mange anvendelser innen naturvitenskapene dukker det opp såkalte periodiske fenomener, slik som bølger eller andre oscillerende dynamiske systemer. Polynomer er ikke periodiske funksjoner og er defor ikke nødvendigvis så gunstige til å tilnærme denne type fenomener.

De trigonometriske funksjonene $\sin(\omega x)$ og $\cos(\omega x)$ for en vilkårlig konstant ω er periodiske og derfor et nye bedre utgangspunkt for å tilnærme andre periodiske funksjoner. Med passende valg for koeffisienten ω kan vi oppnå en hvilket som helst bølgelengde og skalering av funksjonene gir oss kontroll over amplituden. Funksjonene

$$A\sin\omega x + B\cos\omega x$$

danner derfor en utmerket ramme for å beskrive ulike bølgefenomener.

Vi har følgende definisjon av en periodisk funksjon:

Definisjon 7.4.1. *En funksjon f definert på de reelle tallene kalles* ***periodisk*** *med periode T dersom for alle $t \in \mathbb{R}$;*

$$f(t+T) = f(t)$$

Det minste positive tallet T med denne egnskapen kalles ***perioden*** *til f.*

Merk at dersom T er perioden til en funksjon f, så vil også mT være en periode for funksjonen for alle heltall m. Dette kommer av at

$$f(t+2T) = f(t+T+T) = f(t+T) = f(t)$$

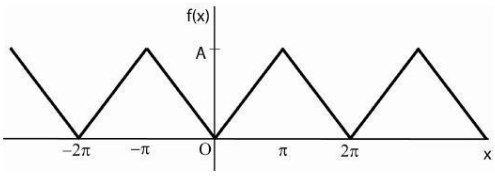
og tilsvarende argument dersom vi erstatter 2-tallet med et vilkårlig annet heltall.

I dette kapitlet kommer vi til å bruke t og x litt om hverandre som argumentet i de trigonometriske funksjonene. Grunnen til at mange periodiske fenomener beskriver størrelser som varierer i tid, og da er det naturlig å bruke variabelen t, mens i andre sammenhenger vil de periodiske fenomenene variere over en strekning. Da er det mer naturlig å bruke x.

Eksempel 7.4.1. *De trigonometriske funksjonene $f(x) = \cos(n\omega x)$ og $g(x) = \sin(n\omega x)$ for en vilkårlig konstant ω er begge periodiske med periode $\frac{2\pi}{\omega}$ for alle heltall n;*

$$\cos(n\omega(x+\tfrac{2\pi}{\omega})) = \cos(n\omega x + 2n\pi) = \cos\omega x$$

Et annet eksempel på en periodisk funksjon er sagtann-kurven (se Figur 7.1).



Figur 7.1. *Sagtann-kurven*

Et begrep vi kommer til å ha nytte av er symmetriegenskapene til en funksjon.

Definisjon 7.4.2. *En funksjon definert på de reelle tallene sies å være **jevn** dersom $f(-x) = f(x)$ for alle x, og **odde** derom $f(-x) = -f(x)$ for alle x.*

Vi har noen enkle regneregler knyttet til jevne og odde funksjoner:

a) Produktet av to jevne funksjoner er jevn.

b) Produktet av to odde funksjoner er jevn.

c) Produktet av en jevn og en odde funksjon er en odde funksjon.

d) Den deriverte av en jevn funksjon er odde.

e) Den deriverte av en odde funksjon er jevn.

f) For en jevn funksjon $g(x)$ har vi

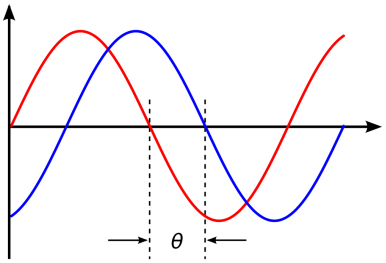
$$\int_{-a}^{a} g(x)\,dx = 2\int_{0}^{a} g(x)\,dx$$

f) For en odde funksjon $g(x)$ har vi

$$\int_{-a}^{a} g(x)\,dx = 0$$

Eksempel 7.4.2. *De trigonometriske funksjonene har motsatte symmetriegenskaper; sinus er odde og cosinus er jevn,*

Trigonometriske funksjoner kan være faseforskjøvet i forhold til hverandre (se Figur 7.2).



Figur 7.2. *Faseforskjøvede trigonometriske funksjoner.*

Faseforskyving uttrykkes ved et koordinatskifte;

$$f(t) = A\cos(\omega t - \phi_0)$$

der ϕ_0 er den såkalte fasevinkelen, og som illustrerer hvor mye cosinus-grafen er forskjøvet langs t-aksen. En interressant observasjon er at om vi legger sammen cosinus-funksjoner med samme periode, men med forskjellig fase, så vil vi fortsatt få en perfekt cosinus-funksjon. Dette prinsippet har en praktfull anvendelse i lydbølger. Lydbølger er periodiske funksjoner komponert av ulike cosinus-funksjoner. Fasefoskyvingen til en cosinus-bølge som sendes ut fra en lydkilde vil avhenge av hvor lang vei lyden går før den når mottageren. Noe av lyden går rett fram, mens andre deler av lyden reflekteres fra vegger etc. Dermed vil det være et mylder av fasefoskyvinger som når mottagerens øre. Imidlertid oppfatter mottageren dette som én lyd.

Det som skjer kan beskrives ved følgende formel:

$$A\cos(\omega t - \phi_0) = B\cos\omega t + C\sin\omega t$$

der $B = A\cos\phi_0$ og $C = A\sin\phi_0$. Dette følger av summeformelen for cosinus; Vi har

$$A\cos(\omega t - \phi_0) = A\cos\omega t\cos\phi_0 + A\sin\omega t\sin\phi_0$$

Formelen kan også anvendes den andre veien; vi har $A = \sqrt{B^2 + C^2}$ og $\phi_0 = \sin^{-1}\frac{C}{A} = \cos^{-1}\frac{B}{A}$. Dermed kan vi addere to ulike faseforskyvinger ved først å skrive dem som en lineær kombinasjon av cosinus og sinus, legge dem sammen, og så bruke formelen baklengs.

Trigonometriske funksjoner er prototypen på periodiske funksjoner. Tidlig på 1800-tallet foreslo en fransk matematiker, *Jean Baptiste Joseph Fourier* at det motsatte også var tilfelle. Det vil si at enhver periodisk funksjon (kontinuerlig og kontinuerlig deriverbar) kan skrives som en sum av (uendelig mange) trigonometriske funksjoner. Dette var en kontroversiell påstand, men ettertiden har vist at Fourier hadde rett.

La $f(t)$ være en periodisk funksjon (som vi i denne sammenhengen alltid vil kreve er kontinuerlig med kontinuerlig derivert) med periode T. La $\omega = \frac{2\pi}{T}$. Da har vi sett at

$$\cos(n\omega t) \quad \text{og} \quad \sin(n\omega t)$$

er periodiske funksjoner med periode T for alle valg av n. Fouriers teorem sier at det motsatte er tilfelle, at gitt $f(t)$ så kan vi på en entydig måte finne koeffisienter $a_i, b_i \in \mathbb{R}$ slik at

$$f(t) = \tfrac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos(n\omega t) + b_n \sin(n\omega t))$$

