Grunnleggende lineæralgebra

Øystein Skartsæterhagen (og bitte litt Nome)

Kapittel 1

Lineære likningssystemer

Grunnlaget for lineær algebra er *lineære likningssystemer*. Vi starter vår reise inn i den lineære algebraen ved å se på noen forskjellige måter å løse likningssystemer på. Til slutt i dette kapitlet skal vi forsikre oss om at vi er helt enige om nøyaktig hva det vil si at et likningssystem er lineært, og innføre en praktisk måte å skrive lineære likningssystemer på.

Forskjellige fremgangsmåter

Her er et eksempel på et lineært likningssystem:

$$(*)\begin{cases} 2x + 3y = 9 \\ -x + 6y = 3 \end{cases}$$

I dette systemet har vi to likninger og to ukjente. En løsning av systemet består av to tall som vi kan sette inn for x og y slik at begge likningene er oppfylt samtidig.

Vi kjenner fra før til flere måter å løse et slikt system på. La oss løse systemet over med noen forskjellige metoder.

Eksempel 1.1. Den kanskje mest åpenbare metoden for å løse et likningssystem er å først løse én likning med hensyn på én av de ukjente, og så sette inn i den andre likningen (eller i *de* andre likningene, hvis vi har et system med mer enn to likninger).

For å løse systemet $(*)$ med denne metoden kan vi først løse den andre likningen med hensyn på x; da får vi:

$$x = 6y - 3$$

Så setter vi dette inn i den første likningen og forenkler:

$$\begin{aligned} 2 \cdot (6y - 3) + 3y &= 9 \\ 12y - 6 + 3y &= 9 \\ 15y &= 15 \\ y &= 1 \end{aligned}$$

Til slutt setter vi denne y-verdien inn i uttrykket vi fant for x, og får:

$$x = 6y - 3 = 6 - 3 = 3$$

Vi har altså funnet ut at for at begge likningene skal være oppfylt, må vi ha at $x = 3$ og $y = 1$. Vi sjekker at dette virkelig er en løsning av $(*)$ ved å sette inn disse verdiene i begge likningene i systemet:

$$\begin{aligned} 2x + 3y &= 2 \cdot 3 + 3 \cdot 1 = 6 + 3 = 9 \\ -x + 6y &= -3 + 6 \cdot 1 = 3 \end{aligned}$$

Vi har nå funnet ut at systemet $(*)$ har nøyaktig én løsning, nemlig $x = 3$ og $y = 1$. △

Metoden i eksempelet over er enkel og grei, men kan bli temmelig tungvint å bruke hvis vi har mer enn to ukjente. Vi ser på en annen løsningsmetode for det samme systemet.

Eksempel 1.2. Vi ganger opp den andre likningen med 2, og deretter legger vi sammen de to likningene:

$$\begin{array}{rl} 2x + 3y = 9 & \text{(første likning)} \\ -2x + 12y = 6 & \text{(andre likning ganget med 2)} \\ \hline 15y = 15 & \text{(sum av likningene over)} \end{array}$$

På denne måten får vi x til å forsvinne, og vi står igjen med en likning med bare y.

Den nye likningen $15y = 15$ kan vi forenkle til $y = 1$. Nå kan vi gange opp denne med -3 og legge sammen med den første likningen fra systemet for å få en likning der y forsvinner:

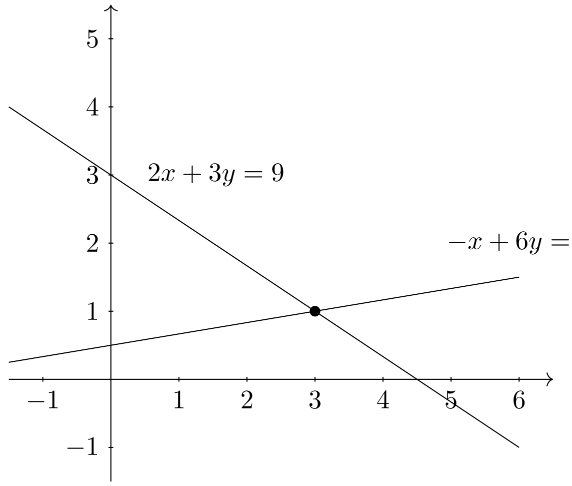
$$\begin{array}{rl} 2x + 3y = 9 & \text{(første likning fra } (*)) \\ -3y = -3 & \text{(ny likning ganget med } -3) \\ \hline 2x = 6 & \text{(sum av likningene over)} \end{array}$$

Nå har vi fått en likning med bare x, og vi forenkler den til $x = 3$. Vi har dermed igjen funnet løsningen $x = 3$ og $y = 6$. △

Vi skal etter hvert komme frem til en generell fremgangsmåte for å løse lineære likningssystemer. Den fremgangsmåten baserer seg på å legge sammen likninger slik som vi gjorde nå.

Før vi går videre løser vi systemet vårt en tredje gang, på en helt annen måte:

Eksempel 1.3. Vi kan også løse systemet $(*)$ grafisk. Vi lager et koordinatsystem med en x-akse og en y-akse. Hver av de to likningene $2x + 3y = 9$ og $-x + 6y = 3$ beskriver en rett linje:



Løsningene av $2x + 3y = 9$ er alle punkter som ligger på den ene linjen, mens løsningene av $-x + 6y = 3$ er alle punkter som ligger på den andre linjen. Den felles løsningen av begge likningene er punktet der de to linjene møtes, nemlig $(3, 1)$. Løsningen er altså $x = 3$ og $y = 1$.

Denne metoden er fin for å visualisere problemet og se løsningen på en intuitiv måte, men ikke nødvendigvis den beste for å finne svaret eksakt. Dessuten blir det vanskelig å tegne hvis vi har mer enn to ukjente (men det kan likevel være nyttig å prøve å se for seg løsningene av for eksempel en likning med tre ukjente på en grafisk måte). △

Hva er et lineært likningssystem?

Et lineært likningssystem er et system av lineære likninger. Men nøyaktig hva mener vi med at en likning er lineær?

Ordet «lineær» kommer fra det latinske «linea», som betyr «linje». Hvis vi har en likning med to ukjente, så kan vi tegne grafen til denne likningen. Vi sier at likningen er lineær hvis grafen dens er en rett linje. I eksempelet over så vi at grafene til likningene $2x + 3y = 9$ og $-x + 6y = 3$ er rette linjer. Men det er også lett å finne likninger som ikke har rettlinjede grafer, for eksempel $y = x^2$ eller $x^2 + y^2 = 4$.

Generelt er det slik at grafen til en likning er en rett linje hvis og bare hvis likningen kan skrives på formen

$$ax + by = c,$$

der a, b og c er konstanter.

Når vi vil se på likninger med mer enn to ukjente, gir det ikke lenger mening å snakke om at grafen blir en rett linje. Men det at likningen kan skrives på formen $ax + by = c$ kan vi lett utvide til å ta med flere ukjente, så det er dette vi bruker i den generelle definisjonen av lineære likninger.

Definisjon. En likning med n ukjente x_1, x_2, …, x_n kalles *lineær* dersom den kan skrives på formen

$$a_1x_1 + a_2x_2 + \cdots a_nx_n = b,$$

der a_1, a_2, …, a_n og b er konstanter. Et *lineært likningssystem* er en samling av én eller flere lineære likninger med de samme ukjente. △

Eksempel 1.4. Her er noen eksempler på lineære likninger:

$$5x_1 + x_2 - 7x_3 + x_4 = 27$$
$$17x - 5y = \pi$$

Og her er noen eksempler på likninger som ikke er lineære:

$$2x^2 + y = 3$$
$$x_1 + 2x_1x_2 + x_3 = 1$$

△

Ekvivalente systemer

Vi sier at to likningssystemer er *ekvivalente* dersom de har samme løsninger. Vi kan løse et likningssystem ved å erstatte det med stadig enklere ekvivalente systemer. Vi tar et eksempel for å se hvordan dette kan gjøres.

Eksempel 1.5. La oss løse det følgende lineære likningssystemet:

$$\begin{cases} x + 2y - 2z = -5 \\ x + 5y + 9z = 33 \\ 2x + 5y - z = 0 \end{cases}$$

Vi vil begynne med å eliminere x-en fra de to siste likningene. Hvis vi trekker den første likningen fra den andre, får vi den nye likningen

$$3y + 11z = 38.$$

Vi bytter ut den andre likningen i systemet med denne nye likningen:

$$\begin{cases} x + 2y - 2z = -5 \\ 3y + 11z = 38 \\ 2x + 5y - z = 0 \end{cases}$$

Dette systemet er ekvivalent med det vi startet med. Hvorfor er det det? Den nye likningen følger fra to av likningene vi hadde fra før. Dermed må enhver løsning av det opprinnelige systemet også være en løsning av det nye systemet.

Men omvendt er det også slik at den opprinnelige midterste likningen (som vi nå har tatt vekk) kan vi få ved å legge sammen de to første likningene i det nye systemet. Dermed må enhver løsning av det nye systemet også være en løsning av det gamle.

Til sammen betyr dette at de to systemene er ekvivalente.

Vi fortsetter å forenkle systemet. Vi eliminerer x fra den siste likningen ved å trekke fra første likning ganget med 2:

$$\begin{cases} x + 2y - 2z = -5 \\ 3y + 11z = 38 \\ y + 3z = 10 \end{cases}$$

Igjen har vi et nytt system som er ekvivalent med det forrige.

Nå vil vi eliminere y fra den siste likningen. Men det er lettere å eliminere y fra den midterste likningen (ved å trekke fra 3 ganger den siste). Likningenes rekkefølge spiller imidlertid ingen rolle, så vi kan bytte om på de to nederste likningene først:

$$\begin{cases} x + 2y - 2z = -5 \\ y + 3z = 10 \\ 3y + 11z = 38 \end{cases}$$

Så eliminerer vi y fra den siste likningen ved å trekke fra andre likning ganget med 3:

$$\begin{cases} x + 2y - 2z = -5 \\ y + 3z = 10 \\ 2z = 8 \end{cases}$$

Nå ser vi fra den siste likningen at vi må ha:

$$z = 4$$

Ved å sette inn det i den midterste likningen får vi:

$$y = 10 - 3 \cdot 4 = -2$$

Til slutt får vi, ved å sette inn y og z i den øverste likningen:

$$x = -5 - 2 \cdot (-2) + 2 \cdot 4 = 7$$

Systemet har altså én løsning:

$$x = 7 \qquad y = -2 \qquad z = 4$$

△

Alle likningssystemene vi skrev opp i dette eksempelet er ekvivalente med hverandre, men det siste er mye enklere å håndtere enn det vi startet med. Prosessen vi utførte for å forenkle systemet kalles *gauss-eliminasjon*, og blir nærmere beskrevet i kapittel 2.

Totalmatrisen til et system

Generelt kan et lineært likningssystem (med m likninger og n ukjente) se slik ut:

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

Når vi skal løse et slikt system, vil vi skrive opp en rekke nye systemer som er ekvivalent med dette, men som er stadig enklere. Da er det unødvendig tungvint å skrive alle de ukjente og alle +-ene og =-tegnene hver eneste gang. Den eneste informasjonen vi trenger å ha med oss gjennom utregningen er koeffisientene (a-ene) og tallene på høyresiden (b-ene).

Totalmatrisen til et likningssystem er en tabell som inneholder akkurat disse tallene:

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right]$$

Den loddrette streken inni matrisen er egentlig ikke nødvendig å ha med, men den er praktisk for å hjelpe oss med å huske at det som står til høyre for streken hører til høyre side av likningene.

Eksempel 1.6. Likningssystemet vi startet med i eksempel 1.5 har følgende totalmatrise:

$$\left[\begin{array}{ccc|c} 1 & 2 & -2 & -5 \\ 1 & 5 & 9 & 33 \\ 2 & 5 & -1 & 0 \end{array}\right]$$

△

Kapittel 2

Gausseliminasjon

Nå skal vi formalisere ideene fra kapittel 1. Vi skal se hvordan vi kan løse et hvilket som helst lineært likningssystem ved å skrive om totalmatrisen til systemet etter bestemte regler.

Reglene for hvordan totalmatrisen kan skrives om kalles *radoperasjoner*, og målet er å få en matrise som er på *trappeform*. Denne prosessen kalles *gausseliminasjon*.

Radoperasjoner

Følgende tre måter å endre en matrise på kalles *radoperasjoner*:

1. Gange alle tallene i en rad med det samme tallet (ikke 0).
2. Legge til (et multiplum av) en rad i en annen.
3. Bytte rekkefølge på radene.

Vi sier at to matriser er *radekvivalente* hvis vi kan komme fra den ene til den andre ved å utføre en eller flere radoperasjoner. Vi bruker notasjonen $M \sim N$ for å si at to matriser M og N er radekvivalente.

Eksempel 2.1. Disse matrisene er radekvivalente, siden vi får den andre matrisen fra den første ved å gange øverste rad med 4:

$$\left[\begin{array}{cc|c} 2 & 5 & 0 \\ 1 & 7 & 4 \end{array}\right] \sim \left[\begin{array}{cc|c} 8 & 20 & 0 \\ 1 & 7 & 4 \end{array}\right]$$

Merk at vi også kan gå motsatt vei: Ved å gange øverste rad i den andre matrisen med $1/4$ får vi tilbake den første matrisen.

Disse to matrisene er også radekvivalente:

$$\left[\begin{array}{cc|c} 3 & 1 & 2 \\ 9 & 5 & 7 \end{array}\right] \sim \left[\begin{array}{cc|c} 3 & 1 & 2 \\ 0 & 2 & 1 \end{array}\right]$$

Her har vi brukt den andre typen radoperasjon: Vi la til -3 ganger øverste rad i nederste rad for å komme fra den første matrisen til den andre. Merk igjen at vi også kan gå motsatt vei: Ved å legge til 3 ganger øverste rad i nederste rad, kommer vi fra den andre matrisen til den første. △

Hele poenget med radoperasjoner er at det å utføre en radoperasjon på en totalmatrise tilsvarer å skrive om likningssystemet til et nytt system som er ekvivalent med det opprinnelige. Vi formulerer dette som et teorem:

Teorem 2.2. *Hvis to likningssystemer har radekvivalente totalmatriser, så er de to likningssystemene ekvivalente.*

Bevis. For å bevise dette, er det nok å vise at det å gjøre en radoperasjon på totalmatrisen til et likningssystem tilsvarer å gjøre en gyldig omskrivning av systemet selv.

Den første typen radoperasjon – å gange alle tallene i en rad med samme tall – tilsvarer å gange med det samme tallet på begge sider av en ligning. Litt mer detaljert: La oss si at

$$a_{i1}\ a_{i2}\ \cdots\ a_{in} \mid b_i$$

er en av radene i totalmatrisen, og at vi ganger opp denne med tallet c slik at vi får:

$$(ca_{i1})\ (ca_{i2})\ \cdots\ (ca_{in}) \mid (cb_i)$$

Dette tilsvarer at vi bytter ut likningen

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i$$

med den nye likningen

$$(ca_{i1})x_1 + (ca_{i2})x_2 + \cdots + (ca_{in})x_n = cb_i.$$

Men det er klart at hvis den opprinnelige likningen var sann, så må også den nye være det. Og siden det ikke tillates at tallet c som vi ganger med er 0, så har vi også det motsatte: Hvis den nye likningen er sann, så må også den opprinnelige være det. Altså gjør vi ingen endring i løsningene av likningssystemet ved å utføre denne typen radoperasjon.

For den andre typen radoperasjon – legge til et multiplum av en rad i en annen – kan vi på tilsvarende måte se at den nye raden vi lager tilsvarer en likning som må være sann hvis de gamle likningene var sanne. Sett at vi legger til c ganger rad i i rad j. Dette tilsvarer at vi ganger opp den i-te likningen med c, og legger til resultatet i den j-te likningen. Alle løsninger av de gamle likningene må da også være løsninger av denne nye likningen. Dessuten kan vi komme tilbake til det gamle systemet (ved å legge til $-c$ ganger rad i i rad j), og dermed må alle løsninger av det nye systemet også være løsninger av det gamle.

Den tredje og siste typen radoperasjon – bytte rekkefølge på radene – gjør åpenbart ingen endringer i løsningene av likningssystemet, siden dette bare tilsvarer å skrive likningene i en annen rekkefølge. □

Eksempel 2.3. Vi gjentar regningen i eksempel 1.5, denne gangen ved å utføre radoperasjoner på totalmatrisen til likningssystemet:

$$\left[\begin{array}{ccc|c} 1 & 2 & -2 & -5 \\ 1 & 5 & 9 & 33 \\ 2 & 5 & -1 & 0 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 2 & -2 & -5 \\ 0 & 3 & 11 & 38 \\ 2 & 5 & -1 & 0 \end{array}\right]$$
$$\sim \left[\begin{array}{ccc|c} 1 & 2 & -2 & -5 \\ 0 & 3 & 11 & 38 \\ 0 & 1 & 3 & 10 \end{array}\right]$$
$$\sim \left[\begin{array}{ccc|c} 1 & 2 & -2 & -5 \\ 0 & 1 & 3 & 10 \\ 0 & 3 & 11 & 38 \end{array}\right]$$
$$\sim \left[\begin{array}{ccc|c} 1 & 2 & -2 & -5 \\ 0 & 1 & 3 & 10 \\ 0 & 0 & 2 & 8 \end{array}\right]$$

Her gjorde vi følgende radoperasjoner: Legge til -1 ganger første rad i andre rad, legge til -2 ganger første rad i tredje rad, bytte andre og tredje rad, og legge til -3 ganger andre rad i tredje rad.

Den siste matrisen her er på det som kalles trappeform, og da er det (som vi så i eksempel 1.5) lett å finne løsningen. Hvis vi vil gjøre det enda lettere, kan vi fortsette med radoperasjoner til vi oppnår det som kalles *redusert trappeform*:

$$\left[\begin{array}{ccc|c} 1 & 2 & -2 & -5 \\ 0 & 1 & 3 & 10 \\ 0 & 0 & 2 & 8 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 2 & -2 & -5 \\ 0 & 1 & 3 & 10 \\ 0 & 0 & 1 & 4 \end{array}\right]$$
$$\sim \left[\begin{array}{ccc|c} 1 & 2 & 0 & 3 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 4 \end{array}\right]$$
$$\sim \left[\begin{array}{ccc|c} 1 & 0 & 0 & 7 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 4 \end{array}\right]$$

Den siste totalmatrisen her svarer til følgende likningssystem:

$$\begin{cases} x = 7 \\ y = -2 \\ z = 4 \end{cases}$$

Her har vi altså kommet helt frem til løsningen. △

Trappeform

Vi vil nå gi en presis definisjon av begrepene «trappeform» og «redusert trappeform». Da trenger vi også et annet begrep, nemlig «pivotelement».

Definisjon. Det første tallet i en rad i en matrise som ikke er 0 kalles *pivotelementet* for den raden. (En rad med bare nuller har ikke noe pivotelement.) △

Eksempel 2.4. Se på følgende matrise:

$$\begin{bmatrix} 3 & -2 & 0 & 2 \\ 0 & 0 & 5 & 12 \\ 1 & 8 & 3 & 7 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Pivotelementene her er tallet 3 i den øverste raden, tallet 5 i den andre raden og tallet 1 i den tredje raden. Den siste raden består av bare nuller, og har derfor ikke noe pivotelement. △

Definisjon. En matrise er på *trappeform* dersom hvert pivotelement er til høyre for alle pivotelementer i tidligere rader, og eventuelle nullrader er helt nederst. △

Eksempel 2.5. Denne matrisen er på trappeform:

$$\begin{bmatrix} 3 & 7 & 6 \\ 0 & 1 & -2 \\ 0 & 0 & 2 \end{bmatrix}$$

Pivotelementene er 3, 1 og 2, og hvert av dem er til høyre for alle de tidligere pivotelementene.

Denne matrisen er også på trappeform:

$$\begin{bmatrix} -2 & 7 & 1 & 5 \\ 0 & 0 & 4 & 9 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Denne matrisen er ikke på trappeform fordi nullradene ikke er samlet nederst:

$$\begin{bmatrix} 3 & 8 & 2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Denne matrisen ser «trappete» ut, men er likevel ikke på trappeform:

$$\begin{bmatrix} 5 & 8 & 7 & 2 \\ 0 & 4 & 1 & 6 \\ 0 & 9 & 2 & 0 \\ 0 & 0 & 3 & 1 \end{bmatrix}$$

Grunnen til at den ikke er på trappeform er at pivotelementet 9 i tredje rad ikke er til høyre for pivotelementet i andre rad, men rett under det isteden. △

Definisjon. En matrise er på *redusert trappeform* hvis den er på trappeform og dessuten oppfyller:

- Alle pivotelementene er 1.
- Alle tall som står over pivotelementer er 0. △

Den siste matrisen i eksempel 2.3 er på redusert trappeform, og der så vi også hva som gjør redusert trappeform nyttig: Løsningen av systemet kan leses av direkte.

