

Lineær Algebra og Vektorrom

Eivind Eriksen

HØGSKOLEN I OSLO, AVDELING FOR INGENIØRUTDANNING

©Eivind Eriksen 2005.

Innhold

Kapittel 1. Lineære likningssystemer	1
1.1. Lineære likningssystemer i to variable	1
1.2. Lineære likningssystemer i tre variable	2
1.3. Generelle lineære likningssystemer	3
1.4. Systematisk løsning av lineære likningssystemer	4
1.5. Gauss eliminasjon	4
1.6. Gauss-Jordan eliminasjon	8
1.7. Konklusjoner	11
Kapittel 2. Matriser og matriseregning	13
2.1. Matriser og vektorer	13
2.2. Addisjon og skalarmultiplikasjon	14
2.3. Euklidske rom	14
2.4. Matrisemultiplikasjon	15
2.5. Lineære likningssystem på matriseform	16
2.6. Invertering av matriser	18
2.7. Determinanter	21
2.8. Bruk av determinanter	26
Kapittel 3. Vektorrom	29
3.1. Vektorrommet \mathbb{R}^3	29
3.2. Lineære underrom av \mathbb{R}^3	33
3.3. Generelle vektorrom	33
3.4. Vektorrommet \mathbb{R}^n	34
3.5. Lineære underrom av \mathbb{R}^n	37
Kapittel 4. Egenverdi-problemer	41
4.1. Lineære transformasjoner	41
4.2. Egenverdier og egenvektorer	43
4.3. Diagonalisering av matriser	46
4.4. Noen anvendelser av diagonalisering	49
4.5. Systemer av lineære differensiallikninger	53

KAPITTEL 1

Lineære likningssystemer

1.1. Lineære likningssystemer i to variable

La oss begynne med å se på et likningssystem som består av to likninger med to ukjente x og y , for eksempel dette systemet:

$$(1.1) \quad \begin{aligned} x + y &= 4 \\ x - y &= 2 \end{aligned}$$

Vi skal først se på hvordan dette likningssystemet kan løses algebraisk, altså ved regning. Vi løser første likning med hensyn på y , og får at $y = -x + 4$. Så setter vi dette inn i andre likning, og får at

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

Første likning gir nå $y = -x + 4 = -3 + 4 = 1$. Likningssystemet har altså en entydig løsning $x = 3$, $y = 1$. Løsningsmetoden som vi har brukt kalles *substitusjonsmetoden*.

Et annet alternativ er å bruke *elimineringsmetoden* for å løse det lineære likningssystemet (1.1). Da går vi fram på følgende måte: Vi multipliserer første likning med -1 , og får at $-x - y = -4$. Siden uttrykkene på høyre og venstre side er like, kan vi legge den nye likningen til andre likning i det opprinnelige systemet. Vi får da

$$\begin{aligned} x - y + (-x - y) &= 2 + (-4) \\ -2y &= -2 \\ y &= 1 \end{aligned}$$

Vitsen med dette er altså at x forsvinner eller blir eliminert fra andre likning. Til slutt ser vi at $y = 1$ gir $x + y = x + 1 = 4$ fra første likning, altså er $x = 3$.

Likningssystemet (1.1) er lineært, fordi hver likning kun inneholder ledd som er polynomer av grad en (en konstant multiplisert med en av variablene) eller grad null (en konstant). Løsningen av hver av de to likningene beskriver en rett linje i xy -planet, gitt ved $y = -x + 4$ for første likning og $y = x - 2$ for andre likning. Dette er grunnen til at likningene kalles *lineære likninger*. Vi ser at de to linjene ikke er parallelle, siden de har forskjellige stigningstall, og dermed skjærer de

hverandre i nøyaktig ett punkt. Dette skjæringspunktet er $(3, 1)$, og svarer til løsningen $x = 3$, $y = 1$ av systemet (1.1).

Ser vi på et vilkårlig lineært likningssystem med to likninger i to variable x og y , vil hver av likningene svare til en rett linje i xy -planet. Vi kan dermed se geometrisk, og helt uten å regne, at det kun er tre muligheter:

- (1) Det er to ikke-parallelle linjer.
- (2) Det er to forskjellige parallelle linjer.
- (3) Det er to like linjer.

I tilfelle (1) får vi nøyaktig ett skjæringspunkt, og dermed nøyaktig en løsning av likningssystemet. I tilfelle (2) får vi ingen skjæringspunkter, og dermed ingen løsning av likningssystemet. I tilfelle (3) får vi en hel linje av skjæringspunkter, og dermed uendelig mange løsninger av likningssystemet. Siden det ikke finnes andre muligheter enn disse tre, så er det for eksempel ikke mulig å ha akkurat to løsninger.

1.2. Lineære likningssystemer i tre variable

Vi har sett at for en lineær likning i to variable x, y , så danner løsningene av likningen en rett linje i xy -planet. Hvis vi istedet ser på en lineær likning i tre variable x, y, z , for eksempel likningen

$$2x + 3y - z = 7,$$

så danner alle løsningene av likningen et plan i xyz -koordinatsystemet, altså i rommet. Dette planet er gitt ved likningen $z = 2x + 3y - 7$. Igjen får vi en geometrisk figur som er helt rett eller lineær. Denne gang blir det et 2-dimensjonalt plan, mens det i tilfellet med to variable ble en 1-dimensjonal linje.

La oss se på et eksempel på et litt mer komplisert lineært likningssystem, som består av tre likninger i tre ukjente x, y, z :

$$(1.2) \quad \begin{array}{rclcl} 3x & - & 8y & + & 10z & = & 22 \\ x & - & 3y & + & 2z & = & 5 \\ 2x & - & 9y & - & 8z & = & -11 \end{array}$$

Igjen skal vi løse dette ved regning, men la oss først forsøke å tenke geometrisk. Hver av de tre likningene svarer til et plan i rommet, og løsningene av ligningssystemet svarer derfor til skjæringspunktene mellom alle tre planene. Hvilke muligheter har vi for skjæringspunkter i dette tilfellet?

La oss løse ligningssystemet algebraisk ved hjelp av substitusjon. Vi løser andre likning med hensyn på x , og får at $x = 3y - 2z + 5$. Vi kunne i utgangspunktet valgt å begynne med en annen likning eller en annen variabel, men velger denne fordi koeffisienten foran x i andre likning er 1 — det gir enklest regning. Vi setter så dette uttrykket inn

i første og tredje likning for å eliminere x fra disse likningene, og får

$$\begin{aligned} 3x - 8y + 10z &= 22 \\ 3(3y - 2z + 5) - 8y + 10z &= 22 \\ y + 4z &= 7 \end{aligned}$$

og

$$\begin{aligned} 2x - 9y - 8z &= -11 \\ 2(3y - 2z + 5) - 9y - 8z &= -11 \\ -3y - 12z &= -21 \end{aligned}$$

Vi løser så den første av disse likningene med hensyn på y , igjen fordi dette gir enklest regning, og får $y = -4z + 7$. Til slutt setter vi dette uttrykket inn i siste likning:

$$\begin{aligned} -3y - 12z &= -21 \\ -3(-4z + 7) - 12z &= -21 \\ -21 &= -21 \end{aligned}$$

Denne likningen gir ingen informasjon om z , den er alltid oppfylt. Det betyr at z kan velges fritt. Men så snart vi har valgt en verdi $z = t$ for denne variabelen, så er x og y bestemt av dette valget:

$$\begin{aligned} z &= t \\ y &= -4z + 7 = -4t + 7 \\ x &= 3y - 2z + 5 = 3(-4t + 7) - 2t + 5 = -14t + 26 \end{aligned}$$

Vi sier at z er en *fri parameter*, og det er vanlig å bruke symboler som s, t for slike frie parametre. Siden vi har en fri parameter, danner løsningene av likningssystemet (1.2) en linje, gitt ved parametriseringen

$$(x, y, z) = (-14t + 26, -4t + 7, t).$$

For andre likningssystemer kan løsningene også ha to eller flere frie parametre. Da danner løsningene et plan eller en rett geometrisk figur av høyere dimensjon. Slike figurer kalles *vektorrom*, og vi skal lære mer om dem senere.

1.3. Generelle lineære likningssystemer

DEFINISJON 1. Et *generelt lineært likningssystem* med m ligninger og n ukjente x_1, x_2, \dots, x_n er et system av likninger som kan skrives på formen

$$(1.3) \quad \begin{array}{cccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ \vdots & & \vdots & & \ddots & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m \end{array}$$

med konstante koeffisienter $a_{ij}, b_i \in \mathbb{R}$ for $1 \leq i \leq m$, $1 \leq j \leq n$. Vi sier at systemet (1.3) er *homogent* hvis $b_1 = b_2 = \dots = b_m = 0$, og *inhomogent* ellers.

Koeffisientene på venstre side i likningssystemet skrives a_{ij} med to indekser i, j . Vi gir alltid disse koeffisientene navn slik at første indeks i er nummeret til likningen koeffisienten står i, mens andre indeks j er nummeret på variabelen koeffisienten står foran. Alle konstantene b_i på høyre side av likningssystemet har kun en indeks i , nummeret på likningen konstanten står i.

En *løsning* av likningssystemet (1.3) er et n -tupplel (x_1, x_2, \dots, x_n) av tall som passer inn i alle m likningene. Vi sier at likningssystemet er *konsistent* hvis det har minst en løsning, og *inkonsistent* hvis det ikke har noen løsninger.

Akkurat som i tilfellet med to lineære likninger i to variable, er det kun tre muligheter for hvordan løsningene av det generelle lineære likningssystemet (1.3) kan se ut:

- (1) Systemet er inkonsistent (har ingen løsninger),
- (2) Systemet er konsistent og har en entydig løsning,
- (3) Systemet har uendelig mange løsninger som kan beskrives ved hjelp av en eller flere frie parametre.

Litt senere skal vi se hvorfor det er slik, og beskrive en generell metode for å finne alle løsningene av et vilkårlig lineært likningssystem.

1.4. Systematisk løsning av lineære likningssystemer

Vi skal beskrive en systematisk metode for å løse generelle lineære likningssystemer med m likninger og n ukjente. Det er to varianter av denne metoden. *Gauss eliminasjon* er den mest effektive metoden om man skal løse store likningssystemer. Den andre metoden, *Gauss-Jordan eliminasjon*, har andre fordeler som vi skal komme tilbake til litt senere.

1.5. Gauss eliminasjon

La oss se på et eksempel på et lineært likningssystem som har en spesiell form, som vi kaller trappeform:

$$(1.4) \quad \begin{array}{rcccc} -x & + & y & + & z & = & 11 \\ & & 3y & - & z & = & 2 \\ & & & & 2z & = & 14 \end{array}$$

Vi kaller det første tallet som er forskjellig fra null i hver likning for den *ledende koeffisienten* i likningen. Et likningssystem som det ovenfor, der de ledende koeffisientene danner en trapp, kalles et likningssystemet på *trappeform*. Det vil si at hver ledende koeffisient kommer lenger til høyre enn de i likningene ovenfor.

Et likningssystemet på trappeform, slik som det i eksempel (1.4), kan lett vint løses ved hjelp av en metode som kalles *baklengs substitusjon*. Den går ut på at vi starter med å løse den siste likningen i systemet, og jobber oss oppover. Vi løser likningssystemet ovenfor på følgende måte: Først finner vi z fra siste likning,

$$\begin{aligned}2z &= 14 \\z &= 7,\end{aligned}$$

så setter vi $z = 7$ inn i andre likning,

$$\begin{aligned}3y - z &= 2 \\3y - 7 &= 2 \\y &= 3,\end{aligned}$$

og til slutt setter vi inn $y = 3$ og $z = 7$ i første likning,

$$\begin{aligned}-x + y + z &= 11 \\-x + 3 + 7 &= 11 \\x &= -1.\end{aligned}$$

Altså har ligningssystemet en entydig løsning $x = -1$, $y = 3$, $z = 7$.

Denne løsningsmetoden baserer seg på at vi har valgt en bestemt rekkefølge på variablene (her er rekkefølgen x, y, z), og at alle likninger skrives opp på en systematisk måte, slik at alle koeffisientene til hver av variablene står rett under hverandre.

Har vi først bestemt oss for en rekkefølge på variablene, er det egentlig ikke nødvendig å skrive opp likningene når vi skal beskrive likningssystemet. Det er nok å skrive opp alle koeffisientene på en systematisk måte:

$$\begin{pmatrix} -1 & 1 & 1 & 11 \\ 0 & 3 & -1 & 2 \\ 0 & 0 & 2 & 14 \end{pmatrix}$$

Vi har skrevet alle koeffisientene i en blokk med tall, med 3 rader og 4 kolonner. En slik blokk kalles en *matrise*, og vi skal lære mye mer om matriser senere. Det er vanlig å sette runde paranteser rundt en matrise, for å vise at en matrise er en avgrenset enhet.

Matrisen ovenfor kalles den *augmenterte matrisen* til det lineære likningssystemet (1.4). Vi ser at hver rad svarer til en likning, de tre første kolonnene svarer til koeffisientene foran variablene x, y, z , og den siste kolonnen svarer til konstantene på høyre side i likningene. For å markere forskjellen mellom de første tre og den siste kolonnen, skriver vi ofte den augmenterte matrisen som

$$(1.5) \quad \left(\begin{array}{ccc|c} -1 & 1 & 1 & 11 \\ 0 & 3 & -1 & 2 \\ 0 & 0 & 2 & 14 \end{array} \right)$$

Legg merke til at når vi først har bestemt oss for rekkefølgen på variablene som inngår, så betyr den augmenterte matrisen (1.5) akkurat det samme som det lineære likningssystemet (1.4). Det er bare en mer kompakt måte å skrive opp likningssystemet på.

DEFINISJON 2. Vi kaller det første tallet forskjellig fra null i hver rad (som ikke er null) i en matrise for en *ledende koeffisient*. En matrise er på *trappeform* hvis følgende betingelser er oppfylt:

- (1) Alle rader som er null står nederst i matrisen,
- (2) Alle tall under en ledende koeffisient er null.

Vi sier at et lineært likningssystem er på trappeform hvis den augmenterte matrisen til systemet er på trappeform. Dette passer med den definisjonen vi gav tidligere, og betyr altså at likningssystemet kan løses ved hjelp av baklengs substitusjon.

Merk at en augmentert matrise på trappeform ikke trenger å ha en ledende koeffisient i hver eneste kolonne på venstre side av den vertikale linjen. For eksempel er

$$\left(\begin{array}{cccc|c} -1 & 1 & 1 & 4 & 11 \\ 0 & 0 & 3 & -1 & 2 \\ 0 & 0 & 0 & 2 & 14 \end{array} \right)$$

en augmentert matrise på trappeform, selv om det ikke er noen ledende koeffisient i andre kolonne (sjekk det!).

La oss løse likningssystemet som denne augmenterte matrisen svarer til, når vi kaller variablene x_1, x_2, x_3, x_4 . At det mangler en ledende koeffisient i andre kolonne, betyr at dette likningssystemet har x_2 som fri parameter. Ved baklengs substitusjon finner vi da følgende løsning: Siste likning gir $2x_4 = 14$, så $x_4 = 7$. Setter vi $x_4 = 7$ inn i andre likning, får vi $3x_3 - x_4 = 2$ eller $3x_3 - 7 = 2$, det vil si $x_3 = 3$. Neste variabel $x_2 = t$, siden x_2 er en fri parameter. Setter vi disse verdiene for x_2, x_3, x_4 inn i først likning, får vi $-x_1 + x_2 + x_3 + 4x_4 = 11$ eller $-x_1 + t + 3 + 7 = 11$, det vil si $x_1 = t - 1$. Altså får vi uendelig mange løsninger, parametrisert ved $(x_1, x_2, x_3, x_4) = (t - 1, t, 3, 7)$.

Merk også at om vi har en ledende koeffisient på høyre side av den vertikale linjen, så betyr det at likningssystemet ikke har noen løsninger. Dette er fordi raden som en slik ledende koeffisient står i svarer til en likning med venstre side

$$\text{V.S.} = 0x_1 + 0x_2 + \cdots + 0x_n = 0,$$

mens høyre side av likningen er forskjellig fra null. En slik likning kan aldri være oppfylt.

Gauss eliminasjon er en generell metode for å redusere en vilkårlig matrise til en matrise på trappeform, og dette gjøres ved hjelp av visse lovlige operasjoner. Men hvilke operasjoner er det som er lovlige?

Når vi gjør operasjoner på den augmenterte matrisen til et lineært likningssystem, så svarer det til å gjøre operasjoner på likningssystemet.

En lovlig operasjon på et likningssystem er et som bevarer løsningsmengden til systemet. Det viser seg at for å løse et vilkårlig lineært likningssystem, så er det nok å bruke følgende typer av operasjoner:

- (1) Bytte om to likninger
- (2) Multiplisere en likning med et tall forskjellig fra null
- (3) Legge til et tall multiplisert med en likning til en annen likning

Skriver vi likningssystemet opp som en augmentert matrise, svarer disse operasjonene på likningene i systemet til bestemte operasjoner på radene i matrisen. Vi kaller disse operasjonene for *elementære radoperasjoner*:

- (1) Bytte om to rader
- (2) Multiplisere en rad med et tall forskjellig fra null
- (3) Legge til et tall multiplisert med en rad til en annen rad

Dette er derfor kun disse operasjonene vi skal bruke når vi reduserer den augmenterte matrisen til et lineært likningssystem til en trappeform.

DEFINISJON 3. *Gauss eliminasjon* er en metode for å redusere en vilkårlig matrise til en matrise på trappeform ved hjelp av elementære radoperasjoner.

La oss først se et eksempel på Gauss eliminasjon før vi beskriver den generelle algoritmen. Vi skal se på det lineære likningssystemet

$$(1.6) \quad \begin{array}{r} x + 2y + z = 4 \\ 3x + 8y + 7z = 20 \\ 2x + 7y + 9z = 23 \end{array}$$

Vi begynner med å skrive ned den augmenterte matrisen som svarer til dette likningssystemet:

$$\left(\begin{array}{ccc|c} 1 & 2 & 1 & 4 \\ 3 & 8 & 7 & 20 \\ 2 & 7 & 9 & 23 \end{array} \right)$$

Neste steg er selve Gauss eliminasjonen, som vi bruker for å redusere denne matrisen til en trappeform. Vi bruker notasjonen $R(i)$ for rad i for å beskrive de elementære radoperasjonene som vi utfører nedenfor:

$$\begin{aligned} \left(\begin{array}{ccc|c} 1 & 2 & 1 & 4 \\ 3 & 8 & 7 & 20 \\ 2 & 7 & 9 & 23 \end{array} \right) &\xrightarrow{R(2):=R(2)-3R(1)} \left(\begin{array}{ccc|c} 1 & 2 & 1 & 4 \\ 0 & 2 & 4 & 8 \\ 2 & 7 & 9 & 23 \end{array} \right) \xrightarrow{R(3):=R(3)-2R(1)} \\ &\left(\begin{array}{ccc|c} 1 & 2 & 1 & 4 \\ 0 & 2 & 4 & 8 \\ 0 & 3 & 7 & 15 \end{array} \right) \xrightarrow{R(3):=R(3)-\frac{3}{2}R(2)} \left(\begin{array}{ccc|c} 1 & 2 & 1 & 4 \\ 0 & 2 & 4 & 8 \\ 0 & 0 & 1 & 3 \end{array} \right) \end{aligned}$$

Trappeformen som vi kommer fram til ovenfor, svarer til følgende lineære likningssystem:

$$\begin{array}{rcl} x + 2y + z & = & 4 \\ & 2y + 4z & = 8 \\ & & z = 3 \end{array}$$

Dette likningssystemet kan vi løse ved hjelp av baklengs substitusjon: Siste likning gir $z = 3$. Setter vi inn $z = 3$ i andre likning, får vi $2y + 12 = 8$, det vil si $y = -2$, og setter vi inn disse verdiene for y og z i første likning, får vi $x - 4 + 3 = 4$, det vil si $x = 5$. Altså har likningssystemet (1.6) en entydig løsning $x = 5$, $y = -2$, $z = 3$, som vi har funnet ved hjelp av Gauss eliminasjon.