De to følgene $\{a_i\}_{i=0}^{\infty}$ og $\{b_i\}_{i=1}^{\infty}$ kalles funksjonens **Fourier-koeffisienter**, og rekka kalles funksjonens **Fourier-rekke**. Vi skal se at Fourier-koeffisientene er entydig gitt av funksjonen. Dette bygger på følgende integralformler:

$$\int_0^T \cos(n\omega t)\, dt = \begin{cases} 0 & \text{if } n \neq 0 \\ T & \text{if } n = 0 \end{cases}$$

$$\int_0^T \sin(n\omega t)\, dt = 0 \quad \forall n \geq 0$$

$$\int_0^T \cos(m\omega t)\cos(n\omega t)\, dt = \begin{cases} 0 & \text{if } n \neq m \\ \frac{T}{2} & \text{if } n = m \end{cases}$$

$$\int_0^T \sin(m\omega t)\sin(n\omega t)\, dt = \begin{cases} 0 & \text{if } n \neq m \\ \frac{T}{2} & \text{if } n = m \end{cases}$$

$$\int_0^T \cos(m\omega t)\sin(n\omega t)\, dt = 0 \quad \forall n \geq 0$$

Alle disse formelene kan bevises ved bruk av de generelle summeformelene for cosinus og sinus:

$$\begin{aligned} \cos x \cos y &= \tfrac{1}{2}(\cos(x-y) + \cos(x+y)) \\ \sin x \sin y &= \tfrac{1}{2}(\cos(x-y) - \cos(x+y)) \\ \sin x \cos y &= \tfrac{1}{2}(\sin(x-y) + \sin(x+y)) \end{aligned}$$

La nå

$$F = \tfrac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos(n\omega t) + b_n \sin(n\omega t))$$

Bruker vi integralformlene gitt over er det enkelt å se at

$$\begin{aligned} \int_0^T F \cos(m\omega t)\, dt &= \frac{1}{2} T a_m \\ \int_0^T F \sin(m\omega t)\, dt &= \frac{1}{2} T b_m \end{aligned}$$

Hvis vi tar utgangspunkt i funksjonen $f(t)$ og setter

$$\begin{aligned} a_n &= \tfrac{2}{T} \int_0^T f(t) \cos(n\omega t)\, dt \quad n = 0, 1, 2, \ldots \\ b_n &= \tfrac{2}{T} \int_0^T f(t) \sin(n\omega t)\, dt \quad n = 1, 2, \ldots \end{aligned}$$

så får vi Fourier-rekka til $f(t)$ gitt ved

$$F = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos(n\omega t) + b_n \sin(n\omega t))$$

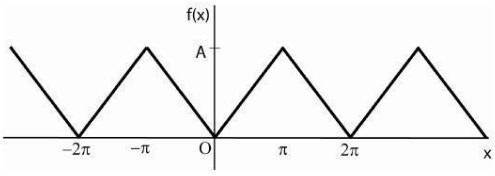
Dersom funksjonen f er periodisk, kontinuerlig og har kontinuerlig derivert, så vil denne rekka konvergere mot $f(t)$. I så fall skriver vi

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos(n\omega t) + b_n \sin(n\omega t))$$

Det vi har oppnådd med dette er å splitte opp en periodisk funksjon i enkle trigonometriske bestanddeler. Hvis vi går tilbake til lyd-anvendelsen, så kan vi si at denne oppskriften beskriver en måte å dekomponere en lyd i de ulike frekvenser som til sammen danner lyden.

Eksempel 7.4.3. *Vi skal finne Fourier-koeffisientene til sagtann-kurven gitt ved*

$$f(t) = \begin{cases} t & \text{når } 0 \le t \le \pi \\ 2\pi - t & \text{når } \pi \le t \le 2\pi \end{cases}$$



Her er $T = 2\pi$, og dermed $\omega = \frac{2\pi}{2\pi} = 1$. En enkel symmetribetraktning gir at

$$b_n = \frac{2}{2\pi} \int_0^{2\pi} f(t) \sin(nt)\, dt = 0 \quad n = 1, 2, \ldots$$

Vi beregner cosinus-koeffisientene:

$$a_n = \frac{2}{2\pi} \left(\int_0^{\pi} t \cos(nt)\, dt + \int_{\pi}^{2\pi} (2\pi - t) \cos(nt)\, dt \right)$$
$$= \begin{cases} \pi & \text{hvis } n = 0 \\ 0 & \text{hvis } n \neq 0 \text{ og jevn} \\ -\frac{4}{\pi n^2} & \text{hvis } n \text{ er odde} \end{cases}$$

Dette gir

$$f(t) = \frac{\pi}{2} - \sum_{k=1}^{\infty} \frac{4}{\pi(2k-1)^2} \cos((2k-1)t)$$

Eksempel 7.4.4. *Funksjonen f er gitt ved*

$$f(t) = \begin{cases} 0 & \text{når } (2k-1)\pi \le t \le 2k\pi \\ 1 & \text{når } 2k\pi < t \le (2k+1)\pi \end{cases}$$

Vi skal finne Fourier-koeffisientene til denne funksjonen. Perioden er $T = 2\pi$ og derfor $\omega = \frac{2\pi}{2\pi} = 1$. Det gir

$$\begin{aligned} a_0 &= \tfrac{1}{\pi}\int_0^{2\pi} f(t)\cos(0t)\,dt \\ &= \tfrac{1}{\pi}\int_\pi^{2\pi} 0\cos 0\,dt + \tfrac{1}{\pi}\int_0^{\pi} 1\cos 0\,dt = 1 \end{aligned}$$

og for $n > 0$;

$$\begin{aligned} a_n &= \tfrac{1}{\pi}\int_0^{2\pi} f(t)\cos(nt)\,dt \\ &= \tfrac{1}{\pi}\int_\pi^{2\pi} 0\cos nt\,dt + \tfrac{1}{\pi}\int_0^{\pi} 1\cos nt\,dt \\ &= \tfrac{1}{\pi}\big[\frac{\sin nt}{n}\big]_0^{\pi} = 0 \end{aligned}$$

Videre får vi for b_n-koeffisientene:

$$\begin{aligned} b_n &= \tfrac{1}{\pi}\int_0^{2\pi} f(t)\sin(nt)\,dt \\ &= \tfrac{1}{\pi}\int_\pi^{2\pi} 0\sin nt\,dt + \tfrac{1}{\pi}\int_0^{\pi} 1\sin nt\,dt \\ &= \tfrac{1}{\pi}\big[\frac{-\cos nt}{n}\big]_0^{\pi} = \frac{-(-1)^n}{n\pi} + \frac{1}{n\pi} \end{aligned}$$

Dette gir Fourier-rekke

$$\begin{aligned} f(t) &= \frac{1}{2} + \sum_{m=0}^{\infty} \frac{2}{(2m+1)\pi}\sin((2m+1)t) \\ &= \frac{1}{2} + \frac{2}{\pi}\big(\sin(t) + \frac{1}{3}\sin(3t) + \frac{1}{5}\sin(5t) + \ldots\big) \end{aligned}$$

Hvis vi tar med 10 ledd og setter inn forskjellige verdier for t, så ser vi at denne funksjonen faktisk gir en god tilnærming av den opprinnelige funksjonen. Her er noen verdier

$$\begin{aligned} f(1) &\approx 0,99177 \\ f(2.5) &\approx 0,97416 \\ f(4) &\approx -0,00319 \end{aligned}$$

Vi kan utnytte informasjonen i eksemplet til å beregne summen av en spesiell rekke. Hvis vi setter $t = \frac{\pi}{2}$ kjenner vi funksjonsverdien $f(\frac{\pi}{2}) = 1$. På den annen side har vi at

$$\sin((2m+1)\frac{\pi}{2}) = \begin{cases} 1 & m \text{ jevn} \\ -1 & m \text{ odde} \end{cases}$$

Det betyr at vi har likheten

$$1 = \frac{1}{2} + \frac{2}{\pi}\big(1 - \frac{1}{3} + \frac{1}{5} - \ldots\big)$$

eller

$$1-\frac{1}{3}+\frac{1}{5}-\frac{1}{7}+\cdots=\frac{\pi}{4}$$

Denne rekken ble for øvrig først oppdaget av James Gregory i 1668, og kalles derfor for Gregory-rekken.