Det å skrive om en matrise til trappeform ved hjelp av radoperasjoner kalles *gausseliminasjon*, oppkalt etter den tyske matematikeren Carl Friedrich Gauss (1777–1855). Noen velger også å ha et eget navn på det å komme frem til *redusert* trappeform, og kaller den prosessen for *Gauss–Jordan-eliminasjon*, oppkalt etter Wilhelm Jordan (1842–1899). Vi tar det ikke så nøye med den forskjellen, og sier «gausseliminasjon» uansett.

(Disse begrepene er uansett historisk sett fullstendig misvisende. Metoden som vi kaller gausseliminasjon var kjent i Kina for flere tusen år siden, og Gauss – som riktignok var et universalgeni og fant opp mengder av flotte ting – har ikke egentlig så mye med den å gjøre.)

Eksistens og entydighet av løsninger

Når vi vil løse et likningssystem, er det noen åpenbare spørsmål vi kan stille:

- Har systemet noen løsning? (*Eksistens*)
- Hvis systemet har løsning: Har det også flere løsninger, eller bare én? (*Entydighet*)

I eksempel 2.3 hadde vi et system med entydig løsning. Vi tar noen flere eksempler for å vise andre ting som kan skje.

Eksempel 2.6. La oss løse følgende system:

$$\begin{cases} x - \ \ 2y = 1 \\ -5x + 10y = -1 \end{cases}$$

Vi setter opp totalmatrisen og gausseliminerer:

$$\left[\begin{array}{cc|c} 1 & -2 & 1 \\ -5 & 10 & -1 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & -2 & 1 \\ 0 & 0 & 4 \end{array}\right]$$

Den siste matrisen svarer til følgende system:

$$\begin{cases} x + 2y = 1 \\ 0x + 0y = 4 \end{cases}$$

Likningen $0x + 0y = 4$ kan også skrives som $0 = 4$, og den kan ikke stemme uansett hva vi setter x og y til å være. Dette systemet har altså ingen løsning. △

Generelt er det slik at hvis vi får en rad i totalmatrisen vår på formen

$$0 \ 0 \ \cdots \ 0 \mid b,$$

der b er et tall som ikke er 0, så har systemet ingen løsning. Denne raden svarer jo til likningen $0 = b$, som ikke kan være sann. Hvis vi har en matrise på trappeform der ingen av radene er på denne formen, så har systemet minst én løsning.

Men et lineært likningssystem kan også ha mer enn én løsning, som vi skal se i det neste eksempelet.

Eksempel 2.7. La oss løse følgende system:

$$\begin{cases} x_1 + 3x_2 + 2x_3 + 3x_4 = 16 \\ x_1 + 3x_2 + 3x_3 + \ \ x_4 = 21 \\ 2x_1 + 6x_2 + 4x_3 + 6x_4 = 32 \end{cases}$$

Vi setter opp totalmatrisen og gausseliminerer:

$$\left[\begin{array}{cccc|c} 1 & 3 & 2 & 3 & 16 \\ 1 & 3 & 3 & 1 & 21 \\ 2 & 6 & 4 & 6 & 32 \end{array}\right] \sim \left[\begin{array}{cccc|c} 1 & 3 & 2 & 3 & 16 \\ 0 & 0 & 1 & -2 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$
$$\sim \left[\begin{array}{cccc|c} 1 & 3 & 0 & 7 & 6 \\ 0 & 0 & 1 & -2 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Den siste matrisen svarer til følgende system:

$$\begin{cases} x_1 + 3x_2 + 7x_4 = 6 \\ \qquad\quad x_3 - 2x_4 = 5 \end{cases}$$

(Her har vi ikke tatt med noen likning for nullraden i matrisen. Det er fordi nullraden står for likningen $0x_1 + 0x_2 + 0x_3 + 0x_4 = 0$, eller med andre ord $0 = 0$. Denne likningen er åpenbart oppfylt uansett hva x_1, x_2, x_3 og x_4 er, så vi trenger ikke ta den med.)

Hvis vi flytter alt unntatt x_1 og x_3 til høyresiden, ser systemet slik ut:

$$\begin{cases} x_1 = -3x_2 - 7x_4 + 6 \\ x_3 = 2x_4 + 5 \end{cases}$$

Vi kan altså finne løsninger av systemet ved å sette x_2 og x_4 til å være hva vi vil, og deretter bruke disse to likhetene til å bestemme x_1 og x_3.

Hvis vi for eksempel velger $x_2 = 0$ og $x_4 = 1$, så får vi følgende løsning:

$$\begin{cases} x_1 = -3 \cdot 0 - 7 \cdot 1 + 6 = -1 \\ x_2 = 0 \\ x_3 = 2 \cdot 1 + 5 = 7 \\ x_4 = 1 \end{cases}$$

For å beskrive alle løsningene av systemet på en ryddig måte, kan vi sette $x_2 = s$ og $x_4 = t$, der s og t står for to vilkårlige tall. Da er alle løsningene gitt ved:

$$\begin{cases} x_1 = -3s - 7t + 6 \\ x_2 = s \\ x_3 = 2t + 5 \\ x_4 = t \end{cases}$$

△

Variabler som vi kan sette til hva vi vil, slik som x_2 og x_4 i eksempelet over, kalles *frie variabler.*

Når vi løser et lineært likningssystem, og har funnet ut at det har minst én løsning, så er det to muligheter. Den ene muligheten er at vi ikke får noen frie variabler (slik som i eksempel 2.3). Da har systemet entydig løsning. Den andre muligheten er at det er en eller flere frie variabler (slik som i eksempel 2.7). Da har systemet uendelig mange løsninger, siden hver av de frie variablene kan settes til å være et hvilket som helst tall.

Dette betyr at det ikke er mulig at vi får for eksempel to løsninger, eller tre løsninger, og så videre. Om det først er mer enn én løsning, må det være uendelig mange.

La oss oppsummere det vi har funnet ut om eksistens og entydighet av løsninger. For ethvert lineært likningssystem må én av følgende være sant:

- Systemet har ingen løsning.
- Systemet har entydig løsning.
- Systemet har uendelig mange løsninger.

Valgfrihet

Når vi gausseliminerer har vi en viss grad av valgfrihet. Det som står fast er hva vi har lov til å gjøre, nemlig de tre typene radoperasjoner, og hva vi vil ende opp med, nemlig (redusert) trappeform. Nøyaktig hvordan vi bruker radoperasjoner for å komme frem kan vi velge selv.

Vi tar et enkelt eksempel for å illustrere dette.

Eksempel 2.8. Anta at vi vil gausseliminere denne totalmatrisen:

$$\left[\begin{array}{cccc|c} 0 & 2 & 4 & 1 & 7 \\ 3 & 8 & 2 & 0 & 4 \\ 5 & 9 & 2 & 4 & 4 \end{array}\right]$$

Her er vi nødt til å bytte øverste rad med en av de to andre for å få pivotelementet i første rad på riktig sted. Men vi velger selv hvilken av de to radene vi vil flytte til toppen. △

Vi har også noe frihet når det gjelder valg av frie variabler.

Eksempel 2.9. I eksempel 2.7 endte vi opp med at de to variablene x_2 og x_4 var frie. Men vi kunne også ha valgt å la x_1 og x_3 være frie, som vi skal se nå.

Vi hadde forenklet systemet til følgende:

$$\begin{cases} x_1 = -3x_2 - 7x_4 + 6 \\ x_3 = 2x_4 + 5 \end{cases}$$

Vi kan løse den andre likningen her for x_4 og få:

$$x_4 = \frac{x_3 - 5}{2}$$

Deretter kan vi sette inn dette i den første likningen og løse for x_2:

$$\begin{aligned} x_2 &= \frac{-x_1 - 7x_4 + 6}{3} \\ &= \frac{-x_1 - \frac{7}{2}(x_3 - 5) + 6}{3} \\ &= -\frac{1}{3}x_1 - \frac{7}{6}x_3 + \frac{47}{6} \end{aligned}$$

Hvis vi nå lar $x_1 = s$ og $x_3 = t$, så har vi følgende generelle løsning:

$$\begin{cases} x_1 = s \\ x_2 = -\dfrac{1}{3}s - \dfrac{7}{6}t + \dfrac{47}{6} \\ x_3 = t \\ x_4 = \dfrac{1}{2}x_3 - \dfrac{5}{2} \end{cases}$$

Dette ser annerledes ut enn det vi fikk i eksempel 2.7, men det beskriver nøyaktig de samme løsningene. (Du kan for eksempel sjekke at hvis vi her setter $s = -1$ og $t = 7$, så får vi den samme løsningen som da vi valgte $x_2 = 0$ og $x_4 = 1$ i eksempel 2.7). △

Kapittel 3

Vektor- og matriselikninger

I denne uken skal vi bruke enkel vektorregning til å analysere lineære likningssystemer. Vi skal ha et spesielt fokus på $\mathbb{R}^3$, for det går an å visualisere; klarer man det, går det lettere å abstrahere til $\mathbb{R}^n$. Senere i kurset skal vi se hvordan noen av konseptene under kan generaliseres, slik at vi kan konstruere teori som kan behandle matematiske emner som tilsynelatende ser veldig forskjellige ut, men følger akkurat de samme lovene.

Vektorregning

Inntil videre skal vi skrive vektorer på høykant

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix},$$

kalt søylevektor. Du kan også tenke på dette som et punkt i $\mathbb{R}^n$. De to viktigste regnereglene for vektorer er skalarmultiplikasjon

$$a\mathbf{x} = \begin{bmatrix} ax_1 \\ ax_2 \\ \vdots \\ ax_n \end{bmatrix}$$

og vektoraddisjon

$$\mathbf{x} + \mathbf{y} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{bmatrix}.$$

En sammensetning av disse to operasjonene

$$a\mathbf{x} + b\mathbf{y} = a\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + b\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} ax_1 + by_1 \\ ax_2 + by_2 \\ \vdots \\ ax_n + by_n \end{bmatrix}$$

kalles en *lineærkombinasjon*. Skalarene a og b kalles vekter. Hvis vi har m vektorer $\mathbf{x}_k$, definerer vi *det lineære spennet*, eller

$$\mathrm{Sp}\{\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_m\}$$

som alle lineærkombinasjoner av vektorene, altså alle vektorer på formen

$$a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + ... + a_m\mathbf{x}_m.$$

Eksempel 3.1.

$$3\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + 2\begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 11 \\ 16 \\ 21 \end{bmatrix}. \qquad \triangle$$

Eksempel 3.2. Spennet til vektorene i eksemplet over, er alle vektorer på formen

$$a\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + b\begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}. \qquad \triangle$$

Vektorlikninger

Ved å ta i bruk lineærkombinasjon, kan vi skrive likningssystemet fra forrige uke

$$\begin{cases} x + 2y - 2z = -5 \\ x + 5y + 9z = 33 \\ 2x + 5y - z = 0 \end{cases}$$

som vektorlikningen

$$x\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} + y\begin{bmatrix} 2 \\ 5 \\ 5 \end{bmatrix} + z\begin{bmatrix} -2 \\ 9 \\ -1 \end{bmatrix} = \begin{bmatrix} -5 \\ 33 \\ 0 \end{bmatrix}.$$

Dette gir oss en ny måte å se likningssystemer på: oppgaven er å finne vektene x, y og z slik at søylene i matrisen lineærkombineres til å bli lik høyresiden.

Eksempel 3.3. Løsningen til systemet over er

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 7 \\ -2 \\ 4 \end{bmatrix}$$

og du kan verifisere at

$$7\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} - 2\begin{bmatrix} 2 \\ 5 \\ 5 \end{bmatrix} + 4\begin{bmatrix} -2 \\ 9 \\ -1 \end{bmatrix} = \begin{bmatrix} -5 \\ 33 \\ 0 \end{bmatrix}. \qquad \triangle$$

Matriselikninger

Produktet av en $n \times n$-matrise og en søylevektor i $\mathbb{R}^n$ defineres som følgende lineærkombinasjon av matrisens søyler

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} =$$

$$x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{nn} \end{bmatrix}.$$

Eksempel 3.4.

$$\begin{bmatrix} 1 & 2 & -2 \\ 1 & 5 & 9 \\ 2 & 5 & -1 \end{bmatrix} \begin{bmatrix} 7 \\ -2 \\ 4 \end{bmatrix} = 7 \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} - 2 \begin{bmatrix} 2 \\ 5 \\ 5 \end{bmatrix} + 4 \begin{bmatrix} -2 \\ 9 \\ -1 \end{bmatrix}. \quad \triangle$$

Nå kan vi skrive likningssystemet fra forrige uke som

$$\begin{bmatrix} 1 & 2 & -2 \\ 1 & 5 & 9 \\ 2 & 5 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -5 \\ 33 \\ 0 \end{bmatrix}.$$

Dersom vi skriver

$$A = \begin{bmatrix} 1 & 2 & -2 \\ 1 & 5 & 9 \\ 2 & 5 & -1 \end{bmatrix},$$

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

og

$$\mathbf{b} = \begin{bmatrix} -5 \\ 33 \\ 0 \end{bmatrix},$$

kan vi innføre den kompakte notasjonen

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

Eksistens og entydighet av løsninger II

likningssystemer deler seg naturlig i tre kategorier; de som har en unik løsning, de som har ingen løsning, og de som har uendelig mange løsninger. Vi skal nå gi en geometrisk illustrasjon av hva som skjer i de forskjellige tilfellene.

Eksempel 3.5. Hvis vi utfører Gauss-eliminasjon på systemet

$$\begin{bmatrix} 2 & 3 & 4 \\ 3 & 4 & 5 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \\ 3 \end{bmatrix},$$

får vi

$$\begin{bmatrix} 2 & 3 & 4 \\ 0 & 1 & 2 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \\ 5 \end{bmatrix}.$$

De to nederste linjene sier at x_2+2x_3 skal være både 2 og 5. Dette er åpenbart umulig, og systemet har ingen løsning. Grunnen er at søylene ligger i samme i plan i $\mathbb{R}^3$, og siden høyresiden ikke ligger i dette planet, er det umulig å skrive den som en lineærkombinasjon av disse vektorene. △

Eksempel 3.6. Hvis vi derimot utfører Gauss-eliminasjon på systemet

$$\begin{bmatrix} 2 & 3 & 4 \\ 3 & 4 & 5 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 9 \\ 12 \\ 15 \end{bmatrix},$$

får vi

$$\begin{bmatrix} 2 & 3 & 4 \\ 0 & 1 & 2 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 9 \\ 3 \\ 3 \end{bmatrix},$$

Nå er to nederste linjene identiske. Søylene i matrisen ligger som kjent i samme plan i $\mathbb{R}^3$, men nå ligger tilfeldigvis høyresiden også i dette planet, og systemet kan derfor løses. Hvis du ønsker å skrive en vektor i et plan som en lineærkombinasjon av tre andre vektorer i samme plan, har du uendelig mange måter å gjøre det på, og derfor har likningssystemet uendelig mange løsninger. △

Eksempel 3.7. likningssystemet

$$\begin{bmatrix} 1 & 2 & -2 \\ 1 & 5 & 9 \\ 2 & 5 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -5 \\ 33 \\ 0 \end{bmatrix}.$$

har som kjent en unik løsning. Vi sier at søylene *spenner ut* $\mathbb{R}^3$, siden alle punkter i $\mathbb{R}^3$ kan skrives som en unik lineærkombinasjon av dem. Merk at søylene i matrisen danner et parallellepiped med volum 2. △

Du kan avgjøre hvorvidt tre vektorer i $\mathbb{R}^3$ ligger i samme plan ved å beregne volumet til parallellepipedet spent ut av de tre vektorene. Dersom volumet blir 0, ligger de i samme plan. Dersom søylene i matrisen A kalles $\mathbf{a}_1$, $\mathbf{a}_2$ og $\mathbf{a}_3$, kalles dette volumet *determinanten* til A, og er gitt ved

$$\det A = \mathbf{a}_1 \cdot \mathbf{a}_2 \times \mathbf{a}_3.$$

Eksempel 3.8.

$$\det \begin{bmatrix} 2 & 3 & 4 \\ 3 & 4 & 5 \\ 4 & 5 & 6 \end{bmatrix} = 0. \qquad \triangle$$

Presise kriterier for når et likningssystem har én, ingen, eller mange løsninger, får vi ikke uten litt mer matematisk maskineri. Men et mentalt bilde av 3×3-systemer kan vi lage oss.

- Hvis matrisens søyler danner et parallellepiped har systemet en unik løsning uansett høyreside.
- Hvis matrisens søyler ligger i samme plan, og høyresiden ikke ligger i dette planet, har systemet ingen løsning.
- Hvis matrisens søyler ligger i samme plan, og høyresiden ligger i dette planet, har systemet uendelig mange løsninger.

En forsmak på lineær uavhengighet

Hvis man har en samling vektorer, sier vi at de er lineært avhengige dersom en av vektorene i samlingen kan skrives som en lineærkombinasjon av de andre, for eksempel dersom tre vektorer i $\mathbb{R}^3$ ligger i samme plan.

Eksempel 3.9. Hvis du utfører gausseliminasjon på systemet

$$\begin{bmatrix} 2 & 3 & 4 & 0 \\ 3 & 4 & 5 & 0 \\ 4 & 5 & 6 & 0 \end{bmatrix}$$

får du

$$\begin{bmatrix} 2 & 3 & 4 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

altså at

$$\begin{cases} 2x + 3y + 4z = 0 \\ \qquad\quad y + 2z = 0 \end{cases}.$$

setter vi $z = s$, får vi $y = -2s$ av den siste likningen, og $x = s$ av den første, slik at

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = s \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$$

er en løsning av systemet for vilkårlige s. Dette betyr at søylene i den opprinnelige matrisen er lineært avhengige. Vi dobbeltsjekker:

$$\begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} - 2 \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix} + \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \qquad \triangle$$

Kapittel 4

Matriser

Nå har vi fått erfaring med å bruke matriser i et par forskjellige sammenhenger. Vi har lært å løse et lineært likningssystem ved å sette opp totalmatrisen til systemet og gausseliminere den (ved hjelp av radoperasjoner på matrisen), og vi har sett at et lineært likningssystem kan skrives på formen

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

der vi har samlet alle koeffisientene i matrisen A. For å få denne likningen til å gi mening, definerte vi hvordan vi kan gange en $m \times n$-matrise med en vektor i $\mathbb{R}^n$ og få ut en vektor i $\mathbb{R}^m$.

I tillegg til dette matrise–vektor-produktet finnes det en hel rekke andre aritmetiske operasjoner vi kan utføre på matriser. I dette kapitlet tar vi en grundig gjennomgang av disse operasjonene.

Definisjoner og notasjon

En *$m \times n$-matrise* er en rektangulær tabell med tall som har m tall i høyden og n tall i bredden:

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

Kolonnene i matrisen er følgende kolonnevektorer:

$$\begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} \quad \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} \quad \cdots \quad \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}$$

Radene i matrisen er følgende radvektorer:

$$\begin{gathered} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \end{bmatrix} \\ \begin{bmatrix} a_{21} & a_{22} & \cdots & a_{2n} \end{bmatrix} \\ \vdots \\ \begin{bmatrix} a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \end{gathered}$$

Eksempel 4.1. Her er et eksempel på en 2×3-matrise:

$$\begin{bmatrix} 5 & 0 & -2 \\ 3 & 1 & 4 \end{bmatrix}$$

Kolonnene i denne matrisen er:

$$\begin{bmatrix} 5 \\ 3 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \text{og} \quad \begin{bmatrix} -2 \\ 4 \end{bmatrix}$$

Radene er:

$$\begin{bmatrix} 5 & 0 & -2 \end{bmatrix} \quad \text{og} \quad \begin{bmatrix} 3 & 1 & 4 \end{bmatrix} \qquad \triangle$$

Noen ganger har vi en liste med kolonnevektorer, si

$$\mathbf{v}_1,\ \mathbf{v}_2,\ \ldots,\ \mathbf{v}_n,$$

og vil lage en matrise som har disse vektorene som kolonner. Den matrisen kan vi skrive slik:

$$\begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \end{bmatrix}$$

Hvis vektorene ligger i $\mathbb{R}^m$, blir dette en $m \times n$-matrise.

Eksempel 4.2. La

$$\mathbf{v}_1 = \begin{bmatrix} 2 \\ 0 \\ 4 \end{bmatrix} \quad \text{og} \quad \mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$$

være to vektorer i $\mathbb{R}^3$. Matrisen $\begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{bmatrix}$ med disse vektorene som kolonner blir da følgende 3×2-matrise:

$$\begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 0 \\ 2 & 4 \end{bmatrix} \qquad \triangle$$

På samme måte kan vi, hvis vi har en liste med radvektorer

$$\mathbf{r}_1,\ \mathbf{r}_2,\ \ldots,\ \mathbf{r}_m,$$

lage en matrise

$$\begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \vdots \\ \mathbf{r}_m \end{bmatrix}$$

med disse vektorene som rader.