Gauss eliminasjon av en vilkårlig matrise er gitt ved følgende framgangsmåte:

- (1) Finn første kolonne som ikke er null. Hvis tallet på øverste plass i denne kolonnen er null, bytt om to rader slik at tallet på denne plassen blir forskjellig fra null. Vi har nå fått en ledende koeffisient på denne plassen.
- (2) Bruk elementære radoperasjoner slik at alle tallene under den ledende koeffisienten fra (1) blir null.
- (3) Vi er nå ferdig med kolonnen som den ledende koeffisienten fra (1) står i. Hvis vi ikke allerede har fått en trappeform, ser vi på kolonnene som står til høyre for den ledende koeffisienten fra (1), og gjentar hele prosedyren.

Legg merke til at denne framgangsmåten kan brukes på en hvilken som helst matrise, og vi vil alltid komme fram til en trappeform. Men det er flere mulige valg underveis, som alle er like riktige, og det er ikke sikkert at man kommer fram til samme trappeform uansett hvordan man velger å gjøre Gauss eliminasjonen. Trappeformen er altså ikke entydig.

1.6. Gauss-Jordan eliminasjon

Gauss-Jordan eliminasjon går ut på å forenkle en matrise ved hjelp av elementære radoperasjoner. Ved Gauss eliminasjon er målet å få matrisen over på trappeform, mens vi ved Gauss-Jordan eliminasjon søker å gjøre matrisen ennå enklere.

DEFINISJON 4. Vi sier at en matrise er på *reduisert trappeform* hvis følgende betingelser er oppfylt:

- (1) Alle rader som er null står nederst i matrisen,
- (2) Alle tall under en ledende koeffisient er null,
- (3) Alle ledende koeffisienter er 1,
- (4) Alle tall over en ledende koeffisient er null.

Vi gjenkjenner (1) og (2) som betingelsene for at en matrise er på trappeform. En redusert trappeform er med andre ord en trappeform som i tillegg oppfylder betingelse (3) og (4) ovenfor. Vi sier at et lineært

likningssystem er på redusert trappeform hvis den augmenterte matrisen til systemet er på trappeform.

DEFINISJON 5. *Gauss-Jordan eliminasjon* er en generell metode for å redusere en vilkårlig matrise til en matrise på redusert trappeform ved hjelp av elementære radoperasjoner.

Siden målet for Gauss-Jordan eliminasjon er å komme fram til en redusert trappeform, som skal tilfredstille noen ekstra betingelser i forhold til en trappeform, er det klart at vi må regne med å gjøre noen ekstra elementære radoperasjoner i forhold til en Gauss eliminasjon for å få dette til. Til gjengjeld er det enda enklere å løse et likningssystem på redusert trappeform enn et på trappeform, fordi vi slipper å gjøre baklengs substitusjon. La oss se noen eksempler på dette.

Først ser vi på et eksempel på en augmentert matrise på redusert trappeform, der hver kolonne på venstre side av den vertikale linjen har en ledende koeffisient:

$$(1.7) \quad \left(\begin{array}{ccc|c} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 7 \end{array} \right)$$

At denne matrisen er på redusert trappeform, kan man sjekke ved å sjekke at betingelse (1) - (4) i definisjonen er oppfylt (gjør det!). Den reduserte trappeformen i dette eksempelet svarer til følgende lineære likningssystem, når vi kaller variablene x, y, z :

$$\begin{array}{rcl} x & = & -1 \\ y & = & 3 \\ z & = & 7 \end{array}$$

Dette betyr at systemet har en entydig løsning $(x, y, z) = (-1, 3, 7)$, som vi kan lese rett ut fra den reduserte trappeformen. Vi trenger altså ikke bruke baklengs substitusjon her.

La oss så se på et annet eksempel på en augmentert matrise på redusert trappeform, der det er noen kolonner på venstre side av den vertikale linjen som er uten ledende koeffisienter:

$$(1.8) \quad \left(\begin{array}{ccc|c} 1 & -5 & 0 & -1 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{array} \right)$$

Kaller vi variablene x, y, z , så betyr den manglende ledende koeffisienten i andre kolonne at $y = t$ er en fri parameter. Siste linje gir $0 = 0$, altså ingen informasjon. Informasjonen i de to første linjene leser vi ut på følgende måte: Første linje betyr at $x - 5y = -1$, altså er $x = 5t - 1$. Andre linje betyr at $z = 3$. Altså har systemet uendelig mange løsninger, parametrisert ved $(x, y, z) = (5t - 1, t, 3)$.

Til slutt ser vi på et eksempel på en augmentert matrise på redusert trappeform, der det er en ledende koeffisient på høyre side av

den vertikale linjen:

$$(1.9) \quad \left(\begin{array}{cccc|c} 1 & -5 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array} \right)$$

I dette tilfellet så svarer raden med en ledende koeffisient på høyre side til en likning av typen $0 = 1$, som aldri er tilfredsstillt. Systemet har derfor ingen løsninger.

Før vi beskriver selve algoritmen for Gauss-Jordan eliminasjon, la oss se på et eksempel der vi løser et lineært likningssystem ved hjelp av Gauss-Jordan eliminasjon. Vi skal se på det lineære likningssystemet

$$(1.10) \quad \boxed{\begin{array}{cccccc} x_1 & + & x_2 & + & x_3 & + & x_4 & = & 12 \\ x_1 & + & 2x_2 & & & + & 5x_4 & = & 17 \\ 3x_1 & + & 2x_2 & + & 4x_3 & - & x_4 & = & 31 \end{array}}$$

Vi begynner med å skrive ned den augmenterte matrisen som svarer til dette likningssystemet:

$$\left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 12 \\ 1 & 2 & 0 & 5 & 17 \\ 3 & 2 & 4 & -1 & 31 \end{array} \right)$$

Neste steg er selve Gauss-Jordan eliminasjonen, som vi bruker for å redusere denne matrisen til en redusert trappeform. Vi bruker samme notasjon for de elementære radoperasjonene som i forrige seksjon:

$$\begin{aligned} & \left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 12 \\ 1 & 2 & 0 & 5 & 17 \\ 3 & 2 & 4 & -1 & 31 \end{array} \right) \xrightarrow{R(2):=R(2)-R(1)} \left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 12 \\ 0 & 1 & -1 & 4 & 5 \\ 3 & 2 & 4 & -1 & 31 \end{array} \right) \\ & \xrightarrow{R(3):=R(3)-3R(1)} \left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 12 \\ 0 & 1 & -1 & 4 & 5 \\ 0 & -1 & 1 & -4 & -5 \end{array} \right) \xrightarrow{R(3):=R(3)+R(2)} \\ & \left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 12 \\ 0 & 1 & -1 & 4 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right) \xrightarrow{R(1):=R(1)-R(2)} \left(\begin{array}{cccc|c} 1 & 0 & 2 & -3 & 7 \\ 0 & 1 & -1 & 4 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right) \end{aligned}$$

Den reduserte trappeformen som vi kommer fram til ovenfor, mangler en ledende koeffisient i tredje og fjerde kolonne, så $x_3 = s$ og $x_4 = t$ er frie parametre. Fra de to første radene i den reduserte trappeformen, leser vi av at $x_1 = -2s + 3t + 7$ og $x_2 = s - 4t + 5$. Dette betyr at likningssystemet (1.10) har uendelig mange løsninger, og de er gitt som $(x_1, x_2, x_3, x_4) = (-2s + 3t + 7, s - 4t + 5, s, t)$, der s og t er frie parametre.

Gauss-Jordan eliminasjon av en vilkårlig matrise er gitt ved følgende framgangsmåte:

- (1) Finn første kolonne som ikke er null. Hvis tallet på øverste plass i denne kolonnen er null, bytt om to rader slik at tallet

på denne plassen blir forskjellig fra null. Vi har nå fått en ledende koeffisient på denne plassen.

- (2) Hvis den ledende koeffisienten fra (1) ikke er 1, multipliser denne raden med et tall slik at koeffisienten blir 1.
- (3) Bruk elementære radoperasjoner slik at alle tallene under den ledende koeffisienten fra (1) blir null.
- (4) Bruk elementære radoperasjoner slik at alle tallene over den ledende koeffisienten fra (1) blir null.
- (5) Vi er nå ferdig med kolonnen som den ledende koeffisienten fra (1) står i. Hvis vi ikke allerede har fått en trappeform, ser vi på kolonnene som står til høyre for den ledende koeffisienten fra (1), og gjentar hele prosedyren.

Legg merke til at denne framgangsmåten kan brukes på en hvilken som helst matrise, og vi vil alltid komme fram til en redusert trappeform. Uansett hvordan vi velger å gjøre Gauss-Jordan eliminasjonen, vil vi komme fram til samme reduserte trappeform. I motsetning til trappeformen vi kommer fram til ved Gauss eliminasjon, er altså den reduserte trappeformen til en matrise entydig. Vi sier at to matriser er *rad-ekvivalente* hvis de har samme reduserte trappeform.

1.7. Konklusjoner

La oss se på et vilkårlig lineært likningssystem med m likninger og n ukjente, og den tilsvarende augmenterte matrisen. Vi kan alltid finne fram til en trappeform til denne matrisen ved hjelp av Gauss eliminasjon, eller den reduserte trappeformen ved hjelp av Gauss-Jordan eliminasjon. Uansett hvilken av de to metodene vi bruker, er svaret vi kommer fram til av en av disse tre typene:

- (1) Alle ledende koeffisienter står på venstre side av den vertikale linjen, og det er en ledende koeffisient for hver kolonne der.
- (2) Alle ledende koeffisienter står på venstre side av den vertikale linjen, men en eller flere kolonner der mangler en ledende koeffisient.
- (3) Det er en ledende koeffisient på høyre side av den vertikale linjen.

Vi har sett at vi får følgende løsninger av det lineære likningssystemet i disse tre tilfellene:

- (1) En entydig løsning,
- (2) Uendelig mange løsninger, gitt av en eller flere frie parametre (en for hver kolonne som mangler ledende koeffisient på venstre side av den vertikale linjen),
- (3) Ingen løsninger.

Vi kan oppsummere disse resultatene i følgende teorem:

TEOREM 1. *Et vilkårlig lineært likningssystem har enten en entydig løsning, uendelig mange løsninger gitt ved en eller flere frie parametre, eller ingen løsninger.*

Til slutt nevner vi homogene lineære likningssystem. Vi husker at slike likningssystemer har $b_1 = b_2 = \dots = b_m = 0$, det vil si at den augmenterte matrisen har en null-kolonne på høyre side av den vertikale linjen. Vi forstår at siden hele denne kolonnen er null, så vil den forbli null uansett hvilke elementære radoperasjoner vi foretar oss. Derfor har trappeformen etter en Gauss eliminasjon, eller den reduserte trappeformen etter en Gauss-Jordan eliminasjon, også en null-kolonne på høyre side av den vertikale linjen.

Dette betyr at tilfelle (3) blant de tre mulige tilfellene ovenfor aldri inntreffer for homogene likningssystemer — vi kan aldri ha en ledende koeffisient på høyre side av den vertikale linjen. Det er fordi et homogent likningssystem alltid har minst en løsning, nemlig *den trivielle løsningen* $x_1 = x_2 = \dots = x_n = 0$.

For homogene likningssystem er det altså kun to muligheter for løsninger: Enten så er vi i situasjon (1) i listen ovenfor, og systemet har kun den trivielle løsningen, eller så er vi i situasjon (2), og systemet har uendelig mange løsninger, gitt ved en eller flere frie parametre (og den trivielle løsningen svarer da til å sette alle de frie parametrene lik null).

KAPITTEL 2

Matriser og matriseregning

Vi har allerede sett en viktig anvendelse av matriser i kapittel 1. Vi brukte augmenterte matriser som en hjelpemiddel til å løse lineære likningssystemer. I dette kapitlet skal vi se nærmere på matriser og matriseregning.

2.1. Matriser og vektorer

DEFINISJON 6. En *matrise* er en rektangulær blokk av reelle tall. Størrelsen av en matrise er gitt ved antall rader og antall kolonner i matrisen. En $m \times n$ -matrise er en matrise med m rader og n kolonner. Legg merke til at radene spesifiseres først, så kolonnene.

Vi bruker gjerne store bokstaver som A , B eller X , Y når vi gir en matrise navn, og vi skriver runde parenteser rundt en blokk av tall for å markere avgrensningen til matrisen. Eksempler på matriser er

$$A = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}, \quad X = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$$

Her er A en 2×2 -matrise, mens X er en 2×3 -matrise. Skal vi gi navn til hvert enkelt tall i matrisen, gir vi dem gjerne navn som a_{ij} for tallet på plass (i, j) , altså i rad i og kolonne j , i matrisen A . Altså er $a_{12} = 3$ og $x_{23} = 6$ i eksemplene ovenfor.

DEFINISJON 7. En *vektor* eller en *kolonnevektor* er en matrise som kun inneholder en kolonne, altså en $n \times 1$ -matrise. En vektor kalles også en n -vektor når vi ønsker å spesifisere størrelsen på vektoren.

Vi bruker gjerne små bokstaver som \mathbf{v} når vi skal gi en vektor navn. Noen ganger skriver vi vektoren \mathbf{v} med pil over navnet v , men i dette kompendiet bruker vi istedet uthevet skrift som notasjon for vektorer. Et eksempel på en vektor er

$$\mathbf{v} = \begin{pmatrix} 1 \\ 0 \\ -3 \\ 2 \end{pmatrix}.$$

Noen ganger skriver vi også denne vektoren som $\mathbf{v} = (1, 0, -3, 2)$.

DEFINISJON 8. En *skalar* er et tall. Grunnen til at vi har et eget navn for dette, er at vi skal kunne skille tall fra matriser og vektorer.

2.2. Addisjon og skalarmultiplikasjon

DEFINISJON 9. Vi kan addere sammen to matriser hvis de har samme størrelse, det vil si samme antall rader og kolonner. Vi adderer da matrisene plass for plass. Vi bruker symbolet $+$ for matriseaddisjon, akkurat som for vanlig addisjon av tall.

La oss ta med noen eksempler på matrise-addisjon. Først ser vi på addisjon av to 2×2 -matriser

$$\begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} + \begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix} = \begin{pmatrix} 2 & 5 \\ 6 & 9 \end{pmatrix},$$

og så et eksempel på et par av matriser som ikke kan adderes på grunn av at de har ulik størrelse

$$\begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} + \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \text{ er ikke definert.}$$

Subtraksjon av matriser foregår på samme måte som addisjon, og vi bruker det vanlige symbolet $-$ for dette.

DEFINISJON 10. Vi kan multiplisere en skalar med en matrise. Vi multipliserer da skalaren inn i matrisen plass for plass. Vi bruker symbolet \cdot for skalarmultiplikasjon av matriser (men akkurat som for vanlig multiplikasjon, så sløyfer vi ofte dette symbolet).

$$7 \begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix} = \begin{pmatrix} 7 & 14 \\ 28 & 35 \end{pmatrix}$$

En spesiell matrise er *nullmatrisen*, matrisen som kun består av tallet null på hver plass. Egentlig fins mange null-matriser, avhengig av størrelsen på matrisen, men for enkelthets skyld skriver vi null-matrisen som 0 uansett størrelse. Vi har at $A + 0 = 0 + A = A$ for en vilkårlig matrise A . Nullmatrisen spiller altså samme rolle for matriseregning som tallet 0 spiller for regning med tall.

Siden vektorer kan betraktes som spesialtilfeller av matriser, så kan vi addere (og subtrahere) vektorer av samme størrelse, og multiplisere vektorer med skalarer.

2.3. Euklidske rom

I dette avsnittet skal vi bruke skrivemåten $\mathbf{v} = (v_1, v_2, \dots, v_n)$ for en vektor \mathbf{v} , istedet for å skrive \mathbf{v} som en kolonnevektor. La oss se på rommet \mathbb{R}^2 som består av alle 2-vektorer $\mathbf{v} = (x, y)$, og rommet \mathbb{R}^3 som består av alle 3-vektorer $\mathbf{v} = (x, y, z)$. Disse rommene er eksempler på Euklidske rom.

DEFINISJON 11. Det *Euklidske n -rommet* er rommet \mathbb{R}^n bestående av alle n -vektorer $\mathbf{v} = (v_1, v_2, \dots, v_n)$.

De viktigste operasjonene på vektorer i Euklidsk n -rom \mathbb{R}^n er addisjon og skalarmultiplikasjon

$$\begin{aligned}\mathbf{v} + \mathbf{w} &= (v_1 + w_1, v_2 + w_2, \dots, v_n + w_n) \\ c\mathbf{v} &= (cv_1, cv_2, \dots, cv_n)\end{aligned}$$

for vektorer $\mathbf{v} = (v_1, v_2, \dots, v_n)$, $\mathbf{w} = (w_1, w_2, \dots, w_n) \in \mathbb{R}^n$ og skalarer $c \in \mathbb{R}$. Disse operasjonene er spesialtilfeller av addisjon og skalarmultiplikasjon av matriser når vi tenker på vektorer som kolonnevektorer, dvs matriser som består av en kolonne.

En annen viktig operasjon er *skalarproduktet* eller *prikkproduktet* av to vektorer i Euklidsk n -rom

$$\mathbf{v} \cdot \mathbf{w} = v_1w_1 + v_2w_2 + \dots + v_nw_n$$

for vektorer $\mathbf{v} = (v_1, v_2, \dots, v_n)$, $\mathbf{w} = (w_1, w_2, \dots, w_n) \in \mathbb{R}^n$. Legg merke til at vi skriver \cdot for denne operasjonen, og at svaret blir et tall, herav navnene prikkprodukt og skalarprodukt.

2.4. Matrisemultiplikasjon

Definisjonen av matrisemultiplikasjon er mye mer komplisert enn de operasjonene vi har sett på så langt. Denne operasjonen skjer *ikke* plass for plass.