Fourier-rekker passer inn i vår formelle setting for ortogonale systemer. Vi tenker på mengden av alle periodiske funksjoner med samme periode som vårt vektorrom V. Videre lar vi basisen $\mathscr{B}$ bestå av alle funksjoner

$$\cos(n\omega t) \quad \text{og} \quad \sin(n\omega t)$$

for alle valg av n. Siden vi har

$$\int_0^T \cos(m\omega t)\cos(n\omega t)\,dt = \begin{cases} 0 & \text{if } n \neq m \\ \frac{T}{2} & \text{if } n = m \end{cases}$$

$$\int_0^T \sin(m\omega t)\sin(n\omega t)\,dt = \begin{cases} 0 & \text{if } n \neq m \\ \frac{T}{2} & \text{if } n = m \end{cases}$$

$$\int_0^T \cos(m\omega t)\sin(n\omega t)\,dt = 0 \quad \forall n \geq 0$$

vil denne basisen være ortogonal.Dermed er det enkelt å beregne Fourier-koeffisientene til en periodisk funksjon ved å dekomponere funksjonen (oppfattet som en vektor i vektorrommet av periodiske funksjoner) med hensyn på basisen $\mathscr{B}$.

Et annet eksempel på et ortogonalt system er det som kalles wavelets og som bl.a. brukes ved digitalisering av bilder. Wavelets er små bølgepakker som egner seg til å beskrive korte lydstøt eller kontraster i et bilde. På 1980-tallet fant man en måte å definere disse waveletene som elementer i et vektorrom, og enda bedre, de dannet en ortogonal basis. Dermed passet hele wavelet-teorien inn i det formelle ortogonalt system-konseptet. Ved å beskrive et bilde gjennom alle dets endringer, som skarpe kanter eller fargeendringer, gjorde ortogonaliteten av waveletene det mulig å på en hensiktsmessig måte å spiltte opp informasjonen i bildet i enkle bestanddeler, som så ved en senere anledning kunne settes sammen og gjenskape bildet mer eller mindre nøyaktig. Effektiviteten i denne prosessen baserte seg helt og holdent på at man har tilgjengelig et ortogonalt system.

7.5 Oppgaver med løsning

Eksempel 7.5.1. *La M_2 være vektorrommet av 2×2-matriser utstyrt med indreproduktet $\langle A,B\rangle = \mathrm{tr}\,(A^TB)$. La*

$$A=\begin{pmatrix}2 & 1\\ -1 & 0\end{pmatrix},\quad B=\begin{pmatrix}1 & -1\\ 1 & 4\end{pmatrix},\quad C=\begin{pmatrix}3 & 0\\ 0 & 4\end{pmatrix}$$

a) Regn ut normen til A, B og C.

b) Vis at A står normalt på B.

c) Verifiser at Pythagoras setning gjelder for triplet A,B,C.

Løsning.

1. **Metodevalg:** Definisjonen av normen til en matrise A er $||A|| = \langle A,A\rangle^{\frac{1}{2}}$, og to elementer i vektorrommet står normalt på hverandre dersom indreproduktet $\langle A,B\rangle = 0$. Pythagoras setning sier at dersom elementene A og B er ortogonale, så gjelder likheten

$$||A||^2+||B||^2 = ||A+B||^2$$

Det betyr at vi må vise at $\langle A,B\rangle = 0$ og at $A+B=C$.

2. **Regning:** Vi har

$$\begin{aligned}||A||^2 &= \mathrm{tr}\,(A^TA) = \mathrm{tr}\left(\begin{pmatrix}2 & -1\\ 1 & 0\end{pmatrix}\begin{pmatrix}2 & 1\\ -1 & 0\end{pmatrix}\right)\\ &= \mathrm{tr}\begin{pmatrix}5 & 2\\ 2 & 1\end{pmatrix} = 6\end{aligned}$$

$$\begin{aligned}||B||^2 &= \mathrm{tr}\,(B^TB) = \mathrm{tr}\left(\begin{pmatrix}1 & 1\\ -1 & 4\end{pmatrix}\begin{pmatrix}1 & -1\\ 1 & 4\end{pmatrix}\right)\\ &= \mathrm{tr}\begin{pmatrix}2 & 3\\ 3 & 17\end{pmatrix} = 19\end{aligned}$$

$$\begin{aligned}||C||^2 &= \mathrm{tr}\,(C^TC) = \mathrm{tr}\left(\begin{pmatrix}3 & 0\\ 0 & 4\end{pmatrix}\begin{pmatrix}3 & 0\\ 0 & 4\end{pmatrix}\right)\\ &= \mathrm{tr}\begin{pmatrix}9 & 0\\ 0 & 16\end{pmatrix} = 25\end{aligned}$$

Det betyr at

$$||A||^+||B||^2 = ||C||^2$$

Videre har vi

$$\begin{aligned}\langle A,B\rangle &= \mathrm{tr}\,(A^TB) = \mathrm{tr}\left(\begin{pmatrix}2 & -1\\ 1 & 0\end{pmatrix}\begin{pmatrix}1 & -1\\ 1 & 4\end{pmatrix}\right)\\ &= \mathrm{tr}\begin{pmatrix}1 & -6\\ 1 & -1\end{pmatrix} = 0\end{aligned}$$

så A og B er ortogonale. Til slutt har vi at

$$A+B=\begin{pmatrix}2 & 1\\ -1 & 0\end{pmatrix}+\begin{pmatrix}1 & -1\\ 1 & 4\end{pmatrix}=\begin{pmatrix}3 & 0\\ 0 & 4\end{pmatrix}=C$$

Eksempel 7.5.2. *Et indreproduktet på $\mathscr{P}_2$ er gitt ved*

$$\langle f(x),g(x)\rangle = f(-1)g(-1)+f(0)g(0)+f(1)g(1)$$

La

$$f(x)=-1+x,\quad g(x)=2+x-2x^2$$

a) Regn ut normen til $f(x)$ og $g(x)$.

b) Vis at $f(x)$ og $g(x)$ er ortogonale polynomer med dette indreproduktet.

c) La R være "parallellogrammet"utspent av $f(x)$ og $g(x)$. Verifiser parallellogramloven

$$||f(x)+g(x)||^2+||f(x)-g(x)||^2=2||f(x)||^2+2||g(x)||^2$$

i dette tilfellet.

Løsning.

1. **Metodevalg:** Vi må regne ut $||f(x)||$, $||g(x)||$, $||f(x)+g(x)||$, $||f(x)-g(x)||$ og $\langle f(x),g(x)\rangle$ og vise at alle formelene stemmer.

2. **Regning:**

$$\begin{aligned}||f(x)|| &= (f(-1)^2+f(0)^2+f(1)^2)^{\frac{1}{2}}\\ &= ((-2)^2+(-1)^2+0^2)^{\frac{1}{2}}=\sqrt{5}\end{aligned}$$

$$\begin{aligned}||g(x)|| &= (g(-1)^2+g(0)^2+g(1)^2)^{\frac{1}{2}}\\ &= ((-1)^2+(2)^2+1^2)^{\frac{1}{2}}=\sqrt{6}\end{aligned}$$

$$\begin{aligned}||f(x)+g(x)|| &= ((-2-1)^2+(-1+2)^2+(0+1)^2)^{\frac{1}{2}}\\ &= \sqrt{11}\end{aligned}$$

$$\begin{aligned}||f(x)-g(x)|| &= ((-2+1)^2+(-1-2)^2+(0-1)^2)^{\frac{1}{2}}\\ &= \sqrt{11}\end{aligned}$$

og

$$\begin{aligned}\langle f(x),g(x)\rangle &= f(-1)g(-1)+f(0)g(0)+f(1)g(1)\\ &= (-2)(-1)+(-1)\cdot 2+0\cdot 1=0\end{aligned}$$

Vi observerer at siden $f(x)$ og $g(x)$ står normalt på hverandre er de to "diagonalene" like lange. Parallellogram-loven uttrykker i dette tilfellet

$$\sqrt{11}^2 + \sqrt{11}^2 = 2\sqrt{5}^2 + 2\sqrt{6}^2$$

eller

$$11 + 11 = 2 \cdot 5 + 2 \cdot 6$$

Kapittel 8

Oppgavesett

Dette kapitlet består av seks oppgavesett med oppgaver fra alle deler av kompendiet.

