

Flervariabel analyse med lineær algebra

Del 1: Et innføringshefte for MAT 1100

av

Tom Lindstrøm og Klara Hveberg

Matematisk institutt
og
Senter for matematikk for anvendelser (CMA)
Universitetet i Oslo
Revidert versjon 2007

Forord

Dette heftet er første del av et kompendium som dekker pensum i lineær algebra og flervariabel analyse i kursene MAT 1100 og MAT 1110 ved Universitetet i Oslo. I de to kapitlene som her foreligger, gjennomgår vi først vektor- og matriseregning i \mathbb{R}^n og deretter den grunnleggende teorien for funksjoner av flere variable (inklusive kjerneregelen, men ikke maks/min-problemer og multippel integrasjon).

Opplegget vårt har to grunntanker. For det første vil vi utnytte samspillet mellom lineær algebra og flervariabel analyse på en bedre måte enn det som er vanlig i “standardopplegget” der disse to emnene undervises i separate kurs. Etter vår mening tjener begge emnene på å bli sett i sammenheng — flervariabel analyse blir begrepsmessig enklere og mer oversiktlig når man kan bruke notasjon og terminologi fra lineæralgebraen, og lineæralgebraen vinner nye anvendelsesområder når den blir knyttet til flervariabel analyse. Vår andre grunntanke er at undervisningen i lineær algebra og flervariabel analyse bør knyttes nærmere til numeriske problemstillinger og bruk av dataverktøy. Siden MAT 1100 er et kurs med grafiske lommeregnerne som eneste elektroniske hjelpemiddel, er ikke denne tanken så synlig i de kapitlene som her foreligger, med den vil komme tydeligere frem både i stoffutvalg og presentasjonsform i de neste kapitlene. Vi har imidlertid prøvd å legge forholdene til rette allerede i de første kapitlene ved å trekke frem begreper som iterasjon, linearisering og egenverdier/egenvektorer på et tidlig stadium. Så selv om numerikken er så godt som fraværende i dette heftet, ser vi det likevel som en del av fakultetets kampanje for å styrke de numeriske sidene av realfagsutdanningen!

Heftet forutsetter at studentene har en solid forståelse av kontinuitet og deriverbarhet av funksjoner av én variabel, f.eks. tilsvarende de første sju kapitlene i Tom Lindstrøms lærebok *Kalkulus* (siden denne boken brukes i de andre delene av MAT 1100, finnes det en del referanser til den i teksten). Siden heftet er del av et større kompendium, finnes det noen få bemerkninger av typen “som vi senere skal se” med henvisninger til senere kapitler.

To ord om notasjon: Vi bruker \square til å markere slutten på et bevis og \clubsuit til å markere slutten på et eksempel.

Endringer i den reviderte utgaven: I den reviderte utgaven (2007) har

vi stokket litt om på teksten. Seksjonene om “Vektorprodukt” og “Determinanter, arealer og volumer” er flyttet fra kapittel 3 til kapittel 1. Til gjengjeld er en del topologisk stoff (åpne og lukkede mengder, følger i \mathbb{R}^n) tatt ut av kapittel 2 (dette stoffet vil isteden bli tatt i et senere kapittel). Vi har også gjort en del mindre endringer som forhåpentligvis gjør teksten enklere å lese. Forøvrig takker vi alle som har funnet trykkfeil i den første utgaven; det er sikkert flere igjen, og finner du noen, så send en e-post til *lindstro@math.uio.no*.

Blindern 19/8-2007

Tom Lindstrøm

Klara Hveberg

Innhold

1	Vektorer og matriser	1
1.1	Algebra for n -tupler	2
1.2	Geometri for n -tupler	6
1.3	Komplekse n -tupler	17
1.4	Vektorproduktet	21
1.5	Matriser	32
1.6	Multiplikasjon av matriser	42
1.7	Identitetsmatriser og inverse matriser	52
1.8	Determinanter, arealer og volumer	58
2	Funksjoner fra \mathbb{R}^n til \mathbb{R}^m	71
2.1	Funksjoner av flere variable	71
2.2	Kontinuerlige funksjoner	75
2.3	Grenseverdier	80
2.4	Derivasjon av skalarfelt	84
2.5	Partiellderiverte av høyere orden	96
2.6	Derivasjon av vektorvaluerte funksjoner	100
2.7	Kjerneregelen	105
2.8	Lineærabildninger	116
2.9	Affinabildninger og lineariseringer	125

Kapittel 1

Vektorer og matriser

Tall spiller en sentral rolle i matematikken — så sentral at mange nok vil si at der det er tall, er det matematikk, og der det ikke er tall, er det ikke matematikk! Fullt så enkelt er det ikke — det finnes mange grener av matematikken der tall spiller en underordnet rolle — men det er likevel ikke til å komme forbi at tall er et av fagets aller viktigste bestanddeler.

I din tidligere matematikkutdanning har du lært å regne med mange slags tall: hele tall, desimaltall, brøker, irrasjonale tall og til og med komplekse tall. I dette kapitlet skal vi gå et skritt videre og regne med tupler av tall, dvs. flere tall på en gang. Du har vært borti dette tidligere når du har regnet med vektorer i planet og i rommet — en vektor $[x, y]$ i planet er et 2-tupple, mens en vektor $[x, y, z]$ i rommet er et 3-tupple. Vi skal nå gå videre og regne med n -tupler for alle naturlige tall n . Hvis du tenker geometrisk, kan dette høres skummelt ut — hvordan skal man kunne forestille seg en 4-dimensjonal vektor $[x, y, z, u]$? Tenker du mer algebraisk, er det ikke noe skummelt i det hele tatt; et 4-tupple $[x, y, z, u]$ er bare en notasjon for å holde styr på fire tall på en praktisk og kortfattet måte. I dette kapitlet skal vi utvikle både algebra og geometri, for selv om det å regne algebraisk med tupler er trygt og ukomplisert, mister man fort oversikten, og den gjenvinner man først når man lærer å tenke på tupler som geometriske objekter.

Vi skal også gå et skritt videre og arbeide med matriser. Dette er rektangulære oppsett av tall som f.eks.

$$\begin{pmatrix} 2 & -1 & 4 \\ 0 & -3 & 2 \end{pmatrix} \quad \text{og} \quad \begin{pmatrix} 3 & 2 \\ -1 & 12 \\ \frac{3}{2} & -2 \end{pmatrix}$$

Ved hjelp av matriser kan vi “transformere” n -tupler på en måte som er viktig i svært mange sammenhenger, både regneteknisk og geometrisk. Matriser og tupler kommer til å spille en sentral rolle også i senere kapitler, dels som nyttige verktøy og dels som selvstendige studieobjekter.

1.1 Algebra for n -tupler

La oss begynne med den grunnleggende definisjonen. Et n -tupple er et uttrykk (a_1, a_2, \dots, a_n) der a_1, a_2, \dots, a_n er reelle tall. Vi ser at $(2, -1, 7, 3)$ er et 4-tupple, mens $(0, 1, \pi, \frac{3}{2}, -7, 3)$ er et 6-tupple. To n -tupler (a_1, a_2, \dots, a_n) og (b_1, b_2, \dots, b_n) regnes som like dersom de inneholder de samme tallene i samme rekkefølge, dvs. hvis $a_1 = b_1, a_2 = b_2, \dots, a_n = b_n$. Legg merke til at $(3, 2, 4) \neq (2, 3, 4)$; selv om tallene er de samme, er rekkefølgen forskjellig.

I dette heftet skal vi bruke bokstaver i **fete typer** som navn på n -tupler, f.eks. $\mathbf{a} = (-2, 3, 0, -17)$. Det er vanskelig å bruke fete typer når man skriver for hånd, og man kan da isteden skrive en pil eller en strek over bokstavene; dvs. $\vec{a} = (-2, 3, 0, -17)$ eller $\bar{a} = (-2, 3, 0, -17)$.

Vi skriver $\mathbf{0}$ for det n -tuplet som har alle komponenter lik 0, altså $\mathbf{0} = (0, 0, \dots, 0)$. Hvis vi har et n -tupple $\mathbf{a} = (a_1, a_2, \dots, a_n)$, skriver vi $-\mathbf{a}$ for n -tuplet $(-a_1, -a_2, \dots, -a_n)$.

Det er en naturlig måte å definere addisjon og subtraksjon av n -tupler på. Dersom $\mathbf{a} = (a_1, a_2, \dots, a_n)$ og $\mathbf{b} = (b_1, b_2, \dots, b_n)$, så er

$$\mathbf{a} + \mathbf{b} = (a_1 + b_1, a_2 + b_2, \dots, a_n + b_n)$$

og

$$\mathbf{a} - \mathbf{b} = (a_1 - b_1, a_2 - b_2, \dots, a_n - b_n)$$

Vi sier at vi adderer og subtraherer *komponentvis*. Legg merke til at vi bare kan addere og subtrahere tupler med like mange komponenter — oppskriften ovenfor gir oss ikke noen måte å addere et 3-tupple og et 7-tupple på. Før vi ser på et eksempel, tar vi med en regneoperasjon til. Dersom s er et tall og $\mathbf{a} = (a_1, a_2, \dots, a_n)$ er et n -tupple, definerer vi produktet av s og \mathbf{a} til å være

$$s\mathbf{a} = (sa_1, sa_2, \dots, sa_n)$$

Vi ganger altså s inn i hver komponent i \mathbf{a} .

Eksempel 1. Vi lar $\mathbf{a} = (-2, 3, 0, -17)$ og $\mathbf{b} = (4, -1, 3, 17)$. Da er

$$\mathbf{a} + \mathbf{b} = (-2 + 4, 3 + (-1), 0 + 3, -17 + 17) = (2, 2, 3, 0)$$

og

$$\mathbf{a} - \mathbf{b} = (-2 - 4, 3 - (-1), 0 - 3, -17 - 17) = (-6, 4, -3, -34)$$

Hvis $s = 3$, får vi

$$s\mathbf{a} = (3 \cdot (-2), 3 \cdot 3, 3 \cdot 0, 3 \cdot (-17)) = (-6, 9, 0, -51)$$



Vi skal innføre en regneoperasjon til. Dersom $\mathbf{a} = (a_1, a_2, \dots, a_n)$ og $\mathbf{b} = (b_1, b_2, \dots, b_n)$ er to n -tupler, definerer vi *skalarproduktet* (også kalt *prikkproduktet*) $\mathbf{a} \cdot \mathbf{b}$ ved

$$\mathbf{a} \cdot \mathbf{b} = a_1 b_1 + a_2 b_2 + \dots + a_n b_n$$

Legg merke til at $\mathbf{a} \cdot \mathbf{b}$ ikke er et n -tupple, men et tall (eller en *skalar* som man ofte sier når man vil understreke at noe er et tall og ikke et n -tupple). Hvis vi lar $\mathbf{a} = (-2, 3, 0, -17)$ og $\mathbf{b} = (4, -1, 3, 17)$ som ovenfor, ser vi at

$$\mathbf{a} \cdot \mathbf{b} = (-2) \cdot 4 + 3 \cdot (-1) + 0 \cdot 3 + (-17) \cdot 17 = -8 - 3 + 0 - 289 = -300$$

Vi har nå sett hvordan vi kan regne med n -tupler, og det er kanskje på tide å ta en kikk på noen eksempler som antyder hvorfor det er et poeng med slike regnestykker. Det første eksemplet viser at n -tupler er naturlige redskap når vi skal holde styr på mer informasjon enn det som kan rommes i et enkelt tall, og at regneoperasjonene svarer til regnestykker det ofte er naturlig å utføre i slike sammenhenger.

Eksempel 2. En forretning har ansatt 7 studenter på timebasis. For å holde styr på hvor mange timer hver student har arbeidet så langt, kan vi bruke et 7-tupple $\mathbf{t} = (t_1, t_2, \dots, t_7)$ der t_1 er antall timer den første studenten har arbeidet, t_2 er antall timer den andre studenten har arbeidet osv. Dersom studentene arbeider mer senere, kan vi på samme måte kode tilleggstimene som et 7-tupple $\mathbf{s} = (s_1, s_2, \dots, s_7)$. Det totale antall timer som studentene har arbeidet, er nå gitt ved $\mathbf{t} + \mathbf{s}$.

Studentene har ulik erfaring og derfor ulik lønn. Hvis student nummer én har en timelønn på p_1 kroner, student nummer to har en timelønn på p_2 kroner osv., kan vi også representere lønnen som et 7-tupple $\mathbf{p} = (p_1, p_2, \dots, p_7)$. Dersom studentene har arbeidet $\mathbf{t} = (t_1, t_2, \dots, t_7)$ timer, er den totale lønnen som forretningen skylder, gitt av skalarproduktet $\mathbf{p} \cdot \mathbf{t} = p_1 t_1 + p_2 t_2 + \dots + p_7 t_7$. Dersom alle studentene får et lønnstillegg på 7 prosent, får vi det nye lønnstuplet ved å gange det gamle med skalar 1.07, altså $1.07\mathbf{p}$.



Vi tar med noen eksempler til som viser hvordan n -tupler brukes til å holde styr på tallmessig informasjon i forskjellige sammenhenger.

Eksempel 3. Tilstanden til en gassbeholder er bestemt av trykket p , temperaturen T og volumet V . Hvis du får i oppdrag å måle tilstanden til beholderen ved forskjellige tidspunkt, kan det være naturlig å bruke 4-tupler $\mathbf{a} = (t, p, T, V)$ der t er tidspunktet for målingen. Forskjellen mellom to målinger \mathbf{a} og \mathbf{b} er da gitt ved differensen $\mathbf{b} - \mathbf{a}$.



Eksempel 4. Et bilde på en fjernsynsskjerm eller en dataskjerm er bygget opp av små lysende punkter (pikslar). Et vanlig format er $1280 \times 1024 = 1310720$ pikslar. I hvert punkt må vi angi styrken til hver av de tre grunnfargene rødt, grønt og blått, så totalt har vi $3 \times 1310720 = 3932160$ tall å holde styr på. En naturlig måte å gjøre dette på er å oppfatte bilder som 3932160-tupler! Dette er ikke noe enestående eksempel — i mange anvendelser er man interessert i tupler med svært mange komponenter. ♣

Her er noen enkle regneregler for n -tupler (det finnes flere):

Setning 1.1.1 (Regneregler for n -tupler) Dersom \mathbf{a} , \mathbf{b} og \mathbf{c} er n -tupler og s og t er reelle tall, gjelder følgende regneregler:

- (a) $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$
- (b) $\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}$
- (c) $s(\mathbf{a} + \mathbf{b}) = s\mathbf{a} + s\mathbf{b}$
- (d) $(s + t)\mathbf{a} = s\mathbf{a} + t\mathbf{a}$
- (e) $\mathbf{c} \cdot (\mathbf{a} + \mathbf{b}) = \mathbf{c} \cdot \mathbf{a} + \mathbf{c} \cdot \mathbf{b}$ og $(\mathbf{a} + \mathbf{b}) \cdot \mathbf{c} = \mathbf{a} \cdot \mathbf{c} + \mathbf{b} \cdot \mathbf{c}$
- (f) $(s\mathbf{a}) \cdot \mathbf{b} = \mathbf{a} \cdot (s\mathbf{b}) = s(\mathbf{a} \cdot \mathbf{b})$
- (g) $\mathbf{a} \cdot \mathbf{a} \geq 0$ med likhet hvis og bare hvis $\mathbf{a} = \mathbf{0}$

Bevis: Alle disse reglene bevises lett ved å regne ut venstre- og høyresiden og kontrollere at svarene stemmer overens. Vi tar (c) og (g) som eksempler: (c) Dersom $\mathbf{a} = (a_1, a_2, \dots, a_n)$ og $\mathbf{b} = (b_1, b_2, \dots, b_n)$, ser vi at venstresiden kan skrives

$$\begin{aligned} s(\mathbf{a} + \mathbf{b}) &= s(a_1 + b_1, a_2 + b_2, \dots, a_n + b_n) \\ &= (s(a_1 + b_1), s(a_2 + b_2), \dots, s(a_n + b_n)) \\ &= (sa_1 + sb_1, sa_2 + sb_2, \dots, sa_n + sb_n) \end{aligned}$$

Tilsvarende kan høyresiden skrives

$$\begin{aligned} s\mathbf{a} + s\mathbf{b} &= (sa_1, sa_2, \dots, sa_n) + (sb_1, sb_2, \dots, sb_n) \\ &= (sa_1 + sb_1, sa_2 + sb_2, \dots, sa_n + sb_n) \end{aligned}$$

Siden de to uttrykkene er like, er (c) bevist.

(g) Vi ser at

$$\mathbf{a} \cdot \mathbf{a} = a_1^2 + a_2^2 + \dots + a_n^2 \geq 0$$

siden kvadrater aldri er negative. Likhet har vi dersom $a_1^2 = 0$, $a_2^2 = 0$, \dots , $a_n^2 = 0$, dvs. dersom $a_1 = 0$, $a_2 = 0$, \dots , $a_n = 0$. \square

Vi har hittil skrevet våre n -tupler liggende

$$\mathbf{a} = (a_1, a_2, \dots, a_n),$$

men vi kan også skrive dem stående

$$\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}.$$

I det første tilfellet kaller vi \mathbf{a} en *radvektor*, mens vi i det andre kaller \mathbf{a} en *søylevektor*. I de fleste situasjoner spiller det ingen rolle om vi skriver n -tuplene på den ene eller andre formen, og vi velger da ofte å skrive dem som radvektorer siden det tar minst plass. Det finnes imidlertid tilfeller der det er viktig å skille mellom radvektorer og søylevektorer, men det skal vi komme tilbake til senere — foreløpig kan du skrive dine vektorer på den måten du måtte ønske. Legg for øvrig merke til at det ofte kan være lettere å få øye på strukturen i et regnestykke når du bruker søylevektorer, f.eks. kan

$$s \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} + t \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} + r \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} sa_1 + tb_1 + rc_1 \\ sa_2 + tb_2 + rc_2 \\ \vdots \\ sa_n + tb_n + rc_n \end{pmatrix}$$

virke mer oversiktlig enn

$$\begin{aligned} & s(a_1, a_2, \dots, a_n) + t(b_1, b_2, \dots, b_n) + r(c_1, c_2, \dots, c_n) = \\ & = (sa_1 + tb_1 + rc_1, sa_2 + tb_2 + rc_2, \dots, sa_n + tb_n + rc_n) \end{aligned}$$

La oss avslutte denne seksjonen med noen flere ord om notasjon. Mengden av alle n -tupler kaller vi \mathbb{R}^n . Når vi skriver $\mathbf{a} \in \mathbb{R}^n$, betyr dette derfor ikke noe annet enn at \mathbf{a} er et n -tupple. Hittil har vi holdt oss til reelle n -tupler, men vi kan selvfølgelig også tenke oss n -tupler (c_1, c_2, \dots, c_n) der komponentene c_1, c_2, \dots, c_n er *komplekse* tall. Mengden av alle slike n -tupler kaller vi \mathbb{C}^n . Vi skal se nærmere på komplekse n -tupler litt senere. Notasjonen kan også gjøres enda mer generell: Dersom A er en hvilken som helst mengde, betegner A^n mengden av alle n -tupler (a_1, a_2, \dots, a_n) der $a_i \in A$ for alle $i = 1, 2, \dots, n$.

Helt til slutt legger vi merke til at et 1-tupple (a_1) ikke er noe annet enn et tall inni en parentes. Parentesen spiller ingen rolle (den er bare med for å avgrense uttrykket), og vi skal derfor regne 1-tuplet (a_1) og tallet a_1 som det samme objektet. Dette betyr at \mathbb{R}^1 og \mathbb{R} er den samme mengden.

Oppgaver til seksjon 1.1

1. Finn $\mathbf{a} + \mathbf{b}$, $\mathbf{a} - \mathbf{b}$, $s\mathbf{a}$ og $\mathbf{a} \cdot \mathbf{b}$ når $\mathbf{a} = (1, -2, 4, -5, 1)$, $\mathbf{b} = (-3, 5, 5, 0, -3)$ og $s = 3$.

2. Finn $\mathbf{a} + \mathbf{b}$, $\mathbf{a} - \mathbf{b}$, $s\mathbf{a}$ og $\mathbf{a} \cdot \mathbf{b}$ når $\mathbf{a} = (7, 0, 4, -2, -5, 4)$, $\mathbf{b} = (0, 2, 1, -6, 0, -1)$ og $s = -4$.

3. Vis at for alle $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ er:

a) $(\mathbf{x} + \mathbf{y}) \cdot (\mathbf{x} + \mathbf{y}) = \mathbf{x} \cdot \mathbf{x} + 2\mathbf{x} \cdot \mathbf{y} + \mathbf{y} \cdot \mathbf{y}$

b) $(\mathbf{x} - \mathbf{y}) \cdot (\mathbf{x} - \mathbf{y}) = \mathbf{x} \cdot \mathbf{x} - 2\mathbf{x} \cdot \mathbf{y} + \mathbf{y} \cdot \mathbf{y}$

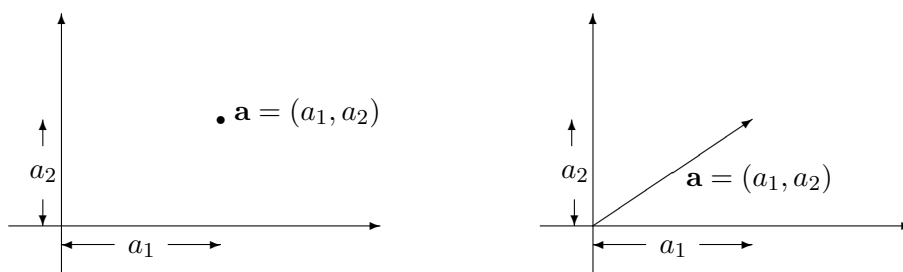
c) $(\mathbf{x} + \mathbf{y}) \cdot (\mathbf{x} - \mathbf{y}) = \mathbf{x} \cdot \mathbf{x} - \mathbf{y} \cdot \mathbf{y}$.

4. Bevis punktene d), e), f) i setning 1.1.1.

5. Et grossistfirma har n vareslag på lager, m_1 enheter av vareslag 1, m_2 enheter av vareslag 2 osv. Verdien av hver enhet er p_1 for vareslag 1, p_2 for vareslag 2 osv. Uttrykk den totale verdien av varelageret som skalarproduktet mellom to n -tupler.

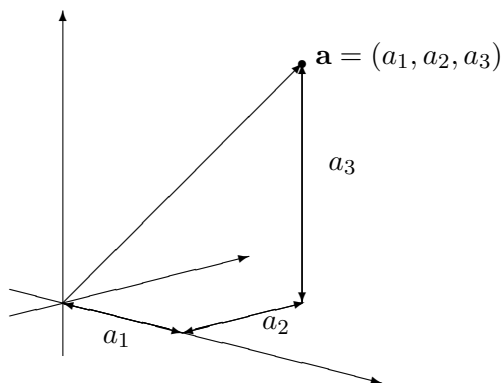
1.2 Geometri for n -tupler

Et 2-tupple er ikke noe annet enn et par (a_1, a_2) . Geometrisk kan vi tenke på et slikt par på to måter — enten som et *punkt* med koordinater a_1 og a_2 , eller som en *vektor* (pil) som starter i origo og ender i dette punktet (se figur 1). I skolematematikken bruker man gjerne forskjellig notasjon ettersom man tenker på paret som et punkt eller som en vektor — et punkt (a_1, a_2) har runde parenteser, mens en vektor $[a_1, a_2]$ har klammeparenteser. Det er ganske tungvint å bruke to forskjellige notasjoner, og vi vil derfor bruke runde parenteser $\mathbf{a} = (a_1, a_2)$ uansett om vi tenker på \mathbf{a} som et punkt eller som en vektor. Hva som er naturlig, fremgår som regel av sammenhengen. Snakker vi om en linje gjennom \mathbf{a} , er det naturlig å tenke på \mathbf{a} som et punkt, men snakker vi om en linje parallell med \mathbf{a} , er det naturlig å tenke på \mathbf{a} som en vektor. Når vi lager figurer, vil vi noen ganger tegne paret (a_1, a_2) som en vektor og andre ganger som et punkt, alt etter hva vi synes passer best i hvert enkelt tilfelle (se figur 1).



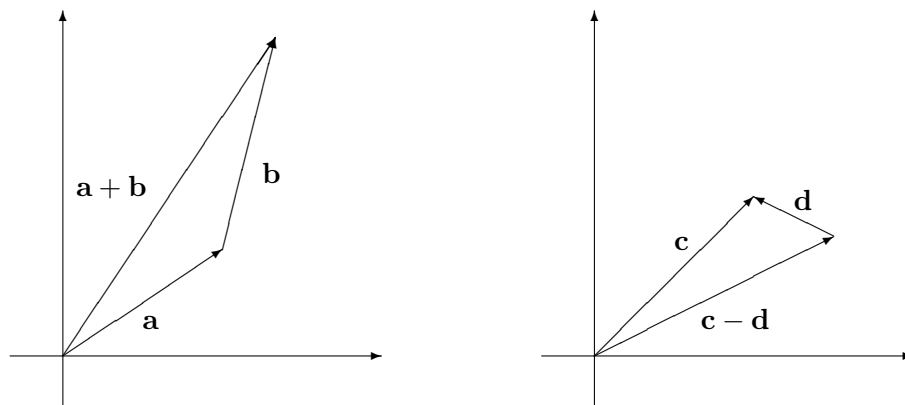
Figur 1: \mathbf{a} som et punkt og som en vektor

På tilsvarende vis kan vi oppfatte 3-tupler som punkter og vektorer i rommet. Figuren nedenfor viser hvordan et 3-tupplel $\mathbf{a} = (a_1, a_2, a_3)$ kan oppfattes som en vektor i rommet.



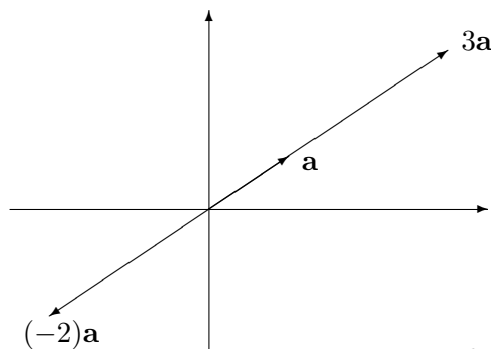
Figur 2: Et 3-tupple som en vektor i rommet

Som du vet fra skolematematikken, har de algebraiske operasjonene vi innførte i forrige seksjon, en geometrisk tolkning når vi tenker på tupler som vektorer i planet eller rommet. Figur 3 viser hvordan vi får frem addisjon og subtraksjon ved å sette sammen vektorer:



Figur 3: Addisjon og subtraksjon av vektorer

Multiplikasjon med en skalar har også en geometrisk tolkning. Dersom vi ganger \mathbf{a} med et *positivt* tall s , beholder vektoren retningen, men blir s ganger så lang. Dersom vi ganger \mathbf{a} med et *negativt* tall s , snur retningen 180° , og den nye vektoren blir $|s|$ ganger så lang som den opprinnelige (se figur 4).



Figur 4: Multiplikasjon med et tall

Det er også andre sammenhenger mellom regneoperasjoner og geometri. Fra skolematematikken vet du for eksempel at to (ikke-null) vektorer \mathbf{a} , \mathbf{b} er parallelle dersom det finnes et tall $s \neq 0$ slik at $\mathbf{a} = s\mathbf{b}$, og at de er ortogonale (dvs. står normalt på hverandre) dersom $\mathbf{a} \cdot \mathbf{b} = 0$. Du vet også at lengden $|\mathbf{a}|$ til vektoren $\mathbf{a} = (a_1, a_2, a_3)$ kan regnes ut fra koordinatene:

$$|\mathbf{a}| = \sqrt{a_1^2 + a_2^2 + a_3^2}$$

og at det er en sammenheng mellom lengden og skalarproduktet:

$$|\mathbf{a}|^2 = \mathbf{a} \cdot \mathbf{a}$$

Det er sammenhengen mellom geometri og algebra som gir liv til vektorregning i to og tre dimensjoner, og det hadde vært nyttig om vi kunne bruke vår geometriske intuisjon på samme måte når vi arbeidet med generelle n -tupler. Dette kan virke som en uoverkommelig oppgave — hvis 2-tupler representerer 2-dimensjonale objekter i planet, og 3-tupler representerer 3-dimensjonale objekter i rommet, så burde 4-tupler representere 4-dimensjonale objekter i et slags 4-dimensjonalt rom? Og, enda verre, 5-tupler burde representere 5-dimensjonale objekter i et 5-dimensjonalt rom, 6-tupler burde representere 6-dimensjonale objekter i et 6-dimensjonalt rom osv? Hvem av oss kan med hånden på hjertet si at de har noen særlig geometrisk intuisjon for det som skjer i 4-, 5- og 6-dimensjonale rom?

Heldigvis behøver vi ikke å ha en slik intuisjon på forhånd, men kan bygge den opp gradvis. Ideen er enkel: vi overfører geometriske begreper fra planet og rommet til det generelle tilfellet ved å bruke de algebraiske beskrivelsene av geometriske egenskaper. Her er et eksempel: At to vektorer \mathbf{a} og \mathbf{b} er ortogonale (dvs. at de står normalt på hverandre), er i utgangspunktet en geometrisk egenskap. Denne egenskapen kan vi beskrive algebraisk ved $\mathbf{a} \cdot \mathbf{b} = 0$. Vi bruker nå denne algebraiske beskrivelsen til å *definere* at to n -tupler \mathbf{a} og \mathbf{b} er ortogonale dersom $\mathbf{a} \cdot \mathbf{b} = 0$ (vi sier da at de står

normalt på hverandre). På denne måten får vi innført det geometriske begrepet ortogonalitet i høyere dimensjoner uten å måtte støtte oss til noen geometrisk intuisjon. Når begrepet først er innført på denne måten, kan vi undersøke i hvilken grad det svarer til våre (geometriske) forestillinger om hva ortogonalitet er. På den måten bygger vi etter hvert opp en intuisjon om ortogonalitet av n -tupler, og denne intuisjonen tar fort en geometrisk form.

La oss begynne med litt terminologi: Mengden \mathbb{R}^n av alle n -tupler kalles det *n -dimensjonale euklidske rommet*, og et n -tupplel \mathbf{a} kalles også en *n -dimensjonal vektor* eller et *n -dimensjonalt punkt*. Som i det 2- og 3-dimensjonale tilfellet skal vi ofte bruke ordet “vektor” når det er naturlig å tenke på \mathbf{a} som et geometrisk objekt med lengde og retning, og vi skal bruke ordet “punkt” når vi er opptatt av noe (f.eks. en linje eller et plan) som går gjennom \mathbf{a} . Logisk sett er det selvfølgelig unødvendig å ha mer enn ett navn på disse objektene, men pedagogisk er det ofte en fordel å kunne bruke et ord som antyder hvilke egenskaper vi er opptatt av i hver enkelt situasjon. Vi har valgt det nøytrale ordet “ n -tupplel” som utgangspunkt for ikke å binde oss for sterkt til den ene eller andre tolkningen.

Hvis $\mathbf{a} = (a_1, a_2)$ er en to-dimensjonal vektor, er lengden gitt ved

$$|\mathbf{a}| = \sqrt{a_1^2 + a_2^2}$$

Tilsvarende er lengden til en tre-dimensjonal vektor $\mathbf{a} = (a_1, a_2, a_3)$ gitt ved

$$|\mathbf{a}| = \sqrt{a_1^2 + a_2^2 + a_3^2}$$

For en n -dimensjonal vektor $\mathbf{a} = (a_1, a_2, \dots, a_n)$ er det derfor naturlig å definere *lengden* (eller *normen* som den også kalles) ved

$$|\mathbf{a}| = \sqrt{a_1^2 + a_2^2 + \dots + a_n^2}$$

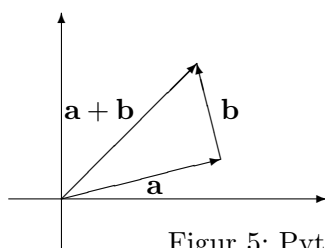
I kapittel 1 definerte vi skalarproduktet av vektorene $\mathbf{a} = (a_1, a_2, \dots, a_n)$ og $\mathbf{b} = (b_1, b_2, \dots, b_n)$ til å være

$$\mathbf{a} \cdot \mathbf{b} = a_1 b_1 + a_2 b_2 + \dots + a_n b_n$$

Vi har den vanlige sammenhengen mellom lengden og skalarproduktet

$$|\mathbf{a}| = \sqrt{\mathbf{a} \cdot \mathbf{a}} \quad \text{eller med andre ord} \quad \mathbf{a} \cdot \mathbf{a} = |\mathbf{a}|^2$$

Vi har allerede definert to n -tupler \mathbf{a} og \mathbf{b} til å være *ortogonale* (eller *stå normalt på hverandre*) dersom $\mathbf{a} \cdot \mathbf{b} = 0$. Ved hjelp av denne definisjonen kan vi formulere en n -dimensjonal versjon av et meget berømt resultat (figur 5 viser den geometriske motivasjonen).



Figur 5: Pythagoras' setning i planet

Setning 1.2.1 (Pythagoras' setning for n -tupler) Dersom $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ er ortogonale, så er

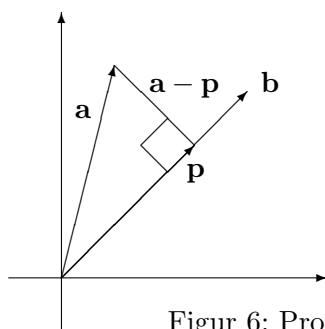
$$|\mathbf{a} + \mathbf{b}|^2 = |\mathbf{a}|^2 + |\mathbf{b}|^2$$

Bevis: Dette er bare et enkelt regnestykke (husk regnereglene for n -tupler fra kapittel 1):

$$\begin{aligned} |\mathbf{a} + \mathbf{b}|^2 &= (\mathbf{a} + \mathbf{b}) \cdot (\mathbf{a} + \mathbf{b}) = \\ &= \mathbf{a} \cdot \mathbf{a} + 2\mathbf{a} \cdot \mathbf{b} + \mathbf{b} \cdot \mathbf{b} = |\mathbf{a}|^2 + 2 \cdot 0 + |\mathbf{b}|^2 = \\ &= |\mathbf{a}|^2 + |\mathbf{b}|^2 \end{aligned}$$

□

Setningen ovenfor er vårt første eksempel på et resultat om n -tupler som er inspirert av en geometrisk observasjon. Vårt neste problem tar utgangspunkt i figur 6. Vi tenker oss at vi er gitt to vektorer \mathbf{a}, \mathbf{b} , og at vi ønsker å finne *projeksjonen* \mathbf{p} av \mathbf{a} ned på \mathbf{b} . Dette betyr at \mathbf{p} er vektoren parallell med \mathbf{b} slik at $\mathbf{a} - \mathbf{p}$ står normalt på \mathbf{b} .

Figur 6: Projeksjonen \mathbf{p} av \mathbf{a} ned på \mathbf{b}

I utgangspunktet er dette en geometrisk problemstilling som bare gir mening for vektorer i planet og rommet, men vi kan bruke vår oversettelsesfilosofi til å gi mening til problemet for generelle n -tupler \mathbf{a} og \mathbf{b} . Siden det n -tuplet \mathbf{p} vi er på jakt etter skal være parallelt med \mathbf{b} , må det finnes et tall t slik at $\mathbf{p} = t\mathbf{b}$, og siden $\mathbf{a} - \mathbf{p}$ skal stå normalt på \mathbf{b} , må vi ha

$$0 = (\mathbf{a} - \mathbf{p}) \cdot \mathbf{b} = (\mathbf{a} - t\mathbf{b}) \cdot \mathbf{b} = \mathbf{a} \cdot \mathbf{b} - t|\mathbf{b}|^2$$

Løser vi denne ligningen med hensyn på t , får vi

$$t = \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{b}|^2},$$

som betyr at

$$\mathbf{p} = t\mathbf{b} = \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{b}|^2} \mathbf{b}$$

Vi får dermed dette resultatet:

Setning 1.2.2 *Anta at \mathbf{a} og \mathbf{b} er to ikke-null vektorer i \mathbb{R}^n . Da er projeksjonen \mathbf{p} av \mathbf{a} ned på \mathbf{b} gitt ved:*

$$\mathbf{p} = \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{b}|^2} \mathbf{b}$$

Lengden til projeksjonen er $|\mathbf{p}| = \frac{|\mathbf{a} \cdot \mathbf{b}|}{|\mathbf{b}|}$.

Bevis: Den første formelen har vi allerede utledet. Den andre kan vi for eksempel finne med følgende regnestykke:

$$|\mathbf{p}| = |t\mathbf{b}| = |t||\mathbf{b}| = \frac{|\mathbf{a} \cdot \mathbf{b}|}{|\mathbf{b}|^2} |\mathbf{b}| = \frac{|\mathbf{a} \cdot \mathbf{b}|}{|\mathbf{b}|}$$

□

La oss kombinere resultatet vi nettopp har bevist med Pythagoras' setning. Siden $\mathbf{a} - \mathbf{p}$ står normalt på \mathbf{b} , må den også stå normalt på \mathbf{p} som er parallell med \mathbf{b} (sjekk dette!). Det betyr at vi kan bruke Pythagoras' setning på vektorene \mathbf{p} , $\mathbf{a} - \mathbf{p}$ og \mathbf{a} (se figur 6 for å få intuitjonen):

$$|\mathbf{a}|^2 = |\mathbf{p}|^2 + |\mathbf{a} - \mathbf{p}|^2$$

Siden $|\mathbf{a} - \mathbf{p}|^2 \geq 0$, betyr dette at

$$|\mathbf{a}|^2 \geq |\mathbf{p}|^2$$

som medfører at

$$|\mathbf{a}| \geq |\mathbf{p}|$$

(husk at både $|\mathbf{a}|$ og $|\mathbf{p}|$ er positive). Ifølge setningen ovenfor er $|\mathbf{p}| = \frac{|\mathbf{a} \cdot \mathbf{b}|}{|\mathbf{b}|}$, og setter vi dette inn i ulikheten, får vi

$$|\mathbf{a}| \geq \frac{|\mathbf{a} \cdot \mathbf{b}|}{|\mathbf{b}|}$$

Ganger vi med $|\mathbf{b}|$ på begge sider, sitter vi igjen med

$$|\mathbf{a}||\mathbf{b}| \geq |\mathbf{a} \cdot \mathbf{b}|$$

Vi har kommet frem til en berømt og meget nyttig ulikhet:

Setning 1.2.3 (Schwarz' ulikhet) For alle $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ gjelder

$$|\mathbf{a} \cdot \mathbf{b}| \leq |\mathbf{a}||\mathbf{b}|$$

Vi har likhet (dvs. $|\mathbf{a} \cdot \mathbf{b}| = |\mathbf{a}||\mathbf{b}|$) hvis og bare hvis \mathbf{a} og \mathbf{b} er parallelle eller minst én av dem er null.

Bevis: I utledningen av ulikheten har vi strengt tatt gått ut i fra at $\mathbf{a}, \mathbf{b} \neq \mathbf{0}$, men ulikheten gjelder åpenbart også om én eller begge vektorer er lik $\mathbf{0}$ (for da er venstresiden i ulikheten lik 0). Det gjenstår dermed bare å sjekke den siste påstanden. Leser du gjennom utledningen av ulikheten en gang til, vil du se at vi har likhet når $|\mathbf{a} - \mathbf{p}| = 0$, dvs. når $\mathbf{a} = \mathbf{p}$. Siden \mathbf{p} er parallell med \mathbf{b} , skjer dette når \mathbf{a} og \mathbf{b} er parallelle. \square

Du husker sikkert fra skolematematikken at


$$\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}||\mathbf{b}| \cos v$$

der v er vinkelen mellom vektorene \mathbf{a} og \mathbf{b} . I utgangspunktet gir denne formelen bare mening når \mathbf{a} og \mathbf{b} er vektorer i planet eller rommet — for generelle n -tupler vet vi jo ikke hva vinkler er. Ved hjelp av Schwarz' ulikhet kan vi nå snu situasjonen på hodet; vi *definerer* rett og slett vinkelen mellom to ikke-null n -tupler \mathbf{a} og \mathbf{b} til å være den vinkelen v mellom 0° og 180° som er slik at $\cos v = \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}||\mathbf{b}|}$. Legg merke til at siden Schwarz' ulikhet garanterer at $-1 \leq \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}||\mathbf{b}|} \leq 1$, så finnes det alltid en slik vinkel v som definisjonen forutsetter. Vi ser også at vi får $\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}||\mathbf{b}| \cos v$.

Hva er så vitsen med et slikt abstrakt og merkelig vinkelbegrep? Kan disse vinklene brukes til noe, og oppfører de seg som de vinklene vi er vant til fra planet og rommet? Dette er fornuftige spørsmål som bare erfaring kan gi svar på. Erfaringen viser at disse vinklene fungerer utmerket, og at de i det store og hele har de samme egenskapene som vinkler i 2 og 3 dimensjoner. Vi skal ikke komme nærmere inn på dette her, men tar med et eksempel på hvordan man finner en vinkel:

Eksempel 1: Finn vinkelen mellom vektorene $\mathbf{a} = (2, -1, 0, 1, 1)$ og $\mathbf{b} = (0, 1, 3, -2, 0)$. Vi har

$$\begin{aligned} \cos v &= \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}||\mathbf{b}|} = \\ &= \frac{(2, -1, 0, 1, 1) \cdot (0, 1, 3, -2, 0)}{|(2, -1, 0, 1, 1)|| (0, 1, 3, -2, 0)|} = \frac{-3}{\sqrt{7}\sqrt{14}} = -\frac{3\sqrt{2}}{14} \end{aligned}$$

Bruker vi en lommeregner, finner vi at $-\frac{3\sqrt{2}}{14} \approx -0.3030$. Dette gir $v \approx \arccos(-0.3030) \approx 107.6^\circ$. 

Et viktig resultat i planet og rommet er trekantulikheten som sier at $|\mathbf{a} + \mathbf{b}| \leq |\mathbf{a}| + |\mathbf{b}|$. Ved hjelp av Schwarz' ulikhet skal vi nå vise at trekantulikheten også gjelder i n dimensjoner.

Setning 1.2.4 (Trekantulikheten) *For alle $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ gjelder*

$$|\mathbf{a} + \mathbf{b}| \leq |\mathbf{a}| + |\mathbf{b}|$$

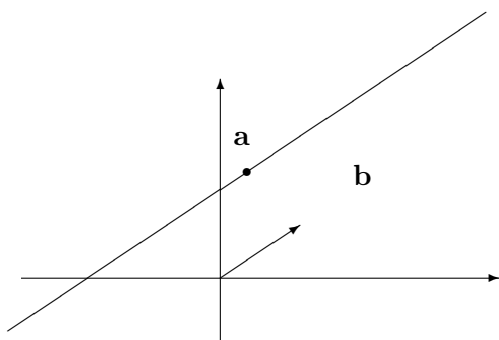
Bevis: Vi har

$$\begin{aligned} |\mathbf{a} + \mathbf{b}|^2 &= (\mathbf{a} + \mathbf{b}) \cdot (\mathbf{a} + \mathbf{b}) \\ &= |\mathbf{a}|^2 + 2 \mathbf{a} \cdot \mathbf{b} + |\mathbf{b}|^2 \leq |\mathbf{a}|^2 + 2 |\mathbf{a}| |\mathbf{b}| + |\mathbf{b}|^2 = (|\mathbf{a}| + |\mathbf{b}|)^2 \end{aligned}$$

der vi har brukt at ifølge Schwarz' ulikhet er $\mathbf{a} \cdot \mathbf{b} \leq |\mathbf{a}| |\mathbf{b}|$. \square

Geometrisk sier trekantulikheten at lengden til den ene siden i en trekant alltid er mindre enn summen av de to andre sidene. Resultatet ovenfor forteller oss at dette også gjelder i høyere dimensjoner. Faktisk spiller trekantulikheten en nøkkelrolle i de fleste forsøk på å generalisere avstandsbegrepet til nye sammenhenger. I dette heftet skal vi ha stor glede av trekantulikheten når vi studerer funksjoner av flere variable.

La oss avslutte denne seksjonen med å se på hvordan vi kan generalisere begrepet *linje* til \mathbb{R}^n . Vi starter i planet. Figur 7 viser en rett linje gjennom punktet \mathbf{a} parallell med vektoren \mathbf{b} .



Figur 7: Rett linje gjennom \mathbf{a} parallell med \mathbf{b}

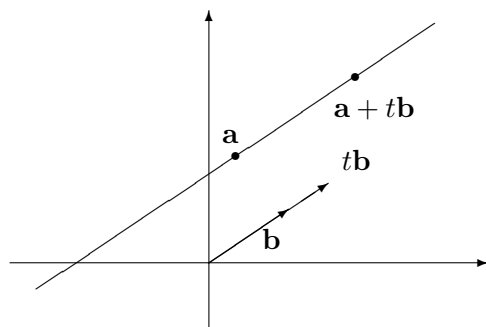
Siden enhver vektor $t\mathbf{b}$ er parallell med \mathbf{b} , ser vi at alle punkter av typen $\mathbf{a} + t\mathbf{b}$ må ligge på linjen (se figur 8). Det er heller ikke så vanskelig å overbevise seg om at ethvert punkt på linjen må være av formen $\mathbf{a} + t\mathbf{b}$ for ett eller annet tall t .

Vi har dermed kommet frem til at de punktene som ligger på den rette linjen, er nøyaktig de som er av typen $\mathbf{a} + t\mathbf{b}$ for et reelt tall t . Vi skriver

gjørne

$$\mathbf{r}(t) = \mathbf{a} + t\mathbf{b}$$

og tenker på $\mathbf{r}(t)$ som et punkt som beveger seg langs linjen når t endrer seg.



Figur 8: Parameterfremstilling av en rett linje

Det er nå lett å generalisere begrepet rett linje til \mathbb{R}^n . Hvis $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$, $\mathbf{b} \neq \mathbf{0}$, så består *den rette linjen gjennom punktet \mathbf{a} og med retningsvektor \mathbf{b}* av alle punkter på formen

$$\mathbf{r}(t) = \mathbf{a} + t\mathbf{b}$$

Bruker vi koordinater, ser vi at hvis $\mathbf{a} = (a_1, a_2, \dots, a_n)$ og $\mathbf{b} = (b_1, b_2, \dots, b_n)$, så blir

$$\mathbf{r}(t) = \mathbf{a} + t\mathbf{b} = (a_1 + tb_1, a_2 + tb_2, \dots, a_n + tb_n)$$

Eksempel 2: Finn en parameterfremstilling til linjen gjennom punktet $\mathbf{a} = (1, 2, 0, -1)$ med retningsvektor $\mathbf{b} = (-1, 2, -3, 1)$, og avgjør om punktet $\mathbf{c} = (2, -1, 1, 4)$ ligger på linjen.

Parameterfremstillingen er

$$\mathbf{r}(t) = \mathbf{a} + t\mathbf{b} = (1 - t, 2 + 2t, -3t, -1 + t)$$

Skal punktet \mathbf{c} ligge på linjen, må det finnes et tall t slik at $\mathbf{r}(t) = \mathbf{c}$, dvs. at følgende ligninger må være oppfylt:

$$1 - t = 2, \quad 2 + 2t = -1, \quad -3t = 1, \quad -1 + t = 4$$

Siden det ikke finnes noe tall t som oppfyller alle disse ligningene, ligger ikke \mathbf{c} på linjen. ♣

La oss avslutte denne seksjonen med et begrep som først vil spille en sentral rolle i senere kapitler, men som det kan være greit å vite om allerede

nå. Anta at vi har vektorer $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ i \mathbb{R}^n . Vi sier at vektoren $\mathbf{v} \in \mathbb{R}^n$ er en *lineærkombinasjon* av $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ dersom det finnes tall s_1, s_2, \dots, s_k slik at

$$\mathbf{v} = s_1 \mathbf{v}_1 + s_2 \mathbf{v}_2 + \dots + s_k \mathbf{v}_k$$

Oppgaver til seksjon 1.2

1. Finn skalarproduktet av $(-2, 3)$ og $(4, 1)$. Finn også vinkelen mellom vektorene.
2. $|\mathbf{a}| = 4$, $|\mathbf{b}| = 5$ og vinkelen mellom \mathbf{a} og \mathbf{b} er 45° . Finn $\mathbf{a} \cdot \mathbf{b}$.
3. Finn vinkelen mellom vektorene $(1, 2, 3)$ og $(-1, 0, 1)$.
4. Regn ut vinkelen mellom $(-1, 2, 6, 2, 4)$ og $(1, 0, 3, 1, 1)$.
5. Finn vinkelen mellom vektorene $\mathbf{a} = (4, 3, 1, 2)$ og $\mathbf{b} = (-1, 3, 2, 0)$. Finn også projeksjonen av \mathbf{a} ned på \mathbf{b} .
6. Hvor lang er projeksjonen av $(-3, 4, 2, 5)$ ned på $(0, 3, 1, 2)$?
7. Skriv $\mathbf{a} = (4, 3)$ som en sum av to vektorer \mathbf{b} og \mathbf{c} der \mathbf{b} er parallell med $\mathbf{d} = (1, 2)$ og \mathbf{c} står normalt på \mathbf{d} .
8. Skriv $\mathbf{a} = (2, 2, 1)$ som en sum av to vektorer \mathbf{b} og \mathbf{c} der \mathbf{b} er parallell med $\mathbf{d} = (1, 0, -1)$ og \mathbf{c} står normalt på \mathbf{d} .
9. Finn vinkelen som hver av vektorene $\mathbf{a} = (\sqrt{3}, 1)$ og $\mathbf{b} = (1, 1)$ danner med x -aksen. Regn ut $\mathbf{a} \cdot \mathbf{b}$ og bruk svaret til å finne et eksakt uttrykk for $\cos(15^\circ)$.
10. Finn to vektorer som begge står normalt på $(3, 2, -1)$ og som ikke er parallelle.
11. Vis at dersom \mathbf{a} står normalt på både \mathbf{b} og \mathbf{c} , så står \mathbf{a} normalt på $\mathbf{b} + \mathbf{c}$.
12. I denne oppgaver er $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$.
 - a) Vis at $|\mathbf{a} + \mathbf{b}|^2 = |\mathbf{a}|^2 + 2 \mathbf{a} \cdot \mathbf{b} + |\mathbf{b}|^2$.
 - b) Finn $\mathbf{a} \cdot \mathbf{b}$ når $|\mathbf{a}| = 6$, $|\mathbf{b}| = 4$ og $|\mathbf{a} + \mathbf{b}| = 3$,
 - c) Anta at $|\mathbf{c}| = 3$, $|\mathbf{d}| = 4$ og $|\mathbf{c} + \mathbf{d}| = 5$. Finn vinkelen mellom \mathbf{c} og \mathbf{d} .
13. Per påstår at han har to vektorer \mathbf{a} og \mathbf{b} slik at $|\mathbf{a}| = 3$, $|\mathbf{b}| = 2$ og $|\mathbf{a} + \mathbf{b}| = 7$. Hvorfor tror du ikke på ham?
14. Kari påstår at hun har to vektorer \mathbf{a} og \mathbf{b} slik at $|\mathbf{a}| = 7$, $|\mathbf{b}| = 2$ og $\mathbf{a} \cdot \mathbf{b} = -16$. Hvorfor tror du ikke på henne?
15. Vis at for alle $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ er $|\mathbf{x}| - |\mathbf{y}| \leq |\mathbf{x} - \mathbf{y}|$. Vis også at $|\mathbf{y}| - |\mathbf{x}| \leq |\mathbf{x} - \mathbf{y}|$, og konkluder med at $||\mathbf{x}| - |\mathbf{y}|| \leq |\mathbf{x} - \mathbf{y}|$.

16. Avstanden $d(\mathbf{a}, \mathbf{b})$ mellom to punkter $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ er lik lengden til vektoren som forbinder dem, dvs. $d(\mathbf{a}, \mathbf{b}) = |\mathbf{b} - \mathbf{a}|$. Bevis at $d(\mathbf{a}, \mathbf{b}) \leq d(\mathbf{a}, \mathbf{c}) + d(\mathbf{c}, \mathbf{b})$ for alle vektorer $\mathbf{a}, \mathbf{b}, \mathbf{c}$. Hva er den geometriske tolkningen av denne ulikheten?

17. Vis at for alle vektorer \mathbf{x} og \mathbf{y} gjelder $|\mathbf{x} + \mathbf{y}|^2 + |\mathbf{x} - \mathbf{y}|^2 = 2|\mathbf{x}|^2 + 2|\mathbf{y}|^2$. Vis at i et parallelogram er summen av kvadratene av sidene lik summen av kvadratene av diagonalene.

18. Finn en parameterfremstilling av linjen som går gjennom punktet $(-1, -1, 2)$ og er parallell med $(2, 3, 1)$.

19. Finn en parameterfremstilling av linjen gjennom $(-3, -2, 5, 8)$ parallell med $(1, -2, -1, 3)$. Sjekk om punktet $(1, -6, 3, 14)$ ligger på linjen.