To matriser A og B kan multipliseres sammen hvis A er en $m \times n$ -matrise og B er en $n \times p$ -matrise. Antall kolonner n i den første matrisen A må altså være det samme som antall rader n i den andre matrisen B . Hvis dette er tilfellet, så blir svaret av matrisemultiplikasjonene en $m \times p$ matrise, som vi skriver $A \cdot B$ eller AB . Hvis dette ikke er tilfelle, så er matrisemultiplikasjonen ikke definert. Skjematisk kan vi skrive dette opp på følgende måte:

$$\begin{array}{ccc} A & B & \longrightarrow A \cdot B \\ m \times n & n \times p & m \times p\end{array}$$

Hvis produktet AB er definert, så regner vi det ut på følgende måte: For å regne ut tallet på plass (i, j) i svaret, altså tallet i rad i og kolonne j i AB , så regner vi ut prikkproduktet av rad i i matrise A med kolonne j i matrise B . På grunn av størrelsene på matrisene, kan vi oppfatte begge disse som n -vektorer, og prikkproduktet eksisterer derfor. Vi får altså følgende formel for tallet på plass (i, j) i AB :

$$(A \cdot B)_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \dots + a_{in}b_{nj}$$

La oss se på noen eksempler. Først multipliserer vi sammen to 2×2 -matriser

$$\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 3 \\ 4 & -1 \end{pmatrix} = \begin{pmatrix} 1 \cdot 0 + 2 \cdot 4 & 1 \cdot 3 + 2 \cdot (-1) \\ 0 \cdot 0 + 1 \cdot 4 & 0 \cdot 3 + 1 \cdot (-1) \end{pmatrix} = \begin{pmatrix} 8 & 1 \\ 4 & -1 \end{pmatrix}$$

og så forsøker vi å multiplisere sammen en 2×3 -matrise med en 2×2 -matrise

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 7 \end{pmatrix} \cdot \begin{pmatrix} 0 & 3 \\ 4 & -1 \end{pmatrix} \text{ er ikke definert.}$$

Hadde vi forsøkt å regne ut den siste matrisemultiplikasjonen, ville vi ha endt opp med en hel rekke med prikk-produkter av en 3-vektor med en 2-vektor (siden første matrise har tre kolonner mens andre matrise har to rader), og slike prikk-produkter er ikke definert.

Strategien for multiplikasjon av matriser er slik: For hver rad i første matrise, prikker vi denne raden med kolonnevektorene i andre matrise, en etter en, og skriver opp hvert av svarene på riktig plass i produkt-matrisen.

Vi ser at operasjonen ikke er symmetrisk. Vi bruker alltid rader fra første matrise og kolonner fra andre matriser. Derfor er rekkefølgen svært viktig når vi ganger sammen matriser. Vi kan uttrykke dette slik: *Multiplikasjon av matriser er ikke kommutative*, vi har $AB \neq BA$.

Vi har en spesiell matrise som kalles *identitetsmatrisen*. Det er matrisen

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

Vi skriver noen ganger I_n istedet for I for å få fram at dette er en $n \times n$ -matrise. Identitetsmatrisen har egenskapen at $AI = IA = A$ for en vilkårlig matrise A . Identitetsmatrisen spiller altså samme rolle i matriseregning som tallet 1 spiller for regning med tall.

Det er en rekke regneregler for matriser, og vi skal ikke liste opp alle her. I utgangspunktet gjelder de fleste regneregler for vanlige tall også for matriser. Eksempler på dette er at $A(B + C) = AB + AC$ (distributiv lov) og at $A(BC) = (AB)C$ (assosiativ lov). Unntaket er at matrisemultiplikasjon ikke er kommutativ, $AB \neq BA$, og alt dette fører med seg. For eksempel har vi for matriser at

$$(A + B)^2 = (A + B)(A + B) = A^2 + AB + BA + B^2$$

Men siden $AB \neq BA$, så er dette *ikke* lik $A^2 + 2AB + B^2$. Når vi regner med matriser, må vi også passe på at størrelsene på matrisene passer overens.

2.5. Lineære likningssystem på matriseform

Så langt har vi ikke sagt så mye om hvorfor matrisemultiplikasjon ble definert slik vi gjorde det i forrige avsnitt. Men eksemplene i dette avsnittet er en klar indikasjon på at det har noe for seg å definere multiplikasjonen slik som vi har gjort det.

La oss se på et generelt lineært likningssystem med n ukjente og m likninger,

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

Vi skal bruke definisjonen av matrisemultiplikasjon til å skrive om dette systemet til en likning som involverer matriser. Koeffisientmatrisen til systemet er gitt ved

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

Vi definerer i tillegg de to vektorene

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \text{ and } \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

Når vi bruker matrisemultiplikasjonen, slik som vi definerte den i forrige avsnitt, får vi at

$$A\mathbf{x} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} = \mathbf{b}$$

Derfor er det lineære likningssystemet ovenfor ekvivalent til den ene matriselikningen

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

I denne likningen er det \mathbf{x} som er den ukjente, og vi får løsningene til likningen uttrykt på vektorform når vi skriver dem på denne måten.

Ser vi for eksempel på et eksempel på et lineært likningssystem fra forrige avsnitt,

$$\begin{array}{cccccc} x_1 & + & x_2 & + & x_3 & + & x_4 & = & 12 \\ x_1 & + & 2x_2 & & & + & 5x_4 & = & 17 \\ 3x_1 & + & 2x_2 & + & 4x_3 & - & x_4 & = & 31 \end{array}$$

husker vi at $(x_1, x_2, x_3, x_4) = (-2s + 3t + 7, s - 4t + 5, s, t)$ var løsningene av systemet, med s og t som frie parametre. Skal vi skrive disse

løsningene på vektorform, får vi

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} -2s + 3t + 7 \\ s - 4t + 5 \\ s \\ t \end{pmatrix} = s \begin{pmatrix} -2 \\ 1 \\ 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 3 \\ -4 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 7 \\ 5 \\ 0 \\ 0 \end{pmatrix}$$

2.6. Invertering av matriser

Invertering av matriser svarer i en viss forstand til divisjon av matriser, på tilsvarende måte som vi kan dele på tallet 2 ved å multiplisere med $2^{-1} = \frac{1}{2}$, som er inversen til tallet 2.

DEFINISJON 12. La A være en $n \times n$ -matrise. En invers matrise til A er en matrise B som er slik at

$$A \cdot B = B \cdot A = I_n$$

Hvis A har en invers matrise B , så er den entydig, og vi skriver den vanligvis A^{-1} . Legg merke til at kun kvadratiske matriser (det vil si matriser med like mange rader som kolonner) kan ha en invers. Hvis en matrise A har en invers, så har A^{-1} samme størrelse som A .

Det er altså slik at noen kvadratiske matriser har en invers, mens andre ikke har det. La oss begynne med noen eksempler. Matrisen

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

har invers

$$A^{-1} = \begin{pmatrix} 1 & 3 \\ 0 & -1 \end{pmatrix},$$

fordi vi har at

$$AA^{-1} = A^{-1}A = \begin{pmatrix} 1 & 3 \\ 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 3 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Derimot har matrisen

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

ingen invers, fordi det ikke finnes noen matriser M^{-1} som passer inn i ligningene $MM^{-1} = M^{-1}M = I_2$.

DEFINISJON 13. Vi sier at en $n \times n$ -matrise A er *invertibel* hvis det fins en invers matrise A^{-1} , og at A er *singulær* hvis det ikke finnes en invers matrise. En ikke-singulær matrise er altså det samme som en invertibel matrise.

Gitt en kvadratisk matrise A , hvordan kan vi finne ut om den har en invers, og eventuelt hvordan den inverse matrisen A^{-1} ser ut? I eksemplene ovenfor er svarene gitt, uten at vi har sagt noe om hvordan vi kom fram til dem. Nedenfor skal vi beskrive to metoder som man kan bruke til å invertere matriser, hvis det er mulig.

Metode 1: Formel. La A være en 2×2 -matrise, da fins det en formel som bestemmer om A har en invers, og hva den inverse matrisen A^{-1} eventuelt er. La matrisen A være gitt ved

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Da kalles tallet $\det(A) = ad - bc$ for *determinanten* til matrisen A , og dette tallet bestemmer om A har en invers. Hvis $ad - bc = 0$, så har A ingen invers matrise A^{-1} , mens hvis $ad - bc \neq 0$, så har A en invers matrise A^{-1} , gitt ved

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

I de to eksemplene som vi så på ovenfor, så er determinantene henholdsvis $\det(A) = 1(-1) - 30 = -1$ og $\det(M) = 10 - 30 = 0$. Det betyr at A har invers gitt ved

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

mens M^{-1} ikke finnes. Legg merke til at matrisene vi har oppgitt her kun gjelder for 2×2 -matriser. Det er i prinsippet mulig å skrive opp formler for $n \times n$ -matriser for en vilkårlig stor n , men disse formlene blir svært raskt svært kompliserte.

Metode 2: Gauss-Jordan eliminasjon. La A være en kvadratisk matrise. Da kan vi samtidig finne ut om A er invertibel, og også hva A^{-1} eventuelt blir, ved å bruke følgende framgangsmåte:

- (1) Skriv opp matrisen $(A | I_n)$
- (2) Reduser denne matrisen til en redusert trappeform ved hjelp av Gauss-Jordan eliminasjon
- (3) Hvis den reduserte trappeformen er av typen $(I_n | B)$, så er A invertibel og $A^{-1} = B$. Hvis den reduserte trappeformen ikke er av denne typen, så er A ikke invertibel.

Litt forenklet kan vi forklare at denne metoden virker på følgende måte: Hvis A^{-1} fins, så er effekten av alle de elementære radoperasjonene den samme som multiplikasjon med A^{-1} fra venstre. Altså blir den reduserte trappeformen

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

Motsatt, hvis den reduserte trappeformen ikke har I_n på venstre side av den vertikale linjen, så betyr det at det ikke fins noe sekvenser av

elementære radoperasjoner som tilsammen bringer A over til I_n . Altså finnes ikke A^{-1} .

La oss ta med et eksempel på bruk av denne metoden for å invertere en matrise. I dette eksempelet ser vi på 3×3 -matrisen

$$A = \begin{pmatrix} 4 & 3 & 2 \\ 5 & 6 & 3 \\ 3 & 5 & 2 \end{pmatrix}$$

Vi følger metoden slik den er beskrevet ovenfor, så vi begynner med å danne matrisen

$$(A | I_3) = \left(\begin{array}{ccc|ccc} 4 & 3 & 2 & 1 & 0 & 0 \\ 5 & 6 & 3 & 0 & 1 & 0 \\ 3 & 5 & 2 & 0 & 0 & 1 \end{array} \right)$$

Så utfører vi Gauss-Jordan eliminasjon på denne matrisen på vanlig måte. Det betyr at vi ikke tar hensyn til den vertikale linjen, og utfører hver av de elementære radoperasjonene på radene som om vi hadde en vanlig 3×6 -matrise. Vi får følgende reduserte trappeform:

$$\left(\begin{array}{ccc|ccc} 4 & 3 & 2 & 1 & 0 & 0 \\ 5 & 6 & 3 & 0 & 1 & 0 \\ 3 & 5 & 2 & 0 & 0 & 1 \end{array} \right) \longrightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 3 & -4 & 3 \\ 0 & 1 & 0 & 1 & -2 & 2 \\ 0 & 0 & 1 & -7 & 11 & -9 \end{array} \right)$$

Vi hoppet her over de enkelte stegene i Gauss-Jordan eliminasjonen. Sjekk at du får samme svar! Til slutt leser vi av den informasjonen som vi får fra den reduserte trappeformen i følge punkt (3) ovenfor. Siden matrisen er på formen $(I_3 | B)$ (det vil si at matrisen foran den vertikale streken er en identitetsmatrise), så er matrisen A inverterbar. Dessuten kan vi lese av A^{-1} fra siste del av matrisen:

$$A^{-1} = \begin{pmatrix} 3 & -4 & 3 \\ 1 & -2 & 2 \\ -7 & 11 & -9 \end{pmatrix}$$

Hadde derimot første del av den reduserte trappeformen ikke vært en identitetsmatrise, så kunne vi ha konkludert at A ikke var inverterbar.

Vi nevner også de viktigste regnereglene for invertering av matriser. Gitt at A og B er invertible matriser av samme størrelser, så har vi

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

Den første regelen her er svært viktig å huske på. Grunnen til at vi må bytte om rekkefølgen av A og B , er at

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AI_nA^{-1} = AA^{-1} = I_n$$

Husk at rekkefølgen er viktig ved matrisemultiplikasjon!

Vi kan bruke invertering av matriser til å løse lineære likningssystem med like mange likninger som ukjente. Vi skriver likningssystemet på matriseform $A\mathbf{x} = \mathbf{b}$. Hvis koeffisientmatrisen A er invertibel, så kan vi løse systemet på følgende måte: $A\mathbf{x} = \mathbf{b}$ gir $A^{-1}(A\mathbf{x}) = A^{-1}\mathbf{b}$, det vil

si $\mathbf{x} = (A^{-1}A)\mathbf{x} = A^{-1}\mathbf{b}$. Altså har likningssystemet en entydig løsning $\mathbf{x} = A^{-1}\mathbf{b}$ i dette tilfellet.

Har vi en matriselikning $AX = B$, der A og B er gitte matriser, mens X er en ukjent matrise som vi skal finne, kan vi gå fram på helt tilsvarende måte: Hvis A er en inverterbar kvadratisk matrise, så har likningssystemet en entydig løsning $X = A^{-1}B$.

Til slutt oppsummerer vi litt om forholdet mellom løsninger av lineære likningssystem og inverterbare matriser. La $A\mathbf{x} = \mathbf{b}$ være et lineært likningssystem med n likninger og n ukjente. Da har vi følgende to muligheter:

A er invertibel	A er ikke invertibel (singulær)
- Den reduserte trappeformen til A er I_n	- Den reduserte trappeformen til A har null-rader
- $A\mathbf{x} = \mathbf{b}$ har entydig løsning $\mathbf{x} = A^{-1}\mathbf{b}$	- $A\mathbf{x} = \mathbf{b}$ har ingen eller uendelig mange løsninger

2.7. Determinanter

Determinanten er en funksjon som kan defineres for en vilkårlig kvadratisk matrise, selv om vi kun definerte den for 2x2-matriser i forrige avsnitt. Determinantens viktigste egenskap er at den bestemmer om matrisen er invertibel.

DEFINISJON 14. La A være en vilkårlig $n \times n$ -matrise. Hvis vi skriver matrisen A på formen

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{12} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

så er determinanten til A gitt ved formelen

$$\det(A) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)},$$

hvor S_n er alle permutasjoner av tallene $1, 2, \dots, n$, og for hver slik permutasjon $\sigma \in S_n$, så er $\operatorname{sgn}(\sigma) = 1$ hvis permutasjonen er jevn (et jevnt antall transposisjoner) og $\operatorname{sgn}(\sigma) = -1$ hvis permutasjonen er odde (et odde antall transposisjoner). En alternativ skrivemåte for determinanten til A er $|A|$.

Det første vi legger merke til er at denne formelen er nokså komplisert! Mengden S_n av alle permutasjoner av tallene $1, 2, \dots, n$ består av $n!$ permutasjoner. Det betyr at i formelen ovenfor er det $n!$ ledd, og hvert ledd er et produkt av n tall fra matrisen. For eksempel er

$2! = 2$, $5! = 120$, $10! = 3628800$ mens $100! \cong 9 \cdot 10^{157}$, så $n!$ og dermed størrelsen på denne formelen vokser svært raskt!

La oss skrive opp hvordan denne formelen blir når $n = 2$ eller $n = 3$. For $n = 2$ er det to permutasjoner av tallene 1 og 2: Først har vi $\sigma = (12)$, det vil si $\sigma(1) = 2$ og $\sigma(2) = 1$. Dette er en jevn permutasjon (null bytter eller transposisjoner). Og så har vi $\sigma = (21)$, det vil si $\sigma(1) = 1$ og $\sigma(2) = 2$. Dette er en odde permutasjon (et bytte eller transposisjon). Så formelen for $n = 2$ blir

$$\det(A) = 1 \cdot a_{11}a_{22} + (-1) \cdot a_{12}a_{21} = a_{11}a_{22} - a_{12}a_{21}$$

Vi ser at vi får en ganske enkel formel med to ledd, som stemmer overens med den vi gav i avsnittet om invertering av matriser.

Så ser vi på tilfellet der $n = 3$. I dette tilfellet har vi tre jevne permutasjoner (123) , (312) og (231) , og tre odde permutasjoner (132) , (213) og (321) . Vi får da følgende formel, med $3! = 6$ ledd:

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

Determinanten til en $n \times n$ -matrise der $n > 3$ kan vi finne på tilsvarende måte, men vi ser at dette blir fort veldig store formler. For $n = 4$ får vi for eksempel $4! = 24$ ledd, for $n = 5$ får vi $5! = 120$ ledd, og for $n = 100$ får vi mer enn $9 \cdot 10^{157}$ ledd!

Metode 1: Utvikling av determinanten langs en rad eller kolonne. Dette er den første metoden for å regne ut determinanter som vi skal se på, og den er egentlig kun en omskriving av formelen i definisjonen. Vi skal beskrive hvordan denne metoden fungerer når vi utvikler determinanten til $n \times n$ -matrisen

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

langs første rad $(a_{11} \ a_{12} \ \dots \ a_{1n})$. Formelen for determinanten blir da

$$\det(A) = a_{11}C_{11} + a_{12}C_{12} + \dots + a_{1n}C_{1n}$$

Tallet C_{ij} som inngår i dette uttrykket kalles *kofaktoren* til a_{ij} , og er definert som

$$C_{ij} = (-1)^{i+j}M_{ij},$$

hvor M_{ij} er determinanten til matrisen vi får ved å ta bort i 'te rad og j 'te kolonne fra den opprinnelige matrisen A , altså raden og kolonnen som a_{ij} står i. Tallene M_{ij} kalles noen ganger for *minorer*.

Ovenfor har vi funnet et uttrykk for $\det(A)$ ved å utvikle determinanten langs første rad. Vi kunne funnet et tilsvarende uttrykk ved å utvikle determinanten langs en annen rad, eller langs en kolonne. I hvert tilfelle blir formelen gitt ved summen av alle tallene i raden eller kolonnen multiplisert med sine tilsvarende kofaktorer.

Legg også merke til at når mens $\det(A)$ er en $n \times n$ -determinant, så er hver av kofaktorene C_{ij} en $(n-1) \times (n-1)$ -determinant. Skal vi skrive opp determinanten til en 3×3 -matrise på denne måten, blir den redusert til tre 2×2 -determinanter. Så det blir fortsatt 6 ledd, men formelen er lettere å huske på denne måten.

Vi tar med et eksempel på utregning av determinanten til en 3×3 -matrise. La A være matrisen

$$A = \begin{pmatrix} 4 & 3 & 2 \\ 5 & 6 & 3 \\ 3 & 5 & 2 \end{pmatrix}$$

fra forrige avsnitt. Vi velger å utvikle determinanten til M langs første kolonne, og får at

$$\begin{aligned} \det(A) &= a_{11}C_{11} + a_{21}C_{21} + a_{31}C_{31} \\ &= 4 \cdot 1 \cdot M_{11} + 5 \cdot (-1) \cdot M_{21} + 3 \cdot 1 \cdot M_{31} \\ &= 4 \cdot 1 \cdot \det \begin{pmatrix} 6 & 3 \\ 5 & 2 \end{pmatrix} + 5 \cdot (-1) \cdot \det \begin{pmatrix} 3 & 2 \\ 5 & 2 \end{pmatrix} + 3 \cdot 1 \cdot \det \begin{pmatrix} 3 & 2 \\ 6 & 3 \end{pmatrix} \\ &= 4(6 \cdot 2 - 3 \cdot 5) - 5(3 \cdot 2 - 2 \cdot 5) + 3(3 \cdot 3 - 2 \cdot 6) \\ &= -12 + 20 - 9 \\ &= -1 \end{aligned}$$

Vi ser at vi finner determinanten $\det(A) = -1$ til 3×3 -matrisen A ved å utvikle den langs første kolonne og regne ut de tre 2×2 -determinantene som dette leder til.