8.1 Oppgavesett

8.1.1 Oppgavesett 1

Oppgave 1. *Et dynamisk system er gitt ved*

$$\mathbf{x}_{n+1} = M \cdot \mathbf{x}_n$$

der M er 3×3-matrisen

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

og $\mathbf{x}_i$ er en 3-søylevektor.

a) *Gitt en tilstandsvektor*

$$\mathbf{x}_0 = \begin{pmatrix} 300 \\ 0 \\ 0 \end{pmatrix}$$

Hva blir $\mathbf{x}_1$?

b) *Regn ut det karakteristiske polynomet til M og finn egenverdiene.*

c) *Finn også de tilhørende egenvektorene.*

d) *Finn en likevektstilstand $\mathbf{u}$ slik at*

$$\mathbf{u} = M \cdot \mathbf{u}$$

og der $u_1 + u_2 + u_3 = 300$.

e) *Skriv tilstandsvektoren $\mathbf{x}_0$ (fra oppgave a)) som en lineær kombinasjon av egenvektorer og vis at*

$$\lim_{n \to \infty} \mathbf{x}_n = \mathbf{u}$$

Oppgave 2. *a) Vi har gitt to refleksjoner i $\mathbb{R}^3$. Den ene tar $\mathbf{v} = (1,1,0)$ på $-\mathbf{v}$, den andre tar $\mathbf{w} = (0,0,1)$ på $-\mathbf{w}$. Finn skjæringslinja mellom de to refleksjonsplanene.*

b) La A være matrisen gitt ved

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

Finn fikspunktene og determinanten til matrisen. Bruk dette til å bestemme hva slags isometri den representerer.

c) Hva er ordenen til isometrien i oppgave b)?

Oppgave 3. *La $\mathscr{P}_2$ være vektorrommet av polynomer av grad mindre enn eller lik 2 utstyrt med indreproduktet*

$$\langle f(x), g(x) \rangle = f(-1)g(-1) + f(0)g(0) + f(1)g(1)$$

a) Vis at polynomene $\frac{1}{\sqrt{2}}x$ og $x^2 - 1$ begge har lengde 1 og at de står normalt på hverandre.

b) Finn en ortonormal basis for polynomer i $\mathscr{P}_2$ relativt til dette indreproduktet.

Oppgave 4. *La D_6 være den dihedrale gruppa*

$$D_6 = \langle S, R \,|\, S^2 = R^6 = Id, SR^5 = RS \rangle$$

Vi definerer en avbildning $\phi : D_6 \to GL(\mathbb{R}^2)$ ved

$$\rho = \phi(R) = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix}$$

og

$$\mu = \phi(S) = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

a) Vis at ρ har orden 6 og μ har orden 2.

b) Vis at vi har relasjonen

$$\mu\rho^5 = \rho\mu$$

og konkluder med at ϕ er en gruppehomomorfi, og dermed en representasjon av D_6 på $\mathbb{R}^2$.

8.1.2 Oppgavesett 2

Oppgave 1. *a) Finn fikspunktene til avbildningen*

$$T(\mathbf{x}) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \mathbf{x} + \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

b) Bruk a) til å avgjøre hva slags isometri T er.

c) Vis enten ved regning eller på annen måte at T har orden 8, det vil si at $T^8 = Id$.

Oppgave 2. *En dynamisk modell er gitt ved*

$$\mathbf{x}_{n+1} = M \cdot \mathbf{x}_n$$

der M er en 2×2-matrise gitt ved

$$M = \begin{pmatrix} 0.5 & 0.1 \\ 0.5 & 0.9 \end{pmatrix}$$

og $\mathbf{x}_i$ er en 2-søylevektor.

a) Regn ut det karakteristiske polynomet til M og finn egenverdier og egenvektorer.

b) Gitt en tilstandsfordeling

$$\mathbf{x} = \begin{pmatrix} 150 \\ 150 \end{pmatrix}$$

Skriv $\mathbf{x}$ som en lineær kombinasjon av egenvektorer.

c) Finn grensen

$$\lim_{n \to \infty} M^n(\mathbf{x})$$

Oppgave 3. *Vi bruker indreproduktet $\langle A, B \rangle = \operatorname{tr}(A^T B)$ på vektorrommet av 2×2-matriser.*

a) Vis at

$$A = \begin{pmatrix} 1 & -2 \\ 1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$$

oppfyller Pythagoras setning;

$$||A||^2 + ||B||^2 = ||A - B||^2$$

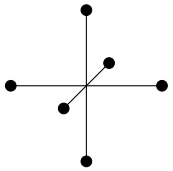
b) Forklar hvorfor $\operatorname{tr}(M^T) = \operatorname{tr}(M)$ for en vilkårlig kvadratisk matrise M.

c) Bruk b) til å vise at

$$\langle M, N \rangle = \langle N, M \rangle$$

for to vilkårlige kvadratiske matriser M og N.

Oppgave 4. *Beskriv symmetriene til den romlige figuren*



hvor de seks endepunktene har koordinater $(\pm 1, 0, 0)$, $(0 \pm 1, 0)$ og $(0, 0, \pm 1)$.

a) La ρ_x være rotasjon mot klokka om x-aksen og ρ_y rotasjon mot klokka om y-aksen. Forklar hvorfor ρ_x og ρ_y utgjør symmetrier av figuren over.

b) La ρ_z være rotasjon mot klokka om z-aksen. Skriv ρ_z som et produkt av ρ_x og ρ_y.

c) Kan du finne en ikke-triviell relasjon mellom symmetriene ρ_x og ρ_y?

8.1.3 Oppgavesett 3

Oppgave 1. *Vi skal sette sammen to stive bevegelser. Den ene, T_1, er en rotasjon med en vinkel $\frac{\pi}{2}$ om punktet $(1,1)$, og den andre, T_2, er refleksjon om linja $y = -x$.*

a) Skriv opp formler på formen $T_i(\mathbf{v}) = A_i\mathbf{v} + \mathbf{b}_i$ for de to avbildningene.

b) Regn ut komposisjonen $T_1 \circ T_2$, og finn eventuelle fikspunkter.

c) Regn også ut komposisjonen $T_2 \circ T_1$. Har denne avbildningen noen fikspunkter?

d) Hva slags avbildning er $T_1 \circ T_2 + T_2 \circ T_1$?

Oppgave 2. *a) La A være en kvadratisk matrise. Vis at produktet $B = A^TA$ er en symmetrisk matrise.*

b) La $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ og A en invertibel $n \times n$-matrise. Vis at

$$\langle \mathbf{u}, \mathbf{v} \rangle = A\mathbf{u} \cdot A\mathbf{v}$$

definerer et indreprodukt ved spesielt å vise at $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$ med likhet hvis og bare hvis $\mathbf{u} = \mathbf{0}$.

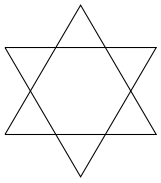
c) Vi kan skrive prikkproduktet

$$\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T\mathbf{v}$$

hvor vi oppfatter vektorene som matriser. La $B = A^TA$ være som over og la $\mathbf{v}$ være en egenvektor for B med egenverdi λ. Vis at $\lambda > 0$ ved å regne ut

$$\lambda \langle \mathbf{v}, \mathbf{v} \rangle$$

Oppgave 3. *a) Beskriv symmetriene til figuren*



b) Finn generatorer og relasjoner til symmetrigruppa til figuren.

c) Finn også konjugasjonsklassene til symmetrigruppa, og avgjør hvor mange elementer hver klasse inneholder.

d) Symmetrigruppa har 4 1-dimensjonale representasjoner, det vil si homomorfier inn i $\mathbb{R}^*$.Beskriv disse 4 ved å gi verdien på hver generator.

Oppgave 4. *La A være matrisen gitt ved*

$$A = \frac{1}{7}\begin{pmatrix} 6 & 3 & 2 \\ 3 & -2 & -6 \\ -2 & 6 & -3 \end{pmatrix}$$

a) Forklar hvorfor A er en ortogonal matrise.

b) Finn en egenvektor med egenverdi 1.

c) Finn fikspunktene til A og bestem hva slags isometri den representerer.