Eksempel 4.3. La

$$\mathbf{r}_1 = \begin{bmatrix} 2 & 1 \end{bmatrix}, \quad \mathbf{r}_2 = \begin{bmatrix} 1 & 0 \end{bmatrix} \quad \text{og} \quad \mathbf{r}_3 = \begin{bmatrix} 2 & 4 \end{bmatrix}$$

være tre radvektorer. Da har vi:

$$\begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \mathbf{r}_3 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 0 \\ 2 & 4 \end{bmatrix} \qquad \triangle$$

Noen ganger vil vi bruke tilsvarende notasjon for å bygge opp en matrise av andre matriser, eller av en kombinasjon av matriser og vektorer.

Eksempel 4.4. La A og B være matriser, og $\mathbf{v}$ en vektor, slik som dette:

$$A = \begin{bmatrix} 3 & 1 \\ 0 & 2 \\ 7 & 1 \end{bmatrix} \qquad B = \begin{bmatrix} 5 & 9 & 2 \\ -1 & 5 & 4 \\ 10 & 0 & 3 \end{bmatrix} \qquad \mathbf{v} = \begin{bmatrix} 6 \\ 8 \\ -5 \end{bmatrix}$$

Da kan vi skrive $\begin{bmatrix} A & B & \mathbf{v} \end{bmatrix}$ for matrisen som inneholder alle tallene fra disse tre satt ved siden av hverandre:

$$\begin{bmatrix} A & B & \mathbf{v} \end{bmatrix} = \begin{bmatrix} 3 & 1 & 5 & 9 & 2 & 6 \\ 0 & 2 & -1 & 5 & 4 & 8 \\ 7 & 1 & 10 & 0 & 3 & -5 \end{bmatrix} \qquad \triangle$$

En $n \times n$-matrise, altså en matrise med like mange rader og kolonner, kaller vi for en *kvadratisk* matrise. For eksempel er matrisen B i eksempel 4.4 en kvadratisk matrise, mens A ikke er det.

Produkt av matrise og vektor

La

$$A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix}$$

være en $m \times n$-matrise med vektorene $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ som kolonner, og la

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

være en vektor i $\mathbb{R}^n$. Vi definerer produktet $A\mathbf{v}$ av A og $\mathbf{v}$ som lineærkombinasjonen av kolonnene i A med tallene i $\mathbf{v}$ som vekter:

$$A\mathbf{v} = \mathbf{a}_1 v_1 + \mathbf{a}_2 v_2 + \cdots + \mathbf{a}_n v_n$$

Merk at produktet $A\mathbf{v}$ bare er definert når bredden av matrisen A er lik høyden av vektoren $\mathbf{v}$.

Eksempel 4.5. Vi regner ut produktet av en 2×3-matrise og en vektor i $\mathbb{R}^3$:

$$\begin{aligned} \begin{bmatrix} 5 & 0 & -2 \\ 3 & 1 & 4 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix} &= \begin{bmatrix} 5 \\ 3 \end{bmatrix} \cdot 2 + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \cdot (-1) + \begin{bmatrix} -2 \\ 4 \end{bmatrix} \cdot 3 \\ &= \begin{bmatrix} 5 \cdot 2 + 0 \cdot (-1) + (-2) \cdot 3 \\ 3 \cdot 2 + 1 \cdot (-1) + 4 \cdot 3 \end{bmatrix} \\ &= \begin{bmatrix} 4 \\ 17 \end{bmatrix} \end{aligned}$$

Merk at resultatet blir en vektor i $\mathbb{R}^2$. $\triangle$

Når vi skal regne ut et produkt $A\mathbf{v}$ av en matrise og en vektor, trenger vi ikke egentlig å skrive opp lineærkombinasjonen av kolonnene i A med vekter fra $\mathbf{v}$ slik som vi gjorde i eksempelet over. Den andre linjen av utregningen i eksempelet viser en mer direkte måte å komme frem på: Tallet som skal være på første posisjon i resultatvektoren får vi ved å gange tallene fra første rad i A med tallene i $\mathbf{v}$, og legge sammen resultatene. Tallet på andre posisjon i resultatvektoren får vi på samme måte fra andre rad i A.

Generelt har vi at dersom

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \quad \text{og} \quad \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix},$$

så kan vi regne ut produktet $A\mathbf{v}$ på følgende måte:

$$A\mathbf{v} = \begin{bmatrix} a_{11}v_1 + a_{12}v_2 + \cdots + a_{1n}v_n \\ a_{21}v_1 + a_{22}v_2 + \cdots + a_{2n}v_n \\ \vdots \\ a_{m1}v_1 + a_{m2}v_2 + \cdots + a_{mn}v_n \end{bmatrix}$$

(Det er ikke vanskelig å se at dette følger fra definisjonen av $A\mathbf{v}$.)

Teorem 4.6. *Hvis A er en $m \times n$-matrise, $\mathbf{v}$ og $\mathbf{w}$ er vektorer i $\mathbb{R}^n$ og c er et tall, så har vi følgende likheter:*

$$A(\mathbf{v} + \mathbf{w}) = A\mathbf{v} + A\mathbf{w} \qquad \text{og} \qquad A(c\mathbf{v}) = c(A\mathbf{v})$$

Eksempel 4.7. La A og $\mathbf{v}$ være følgende matrise og vektor:

$$A = \begin{bmatrix} 5 & 0 & -2 \\ 3 & 1 & 4 \end{bmatrix} \quad \text{og} \quad \mathbf{v} = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}$$

I eksempel 4.5 regnet vi ut produktet $A\mathbf{v}$. Men hva om vi har lyst til å gange A med vektoren $(8, -4, 12)$? Siden denne vektoren er lik $4 \cdot \mathbf{v}$ kan vi bruke teorem 4.6 og få:

$$A \begin{bmatrix} 8 \\ -4 \\ 12 \end{bmatrix} = A \cdot (4 \cdot \mathbf{v}) = 4 \cdot (A \cdot \mathbf{v}) = 4 \cdot \begin{bmatrix} 4 \\ 17 \end{bmatrix} = \begin{bmatrix} 16 \\ 68 \end{bmatrix}$$

Altså: A ganget med vektoren $4\mathbf{v}$ er det samme som 4 ganger vektoren $A\mathbf{v}$. $\triangle$

Det er verdt å merke seg hva som skjer hvis vi ganger en matrise med en vektor der nøyaktig ett av tallene er 1, og resten er 0. La oss teste dette med en eksempelmatrise:

Eksempel 4.8. Vi ganger 2×3-matrisen

$$A = \begin{bmatrix} 5 & 0 & -2 \\ 3 & 1 & 4 \end{bmatrix}$$

med de tre vektorene

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad \text{og} \quad \mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

og får følgende:

$$A\mathbf{e}_1 = \begin{bmatrix} 5 \\ 3 \end{bmatrix} \qquad A\mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad A\mathbf{e}_3 = \begin{bmatrix} -2 \\ 4 \end{bmatrix}$$

Resultatene ble altså de tre kolonnene i A. $\triangle$

Generelt har vi at hvis A er en $m \times n$-matrise, og $\mathbf{e}_1$, $\mathbf{e}_2, \ldots, \mathbf{e}_n$ er vektorene i $\mathbb{R}^n$ som er slik at $\mathbf{e}_i$ har et 1-tall i sin i-te koordinat og bare 0-er ellers, så er

$$A\mathbf{e}_1, \quad A\mathbf{e}_2, \quad \ldots, \quad A\mathbf{e}_n$$

nøyaktig samme vektorer som kolonnene i A.

Sum og skalering av matriser

La A og B være to $m \times n$-matriser:

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \qquad B = \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ b_{21} & \cdots & b_{2n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mn} \end{bmatrix}$$

Vi definerer summen $A + B$ på den mest åpenbare måten – vi legger sammen tallene fra de to matrisene i hver posisjon, og får en ny $m \times n$-matrise:

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

Produktet av et tall og en matrise defineres også på den åpenbare måten – vi ganger med tallet på hver plass i matrisen:

$$cA = \begin{bmatrix} c \cdot a_{11} & c \cdot a_{12} & \cdots & c \cdot a_{1n} \\ c \cdot a_{21} & c \cdot a_{22} & \cdots & c \cdot a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c \cdot a_{m1} & c \cdot a_{m2} & \cdots & c \cdot a_{mn} \end{bmatrix}$$

Teorem 4.9. *Hvis A og B er $m \times n$-matriser, $\mathbf{v}$ er en vektor i $\mathbb{R}^n$ og c er et tall, så har vi følgende likheter:*

$$(A + B)\mathbf{v} = A\mathbf{v} + B\mathbf{v} \qquad \textit{og} \qquad (cA)\mathbf{v} = c(A\mathbf{v})$$

Vektorer som matriser

Hittil har vi snakket om vektorer og matriser som to forskjellige slags ting, men vi kan også velge å se på vektorer som et spesialtilfelle av matriser der enten høyden eller bredden er 1.

En kolonnevektor

$$\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{bmatrix}$$

kan vi tenke på som en $m \times 1$-matrise, og en radvektor

$$\begin{bmatrix} w_1 & w_2 & \cdots & w_n \end{bmatrix}$$

kan vi tenke på som en $1 \times n$-matrise.

Eksempel 4.10. La $\mathbf{v}$ og $\mathbf{w}$ være henholdsvis en kolonnevektor og en radvektor:

$$\mathbf{v} = \begin{bmatrix} 5 \\ -2 \\ 4 \end{bmatrix} \qquad \mathbf{w} = \begin{bmatrix} 2 & 3 & 1 \end{bmatrix}$$

Vektoren $\mathbf{v}$ kan vi også se på som en 3×1-matrise, og vektoren $\mathbf{w}$ kan vi se på som en 1×3-matrise.

Hvis vi velger å tenke på $\mathbf{w}$ som en matrise og $\mathbf{v}$ som en vektor, så kan vi gange dem sammen med den vanlige regelen for produkt av matrise og vektor:

$$\begin{aligned} \mathbf{w} \cdot \mathbf{v} &= \begin{bmatrix} 2 & 3 & 1 \end{bmatrix} \cdot \begin{bmatrix} 5 \\ -2 \\ 4 \end{bmatrix} \\ &= \begin{bmatrix} 2 \cdot 5 + 3 \cdot (-2) + 1 \cdot 4 \end{bmatrix} = \begin{bmatrix} 8 \end{bmatrix} \end{aligned}$$

Resultatet blir vektoren $\begin{bmatrix} 8 \end{bmatrix}$ i $\mathbb{R}^1$. En vektor i $\mathbb{R}^1$ består av kun ett tall, og vi vil vanligvis si at en slik vektor er det samme som det ene tallet. Med andre ord kan vi sløyfe klammene og ganske enkelt skrive:

$$\mathbf{w} \cdot \mathbf{v} = 8. \qquad \triangle$$

Generelt har vi at produktet av en radvektor og en kolonnevektor er gitt ved følgende uttrykk:

$$\begin{bmatrix} w_1 & w_2 & \cdots & w_n \end{bmatrix} \cdot \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = w_1 v_1 + w_2 v_2 + \cdots + w_n v_n$$

(Merk at det må være like mange tall i radvektoren som i kolonnevektoren for at vi skal kunne gange dem sammen.)

Matrisemultiplikasjon

Nå kommer vi til den mest spennende av de aritmetiske operasjonene vi kan gjøre med matriser, nemlig *matrisemultiplikasjon*.

Det er litt mer komplisert å beskrive hvordan vi multipliserer matriser enn hvordan vi summerer dem. Det er imidlertid en god grunn til at det er slik. Vi kunne valgt å definere multiplikasjon av matriser på tilsvarende måte som sum, men det ville ikke blitt spesielt nyttig. Vi vil nemlig at matrisemultiplikasjon skal oppføre seg pent sammen med multiplikasjon av matriser med vektorer. Spesielt vil vi at følgende likhet skal holde:

$$(AB)\mathbf{v} = A(B\mathbf{v})$$

Vi vil altså at vi fritt skal kunne flytte parentesene, akkurat slik vi kan gjøre med et produkt av tre tall.

Hvordan kan vi definere produkt av matriser slik at dette fungerer? La oss først se på et eksempel.

Eksempel 4.11. La A og B være følgende to 2×2-matriser:

$$A = \begin{bmatrix} 3 & -5 \\ 2 & 7 \end{bmatrix} \qquad B = \begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix}$$

Vi har lyst til å finne matrisen AB som skal være slik at $(AB)\mathbf{v} = A(B(\mathbf{v}))$ for alle vektorer $\mathbf{v}$ i $\mathbb{R}^2$.

Nå er det lurt å se på de to spesielle vektorene

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{og} \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Vi vet fra tidligere at hvis vi ganger en 2×2-matrise med en av disse vektorene, så får vi ut den første eller den andre kolonnen i matrisen.

Vi regner ut:

$$B\begin{bmatrix}1\\0\end{bmatrix} = \begin{bmatrix}4\\2\end{bmatrix}$$
$$A\left(B\begin{bmatrix}1\\0\end{bmatrix}\right) = \begin{bmatrix}3 & -5\\2 & 7\end{bmatrix}\begin{bmatrix}4\\2\end{bmatrix} = \begin{bmatrix}2\\22\end{bmatrix}$$

Det vi er ute etter er at AB skal være slik at

$$(AB)\mathbf{v} = A(B\mathbf{v})$$

for alle vektorer $\mathbf{v}$. Spesielt må vi da ha:

$$(AB)\begin{bmatrix}1\\0\end{bmatrix} = A\left(B\begin{bmatrix}1\\0\end{bmatrix}\right) = \begin{bmatrix}2\\22\end{bmatrix}$$

Men det å gange med vektoren

$$\begin{bmatrix}1\\0\end{bmatrix}$$

er det samme som å plukke ut første kolonne av matrisen, så vi har nå funnet ut at første kolonne i matrisen AB må være:

$$\begin{bmatrix}2\\22\end{bmatrix}$$

På samme måte finner vi andre kolonne i AB:

$$B\begin{bmatrix}0\\1\end{bmatrix} = \begin{bmatrix}3\\1\end{bmatrix}$$
$$A\left(B\begin{bmatrix}0\\1\end{bmatrix}\right) = \begin{bmatrix}3 & -5\\2 & 7\end{bmatrix}\begin{bmatrix}3\\1\end{bmatrix} = \begin{bmatrix}4\\13\end{bmatrix}$$
$$(AB)\begin{bmatrix}0\\1\end{bmatrix} = A\left(B\begin{bmatrix}0\\1\end{bmatrix}\right) = \begin{bmatrix}4\\13\end{bmatrix}$$

Dette betyr at andre kolonne i AB må være:

$$\begin{bmatrix}4\\13\end{bmatrix}$$

Dermed kommer vi frem til at produktet av A og B er:

$$AB = \begin{bmatrix}2 & 4\\22 & 13\end{bmatrix} \qquad \triangle$$

La oss nå generalisere det vi gjorde i eksempelet. Foreløpig ser vi på generelle 2×2-matriser, og så tar vi det helt generelle tilfellet, med matriser av vilkårlig størrelse, etterpå.

La A og B være to 2×2-matriser, og la

$$\mathbf{e}_1 = \begin{bmatrix}1\\0\end{bmatrix} \quad \text{og} \quad \mathbf{e}_2 = \begin{bmatrix}0\\1\end{bmatrix}$$

være de to spesielle vektorene vi brukte i eksempelet over. La $\mathbf{b}_1$ og $\mathbf{b}_2$ være kolonnene i B, slik at vi har:

$$B = \begin{bmatrix}\mathbf{b}_1 & \mathbf{b}_2\end{bmatrix}$$

På samme måte som i eksempelet får vi nå:

$$(AB)\mathbf{e}_1 = A(B\mathbf{e}_1) = A\mathbf{b}_1$$
$$(AB)\mathbf{e}_2 = A(B\mathbf{e}_2) = A\mathbf{b}_2$$

Dette betyr at første kolonne i matrisen AB må være $A\mathbf{b}_1$, og andre kolonne må være $A\mathbf{b}_2$. Vi får altså:

$$AB = \begin{bmatrix}A\mathbf{b}_1 & A\mathbf{b}_2\end{bmatrix}$$

Hvis vi lar $\mathbf{a}_1$ og $\mathbf{a}_2$ være radene i A, så gir dette oss at:

$$AB = \begin{bmatrix}\mathbf{a}_1\mathbf{b}_1 & \mathbf{a}_1\mathbf{b}_2\\ \mathbf{a}_2\mathbf{b}_1 & \mathbf{a}_2\mathbf{b}_2\end{bmatrix}$$

For å virkelig gjøre dette detaljert, kan vi skrive opp nøyaktig hvordan vi finner hvert tall i AB ut fra hvert enkelt av tallene i A og B. Hvis

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

så får vi:

$$AB = \begin{bmatrix}a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22}\\a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{21}b_{22}\end{bmatrix}$$

Merk hvordan dette siste uttrykket er bygd opp. Når vi skal finne tallet som skal stå på en bestemt posisjon i AB, går vi bortover den tilsvarende *raden* i A og samtidig nedover den tilsvarende *kolonnen* i B. Vi ganger sammen tallene vi finner i A med de vi finner i B, og legger sammen disse produktene.

Alt det vi gjorde nå fungerer helt tilsvarende når vi går til større matriser enn 2×2. Men for at det skal gå an å gange sammen to matriser A og B, må de være «kompatible» i størrelse. Vi finner produktet AB ved å kombinere rader fra A med kolonner fra B, og for at dette skal gå an, må lengden av radene i A være lik lengden av kolonnene i B. Det vil si at bredden til A må være lik høyden til B.

Basert på det vi har gjort nå lager vi en generell definisjon av matrisemultiplikasjon.

Definisjon. La A være en $m \times n$-matrise med rader $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$, og la B være en $n \times p$-matrise med kolonner $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_p$:

$$A = \begin{bmatrix}\mathbf{a}_1\\\mathbf{a}_2\\\vdots\\\mathbf{a}_m\end{bmatrix} \qquad B = \begin{bmatrix}\mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_p\end{bmatrix}$$

Da er produktet av A og B en $m \times p$-matrise definert ved:

$$AB = \begin{bmatrix}\mathbf{a}_1\mathbf{b}_1 & \mathbf{a}_1\mathbf{b}_2 & \cdots & \mathbf{a}_1\mathbf{b}_p\\ \mathbf{a}_2\mathbf{b}_1 & \mathbf{a}_2\mathbf{b}_2 & \cdots & \mathbf{a}_2\mathbf{b}_p\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{a}_m\mathbf{b}_1 & \mathbf{a}_m\mathbf{b}_2 & \cdots & \mathbf{a}_m\mathbf{b}_p\end{bmatrix} \qquad \triangle$$

Hvis vi lar A og B være som i definisjonen, kan vi også skrive produktet slik:

$$AB = \begin{bmatrix}A\mathbf{b}_1 & A\mathbf{b}_2 & \cdots & A\mathbf{b}_p\end{bmatrix}$$

Eksempel 4.12. La A og B og C være følgende matriser:

$$A = \begin{bmatrix}5 & 0 & -2\\3 & 1 & 4\end{bmatrix} \qquad B = \begin{bmatrix}2 & 1\\1 & 0\\2 & 4\end{bmatrix} \qquad C = \begin{bmatrix}1 & 2\\3 & 4\end{bmatrix}$$

Siden A er en 2×3-matrise og B er en 3×2-matrise, er AB en 2×2-matrise. Vi regner ut denne matrisen ved å bruke definisjonen:

$$
\begin{aligned}
AB &= \begin{bmatrix} 5 \cdot 2 + 0 \cdot 1 + (-2) \cdot 2 & 5 \cdot 1 + 0 \cdot 0 + (-2) \cdot 4 \\ 3 \cdot 2 + 1 \cdot 1 + 4 \cdot 2 & 3 \cdot 1 + 1 \cdot 0 + 4 \cdot 4 \end{bmatrix} \\
&= \begin{bmatrix} 6 & -3 \\ 15 & 19 \end{bmatrix}
\end{aligned}
$$

Hvis vi ganger sammen de samme to matrisene i motsatt rekkefølge, får vi en 3×3-matrise:

$$
\begin{aligned}
BA &= \begin{bmatrix} 2 \cdot 5 + 1 \cdot 3 & 2 \cdot 0 + 1 \cdot 1 & 2 \cdot (-2) + 1 \cdot 4 \\ 1 \cdot 5 + 0 \cdot 3 & 1 \cdot 0 + 0 \cdot 1 & 1 \cdot (-2) + 0 \cdot 4 \\ 2 \cdot 5 + 4 \cdot 3 & 2 \cdot 0 + 4 \cdot 1 & 2 \cdot (-2) + 4 \cdot 4 \end{bmatrix} \\
&= \begin{bmatrix} 13 & 1 & 0 \\ 5 & 0 & -2 \\ 22 & 4 & 12 \end{bmatrix}
\end{aligned}
$$

Produktet av matrisene B og C blir en 3×2-matrise:

$$
BC = \begin{bmatrix} 5 & 8 \\ 1 & 2 \\ 14 & 20 \end{bmatrix}
$$

Men produktet CB er ikke definert, siden C er en 2×2-matrise og B en 3×2-matrise.