Et nyttig triks for å lettest mulig regne ut en determinant, er å utvikle den langs en rad eller kolonne som inneholder mange nuller. Dette er nyttig fordi for hvert tall som er null i raden eller kolonnen, så blir det unødvendig å regne ut den tilsvarende kofaktoren (den skal jo multipliseres med null, og forsvinner derfor uansett).

La oss se på et eksempel til. Vi skal regne ut determinanten til matrisen M , som denne gang er en 4×4 -matrise gitt ved

$$M = \begin{pmatrix} 7 & 3 & -1 & 4 \\ 0 & 3 & -1 & 1 \\ 0 & 0 & -1 & 4 \\ 0 & 0 & 0 & 11 \end{pmatrix}$$

Matriser som har denne formen, der alt under hoved-diagonalen er null, kalles *øvre triangulære matriser* (mens matriser der alt over hoved-diagonalen er null kalles *nedre triangulære matriser*). Vi velger å utvikle $\det(M)$ langs første kolonne, siden denne kolonnen har mange nuller. Vi får da følgende resultat:

$$\det(M) = 7 \cdot \det \begin{pmatrix} 3 & -1 & 1 \\ 0 & -1 & 4 \\ 0 & 0 & 11 \end{pmatrix} + 0 \cdot C_{21} + 0 \cdot C_{31} + 0 \cdot C_{41}$$

Vi ser at de tre siste kofaktorene er unødvendige å regne ut, og det forenkler beregningen mye! Den første kofaktoren er en 3×3 -determinant. Igjen regner vi ut denne determinanten ved å utvikle den langs første kolonne, som inneholder mange nuller:

$$\det(M) = 7(3 \cdot \det \begin{pmatrix} -1 & 4 \\ 0 & 11 \end{pmatrix}) + 0 \cdot C_{21} + 0 \cdot C_{31}$$

Igjen ser vi at de to siste kofaktorene C_{21} og C_{31} i 3×3 -matrisen ikke er nødvendige å regne ut. Så vi sitter igjen med følgende forholdsvis enkle beregning:

$$\det(M) = 7 \cdot 3 \cdot ((-1) \cdot 11 - 4 \cdot 0) = 7 \cdot 3 \cdot (-1) \cdot 11 = -231$$

Vi kan konkludere at i matriser som inneholder mange nuller, så blir det mye enklere å regne ut determinanter — og det lønner seg å velge utviklingen av determinanten på en lur måte. Vi ser også at determinanten i dette eksempelet er produktet av tallene på diagonalen. Dette gjelder faktisk for alle triangulære matriser:

TEOREM 2. *La A være en kvadratisk matrise. Hvis A er øvre eller nedre triangulær, så er*

$$\det(A) = a_{11} \cdot a_{22} \cdots a_{nn}$$

altså produktet av elementene på hoved-diagonalen.

Metode 2: Determinanter ved hjelp av Gauss eliminasjon.

La A være en $n \times n$ -matrise. Hvis vi utfører Gauss eliminasjon på A , kommer vi fram til en trappeform, som vi kaller T . Determinanten til T er alltid lett å regne ut. Det skyldes at enhver trappeform som er kvadratisk må være øvre triangulær, og derfor kan vi i følge teoremet ovenfor regne ut $\det(T)$ som produktet av tallene på diagonalen til T . Men hva er forholdet mellom determinanten til den opprinnelige matrisen A og determinanten til trappeformen T ? Det kommer an på hvilke elementære radoperasjoner vi har utført for å komme fra A til T ! Vi har følgende tabell over effekten som hver av de elementære radoperasjonene har på determinanten:

- Bytte om to rader	- Determinanten skifter fortegn
- Legge til et multiplum av en rad til en annen rad	- Ingen endring av determinanten
- En rad multipliseres med et tall $c \neq 0$	- Determinanten multipliseres med tallet c

Vi får dermed formelen $\det(T) = \det(A) \cdot (-1)^r \cdot c_1 \cdots c_s$, der $\det(T)$ er determinanten til trappeformen vi kommer fram til, r er antall bytter av rader vi gjør i Gauss eliminasjonen, og c_1, \dots, c_s er de tallene vi multipliserer rader med i Gauss eliminasjonen. Legg merke til at vi bare skal ta med tallet c_i hvis vi multipliserer en rad med c_i — hvis vi

derimot tar et multiplum av en rad og legger til en annen rad trenger vi ikke å ta med dette i formelen ovenfor. Til slutt må vi løse ut $\det(A)$ fra denne likningen. Husk at det beste er om vi klarer å gjennomføre Gauss eliminasjonen uten å gjøre noen bytter av rader og uten å multiplisere noen rader med tall $c \neq 0$. Da blir $\det(A) = \det(T)$.

Vi skal se på et eksempel på bruk av denne metoden for å regne ut en determinant. La oss returnere til 3×3 -matrisen som vi så på tidligere, gitt ved

$$A = \begin{pmatrix} 4 & 3 & 2 \\ 5 & 6 & 3 \\ 3 & 5 & 2 \end{pmatrix}$$

Vi bruker nå Gauss eliminasjon for å redusere A til en trappeformen, og bruker færrest mulig operasjoner av typen bytte av rader og multiplikasjon av en rad med et tall $\neq 0$:

$$\begin{aligned} A &= \begin{pmatrix} 4 & 3 & 2 \\ 5 & 6 & 3 \\ 3 & 5 & 2 \end{pmatrix} \xrightarrow{R(1):=R(1)-R(2)} \begin{pmatrix} -1 & -3 & -1 \\ 5 & 6 & 3 \\ 3 & 5 & 2 \end{pmatrix} \xrightarrow{R(2):=R(2)+5R(1)} \\ &\begin{pmatrix} -1 & -3 & -1 \\ 0 & -9 & -2 \\ 3 & 5 & 2 \end{pmatrix} \xrightarrow{R(3):=R(3)+3R(1)} \begin{pmatrix} -1 & -3 & -1 \\ 0 & -9 & -2 \\ 0 & -4 & -1 \end{pmatrix} \xrightarrow{R(2):=R(2)-2R(3)} \\ &\begin{pmatrix} -1 & -3 & -1 \\ 0 & -1 & 0 \\ 0 & -4 & -1 \end{pmatrix} \xrightarrow{R(3):=R(3)-4R(2)} \begin{pmatrix} -1 & -3 & -1 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} = T \end{aligned}$$

Vi har altså kommet fram til en trappeform T med determinant

$$\det(T) = (-1)(-1)(-1) = -1$$

Siden vi ikke har brukt noen radbytter og ikke har multiplisert noen rader med tall $\neq 0$, har vi at $\det(A) = \det(T) = -1$. Vi ser at dette stemmer overens med resultatet vi fant ved hjelp av metode 1.

De to metodene for å regne ut determinanter som vi har beskrevet ovenfor har sine fordeler og ulemper. Hvis matrisene er forholdsvis små, så vil metode 1 ofte være grei å bruke. Men hvis vi skal regne ut determinanten til en stor matrise, så vil metode 2 lønne seg. Grovt regnet er antall operasjoner som skal til for å regne ut determinanten til en $n \times n$ -matrise med metode 1 i størrelsesorden $n!$ mens det med metode 2 er i størrelsesorden n^3 operasjoner. Og $n!$ blir svært mye større enn n^3 når n er stor. For eksempel er $100! \cong 9 \cdot 10^{157}$ mens $100^3 = 1000000 = 1 \cdot 10^6!$

Til slutt tar vi med litt om regneregler for determinanter. Hvis A, B er $n \times n$ -matriser og $c \in \mathbb{R}$ er et tall, så har vi at

- (1) $\det(AB) = \det(A)\det(B)$
- (2) $\det(A^{-1}) = 1/\det(A)$ hvis $\det(A) \neq 0$,
- (3) $\det(A^n) = \det(A)^n$
- (4) $\det(cA) = c^n \det(A)$

Legg spesielt merke til at det ikke er noen enkel formel for determinanten $\det(A + B)$ til en sum av matriser, og at $\det(cA) = c^n \det(A)$ (det vil si at vi får en faktor c for hver rad i matrisen A).

2.8. Bruk av determinanter

Den viktigste egenskapen til determinanten er at en $n \times n$ -matrise A er invertibel hvis og bare hvis $\det(A) \neq 0$, som nevnt i forrige avsnitt. Men determinanter kan også brukes i andre sammenhenger, og vi skal se på to viktige anvendelser av determinanten.

Cramers regel. La oss se på et lineært likningssystem $A\mathbf{x} = \mathbf{b}$ med n ukjente og n likninger. Da vet vi at systemet har en entydig løsning om $\det(A) \neq 0$. Cramers regel gjør oss i stand til å regne ut denne løsningen ved hjelp av determinanter.

La oss skrive opp koeffisientmatrisen A til likningssystemet som en matrise som består av n kolonne-vektorer,

$$A = (\mathbf{a}_1 \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_n)$$

Da sier *Cramers regel* at hvis $\det(A) \neq 0$, så har likningssystemet $A\mathbf{x} = \mathbf{b}$ en entydig løsning \mathbf{x} , der

$$x_i = \frac{\det(\mathbf{a}_1 \cdots \mathbf{a}_{i-1} \quad \mathbf{b} \quad \mathbf{a}_{i+1} \cdots \mathbf{a}_n)}{\det(A)}$$

for $1 \leq i \leq n$.

Husk på at selv om formelen i Cramers regel ser enkel ut, så må vi regne ut $n + 1$ $n \times n$ -determinanter for å finne løsningen, og det gir mye regning! La oss se på et eksempel med tre likninger og tre ukjente, nemlig likningssystemet

$$\begin{aligned} 3x - 8y + 10z &= 22 \\ x - 3y + 2z &= 5 \\ 4x - 9y - 8z &= -11 \end{aligned}$$

I dette eksempelet har vi

$$A = \begin{pmatrix} 3 & -8 & 10 \\ 1 & -3 & 2 \\ 4 & -9 & -8 \end{pmatrix} \quad \text{og} \quad \mathbf{b} = \begin{pmatrix} 22 \\ 5 \\ -11 \end{pmatrix}$$

I tillegg til $\det(A)$, må vi derfor også regne ut determinanten til de tre matrisene

$$A_1 = \begin{pmatrix} 22 & -8 & 10 \\ 5 & -3 & 2 \\ -11 & -9 & -8 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 3 & 22 & 10 \\ 1 & 5 & 2 \\ 4 & -11 & -8 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 3 & -8 & 22 \\ 1 & -3 & 5 \\ 4 & -9 & -11 \end{pmatrix}$$

for å finne løsningen av likningssystemet ved hjelp av Cramers regel, altså totalt fire 3×3 -determinanter. Dette blir mye arbeid! Etter å ha regnet lenge, finner vi ut at

$$\det(A) = 28, \quad \det(A_1) = 0, \quad \det(A_2) = -12, \quad \det(A_3) = 178$$

og dermed ser vi at likningssystemet har en entydig løsning

$$\begin{aligned}x &= \det(A_1)/\det(A) = 0/28 = 0 \\y &= \det(A_2)/\det(A) = -12/28 = -3/7 \\z &= \det(A_3)/\det(A) = 178/28 = 89/14\end{aligned}$$

ved hjelp av Cramers regel. Generelt har Cramers regel noen teoretiske fordeler, men Gauss eller Gauss-Jordan eliminasjon er en mye raskere metode, spesielt når likningssystemene blir store.

Adjungerte matriser. La A være en kvadratisk $n \times n$ -matrise. Da er *kofaktormatrisen* til A matrisen som består av alle kofaktorene til A , mer bestemt matrisen som har kofaktoren C_{ij} på plass (i, j) . Med andre ord er kofaktormatrisen (C_{ij}) til A gitt som matrisen

$$(C_{ij}) = \begin{pmatrix} C_{11} & C_{12} & \cdots & C_{1n} \\ C_{21} & C_{22} & \cdots & C_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ C_{n1} & C_{n2} & \cdots & C_{nn} \end{pmatrix}$$

Denne kofaktormatrisen er viktig fordi den er nært knyttet til den inverse matrisen til A , hvis den finnes. For å regne den ut, må vi altså regne ut en rekke med $(n-1) \times (n-1)$ -determinanter.

For å beskrive sammenhengen, må vi først definere transponerte matriser. Hvis A er en vilkårlig $m \times n$ -matrise, så definerer vi den *transponerte matrisen* A^t til å være $n \times m$ -matrisen vi får når vi spegler A om hoveddiagonalen. Med andre ord, hvis A er matrisen

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

så er den transponerte matrisen A^t gitt som

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

Altså er kolonnene i A blitt til rader i A^t , og radene i A er blitt til kolonner i A^t .

La oss ta med et eksempel på transponering av matriser. Hvis A er 3×3 -matrisen

$$A = \begin{pmatrix} 3 & -8 & 10 \\ 1 & -3 & 2 \\ 4 & -9 & -8 \end{pmatrix}$$

så blir den transponerte matrisen A^t matrisen

$$A^t = \begin{pmatrix} 3 & 1 & 4 \\ -8 & -3 & -9 \\ 10 & 2 & -8 \end{pmatrix}.$$

Vi tar også med noen regneregler for transponerte matriser. For vilkårlige matriser A, B med kompatible størrelser og for en vilkårlig skalar c har vi

- (1) $(A + B)^t = A^t + B^t$
- (2) $(cA)^t = cA^t$
- (3) $(A^t)^t = A$
- (4) $(AB)^t = B^t A^t$
- (5) $\det(A^t) = \det(A)$

Legg spesielt merke til den siste regneregelen. Den skyldes at når vi utvikler determinanten til A langs første rad, så får vi akkurat samme resultat som når vi utvikler determinanten til A^t langs første kolonne.

Vi kan nå formulere sammenhengen mellom den inverse matrisen til A (når den fins) og kofaktormatrisen til A . Hvis A er en invertibel $n \times n$ -matrise, så har vi at

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

der $\operatorname{adj}(A) = (C_{ij})^t$ er den transponerte matrisen til kofaktormatrisen til A . Matrisen $\operatorname{adj}(A)$ kalles *den adjungerte matrisen* til A .

La oss se på et eksempel på utregning av A^{-1} ved hjelp av kofaktormatriser. Igjen ser vi på eksempelet

$$A = \begin{pmatrix} 3 & -8 & 10 \\ 1 & -3 & 2 \\ 4 & -9 & -8 \end{pmatrix}$$

Vi vet at $\det(A) = 28$, og dermed er A invertibel. For å regne ut A^{-1} bruker vi formelen ovenfor, og vi må dermed regne ut alle kofaktorene til matrisen A . Det blir totalt 9 ulike 2×2 -determinanter, og litt regning gir at

$$(C_{ij}) = \begin{pmatrix} 42 & 16 & 3 \\ -154 & -64 & 59 \\ 14 & 4 & -1 \end{pmatrix}$$

Formelen for A^{-1} gir dermed at

$$A^{-1} = \frac{1}{\det(A)} \operatorname{adj}(A) = \frac{1}{28} \begin{pmatrix} 42 & -154 & 14 \\ 16 & -64 & 4 \\ 3 & 59 & -1 \end{pmatrix}$$

Generelt har formelen for A^{-1} noen teoretiske fordeler, men Gauss-Jordan eliminasjon av $(A|I_n)$ er en mye raskere metode for å finne A^{-1} , spesielt når matrisen A er stor.

KAPITTEL 3

Vektorrom

3.1. Vektorrommet \mathbb{R}^3

I seksjon ?? diskuterte vi det Euklidske rommet \mathbb{R}^3 , og det skal vi nå se litt nærmere på. Det er definert som

$$\mathbb{R}^3 = \{(a, b, c) : a, b, c \in \mathbb{R}\},$$

det vil si alle 3-tupler (a, b, c) slik at $a, b, c \in \mathbb{R}$ er reelle tall. Vi tenker gjerne på (a, b, c) som et punkt i det tre-dimensjonale rommet, med koordinater $x = a$, $y = b$, $z = c$. Men vi kan også tenke på (a, b, c) som en *vektor*, det vil si en pil med startpunkt i origo og endepunkt i punktet (a, b, c) . Vi skriver gjerne $\mathbf{v} = (a, b, c)$ når vi tenker på (a, b, c) som en vektor. Vi nevner også at vi noen ganger skriver en vektor som en kolonne-vektor,

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

Det er altså bare en annen måte å skrive vektoren $\mathbf{v} = (a, b, c)$ på.

På rommet \mathbb{R}^3 har vi to viktige operasjoner som vi beskrev i seksjon ??, nemlig vektor-addisjon og skalar-multiplikasjon. Vektor-addisjon beskriver hvordan man adderer sammen to vektorer i \mathbb{R}^3 , og skalar-multiplikasjon beskriver hvordan man multipliserer en skalar (et tall) med en vektor i \mathbb{R}^3 . Vi husker at disse operasjonene ble definert plass for plass, slik at

$$\begin{aligned} (u_1, u_2, u_3) + (v_1, v_2, v_3) &= (u_1 + v_1, u_2 + v_2, u_3 + v_3) \\ r(v_1, v_2, v_3) &= (rv_1, rv_2, rv_3) \end{aligned}$$

når $(u_1, u_2, u_3), (v_1, v_2, v_3) \in \mathbb{R}^3$ er vektorer og $r \in \mathbb{R}$ er en skalar. I begge operasjonene blir resultatet en vektor i \mathbb{R}^3 .

Vi kan også beskrive disse operasjonene *geometrisk*, når vektorene \mathbf{u} og \mathbf{v} er representert som piler. For å finne $\mathbf{u} + \mathbf{v}$, parallell-forskyver vi \mathbf{v} slik at denne pilen starter der \mathbf{u} slutter, og summen $\mathbf{u} + \mathbf{v}$ kan man da finne igjen som pilen som starter ved starten av \mathbf{u} og slutter i enden av \mathbf{v} . Skal vi finne $r\mathbf{v}$, så danner vi en vektor med samme retning som \mathbf{v} men som er strukket med en faktor $|r|$, og som er snudd hvis $r < 0$.