8.1.4 Oppgavesett 4

Oppgave 1. *a) Vi setter sammen en refleksjon om x-aksen med en refleksjon om y-aksen. Hva får vi? Hva får vi dersom vi bytter om på rekkefølgen?*

b) Hva får vi om vi setter sammen en refleksjon om den vertikale linja $x = a$ med en refleksjon om den horisontale linja $y = b$?

c) Vi kaller refleksjonene i b) for $S_{x=a}$ og $S_{y=b}$. La $T_{(a,b)}$ være translasjonen

$$T_{(a,b)}(\mathbf{v}) = \mathbf{v} + (a,b)$$

Vis at

$$S_{x=a} \circ S_{y=b} = T_{(2a,2b)} \circ S_{x=0} \circ S_{y=0}$$

Oppgave 2. *Vi lar*

$$C_n = \langle a \,|\, a^n = Id \rangle$$

betegne en gruppe generert av ett element av orden n. Dette er den sykliske n-gruppa.

a) Vis at ordenen til C_n er n.

b) Vis at gruppa C_n er abelsk.

c) La m være et helt tall som deler n, vi skriver $n = mk$. Vis at

$$\langle a^m \,|\, a^{mk} = Id \rangle$$

er en undergruppe av $\langle a \,|\, a^n = Id \rangle$ av orden k.

d) Et element $g \in G$ i en gruppe er en generator for gruppa dersom $G = \langle g \rangle$. Anta at a er en generator for den sykliske gruppa C_8. Avgjør hvilke potenser a^m som også genererer hele gruppa.

Oppgave 3. *La A være matrisen gitt ved*

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

og

$$\mathbf{u} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \ \mathbf{v} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \ \mathbf{w} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

a) Vis at

$$\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$$

danner en lineært uavhengig mengde.

b) Regn ut determinanten til A og bruk den til å beregne trippelproduktet

$$[A\mathbf{u}, A\mathbf{v}, A\mathbf{w}]$$

Oppgave 4. *La V være vektorrommet utspent av*

$$\mathscr{B} = \{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 2 \end{pmatrix} \}$$

a) Vis at mengden $\mathscr{B}$ er lineært uavhengig.

b) Forklar hvorfor $\mathscr{B}$ danner en basis for $\mathbb{R}^2$.

c) La

$$\mathscr{S} = \{\begin{pmatrix}1\\0\end{pmatrix}, \quad \begin{pmatrix}0\\1\end{pmatrix}\}$$

Forklar hvorfor $\mathscr{B}$ og $\mathscr{S}$ genererer samme vektorrom.

d) Skriv opp matrisen til en lineæravbildning som avbilder $\begin{pmatrix}1\\1\end{pmatrix}$ på $\begin{pmatrix}1\\0\end{pmatrix}$ og $\begin{pmatrix}1\\2\end{pmatrix}$ på $\begin{pmatrix}0\\1\end{pmatrix}$.

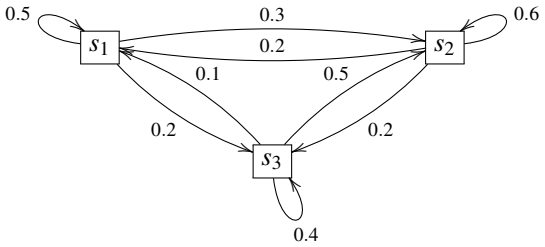
e) La

$$\begin{pmatrix}1\\1\end{pmatrix}, \quad \begin{pmatrix}1\\2\end{pmatrix}$$

være søylene i en 2×2-matrise. Hva er rangen til denne matrisen? Og dimensjonen til nullrommet?

8.1.5 Oppgavesett 5

Oppgave 1. *I denne oppgaven betrakter vi et dynamisk system illustrert ved boksdiagrammet*



a) Finn overgangsmatrisen M til systemet, og forklar hva vi mener med en stokastisk matrise. Vis at M er stokastisk.

b) Regn ut det karakteristiske polynomet til M. Siden M er stokastisk kjenner vi a priori den største egenverdien til M. Bruk det til å regne ut de andre egenverdiene.

c) Regn også ut egenvektorene til M. Danner egenvektorene en basis for $\mathbb{R}^3$? Gi en kort begrunnelse for svaret.

d) Vi tar utgangspunkt i en vektor $\mathbf{v}_0 = \begin{pmatrix}3\\-1\\2\end{pmatrix}$. Skriv denne vektoren som en lineærkombinasjon av egenvektorer.

e) Finn et uttrykk for $M^n\mathbf{v}_0$ for et vilkårlig positivt heltall n. Hva skjer når $n\to\infty$?

f) La $W \subset \mathbb{R}^3$ generert av vektorene

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad og \quad \mathbf{v}_2 = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$$

Hva skjer med vektorer i underrommet W når vi multipliserer dem med M mange ganger?

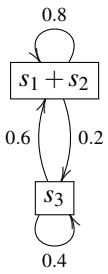
g) Hva er rangen til M?

h) Vi danner et nytt system ved å slå sammen tilstandene 1 og 2 i det opprinnelige systemet. Det betyr at vi erstatter en vektor $\mathbf{u} \in \mathbb{R}^3$ med en ny vektor i $\mathbb{R}^2$. Forklar hvorfor denne prosessen kan beskrives av en 2×3-matrise

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

Vi kaller denne prosessen for en projeksjon.

i) Det nye systemet kan illustreres med følgende boks-diagram



Vi lar M_1 være overgangsmatrisen for det nye systemet. Vis at overgangsmatrisene i dette tilfellet er kompatible med projeksjonen, i betydningen av at

$$M_1 \cdot P = P \cdot M$$

j) Regn ut det karakteristiske polynomet for M_1 og finn egenverdiene. Er det noen sammenheng mellom egenverdiene og egenvektorene for de to matrisene?

Generelt vil vi ikke overgangsmatrisene og projeksjonen være kompatible. Dersom verdiene på pilene fra s_1 og s_2 til s_3 hadde vært forskjellige, ville overgangs-koeffisientene til det projiserte systemet være avhengig av fordelingen mellom s_1 og s_2, og ikke bare summen av dem. I vårt tilfelle er overgangene fra s_1 og s_2 til s_3 like med verdi 0.2. Dermed vil fordelingen mellom klassene 1 og 2 ikke ha noen betydning for videre forløp når vi slår dem sammen.

Oppgave 2. *I denne oppgaven lar vi $\mathscr{P}_n$ bety vektorrommet av polynomer av grad mindre enn eller lik n.*

a) La $D : \mathscr{P}_3 \to \mathscr{P}_2$ være derivasjonsavbildningen gitt ved $D(f(x)) = f'(x)$. Forklar hvorfor vi lar avbildningen gå fra $\mathscr{P}_3$ til $\mathscr{P}_2$.

b) La $\mathscr{B}_n = \{1, x, x^2, \ldots, x^n\}$ være en basis for $\mathscr{P}_n$. Forklar hvorfor matrisen til D relativt til basisene $\mathscr{B}_3$ og $\mathscr{B}_2$ er gitt ved

$$[D] = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}$$

c) Finn rangen og dimensjonen til nullrommet til D.

Vi definerer integralavbildningen $I : \mathscr{P}_2 \to \mathscr{P}_3$ ved

$$I(f(x)) = \int_0^x f(t)\,dt$$

d) Forklar hvorfor matrisen til I relativt til basisene $\mathscr{B}_2$ og $\mathscr{B}_3$ er gitt ved

$$[I] = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{3} \end{pmatrix}$$

e) Vis at den sammensatte avbildningen $D \circ I$ er identiteten på $\mathscr{P}_2$.

f) Hva blir den sammensatte avbildningen $I \circ D : \mathscr{P}_3 \to \mathscr{P}_3$? Beskriv nullrommet til denne avbildningen?