20. Finn en parameterfremstilling av linjen som går gjennom punktene $(2, -1, 3)$ og $(3, 8, -2)$.

21. Finn en parameterfremstilling av linjen som går gjennom punktene $(7, -3, 2, 4, -2)$ og $(2, 1, -1, -1, 5)$.

22. Finn en parameterfremstilling for linjen som går gjennom $(5, -2)$ og som står normalt på $(-1, 2)$.

23. Finn en parameterfremstilling for linjen i planet som har ligning $2x + 3y = 6$.

24. En linje i planet har parameterfremstilling $(-3 + 2t, 2 - t)$. Finn en ligning av typen $y = ax + b$ for denne linjen.

25. To skip er på kryssende kurs. Ved tiden $t = 0$ er det ene skipet i punktet $(0, 4)$, og det andre skipet i punktet $(39, 14)$ (alle avstander er målt i nautiske mil.) Det første skipet beveger seg parallelt med vektoren $(3, 4)$ med en fart av 15 knop (1 knop = 1 nautisk mil per time). Det andre skipet beveger seg parallelt med vektoren $(-12, 5)$ med en fart av 13 knop.

a) Hvor vil kursene krysse hverandre?

b) Vil skipene kollidere?

26. To fly er i det samme området. Ved tiden $t = 0$ er det ene flyet i punktet $(0, 0, 2000)$ og flyr med en fart på 150m/s parallelt med vektoren $(2, 2, 1)$. Det andre flyet er ved tiden $t = 0$ i punktet $(5000, -1000, 4000)$ og 20 sekunder senere i punktet $(4400, 2000, 4000)$. Flyet følger en rett linje og holder konstant hastighet.

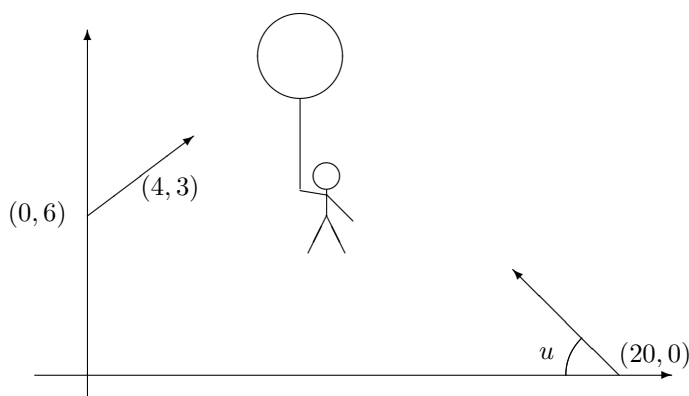
a) Vil kursene til de to flyene skjære hverandre?

b) Vil flyene kollidere?

27. I sin evige jakt etter honning forsøker Ole Brumm å invadere et tre ved hjelp av en ballong. Plutselig blir ballongen tatt av et vindkast og farer av sted med Ole Brumm. Etter å ha tenkt seg om et øyeblikk, innser Kristoffer Robin at hans eneste sjanse til å redde vennen er å skyte istykker ballongen med lekegeværet sitt. Figuren nedenfor viser en skisse av situasjonen.

Når vindkastet kommer ved tiden $t = 0$, befinner ballongen seg i punktet $(0, 6)$. Den blir ført av gårde med en fart av 5 m/s i retningen $(4, 3)$. Ved tiden $t = 2$ skyter Kristoffer Robin mot ballongen fra sin posisjon $(20, 0)$. Vinkelen mellom geværet og underlaget er u , og vi regner med at kulen beveger seg rettlinjet med en fart av 70 m/s . Alle avstander er målt i meter og tiden er målt i sekunder.

- Forklar at ballongens posisjon ved tiden t er $(4t, 6 + 3t)$.
- Vis at kulens posisjon ved tiden t er $(20 - 70(t - 2)\cos u, 70(t - 2)\sin u)$.
- Hvilken vinkel u må Kristoffer Robin holde geværet i for å treffe midt i ballongen? Hvor langt er det ned til bakken når ballongen blir truffet?



28. I denne oppgaven skal vi se på et annet bevis for Schwarz' ulikhet.

- Vis at for alle $a, b \in \mathbb{R}$ og alle $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ er

$$0 \leq |a\mathbf{x} \pm b\mathbf{y}|^2 = a^2|\mathbf{x}|^2 \pm 2ab\mathbf{x} \cdot \mathbf{y} + b^2|\mathbf{y}|^2$$

- Velg $a = |\mathbf{y}|$, $b = |\mathbf{x}|$ i ulikhetene ovenfor og utled Schwarz' ulikhet.

1.3 Komplekse n -tupler

Hittil har vi bare sett på n -tupler $\mathbf{a} = (a_1, a_2, \dots, a_n)$ der komponentene a_1, a_2, \dots, a_n er reelle tall. Vi skal nå ta en rask titt på det komplekse tilfellet. Som nevnt tidligere kalles mengden av alle komplekse n -tupler for \mathbb{C}^n . Addisjon og subtraksjon av komplekse n -tupler foregår komponentvis akkurat som i det reelle tilfellet. Også multiplikasjon med skalar (som nå godt kan være kompleks) foregår akkurat som før. La oss se på et eksempel:

Eksempel 1: Regn ut $s\mathbf{a} + t\mathbf{b}$ når $s = 1 + i$, $t = i$, $\mathbf{a} = \begin{pmatrix} 2 \\ i \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} 1 - i \\ 1 + i \end{pmatrix}$. Vi får (husk at $i^2 = -1$):

$$s\mathbf{a} + t\mathbf{b} = (1 + i) \begin{pmatrix} 2 \\ i \end{pmatrix} + i \begin{pmatrix} 1 - i \\ 1 + i \end{pmatrix} = \begin{pmatrix} (1 + i)2 \\ (1 + i)i \end{pmatrix} + \begin{pmatrix} i(1 - i) \\ i(1 + i) \end{pmatrix} =$$

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



Når vi skal definere *normen* (dvs. *lengden*) til et komplekst n -tupple, må vi være litt forsiktig. Dersom $\mathbf{a} = (z_1, z_2, \dots, z_n)$, setter vi

$$|\mathbf{a}| = \sqrt{|z_1|^2 + |z_2|^2 + \dots + |z_n|^2}$$

Vi tar altså tallverdien til komponentene før vi kvadrerer dem. Grunnen til dette er at kvadratet til et komplekst tall ikke behøver å være positivt og derfor gir et dårlig mål på størrelse.

Eksempel 2: Finn normen til vektoren $\mathbf{a} = (2+i, 4+i, 1-3i)$. Husk at dersom $z = a+ib$, så er $|z| = \sqrt{a^2+b^2}$ og dermed $|z|^2 = a^2+b^2$. Dette gir

$$\begin{aligned} |\mathbf{a}| &= \sqrt{|2+i|^2 + |4+i|^2 + |1-3i|^2} = \\ &= \sqrt{(2^2+1^2) + (4^2+1^2) + (1^2+(-3)^2)} = \\ &= \sqrt{4+1+16+1+1+9} = \sqrt{32} = 4\sqrt{2} \end{aligned}$$

Lengden til vektoren er altså $4\sqrt{2}$.



For å få til det riktige samspillet mellom normen og skalarproduktet, må vi også gjøre en liten justering i definisjonen av skalarprodukt. Dersom $\mathbf{a} = (z_1, z_2, \dots, z_n)$ og $\mathbf{b} = (w_1, w_2, \dots, w_n)$, definerer vi

$$\mathbf{a} \cdot \mathbf{b} = z_1 \overline{w_1} + z_2 \overline{w_2} + \dots + z_n \overline{w_n}$$

Vi komplekskonjugerer altså den andre faktoren i skalarproduktet. Siden $z\overline{z} = |z|^2$, ser vi at

$$\mathbf{a} \cdot \mathbf{a} = z_1 \overline{z_1} + z_2 \overline{z_2} + \dots + z_n \overline{z_n} = |z_1|^2 + |z_2|^2 + \dots + |z_n|^2 = |\mathbf{a}|^2$$

Vi har altså den vanlige sammenhengen mellom norm og skalarprodukt:

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

La oss ta med et eksempel på hvordan man regner ut et komplekst skalarprodukt.

Eksempel 3: Regn ut $\mathbf{a} \cdot \mathbf{b}$ når $\mathbf{a} = (1+i, -2, 1+3i)$ og $\mathbf{b} = (2+2i, 1-2i, 3+4i)$. Vi får

$$\begin{aligned} \mathbf{a} \cdot \mathbf{b} &= (1+i)(2-2i) + (-2)(1+2i) + (1+3i)(3-4i) = \\ &= 2-2i+2i+2-2-4i+3-4i+9i+12 = 17+i \end{aligned}$$

Skalarproduktet av to komplekse vektorer er altså et komplekst tall. ♣

På grunn av komplekskonjugasjonen i annen faktor, har vi ikke lenger at $\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}$ for komplekse vektorer (den *kommutive lov* gjelder altså ikke for komplekse skalarprodukter). Bruker vi regnereglene for konjugasjon (se *Kalkulus*, setning 3.1.5), ser vi imidlertid at

$$\begin{aligned}\overline{\mathbf{b} \cdot \mathbf{a}} &= \overline{w_1 \bar{z}_1 + w_2 \bar{z}_2 + \cdots + w_n \bar{z}_n} = \\ &= \overline{w_1} \bar{\bar{z}_1} + \overline{w_2} \bar{\bar{z}_2} + \cdots + \overline{w_n} \bar{\bar{z}_n} = \\ &= \overline{w_1} z_1 + \overline{w_2} z_2 + \cdots + \overline{w_n} z_n = \\ &= z_1 \overline{w_1} + z_2 \overline{w_2} + \cdots + z_n \overline{w_n} = \mathbf{a} \cdot \mathbf{b}\end{aligned}$$

Bytter vi om på faktorenes rekkefølge, så komplekskonjugerer vi altså resultatet!

Siden vi er så vant til at faktorenes rekkefølge ikke spiller noen rolle, er det lett å bli lurt av det komplekse skalarproduktet. Dersom vi skal regne ut $(\mathbf{a} + \mathbf{b}) \cdot (\mathbf{a} + \mathbf{b})$, er det f.eks. fristende å bruke første kvadratsetning til å skrive svaret

$$\mathbf{a} \cdot \mathbf{a} + 2\mathbf{a} \cdot \mathbf{b} + \mathbf{b} \cdot \mathbf{b},$$

men dette blir ikke riktig! Ganger vi nemlig ut parentesene litt forsiktig, ser vi at

$$(\mathbf{a} + \mathbf{b}) \cdot (\mathbf{a} + \mathbf{b}) = \mathbf{a} \cdot \mathbf{a} + \mathbf{a} \cdot \mathbf{b} + \mathbf{b} \cdot \mathbf{a} + \mathbf{b} \cdot \mathbf{b}$$

og siden $\mathbf{a} \cdot \mathbf{b}$ og $\mathbf{b} \cdot \mathbf{a}$ *ikke* er like, kan vi ikke slå sammen de to midterste leddene på vanlig måte. Siden tallene $\mathbf{a} \cdot \mathbf{b}$ og $\mathbf{b} \cdot \mathbf{a}$ er komplekskonjugerte, har vi imidlertid at $\mathbf{a} \cdot \mathbf{b} + \mathbf{b} \cdot \mathbf{a} = 2\operatorname{Re}(\mathbf{a} \cdot \mathbf{b})$ (der Re står for realdel), og dermed kan vi skrive

$$\begin{aligned}(\mathbf{a} + \mathbf{b}) \cdot (\mathbf{a} + \mathbf{b}) &= \mathbf{a} \cdot \mathbf{a} + 2\operatorname{Re}(\mathbf{a} \cdot \mathbf{b}) + \mathbf{b} \cdot \mathbf{b} = \\ &= |\mathbf{a}|^2 + 2\operatorname{Re}(\mathbf{a} \cdot \mathbf{b}) + |\mathbf{b}|^2\end{aligned}$$

Dette regnestykket viser at vi må være litt forsiktige når vi overfører standard regneprosedyrer til komplekse skalarprodukt.

Vi skriver opp de grunnleggende regnereglene for komplekse n -tupler (sammenlign med setning 1.1.1 for reelle n -tupler):

Setning 1.3.1 (Regneregler for komplekse n -tupler.) Dersom \mathbf{a} , \mathbf{b} og \mathbf{c} er n -tupler og s og t er komplekse tall, gjelder følgende regneregler:

- (a) $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$
- (b) $\mathbf{a} \cdot \mathbf{b} = \overline{\mathbf{b} \cdot \mathbf{a}}$

$$(c) \quad s(\mathbf{a} + \mathbf{b}) = s\mathbf{a} + s\mathbf{b}$$

$$(d) \quad (s + t)\mathbf{a} = s\mathbf{a} + t\mathbf{a}$$

$$(e) \quad \mathbf{c} \cdot (\mathbf{a} + \mathbf{b}) = \mathbf{c} \cdot \mathbf{a} + \mathbf{c} \cdot \mathbf{b} \text{ og } (\mathbf{a} + \mathbf{b}) \cdot \mathbf{c} = \mathbf{a} \cdot \mathbf{c} + \mathbf{b} \cdot \mathbf{c}$$

$$(f) \quad (s\mathbf{a}) \cdot \mathbf{b} = s(\mathbf{a} \cdot \mathbf{b}) \text{ og } \mathbf{a} \cdot (s\mathbf{b}) = \overline{s}(\mathbf{a} \cdot \mathbf{b})$$

$$(g) \quad \mathbf{a} \cdot \mathbf{a} \geq 0 \text{ med likhet hvis og bare hvis } \mathbf{a} = \mathbf{0}$$

Bevis: Med unntak av b) (som vi nettopp har bevist) og andre del av f), er dette nøyaktig de samme reglene som i setning 1.1.1, og bevisene er også de samme. Vi nøyer oss derfor med å vise andre del av f). Hvis $\mathbf{a} = (z_1, z_2, \dots, z_n)$ og $\mathbf{b} = (w_1, w_2, \dots, w_n)$, ser vi at $s\mathbf{b} = (sw_1, sw_2, \dots, sw_n)$, og dermed er

$$\begin{aligned} \mathbf{a} \cdot (s\mathbf{b}) &= z_1 \overline{sw_1} + z_2 \overline{sw_2} + \dots + z_n \overline{sw_n} = \\ &= \overline{s}(z_1 \overline{w_1} + z_2 \overline{w_2} + \dots + z_n \overline{w_n}) = \overline{s}(\mathbf{a} \cdot \mathbf{b}) \end{aligned}$$

□

Vi kan innføre geometriske begreper for komplekse n -tupler akkurat som for reelle selv om visualiseringen blir enda vanskeligere i dette tilfellet. Vi sier f.eks. at to vektorer $\mathbf{a}, \mathbf{b} \in \mathbb{C}^n$ er *ortogonale* dersom $\mathbf{a} \cdot \mathbf{b} = 0$. Legg merke til at dette medfører at $\mathbf{b} \cdot \mathbf{a} = 0$ siden $\mathbf{b} \cdot \mathbf{a} = \overline{\mathbf{a} \cdot \mathbf{b}} = \overline{0} = 0$. Argumentkjeden som ga oss Pythagoras' setning, Schwarz' ulikhet og trekantulikheten i forrige seksjon, fungerer med små justeringer også i det komplekse tilfellet, og vi nøyer oss med å skrive opp resultatet:

Setning 1.3.2 *For komplekse vektorer \mathbf{a}, \mathbf{b} gjelder:*

(i) (Pythagoras' setning) *Dersom $\mathbf{a}, \mathbf{b} \in \mathbb{C}^n$ er ortogonale, så er*

$$|\mathbf{a}|^2 + |\mathbf{b}|^2 = |\mathbf{a} + \mathbf{b}|^2$$

(ii) (Schwarz' ulikhet) *For alle $\mathbf{a}, \mathbf{b} \in \mathbb{C}^n$ er*

$$|\mathbf{a} \cdot \mathbf{b}| \leq |\mathbf{a}||\mathbf{b}|$$

(iii) (Trekantulikheten) *For alle $\mathbf{a}, \mathbf{b} \in \mathbb{C}^n$ er*

$$|\mathbf{a} + \mathbf{b}| \leq |\mathbf{a}| + |\mathbf{b}|$$

□

Oppgaver til seksjon 1.3

1. Regn ut $s\mathbf{x} + t\mathbf{y}$ når $s = i$, $t = 1 + 2i$, $\mathbf{x} = \begin{pmatrix} -4i \\ 2 - i \end{pmatrix}$ og $\mathbf{y} = \begin{pmatrix} 2 + i \\ 2i \end{pmatrix}$.
2. Finn lengden til vektorene $\mathbf{a} = (3 + 2i, -1 + i)$ og $\mathbf{b} = (i, 2 + 3i, -2 - i)$.
3. Regn ut skalarproduktet $\mathbf{x} \cdot \mathbf{y}$ når $\mathbf{x} = (1 + 3i, -2i, 2 + 3i)$ og $\mathbf{y} = (2, 1 + 2i, -1 + i)$.
4. Vis at for alle $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$ er

$$|\mathbf{x} - \mathbf{y}|^2 = |\mathbf{x}|^2 - 2 \operatorname{Re}(\mathbf{x} \cdot \mathbf{y}) + |\mathbf{y}|^2$$

og

$$(\mathbf{x} + \mathbf{y}) \cdot (\mathbf{x} - \mathbf{y}) = |\mathbf{x}|^2 - 2 \operatorname{Im}(\mathbf{x} \cdot \mathbf{y}) - |\mathbf{y}|^2$$

hvor $\operatorname{Re}(z)$ og $\operatorname{Im}(z)$ betegner hhv. realdelen og imaginærdelen til z .

5. Bevis setning 1.3.2 ved å gå gjennom beviset for de tilsvarende resultatene i seksjon 1.2 og se hvilke modifikasjoner som må gjøres.

1.4 Vektorproduktet

De regneoperasjonene vi hittil har sett på, er definert for vektorer av alle dimensjoner. I denne seksjonen skal vi se studere en operasjon — *vektorproduktet* — som bare er definert for tredimensjonale vektorer. Siden 3 er den fysiske romdimensjonen, brukes vektorproduktet ofte i geometriske problemstillinger. Det brukes også mye i fysikk og mekanikk — tok du fysikk i videregående skole, har du sikkert støtt på “høyrehåndsregler” i en del sammenhenger, og bak enhver slik høyrehåndsregel skjuler det seg et vektorprodukt.

Det er to måter å definere vektorproduktet på, en geometrisk og en algebraisk, og det er samspillet mellom disse to betraktningsmåtene som gir vektorproduktet slagkraft. Vi skal ta utgangspunkt i den algebraiske definisjonen. Før vi begynner, minner vi om at man i tre dimensjoner gjerne skriver enhetsvektorene langs aksene på denne måten:

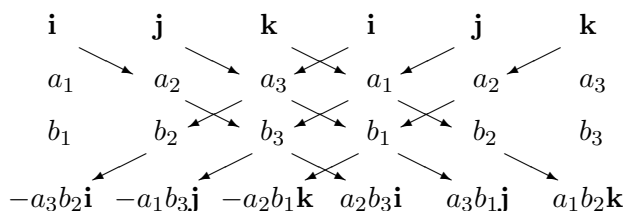
$$\mathbf{i} = (1, 0, 0), \quad \mathbf{j} = (0, 1, 0), \quad \mathbf{k} = (0, 0, 1).$$

Gitt to vektorer $\mathbf{a} = (a_1, a_2, a_3)$ og $\mathbf{b} = (b_1, b_2, b_3)$ i \mathbb{R}^3 definerer vi nå *vektorproduktet* (også kalt *kryssproduktet*) $\mathbf{a} \times \mathbf{b}$ ved:

$$\begin{aligned} \mathbf{a} \times \mathbf{b} &= (a_2b_3 - a_3b_2, a_3b_1 - a_1b_3, a_1b_2 - a_2b_1) \\ &= (a_2b_3 - a_3b_2)\mathbf{i} + (a_3b_1 - a_1b_3)\mathbf{j} + (a_1b_2 - a_2b_1)\mathbf{k} \end{aligned}$$

Denne formelen kan være vanskelig å huske, men det finnes huskereglene. Én slik regel er vist i skjemaet nedenfor. Vi multipliserer langs pilene og gir

resultatet positiv verdi dersom pilene går fra venstre mot høyre og negativ verdi dersom de går fra høyre mot venstre (kjenner du en annen huskeregel fra før, kan du trygt bruke den).



Figur 1: Huskeregel for vektorproduktet

La oss regne ut et vektorprodukt:

Eksempel 1: Finn vektorproduktet av $\mathbf{a} = (3, -1, 2)$ og $\mathbf{b} = (4, -2, 5)$. Vi får:

$$\begin{aligned}\mathbf{a} \times \mathbf{b} &= ((-1) \cdot 5 - 2 \cdot (-2))\mathbf{i} + (2 \cdot 4 - 3 \cdot 5)\mathbf{j} + (3 \cdot (-2) - (-1) \cdot 4)\mathbf{k} \\ &= (-1, -7, -2)\end{aligned}$$



La oss se hva som skjer dersom vi regner ut $\mathbf{b} \times \mathbf{a}$ istedenfor $\mathbf{a} \times \mathbf{b}$:

$$\begin{aligned}\mathbf{b} \times \mathbf{a} &= (b_2a_3 - b_3a_2, b_3a_1 - b_1a_3, b_1a_2 - b_2a_1) \\ &= -(a_2b_3 - a_3b_2, a_3b_1 - a_1b_3, a_1b_2 - a_2b_1) = -(\mathbf{a} \times \mathbf{b})\end{aligned}$$

Akkurat som skalarproduktet for komplekse vektorer er altså vektorproduktet ikke-kommutativt, men vi har en formel som gjør at vi lett kan regne ut $\mathbf{b} \times \mathbf{a}$ når vi kjenner $\mathbf{a} \times \mathbf{b}$.

Her er en liste over de grunnleggende egenskapene til vektorproduktet:

Setning 1.4.1 For vektorer $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^3$ gjelder:

- (a) $\mathbf{a} \times \mathbf{b} = -(\mathbf{b} \times \mathbf{a})$
- (b) $\mathbf{a} \times (\mathbf{b} + \mathbf{c}) = \mathbf{a} \times \mathbf{b} + \mathbf{a} \times \mathbf{c}$ og $(\mathbf{a} + \mathbf{b}) \times \mathbf{c} = \mathbf{a} \times \mathbf{c} + \mathbf{b} \times \mathbf{c}$
- (c) $\mathbf{a} \times (s\mathbf{b}) = s(\mathbf{a} \times \mathbf{b})$ og $(s\mathbf{a}) \times \mathbf{b} = s(\mathbf{a} \times \mathbf{b})$ der $s \in \mathbf{R}$
- (d) $\mathbf{a} \times \mathbf{b}$ står ortogonalt på både \mathbf{a} og \mathbf{b}
- (e) (Lagranges identitet) $|\mathbf{a} \times \mathbf{b}|^2 = |\mathbf{a}|^2|\mathbf{b}|^2 - (\mathbf{a} \cdot \mathbf{b})^2$

Bevis: Punkt a) har vi allerede bevist og de andre bevisene er av samme type — vi skriver vektorene på koordinatform, regner ut og ser at det stemmer.

Vi tar c), d) og e) som eksempler:

c) Hvis $\mathbf{a} = (a_1, a_2, a_3)$ og $\mathbf{b} = (b_1, b_2, b_3)$, så er $s\mathbf{a} = (sa_1, sa_2, sa_3)$. Vi får:

$$\begin{aligned} (s\mathbf{a}) \times \mathbf{b} &= ((sa_2)b_3 - (sa_3)b_2, (sa_3)b_1 - (sa_1)b_3, (sa_1)b_2 - (sa_2)b_1) \\ &= s(a_2b_3 - a_3b_2, a_3b_1 - a_1b_3, a_1b_2 - a_2b_1) = s(\mathbf{a} \times \mathbf{b}) \end{aligned}$$

Den andre likheten i c) går på samme måte.

d) For å vise at \mathbf{a} står ortogonalt på $\mathbf{a} \times \mathbf{b}$, må vi vise at $\mathbf{a} \cdot (\mathbf{a} \times \mathbf{b}) = \mathbf{0}$. Vi får:

$$\begin{aligned} \mathbf{a} \cdot (\mathbf{a} \times \mathbf{b}) &= a_1(a_2b_3 - a_3b_2) + a_2(a_3b_1 - a_1b_3) + a_3(a_1b_2 - a_2b_1) \\ &= a_1a_2b_3 - a_1a_3b_2 + a_2a_3b_1 - a_2a_1b_3 + a_3a_1b_2 - a_3a_2b_1 = 0 \end{aligned}$$

En helt tilsvarende regning viser at \mathbf{b} står ortogonalt på $\mathbf{a} \times \mathbf{b}$.

e) Vi skriver $\mathbf{a} = (a_1, a_2, a_3)$, $\mathbf{b} = (b_1, b_2, b_3)$ og regner ut begge sider (du er ikke forpliktet til å føle at dette er spesielt festlig):

$$\begin{aligned} |(\mathbf{a} \times \mathbf{b})|^2 &= (a_2b_3 - a_3b_2)^2 + (a_3b_1 - a_1b_3)^2 + (a_1b_2 - a_2b_1)^2 \\ &= a_2^2b_3^2 - 2a_2a_3b_2b_3 + a_3^2b_2^2 + a_3^2b_1^2 - 2a_1a_3b_1b_3 + a_1^2b_3^2 + a_1^2b_2^2 - 2a_1a_2b_1b_2 + a_2^2b_1^2 \end{aligned}$$

og

$$\begin{aligned} |\mathbf{a}|^2|\mathbf{b}|^2 - (\mathbf{a} \cdot \mathbf{b})^2 &= (a_1^2 + a_2^2 + a_3^2)(b_1^2 + b_2^2 + b_3^2) - (a_1b_1 + a_2b_2 + a_3b_3)^2 \\ &= a_1^2b_1^2 + a_1^2b_2^2 + a_1^2b_3^2 + a_2^2b_1^2 + a_2^2b_2^2 + a_2^2b_3^2 + a_3^2b_1^2 + a_3^2b_2^2 + a_3^2b_3^2 \\ &\quad - a_1^2b_1^2 - a_2^2b_2^2 - a_3^2b_3^2 - 2a_1a_2b_1b_2 - 2a_1a_3b_1b_3 - 2a_2a_3b_2b_3 \\ &= a_1^2b_2^2 + a_1^2b_3^2 + a_2^2b_1^2 + a_2^2b_3^2 + a_3^2b_1^2 + a_3^2b_2^2 - 2a_1a_2b_1b_2 - 2a_1a_3b_1b_3 - 2a_2a_3b_2b_3 \end{aligned}$$

Bortsett fra rekkefølgen på leddene er dette det samme uttrykket som vi fikk ovenfor. Dermed er e) bevist. \square

Bemerkning: Legg merke til at det ikke finnes noen assosiativ lov i listen ovenfor — generelt er nemlig $(\mathbf{a} \times \mathbf{b}) \times \mathbf{c} \neq \mathbf{a} \times (\mathbf{b} \times \mathbf{c})$. Som et eksempel lar vi $\mathbf{a} = (1, 1, 0)$, $\mathbf{b} = (1, 0, 0)$ og $\mathbf{c} = (0, 0, 1)$. Da er

$$(\mathbf{a} \times \mathbf{b}) \times \mathbf{c} = ((1, 1, 0) \times (1, 0, 0)) \times (0, 0, 1) = (0, 0, -1) \times (0, 0, 1) = \mathbf{0}$$

mens

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (1, 1, 0) \times ((1, 0, 0) \times (0, 0, 1)) = (1, 1, 0) \times (0, -1, 0) = (0, 0, -1)$$

At $(\mathbf{a} \times \mathbf{b}) \times \mathbf{c} \neq \mathbf{a} \times (\mathbf{b} \times \mathbf{c})$ betyr at uttrykket $\mathbf{a} \times \mathbf{b} \times \mathbf{c}$ ikke gir noen mening — vi må ha med parenteser for å presisere hvilken rekkefølge produktene skal utføres i.

Som allerede nevnt, er det også en geometrisk måte å beskrive vektorproduktet på. For å finne frem til denne geometriske beskrivelsen, tar vi utgangspunkt i punkt e) i setningen ovenfor:

$$|\mathbf{a} \times \mathbf{b}|^2 = |\mathbf{a}|^2 |\mathbf{b}|^2 - (\mathbf{a} \cdot \mathbf{b})^2$$

Siden vi allerede vet at $\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}| |\mathbf{b}| \cos v$, der v er vinkelen mellom \mathbf{a} og \mathbf{b} , så er

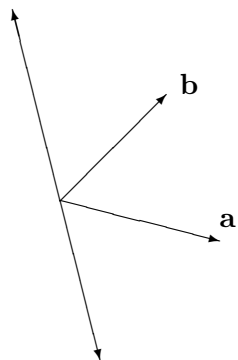
$$|\mathbf{a} \times \mathbf{b}|^2 = |\mathbf{a}|^2 |\mathbf{b}|^2 - |\mathbf{a}|^2 |\mathbf{b}|^2 \cos^2 v = |\mathbf{a}|^2 |\mathbf{b}|^2 \sin^2 v$$

der vi har benyttet at $1 - \cos^2 v = \sin^2 v$. Altså er

$$|\mathbf{a} \times \mathbf{b}| = |\mathbf{a}| |\mathbf{b}| \sin v$$

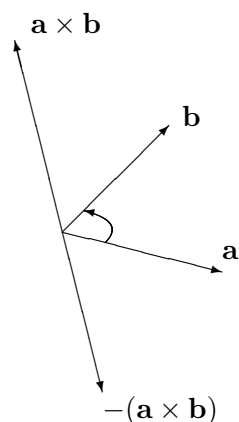
(husk at siden $0^\circ \leq v \leq 180^\circ$, er $\sin v$ aldri negativ). Dermed vet vi hvor lang vektoren $\mathbf{a} \times \mathbf{b}$ er. Legg spesielt merke til at $\mathbf{a} \times \mathbf{b} = \mathbf{0}$ hvis og bare hvis $\sin v = 0$, dvs. dersom \mathbf{a} og \mathbf{b} er parallelle.

Fra punkt d) i setning 1.4.1 vet vi også noe om retningen til $\mathbf{a} \times \mathbf{b}$, nemlig at $\mathbf{a} \times \mathbf{b}$ står normalt på både \mathbf{a} og \mathbf{b} . Nå finnes det to motsatt rettede vektorer som har lengde $|\mathbf{a}| |\mathbf{b}| \sin v$ og står normalt på både \mathbf{a} og \mathbf{b} (se figur 2).



Figur 2: To like lange vektorer som står normalt på både \mathbf{a} og \mathbf{b}

For å vite hvilken av disse to vektorene som er $\mathbf{a} \times \mathbf{b}$, bruker vi *høyrehåndsregelen*: Vi legger høyre hånd med fingrene pekende den korteste veien fra \mathbf{a} til \mathbf{b} mens vi spriker med tommelen. Da er $\mathbf{a} \times \mathbf{b}$ den av de to normalvektorene som peker i tommelens retning (se figur 3 der den krumme pila viser den retningen fingrene peker). Legg merke til at når vi regner ut $\mathbf{b} \times \mathbf{a}$, skal fingrene spenne over den samme vinkelen, men i motsatt retning (fra \mathbf{b} mot \mathbf{a}), og tommelen kommer derfor til å peke motsatt vei. Dette er den geometriske forklaringen på regelen $\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a}$.

Figur 3: Vektorene $\mathbf{a} \times \mathbf{b}$ og $-(\mathbf{a} \times \mathbf{b})$

La oss oppsummere det vi har kommet frem til:

Setning 1.4.2 *La \mathbf{a} og \mathbf{b} være to vektorer i \mathbb{R}^3 og kall vinkelen mellom dem v . Da har vektorproduktet $\mathbf{a} \times \mathbf{b}$ lengde $|\mathbf{a}||\mathbf{b}|\sin v$ og står normalt på både \mathbf{a} og \mathbf{b} . Retningen til $\mathbf{a} \times \mathbf{b}$ er gitt ved høyrehåndsregelen.*

**Bevis:* Vi har bevist alt bortsett fra høyrehåndsregelen. Beviset vi skal gi for denne regelen kan se litt umatematisk og skissemessig (og vanskelig!) ut, men det kan uten store endringer bygges ut til et fullverdig bevis. Du kan godt hoppe over dette beviset uten å få problemer med det som kommer senere.

Vi skal først bevise høyrehåndsregelen for det spesialtilfellet der \mathbf{a} og \mathbf{b} ligger i xy -planet, og \mathbf{a} peker langs den positive x -aksen. Det betyr at \mathbf{a} har koordinater $\mathbf{a} = (a_1, 0, 0)$ der $a_1 > 0$, og \mathbf{b} har koordinater $\mathbf{b} = (b_1, b_2, 0)$. I dette tilfellet vil $\mathbf{a} \times \mathbf{b}$ være parallell med z -aksen, og vi må undersøke når den peker i positiv retning. Bruker vi høyrehåndsregelen, får vi at $\mathbf{a} \times \mathbf{b}$ skal peke langs den positive z -aksen dersom \mathbf{b} ligger i første eller annen kvadrant av xy -planet (dvs. når $b_2 > 0$), og langs den negative z -aksen dersom \mathbf{b} ligger i tredje eller fjerde kvadrant (dvs. når $b_2 < 0$). Bruker vi isteden formelen for vektorproduktet, ser vi at

$$\mathbf{a} \times \mathbf{b} = 0\mathbf{i} + 0\mathbf{j} + a_1b_2\mathbf{k} = a_1b_2\mathbf{k}$$

Denne vektoren peker langs den positive eller negative z -aksen ettersom b_2 er positiv eller negativ (husk at $a_1 > 0$). Dette betyr at formelen og høyrehåndsregelen gir samme resultat, og dermed er høyrehåndsregelen bevist i dette tilfellet.

Vi er nå rede til å se på det generelle tilfellet $\mathbf{a} = (a_1, a_2, a_3)$, $\mathbf{b} = (b_1, b_2, b_3)$. Velg et par av vektorer \mathbf{a}_0 , \mathbf{b}_0 i xy -planet slik at \mathbf{a} peker langs den positive x -aksen, og slik at \mathbf{a}_0 , \mathbf{b}_0 er en "kopi" av paret \mathbf{a} , \mathbf{b} . Med dette

mener vi at \mathbf{a}_0 er like lang som \mathbf{a} , \mathbf{b}_0 er like lang som \mathbf{b} , og at vinkelen fra \mathbf{a}_0 til \mathbf{b}_0 er lik vinkelen fra \mathbf{a} til \mathbf{b} . La $\mathbf{c}_0 = \mathbf{a}_0 \times \mathbf{b}_0$. Etter det vi allerede har vist, gjelder høyrehåndsregelen for \mathbf{a}_0 , \mathbf{b}_0 og \mathbf{c}_0 .

La oss nå tenke på vektortriplet $(\mathbf{a}_0, \mathbf{b}_0, \mathbf{c}_0)$ som en materiell gjenstand, f.eks. tre sammensveisede biter av ståltråd. Vi flytter nå denne gjenstanden med en kontinuerlig bevegelse, uten å deformere den på noen måte, slik at \mathbf{a}_0 ender opp som \mathbf{a} , og \mathbf{b}_0 ender opp som \mathbf{b} . La $\mathbf{a}(t)$ være posisjonen til \mathbf{a}_0 etter t sekunder av denne bevegelsen, og la $\mathbf{b}(t)$ og $\mathbf{c}(t)$ være de tilsvarende posisjonene til \mathbf{b}_0 og \mathbf{c}_0 . Hvis bevegelsen tar T sekunder, er dermed $\mathbf{a} = \mathbf{a}(T)$ og $\mathbf{b} = \mathbf{b}(T)$.

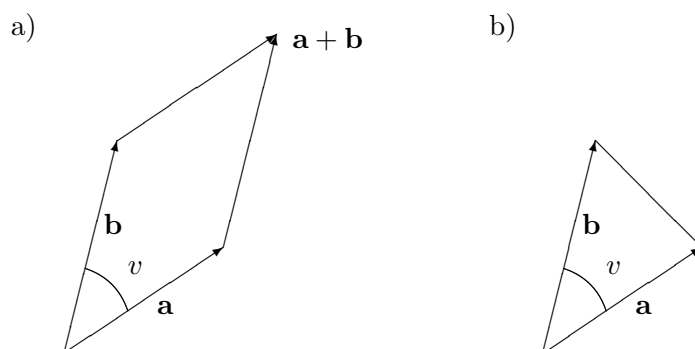
Dersom $\mathbf{c}(T) = \mathbf{a}(T) \times \mathbf{b}(T)$, er høyrehåndsregelen oppfylt for triplet $\mathbf{a}(T) = \mathbf{a}$, $\mathbf{b}(T) = \mathbf{b}$, $\mathbf{c}(T) = \mathbf{c}$. Vi skal derfor anta at $\mathbf{c}(T) = -\mathbf{a}(T) \times \mathbf{b}(T)$ (den eneste andre muligheten) og vise at dette fører til en selvmotsigelse. La t_0 være det første tidspunktet der $\mathbf{c}(t)$ skifter fra å være lik $\mathbf{a}(t) \times \mathbf{b}(t)$ til å være lik $-\mathbf{a}(t) \times \mathbf{b}(t)$ (formelt er $t_0 = \inf\{t : \mathbf{c}(t) = -\mathbf{a}(t) \times \mathbf{b}(t)\}$). Siden $\mathbf{c}(t)$ beveger seg kontinuerlig, betyr dette at $\mathbf{a}(t) \times \mathbf{b}(t)$ må gjøre et sprang ved tidspunktet t_0 . Men det er umulig siden $\mathbf{a}(t) \times \mathbf{b}(t)$ vil bevege seg kontinuerlig når $\mathbf{a}(t)$ og $\mathbf{b}(t)$ gjør det (tenk på det algebraiske uttrykket for vektorproduktet). Dermed har vi fått vår selvmotsigelse, og beviset er fullført. \square

Vi skal nå se på noen av de tingene vektorproduktet kan brukes til. Først et enkelt eksempel.

Eksempel 2: Finn en vektor som står ortogonalt på både $\mathbf{a} = (1, -2, 3)$ og $\mathbf{b} = (4, -1, -2)$. Vi regner rett og slett ut vektorproduktet:

$$\mathbf{a} \times \mathbf{b} = (1, -2, 3) \times (4, -1, -2) = (7, 14, 7)$$

Legg merke til at siden $(7, 14, 7) = 7(1, 2, 1)$, kan vi forenkle løsningen til $(1, 2, 1)$. \clubsuit

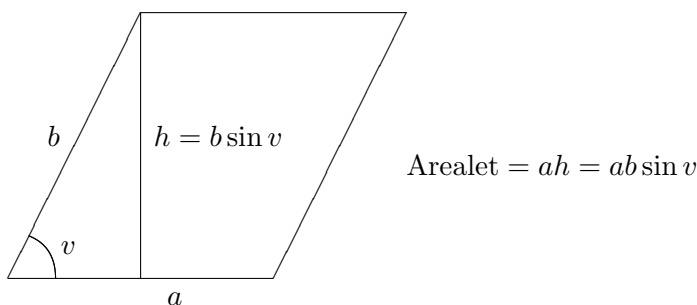


Figur 4: Parallelogrammet og trekanten utspent av \mathbf{a} og \mathbf{b}

Det neste vi skal se på, er hvordan vektorproduktet kan brukes til å regne ut arealer. To vektorer \mathbf{a} og \mathbf{b} utspenner på en naturlig måte et parallelogram (se figur 4a). Halvparten av dette parallelogrammet (se figur 4b), utgjør trekanten utspent av \mathbf{a} og \mathbf{b} .

Setning 1.4.3 *Arealet til parallelogrammet utspent av vektorene \mathbf{a} og \mathbf{b} er lik $|\mathbf{a} \times \mathbf{b}|$. Arealet til trekanten utspent av \mathbf{a} og \mathbf{b} er $\frac{1}{2}|\mathbf{a} \times \mathbf{b}|$*

Bevis: Enkel geometri forteller oss at arealet til et parallelogram er produktet av de to sidene ganget med sinus til den mellomliggende vinkelen (se figur 5). For vårt parallelogram blir dette $|\mathbf{a}||\mathbf{b}|\sin v$, som vi vet er lik $|\mathbf{a} \times \mathbf{b}|$. Siden arealet av trekanten er halvparten av arealet til parallelogrammet, får vi også formelen for arealet av trekanten. \square



Figur 5: Arealet til et parallelogram

Eksempel 3: Finn arealet til trekanten med hjørner i punktene $\mathbf{a} = (2, -7, 3)$, $\mathbf{b} = (-2, 3, 2)$ og $\mathbf{c} = (2, 2, 2)$. Denne trekanten har samme areal som den utspent av vektorene $\mathbf{c} - \mathbf{a}$ og $\mathbf{b} - \mathbf{a}$ (hvorfor?) Siden $\mathbf{c} - \mathbf{a} = (0, 9, -1)$ og $\mathbf{b} - \mathbf{a} = (-4, 10, -1)$, får vi:

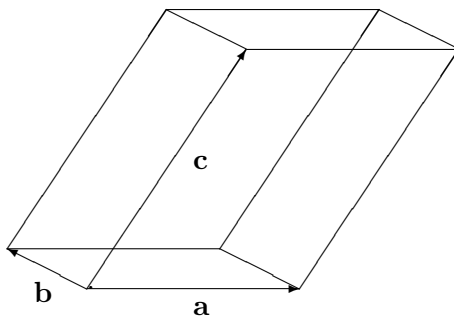
$$(\mathbf{c} - \mathbf{a}) \times (\mathbf{b} - \mathbf{a}) = (0, 9, -1) \times (-4, 10, -1) = (1, 4, 36)$$

Dermed er arealet til trekanten lik

$$\frac{1}{2}|(1, 4, 36)| = \frac{1}{2}\sqrt{1^2 + 4^2 + 36^2} = \frac{1}{2}\sqrt{1313} \quad \clubsuit$$

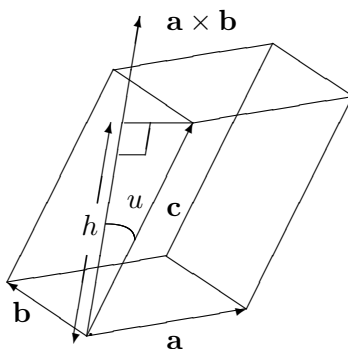
Eksemplet ovenfor viser noe av styrken ved å bruke vektorregning til å regne ut geometriske størrelser; det spiller ikke noen rolle hvor komplisert punktene ligger i forhold til hverandre — vi bare kopler inn den generelle formelen og ut faller svaret. Hadde vi prøvd å finne arealet med tradisjonelle geometriske metoder, hadde vi fort druknet i finurlige tegninger og kompliserte beregninger. Ulempen ved å bruke vektorregning er at vi ofte mister kontakten med det geometriske bildet — regningen viser oss at noe er riktig, men vi skjønner ikke riktig hvorfor.

Vektorproduktet kan også brukes til å regne ut volumer. Tre vektorer \mathbf{a} , \mathbf{b} , \mathbf{c} i rommet definerer på en naturlig måte et romlegeme, et *parallelepiped*, som vist på figur 6.



Figur 6: Parallelepipedet utspent av \mathbf{a} , \mathbf{b} og \mathbf{c}

Fra skolen vet vi at volumet til et parallelepiped er arealet av grunnflaten ganget med høyden. Sier vi at grunnflaten er parallelogrammet utspent av \mathbf{a} og \mathbf{b} , vet vi at arealet til grunnflaten er $|\mathbf{a} \times \mathbf{b}|$. På figur 7 har vi kalt vinkelen mellom $\mathbf{a} \times \mathbf{b}$ og den tredje vektoren \mathbf{c} for u .



Figur 7: Volumet til et parallelepiped

Siden $\mathbf{a} \times \mathbf{b}$ står normalt på grunnflaten, blir høyden h lik $|\mathbf{c}| \cos u$ (vi må ha med tallverdien rundt cosinus i tilfelle u er større en 90°). Volumet til parallelepipedet er derfor $|\mathbf{a} \times \mathbf{b}| |\mathbf{c}| \cos u$. Men dette uttrykket er jo lik $|(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}|$ (husk den geometriske beskrivelsen av skalarproduktet). Dermed har vi vist:

Setning 1.4.4 *Volumet til parallelepipedet utspent av vektorene \mathbf{a} , \mathbf{b} , \mathbf{c} er $|(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}|$*

Bemerkning: Når vi skal regne ut volumet til parallelepipedet utspent av \mathbf{a} , \mathbf{b} og \mathbf{c} , spiller selvfølgelig ikke rekkefølgen av de tre vektorene noen rolle.

Volumet kan derfor skrives som både $|(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}|$, $|(\mathbf{a} \times \mathbf{c}) \cdot \mathbf{b}|$, $|(\mathbf{b} \times \mathbf{a}) \cdot \mathbf{c}|$, $|(\mathbf{b} \times \mathbf{c}) \cdot \mathbf{a}|$, $|(\mathbf{c} \times \mathbf{a}) \cdot \mathbf{b}|$, og $|(\mathbf{c} \times \mathbf{b}) \cdot \mathbf{a}|$. Disse seks uttrykkene må derfor være like. Hvis du orker, kan du sjekke dette ved direkte utregning.

Eksempel 4: Finn volumet til parallellepipedet som utspennes av vektorene $(4, 0, 3)$, $(-1, 2, -3)$ og $(0, 2, 1)$. Ifølge setningen er dette volumet gitt ved $|((4, 0, 3) \times (-1, 2, -3)) \cdot (0, 2, 1)|$. Vi regner først ut

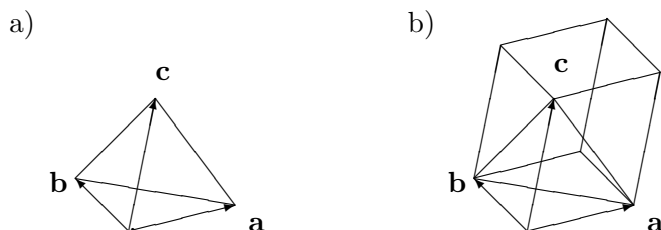
$$(4, 0, 3) \times (-1, 2, -3) = (-6, 9, 8)$$

Deretter tar vi

$$|(-6, 9, 8) \cdot (0, 2, 1)| = |0 + 18 + 8| = 26$$

Volumet er altså 26. ♣

Tre vektorer \mathbf{a} , \mathbf{b} , \mathbf{c} utspenner også en pyramide (se figur 8a). For å finne volumet til denne pyramiden husker vi at volumet til en generell pyramide er $\frac{1}{3}gh$, der h er høyden og g er arealet til grunnflaten.



Figur 8: Pyramiden og parallellepipedet utspent av \mathbf{a} , \mathbf{b} og \mathbf{c}

Sammenligner vi pyramiden utspent av \mathbf{a} , \mathbf{b} , \mathbf{c} med parallellepipedet utspent av de samme vektorene (figur 8b), ser vi at høydene er like, men at grunnflaten til pyramiden er halvparten av grunnflaten til parallellepipedet. Det betyr at volumet til pyramiden må være en seksdel av volumet til parallellepipedet. Dermed har vi:

Korollar 1.4.5 *Volumet av pyramiden utspent av de tre vektorene \mathbf{a} , \mathbf{b} , \mathbf{c} er $\frac{1}{6}|(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}|$*

Eksempel 5: Finn volumet til pyramiden med hjørner i punktene $\mathbf{a} = (-1, 2, -3)$, $\mathbf{b} = (1, 4, 1)$, $\mathbf{c} = (0, 4, 7)$ og $\mathbf{d} = (3, 0, 5)$. Denne pyramiden har samme volum som pyramiden utspent av vektorene $\mathbf{b} - \mathbf{a}$, $\mathbf{c} - \mathbf{a}$ og $\mathbf{d} - \mathbf{a}$ (forklar hvorfor!) Siden

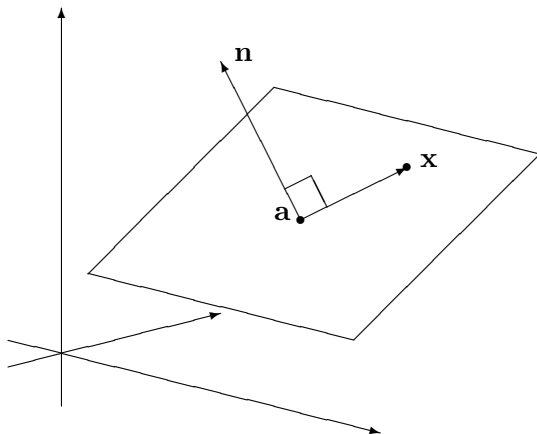
$$\mathbf{b} - \mathbf{a} = (2, 2, 4) \quad \mathbf{c} - \mathbf{a} = (1, 2, 10) \quad \mathbf{d} - \mathbf{a} = (4, -2, 8)$$

får vi:

$$\begin{aligned}\text{Volum} &= \frac{1}{6} |((2, 2, 4) \times (1, 2, 10)) \cdot (4, -2, 8)| \\ &= \frac{1}{6} |(12, -16, 2) \cdot (4, -2, 8)| = \frac{96}{6} = 16\end{aligned}$$



Helt til slutt i dette avsnittet skal vi se hvordan vi kan bruke vektorproduktet til å finne ligningen til et plan. Den enkleste måten å beskrive et plan på, er som regel å angi en normalvektor $\mathbf{n} = (n_1, n_2, n_3)$ pluss et punkt $\mathbf{a} = (a_1, a_2, a_3)$ som planet går gjennom. Planet består da av alle de punkter $\mathbf{x} = (x, y, z)$ slik at $\mathbf{x} - \mathbf{a}$ står normalt på \mathbf{n} (se figur 9).



Figur 9: \mathbf{x} ligger i planet gjennom \mathbf{a} normalt på \mathbf{n}

At $\mathbf{x} - \mathbf{a}$ står normalt på \mathbf{n} er ekvivalent med at

$$0 = \mathbf{n} \cdot (\mathbf{x} - \mathbf{a}) = n_1x + n_2y + n_3z - n_1a_1 - n_2a_2 - n_3a_3$$

eller med andre ord

$$n_1x + n_2y + n_3z = n_1a_1 + n_2a_2 + n_3a_3$$

Dette kaller vi *ligningen* til planet. Legg merke til at koeffisientene til x , y og z rett og slett er koordinatene til \mathbf{n}

Eksempel 6: Finn ligningen til planet som går gjennom $\mathbf{a} = (-3, 1, 2)$ og står normalt på $\mathbf{n} = (-4, 2, -1)$. Undersøk om punktet $(-2, 4, 3)$ ligger i dette planet.

Vi skal finne alle vektorer $\mathbf{x} = (x, y, z)$ slik at $0 = \mathbf{n} \cdot (\mathbf{x} - \mathbf{a})$, dvs.

$$0 = (-4, 2, -1) \cdot (x - (-3), y - 1, z - 2)$$

$$= (-4)(x + 3) + 2(y - 1) - 1(z - 2) = -4x + 2y - z - 12$$

Altså blir ligningen

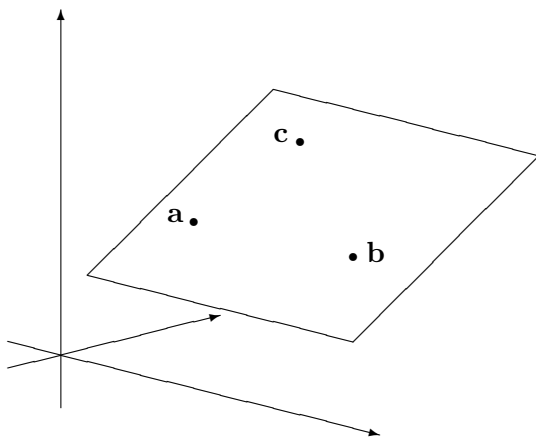
$$-4x + 2y - z = 12$$

For å undersøke om punktet $(-2, 4, 3)$ ligger i planet, sjekker vi om det passer i ligningen. Setter vi inn på venstre side, får vi:

$$-4 \cdot (-2) + 2 \cdot 4 - 3 = 13 \neq 12$$

som viser at punktet ikke ligger i planet. ♣

En annen måte å beskrive et plan på, er å spesifisere tre punkter \mathbf{a} , \mathbf{b} , \mathbf{c} som planet går gjennom, se figur 10 (sørg for at punktene ikke ligger på samme rette linje!). For å finne ligningen til planet, må vi da først finne en normalvektor. Det er ikke så vanskelig; en normalvektor til planet må stå normalt på begge vektorene $\mathbf{b} - \mathbf{a}$ og $\mathbf{c} - \mathbf{a}$ (siden begge disse vektorene i sin helhet ligger i planet), og kryssproduktet $\mathbf{n} = (\mathbf{b} - \mathbf{a}) \times (\mathbf{c} - \mathbf{a})$ er derfor et naturlig valg.



Figur 10: Planet gjennom \mathbf{a} , \mathbf{b} og \mathbf{c}

Eksempel 7: Finn ligningen til planet som går gjennom $\mathbf{a} = (1, -2, 1)$, $\mathbf{b} = (3, -2, 2)$ og $\mathbf{c} = (2, 0, 1)$.