Vi kunne også regnet ut for eksempel CA, men AC er ikke definert. △

Vi kan nå legge merke til at matrisemultiplikasjon – i motsetning til multiplikasjon av vanlige tall – ikke er kommutativt. Det vil si at faktorenes rekkefølge spiller en rolle: AB er ikke nødvendigvis det samme som BA.

Vi må altså passe på at vi ikke av gammel vane bytter om faktorene når vi jobber med et produkt av matriser. Men mange andre regneregler fungerer like bra med matriser som med tall. Vi tar et teorem med noen regneregler for matrisemultiplikasjon.

Teorem 4.13. *La A, B og C være matriser, $\mathbf{v}$ en vektor, og c et tall. I hver del av teoremet antar vi at størrelsene på matrisene og vektoren er slik at alle operasjonene som brukes er definert.*

(a) Matrisemultiplikasjon er en assosiativ operasjon:

$$A(BC) = (AB)C$$

Et spesialtilfelle av dette er følgende:

$$(AB)\mathbf{v} = A(B\mathbf{v})$$

(b) Å skalere et matriseprodukt er det samme som å skalere én av faktorene og deretter multiplisere:

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

(c) Matrisemultiplikasjon distribuerer over addisjon av matriser:

$$A(B+C) = AB+AC \quad \textit{og} \quad (A+B)C = AC+BC$$

Transponering

Vi har hittil snakket om aritmetiske operasjoner på matriser som tilsvarer operasjoner vi kan gjøre med tall: addisjon og multiplikasjon. Operasjonen *transponering*, derimot, er helt spesiell for matriser, og går ut på at vi bytter om rader og kolonner.

Definisjon. La

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

være en $m \times n$-matrise. Den *transponerte* av A er $n \times m$-matrisen

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

der radene og kolonnene i A er byttet om. △

Eksempel 4.14. Hvis vi lar A være matrisen

$$
A = \begin{bmatrix} 5 & 0 & -2 \\ 3 & 1 & 4 \end{bmatrix},
$$

så er den transponerte av A gitt ved:

$$
A^\top = \begin{bmatrix} 5 & 3 \\ 0 & 1 \\ -2 & 4 \end{bmatrix}
$$

Hvis vi transponerer denne matrisen igjen, så kommer vi tilbake til utgangspunktet:

$$(A^\top)^\top = A$$ △

Vi tar med noen regneregler for transponering.

Teorem 4.15. *For enhver matrise A har vi at å transponere to ganger gir den opprinnelige matrisen:*

$$(A^\top)^\top = A$$

Hvis A og B er matriser sik at produktet AB er definert, så er den transponerte av produktet lik produktet av de transponerte, i motsatt rekkefølge:

$$(AB)^\top = B^\top \cdot A^\top$$

Identitetsmatriser

Når vi ser på multiplikasjon av tall, så er det ett bestemt tall som oppfører seg helt spesielt, nemlig tallet 1. Tallet 1 oppfører seg som et *identitetselement* med hensyn på multiplikasjon. Det betyr at å gange med 1 ikke endrer noe. Hvis vi starter med et hvilket som helst tall, og ganger det med 1, så får vi bare det samme tallet som resultat:

$$a \cdot 1 = a$$

Finnes det noe tilsvarende som dette i verdenen av matriser og matrisemultiplikasjon? Med andre ord: Finnes det en matrise I slik at

$$A \cdot I = A$$

for alle matriser A?

Det er ganske tydelig at det ikke gir mening å stille akkurat det spørsmålet, for det går ikke an å finne én matrise I som er slik at produktet AI er definert for alle matriser A av alle mulige størrelser.

Så vi må begrense oss til å se på matriser av én størrelse om gangen. Dessuten må vi huske på at matriseproduktet ikke er kommutativt, så AI og IA er ikke nødvendigvis det samme. For å si at I er et identitetselement vil vi at både AI og IA skal bli lik A.

Det viser seg at å finne identitetselementer bare er mulig hvis vi dessuten begrenser oss til kvadratiske matriser. Da kan vi reformulere spørsmålet vårt slik:

Finnes det en $n \times n$-matrise I_n som er slik at

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

for alle $n \times n$-matriser A?

Det er ikke vanskelig å se at denne matrisen oppfyller egenskapene vi er ute etter:

$$I_n = \underbrace{\begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}}_{n \times n\text{-matrise}}$$

Matrisen I_n er en $n \times n$-matrise der det er 1-tall langs diagonalen mellom øverste venstre hjørne og nederste høyre hjørne, og bare 0 ellers. Den kalles *identitetsmatrisen* av størrelse n.

Eksempel 4.16. Identitetsmatrisen av størrelse 2 er:

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Vi sjekker enkelt at vi kan gange en hvilken som helst 2×2-matrise med I_2, til venstre eller høyre, uten at noe endres:

$$\begin{aligned} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} &= \begin{bmatrix} a_{11} \cdot 1 + a_{12} \cdot 0 & a_{11} \cdot 0 + a_{12} \cdot 1 \\ a_{21} \cdot 1 + a_{22} \cdot 0 & a_{21} \cdot 0 + a_{22} \cdot 1 \end{bmatrix} \\ &= \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \\ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} &= \begin{bmatrix} 1 \cdot a_{11} + 0 \cdot a_{21} & 1 \cdot a_{12} + 0 \cdot a_{22} \\ 0 \cdot a_{11} + 1 \cdot a_{21} & 0 \cdot a_{12} + 1 \cdot a_{22} \end{bmatrix} \\ &= \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \end{aligned}$$

Vi ser altså at identitetsmatrisen I_2 oppfører seg som et identitetselement med hensyn på multiplikasjon av 2×2-matriser. △

Vi ser også at det å gange en identitetsmatrise med en vektor ikke endrer vektoren. Altså: Hvis $\mathbf{v}$ er en vektor i $\mathbb{R}^n$, så er

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

Mer generelt har vi at om vi ganger en identitetsmatrise med en hvilken som helst matrise som den kan ganges med, så får vi den matrisen som resultat. Altså: Hvis A er en $m \times n$-matrise, så har vi:

$$I_m \cdot A = A = A \cdot I_n$$

Potenser av matriser

Hvis A er en kvadratisk matrise, så kan vi gange A med seg selv. Vi definerer potenser av A på tilsvarende måte som potenser av tall:

$$\begin{aligned} A^2 &= A \cdot A, \\ A^3 &= A \cdot A \cdot A, \end{aligned}$$

og så videre. Generelt definerer vi at A opphøyd i n-te er produktet av A med seg selv n ganger:

$$A^n = \underbrace{A \cdot A \cdot \cdots \cdot A}_{n \text{ ganger}}$$

Et spesielt tilfelle er å opphøye i 0-te. For tall har vi definert at $a^0 = 1$. Men vi vet jo at for matriser spiller identitetsmatrisen den samme rollen som 1 gjør for tall. Derfor definerer vi at en $n \times n$-matrise opphøyd i 0-te blir identitetsmatrisen av størrelse n:

$$A^0 = I_n$$

Inverser

For multiplikasjon av tall har vi identitetselementet 1, og vi har dessuten *inverser*. Gitt et tall a finnes et tall b, inversen til a, som er slik at

$$a \cdot b = 1.$$

Inversen til a er selvfølgelig bare tallet $1/a$. For eksempel: Inversen til tallet 5 er $1/5$, og inversen til $3/4$ er $4/3$.

Kan vi på tilsvarende måte finne inverser til matriser? Igjen begrenser vi oss til å se på kvadratiske matriser, og spørsmålet blir: Gitt en $n \times n$-matrise A, finnes det en matrise B som er slik at likhetene

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

er oppfylt?

Vi tar et eksempel for å se hvordan noen slike matriser kan se ut.

Eksempel 4.17. La A og B være følgende 2×2-matriser:

$$A = \begin{bmatrix} 4 & -2 \\ 3 & -1 \end{bmatrix} \quad \text{og} \quad B = \begin{bmatrix} -1/2 & 1 \\ -3/2 & 2 \end{bmatrix}$$

Da kan vi regne ut at

$$A \cdot B = \begin{bmatrix} 4 & -2 \\ 3 & -1 \end{bmatrix} \begin{bmatrix} -1/2 & 1 \\ -3/2 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2$$

og

$$B \cdot A = \begin{bmatrix} -1/2 & 1 \\ -3/2 & 2 \end{bmatrix} \begin{bmatrix} 4 & -2 \\ 3 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2.$$

Disse matrisene oppfyller altså likhetene

$$A \cdot B = I_2 = B \cdot A.$$

△

Vi definerer begrepet «invers» ved å bruke disse likhetene.

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

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

En matrise er *inverterbar* hvis den har en invers. △

Denne definisjonen gir opphav til noen åpenbare spørsmål:

- Finnes det kvadratiske matriser som ikke har noen invers? (Vi har gitt et eget navn, «inverterbar», til matriser som har invers. Dette hinter ganske sterkt om at det bør finnes matriser som ikke har invers også.)
- Kan en matrise ha mer enn én invers?

Vi kan ganske enkelt besvare det første spørsmålet ved å se på et eksempel.

Eksempel 4.18. Er matrisen

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

inverterbar?

La oss skje hva som skjer hvis vi antar at det finnes en matrise

$$B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

som er en invers til A. Da må vi ha at $AB = I_2$, altså:

$$\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Men produktet på venstre side her er:

$$\begin{bmatrix} b_{11} + b_{21} & b_{12} + b_{22} \\ 0 & 0 \end{bmatrix}$$

Dermed får vi (ved å se på tallene nederst til høyre i matrisene som skal være like) at $0 = 1$, noe som ikke er mulig. Det kan altså ikke finnes noen slik B som er en invers til A, så A er ikke inverterbar. △

La oss nå ta for oss spørsmålet om hvorvidt samme matrise kan ha flere inverser. Svaret er at det kan den ikke, og det kan vi vise ganske enkelt ut fra definisjonen av invers.

Teorem 4.19. *Hvis en matrise er inverterbar, så har den nøyaktig én invers.*

Bevis. La A være en inverterbar $n \times n$-matrise. Anta at B er en invers til A, og at C også er en invers til A; det vil si at

$$A \cdot B = I_n = B \cdot A \qquad \text{og} \qquad A \cdot C = I_n = C \cdot A.$$

Vi vil vise at B og C ikke kan være forskjellige, altså at vi nå må ha $B = C$.

La oss ta utgangspunkt i produktet BAC. Dette kan vi skrive som enten $(BA) \cdot C$ eller $B \cdot (AC)$, og i hvert tilfelle får vi (ved å bruke likhetene over) at uttrykket i parentes blir identitetsmatrisen. Vi setter dette sammen og får:

$$C = I_n \cdot C = (BA) \cdot C = B \cdot (AC) = B \cdot I_n = B$$

Vi har altså kommet frem til at $B = C$, så inversen er entydig. □

Nå som vi vet at en matrise A ikke kan ha mer enn én invers, kan vi slutte å snakke om «en invers til A» i ubestemt form. Isteden sier vi «inversen til A» i bestemt form. Vi definerer dessuten en notasjon for inversen. Hvis A er en inverterbar matrise, så skriver vi A^{-1} for inversen til A.

Eksempel 4.20. I eksempel 4.17 er matrisen B en invers til A; vi har altså at $A^{-1} = B$. Vi får dessuten at A er en invers til B, slik at $B^{-1} = A$. △

Hvorfor er inverser interessante? Én grunn er at de kan fortelle oss noe om løsninger av likninger.

Hvis vi skal løse en likning

$$ax = b$$

der a og b er tall, så vil vi selvfølgelig dele på a for å få x alene på venstresiden. Det er det samme som å gange med inversen til a.

Når vi skal løse en matriselikning

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

så kan vi ikke dele på A. Men hvis A er inverterbar, så kan vi nesten gjøre det likevel, for da kan vi gange med inversen til A. Det gjør at vi kan konkludere med at likningen er løsbar – uansett hva høyresidevektoren $\mathbf{b}$ er – og dessuten at løsningen må være entydig. Vi skriver opp dette som et teorem.

Teorem 4.21. *La A være en $n \times n$-matrise, og $\mathbf{b}$ en vektor i $\mathbb{R}^n$. Hvis A er inverterbar, så har likningen $A\mathbf{x} = \mathbf{b}$ entydig løsning, og løsningen er $\mathbf{x} = A^{-1} \cdot \mathbf{b}$.*

Bevis. Vi sjekker ved innsetting at $\mathbf{x} = A^{-1} \cdot \mathbf{b}$ er en løsning av likningen. Vi har:

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

Det betyr at $\mathbf{x} = A^{-1} \cdot \mathbf{b}$ er en løsning.

Nå må vi sjekke at den er entydig. Fra likningen $A\mathbf{x} = \mathbf{b}$ får vi ved å gange til venstre med A^{-1} på begge sider av likhetstegnet:

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

Men siden $A^{-1}A = I_n$ kan venstresiden her forenkles til $I_n\mathbf{x}$, som bare er $\mathbf{x}$. Dermed har vi:

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

Det betyr at dette er den eneste løsningen av likningen, og beviset er ferdig. □

Eksempel 4.22. La oss se på følgende likning:

$$\begin{bmatrix} 4 & -2 \\ 3 & -1 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 2 \\ 5 \end{bmatrix}$$

Fra eksempel 4.17 vet vi at matrisen på venstresiden av denne likningen er inverterbar, og at inversen er:

$$\begin{bmatrix} -1/2 & 1 \\ -3/2 & 2 \end{bmatrix}$$

Da sier teorem 4.21 at likningen har entydig løsning, og at løsningen er:

$$\mathbf{x} = \begin{bmatrix} -1/2 & 1 \\ -3/2 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 5 \end{bmatrix} = \begin{bmatrix} 4 \\ 7 \end{bmatrix} \qquad \triangle$$

Nå vet vi ganske mye om inverser av matriser – men ikke hvordan vi regner ut en invers. Vi kommer til dette snart, men først skal vi se på en generell teknikk for å løse flere likningssystemer samtidig. Denne teknikken skal vi deretter bruke til å utlede en metode for å regne ut inverser.

Samtidig løsning av flere systemer

Vi husker at et lineært likningssystem kan skrives som en matriselikning $A\mathbf{x} = \mathbf{b}$, og at vi løser det ved å gausseliminere totalmatrisen $\left[\begin{array}{c|c} A & \mathbf{b} \end{array}\right]$.

Anta nå at vi vil løse flere systemer

$$\begin{aligned} A\mathbf{x}_1 &= \mathbf{b}_1 \\ A\mathbf{x}_2 &= \mathbf{b}_2 \\ &\vdots \\ A\mathbf{x}_t &= \mathbf{b}_t \end{aligned}$$

med samme koeffisientmatrise A, men forskjellige høyresidevektorer $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_t$. Det kan vi selvsagt gjøre ved å utføre gausseliminasjon på totalmatrisene til alle systemene:

$$\left[\begin{array}{c|c} A & \mathbf{b}_1 \end{array}\right] \quad \left[\begin{array}{c|c} A & \mathbf{b}_2 \end{array}\right] \quad \cdots \quad \left[\begin{array}{c|c} A & \mathbf{b}_t \end{array}\right]$$

Men da gjør vi egentlig den samme gausseliminasjonen mange ganger. Det eneste som blir forskjellig er hva vi får i siste kolonne. Vi kan spare oss for arbeid ved å slå sammen totalmatrisene til den ene matrisen

$$\left[\begin{array}{c|cccc} A & \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_t \end{array}\right],$$

og gausseliminere den.

Eksempel 4.23. La A være følgende matrise:

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

Vi vil løse disse tre systemene:

$$A\mathbf{x}_1 = \begin{bmatrix} -2 \\ 7 \end{bmatrix} \qquad A\mathbf{x}_2 = \begin{bmatrix} 10 \\ 1 \end{bmatrix} \qquad A\mathbf{x}_3 = \begin{bmatrix} -4 \\ 0 \end{bmatrix}$$

Vi lager en kombinert totalmatrise for alle systemene, og gausseliminerer den:

$$\begin{aligned} \left[\begin{array}{cc|ccc} 2 & -2 & -2 & 10 & -4 \\ 1 & 3 & 7 & 1 & 0 \end{array}\right] &\sim \left[\begin{array}{cc|ccc} 1 & 3 & 7 & 1 & 0 \\ 2 & -2 & -2 & 10 & -4 \end{array}\right] \\ &\sim \left[\begin{array}{cc|ccc} 1 & 3 & 7 & 1 & 0 \\ 0 & -8 & -16 & 8 & -4 \end{array}\right] \\ &\sim \left[\begin{array}{cc|ccc} 1 & 3 & 7 & 1 & 0 \\ 0 & 1 & 2 & -1 & 1/2 \end{array}\right] \\ &\sim \left[\begin{array}{cc|ccc} 1 & 0 & 1 & 4 & -3/2 \\ 0 & 1 & 2 & -1 & 1/2 \end{array}\right] \end{aligned}$$

Den siste matrisen her er på redusert trappeform, og nå kan vi finne løsningene av de tre systemene ved å se på de tre høyresidene i denne matrisen:

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \qquad \mathbf{x}_2 = \begin{bmatrix} 4 \\ -1 \end{bmatrix} \qquad \mathbf{x}_3 = \begin{bmatrix} -3/2 \\ 1/2 \end{bmatrix} \qquad \triangle$$

Beregning av inverser

La oss nå se på hvordan vi kan regne ut inverser. Anta at vi har en $n \times n$-matrise A. Vi vil finne ut om den er inverterbar, og i så fall vil vi finne inversmatrisen A^{-1}.

Vi ser på likningen

$$AX = I_n,$$

der X er en ukjent $n \times n$-matrise. Hvis vi lar

$$\mathbf{x}_1,\ \mathbf{x}_2,\ \ldots,\ \mathbf{x}_n$$

være kolonnene i X, altså

$$X = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_n \end{bmatrix},$$

så kan vi skrive produktet AX slik:

$$AX = \begin{bmatrix} A\mathbf{x}_1 & A\mathbf{x}_2 & \cdots & A\mathbf{x}_n \end{bmatrix}$$

La oss gi navn til kolonnene i identitetsmatrisen I_n også:

$$I_n = \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \ldots & \mathbf{e}_n \end{bmatrix}$$

Det vil si at $\mathbf{e}_i$ er den vektoren i $\mathbb{R}^n$ som har et 1-tall på posisjon i, og bare 0-er ellers.

Nå kan vi, ved å se på hver kolonne, skrive om likningen $AX = I_n$ til disse n likningene:

$$\begin{aligned} A\mathbf{x}_1 &= \mathbf{e}_1 \\ A\mathbf{x}_2 &= \mathbf{e}_2 \\ &\vdots \\ A\mathbf{x}_n &= \mathbf{e}_n \end{aligned}$$

Dermed kan vi bruke teknikken vi beskrev over for å løse flere likningssystemer samtidig. Da må vi gausseliminere matrisen

$$\left[\begin{array}{c|cccc} A & \mathbf{e}_1 & \mathbf{e}_2 & \cdots & \mathbf{e}_n \end{array}\right]$$

for å løse disse systemene.

La oss ta et eksempel for å se hvordan dette blir i praksis.

Eksempel 4.24. Vi vil forsøke å invertere følgende matrise:

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

Vi følger ideene beskrevet over, så vi finner en matrise X slik at $AX = I_3$ ved å løse følgende tre systemer:

$$A\mathbf{x}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad A\mathbf{x}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad A\mathbf{x}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Vi setter opp den kombinerte totalmatrisen for de tre systemene og gausseliminerer:

$$\left[\begin{array}{ccc|ccc} 1 & -1 & -3 & 1 & 0 & 0 \\ 0 & 1 & 3 & 0 & 1 & 0 \\ 2 & -2 & -1 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & -1 & -3 & 1 & 0 & 0 \\ 0 & 1 & 3 & 0 & 1 & 0 \\ 0 & 0 & 5 & -2 & 0 & 1 \end{array}\right]$$
$$\sim \left[\begin{array}{ccc|ccc} 1 & -1 & -3 & 1 & 0 & 0 \\ 0 & 1 & 3 & 0 & 1 & 0 \\ 0 & 0 & 1 & -\frac{2}{5} & 0 & \frac{1}{5} \end{array}\right]$$
$$\sim \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 3 & 0 & 1 & 0 \\ 0 & 0 & 1 & -\frac{2}{5} & 0 & \frac{1}{5} \end{array}\right]$$
$$\sim \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & \frac{6}{5} & 1 & -\frac{3}{5} \\ 0 & 0 & 1 & -\frac{2}{5} & 0 & \frac{1}{5} \end{array}\right]$$

Vi får altså følgende løsninger:

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 6/5 \\ -2/5 \end{bmatrix} \quad \mathbf{x}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \quad \mathbf{x}_3 = \begin{bmatrix} 0 \\ -3/5 \\ 1/5 \end{bmatrix}$$

Nå finner vi matrisen X ved å bruke $\mathbf{x}_1$, $\mathbf{x}_2$ og $\mathbf{x}_3$ som kolonner:

$$X = \begin{bmatrix} 1 & 1 & 0 \\ 6/5 & 1 & -3/5 \\ -2/5 & 0 & 1/5 \end{bmatrix}$$

Vi har funnet X ved å løse likningen $AX = I_3$, så vi vet at dette stemmer. Det kan likevel være lurt å sjekke det ved å gange sammen A og X, for å være sikre på at vi ikke har regnet feil.