Denne oppgaven gir en illustrasjon av fundamentalteoremet for differensial- og integralregningen. For enkelthet skyld forholder vi oss til polynomer av lav grad, men vi kunne gjerne ha betraktet deriverbare funksjoner på et intervall. Vi ville da også fått at $D \circ I$ er identiteten, mens $I \circ D$ har konstantfunksjonene som sitt nullrom, og forøvrig opptrer som identiteten.

8.1.6 Oppgavesett 6

Oppgave 1. *Indreproduktet på vektorrommet $M_{2\times 2}$ av 2×2-matriser er gitt ved*

$$\langle A, B \rangle = \operatorname{tr}(A^T B)$$

for $A, B \in M_{2\times 2}$. La

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

være en rotasjon med vinkel θ og

$$S = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

en refleksjon om x-aksen.

a) Regn ut normen til R_θ og S.

b) Vis at indreproduktet av to rotasjoner $R_{\theta_1} R_{\theta_2}$ er gitt ved

$$\langle R_{\theta_1}, R_{\theta_2} \rangle = 2\cos(\theta_1 - \theta_2)$$

(Hint: $\cos(\theta_1 - \theta_2) = \cos\theta_1 \cos\theta_2 + \sin\theta_1 \sin\theta_2$*)*

c) Vis at R_θ og S står normalt på hverandre med dette indreproduktet.

Vi kan finne et uttrykk for en vilkårlig refleksjon gjennom origo ved å se på produktet av S med en rotasjon R_θ. La $S_\theta = R_\theta S$.

d) Regn ut normen $||S_\theta||$ til S_θ.

e) Vis at

$$\langle S_{\theta_1}, S_{\theta_2} \rangle = \langle R_{\theta_1}, R_{\theta_2} \rangle$$

og at

$$\langle R_{\theta_1}, S_{\theta_2} \rangle = 0$$

Oppgave 2. *Vi skal se på symmetriene til en kube K med hjørner $(\pm 1, \pm 1, \pm 1)$. Kuben har 6 flater som alle er kvadrater. Det finnes derfor $24 = 6 \cdot 4$ symmetrier. Vi skal finne alle disse symmetriene. La ρ være en rotasjon med vinkel $\frac{\pi}{2}$ mot klokka og med rotasjonsakse $(0,0,1)$. La videre μ være en rotasjon med vinkel $\frac{\pi}{2}$ mot klokka og med rotasjonsakse $(0,1,0)$.*

a) Skriv ρ og μ på matriseform.

b) Hva er ordenen til ρ og μ?

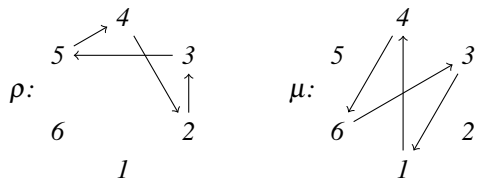
c) Produktene $\rho\mu$ og $\mu\rho$ er også rotasjoner. Finn rotasjonsaksene til de to rotasjonene. Finn også deres orden.

Symmetriene til K kan bestemmes skrittvis: Først bestemmer vi hvilken sideflate som ligger underst, deretter holder vi denne sideflaten på plass og roterer rundt normalvektoren til denne flaten. Hvilken sideflate som ligger i bunnen avgjøres gjennom å se på hvilken av enhetsvektorene $\pm\mathbf{e}_1$, $\pm\mathbf{e}_2$, $\pm\mathbf{e}_3$ som avbildes på $(0,0,-1)$.

Som et enkelt hjelpemiddel nummererer vi sideflatene som på en terning:

$$1 \longleftrightarrow (0,0,-1) \qquad 4 \longleftrightarrow (-1,0,0)$$
$$2 \longleftrightarrow (0,-1,0) \qquad 5 \longleftrightarrow (0,1,0)$$
$$3 \longleftrightarrow (1,0,0) \qquad 6 \longleftrightarrow (0,0,1)$$

Da kan vi illustrere ρ og μ ved en figur som indikerer hvordan avbildningene permuterer sideflatene:



d) Finn ut hvilken av enhetsvektorene som avbildes på $(0,0,-1)$ ved avbildningene ρ og μ.

e) Forklar hvordan man ved en kombinasjon av ρ og μ kan få en hvilken som helst av enhetsvektorene til å bli avbildet på $(0,0,-1)$.

f) Lag en tilsvarende figur som over for komposisjonen $\mu\rho$.

En konsekvens av e) er at vi kan skrive alle symmetrier som produkter av ρ-er og μ-er, men det finnes noen relasjoner mellom dem.

g) Vis ved å multiplisere ut matrisene at $\mu\rho^2 = \rho^2\mu^3$. Forklar også hvordan vi kan se dette ut fra figurene over.

8.2 Fasit til oppgavesett

Sett 1:

Oppgave 3. *a)* $\begin{pmatrix}180\\30\\90\end{pmatrix}$

b)

$$-\lambda^3+1.5\lambda^2-0.54\lambda+0.04= \\ -(\lambda-1)(\lambda-0.4)(\lambda-0.1)$$

c) $\begin{pmatrix}1\\1\\1\end{pmatrix}, \begin{pmatrix}1\\-1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\-1\end{pmatrix}$

d) $\mathbf{u}=\begin{pmatrix}100\\100\\100\end{pmatrix}$

e)

$$\begin{pmatrix}300\\0\\0\end{pmatrix}=\begin{pmatrix}100\\100\\100\end{pmatrix}+\begin{pmatrix}200\\-200\\0\end{pmatrix}+\begin{pmatrix}0\\100\\-100\end{pmatrix} \\ \mapsto \begin{pmatrix}100\\100\\100\end{pmatrix}$$

Oppgave 4. *a) Linja langs* $(1,-1,0)$ *b) Fikspunkt* $(0,0,0)$*, determinant =-1, refleksjon og rotasjon (inversjon) c) 4*

Oppgave 5. *b)* $\{\frac{1}{\sqrt{2}}x, x^2-1, \frac{1}{\sqrt{2}}x^2\}$

Oppgave 6. $\chi(\rho)=1$, $\chi(\mu)=0$

Sett 2:

Oppgave 1. *a)* $(\frac{\sqrt{2}+1}{2}, \frac{1}{2})$ *b) Rotasjon* $\frac{\pi}{4}$ *om* $\frac{1}{2}\begin{pmatrix}\sqrt{2}+1\\1\end{pmatrix}$

Oppgave 2. *a)* $(\lambda-1)(\lambda-0.4)$, $\mathbf{v}_1=(1,5)$ $\mathbf{v}_2=(1,-1)$
b) $\begin{pmatrix}150\\150\end{pmatrix}=\begin{pmatrix}50\\250\end{pmatrix}+\begin{pmatrix}100\\-100\end{pmatrix}$ *c)* $\begin{pmatrix}50\\250\end{pmatrix}$

Oppgave 3. *b) M og* M^T *har samme diagonalelementer*

Oppgave 4. *b)* $\rho_z=\rho_x^{-1}\rho_y^{-1}\rho_x$
c) $\rho_x^{-1}\rho_y^{-1}\rho_x=\rho_y^{-1}\rho_x\rho_y$

Sett 3:

Oppgave 1. *a)* $T_1(\mathbf{v}) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\mathbf{v} + \begin{pmatrix} 2 \\ 0 \end{pmatrix}$,
$T_2(\mathbf{v}) = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}\mathbf{v}$
b) $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\mathbf{v} + \begin{pmatrix} 2 \\ 0 \end{pmatrix}$, *ingen fikspunkter*
c) $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}\mathbf{v} + \begin{pmatrix} 0 \\ -2 \end{pmatrix}$, *ingen fikspunkter*
d) $()T_1T_2 + T_2T_1)(\mathbf{v}) = \begin{pmatrix} 2 \\ -2 \end{pmatrix}$, *fikspunkt* $\begin{pmatrix} 2 \\ -2 \end{pmatrix}$

Oppgave 2.