Siden $\mathbf{b} - \mathbf{a} = (2, 0, 1)$ og $\mathbf{c} - \mathbf{a} = (1, 2, 0)$, så er en normalvektor gitt ved

$$\mathbf{n} = (\mathbf{b} - \mathbf{a}) \times (\mathbf{c} - \mathbf{a}) = (2, 0, 1) \times (1, 2, 0) = (-2, 1, 4)$$

Siden normalvektoren er $(-2, 1, 4)$, vet vi at planligningen har formen

$$-2x + y + 4z = d$$

for et tall d . For å finne d , kan vi f.eks. sette inn koordinatene til \mathbf{a} i ligningen:

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

Altså er ligningen til planet

$$-2x + y + 4z = 0$$



Oppgaver til seksjon 1.4

1. Regn ut $\mathbf{a} \times \mathbf{b}$ når

a) $\mathbf{a} = (-1, 3, 2)$ $\mathbf{b} = (-2, 1, 7)$ b) $\mathbf{a} = (4, -3, 1)$ $\mathbf{b} = (-6, 1, 0)$

2. Finn arealet til parallellogrammet utspent av $\mathbf{a} = (-2, 3, 1)$ og $\mathbf{b} = (4, 0, -2)$.

3. En trekant har hjørner i punktene $(0, -1, 2)$, $(2, -1, 4)$ og $(3, 0, 4)$. Finn arealet.

4. Finn en vektor som står normalt på både $(2, 0, -3)$ og $(-1, 3, 4)$.

5. Regn ut $\mathbf{i} \times \mathbf{j}$, $\mathbf{j} \times \mathbf{k}$, $\mathbf{k} \times \mathbf{i}$.

6. Finn volumet til parallellepipedet utspent av $(3, -2, -2)$, $(0, 0, 4)$ og $(-3, 2, 1)$.

7. En pyramide har hjørner i punktene $(2, -1, 2)$, $(0, 5, -3)$, $(2, 4, 6)$ og $(3, -2, 4)$. Finn volumet.

8. Finn en ligning for planet som går gjennom punktene $\mathbf{a} = (1, 1, -1)$, $\mathbf{b} = (0, -2, -6)$, $\mathbf{c} = (2, 3, 3)$.

9. Finn en ligning for planet som går gjennom punktene $\mathbf{a} = (1, 2, 1)$, $\mathbf{b} = (2, 3, 0)$, $\mathbf{c} = (2, 1, -1)$.

10. Anta at alle hjørnene i et parallellepiped har heltallige koeffisienter. Vis at volumet er et helt tall.

11. Bevis setning 1.4.1b).

1.5 Matriser

Hittil har vi bare studert n -tupler, men tiden er nå inne for å introdusere matriser. En $m \times n$ -matrise

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

er et rektangulært oppsett av tall med m rader (linjer) og n søyler. To eksempler er 2×3 -matrisen

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

og 3×3 -matrisen

$$C = \begin{pmatrix} 2 & -3 & 1 \\ \frac{1}{2} & -2 & \pi \\ 1 & -\frac{3}{4} & 0 \end{pmatrix}$$

Tallene $a_{11}, a_{12}, \dots, a_{mn}$ som inngår i matrisen A , kaller vi *elementene* i A . Legg merke til hvordan vi nummerer elementene: a_{34} er elementet i rad 3 og søyle 4 (strengt tatt burde vi ha skrevet et komma mellom 3-tallet og 4-tallet for å gjøre det klart at det ikke er element nummer 34 det er snakk om, men man blir fort lei av å skrive alle kommaene, og det er derfor vanlig å droppe dem dersom det ikke kan oppstå misforståelser). Legg også merke til at vi kan oppfatte vektorer som matriser. En radvektor

$$\mathbf{a} = (a_1, a_2, \dots, a_n)$$

er en $1 \times n$ -matrise, mens en søylevektor

$$\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$$

er en $n \times 1$ -matrise. Som matriser betraktet er altså radvektoren og søylevektoren av forskjellig type, og når vi blander vektorer og matriser er det derfor viktig å holde styr på om \mathbf{a} skal oppfattes som en rad- eller en søylevektor. Vi skal komme tilbake til dette etter hvert. Legg forøvrig merke til at en 1×1 -matrise (a_{11}) bare er et tall med en parentes rundt. Parentesen spiller ingen rolle, og vi skal derfor si at en 1×1 -matrise bare er et tall a_{11} .

Akkurat som for n -tupler kan vi definere addisjon og subtraksjon komponentvis. Dersom vi har to $m \times n$ -matriser

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \quad \text{og} \quad B = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{pmatrix}$$

så definerer vi

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

og

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

Siden alle regneoperasjoner foregår komponentvis, vil de vanlige regnereglene for addisjon og subtraksjon også gjelde for matriser.

Vi kan også multiplisere en matrise med et tall s ved å gange tallet inn i hver komponent:

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

La oss ta et eksempel der vi kombinerer flere av regneoperasjonene:

Eksempel 1: Regn ut $3A - 2B$ når

$$A = \begin{pmatrix} 1 & -1 & 2 \\ 4 & 0 & -1 \end{pmatrix} \quad \text{og} \quad B = \begin{pmatrix} 2 & -3 & 0 \\ 6 & -2 & 3 \end{pmatrix}$$

Vi får

$$3A - 2B = \begin{pmatrix} 3 & -3 & 6 \\ 12 & 0 & -3 \end{pmatrix} - \begin{pmatrix} 4 & -6 & 0 \\ 12 & -4 & 6 \end{pmatrix} = \begin{pmatrix} -1 & 3 & 6 \\ 0 & 4 & -9 \end{pmatrix}$$



Transponering er en viktig operasjon for matriser som ikke finnes for tall. Vi transponerer en matrise ved å bytte om rader og søyler. Den *transponerte* til matrisen A ovenfor er derfor

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

Legg merke til at når A er en $m \times n$ -matrise, så er A^T en $n \times m$ -matrise.

Eksempel 2: Den transponerte til 2×3 -matrisen

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

er

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

som er en 3×2 -matrise.



Transponering kan lett kombineres med addisjon, subtraksjon og multiplikasjon med skalar. Du kan sjekke at

$$(A + B)^T = A^T + B^T, \quad (A - B)^T = A^T - B^T, \quad (cA)^T = cA^T$$

Hvis vi transponerer den transponerte, kommer vi tilbake til utgangspunktet:

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

Legg også merke til at rad- og søylevektorer er transponerte av hverandre. Oppfatter vi søylevektoren

$$\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$$

som en matrise, blir den transponerte en radvektor

$$\mathbf{a}^T = (a_1, a_2, \dots, a_n)$$

De grunnleggende regneoperasjonene for matriser som vi nå har sett på, er ganske enkle, men man kan jo lure på hva det hele er godt for — hvorfor innfører vi egentlig matriser når vi allerede har n -tupler til å holde styr på lister av tall? Ett svar er at en del lister kommer naturlig i rektangelform, og at det er greit å beholde denne formen for ikke å miste oversikt over informasjonen. Opplysningene vi legger inn i et regneark vil for eksempel ofte ha matrisiform. Det neste eksemplet viser en situasjon som ofte forekommer, og der elementene i matrisen viser fordelingen mellom forskjellige muligheter.

Eksempel 3: En fruktpresse mottar epler fra fire forskjellige produsenter. Eplene blir sortert i tre kategorier: god, middels, dårlig. Erfaringene viser at produsentene har forskjellig kvalitet på sine produkter. Produsent 1 leverer 50% av god kvalitet, 30% av middels kvalitet og 20% av dårlig kvalitet. Tallene for de andre produsentene er: Produsent 2: 30%, 40%, 30%; produsent 3: 25%, 40%, 35%, produsent 4: 20%, 60%, 20%. Dersom vi gir hver produsent en søyle, kan vi sette opp denne informasjonen som en 3×4 -matrise (husk at prosent betyr “hundredel”):

$$A = \begin{pmatrix} 0.5 & 0.3 & 0.25 & 0.2 \\ 0.3 & 0.4 & 0.4 & 0.6 \\ 0.2 & 0.3 & 0.35 & 0.2 \end{pmatrix}$$

Anta nå at fruktpressen mottar leveringer fra hver av produsentene: 4 tonn fra produsent 1, 5 tonn fra produsent 2, 3 tonn fra produsent 3 og 6 tonn fra produsent 4. Vi setter opp denne leveransen som en søylevektor

$$\mathbf{b} = \begin{pmatrix} 4 \\ 5 \\ 3 \\ 6 \end{pmatrix}$$

Vi ønsker å finne ut hvor mange tonn vi har fått av hver kvalitet. Observer at dersom vi ganger hvert av tallene i første rad i A med tilsvarende tall i \mathbf{b} , og så legger sammen, får vi antall tonn av beste kvalitet:

$$0.5 \cdot 4 + 0.3 \cdot 5 + 0.25 \cdot 3 + 0.2 \cdot 6 = 2 + 1.5 + 0.75 + 1.2 = 5.45$$

Tilsvarende får vi antall tonn av nest beste kvalitet ved å gange tallene i annen rad i A med tilsvarende tall i \mathbf{b} , og så legge sammen:

$$0.3 \cdot 4 + 0.4 \cdot 5 + 0.4 \cdot 3 + 0.6 \cdot 6 = 1.2 + 2 + 1.2 + 3.6 = 8$$

Til slutt får vi antall tonn av dårligste kvalitet ved å gange tallene i nederste rad i A med tilsvarende tall i \mathbf{b} og så legge sammen:

$$0.2 \cdot 4 + 0.3 \cdot 5 + 0.35 \cdot 3 + 0.2 \cdot 6 = 0.8 + 1.5 + 1.05 + 1.2 = 4.55$$

Legg merke til at i hvert av disse regnestykkene har vi regnet ut et slags skalarprodukt mellom en rad i matrisen A og vektoren \mathbf{b} (vi sier et “slags” skalarprodukt siden det er et litt uortodokst produkt mellom en radvektor og en søylevektor). Legg også merke til at vi kan tenke på resultatet av regnestykkene som en ny vektor

$$\mathbf{c} = \begin{pmatrix} 5.45 \\ 8 \\ 4.55 \end{pmatrix}$$

som forteller oss hvor mange tonn vi har av hver kvalitet. ♣

Vi skal se på et eksempel til av lignende type.

Eksempel 4: Et kjøpesenter har tre stativ X , Y og Z hvor du kan hente og avlevere handlevogner. Av de vognene som starter dagen i stativ X , vil 70% avslutte den på samme sted, 20% vil ha endt opp i Y , og 10% i Z . Av de vognene som startet dagen i stativ Y , vil 30% avslutte dagen i stativ X , mens henholdsvis 50% og 20% vil havne i stativene Y og Z . De tilsvarende tallene for vogner som starter i Z , er at 40% ender dagen i X , 20% i Y og 40% i Z . Vi kan ordne disse tallene i en matrise A der første søyle gir fordelingen av de vognene som startet i X , andre søyle gir fordelingen av de vognene som startet i Y og tredje søyle gir fordelingen av vognene som startet i Z :

$$A = \begin{pmatrix} 0.7 & 0.3 & 0.4 \\ 0.2 & 0.5 & 0.2 \\ 0.1 & 0.2 & 0.4 \end{pmatrix}$$

Anta nå at vi startet dagen med 100 vogner i X , 70 vogner i Y og 30 vogner i Z , og la oss skrive dette som en søylevektor

$$\mathbf{b} = \begin{pmatrix} 100 \\ 70 \\ 30 \end{pmatrix}$$

Vi ønsker å finne ut hvor mange vogner som befinner seg på hvert sted ved slutten av dagen. Antall vogner i X får vi ved å gange tallene i første rad i A med tilsvarende komponent i vektoren \mathbf{b} , og så legge sammen:

$$0.7 \cdot 100 + 0.3 \cdot 70 + 0.4 \cdot 30 = 70 + 21 + 12 = 103$$

Antall vogner i Y får vi tilsvarende ved å gange hvert tall i annen rad i A med tilsvarende tall i \mathbf{b} , og så legge sammen:

$$0.2 \cdot 100 + 0.5 \cdot 70 + 0.2 \cdot 30 = 20 + 35 + 6 = 61$$

Antall vogner i Z ved slutten av dagen, får vi så ved å gange hvert tall i tredje rad i A med tilsvarende tall i \mathbf{b} , og legge sammen:

$$0.1 \cdot 100 + 0.2 \cdot 70 + 0.4 \cdot 30 = 10 + 14 + 12 = 36$$

Vi kan skrive opp resultatet som en ny vektor

$$\mathbf{c} = \begin{pmatrix} 103 \\ 61 \\ 36 \end{pmatrix}$$

som gir oss fordelingen av handlevogner ved slutten av dagen. Legg merke til at vi i dette eksemplet har gjort akkurat samme operasjoner som i det forrige; vi har tatt skalarproduktene mellom radene i A og søylevektoren \mathbf{b} .



I begge eksemplene ovenfor gjennomførte vi samme type regneoperasjoner. Vi startet med en matrise A og en søylevektor \mathbf{b} , og laget en ny søylevektor \mathbf{c} der komponentene fremkom som skalarprodukt av radene i A og vektoren \mathbf{b} . Denne regneoperasjonen er så vanlig at det er greit å ha et eget navn på den.

Definisjon 1.5.1 (Multiplikasjon av matrise og søylevektor) Anta at A er en $m \times n$ -matrise og at \mathbf{b} er en n -dimensjonal søylevektor:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \quad \text{og} \quad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

Produktet av A og \mathbf{b} er da den m -dimensjonale søylevektoren $\mathbf{c} = A\mathbf{b}$ gitt ved

$$\mathbf{c} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{pmatrix} = \begin{pmatrix} a_{11}b_1 + a_{12}b_2 + \cdots + a_{1n}b_n \\ a_{21}b_1 + a_{22}b_2 + \cdots + a_{2n}b_n \\ \vdots \\ a_{m1}b_1 + a_{m2}b_2 + \cdots + a_{mn}b_n \end{pmatrix}$$

Den i -te komponentene i $\mathbf{c} = A\mathbf{b}$ fremkommer altså ved at vi tar skalarproduktet av den i -te raden i A med vektoren \mathbf{b} .

Legg merke til at produktet $A\mathbf{b}$ bare er definert når A og \mathbf{b} passer sammen størrelsesmessig; \mathbf{b} må ha like mange komponenter som A har søyler. Observer også at produktet $A\mathbf{b}$ er en søylevektor med like mange rader som A .

Eksempel 5: Finn $A\mathbf{b}$ når

$$A = \begin{pmatrix} 2 & 1 \\ 4 & -1 \\ -2 & 5 \end{pmatrix} \quad \text{og} \quad \mathbf{b} = \begin{pmatrix} 3 \\ 6 \end{pmatrix}$$

Vi får

$$A\mathbf{b} = \begin{pmatrix} 2 \cdot 3 + 1 \cdot 6 \\ 4 \cdot 3 + (-1) \cdot 6 \\ (-2) \cdot 3 + 5 \cdot 6 \end{pmatrix} = \begin{pmatrix} 12 \\ 6 \\ 24 \end{pmatrix}$$



Vi skal komme tilbake til regneregler for produkter senere, men skriver opp noen av de enkleste og vanligste (i formlene nedenfor er s et tall):

$$(A + B)\mathbf{b} = A\mathbf{b} + B\mathbf{b}, \quad (sA)\mathbf{b} = s(A\mathbf{b})$$

$$A(\mathbf{b} + \mathbf{c}) = A\mathbf{b} + A\mathbf{c}, \quad A(s\mathbf{b}) = s(A\mathbf{b})$$

Det er god trening å sjekke at disse reglene holder. Vær oppmerksom på at det også er en del ting du kan gjøre med vanlige produkter, som du *ikke* kan gjøre med produktet ovenfor, f.eks. kan du ikke bytte om på rekkefølgen av faktorene ($\mathbf{b}A$ gir ikke mening). Du kan heller ikke forkorte A i uttrykk av typen $A\mathbf{b} = A\mathbf{c}$ eller \mathbf{b} i uttrykk av typen $A\mathbf{b} = B\mathbf{b}$ — å forkorte betyr egentlig å gange med et inverst element på begge sider, og vi vet foreløpig ikke noe om inverse vektorer og matriser (det viser seg at vektorer ikke har inverser, men at noen matriser har det!)

La oss ta en ny kikk på eksempel 3 og 4 i lys av de begrepene vi nå har innført:

Eksempel 3 og 4 på nytt: I eksempel 3 (fruktpressen) ble leveransene fra de fire produsentene kodet som et 4-tupple \mathbf{b} , mens 3-tuplet \mathbf{c} fortalte oss

hvor mange tonn vi mottok av hver kvalitet. Vi regnet ut elementene i \mathbf{c} ved å ta skalarproduktet av radene i matrisen A med vektoren \mathbf{b} — med andre ord:

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

Som vi skal komme tilbake til senere, kan vi tenke på multiplikasjon med A som en “transformasjon” — vektoren \mathbf{b} forteller oss hvor mye frukt vi får fra hver produsent, og multiplikasjon med A “transformerer” denne kunnskapen til kunnskap om hvor mye vi mottar av hver kvalitet (representert ved vektoren \mathbf{c}).

Noe lignende skjer i eksempel 4 (handlevognene). Her representerer vektor \mathbf{b} fordelingen av handlevognar ved begynnelsen av dagen, mens vektor \mathbf{c} representerer fordelingen ved slutten av dagen. Vi regnet ut elementene i \mathbf{c} ved å ta skalarproduktet av radene i matrisen A med vektoren \mathbf{b} — med andre ord:

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

Igjen kan vi tenke på multiplikasjon med A som en transformasjon som transformerer kunnskap om hvor handlevognene er ved begynnelsen av dagen til kunnskap om hvor de er ved slutten av dagen. ♣

La oss avslutte med en liten observasjon som vi kommer til å få glede av senere. Anta at

$$\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \quad \text{og} \quad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

er to søylevektorer. Transponerer vi \mathbf{a} , får vi radvektoren

$$\mathbf{a}^T = (a_1, a_2, \dots, a_n)$$

som vi kan oppfatte som en $1 \times n$ -matrise. Vi kan nå ta produktet av matrisen \mathbf{a}^T og vektoren \mathbf{b} . Dette gir oss en 1×1 -matrise (dvs. et tall):

$$\mathbf{a}^T \mathbf{b} = a_1 b_1 + a_2 b_2 + \dots + a_n b_n = \mathbf{a} \cdot \mathbf{b}$$

Matriseproduktet $\mathbf{a}^T \mathbf{b}$ er altså lik skalarproduktet $\mathbf{a} \cdot \mathbf{b}$. Dette kan virke som en kuriositet, men viser seg å være viktig i en del sammenhenger.

Oppgaver til seksjon 1.5

1. Regn ut $2A$, $(-3)B$, $A + B$ og $A - B$ når

$$A = \begin{pmatrix} 3 & -2 & 4 \\ -1 & 0 & 6 \end{pmatrix} \quad \text{og} \quad B = \begin{pmatrix} 2 & 2 & -1 \\ 4 & 0 & -6 \end{pmatrix}$$

2. Regn ut $4A - 3B$ når

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

3. Finn de transponerte:

$$A = \begin{pmatrix} 1 & -2 & 6 \\ -1 & 8 & 6 \end{pmatrix} \quad \text{og} \quad B = \begin{pmatrix} 2 & 2 & -1 \\ 4 & 0 & -6 \\ 1 & 7 & -3 \end{pmatrix}$$

4. Finn de transponerte til A , B og $4A - B$ når

$$A = \begin{pmatrix} 1 & 0 \\ -5 & 7 \\ 2 & -1 \end{pmatrix} \quad \text{og} \quad B = \begin{pmatrix} 4 & -1 \\ -1 & 3 \\ 2 & -6 \end{pmatrix}$$

5. Regn ut $A\mathbf{x}$ når:

a) $A = \begin{pmatrix} 1 & 0 & -3 \\ -2 & -3 & 2 \end{pmatrix}$ og $\mathbf{x} = \begin{pmatrix} -2 \\ 3 \\ -1 \end{pmatrix}$.

b) $A = \begin{pmatrix} 2 & 0 \\ 3 & 1 \\ 6 & -2 \end{pmatrix}$ og $\mathbf{x} = \begin{pmatrix} 3 \\ -2 \end{pmatrix}$.

c) $A = \begin{pmatrix} 2 & 1 & 0 \\ -3 & 4 & -2 \\ 1 & -3 & 2 \end{pmatrix}$ og $\mathbf{x} = \begin{pmatrix} 4 \\ 0 \\ 3 \end{pmatrix}$.

6. Regn ut hvor mange tonn vi får av hver kvalitet i eksempel 3 dersom produsent 1 leverer 10 tonn, produsent 2 leverer 5 tonn, produsent 3 leverer 8 tonn og produsent 4 leverer 6 tonn.

7. Hvordan vil handlevognene i eksempel 4 fordele seg ved slutten av dagen dersom vi begynner med 50 vogner i stativ X , 70 i stativ Y og 80 i stativ Z ?

8. Bevis regnereglene for transponering som står rett etter eksempel 2.

9. Bevis regnereglene for multiplikasjon av matrise og vektor som står rett etter eksempel 5.

10. Et oljemottak får olje fra 3 oljefelt. Oljen inndeles i fem kvaliteter A, B, C, D, E. Fra oljefelt 1 er fordelingen: A: 10%, B: 20%, C: 30%, D: 30%, E: 10%. Fra oljefelt 2 er fordelingen: A: 0%, B: 30%, C: 30%, D: 30%, E: 10%. Fra oljefelt 3 er fordelingen: A: 20%, B: 20%, C: 10%, D: 10%, E: 40%. Anta at mottaket får inn x enheter fra felt 1, y enheter fra felt 2 og z enheter fra felt 3. Finn en matrise som kan brukes til å regne ut hvor mange enheter man får av hver kvalitet. Hva blir resultatet når $x = 10$, $y = 12$ og $z = 8$?

11. En smittsom sykdom sprer seg i et land. Helsemyndighetene deler befolkningen i tre grupper: smitteutsatte, syke, immune. Fra en uke til den neste regner man at 5% av de smitteutsatte blir syke, mens 1% av dem blir immune uten å ha vært

syke. Av de syke blir 80% immune, mens resten fortsatt er syke. Én prosent av de immune mister immuniteten og blir smitteutsatte, mens resten fortsatt er immune. La x_n, y_n, z_n være den andelen av befolkningen som er hhv. smitteutsatt, syk og

immun etter n uker. La $\mathbf{v}_n = \begin{pmatrix} x_n \\ y_n \\ z_n \end{pmatrix}$.

- Finn en matrise A slik at $A\mathbf{v}_n = \mathbf{v}_{n+1}$.
- I uke 0 er 10% av befolkningen syke, mens resten er smitteutsatte. Finn fordelingen av smitteutsatte, syke og immune i uke 1 og 2.

12. En forskergruppe som studerer et bestemt dyreslag, deler bestanden inn i fire grupper:

- unge, dvs. dyr født samme vår
- unge voksne, dvs. dyr født året før
- voksne, dvs. dyr født to år før
- gamle, dvs. dyr født mer enn to år før

Statistikken viser at bare 5% av de unge overlever til året etter. Av de unge voksne overlever 50% til året etter. I tillegg gir en ung voksen i gjennomsnitt opphav til 20 unger som blir født året etter. Blant de voksne overlever 30% til året etter. I tillegg gir en voksen i gjennomsnitt opphav til 50 unger som blir født året etter. Av de gamle overlever 10% til året etter. I tillegg gir en gammel i gjennomsnitt opphav til 10 unger som blir født året etter. La x_n, y_n, z_n, u_n være hhv. antall unge, antall unge voksne, antall voksne og antall gamle etter n år, og skriv

$$\mathbf{v}_n = \begin{pmatrix} x_n \\ y_n \\ z_n \\ u_n \end{pmatrix}$$

- Finn en matrise A slik at $\mathbf{v}_{n+1} = A\mathbf{v}_n$.
- Anta at det et år settes ut 100 voksne dyr i et terreng der det ikke er noen dyr fra før av. Hvor mange dyr av hver kategori vil det være i dette terrenget to år senere?

13. Per, Pål og Espen leker sisten. Hver gang Per eller Pål har sisten, er det 90% sjanse for at Espen får sisten neste gang, mens hver gang Espen har sisten, har Per og Pål 50% sjanse hver for å få sisten neste gang. Anta at x_n er sannsynligheten for at Per er den n -te som har sisten, at y_n er sannsynligheten for at Pål er den n -te som har sisten, og at z_n er sannsynligheten for at Espen er den n -te som har

sisten. La $\mathbf{v}_n = \begin{pmatrix} x_n \\ y_n \\ z_n \end{pmatrix}$.

- Finn en matrise A slik at $\mathbf{v}_{n+1} = A\mathbf{v}_n$.
- Anta at Pål har sisten først. Hva er sannsynligheten for at henholdsvis Per, Pål og Espen er den tredje som har sisten?

1.6 Multiplikasjon av matriser

I forrige seksjon så vi at addisjon og subtraksjon av matriser er lett; vi bare adderer og subtraherer komponentvis. Det er fristende å definere multiplikasjon på samme måte; hvis

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \quad \text{og} \quad B = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{pmatrix}$$

så bør kanskje produktet være

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

Det er ikke noe galt med dette produktet (det kalles *Hadamard-produktet* og er nyttig for noen formål), men det er en annen type matriseprodukt som er langt viktigere. Ved første øyekast ser dette produktet litt merkelig ut, og for å motivere det skal vi gå tilbake til produktet $A\mathbf{b}$ fra slutten av forrige seksjon. (Dersom du synes denne motivasjonen er vanskelig, kan du trøste deg med at det kommer en oppsummering i definisjon 1.6.1.)

La oss se litt nærmere på produktet $\mathbf{c} = A\mathbf{b}$ mellom en matrise A og en vektor \mathbf{b} . Som allerede nevnt, er det ofte smart å tenke på dette som en *transformasjon*; vi starter med en vektor \mathbf{b} , og matrisen A transformerer \mathbf{b} til en ny vektor \mathbf{c} . I eksempel 3 i forrige seksjon (fruktpressen), så vi f.eks. hvordan A transformerer kunnskap om hvor mange tonn epler vi får fra hver produsent (kodet opp i vektoren \mathbf{b}) til kunnskap om hvor mange tonn epler vi har av hver kvalitet (kodet opp i vektoren $\mathbf{c} = A\mathbf{b}$). Noe tilsvarende skjedde i eksempel 4 (handlevognene); i dette tilfellet transformerer A kunnskap om hvor vognene er ved begynnelsen av dagen (gitt ved vektoren \mathbf{b}), til kunnskap om hvor de er ved slutten av dagen (gitt ved vektoren $\mathbf{c} = A\mathbf{b}$). Legg merke til at i begge eksemplene bruker vi den samme matrisen A uansett hvilken input-vektor \mathbf{b} vi har — det er slik gjenbruk som ofte gjør matriser nyttige.

Vi kan illustrere situasjonen med diagrammet nedenfor; en $m \times n$ -matrise A transformerer vektorer \mathbf{x} i \mathbb{R}^n til vektorer $\mathbf{y} = A\mathbf{x}$ i \mathbb{R}^m :

$$\mathbf{x} \in \mathbb{R}^n \xrightarrow{A} \mathbf{y} = A\mathbf{x} \in \mathbb{R}^m$$

I mange situasjoner må vi foreta flere transformasjoner etter hverandre. Neste diagram viser en slik situasjon; først transformeres $\mathbf{x} \in \mathbb{R}^k$ til $\mathbf{y} = B\mathbf{x} \in \mathbb{R}^n$ ved hjelp av matrisen B , og deretter transformeres $\mathbf{y} \in \mathbb{R}^n$ til $\mathbf{z} = A\mathbf{y} = A(B\mathbf{x}) \in \mathbb{R}^m$ ved hjelp av matrisen A :

$$\mathbf{x} \in \mathbb{R}^k \xrightarrow{B} \mathbf{y} = B\mathbf{x} \in \mathbb{R}^n \xrightarrow{A} \mathbf{z} = A\mathbf{y} = A(B\mathbf{x}) \in \mathbb{R}^m$$

Det er naturlig å spørre om det finnes en matrise C som tar oss direkte fra \mathbf{x} til \mathbf{z} uten å gå veien om \mathbf{y} , dvs. slik at $\mathbf{z} = C\mathbf{x}$. Neste diagram viser hvordan en slik C vil fungere:

$$\begin{array}{ccccc} \mathbf{x} \in \mathbb{R}^k & \xrightarrow{B} & \mathbf{y} = B\mathbf{x} \in \mathbb{R}^n & \xrightarrow{A} & \mathbf{z} = A\mathbf{y} = A(B\mathbf{x}) \in \mathbb{R}^m \\ & \searrow & & \nearrow & \\ & & C & & \end{array}$$

Det viser seg at det eksisterer en slik matrise C , og vi skal nå finne den. Før vi begynner kan det være greit å bestemme størrelsen til de tre matrisene våre A , B og C . Vi ser at A transformerer vektorer i \mathbb{R}^n til vektorer i \mathbb{R}^m , og det betyr at A er en $m \times n$ -matrise. Tilsvarende transformerer B vektorer i \mathbb{R}^k til vektorer i \mathbb{R}^n , og det betyr at B er en $n \times k$ -matrise. Den ukjente matrisen C skal transformere vektorer i \mathbb{R}^k til vektorer i \mathbb{R}^m , og må derfor være en $m \times k$ -matrise. Vi har derfor

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \quad \text{og} \quad B = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1k} \\ b_{21} & b_{22} & \cdots & b_{2k} \\ \vdots & \vdots & \cdots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nk} \end{pmatrix}$$

som våre kjente matriser, og

$$C = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1k} \\ c_{21} & c_{22} & \cdots & c_{2k} \\ \vdots & \vdots & \cdots & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mk} \end{pmatrix}$$

som vår ukjente matrise. Hvis \mathbf{x} er vektoren

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{pmatrix},$$

ser vi at

$$\mathbf{y} = B\mathbf{x} = \begin{pmatrix} b_{11}x_1 + b_{12}x_2 + \cdots + b_{1k}x_k \\ b_{21}x_1 + b_{22}x_2 + \cdots + b_{2k}x_k \\ \vdots \\ b_{n1}x_1 + b_{n2}x_2 + \cdots + b_{nk}x_k \end{pmatrix}$$

Ser vi på den j -te raden i denne ligningen, får vi $y_j = b_{j1}x_1 + b_{j2}x_2 + \cdots + b_{jk}x_k$. Vi har også

$$\mathbf{z} = A\mathbf{y} = \begin{pmatrix} a_{11}y_1 + a_{12}y_2 + \cdots + a_{1n}y_n \\ a_{21}y_1 + a_{22}y_2 + \cdots + a_{2n}y_n \\ \vdots \\ a_{m1}y_1 + a_{m2}y_2 + \cdots + a_{mn}y_n \end{pmatrix}$$

og setter vi $y_j = b_{j1}x_1 + b_{j2}x_2 + \cdots + b_{jk}x_k$ inn i dette uttrykket, ser vi at den i -te komponenten til \mathbf{z} er gitt ved

$$z_i = a_{i1}y_1 + a_{i2}y_2 + \cdots + a_{in}y_n = a_{i1}(b_{11}x_1 + b_{12}x_2 + \cdots + b_{1k}x_k) + \\ + a_{i2}(b_{21}x_1 + b_{22}x_2 + \cdots + b_{2k}x_k) + \cdots + a_{in}(b_{n1}x_1 + b_{n2}x_2 + \cdots + b_{nk}x_k)$$

Samler vi alle x_1 -ledd for seg, alle x_2 -ledd for seg osv, får vi:

$$z_i = (a_{i1}b_{11} + a_{i2}b_{21} + \cdots + a_{in}b_{n1})x_1 + \\ + (a_{i1}b_{12} + a_{i2}b_{22} + \cdots + a_{in}b_{n2})x_2 + \cdots \\ + (a_{i1}b_{1k} + a_{i2}b_{2k} + \cdots + a_{in}b_{nk})x_k$$

På den annen side: hvis vi tenker oss at \mathbf{z} fremkommer direkte fra \mathbf{x} ved bruk av matrisen C , har vi

$$\mathbf{z} = C\mathbf{x} = \begin{pmatrix} c_{11}x_1 + c_{12}x_2 + \cdots + c_{1k}x_k \\ c_{21}x_1 + c_{22}x_2 + \cdots + c_{2k}x_k \\ \vdots \\ c_{m1}x_1 + c_{m2}x_2 + \cdots + c_{mk}x_k \end{pmatrix}$$

dvs. at den i -te komponenten til z er gitt ved

$$z_i = c_{i1}x_1 + c_{i2}x_2 + \cdots + c_{ik}x_k$$

Skal de to uttrykkene for z_i være like, må vi ha

$$c_{i1} = a_{i1}b_{11} + a_{i2}b_{21} + \cdots + a_{in}b_{n1}$$

$$c_{i2} = a_{i1}b_{12} + a_{i2}b_{22} + \cdots + a_{in}b_{n2}$$

osv. ned til

$$c_{ik} = a_{i1}b_{1k} + a_{i2}b_{2k} + \cdots + a_{in}b_{nk}$$

Generelt har vi altså

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$$

Legg merke til at dette er skalarproduktet av den i -te raden i A med den j -te søylen i B .

La oss oppsummere. Vi har vist at den matrisen

$$C = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1k} \\ c_{21} & c_{22} & \cdots & c_{2k} \\ \vdots & \vdots & \cdots & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mk} \end{pmatrix}$$

som i én operasjon utfører den samme transformasjonen som B etterfulgt av A , er gitt ved

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$$

Siden det å transformere vektorer er det viktigste matriser gjør, tar vi denne formelen som utgangspunkt for vår definisjon av matriseprodukt.

Definisjon 1.6.1 Anta at A er en $m \times n$ -matrise og at B er en $n \times k$ -matrise. Da er matriseproduktet $C = AB$ definert som $m \times k$ -matrisen C med komponenter

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$$

Vi får altså den ij -te komponenten i C ved å ta skalarproduktet av den i -te raden i A med den j -te søylen i B .

Figuren viser grafisk hvordan vi finner det ij -te elementet i produktmatrisen C : Vi tar skalarproduktet av den i -te raden i A med den j -te søylen i B :

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ \boxed{a_{i1} \quad a_{i2} \quad \cdots \quad a_{in}} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & \cdots & \boxed{b_{1j}} & \cdots & b_{1k} \\ b_{21} & b_{22} & \cdots & \boxed{b_{2j}} & \cdots & b_{2k} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ b_{n1} & b_{n2} & \cdots & \boxed{b_{nj}} & \cdots & b_{nk} \end{pmatrix}$$

Legg merke til at matriseproduktet AB bare er definert når A og B passer sammen i størrelse: radene i A må være like lange som søylene i B . Dette betyr at den “siste” dimensjonen n i $m \times n$ -matrisen A er lik den “første” dimensjonen n i $n \times k$ -matrisen B . Legg merke til at hvis vi stryker de to n -ene i $m \times n$ og $n \times k$, sitter vi igjen med størrelsen $m \times k$ til produktmatrisen C .

Eksempel 1: Regn ut AB når

$$A = \begin{pmatrix} 2 & 3 \\ 1 & -1 \\ 4 & -2 \end{pmatrix} \quad \text{og} \quad B = \begin{pmatrix} 4 & -3 & 2 \\ -1 & 4 & 5 \end{pmatrix}$$

Siden A er en 3×2 og B er 2×3 -matrise, eksisterer produktet og er en 3×3 -matrise. Vi får

$$AB = \begin{pmatrix} 2 \cdot 4 + 3 \cdot (-1) & 2 \cdot (-3) + 3 \cdot 4 & 2 \cdot 2 + 3 \cdot 5 \\ 1 \cdot 4 + (-1) \cdot (-1) & 1 \cdot (-3) + (-1) \cdot 4 & 1 \cdot 2 + (-1) \cdot 5 \\ 4 \cdot 4 + (-2) \cdot (-1) & 4 \cdot (-3) + (-2) \cdot 4 & 4 \cdot 2 + (-2) \cdot 5 \end{pmatrix} =$$

$$= \begin{pmatrix} 5 & 6 & 19 \\ 5 & -7 & -3 \\ 18 & -20 & -2 \end{pmatrix}$$



Legg merke til at selv i de tilfellene hvor AB er definert, kan vi ikke nødvendigvis regne ut BA . Her prøver vi nemlig å multiplisere en $n \times k$ -matrise med en $m \times n$ -matrise, og det er bare mulig når $k = m$. Selv i de tilfellene hvor både AB og BA er definert, er de som regel forskjellige. Her er et eksempel.

Eksempel 2: Regn ut AB og BA når

$$A = \begin{pmatrix} 2 & 1 \\ -1 & 4 \end{pmatrix} \quad \text{og} \quad B = \begin{pmatrix} -3 & 2 \\ 4 & 5 \end{pmatrix}$$

Vi får

$$AB = \begin{pmatrix} 2 \cdot (-3) + 1 \cdot 4 & 2 \cdot 2 + 1 \cdot 5 \\ (-1) \cdot (-3) + 4 \cdot 4 & (-1) \cdot 2 + 4 \cdot 5 \end{pmatrix} = \begin{pmatrix} -2 & 9 \\ 19 & 18 \end{pmatrix}$$

og

$$BA = \begin{pmatrix} (-3) \cdot 2 + 2 \cdot (-1) & (-3) \cdot 1 + 2 \cdot 4 \\ 4 \cdot 2 + 5 \cdot (-1) & 4 \cdot 1 + 5 \cdot 4 \end{pmatrix} = \begin{pmatrix} -8 & 5 \\ 3 & 24 \end{pmatrix}$$

Vi har altså $AB \neq BA$.



Regnereglen $ab = ba$ (som altså ikke gjelder for matriser!) kalles for *den kommutative lov*. Siden matrisemultiplikasjon ikke oppfyller denne regelen, sier vi at matriseproduktet er *ikke-kommutativt*. Vi har tidligere sett at vektorproduktet ikke er kommutativt (det er heller ikke skalarproduktet for komplekse n -tupler), men der har vi i hvert fall en grei regel for sammenhengen mellom $\mathbf{a} \times \mathbf{b}$ og $\mathbf{b} \times \mathbf{a}$. Det har vi ikke for multiplikasjon av matriser; kjenner vi AB , har vi fortsatt ingen anelse om hva BA er!

Det tar tid å vende seg til ikke-kommutative operasjoner, og det eneste rådet vi kan gi er: Vær forsiktig og begrunn hvert eneste skritt når du regner med slike operasjoner! Heldigvis er det mange regneregler som fortsatt gjelder for matriseprodukter:

Setning 1.6.2 (Regneregler for matrisemultiplikasjon) I hvert av punktene nedenfor antar vi at A , B og C er matriser slik at regneoperasjonene er definert.

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

$$(ii) A(B + C) = AB + AC$$

$$(iii) (B + C)A = BA + CA$$

$$(iv) (sA)B = A(sB) = s(AB) \text{ for alle tall } s$$

$$(v) (AB)^T = B^T A^T$$

Bevis: Bevisene går rett og slett ut på å gange ut venstre og høyre side av likhetene og sjekke at svarene er like. Vi tar (i) og (v) som eksempler — de andre punktene er atskillig enklere enn (i):

* (i) (Dette punktet er kronglete og du mister ikke noe i fortsettelsen hvis du hopper over beviset.) Vi skal bevise at de to matrisene $X = (AB)C$ og $Y = A(BC)$ er like. Før man begynner, er det lurt å være helt sikker på hva som er forskjellen på de to uttrykkene. Poenget er at vi alltid gjennomfører operasjonene inni parentesene først. Når vi regner ut $X = (AB)C$, multipliserer vi altså først sammen A og B til en matrise $D = AB$, og så multipliserer vi D fra høyre med C slik at vi får $X = DC$. Når vi regner ut $Y = A(BC)$, ganger vi først sammen B og C til en matrise $E = BC$, og så ganger vi denne fra venstre med A slik at vi får $Y = AE$. Vår oppgave er altså å vise at $DC = AE$. Formulert på denne måten er ikke likheten så selvfølgelig som den kan se ut ved første blick.

Før vi begynner på regningene, er det greit å se litt på størrelsen til de involverte matrisene. For at produktene $(AB)C$ og $A(BC)$ skal gi mening, må det finnes tall m, n, k, l slik at A er en $m \times n$ -matrise, B en $n \times k$ -matrise og C en $k \times l$ -matrise. Dermed blir $D = AB$ en $m \times k$ -matrise og $E = BC$ en $n \times l$ -matrise. Legg merke til at både $X = DC$ og $Y = AE$ er $m \times l$ -matriser, så størrelsesmessig er i hvert fall de to matrisene like.

Vi er nå klare til å gjøre regningene våre. Vi observerer først at det ij -te elementet i $X = DC$ er lik

$$x_{ij} = d_{i1}c_{1j} + d_{i2}c_{2j} + \cdots + d_{ik}c_{kj} = \sum_{p=1}^k d_{ip}c_{pj}$$

Siden $D = AB$ er videre

$$d_{ip} = a_{i1}b_{1p} + a_{i2}b_{2p} + \cdots + a_{in}b_{np} = \sum_{q=1}^n a_{iq}b_{qp}$$

Setter vi dette inn i uttrykket ovenfor, får vi

$$x_{ij} = \sum_{p=1}^k d_{ip} c_{pj} = \sum_{p=1}^k \left(\sum_{q=1}^n a_{iq} b_{qp} \right) c_{pj} = \sum_{p=1}^k \sum_{q=1}^n a_{iq} b_{qp} c_{pj}$$

dvs. at x_{ij} er summen av alle mulige kombinasjoner $a_{iq} b_{qp} c_{pj}$.

På tilsvarende måte regner vi ut $Y = AE$. Det ij -te elementet i Y er

$$y_{ij} = a_{i1} e_{1j} + a_{i2} e_{2j} + \cdots + a_{in} e_{nj} = \sum_{q=1}^n a_{iq} e_{qj}$$

Siden $E = BC$ er videre

$$e_{qj} = b_{q1} c_{1j} + b_{q2} c_{2j} + \cdots + b_{qk} c_{kj} = \sum_{p=1}^k b_{qp} c_{pj}$$

Setter vi dette inn i uttrykket ovenfor, får vi

$$y_{ij} = \sum_{q=1}^n a_{iq} e_{qj} = \sum_{q=1}^n a_{iq} \left(\sum_{p=1}^k b_{qp} c_{pj} \right) = \sum_{q=1}^n \sum_{p=1}^k a_{iq} b_{qp} c_{pj}$$

Dette viser at y_{ij} er lik summen av alle mulige kombinasjoner $a_{iq} b_{qp} c_{pj}$, akkurat som x_{ij} er. Altså er $x_{ij} = y_{ij}$, og beviset for (i) er ferdig.

(v) Anta at A er en $m \times n$ -matrise og at B er en $n \times k$ -matrise. Da er A^T en $n \times m$ -matrise og B^T en $k \times n$ -matrise, så både $(AB)^T$ og $B^T A^T$ er $k \times m$ -matriser. Det ij -te elementet i $C = (AB)^T$ er lik det ji -te elementet i AB , dvs:

$$c_{ij} = a_{j1} b_{1i} + a_{j2} b_{2i} + \cdots + a_{jn} b_{ni}$$

På den annen side fremkommer det ij -te elementet i $D = B^T A^T$ ved at vi tar skalarproduktet av den i -te raden i B^T med den j -te søylen i A^T . Men den i -te raden i B^T er den i -te søylen i B , og den j -te søylen i A^T , er den j -te raden i A , så dette blir skalarproduktet mellom den j -te raden i A og den i -søylen i B , dvs.

$$d_{ij} = a_{j1} b_{1i} + a_{j2} b_{2i} + \cdots + a_{jn} b_{ni}$$

Altså er $c_{ij} = d_{ij}$, og beviset er fullført. \square

Regnereglene ovenfor gjør det enklere å regne med matriser, men vi må som sagt være forsiktige med en del ting — spesielt at vi ikke bytter om på rekkefølgen av matrisene (husk at vanligvis er AB ulik BA). Vi kan heller ikke forkorte matriser på vanlig måte — å forkorte betyr egentlig å

multiplisere med et invert element, og foreløpig vet vi ingen ting om inverse matriser.

La oss ta med oss en liten observasjon. Dersom A er en $m \times n$ -matrise og \mathbf{b} er en søylevektor med n -komponenter, kan vi danne produktet $A\mathbf{b}$ på to måter — vi kan enten tenke på det som produktet av matrisen A og vektoren \mathbf{b} slik vi definerte det i forrige seksjon, eller vi kan tenke på det som produktet av $m \times n$ -matrisen A med $n \times 1$ -matrisen

$$\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

slik vi har definert det i denne seksjonen. Det er lett å se at disse to måtene gir akkurat samme svar. Dette betyr at vi kan bruke regnereglene ovenfor også når noen av matrisene er vektorer.

Eksempel 3: Vi har en $m \times n$ -matrise A og en $n \times k$ -matrise B . Vi lar B transformere en søylevektor $\mathbf{x} \in \mathbb{R}^k$ ved $\mathbf{y} = B\mathbf{x}$. Så lar vi A transformere \mathbf{y} slik at $\mathbf{z} = A\mathbf{y}$. Det samlede resultatet er:

$$\mathbf{z} = A\mathbf{y} = A(B\mathbf{x})$$

Ved hjelp av regneregel (i) ovenfor kan vi flytte parentesene:

$$\mathbf{z} = A\mathbf{y} = A(B\mathbf{x}) = (AB)\mathbf{x}$$

Dette viser noe vi allerede vet, nemlig at når vi transformerer en vektor to ganger ved først å bruke B og deretter A , så er det det samme som å transformere én gang ved hjelp av produktet AB . ♣

Vi avslutter denne seksjonen med en ulikhet som vi vil få bruk for i neste kapittel. Hvis

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

er en $m \times n$ -matrise, definerer vi *normen til A* til å være

$$\|A\| = \sqrt{\sum_{i=1}^m \sum_{j=1}^n a_{ij}^2}$$

(vi legger altså sammen kvadratene a_{ij}^2 av alle leddene og tar kvadratroten til summen).

Setning 1.6.3 Hvis A er en $m \times n$ -matrise og $\mathbf{x} \in \mathbb{R}^n$, så er

$$|A\mathbf{x}| \leq \|A\| |\mathbf{x}|$$

Med andre ord: Når vi ganger vektoren \mathbf{x} med matrisen A , så øker lengden maksimalt med en faktor $\|A\|$.

Bevis: La $\mathbf{y} = A\mathbf{x}$. Da er

$$|A\mathbf{x}| = \sqrt{y_1^2 + y_2^2 + \cdots + y_m^2}$$

der y_1, y_2, \dots, y_m er komponentene til \mathbf{y} . Dersom $\mathbf{a}_i = (a_{i1}, a_{i2}, \dots, a_{in})$ er den i -te linjen i A , vet vi at $y_i = \mathbf{a}_i \cdot \mathbf{x}$. Ifølge Schwarz' ulikhet er dermed $|y_i| \leq |\mathbf{a}_i| |\mathbf{x}|$. Kvadrerer vi, får vi $y_i^2 \leq |\mathbf{a}_i|^2 |\mathbf{x}|^2$. Setter vi dette inn i uttrykket for $A\mathbf{x}$, ser vi at

$$\begin{aligned} |A\mathbf{x}| &= \sqrt{y_1^2 + y_2^2 + \cdots + y_m^2} \leq \\ &\leq \sqrt{|\mathbf{a}_1|^2 |\mathbf{x}|^2 + |\mathbf{a}_2|^2 |\mathbf{x}|^2 + \cdots + |\mathbf{a}_m|^2 |\mathbf{x}|^2} = \\ &= \sqrt{|\mathbf{a}_1|^2 + |\mathbf{a}_2|^2 + \cdots + |\mathbf{a}_m|^2} |\mathbf{x}| = \|A\| |\mathbf{x}| \end{aligned}$$

der vi har brukt at

$$\sqrt{|\mathbf{a}_1|^2 + |\mathbf{a}_2|^2 + \cdots + |\mathbf{a}_m|^2} = \sqrt{\sum_{i=1}^m \sum_{j=1}^n a_{ij}^2} = \|A\|$$

□

Oppgaver til seksjon 1.6

1. Regn ut AB og BA når

a) $A = \begin{pmatrix} 1 & -2 \\ 3 & 1 \end{pmatrix}$ og $B = \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}$

b) $A = \begin{pmatrix} 1 & -1 & 0 \\ -2 & 0 & 1 \\ -1 & 2 & 1 \end{pmatrix}$ og $B = \begin{pmatrix} 0 & 2 & 1 \\ -1 & -2 & 0 \\ 3 & -1 & 2 \end{pmatrix}$

2. Regn ut AB når

$$A = \begin{pmatrix} 1 & -2 & 3 \\ 0 & -1 & 2 \end{pmatrix} \text{ og } B = \begin{pmatrix} 2 & 1 \\ 0 & -3 \\ 1 & 0 \end{pmatrix}.$$

3. Regn ut AB når

$$A = \begin{pmatrix} 1 & -2 \\ 3 & 0 \\ -1 & 2 \end{pmatrix} \text{ og } B = \begin{pmatrix} 2 & 1 & 0 \\ -3 & 1 & 1 \end{pmatrix}.$$

4. a) Hvilken dimensjon har produktmatrisen AB hvis A er en 8×6 -matrise og B er en 6×9 -matrise?

b) Hvilken dimensjon har matrisen B hvis A er en 4×3 -matrise og produktmatrisen AB er en 4×5 -matrise?

c) Hvor mange søyler har matrisen B hvis produktet AB er en 5×7 -matrise?

5. Gitt følgende matriser:

$$A = \begin{pmatrix} 1 & 9 \\ 0 & 8 \\ 7 & 4 \end{pmatrix}, B = \begin{pmatrix} 2 & 8 \\ 9 & 6 \end{pmatrix} \text{ og } C = \begin{pmatrix} 3 & 0 & 5 \\ 0 & 9 & 4 \end{pmatrix}$$

Regn ut følgende uttrykk hvis det er definert (hvis uttrykket ikke er definert, skal du begrunne hvorfor):

- a) AB
- b) AC
- c) $A(B + C)$
- d) $(BC)^T$
- e) $B^T C^T$
- f) $(A + C^T)B$
- g) $B(A^T - 2C)$

6. Gitt matrisene

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 2 \end{pmatrix}, B = \begin{pmatrix} 1 & 1 \\ 3 & 4 \end{pmatrix}, C = \begin{pmatrix} 2 & 5 \\ 3 & 4 \end{pmatrix} \text{ og } D = \begin{pmatrix} 3 & 7 \\ 0 & 0 \end{pmatrix}$$

Sjekk at $AB = AC$ selv om $B \neq C$. Sjekk også at $AD = 0$ selv om $A \neq 0$ og $D \neq 0$.

7. Gitt matrisene $D = \begin{pmatrix} 3 & 1 & 4 \\ 1 & 5 & 9 \\ 2 & 6 & 5 \end{pmatrix}$ og $E = \begin{pmatrix} 2 & 7 & 1 \\ 8 & 2 & 8 \\ 1 & 8 & 2 \end{pmatrix}$.

Finn første rad i produktmatrisen DE . Finn deretter andre søyle i produktmatrisen DE .

8. a) Hvis første og andre søyle i matrisen B er like, hva kan du si om første og andre søyle i produktmatrisen AB (dersom AB er definert)?

b) Hvis andre søyle i matrisen B bare består av nuller, hva kan du si om andre søyle i produktmatrisen AB ?

9. Finn normen $\|A\|$ til matrisen $A = \begin{pmatrix} 1 & -1 & 3 \\ 2 & 1 & 0 \end{pmatrix}$.

10. Dersom A er en $n \times n$ -matrise (dvs. har like mange rader og søyler) definerer vi potensene A^n på vanlig måte: $A^2 = AA$, $A^3 = AA^2$, $A^4 = AA^3$ osv. Regn ut A^2 og A^3 når $A = \begin{pmatrix} 1 & -2 \\ -1 & 1 \end{pmatrix}$.

11. La $A = \begin{pmatrix} 1 & -1 & 2 \\ 0 & 1 & 3 \end{pmatrix}$, $B = \begin{pmatrix} 2 & -1 \\ -1 & 0 \\ 3 & -1 \end{pmatrix}$, $\mathbf{x} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$.

- a) Finn $\mathbf{y} = B\mathbf{x}$.
- b) Regn ut $C = AB$.
- c) Regn ut både $A\mathbf{y}$ og $C\mathbf{x}$. Sammenlign resultatene.

12. I en barnehage kjører barna sanden frem og tilbake i trillebår mellom tre sandkasser A , B og C (og mister noe på veien!). Av den sanden som er i sandkasse A ved begynnelsen av dagen, vil 70% være i A ved slutten av dagen, 15% vil være i B , 10% i C (og resten vil være mistet). Av den sanden som starter dagen i B , vil 75% være i B ved slutten av dagen, 5% vil være i A og 10% i C . Av den sanden som startet i C , vil 70% være i C ved slutten av dagen, 10% vil være i A og 15% i B .

- a) Finn en matrise M slik at hvis x, y, z er henholdsvis antall liter sand i A , B og C ved begynnelsen av dagen, og $\mathbf{u} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$, så angir komponentene til $\mathbf{v} = M\mathbf{u}$ hvor mange liter sand det er i hver kasse ved slutten av dagen.
- b) Etter at barna har gått hjem, forsøker personalet å fordele sanden på nytt. De flytter 20% av sanden i B til A og 5% av sanden i C til B . Finn en matrise N slik at komponentene til $\mathbf{w} = N\mathbf{v}$ angir hvor mye sand det nå er i hver sandkasse.
- c) Regn ut $K = NM$. Anta at fordeling ved begynnelsen av dagen var $x = 200$, $y = 300$, $z = 400$. Finn fordelingen på slutten av dagen etter at personalet har omfordelt sanden.
- d) Hvor mye sand vil det være i hver kasse på slutten av neste dag dersom den forløper på tilsvarende måte?