Hvis du prøver å gange dem sammen motsatt vei, vil du oppdage at vi også har $XA = I_3$. Dette betyr at X er inversen til A, altså at $A^{-1} = X$. △

I eksempelet løste vi likningen $AX = I_3$, og det viste seg at matrisen X som vi fant også oppfylte likheten $XA = I_3$, slik at vi kunne konkludere med at $A^{-1} = X$.

Dette var ikke en tilfeldighet – det er faktisk alltid nok å løse likningen $AX = I_n$ for å finne inversen til A. Vi skal bevise dette, men vi tar først et lemma (hjelperesultat) som vi skal bruke i beviset vårt.

Lemma 4.25. *La A og B være $n \times n$-matriser. Dersom*

$$\left[\begin{array}{c|c} A & I_n \end{array}\right] \sim \left[\begin{array}{c|c} I_n & B \end{array}\right],$$

så er $AB = I_n$.

Bevis. Vi viste over (i diskusjonen før eksempel 4.24) at vi kan løse likningen $AX = I_n$ ved å gausseliminere matrisen

$$\left[\begin{array}{c|c} A & I_n \end{array}\right].$$

Nå har vi antatt at

$$\left[\begin{array}{c|c} A & I_n \end{array}\right] \sim \left[\begin{array}{c|c} I_n & B \end{array}\right],$$

og siden den andre matrisen her er på redusert trappeform, er det den vi ender opp med når vi gausseliminerer. Det vil si at $X = B$ er løsningen av likningen $AX = I_n$, altså har vi $AB = I_n$. □

Nå er vi klare for å bevise at metoden vår for å finne inverser fungerer.

Teorem 4.26. *La A være en $n \times n$-matrise.*

(a) A er inverterbar hvis og bare hvis $A \sim I_n$.

(b) Hvis A er inverterbar, så kan vi finne inversen ved å gausseliminere matrisen

$$\left[\begin{array}{c|c} A & I_n \end{array}\right]$$

til redusert trappeform og lese av høyre halvdel av den resulterende matrisen. Med andre ord: Resultatet av gausseliminasjonen blir følgende matrise:

$$\left[\begin{array}{c|c} I_n & A^{-1} \end{array}\right]$$

Bevis. Når vi gausseliminerer matrisen

$$\left[\begin{array}{c|c} A & I_n \end{array}\right]$$

til redusert trappeform, må venstre halvdel av den resulterende matrisen enten bli I_n, eller en matrise med minst én nullrad. Men i høyre halvdel kan det ikke bli noen nullrader, for enhver rad vi får etter å ha gjort radoperasjoner på I_n er på formen

$$a_1\mathbf{r}_1 + a_2\mathbf{r}_2 + \cdots + a_n\mathbf{r}_n$$

der $\mathbf{r}_1$, $\mathbf{r}_2$, ..., $\mathbf{r}_n$ er radene i I_n og minst én a_i er ulik 0. Dette betyr at hvis vi får en nullrad i venstre halvdel av trappeformmatrisen, så har likningen $AX = I_n$ ingen løsning, og dermed er A ikke inverterbar. Dermed har vi vist én halvdel av påstanden i del (a): Hvis A er inverterbar, så må vi ha $A \sim I_n$.

La oss nå anta at $A \sim I_n$. Vi vil vise at da er A inverterbar, og at metoden beskrevet i del (b) gir riktig svar. La B være matrisen vi får som svar ved å bruke denne metoden. Det vil si at følgende matriser er begynnelsen og slutten av gausseliminasjonen vi utfører:

$$\left[\begin{array}{c|c} A & I_n \end{array}\right] \sim \left[\begin{array}{c|c} I_n & B \end{array}\right],$$

Nå sier lemma 4.25 at $AB = I_n$.

La oss stokke litt om på kolonnene i matrisen

$$\left[\begin{array}{c|c} A & I_n \end{array}\right]$$

og isteden se på følgende matrise:

$$\left[\begin{array}{c|c} I_n & A \end{array}\right]$$

Hvis vi utfører akkurat de samme radoperasjonene på denne som vi gjorde i gausseliminasjonen av den første matrisen, så får vi akkurat samme resultat, men med tilsvarende omstokking av kolonnene, altså:

$$\left[\begin{array}{c|c} B & I_n \end{array}\right]$$

Dette betyr at disse matrisene er radekvivalente:

$$\left[\begin{array}{c|c} B & I_n \end{array}\right] \sim \left[\begin{array}{c|c} I_n & A \end{array}\right]$$

Ved å bruke lemma 4.25 igjen, på denne siste radekvivalensen, får vi at $BA = I_n$.

Vi har altså vist at vi har

$$AB = I_n = BA,$$

som betyr at B er inversen til A. Det vil si at vi har bevist andre halvdel av del (a) (hvis $A \sim I_n$, så er A inverterbar), og vi har bevist at metoden i del (b) gir riktig svar. $\square$

Det alt dette betyr i praksis er at hvis vi har en matrise A som vi har lyst til å invertere, hvis det er mulig, så setter vi opp matrisen

$$\left[\begin{array}{c|c} A & I_n \end{array}\right]$$

og gausseliminerer. Da er det to muligheter. Enten får vi en nullrad i venstre halvdel, og da er A ikke inverterbar. Eller så kommer vi frem til redusert trappeform uten noen nullrad i venstre halvdel, og da har vi matrisen

$$\left[\begin{array}{c|c} I_n & A^{-1} \end{array}\right]$$

der inversen til A kan leses av i høyre halvdel.

Kapittel 5

Lineær uavhengighet

Definisjonen av lineær uavhengighet

Vi starter med et eksempel:

Eksempel 5.1. La $\mathbf{u}$, $\mathbf{v}$ og $\mathbf{w}$ være følgende tre vektorer i $\mathbb{R}^3$:

$$\mathbf{u} = \begin{bmatrix} 4 \\ 4 \\ 9 \end{bmatrix} \qquad \mathbf{v} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \qquad \mathbf{w} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Disse vektorene oppfyller følgende likhet (det er lett å sjekke):

$$\mathbf{u} = -5 \cdot \mathbf{v} + 9 \cdot \mathbf{w}$$

En slik lineær likhet som knytter sammen vektorer tenker vi på som en «avhengighet» mellom vektorene, og vi sier at vektorene $\mathbf{u}$, $\mathbf{v}$ og $\mathbf{w}$ er lineært avhengige fordi det finnes en slik sammenheng mellom dem.

Vi kan også skrive likheten vår på følgende måte ved å sette alle vektorene på samme side av likhetstegnet:

$$\mathbf{u} + 5 \cdot \mathbf{v} - 9 \cdot \mathbf{w} = \mathbf{0}$$

Det vi har gjort nå er å skrive nullvektoren som en lineærkombinasjon av $\mathbf{u}$, $\mathbf{v}$ og $\mathbf{w}$. △

Hvis vi har en liste $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ med vektorer, så er det klart at nullvektoren $\mathbf{0}$ er en lineærkombinasjon av disse, fordi vi har likheten

$$0 \cdot \mathbf{v}_1 + 0 \cdot \mathbf{v}_2 + \cdots + 0 \cdot \mathbf{v}_n = \mathbf{0}$$

der vi har satt alle vektene til å være 0.

Men spørsmålet vi kan stille er: Kan vi også skrive $\mathbf{0}$ som en lineærkombinasjon av vektorene våre på en annen måte, der ikke alle vektene er 0? I eksempelet over kunne vi det, men i andre tilfeller er det ikke mulig. Dette er det vi vil bruke som definerende egenskap for å si at noen gitte vektorer enten er lineært avhengige eller lineært uavhengige.

Definisjon. La $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ være vektorer i $\mathbb{R}^m$. Disse vektorene er *lineært uavhengige* dersom likningen

$$\mathbf{v}_1 \cdot x_1 + \mathbf{v}_2 \cdot x_2 + \cdots + \mathbf{v}_n \cdot x_n = \mathbf{0}$$

ikke har andre løsninger enn den trivielle løsningen $x_1 = x_2 = \cdots = x_n = 0$.

I motsatt tilfelle kalles de *lineært avhengige*. △

Eksempel 5.2. Se på vektorene

$$\mathbf{v}_1 = \begin{bmatrix} 3 \\ -2 \\ 0 \end{bmatrix} \qquad \text{og} \qquad \mathbf{v}_2 = \begin{bmatrix} 0 \\ 0 \\ 4 \end{bmatrix}$$

i $\mathbb{R}^3$. Likningen

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

medfører at $3x_1 = 0$ og $4x_2 = 0$, altså $x_1 = x_2 = 0$, så den har bare den trivielle løsningen. Dermed er $\mathbf{v}_1$ og $\mathbf{v}_2$ lineært uavhengige. △

Eksempel 5.3. Vektorene $\mathbf{u}$, $\mathbf{v}$ og $\mathbf{w}$ i eksempel 5.1 er lineært avhengige, siden likningen

$$\mathbf{u}x_1 + \mathbf{v}x_2 + \mathbf{w}x_3 = \mathbf{0}$$

har ikketrivielle løsninger (for eksempel løsningen $x_1 = 1$, $x_2 = 5$ og $x_3 = -9$). △

Lineær uavhengighet for to vektorer

Hvis vi ser på bare to vektorer, er det ikke vanskelig å sjekke om de er lineært uavhengige eller ikke.

Eksempel 5.4. Vi undersøker om følgende vektorer er lineært uavhengige:

$$\mathbf{v}_1 = \begin{bmatrix} 8 \\ 2 \\ -12 \end{bmatrix} \qquad \text{og} \qquad \mathbf{v}_2 = \begin{bmatrix} 4 \\ 1 \\ -6 \end{bmatrix}$$

Siden vi har $\mathbf{v}_1 = 2 \cdot \mathbf{v}_2$, får vi at nullvektoren kan skrives som en ikketriviell lineærkombinasjon av $\mathbf{v}_1$ og $\mathbf{2}$ på denne måten:

$$1 \cdot \mathbf{v}_1 - 2 \cdot \mathbf{v}_2 = \mathbf{0}$$

Vektorene $\mathbf{v}_1$ og $\mathbf{v}_2$ er derfor lineært avhengige. △

I dette eksempelet hadde vi at den ene vektoren kunne skrives som et tall ganger den andre, og ut fra det fant vi at vektorene var lineært avhengige. Vi viser at dette generelt er nok til å bestemme om to vektorer er lineært uavhengige eller ikke.

Teorem 5.5. *To vektorer* $\mathbf{u}$ *og* $\mathbf{v}$ *er lineært uavhengige hvis og bare hvis ingen av dem er lik en skalar ganger den andre.*

Bevis. Påstanden i teoremet kan også formuleres slik: Vektorene er lineært *avhengige* hvis og bare hvis en av dem er lik en skalar ganger den andre. Vi viser dette.

Anta først at $\mathbf{u}$ og $\mathbf{v}$ er lineært avhengige. Da finnes to tall a og b slik at

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

og minst én av a og b er ulik 0. Hvis $a \neq 0$, får vi

$$\mathbf{u} = -\frac{b}{a}\mathbf{v}.$$

Hvis $b \neq 0$, får vi

$$\mathbf{v} = -\frac{a}{b}\mathbf{u}.$$

I begge tilfeller har vi at én av vektorene er en skalar ganger den andre.

Nå viser vi den motsatt implikasjonen. Anta derfor at én av vektorene er en skalar ganger den andre. Hvis $\mathbf{u} = c \cdot \mathbf{v}$ for en skalar c, så får vi:

$$\mathbf{u} \cdot 1 + \mathbf{v} \cdot (-c) = \mathbf{0}.$$

Hvis $\mathbf{v} = d \cdot \mathbf{u}$ for en skalar d, så får vi:

$$\mathbf{u} \cdot d + \mathbf{v} \cdot (-1) = \mathbf{0}.$$

I begge tilfeller har vi en ikketriviell løsning av likningen $\mathbf{u} \cdot x_1 + \mathbf{v} \cdot x_2 = \mathbf{0}$, og det betyr at vektorene er lineært avhengige. □

Teoremet sier altså at to vektorer er lineært uavhengige hvis og bare hvis de ikke ligger på en rett linje gjennom origo.

Hvordan sjekke lineær uavhengighet

Det å sjekke om vektorer $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ er lineært uavhengige, er (fra definisjonen) det samme som å sjekke om likningen

$$\mathbf{v}_1 \cdot x_1 + \mathbf{v}_2 \cdot x_2 + \cdots + \mathbf{v}_n \cdot x_n = \mathbf{0}$$

har uendelig mange løsninger, eller bare én. Denne likningen kan vi selvfølgelig løse på vanlig måte, ved å gausseliminere totalmatrisen:

$$\left[\begin{array}{cccc|c} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n & \mathbf{0} \end{array}\right]$$

Eksempel 5.6. Er disse vektorene lineært uavhengige?

$$\mathbf{u} = \begin{bmatrix} 3 \\ 9 \\ 3 \\ 3 \end{bmatrix}, \qquad \mathbf{v} = \begin{bmatrix} 2 \\ 7 \\ 2 \\ 4 \end{bmatrix} \qquad \text{og} \qquad \mathbf{w} = \begin{bmatrix} 8 \\ 31 \\ 12 \\ 22 \end{bmatrix}$$

Vi setter opp totalmatrisen for likningssystemet

$$\mathbf{u} \cdot x + \mathbf{v} \cdot y + \mathbf{w} \cdot z = \mathbf{0},$$

og gausseliminerer den:

$$\left[\begin{array}{ccc|c} 3 & 2 & 8 & 0 \\ 9 & 7 & 31 & 0 \\ 3 & 2 & 12 & 0 \\ 3 & 4 & 22 & 0 \end{array}\right] \sim \left[\begin{array}{ccc|c} 3 & 2 & 8 & 0 \\ 0 & 1 & 7 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Vi kunne fortsatt videre til redusert trappeform, men allerede her er det tydelig at vi får kun én løsning: $x = y = z = 0$. Dette betyr at vektorene $\mathbf{u}$, $\mathbf{v}$ og $\mathbf{w}$ er lineært uavhengige. △

Det essensielle vi trenger å få ut av gausseliminasjonen for å sjekke lineær uavhengighet, er om det blir noen frie variabler eller ikke. Merk også at vi ikke egentlig trenger å ta med høyresidevektoren i gausseliminasjonen. I og med at det er bare 0-er der fra begynnelsen av, kan det aldri bli noe annet enn 0 der, uansett hvilke radoperasjoner vi utfører.

Vi skriver opp et teorem basert på det vi har observert nå.

Teorem 5.7. *La A være en matrise. Følgende påstander er ekvivalente:*

1. *Kolonnene i A er lineært uavhengige.*
2. *Likningen $A\mathbf{x} = \mathbf{0}$ har bare den trivielle løsningen $\mathbf{x} = \mathbf{0}$.*
3. *Vi får ingen frie variabler når vi løser $A\mathbf{x} = \mathbf{0}$.*
4. *Når vi gausseliminerer A, får vi et pivotelement i hver kolonne.*

Bevis. Påstand 2 er bare en omskrivning av definisjonen av lineær uavhengighet til en matriselikning. Påstand 3 forklarer hvordan vi kan se at likningen $A\mathbf{x} = \mathbf{0}$ ikke har mer enn én løsning (husk at vi vet at den alltid har $\mathbf{x} = \mathbf{0}$ som løsning). Påstand 4 er en omformulering av påstand 3, der vi utnytter at vi vet at den siste kolonnen i totalmatrisen uansett bare består av 0-er, så vi trenger ikke ta den med i gausseliminasjonen. □

Teorem 5.7 gir oss en grei metode for å sjekke lineær uavhengighet. Hvis vi har vektorer

$$\mathbf{v}_1, \ \mathbf{v}_2, \ \ldots, \ \mathbf{v}_n$$

i $\mathbb{R}^m$, kan vi finne ut om de er lineært uavhengige på denne måten:

1. Lag en matrise $A = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \ldots & \mathbf{v}_n \end{bmatrix}$ med disse vektorene som kolonner.
2. Gausseliminer A til trappeform.
3. Hvis hver kolonne inneholder et pivotelement, er vektorene lineært uavhengige. Ellers er de lineært avhengige.

Når du gjør dette, bør du imidlertid ikke bare følge denne oppskriften slavisk (og du må for all del ikke pugge slike fremgangsmåter som dette), men huske på at det du egentlig gjør er å sjekke om likningen

$$\mathbf{v}_1 \cdot x_1 + \mathbf{v}_2 \cdot x_2 + \cdots + \mathbf{v}_n \cdot x_n = \mathbf{0}$$

har entydig løsning eller ikke.

Eksempel 5.8. Er disse vektorene lineært uavhengige?

$$\mathbf{v}_1 = \begin{bmatrix} 5 \\ 10 \\ 5 \\ 0 \end{bmatrix}, \qquad \mathbf{v}_2 = \begin{bmatrix} 3 \\ 7 \\ 5 \\ 4 \end{bmatrix} \qquad \text{og} \qquad \mathbf{v}_3 = \begin{bmatrix} 2 \\ 6 \\ 6 \\ 8 \end{bmatrix}$$

Vi gausseliminerer matrisen $\begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{bmatrix}$:

$$\begin{bmatrix} 5 & 3 & 2 \\ 10 & 7 & 6 \\ 5 & 5 & 6 \\ 0 & 4 & 8 \end{bmatrix} \sim \begin{bmatrix} 5 & 3 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Siste kolonne har ikke noe pivotelement. Dermed er $\mathbf{v}_1$, $\mathbf{v}_2$ og $\mathbf{v}_3$ lineært avhengige. △

Eksempel 5.9. Er disse vektorene lineært uavhengige?

$$\mathbf{v}_1 = \begin{bmatrix} 8 \\ 7 \\ 4 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 14 \\ -2 \\ 5 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{v}_4 = \begin{bmatrix} 7 \\ 5 \\ 11 \end{bmatrix}$$

For å sjekke dette kan vi gausseliminere denne matrisen:

$$\begin{bmatrix} 8 & 14 & 3 & 7 \\ 7 & -2 & 1 & 5 \\ 4 & 5 & 0 & 11 \end{bmatrix}$$

Men vi trenger ikke egentlig å utføre gausseliminasjonen. Vi ser med en gang at uansett hva som skjer, så kan vi ikke få mer enn tre pivotelementer (ett i hver rad). Dermed kan det ikke bli pivotelementer i alle de fire kolonnene, så vektorene $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ og $\mathbf{v}_4$ er lineært avhengige. △

Som vi så i dette siste eksempelet, trenger vi ikke alltid å gausseliminere for å finne ut om vektorer er lineært uavhengige eller ikke. Noen ganger kan vi se det på enklere måter.

Vi lister opp noen forskjellige betingelser som kan være nyttige å se etter for å oppdage at vektorer er lineært *avhengige*.

Teorem 5.10. *Gitt n vektorer $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ i $\mathbb{R}^m$. Hvis*

1. *en av vektorene er en lineærkombinasjon av de andre, eller*
2. *en av vektorene er $\mathbf{0}$, eller*
3. *$n > m$,*

så er vektorene lineært avhengige.

Bevis. Anta først at én vektor $\mathbf{v}_k$ er en lineærkombinasjon av de andre:

$$\mathbf{v}_k = \sum_{i \neq k} a_i \mathbf{v}_i$$

Da kan vi sette $a_k = -1$ og få:

$$\sum_{i=1}^{n} a_i \mathbf{v}_i = \mathbf{0}$$

Her har vi skrevet nullvektoren som en ikketriviell lineærkombinasjon av vektorene våre (vi vet ikke hva alle a_i-ene er, men vi vet i hvert fall at én av dem, a_k, ikke er 0). Det betyr at vektorene er lineært avhengige.

Nå går vi videre til å se på den andre antagelsen i teoremet, så vi antar at én av vektorene i listen er nullvektoren. Hvis $\mathbf{v}_k = \mathbf{0}$, så kan vi definere n tall $a_1, a_2, \ldots, a_n$ ved:

$$a_i = \begin{cases} 0 & \text{hvis } i \neq k \\ 1 & \text{hvis } i = k \end{cases}$$

Da får vi at

$$\sum_{i=1}^{n} a_i \mathbf{v}_i = \mathbf{0},$$

og vektorene er lineært avhengige.

Til slutt ser vi på den tredje antagelsen. Akkurat som i eksempel 5.9 får vi her at når vi gausseliminerer matrisen

$$\begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \ldots & \mathbf{v}_n \end{bmatrix},$$

så får vi maksimalt m pivotelementer (ett i hver rad), men vi trenger n pivotelementer (ett i hver kolonne) for at vektorene skal være lineært uavhengige. Når $n > m$ går ikke det an, så da er vektorene lineært avhengige. □

Den første betingelsen i teorem 5.10 er ikke bare tilstrekkelig for å få lineær avhengighet, den er ekvivalent med at vektorene er lineært avhengige. Vi viser dette også.