Oppgave 3. *Rotasjon av orden 6, 6 refleksjoner*
b) ρ, μ, $\rho^6 = \mu^2 = Id$, $\mu\rho^5 = \rho\mu$
c) $\{Id\}$, $\{\rho, \rho^5\}$, $\{\rho^2, \rho^4\}$, $\{\rho^3\}$, $\{\mu, \rho^2\mu, \rho^4\mu\}$, $\{\rho\mu, \rho^3\mu, \rho^5\mu\}$
d) $\rho, \mu \mapsto \pm 1$

Oppgave 4. *b)* $\begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix}$, *c) linja langs* $\begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix}$, *rotasjon*

Sett 4:

Oppgave 1. *a) Rotasjon* π, *motsatt* $-\pi$
b) Rotasjon π *om* (a, b)

Oppgave 2. *d) a,* a^7

Oppgave 3. *b) -4*

Oppgave 4. *d)* $\begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$ *e) rang 2, nullrom 0*

Sett 5:

Oppgave 1. *a)*

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

Alle elementene i en stokastisk matrise (kvadratisk) er ikke-negative, og kolonnesummene er 1.

b)

$$\begin{aligned} \chi_M(\lambda) &= \begin{vmatrix} 0.5-\lambda & 0.2 & 0.1 \\ 0.3 & 0.6-\lambda & 0.5 \\ 0.2 & 0.2 & 0.4-\lambda \end{vmatrix} \\ &= -\lambda^3 + 1.5\lambda^2 - 0.5\lambda + 0.06 \\ &= -(\lambda - 1)(\lambda^2 - 0.5\lambda + 0.06) \end{aligned}$$

Vi løser $\lambda^2 - 0.5\lambda + 0.06 = 0$ *ved abc-formelen;*

$$\lambda = \frac{0.5 \pm \sqrt{0.25 - 4 * 0.06}}{2} = 0.25 \pm 0.05$$

som gir $\lambda_2 = 0.2$ *og* $\lambda_3 = 0.3$.

c) For $\lambda_1 = 1$*;*

$$\begin{pmatrix} 0.5-1 & 0.2 & 0.1 \\ 0.3 & 0.6-1 & 0.5 \\ 0.2 & 0.2 & 0.4-1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

gir egenvektor

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$$

For $\lambda_2 = 0.2$*;*

$$\begin{pmatrix} 0.5-0.2 & 0.2 & 0.1 \\ 0.3 & 0.6-0.2 & 0.5 \\ 0.2 & 0.2 & 0.4-0.2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

gir egenvektor

$$\mathbf{v}_2 = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$$

For $\lambda_3 = 0.3$*;*

$$\begin{pmatrix} 0.5-0.3 & 0.2 & 0.1 \\ 0.3 & 0.6-0.3 & 0.5 \\ 0.2 & 0.2 & 0.4-0.3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

gir egenvektor

$$\mathbf{v}_3 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$$

Egenvektorene korresponderer til tre forskjellige egenverdier. Da vet vi at egenvektorene er lineært uavhengig. Det er tre av dem og de ligger i $\mathbb{R}^3$*. Derfor vil de danne en basis.*

d)

$$\mathbf{v}_0 = \begin{pmatrix} 3 \\ -1 \\ 2 \end{pmatrix} = a \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} + b \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} + c \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$$

gir $a+b+c = 3$*,* $2a-2b-c = -1$ *og* $a+b = 2$*. Det betyr at*

$$\mathbf{v}_0 = \begin{pmatrix} 3 \\ -1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$$

e)

$$\begin{aligned} M^n\mathbf{v}_0 &= M^n\begin{pmatrix}1\\2\\1\end{pmatrix} + M^n\begin{pmatrix}1\\-2\\1\end{pmatrix} + M^n 1 - 10 \\ &= 1^n\begin{pmatrix}1\\2\\1\end{pmatrix} + (0.2)^n\begin{pmatrix}1\\-2\\1\end{pmatrix} + (0.3)^n 1 - 10 \\ &= \begin{pmatrix}1+(0.2)^n+(0.3)^n\\2-(0.2)^n-(0.3)^n\\1+(0.2)^n\end{pmatrix} \to \begin{pmatrix}1\\2\\1\end{pmatrix} \end{aligned}$$

når $n \to \infty$.

f) Vi har

$$\mathbf{v}_2 = \begin{pmatrix}0\\-1\\1\end{pmatrix} = -\begin{pmatrix}1\\-1\\0\end{pmatrix} + \begin{pmatrix}1\\-2\\1\end{pmatrix}$$

Det betyr at W *består av vektorer av typen*

$$\mathbf{v} = a\mathbf{v}_1 + b\mathbf{v}_2 = (a-b)\begin{pmatrix}1\\-1\\0\end{pmatrix} + b\begin{pmatrix}1\\-2\\1\end{pmatrix}$$

og derfor

$$\begin{aligned} M^n\mathbf{v} &= (a-b)(0.2)^n\begin{pmatrix}1\\-1\\0\end{pmatrix} + b(0.3)^n\begin{pmatrix}1\\-2\\1\end{pmatrix} \\ &\to 0 \quad \textit{når } n \to \infty \end{aligned}$$

g) $\det M = 0.06 \neq 0$*, så maksimal rang, dvs. 3*

h) Matrisen uttrykker det å slå sammen de to første klassene;

$$\begin{pmatrix}1&1&0\\0&0&1\end{pmatrix}\begin{pmatrix}s_1\\s_2\\s_3\end{pmatrix} = \begin{pmatrix}s_1+s_2\\s_3\end{pmatrix}$$

i)

$$\begin{aligned} M_1 \cdot P &= \begin{pmatrix}0.8&0.6\\0.2&0.4\end{pmatrix}\begin{pmatrix}1&1&0\\0&0&1\end{pmatrix} \\ &= \begin{pmatrix}0.8&0.8&0.6\\0.2&0.2&0.4\end{pmatrix} \end{aligned}$$

$$\begin{aligned} P \cdot M &= \begin{pmatrix}1&1&0\\0&0&1\end{pmatrix} \cdot \begin{pmatrix}0.5&0.2&0.1\\0.3&0.6&0.5\\0.2&0.2&0.4\end{pmatrix} \\ &= \begin{pmatrix}0.8&0.8&0.6\\0.2&0.2&0.4\end{pmatrix} \end{aligned}$$

j)

$$\lambda^2 - 1.2\lambda + 0.2$$

med røtter $\lambda_1 = 1$ *og* $\lambda_2 = 0.2$*. Det er en sammenheng. De nye egenvektorene blir*

$$P\begin{pmatrix}1\\2\\1\end{pmatrix} = \begin{pmatrix}3\\1\end{pmatrix}$$
$$P\begin{pmatrix}1\\-2\\1\end{pmatrix} = \begin{pmatrix}-1\\1\end{pmatrix} P\begin{pmatrix}1\\-1\\0\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix}$$

hvor den tredje ikke er noen egenvektor. Egenverdiene følger med.

Oppgave 2. *a) Avbildningen går fra* $\mathscr{P}_3$ *til* $\mathscr{P}_2$ *fordi derivasjon reduserer graden til et polynom med 1.*

b)

$$D(1) = 0, \quad D(x) = 1, \quad D(x^2) = 2x, \quad D(x^3) = 3x^2$$

gir

$$[D] = \begin{pmatrix}(0)_{\mathscr{B}_2} & (1)_{\mathscr{B}_2} & (2x)_{\mathscr{B}_2} & (3x^2)_{\mathscr{B}_2}\end{pmatrix} = \begin{pmatrix}0 & 1 & 0 & 0\\ 0 & 0 & 2 & 0\\ 0 & 0 & 0 & 3\end{pmatrix}$$

c) Nullrommet er gitt ved konstantene, dvs. dimensjon 1. Det gir rang $= 4 - 1 = 3$*.*

d)

$$I(1) = x, \quad I(x) = \frac{1}{2}x^2, \quad I(x^2) = \frac{1}{3}x^3$$

gir

$$[I] = \begin{pmatrix}(1)_{\mathscr{B}_3} & (\frac{1}{2}x^2)_{\mathscr{B}_2} & (\frac{1}{3}x^3)_{\mathscr{B}_2}\end{pmatrix} = \begin{pmatrix}0 & 0 & 0\\ 1 & 0 & 0\\ 0 & \frac{1}{2} & 0\\ 0 & 0 & \frac{1}{3}\end{pmatrix}$$

e) Multipliser ut $[D] \cdot [I]$*.*

f)

$$[I] \cdot [D] = \begin{pmatrix}0 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}$$

Nullrommet er gitt ved

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

som svarer til de konstante polynomene.