13. Bevis punktene (ii), (iii) og (iv) i setning 1.6.2.

1.7 Identitetsmatriser og inverse matriser

I denne seksjonen skal vi bare arbeide med *kvadratiske* matriser, dvs. matriser med like mange rader som søyler. Matrisene vil altså være $n \times n$ -matriser for et helt tall n . En spesiell slik matrise er *identitetsmatrisen*

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

som har 1-ere på diagonalen og 0-er overalt ellers. Ganger du en annen $n \times n$ -matrise A med I_n (gjør det!), ser du at

$$AI_n = A \quad \text{og} \quad I_n A = A$$

Uansett om vi multipliserer A med I_n fra høyre eller venstre, får vi altså A tilbake. Blant tall er det bare 1 som har en tilsvarende egenskap; ganger vi et tall med 1, får vi tallet tilbake. Identitetsmatrisen I_n spiller derfor mye av den samme rollen for matrisemultiplikasjon som 1 gjør for vanlig multiplikasjon.

Et tall $a \neq 0$ har alltid et invertert tall a^{-1} . Dette tallet er definert ved at produktet av a og a^{-1} er lik 1, dvs. $aa^{-1} = 1$. Vi kan på tilsvarende måte definere inverse matriser:

Definisjon 1.7.1 Anta at A er en $n \times n$ -matrise. En $n \times n$ -matrise X kalles en invers matrise til A dersom

$$AX = XA = I_n$$

(siden matrisemultiplikasjon ikke er kommutativ, krever vi å få I_n som svar uansett hvilken side vi multipliserer fra).

Som allerede nevnt har alle tall bortsett fra 0 en invers. For matriser er det mer komplisert; det finnes mange matriser som ikke har invers! Vi skal komme tilbake til spørsmålet om når en matrise har en invers etter hvert, men foreløpig nøyer vi oss med å vise noe enklere — nemlig at ingen matrise har mer enn én invers.

Setning 1.7.2 En $n \times n$ -matrise har høyst én invers.

Bevis: Anta at både X og Y er inverser til A . Da har vi

$$X = I_n X = (YA)X = Y(AX) = YI_n = Y$$

Altså er de to inversene like. ♣

Nå som vi vet at det finnes høyst én invers, kan vi være mer konkrete i språkbruken.

Definisjon 1.7.3 En $n \times n$ -matrise A kalles inverterbar dersom den har en invers, og den inverse matrisen betegnes med A^{-1} . En matrise som ikke er inverterbar, kalles singulær.

Selv om vi ikke skal utvikle noen teori for hvordan man finner inverse matriser på det nåværende tidspunkt, er det instruktivt å se på noen enkle eksempler.

Eksempel 1: Finn den inverse matrisen til

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

Vi må finne en 2×2 -matrise $X = \begin{pmatrix} x & y \\ z & u \end{pmatrix}$ slik at $AX = I_2$ og $XA = I_2$.

Siden

$$AX = \begin{pmatrix} 3x - z & 3y - u \\ x + 2z & y + 2u \end{pmatrix}$$

kan ligningen $AX = I_2$ skrives

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

Dette gir oss fire ligninger med fire ukjente

$$3x - z = 1$$

$$3y - u = 0$$

$$x + 2z = 0$$

$$y + 2u = 1$$

Dette ligningssystemet er lett å løse (legg merke til at vi har to ligninger som bare inneholder x og z , og to som bare inneholder y og u), og vi får $x = \frac{2}{7}$, $y = \frac{1}{7}$, $z = -\frac{1}{7}$ og $u = \frac{3}{7}$. Dette betyr at

$$X = \begin{pmatrix} \frac{2}{7} & \frac{1}{7} \\ -\frac{1}{7} & \frac{3}{7} \end{pmatrix}$$

tilfredsstiller ligningen $AX = I_2$. Vi må også sjekke at X tilfredsstiller det andre kravet til en invers matrise, nemlig at $XA = I_2$:

$$\begin{aligned} XA &= \begin{pmatrix} \frac{2}{7} & \frac{1}{7} \\ -\frac{1}{7} & \frac{3}{7} \end{pmatrix} \begin{pmatrix} 3 & -1 \\ 1 & 2 \end{pmatrix} = \\ &= \begin{pmatrix} \frac{2}{7} \cdot 3 + \frac{1}{7} \cdot 1 & \frac{2}{7} \cdot (-1) + \frac{1}{7} \cdot 2 \\ -\frac{1}{7} \cdot 3 + \frac{3}{7} \cdot 1 & -\frac{1}{7} \cdot (-1) + \frac{3}{7} \cdot 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \end{aligned}$$

Siden vi nå har vist at A er inverterbar og at

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

er oppgaven vår fullført. ♣

Bemerkning: Det kan virke som vi har flaks på slutten av eksemplet ovenfor når vi sjekker at $XA = I_2$ — vi har laget X slik at $AX = I_2$, men siden matrisemultiplikasjon ikke er kommutativ, er det ingen grunn til å tro at dette vil medføre at $XA = I_2$. Det viser seg imidlertid at det alltid er slik; dersom $n \times n$ -matrisen X er en *ensidig invers* til $n \times n$ -matrisen A (dvs. at vi enten har $AX = I_n$ eller $XA = I_n$), så er A inverterbar og $X = A^{-1}$. Det er imidlertid forbausende vanskelig å bevise dette, og vi må utsette det til en senere anledning.

La oss nå se på en matrise som *ikke* har en invers:

Eksempel 2: La

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

Vi forsøker å finne en invers matrise

$$X = \begin{pmatrix} x & y \\ z & u \end{pmatrix}$$

på samme måte som i forrige eksempel, nemlig ved å kreve at $AX = I_2$. Siden

$$AX = \begin{pmatrix} x + 2z & y + 2u \\ 2x + 4z & 2y + 4u \end{pmatrix}$$

gir dette ligningssystemet

$$\begin{aligned} x + 2z &= 1 \\ y + 2u &= 0 \\ 2x + 4z &= 0 \\ 2y + 4u &= 1 \end{aligned}$$

Det er lett å se at dette ligningssystemet ikke har løsninger; ganger vi den første ligningen med 2, får vi $2x + 4z = 2$ som er i åpenbar strid med den tredje ligningen. ♣

De to eksemplene ovenfor viser at det er en nær sammenheng mellom invertering av matriser og lineære ligningssystemer. Vi skal komme tilbake til denne sammenheng senere når vi skal utvikle effektive metoder for å finne inverse matriser. Foreløpig nøyer vi oss med å se på noen enkle regneregler for inverse matriser:

Setning 1.7.4 Anta at A og B er inverterbare $n \times n$ -matriser. Da er

- (i) sA inverterbar for alle tall $s \neq 0$, og $(sA)^{-1} = s^{-1}A^{-1}$
- (ii) AB inverterbar, og $(AB)^{-1} = B^{-1}A^{-1}$
- (iii) A^T inverterbar, og $(A^T)^{-1} = (A^{-1})^T$.
- (iv) A^{-1} er inverterbar og $(A^{-1})^{-1} = A$.

Bevis: Vi konsentrerer oss om (ii), og overlater de (enklere!) (i), (iii) og (iv) til leserne. For å bevise (ii) må vi sjekke at

$$(AB)(B^{-1}A^{-1}) = I_n \quad \text{og} \quad (B^{-1}A^{-1})(AB) = I_n$$

Dette er en liten herjing i parentesflytting (husk setning 1.6.2(i)):

$$\begin{aligned} (AB)(B^{-1}A^{-1}) &= ((AB)B^{-1})A^{-1} = (A(BB^{-1}))A^{-1} = \\ &= (AI_n)A^{-1} = AA^{-1} = I_n \end{aligned}$$

Helt tilsvarende får vi:

$$\begin{aligned} (B^{-1}A^{-1})(AB) &= B^{-1}(A^{-1}(AB)) = B^{-1}((A^{-1}A)B) = \\ &= B^{-1}(I_n B) = B^{-1}B = I_n \end{aligned}$$

□

Bemerkning: Legg merke til at vi ikke har noen regler for de inverse til $A + B$ og $A - B$. Disse matrisene behøver ikke være inverterbare selv om A og B er det, og selv i de tilfellene hvor de er inverterbare, finnes det ikke noen enkel måte å finne den inverse på.

Dersom en matrise A er inverterbar, kan vi forkorte den bort i ligninger av typen $AX = AB$ og få $X = B$. Grunnen er at vi kan multiplisere fra venstre med A^{-1} på begge sider av ligningen:

$$\begin{aligned} AX = AB &\implies A^{-1}(AX) = A^{-1}(AB) \implies (A^{-1}A)X = (A^{-1}A)B \implies \\ &\implies I_n X = I_n B \implies X = B \end{aligned}$$

På tilsvarende måte kan vi forkorte med A i ligningen $XA = BA$ og få $X = B$ (i dette tilfellet må vi gange ligningen fra høyre med A^{-1}). Vi kan imidlertid *ikke* forkorte med A i ligningen $AX = BA$ — uansett om vi ganger med A^{-1} fra venstre eller høyre, vil det være en A vi ikke greier å forkorte bort.

I forrige seksjon tenkte vi på matriser som transformasjoner; når vi ganger en vektor \mathbf{x} med en matrise A , transformerer vi \mathbf{x} til en ny vektor $\mathbf{y} = A\mathbf{x}$.

Dersom A er inverterbar, kan vi gange den siste ligningen fra venstre med den inverse matrisen A^{-1} . Vi får da

$$A^{-1}\mathbf{y} = A^{-1}(A\mathbf{x}) = (A^{-1}A)\mathbf{x} = I_n\mathbf{x} = \mathbf{x}$$

Dette viser at A^{-1} er den omvendte transformasjonen til A — hvis A transformerer \mathbf{x} til \mathbf{y} , vil A^{-1} transformere \mathbf{y} til \mathbf{x} . Denne ideen om A^{-1} som den omvendte transformasjonen til A er viktig når man bruker matriser i praksis.

$$\mathbf{x} = A^{-1}\mathbf{y} \begin{array}{c} \xrightarrow{A} \\ \xleftarrow{A^{-1}} \end{array} \mathbf{y} = A\mathbf{x}$$

Siden det finnes mange transformasjoner vi ikke kan vende om på, gir dette bildet oss en bedre forståelse av hvorfor det finnes matriser som ikke er inverterbare. Hvis det f.eks. finnes to forskjellige vektorer \mathbf{x} og \mathbf{x}' som transformeres til den samme vektoren \mathbf{y} (dvs. vi har både $\mathbf{y} = A\mathbf{x}$ og $\mathbf{y} = A\mathbf{x}'$), så kan ikke transformasjonen vendes om; vi kan ikke starte med \mathbf{y} og transformere den til både \mathbf{x} og \mathbf{x}' !

Oppgaver til seksjon 1.7

1. La $A = \begin{pmatrix} 1 & 3 \\ -2 & 1 \end{pmatrix}$. Kontroller at $AI_2 = A$ og $I_2A = A$ ved å gjennomføre utregningene.

2. La $A = \begin{pmatrix} 1 & 0 & -2 \\ -3 & 1 & 4 \\ 2 & -3 & 4 \end{pmatrix}$. Vis at $B = \begin{pmatrix} 8 & 3 & 1 \\ 10 & 4 & 1 \\ \frac{7}{2} & \frac{3}{2} & \frac{1}{2} \end{pmatrix}$ er den inverse matrisen til A ved å regne ut AB og BA .

3. Avgjør om følgende matriser er inverterbare eller singulære:

$$A = \begin{pmatrix} 3 & -7 \\ 1 & -2 \end{pmatrix}, B = \begin{pmatrix} 5 & 0 \\ 0 & 4 \end{pmatrix}, C = \begin{pmatrix} 2 & 0 \\ 6 & 0 \end{pmatrix} \text{ og } D = \begin{pmatrix} 2 & 4 \\ 3 & 4 \end{pmatrix}$$

4. En inverterbar matrise A har en invers matrise som er gitt ved $A^{-1} = \begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix}$. Finn matrisen A .

5. Gitt to inverterbare matriser A og B , hvor $A^{-1} = \begin{pmatrix} 1 & 4 \\ 2 & 9 \end{pmatrix}$ og $B^{-1} = \begin{pmatrix} 8 & 3 \\ 2 & 1 \end{pmatrix}$. Finn $(AB)^{-1}$.

6. a) La $\mathbf{a} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$. Vis at $I_2\mathbf{a} = \mathbf{a}$.

b) La $\mathbf{b} = \begin{pmatrix} -3 \\ 2 \\ -1 \end{pmatrix}$. Vis at $I_3 \mathbf{b} = \mathbf{b}$.

c) Vis at $I_n \mathbf{c} = \mathbf{c}$ for alle $\mathbf{c} \in \mathbb{R}^n$.

7. a) Anta at $A = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$ der $a, b \neq 0$. Vis at A er inverterbar og at $A^{-1} = \begin{pmatrix} a^{-1} & 0 \\ 0 & b^{-1} \end{pmatrix}$.

b) Anta at $B = \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix}$ der $a, b, c \neq 0$. Vis at B er inverterbar og at $B^{-1} = \begin{pmatrix} a^{-1} & 0 & 0 \\ 0 & b^{-1} & 0 \\ 0 & 0 & c^{-1} \end{pmatrix}$.

c) Formuler et tilsvarende resultat for $n \times n$ -matriser.

8. Vis at dersom A og B er inverterbare, så er den inverse til $(AB)^T$ lik $(A^{-1})^T(B^{-1})^T$.

9. Bevis punktene (i), (iii) og (iv) i setning 1.6.4.

10. a) Vis at 2×2 -matrisen $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ er inverterbar hvis og bare hvis $ad - bc \neq 0$, og at den inverse matrisen i så fall er gitt ved $A^{-1} = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$.

b) Bruk formelen fra punkt a) til å finne den inverse til matrisen $A = \begin{pmatrix} 2 & -5 \\ -1 & 3 \end{pmatrix}$.

c) Bruk matrisen du fant i punkt b) til å løse ligningssystemet

$$2x - 5y = 3$$

$$-x + 3y = 2$$

Hint: Systemet kan skrives $\begin{pmatrix} 2 & -5 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$.

1.8 Determinanter, arealer og volumer

Til enhver $n \times n$ -matrise

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

hører det et tall som kalles *determinanten* til A , og som betegnes med

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

Generelle $n \times n$ -determinanter kan være litt vanskelige å forstå seg på, så vi skal derfor begynne med 2×2 - og 3×3 -determinanter og deres geometriske egenskaper.

Hvis A er en 2×2 -matrise

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

så er *determinanten* til A definert ved

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

Legg merke til at uttrykket $ad - bc$ fremkommer fra diagonalene i matrisen $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$; ganger vi sammen tallene i den ene diagonalen, får vi ad , og ganger vi sammen tallene i den andre diagonalen, får vi bc .

La oss regne ut en determinant.

Eksempel 1: Vi ser at

$$\begin{vmatrix} -3 & -5 \\ 4 & 2 \end{vmatrix} = (-3) \cdot 2 - (-5) \cdot 4 = -6 + 20 = 14$$



Dersom vi har to vektorer $\mathbf{a} = (a_1, a_2)$, $\mathbf{b} = (b_1, b_2)$, kan vi lage en 2×2 -determinant $\det(\mathbf{a}, \mathbf{b})$ ved å legge inn vektorene som radene i en matrise på denne måten:

$$\det(\mathbf{a}, \mathbf{b}) = \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} = a_1 b_2 - a_2 b_1$$

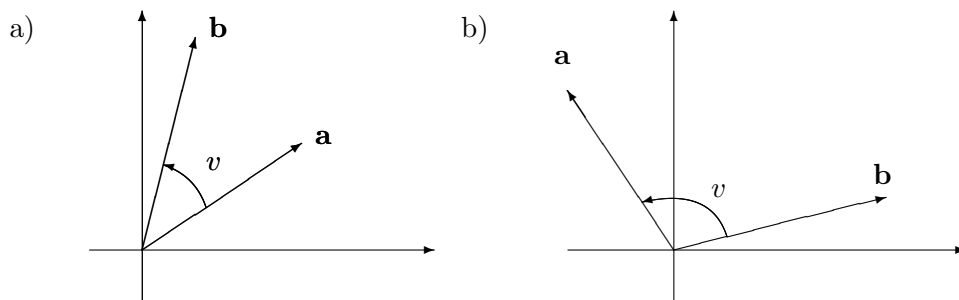
Legg merke til at dersom vi bytter om rekkefølgen på vektorene \mathbf{a} og \mathbf{b} , så skifter determinanten fortegn:

$$\det(\mathbf{b}, \mathbf{a}) = \begin{vmatrix} b_1 & b_2 \\ a_1 & a_2 \end{vmatrix} = b_1 a_2 - b_2 a_1 = -(a_1 b_2 - a_2 b_1) = -\det(\mathbf{a}, \mathbf{b})$$

Som vi snart skal se, har både fortegnet og størrelsen til $\det(\mathbf{a}, \mathbf{b})$ en geometrisk betydning.

Bemerkning: Man kan lure på hvorfor vi legger inn vektorene \mathbf{a} og \mathbf{b} som rader og ikke som søyler. Det viser seg at dette ikke spiller noen rolle siden $\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}$, men vi har valgt å ta utgangspunkt i rader siden det passer best til anvendelsene i senere kapitler (for andre anvendelser hadde det passet vel så bra å bruke søyler!).

For å forstå den geometriske tolkningen av 2×2 -determinanter, er det nyttig å vite litt om orientering av vektorpar. To vektorer \mathbf{a} og \mathbf{b} bestemmer en vinkel v mellom 0° og 180° som vist på figur 1. Vi kaller dette *vinkelen mellom \mathbf{a} og \mathbf{b}* . Legg merke til at dersom vi beveger oss i positiv omløpsretning, vil denne vinkelen noen ganger starte i \mathbf{a} og ende i \mathbf{b} (se figur 1a) og andre ganger starte i \mathbf{b} og ende i \mathbf{a} (se figur 1b). I det første tilfellet sier vi at paret (\mathbf{a}, \mathbf{b}) er *positivt orientert*, i det andre tilfellet at det er *negativt orientert*. Her er åpenbart rekkefølgen til vektorene viktig — \mathbf{a} er første vektor og \mathbf{b} er andre vektor. Bytter vi om rekkefølgen av vektorene, bytter vi også orientering.

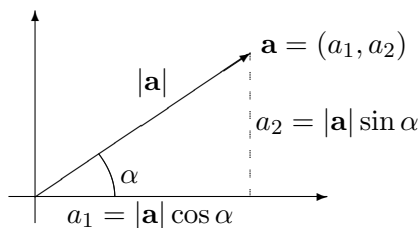


Figur 1: Vinkelen v mellom \mathbf{a} og \mathbf{b}

For å studere den geometriske tolkningen av determinanten

$$\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}$$

lønner det seg å skrive vektorene $\mathbf{a} = (a_1, a_2)$ og $\mathbf{b} = (b_1, b_2)$ på polarform.



Figur 2: Vektoren \mathbf{a} på polarform

Lar vi α være vinkelen fra den positive x -aksen til vektoren \mathbf{a} , så er

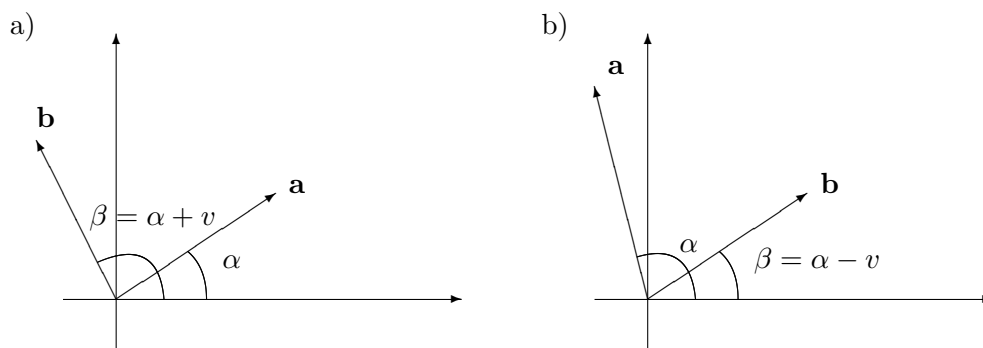
$$\mathbf{a} = (|\mathbf{a}| \cos \alpha, |\mathbf{a}| \sin \alpha)$$

(se figur 2 og husk det du har lært om polarform til komplekse tall). På tilsvarende måte er

$$\mathbf{b} = (|\mathbf{b}| \cos \beta, |\mathbf{b}| \sin \beta)$$

der β er vinkelen fra den positive x -aksen til vektoren \mathbf{b} .

Vi lar fortsatt v være vinkelen mellom vektorene \mathbf{a} og \mathbf{b} som på figur 1. Dersom paret (\mathbf{a}, \mathbf{b}) er positivt orientert, ser vi fra figur 3a at $\beta = \alpha + v$. Dersom paret (\mathbf{a}, \mathbf{b}) er negativt orientert, ser vi fra figur 3b at vi $\beta = \alpha - v$.



Figur 3: Sammenhengen $\beta = \alpha \pm v$

Vi har altså

$$\beta = \alpha \pm v$$

der fortegnet er pluss eller minus ettersom paret (\mathbf{a}, \mathbf{b}) er positivt eller negativt orientert.

La oss nå finne determinanten uttrykt ved α og β :

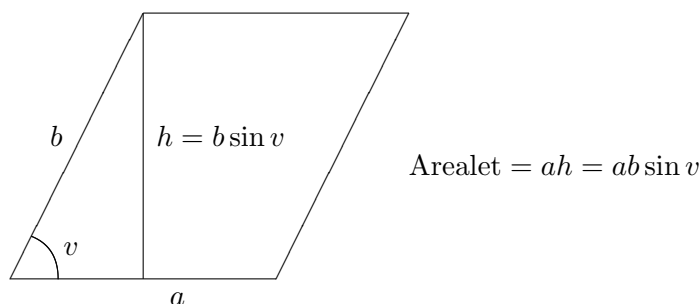
$$\begin{aligned} \det(\mathbf{a}, \mathbf{b}) &= \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} = \begin{vmatrix} |\mathbf{a}| \cos \alpha & |\mathbf{a}| \sin \alpha \\ |\mathbf{b}| \cos \beta & |\mathbf{b}| \sin \beta \end{vmatrix} \\ &= |\mathbf{a}| \cos \alpha |\mathbf{b}| \sin \beta - |\mathbf{a}| \sin \alpha |\mathbf{b}| \cos \beta \\ &= |\mathbf{a}| |\mathbf{b}| (\sin \beta \cos \alpha - \cos \beta \sin \alpha) \\ &= |\mathbf{a}| |\mathbf{b}| \sin(\beta - \alpha) \end{aligned}$$

der vi i den siste overgangen har brukt formelen for sinus til en differens. Siden $\beta = \alpha \pm v$, får vi

$$\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} = |\mathbf{a}| |\mathbf{b}| \sin(\pm v) = \pm |\mathbf{a}| |\mathbf{b}| \sin v$$

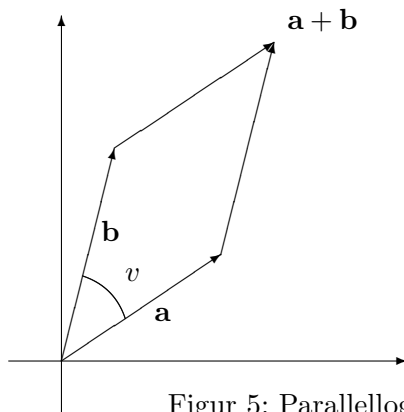
Siden $\sin v$ aldri er negativ (v ligger per definisjon i intervallet $[0^\circ, 180^\circ]$ der sinus er positiv), vil $\det(\mathbf{a}, \mathbf{b})$ altså være positiv dersom paret (\mathbf{a}, \mathbf{b})

er positivt orientert, og negativ dersom dette paret er negativt orientert. Fortegnet til determinanten $\det(\mathbf{a}, \mathbf{b})$ gjenspeiler altså orienteringen til paret (\mathbf{a}, \mathbf{b}) . Legg forøvrig merke til at $\det(\mathbf{a}, \mathbf{b})$ er 0 dersom v er 0° eller 180° , det vil si når \mathbf{a} og \mathbf{b} er parallelle.



Figur 4: Arealet til et parallelogram

Etter at vi nå har funnet ut hva fortegnet til determinanten betyr, er det på tide å se på absoluttverdien. Aller først repeterer vi formelen for arealet til et parallelogram. Som det fremgår fra figur 4, er dette arealet gitt ved $A = ab \sin v$, der a og b er lengdene til sidene, og der v er vinkelen mellom dem (på figuren er vinkel v spiss, men det er lett å se at resultatet også holder dersom vinkelen er stump).



Figur 5: Parallelogrammet utspent av \mathbf{a} og \mathbf{b}

Arealet til parallelogrammet utspent av vektorene $\mathbf{a} = (a_1, a_2)$ og $\mathbf{b} = (b_1, b_2)$ (se figur 5) er derfor lik $|\mathbf{a}||\mathbf{b}|\sin v = \pm \det(\mathbf{a}, \mathbf{b})$ der fortegnet er pluss eller minus ettersom (\mathbf{a}, \mathbf{b}) er positivt eller negativt orientert. Det betyr at arealet er lik tallverdien til determinanten. La oss oppsummere resultatene.

Setning 1.8.1 Determinanten

$$\det(\mathbf{a}, \mathbf{b}) = \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}$$

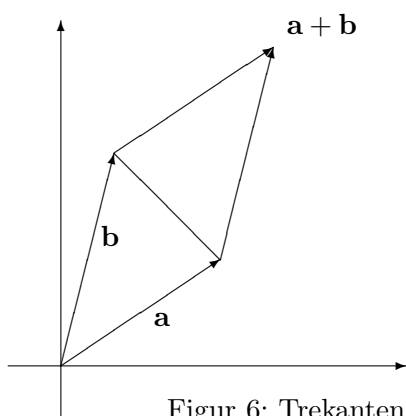
er positiv dersom vektorparet (\mathbf{a}, \mathbf{b}) er positivt orientert og negativ dersom paret er negativt orientert. Arealet til parallellogrammet utspent av \mathbf{a} og \mathbf{b} er lik tallverdien til determinanten.

Bemerkning: Matematikere sier at determinanten gir oss arealet med fortegn (eller orientering). Det kan virke merkelig å knytte fortegn til areal, men spesielt når man skal studere arealet til flater, viser det seg viktig å holde styr på retningen — det er i mange sammenhenger viktig å vite hva man skal regne som flatens “overside/underside” eller “utside/innside”. Som vi skal se senere i denne seksjonen, kan sammenhengen mellom determinant og “areal med fortegn” generaliseres til tre dimensjoner.

Eksempel 2: Finn arealet utspent av vektorene $\mathbf{a} = (3, -7)$ og $\mathbf{b} = (-4, 5)$. Vi får

$$\det(\mathbf{a}, \mathbf{b}) = \begin{vmatrix} 3 & -7 \\ -4 & 5 \end{vmatrix} = 3 \cdot 5 - (-7) \cdot (-4) = 15 - 28 = -13$$

Arealet er dermed $|-13| = 13$. Siden $\det(\mathbf{a}, \mathbf{b})$ er negativ, er paret (\mathbf{a}, \mathbf{b}) negativt orientert, dvs. at vinkelen fra \mathbf{a} til \mathbf{b} er større enn 180° . ♣



Figur 6: Trekanten med sider \mathbf{a} og \mathbf{b}

Determinanter kan også brukes til å regne ut arealet til trekanter. Arealet til trekanten med sider \mathbf{a} og \mathbf{b} er halvparten av arealet til parallellogrammet utspent av disse vektorene (se figur 6).

Vi har derfor følgende resultat:

Korollar 1.8.2 Arealet til trekanten med sider \mathbf{a} og \mathbf{b} er $\frac{1}{2} \cdot |\det(\mathbf{a}, \mathbf{b})|$

Eksempel 3: Finn arealet til trekanten med hjørner i punktene $\mathbf{c} = (-1, 2)$, $\mathbf{d} = (4, 8)$ og $\mathbf{e} = (2, -3)$. Vi regner ut

$$\mathbf{a} = \mathbf{d} - \mathbf{c} = (4, 8) - (-1, 2) = (5, 6)$$

$$\mathbf{b} = \mathbf{e} - \mathbf{c} = (2, -3) - (-1, 2) = (3, -5)$$

Trekanten vi er på jakt etter, har samme areal som trekanten med sider \mathbf{a} og \mathbf{b} (hvorfor?). Dermed er

$$\text{Areal} = \frac{1}{2} \cdot |\det(\mathbf{a}, \mathbf{b})| = \frac{1}{2} \cdot \begin{vmatrix} 5 & 6 \\ 3 & -5 \end{vmatrix} = \frac{1}{2} \cdot |-25 - 18| = \frac{43}{2} \quad \clubsuit$$

Før vi går over til 3×3 -determinanter, tar vi med et resultat til. Dette resultatet kan virke litt underlig på det nåværende tidspunkt, men det skal bli en viktig inspirasjonskilde når vi studerer generelle $n \times n$ -determinanter i et senere kapittel.

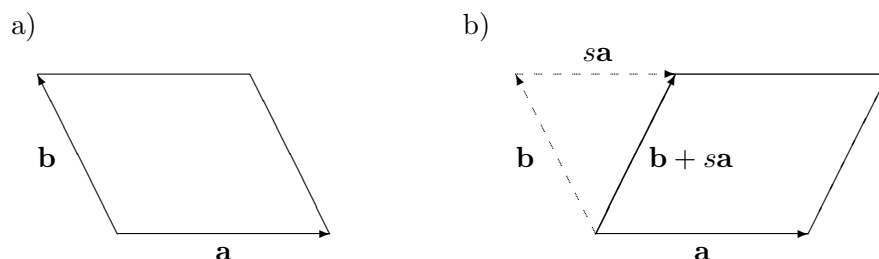
Setning 1.8.3 For 2×2 -matriser gjelder:

- (i) $\det(I_2) = 1$ (husk at I_2 er identitetsmatrisen $I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$).
- (ii) Dersom vi bytter om to rader, så bytter determinanten fortegn (dvs. $\det(\mathbf{a}, \mathbf{b}) = -\det(\mathbf{b}, \mathbf{a})$).
- (iii) Dersom vi ganger alle elementene i en rad med et tall s , så forandrer også matrisen seg med en faktor s (dvs. $\det(s\mathbf{a}, \mathbf{b}) = s \det(\mathbf{a}, \mathbf{b})$ og $\det(\mathbf{a}, s\mathbf{b}) = s \det(\mathbf{a}, \mathbf{b})$).
- (iv) Dersom vi adderer et tall ganger en rad til en av de andre radene, endrer ikke determinanten verdi (dvs. $\det(\mathbf{a} + s\mathbf{b}, \mathbf{b}) = \det(\mathbf{a}, \mathbf{b})$ og $\det(\mathbf{a}, \mathbf{b} + s\mathbf{a}) = \det(\mathbf{a}, \mathbf{b})$).

Bevis: Alle punktene kan vises ved direkte utregning, men vi vil gjerne forstå dem geometrisk selv om det i noen tilfeller er litt mer omstendelig:

- (i) Parallelogrammet utspent av radvektorene $\mathbf{a} = (1, 0)$ og $\mathbf{b} = (0, 1)$ er et kvadrat med side 1. Arealet er opplagt 1, og siden paret (\mathbf{a}, \mathbf{b}) er positivt orientert, er $\det(I_2) = 1$.
- (ii) Bytter vi om på radene, bytter vi orientering på paret (\mathbf{a}, \mathbf{b}) , og determinanten bytter dermed fortegn.
- (iii) Ganger vi den ene vektoren med et positivt tall s , blir enten grunnlinjen eller høyden i det utspente parallelogrammet ganget med s , og arealet øker/avtar derfor med en faktor s . Ganger vi med et negativt tall, endres høyden eller grunnlinjen med en faktor $|s|$, men i tillegg bytter vektorparet orientering slik at også i dette tilfellet endrer determinanten seg med en faktor s .

(iv) Her trenger vi en liten figur. Figur 7a viser parallelogrammet utspent av \mathbf{a} og \mathbf{b} , mens figur 7b viser parallelogrammet utspent av \mathbf{a} og $\mathbf{b} + s\mathbf{a}$. De to parallelogrammene har samme grunnlinje og høyde, og derfor samme areal. Det er også lett å se at uansett hvor stor s er, så har parene (\mathbf{a}, \mathbf{b}) og $(\mathbf{a}, \mathbf{b} + s\mathbf{a})$ samme orientering. Altså er $\det(\mathbf{a}, \mathbf{b}) = \det(\mathbf{a}, \mathbf{b} + s\mathbf{a})$. \square



Figur 7: Parallelogrammene utspent av henholdsvis \mathbf{a}, \mathbf{b} og $\mathbf{a}, \mathbf{b} + s\mathbf{a}$

3×3 -determinanter

Determinanten til en 3×3 -matrise

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

er definert ved

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

der 2×2 -determinantene på høyre side regnes ut på vanlig måte. Legg merke til hvordan disse 2×2 -determinantene fremkommer fra den opprinnelige determinanten — for å finne den 2×2 -determinanten som ganges med a_{11} , stryker vi den linjen og den søylen som går gjennom a_{11} (se figur 8), for å finne den 2×2 -determinanten som ganges med a_{12} , stryker vi den linjen og den søylen som går gjennom a_{12} , osv. Legg også merke til at fortegnene til leddene på høyre side veksler mellom pluss og minus.

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

Figur 8: 2×2 -determinanten som skal ganges med a_{11}

La oss regne ut en 3×3 -determinant.

Eksempel 4: Regn ut

$$\begin{vmatrix} 2 & 3 & -1 \\ 5 & -4 & 0 \\ -3 & 1 & 2 \end{vmatrix} = 2 \begin{vmatrix} -4 & 0 \\ 1 & 2 \end{vmatrix} - 3 \begin{vmatrix} 5 & 0 \\ -3 & 2 \end{vmatrix} + (-1) \begin{vmatrix} 5 & -4 \\ -3 & 1 \end{vmatrix} \\ = 2((-4) \cdot 2 - 0 \cdot 1) - 3(5 \cdot 2 - 0 \cdot (-3)) + (-1)(5 \cdot 1 - (-4)(-3)) \\ = 2(-8) - 3 \cdot 10 + (-1)(-7) = -16 - 30 + 7 = -39$$



Det er en nær sammenheng mellom 3×3 -determinanter og kryssproduktet. Som et første eksempel har vi følgende huskeregel for kryssproduktet:

$$\begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} =$$

$$= (a_2b_3 - a_3b_2)\mathbf{i} + (a_3b_1 - a_1b_3)\mathbf{j} + (a_1b_2 - a_2b_1)\mathbf{k} = \mathbf{a} \times \mathbf{b}$$

Siden vi bare har definert determinanten når elementene i første rad er tall (og ikke vektorer), gir det første skrittet i denne utregningen egentlig ikke mening, men resultatet er likevel en grei huskeregel.

Vi har tidligere sett at 2×2 -determinanter kan brukes til å regne ut arealer og til å bestemme orienteringen til vektorpar (\mathbf{a}, \mathbf{b}) . På tilsvarende måte kan vi bruke 3×3 -determinanter til å regne ut volumer og til å bestemme orienteringen til vektortripler $(\mathbf{a}, \mathbf{b}, \mathbf{c})$. Før vi begynner, kan det være greit å bli enig om notasjonen. Dersom $\mathbf{a} = (a_1, a_2, a_3)$, $\mathbf{b} = (b_1, b_2, b_3)$, $\mathbf{c} = (c_1, c_2, c_3)$, skriver vi

$$\det(\mathbf{a}, \mathbf{b}, \mathbf{c}) = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

Vi observerer så at

$$\det(\mathbf{a}, \mathbf{b}, \mathbf{c}) = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = a_1 \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} - a_2 \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} + a_3 \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix} \\ = a_1(b_2c_3 - b_3c_2) - a_2(b_1c_3 - b_3c_1) + a_3(b_1c_2 - b_2c_1) = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$$

Sammenholder vi dette med setning 1.6.4 og korollar 1.6.5, får vi:

Setning 1.8.4 *Volumet av parallellepipedet utspent av vektorene $\mathbf{a}, \mathbf{b}, \mathbf{c}$ er $|\det(\mathbf{a}, \mathbf{b}, \mathbf{c})|$. Volumet av pyramiden utspent av $\mathbf{a}, \mathbf{b}, \mathbf{c}$ er $\frac{1}{6}|\det(\mathbf{a}, \mathbf{b}, \mathbf{c})|$.*

Legg merke til at $\det(\mathbf{a}, \mathbf{b}, \mathbf{c}) = 0$ hvis volumet til parallellepipedet er 0. Det skjer hvis vektorene \mathbf{a} , \mathbf{b} og \mathbf{c} ligger i samme plan gjennom origo. På det nåværende tidspunkt kan dette virke som en uvesentlig observasjon, men det viser seg faktisk å være en av hovedårsakene til determinantenes betydning.

Hva så med orienteringen? Først må vi definere når et trippel $(\mathbf{a}, \mathbf{b}, \mathbf{c})$ er positivt og negativt orientert: To ikke-parallele vektorer \mathbf{a} og \mathbf{b} definerer sammen med origo et plan (planet gjennom punktene $\mathbf{0}$, \mathbf{a} og \mathbf{b}). Dette planet deler rommet i to halvdelar. Dersom \mathbf{c} ligger på samme side av planet som kryssproduktet $\mathbf{a} \times \mathbf{b}$, sier vi at tripplet $(\mathbf{a}, \mathbf{b}, \mathbf{c})$ er *positivt orientert*. Dersom \mathbf{c} ligger på den andre siden av planet, sier vi at tripplet er *negativt orientert*. Bruker vi den geometriske tolkningen av skalarproduktet, ser vi at tripplet $(\mathbf{a}, \mathbf{b}, \mathbf{c})$ er positivt orientert hvis og bare hvis $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}$ er positiv (for da er vinkelen mellom \mathbf{c} og $\mathbf{a} \times \mathbf{b}$ mindre enn 90°). Det er lett (men ikke særlig spennende) å sjekke at $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = \det(\mathbf{a}, \mathbf{b}, \mathbf{c})$ (det er ikke noe mystisk i dette — både $|(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}|$ og $|\det(\mathbf{a}, \mathbf{b}, \mathbf{c})|$ er lik volumet til parallellepipedet utspent av \mathbf{a} , \mathbf{b} og \mathbf{c} , så alt vi sjekker er at fortegnet er det samme). Dette betyr at $(\mathbf{a}, \mathbf{b}, \mathbf{c})$ er positivt orientert hvis og bare hvis $\det(\mathbf{a}, \mathbf{b}, \mathbf{c})$ er positiv. Vi har altså den samme forbindelsen mellom positiv orientering og positiv determinant som i det to-dimensjonale tilfellet.

Vi tar med en tredimensjonal versjon av setning 1.8.3.

Setning 1.8.5 *For 3×3 -matriser gjelder:*

- (i) $\det(I_3) = 1$ (husk at I_3 er identitetsmatrisen $I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$).
- (ii) Dersom vi bytter om to rader, så bytter determinanten fortegn (det vil f.eks. si at $\det(\mathbf{a}, \mathbf{b}, \mathbf{c}) = -\det(\mathbf{c}, \mathbf{b}, \mathbf{a})$).
- (iii) Dersom vi ganger alle elementene i en rad med et tall s , så forandrer også matrisen seg med en faktor s (det vil f.eks. si at $\det(\mathbf{a}, s\mathbf{b}, \mathbf{c}) = s \det(\mathbf{a}, \mathbf{b}, \mathbf{c})$).
- (iv) Dersom vi adderer et tall ganger en rad til en av de andre radene, endrer ikke determinanten verdi (det vil f.eks. si at $\det(\mathbf{a}, \mathbf{b}, \mathbf{c} + s\mathbf{a}) = \det(\mathbf{a}, \mathbf{b}, \mathbf{c})$).

Bevis: Vi skal ikke gjennomgå punktene i detalj, bare se på hovedideene. Punkt (i) kan du bevise enten ved direkte utregning eller ved å observere at parallelogrammet utspent av vektorene $\mathbf{a} = (1, 0, 0)$, $\mathbf{b} = (0, 1, 0)$ og $\mathbf{c} = (0, 0, 1)$ er en terning med side 1. I punkt (ii) vet vi allerede at tallverdien til determinanten er uforandret om vi bytter om på radene (fordi begge er lik volumet til det samme parallellepipedet), og alt du behøver å sjekke er at orienteringen snur når du bytter om to vektorer (bruk høyrehåndsregelen).

Punkt (iii) følger på samme måte som i det todimensjonale tilfellet; ganger du en av sidekantene i et parallellepiped med $s > 0$, øker også volumet med en faktor s , men ganger du med en faktor $s < 0$, må du også ta hensyn til at orienteringen snur. For å bevise punkt (iv) kan du bruke akkurat samme figur som i det todimensjonale tilfellet (figur 7) — den eneste forskjellen er at det nå inngår en vektor til. Denne vektoren stikker på skrå ut av (eller inn i) papiret og inngår ikke i beregningene på noen forstyrrende måte (det eneste den bidrar med er den felles høyden i parallellogrammene).

$n \times n$ -determinanter

Vi skal ikke studere generelle $n \times n$ -determinanter for alvor i dette kapitlet, men det kan være morsomt å vite hvordan de regnes ut. Gitt en 4×4 -matrise

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix}$$

definerer vi *determinanten* $\det(A)$ ved

$$\begin{aligned} \det(A) = & a_{11} \begin{vmatrix} a_{22} & a_{23} & a_{24} \\ a_{32} & a_{33} & a_{34} \\ a_{42} & a_{43} & a_{44} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} & a_{24} \\ a_{31} & a_{33} & a_{34} \\ a_{41} & a_{43} & a_{44} \end{vmatrix} + \\ & + a_{13} \begin{vmatrix} a_{21} & a_{22} & a_{24} \\ a_{31} & a_{32} & a_{34} \\ a_{41} & a_{42} & a_{44} \end{vmatrix} - a_{14} \begin{vmatrix} a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{vmatrix} \end{aligned}$$

Sammenligner du denne definisjonen med definisjonen av 3×3 -determinanter, vil du oppdage det generelle mønsteret som gjør at vi kan gå videre og definere 5×5 -determinanter, 6×6 -determinanter osv.

Vi har tidligere sett at *tallverdien* til 2×2 -determinanten $\det(\mathbf{a}, \mathbf{b})$ gir oss arealet utspent av vektorene \mathbf{a} og \mathbf{b} , mens *fortegnet* til determinanten forteller oss om orienteringen til paret (\mathbf{a}, \mathbf{b}) . På tilsvarende vis vet vi at tallverdien til en 3×3 -determinant $\det(\mathbf{a}, \mathbf{b}, \mathbf{c})$ gir oss volumet utspent av vektorene \mathbf{a} , \mathbf{b} og \mathbf{c} , mens fortegnet forteller oss om orienteringen til triplet $(\mathbf{a}, \mathbf{b}, \mathbf{c})$. Vi kan bruke disse observasjonene til å definere volum og orientering i høyere dimensjoner. Gitt n vektorer $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ definerer vi *volumet utspent av disse vektorene* til å være tallverdien til $n \times n$ -determinanten $\det(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)$. Vi sier at n -tuplet $(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)$ (legg merke til at dette er et n -tupple av vektorer) er *positivt* (henholdsvis *negativt*) orientert dersom determinanten er positiv (henholdsvis negativ). Vi skal ikke gå nærmere inn på volum og orientering i dette heftet, men vi skal komme tilbake til generelle determinanter i et senere kapittel.

Bemerkning: Som vi skal se i et senere kapittel, er det en nær sammenheng mellom determinanter og inverterbarhet — det viser seg at en kvadratisk matrise A er inverterbar hvis og bare hvis $\det(A) \neq 0$. Dette er kanskje ikke så overraskende siden begge begrepene har med degenerasjon å gjøre: En kvadratisk matrise A er singular (ikke inverterbar), dersom det finnes to forskjellige vektorer \mathbf{x} og \mathbf{x}' slik at $A\mathbf{x} = A\mathbf{x}'$, og den har determinant lik 0 dersom radene ikke spenner ut et volum. Det viser seg at disse to formene for degenerasjon er ekvivalente.

Oppgaver til seksjon 1.8

1. Regn ut determinantene

$$\text{a) } \begin{vmatrix} 4 & 3 \\ -2 & 2 \end{vmatrix} \quad \text{b) } \begin{vmatrix} 8 & -2 \\ 3 & 4 \end{vmatrix} \quad \text{c) } \begin{vmatrix} 6 & -3 \\ -2 & 1 \end{vmatrix}$$

2. Finn arealet til parallelogrammet utspent av $\mathbf{a} = (1, 3)$ og $\mathbf{b} = (4, 1)$.

3. En trekant har hjørner i punktene $(-1, 2)$, $(4, 3)$, $(1, 7)$. Finn arealet.

4. En firkant har hjørner i punktene $(0, 1)$, $(5, 1)$, $(1, 7)$ og $(7, 4)$. Finn arealet.

5. Avgjør om parene (\mathbf{a}, \mathbf{b}) er positivt eller negativt orientert:

$$\text{a) } \mathbf{a} = (3, -1) \quad \mathbf{b} = (-7, 2) \quad \text{b) } \mathbf{a} = (-1, 5) \quad \mathbf{b} = (3, 2)$$

6. Vis at $\det(\mathbf{a}, \mathbf{b}) = 0$ hvis og bare hvis vektorene $\mathbf{a}, \mathbf{b} \in \mathbb{R}^2$ er parallelle eller (minst) én av dem er $\mathbf{0}$.

7. Vis at $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = \begin{vmatrix} a & c \\ b & d \end{vmatrix}$, dvs. at vi får den samme determinanten om vi bytter om linjer og søyler.

8. Alle hjørnene til et parallelogram har heltallige koordinater. Vis at arealet er et helt tall.

9. Anta at $\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \neq 0$

a) Vis at ligningssystemet $a_1x + b_1y = c_1$, $a_2x + b_2y = c_2$ har løsningen

$$x = \frac{\begin{vmatrix} c_1 & b_1 \\ c_2 & b_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}} \quad y = \frac{\begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}}$$

b) Hva skjer med ligningssystemet når $\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = 0$?

10. Regn ut determinantene:

$$\text{a) } \begin{vmatrix} 3 & -2 & -1 \\ 1 & 4 & 3 \\ 2 & 1 & 7 \end{vmatrix} \quad \text{b) } \begin{vmatrix} -2 & 4 & 0 \\ -2 & 3 & 3 \\ 1 & 0 & 4 \end{vmatrix} \quad \text{c) } \begin{vmatrix} 1 & 2 & 3 \\ -2 & 5 & 4 \\ 3 & -3 & -1 \end{vmatrix}$$

11. Finn volumet til parallellepipedet utspent av $(-1, 0, 2)$, $(3, -1, 3)$ og $(4, 0, -1)$.

12. Finn volumet til pyramiden med hjørner i punktene $(2, 2, 2)$, $(-1, 2, 3)$, $(3, 4, 2)$ og $(7, 2, 2)$.

13. Avgjør om triplet $(\mathbf{a}, \mathbf{b}, \mathbf{c})$ er positivt eller negativt orientert når $\mathbf{a} = (-1, 2, 3)$, $\mathbf{b} = (0, 2, 4)$ og $\mathbf{c} = (7, -1, 2)$.

14. Vis at $\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$, dvs. at determinanten er den samme om vi bytter om søyler og linjer.

15. Vis at dersom \mathbf{a} , \mathbf{b} og \mathbf{c} er ortogonale, så er $\det(\mathbf{a}, \mathbf{b}, \mathbf{c}) = |\mathbf{a}||\mathbf{b}||\mathbf{c}|$.

16. Regn ut determinanten til 4×4 -matrisen

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

17. I denne oppgaven er \mathbf{a} , \mathbf{b} , \mathbf{c} og \mathbf{d} tredimensjonale vektorer.

a) Vis at dersom to av vektorene \mathbf{a} , \mathbf{b} , \mathbf{c} er like, så er $\det(\mathbf{a}, \mathbf{b}, \mathbf{c}) = 0$

b) Vis at for alle vektorer \mathbf{a} , \mathbf{b} , \mathbf{c} , \mathbf{d} og alle skalarer s , t gjelder

$$\det(s\mathbf{a} + t\mathbf{d}, \mathbf{b}, \mathbf{c}) = s \det(\mathbf{a}, \mathbf{b}, \mathbf{c}) + t \det(\mathbf{d}, \mathbf{b}, \mathbf{c})$$

c) Vi sier at en vektor \mathbf{a} er en *lineærkombinasjon* av vektorene \mathbf{b} , \mathbf{c} dersom det finnes skalarer s , t slik at $\mathbf{a} = s\mathbf{b} + t\mathbf{c}$. Bruk a) og b) til å vise at dersom \mathbf{a} er en lineærkombinasjon av \mathbf{b} og \mathbf{c} , så er $\det(\mathbf{a}, \mathbf{b}, \mathbf{c}) = 0$.

d) Gi en geometrisk forklaring på resultatet i c).

18. Bevis setning 1.8.3 ved regning (dvs. regn ut begge sider av likhetene og se at de stemmer).

19. Bevis setning 1.8.5 ved regning (dvs. regn ut begge sider av likhetene og se at de stemmer).

20. Vis at en 2×2 -matrise A er inverterbar hvis og bare hvis $\det(A) \neq 0$. (*Hint:* Mesteparten av jobben er gjort i oppgave 1.7.10.)

Kapittel 2

Funksjoner fra \mathbb{R}^n til \mathbb{R}^m

I dette kapitlet skal vi studere funksjoner \mathbf{F} fra \mathbb{R}^n til \mathbb{R}^m . En slik funksjon er bare en regel som til hvert n -tupple $\mathbf{x} = (x_1, x_2, \dots, x_n)$ tilordner et m -tupple $\mathbf{y} = \mathbf{F}(\mathbf{x})$. Dersom $n > 1$ kaller vi \mathbf{F} en *funksjon av flere variable*, og det er slike funksjoner vi skal studere i dette kapitlet. Funksjoner av flere variable har mange likhetstrekk med de funksjonene av én variabel som du kjenner fra før, men de har også en del nye og litt uvante egenskaper. Spesielt blir geometrien mer komplisert når vi får mer enn én variabel å arbeide med.

2.1 Funksjoner av flere variable

Som allerede nevnt er en funksjon \mathbf{F} fra \mathbb{R}^n til \mathbb{R}^m bare en regel som til hver $\mathbf{x} \in \mathbb{R}^n$ gir oss en $\mathbf{y} = \mathbf{F}(\mathbf{x})$ i \mathbb{R}^m . Ofte er disse reglene gitt ved en formel, f.eks. kan en funksjon fra \mathbb{R}^4 til \mathbb{R}^3 være gitt ved

$$\mathbf{F}(x, y, z, u) = (2x^2z, xy^2z, ye^{x^2y+\sin y})$$

Gitt et 4-tupple $\mathbf{x} = (x, y, z, u)$ forteller denne formelen oss hvordan vi kan regne ut 3-tuplet $\mathbf{F}(\mathbf{x})$. Som du ser, veksler vi mellom skrivemåtene $\mathbf{F}(\mathbf{x})$ og $\mathbf{F}(x_1, x_2, \dots, x_n)$; den første er ofte mest praktisk når vi snakker om en generell, uspesifisert funksjon, mens den siste som regel er greiest når vi snakker om en bestemt funksjon gitt ved en formel. Husk at $\mathbb{R}^1 = \mathbb{R}$ slik at teorien vår også dekker funksjoner $F : \mathbb{R}^n \rightarrow \mathbb{R}$ der \mathbf{x} er en vektor, men $F(\mathbf{x})$ er et tall.

Vi husker at vanlige funksjoner $f(x)$ ikke alltid er definert for alle reelle tall x , og på samme måte vil heller ikke disse nye funksjonene $\mathbf{F}(\mathbf{x})$ nødvendigvis være definert for alle vektorer $\mathbf{x} \in \mathbb{R}^n$. Den mengden $A \subset \mathbb{R}^n$ av vektorer \mathbf{x} som $\mathbf{F}(\mathbf{x})$ er definert for, kaller vi *definisjonsmengden* eller *definisjonsområdet* til \mathbf{F} . Vi skal av og til bruke symbolet $D_{\mathbf{F}}$ for definisjonsområdet til \mathbf{F} . Når vi skriver

$$\mathbf{F} : A \rightarrow \mathbb{R}^m$$

mener vi at \mathbf{F} er en funksjon definert på mengden A med verdier i \mathbb{R}^m . Hvis en funksjon er gitt ved en formel, og definisjonsområdet ikke er spesifisert, regner vi med at funksjonen er definert overalt hvor formelen gir mening.