Teorem 5.11. *Vektorene $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ i $\mathbb{R}^m$ er lineært uavhengige hvis og bare hvis ingen av dem kan skrives som en lineærkombinasjon av de andre.*

Bevis. Påstanden i teoremet er det samme som å si at vektorene er lineært avhengige hvis og bare hvis en av dem kan skrives som en lineærkombinasjon av de andre.

I teorem 5.10 viste vi at dersom en av vektorene er en lineærkombinasjon av de andre, så er de lineært avhengige. Det gjenstår å vise at dersom vektorene er lineært avhengige, så er en av dem en lineærkombinasjon av de andre.

Anta at vektorene er lineært avhengige, altså at vi har

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n = \mathbf{0},$$

der minst én av a_i-ene er ulik 0. Velg en k slik at $a_k \neq 0$. Da har vi:

$$a_k\mathbf{v}_k = \sum_{i \neq k} (-a_i)\mathbf{v}_i$$

Siden $a_k \neq 0$ kan vi dele på a_k og få:

$$\mathbf{v}_k = \sum_{i \neq k} \frac{-a_i}{a_k} \cdot \mathbf{v}_i$$

Dermed er vektoren $\mathbf{v}_k$ en lineærkombinasjon av de andre vektorene i listen. □

Like mange vektorer som dimensjonen

Til slutt ser vi på hva vi kan si om lineær uavhengighet hvis vi ser på n vektorer i $\mathbb{R}^n$. Nå har vi altså like mange vektorer som dimensjonen til rommet vektorene bor i.

Hvis vi har to vektorer i $\mathbb{R}^2$, så kan vi skille mellom følgende tre tilfeller:

1. Begge vektorene er nullvektoren. Da utspenner de bare mengden bestående av nullvektoren, og de er lineært avhengige.
2. Minst én av vektorene er ulik $\mathbf{0}$, og vektorene ligger på samme linje. Da utspenner de denne linjen, og de er lineært avhengige.
3. Vektorene ligger ikke på samme linje, de peker altså i hver sin retning. Da utspenner de hele planet, og de er lineært uavhengige.

Hvis vi har tre vektorer i $\mathbb{R}^3$, så kan vi skille mellom fire tilfeller:

1. Alle er nullvektoren. Da utspenner de bare mengden bestående av nullvektoren, og de er lineært avhengige.
2. Minst én av vektorene er ulik $\mathbf{0}$, og vektorene ligger på samme linje. Da utspenner de denne linjen, og de er lineært avhengige.
3. Vektorene ligger ikke på samme linje, men det finnes et plan i $\mathbb{R}^3$ som inneholder alle tre. Da utspenner de dette planet, og de er lineært avhengige.
4. Vektorene ligger ikke i samme plan. Da utspenner de hele $\mathbb{R}^3$, og de er lineært uavhengige.

Generelt har vi at n vektorer i $\mathbb{R}^n$ enten er lineært avhengige og utspenner en mengde som er mindre enn $\mathbb{R}^n$, eller så er de lineært uavhengige og utspenner hele $\mathbb{R}^n$.

Teorem 5.12. *Hvis vi har n vektorer $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ i $\mathbb{R}^n$, så er de lineært uavhengige hvis og bare hvis de utspenner hele $\mathbb{R}^n$, altså hvis og bare hvis*

$$\mathrm{Sp}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\} = \mathbb{R}^n.$$

Bevis. La

$$A = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \ldots & \mathbf{v}_n \end{bmatrix}$$

være $n \times n$-matrisen med vektorene våre som kolonner. Vi vet at vektorene er lineært uavhengige hvis og bare hvis vi får pivotelementer i alle kolonner når vi gausseliminerer A. Men siden A er kvadratisk, er dette det samme som at vi får pivotelementer i alle rader. Det er igjen ekvivalent med at

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

har løsning for alle vektorer $\mathbf{b}$ i $\mathbb{R}^n$, som er det samme som at kolonnene i A utspenner $\mathbb{R}^n$. $\square$

Kapittel 6

Determinanter

En matrise inneholder mange tall og dermed mye informasjon – så mye at det kan være litt overveldende. Vi kan kondensere ned all informasjonen i en kvadratisk matrise til ett enkelt tall som kalles *determinanten* til matrisen. Dette ene tallet sier oss en hel del om hvordan matrisen oppfører seg.

Vi har to forskjellige notasjoner for determinanter. Hvis

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

er en $n \times n$-matrise, så skriver vi enten

$$\det A \qquad \text{eller} \qquad \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

for determinanten til A. De to notasjonene betyr akkurat det samme. Den første er praktisk å bruke når vi har en variabel som står for matrisen vi vil ta determinanten av; den andre når vi vil liste opp alle tallene i matrisen.

Determinanter for 2×2-matriser

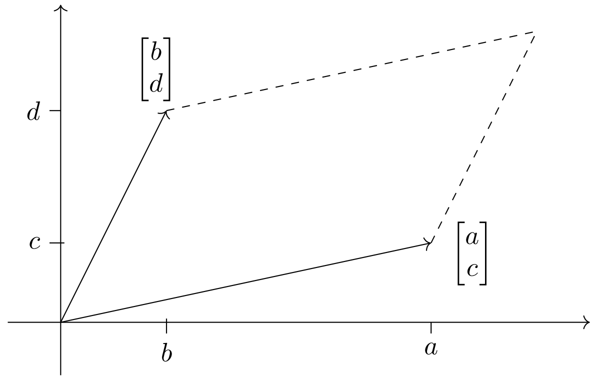
For en 2×2-matrise

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

er determinanten definert ved:

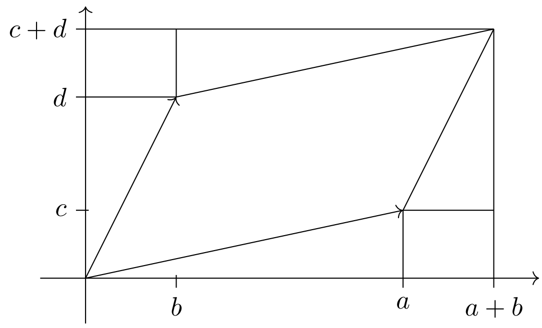
$$\det A = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

Determinanten har en fin geometrisk tolkning. Vi ser på de to kolonnene i A som vektorer i $\mathbb{R}^2$, tegner dem som piler i planet, og lager et parallellogram med disse som to av sidene. Dette paralellogrammet kan vi kalle parallellogrammet utspent av de to vektorene.



Parallellogrammet utspent av kolonnene i A

La oss beregne arealet av dette parallellogrammet. Vi tegner på noen hjelpelinjer:



Parallellogramarealberegningshjelpefigur

Vi kan finne arealet av parallellogrammet ved å starte med arealet av det store rektangelet, som er

$$(a + b)(c + d),$$

og trekke fra arealene av de to små rektanglene og de fire trekantene som omgir parallellogrammet. Vi ser at hvert av de små rektanglene har areal bc, at de to trekantene øverst og nederst til sammen har areal ac, og at trekantene til venstre og høyre til sammen har areal bd. Det betyr at arealet av parallellogrammet er:

$$\begin{aligned} &(a + b)(c + d) - 2bc - ac - bd \\ &\quad = ac + ad + bc + bd - 2bc - ac - bd \\ &\quad = ad - bc \\ &\quad = \det A \end{aligned}$$

Vi kommer altså frem til at arealet av parallellogrammet utspent av kolonnene i A er lik determinanten til A. Denne utregningen var imidlertid litt avhengig av hvordan disse to kolonnevektorene er plassert i forhold til hverandre i planet. Hvis vi hadde byttet plass på kolonnene, så ville vi isteden fått $bc - ad$ som areal. Da ville altså determinanten vært negativ.

Det som holder i alle tilfeller, er at arealet av parallellogrammet utspent av kolonnene i A er lik absoluttverdien til determinanten:

$$|\det A|$$

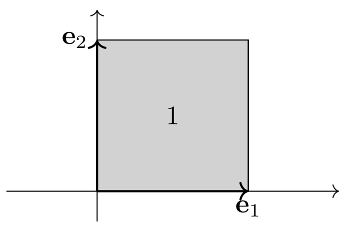
Så determinanten gir oss arealet til et parallellogram. Men hva forteller det oss om matrisen A?

Vi kan tenke på A som en transformasjon av planet der hver vektor $\mathbf{v}$ i $\mathbb{R}^2$ sendes til vektoren $A\mathbf{v}$. Determinanten sier noe om hvordan planet endres under denne transformasjonen.

La

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad \text{og} \qquad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

være de to enhetsvektorene i $\mathbb{R}^2$, og se på kvadratet utspent av disse:



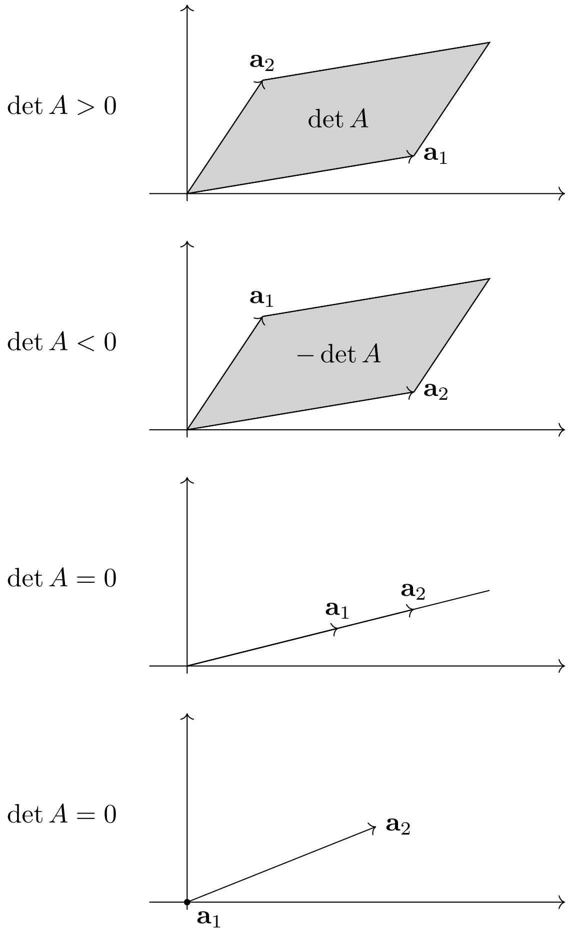
Enhetskvadratet

Nå vil vi se på hva som skjer med dette kvadratet dersom vi sender hvert punkt $\mathbf{v}$ i $\mathbb{R}^2$ til $A\mathbf{v}$, der

$$A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 \end{bmatrix}$$

er en 2×2-matrise med $\mathbf{a}_1$ og $\mathbf{a}_2$ som kolonner. Da sendes vektoren $\mathbf{e}_1$ til $\mathbf{a}_1$, og vektoren $\mathbf{e}_2$ sendes til $\mathbf{a}_2$. Alle vektorene som ligger inni enhetskvadratet i forrige figur sendes til vektorer som ligger inni parallellogrammet utspent av $\mathbf{a}_1$ og $\mathbf{a}_2$.

Vi skisserer noen forskjellige muligheter, for forskjellige valg av matrisen A:



I den første figuren har vi en matrise med positiv determinant. Da gjør transformasjonen $\mathbf{v} \mapsto A\mathbf{v}$ at enhetskvadratet skaleres til et parallellogram med areal $\det A$. Hvis $\det A > 1$ betyr dette at planet «blåses opp»; hvis $0 < \det A < 1$ betyr det at planet «krympes sammen».

I den andre figuren har vi en matrise med negativ determinant. Da er situasjonen helt lik som i den første figuren, bortsett fra at det er $(-\det A)$ som er arealet. Da får vi at med $\det A < -1$ blir planet «blåst opp», og med $-1 < \det A < 0$ blir det «krympet sammen».

I den tredje og den fjerde figuren har vi situasjoner der determinanten er 0. Det vil si at parallellogrammet utspent av kolonnene i A har areal 0. Det blir altså ikke et virkelig parallellogram i disse tilfellene; det har kollapset til et «degenerert» parallellogram som er bare en linje. På samme måte vil transformasjonen $\mathbf{v} \mapsto A\mathbf{v}$ i disse tilfellene kollapse hele planet ned til linjen utspent av $\mathbf{a}_1$ og $\mathbf{a}_2$.

Determinanter for 3×3-matriser

For en 3×3-matrise

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

er determinanten definert ved:

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

Her er det også en geometrisk tolkning. Vi kan tegne opp kolonnene i A som piler i $\mathbb{R}^3$, og lage et parallellepiped med disse pilene som tre av sidene. Dette kaller vi for parallellepipedet utspent av de tre vektorene. Da har vi at volumet av parallellepipedet utspent av kolonnene i A er lik absoluttverdien av determinanten til A.

Den generelle definisjonen

Vi definerer determinanten til en vilkårlig stor kvadratisk matrise etter samme mønster som determinanten til en 3×3-matrise.

Definisjon. La

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

være en $n \times n$-matrise. *Determinanten* til A, som har notasjonen $\det A$, defineres på følgende måte.

1. Hvis $n = 1$, så har vi at $A = \begin{bmatrix} a_{11} \end{bmatrix}$, og da definerer vi at $\det A = a_{11}$.

2. Hvis $n > 1$, innfører vi først noen hjelpevariabler. For hver i og j fra 1 til n setter vi A_{ij} til å være $(n-1) \times (n-1)$-matrisen vi får ved å fjerne rad i og kolonne j fra A, og vi setter

   $$C_{ij} = (-1)^{i+j} \det A_{ij}$$

   til å være determinanten til denne matrisen, med et fortegn som avhenger av i og j. Determinanten til A defineres ved:

   $$\det A = \sum_{j=1}^{n} a_{1j} C_{1j} \qquad \triangle$$

Tallene C_{ij} i definisjonen kalles *kofaktorer* av A.

Det er ikke vanskelig å se at hvis vi setter inn en 2×2-matrise eller en 3×3-matrise i denne generelle definisjonen, så får vi bare de vanlige reglene for determinanter av 2×2- og 3×3-matriser.

For å få litt erfaring med å bruke definisjonen på større matriser regner vi ut determinanten av en 4×4-matrise.

Eksempel 6.1. La A være følgende 4×4-matrise:

$$A = \begin{bmatrix} 3 & 0 & 2 & 4 \\ 1 & 2 & 1 & 0 \\ 2 & 0 & 1 & 0 \\ 7 & 0 & 5 & 3 \end{bmatrix}$$

Vi regner ut determinanten til A. Fra definisjonen får vi:

$$\det A = 3 \begin{vmatrix} 2 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 5 & 3 \end{vmatrix} - 0 \begin{vmatrix} 1 & 1 & 0 \\ 2 & 1 & 0 \\ 7 & 5 & 3 \end{vmatrix} + 2 \begin{vmatrix} 1 & 2 & 0 \\ 2 & 0 & 0 \\ 7 & 0 & 3 \end{vmatrix} - 4 \begin{vmatrix} 1 & 2 & 1 \\ 2 & 0 & 1 \\ 7 & 0 & 5 \end{vmatrix}$$

Vi regner ut hver av 3×3-determinantene som trengs (merk at vi ikke trenger å regne ut den andre, for den skal uansett ganges med 0):

$$\begin{vmatrix} 2 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 5 & 3 \end{vmatrix} = 2 \begin{vmatrix} 1 & 0 \\ 5 & 3 \end{vmatrix} - 1 \begin{vmatrix} 0 & 0 \\ 0 & 3 \end{vmatrix} + 0 \begin{vmatrix} 0 & 1 \\ 0 & 5 \end{vmatrix} = 6$$

$$\begin{vmatrix} 1 & 2 & 0 \\ 2 & 0 & 0 \\ 7 & 0 & 3 \end{vmatrix} = 1 \begin{vmatrix} 0 & 0 \\ 0 & 3 \end{vmatrix} - 2 \begin{vmatrix} 2 & 0 \\ 7 & 3 \end{vmatrix} + 0 \begin{vmatrix} 2 & 0 \\ 7 & 0 \end{vmatrix} = -12$$

$$\begin{vmatrix} 1 & 2 & 1 \\ 2 & 0 & 1 \\ 7 & 0 & 5 \end{vmatrix} = 1 \begin{vmatrix} 0 & 1 \\ 0 & 5 \end{vmatrix} - 2 \begin{vmatrix} 2 & 1 \\ 7 & 5 \end{vmatrix} + 1 \begin{vmatrix} 2 & 0 \\ 7 & 0 \end{vmatrix} = -6$$

Ved å sette inn disse i uttrykket for $\det A$ får vi:

$$\det A = 3 \cdot 6 - 0 + 2 \cdot (-12) - 4 \cdot (-6) = 18$$

Vi har altså regnet ut at $\det A = 18$. $\triangle$

Ved hjelp av definisjonen kan vi regne ut determinanten til en hvilken som helst kvadratisk matrise, men det kan bli veldig mye jobb. I eksempelet så vi at determinanten til en 4×4-matrise er definert ut fra determinantene til fire 3×3-matriser, og hver av disse er igjen definert ut fra determinantene til tre 2×2-matriser. Hvis vi går til større matriser, blir arbeidsmengden fort enormt stor.

I løpet av dette kapitlet skal vi se på noen lure teknikker for å regne ut determinanter på mindre arbeidskrevende måter.

Kofaktorekspansjon

I definisjonen av determinanten går vi gjennom første rad i matrisen, og ser på tallene

$$a_{11}, \ a_{12}, \ \ldots, \ a_{1n}.$$

Hvert tall a_{1j} ganges med den tilhørende kofaktoren C_{1j}, og til slutt summerer vi alle disse produktene.

Det er imidlertid ikke nødvendig å gå langs første rad når vi gjør dette. Det fungerer like bra å gå langs en annen rad og følge det samme systemet, og resultatet blir det samme. Det går dessuten an å gå langs en hvilken som helst kolonne med samme system. Vi oppsummerer dette i følgende teorem.

Teorem 6.2. *La A være en $n \times n$-matrise, der $n > 1$, og la A_{ij} og C_{ij} være som i definisjonen av determinant. Da har vi*

$$\det A = \sum_{j=1}^{n} a_{kj} C_{kj} = \sum_{i=1}^{n} a_{il} C_{il}$$

for alle k og l slik at $1 \leq k \leq n$ og $1 \leq l \leq n$.

Å regne ut determinanten ved å beregne en sum av tall fra matrisen ganget med kofaktorer slik som i teoremet kalles *kofaktorekspansjon*. Vi sier at vi gjør kofaktorekspansjon langs rad k eller langs kolonne l.

La oss nå regne ut den samme determinanten som i eksempel 6.1, men på en lurere måte.

Eksempel 6.3. Vi lar igjen A være denne 4×4-matrisen:

$$A = \begin{bmatrix} 3 & 0 & 2 & 4 \\ 1 & 2 & 1 & 0 \\ 2 & 0 & 1 & 0 \\ 7 & 0 & 5 & 3 \end{bmatrix}$$

Vi kan observere at den andre kolonnen inneholder nesten bare nuller, så det er lurt å gjøre kofaktorekspansjon langs den. Da får vi:

$$\begin{aligned} \det A &= -0 \begin{vmatrix} 1 & 1 & 0 \\ 2 & 1 & 0 \\ 7 & 5 & 3 \end{vmatrix} + 2 \begin{vmatrix} 3 & 2 & 4 \\ 2 & 1 & 0 \\ 7 & 5 & 3 \end{vmatrix} \\ &\quad - 0 \begin{vmatrix} 3 & 2 & 4 \\ 1 & 1 & 0 \\ 7 & 5 & 3 \end{vmatrix} + 0 \begin{vmatrix} 3 & 2 & 4 \\ 1 & 1 & 0 \\ 2 & 1 & 0 \end{vmatrix} \\ &= 2 \cdot \left(4 \cdot \begin{vmatrix} 2 & 1 \\ 7 & 5 \end{vmatrix} + 3 \cdot \begin{vmatrix} 3 & 2 \\ 2 & 1 \end{vmatrix} \right) \\ &= 2 \cdot (4 \cdot 3 + 3 \cdot (-1)) = 18 \end{aligned}$$

Vi fikk samme resultat som i eksempel 6.1, men med mindre arbeid, siden vi bare trengte å regne ut én 3×3-determinant. △

Vi må passe på at vi får fortegnene i kofaktorene riktig. Når vi gjør kofaktorekspansjon langs første rad (slik som i definisjonen) eller langs første kolonne, starter vi alltid med positivt fortegn i det første leddet. Men når vi ekspanderer langs en annen rad eller kolonne, kan det hende vi må starte med negativt fortegn. Det kan være nyttig å bruke følgende diagram som en huskeregel for hvilket fortegn vi skal ha i de forskjellige kofaktorene:

$$\begin{bmatrix} + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

Determinanter og radoperasjoner

Hvis vi utfører en radoperasjon på en matrise, så får vi en ny matrise. Den matrisen har ikke nødvendigvis samme determinant som den opprinnelige, men det viser seg at determinanten endrer seg på ganske kontrollerte måter når vi utfører radoperasjoner. Dette kan vi utnytte for å spare oss for en del arbeid når vi skal regne ut determinanter, spesielt hvis vi har store matriser.