Sett 6:

Oppgave 1. *a)*

$$\begin{aligned}\langle R_\theta, R_\theta\rangle^{\frac{1}{2}} &= \operatorname{tr}((R_\theta)^T R_\theta)^{\frac{1}{2}} \\ &= \operatorname{tr}\begin{pmatrix}\cos\theta & \sin\theta \\ -\sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}\cos\theta & -\sin\theta \\ \sin\theta & \cos\theta\end{pmatrix}^{\frac{1}{2}} \\ &= \operatorname{tr}\begin{pmatrix}1 & 0 \\ 0 & 1\end{pmatrix}^{\frac{1}{2}} = \sqrt{2}\end{aligned}$$

$$\begin{aligned}\langle S, S\rangle^{\frac{1}{2}} = \operatorname{tr}(S^T S)^{\frac{1}{2}} &= \operatorname{tr}\begin{pmatrix}1 & 0 \\ 0 & -1\end{pmatrix}\begin{pmatrix}1 & 0 \\ 0 & -1\end{pmatrix}^{\frac{1}{2}} \\ &= \operatorname{tr}\begin{pmatrix}1 & 0 \\ 0 & 1\end{pmatrix}^{\frac{1}{2}} = \sqrt{2}\end{aligned}$$

b)

$$\begin{aligned}\langle R_{\theta_1}, R_{\theta_2}\rangle &= \operatorname{tr}((R_{\theta_1})^T R_{\theta_2}) \\ &= \operatorname{tr}\begin{pmatrix}\cos\theta_1 & \sin\theta_1 \\ -\sin\theta_1 & \cos\theta_1\end{pmatrix}\begin{pmatrix}\cos\theta_2 & -\sin\theta_2 \\ \sin\theta_2 & \cos\theta_2\end{pmatrix} \\ &= 2(\cos\theta_1\cos\theta_2 + \sin\theta_1\sin\theta_2) \\ &= 2\cos(\theta_1 - \theta_2)\end{aligned}$$

c)

$$\begin{aligned}\langle R_\theta, S\rangle &= \operatorname{tr}((R_\theta)^T S) \\ &= \operatorname{tr}\begin{pmatrix}\cos\theta & \sin\theta \\ -\sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}1 & 0 \\ 0 & -1\end{pmatrix} \\ &= \cos\theta - \cos\theta = 0\end{aligned}$$

d)

$$\begin{aligned}||S_\theta||^2 = \langle S_\theta, S_\theta\rangle &= \operatorname{tr}((R_\theta S)^T R_\theta S) \\ &= \operatorname{tr}(S^T (R_\theta)^T R_\theta S) = \operatorname{tr}(S^T S) \\ &= ||S||^2 = 2\end{aligned}$$

e)

$$\begin{aligned}\langle S_{\theta_1}, S_{\theta_2}\rangle &= \operatorname{tr}(S^T (R_{\theta_1})^T R_{\theta_2} S) \\ &= \operatorname{tr}(SS^T (R_{\theta_1})^T R_{\theta_2}) \\ &= \operatorname{tr}((R_{\theta_1})^T R_{\theta_2}) = \langle R_{\theta_1}, R_{\theta_2}\rangle\end{aligned}$$

$$
\begin{aligned}
\langle R_{\theta_1}, S_{\theta_2} \rangle &= \operatorname{tr}\left((R_{\theta_1})^T R_{\theta_2} S\right) \\
&= \operatorname{tr}\begin{pmatrix} \cos\theta_1 & \sin\theta_1 \\ -\sin\theta_1 & \cos\theta_1 \end{pmatrix} \\
&\qquad \begin{pmatrix} \cos\theta_2 & -\sin\theta_2 \\ \sin\theta_2 & \cos\theta_2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \\
&= \operatorname{tr}\begin{pmatrix} \cos\theta_1 & \sin\theta_1 \\ -\sin\theta_1 & \cos\theta_1 \end{pmatrix} \begin{pmatrix} \cos\theta_2 & \sin\theta_2 \\ \sin\theta_2 & -\cos\theta_2 \end{pmatrix} \\
&= \cos\theta_1\cos\theta_2 + \sin\theta_1\sin\theta_2 \\
&\qquad - \sin\theta_1\sin\theta_2 - \cos\theta_1\cos\theta_2 \\
&= 0
\end{aligned}
$$

Oppgave 2. *a)*

$$
\rho = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}
$$

$$
\mu = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix}
$$

b) ρ *og* μ *har begge orden 4.*

c)

$$
\begin{aligned}
\rho\mu &= \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix} \\
&= \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & 1 \\ -1 & 0 & 0 \end{pmatrix}
\end{aligned}
$$

En egenvektor for egenverdien $\lambda = 1$ *er gitt ved* $\mathbf{v} = (1,-1,-1)$ *som også er rotasjonsaksen.*

$$
\begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & 1 \\ -1 & 0 & 0 \end{pmatrix}^3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}
$$

Det vil si orden 3.

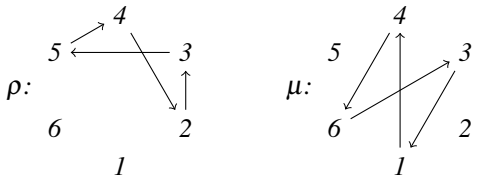
$$
\begin{aligned}
\mu\rho &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}
\end{aligned}
$$

Rotasjonsakse $(1,1,1)$ *og orden 3.*

Husk:

$$1 \longleftrightarrow (0,0,-1) \qquad 4 \longleftrightarrow (-1,0,0)$$
$$2 \longleftrightarrow (0,-1,0) \qquad 5 \longleftrightarrow (0,1,0)$$
$$3 \longleftrightarrow (1,0,0) \qquad 6 \longleftrightarrow (0,0,1)$$

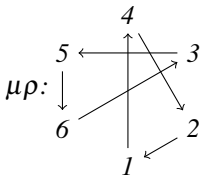
og



d) $\rho((0,0,-1)) = (0,0,-1)$ *og* $\mu((1,0,0)) = (0,0,-1)$

e) $(0,0,-1)$ *er nr. 1, og* μ *tar 3 på 1. 5 går til 3 ved* ρ^3 *og 2 går til 3 ved* ρ*. De andre går til 3 ved en passende potens av* μ*. Så alle kan komme til 3 og derfor til 1.*

f)



g)

$$\mu\rho^2 = \mu\rho\rho = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

$$\rho^2\mu^3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}^2 \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix}^3 = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

For $\mu\rho^2$ har vi

$$
\begin{aligned}
1 &\to 1 \to 1 \to 4 \\
2 &\to 3 \to 5 \to 5 \\
3 &\to 5 \to 4 \to 6 \\
4 &\to 2 \to 3 \to 1 \\
5 &\to 4 \to 2 \to 2 \\
6 &\to 6 \to 6 \to 3
\end{aligned}
$$

For $\rho^2\mu^3$ har vi

$$
\begin{aligned}
1 &\to 4 \to 6 \to 3 \to 5 \to 4 \\
2 &\to 2 \to 2 \to 2 \to 3 \to 5 \\
3 &\to 1 \to 4 \to 6 \to 6 \to 6 \\
4 &\to 6 \to 3 \to 1 \to 1 \to 1 \\
5 &\to 5 \to 5 \to 5 \to 4 \to 2 \\
6 &\to 3 \to 1 \to 4 \to 2 \to 3
\end{aligned}
$$

som gir samme sluttresultat.