En funksjon definert på en delmengde av \mathbb{R}^n kaller vi en funksjon av n -variable; f.eks. er

$$\mathbf{G}(x, y) = (x^y, \ln(1 - xy), \sin(x^2y))$$

en funksjon av to variable med verdier i \mathbb{R}^3 , mens

$$g(x, y, z) = x^2 + ye^{z+y^2}$$

er en funksjon av tre variable med verdier i \mathbb{R} (når funksjonene tar verdier i \mathbb{R} skriver vi dem gjerne uten fete typer og ofte med små bokstaver).

En funksjon som tar verdier i \mathbb{R} kalles gjerne et *skalarfelt* når vi vil understreke at verdiene er tall og ikke vektorer. Legg merke til at en funksjon $\mathbf{F} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ kan skrives

$$\mathbf{F}(x_1, x_2, \dots, x_n) = (F_1(x_1, x_2, \dots, x_n), \dots, F_m(x_1, x_2, \dots, x_n))$$

der F_1, \dots, F_m er skalarfelt. Vi kaller F_1, \dots, F_m for *komponentene* til \mathbf{F} . Funksjonen \mathbf{G} ovenfor har altså komponentene.

$$G_1(x, y) = x^y, \quad G_2(x, y) = \ln(1 - xy), \quad G_3(x, y) = \sin(x^2y)$$

Vi skal ofte gjøre bruk av at funksjoner med verdier i \mathbb{R}^m er bygget opp av skalarfelt på denne måten, for eksempel ved at vi først innfører begreper og beviser resultater for skalarfelt og så utvider til funksjoner med verdier i \mathbb{R}^m .

Det er naturlig å spørre seg selv om hvorfor man må studere funksjoner av flere variable — greier det seg ikke med de funksjonene av bare én variabel som man kjenner fra før? Det viser seg imidlertid at funksjoner av flere variable dukker naturlig opp i svært mange sammenhenger. Her er et lite utvalg:

- BMI (body mass index) brukes ofte som en indikator på overvekt og undervekt. For å finne din BMI, tar du vekten din v (målt i kilo) og deler på kvadratet av høyden din h (målt i meter). Vi kan tenke på BMI som en funksjon f av to variable med verdier i \mathbb{R} :

$$f(v, h) = \frac{v}{h^2}$$

- Når en gjenstand varmes opp, vil temperaturen avhenge av når og hvor vi måler. Det er naturlig å angi temperaturen i punktet (x, y, z) ved tiden t som $T(x, y, z, t)$. Dette er en funksjon av fire variable med verdier i \mathbb{R} .

- Gravitasjonskraften fra jorden på en gjenstand i verdensrommet avhenger av posisjonen til gjenstanden. Tenker vi på kraften som en vektor $\mathbf{F}(\mathbf{x})$ (vi er interessert i både størrelse og retning) som er avhengig av posisjonen $\mathbf{x} = (x, y, z)$, får vi en funksjon av tre variable med verdier i \mathbb{R}^3 . Dersom vi plasserer jorden i origo, kan \mathbf{F} i dette tilfellet skrives

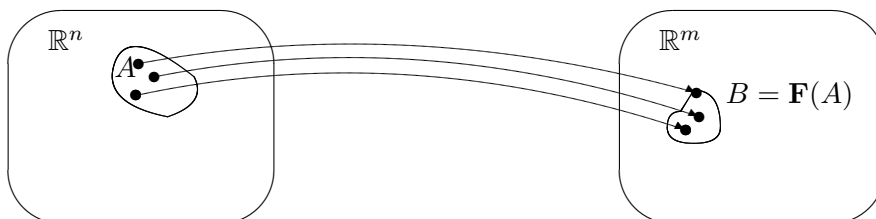
$$\mathbf{F}(\mathbf{x}) = -\gamma \frac{Mm}{|\mathbf{x}|^3} \mathbf{x}$$

der γ er en naturkonstant (gravitasjonskonstanten), M er massen til jorden og m massen til gjenstanden.

- Meteorologer arbeider med vind i atmosfæren. Vinden i et punkt med koordinater (x, y, z) vil være en vektor $\mathbf{v}(x, y, z)$ (vi er interessert i både vindstyrken og retningen). Vi kan tenke på dette som en funksjon av tre variable med verdier i \mathbb{R}^3 .
- I eksempel 3 i seksjon 1.5 så vi på fire produsenter som leverte epler av tre ulike kvaliteter til en fruktpresse. Dersom produsentene leverte hhv. x , y , z og u tonn epler, kunne vi regne ut hvor mange tonn vi ville få av hver kvalitet. Dette gir oss en funksjon av fire variable med verdier i \mathbb{R}^3 . Vi kan skrive funksjonen som $\mathbf{F}(\mathbf{x}) = A\mathbf{x}$, eller med koordinater

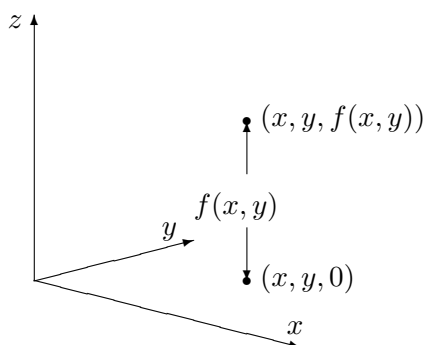
$$\mathbf{F}(x, y, z, u) = \begin{pmatrix} 0.5x + 0.3y + 0.25z + 0.2u \\ 0.3x + 0.4y + 0.4z + 0.6u \\ 0.2x + 0.3y + 0.35z + 0.2u \end{pmatrix}$$

Et av de viktigste verktøyene når man studerer funksjoner av én variabel er grafisk fremstilling. Dessverre er det ikke så lett å gi realistiske grafiske fremstillinger av funksjoner fra \mathbb{R}^n til \mathbb{R}^m . Noen ganger kan vi imidlertid ha stor glede av mer stiliserte fremstillinger som på figuren nedenfor. Den illustrerer hvordan en funksjon \mathbf{F} fra \mathbb{R}^n til \mathbb{R}^m sender alle punkter i en mengde $A \subset \mathbb{R}^n$ på en mengde $B \subset \mathbb{R}^m$. Mengden B kalles *bildet av A under F* og betegnes ofte med $B = \mathbf{F}(A)$. Du vil se mange figurer av denne typen utover i kapitlet.



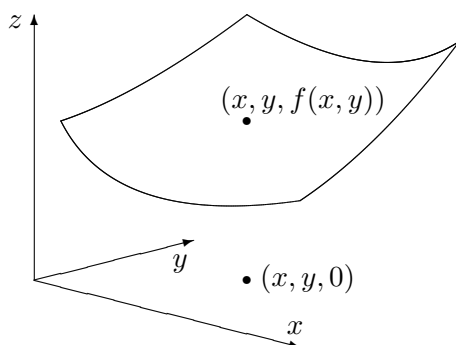
Figur 1: Funksjon fra \mathbb{R}^n til \mathbb{R}^m

Det er imidlertid ett tilfelle der det går an å lage gode, realistiske fremstillinger av en funksjon av flere variable, og det er når funksjonen går fra \mathbb{R}^2 til \mathbb{R} . Vi ser altså på en funksjon $z = f(x, y)$. For å tegne funksjonsgrafen lager vi først et tre-dimensjonalt koordinatsystem som vist på figur 2.



Figur 2: Plotting av skalarfelt

Gitt variabelverdier x og y , finner vi punktet $(x, y, 0)$ i xy -planet. Vi flytter oss nå loddrett (dvs. parallelt med z -aksen) til vi finner punktet $(x, y, f(x, y))$. Dette er det første punktet på funksjonsgrafen vår. Gjentar vi denne prosedyren for stadig flere variabelverdier (x, y) , vokser grafen etterhvert frem som en flate i rommet (se figur 3).



Figur 3: Grafisk fremstilling av skalarfelt

Selv om denne prosedyren på en grei måte forklarer hva grafen til et skalarfelt er, så er den i praksis ubrukelig som en oppskrift for hvordan man tegner grafen. Prøver du den, selv på en enkel funksjon, oppdager du fort at du helt mister romfølelsen i bildet. I neste kapittel skal vi se på mer praktiske metoder for å tegne slike funksjonsgrafer — det som er viktig i dette kapitlet, er at du vet hvordan du kan tenke på grafen til f som en flate.

Oppgaver til seksjon 2.1

1. Finn definisjonsområdet til funksjonen:

a) $f(x, y) = \frac{1}{x^2 + 4y^2}$

b) $f(x, y) = \frac{1}{x^2 - y^2}$

c) $f(x, y) = \ln(x + y)$

d) $f(x, y) = \tan(x - y)$

e) $f(x, y, z) = \frac{1}{x^2 + y^2 + z^2 - 25}$

2.2 Kontinuerlige funksjoner

I seksjon 5.1 i *Kalkulus* studerte vi ϵ - δ -definisjonen av kontinuitet. Denne definisjonen er sannsynligvis ikke like populær blant alle, men den har mange fordeler, blant annet at den lett kan generaliseres til nye situasjoner. I denne seksjonen skal vi se hvordan den generaliseres til funksjoner av flere variable. Før vi begynner, trenger vi å vite litt om avstander og kuler i \mathbb{R}^n .

Akkurat som i planet og rommet lar vi *avstanden* mellom to punkter \mathbf{a} og \mathbf{x} i \mathbb{R}^n være lik lengden til vektoren $\mathbf{x} - \mathbf{a}$ som forbinder dem, det vil si

$$d(\mathbf{a}, \mathbf{x}) = |\mathbf{x} - \mathbf{a}| = \sqrt{(x_1 - a_1)^2 + (x_2 - a_2)^2 + \cdots + (x_n - a_n)^2}$$

(tenk deg at d står for *distance*). Mengden

$$B(\mathbf{a}, r) = \{\mathbf{x} \in \mathbb{R}^n : |\mathbf{x} - \mathbf{a}| < r\}$$

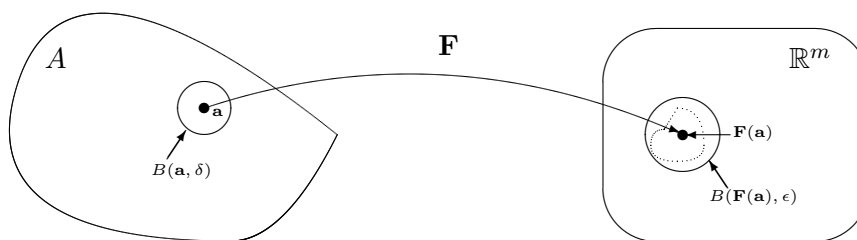
består av de punktene i \mathbb{R}^n som har avstand mindre enn r til punktet \mathbf{a} . Vi kaller $B(\mathbf{a}, r)$ *kulen om \mathbf{a} med radius r* . Legg merke til at i \mathbb{R}^3 er dette virkelig en (åpen) kule i tradisjonell forstand, mens det i \mathbb{R}^2 er en (åpen) sirkelskive og i \mathbb{R} et åpent intervall. Vi velger å bruke “kule” som et fellesord i alle dimensjoner, selv om det til å begynne med kan virke litt uvant når vi arbeider i planet eller på tallinjen. De fleste illustrasjonene våre vil være i planet, og der vil kuler fremstå som sirkler.

Vi kan nå definere kontinuitet for funksjoner av flere variable:

Definisjon 2.2.1 Anta at $A \subset \mathbb{R}^n$, og at $\mathbf{a} \in A$. En funksjon $\mathbf{F} : A \rightarrow \mathbb{R}^m$ er kontinuerlig i \mathbf{a} dersom det til enhver $\epsilon > 0$ finnes en $\delta > 0$ slik at

$$|\mathbf{F}(\mathbf{x}) - \mathbf{F}(\mathbf{a})| < \epsilon \text{ for alle } \mathbf{x} \in A \text{ slik at } |\mathbf{x} - \mathbf{a}| < \delta$$

Figuren nedenfor illustrerer definisjonen: Gitt en kule $B(\mathbf{F}(\mathbf{a}), \epsilon)$ om punktet $\mathbf{F}(\mathbf{a})$, kan vi finne en kule $B(\mathbf{a}, \delta)$ om punktet \mathbf{a} slik at bildet av $B(\mathbf{a}, \delta)$ (markert med den stiplede kurven på figuren) ligger helt inni $B(\mathbf{F}(\mathbf{a}), \epsilon)$.

Figur 1: Kontinuitet i punktet \mathbf{a}

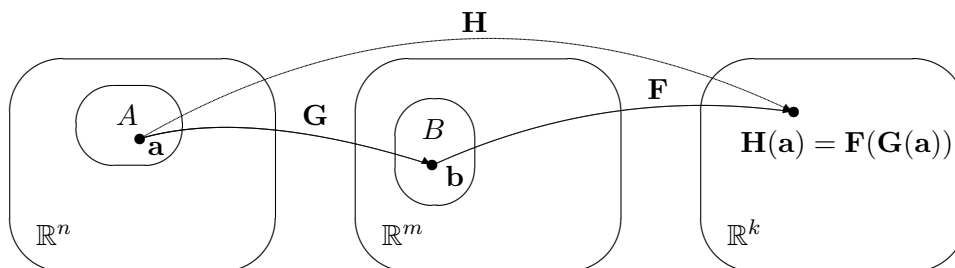
Siden kontinuitet er definert på akkurat samme måte som for funksjoner av én variabel, har vi de samme reglene med (nesten) de samme bevisene.

Setning 2.2.2 Anta at $A \subset \mathbb{R}^n$, og at funksjonene $\mathbf{F}, \mathbf{G} : A \rightarrow \mathbb{R}^m$ er kontinuerlige i $\mathbf{a} \in A$. Da er $\mathbf{F} + \mathbf{G}$, $\mathbf{F} - \mathbf{G}$ og $\mathbf{F} \cdot \mathbf{G}$ kontinuerlige i \mathbf{a} . Det er også $\frac{\mathbf{F}}{\mathbf{G}}$ forutsatt at \mathbf{F} og \mathbf{G} tar verdier i \mathbb{R} (slik at divisjon gir mening) og $\mathbf{G}(\mathbf{a}) \neq 0$. \square

Beviset overlates til leserne, se setning 5.1.5 i *Kalkulus* for hjelp (du får også bruk for trekantulikheten og Schwarz' ulikhet fra seksjon 1.2).

Setning 2.2.3 Anta at vi har to mengder $A \subset \mathbb{R}^n$, $B \subset \mathbb{R}^m$, og to funksjoner $\mathbf{G} : A \rightarrow B$, $\mathbf{F} : B \rightarrow \mathbb{R}^k$ (se figuren nedenfor). Dersom \mathbf{G} er kontinuerlig i punktet \mathbf{a} , og \mathbf{F} er kontinuerlig i punktet $\mathbf{b} = \mathbf{G}(\mathbf{a})$, så er den sammensatte funksjonen $\mathbf{H}(x) = \mathbf{F}(\mathbf{G}(x))$ kontinuerlig i \mathbf{a} . \square

Beviset overlates til leserne, se setning 5.1.7 i *Kalkulus* for hjelp. Figur 2 viser hvordan \mathbf{G} , \mathbf{F} og \mathbf{H} virker.



Figur 2: Sammensetning av funksjoner

Når vi skal bruke reglene ovenfor til å vise at en funksjon $\mathbf{F} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ er kontinuerlig, lønner det seg ofte å skrive den på komponentform:

$$\mathbf{F}(\mathbf{x}) = (F_1(\mathbf{x}), F_2(\mathbf{x}), \dots, F_m(\mathbf{x}))$$

Det neste resultatet forteller oss nemlig at \mathbf{F} er kontinuerlig hvis og bare hvis hver komponent F_i er kontinuerlig.

Setning 2.2.4 Anta at $\mathbf{F} : A \rightarrow \mathbb{R}^m$ er en funksjon av n variable med komponenter

$$\mathbf{F}(\mathbf{x}) = (F_1(\mathbf{x}), F_2(\mathbf{x}), \dots, F_m(\mathbf{x}))$$

Da er \mathbf{F} kontinuert i et punkt $\mathbf{a} \in A$ hvis og bare hvis hver komponent F_i er kontinuert i \mathbf{a} .

Bevis: La oss først anta at alle komponentene F_1, F_2, \dots, F_m er kontinuerte i \mathbf{a} . Vi må vise at for enhver $\epsilon > 0$, finnes det en $\delta > 0$ slik at hvis $|\mathbf{x} - \mathbf{a}| < \delta$, så er $|\mathbf{F}(\mathbf{x}) - \mathbf{F}(\mathbf{a})| < \epsilon$. Siden hver komponent F_i er kontinuert i \mathbf{a} , finnes det en $\delta_i > 0$ slik at $|F_i(\mathbf{x}) - F_i(\mathbf{a})| < \frac{\epsilon}{\sqrt{m}}$ når $|\mathbf{x} - \mathbf{a}| < \delta_i$. Vi velger δ til å være det minste av tallene $\delta_1, \delta_2, \dots, \delta_m$. Hvis $|\mathbf{x} - \mathbf{a}| < \delta$, har vi da

$$\begin{aligned} |\mathbf{F}(\mathbf{x}) - \mathbf{F}(\mathbf{a})| &= \sqrt{(F_1(\mathbf{x}) - F_1(\mathbf{a}))^2 + \dots + (F_m(\mathbf{x}) - F_m(\mathbf{a}))^2} \leq \\ &\leq \sqrt{\left(\frac{\epsilon}{\sqrt{m}}\right)^2 + \dots + \left(\frac{\epsilon}{\sqrt{m}}\right)^2} = \sqrt{m \left(\frac{\epsilon^2}{m}\right)} = \epsilon \end{aligned}$$

Dette viser at \mathbf{F} er kontinuert i \mathbf{a} .

Anta nå omvendt at \mathbf{F} er kontinuert i \mathbf{a} , og at vi skal vise at komponenten F_i er kontinuert i \mathbf{a} . Gitt $\epsilon > 0$, må vi finne en $\delta > 0$ slik at $|F_i(\mathbf{x}) - F_i(\mathbf{a})| < \epsilon$ når $|\mathbf{x} - \mathbf{a}| < \delta$. Siden \mathbf{F} er kontinuert i \mathbf{a} , finnes det en $\delta > 0$ slik at $|\mathbf{F}(\mathbf{x}) - \mathbf{F}(\mathbf{a})| < \epsilon$ når $|\mathbf{x} - \mathbf{a}| < \delta$. Men dermed er

$$\begin{aligned} |F_i(\mathbf{x}) - F_i(\mathbf{a})| &\leq \sqrt{(F_1(\mathbf{x}) - F_1(\mathbf{a}))^2 + \dots + (F_m(\mathbf{x}) - F_m(\mathbf{a}))^2} = \\ &= |\mathbf{F}(\mathbf{x}) - \mathbf{F}(\mathbf{a})| < \epsilon \end{aligned}$$

når $|\mathbf{x} - \mathbf{a}| < \delta$. Dette viser at F_i er kontinuert i \mathbf{a} . □

De enkleste funksjonene av flere variable (bortsett fra konstantfunksjonene) er de som bare gir oss en av variablene som output, f.eks.

$$f(x, y, z, u) = y \quad \text{eller} \quad g(x, y, z) = z$$

Generelt kaller vi

$$k_i(x_1, x_2, \dots, x_n) = x_i$$

den i -te *koordinatfunksjonen*. Det er lett å se at disse koordinatfunksjonene er kontinuerte, og det neste eksemplet forteller oss hvordan vi kan bruke dem og resultatene ovenfor til å vise at mer kompliserte funksjoner er kontinuerte.

Eksempel 1: Vis at funksjonen $\mathbf{F} : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ definert ved

$$\mathbf{F}(x, y, z) = \left(x^2 \sin(yz), \frac{e^{x^2+z}}{y^2+1} \right)$$

er kontinuerlig i punktet $\mathbf{a} = (1, 0, -1)$. Ifølge setning 2.2.4 er det nok å vise at begge funksjonene

$$F_1(x, y, z) = x^2 \sin(yz) \quad \text{og} \quad F_2(x, y, z) = \frac{e^{x^2+z}}{y^2+1}$$

er kontinuerlig i \mathbf{a} . Vi begynner med $F_1(x, y, z)$. Siden koordinatfunksjonene $k_2(x, y, z) = y$ og $k_3(x, y, z) = z$ er kontinuerlige, er produktet av dem $f(x, y, z) = yz$ også kontinuerlig ifølge setning 2.2.2. Siden sinus er en kontinuerlig funksjon, må den sammensatte funksjonen $\sin(yz)$ også være kontinuerlig ifølge setning 2.2.3. På tilsvarende måte ser vi at funksjonen $g(x, y, z) = x^2$ er kontinuerlig (fordi den er lik $k_1(x, y, z) \cdot k_1(x, y, z)$ som er et produkt av to kontinuerlige funksjoner), og dermed er produktet $F_1(x, y, z) = x^2 \sin(yz)$ kontinuerlig.

Vi behandler F_2 på samme måte: Funksjonen $h(x, y, z) = x^2 + z$ er kontinuerlig siden den er bygget opp fra kontinuerlige koordinatfunksjoner ved hjelp av multiplikasjon og addisjon. Siden eksponentialfunksjonen er kontinuerlig, er da $p(x, y, z) = e^{x^2+z}$ kontinuerlig. Ved et tilsvarende resonnement ser vi at nevneren $y^2 + 1$ er kontinuerlig, og siden den er forskjellig fra 0, må brøken

$$F_2(x, y, z) = \frac{e^{x^2+z}}{y^2+1}$$

være kontinuerlig. ♣

Ved hjelp av teknikken i dette eksemplet er det lett å vise at funksjoner bygget opp ved hjelp av potensfunksjoner, eksponentialfunksjoner, logaritmer, trigonometriske funksjoner og arcusfunksjoner er kontinuerlige (men husk å sjekke at de er definert i punktet du er interessert i!)

Hittil har teorien for kontinuerlige funksjoner av flere variable vært forbløffende lik teorien for kontinuerlige funksjoner av én variabel. På grunn av den rikere geometrien finnes det imidlertid fenomener for flervariabel funksjoner som ikke har noen motsvarighet for funksjoner av én variabel. Her er et eksempel.

Eksempel 2: Funksjonen $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ er gitt ved

$$f(x, y) = \begin{cases} \frac{x^2 y}{x^4 + y^2} & \text{for } (x, y) \neq (0, 0) \\ 0 & \text{for } (x, y) = (0, 0) \end{cases}$$

Vi skal undersøke om f er kontinuerlig i $(0, 0)$. Uformelt sier en gjerne at f er kontinuerlig $(0, 0)$ dersom $f(x, y)$ nærmer seg $f(0, 0)$ når (x, y) nærmer seg $(0, 0)$. La oss derfor se hva som skjer når (x, y) nærmer seg $(0, 0)$ på forskjellige måter.

La oss først se på hva som skjer når (x, y) nærmer seg origo langs en skrålinje $y = cx$ der $c \neq 0$; dvs. vi har punkter (x, cx) der x går mot null. Setter vi inn i funksjonsuttrykket, får vi

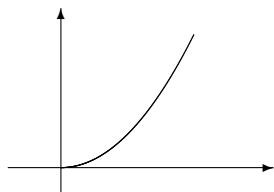
$$f(x, y) = f(x, cx) = \frac{x^2(cx)}{x^4 + cx^2} = \frac{cx}{x^2 + c} \rightarrow \frac{0}{c} = 0$$

Tilsvarende ser vi at når (x, y) nærmer seg $(0, 0)$ langs x -aksen, så er punktene på formen $(x, 0)$ der $x \rightarrow 0$, og vi får

$$f(x, y) = f(x, 0) = \frac{x^2 \cdot 0}{x^4 + 0^2} = 0$$

Akkurat det samme skjer om vi lar (x, y) nærme seg $(0, 0)$ langs y -aksen.

Alt dette tyder på at f er kontinuerlig i $(0, 0)$, men det finnes jo andre måter å nærme seg et punkt på enn å følge en rett linje. La oss prøve å nærme oss $(0, 0)$ langs parabelen $y = x^2$ (se figuren). Vi ser altså på punkter (x, x^2) der x går mot 0.



Figur 3. Punkter som nærmer seg $(0, 0)$ langs parabelen $y = x^2$

Setter vi inn i funksjonsuttrykket, får vi

$$f(x, y) = f(x, x^2) = \frac{x^2 \cdot x^2}{x^4 + (x^2)^2} = \frac{1}{2}$$

Dette viser at f *ikke* er kontinuerlig i $(0, 0)$ siden vi har punkter med verdien $\frac{1}{2}$ så nær $(0, 0)$ vi måtte ønske. Funksjonen virker altså å være kontinuerlig så lenge vi beveger oss langs rette linjer, men er det likevel ikke! ♣

Eksemplet ovenfor viser at det kan være vanskelig å få god oversikt over hvordan en funksjon av flere variable oppfører seg, og at det kanskje finnes flere geometriske muligheter enn det vi kan forestille oss. Det er i slike sammenhenger vi virkelige får nytte av abstrakte definisjoner av ϵ - δ -typen; de gir oss muligheten til å føre vanntette bevis selv i tilfeller der vi ikke er sikre på om vi har fått full oversikt over alle geometriske snurrepiperier!

Hittil har vi bare snakket om kontinuitet i et punkt. Vi avslutter denne seksjonen med definisjonen av en kontinuerlig funksjon — den er helt tilsvarende definisjonen for funksjoner av én variabel.

Definisjon 2.2.5 En funksjon F kalles kontinuerlig dersom den er kontinuerlig i alle punkter i sitt definisjonsområde.

Oppgaver til seksjon 2.2

1. Vis at funksjonen f er kontinuerlig

- | | |
|-------------------------------------|-----------------------------------|
| a) $f(x, y) = x + y$ | d) $f(x, y) = e^{-x} \sin(x + y)$ |
| b) $f(x, y) = x^2 y + y$ | e) $f(x, y, z) = x^2 + y^2 + z^2$ |
| c) $f(x, y) = \frac{xy}{1+x^2+y^2}$ | |

2. Vis at funksjonene er kontinuerlige:

- a) $\mathbf{F}(x, y, z) = (x^2 z + y, x^2 \sin xyz, x^3)$
 b) $\mathbf{G}(x, y, z, u) = (e^{xu+z^2}, z \cos xy^2 u)$
 c) $\mathbf{H}(x, y, z, u) = (x^y e^{xz^2}, z + u^2, x^2 + 3yzu)$
 d) $\mathbf{K}(x, y, z, u, v) = (\sin(xy + z^2 v), 2uv)$

3. Vis at koordinatfunksjonene $k_i(x_1, x_2, \dots, x_n) = x_i$ er kontinuerlige ved å bruke definisjonen av kontinuitet.

4. a) Anta det finnes en konstant $M \in \mathbb{R}$ slik at $|\mathbf{F}(\mathbf{x}) - \mathbf{F}(\mathbf{y})| \leq M|\mathbf{x} - \mathbf{y}|$ for alle \mathbf{x}, \mathbf{y} i definisjonsområdet $D_{\mathbf{F}}$. Vis at \mathbf{F} er kontinuerlig.

b) Anta at $\mathbf{F}(\mathbf{x}) = A\mathbf{x}$ for en matrise A . Vis at \mathbf{F} er kontinuerlig. (*Hint:* Husk setning 1.6.3.)

5. I denne oppgaven har du bruk for trekantulikheten som sier at hvis $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, så er $|\mathbf{x} + \mathbf{y}| \leq |\mathbf{x}| + |\mathbf{y}|$.

a) Vis at $||\mathbf{x}| - |\mathbf{y}|| \leq |\mathbf{x} - \mathbf{y}|$ for alle $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

b) La $\mathbf{a} \in \mathbb{R}^n$. Vis at funksjonen $f(\mathbf{x}) = |\mathbf{x} - \mathbf{a}|$ er kontinuerlig.

c) Vis at funksjon $g(\mathbf{x}) = \frac{1}{|\mathbf{x} - \mathbf{a}|}$ er kontinuerlig der den er definert.

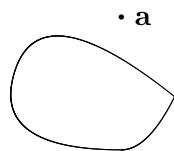
6. Bevis setning 2.2.2.

7. Bevis setning 2.2.3.

2.3 Grenseverdier

I teorien for funksjoner av én variabel opererer vi både med ensidige og tosidige grenser. Ensidige grenser er blant annet nyttig når vi skal avgjøre om en funksjon er kontinuerlig i enden av sitt definisjonsområde — vi bruker grenseverdien $\lim_{x \rightarrow a^+} f(x)$ for å undersøke om f er kontinuerlig i det venstre endepunktet a av intervallet $[a, b]$, og grenseverdien $\lim_{x \rightarrow b^-} f(x)$ for å undersøke om f er kontinuerlig i det høyre endepunktet b . Med flere variable er det så mange måter å nærme seg et randpunkt på at vi ikke kan ha ett grensebegrep for hver måte — vi må finne frem til et felles begrep som dekker alle tilfeller.

La oss først se litt på hvilke punkter det er naturlig å regne ut grenseverdier i. Vi ser på en funksjon $\mathbf{F} : A \rightarrow \mathbb{R}^m$ definert på en delmengde A av \mathbb{R}^n . Figur 1 viser én mulighet; her består A av en sammenhengende mengde pluss et isolert punkt \mathbf{a} .



Figur 1

Siden vi ikke kan nærme oss \mathbf{a} innenfor definisjonsmengden til \mathbf{F} , er det ikke rimelig å definere grenseverdien til \mathbf{F} i punktet \mathbf{a} . Figur 2 viser en annen situasjon.



Figur 2

På denne figuren er omkretsen (“randen”) til A stiplet for å markere at punktene der ikke hører med til mengden. Dermed er \mathbf{a} ikke med i definisjonsmengden A , men det går fint an å nærme seg \mathbf{a} fra A . I dette tilfellet er det rimelig å definere grenseverdien til \mathbf{F} i punktet \mathbf{a} . Konklusjonen på disse observasjonene må bli at det er rimelig å definere grenseverdien til \mathbf{F} i punktet \mathbf{a} dersom det går an å nærme seg \mathbf{a} med punkter ($\neq \mathbf{a}$) fra definisjonsmengden til \mathbf{F} , uansett om punktet \mathbf{a} selv ligger i definisjonsmengden eller ikke. Slike punkter kaller vi *oppbopningspunkter* for A . Her er den presise definisjonen:

Definisjon 2.3.1 La A være en delmengde av \mathbb{R}^n . Et punkt $\mathbf{a} \in \mathbb{R}^n$ kalles et oppbopningspunkt for A dersom enhver kule $B(\mathbf{a}, r)$ om \mathbf{a} inneholder uendelig mange punkter fra A .

Vi er nå klare til å definere grenseverdien i et oppbopningspunkt:

Definisjon 2.3.2 La $\mathbf{F} : A \rightarrow \mathbb{R}^m$ være en funksjon av n variable og anta at $\mathbf{a} \in \mathbb{R}^n$ er et oppbopningspunkt for A . Vi sier at $\mathbf{b} \in \mathbb{R}^m$ er grenseverdien for \mathbf{F} i punktet \mathbf{a} dersom det for hver $\epsilon > 0$ finnes en $\delta > 0$ slik at

$$|\mathbf{F}(\mathbf{x}) - \mathbf{b}| < \epsilon \text{ for alle } \mathbf{x} \in A \text{ slik at } 0 < |\mathbf{x} - \mathbf{a}| < \delta$$

Vi skriver $\lim_{\mathbf{x} \rightarrow \mathbf{a}} \mathbf{F}(\mathbf{x}) = \mathbf{b}$. Dersom \mathbf{a} ikke er et oppbopningspunkt for A , er grenseverdien ikke definert.

Denne definisjonen ligner på den du finner for funksjoner av én variabel i definisjon 5.4.1 i *Kalkulus*, men skiller seg på et viktig punkt: Vi insisterer

ikke lenger på at \mathbf{F} skal være definert i alle punkter i nærheten av \mathbf{a} , men kompenserer for dette ved bare å kreve at $|\mathbf{F}(\mathbf{x}) - \mathbf{b}| < \epsilon$ skal holde for punkter \mathbf{x} som er med i definisjonsområdet A til \mathbf{F} .

Akkurat som kontinuitet kan grenseverdier studeres komponentvis:

Setning 2.3.3 La $\mathbf{F} : A \rightarrow \mathbb{R}^m$ være en funksjon av n variable, og anta at $\mathbf{a} \in \mathbb{R}^n$ er et opphopningspunkt for A . Anta at komponentene til \mathbf{F} er gitt ved

$$\mathbf{F}(\mathbf{x}) = (F_1(\mathbf{x}), F_2(\mathbf{x}), \dots, F_m(\mathbf{x}))$$

og la \mathbf{b} være en vektor med komponenter $\mathbf{b} = (b_1, b_2, \dots, b_m)$. Da er

$$\lim_{x \rightarrow \mathbf{a}} \mathbf{F}(\mathbf{x}) = \mathbf{b}$$

hvis og bare hvis

$$\lim_{x \rightarrow \mathbf{a}} F_i(\mathbf{x}) = b_i \text{ for alle } i$$

□

Beviset overlates til leserne. Ideen er den samme som i beviset for setning 2.2.4.

Vi har også de vanlige regnereglerne for grenseverdier til summer, differenser, produkter og brøker:

Setning 2.3.4 (Regneregler for grenseverdier) Anta at $\mathbf{F}, \mathbf{G} : A \rightarrow \mathbb{R}^m$ er to funksjoner av n variable og at $\mathbf{a} \in A$ er et opphopningspunkt for A . Dersom $\lim_{x \rightarrow \mathbf{a}} \mathbf{F}(\mathbf{x}) = \mathbf{A}$ og $\lim_{x \rightarrow \mathbf{a}} \mathbf{G}(\mathbf{x}) = \mathbf{B}$, så er:

a) $\lim_{x \rightarrow \mathbf{a}} (\mathbf{F}(\mathbf{x}) + \mathbf{G}(\mathbf{x})) = \mathbf{A} + \mathbf{B}$.

b) $\lim_{x \rightarrow \mathbf{a}} (\mathbf{F}(\mathbf{x}) - \mathbf{G}(\mathbf{x})) = \mathbf{A} - \mathbf{B}$.

c) $\lim_{x \rightarrow \mathbf{a}} (\mathbf{F}(\mathbf{x}) \cdot \mathbf{G}(\mathbf{x})) = \mathbf{A} \cdot \mathbf{B}$.

d) $\lim_{x \rightarrow \mathbf{a}} \frac{\mathbf{F}(\mathbf{x})}{\mathbf{G}(\mathbf{x})} = \frac{\mathbf{A}}{\mathbf{B}}$ forutsatt at \mathbf{F} og \mathbf{G} tar verdier i \mathbb{R} og $\mathbf{B} \neq 0$.

□

Også disse bevisene overlates til leserne.

Det neste resultatet tar seg av sammenhengen mellom grenseverdier og kontinuitet. Legg merke til at det bare gjelder for kontinuitet i opphopningspunkter (men kontinuitet i isolerte punkter er uansett en ganske kunstig sak).

Setning 2.3.5 La $\mathbf{F} : A \rightarrow \mathbb{R}^m$ være en funksjon av n variable, og anta at $\mathbf{a} \in A$ er et opphopningspunkt for A . Da er \mathbf{F} kontinuerlig i \mathbf{a} hvis og bare hvis $\lim_{\mathbf{x} \rightarrow \mathbf{a}} \mathbf{F}(\mathbf{x}) = \mathbf{F}(\mathbf{a})$

Bevis: Sammenlign definisjonen av kontinuitet med definisjonen av grenseverdi. □

Resultatet ovenfor er nyttig når vi skal regne ut enkle grenseverdier:

Eksempel 1: Finn grenseverdien til

$$F(x, y) = (x^2 y, e^{-xy} \sin(\pi x))$$

når $(x, y) \rightarrow (1, -2)$.

Vi ser at funksjonen er kontinuert i $(1, -2)$, så

$$\lim_{(x,y) \rightarrow (1,-2)} F(x, y) = F(1, -2) = (1^2 \cdot (-2), e^{-1 \cdot (-2)} \sin(\pi \cdot 1)) = (-2, 0)$$



I noen eksempler må vi forenkle uttrykket før vi går til grensen:

Eksempel 2: Finn grenseverdien til

$$f(x, y) = \frac{x^2 - y^2}{x - y}$$

når $(x, y) \rightarrow (0, 0)$.

Siden $x^2 - y^2 = (x + y)(x - y)$ har vi

$$\begin{aligned} \lim_{(x,y) \rightarrow (0,0)} f(x, y) &= \lim_{(x,y) \rightarrow (0,0)} \frac{x^2 - y^2}{x - y} = \\ &= \lim_{(x,y) \rightarrow (0,0)} \frac{(x + y)(x - y)}{x - y} = \lim_{(x,y) \rightarrow (0,0)} (x + y) = 0 \end{aligned}$$



Oppgaver til seksjon 2.3

1. Finn grenseverdiene

- a) $\lim_{(x,y) \rightarrow (2,3)} (x^3 + 2xy)$
- b) $\lim_{(x,y) \rightarrow (1, \frac{\pi}{2})} x^2 \sin(xy)$
- c) $\lim_{(x,y) \rightarrow (1,0)} \frac{e^{x+y}}{x^2 + 3y}$
- d) $\lim_{(x,y) \rightarrow (0,0)} \frac{\sin(xy)}{xy} \cdot \cos(x + y)$

2. Anta at $A \subset \mathbb{R}^n$ og at $\mathbf{a} \in \mathbb{R}^n$. Anta at enhver kule $B(\mathbf{a}, \epsilon)$ om \mathbf{a} inneholder minst ett element fra A forskjellig fra \mathbf{a} . Vis at \mathbf{a} er et opphopningspunkt for A .

3. Bevis setning 2.3.3.

4. Bevis setning 2.3.4

5. Bevis setning 2.3.5

2.4 Derivasjon av skalarfelt

Vi skal nå begynne å se på derivasjon av funksjoner av flere variable. For at det ikke skal være altfor mange komponenter å holde styr på, skal vi først derivere skalarfelt, dvs. funksjoner $f : \mathbb{R}^n \rightarrow \mathbb{R}$ der verdien $f(\mathbf{x})$ er tall. Før vi setter igang for alvor, trenger vi en definisjon: Hvis A er en delmengde av \mathbb{R}^n , kalles $\mathbf{a} \in A$ et *indre punkt* i A dersom det finnes en $\epsilon > 0$ slik at $B(\mathbf{a}, \epsilon) \subset A$. Dette betyr at dersom vi starter i \mathbf{a} , kan vi gå et lite stykke i en hvilken som helst retning uten å forlate A .

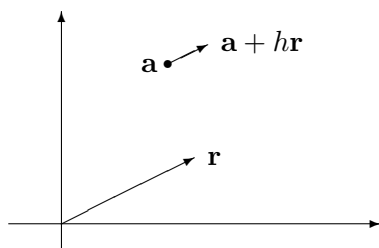
For en funksjon $y = f(x)$ av én variabel forteller den deriverte $f'(x)$ oss hvor fort funksjonen vokser i punktet x — går vi et lite skritt med lengde h langs x -aksen, vil funksjonsverdien øke med (omtrent) $f'(x)h$. For funksjoner av flere variable er situasjonen mer komplisert; vi har flere akser å bevege oss langs, og vi kan ikke regne med at funksjonen stiger like mye uansett hvilken retning vi går i. Før vi regner ut stigningstallet til funksjonen, må vi derfor spesifisere hvilken retning vi er interessert i. Dette er idéen bak begrepet retningsderivert:

Definisjon 2.4.1 Anta at funksjonen $f : A \rightarrow \mathbb{R}$ er definert på en delmengde A av \mathbb{R}^n og at \mathbf{a} er et indre punkt i A . Tenk på $\mathbf{r} \in \mathbb{R}^n$ som en vektor. Den retningsderiverte til f i punktet \mathbf{a} og retningen \mathbf{r} er gitt ved

$$f'(\mathbf{a}; \mathbf{r}) = \lim_{h \rightarrow 0} \frac{f(\mathbf{a} + h\mathbf{r}) - f(\mathbf{a})}{h}$$

forutsatt at denne grensen eksisterer.

Figur 1 viser ideen bak definisjonen.



Figur 1

Punktene $\mathbf{a} + h\mathbf{r}$ er de punktene vi kommer til hvis vi starter i \mathbf{a} og går i retning \mathbf{r} . Differansen $f(\mathbf{a} + h\mathbf{r}) - f(\mathbf{a})$ forteller oss hvor mye funksjonen øker når vi beveger oss i denne retning, og brøken

$$\frac{f(\mathbf{a} + h\mathbf{r}) - f(\mathbf{a})}{h}$$

er økningen per lengdeenhet når vi bruker $|\mathbf{r}|$ som måleenhet.

Legg merke til at vi bare har definert retningsderiverte i indre punkter. Det garanterer at uttrykket $f(\mathbf{a} + h\mathbf{r}) - f(\mathbf{a})$ alltid gir mening bare vi velger

h liten nok. I punkter som ikke er indre, kan man ofte definere retningsderiverte i noen retninger, men ikke i andre. Vi skal ikke komme inn på dette her, men konsentrere oss om indre punkter.

Eksempel 1: La $f(x, y) = x^2 + xy$. Vi skal beregne den retningsderiverte $f'(\mathbf{a}; \mathbf{r})$ når $\mathbf{a} = (1, 0)$, $\mathbf{r} = (2, 1)$. Først observerer vi at

$$\mathbf{a} + h\mathbf{r} = (1, 0) + h(2, 1) = (1 + 2h, h),$$

som gir $f(\mathbf{a} + h\mathbf{r}) = (1 + 2h)^2 + (1 + 2h)h = 1 + 5h + 6h^2$. Tilsvarende er $f(\mathbf{a}) = f(1, 0) = 1^2 + 1 \cdot 0 = 1$. Vi får

$$\begin{aligned} f'(\mathbf{a}; \mathbf{r}) &= \lim_{h \rightarrow 0} \frac{f(\mathbf{a} + h\mathbf{r}) - f(\mathbf{a})}{h} = \lim_{h \rightarrow 0} \frac{(1 + 5h + 6h^2) - 1}{h} \\ &= \lim_{h \rightarrow 0} \frac{5h + 6h^2}{h} = \lim_{h \rightarrow 0} (5 + 6h) = 5. \end{aligned}$$

Hva betyr dette resultatet? Legg merke til at lengden til vektoren \mathbf{r} er $|\mathbf{r}| = \sqrt{2^2 + 1^2} = \sqrt{5}$. Dersom vi går et lite stykke $h\sqrt{5}$ i retningen til vektoren $\mathbf{r} = (2, 1)$, vil funksjonsverdien stige med (omtrent) $f'(\mathbf{a}; \mathbf{r}) \cdot h = 5 \cdot h$. ♣

Det er lettest å forstå hva den retningsderiverte er dersom vektoren \mathbf{r} har lengde 1 – da er $f'(\mathbf{a}; \mathbf{r})$ rett og slett stigningstallet til funksjonen i retning \mathbf{r} når vi måler med vanlige enheter. Utifra dette kan det være fristende å forutsette at $|\mathbf{r}| = 1$ når vi regner med retningsderiverte, men dette viser seg å være upraktisk, blant annet fordi enhetsvektorer ofte inneholder stygge kvadratrøtter.

Så langt kan det se ut som om vi må bygge opp en ny derivasjonsteori helt fra bunnen av for å kunne beregne retningsderiverte til funksjoner av flere variable. Det er heldigvis ikke nødvendig; ved hjelp av såkalte partiellderiverte kan vi føre mye av teorien tilbake til vanlig derivasjon av funksjoner av én variabel. Før vi definerer partiellderiverte, er det lurt å bli enig om litt notasjon.

Den i -te enhetsvektoren \mathbf{e}_i i \mathbb{R}^n er vektoren

$$\mathbf{e}_i = (0, 0, \dots, 0, \underset{\substack{\uparrow \\ i\text{-te plass}}}{1}, 0, \dots, 0)$$

parallel med den i -te koordinataksen.

Definisjon 2.4.2 La $f: A \rightarrow \mathbb{R}$ være en funksjon av n variable, og la \mathbf{a} være et indre punkt i A . Den i -te partiellderiverte $\frac{\partial f}{\partial x_i}(\mathbf{a})$ er den retningsderiverte av f i retning av den i -te enhetsvektoren \mathbf{e}_i ; det vil si

$$\frac{\partial f}{\partial x_i}(\mathbf{a}) = f'(\mathbf{a}; \mathbf{e}_i)$$

Andre notasjoner for $\frac{\partial f}{\partial x_i}(\mathbf{a})$ er $D_i f(\mathbf{a})$ og $f_{x_i}(\mathbf{a})$. De partiellderiverte er altså stigningstallene til funksjonen parallelt med koordinataksene. Skriver vi ut definisjonen i detalj, ser vi at

$$\begin{aligned}\frac{\partial f}{\partial x_i}(\mathbf{a}) &= f'(\mathbf{a}; \mathbf{e}_i) = \lim_{h \rightarrow 0} \frac{f(\mathbf{a} + h\mathbf{e}_i) - f(\mathbf{a})}{h} = \\ &= \lim_{h \rightarrow 0} \frac{f(a_1, a_2, \dots, a_i + h, \dots, a_n) - f(a_1, a_2, \dots, a_i, \dots, a_n)}{h}\end{aligned}$$

Det siste uttrykket har en slående likhet med definisjonen av vanlig derivert. Underslår vi de variablene hvor det ikke skjer noen endring, ser vi at

$$\frac{\partial f}{\partial x_i}(\mathbf{a}) = \lim_{h \rightarrow 0} \frac{f(a_i + h) - f(a_i)}{h}$$

Dette betyr at vi kan finne den partiellderiverte $\frac{\partial f}{\partial x_i}$ ved å derivere uttrykket $y = f(x_1, \dots, x_i, \dots, x_n)$ med hensyn på x_i mens vi later som om alle de andre variablene er konstanter.

Eksempel 2: Finn de partiellderiverte $\frac{\partial f}{\partial x}$ og $\frac{\partial f}{\partial y}$ til funksjonen

$$f(x, y) = x^2 + xy^3 + \sin(xy).$$

For å finne $\frac{\partial f}{\partial x}$ deriverer vi uttrykket med hensyn på x mens vi later som om y er en konstant:

$$\frac{\partial f}{\partial x}(x, y) = 2x + y^3 + \cos(xy) \cdot y$$

For å finne $\frac{\partial f}{\partial y}$ deriverer vi med hensyn på y mens vi holder x konstant:

$$\frac{\partial f}{\partial y}(x, y) = 0 + 3xy^2 + \cos(xy)x = 3xy^2 + x \cos(xy)$$



En funksjon $f(x_1, x_2, \dots, x_n)$ av n variable har n partiellderiverte $\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n}$. Vi kan sette sammen disse til en vektor

$$\left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n} \right)$$

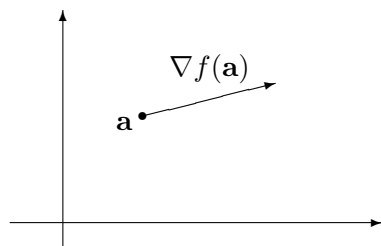
Denne vektoren er så viktig at den har fått sitt eget navn og sitt eget symbol.

Definisjon 2.4.3 Anta at de partiellderiverte til f eksisterer i punktet $\mathbf{a} \in \mathbb{R}^n$. Da kalles

$$\nabla f(\mathbf{a}) = \left(\frac{\partial f}{\partial x_1}(\mathbf{a}), \frac{\partial f}{\partial x_2}(\mathbf{a}), \dots, \frac{\partial f}{\partial x_n}(\mathbf{a}) \right)$$

gradienten til f i punktet \mathbf{a} .

Det er ofte lurt å tenke på gradienten $\nabla f(\mathbf{a})$ som en vektor som starter i punktet \mathbf{a} slik som vist på figuren nedenfor. Som vi snart skal se (setning 2.4.7), får gradienten da en geometrisk betydning — den peker i den retningen hvor funksjonen vokser raskest, og lengden $|\nabla f(\mathbf{a})|$ er lik stigningstallet i denne retningen.



Figur 2

Eksempel 3: Finn gradienten til

$$f(x, y, z) = x^2 y e^{xz}$$

i punktet $\mathbf{a} = (1, -2, 0)$.

Vi må først finne de partiellderiverte. Deriverer vi mhp. x som om y og z er konstanter, får vi

$$\frac{\partial f}{\partial x} = 2xye^{xz} + x^2 ye^{xz} z = xye^{xz}(2 + xz)$$

Deriverer vi mhp. y som om x og z er konstanter, får vi tilsvarende

$$\frac{\partial f}{\partial y} = x^2 e^{xz}$$

Til slutt deriverer vi mhp. z som om x og y er konstanter:

$$\frac{\partial f}{\partial z} = x^2 ye^{xz} x = x^3 ye^{xz}$$

Gradienten i et generelt punkt er dermed

$$\nabla f(x, y, z) = (xye^{xz}(2 + xz), x^2 e^{xz}, x^3 ye^{xz})$$

I vårt punkt $\mathbf{a} = (1, -2, 0)$ får vi

$$\begin{aligned} \nabla f(1, -2, 0) &= (1 \cdot (-2) \cdot e^{1 \cdot 0} \cdot (2 + 1 \cdot 0), 1^2 \cdot e^{1 \cdot 0}, 1^3 \cdot (-2) \cdot e^{1 \cdot 0}) = \\ &= (-4, 1, -2) \end{aligned}$$



Vi har nå sett hvordan vi kan bruke våre vanlige derivasjonsregler til å regne ut partiellderiverte — vi bare deriverer som om de andre variablene

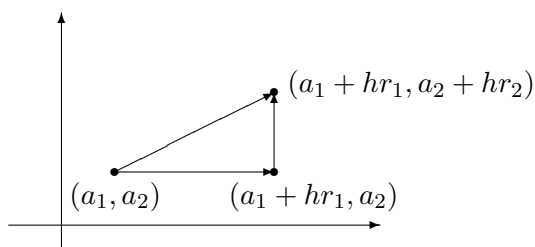
var konstanter. Neste post på programmet er å vise hvordan vi kan regne ut retningsderiverte ved hjelp av partiellderiverte.

La oss begynne med å se på hva som skjer i to dimensjoner. Anta at vi ønsker å derivere f i punktet $\mathbf{a} = (a_1, a_2)$ i retningen $\mathbf{r} = (r_1, r_2)$. Dersom h er en liten størrelse, vet vi at

$$f(a_1 + hr_1, a_2 + hr_2) - f(a_1, a_2) = f'(\mathbf{a}; \mathbf{r}) \cdot h \quad (2.4.1)$$

når vi ser bort fra en størrelse som er liten sammenlignet med h .

Vi kan også beregne differansen $f(a_1 + hr_1, a_2 + hr_2) - f(a_1, a_2)$ på en annen måte. Istedenfor å gå direkte langs vektoren \mathbf{r} velger vi å gå parallelt med koordinataksene som vist på figur 2.