Det er ikke vanskelig (men litt tidkrevende) å sjekke at operasjonene oppfyller følgende regneregler for alle vektorer $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^3$ og alle

skalarer $r, s \in \mathbb{R}$:

$$\begin{aligned} \mathbf{u} + \mathbf{v} &= \mathbf{v} + \mathbf{u} \\ \mathbf{u} + (\mathbf{v} + \mathbf{w}) &= (\mathbf{u} + \mathbf{v}) + \mathbf{w} \\ \mathbf{u} + \mathbf{0} &= \mathbf{u} \\ \mathbf{u} + (-\mathbf{u}) &= \mathbf{0} \\ r(\mathbf{u} + \mathbf{v}) &= r\mathbf{u} + r\mathbf{v} \\ (r + s)\mathbf{u} &= r\mathbf{u} + s\mathbf{u} \\ r(s\mathbf{u}) &= (rs)\mathbf{u} \\ 1\mathbf{u} &= \mathbf{u} \end{aligned}$$

Her er $\mathbf{0} = (0, 0, 0)$ og $-\mathbf{u} = (-1)\mathbf{u}$. Altså er det Euklidske rommet \mathbb{R}^3 sammen med operasjonene vektor-addisjon og skalar-multiplikasjon et *vektorrom*.

La oss for et øyeblikk forlate \mathbb{R}^3 , og se på det Euklidske rommet \mathbb{R}^2 . Det består av alle punkter $(a, b) \in \mathbb{R}^2$ i planet, og disse kan oppfattes som vektorer i planet akkurat som punktene $(a, b, c) \in \mathbb{R}^3$ kan oppfattes som vektorer i rommet. Vi definerte vektor-addisjon og skalar-multiplikasjon for vektorer i \mathbb{R}^2 i seksjon ??, og man kan sjekke at disse operasjonene på \mathbb{R}^2 oppfyller akkurat samme regneregler som de tilsvarende operasjonene på \mathbb{R}^3 . Det betyr at \mathbb{R}^2 med operasjonene vektor-addisjon og skalar-multiplikasjon er et vektorrom.

La $\mathbf{u}, \mathbf{v}, \mathbf{w}$ være tre vektorer i \mathbb{R}^3 . Vi definerer en *lineær-kombinasjon* av vektorene $\mathbf{u}, \mathbf{v}, \mathbf{w}$ til å være et uttrykk av formen

$$a\mathbf{u} + b\mathbf{v} + c\mathbf{w}$$

for skalarer $a, b, c \in \mathbb{R}$. Hvis vi skriver $\mathbf{u} = (u_1, u_2, u_3)$, $\mathbf{v} = (v_1, v_2, v_3)$ og $\mathbf{w} = (w_1, w_2, w_3)$, så blir denne lineær-kombinasjonen

$$a\mathbf{u} + b\mathbf{v} + c\mathbf{w} = (au_1 + bv_1 + cw_1, au_2 + bv_2 + cw_2, au_3 + bv_3 + cw_3)$$

En *relasjon* er en slik lineær-kombinasjon som er null, dvs

$$a\mathbf{u} + b\mathbf{v} + c\mathbf{w} = \mathbf{0}$$

Uansett hvordan vektorene $\mathbf{u}, \mathbf{v}, \mathbf{w}$ ser ut, så er $0\mathbf{u} + 0\mathbf{v} + 0\mathbf{w} = \mathbf{0}$, og dette kalles en *triviell relasjon*. Men spørsmålet er om det fins noen andre relasjoner.

Hvis det finnes en ikke-triviell relasjon mellom $\mathbf{u}, \mathbf{v}, \mathbf{w}$, altså en relasjon

$$a\mathbf{u} + b\mathbf{v} + c\mathbf{w} = \mathbf{0}$$

slik at minst ett av tallene $a, b, c \in \mathbb{R}$ er forskjellig fra null, så sier vi at $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ er *lineært avhengige vektorer*. Alle ikke-trivielle relasjoner kalles i så fall avhengighetsrelasjoner.

For eksempel så fins det en ikke-triviell relasjon mellom vektorene $\mathbf{u} = (1, 2, -3)$, $\mathbf{v} = (3, 1, -2)$ og $\mathbf{w} = (5, -5, 6)$. I dette eksempelet er

det nemlig slik at

$$4\mathbf{u} - 3\mathbf{v} + \mathbf{w} = \mathbf{0}$$

(sjekk dette!). Altså er vektorene $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ lineært avhengige vektorer i \mathbb{R}^3 , og relasjonen ovenfor er en avhengighetsrelasjon. Men det fins mange andre slike relasjoner, for eksempel er $12\mathbf{u} - 9\mathbf{v} + 3\mathbf{w} = \mathbf{0}$.

Hvis derimot den trivielle relasjonen er den eneste relasjonen vi kan finne mellom vektorene $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$, så sier vi at disse vektorene er *lineært uavhengige vektorer*. For eksempel er vektorene $\mathbf{u} = (1, 3, 2)$, $\mathbf{v} = (2, 8, 7)$ og $\mathbf{w} = (1, 7, 9)$ lineært uavhengige vektorer. Det er fordi det ikke finnes noen ikke-trivielle relasjoner

$$a\mathbf{u} + b\mathbf{v} + c\mathbf{w} = \mathbf{0}$$

altså relasjoner der minst ett av tallene a, b, c er forskjellig fra null.

Men hvordan kan vi egentlig finne ut om tre vektorer er lineært uavhengi eller ikke? Det korte svaret på dette spørsmålet er gitt av følgende resultat:

TEOREM 3. *La $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^3$ være tre vektorer i \mathbb{R}^3 , og la A være matrisen som har disse vektorene som kolonner. Da er $\det(A) \neq 0$ hvis $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ er lineært uavhengige, og $\det(A) = 0$ hvis $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ er lineært avhengige.*

Vi skal se litt nærmere på hvorfor det er slik, og også hvordan vi kan finne alle avhengighetsrelasjonene hvis vektorene er lineært avhengige. Hvis vi skriver $\mathbf{u} = (u_1, u_2, u_3)$, $\mathbf{v} = (v_1, v_2, v_3)$ og $\mathbf{w} = (w_1, w_2, w_3)$, så er matrisen A som har disse tre vektorene som kolonner gitt ved

$$A = \begin{pmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{pmatrix}$$

og vi ser at

$$\begin{pmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = a \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} + b \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} + c \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = a\mathbf{u} + b\mathbf{v} + c\mathbf{w}$$

Hvis vi skriver $\mathbf{x} = (a, b, c)$, så er det å jakte på relasjoner det samme som å løse det lineære likningssystemet $A\mathbf{x} = \mathbf{0}$. Dette er et homogent likningssystem, som alltid har den trivielle løsningen $\mathbf{x} = \mathbf{0}$, det vil si $(a, b, c) = (0, 0, 0)$. Denne løsningen svarer til den trivielle relasjonen mellom vektorene.

Vi vet altså at hvis $\det(A) \neq 0$ så har likningssystemet kun denne ene løsningen, og altså kun den trivielle relasjonen. Så i dette tilfellet er \mathbf{u}, \mathbf{v} og \mathbf{w} lineært uavhengige vektorer. Vi vet også at hvis $\det(A) = 0$, så har systemet uendelig mange løsninger, som vi kan beskrive ved hjelp av en eller flere frie parametre. Altså fins det uendelig mange relasjoner, og det betyr at \mathbf{u}, \mathbf{v} og \mathbf{w} er lineært avhengige vektorer. Hvis vi vil finne avhengighetsrelasjonene, så må vi løse likningssystemet $A\mathbf{x} = \mathbf{0}$.

La oss gå tilbake til eksempelet der $\mathbf{u} = (1, 2, -3)$, $\mathbf{v} = (3, 1, -2)$ og $\mathbf{w} = (5, -5, 6)$. I dette tilfellet er

$$A = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 1 & -5 \\ -3 & -2 & 6 \end{pmatrix}$$

Vi får dermed at $\det(A) = 0$, og dermed vet vi at de tre vektorene har uendelig mange avhengighetsrelasjoner og dermed er lineært avhengige. For å finne avhengighetsrelasjonene, løser vi likningssystemet $A\mathbf{x} = \mathbf{0}$. Vi bruker Gauss-Jordan eliminasjon for å finne den reduserte trappeformen til A , som er

$$\begin{pmatrix} 1 & 0 & -4 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix}$$

og dette betyr at løsningene er gitt ved $a = 4s$, $b = -3s$ og $c = s$, eller $(a, b, c) = s \cdot (4, -3, 1)$. Vi ser at $s = 0$ gir den trivielle løsningen, mens alle andre verdier av s gir avhengighetsrelasjoner. For eksempel gir $s = 1$ at $4\mathbf{u} - 3\mathbf{v} + \mathbf{w} = \mathbf{0}$, og $s = 3$ gir $12\mathbf{u} - 9\mathbf{v} + 3\mathbf{w} = \mathbf{0}$.

Vi definerer en *basis for* \mathbb{R}^3 til å være en samling av tre vektorer som er lineært uavhengige. Det finnes mange forskjellige basiser for \mathbb{R}^3 , for eksempel vektorene $\mathbf{u} = (1, 3, 2)$, $\mathbf{v} = (2, 8, 7)$ og $\mathbf{w} = (1, 7, 9)$. Men den enkleste basisen for \mathbb{R}^3 , som ofte kalles *standard basis*, består av vektorene $\mathbf{i} = (1, 0, 0)$, $\mathbf{j} = (0, 1, 0)$ og $\mathbf{k} = (0, 0, 1)$.

Enhver basis for \mathbb{R}^3 er slik at enhver vektor i \mathbb{R}^3 kan skrives som en lineær-kombinasjon av basis-vektorene. For eksempel kan vektoren $\mathbf{t} = (7, 5, 3)$ skrives som lineær-kombinasjonen

$$\mathbf{t} = (7, 5, 3) = 7\mathbf{i} + 5\mathbf{j} + 3\mathbf{k}$$

av vektorene i standard basis. Siden vektorene $\mathbf{u} = (1, 3, 2)$, $\mathbf{v} = (2, 8, 7)$ og $\mathbf{w} = (1, 7, 9)$ også danner en basis for \mathbb{R}^3 , må \mathbf{t} også kunne skrives som en lineær-kombinasjon av disse vektorene. For å finne hvilken lineær-kombinasjon dette er, må vi løse likningssystemet

$$\begin{pmatrix} 1 & 2 & 1 \\ 3 & 8 & 7 \\ 2 & 7 & 9 \end{pmatrix} \mathbf{x} = \mathbf{t}$$

Siden vektorene er lineært uavhengige, har koeffisientmatrisen determinant forskjellig fra null, og systemet har en entydig løsning. Vi kan for eksempel finne denne løsningen ved hjelp av Gauss-Jordan eliminasjon. Dette gir $\mathbf{x} = (62, -34, 13)$, og altså kan \mathbf{t} skrives som lineær-kombinasjonen

$$\mathbf{t} = (7, 5, 3) = 62\mathbf{u} - 34\mathbf{v} + 13\mathbf{w}$$

3.2. Lineære underrom av \mathbb{R}^3

La $V \subseteq \mathbb{R}^3$ være en delmengde av \mathbb{R}^3 . Altså er V en samling av vektorer fra \mathbb{R}^3 . Vi sier at V er et *lineært underrom* hvis følgende betingelser er oppfylt:

- (1) Hvis \mathbf{u} og \mathbf{v} er i V , så er $\mathbf{u} + \mathbf{v}$ i V
- (2) Hvis \mathbf{u} er i V og $r \in \mathbb{R}$ er en skalar, så er $r\mathbf{u}$ i V

Vi ser med en gang at $V = \mathbf{0}$, delmengden som kun består av vektoren $\mathbf{0}$, er et lineært underrom av \mathbb{R}^3 , og at $V = \mathbb{R}^3$, hele \mathbb{R}^3 , også er et lineært underrom av \mathbb{R}^3 . Det er det minste og det største lineære underrom av \mathbb{R}^3 .

Hvilke andre lineære underrom av $V \subseteq \mathbb{R}^3$ finnes der? Det er to muligheter:

- (1) V inneholder to lineært uavhengige vektorer
- (2) V inneholder ikke to lineært uavhengige vektorer

Siden $V \neq \mathbf{0}$, kan vi alltid finne en vektor $\mathbf{v} \neq \mathbf{0}$ i V . I det andre tilfellet er alle andre vektorer i V et multippel av denne vektoren, og derfor ligger alle vektorene i V langs linjen gjennom origo utspent av vektoren \mathbf{v} . Vi sier da at V er en linje i \mathbb{R}^3 .

I det første tilfelle kan vi finne en annen vektor \mathbf{w} i V slik at \mathbf{v} og \mathbf{w} er lineært uavhengige. Alle andre vektorer i V vil da være en lineærkombinasjon av disse to vektorene, og derfor ligger alle vektorene i V i et plan gjennom origo utspent av \mathbf{v} og \mathbf{w} . Vi sier da at V er et plan i \mathbb{R}^3 .

THEOREM 4. *La $A\mathbf{x} = \mathbf{0}$ være et homogent lineært likningssystem med tre ukjente. Da er løsningsmengden $\{\mathbf{x} \in \mathbb{R}^3 : A\mathbf{x} = \mathbf{0}\}$ et lineært underrom av \mathbb{R}^3 .*

3.3. Generelle vektorrom

Et *vektorrom* består av en mengde V av vektorer, og to operasjoner på V som kalles *vektor-addisjon* og *skalar-multiplikasjon*. Vektor-addisjon er en operasjon der vi summerer to vektorer fra V , og skalar-multiplikasjon er en operasjon hvor vi multipliserer en skalar (et tall) med en vektor fra V , og begge operasjonene gir en ny vektor i V som svar. I tillegg kreves det at disse to operasjonene oppfyller følgende

betingelser:

$$\begin{aligned} \mathbf{u} + \mathbf{v} &= \mathbf{v} + \mathbf{u} \\ \mathbf{u} + (\mathbf{v} + \mathbf{w}) &= (\mathbf{u} + \mathbf{v}) + \mathbf{w} \\ \mathbf{u} + \mathbf{0} &= \mathbf{u} \\ \mathbf{u} + (-\mathbf{u}) &= \mathbf{0} \\ r(\mathbf{u} + \mathbf{v}) &= r\mathbf{u} + r\mathbf{v} \\ (r + s)\mathbf{u} &= r\mathbf{u} + s\mathbf{u} \\ r(s\mathbf{u}) &= (rs)\mathbf{u} \\ 1\mathbf{u} &= \mathbf{u} \end{aligned}$$

for alle vektorer $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ og alle skalarer $r, s \in \mathbb{R}$. Men det er *ikke* noe krav at V skal bestå av vanlige vektorer, eller at operasjonene skal være definert ved hjelp av vanlig (komponentvis) addisjon og skalar-multiplikasjon.

Eksempler på vektorrom er \mathbb{R}^2 og \mathbb{R}^3 , som vi allerede har sett på, og mer generelt så er \mathbb{R}^n (med de vanlige operasjonene) et vektorrom for $n = 1, 2, 3, 4, \dots$. Vi se på \mathbb{R}^n som vektorrom i neste seksjon.

Men før vi kommer til vektorrommet \mathbb{R}^n , så skal vi se på et annet eksempel på et vektorrom som er av en helt annen type. La oss skrive $C(\mathbb{R})$ for mengden av alle kontinuerlige funksjoner $f : \mathbb{R} \rightarrow \mathbb{R}$, det vil si alle funksjoner $f = f(x)$ i en variabel som er definert og kontinuerlig for alle verdier av x .

For å se at $C(\mathbb{R})$ er et vektorrom, må vi definere to operasjoner på mengden $C(\mathbb{R})$ som svarer til addisjon og skalar-multiplikasjon av vanlige vektorer. For addisjon velger vi å addere sammen to funksjoner, slik at hvis for eksempel $f(x) = x^2$ og $g(x) = \sin(x)$, så er funksjonen $f + g$ gitt ved $(f + g)(x) = x^2 + \sin(x)$. For skalarmultiplikasjon velger vi å multiplisere en skalar med en funksjon, slik at hvis $r = 3$ og $f(x) = x^2$, så er rf funksjonen gitt ved $(rf)(x) = 3x^2$. Det viser seg at disse operasjonene oppfyller alle kravene ovenfor, så derfor er $C(\mathbb{R})$ et vektorrom. Vektorrom av denne typen kalles et *funksjonsrom*.

3.4. Vektorrommet \mathbb{R}^n

I seksjon ?? diskuterte vi det Euklidske rommet \mathbb{R}^n når n et positivt heltall, og det skal vi nå se litt nærmere på. Det er definert som

$$\mathbb{R}^n = \{(u_1, u_2, \dots, u_n) : u_1, u_2, \dots, u_n \in \mathbb{R}\},$$

det vil si alle n -tupler (u_1, u_2, \dots, u_n) slik at hver komponent $u_i \in \mathbb{R}$ er et reelt tall.

Det er lett å tenke seg at et tallpar (a, b) i \mathbb{R}^2 representerer et punkt i planet med koordinater $x = a$ og $y = b$, og tilsvarende at et trippel (a, b, c) i \mathbb{R}^3 representerer et punkt i rommet med koordinater $x = a, y = b$, og $z = c$. Men hvis for eksempel $n = 5$, så er det

vanskligere å forestille seg et 5-tupplel $(u_1, u_2, u_3, u_4, u_5)$ *geometrisk* som et punkt i det 5-dimensjonale rommet. Å regne med 5-tupler *algebraisk* er likevel ikke så mye vanskeligere enn tallpar eller tripler, for eksempel er det ikke noe problem for oss å legge sammen to 5-tupler, vi gjør det plass for plass akkurat som for tripler i \mathbb{R}^3 .

Det er ofte svært nyttig å bruke \mathbb{R}^n til å modellere ting som avhenger av n parametre, ofte flere enn tre. For eksempel kunne vi tenke oss at en majones-produsent vil produsere en ny og revolusjonerende type majones. Blant 100 forskjellige stoffer som det er aktuelt å ha i den nye majonesen, vil produsenten finne den kombinasjonen som folk liker best. Vi kan da tenke oss at et 100-tupplel (u_1, u_2, \dots, u_n) representerer en bestemt måte å sette sammen majonesen på, der man tar med u_1 gram av stoff nummer 1, u_2 gram av stoff nummer 2, og så videre. Dermed representerer vektorene i vektorrommet \mathbb{R}^{100} alle de mulige sammensetningene av den nye typen majones. Selv om det ikke er mulig å forestille seg punkter i et 100-dimensjonalt rom, så kan man bruke en slik representasjon for å regne på forskjellige typer majoneser, eller andre fenomener som avhenger av mange parametre!

Men tilbake til det Euklidske rommet \mathbb{R}^n , som altså består av alle n -tupler (u_1, \dots, u_n) med $u_i \in \mathbb{R}$. Vi kan i prinsippet tenke på et slikt tupplel som en vektor, det vil si en pil fra origo til punktet (u_1, \dots, u_n) i \mathbb{R}^n , selv dette er vanskelig å forestille seg geometrisk. Akkurat som for \mathbb{R}^3 , kan vi også skrive en vektor i \mathbb{R}^n både som et n -tupplel (u_1, \dots, u_n) , eller som en kolonnevektor.