Teorem 6.4. *La A være en $n \times n$-matrise, og la B være en matrise vi får ved å utføre en radoperasjon på A. Da har vi følgende sammenheng mellom determinantene til A og B, basert på hvilken type radoperasjon vi utførte:*

Radoperasjon	*Resultat*
Gange en rad med et tall k	$\det B = k \cdot \det A$
Legge til et multiplum av én rad i en annen	$\det B = \det A$
Bytte om to rader	$\det B = -\det A$

La oss bruke dette teoremet til å beregne en determinant.

Eksempel 6.5. Vi regner ut $\det A$, der

$$A = \begin{bmatrix} 3 & 3 & 12 \\ 2 & 2 & 13 \\ 4 & 2 & 19 \end{bmatrix}$$

Vi får:

$$\begin{aligned} \det A &= \begin{vmatrix} 3 & 3 & 12 \\ 2 & 2 & 13 \\ 4 & 2 & 19 \end{vmatrix} = 3 \cdot \begin{vmatrix} 1 & 1 & 4 \\ 2 & 2 & 13 \\ 4 & 2 & 19 \end{vmatrix} \\ &= 3 \cdot \begin{vmatrix} 1 & 1 & 4 \\ 0 & 0 & 5 \\ 0 & -2 & 3 \end{vmatrix} = -3 \cdot \begin{vmatrix} 1 & 1 & 4 \\ 0 & -2 & 3 \\ 0 & 0 & 5 \end{vmatrix} \\ &= -3 \cdot 1 \cdot \begin{vmatrix} -2 & 3 \\ 0 & 5 \end{vmatrix} \\ &= -3 \cdot 1 \cdot (-2) \cdot 5 = 30 \end{aligned}$$

Her startet vi med å gjøre radoperasjoner på matrisen, samtidig som vi holdt styr på hvordan determinanten endret seg.

Først ganget vi øverste rad med $1/3$. Det medførte at determinanten til den nye matrisen ble $1/3$ ganger determinanten til A, så vi måtte gange den med 3 for at tallene skal bli like. En hendig måte å huske hvordan dette fungerer er å tenke på det som å sette et tall utenfor parentes. På samme måte som vi kan trekke ut et 3-tall fra en parentes og få

$$(3 + 3 + 12) = 3 \cdot (1 + 1 + 12),$$

kan vi trekke ut et 3-tall fra en rad i en determinant.

Etterpå trakk vi fra multipler av første rad i de to andre radene, men det medførte ingen endring av determinanten.

Så byttet vi de to nederste radene, og det gjorde at determinanten skifter fortegn.

Til slutt gjorde vi kofaktorekspansjon langs den første kolonnen. Siden vi ved å utføre radoperasjoner hadde sørget for å få bare nuller under diagonalen, ble kofaktorekspansjonen enkel og grei. △

Triangulære matriser

Vi sier at en $n \times n$-matrise er *øvre triangulær* hvis alle tall under diagonalen er 0, altså hvis den er på følgende form:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{bmatrix}$$

Tilsvarende sier vi at en $n \times n$-matrise er *nedre triangulær* hvis alle tall over diagonalen er 0, altså hvis den er på følgende form:

$$\begin{bmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ a_{21} & a_{22} & 0 & \cdots & 0 \\ a_{31} & a_{32} & a_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{bmatrix}$$

Eksempel 6.6. I eksempel 6.5 brukte vi radoperasjoner til å skrive om matrisen vår til følgende:

$$\begin{bmatrix} 1 & 1 & 4 \\ 0 & -2 & 3 \\ 0 & 0 & 5 \end{bmatrix}$$

Denne matrisen er øvre triangulær. △

Hvis vi skal finne determinanten til en øvre triangulær matrise er det praktisk å kofaktorekspandere langs første kolonne. Da får vi bare ett ledd i ekspansjonen, nemlig tallet øverst til venstre i matrisen ganget med determinanten til matrisen der første rad og kolonne er fjernet. Denne matrisen er igjen øvre triangulær. Så fortsetter vi med kofaktorekspansjon langs første kolonne i hvert steg nedover. Det vi ender opp med til slutt er å bare gange sammen alle tallene på diagonalen.

Vi oppsummerer dette i et teorem.

Teorem 6.7. *La A være en (øvre eller nedre) triangulær $n \times n$-matrise. Da er determinanten til A lik produktet av tallene på diagonalen i A:*

$$\det A = a_{11} \cdot a_{22} \cdot a_{33} \cdots a_{nn}$$

Eksempel 6.8. Ved å bruke teoremet kan vi skrive opp determinanten til en triangulær matrise direkte, uten å måtte beregne andre determinanter først:

$$\begin{vmatrix} 1 & 1 & 4 \\ 0 & -2 & 3 \\ 0 & 0 & 5 \end{vmatrix} = 1 \cdot (-2) \cdot 5$$

△

Hvis vi skal beregne determinanten til en stor og stygg matrise, er det en god strategi å bruke radoperasjoner for å skrive om matrisen til øvre triangulær form (og holde orden på hvordan determinanten endrer seg ved hjelp av teorem 6.4), og så regne ut determinanten til den triangulære matrisen ved hjelp av teorem 6.7.

Flere regneregler

Vi tar med et par regneregler til for determinanter.

Teorem 6.9. *Determinanten til et produkt av to matriser er produktet av determinantene. Altså: Hvis A og B er to $n \times n$-matriser, så er*

$$\det(AB) = (\det A)(\det B).$$

Teorem 6.10. *Determinanten endrer seg ikke når vi transponerer matrisen. Altså: Hvis A er en $n \times n$-matrise, så er*

$$\det A = \det A^\top.$$

Karakterisering av inverterbarhet

Vi kan bruke determinanten til å sjekke om en matrise er inverterbar eller ikke. Dette kan dessuten knyttes sammen med hvorvidt kolonnene i matrisen vår er lineært uavhengige, og hva det utspenner.

Følgende teorem gir en presis sammenheng mellom inverterbarhet, determinant, lineær uavhengighet og mengden utspent av kolonnene.

Teorem 6.11. *La A være en $n \times n$-matrise. Følgende påstander er ekvivalente:*

1. *A er inverterbar.*
2. $\det A \neq 0$.
3. *Kolonnene i A er lineært uavhengige.*
4. *Kolonnene i A utspenner $\mathbb{R}^n$.*

Kapittel 7

Egenverdier og egenvektorer

Det er ofte hensiktsmessig å tenke på en matrise ikke bare som en tabell med tall, men som en transformasjon av vektorer. Hvis A er en $m \times n$-matrise, så gir A en transformasjon

$$\mathbf{v} \mapsto A\mathbf{v}$$

fra $\mathbb{R}^n$ til $\mathbb{R}^m$. Vi kan anvende A på en vektor $\mathbf{v}$ i $\mathbb{R}^n$, og den vektoren transformeres til en vektor $A\mathbf{v}$ i $\mathbb{R}^m$.

Hvis A er en $n \times n$-matrise, altså en kvadratisk matrise, så sender den vektorer i $\mathbb{R}^n$ til andre vektorer som også er i $\mathbb{R}^n$. Generelt kan vektoren $A\mathbf{v}$ være veldig forskjellig fra $\mathbf{v}$, men noen ganger er den ikke det. Hvis virkningen av A på $\mathbf{v}$ er det samme som å bare gange opp $\mathbf{v}$ med et tall, så kalles $\mathbf{v}$ en *egenvektor* for A, og tallet kalles en *egenverdi*.

I dette kapitlet skal vi se hvordan vi kan finne alle egenverdiene og egenvektorene til en matrise, og vi skal se noen interessante egenskaper de har.

Definisjon av egenverdier og egenvektorer

Vi starter med et enkelt eksempel, slik at vi har et konkret tilfelle vi kan ha i tankene når vi kommer til definisjonen.

Eksempel 7.1. La A være følgende 2×2-matrise:

$$A = \begin{bmatrix} 1 & 3 \\ 4 & -3 \end{bmatrix}$$

Vi vil se på hva som skjer med punkter i planet når vi ganger dem med A, altså når vi sender en vektor $\mathbf{x}$ til vektoren $A\mathbf{x}$.

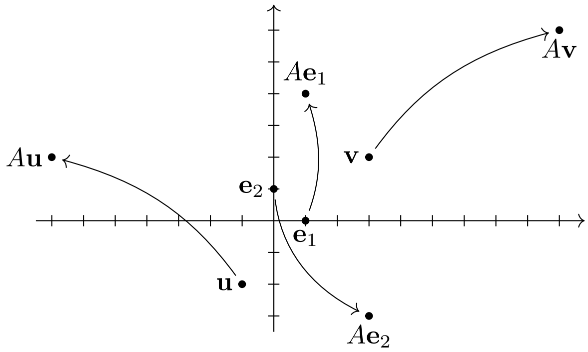
Vi velger følgende fire vektorer og ser hva A gjør med dem:

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \mathbf{u} = \begin{bmatrix} -1 \\ -2 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

Vi får:

$$A\mathbf{e}_1 = \begin{bmatrix} 1 \\ 4 \end{bmatrix}, \; A\mathbf{e}_2 = \begin{bmatrix} 3 \\ -3 \end{bmatrix}, \; A\mathbf{u} = \begin{bmatrix} -7 \\ 2 \end{bmatrix}, \; A\mathbf{v} = \begin{bmatrix} 9 \\ 6 \end{bmatrix}$$

La oss tegne opp de fire vektorene, samt vektorene A sender dem til, som punkter i planet.



Matrisen A kaster vektorene rundt i planet

Vi ser at matrisen sender de fire eksempelvektorene våre i forskjellige retninger. Men akkurat vektoren

$$\mathbf{v} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

er interessant. Det som skjer med den er at den blir sendt til

$$\begin{bmatrix} 9 \\ 6 \end{bmatrix},$$

men det er jo det samme som $3 \cdot \mathbf{v}$. Virkningen av matrisen A på akkurat denne vektoren er altså bare å skalere den opp med 3:

$$A\mathbf{v} = 3\mathbf{v} \qquad \triangle$$

Teorien om egenverdier og egenvektorer handler om å identifisere slike situasjoner som den vi så i eksempelet, der virkningen av en matrise på en vektor blir det samme som å bare gange opp vektoren med et tall.

Definisjon. La A være en $n \times n$-matrise, λ et tall og $\mathbf{v}$ en vektor i $\mathbb{R}^n$ som ikke er nullvektoren. Hvis

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

så sier vi at tallet λ er en *egenverdi* for matrisen A, og at vektoren $\mathbf{v}$ er en *egenvektor* for A som hører til egenverdien λ. $\triangle$

Et par merknader – én matematisk og én språklig – er på sin plass etter denne definisjonen.

Merk. Hvorfor sier vi at $\mathbf{v}$ ikke skal være nullvektoren? Jo, for hvis vi setter inn nullvektoren for $\mathbf{v}$, så får vi likningen

$$A \cdot \mathbf{0} = \lambda \cdot \mathbf{0},$$

som er oppfylt for alle matriser A og alle tall λ. Hvis vi hadde tillatt nullvektoren som en egenvektor, så ville vi altså fått at alle tall er egenverdier for alle matriser. Da blir egenverdibegrepet ganske meningsløst. △

Merk. Det er vanlig å bruke den greske bokstaven λ som variabelnavn for egenverdier. Vi kunne i og for seg brukt en hvilken som helst annen bokstav, men siden det er λ folk vanligvis bruker, så gjør vi det. Navnet på bokstaven uttales «lambda», og den tilsvarer bokstaven l i det latinske alfabetet. Hvis vi for eksempel skriver navnet på filosofen Platon på hans eget språk, så er λ den andre bokstaven: Πλατων. △

Eksempel 7.2. Vi ser igjen på matrisen A fra eksempel 7.1. Vi så at vektoren

$$\mathbf{v} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

oppfylte likheten

$$A\mathbf{v} = 3\mathbf{v}.$$

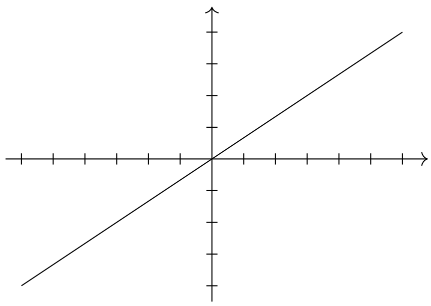
Det betyr at 3 er en egenverdi for matrisen A, og at $\mathbf{v}$ er en egenvektor som hører til egenverdien 3.

Finnes det flere egenvektorer? Hvis vi ser på en vektor som er parallell med $\mathbf{v}$, altså som er på formen $\mathbf{w} = c \cdot \mathbf{v}$ der c er et tall, så får vi:

$$A\mathbf{w} = A \cdot (c\mathbf{v}) = c \cdot (A\mathbf{v}) = c \cdot (3\mathbf{v}) = 3 \cdot \mathbf{w}$$

Enhver slik vektor er altså en egenvektor som hører til egenverdien 3, forutsatt at den ikke er nullvektoren.

Vi har altså funnet ut at A i hvert fall har én egenverdi, nemlig 3, og uendelig mange egenvektorer som hører til denne egenverdien, nemlig alle vektorene på denne linjen (unntatt nullvektoren):



Det vi foreløpig ikke vet, er om det kan finnes enda flere egenvektorer, og om A har flere egenverdier enn 3. Vi skal vende tilbake til dette eksempelet om en stund og finne ut av dette, etter at vi har kommet frem til en generell metode for å finne alle egenverdiene og egenvektorene til en matrise. △

I eksempelet så vi at vi ut fra én egenvektor kunne finne uendelig mange egenvektorer som hørte til den samme egenverdien. Vi formulerer dette generelt som et teorem. (Beviset får du enkelt ved å generalisere det vi gjorde i eksempelet.)

Teorem 7.3. *Anta at λ er en egenverdi for en $n \times n$-matrise A, og at $\mathbf{v}$ er en tilhørende egenvektor. Da er alle multipler $c\mathbf{v}$ av vektoren $\mathbf{v}$, der c er et tall som ikke er 0, også egenvektorer som hører til egenverdien λ. Med andre ord er alle vektorer i mengden* $\mathrm{Sp}\{\mathbf{v}\}$, *unntatt nullvektoren, egenvektorer som hører til egenverdien λ.*

En annen ting som vi enkelt kan se ut fra definisjonen av egenverdier og egenvektorer, er hva som skal til for at en matrise skal ha 0 som egenverdi. Hvis vi setter inn 0 for λ i likheten

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

så får vi $A\mathbf{v} = \mathbf{0}$. Det vil si at en matrise A har 0 som egenverdi hvis og bare hvis likningen

$$A\mathbf{x} = \mathbf{0}$$

har ikketrivielle løsninger, og dette er igjen sant hvis og bare hvis A ikke er inverterbar. Vi formulerer dette også som et teorem.

Teorem 7.4. *En $n \times n$-matrise A har 0 som egenverdi hvis og bare hvis den ikke er inverterbar.*

Noen geometriske eksempler

Før vi utleder den generelle fremgangsmåten for å regne ut egenverdier og egenvektorer, tar vi noen enkle eksempler der vi kan se geometrisk hva egenverdiene og egenvektorene må være.

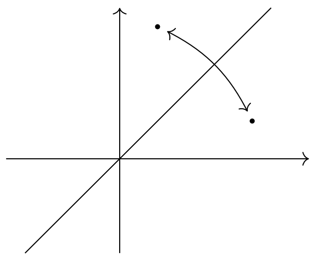
Eksempel 7.5. La A være matrisen

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

Da er

$$A \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} v_2 \\ v_1 \end{bmatrix}$$

for enhver vektor (v_1, v_2) i $\mathbb{R}^2$. Virkningen av A er altså en refleksjon om diagonallinjen som går gjennom origo og punktet $(1, 1)$.



Matrisen A reflekterer vektorene om diagonalen

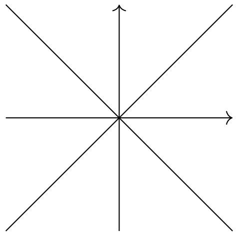
Vi ser lett at alle punkter på denne diagonalen sendes til seg selv, så disse oppfyller likheten

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

og er altså egenvektorer tilhørende egenverdien 1. Når vi tenker oss litt mer om, finner vi dessuten ut at alle punkter på den omvendte diagonalen reflekteres gjennom origo slik at de oppfyller likheten

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

og er egenvektorer tilhørende egenverdien -1. Da har vi funnet at alle vektorene på disse to diagonallinjene (unntatt nullvektoren, selvsagt) er egenvektorer:



Egenvektorene er på diagonalene

Men for alle andre vektorer $\mathbf{v}$ i planet gjør refleksjonen av $A\mathbf{v}$ peker i en annen retning enn $\mathbf{v}$, så det finnes ikke flere egenvektorer. △

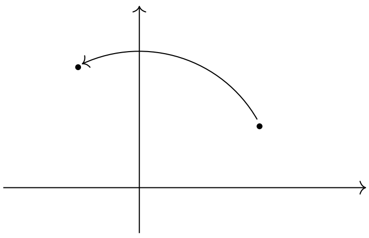
Eksempel 7.6. La A være matrisen

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

Da er

$$A \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} -v_2 \\ v_1 \end{bmatrix}$$

for enhver vektor (v_1, v_2) i $\mathbb{R}^2$. Virkningen av A er altså å rotere planet med 90°.



Matrisen A roterer planet

Dermed kan ikke A ha noen egenvektorer, siden enhver vektor (utenom nullvektoren) sendes til en som peker i en annen retning. △

Hvordan finne egenverdier/-vektorer

Gitt en $n \times n$-matrise A, hvordan kan vi finne alle egenverdiene og egenvektorene dens?

Ut fra definisjonen er vi på jakt etter tall λ og vektorer $\mathbf{v} \neq \mathbf{0}$ som oppfyller likheten

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

Vi har altså en likning med både λ og $\mathbf{v}$ som ukjente, og den ser ved første øyekast ganske uhåndterlig ut. Men vi kan trikse litt med den. Vi kan først flytte alt over til venstre side:

$$A\mathbf{v} - \lambda\mathbf{v} = \mathbf{0}$$

Nå fremstår det som veldig fristende å sette $\mathbf{v}$-en utenfor parentes, altså å skrive $(A - \lambda)\mathbf{v}$. Men det går ikke an, for uttrykket $A - \lambda$, altså en matrise minus et tall, gir ikke mening.

Nå kan vi bruke et lurt triks: Vi ganger inn identitetsmatrisen I_n. Vi vet at $I_n\mathbf{v}$ bare blir $\mathbf{v}$ uansett hva vektoren $\mathbf{v}$ er, så vi kan skrive om likningen til:

$$A\mathbf{v} - \lambda I_n\mathbf{v} = \mathbf{0}$$

Det vi har oppnådd nå er at vi har en $n \times n$-matrise ganget med $\mathbf{v}$ i hvert ledd, og da kan vi sette $\mathbf{v}$ utenfor parentes:

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

Nå ser vi at λ er en egenverdi for A hvis og bare hvis likningen

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

har en ikketriviell løsning. Men dette er bare et vanlig lineært likningssystem med

$$A - \lambda I_n$$

som koeffisientmatrise, og vi vet fra før at et slikt system har ikketrivielle løsninger hvis og bare hvis

$$\det(A - \lambda I_n) = 0.$$

Her har vi endt opp med en likning med bare λ som ukjent. Vi kan altså løse denne for å finne egenverdiene, uten at vi samtidig må tenke på hva de tilhørende egenvektorene skal være.

Vi oppsummerer det vi har funnet ut i et teorem.

Teorem 7.7. *La A være en $n \times n$-matrise.*

(a) Egenverdiene til A er alle løsninger λ av likningen

$$\det(A - \lambda I_n) = 0.$$

(b) Hvis λ er en egenverdi for A, så er de tilhørende egenvektorene gitt ved alle ikketrivielle løsninger av likningen

$$(A - \lambda I_n) \cdot \mathbf{x} = \mathbf{0}.$$

Uttrykket

$$\det(A - \lambda I_n),$$

som står på venstresiden av likningen vi løser for å finne egenverdiene, blir et n-tegradspolynom i λ. Vi kaller det for det *karakteristiske polynomet* til A.

Eksempel 7.8. Nå kan vi bruke teorem 7.7 til å finne alle egenverdiene og egenvektorene til matrisen

$$A = \begin{bmatrix} 1 & 3 \\ 4 & -3 \end{bmatrix}$$

fra eksempel 7.1.