Figur 3

Vi ser at

$$\begin{aligned} f(a_1 + hr_1, a_2 + hr_2) - f(a_1, a_2) &= \\ &= \overbrace{f(a_1 + hr_1, a_2 + hr_2) - f(a_1 + hr_1, a_2)}^{\text{økning parallelt med } y\text{-aksen}} + \\ &\quad \overbrace{f(a_1 + hr_1, a_2) - f(a_1, a_2)}^{\text{økning parallelt med } x\text{-aksen}} \end{aligned}$$

Ser vi bort fra feil som er små sammenlignet med h , får vi videre

$$\begin{aligned} f(a_1 + hr_1, a) - f(a_1, a_2) &= \frac{\partial f}{\partial x}(\mathbf{a}) \cdot hr_1 \\ f(a_1 + hr_1, a + hr_2) - f(a_1 + hr_1, a_2) &= \frac{\partial f}{\partial y}(\mathbf{a}) \cdot hr_2 \end{aligned}$$

Dette betyr at

$$f(a_1 + hr_1, a_2 + hr_2) - f(a_1, a_2) = \frac{\partial f}{\partial x}(\mathbf{a})hr_1 + \frac{\partial f}{\partial y}(\mathbf{a})hr_2 \quad (2.4.2)$$

(fortsatt med en feil som er liten sammenlignet med h). Sammenligner vi de to uttrykkene (2.4.1) og (2.4.2) vi nå har for $f(a_1 + hr_1, a_2 + hr_2) - f(a_1, a_2)$, ser vi at

$$f'(\mathbf{a}; \mathbf{r}) = \frac{\partial f}{\partial x}(\mathbf{a})r_1 + \frac{\partial f}{\partial y}(\mathbf{a})r_2$$

Gjennomfører vi et tilsvarende argument i n variable, får vi formelen

$$f'(\mathbf{a}; \mathbf{r}) = \frac{\partial f}{\partial x_1}(\mathbf{a})r_1 + \frac{\partial f}{\partial x_2}(\mathbf{a})r_2 + \dots + \frac{\partial f}{\partial x_n}(\mathbf{a})r_n$$

der $\mathbf{r} = (r_1, r_2, \dots, r_n)$. Husker vi at gradienten til f er gitt ved

$$\nabla f(\mathbf{a}) = \left(\frac{\partial f}{\partial x_1}(\mathbf{a}), \frac{\partial f}{\partial x_2}(\mathbf{a}), \dots, \frac{\partial f}{\partial x_n}(\mathbf{a}) \right),$$

kan vi skrive formelen ovenfor som et skalarprodukt

$$f'(\mathbf{a}; \mathbf{r}) = \nabla f(\mathbf{a}) \cdot \mathbf{r} \quad (2.4.3)$$

Denne formelen gir oss en effektiv måte å regne ut retningsderiverte på; vi finner alle de partiellderiverte, setter dem sammen til en gradient, og tar skalarproduktet mellom gradienten og vektoren \mathbf{r} . Beregningen som ledet oss frem til (2.4.3), er imidlertid ingen streng utledning i matematisk forstand (vi har skrevet eksakt likhet = en rekke steder der vi bare hadde omtrentlig likhet), og det viser seg at det finnes funksjoner f som ikke oppfyller (2.4.3) til tross for at både den retningsderiverte og de partiellderiverte eksisterer (et eksempel på dette er funksjonen i eksempel 2 i seksjon 2.2). Disse funksjonene oppfører seg imidlertid så merkelig at vi ønsker å utelukke dem fra teorien vår. Vi skal derfor innføre et begrep *deriverbar funksjon* som fanger opp funksjoner med den oppførselen vi ønsker oss. Utgangspunktet for definisjonen er at vi ønsker at $\nabla f(\mathbf{a}) \cdot \mathbf{r}$ skal være en god tilnærming til funksjonsdifferansen $f(\mathbf{a} + \mathbf{r}) - f(\mathbf{a})$ når \mathbf{r} er liten. Mer presist ønsker vi at “feilledet”

$$\sigma(\mathbf{r}) = f(\mathbf{a} + \mathbf{r}) - f(\mathbf{a}) - \nabla f(\mathbf{a}) \cdot \mathbf{r}$$

skal bli mindre og mindre sammenlignet med størrelsen til \mathbf{r} , dvs. at

$$\lim_{\mathbf{r} \rightarrow \mathbf{0}} \frac{\sigma(\mathbf{r})}{|\mathbf{r}|} = 0$$

Vi får altså denne definisjonen:

Definisjon 2.4.4 Anta at $f: A \rightarrow \mathbb{R}$ er definert på en delmengde A av \mathbb{R}^n og at \mathbf{a} er et indre punkt i A . Anta videre at alle de partiellderiverte til f eksisterer i punktet \mathbf{a} . Vi sier at f er *deriverbar* i \mathbf{a} dersom funksjonen

$$\sigma(\mathbf{r}) = f(\mathbf{a} + \mathbf{r}) - f(\mathbf{a}) - \nabla f(\mathbf{a}) \cdot \mathbf{r}$$

går mot 0 hurtigere enn $|\mathbf{r}|$, dvs.

$$\lim_{\mathbf{r} \rightarrow \mathbf{0}} \frac{\sigma(\mathbf{r})}{|\mathbf{r}|} = 0$$

Kommentar: Man kan lure på hvorfor vi ikke rett å slett sier at f er deriverbar i \mathbf{a} dersom $f'(\mathbf{a}; \mathbf{r}) = \nabla f(\mathbf{a}) \cdot \mathbf{r}$ holder for alle \mathbf{r} . Det viser seg imidlertid at definisjonen ovenfor gir en glattere teori hvor delene passer bedre sammen. Det neste resultatet sier dessuten at vi får likheten $f'(\mathbf{a}; \mathbf{r}) = \nabla f(\mathbf{a}) \cdot \mathbf{r}$ uansett.

Setning 2.4.5 Anta at $f: A \rightarrow \mathbb{R}$ er deriverbar i \mathbf{a} . Da er $f'(\mathbf{a}; \mathbf{r}) = \nabla f(\mathbf{a}) \cdot \mathbf{r}$ for alle \mathbf{r} .

Bevis: Vi har

$$\begin{aligned} f(\mathbf{a}; \mathbf{r}) &= \lim_{h \rightarrow 0} \frac{f(\mathbf{a} + h\mathbf{r}) - f(\mathbf{a})}{h} = \lim_{h \rightarrow 0} \frac{\nabla f(\mathbf{a}) \cdot (h\mathbf{r}) + \sigma(h\mathbf{r})}{h} = \\ &= \lim_{h \rightarrow 0} \left(\nabla f(\mathbf{a}) \cdot \mathbf{r} + |\mathbf{r}| \frac{\sigma(h\mathbf{r})}{h|\mathbf{r}|} \right) = \nabla f(\mathbf{a}) \cdot \mathbf{r} \end{aligned}$$

siden $\lim_{h \rightarrow 0} \frac{\sigma(h\mathbf{r})}{h|\mathbf{r}|} = 0$. □

Vi tar med et eksempel på hvordan formel (2.4.3) kan brukes til å beregne retningsderiverte.

Eksempel 4: La oss anta at funksjonen

$$f(x, y, z) = x^2y + e^{-yz}$$

er deriverbar. Vi skal finne den retningsderiverte $f'(\mathbf{a}; \mathbf{r})$ der $\mathbf{a} = (1, 1, 1)$ og $\mathbf{r} = (1, -1, 1)$. La oss først finne gradienten til f :

$$\frac{\partial f}{\partial x} = 2xy; \quad \frac{\partial f}{\partial y} = x^2 - ze^{-yz}; \quad \frac{\partial f}{\partial z} = -ye^{-yz}$$

Dette gir $\nabla f = (2xy, x^2 - ze^{-yz}, -ye^{-yz})$ og $\nabla f(\mathbf{a}) = \nabla f(1, 1, 1) = (2, 1 - e^{-1}, -e^{-1})$. Følgelig er

$$f'(\mathbf{a}; \mathbf{r}) = \nabla f(\mathbf{a}) \cdot \mathbf{r} = (2, 1 - e^{-1}, -e^{-1}) \cdot (1, -1, 1) = 1$$



For å kunne bruke formel (2.4.3) trenger vi å vite at våre funksjoner er deriverbare. Det neste resultatet gir oss den informasjonen vi vanligvis trenger. Beviset er ganske langt og komplisert, og egner seg nok best for de ivrigste og flittigste. Et lite ord om terminologi: Vi sier at en funksjon er definert i en *omegn* om \mathbf{a} dersom det finnes en kule $B(\mathbf{a}, \epsilon)$ om \mathbf{a} der funksjonen er definert (den kan godt være definert på et større område — poenget er at vi i hvert fall vil sikre oss at den er definert for alle punkter tilstrekkelig nær \mathbf{a}).

Teorem 2.4.6 La $f : A \rightarrow \mathbb{R}^m$ være en funksjon av n variable. Anta at alle de partiellderiverte $\frac{\partial f}{\partial x_i}$ er definert i en omegn om $\mathbf{a} \in A$, og at de er kontinuertlige i \mathbf{a} . Da er f deriverbar i \mathbf{a} .

Bevis: For at ikke notasjonen skal bli for overveldende, skal vi nøye oss med å bevise setningen for en funksjon $f(x_1, x_2)$ av to variable. Beviset er en litt oppstrammet variant av det argumentet som ledet oss til formel (2.4.3).

Vi velger \mathbf{r} så liten at $\mathbf{a} + \mathbf{r}$ ligger innenfor det området hvor de partiellderiverte eksisterer. Vi lar $\mathbf{a} = (a_1, a_2)$, $\mathbf{r} = (r_1, r_2)$ og observerer at

$$\begin{aligned} f(\mathbf{a} + \mathbf{r}) - f(\mathbf{a}) &= f(a_1 + r_1, a_2 + r_2) - f(a_1, a_2) \\ &= f(a_1 + r_1, a_2 + r_2) - f(a_1 + r_1, a_2) + \\ &\quad + f(a_1 + r_1, a_2) - f(a_1, a_2) \end{aligned}$$

Dette er samme type omskrivning som vi foretok da vi regnet oss frem til (2.4.3). Hvis vi tenker på $x \rightarrow f(x, a_2)$ som en funksjon av én variabel, forteller middelverdisetningen oss at det finnes et punkt c mellom a_1 og $a_1 + r_1$ slik at

$$f(a_1 + r_1, a_2) - f(a_1, a_2) = \frac{\partial f}{\partial x_1}(c, a_2) r_1$$

Helt tilsvarende kan vi finne et punkt d mellom a_2 og $a_2 + r_2$ slik at

$$f(a_1 + r_1, a_2 + r_2) - f(a_1 + r_1, a_2) = \frac{\partial f}{\partial x_2}(a_1 + r_1, d) r_2$$

Kombinerer vi de resultatene vi nå har, ser vi at

$$f(\mathbf{a} + \mathbf{r}) - f(\mathbf{a}) = \frac{\partial f}{\partial x_1}(c, a_2) r_1 + \frac{\partial f}{\partial x_2}(a_1 + r_1, d) r_2$$

Trekker vi fra $\nabla f(\mathbf{a}) \cdot \mathbf{r} = \frac{\partial f}{\partial x_1}(a_1, a_2) r_1 + \frac{\partial f}{\partial x_2}(a_1, a_2) r_2$ på begge sider, får vi

$$\begin{aligned} f(\mathbf{a} + \mathbf{r}) - f(\mathbf{a}) - \nabla f(\mathbf{a}) \cdot \mathbf{r} &= \\ &= \frac{\partial f}{\partial x_1}(c, a_2) r_1 + \frac{\partial f}{\partial x_2}(a_1 + r_1, d) r_2 - \frac{\partial f}{\partial x_1}(a_1, a_2) r_1 - \frac{\partial f}{\partial x_2}(a_1, a_2) r_2 = \\ &= \left(\frac{\partial f}{\partial x_1}(c, a_2) - \frac{\partial f}{\partial x_1}(a_1, a_2) \right) r_1 + \left(\frac{\partial f}{\partial x_2}(a_1 + r_1, d) - \frac{\partial f}{\partial x_2}(a_1, a_2) \right) r_2 \end{aligned}$$

Sammenligner vi dette med definisjon 2.4.4, ser vi at

$$\sigma(\mathbf{r}) = \left(\frac{\partial f}{\partial x_1}(c, a_2) - \frac{\partial f}{\partial x_1}(a_1, a_2) \right) r_1 + \left(\frac{\partial f}{\partial x_2}(a_1 + r_1, d) - \frac{\partial f}{\partial x_2}(a_1, a_2) \right) r_2$$

Vår oppgave er å vise at $\lim_{\mathbf{r} \rightarrow \mathbf{0}} \frac{1}{|\mathbf{r}|} \sigma(\mathbf{r}) = \mathbf{0}$. Siden $|r_1|, |r_2| \leq |\mathbf{r}|$, får vi

$$|\sigma(\mathbf{r})| \leq \left(\left| \frac{\partial f}{\partial x_1}(c, a_2) - \frac{\partial f}{\partial x_1}(a_1, a_2) \right| + \left| \frac{\partial f}{\partial x_2}(a_1 + r_1, d) - \frac{\partial f}{\partial x_2}(a_1, a_2) \right| \right) |\mathbf{r}|$$

Deler vi på $|\mathbf{r}|$, ser vi at

$$\frac{|\sigma(\mathbf{r})|}{|\mathbf{r}|} \leq \left(\left| \frac{\partial f}{\partial x_1}(c, a_2) - \frac{\partial f}{\partial x_1}(a_1, a_2) \right| + \left| \frac{\partial f}{\partial x_2}(a_1 + r_1, d) - \frac{\partial f}{\partial x_2}(a_1, a_2) \right| \right)$$

Dette uttrykket går mot 0 fordi de partiellderiverte er kontinuerlige i $\mathbf{a} = (a_1, a_2)$, og (c, a_2) og $(a_1 + r_1, d)$ nærmer seg (a_1, a_2) når $\mathbf{r} \rightarrow \mathbf{0}$. Dermed er teoremet bevist. \square

Sammen med setning 2.4.5 forteller setningen ovenfor oss at så lenge de partiellderiverte er kontinuerlige, kan vi trygt bruke formelen

$$f'(\mathbf{a}; \mathbf{r}) = \nabla f(\mathbf{a}) \cdot \mathbf{r}.$$

Vi har allerede nevnt den geometriske tolkningen av gradienten — at gradienten i punktet \mathbf{a} peker i den retningen hvor funksjonen vokser hurtigst, og at stigningstallet i denne retningen er lik lengden til gradienten. Vi har nå de redskapene som trengs til å bevise dette.

Setning 2.4.7 *Anta at f er deriverbar i \mathbf{a} . Da peker gradienten $\nabla f(\mathbf{a})$ i den retningen hvor f vokser hurtigst i punktet \mathbf{a} , og stigningstallet til f i denne retningen er $|\nabla f(\mathbf{a})|$.*

Bevis: Hvis \mathbf{u} er en enhetsvektor (dvs. $|\mathbf{u}| = 1$), så forteller den retningsderiverte $f'(\mathbf{a}; \mathbf{u}) = \nabla f(\mathbf{a}) \cdot \mathbf{u}$ oss hvor fort funksjonen vokser i den retningen \mathbf{u} peker. Funksjonen f vokser derfor hurtigst i den retningen \mathbf{u} hvor $\nabla f(\mathbf{a}) \cdot \mathbf{u}$ er størst. Siden \mathbf{u} er en enhetsvektor, forteller Schwarz' ulikhet (setning 1.2.3) oss at

$$|\nabla f(\mathbf{a}) \cdot \mathbf{u}| \leq |\nabla f(\mathbf{a})| |\mathbf{u}| = |\nabla f(\mathbf{a})|$$

med likhet bare hvis \mathbf{u} og $\nabla f(\mathbf{a})$ er parallelle. Dette betyr at $|\nabla f(\mathbf{a}) \cdot \mathbf{u}|$ er størst når \mathbf{u} og $\nabla f(\mathbf{a})$ er parallelle. Nå er det to enhetsvektorer som er parallelle med $\nabla f(\mathbf{a})$ — en som peker samme vei som $\nabla f(\mathbf{a})$, og en som peker motsatt vei. Det er lett å se at $\nabla f(\mathbf{a}) \cdot \mathbf{u} = |\nabla f(\mathbf{a})|$ når \mathbf{u} og $\nabla f(\mathbf{a})$ peker samme vei, og at $\nabla f(\mathbf{a}) \cdot \mathbf{u} = -|\nabla f(\mathbf{a})|$ når \mathbf{u} og $\nabla f(\mathbf{a})$ peker motsatt vei. Altså har f sitt største stigningstall $|\nabla f(\mathbf{a})|$ i den retningen som $\nabla f(\mathbf{a})$ peker. \square

Eksempel 5: I hvilken retning vokser funksjonen $f(x, y) = x^3 y \sin(\pi xy)$ hurtigst når vi står i punktet $(1, \frac{1}{2})$?

Vi må først finne gradienten. De partiellderiverte er

$$\frac{\partial f}{\partial x} = 3x^2 y \sin(\pi xy) + x^3 y \cos(\pi xy)(\pi y) = 3x^2 y \sin(\pi xy) + \pi x^3 y^2 \cos(\pi xy)$$

$$\frac{\partial f}{\partial y} = x^3 \sin(\pi xy) + x^3 y \cos(\pi xy)(\pi x) = x^3 \sin(\pi xy) + \pi x^4 y \cos(\pi xy)$$

Setter vi inn $(x, y) = (1, \frac{1}{2})$, får vi

$$\frac{\partial f}{\partial x} \left(1, \frac{1}{2} \right) = 3 \cdot 1^2 \cdot \frac{1}{2} \cdot \sin(\pi \cdot 1 \cdot \frac{1}{2}) + \pi \cdot 1^3 \cdot \left(\frac{1}{2} \right)^2 \cdot \cos(\pi \cdot 1 \cdot \frac{1}{2}) = \frac{3}{2}$$

og

$$\frac{\partial f}{\partial y} \left(1, \frac{1}{2} \right) = 1^3 \cdot \sin(\pi \cdot 1 \cdot \frac{1}{2}) + \pi \cdot 1^4 \cdot \frac{1}{2} \cdot \cos(\pi \cdot 1 \cdot \frac{1}{2}) = 1$$

Dermed er

$$\nabla f \left(1, \frac{1}{2} \right) = \left(\frac{3}{2}, 1 \right)$$

Funksjonen vokser altså raskest i retningen $(\frac{3}{2}, 1)$, og stigningstallet i denne retningen er

$$|\nabla f \left(1, \frac{1}{2} \right)| = \sqrt{\left(\frac{3}{2} \right)^2 + 1^2} = \sqrt{\frac{9}{4} + 1} = \frac{1}{2} \sqrt{13}$$



Vi tar med et litt mer teoretisk resultat som vi får bruk for siden. For funksjoner av én variabel vet vi at dersom den deriverte eksisterer i et punkt, så er funksjonen kontinuert i punktet. For funksjoner av flere variable er sammenhengen litt mer subtil — det kan faktisk hende at alle de retningsderiverte eksisterer i et punkt, men at funksjonen likevel ikke er kontinuert i punktet (se oppgave 7). Er funksjonen deriverbar i betydningen vi innførte i definisjon 2.4.4, er vi imidlertid på den sikre siden.

Setning 2.4.8 *Anta at $f : A \rightarrow \mathbb{R}$ er en funksjon av n variable. Dersom f er deriverbar i et punkt $\mathbf{a} \in A$, så er f kontinuert i \mathbf{a} .*

Bevis: Ifølge setning 2.3.5 er det nok å vise at $\lim_{\mathbf{x} \rightarrow \mathbf{a}} f(\mathbf{x}) = f(\mathbf{a})$. Setter vi $\mathbf{r} = \mathbf{x} - \mathbf{a}$, er dette det samme som å vise at $\lim_{\mathbf{r} \rightarrow \mathbf{0}} f(\mathbf{a} + \mathbf{r}) = f(\mathbf{a})$. Siden f er deriverbar i \mathbf{a} , vet vi at

$$f(\mathbf{a} + \mathbf{r}) = f(\mathbf{a}) + \nabla f(\mathbf{a}) \cdot \mathbf{r} + \sigma(\mathbf{r})$$

der $\frac{\sigma(\mathbf{r})}{|\mathbf{r}|} \rightarrow 0$ når $\mathbf{r} \rightarrow \mathbf{0}$. Dette betyr spesielt at $\sigma(\mathbf{r}) \rightarrow 0$, og dermed er

$$\lim_{\mathbf{r} \rightarrow \mathbf{0}} f(\mathbf{a} + \mathbf{r}) = \lim_{\mathbf{r} \rightarrow \mathbf{0}} (f(\mathbf{a}) + \nabla f(\mathbf{a}) \cdot \mathbf{r} + \sigma(\mathbf{r})) = f(\mathbf{a}) + 0 + 0 = f(\mathbf{a})$$

□

Helt til slutt skal vi se på eksempel som viser en typisk anvendelse av gradienter og partiellderiverte.

Eksempel 6: Anta at en tynn gass oppbevares i en beholder der vi kan justere volumet V og temperaturen T . Det viser seg da at trykket P er proporsjonalt med temperaturen T og omvendt proporsjonalt med volumet V , dvs. $P = k \frac{T}{V}$ der k er en konstant som blant annet avhenger av hvor mye gass det er i beholderen (dette forutsetter egentlig at temperaturen måles i grader Kelvin og ikke Celsius, men det behøver vi ikke bry oss om her). Vi kan tenke på trykket P som en funksjon av T og V :

$$P(T, V) = k \frac{T}{V}$$

Partiellderiverer vi dette uttrykket, får vi

$$\frac{\partial P}{\partial T} = k \frac{1}{V} \quad \text{og} \quad \frac{\partial P}{\partial V} = -k \frac{T}{V^2}$$

Gradienten er dermed

$$\nabla P(T, V) = \left(k \frac{1}{V}, -k \frac{T}{V^2} \right)$$

Det er ofte naturlig å spørre hvor mye trykket endrer seg dersom vi gir temperaturen et lite tillegg ΔT og volumet et lite tillegg ΔV . Det eksakte uttrykket for dette tillegget er selvfølgelig

$$\Delta P = P(T + \Delta T, V + \Delta V) - P(T, V),$$

men denne differansen er tung å arbeide med. Vi vet imidlertid at den har en god tilnærming i skalarproduktet av gradienten ∇P med tilvekstvektoren $(\Delta T, \Delta V)$:

$$\begin{aligned} \Delta P &= P(T + \Delta T, V + \Delta V) - P(T, V) \approx \nabla P(T, V) \cdot (\Delta T, \Delta V) = \\ &= \frac{\partial P}{\partial T} \Delta T + \frac{\partial P}{\partial V} \Delta V = k \frac{1}{V} \Delta T - k \frac{T}{V^2} \Delta V \end{aligned}$$

Ved hjelp av denne formelen er det lett å anslå hvor mye trykket endrer seg når vi regulerer volumet og temperaturen.

La oss nå anta at vi regulerer temperaturen og volumet kontinuerlig, og at endringene ΔP , ΔT og ΔV har foregått i løpet av et lite tidsintervall Δt . Endringen per tidsenhet er da

$$\frac{\Delta P}{\Delta t} \approx k \frac{1}{V} \frac{\Delta T}{\Delta t} - k \frac{T}{V^2} \frac{\Delta V}{\Delta t}$$

med bedre tilnærming dess mindre Δt er. Lar vi $\Delta t \rightarrow 0$, får vi

$$P'(t) = k \frac{1}{V} T'(t) - k \frac{T}{V^2} V'(t)$$

Ved hjelp av denne formelen kan vi regne ut hvor fort trykket endrer seg dersom vi kjenner endringshastighetene til temperaturen og volumet. Formelen er en forsmak på *kjerneregelen for funksjoner av flere variable*. Vi skal se nærmere på denne regelen i seksjon 2.7. ♣

Oppgaver til seksjon 2.4

1. Finn de partiellderiverte til f .

- | | |
|------------------------------------|--|
| a) $f(x, y) = x^3y + 3xy^4$ | e) $f(x, y, z) = (x + y)e^{-z}$ |
| b) $f(x, y) = \frac{x^2 + x^3}{y}$ | f) $f(x, y, z) = \frac{z^2 \tan x}{1 + y^2}$ |
| c) $f(x, y) = \cos(x + y^2)$ | g) $f(x, y, z) = z \arctan(x + y)$ |
| d) $f(x, y) = x^2 \ln(xy^2)$ | h) $f(x, y, z, u) = (z^2 + u)e^{-x+3y}$ |

2. Finn gradienten til funksjonen:

- a) $f(x, y) = x^2y$
- b) $f(x, y, z) = x \cos(xy^2z)$
- c) $f(u, v, w) = we^{u \cos v}$
- d) $f(z_1, z_2, z_3) = z_3 \arctan(z_1z_2) + e^{z_3}$

3. Finn den retningsderiverte $f'(\mathbf{a}; \mathbf{r})$:

- a) $f(x, y) = 3xy + y^2$; $\mathbf{a} = (1, 2)$; $\mathbf{r} = (3, -1)$
- b) $f(x, y) = \ln(x + y^2)$; $\mathbf{a} = (1, 0)$; $\mathbf{r} = (-1, 1)$
- c) $f(x, y, z) = x^2y + z^2$; $\mathbf{a} = (1, 0, 1)$; $\mathbf{r} = (1, 1, -1)$
- d) $f(x, y, z) = z \sin(xy)$; $\mathbf{a} = (\frac{\pi}{2}, 1, 0)$; $\mathbf{r} = (2, 0, -1)$

4. I hvilken retning vokser funksjonen hurtigst i det angitte punktet

- a) $f(x, y) = -x^2y + 7y^3$, $\mathbf{a} = (4, -3)$
- b) $f(x, y, z) = (x^2 - y^2)e^z$; $\mathbf{a} = (1, -1, 3)$
- c) $f(x, y, z, u) = xuz^2 - y^2zu$; $\mathbf{a} = (1, 0, -2, 3)$

5. Volumet til en sylinder med radius r og høyde h er $V = \pi r^2 h$. Når høyden og radien varierer, kan vi tenke på dette som en funksjon i to variable $V(r, h) = \pi r^2 h$. Forklar at når radien endrer seg fra r til $r + \Delta r$ og høyden endrer seg fra h til $h + \Delta h$, så er endringen i V tilnærmet gitt ved

$$\Delta V \approx \frac{\partial V}{\partial r} \cdot \Delta r + \frac{\partial V}{\partial h} \cdot \Delta h = 2\pi r h \Delta r + \pi r^2 \Delta h.$$

Anta at du har en sylinder hvor du vet at radien ligger mellom 2 m og 2.05 m og hvor høyden ligger mellom 5 m og 5.05 m. Bruk formelen ovenfor til å anslå usikkerheten i volumet.

6. BMI (body mass index) er en indikator for undervekt og overvekt. For å finne din BMI tar du vekten din (målt i kilo) og deler på kvadratet av høyden din (målt i meter). Du kan tenke på BMI som en funksjon av to variable

$$f(v, h) = \frac{v}{h^2}$$

a) Vis at dersom Δv og Δh er små endringer i vekt og høyde, så er endringen i BMI gitt ved

$$\Delta f(v, h) \approx \frac{\Delta v}{h^2} - 2 \frac{v}{h^3} \Delta h$$

b) En tommelfingerregel sier at for hver ekstra centimeter du har i høyde, kan du “tåle” en ekstra kilo uten at BMI-en din endrer seg særlig. Bruk formelen i a) til å undersøke hvor godt dette passer for personer med forskjellig vekt og høyde.

7. Vi skal se mer på funksjonen $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ fra eksempel 2 i seksjon 2.2. Husk at denne funksjonen er gitt ved

$$f(x, y) = \begin{cases} \frac{x^2 y}{x^4 + y^2} & \text{for } (x, y) \neq (0, 0) \\ 0 & \text{for } (x, y) = (0, 0) \end{cases}$$

a) Vis at $\frac{\partial f}{\partial x}(0, 0) = 0$ og $\frac{\partial f}{\partial y}(0, 0) = 0$. Hva er $\nabla f(0, 0)$?

b) Bruk definisjonen av retningsderivert til å vise at $f(\mathbf{0}; \mathbf{r}) = \frac{r_1^2}{r_2}$ der $\mathbf{r} = (r_1, r_2)$, $r_2 \neq 0$.

c) Vis at for denne funksjonen gjelder ikke likheten $f(\mathbf{0}; \mathbf{r}) = \nabla f(\mathbf{0}) \cdot \mathbf{r}$.

d) Vis at alle de retningsderiverte til f eksisterer i $\mathbf{0}$, men at funksjonen hverken er kontinuerlig eller deriverbar i punktet.

2.5 Partiellderivate av høyere orden

Fra teorien for funksjoner av en variabel vet vi at det ofte er nyttig eller nødvendig å derivere mer enn én gang. Også i flervariabel teori er det ofte nyttig å arbeide med annenderiverte, tredjederiverte osv. Den store forskjellen er at vi har så mange flere måter å derivere på.

Eksempel 1: La $f(x, y) = x^2 y^3 + y^2$. Vi har to partiellderivate av første orden

$$\frac{\partial f}{\partial x} = 2xy^3 \quad \text{og} \quad \frac{\partial f}{\partial y} = 3x^2 y^2 + 2y$$

Når vi skal regne ut annenderiverte, har vi mange valg. Vi kan for eksempel derivere $\frac{\partial f}{\partial x}$ med hensyn på x en gang til:

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial x} \right) = \frac{\partial}{\partial x} (2xy^3) = 2y^3.$$

Vi kan også derivere $\frac{\partial f}{\partial x}$ med hensyn på y :

$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial x} \right) = \frac{\partial}{\partial y} (2xy^3) = 6xy^2.$$

I tillegg kan vi derivere $\frac{\partial f}{\partial y}$ med hensyn på både x og y :

$$\begin{aligned}\frac{\partial^2 f}{\partial x \partial y} &= \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial y} \right) = \frac{\partial}{\partial x} (3x^2 y^2 + 2y) = 6xy^2 \\ \frac{\partial^2 f}{\partial y^2} &= \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial y} \right) = \frac{\partial}{\partial y} (3x^2 y^2 + 2y) = 6x^2 y + 2\end{aligned}$$

Vi har altså fire annenordens partiellderivate for denne funksjonen: $\frac{\partial^2 f}{\partial x^2}$, $\frac{\partial^2 f}{\partial y \partial x}$, $\frac{\partial^2 f}{\partial x \partial y}$ og $\frac{\partial^2 f}{\partial y^2}$. ♣

Den generelle notasjonen skulle fremgå av eksemplet ovenfor —

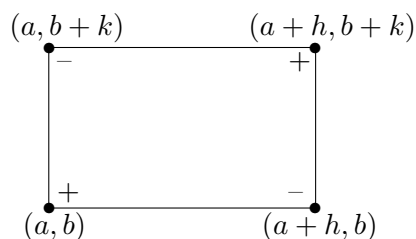
$$\frac{\partial^n f}{\partial x_{i_n} \dots \partial x_{i_2} \partial x_{i_1}}$$

er den funksjonen vi får ved å derivere funksjonen f n ganger, først med hensyn på variabelen x_{i_1} , så med hensyn på variabelen x_{i_2} osv.

I eksemplet ovenfor så vi at de to “blandede” partiellderivate $\frac{\partial^2 f}{\partial x \partial y}$ og $\frac{\partial^2 f}{\partial y \partial x}$ var like. Dette er ikke en universell regel; det finnes funksjoner f slik at $\frac{\partial^2 f}{\partial x \partial y}$ og $\frac{\partial^2 f}{\partial y \partial x}$ er forskjellige, men for de fleste vi støter på i praksis, vil de blandede partiellderivate være like. Den neste setningen viser at dette gjelder dersom de annenordens partiellderivate eksisterer i en omegn rundt punktet \mathbf{a} og er kontinuerlige i \mathbf{a} . Beviset er krevende og minner om beviset for teorem 2.4.6. I oppgave 4 finner du et eksempel på en funksjon der de blandede partiellderivate *ikke* er like.

Setning 2.5.1 La $f(x_1, \dots, x_n)$ være en funksjon av n variable. Anta at $\frac{\partial^2 f}{\partial x_i \partial x_j}$ og $\frac{\partial^2 f}{\partial x_j \partial x_i}$ eksisterer i en omegn om punktet \mathbf{a} og er kontinuerlige i \mathbf{a} . Da er $\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{a}) = \frac{\partial^2 f}{\partial x_j \partial x_i}(\mathbf{a})$.

Bevis: For å forenkle notasjonen antar vi at $f(x, y)$ er en funksjon av to variable, og at $\frac{\partial^2 f}{\partial x \partial y}$ og $\frac{\partial^2 f}{\partial y \partial x}$ eksisterer i en omegn om punktet (a, b) og er kontinuerlige i (a, b) . Anta at tallene h, k er så små at hele rektangelet i figur 1 ligger i det området der de blandede partiellderivate eksisterer.



Figur 1

La

$$\Delta(h, k) = f(a + h, b + k) - f(a, b + k) - f(a + h, b) + f(a, b)$$

der vi har kombinert funksjonsverdiene i hjørnene på rektanget vårt ved å bruke fortegnene vist på figuren. Vi skal vise at grenseverdien

$$\lim_{(h,k) \rightarrow (0,0)} \frac{\Delta(h, k)}{h \cdot k}$$

er lik både $\frac{\partial^2 f}{\partial x \partial y}(a, b)$ og $\frac{\partial^2 f}{\partial y \partial x}(a, b)$.

Først skritt er å bruke middelverdisetningen på funksjonen

$$g(x) = f(x, b + k) - f(x, b).$$

Vi får

$$g(a + h) - g(a) = g'(c) \cdot h$$

for en c mellom a og $a + h$. Setter vi inn den opprinnelige funksjonen, ser vi at

$$\begin{aligned} f(a + h, b + k) - f(a + h, b) - f(a, b + k) + f(a, b) &= \\ &= \left[\frac{\partial f}{\partial x}(c, b + k) - \frac{\partial f}{\partial x}(c, b) \right] h \end{aligned}$$

Dette kan også skrives

$$\Delta(h, k) = \left[\frac{\partial f}{\partial x}(c, b + k) - \frac{\partial f}{\partial x}(c, b) \right] h$$

Neste steg er å bruke middelverdisetningen på funksjonen

$$G(y) = \frac{\partial f}{\partial x}(c, y).$$

Vi får

$$G(b + k) - G(b) = G'(d) \cdot k$$

for en d mellom b og $b + k$. Dette kan også skrives

$$\frac{\partial f}{\partial x}(c, b + k) - \frac{\partial f}{\partial x}(c, b) = \frac{\partial^2 f}{\partial y \partial x}(c, d) \cdot k$$

Kombinerer vi formlene våre, ser vi at

$$\Delta(h, k) = \frac{\partial^2 f}{\partial y \partial x}(c, d) \cdot hk$$

Siden $\frac{\partial^2 f}{\partial y \partial x}$ er kontinuert i (a, b) , vil $\frac{\partial^2 f}{\partial y \partial x}(c, d) \rightarrow \frac{\partial^2 f}{\partial y \partial x}(a, b)$ når $(h, k) \rightarrow 0$. Følgelig er

$$\lim_{(h,k) \rightarrow (0,0)} \frac{\Delta(h, k)}{hk} = \frac{\partial^2 f}{\partial y \partial x}(a, b)$$

For å vise at også $\lim_{(h,k) \rightarrow (0,0)} \frac{\Delta(h,k)}{hk} = \frac{\partial^2 f}{\partial x \partial y}(a,b)$, bytter vi om på rollene til variablene x og y i argumentet ovenfor. Vi starter med å bruke middelverdisetningen på funksjonen

$$\gamma(y) = f(a+h, y) - f(a, y)$$

og fortsetter på akkurat samme måte som ovenfor. Detaljene overlates til leserne. \square

At blandede partiellderivate av annen orden er like, medfører også at blandede partiellderivate av høyere orden er like dersom de inneholder like mange derivasjoner med hensyn på hver variabel. Dersom f har kontinuerlige fjerdederivate, kan vi for eksempel vise at

$$\frac{\partial^4 f}{\partial x \partial y \partial z \partial x} = \frac{\partial^4 f}{\partial z \partial y \partial x \partial x}$$

på følgende måte:

$$\frac{\partial^4 f}{\partial x \partial y \partial z \partial x} = \frac{\partial^4 f}{\partial x \partial z \partial y \partial x} = \frac{\partial^4 f}{\partial z \partial x \partial y \partial x} = \frac{\partial^4 f}{\partial z \partial y \partial x \partial x}$$

Overbevis deg selv om at du kan begrunne disse overgangene.

Oppgaver til seksjon 2.5

1. Regn ut de annenordens partiellderivate:

- a) $f(x, y) = 3x^2y + 2y^2x$
- b) $f(x, y) = x \sin y$
- c) $f(x, y) = x^2 e^{x-y}$
- d) $f(x, y, z) = x^2 z - y^2 z^2$

2. Regn ut de partiellderivate:

- a) $\frac{\partial^3 f}{\partial x \partial z \partial x}$ når $f(x, y, z) = x^2 y e^{xz}$.
- b) $\frac{\partial^4 f}{\partial y \partial z \partial x \partial z}$ når $f(x, y, z) = x^2 y^3 \cos xyz$.

3. Gjennomfør den siste delen av beviset for setning 2.5.1 (dvs. at $\lim_{(h,k) \rightarrow (0,0)} \frac{\Delta(h,k)}{hk} = \frac{\partial^2 f}{\partial x \partial y}(a,b)$).

4. I denne oppgaven skal vi se på en funksjon $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ slik at $\frac{\partial^2 f}{\partial x \partial y}(0,0) \neq \frac{\partial^2 f}{\partial y \partial x}(0,0)$. Funksjonen er gitt ved

$$f(x, y) = \begin{cases} \frac{x^3 y - x y^3}{x^2 + y^2} & \text{når } (x, y) \neq (0, 0) \\ 0 & \text{når } (x, y) = (0, 0) \end{cases}$$

a) Vis at $f(x, 0) = 0$ for alle x og at $f(0, y) = 0$ for alle y . Bruk dette til å vise at $\frac{\partial f}{\partial x}(0, 0) = 0$ og $\frac{\partial f}{\partial y}(0, 0) = 0$.

b) Vis at for $(x, y) \neq (0, 0)$ er

$$\frac{\partial f}{\partial x}(x, y) = \frac{y(x^4 + 4x^2y^2 - y^4)}{(x^2 + y^2)^2}$$

$$\frac{\partial f}{\partial y}(x, y) = -\frac{x(y^4 + 4x^2y^2 - x^4)}{(x^2 + y^2)^2}$$

c) Vis at $\frac{\partial^2 f}{\partial y \partial x}(0, 0) = -1$ ved å bruke

$$\frac{\partial^2 f}{\partial y \partial x}(0, 0) = \lim_{h \rightarrow 0} \frac{\frac{\partial f}{\partial x}(0, h) - \frac{\partial f}{\partial x}(0, 0)}{h}$$

Vis på tilsvarende måte at $\frac{\partial^2 f}{\partial x \partial y}(0, 0) = 1$.

2.6 Derivasjon av vektorvaluerte funksjoner

Hittil har vi sett på derivasjon av funksjoner som tar verdier i \mathbb{R} . Vi skal nå utvide teorien vår til også å omfatte funksjoner som tar verdier i \mathbb{R}^m der m er større enn 1, altså til *vektorvaluerte* funksjoner. Husk at en funksjon $\mathbf{F} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ kan skrives på komponentform

$$\mathbf{F}(\mathbf{x}) = \begin{pmatrix} F_1(\mathbf{x}) \\ F_2(\mathbf{x}) \\ \vdots \\ F_m(\mathbf{x}) \end{pmatrix}$$

der F_1, F_2, \dots, F_m er funksjoner med verdier i \mathbb{R} (i denne seksjonen er det lurt å tenke på de fleste vektorer som søylevektorer siden vi skal multiplisere dem med matriser). Dersom alle disse funksjonen lar seg derivere, kan vi samle alle de partiellderiverte i en stor matrise, den såkalte *Jacobi-matrisen* til \mathbf{F} :

$$\mathbf{F}'(\mathbf{a}) = \begin{pmatrix} \frac{\partial F_1}{\partial x_1}(\mathbf{a}) & \frac{\partial F_1}{\partial x_2}(\mathbf{a}) & \cdots & \frac{\partial F_1}{\partial x_n}(\mathbf{a}) \\ \frac{\partial F_2}{\partial x_1}(\mathbf{a}) & \frac{\partial F_2}{\partial x_2}(\mathbf{a}) & \cdots & \frac{\partial F_2}{\partial x_n}(\mathbf{a}) \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial F_m}{\partial x_1}(\mathbf{a}) & \frac{\partial F_m}{\partial x_2}(\mathbf{a}) & \cdots & \frac{\partial F_m}{\partial x_n}(\mathbf{a}) \end{pmatrix}$$

Legg merke til hvordan denne matrisen er bygget opp: I første linje har vi alle partiellderiverte av førstekomponent F_1 , i annen linje alle partiellderiverte

av annenkomponent F_2 osv. Sagt på en annen måte: Første linje i Jacobi-matrisen er gradienten til F_1 , andre linje er gradienten til F_2 osv. Dersom \mathbf{F} er et skalarfelt (og altså tar verdier i \mathbb{R}), har Jacobi-matrisen bare én linje og er identisk med gradienten $\nabla \mathbf{F}$.

Jacobi-matrisen spiller på mange måter den samme rollen for funksjoner av flere variable som den deriverte gjør for funksjoner av én variabel, og vi har derfor valgt en notasjon $\mathbf{F}'(\mathbf{a})$ som ligner på den notasjonen vi er vant til.

Hvis man først har lært seg å partiellderivere, er det ingen kunst (men en del arbeid!) å finne en Jacobi-matrise.

Eksempel 1: Finn Jacobi-matrisen til funksjonen

$$\mathbf{F}(x, y) = \begin{pmatrix} xy^3 \\ e^{x+y^2} \\ 3x^2y \end{pmatrix}$$

I dette tilfellet er

$$F_1(x, y) = xy^3, \quad F_2(x, y) = e^{x+y^2} \quad \text{og} \quad F_3(x, y) = 3x^2y$$

Vi partiellderiverer og får

$$\begin{aligned} \frac{\partial F_1}{\partial x} &= y^3 & \text{og} & \quad \frac{\partial F_1}{\partial y} = 3xy^2 \\ \frac{\partial F_2}{\partial x} &= e^{x+y^2} & \text{og} & \quad \frac{\partial F_2}{\partial y} = 2ye^{x+y^2} \\ \frac{\partial F_3}{\partial x} &= 6xy & \text{og} & \quad \frac{\partial F_3}{\partial y} = 3x^2 \end{aligned}$$

Jacobi-matrisen blir dermed

$$\mathbf{F}'(x, y) = \begin{pmatrix} y^3 & 3xy^2 \\ e^{x+y^2} & 2ye^{x+y^2} \\ 6xy & 3x^2 \end{pmatrix}$$



Vi skal nå forsøke å finne ut hvordan vi bør definere deriverbarhet for vektorvaluerte funksjoner \mathbf{F} . La oss gå tilbake til den generelle Jacobi-matrisen

$$\mathbf{F}'(\mathbf{a}) = \begin{pmatrix} \frac{\partial F_1}{\partial x_1}(\mathbf{a}) & \frac{\partial F_1}{\partial x_2}(\mathbf{a}) & \cdots & \frac{\partial F_1}{\partial x_n}(\mathbf{a}) \\ \frac{\partial F_2}{\partial x_1}(\mathbf{a}) & \frac{\partial F_2}{\partial x_2}(\mathbf{a}) & \cdots & \frac{\partial F_2}{\partial x_n}(\mathbf{a}) \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial F_m}{\partial x_1}(\mathbf{a}) & \frac{\partial F_m}{\partial x_2}(\mathbf{a}) & \cdots & \frac{\partial F_m}{\partial x_n}(\mathbf{a}) \end{pmatrix}$$

Ganger vi denne med en søylevektor

$$\mathbf{r} = \begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{pmatrix}$$

får vi

$$\mathbf{F}'(\mathbf{a})\mathbf{r} = \begin{pmatrix} \frac{\partial F_1}{\partial x_1}(\mathbf{a}) & \frac{\partial F_1}{\partial x_2}(\mathbf{a}) & \cdots & \frac{\partial F_1}{\partial x_n}(\mathbf{a}) \\ \frac{\partial F_2}{\partial x_1}(\mathbf{a}) & \frac{\partial F_2}{\partial x_2}(\mathbf{a}) & \cdots & \frac{\partial F_2}{\partial x_n}(\mathbf{a}) \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial F_m}{\partial x_1}(\mathbf{a}) & \frac{\partial F_m}{\partial x_2}(\mathbf{a}) & \cdots & \frac{\partial F_m}{\partial x_n}(\mathbf{a}) \end{pmatrix} \begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{pmatrix} = \begin{pmatrix} \nabla F_1(\mathbf{a}) \cdot \mathbf{r} \\ \nabla F_2(\mathbf{a}) \cdot \mathbf{r} \\ \vdots \\ \nabla F_m(\mathbf{a}) \cdot \mathbf{r} \end{pmatrix}$$

der vi har brukt at første rad i $\mathbf{F}'(\mathbf{a})$ er gradienten til F_1 , andre rad er gradienten til F_2 osv. Dersom F_1, F_2, \dots, F_m er deriverbare funksjoner, er

$$\nabla F_1(\mathbf{a}) \cdot \mathbf{r} \approx F_1(\mathbf{a} + \mathbf{r}) - F_1(\mathbf{a})$$

$$\nabla F_2(\mathbf{a}) \cdot \mathbf{r} \approx F_2(\mathbf{a} + \mathbf{r}) - F_2(\mathbf{a})$$

$$\vdots \quad \vdots \quad \vdots \quad \vdots$$

$$\nabla F_m(\mathbf{a}) \cdot \mathbf{r} \approx F_m(\mathbf{a} + \mathbf{r}) - F_m(\mathbf{a})$$

med bedre tilnærming dess mindre \mathbf{r} er. Altså er

$$\mathbf{F}'(\mathbf{a})\mathbf{r} \approx \begin{pmatrix} F_1(\mathbf{a} + \mathbf{r}) - F_1(\mathbf{a}) \\ F_2(\mathbf{a} + \mathbf{r}) - F_2(\mathbf{a}) \\ \vdots \\ F_m(\mathbf{a} + \mathbf{r}) - F_m(\mathbf{a}) \end{pmatrix} = \mathbf{F}(\mathbf{a} + \mathbf{r}) - \mathbf{F}(\mathbf{a})$$

Dette gir oss en indikasjon på hvordan vi skal definere deriverbarhet for vektorvaluerte funksjoner — vi ønsker at avviket i tilnærmingen

$$\mathbf{F}(\mathbf{a} + \mathbf{r}) - \mathbf{F}(\mathbf{a}) \approx \mathbf{F}'(\mathbf{a})\mathbf{r}$$

skal være lite sammenlignet med størrelsen til \mathbf{r} :

Definisjon 2.6.1 Anta at $\mathbf{F} : A \rightarrow \mathbb{R}^m$ er en funksjon av n variable og at \mathbf{a} er et indre punkt i A . Vi sier at \mathbf{F} er deriverbar i \mathbf{a} dersom funksjonen

$$\boldsymbol{\sigma}(\mathbf{r}) = \mathbf{F}(\mathbf{a} + \mathbf{r}) - \mathbf{F}(\mathbf{a}) - \mathbf{F}'(\mathbf{a})\mathbf{r}$$

går mot null fortere enn $|\mathbf{r}|$, dvs. at

$$\lim_{\mathbf{r} \rightarrow \mathbf{0}} \frac{1}{|\mathbf{r}|} \boldsymbol{\sigma}(\mathbf{r}) = \mathbf{0}$$

(legg merke til at $\boldsymbol{\sigma}$ nå er en vektorvaluert funksjon med verdier i \mathbb{R}^m).

Mange vil kanskje frykte at vi nå må begynne helt forfra med å utvikle en teori for deriverbarhet akkurat som vi gjorde for skalarfelt i seksjon 2.4, men takket være den neste setningen er det unødvendig.

Setning 2.6.2 *En funksjon $\mathbf{F} : A \rightarrow \mathbb{R}^m$ er deriverbar i et indre punkt $\mathbf{a} \in A$ hvis og bare hvis hver komponent F_i er deriverbar i \mathbf{a} .*

Bevis: Vi har

$$\begin{aligned}\boldsymbol{\sigma}(\mathbf{r}) &= \mathbf{F}(\mathbf{a} + \mathbf{r}) - \mathbf{F}(\mathbf{a}) - \mathbf{F}'(\mathbf{a})\mathbf{r} = \\ &= \begin{pmatrix} F_1(\mathbf{a} + \mathbf{r}) - F_1(\mathbf{a}) - \nabla F_1(\mathbf{a}) \cdot \mathbf{r} \\ F_2(\mathbf{a} + \mathbf{r}) - F_2(\mathbf{a}) - \nabla F_2(\mathbf{a}) \cdot \mathbf{r} \\ \vdots \\ F_m(\mathbf{a} + \mathbf{r}) - F_m(\mathbf{a}) - \nabla F_m(\mathbf{a}) \cdot \mathbf{r} \end{pmatrix} = \begin{pmatrix} \sigma_1(\mathbf{r}) \\ \sigma_2(\mathbf{r}) \\ \vdots \\ \sigma_m(\mathbf{r}) \end{pmatrix}\end{aligned}$$

Dermed er

$$\frac{1}{|\mathbf{r}|}\boldsymbol{\sigma}(\mathbf{r}) = \begin{pmatrix} \frac{\sigma_1(\mathbf{r})}{|\mathbf{r}|} \\ \frac{\sigma_2(\mathbf{r})}{|\mathbf{r}|} \\ \vdots \\ \frac{\sigma_m(\mathbf{r})}{|\mathbf{r}|} \end{pmatrix}$$

og vi vet fra setning 2.3.3 at $\frac{1}{|\mathbf{r}|}\boldsymbol{\sigma}(\mathbf{r}) \rightarrow \mathbf{0}$ hvis og bare hvis $\frac{\sigma_i(\mathbf{r})}{|\mathbf{r}|} \rightarrow 0$ for alle i . Dette betyr at \mathbf{F} er deriverbar hvis og bare hvis hver F_i er deriverbar. \square

Kombinerer vi dette resultatet med setning 2.4.6, får vi:

Korollar 2.6.3 *Anta at $\mathbf{F} : A \rightarrow \mathbb{R}^m$ er en funksjon av n variable og at \mathbf{a} er et indre punkt i A . Dersom alle komponentene $\frac{\partial F_i}{\partial x_j}$ i Jacobi-matrisen er definert i en omegn rundt \mathbf{a} og er kontinuertlige i \mathbf{a} , så er \mathbf{F} deriverbar i \mathbf{a} .*

Bevis: Ifølge setning 2.4.6 er hver av komponentene F_i til \mathbf{F} deriverbar i \mathbf{a} , og ifølge setningen ovenfor er da \mathbf{F} deriverbar i \mathbf{a} . \square

Eksempel 2: Vis at funksjonen

$$\mathbf{F}(x, y) = \begin{pmatrix} xy^3 \\ e^{x+y^2} \\ 3x^2y \end{pmatrix}$$

fra eksempel 1 er deriverbar.

Vi har allerede regnet ut Jacobi-matrisen til \mathbf{F} , og komponentene er åpenbart kontinuertlige overalt. Altså er \mathbf{F} deriverbar ifølge korollaret. \square

Den neste setningen kan se litt underlig ut, men vi kommer til å ha stor glede av den. Setningen sier at ingen annen matrise B kan “stjele jobben” til Jacobi-matrisen $\mathbf{F}'(\mathbf{a})$ — dersom B tilfredsstiller en betingelse av samme type som $\mathbf{F}'(\mathbf{a})$ oppfyller i definisjon 2.6.1, så må B være lik $\mathbf{F}'(\mathbf{a})$. Det finnes også andre tolkninger av setningen som vi skal komme tilbake til senere.

Setning 2.6.4 La $\mathbf{F} : A \rightarrow \mathbb{R}^m$ være en funksjon av n variable og la \mathbf{a} være et indre punkt i A . Anta at det finnes en $m \times n$ -matrise B slik at funksjonen

$$\hat{\sigma}(\mathbf{r}) = \mathbf{F}(\mathbf{a} + \mathbf{r}) - \mathbf{F}(\mathbf{a}) - B\mathbf{r}$$

tilfredsstiller

$$\lim_{\mathbf{r} \rightarrow \mathbf{0}} \frac{1}{|\mathbf{r}|} \hat{\sigma}(\mathbf{r}) = \mathbf{0}.$$

Da er \mathbf{F} deriverbar i \mathbf{a} og

$$\mathbf{F}'(\mathbf{a}) = B$$

Bevis: Vi velger $\mathbf{r} = r\mathbf{e}_i$ der

$$\mathbf{e}_i = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$$

er den i -te enhetsvektoren og der $r \in \mathbb{R}$. Da er

$$\mathbf{F}(\mathbf{a} + r\mathbf{e}_i) - \mathbf{F}(\mathbf{a}) = B(r\mathbf{e}_i) + \hat{\sigma}(r\mathbf{e}_i)$$

Deler vi på r og bruker at $|r\mathbf{e}_i| = r$, får vi

$$\frac{\mathbf{F}(\mathbf{a} + r\mathbf{e}_i) - \mathbf{F}(\mathbf{a})}{r} = B\mathbf{e}_i + \frac{\hat{\sigma}(r\mathbf{e}_i)}{|r\mathbf{e}_i|}$$

Når $r \rightarrow 0$, vil det siste leddet på høyre side gå mot $\mathbf{0}$ ifølge antagelsen, så høyresiden går mot

$$B\mathbf{e}_i = \begin{pmatrix} b_{1i} \\ b_{2i} \\ \vdots \\ b_{mi} \end{pmatrix}$$

Da må venstresiden konvergere mot det samme, og vi får

$$\lim_{r \rightarrow 0} \frac{\mathbf{F}(\mathbf{a} + r\mathbf{e}_i) - \mathbf{F}(\mathbf{a})}{r} = \begin{pmatrix} \lim_{r \rightarrow 0} \frac{F_1(\mathbf{a} + r\mathbf{e}_i) - F_1(\mathbf{a})}{r} \\ \lim_{r \rightarrow 0} \frac{F_2(\mathbf{a} + r\mathbf{e}_i) - F_2(\mathbf{a})}{r} \\ \vdots \\ \lim_{r \rightarrow 0} \frac{F_m(\mathbf{a} + r\mathbf{e}_i) - F_m(\mathbf{a})}{r} \end{pmatrix} = \begin{pmatrix} b_{1i} \\ b_{2i} \\ \vdots \\ b_{mi} \end{pmatrix}$$

Ved definisjonen av partiellderivate betyr dette at $\frac{\partial F_1}{\partial x_i} = b_{1i}$, $\frac{\partial F_2}{\partial x_i} = b_{2i}$ og så videre. Altså eksisterer alle de partiellderivate og $\mathbf{F}'(\mathbf{a}) = B$. Deriverbarhet følger nå av at $\lim_{\mathbf{r} \rightarrow \mathbf{0}} \frac{1}{|\mathbf{r}|} \hat{\boldsymbol{\sigma}}(\mathbf{r}) = \mathbf{0}$. \square

Oppgaver til seksjon 2.6

1. Finn Jacobi-matrisen til funksjonen.

a) $\mathbf{F}(x, y) = (x^2y, x + y^2)$.

b) $\mathbf{F}(x, y, z) = (e^{x^2y+z}, xyz^2)$.

c) $\mathbf{F}(x, y) = (x \arctan(xy), x \ln y, xy \cos y^2)$.

d) $\mathbf{F}(x, y, z, u) = (xy \sin(xu^2), z^2u)$.