Vi definerte operasjonene vektor-addisjon og skalar-multiplikasjon for \mathbb{R}^n i seksjon ???. Vektor-addisjon ble definert plass for plass, slik at

$$(u_1, u_2, \dots, u_n) + (v_1, v_2, \dots, v_n) = (u_1 + v_1, u_2 + v_2, \dots, u_n + v_n)$$

for alle vektorer $\mathbf{u} = (u_1, u_2, \dots, u_n)$ og $\mathbf{v} = (v_1, v_2, \dots, v_n)$ i \mathbb{R}^n . Tilsvarende ble skalar-multiplikasjon definert plass for plass, slik at

$$r(u_1, u_2, \dots, u_n) = (ru_1, ru_2, \dots, ru_n)$$

for alle skalarer $r \in \mathbb{R}$ og alle vektorer $\mathbf{u} = (u_1, u_2, \dots, u_n)$ i \mathbb{R}^n . Man kan, akkurat som for \mathbb{R}^3 , sjekke at disse operasjonene tilfredsstill de åtte regnereglerne i definisjonen av vektorrom. Dette betyr altså at \mathbb{R}^n er et vektorrom.

Vi husker at vi definerte begrepene *lineær uavhengighet* for en mengde av tre vektorer i vektorrommet \mathbb{R}^3 , og også begrepet *basis*. Nå skal vi gjøre dette litt mer generelt, for m vektorer i \mathbb{R}^n .

Gitt en mengde $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m$ av m vektorer i \mathbb{R}^n . Vi kaller et uttrykk på formen

$$a_1 \mathbf{u}_1 + a_2 \mathbf{u}_2 + \dots + a_m \mathbf{u}_m,$$

der $a_1, \dots, a_m \in \mathbb{R}$ er skalarer, for en *lineær-kombinasjon* av vektorene $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m$. En *relasjon* er en slik lineær-kombinasjon som er nullvektoren, det vil si en lineær-kombinasjon slik at

$$a_1 \mathbf{u}_1 + a_2 \mathbf{u}_2 + \dots + a_m \mathbf{u}_m = \mathbf{0}$$

Den trivielle relasjonen $0\mathbf{u}_1 + 0\mathbf{u}_2 + \dots + 0\mathbf{u}_m = \mathbf{0}$ finnes alltid, og alle andre relasjoner (det vil si de relasjonene der minst en $a_i \neq 0$) kalles ikke-trivielle.

Vektorene $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m$ er lineært avhengige hvis en av vektorene kan skrives som en lineær-kombinasjon av de andre. En bedre måte å uttrykke dette på, er å si at vektorene $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m$ er *lineært avhengige* hvis det finnes minst en ikke-triviell relasjon mellom dem. Hvis det ikke finnes noen ikke-trivielle relasjoner, så er $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m$ *lineært uavhengige* vektorer.

THEOREM 5. La $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m$ være m vektorer i \mathbb{R}^n , og la A være matrisen

$$A = \begin{pmatrix} u_{11} & u_{12} & \dots & u_{1m} \\ u_{21} & u_{22} & \dots & u_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ u_{n1} & u_{n2} & \dots & u_{nm} \end{pmatrix}$$

med disse vektorene som kolonner. Vektorene $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m$ er lineært uavhengige hvis likningssystemet $A\mathbf{x} = \mathbf{0}$ har en entydig løsning, og lineært avhengige hvis systemet har uendelig mange løsninger.

Hvis $m = n$, slik at matrisen A blir kvadratisk, så ser vi at systemet $A\mathbf{x} = \mathbf{0}$ har en entydig løsning hvis $\det(A) \neq 0$, og at det har uendelig mange løsninger hvis $\det(A) = 0$. Derfor er $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m$ lineært uavhengige hvis $\det(A) \neq 0$ og lineært avhengige hvis $\det(A) = 0$. Men husk at dette kriteriet kun kan brukes om n vektorer i \mathbb{R}^n , og altså ikke kan brukes om $m \neq n$. I så fall må vi se på likningssystemet $A\mathbf{x} = \mathbf{0}$.

Et annet viktig begrep er *det lineære spennet* eller *spennet* til vektorene $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m$. Dette er mengden av alle vektorer \mathbf{v} som kan skrives som en lineær-kombinasjon

$$\mathbf{v} = a_1 \mathbf{u}_1 + a_2 \mathbf{u}_2 + \dots + a_m \mathbf{u}_m$$

av vektorene $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m$. Hvis enhver vektor i \mathbb{R}^n kan skrives som en lineær-kombinasjon av vektorene $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m$, så sier vi at disse vektorene *utspenner* vektorrommet \mathbb{R}^n .

Vi definerer en *basis* for vektorrommet \mathbb{R}^n til å være en mengde $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m$ av vektorer i \mathbb{R}^n som både er lineært uavhengige og som utspenner hele \mathbb{R}^n . I så fall kan enhver vektor \mathbf{v} i \mathbb{R}^n skrives som

$$\mathbf{v} = a_1 \mathbf{u}_1 + a_2 \mathbf{u}_2 + \dots + a_m \mathbf{u}_m$$

på en entydig måte.

THEOREM 6. La $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m$ være m vektorer i \mathbb{R}^n .

- (1) Hvis $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m$ er lineært uavhengige vektorer i \mathbb{R}^n , så er $m \leq n$.
- (2) Hvis $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m$ er vektorer som utspenner hele \mathbb{R}^n , så er $m \geq n$.
- (3) Hvis $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m$ er en basis for \mathbb{R}^n , så er $m = n$.

Dette teoremet sier altså at enhver basis for \mathbb{R}^n består av akkurat n vektorer. Følgende regel er den enkleste å bruke hvis vi skal finne en basis for \mathbb{R}^n , eller sjekke om en oppgitt mengde er en basis. **En mengde vektorer i \mathbb{R}^n danner en basis hvis følgende krav er oppfylt:**

- (1) Mengden består av n vektorer i \mathbb{R}^n .
- (2) Vektorene er lineært uavhengige.

Vi kaller antallet vektorer i en basis for \mathbb{R}^n for *dimensjonen* til vektorrommet \mathbb{R}^n . Altså er dimensjonen til \mathbb{R}^3 lik 3 og dimensjonen til \mathbb{R}^n lik n . Vi skriver dette som $\dim(\mathbb{R}^3) = 3$ og $\dim(\mathbb{R}^n) = n$.

3.5. Lineære underrom av \mathbb{R}^n

La $V \subseteq \mathbb{R}^n$ være en delmengde av \mathbb{R}^n . Altså er V en samling av vektorer fra \mathbb{R}^n . Vi sier at V er et *lineært underrom* hvis følgende betingelser er oppfylt:

- (1) Hvis \mathbf{u} og \mathbf{v} er i V , så er $\mathbf{u} + \mathbf{v}$ i V
- (2) Hvis \mathbf{u} er i V og $r \in \mathbb{R}$ er en skalar, så er $r\mathbf{u}$ i V

Vi ser med en gang at $V = \mathbf{0}$, delmengden som kun består av vektoren $\mathbf{0}$, er et lineært underrom av \mathbb{R}^n , og at $V = \mathbb{R}^n$ også er et lineært underrom av \mathbb{R}^n . Disse underrommene er henholdsvis det minste og det største lineære underrommet av \mathbb{R}^n . Men det finnes en rekke andre lineære underrom V av \mathbb{R}^n .

Ethvert lineært underrom V av \mathbb{R}^n er et vektorrom. Det følger nemlig av definisjonen av lineære underrom at V arver operasjonene vektoraddisjon og skalar-multiplikasjon fra \mathbb{R}^n , og disse operasjonene oppfyller derfor automatisk de åtte regnereglene fra definisjonen av vektorrom.

La V være et underrom av \mathbb{R}^n . En *basis* for V er en mengde vektorer $\{\mathbf{u}_1, \dots, \mathbf{u}_r\}$ i V slik at

- (1) $\{\mathbf{u}_1, \dots, \mathbf{u}_r\}$ utspenner V ,
- (2) $\{\mathbf{u}_1, \dots, \mathbf{u}_r\}$ er lineært uavhengige.

Dimensjonen til underrommet V er antall vektorer i en basis, som vi har kalt r ovenfor, og vi skriver $\dim(V)$ for dette antallet. For \mathbb{R}^n vet vi at $\dim(\mathbb{R}^n) = n$ allerede før vi begynner å regne, og vi har sett hvordan dette gjør det enklere å sjekke om en mengde vektorer danner en basis. Men hvordan kan vi vite hva $\dim(V)$ er, altså hva som er det riktige antallet vektorer i en basis for et lineært underrom V ?

I praksis er enhver delmengde V av \mathbb{R}^n som er definert ved en eller flere førstegrads-likninger uten konstantledd et lineært underrom. Dette kan vi uttrykke ved hjelp av matriser:

TEOREM 7. *La $A\mathbf{x} = \mathbf{0}$ være et homogent lineært likningssystem med n ukjente. Da er løsningsmengden $V = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = \mathbf{0}\}$ et lineært underrom av \mathbb{R}^n .*

Underrommet V består altså av alle løsningene $\mathbf{x} = (x_1, x_2, \dots, x_n)$ i \mathbb{R}^n av $A\mathbf{x} = \mathbf{0}$. Vi vet at hvis vi løser likningssystemet ved hjelp av Gauss eller Gauss-Jordan eliminasjon, så vil trappeformen til A ha ledende koeffisienter i noen av kolonnene, mens de resterende kolonnene svarer til frie parametre. Vi definerer *rang*en til matrisen A til å være antall ledende koeffisienter i trappeformen til A , og vi skriver $\text{rk}(A)$ for dette tallet (rk står for *rank* på engelsk). Da blir antall frie parametre i løsningsmengden $n - \text{rk}(A)$. Vi får derfor følgende formel:

TEOREM 8. *La $V = \{\mathbf{x} : A\mathbf{x} = \mathbf{0}\}$ være et lineært underrom av \mathbb{R}^n . Da er $\dim(V) = n - \text{rk}(A)$.*

La oss se på et eksempel. Vi tar for oss det lineære likningssystemet som skrives $A\mathbf{x} = \mathbf{0}$ på matriseform, og som er

$$\begin{array}{cccccc} x_1 & + & x_2 & + & x_3 & + & x_4 & = & 0 \\ x_1 & + & 2x_2 & & & + & 5x_4 & = & 0 \\ 3x_1 & + & 2x_2 & + & 4x_3 & - & x_4 & = & 0 \end{array}$$

når vi skriver det ut. Mengden $V = \{\mathbf{x} \in \mathbb{R}^4 : A\mathbf{x} = \mathbf{0}\}$ av løsninger av dette systemet er en delmengde av \mathbb{R}^4 , og man kan sjekke at dette er et lineært underrom (gjør det!). Legg merke til at hver av de fire likningene er førstegrads-likninger uten konstantledd. La oss regne ut hvordan V ser ut. Da trenger vi å løse systemet, og vi kan for eksempel bruke Gauss-Jordan eliminasjon til å finne den reduserte trappeformen til den augmenterte matrisen,

$$\left(\begin{array}{cccc|c} 1 & 0 & 2 & -3 & 0 \\ 0 & 1 & -1 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right)$$

Siden trappeformen har to ledende koeffisienter, er $\text{rk}(A) = 2$, og dette gir $\dim(V) = 4 - 2 = 2$. Vi vet altså at en basis må bestå av to vektorer.

La oss finne en basis ved regning. Fra trappeformen ser vi at $x_3 = s$ og $x_4 = t$ er frie parametre, og vi får at $x_1 = -2s + 3t$ og $x_2 = s - 4t$. Vi skriver løsningen på vektorform som

$$\begin{aligned} \mathbf{x} &= (x_1, x_2, x_3, x_4) = (-2s + 3t, s - 4t, s, t) \\ &= s \cdot (-2, 1, 1, 0) + t \cdot (3, -4, 0, 1) \end{aligned}$$

Dette betyr at vektorrommet V av løsninger av dette likningssystemet er alle mulige lineær-kombinasjoner av løsningene $\mathbf{u}_1 = (-2, 1, 0, 1)$ og $\mathbf{u}_2 = (3, -4, 0, 1)$. Vektorene \mathbf{u}_1 og \mathbf{u}_2 utspenner da per definisjon V ,

og de er også lineært uavhengige (sjekk det!). Altså er $\{\mathbf{u}_1, \mathbf{u}_2\}$ en basis for V , og $\dim(V) = 2$.

Det er verdt å merke seg at når vi finner løsningene av et homogent lineært likningsystem ved hjelp av Gauss eller Gauss-Jordan eliminasjon, slik som vi har gjort ovenfor, så blir de enkelte løsningene man finner alltid lineært uavhengige. Det er dette som gjør at dimensjonen til V er antallet frie parametre.

KAPITTEL 4

Eigenverdi-problemer

4.1. Lineære transformasjoner

Vanlige funksjoner i en eller flere variable er vi vant til å skrive som $f(x) = 2x - 4$ eller $g(x, y) = x^2 + y^2$. Når vi skal få fram at $f = f(x)$ er en funksjon i en variabel og at $g = g(x, y)$ er en funksjon i to variable, skriver vi ofte $f : \mathbb{R} \rightarrow \mathbb{R}$ og $g : \mathbb{R}^2 \rightarrow \mathbb{R}$. Skrivemåten $f : \mathbb{R}^n \rightarrow \mathbb{R}$ betyr altså at $f = f(x_1, \dots, x_n)$ er en funksjon i n variable og at funksjonsverdien $y = f(x_1, \dots, x_n) \in \mathbb{R}$ er et reelt tall for ethvert n -tupple $(x_1, \dots, x_n) \in \mathbb{R}^n$.

Mer generelt kan vi se på *transformasjoner* $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$. Et eksempel på en slik transformasjon er $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ gitt ved

$$T(x, y) = (x, -y)$$

Legg merke til at funksjonsverdien i dette tilfellet ikke lenger er et reelt tall, men et tallpar i \mathbb{R}^2 . Transformasjonen sender i dette tilfellet en vektor i \mathbb{R}^2 til en annen vektor i \mathbb{R}^2 , nemlig speilingen av den opprinnelige vektoren om x -aksen.

Vi definerer en *lineær transformasjon* $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ til å være en transformasjon som oppfyller følgende betingelser:

- (1) $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$,
- (2) $T(r\mathbf{u}) = rT(\mathbf{u})$.

I praksis er $T(x_1, \dots, x_n) = (f_1(x_1, \dots, x_n), \dots, f_m(x_1, \dots, x_n))$ gitt ved m funksjoner f_1, \dots, f_m . I eksempelet $T(x, y) = (x, -y)$ er disse funksjonene $f_1(x, y) = x$ og $f_2(x, y) = -y$. Betingelsene (1) og (2) for at en transformasjon er lineær betyr at alle funksjonene f_1, \dots, f_m er førstegradsfunksjoner (lineære funksjoner) uten konstantledd.

I eksempelet $T(x, y) = (x, -y)$ er begge funksjonene førstegradsfunksjoner uten konstantledd, så speilingen $T(x, y) = (x, -y)$ er en lineær transformasjon. Transformasjonen $S(x, y) = (x - y + 1, x + 4y)$ er derimot ikke lineær fordi $x - y + 1$ har konstantleddet 1, mens $R(x, y) = (x - y, x^2 + y)$ ikke er lineær fordi $x^2 + y$ ikke er en førstegradsfunksjon.

Vi ser at hvis $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ er en generell lineær transformasjon, så er den gitt ved funksjoner f_1, \dots, f_m som har formen

$$\begin{aligned} f_1(x_1, \dots, x_n) &= a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ f_2(x_1, \dots, x_n) &= a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ &\vdots \\ f_m(x_1, \dots, x_n) &= a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \end{aligned}$$

siden hver av disse funksjonene altså må være førstegradsfunksjoner uten konstantledd. En lineær transformasjon er det derfor gitt ved en $m \times n$ -matrise A , gitt ved

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

som kalles matrisen til lineær transformasjonen $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$. For eksempel er

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

matrisen til speilingen $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ gitt ved $T(x, y) = (x, -y)$.

Det finnes endel viktige typer lineær transformasjoner i \mathbb{R}^2 og \mathbb{R}^3 som er spesielt viktige og som man kan forestille seg geometrisk. Vi nevner følgende typer:

- Speiling
- Projeksjon
- Rotasjon
- Strekking

Vi har allerede sett på en speiling i \mathbb{R}^2 , og vi skal også ta med noen andre eksempler på lineær transformasjoner i \mathbb{R}^2 .

La oss begynne med en projeksjon. Et eksempel på en slik lineær transformasjon er $T(x, y) = (x, 0)$, projeksjonen på x -aksen, som har matrise

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

Legg merke til at mens en speiling kjennetegnes ved at matrisen A oppfyller $A^2 = I$, så kjennetegnes en projeksjon ved at matrisen A oppfyller $A^2 = A$ (sjekk dette i eksemplene ovenfor!).

Et eksempel på en strekking i \mathbb{R}^2 er for eksempel $T(x, y) = (4x, 3y)$. Denne lineær transformasjonen strekker enhver vektor med faktor 4 i x -retning og faktor 3 i y -retning, og den har matrise

$$\begin{pmatrix} 4 & 0 \\ 0 & 3 \end{pmatrix}$$

Et annet eksempel er $T(x, y) = (0.1x, 2y)$. Her blir enhver vektor strukket med en faktor 0.1 (altså blir den presset sammen) i x -retning, mens den blir strukket med faktor 2 i y -retning.

Til slutt tar vi med rotasjon av \mathbb{R}^2 om origo med en vinkel θ (mot urviserens retning). Denne lineær transformasjonen er gitt ved formelen $T(x, y) = (\cos \theta \cdot x - \sin \theta \cdot y, \sin \theta \cdot x + \cos \theta \cdot y)$. Matrisen til denne lineær transformasjonen er

$$\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$$

Er for eksempel $\theta = \frac{\pi}{2}$, så blir dette lineær transformasjonen som roterer \mathbb{R}^2 med en vinkel på 90° mot urviseren, og formelen ovenfor viser at

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

er matrisen til denne lineær transformasjonen, som altså er gitt ved $T(x, y) = (-y, x)$.

4.2. Egenverdier og egenvektorer

La A være en gitt $n \times n$ -matrise. Vi definerer en *egenverdi* for A til å være en skalar $\lambda \in \mathbb{R}$ slik at likningen

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

har minst en ikke-triviell løsning $\mathbf{x} \neq \mathbf{0}$. Legg merke til at både $A\mathbf{x}$ og $\lambda\mathbf{x}$ gir en vektor i \mathbb{R}^n som svar, selv om A er en $n \times n$ -matrise og $\lambda \in \mathbb{R}$ er en skalar. Derfor gir likningen ovenfor mening. Legg også merke til at $A\mathbf{x} = \lambda\mathbf{x}$ alltid har den trivielle løsningen $\mathbf{x} = \mathbf{0}$, uansett hvilke verdier A og λ har.

Hvis λ er en egenverdi for A , så kalles alle løsningene av likningen $A\mathbf{x} = \lambda\mathbf{x}$ (inkludert den trivielle løsningen $\mathbf{x} = \mathbf{0}$) for *egenvektorene* til A for egenverdien λ . Samlingen av alle disse egenvektorer kalles *egenrommet* til A for egenverdien λ , og skrives $E_\lambda = \{\mathbf{x} : A\mathbf{x} = \lambda\mathbf{x}\}$. For hver egenverdi λ for A , så er E_λ et lineært underrom av \mathbb{R}^n .