Vi finner egenverdiene ved å løse likningen

$$\det(A - \lambda I_2) = 0,$$

der venstresiden er det karakteristiske polynomet til A. La oss først se hvordan matrisen $A - \lambda I_2$ ser ut:

$$A - \lambda I_2 = \begin{bmatrix} 1 & 3 \\ 4 & -3 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} = \begin{bmatrix} 1-\lambda & 3 \\ 4 & -3-\lambda \end{bmatrix}$$

Det karakteristiske polynomet blir:

$$\begin{aligned} \det(A - \lambda I_2) &= \begin{vmatrix} 1-\lambda & 3 \\ 4 & -3-\lambda \end{vmatrix} \\ &= (1-\lambda)(-3-\lambda) - 3 \cdot 4 \\ &= \lambda^2 + 2\lambda - 15 \end{aligned}$$

Det betyr at vi kan løse andregradslikningen

$$\lambda^2 + 2\lambda - 15 = 0$$

for å finne egenverdiene. Vi løser den på vanlig måte og får:

$$\lambda = \frac{-2 \pm \sqrt{2^2 - 4 \cdot (-15)}}{2} = -1 \pm 4$$

Vi får altså to egenverdier: 3 og -5.

Vi finner alle egenvektorer som hører til egenverdien 3 ved å løse likningen $(A - 3I_2)\mathbf{x} = \mathbf{0}$. Vi kan løse denne likningen ved å gausseliminere matrisen $(A - 3I_2)$:

$$\begin{bmatrix} -2 & 3 \\ 4 & -6 \end{bmatrix} \sim \begin{bmatrix} -2 & 3 \\ 0 & 0 \end{bmatrix}$$

Vi får én fri variabel, og løsningene blir

$$\mathbf{x} = \begin{bmatrix} 3 \\ 2 \end{bmatrix} \cdot t$$

for alle tall t. Egenvektorene som hører til egenverdien 3 er altså alle vektorer i

$$\mathrm{Sp}\left\{ \begin{bmatrix} 3 \\ 2 \end{bmatrix} \right\},$$

unntatt nullvektoren.

Vi finner alle egenvektorer som hører til egenverdien -5 ved å løse likningen $(A + 5I_2)\mathbf{x} = \mathbf{0}$. Vi kan løse denne likningen ved å gausseliminere matrisen $(A + 5I_2)$:

$$\begin{bmatrix} 6 & 3 \\ 4 & 2 \end{bmatrix} \sim \begin{bmatrix} 2 & 1 \\ 0 & 0 \end{bmatrix}$$

Vi får én fri variabel, og løsningene blir

$$\mathbf{x} = \begin{bmatrix} 1 \\ -2 \end{bmatrix} \cdot t$$

for alle tall t. Egenvektorene som hører til egenverdien -5 er altså alle vektorer i

$$\mathrm{Sp}\left\{ \begin{bmatrix} 1 \\ -2 \end{bmatrix} \right\},$$

unntatt nullvektoren. △

Egenrom

Vi har sett at egenverdiene til en matrise er noen enkeltverdier, mens egenvektorene er uendelig mange (dersom matrisen har egenverdier og egenvektorer). I eksempel 7.8 beskrev vi egenvektorene tilhørende en gitt egenvektor ved å si «alle vektorer i (…) unntatt nullvektoren». Vi innfører nå et nytt begrep som gjør det litt enklere å beskrive alle egenvektorene til en egenverdi.

Definisjon. La A være en $n \times n$-matrise, og anta at λ er en egenverdi for A. Da er *egenrommet* til λ mengden av alle egenvektorer som hører til λ, samt nullvektoren; altså mengden

$$\{\, \mathbf{v} \in \mathbb{R}^n \mid A\mathbf{v} = \lambda\mathbf{v} \,\}.$$

△

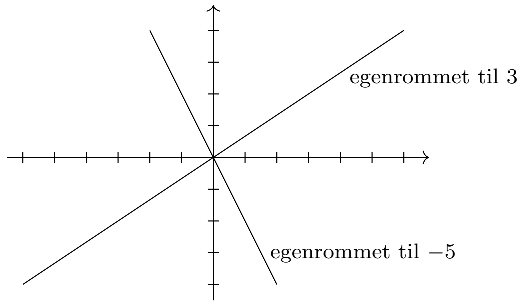
Eksempel 7.9. I eksempel 7.8 kunne vi sagt at egenrommet til egenvektoren 3 er

$$\mathrm{Sp}\left\{ \begin{bmatrix} 3 \\ 2 \end{bmatrix} \right\},$$

og at egenrommet til egenverdien -5 er

$$\mathrm{Sp}\left\{ \begin{bmatrix} 1 \\ -2 \end{bmatrix} \right\}.$$

Hvert av disse to egenrommene er en linje i planet:



△

La oss nå ta et litt større eksempel.

Eksempel 7.10. Vi finner egenverdiene til matrisen

$$A = \begin{bmatrix} -8 & 0 & 6 \\ 12 & 4 & -6 \\ -20 & 0 & 14 \end{bmatrix},$$

og de tilhørende egenrommene.

Det karakteristiske polynomet til A er:

$$\begin{aligned} \det(A - \lambda I_3) &= \begin{vmatrix} -8-\lambda & 0 & 6 \\ 12 & 4-\lambda & -6 \\ -20 & 0 & 14-\lambda \end{vmatrix} \\ &= (4-\lambda) \cdot \begin{vmatrix} -8-\lambda & 6 \\ -20 & 14-\lambda \end{vmatrix} \\ &= (4-\lambda)\big((-8-\lambda)(14-\lambda) + 6 \cdot 20\big) \\ &= (4-\lambda)(\lambda^2 - 6\lambda + 8) \end{aligned}$$

Vi finner altså egenverdiene til A ved å løse tredjegradslikningen

$$(4-\lambda)(\lambda^2 - 6\lambda + 8) = 0.$$

Denne likningen er ekvivalent med at

$$4 - \lambda = 0 \qquad \text{eller} \qquad \lambda^2 - 6\lambda + 8 = 0.$$

Andregradslikningen $\lambda^2 - 6\lambda + 8 = 0$ har løsninger

$$\lambda = \frac{6 \pm \sqrt{6^2 - 4 \cdot 8}}{2} = 3 \pm 1,$$

så vi får to egenverdier: 2 og 4.

Vi finner egenrommene ved å løse likningene

$$(A - 2I_3)\mathbf{x} = \mathbf{0} \qquad \text{og} \qquad (A - 4I_3)\mathbf{x} = \mathbf{0}.$$

Vi tar ikke med all utregningen her, men du bør gjøre den selv. Resultatet blir at egenrommet til egenverdien 2 er

$$\mathrm{Sp}\left\{ \begin{bmatrix} 3 \\ -3 \\ 5 \end{bmatrix} \right\},$$

og egenrommet til egenverdien 4 er

$$\mathrm{Sp}\left\{\begin{bmatrix}0\\1\\0\end{bmatrix},\begin{bmatrix}1\\0\\2\end{bmatrix}\right\}.$$

Egenrommet til 2 er altså en linje i $\mathbb{R}^3$, mens egenrommet til 4 er et plan. △

Diagonalmatriser

La oss se på noen eksempler på matriser der det er veldig lett å se hva egenverdiene er.

Eksempel 7.11. Har identitetsmatrisen I_n noen egenverdier? Vi vet at

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

for alle vektorer $\mathbf{v}$ i $\mathbb{R}^n$. Dermed ser vi at I_n har tallet 1 som egenverdi, med hele $\mathbb{R}^n$ som det tilhørende egenrommet. △

Eksempel 7.12. La A være følgende 2×2-matrise:

$$A = \begin{bmatrix}9 & 0\\0 & 9\end{bmatrix}$$

Da har vi at

$$A\mathbf{v} = 9\mathbf{v}$$

for alle $\mathbf{v}$ i $\mathbb{R}^2$. Det betyr at 9 er en egenverdi for A, og det tilhørende egenrommet er hele $\mathbb{R}^2$. △

Eksempel 7.13. La A være følgende 3×3-matrise:

$$A = \begin{bmatrix}7 & 0 & 0\\0 & -3 & 0\\0 & 0 & 1\end{bmatrix}$$

Da har vi at

$$A \cdot \begin{bmatrix}v_1\\v_2\\v_3\end{bmatrix} = \begin{bmatrix}7v_1\\-3v_2\\1v_3\end{bmatrix}$$

for en vektor (v_1, v_2, v_3) i $\mathbb{R}^3$. Da ser vi lett at de tre enhetsvektorene

$$\begin{bmatrix}1\\0\\0\end{bmatrix}, \quad \begin{bmatrix}0\\1\\0\end{bmatrix} \quad \text{og} \quad \begin{bmatrix}0\\0\\1\end{bmatrix}$$

er egenvektorer, med 7, -3 og 1 som tilhørende egenverdier. Det er også lett å se at hvis vi har en vektor (v_1, v_2, v_3) der minst to av komponentene v_1, v_2 og v_3 ikke er 0, så kan den ikke være en egenvektor, siden komponentene ganges opp med ulike tall når vi ganger vektoren med A.

Vi ser altså at matrisen har egenverdiene 7, -3 og 1, med

$$\mathrm{Sp}\left\{\begin{bmatrix}1\\0\\0\end{bmatrix}\right\}, \quad \mathrm{Sp}\left\{\begin{bmatrix}0\\1\\0\end{bmatrix}\right\} \quad \text{og} \quad \mathrm{Sp}\left\{\begin{bmatrix}0\\0\\1\end{bmatrix}\right\}$$

som tilhørende egenrom. △

I alle disse tre eksemplene hadde vi matriser der de eneste tallene som ikke er 0 er på diagonalen. Vi gir et navn til slike matriser.

Definisjon. En *diagonalmatrise* er en kvadratisk matrise der alle tall utenfor diagonalen er 0, altså en matrise på følgende form:

$$\begin{bmatrix}a_{11} & 0 & 0 & \cdots & 0\\0 & a_{22} & 0 & \cdots & 0\\0 & 0 & a_{33} & \cdots & 0\\\vdots & \vdots & \vdots & \ddots & \vdots\\0 & 0 & 0 & \vdots & a_{nn}\end{bmatrix}$$

△

På samme måte som i eksemplene over kan vi lett finne egenverdiene til enhver diagonalmatrise.

Teorem 7.14. *Egenverdiene til en diagonalmatrise er tallene på diagonalen.*

Lineær uavhengighet av egenvektorer

Vi skal nå ta en ganske omfattende diskusjon om hvorvidt egenvektorer er lineært uavhengige av hverandre, og hva vi kan si om vektorer som ligger i mengden utspent av noen gitte egenvektorer.

Hele diskusjonen kommer til å bli konkludert med et fint og flott teorem (teorem 7.15), men la oss ikke se på konklusjonen helt med en gang. For å prøve å forstå hvordan man kunne kommet frem til det teoremet på egen hånd, skal vi bygge oss opp mot det i små steg. Vi begynner med å se på én egenvektor, og deretter to egenvektorer som hører til forskjellige egenverdier. Vi ser hva vi kan si om disse, og etter hvert kommer vi til å nærme oss et generelt resultat.

Gjennom alt det vi skal gjøre nå lar vi A være en $n \times n$-matrise.

Vi vet fra før (teorem 7.3) at hvis vi har en egenvektor $\mathbf{v}$ som hører til en egenverdi λ, så er enhver vektor i $\mathrm{Sp}\{\mathbf{v}\}$, utenom nullvektoren, også en egenvektor som hører til λ. Vi har altså:

Hvis $\mathbf{w}$ *er en skalar ganger en egenvektor* $\mathbf{v}$*, og ikke er lik nullvektoren, så er* $\mathbf{w}$ *en egenvektor tilhørende samme egenverdi som* $\mathbf{v}$*.*

Dette enkle og greie resultatet skal vi utnytte når vi nå går videre til å se på flere egenvektorer som hører til forskjellige egenverdier.

La oss se på hva vi kan si hvis vi har to egenvektorer

$$\mathbf{v}_1 \text{ og } \mathbf{v}_2$$

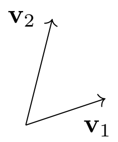
som hører til to forskjellige egenverdier

$$\lambda_1 \text{ og } \lambda_2.$$

Vi vet at alle vektorer i $\mathrm{Sp}\{\mathbf{v}_1\}$ er egenvektorer som hører til egenverdien λ_1. Dermed er det klart at $\mathbf{v}_2$ ikke kan være i $\mathrm{Sp}\{\mathbf{v}_1\}$, siden $\mathbf{v}_2$ hører til egenverdien λ_2. På samme måte ser vi at $\mathbf{v}_1$ ikke kan være i $\mathrm{Sp}\{\mathbf{v}_2\}$. Dermed ser vi (ved å bruke teorem 5.5, eller ved å tenke selv) at vektorene $\mathbf{v}_1$ og $\mathbf{v}_2$ er lineært uavhengige. Vi har altså vist:

To egenvektorer som hører til forskjellige egenverdier må være lineært uavhengige.

Dette betyr at vi kan se for oss de to vektorene $\mathbf{v}_1$ og $\mathbf{v}_2$ som to piler som peker i forskjellige retninger:



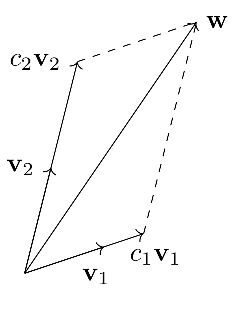
De to egenvektorene våre

La oss nå se hva vi kan si om en vektor $\mathbf{w}$ i $\mathrm{Sp}\{\mathbf{v}_1, \mathbf{v}_2\}$, altså i planet utspent av $\mathbf{v}_1$ og $\mathbf{v}_2$. En slik vektor $\mathbf{w}$ er en lineærkombinasjon av $\mathbf{v}_1$ og $\mathbf{v}_2$, så det finnes tall c_1 og c_2 slik at

$$\mathbf{w} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2.$$

For det første: Hvis $c_2 = 0$, så er $\mathbf{w}$ bare et tall ganger $\mathbf{v}_1$, og da vet vi at den er en egenvektor som hører til egenverdien λ_1. På samme måte får vi at hvis $c_1 = 0$, så er $\mathbf{w}$ en egenvektor som hører til egenverdien λ_2. Men hva kan vi si dersom både c_1 og c_2 er forskjellige fra 0, altså dersom $\mathbf{w}$ ikke ligger i $\mathrm{Sp}\{\mathbf{v}_1\}$ eller i $\mathrm{Sp}\{\mathbf{v}_2\}$?

Vi tegner inn $\mathbf{w}$ på tegningen vår, og får med hvordan den er en lineærkombinasjon av $\mathbf{v}_1$ og $\mathbf{v}_2$:

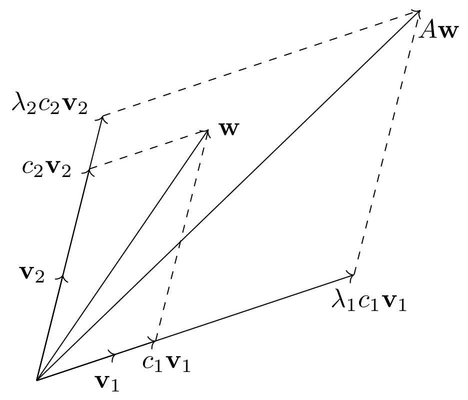


Vektoren $\mathbf{w}$ *er en lineærkombinasjon av* $\mathbf{v}_1$ *og* $\mathbf{v}_2$

Nå vil vi finne ut om $\mathbf{w}$ er en egenvektor eller ikke, altså om det finnes et tall λ slik at $A\mathbf{w} = \lambda\mathbf{w}$. Vi ser på uttrykket $A\mathbf{w}$. Siden vi vet at $\mathbf{v}_1$ og $\mathbf{v}_2$ er egenvektorer tilhørende egenverdiene λ_1 og λ_2, får vi:

$$A\mathbf{w} = Ac_1\mathbf{v}_1 + Ac_2\mathbf{v}_2 = \lambda_1 c_1\mathbf{v}_1 + \lambda_2 c_2\mathbf{v}_2$$

Vi tegner inn $A\mathbf{w}$ også på tegningen vår:



Vektoren $A\mathbf{w}$ *må peke i en annen retning enn* $\mathbf{w}$ *fordi* $\lambda_1 \neq \lambda_2$

Denne siste figuren viser at $\mathbf{w}$ ikke kan være en egenvektor. Den eneste muligheten for å få $A\mathbf{w}$ til å peke i samme retning som $\mathbf{w}$ er å anta at $\lambda_1 = \lambda_2$, og vi har jo antatt akkurat det motsatte, nemlig at $\lambda_1 \neq \lambda_2$.

Denne figuren er det man bør se for seg for å forstå hva som skjer, men vi kan vise det samme på en mer presis og rent algebraisk måte som ikke avhenger av figuren.

Hvis vi ganger $\mathbf{w}$ med et tall λ, så får vi:

$$\lambda\mathbf{w} = \lambda c_1\mathbf{v}_1 + \lambda c_2\mathbf{v}_2$$

Og vi så at hvis vi ganger A med $\mathbf{w}$, så får vi:

$$A\mathbf{w} = \lambda_1 c_1\mathbf{v}_1 + \lambda_2 c_2\mathbf{v}_2$$

Hvis $\mathbf{w}$ er en egenverdi, så finnes det en λ slik at $A\mathbf{w} = \lambda\mathbf{w}$, og dermed:

$$\lambda_1 c_1\mathbf{v}_1 + \lambda_2 c_2\mathbf{v}_2 = \lambda c_1\mathbf{v}_1 + \lambda c_2\mathbf{v}_2$$

Siden $\mathbf{v}_1$ og $\mathbf{v}_2$ er lineært uavhengige, betyr dette at

$$\lambda_1 c_1 = \lambda c_1 \qquad \text{og} \qquad \lambda_2 c_2 = \lambda c_2.$$

Men vi har antatt at både c_1 og c_2 er forskjellige fra 0 og at $\lambda_1 \neq \lambda_2$, og da er dette umulig. Det vil si at $\mathbf{w}$ ikke er en egenvektor.

Nå har vi vist:

Hvis $\mathbf{w}$ *er en lineærkombinasjon av to egenvektorer som hører til forskjellige egenverdier, så er det to muligheter: Enten er* $\mathbf{w}$ *lik en skalar ganger én av de to egenvektorene, eller så er* $\mathbf{w}$ *ikke en egenvektor.*

Nå går vi videre til å se på tre egenvektorer

$$\mathbf{v}_1,\ \mathbf{v}_2 \text{ og } \mathbf{v}_3$$

som hører til forskjellige egenverdier

$$\lambda_1,\ \lambda_2 \text{ og } \lambda_3.$$

Da kan vi bruke det vi nettopp viste til å konkludere med at ingen av disse egenvektorene kan være en lineærkombinasjon av de andre to. For hvis $\mathbf{v}_3$ skulle vært i $\mathrm{Sp}\{\mathbf{v}_1, \mathbf{v}_2\}$, så ville vi fått at den enten er nullvektoren, eller en skalar ganger $\mathbf{v}_1$ eller $\mathbf{v}_2$ (men da ville den hatt λ_1 eller λ_2 som tilhørende egenverdi istedenfor λ_3), eller så ville den ikke vært en egenvektor i det hele tatt. Alle disse alternativene er utelukket, siden vi har antatt at $\mathbf{v}_3$ er en egenvektor tilhørende egenverdien λ_3.

På samme måte får vi at $\mathbf{v}_2$ ikke kan være i $\mathrm{Sp}\{\mathbf{v}_1, \mathbf{v}_3\}$ og at $\mathbf{v}_1$ ikke kan være i $\mathrm{Sp}\{\mathbf{v}_2, \mathbf{v}_3\}$.

Siden vi har vist at ingen av de tre vektorene er en lineærkombinasjon av de andre to, får vi (fra teorem 5.11) at de er lineært uavhengige. Vi har altså vist:

Tre egenvektorer som hører til forskjellige egenverdier må være lineært uavhengige.

Men vi trenger ikke å gi oss der. Ved et helt tilsvarende argument som det vi hadde i tilfellet med to vektorer kan vi vise:

Hvis $\mathbf{w}$ *er en lineærkombinasjon av tre egenvektorer som hører til forskjellige egenverdier, så er det to muligheter: Enten er* $\mathbf{w}$ *lik en skalar ganger én av de tre egenvektorene, eller så er* $\mathbf{w}$ *ikke en egenvektor.*

Hvis du fremdeles henger med, så ser du antagelig hvordan vi nå kan gå videre til:

Fire egenvektorer som hører til forskjellige egenverdier må være lineært uavhengige.

Herfra kan vi fortsette på akkurat samme måte, og vi får de samme resultatene for enhver liste av vilkårlig mange egenvektorer som hører til forskjellige egenverdier. Da har vi kommet til slutten på vår lange diskusjon om lineær uavhengighet for egenvektorer, og vi oppsummerer det vi har funnet ut i et teorem.

Teorem 7.15. *La* $\lambda_1, \lambda_2, \ldots, \lambda_t$ *være forskjellige egenverdier for en matrise* A*, og la* $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_t$ *være egenvektorer som hører til hver av disse egenverdiene. Da har vi:*

(a) Vektorene $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_t$ *er lineært uavhengige.*

(b) I mengden

$$\mathrm{Sp}\{\mathbf{v}_1,\ \mathbf{v}_2,\ \cdots,\ \mathbf{v}_t\}$$

utspent av egenvektorene vi ser på finnes det ingen andre egenvektorer enn de som er et multippel $c \cdot \mathbf{v}_i$ *av én av vektorene i listen.*