2.7 Kjerneregelen

For funksjoner av én variabel sier kjerneregelen at den deriverte til den sammensatte funksjonen $h(x) = f(g(x))$ er gitt ved $h'(x) = f'(g(x))g'(x)$. Kjerneregelen for funksjoner av flere variable har akkurat samme form; dersom \mathbf{H} er den sammensatte funksjonen $\mathbf{H}(\mathbf{x}) = \mathbf{F}(\mathbf{G}(\mathbf{x}))$, så er

$$\mathbf{H}'(\mathbf{x}) = \mathbf{F}'(\mathbf{G}(\mathbf{x}))\mathbf{G}'(\mathbf{x}) \quad (2.7.1)$$

Selv om formen er den samme, er innholdet mer komplisert — uttrykkene $\mathbf{H}'(\mathbf{x})$, $\mathbf{F}'(\mathbf{G}(\mathbf{x}))$ og $\mathbf{G}'(\mathbf{x})$ er nå matriser, og formelen sier at Jacobi-matrisen til \mathbf{H} i punktet \mathbf{x} er lik matriseproduktet av Jacobi-matrisen til \mathbf{F} i punktet $\mathbf{G}(\mathbf{x})$ og Jacobi-matrisen til \mathbf{G} i punktet \mathbf{x} .

Formel (2.7.1) uttrykker kjerneregelen i en kortfattet form som er lett å huske, men som ikke er så praktisk å bruke i utregninger. Vi skal derfor nærme oss kjerneregelen fra en litt annen synsvinkel og heller komme tilbake til matriseformuleringen i (2.7.1) etter hvert. Vi tenker oss at vi har en funksjon (et skalarfelt) $f(u_1, u_2, \dots, u_m)$ av m variable. For hver variabel u_i

substituerer vi en funksjon $g_i(x_1, x_2, \dots, x_n)$ slik at vi får en sammensatt funksjon

$$h(x_1, x_2, \dots, x_n) = f(g_1(x_1, x_2, \dots, x_n), g_2(x_1, x_2, \dots, x_n), \dots, g_m(x_1, x_2, \dots, x_n))$$

av n variable. Spørsmålet er: dersom vi kjenner alle de partiellderivate til de opprinnelige funksjonene f, g_1, g_2, \dots, g_m , kan vi da finne de partiellderivate til den sammensatte funksjonen h ? Svaret er ja, men før vi skriver opp formelen som viser oss hvordan vi kan regne ut de partiellderivate til h , er det lurt å komprimere notasjonen litt slik at vi slipper å arbeide med altfor lange uttrykk. Dersom vi bruker vektornotasjon og skriver

$$\mathbf{u} = (u_1, u_2, \dots, u_m)$$

$$\mathbf{x} = (x_1, x_2, \dots, x_n)$$

og

$$\mathbf{g}(\mathbf{x}) = (g_1(x_1, x_2, \dots, x_n), g_2(x_1, x_2, \dots, x_n), \dots, g_m(x_1, x_2, \dots, x_n)),$$

kan definisjonen av h skrives

$$h(\mathbf{x}) = f(\mathbf{g}(\mathbf{x}))$$

Dersom \mathbf{g} er deriverbar i punktet \mathbf{x} og f er deriverbar i punktet $\mathbf{u} = \mathbf{g}(\mathbf{x})$, viser det seg at de partiellderivate til h er gitt ved

$$\frac{\partial h}{\partial x_i}(\mathbf{x}) = \frac{\partial f}{\partial u_1}(\mathbf{u}) \frac{\partial g_1}{\partial x_i}(\mathbf{x}) + \frac{\partial f}{\partial u_2}(\mathbf{u}) \frac{\partial g_2}{\partial x_i}(\mathbf{x}) + \dots + \frac{\partial f}{\partial u_m}(\mathbf{u}) \frac{\partial g_m}{\partial x_i}(\mathbf{x}) \quad (2.7.2)$$

Denne formelen kaller vi *kjerneregelen på komponentform*, mens (2.7.1) kalles *kjerneregelen på matriseform* — vi skal senere se at selv om formen er forskjellig, så uttrykker de to formlene det samme.

Legg merke til at vi i (2.7.2) evaluerer de partiellderivate til f i punktet $\mathbf{u} = \mathbf{g}(\mathbf{x})$; det betyr at etter at vi har regnet ut de partiellderivate til f , må vi erstatte u_1 med $g_1(x_1, \dots, x_n)$, u_2 med $g_2(x_1, \dots, x_n)$ osv. Legg også merke til at det er et system i hvilke derivasjoner vi gjør i formel (2.7.2); vi deriverer med hensyn på alle “mellomvariablene” u_1, u_2, \dots, u_m , men bare med hensyn på én av “grunnvariablene” x_1, x_2, \dots, x_n , nemlig den x_i -en som inngår i den opprinnelige derivasjonen av h .

Ting blir klarere hvis vi ser på et eksempel:

Eksempel 1: Vi skal se på tilfellet der

$$f(u_1, u_2) = 2u_1u_2^2$$

og

$$g_1(x_1, x_2, x_3) = x_1x_2 \sin x_3$$

$$g_2(x_1, x_2, x_3) = 3x_1^2 x_2 x_3$$

Vi finner den sammensatte funksjonen $h = f(\mathbf{g}(\mathbf{x}))$ ved å substituere $u_1 = g_1(x_1, x_2, x_3) = x_1 x_2 \sin x_3$ og $u_2 = g_2(x_1, x_2, x_3) = 3x_1^2 x_2 x_3$ inn i uttrykket for f :

$$\begin{aligned} h(x_1, x_2, x_3) &= f(u_1, u_2) = 2u_1 u_2^2 = \\ &= 2(x_1 x_2 \sin x_3)(3x_1^2 x_2 x_3)^2 = 18x_1^5 x_2^3 x_3^2 \sin x_3 \end{aligned}$$

Vi kan selvfølgelig finne de partiellderiverte til h ved å derivere dette uttrykket på vanlig måte, men la oss se hvordan vi kan bruke kjernerregelen isteden. Ifølge (2.7.2) er

$$\frac{\partial h}{\partial x_1}(\mathbf{x}) = \frac{\partial f}{\partial u_1}(\mathbf{u}) \frac{\partial g_1}{\partial x_1}(\mathbf{x}) + \frac{\partial f}{\partial u_2}(\mathbf{u}) \frac{\partial g_2}{\partial x_1}(\mathbf{x})$$

Regner vi ut de partiellderiverte til f , g_1 og g_2 og setter inn i dette uttrykket, får vi

$$\frac{\partial h}{\partial x_1}(\mathbf{x}) = (2u_2^2)(x_2 \sin x_3) + (4u_1 u_2)(6x_1 x_2 x_3)$$

Helt til slutt setter vi inn uttrykkene $u_1 = g_1(x_1, x_2, x_3) = x_1 x_2 \sin x_3$ og $u_2 = g_2(x_1, x_2, x_3) = 3x_1^2 x_2 x_3$:

$$\begin{aligned} \frac{\partial h}{\partial x_1}(\mathbf{x}) &= (2(3x_1^2 x_2 x_3)^2)(x_2 \sin x_3) + (4(x_1 x_2 \sin x_3)(3x_1^2 x_2 x_3))(6x_1 x_2 x_3) = \\ &= 18x_1^4 x_2^3 x_3^2 \sin x_3 + 72x_1^4 x_2^3 x_3^2 \sin x_3 = 90x_1^4 x_2^3 x_3^2 \sin x_3 \end{aligned}$$

Det er god trening i å finne $\frac{\partial h}{\partial x_2}(\mathbf{x})$ og $\frac{\partial h}{\partial x_3}(\mathbf{x})$ på tilsvarende måte. ♣

Bemerkning: Hvis du ser nøyer på eksemplet ovenfor, vil du oppdage at vi har regnet ut $\frac{\partial h}{\partial x_1}(\mathbf{x})$ på en svært tungvinn måte — det går mye raske-
re å bare derivere uttrykket $h(x_1, x_2, x_3) = 18x_1^5 x_2^3 x_3^2 \sin x_3$ på vanlig måte! Dette er ganske typisk; kjernerregelen i flere variable er faktisk ikke en særlig effektiv metode for å regne ut partiellderiverte til funksjoner gitt ved formler — regelens styrke ligger isteden i at den er et utmerket redskap til å utlede *generelle* sammenhenger mellom varierende størrelser. Før vi ser nærmere på dette, kan det være lurt å ta med enda et konkret eksempel slik at vi blir litt bedre kjent med kjernerregelen.

Eksempel 2: I eksemplet ovenfor het de variable u_1, u_2 og x_1, x_2, x_3 akkurat som i vår generelle formel (2.7.2). Når man bruker kjernerregelen i andre fag, har de variable vanligvis helt andre navn, og det er viktig at vi da greier å kjenne igjen de forskjellige delene av kjernerregelen. La oss nå tenke oss at vi har en funksjon

$$f(s, t) = e^{st^2}$$

og at s og t er funksjoner av underliggende variable u og v

$$s = S(u, v) = 3uv^3 \quad \text{og} \quad t = T(u, v) = \cos(uv)$$

Vi vil bruke kjerneregelen til å finne de partiellderiverte til den sammensatte funksjonen

$$h(u, v) = f(S(u, v), T(u, v))$$

I dette tilfellet sier formel (2.7.2)

$$\frac{\partial h}{\partial u}(u, v) = \frac{\partial f}{\partial s}(s, t) \frac{\partial S}{\partial u}(u, v) + \frac{\partial f}{\partial t}(s, t) \frac{\partial T}{\partial u}(u, v)$$

og

$$\frac{\partial h}{\partial v}(u, v) = \frac{\partial f}{\partial s}(s, t) \frac{\partial S}{\partial v}(u, v) + \frac{\partial f}{\partial t}(s, t) \frac{\partial T}{\partial v}(u, v)$$

(husk at vi skal derivere med hensyn på begge “mellomvariablene” s og t , men bare med hensyn på den “opprinnelige” grunnvariablen’ u eller v). Regner vi ut de partiellderiverte og setter inn $s = S(u, v) = 3uv^3$ og $t = T(u, v) = \cos(uv)$, får vi

$$\begin{aligned} \frac{\partial h}{\partial u}(u, v) &= e^{st^2} 3v^3 + 2e^{st}(-\sin(uv)v) = \\ &= e^{3uv^3} (\cos(uv))^2 3v^3 - 2e^{3uv^3} \cos(uv) \sin(uv)v = \\ &= ve^{3uv^3} \cos(uv) \left(3v^2 \cos(uv) - 2 \sin(uv) \right) \end{aligned}$$

Den andre partiellderiverte $\frac{\partial h}{\partial v}(u, v)$ finner vi på tilsvarende måte. □

Vi skal ta med et par eksempler på hvordan kjerneregelen brukes i andre fag. Før vi gjør det, kan det være greit med noen ord om notasjon. Når vi bruker kjerneregelen i praktiske situasjoner, blir uttrykkene ofte lange og uoversiktlige. For å lette lesbarheten er det vanlig å utelate punktene som funksjonene evalueres i — man skriver altså

$$\frac{\partial h}{\partial x_i} = \frac{\partial f}{\partial u_1} \frac{\partial g_1}{\partial x_i} + \frac{\partial f}{\partial u_2} \frac{\partial g_2}{\partial x_i} + \dots + \frac{\partial f}{\partial u_m} \frac{\partial g_m}{\partial x_i}$$

istedenfor

$$\frac{\partial h}{\partial x_i}(\mathbf{x}) = \frac{\partial f}{\partial u_1}(\mathbf{u}) \frac{\partial g_1}{\partial x_i}(\mathbf{x}) + \frac{\partial f}{\partial u_2}(\mathbf{u}) \frac{\partial g_2}{\partial x_i}(\mathbf{x}) + \dots + \frac{\partial f}{\partial u_m}(\mathbf{u}) \frac{\partial g_m}{\partial x_i}(\mathbf{x})$$

og setter først inn verdiene til \mathbf{x} og \mathbf{u} når man har behov for dem.

Eksempel 3: I økonomi er utbyttet til et firma avhengig av forskjellige faktorer. For en bonde kan vi tenke oss at utbyttet er avhengig av tre faktorer: arbeidsinnsatsen a , den løpende kapitalinvesteringen k , og arealet av dyrkbar mark m . Utbyttet U er dermed en funksjon

$$U = F(a, k, m)$$

Vi tenker oss at de tre faktorene varierer med tiden: $a = A(t)$, $k = K(t)$ og $m = M(t)$. Kjernerregelen forteller oss nå hvordan utbyttet U varierer med tiden:

$$\frac{\partial U}{\partial t} = \frac{\partial F}{\partial a} \frac{\partial A}{\partial t} + \frac{\partial F}{\partial k} \frac{\partial K}{\partial t} + \frac{\partial F}{\partial m} \frac{\partial M}{\partial t}$$

der vi må huske at vi skal evaluere de partiellderiverte $\frac{\partial F}{\partial a}$, $\frac{\partial F}{\partial k}$, $\frac{\partial F}{\partial m}$ i punktet $(A(t), K(t), M(t))$. Legg merke til at i denne formelen er U , A , K og M bare avhengige av én variabel t , og vi kan derfor skrive vanlige deriverte istedenfor partiellderiverte:

$$U'(t) = \frac{\partial F}{\partial a} A'(t) + \frac{\partial F}{\partial k} K'(t) + \frac{\partial F}{\partial m} M'(t)$$

I økonomiske modeller antar man ofte at utbyttefunksjonen F er en *Cobb-Douglasfunksjon*, dvs. at den har formen

$$F(a, k, m) = Ca^\alpha k^\beta m^\gamma$$

der C , α , β og γ er positive konstanter. Deriverer vi dette uttrykket, får vi

$$\frac{\partial F}{\partial a} = C\alpha a^{\alpha-1} k^\beta m^\gamma, \quad \frac{\partial F}{\partial k} = C\beta a^\alpha k^{\beta-1} m^\gamma, \quad \frac{\partial F}{\partial m} = C\gamma a^\alpha k^\beta m^{\gamma-1}$$

som innsatt i uttrykket ovenfor gir (etter litt opprydning):

$$U'(t) = CA(t)^\alpha K(t)^\beta M(t)^\gamma \left(\alpha \frac{A'(t)}{A(t)} + \beta \frac{K'(t)}{K(t)} + \gamma \frac{M'(t)}{M(t)} \right)$$

I en økonomisk situasjon med begrensede ressurser vil man ofte overveie å styrke noen innsatsområder på bekostning av andre. Bonden vår kan f.eks. overveie å utvide det dyrkede arealet m ved å kjøpe tilleggsjord, men er da tvunget til å redusere kapitalinvesteringen k . Om slike omprioriteringer er lønnsomme eller ikke, viser seg i fortegnet til den deriverte $U'(t)$ av utbyttefunksjonen. Ifølge formelen ovenfor er dette fortegnet bestemt av fortegnet til uttrykket

$$\alpha \frac{A'(t)}{A(t)} + \beta \frac{K'(t)}{K(t)} + \gamma \frac{M'(t)}{M(t)}$$

($CA(t)^\alpha K(t)^\beta M(t)^\gamma$ er alltid positiv). Fortegnet til den deriverte av utbyttefunksjonen er altså bestemt av en kombinasjon av de *relative veksthastighetene* $\frac{A'(t)}{A(t)}$, $\frac{K'(t)}{K(t)}$ og $\frac{M'(t)}{M(t)}$ og Cobb-Douglass eksponentene α , β og γ . For en

økonom med trening i å tolke Cobb-Douglaseksponenter gir et slikt uttrykk økonomisk mening. \square

Eksempel 4: Når en gass oppbevares i en beholder, er trykket P en funksjon av temperaturen T og volumet V ; vi har altså

$$P = f(T, V)$$

Hvilken funksjon f det er naturlig å bruke, avhenger både av hvilken gass vi ser på, og av hvor stor nøyaktighet vi ønsker. Foreløpig lar vi f være uspesifisert slik at regningene våre gjelder for alle modeller. Dersom volumet og trykket varierer med tiden t , vil også trykket variere med tiden

$$P(t) = f(T(t), V(t))$$

Vi kan finne den deriverte av denne funksjonen ved å bruke kjerneregelen (husk at $P(t)$, $T(t)$ og $V(t)$ bare avhenger av én variabel t slik at vi kan bruke vanlige deriverte $P'(t)$, $T'(t)$ og $V'(t)$ istedenfor partiellderivate $\frac{\partial P}{\partial t}$, $\frac{\partial T}{\partial t}$ og $\frac{\partial V}{\partial t}$):

$$P'(t) = \frac{\partial f}{\partial T} T'(t) + \frac{\partial f}{\partial V} V'(t)$$

Dersom gassen er tynn og kreftene mellom gasspartiklene er svake (fysikerne kaller dette en *ideell gass*), kan man anta at trykket P er proporsjonalt med temperaturen T og omvendt proporsjonalt med volumet V , dvs. $P = k \frac{T}{V}$ der k er en konstant som blant annet avhenger av hvor mye gass det er i beholderen. Vi har altså

$$P = f(T, V) = k \frac{T}{V}$$

I dette tilfellet er

$$\frac{\partial f}{\partial T} = k \frac{1}{V} \quad \text{og} \quad \frac{\partial f}{\partial V} = -k \frac{T}{V^2}$$

Setter vi dette inn i formelen for $P'(t)$, får vi

$$P'(t) = k \frac{1}{V} T'(t) - k \frac{T}{V^2} V'(t)$$

Vi har tidligere regnet oss frem til denne formelen på en litt annen måte (se eksempel 5 i seksjon 2.4).

Regning med partiellderivate og kjerneregelen står sentralt i termodynamikken, og utregningene ovenfor er bare et enkelt eksempel. \clubsuit

Vi begynte denne seksjonen med å skrive opp kjerneregelen på matriseform:

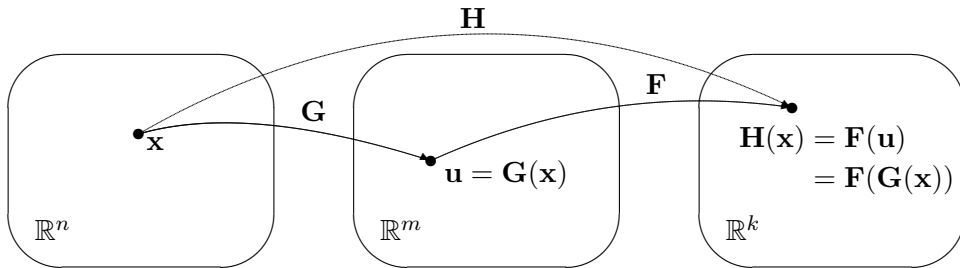
$$\mathbf{H}'(\mathbf{x}) = \mathbf{F}'(\mathbf{G}(\mathbf{x}))\mathbf{G}'(\mathbf{x}),$$

men siden har vi bare arbeidet med regelen på komponentform:

$$\frac{\partial h}{\partial x_i}(\mathbf{x}) = \frac{\partial f}{\partial u_1}(\mathbf{u}) \frac{\partial g_1}{\partial x_i}(\mathbf{x}) + \frac{\partial f}{\partial u_2}(\mathbf{u}) \frac{\partial g_2}{\partial x_i}(\mathbf{x}) + \dots + \frac{\partial f}{\partial u_m}(\mathbf{u}) \frac{\partial g_m}{\partial x_i}(\mathbf{x})$$

La oss nå se litt på sammenhengen mellom de to formene.

Når vi arbeider med kjerneregelen på matrisiform, tenker vi oss at vi starter med to funksjoner $\mathbf{G} : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $\mathbf{F} : \mathbb{R}^m \rightarrow \mathbb{R}^k$ og at vi danner den sammensatte funksjonen $\mathbf{H} : \mathbb{R}^n \rightarrow \mathbb{R}^k$ ved $\mathbf{H}(\mathbf{x}) = \mathbf{F}(\mathbf{G}(\mathbf{x}))$ som vist på figuren.



Skriver vi ut komponentene til $\mathbf{H}'(\mathbf{x})$, $\mathbf{F}'(\mathbf{G}(\mathbf{x}))$, og $\mathbf{G}'(\mathbf{x})$, får kjerneregelen $\mathbf{H}'(\mathbf{x}) = \mathbf{F}'(\mathbf{G}(\mathbf{x}))\mathbf{G}'(\mathbf{x})$ denne formen:

$$\begin{pmatrix} \frac{\partial H_1}{\partial x_1}(\mathbf{x}) & \frac{\partial H_1}{\partial x_2}(\mathbf{x}) & \dots & \frac{\partial H_1}{\partial x_n}(\mathbf{x}) \\ \frac{\partial H_2}{\partial x_1}(\mathbf{x}) & \frac{\partial H_2}{\partial x_2}(\mathbf{x}) & \dots & \frac{\partial H_2}{\partial x_n}(\mathbf{x}) \\ \vdots & \vdots & \dots & \vdots \\ \frac{\partial H_k}{\partial x_1}(\mathbf{x}) & \frac{\partial H_k}{\partial x_2}(\mathbf{x}) & \dots & \frac{\partial H_k}{\partial x_n}(\mathbf{x}) \end{pmatrix} =$$

$$= \begin{pmatrix} \frac{\partial F_1}{\partial u_1}(\mathbf{G}(\mathbf{x})) & \frac{\partial F_1}{\partial u_2}(\mathbf{G}(\mathbf{x})) & \dots & \frac{\partial F_1}{\partial u_m}(\mathbf{G}(\mathbf{x})) \\ \frac{\partial F_2}{\partial u_1}(\mathbf{G}(\mathbf{x})) & \frac{\partial F_2}{\partial u_2}(\mathbf{G}(\mathbf{x})) & \dots & \frac{\partial F_2}{\partial u_m}(\mathbf{G}(\mathbf{x})) \\ \vdots & \vdots & \dots & \vdots \\ \frac{\partial F_k}{\partial u_1}(\mathbf{G}(\mathbf{x})) & \frac{\partial F_k}{\partial u_2}(\mathbf{G}(\mathbf{x})) & \dots & \frac{\partial F_k}{\partial u_m}(\mathbf{G}(\mathbf{x})) \end{pmatrix} \begin{pmatrix} \frac{\partial G_1}{\partial x_1}(\mathbf{x}) & \frac{\partial G_1}{\partial x_2}(\mathbf{x}) & \dots & \frac{\partial G_1}{\partial x_n}(\mathbf{x}) \\ \frac{\partial G_2}{\partial x_1}(\mathbf{x}) & \frac{\partial G_2}{\partial x_2}(\mathbf{x}) & \dots & \frac{\partial G_2}{\partial x_n}(\mathbf{x}) \\ \vdots & \vdots & \dots & \vdots \\ \frac{\partial G_m}{\partial x_1}(\mathbf{x}) & \frac{\partial G_m}{\partial x_2}(\mathbf{x}) & \dots & \frac{\partial G_m}{\partial x_n}(\mathbf{x}) \end{pmatrix}$$

Den ij -te komponenten i den første matrisen er $\frac{\partial H_i}{\partial x_j}(\mathbf{x})$. Denne komponenten må være skalarproduktet av den i -te linjen i den andre matrisen og den j -te søylen i den tredje matrisen, altså:

$$\frac{\partial H_i}{\partial x_j}(\mathbf{x}) = \frac{\partial F_i}{\partial u_1}(\mathbf{G}(\mathbf{x})) \frac{\partial G_1}{\partial x_j}(\mathbf{x}) + \frac{\partial F_i}{\partial u_2}(\mathbf{G}(\mathbf{x})) \frac{\partial G_2}{\partial x_j}(\mathbf{x}) + \dots + \frac{\partial F_i}{\partial u_m}(\mathbf{G}(\mathbf{x})) \frac{\partial G_m}{\partial x_j}(\mathbf{x})$$

Dette er ikke noe annet enn kjerneregelen på komponentform anvendt på funksjonen $H_i(\mathbf{x}) = F_i(\mathbf{G}(\mathbf{x}))$. Vi har dermed vist at kjerneregelen på komponentform rett og slett er det vi får når vi ganger ut matriseproduktet i formelen $\mathbf{H}'(\mathbf{x}) = \mathbf{F}'(\mathbf{G}(\mathbf{x}))\mathbf{G}'(\mathbf{x})$ og sjekker hva som skjer med hver enkelt komponent. Når vi nå skal bevise kjerneregelen, holder det derfor å bevise kjerneregelen på matriseform.

Teorem 2.7.1 (Kjerneregelen på matriseform) *Anta at vi har to mengder $A \subset \mathbb{R}^n$, $B \subset \mathbb{R}^m$ og to funksjoner $\mathbf{G} : A \rightarrow B$, $\mathbf{F} : B \rightarrow \mathbb{R}^k$. Dersom \mathbf{G} er deriverbar i punktet $\mathbf{a} \in A$ og \mathbf{F} er deriverbar i punktet $\mathbf{b} = \mathbf{G}(\mathbf{a})$, så er den sammensatte funksjonen $\mathbf{H}(\mathbf{x}) = \mathbf{F}(\mathbf{G}(\mathbf{x}))$ deriverbar i \mathbf{a} , og Jacobi-matrisen til \mathbf{H} er gitt ved*

$$\mathbf{H}'(\mathbf{a}) = \mathbf{F}'(\mathbf{G}(\mathbf{a}))\mathbf{G}'(\mathbf{a})$$

Bevis: Dette beviset er ikke så vanskelig når man har forstått ideen, men det er langt og litt kronglete. Ideen er å bruke setning 2.6.4 med $\mathbf{F}'(\mathbf{G}(\mathbf{a}))\mathbf{G}'(\mathbf{a})$ i rollen som matrisen B . Alt vi da behøver å vise, er at funksjonen

$$\hat{\sigma}(\mathbf{r}) = \mathbf{H}(\mathbf{a} + \mathbf{r}) - \mathbf{H}(\mathbf{a}) - (\mathbf{F}'(\mathbf{G}(\mathbf{a}))\mathbf{G}'(\mathbf{a}))\mathbf{r} \quad (2.7.3)$$

tilfredsstiller kravet

$$\lim_{\mathbf{r} \rightarrow \mathbf{0}} \frac{1}{|\mathbf{r}|} \hat{\sigma}(\mathbf{r}) = \mathbf{0} \quad (2.7.4)$$

For å vise dette, må vi bruke at \mathbf{G} og \mathbf{F} er deriverbare. Siden \mathbf{G} er deriverbar i \mathbf{a} , vet vi at funksjonen

$$\sigma_1(\mathbf{r}) = \mathbf{G}(\mathbf{a} + \mathbf{r}) - \mathbf{G}(\mathbf{a}) - \mathbf{G}'(\mathbf{a})\mathbf{r} \quad (2.7.5)$$

tilfredsstiller kravet

$$\lim_{\mathbf{r} \rightarrow \mathbf{0}} \frac{1}{|\mathbf{r}|} \sigma_1(\mathbf{r}) = \mathbf{0} \quad (2.7.6)$$

Siden \mathbf{F} er deriverbar i $\mathbf{b} = \mathbf{G}(\mathbf{a})$, vet vi tilsvarende at funksjonen

$$\begin{aligned} \sigma_2(\mathbf{s}) &= \mathbf{F}(\mathbf{b} + \mathbf{s}) - \mathbf{F}(\mathbf{b}) - \mathbf{F}'(\mathbf{b})\mathbf{s} = \\ &= \mathbf{F}(\mathbf{G}(\mathbf{a}) + \mathbf{s}) - \mathbf{F}(\mathbf{G}(\mathbf{a})) - \mathbf{F}'(\mathbf{G}(\mathbf{a}))\mathbf{s} \end{aligned} \quad (2.7.7)$$

tilfredsstiller kravet

$$\lim_{\mathbf{s} \rightarrow \mathbf{0}} \frac{1}{|\mathbf{s}|} \sigma_2(\mathbf{s}) = \mathbf{0} \quad (2.7.8)$$

Vi setter $\mathbf{s} = \mathbf{G}(\mathbf{a} + \mathbf{r}) - \mathbf{G}(\mathbf{a})$. Da er $\mathbf{G}(\mathbf{a} + \mathbf{r}) = \mathbf{G}(\mathbf{a}) + \mathbf{s}$, og fra (2.7.5) ser vi dessuten at $\mathbf{s} = \mathbf{G}'(\mathbf{a})\mathbf{r} + \sigma_1(\mathbf{r})$. Dermed er

$$\mathbf{H}(\mathbf{a} + \mathbf{r}) = \mathbf{F}(\mathbf{G}(\mathbf{a} + \mathbf{r})) = \mathbf{F}(\mathbf{G}(\mathbf{a}) + \mathbf{s}) = \mathbf{F}(\mathbf{G}(\mathbf{a})) + \mathbf{F}'(\mathbf{G}(\mathbf{a}))\mathbf{s} + \sigma_2(\mathbf{s})$$

der vi i siste trinn har brukt (2.7.7). Siden $\mathbf{s} = \mathbf{G}'(\mathbf{a})\mathbf{r} + \boldsymbol{\sigma}_1(\mathbf{r})$, gir dette videre

$$\begin{aligned}\mathbf{H}(\mathbf{a} + \mathbf{r}) &= \mathbf{F}(\mathbf{G}(\mathbf{a})) + \mathbf{F}'(\mathbf{G}(\mathbf{a}))\left(\mathbf{G}'(\mathbf{a})\mathbf{r} + \boldsymbol{\sigma}_1(\mathbf{r})\right) + \boldsymbol{\sigma}_2(\mathbf{s}) = \\ &= \mathbf{H}(\mathbf{a}) + \mathbf{F}'(\mathbf{G}(\mathbf{a}))\mathbf{G}'(\mathbf{a})\mathbf{r} + \mathbf{F}'(\mathbf{G}(\mathbf{a}))\boldsymbol{\sigma}_1(\mathbf{r}) + \boldsymbol{\sigma}_2(\mathbf{s})\end{aligned}$$

Sammenligner vi dette med (2.7.3), ser vi at

$$\hat{\boldsymbol{\sigma}}(\mathbf{r}) = \mathbf{F}'(\mathbf{G}(\mathbf{a}))\boldsymbol{\sigma}_1(\mathbf{r}) + \boldsymbol{\sigma}_2(\mathbf{s})$$

Ifølge (2.7.4) er vår oppgave dermed å vise at

$$\lim_{\mathbf{r} \rightarrow \mathbf{0}} \frac{1}{|\mathbf{r}|} \hat{\boldsymbol{\sigma}}(\mathbf{r}) = \lim_{\mathbf{r} \rightarrow \mathbf{0}} \frac{1}{|\mathbf{r}|} \left(\mathbf{F}'(\mathbf{G}(\mathbf{a}))\boldsymbol{\sigma}_1(\mathbf{r}) + \boldsymbol{\sigma}_2(\mathbf{s}) \right) = \mathbf{0} \quad (2.7.9)$$

Vi tar ett ledd av gangen. Fra setning 1.5.3 vet vi at

$$|\mathbf{F}'(\mathbf{G}(\mathbf{a}))\boldsymbol{\sigma}_1(\mathbf{r})| \leq \|\mathbf{F}'(\mathbf{G}(\mathbf{a}))\| |\boldsymbol{\sigma}_1(\mathbf{r})|$$

der $\|\mathbf{F}'(\mathbf{G}(\mathbf{a}))\|$ bare er et fast tall (normen til matrisen $\mathbf{F}'(\mathbf{G}(\mathbf{a}))$). Dermed har vi

$$\left| \frac{1}{|\mathbf{r}|} \left(\mathbf{F}'(\mathbf{G}(\mathbf{a}))\boldsymbol{\sigma}_1(\mathbf{r}) \right) \right| \leq \|\mathbf{F}'(\mathbf{G}(\mathbf{a}))\| \frac{|\boldsymbol{\sigma}_1(\mathbf{r})|}{|\mathbf{r}|} \rightarrow 0$$

siden $\frac{|\boldsymbol{\sigma}_1(\mathbf{r})|}{|\mathbf{r}|} \rightarrow 0$ når $\mathbf{r} \rightarrow \mathbf{0}$. Dette viser at den første delen av uttrykket i (2.7.9) går mot $\mathbf{0}$. Den andre delen

$$\lim_{\mathbf{r} \rightarrow \mathbf{0}} \frac{1}{|\mathbf{r}|} \boldsymbol{\sigma}_2(\mathbf{s}) = \mathbf{0}$$

er litt verre. Vi legger først merke til at siden $\boldsymbol{\sigma}_2(\mathbf{0}) = \mathbf{0}$, har vi ingen problemer med tilfellet $\mathbf{s} = \mathbf{0}$. Vi kan derfor anta at $\mathbf{s} \neq \mathbf{0}$, og multiplisere uttrykket $\frac{1}{|\mathbf{r}|} \boldsymbol{\sigma}_2(\mathbf{s})$ med $|\mathbf{s}|$ i teller og nevner:

$$\lim_{\mathbf{r} \rightarrow \mathbf{0}} \frac{1}{|\mathbf{r}|} \boldsymbol{\sigma}_2(\mathbf{s}) = \lim_{\mathbf{r} \rightarrow \mathbf{0}} \frac{|\mathbf{s}|}{|\mathbf{r}|} \cdot \frac{\boldsymbol{\sigma}_2(\mathbf{s})}{|\mathbf{s}|}$$

Når $\mathbf{r} \rightarrow \mathbf{0}$, vil $\mathbf{s} = \mathbf{G}(\mathbf{a})\mathbf{r} + \boldsymbol{\sigma}_1(\mathbf{r}) \rightarrow \mathbf{0}$, og dette medfører at

$$\lim_{\mathbf{r} \rightarrow \mathbf{0}} \frac{\boldsymbol{\sigma}_2(\mathbf{s})}{|\mathbf{s}|} = \lim_{\mathbf{s} \rightarrow \mathbf{0}} \frac{\boldsymbol{\sigma}_2(\mathbf{s})}{|\mathbf{s}|} = \mathbf{0}$$

ifølge (2.7.8). Det gjenstår derfor å vise at faktoren $\frac{|\mathbf{s}|}{|\mathbf{r}|}$ er begrenset og ikke kan gå mot uendelig. Bruker vi at $\mathbf{s} = \mathbf{G}'(\mathbf{a})\mathbf{r} + \boldsymbol{\sigma}_1(\mathbf{r})$, får vi

$$\frac{|\mathbf{s}|}{|\mathbf{r}|} = \frac{|\mathbf{G}'(\mathbf{a})\mathbf{r} + \boldsymbol{\sigma}_1(\mathbf{r})|}{|\mathbf{r}|} \leq \frac{|\mathbf{G}'(\mathbf{a})\mathbf{r}|}{|\mathbf{r}|} + \frac{|\boldsymbol{\sigma}_1(\mathbf{r})|}{|\mathbf{r}|} \leq \|\mathbf{G}'(\mathbf{a})\| + \frac{|\boldsymbol{\sigma}_1(\mathbf{r})|}{|\mathbf{r}|}$$

som er begrenset siden $\frac{\sigma_1(\mathbf{r})}{|\mathbf{r}|} \rightarrow \mathbf{0}$. Dermed har vi vist at

$$\lim_{\mathbf{r} \rightarrow \mathbf{0}} \frac{1}{|\mathbf{r}|} \hat{\sigma}(\mathbf{r}) = \lim_{\mathbf{r} \rightarrow \mathbf{0}} \frac{1}{|\mathbf{r}|} \left(\mathbf{F}'(\mathbf{G}(\mathbf{a}))\sigma_1(\mathbf{r}) + \sigma_2(\mathbf{s}) \right) = \mathbf{0}$$

og beviset er fullført. \square

Bemerkning: Legg merke til at kjerneregelen garanterer at \mathbf{H} er deriverbar i \mathbf{a} så lenge \mathbf{G} er deriverbar i \mathbf{a} og \mathbf{F} er deriverbar i $\mathbf{b} = \mathbf{G}(\mathbf{a})$. Siden deriverbarhet kan være vanskelig å sjekke, er denne garantien ofte til stor hjelp i teoretisk arbeid.

Vi skriver også opp den presise formuleringen av kjerneregelen på komponentform:

Teorem 2.7.2 (Kjerneregelen på komponentform) *Anta at vi har to mengder $A \subset \mathbb{R}^n$, $B \subset \mathbb{R}^m$ og to funksjoner $\mathbf{G} : A \rightarrow B$, $\mathbf{F} : B \rightarrow \mathbb{R}^k$. Dersom \mathbf{G} er deriverbar i punktet $\mathbf{a} \in A$ og \mathbf{F} er deriverbar i punktet $\mathbf{b} = \mathbf{G}(\mathbf{a})$, så er den sammensatte funksjonen $\mathbf{H}(\mathbf{x}) = \mathbf{F}(\mathbf{G}(\mathbf{x}))$ deriverbar i \mathbf{a} , og de partiellderiverte til \mathbf{H} er gitt ved*

$$\begin{aligned} \frac{\partial H_i}{\partial x_j}(\mathbf{a}) &= \sum_{p=1}^m \frac{\partial F_i}{\partial u_p}(\mathbf{G}(\mathbf{a})) \frac{\partial G_p}{\partial x_j}(\mathbf{a}) = \\ &= \frac{\partial F_i}{\partial u_1}(\mathbf{G}(\mathbf{a})) \frac{\partial G_1}{\partial x_j}(\mathbf{a}) + \frac{\partial F_i}{\partial u_2}(\mathbf{G}(\mathbf{a})) \frac{\partial G_2}{\partial x_j}(\mathbf{a}) + \cdots + \frac{\partial F_i}{\partial u_m}(\mathbf{G}(\mathbf{a})) \frac{\partial G_m}{\partial x_j}(\mathbf{a}) \end{aligned}$$

Bevis: Som vi allerede har sett, følger denne formuleringen direkte fra matriseformen når man utfører matrisemultiplikasjonen og sammenligner koeffisientene på begge sider av likhetstegnet. \square

Oppgaver til seksjon 2.7

1. La $f(u, v) = u^2 + v$, $g(x, y) = 2xy$, $h(x, y) = x + y^2$. Bruk kjerneregelen til å finne de partiellderiverte av $k(x, y) = f(g(x, y), h(x, y))$.

2. La $f(u, v) = u \cdot e^{-v}$, $g(x, y, z) = 2xy + z$, $h(x, y, z) = 2y(z + x)$. Bruk kjerneregelen til å finne de partiellderiverte av $k(x, y, z) = f(g(x, y, z), h(x, y, z))$.

3. Bruk kjerneregelen til å regne ut $\frac{\partial h}{\partial x_2}$ og $\frac{\partial h}{\partial x_3}$ i eksempel 1.

4. Bruk kjerneregelen til å regne ut $\frac{\partial h}{\partial v}$ i eksempel 2.

5. Vi har to funksjoner $\mathbf{G} : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ og $\mathbf{F} : \mathbb{R}^3 \rightarrow \mathbb{R}^2$. Anta at $\mathbf{G}(1, -2) = (1, 2, 3)$ og at

$$\mathbf{G}'(1, -2) = \begin{pmatrix} 1 & -2 \\ 3 & 1 \\ 2 & -1 \end{pmatrix}, \quad \mathbf{F}'(1, 2, 3) = \begin{pmatrix} 2 & 1 & 4 \\ 0 & 2 & 2 \end{pmatrix}$$

Finn Jacobi-matrisen til den sammensatte funksjonen $\mathbf{H}(\mathbf{x}) = \mathbf{F}(\mathbf{G}(\mathbf{x}))$ i punktet $(1, -2)$.

6. Vi har to funksjoner $\mathbf{G} : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ og $\mathbf{F} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$. Anta at $\mathbf{G}(-1, -2, 1) = (2, 4)$ og at

$$\mathbf{G}'(-1, -2, 1) = \begin{pmatrix} 4 & -2 & 0 \\ 1 & 3 & -1 \end{pmatrix}, \quad \mathbf{F}'(2, 4) = \begin{pmatrix} 2 & -3 \\ 0 & 2 \end{pmatrix}$$

Finn Jacobi-matrisen til den sammensatte funksjonen $\mathbf{H}(\mathbf{x}) = \mathbf{F}(\mathbf{G}(\mathbf{x}))$ i punktet $(-1, -2, 1)$.

7. To vareslag konkurrerer om det samme markedet. Etterspørselen E_1 etter det første vareslaget varierer med prisene p_1 og p_2 på begge vareslagene.

Vi har altså en funksjon $E_1 = E_1(p_1, p_2)$. Anta at vi vet hvordan prisene $p_1 = p_1(t)$ og $p_2 = p_2(t)$ varierer med tiden. Vis at etterspørselens variasjon med tiden kan uttrykkes ved

$$\frac{dE_1}{dt} = \frac{\partial E_1}{\partial p_1} p'_1(t) + \frac{\partial E_1}{\partial p_2} p'_2(t)$$

8. Temperaturen T i et område avhenger av posisjonen; vi kan tenke oss at den er gitt som en funksjon $T = f(x, y)$ av to variable der x og y er vanlige koordinater. Vi innfører nå polarkoordinater r og θ på vanlig måte slik at $x = r \cos \theta$, $y = r \sin \theta$. Vi får da temperaturen som en funksjon $T = f(r \cos \theta, r \sin \theta)$ av r og θ .

a) Vis at

$$\begin{aligned} \frac{\partial T}{\partial r} &= \frac{\partial f}{\partial x} \cos \theta + \frac{\partial f}{\partial y} \sin \theta \\ \frac{\partial T}{\partial \theta} &= -\frac{\partial f}{\partial x} r \sin \theta + \frac{\partial f}{\partial y} r \cos \theta \end{aligned}$$

b) En radiomerket fugl beveger seg i området. Radiosignalene viser hvordan avstanden r og vinkelen θ varierer med tiden; vi har $r = g(t)$ og $\theta = h(t)$. Vis at temperaturendringene fuglen opplever er gitt ved

$$T'(t) = \left(\frac{\partial f}{\partial x} \cos \theta + \frac{\partial f}{\partial y} \sin \theta \right) g'(t) + \left(-\frac{\partial f}{\partial x} r \sin \theta + \frac{\partial f}{\partial y} r \cos \theta \right) h'(t)$$

der vi må sette inn $r = g(t)$, $\theta = h(t)$, $x = g(t) \cos h(t)$, $y = g(t) \sin h(t)$.

9. La $f : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$ være en deriverbar funksjon av $n+1$ variable og anta at det finnes en deriverbar funksjon $g : \mathbb{R}^n \rightarrow \mathbb{R}$ slik at

$$f(x_1, x_2, \dots, x_n, g(x_1, x_2, \dots, x_n)) = 0$$

for alle x_1, x_2, \dots, x_n . (Tenk deg at $y = g(x_1, x_2, \dots, x_n)$ er det uttrykket du får dersom du løser ligningen $f(x_1, x_2, \dots, x_n, y) = 0$ med hensyn på y).

a) Vis at

$$\frac{\partial g}{\partial x_i}(x_1, x_2, \dots, x_n) = -\frac{\frac{\partial f}{\partial x_i}(x_1, x_2, \dots, x_n, g(x_1, x_2, \dots, x_n))}{\frac{\partial f}{\partial x_{n+1}}(x_1, x_2, \dots, x_n, g(x_1, x_2, \dots, x_n))}$$

- b) La $f(x, y) = x^2 + y^2 - R^2$ der R er en positiv konstant, og anta at $z = g(x, y)$ er en deriverbar funksjon slik at $f(x, g(x)) = 0$. Vis at

$$g'(x) = -\frac{x}{g(x)}$$

Gi en geometrisk tolkning av resultatet.

- c) La $f(x, y, z) = x^2 + y^2 + z^2 - R^2$ der R er en positiv konstant og anta at $z = g(x, y)$ er en deriverbar funksjon slik at $f(x, y, g(x, y)) = 0$. Vis at

$$\frac{\partial g}{\partial x}(x, y) = -\frac{x}{g(x, y)}$$

og

$$\frac{\partial g}{\partial y}(x, y) = -\frac{y}{g(x, y)}$$

Gi en geometrisk tolkning av resultatet.

2.8 Lineæravbildninger

I de to siste seksjonene i dette kapitlet skal vi se på to spesielle og nært beslektede typer av funksjoner fra \mathbb{R}^n to \mathbb{R}^m — *lineæravbildninger* og *affin-avbildninger*. Disse funksjonstypene knytter teorien i dette kapitlet nærmere til teorien for matriser og vektorer i forrige kapittel.

Vi begynner med definisjonen av lineæravbildninger.

Definisjon 2.8.1 En funksjon $\mathbf{T} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ kalles en lineæravbildning dersom vi for alle $c \in \mathbb{R}$ og alle $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ har:

- (i) $\mathbf{T}(c\mathbf{x}) = c\mathbf{T}(\mathbf{x})$
(ii) $\mathbf{T}(\mathbf{x} + \mathbf{y}) = \mathbf{T}(\mathbf{x}) + \mathbf{T}(\mathbf{y})$

De aller fleste funksjoner fra \mathbb{R}^n til \mathbb{R}^m er *ikke* lineæravbildninger, men disse funksjonene er allikevel så viktige at det er en hel gren av matematikken som hovedsakelig handler om dem — denne grenen kalles *lineær algebra*.

La oss begynne med en enkel og nyttig generalisering av definisjonen.

Setning 2.8.2 Anta at $\mathbf{T} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ er en lineæravbildning. Da er

$$\mathbf{T}(c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_k\mathbf{x}_k) = c_1\mathbf{T}(\mathbf{x}_1) + c_2\mathbf{T}(\mathbf{x}_2) + \cdots + c_k\mathbf{T}(\mathbf{x}_k)$$

for alle tall $c_1, c_2, \dots, c_k \in \mathbb{R}$ og alle vektorer $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k \in \mathbb{R}^n$.

Bevis: Vi spalter av ett og ett ledd. Siden vi kan oppfatte $c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_k\mathbf{x}_k$ som en sum av to ledd $c_1\mathbf{x}_1$ og $c_2\mathbf{x}_2 + \cdots + c_k\mathbf{x}_k$, har vi

$$\begin{aligned} \mathbf{T}(c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_k\mathbf{x}_k) &= \mathbf{T}(c_1\mathbf{x}_1) + \mathbf{T}(c_2\mathbf{x}_2 + \cdots + c_k\mathbf{x}_k) = \\ &= c_1\mathbf{T}(\mathbf{x}_1) + \mathbf{T}(c_2\mathbf{x}_2 + \cdots + c_k\mathbf{x}_k) \end{aligned}$$

der vi har brukt de to punktene i definisjonen av lineæravbildning. Vi kan nå spalte av leddet $c_2\mathbf{x}_2$ på akkurat samme måte, og fortsetter vi slik, står vi til slutt igjen med

$$\mathbf{T}(c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_k\mathbf{x}_k) = c_1\mathbf{T}(\mathbf{x}_1) + c_2\mathbf{T}(\mathbf{x}_2) + \cdots + c_k\mathbf{T}(\mathbf{x}_k)$$

□

Det neste resultatet viser oss at lineæravbildninger finnes:

Setning 2.8.3 *Anta at A er en $m \times n$ -matrise. Da er funksjonen $\mathbf{T} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ definert ved*

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

en lineæravbildning.

Bevis: Etter regnereglerne for matrisemultiplikasjon er $\mathbf{T}(c\mathbf{x}) = A(c\mathbf{x}) = c(A\mathbf{x}) = c\mathbf{T}(\mathbf{x})$ og $\mathbf{T}(\mathbf{x} + \mathbf{y}) = A(\mathbf{x} + \mathbf{y}) = A\mathbf{x} + A\mathbf{y} = \mathbf{T}(\mathbf{x}) + \mathbf{T}(\mathbf{y})$. □

Eksempel 1: Hvis A er 2×3 -matrisen

$$A = \begin{pmatrix} 2 & -1 & 3 \\ 1 & -4 & 2 \end{pmatrix},$$

får vi en lineæravbildning $\mathbf{T} : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ ved

$$\mathbf{T}(x, y, z) = \begin{pmatrix} 2 & -1 & 3 \\ 1 & -4 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x - y + 3z \\ x - 4y + 2z \end{pmatrix}$$

♣

Den neste setningen er nok mer overraskende — den sier at det ikke finnes andre lineæravbildninger enn de som er gitt av matriser!

Setning 2.8.4 *Anta at $\mathbf{T} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ er en lineæravbildning. Da finnes det en $m \times n$ -matrise A slik at*

$$\mathbf{T}(\mathbf{x}) = A\mathbf{x} \quad \text{for alle } \mathbf{x} \in \mathbb{R}^n$$

Matrisen A er gitt ved at den j -te søylen er lik $\mathbf{T}(\mathbf{e}_j)$ der \mathbf{e}_j er den j -te enhetsvektoren

$$\mathbf{e}_j = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \longleftarrow j\text{-te komponent}$$

Bevis: Vi begynner med å sette navn på komponentene til $\mathbf{T}(\mathbf{e}_j)$ (som du vil se, er navnvalget inspirert av resultatet vi skal frem til!):

$$\mathbf{T}(\mathbf{e}_j) = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{ij} \\ \vdots \\ a_{mj} \end{pmatrix}$$

Siden enhver vektor $\mathbf{x} \in \mathbb{R}^n$ kan uttrykkes ved hjelp av enhetsvektorene på denne måten

$$\begin{aligned} \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} &= x_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \cdots + x_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} = \\ &= x_1 \mathbf{e}_1 + x_2 \mathbf{e}_2 + \cdots + x_n \mathbf{e}_n, \end{aligned}$$

har vi ifølge setningen ovenfor

$$\begin{aligned} \mathbf{T}(\mathbf{x}) &= \mathbf{T}(x_1 \mathbf{e}_1 + x_2 \mathbf{e}_2 + \cdots + x_n \mathbf{e}_n) = x_1 \mathbf{T}(\mathbf{e}_1) + x_2 \mathbf{T}(\mathbf{e}_2) + \cdots + x_n \mathbf{T}(\mathbf{e}_n) = \\ &= x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} + \cdots + x_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} = \\ &= \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix} = A\mathbf{x} \end{aligned}$$

der

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

Dermed er setningen bevist. \square

Vi skal nå se hvordan vi kan bruke resultatet ovenfor på refleksjoner og rotasjoner i planet. Dette er viktige eksempler som du kommer til å få glede av i fremtiden.

Eksempel 2: Vi skal nå se på avbildningen $\mathbf{T} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ som avbilder enhver vektor på sitt speilbilde om x -aksen. Det er lett å overbevise seg om at T er en lineærabildning, og vi skal bruke setningen ovenfor til å finne den tilhørende matrisen. Siden speilingen er om x -aksen, har vi

$$\mathbf{T}(\mathbf{e}_1) = \mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{og} \quad \mathbf{T}(\mathbf{e}_2) = -\mathbf{e}_2 = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$$

Ifølge setningen ovenfor får vi matrisen til T ved å bruke disse vektorene som søylevektorer:

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

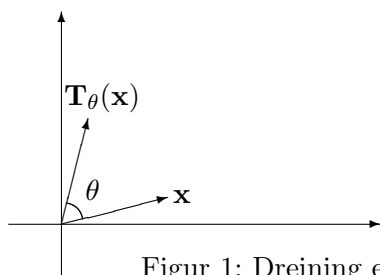
Hvis $\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}$, er dermed

$$\mathbf{T}(\mathbf{x}) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ -y \end{pmatrix}$$



I eksemplet ovenfor er avbildningen så enkel at det ikke ville by på noe problem å skrive opp formelen for $\mathbf{T}(\mathbf{x})$ direkte uten å gå veien om $\mathbf{T}(\mathbf{e}_1)$ og $\mathbf{T}(\mathbf{e}_2)$. I litt mer kompliserte eksempler blir imidlertid regningene atskillig mer oversiktlige om vi går veien om $\mathbf{T}(\mathbf{e}_1)$ og $\mathbf{T}(\mathbf{e}_2)$.

Eksempel 3: Avbildningen $\mathbf{T}_\theta : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ er gitt ved at den dreier enhver vektor \mathbf{x} en vinkel θ i positiv omløpsretning (se figur 1).