I dette avsnittet skal vi lære hvordan vi kan finne egenverdier og egenvektorer til en gitt kvadratisk matrise. Senere skal vi lære hvordan vi kan bruke dette til å *diagonalisere* en matrise, en standard teknikk som brukes i mange anvendelser – for eksempel til løsning av *lineære differensiallikninger med konstante koeffisienter*. Vi kommer tilbake til slike praktiske anvendelser senere.

For å se på hvordan vi finner egenverdier og egenvektorer til en gitt kvadratisk matrise A , la oss begynne med eksempelet

$$A = \begin{pmatrix} 5 & -6 \\ 2 & -2 \end{pmatrix}$$

For å finne egenverdiene til A , må vi altså se på når likningen $A\mathbf{x} = \lambda\mathbf{x}$ har en ikke-triviell løsning. Det første vi legger merke til, er at dette er

et lineært likningssystem med to likninger og to ukjente. Vi kan nemlig skrive dette systemet som

$$\begin{pmatrix} 5 & -6 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \lambda \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

det vil si som

$$\begin{aligned} 5x_1 - 6x_2 &= \lambda x_1 \\ 2x_1 - 2x_2 &= \lambda x_2 \end{aligned}$$

Flytter vi alle leddene over på venstre side, får vi det lineære homogene likningssystemet

$$\begin{aligned} (5 - \lambda)x_1 - 6x_2 &= 0 \\ 2x_1 + (-2 - \lambda)x_2 &= 0 \end{aligned}$$

Vi tenker altså på dette som et likningssystem med ukjente x_1 og x_2 , med koeffisientmatrise som avhenger av λ . Et slikt likningssystem har en ikke-triviell løsning $\mathbf{x} \neq \mathbf{0}$ hvis og bare hvis

$$\det \begin{pmatrix} 5 - \lambda & -6 \\ 2 & -2 - \lambda \end{pmatrix} = 0$$

Regner vi dette ut, får vi likningen $(5 - \lambda)(-2 - \lambda) - 2(-6) = 0$, det vil si $\lambda^2 - 3\lambda + 2 = 0$. Egenverdiene til A er altså eksakt løsningene av denne likningen, og vi finner at disse løsningene er $\lambda = 1$ og $\lambda = 2$.

Mer generelt, om A er en vilkårlig kvadratisk matrise, så kan vi omforme likningen $A\mathbf{x} = \lambda\mathbf{x}$ til formen $A\mathbf{x} - \lambda\mathbf{x} = \mathbf{0}$. Det er nå fristende å sette \mathbf{x} utenfor en parentes, men dette er ikke uten videre mulig, fordi A er en matrise mens λ er en skalar. Men vi kan skrive dette om som

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

Dette er et homogent lineært likningssystem på matrisiform, som har koeffisientmatrise $A - \lambda I$, og vi vet at dette systemet har ikke-trivielle løsninger hvis og bare hvis

$$\det(A - \lambda I) = 0$$

Denne likningen kalles den *karakteristiske likningen* til A , og løsningene av denne likningen er eksakt egenverdiene til A . Legg merke til at det var akkurat denne likningen vi kom fram til i eksempelet ovenfor.

Hvis A er en $n \times n$ -matrise, så blir den karakteristiske likningen $\det(A - \lambda I) = 0$ en n 'te-gradslikning med λ som den ukjente. Faktisk blir det en likning på formen

$$(-1)^n \lambda^n + c_{n-1} \cdot \lambda^{n-1} + \cdots + c_1 \cdot \lambda + c_0 = 0$$

der c_{n-1}, \dots, c_0 er koeffisienter som avhenger av matrisen A . Det finnes faktisk formler for disse uttrykkene, men noen av dem er nokså kompliserte. Hvis vi setter inn $\lambda = 0$ både i uttrykket ovenfor og i den karakteristiske likningen, ser vi at $c_0 = \det(A)$ holder helt generelt.

Vi konkluderer at det i praksis kan bli vanskelig både å finne den karakteristiske likningen og å løse den når A er en stor matrise.

La oss se på et eksempel til, der $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. I dette tilfellet blir den karakteristiske likningen

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

det vil si $\lambda^2 + 1 = 0$ eller $\lambda^2 = -1$. Denne likningen har ingen reelle løsninger, og derfor har A ingen reelle egenverdier og ingen reelle egenvektorer i dette tilfellet. Det er mulig å finne løsninger på likningen blant såkalte *komplekse tall*, men det ligger utenfor pensum i dette kurset.

Vi har så langt sett på hvordan man finner egenverdiene til en kvadratisk matrise A ved å sette opp den karakteristiske likningen og løse denne. Vi skal nå gå videre og se på hvordan vi finner de tilsvarende egenvektorene. Vi husker at for en gitt egenverdi λ , så er egenvektorene til A for λ alle løsninger av likningen $A\mathbf{x} = \lambda\mathbf{x}$. Har vi først funnet alle egenverdiene til A , kan vi derfor finne alle egenvektorene ved å løse likningssystemet $(A - \lambda I)\mathbf{x} = \mathbf{0}$ for hver eneste egenverdi λ etter tur.

La oss se på eksempelet $A = \begin{pmatrix} 5 & -6 \\ 2 & -2 \end{pmatrix}$. Vi fant tidligere ut at $\lambda = 1$ og $\lambda = 2$ er de eneste egenverdiene til A i dette eksempelet. Altså har likningssystemet $A\mathbf{x} = \lambda\mathbf{x}$, som vi skrev om til formen

$$(5 - \lambda)x_1 - 6x_2 = 0$$

$$2x_1 + (-2 - \lambda)x_2 = 0$$

ikke-trivielle løsninger kun for $\lambda = 1$ og $\lambda = 2$.

For å finne egenvektorene for $\lambda = 1$, setter vi inn $\lambda = 1$ i systemet ovenfor og løser det på vanlig måte. Vi får

$$4x_1 - 6x_2 = 0$$

$$2x_1 - 3x_2 = 0$$

Vi kan for eksempel løse dette systemet ved hjelp av Gauss-Jordan eliminasjon. Vi får løsningen $\mathbf{x} = t \cdot \left(\frac{3}{2}, 1\right)$. Dermed blir egenrommet E_1 gitt ved

$$E_1 = \left\{ t \cdot \left(\frac{3}{2}, 1\right) \right\}$$

og vektoren $\mathbf{x} = \left(\frac{3}{2}, 1\right)$ er en basis for E_1 .

For å finne egenvektorene for $\lambda = 2$, setter vi inn $\lambda = 2$ i systemet, og får

$$3x_1 - 6x_2 = 0$$

$$2x_1 - 4x_2 = 0$$

Igjen løser vi systemet ved hjelp av Gauss-Jordan eliminasjon, og får løsningen $\mathbf{x} = t \cdot (2, 1)$. Dermed blir egenrommet E_2 gitt ved

$$E_2 = \{ t \cdot (2, 1) \}$$

og $\mathbf{x} = (2, 1)$ er en basis for E_2 .

Vi kan oppsummere den generelle metoden for å finne egenverdier og egenvektorer for en gitt kvadratisk matrise A slik:

- (1) Vi skriver opp den karakteristiske ligningen $\det(A - \lambda I) = 0$.
- (2) Vi finner alle egenverdiene til A ved å løse denne likningen.
- (3) For hver egenverdi λ av A , setter vi inn λ i likningssystemet $(A - \lambda I)\mathbf{x} = \mathbf{0}$ og løser det. Egenrommet E_λ av alle løsningene er egenvektorene til A med egenverdi λ .

Husk at en egenverdi for en $n \times n$ -matrise A er en skalar, mens en egenvektor for A er en vektor i \mathbb{R}^n !

4.3. Diagonalisering av matriser

La A være en kvadratisk matrise. Vi sier at A er *diagonaliserbar* hvis det fins en invertibel matrise P slik at $P^{-1}AP$ er en diagonal matrise. I så fall skriver vi ofte $D = P^{-1}AP$ for diagonaliseringen av A . Ved å multiplisere likningen med P fra venstre og P^{-1} fra høyre, kan vi også skrive denne likningen på formen $A = PDP^{-1}$.

Ikke alle kvadratiske matriser A er diagonaliserbare. Faktisk så har diagonaliseringer mye med egenverdier og egenvektorer å gjøre. I denne seksjonen skal vi se hvordan vi kan bruke egenverdier og egenvektorer til å diagonalisere en matrise, hvis det er mulig. Litt senere skal vi se hva vi kan bruke dette til.

La oss begynne med et eksempel, nemlig matrisen A fra forrige avsnitt, gitt ved

$$A = \begin{pmatrix} 5 & -6 \\ 2 & -2 \end{pmatrix}$$

Vi fant i forrige avsnitt ut at egenverdiene til A er 1 og 2, og en basis for de tilsvarende egenrommene er gitt ved vektorene $\mathbf{v}_1 = (\frac{3}{2}, 1)$ og $\mathbf{v}_2 = (2, 1)$. La oss danne matrisen P med egenvektorene \mathbf{v}_1 og \mathbf{v}_2 som kolonner,

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

Da er P en invertibel matrise, og regner vi ut uttrykket $D = P^{-1}AP$, finner vi at

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

Altså er D den diagonale matrisen med egenverdiene til A på diagonalen, og rekkefølgen til egenverdiene svarer til rekkefølgen til vektorene \mathbf{v}_1 og \mathbf{v}_2 i matrisen P . Konklusjonen er at A er diagonaliserbar, og at D er en diagonalisering av A .

Det er ikke tilfeldig at vi kan finne en diagonalisering av A ved å bruke egenverdiene og egenvektorene til A i dette eksempelet. Det samme kan vi gjøre i alle tilfeller der matrisen A har *nok* egenvektorer:

TEOREM 9. La A være en $n \times n$ -matrise, og velg en basis for egenrommet E_λ for hver egenverdi λ av A . Hvis vi totalt har n basisvektorer $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ for disse egenrommene, med tilsvarende egenverdier $\lambda_1, \lambda_2, \dots, \lambda_n$, så er

$$D = P^{-1}AP = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

en diagonalisering av A , med $P = (\mathbf{v}_1 | \mathbf{v}_2 | \dots | \mathbf{v}_n)$, matrisen som består av $\mathbf{v}_1, \dots, \mathbf{v}_n$ som kolonner. Hvis vi totalt har færre enn n basisvektorer for disse egenrommene, så kan A ikke diagonaliseres.

Vi skal se på hvorfor det er slik. Hvis vi danner matrisen P slik som i teoremet ovenfor, med egenvektorene $\mathbf{v}_1, \dots, \mathbf{v}_n$ som kolonner, så blir

$$A \cdot P = A \cdot (\mathbf{v}_1 | \dots | \mathbf{v}_n) = (A\mathbf{v}_1 | \dots | A\mathbf{v}_n) = (\lambda_1\mathbf{v}_1 | \dots | \lambda_n\mathbf{v}_n)$$

per definisjon av egenverdier og egenvektorer, mens vi har at

$$P \cdot D = (\mathbf{v}_1 | \dots | \mathbf{v}_n) \cdot \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix} = (\lambda_1\mathbf{v}_1 | \dots | \lambda_n\mathbf{v}_n)$$

Altså er $AP = PD$, og multiplikasjon fra venstre med P^{-1} gir dermed at $D = P^{-1}AP$. Legg merke til at noen av egenverdiene $\lambda_1, \dots, \lambda_n$ kan være like. Hvis en egenverdi λ opptrer r ganger blant $\lambda_1, \dots, \lambda_n$, så svarer det til egenrommet E_λ har en basis som består av r egenvektorer.

Men legg også merke til at argumentet ovenfor kun kan brukes hvis det totalt finnes n basisvektorer for de forskjellige egenrommene til A , der n er antall kolonner og rader i matrisen A . Hvis ikke, så finnes det jo ikke nok kolonner i P til å danne en kvadratisk matrise! For eksempel har vi tidligere sett at matrisen $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ ikke har noen reelle egenverdier. Denne matrisen har derfor ingen reelle egenvektorer heller, og kan ikke diagonaliseres.

La oss se på et annet eksempel, nemlig matrisen $A = \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}$. I dette tilfellet får vi den karakteristiske likningen

$$\det \begin{pmatrix} 3 - \lambda & 1 \\ 0 & 3 - \lambda \end{pmatrix} = (3 - \lambda)^2 = 0$$

som kun gir en egenverdi $\lambda = 3$. La oss regne ut de tilhørende egenvektorene. Vi ser da på likningssystemet vi får når vi setter inn $\lambda = 3$, altså

$$0x_1 + 1x_2 = 0$$

$$0x_1 + 0x_2 = 0$$

Dette gir løsning $x_2 = 0$, som vi skriver $\mathbf{x} = t \cdot (1, 0)$ på vektorform. Altså er $E_3 = \{t \cdot (1, 0)\}$. Tar vi med en basis for egenvektorer for hver egenverdi, får vi i dette tilfellet kun med vektoren $\mathbf{v}_1 = (1, 0)$. Siden vi har færre enn $n = 2$ basis-vektorer, så kan A ikke diagonaliseres.

Til slutt tar vi med et litt mer komplisert eksempel, der matrisen A er 3×3 -matrisen

$$A = \begin{pmatrix} 3 & 0 & 0 \\ -4 & 6 & 2 \\ 16 & -15 & -5 \end{pmatrix}$$

I dette tilfellet er den karakteristiske likningen gitt ved

$$\det(A - \lambda I) = \det \begin{pmatrix} 3 - \lambda & 0 & 0 \\ -4 & 6 - \lambda & 2 \\ 16 & -15 & -5 - \lambda \end{pmatrix} = 0$$

og dette gir $\lambda(\lambda - 1)(3 - \lambda) = 0$. Det følger derfor at egenverdiene til A er $\lambda = 0, 1, 3$. I hvert av de tre tilfellene må vi finne en basis for egenrommet E_λ , og etter litt regning finner vi at

$$E_0 = \left\{ t \cdot \left(0, -\frac{1}{3}, 1 \right) \right\}$$

$$E_1 = \left\{ t \cdot \left(0, -\frac{2}{5}, 1 \right) \right\}$$

$$E_3 = \left\{ t \cdot \left(\frac{1}{2}, 0, 1 \right) \right\}$$

Egenrommene har derfor basis-vektorer $\mathbf{v}_1 = (0, -\frac{1}{3}, 1)$, $\mathbf{v}_2 = (0, -\frac{2}{5}, 1)$ og $\mathbf{v}_3 = (\frac{1}{2}, 0, 1)$. I alt har vi $n = 3$ basis-vektorer, så vi danner matrisen

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

og regner vi ut $D = P^{-1}AP$. Vi vet da at

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

som er en diagonalisering av A .

TEOREM 10. *La A være en $n \times n$ -matrise. I hvert av de to tilfellene nedenfor kan A diagonaliseres:*

- (1) *Når A har n forskjellige reelle egenverdier.*
- (2) *Når A er en symmetrisk matrise, det vil si $A^t = A$.*

Til slutt nevner vi to viktige resultater som har med egenverdier å gjøre. Disse resultatene gjelder faktisk for enhver kvadratiske matrise A , men de er begge lettest å forstå og å bevise når matrisen A kan diagonaliseres.

TEOREM 11 (Cayley-Hamilton). *La A være en $n \times n$ -matrise, og la den karakteristiske likningen til A være gitt som*

$$(-1)^n \lambda^n + c_{n-1} \lambda^{n-1} + \dots + c_1 \lambda + c_0 = 0$$

Da oppfyller matrisen A den karakteristiske likningen, det vil si at

$$(-1)^n A^n + c_{n-1} A^{n-1} + \dots + c_1 A + c_0 I = 0$$

Vi tar med et eksempel på hvordan dette resultatet kan brukes. La A være matrisen

$$A = \begin{pmatrix} 3 & 0 & 0 \\ -4 & 6 & 2 \\ 16 & -15 & -5 \end{pmatrix}$$

fra eksempelet ovenfor. Vi regnet ut den karakteristiske likningen til A , som er gitt som

$$\det(A - \lambda I) = -\lambda^3 + 4\lambda^2 - 3\lambda = \lambda(\lambda - 1)(3 - \lambda) = 0$$

Cayley-Hamiltons teorem gir derfor at $-A^3 + 4A^2 - 3A = 0$. Hvis vi skriver dette om, får vi at $A^3 = 4A^2 - 3A$ (sjekk dette!).

TEOREM 12. *La A være en $n \times n$ -matrise, og la $\lambda_1, \dots, \lambda_n$ være egenverdiene til A . Da er*

$$\det(A) = \lambda_1 \cdot \lambda_2 \cdots \lambda_n$$

Legg merke til at vi i denne formelen må ta med *alle* egenverdiene til A . Hvis en egenverdi opptrer flere ganger, må vi ta med alle kopiene av den. For eksempel hadde vi et eksempel tidligere, $A = \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}$, der den karakteristiske likningen var $(\lambda - 3)^2 = 0$. I dette tilfellet må vi ta med egenverdien $\lambda = 3$ to ganger for at formelen skal bli korrekt. Determinanten til A er altså

$$\det(A) = \lambda_1 \cdot \lambda_2 = 3 \cdot 3 = 9$$

i dette tilfellet. Vi skal kun bruke formelen for determinanten i de tilfellene der alle egenverdiene blir reelle.

4.4. Noen anvendelser av diagonalisering

Vi ser på følgende problem: Hvis A er en kvadratisk matrise, så kan vi regne ut $A^2 = A \cdot A$, $A^3 = A \cdot A \cdot A$, og så videre. For å regne ut A^m når m er et stort heltall blir nokså mye arbeid hvis vi skal gjøre det på denne måten, ved ganske enkelt å multiplisere sammen matrisen A med seg selv mange ganger.

Anta at A er en diagonaliserbar matrise, slik at $D = P^{-1}AP$ er en diagonal matrise. I dette tilfellet blir det mye enklere å regne ut A^m .

Det er to grunner til det: For det første, hvis

$$D = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

er en diagonal matrise, så er det ganske enkelt å regne ut D^m for en hvilken som helst m : Vi får rett og slett

$$D^m = \begin{pmatrix} \lambda_1^m & 0 & \dots & 0 \\ 0 & \lambda_2^m & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n^m \end{pmatrix}$$

Grunnen til denne forenklingen er altså at det er lett å multiplisere sammen diagonale matriser.

For det andre, om $D = P^{-1}AP$, så blir $A = PDP^{-1}$. Vi finner da følgende formel for A^m :

$$\begin{aligned} A^m &= (PDP^{-1})^m \\ &= (PDP^{-1})(PDP^{-1}) \dots (PDP^{-1}) \\ &= PDD \dots DP^{-1} \\ &= PD^m P^{-1} \end{aligned}$$

Vi får altså at $A^m = PD^m P^{-1}$. Så vi kan rett og slett regne ut D^m (som er mye lettere enn å regne ut A^m), og multiplisere med P og P^{-1} på hver sin side, så finner vi A^m !

La oss se på et eksempel, nemlig matrisen A fra forrige avsnitt gitt ved

$$A = \begin{pmatrix} 3 & 0 & 0 \\ -4 & 6 & 2 \\ 16 & -15 & -5 \end{pmatrix}$$

I dette tilfellet fant vi en diagonalisering $D = P^{-1}AP$, der matrisene D og P var gitt ved hjelp av egenverdier og egenvektorer som

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

og

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

For å finne A^m må vi regne ut $A^m = PD^mP^{-1}$. Det er nokså lett å finne D^m , som er gitt ved formelen

$$D^m = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3^m \end{pmatrix}$$

Vi trenger også matrisen P^{-1} , som vi finner ved hjelp av Gauss-Jordan eliminasjon:

$$P^{-1} = \begin{pmatrix} -12 & 15 & 6 \\ 10 & -15 & -5 \\ 2 & 0 & 0 \end{pmatrix}$$

Til slutt regner vi ut matrisen A^m for $m > 0$ ved hjelp av formelen

$$\begin{aligned} A^m &= PD^mP^{-1} \\ &= \begin{pmatrix} 0 & 0 & \frac{1}{2} \\ -\frac{1}{3} & -\frac{2}{5} & 0 \\ 1 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3^m \end{pmatrix} \cdot \begin{pmatrix} -12 & 15 & 6 \\ 10 & -15 & -5 \\ 2 & 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 3^m & 0 & 0 \\ -4 & 6 & 2 \\ 2 \cdot 3^m + 10 & -15 & -5 \end{pmatrix} \end{aligned}$$

Vil vi for eksempel finne A^{10} , setter vi inn $m = 10$, og får at

$$A^{10} = \begin{pmatrix} 3^{10} & 0 & 0 \\ -4 & 6 & 2 \\ 2 \cdot 3^{10} + 10 & -15 & -5 \end{pmatrix} = \begin{pmatrix} 59049 & 0 & 0 \\ -4 & 6 & 2 \\ 118108 & -15 & -5 \end{pmatrix}$$

Vi skal til slutt forklare hvordan man kan bruke denne metoden til å løse mange praktiske problemer. De problemene som vi skal se på, kan alle beskrives på denne måten:

- (1) Tilstanden ved ethvert tidspunkt kan beskrives av en *tilstandsvektor* $\mathbf{x} \in \mathbb{R}^n$, altså ved hjelp av n variable x_1, x_2, \dots, x_n .
- (2) Ved starttidspunktet $t = 0$ er tilstanden beskrevet av vektoren \mathbf{x}_0 , og etter i tidsperioder av vektoren \mathbf{x}_i .
- (3) Forandringen fra en tidsperiode til den neste er beskrevet ved hjelp av en $n \times n$ -matrise A , som kalles *overgangsmatrisen* til systemet, slik at $\mathbf{x}_{i+1} = A\mathbf{x}_i$ for alle $i \geq 0$.

For et hvilket som helst system som kan beskrives slik, ser vi at tilstanden etter m tidsperioder er gitt ved

$$\mathbf{x}_m = A^m \mathbf{x}_0$$

Dette forklarer hvordan vi kan bruke metodene for å regne ut A^m til å beskrive slike systemer.

Før vi går videre, la oss se på et konkret eksempel på et system som kan beskrives på denne måten. Vi tenker at at vi innenfor et begrenset område har R_0 rever og K_0 kaniner ved starttidspunktet, og R_i rever

og K_i kaniner i måneder etter starttidspunktet. Vi antar at sammenhengen mellom antall rever og kaniner fra en måned til den neste er gitt ved følgende *modell*:

$$\begin{aligned}R_{i+1} &= 0.4R_i + 0.3K_i \\K_{i+1} &= -rR_i + 1.2K_i\end{aligned}$$

Setter vi $\mathbf{x}_i = (R_i, K_i)$, så blir overgangsmatrisen til dette systemet gitt ved

$$A = \begin{pmatrix} 0.4 & 0.3 \\ -r & 1.2 \end{pmatrix}$$

La oss regne på tilfellet $r = 0.4$, det vil si at hver rev spiser 0.4 kaniner per måned i gjennomsnitt. Hva skjer med tilstanden $\mathbf{x}_m = A^m \mathbf{x}_0$ da etterhvert som tiden går, det vil si etterhvert som m vokser?

For å svare på dette spørsmålet må vi regne ut A^m , og vi gjør dette ved hjelp av samme metode som vi har brukt tidligere. Vi finner først den karakteristiske likningen til A , som blir

$$\det \begin{pmatrix} 0.4 - \lambda & 0.3 \\ -0.4 & 1.2 - \lambda \end{pmatrix} = \lambda^2 - 1.6\lambda + 0.6 = 0$$

Dette gir egenverdier $\lambda_1 = 1$ og $\lambda_2 = 0.6$. Vi regner så ut egenrommene. For λ_1 får vi likningssystemet

$$\begin{pmatrix} -0.6 & 0.3 \\ -0.4 & 0.2 \end{pmatrix} \mathbf{x} = \mathbf{0}$$

og dette gir egenvektoren $\mathbf{v}_1 = (1, 2)$ som basis for egenrommet E_1 . For λ_2 får vi likningssystemet

$$\begin{pmatrix} -0.2 & 0.3 \\ -0.4 & 0.6 \end{pmatrix} \mathbf{x} = \mathbf{0}$$

og dette gir egenvektoren $\mathbf{v}_2 = (3, 2)$ som basis for egenrommet $E_{0.6}$. Så A er diagonaliserbar, og vi lar $P = \begin{pmatrix} 1 & 3 \\ 2 & 2 \end{pmatrix}$. Dette gir invers

$$P^{-1} = -\frac{1}{4} \begin{pmatrix} 2 & -3 \\ -2 & 1 \end{pmatrix}$$

og $D = P^{-1}AP$ er gitt ved

$$D = \begin{pmatrix} 1 & 0 \\ 0 & 0.6 \end{pmatrix}$$

Til slutt regner vi ut $A^m = PD^mP^{-1}$, som er gitt ved følgende uttrykk:

$$A^m = \frac{1}{4} \begin{pmatrix} 6 \cdot 0.6^m - 2 & -3 \cdot 0.6^m + 3 \\ 4 \cdot 0.6^m - 4 & -2 \cdot 0.6^m + 6 \end{pmatrix}$$

Etter lang tid, det vil si når m er blitt stor, så er 0.6^m svært lite. For store verdier av m kan vi derfor sette $0.6^m \approx 0$, og dette gir

$$\mathbf{x}_m = A^m \mathbf{x}_0 \approx \frac{1}{4} \begin{pmatrix} -2 & 3 \\ -4 & 6 \end{pmatrix} \begin{pmatrix} R_0 \\ K_0 \end{pmatrix} = \frac{1}{4} (3K_0 - 2R_0) \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

Dette betyr at hvis $3K_0 - 2R_0 > 0$, så vil antallet rever og kaniner i området etterhvert nærme seg en likevektssituasjon, der det er dobbelt så mange kaniner som rever i området.

Gjør vi en tilsvarende regning for andre verdier av r , så kan vi vise at hvis revene blir grådige og spiser mer enn 0.4 kaniner per rev i gjennomsnitt per måned (det vil si $r > 0.4$), så vil både antall rever og antall kaniner avta mot null etterhvert som tiden går. Hvis revene spiser mindre enn 0.4 kaniner per rev i gjennomsnitt per måned (det vil si at $r < 0.4$), så vil antall rever og antall kaniner i følge modellen vokse mot uendelig etter hvert som tiden går. Sjekk dette ved å regne på tilfellene $r = 0.5$ og $r = 0.325$!

4.5. Systemer av lineære differensiallikninger

Vi nevner også en annen viktig anvendelse. Vi kan bruke egenverdier og egenvektorer til å løse systemer av lineære differensiallikninger. Vi skal ta for oss løsningsmetoder for homogene systemer av første ordens lineære differensiallikninger med konstante koeffisienter.

La oss starte med følgende problem: Vi har tre beholdere B_1 , B_2 og B_3 som alle inneholder saltvannsoppløsninger, og volumet av disse oppløsningene er henholdsvis $V_1 = 20$, $V_2 = 40$ og $V_3 = 50$ (liter). Vi antar at det renner rent vann inn i beholder B_1 , og saltvannsoppløsning fra beholder B_1 til beholder B_2 , fra beholder B_2 til beholder B_3 , og ut av beholder B_3 , og at alle væskestrømmene har konstant hastighet $r = 10$ (liter per minutt). Vi antar også at det er en røremekanisme i hver beholder, slik at saltet er jevnt fordelt i saltvannsoppløsningene til enhver tid. Hvis vi kaller saltmengden (kg) i de tre beholderene for $x_1(t)$, $x_2(t)$ og $x_3(t)$ etter t minutter, og $x_1(0) = 15$, $x_2(0) = x_3(0) = 0$ ved starttidspunktet, finn saltmengdene $x_1(t)$, $x_2(t)$ og $x_3(t)$.

Vi setter opp likninger for å bestemme x_1 , x_2 og x_3 . I løpet av tidsrommet Δt , er endringene i saltmengdene gitt ved:

$$\begin{aligned} \Delta x_1 &= \left(0 - r \frac{x_1}{V_1}\right) \Delta t \\ \Delta x_2 &= \left(r \frac{x_1}{V_1} - r \frac{x_2}{V_2}\right) \Delta t \\ \Delta x_3 &= \left(r \frac{x_2}{V_2} - r \frac{x_3}{V_3}\right) \Delta t \end{aligned}$$

Vi setter $k_i = \frac{r}{V_i}$ for $i = 1, 2, 3$, og får dermed likningene

$$\begin{aligned}x'_1 &= -k_1x_1 \\x'_2 &= k_1x_1 - k_2x_2 \\x'_3 &= k_2x_2 - k_3x_3\end{aligned}$$

Vi får altså et system av første ordens differensiallikninger, siden hver likning inneholder en av de deriverte x'_i . Koeffisientene k_i er gitt ved $k_1 = 0.5$, $k_2 = 0.25$ og $k_3 = 0.2$, og er altså konstanter. Vi har ingen konstantledd. Tilsammen gir dette oss et homogent system av første ordens lineære differensiallikninger. Vi kan skrive systemet på matriseform som

$$(4.1) \quad \begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \end{pmatrix} = \begin{pmatrix} -k_1 & 0 & 0 \\ k_1 & -k_2 & 0 \\ 0 & k_2 & -k_3 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

Vi kan skrive dette systemet enda kortere som $\mathbf{x}' = A\mathbf{x}$, hvor A er koeffisientmatrisen til systemet.

Vi skal nå se litt på hvordan vi kan løse dette systemet, det vil si hvordan vi kan finne uttrykk for funksjonene $x_1(t)$, $x_2(t)$ og $x_3(t)$ som passer inn i disse likningene. Den *generelle løsningen* er de mest generelle funksjonene som er løsninger. Når vi så har funnet den generelle løsningen, må vi forsøke å få den til å passe med *initial-betingelsene*, det vil si opplysningene om at $x_1(0) = 15$ og $x_2(0) = x_3(0) = 0$.

For å finne fram til en løsningsmetode, kan vi se på et helt generelt system $\mathbf{x}' = A\mathbf{x}$, der A er en $n \times n$ -matrise med konstante koeffisienter. Hvis vi skriver

$$\mathbf{x} = \mathbf{v} \cdot e^{\lambda t}$$

der $\lambda \in \mathbb{R}$ er en skalar og \mathbf{v} er en vektor i \mathbb{R}^n (altså en n -vektor med konstante koeffisienter), og forsøker å sette dette uttrykket inn i likningen $\mathbf{x}' = A\mathbf{x}$, så får vi

$$\mathbf{x}' = (\mathbf{v} \cdot e^{\lambda t})' = \mathbf{v} \cdot (\lambda e^{\lambda t}) = (\lambda \mathbf{v})e^{\lambda t}$$

og

$$A\mathbf{x} = A \cdot (\mathbf{v} \cdot e^{\lambda t}) = (A\mathbf{v})e^{\lambda t}$$

Setter vi disse to uttrykkene lik hverandre, kan vi forkorte faktoren $e^{\lambda t}$ på begge sider, siden denne faktoren aldri kan være null. Da får vi likningen

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

Vi ser at denne likningen er oppfylt hvis enten $\mathbf{v} = \mathbf{0}$ eller hvis λ er en egenverdi for A med egenvektor \mathbf{v} . Det betyr altså at $\mathbf{x} = \mathbf{v}e^{\lambda t}$ er en ikke-triviell løsning av systemet $\mathbf{x}' = A\mathbf{x}$ hvis λ er en egenverdi og \mathbf{v} er en egenvektor for A (mens $\mathbf{v} = \mathbf{0}$ gir en triviell løsning).

Vi forsøker nå å bruke dette til løse systemet (4.1) ovenfor. Vi må derfor finne egenverdier og egenvektorer til matrisen

$$A = \begin{pmatrix} -k_1 & 0 & 0 \\ k_1 & -k_2 & 0 \\ 0 & k_2 & -k_3 \end{pmatrix} = \begin{pmatrix} -0.5 & 0 & 0 \\ 0.5 & -0.25 & 0 \\ 0 & 0.25 & -0.2 \end{pmatrix}$$

Dette gir karakteristisk likning

$$\det(A - \lambda I) = (-0.5 - \lambda)(-0.25 - \lambda)(-0.2 - \lambda) = 0$$

og dermed egenverdier $\lambda_1 = -0.5$, $\lambda_2 = -0.25$ og $\lambda_3 = -0.2$. Etter litt regning finner vi også en egenvektor \mathbf{v}_i for hver av disse tre egenverdiene som er basis for det tilsvarende egenrommet:

$$\mathbf{v}_1 = (3, -6, 5)$$

$$\mathbf{v}_2 = (0, 1, -5)$$

$$\mathbf{v}_3 = (0, 0, 1)$$

Da vet vi at følgende tre vektorer alle gir løsninger av systemet av differensial-likninger:

$$\mathbf{u}_1 = \mathbf{v}_1 \cdot e^{\lambda_1 t} = (3, -6, 5)e^{-0.5t} = (3e^{-0.5t}, -6e^{-0.5t}, 5e^{-0.5t})$$

$$\mathbf{u}_2 = \mathbf{v}_2 \cdot e^{\lambda_2 t} = (0, 1, -5)e^{-0.25t} = (0, e^{-0.25t}, -5e^{-0.25t})$$

$$\mathbf{u}_3 = \mathbf{v}_3 \cdot e^{\lambda_3 t} = (0, 0, 1)e^{-0.2t} = (0, 0, e^{-0.2t})$$

Sjekk dette ved å sette inn hver av vektorene \mathbf{u}_1 , \mathbf{u}_2 og \mathbf{u}_3 for \mathbf{x} i systemet (4.1)! I hvert av de tre tilfellene må vi altså sjekke at alle tre likningene er oppfylt. Dette betyr at \mathbf{u}_1 , \mathbf{u}_2 og \mathbf{u}_3 hver for seg er en løsning av systemet (4.1).

Vårt system av differensial-likninger har tre likninger og tre ukjente. Vi har funnet tre forskjellige løsninger av dette systemet. Den generelle teorien forteller oss da at hvis de tre løsningene vi har funnet er *lineært uavhengige*, så finner vi den generelle løsningen av systemet (4.1) ved å ta alle mulige lineærkombinasjoner $\mathbf{x} = C_1\mathbf{u}_1 + C_2\mathbf{u}_2 + C_3\mathbf{u}_3$ av de tre løsningene vi har funnet. Altså blir den generelle løsningen i vårt eksempel

$$\begin{aligned} \mathbf{x} &= C_1(3e^{-0.5t}, -6e^{-0.5t}, 5e^{-0.5t}) + C_2(0, e^{-0.25t}, -5e^{-0.25t}) \\ &\quad + C_3(0, 0, e^{-0.2t}) \\ &= (3C_1e^{-0.5t}, -6C_1e^{-0.5t} + C_2e^{-0.25t}, 5C_1e^{-0.5t} \\ &\quad - 5C_2e^{-0.25t} + C_3e^{-0.2t}) \end{aligned}$$

der C_1 , C_2 og C_3 er fritt valgte konstanter. Det betyr altså at en hvilken som helst løsning av (4.1) passer inn i dette uttrykket for et eller annet valg av C_1 , C_2 og C_3 .

Den observante leser vil ha lagt merke til at vi hoppet over å sjekke om de tre løsningene \mathbf{u}_1 , \mathbf{u}_2 og \mathbf{u}_3 er lineært uavhengige. Det finnes

mange metoder for å sjekke dette, og et par av dem står beskrevet i boken. Men vi skal nøye oss med å si at hvis vi finner n basisvektorer $\mathbf{v}_1, \dots, \mathbf{v}_n$ for egenrommene til en $n \times n$ -matrise, og danner løsningene fra disse egenvektorene slik vi har gjort ovenfor, så blir de n løsningene automatisk lineært uavhengige.

TEOREM 13. La $\mathbf{x}' = A\mathbf{x}$ være et system av differensial-likninger gitt ved en $n \times n$ -matrise A med konstante koeffisienter. Hvis vi til sammen kan finne n basisvektorer $\mathbf{v}_1, \dots, \mathbf{v}_n$ for de forskjellige egenrommene for A , og $\lambda_1, \dots, \lambda_n$ er de tilsvarende egenverdiene, så er den generelle løsningen av $\mathbf{x}' = A\mathbf{x}$ gitt ved

$$\mathbf{x} = C_1 \mathbf{u}_1 + \dots + C_n \mathbf{u}_n$$

hvor $\mathbf{u}_i = \mathbf{v}_i \cdot e^{\lambda_i t}$ for $i = 1, 2, \dots, n$.

Vi har altså funnet den generelle løsningen av vårt opprinnelige problem om saltinnholdet i de tre beholderene etter hvert som tiden går. Men vi har ennå ikke brukt opplysningen om at $x_1(0) = 15$ og $x_2(0) = x_3(0) = 0$. Dette kalles en *initialbetingelse* eller en *startbetingelse*, og kan brukes til å bestemme konstantene C_1 , C_2 og C_3 . Vi regner først ut $x_i(0)$ fra den generelle løsningen:

$$\begin{aligned} x_1(0) &= 3C_1 \\ x_2(0) &= -6C_1 + C_2 \\ x_3(0) &= 5C_1 - 5C_2 + C_3 \end{aligned}$$

Setter vi inn verdiene fra initial-betingelsen:

$$\begin{aligned} 15 &= 3C_1 \\ 0 &= -6C_1 + C_2 \\ 0 &= 5C_1 - 5C_2 + C_3 \end{aligned}$$

Vi får altså ett lineært likningssystem, og i dette tilfellet er det ikke vanskelig å løse. Fra første likning får vi $C_1 = 5$. Setter vi dette inn i andre likning, får vi $C_2 = 30$, og siste likning gir da $C_3 = 125$. Altså blir løsningen

$$\begin{aligned} x_1(t) &= 15e^{-0.5t} \\ x_2(t) &= -30e^{-0.5t} + 30e^{-0.25t} \\ x_3(t) &= 25e^{-0.5t} - 150e^{-0.25t} + 125e^{-0.2t} \end{aligned}$$

Dette beskriver saltinnholdet i de tre beholderene etter tiden t minutter. Vi ser at saltinnholdet i beholder 1 vil synke etterhvert som tiden går. Men det er vanskeligere å se hvordan saltinnholdet i beholder 2 og 3 vil utvikle seg. Forsøk å tegne grafene ved hjelp av kalkulator eller PC for å se hvordan de utvikler seg!