Figur 1: Dreining en vinkel θ

Det er ikke vanskelig å overbevise seg om at \mathbf{T}_θ er en lineærabildning (f.eks. sier regelen $\mathbf{T}_\theta(\mathbf{x} + \mathbf{y}) = \mathbf{T}_\theta(\mathbf{x}) + \mathbf{T}_\theta(\mathbf{y})$ i dette tilfellet at dersom vi først legger sammen to vektorer \mathbf{x} og \mathbf{y} , og så dreier resultatet en vinkel θ , så får vi det samme som om vi først dreier begge vektorene en vinkel θ , og så legger sammen de nye vektorene). For å finne matrisen til \mathbf{T}_θ må vi beregne $\mathbf{T}_\theta(\mathbf{e}_1)$ og $\mathbf{T}_\theta(\mathbf{e}_2)$. Dreier vi enhetsvektoren \mathbf{e}_1 en vinkel θ , ender den i punktet $\begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix}$ (dette er bare definisjonen av cosinus og sinus til

generelle vinkler). Dette betyr altså at

$$\mathbf{T}_\theta(\mathbf{e}_1) = \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix}$$

Siden \mathbf{e}_2 ligger en vinkel $\pi/2$ foran \mathbf{e}_1 når vi dreier i positiv retning, får vi

$$\mathbf{T}_\theta(\mathbf{e}_2) = \begin{pmatrix} \cos(\theta + \frac{\pi}{2}) \\ \sin(\theta + \frac{\pi}{2}) \end{pmatrix} = \begin{pmatrix} -\sin(\theta) \\ \cos(\theta) \end{pmatrix}$$

(for å få til den siste overgangen kan du enten resonnerer geometrisk eller bruke formlene for sinus og cosinus til en sum). Matrisen til lineærabildningen \mathbf{T}_θ er altså

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

Dersom vi dreier en vektor $\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}$ en vinkel θ i positiv retning, får vi dermed en ny vektor

$$\mathbf{x}' = A_\theta \mathbf{x} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \cos \theta - y \sin \theta \\ x \sin \theta + y \cos \theta \end{pmatrix}$$

Vi har nå løst den opprinnelige oppgaven vår som var å finne matrisen til lineærabildningen \mathbf{T}_θ . La oss gå litt videre for å se hva som skjer når vi kombinerer to rotasjoner. Dersom vi dreier en annen vinkel ϕ , får vi selvfølgelig matrisen

$$A_\phi = \begin{pmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{pmatrix}$$

La oss nå anta at vi først dreier en vinkel θ og deretter en vinkel ϕ . I alt har vi da dreiet en vinkel $\phi + \theta$ tilsvarende matrisen

$$A_{\phi+\theta} = \begin{pmatrix} \cos(\phi + \theta) & -\sin(\phi + \theta) \\ \sin(\phi + \theta) & \cos(\phi + \theta) \end{pmatrix}$$

På den annen side vet vi at når vi gjør to transformasjoner etter hverandre, svarer dette til å multiplisere de tilhørende matrisene med hverandre. Med andre ord må

$$A_{\phi+\theta} = A_\phi A_\theta$$

Skriver vi ut komponentene, ser vi at

$$A_{\phi+\theta} = \begin{pmatrix} \cos(\phi + \theta) & -\sin(\phi + \theta) \\ \sin(\phi + \theta) & \cos(\phi + \theta) \end{pmatrix}$$

og

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

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

Sammenligner du komponentene i de to uttrykkene, vil du gjenkjenne formelene for sinus og cosinus til en sum. Vi har altså brukt matriser til å gi et nytt bevis for disse formelene. ♣

Eigenverdier

Det er ofte et omstendelig arbeid å regne ut $\mathbf{T}(\mathbf{x})$ for en lineærvbildning \mathbf{T} og en vektor \mathbf{x} . For noen vektorer går det imidlertid raskt — alt vi behøver å gjøre, er å gange vektoren med et tall λ . Slike vektorer kalles *egenvektorer*. Egenvektorer kan bare finnes når \mathbf{T} avbilder et rom inn i seg selv, altså når den går fra et rom \mathbb{R}^n til det *samme* rommet \mathbb{R}^n . Her er den formelle definisjonen:

Definisjon 2.8.5 Anta at $\mathbf{T} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ er en lineærvbildning. Vi kaller $\mathbf{x} \neq \mathbf{0}$ en egenvektor for \mathbf{T} dersom det finnes et tall λ slik at

$$\mathbf{T}(\mathbf{x}) = \lambda \mathbf{x}$$

Tallet λ kaller vi egenverdien til \mathbf{x} .

Bemerkning: I definisjonen ovenfor har vi knyttet egenvektorer til lineærvbildninger, men vi kunne like godt ha knyttet dem til matriser. Vektoren $\mathbf{x} \neq \mathbf{0}$ er en egenvektor for matrisen A dersom det finnes et tall λ slik at $A\mathbf{x} = \lambda\mathbf{x}$.

Eksempel 4: La oss undersøke om $\mathbf{a} = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$ er en egenvektor for lineærvbildningen $\mathbf{T}(\mathbf{x}) = A\mathbf{x}$ der $A = \begin{pmatrix} 1 & -8 \\ -2 & 1 \end{pmatrix}$. Vi har

$$\mathbf{T}(\mathbf{a}) = \begin{pmatrix} 1 & -8 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 10 \\ -5 \end{pmatrix} = 5\mathbf{a}$$

så \mathbf{a} er en egenvektor med egenverdi 5.

Gjør vi tilsvarende beregninger for $\mathbf{b} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$, får vi

$$\mathbf{T}(\mathbf{b}) = \begin{pmatrix} 1 & -8 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} -6 \\ -3 \end{pmatrix} = (-3)\mathbf{b}$$

Altså er \mathbf{b} en egenvektor med egenverdi -3 .

Anta nå at en tredje vektor \mathbf{c} kan skrives som en lineærkombinasjon av egenvektorene \mathbf{a} og \mathbf{b} , dvs. at det finnes tall x og y slik at $\mathbf{c} = x\mathbf{a} + y\mathbf{b}$ (det viser seg faktisk at alle vektorer i \mathbb{R}^2 kan skrives på denne måten). Da er

$$\mathbf{T}(\mathbf{c}) = \mathbf{T}(x\mathbf{a} + y\mathbf{b}) = x\mathbf{T}(\mathbf{a}) + y\mathbf{T}(\mathbf{b}) = 5x\mathbf{a} - 3y\mathbf{b}$$

Det viser seg altså at ikke bare egenvektorene selv, men også deres lineærkombinasjoner kan behandles på en enkel måte. ♣

Observasjonen i slutten av eksemplet ovenfor er viktig. Det viser seg at for “de fleste” lineærabildninger $\mathbf{T} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ finnes det egenvektorer $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ slik at *enhver* vektor \mathbf{v} kan skrives som en lineærkombinasjon

$$\mathbf{v} = x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2 + \dots + x_n \mathbf{v}_n$$

Hvis egenverdiene er henholdsvis $\lambda_1, \lambda_2, \dots, \lambda_n$, får vi da

$$\begin{aligned} \mathbf{T}(\mathbf{v}) &= \mathbf{T}(x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2 + \dots + x_n \mathbf{v}_n) = \\ &= x_1 \mathbf{T}(\mathbf{v}_1) + x_2 \mathbf{T}(\mathbf{v}_2) + \dots + x_n \mathbf{T}(\mathbf{v}_n) = \\ &= x_1 \lambda_1 \mathbf{v}_1 + x_2 \lambda_2 \mathbf{v}_2 + \dots + x_n \lambda_n \mathbf{v}_n \end{aligned}$$

Bruker vi \mathbf{T} på begge side av dette uttrykket, får vi på tilsvarende måte

$$\mathbf{T}^2(\mathbf{v}) = \mathbf{T}(\mathbf{T}(\mathbf{v})) = x_1 \lambda_1^2 \mathbf{v}_1 + x_2 \lambda_2^2 \mathbf{v}_2 + \dots + x_n \lambda_n^2 \mathbf{v}_n$$

Fortsetter vi på denne måten, får vi generelt

$$\mathbf{T}^k(\mathbf{v}) = x_1 \lambda_1^k \mathbf{v}_1 + x_2 \lambda_2^k \mathbf{v}_2 + \dots + x_n \lambda_n^k \mathbf{v}_n$$

Denne formelen forteller oss at dersom vi kjenner egenvektorene og egenverdiene til en lineærabildning \mathbf{T} , så har vi god oversikt både over \mathbf{T} selv og over alle dens potenser. I mange anvendelser forteller $\mathbf{T}^k(\mathbf{x})$ hvordan et system utvikler seg når vi starter i en tilstand \mathbf{x} og lar tiden gå ($\mathbf{T}^1(\mathbf{x})$ er tilstanden etter ett tidsintervall, $\mathbf{T}^2(\mathbf{x})$ tilstanden etter to tidsintervaller osv.) Formelen ovenfor forteller oss da at veksten til systemet hovedsakelig er bestemt av den største egenverdien (eller, for å være helt presis, den egenverdien som har størst tallverdi).

Vi skal foreløpig ikke komme nærmere inn på hvordan man kan finne egenverdier i praksis. Det er likevel viktig å vite litt om begrepet og dets anvendelser siden det vil gjøre det lettere å forstå hensikten med mye av teorien som kommer senere.

Bemerkning: Ser du nøyer på definisjon 2.8.5, vil du se at den er litt upresis når det gjelder hva slags tall egenverdien λ skal være og hva slags vektor egenvektoren \mathbf{x} skal være. Det viser seg at det finnes reelle matriser som har komplekse egenverdier og egenvektorer. Om man vil “regne med” disse avhenger av problemstillingen man ser på — i noen tilfeller er det nyttig å ha dem med, i andre tilfeller må man ekskludere dem. Situasjonen minner om den vi har for ligninger av n -te grad; noen ganger er det nyttig å ha med de komplekse løsningene og andre ganger ikke.

Eksempel 5: Dersom θ ikke er multiplum av π , viser det seg at rotasjonsmatrisene

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

ikke har reelle egenvektorer og egenverdier. De har imidlertid de komplekse egenvektorene

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ -i \end{pmatrix} \quad \text{og} \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ i \end{pmatrix}$$

med tilhørende egenverdier $\lambda_1 = \cos \theta + i \sin \theta$ og $\lambda_2 = \cos \theta - i \sin \theta$. La oss sjekke dette for λ_1 og \mathbf{v}_1 :

$$\begin{aligned} A_\theta \mathbf{v}_1 &= \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} 1 \\ -i \end{pmatrix} = \begin{pmatrix} \cos \theta + i \sin \theta \\ \sin \theta - i \cos \theta \end{pmatrix} = \\ &= (\cos \theta + i \sin \theta) \begin{pmatrix} 1 \\ -i \end{pmatrix} = \lambda_1 \mathbf{v}_1 \end{aligned}$$

Du kan på tilsvarende måte sjekke at $A_\theta \mathbf{v}_2 = \lambda_2 \mathbf{v}_2$.



Oppgaver til seksjon 2.8

1. Finn matrisen til lineærabildningen $\mathbf{T} : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ gitt ved

$$\mathbf{T}(x, y, z) = \begin{pmatrix} 2x - y + z \\ -x + y - 3z \end{pmatrix}$$

2. En lineærabildning $\mathbf{T} : \mathbb{R}^2 \rightarrow \mathbb{R}^4$ tilfredsstiller

$$\mathbf{T}(\mathbf{e}_1) = \begin{pmatrix} -1 \\ 2 \\ -3 \\ 4 \end{pmatrix} \quad \mathbf{T}(\mathbf{e}_2) = \begin{pmatrix} 0 \\ -2 \\ 4 \\ 7 \end{pmatrix}$$

Finn matrisen til \mathbf{T} .

3. La $\mathbf{a}, \mathbf{b} \in \mathbb{R}^2$. Lineærabildningen $\mathbf{T} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ tilfredsstiller $\mathbf{T}(\mathbf{a}) = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$,

$\mathbf{T}(\mathbf{b}) = \begin{pmatrix} 0 \\ 3 \end{pmatrix}$. Finn $\mathbf{T}(3\mathbf{a} - 2\mathbf{b})$.

4. Lineærabildningen $\mathbf{T} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ avbilder ethvert punkt på sitt speilbilde om y -aksen. Finn matrisen til \mathbf{T} .

5. Lineærabildningen $\mathbf{T} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ fordobler alle annenkomponenter, men endrer ikke førstekomponenter. Finn matrisen til \mathbf{T} .

6. Lineærabildningen $\mathbf{T} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ gjør alle vektorer dobbelt så lange og dreier dem en vinkel θ i positiv retning. Finn matrisen til \mathbf{T} .

7. Lineæravbildningen $\mathbf{T} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ avbilder alle vektorer på sin projeksjon ned i xy -planet. Finn matrisen til \mathbf{T} .

8. Lineæravbildningen $\mathbf{T} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ speiler alle vektorer om x -aksen og dreier dem deretter en vinkel θ i positiv retning. Finn matrisen til \mathbf{T} (det kan være lurt å tenke på \mathbf{T} som sammensetningen av to enklere avbildninger).

9. La A_θ være rotasjonsmatrisen i eksempel 3. Forklar at $A_{-\theta}$ er den inverse matrisen til A_θ uten å regne. Kontroller ved å regne ut $A_\theta A_{-\theta}$.

10. \mathcal{L} er linjen gjennom origo som fremkommer når vi dreier x -aksen en vinkel θ i positiv retning. Lineæravbildningen $\mathbf{T} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ speiler alle punkter om linjen \mathcal{L} . La A_ϕ være matrisen til avbildningen som dreier alle vektorer en vinkel ϕ i positiv retning, og la B være matrisen til avbildningen som speiler alle punkter om x -aksen. Forklar at matrisen C til T er gitt ved

$$C = A_\theta B A_{-\theta}$$

og bruk denne formelen til å finne C .

11. La $\mathbf{a} = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$ og $\mathbf{b} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$.

a) Finn tall x, y, z, u slik at $\mathbf{e}_1 = x\mathbf{a} + y\mathbf{b}$ og $\mathbf{e}_2 = z\mathbf{a} + u\mathbf{b}$.

b) Lineæravbildningen $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ tilfredsstiller $T(\mathbf{a}) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $T(\mathbf{b}) = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. Finn $T(\mathbf{e}_1)$ og $T(\mathbf{e}_2)$.

c) Finn matrisen til \mathbf{T} .

12. a) Finn tall x, y slik at $\mathbf{e}_1 = x\mathbf{a} + y\mathbf{b}$ der \mathbf{a}, \mathbf{b} er som i eksempel 4. Finn $\mathbf{T}^4(\mathbf{e}_1)$ der \mathbf{T} er lineæravbildningen i eksemplet.

b) Finn tall u, v slik at $\mathbf{e}_2 = u\mathbf{a} + v\mathbf{b}$. Finn $\mathbf{T}^4(\mathbf{e}_2)$. Hva er matrisen til \mathbf{T}^4 ?

13. Fullfør eksempel 5 ved å sjekke at \mathbf{v}_2 er en egenvektor med egenverdi λ_2 .

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

a) Vis at $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ er en egenvektor for A med egenverdi $\lambda_1 = 3$.

b) Vis at $\mathbf{v}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ er en egenvektor for A med egenverdi $\lambda_1 = -1$.

c) La $\mathbf{a} = \begin{pmatrix} 3 \\ -1 \end{pmatrix}$. Finn tall x, y slik at $\mathbf{a} = x\mathbf{v}_1 + y\mathbf{v}_2$. Regn ut $A^{10}\mathbf{a}$.

15. Anta at $\mathbf{F} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ tilfredsstiller

$$\mathbf{F}(c\mathbf{x} + d\mathbf{y}) = c\mathbf{F}(\mathbf{x}) + d\mathbf{F}(\mathbf{y})$$

for alle $c, d \in \mathbb{R}$ og alle $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Vis at \mathbf{F} er en lineærabildning.

16. Anta at $\mathbf{a}_1, \mathbf{a}_2 \in \mathbb{R}^2$ ($\mathbf{a}_1, \mathbf{a}_2 \neq \mathbf{0}$) ikke er parallelle, og la $\mathbf{b}_1, \mathbf{b}_2$ være to vektorer i \mathbb{R}^2 . Vis at de finnes nøyaktig én lineærabildning $\mathbf{T} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ slik at $\mathbf{T}(\mathbf{a}_1) = \mathbf{b}_1$ og $\mathbf{T}(\mathbf{a}_2) = \mathbf{b}_2$.

2.9 Affinavbildninger og lineariseringer

Som vi så i eksemplene i forrige seksjon, er lineærabildninger ofte nyttige når vi skal beskrive geometriske transformasjoner som refleksjoner og rotaasjoner. De har imidlertid en stor svakhet; siden en lineærabildning alltid tilfredsstiller $\mathbf{T}(\mathbf{0}) = \mathbf{0}$, kan ikke lineærabildninger brukes til å forskyve figurer i planet. Vi skal nå utvide klassen av avbildninger slik at vi også kan behandle forskyvninger (eller *translasjoner* som matematikere liker å kalle dem).

Definisjon 2.9.1 En funksjon $\mathbf{F} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ kalles en affinavbildning dersom det finnes en $m \times n$ -matrise A og en vektor $\mathbf{c} \in \mathbb{R}^m$ slik at

$$\mathbf{F}(\mathbf{x}) = A\mathbf{x} + \mathbf{c} \quad \text{for alle } \mathbf{x} \in \mathbb{R}^n$$

Vi kaller A matrisen til \mathbf{F} og \mathbf{c} konstantleddet til \mathbf{F} .

Vi ser at lineærabildninger rett og slett er affinavbildninger med konstantledd \mathbf{c} lik $\mathbf{0}$. Vi ser også at translasjonen $\mathbf{F}(\mathbf{x}) = \mathbf{x} + \mathbf{c}$ som forskyver alle vektorer en distanse \mathbf{c} , er en affinavbildning siden den kan skrives

$$\mathbf{F}(\mathbf{x}) = I_n \mathbf{x} + \mathbf{c}$$

(husk at I_n er $n \times n$ -identitetsmatrisen og at $I_n \mathbf{x} = \mathbf{x}$ for alle $\mathbf{x} \in \mathbb{R}^n$).

En viktig egenskap ved affinavbildninger er at de avbilder rette linjer på rette linjer. La oss være helt sikre på at vi skjønner hva dette betyr. Anta at \mathcal{L} er den rette linjen i \mathbb{R}^n som går gjennom punktet \mathbf{a} og har retningsvektor \mathbf{b} (det betyr at \mathcal{L} er samlingen av alle punkter på formen $\mathbf{r}(t) = \mathbf{a} + t\mathbf{b}$ slik vi så i seksjon 1.2). Dersom \mathbf{F} er en kontinuerlig funksjon fra \mathbb{R}^n til \mathbb{R}^m , kan vi bruke \mathbf{F} på alle punktene som ligger på linjen \mathcal{L} . Vi får da en samling av punkter i \mathbb{R}^m som vi kaller *bildet av \mathcal{L} under \mathbf{F}* . Vanligvis vil dette bildet være en kurve i \mathbb{R}^m . Når \mathbf{F} er en affinavbildning, er denne kurven en rett linje.

Setning 2.9.2 Anta at $\mathbf{F}(\mathbf{x}) = A\mathbf{x} + \mathbf{c}$ er en affinavbildning fra \mathbb{R}^n til \mathbb{R}^m , og la $\mathbf{r}(t) = \mathbf{a} + t\mathbf{b}$ være parameterfremstillingen til en linje \mathcal{L} i \mathbb{R}^n . Dersom $A\mathbf{b} \neq \mathbf{0}$, vil bildet av \mathcal{L} under \mathbf{F} være linjen i \mathbb{R}^m som går gjennom punktet $\mathbf{F}(\mathbf{a})$ og har retningsvektor $A\mathbf{b}$.

Bevis: Vi ser at

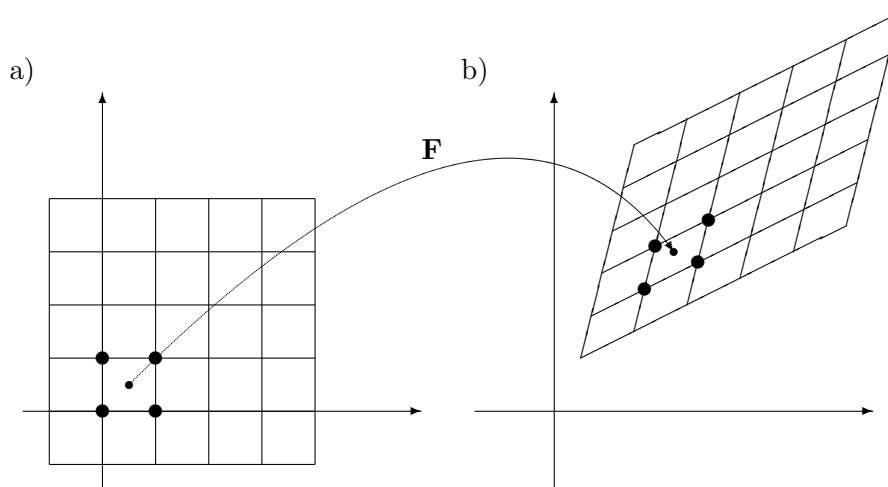
$$\mathbf{F}(\mathbf{r}(t)) = A(\mathbf{a} + t\mathbf{b}) + \mathbf{c} = A\mathbf{a} + \mathbf{c} + t(A\mathbf{b}) = \mathbf{F}(\mathbf{a}) + t(A\mathbf{b})$$

som er parameterframstillingen til en rett linje som går gjennom punktet $\mathbf{F}(\mathbf{a})$ og har retningsvektor $A\mathbf{b}$. \square

Bemerkning. Dersom $A\mathbf{b} = \mathbf{0}$, ser vi at bildet av \mathcal{L} degenererer til ett eneste punkt. Legg også merke til at dersom vi bruker \mathbf{F} på to parallelle linjer (dvs. to linjer med samme retningsvektor \mathbf{b}), så blir også de resulterende linjene parallelle (fordi de får samme retningsvektor $A\mathbf{b}$).

Determinanten som forstørrelsesfaktor

Som vi akkurat har sett, er bildet av en rett linje under en affinavbildning selv en rett linje. Vi vet også at affinavbildninger avbilder parallelle linjer på parallelle linjer. Figur 1 illustrerer dette for en affinavbildning $\mathbf{F} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$; rutenettet i punkt a) avbildes på det forskjøvede og fordreide rutenettet i punkt b). Kvadratet med markerte hjørner i a) avbildes på parallelogrammet med markerte hjørner i b).



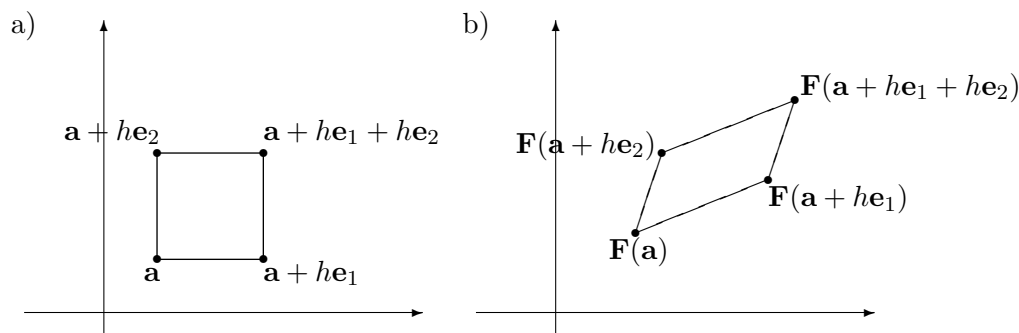
Figur 1: En affinavbildning \mathbf{F} anvendt på et rutenett

Et spørsmål som ofte dukker opp, er hvor mye avbildningen F forstørrer eller forminsker arealet: Hvor stort er arealet til parallelogrammet vi ender opp med, sammenlignet med arealet til kvadratet vi startet med? Figur 2 viser et kvadrat før og etter vi har brukt \mathbf{F} på det. Arealet til kvadratet i punkt a) er h^2 . Parallelogrammet i b) er utspent av vektorene $\mathbf{F}(\mathbf{a} + h\mathbf{e}_1) - \mathbf{F}(\mathbf{a})$ og $\mathbf{F}(\mathbf{a} + h\mathbf{e}_2) - \mathbf{F}(\mathbf{a})$. La oss se nærmere på disse størrelsene.

Siden $\mathbf{F} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ er en affinavbildning, er den på formen

$$\mathbf{F}(\mathbf{x}) = A\mathbf{x} + \mathbf{c}$$

der $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ er en 2×2 -matrise og $\mathbf{c} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$ er en vektor i \mathbb{R}^2 (siden det er matrisemultiplikasjon involvert, må vi skrive vektorene som søylevektorer). Vi ser nå at



Figur 2: Bildet av et kvadrat under \mathbf{F}

$$\mathbf{F}(\mathbf{a} + h\mathbf{e}_1) - \mathbf{F}(\mathbf{a}) = (A(\mathbf{a} + h\mathbf{e}_1) + \mathbf{c}) - (A\mathbf{a} + \mathbf{c}) = hA\mathbf{e}_1 = h \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}$$

$$\mathbf{F}(\mathbf{a} + h\mathbf{e}_2) - \mathbf{F}(\mathbf{a}) = (A(\mathbf{a} + h\mathbf{e}_2) + \mathbf{c}) - (A\mathbf{a} + \mathbf{c}) = hA\mathbf{e}_2 = h \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}$$

Parallelogrammet er derfor utspent av vektorene (ha_{11}, ha_{21}) og (ha_{12}, ha_{22}) , og har — ifølge setning 1.8.1 — areal

$$\left| \begin{pmatrix} ha_{11} & ha_{21} \\ ha_{12} & ha_{22} \end{pmatrix} \right| = |h^2 a_{11} a_{22} - h^2 a_{12} a_{21}| = h^2 |\det(A)|$$

Vi ser altså at arealet har endret seg med en faktor $|\det(A)|$; tallverdien til determinanten er *forstørrelsesfaktoren* til affinavbildningen F .

Vi kan gjennomføre akkurat det samme resonnementet i det tredimensjonale tilfellet ved å dele rommet opp i små terninger med sider parallelle med aksene. En affinavbildning $\mathbf{F}(\mathbf{x}) = A\mathbf{x} + \mathbf{c}$ vil avbilde disse terningene på parallelepipedene, og volmet til parallelepipedene vil være $|\det(A)|$ ganger volumet til terningene. La oss oppsummere resultatene.

Setning 2.9.3 Dersom $\mathbf{F} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ er en affinavbildning med matrise A , så forstørrer \mathbf{F} arealer med en faktor $|\det(A)|$. Dersom $\mathbf{F} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ er en affinavbildning med matrise A , så forstørrer \mathbf{F} volumer med en faktor $|\det(A)|$. \square

Bemerkning: Det kan se ut som vi tar i litt vel kraftig i setningen ovenfor — strengt tatt har vi vel bare vist at $|\det(A)|$ er forstørrelsesfaktoren for arealet til kvadrater, og ikke for mer generelle arealer? Det viser seg imidlertid at alle andre mengder vi kan definere arealet til, kan tilnærmes med kvadrater,

og resultatet gjelder derfor generelt. Tilsvarende gjelder for terninger i det tredimensjonale tilfellet.

Det er naturlig å spørre om fortegnet til determinanten har en geometrisk tolkning også for avbildninger. Det har det: Fortegnet er positivt dersom F bevarer orienteringen til enhetsvektorene \mathbf{i} og \mathbf{j} (\mathbf{i} , \mathbf{j} og \mathbf{k} i det tredimensjonale tilfellet) og negativt om F reverserer orienteringen.

Linearisering

Til slutt i denne seksjonen skal vi se hvordan vi kan bruke affinavbildninger til å generalisere tangentbegrepet fra funksjoner $f : \mathbb{R} \rightarrow \mathbb{R}$ til funksjoner $\mathbf{F} : \mathbb{R}^n \rightarrow \mathbb{R}^m$. Vi skal først se på Jacobi-matrisen til en affinavbildning \mathbf{F} . Anta at \mathbf{F} er gitt ved $\mathbf{F}(\mathbf{x}) = A\mathbf{x} + \mathbf{c}$ der

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

og

$$\mathbf{c} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{pmatrix}$$

Skriver vi ut definisjonen $\mathbf{F}(\mathbf{x}) = A\mathbf{x} + \mathbf{c}$ på komponentform, får vi

$$\begin{aligned} \mathbf{F}(\mathbf{x}) &= \begin{pmatrix} F_1(\mathbf{x}) \\ F_2(\mathbf{x}) \\ \vdots \\ F_m(\mathbf{x}) \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{pmatrix} = \\ &= \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n + c_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n + c_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n + c_m \end{pmatrix} \end{aligned}$$

Deriverer vi den i -te komponenten F_i med hensyn på den j -te variabelen x_j , får vi

$$\frac{\partial F_i}{\partial x_j} = \frac{\partial}{\partial x_j}(a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n + c_i) = a_{ij}$$

Dette betyr at Jacobi-matrisen til en affinavbildning \mathbf{F} rett og slett er matrisen A til \mathbf{F} .

Setning 2.9.4 Anta at affinavbildningen $\mathbf{F} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ er gitt ved $\mathbf{F}(\mathbf{x}) = A\mathbf{x} + \mathbf{c}$. Da er Jacobi-matrisen til \mathbf{F} lik matrisen A til \mathbf{F} . \square

Vi er nå klare til å studere hvordan affinavbildninger kan brukes til å tilnærme mer generelle funksjoner. Husk at hvis $f : \mathbb{R} \rightarrow \mathbb{R}$ er en funksjon av én variabel, er tangenten $T_a f$ i punktet a gitt ved

$$T_a f(x) = f(a) + f'(a)(x - a)$$

Starter vi isteden med en funksjon $\mathbf{F} : \mathbb{R}^n \rightarrow \mathbb{R}^m$, kan vi på tilsvarende måte definere en funksjon $T_a \mathbf{F}$ ved å etterligne uttrykket ovenfor:

$$T_a \mathbf{F}(\mathbf{x}) = \mathbf{F}(\mathbf{a}) + \mathbf{F}'(\mathbf{a})(\mathbf{x} - \mathbf{a})$$

Ganger vi ut parentesen, får vi

$$T_a \mathbf{F}(\mathbf{x}) = \mathbf{F}(\mathbf{a}) - \mathbf{F}'(\mathbf{a})\mathbf{a} + \mathbf{F}'(\mathbf{a})\mathbf{x}$$

som viser at $T_a \mathbf{F}$ er en affinavbildning med matrise $\mathbf{F}'(\mathbf{a})$ og konstantledd $\mathbf{F}(\mathbf{a}) - \mathbf{F}'(\mathbf{a})\mathbf{a}$.

Definisjon 2.9.5 Anta at $\mathbf{F} : A \rightarrow \mathbb{R}^m$ er en funksjon av n variable som er deriverbar i punktet \mathbf{a} . Affinavbildningen $T_a \mathbf{F} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ gitt ved

$$T_a \mathbf{F}(\mathbf{x}) = \mathbf{F}(\mathbf{a}) + \mathbf{F}'(\mathbf{a})(\mathbf{x} - \mathbf{a})$$

kalles lineariseringen til \mathbf{F} i punktet \mathbf{a} .

Legg merke til at $T_a \mathbf{F}(\mathbf{a}) = \mathbf{F}(\mathbf{a})$. Siden vi allerede har observert at matrisen til affinavbildningen $T_a \mathbf{F}$ er $\mathbf{F}'(\mathbf{a})$, vet vi fra setning 2.9.3 at Jacobi-matrisen til $T_a \mathbf{F}$ er $\mathbf{F}'(\mathbf{a})$. Dette betyr at lineariseringen $T_a \mathbf{F}$ har samme verdi og samme deriverte i punktet \mathbf{a} som den opprinnelige funksjonen \mathbf{F} . Dette er akkurat samme egenskap som tangenten har for funksjoner av én variabel. En annen viktig (og nært beslektet) egenskap ved tangenten er at den er linjen som smyer seg tettest inntil funksjonsgrafen til f i nærheten av punktet a . Vi skal snart vise en tilsvarende egenskap for lineariseringen — at den er affinavbildningen som ligger tettest opptil \mathbf{F} i nærheten av \mathbf{a} . Men før vi gjør dette, kan det være lurt å kikke på et eksempel.

Eksempel 1: Vi skal finne lineariseringen til funksjonen $\mathbf{F} : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ gitt ved

$$\mathbf{F}(x, y, z) = \begin{pmatrix} x^2 y \\ 2xyz \end{pmatrix}$$

i punktet $\mathbf{a} = \begin{pmatrix} 1 \\ -2 \\ -1 \end{pmatrix}$. La oss først regne ut Jacobi-matrisen til \mathbf{F} :

$$\mathbf{F}'(x, y, z) = \begin{pmatrix} \frac{\partial F_1}{\partial x} & \frac{\partial F_1}{\partial y} & \frac{\partial F_1}{\partial z} \\ \frac{\partial F_2}{\partial x} & \frac{\partial F_2}{\partial y} & \frac{\partial F_2}{\partial z} \end{pmatrix} = \begin{pmatrix} 2xy & x^2 & 0 \\ 2yz & 2xz & 2xy \end{pmatrix}$$

I punktet \mathbf{a} har vi dermed

$$\begin{aligned}\mathbf{F}'(1, -2, -1) &= \begin{pmatrix} 2 \cdot 1 \cdot (-2) & 1^2 & 0 \\ 2 \cdot (-2) \cdot (-1) & 2 \cdot 1 \cdot (-1) & 2 \cdot 1 \cdot (-2) \end{pmatrix} = \\ &= \begin{pmatrix} -4 & 1 & 0 \\ 4 & -2 & -4 \end{pmatrix}\end{aligned}$$

Vi har også

$$\mathbf{F}(1, -2, -1) = \begin{pmatrix} 1^2 \cdot (-2) \\ 2 \cdot 1 \cdot (-2) \cdot (-1) \end{pmatrix} = \begin{pmatrix} -2 \\ 4 \end{pmatrix}$$

Vi kan nå regne ut lineariseringen. Med $\mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ får vi:

$$\begin{aligned}T_{\mathbf{a}}\mathbf{F}(\mathbf{x}) &= \mathbf{F}(\mathbf{a}) + \mathbf{F}'(\mathbf{a})(\mathbf{x} - \mathbf{a}) = \\ &= \begin{pmatrix} -2 \\ 4 \end{pmatrix} + \begin{pmatrix} -4 & 1 & 0 \\ 4 & -2 & -4 \end{pmatrix} \left(\begin{pmatrix} x \\ y \\ z \end{pmatrix} - \begin{pmatrix} 1 \\ -2 \\ -1 \end{pmatrix} \right) = \\ &= \begin{pmatrix} -2 \\ 4 \end{pmatrix} + \begin{pmatrix} -4x + y \\ 4x - 2y - 4z \end{pmatrix} - \begin{pmatrix} -6 \\ 12 \end{pmatrix} = \begin{pmatrix} -4x + y + 4 \\ 4x - 2y - 4z - 8 \end{pmatrix}\end{aligned}$$

Altså er

$$T_{\mathbf{a}}\mathbf{F}(x, y, z) = \begin{pmatrix} -4x + y + 4 \\ 4x - 2y - 4z - 8 \end{pmatrix}$$



La oss vende tilbake til resultatet vi annonserte ovenfor, det som sier at lineariseringen $T_{\mathbf{a}}\mathbf{F}$ er den affinavbildningen som ligger tettest opptil \mathbf{F} i nærheten av \mathbf{a} . Et punkt \mathbf{x} i nærheten av \mathbf{a} kan vi skrive $\mathbf{x} = \mathbf{a} + \mathbf{r}$ der \mathbf{r} er liten. Vi ønsker å vise at differensen $\mathbf{F}(\mathbf{a} + \mathbf{r}) - T_{\mathbf{a}}\mathbf{F}(\mathbf{a} + \mathbf{r})$ er liten sammenlignet med $|\mathbf{r}|$ for små \mathbf{r} , og at tilnærmingen blir bedre og bedre når \mathbf{r} går mot $\mathbf{0}$. En naturlig måte å formulere dette på er å kreve at

$$\lim_{\mathbf{r} \rightarrow \mathbf{0}} \frac{1}{|\mathbf{r}|} \left(\mathbf{F}(\mathbf{a} + \mathbf{r}) - T_{\mathbf{a}}\mathbf{F}(\mathbf{a} + \mathbf{r}) \right) = \mathbf{0}$$

Bruker vi at

$$T_{\mathbf{a}}\mathbf{F}(\mathbf{a} + \mathbf{r}) = \mathbf{F}(\mathbf{a}) + \mathbf{F}'(\mathbf{a})((\mathbf{a} + \mathbf{r}) - \mathbf{a}) = \mathbf{F}(\mathbf{a}) + \mathbf{F}'(\mathbf{a})\mathbf{r}$$

ser vi at

$$\mathbf{F}(\mathbf{a} + \mathbf{r}) - T_{\mathbf{a}}\mathbf{F}(\mathbf{a} + \mathbf{r}) = \mathbf{F}(\mathbf{a} + \mathbf{r}) - \mathbf{F}(\mathbf{a}) - \mathbf{F}'(\mathbf{a})\mathbf{r} = \boldsymbol{\sigma}(\mathbf{r})$$

der $\boldsymbol{\sigma}$ er som i definisjon 2.6.1. Nå er det ikke så vanskelig å formulere og bevise resultatet.

Teorem 2.9.6 Anta at $\mathbf{F} : A \rightarrow \mathbb{R}^m$ er en funksjon av n variable som er deriverbar i punktet \mathbf{a} , og la $T_{\mathbf{a}}\mathbf{F}$ være lineariseringen til \mathbf{F} i \mathbf{a} . Da er

$$\lim_{\mathbf{r} \rightarrow \mathbf{0}} \frac{1}{|\mathbf{r}|} \left(\mathbf{F}(\mathbf{a} + \mathbf{r}) - T_{\mathbf{a}}\mathbf{F}(\mathbf{a} + \mathbf{r}) \right) = \mathbf{0}$$

Det finnes ingen annen affinavbildning $\mathbf{G} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ slik at

$$\lim_{\mathbf{r} \rightarrow \mathbf{0}} \frac{1}{|\mathbf{r}|} \left(\mathbf{F}(\mathbf{a} + \mathbf{r}) - \mathbf{G}(\mathbf{a} + \mathbf{r}) \right) = \mathbf{0}$$

Bevis: Det følger av definisjonen av deriverbarhet (definisjon 2.6.1) og regningene ovenfor at

$$\lim_{\mathbf{r} \rightarrow \mathbf{0}} \frac{1}{|\mathbf{r}|} \left(\mathbf{F}(\mathbf{a} + \mathbf{r}) - T_{\mathbf{a}}\mathbf{F}(\mathbf{a} + \mathbf{r}) \right) = \lim_{\mathbf{r} \rightarrow \mathbf{0}} \frac{1}{|\mathbf{r}|} \boldsymbol{\sigma}(\mathbf{r}) = \mathbf{0}$$

Dermed er første del av teoremet bevist. For å bevise den andre delen, antar vi at $\mathbf{G}(\mathbf{x})$ er en affinavbildning slik at

$$\lim_{\mathbf{r} \rightarrow \mathbf{0}} \frac{1}{|\mathbf{r}|} \left(\mathbf{F}(\mathbf{a} + \mathbf{r}) - \mathbf{G}(\mathbf{a} + \mathbf{r}) \right) = \mathbf{0}$$

Vi må vise at $\mathbf{G} = T_{\mathbf{a}}\mathbf{F}$. Siden \mathbf{G} er en affinavbildning, vet vi at $\mathbf{G}(\mathbf{x}) = A\mathbf{x} + \mathbf{c}$ for en matrise A og en vektor \mathbf{c} . Setter vi dette inn i uttrykket ovenfor, får vi

$$\lim_{\mathbf{r} \rightarrow \mathbf{0}} \frac{1}{|\mathbf{r}|} \left(\mathbf{F}(\mathbf{a} + \mathbf{r}) - \mathbf{G}(\mathbf{a} + \mathbf{r}) \right) = \lim_{\mathbf{r} \rightarrow \mathbf{0}} \frac{1}{|\mathbf{r}|} \left(\mathbf{F}(\mathbf{a} + \mathbf{r}) - A\mathbf{a} - A\mathbf{r} - \mathbf{c} \right) = \mathbf{0}$$

Skal dette være mulig, må $\mathbf{F}(\mathbf{a} + \mathbf{r}) - A\mathbf{a} - A\mathbf{r} - \mathbf{c} \rightarrow \mathbf{0}$ når $\mathbf{r} \rightarrow \mathbf{0}$, og det er bare tilfellet hvis $\mathbf{F}(\mathbf{a}) = A\mathbf{a} + \mathbf{c}$. Setter vi $A\mathbf{a} + \mathbf{c} = \mathbf{F}(\mathbf{a})$ inn i det siste uttrykket ovenfor, ser vi at

$$\lim_{\mathbf{r} \rightarrow \mathbf{0}} \frac{1}{|\mathbf{r}|} \left(\mathbf{F}(\mathbf{a} + \mathbf{r}) - \mathbf{F}(\mathbf{a}) - A\mathbf{r} \right) = \mathbf{0}$$

og ifølge setning 2.6.4 er det bare mulig dersom $A = \mathbf{F}'(\mathbf{a})$. Siden vi allerede vet at $\mathbf{F}(\mathbf{a}) = A\mathbf{a} + \mathbf{c}$, følger det at $\mathbf{c} = \mathbf{F}(\mathbf{a}) - A\mathbf{a} = \mathbf{F}(\mathbf{a}) - \mathbf{F}'(\mathbf{a})\mathbf{a}$. Dermed er

$$\mathbf{G}(\mathbf{x}) = A\mathbf{x} + \mathbf{c} = \mathbf{F}'(\mathbf{a})\mathbf{x} + \mathbf{F}(\mathbf{a}) - \mathbf{F}'(\mathbf{a})\mathbf{a} = \mathbf{F}(\mathbf{a}) + \mathbf{F}'(\mathbf{a})(\mathbf{x} - \mathbf{a}) = T_{\mathbf{a}}\mathbf{F}(\mathbf{x})$$

□

Det som er viktigst å ta med seg fra teoremet ovenfor, er at lineariseringen $T_{\mathbf{a}}\mathbf{F}$ er en svært god tilnærming til \mathbf{F} i området rundt \mathbf{a} . Dette skal vi få bruk for både når vi studerer Newtons metode i flere variable, og når vi

studerer skifte av variable i multiple integraler (dvs. integraler av funksjoner av flere variable).

Bemerkning: Du har kanskje lurt på hvorfor $T_{\mathbf{a}}\mathbf{F}$ kalles en *linearisering* og ikke en *affinisering* — $T_{\mathbf{a}}\mathbf{F}$ er tross alt en affinavbildning og ikke (bortsett fra i helt spesielle tilfeller) en lineæravbildning. Dessverre er bruken av ordparet lineær/affin svært usystematisk og forvirrende; ordene brukes om hverandre på en måte som ikke er lett å få oversikt over. Et annet eksempel på denne vaklingen får du om du betrakter en funksjon $f : \mathbb{R} \rightarrow \mathbb{R}$ på formen $f(x) = ax + b$. Når man studerer funksjoner av én variabel, er alle enige om å kalle dette en lineær funksjon, men ifølge terminologien vi nå har innført, er den bare en lineæravbildning hvis $c = 0$; i alle andre tilfeller må vi nøye oss med å kalle den en affinavbildning! Heldigvis er det sjelden at den vaklende terminologibruken skaper misforståelser, men det er greit å vite om den slik at man kan være litt på vakt.

Opppgaver til seksjon 2.9

1. Finn matrisen og konstantleddet til affinavbildningen

$$\mathbf{F}(x, y, z) = \begin{pmatrix} 2x - 3y + z - 7 \\ -x + z - 2 \end{pmatrix}$$

2. \mathcal{L} er linjen i \mathbb{R}^3 med parametrisering

$$\mathbf{r}(t) = \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix} + t \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$$

og $\mathbf{F} : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ er affinavbildningen

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

Finn en parametrisering av bildet av \mathcal{L} under \mathbf{F} .

3. Finn lineariseringen til funksjonen $\mathbf{F} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ gitt ved

$$\mathbf{F}(x, y) = \begin{pmatrix} x^2 y \\ xy + x \end{pmatrix}$$

i punktet $\mathbf{a} = (-2, 1)$.

4. Finn lineariseringen til funksjonen $\mathbf{F} : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ gitt ved

$$\mathbf{F}(x, y) = \begin{pmatrix} x \sin(xy) \\ xe^y \\ 2x^3 + y \end{pmatrix}$$

i punktet $\mathbf{a} = (2, 0)$.

5. En affinavbildning $\mathbf{F} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ tilfredsstiller $\mathbf{F}(0, 0) = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, $\mathbf{F}(1, 0) = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$ og $\mathbf{F}(0, 1) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$. Finn matrisen og konstantleddet til \mathbf{F} .

6. Affinavbildningen $\mathbf{F} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ dreier enhver vektor en vinkel $\frac{\pi}{4}$ i positiv retning, og flytter den deretter en distanse $(3, -1)$. Finn matrisen og konstantleddet til \mathbf{F} .

7. a) Affinavbildningen $\mathbf{F} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ avbilder ethvert punkt på sitt speilbilde om den vertikale linjen $x = 3$. Finn matrisen og konstantleddet til \mathbf{F} .

b) Affinavbildningen $\mathbf{G} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ avbilder ethvert punkt på sitt speilbilde om den horisontale linjen $y = -2$. Finn matrisen og konstantleddet til \mathbf{F} .

8. Affinavbildningen $\mathbf{F} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ avbilder ethvert punkt på sitt speilbilde om linjen $y = x + 1$. Finn matrisen og konstantleddet til \mathbf{F} .

9. Anta at punktene $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3 \in \mathbb{R}^2$ ikke ligger på samme rette linje, og la $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$ være tre punkter i \mathbb{R}^2 . Vis at det finnes nøyaktig én affinavbildning $\mathbf{F} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ slik at $\mathbf{F}(\mathbf{a}_1) = \mathbf{b}_1$, $\mathbf{F}(\mathbf{a}_2) = \mathbf{b}_2$ og $\mathbf{F}(\mathbf{a}_3) = \mathbf{b}_3$.

Fasit

Finner du feil i fasiten, så send en e-post til *lindstro@math.uio.no*.

Seksjon 1.1

Oppgave 1: $\mathbf{a}+\mathbf{b} = (-2, 3, 9, -5, -2)$, $\mathbf{a}-\mathbf{b} = (4, -7, -1, -5, 4)$, $s\mathbf{a} = (3, -6, 12, -15, 3)$,
 $\mathbf{a} \cdot \mathbf{b} = 4$

Oppgave 2: $\mathbf{a}+\mathbf{b} = (7, 2, 5, -8, -5, 3)$, $\mathbf{a}-\mathbf{b} = (7, -2, 3, 4, -5, 5)$, $s\mathbf{a} = (-28, 0, -16, 8, 20, -16)$,
 $\mathbf{a} \cdot \mathbf{b} = 12$

Oppgave 5: $\mathbf{m} \cdot \mathbf{p}$ der $\mathbf{m} = (m_1, m_2, \dots, m_n)$ og $\mathbf{p} = (p_1, p_2, \dots, p_n)$.

Seksjon 1.2

Oppgave 1: Skalarprodukt: -5 , vinkel: $\theta \approx 109.65^\circ$

Oppgave 2: $10\sqrt{2}$

Oppgave 3: ca. 67.8°

Oppgave 4: ca. 31.8° .

Oppgave 5: ca. 70° , $\mathbf{p} = \frac{1}{2}\mathbf{b}$

Oppgave 6: $\frac{12}{7}\sqrt{14}$

Oppgave 7: $(4, 3) = (2, 4) + (2, -1)$

Oppgave 8: $(2, 2, 1) = (\frac{1}{2}, 0, -\frac{1}{2}) + (\frac{3}{2}, 2, \frac{3}{2})$

Oppgave 9: 30° og 45° , $\cos 15^\circ = \frac{\sqrt{6}+\sqrt{2}}{4}$

Oppgave 10: Utallige muligheter, f.eks. $(1, 0, 3)$ og $(0, 1, 2)$. Sjekk dine svar ved å ta skalarproduktet med $(3, 2, -1)$.

Oppgave 12: b) $-\frac{43}{2}$, c) 90°

Oppgave 13: Umulig ifølge trekantulikheten.

Oppgave 14: Umulig ifølge Schwarz' ulikhet.

Oppgave 18: For eksempel $\mathbf{r}(t) = (-1 + 2t, -1 + 3t, 2 + t)$ (det er mange muligheter)

Oppgave 19: For eksempel $\mathbf{r}(t) = (-3 + t, -2 - 2t, 5 - t, 8 + 3t)$ (det er mange muligheter). Nei

Oppgave 20: For eksempel $\mathbf{r}(t) = (2 + t, -1 + 9t, 3 - 5t)$ (det er mange muligheter).

Oppgave 21: For eksempel $\mathbf{r}(t) = (7 - 5t, -3 + 4t, 2 - 3t, 4 - 5t, -2 + 7t)$ (det er mange muligheter).

Oppgave 22: For eksempel $\mathbf{r}(t) = (5 + 2t, -2 + t)$ (det er mange muligheter)

Oppgave 23: For eksempel $\mathbf{r}(t) = (3t, 2 - 2t)$ (det er mange muligheter)

Oppgave 24: $y = -\frac{1}{2}x + \frac{1}{2}$

Oppgave 25: a) Skjæringspunkt $(15, 24)$.

b) Skipene kolliderer ikke.

Oppgave 26:

a) Ja, kursene til flyene krysser hverandre i punktet (4000, 4000, 4000).

b) Nei, flyene kolliderer ikke. De kommer til møtestedet etter henholdsvis 40 sekunder og $\frac{100}{3}$ sekunder.

Oppgave 27: c) ca. 49° og ca. 12.7 m.

Seksjon 1.3

Oppgave 1: $\begin{pmatrix} 4 + 5i \\ -3 + 4i \end{pmatrix}$.

Oppgave 2: $|\mathbf{a}| = \sqrt{15}$, $|\mathbf{b}| = \sqrt{19}$

Oppgave 3: $-1 - i$

Seksjon 1.4

Oppgave 1: a) (19, 3, 5) b) (-1, -6, -14)

Oppgave 2: $6\sqrt{5}$

Oppgave 3: $\sqrt{3}$

Oppgave 4: (9, -5, 6)

Oppgave 5: $\mathbf{i} \times \mathbf{j} = \mathbf{k}$, $\mathbf{j} \times \mathbf{k} = \mathbf{i}$, $\mathbf{k} \times \mathbf{i} = \mathbf{j}$

Oppgave 6: 0

Oppgave 7: $\frac{7}{2}$

Oppgave 8: $2x + y - z = 4$

Oppgave 9: $3x - y + 2z = 3$

Oppgave 10: Hint: Bruk setning 1.4.4.

Seksjon 1.5

Oppgave 1: $2A = \begin{pmatrix} 6 & -4 & 8 \\ -2 & 0 & 12 \end{pmatrix}$, $-3B = \begin{pmatrix} -6 & -6 & 3 \\ -12 & 0 & 18 \end{pmatrix}$

$A + B = \begin{pmatrix} 5 & 0 & 3 \\ 3 & 0 & 0 \end{pmatrix}$, $A - B = \begin{pmatrix} 1 & -4 & 5 \\ -5 & 0 & 12 \end{pmatrix}$

Oppgave 2: $4A - 3B = \begin{pmatrix} -5 & -18 \\ 14 & -1 \\ 3 & -4 \end{pmatrix}$.

Oppgave 3: $A^T = \begin{pmatrix} 1 & -1 \\ -2 & 8 \\ 6 & 6 \end{pmatrix}$, $B^T = \begin{pmatrix} 2 & 4 & 1 \\ 2 & 0 & 7 \\ -1 & -6 & -3 \end{pmatrix}$

Oppgave 4: $A^T = \begin{pmatrix} 1 & -5 & 2 \\ 0 & 7 & -1 \end{pmatrix}$, $B^T = \begin{pmatrix} 4 & -1 & 2 \\ -1 & 3 & -6 \end{pmatrix}$

$(4A - B)^T = \begin{pmatrix} 0 & -19 & 6 \\ 1 & 25 & 2 \end{pmatrix}$

Oppgave 5: a) $\begin{pmatrix} 1 \\ -7 \end{pmatrix}$

b) $\begin{pmatrix} 6 \\ 7 \\ 22 \end{pmatrix}$

c) $\begin{pmatrix} 8 \\ -18 \\ 10 \end{pmatrix}$

Oppgave 6: I: 9.7 tonn, II: 11.8 tonn, III: 7.5 tonn

Oppgave 7: X: 88 vogner, Y: 61 vogner, Z: 51 vogner

Oppgave 10: $\begin{pmatrix} 0.1 & 0 & 0.2 \\ 0.2 & 0.3 & 0.2 \\ 0.3 & 0.3 & 0.1 \\ 0.3 & 0.3 & 0.1 \\ 0.1 & 0.1 & 0.4 \end{pmatrix}$

A: 2.6 enheter, B: 7.2 enheter, C: 7.4 enheter, D: 7.4 enheter, E: 5.4 enheter

Oppgave 11: a) $\begin{pmatrix} 0.94 & 0 & 0.01 \\ 0.05 & 0.2 & 0 \\ 0.01 & 0.8 & 0.99 \end{pmatrix}$

b) $\mathbf{v}_1 = \begin{pmatrix} 0.846 \\ 0.065 \\ 0.089 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} 0.79613 \\ 0.0553 \\ 0.14857 \end{pmatrix}$ (vi tar med flere siffer enn det som er rimelig for å gjøre svaret lett å kontrollere).

Oppgave 12: a) $\begin{pmatrix} 0 & 20 & 50 & 10 \\ 0.05 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & 0 \\ 0 & 0 & 0.3 & 0.1 \end{pmatrix}$

b) Unge: 300, unge voksne 250, voksne 0, eldre 3.

Oppgave 13: a) $\begin{pmatrix} 0 & 0.1 & 0.5 \\ 0.1 & 0 & 0.5 \\ 0.9 & 0.9 & 0 \end{pmatrix}$

b) Per: 45%, Pål: 46%, Espen 9%.

Seksjon 1.6

Oppgave 1. a) $AB = \begin{pmatrix} 0 & -5 \\ 7 & -1 \end{pmatrix}$, $BA = \begin{pmatrix} -1 & -5 \\ 7 & 0 \end{pmatrix}$

b) $AB = \begin{pmatrix} 1 & 4 & 1 \\ 3 & -5 & 0 \\ 1 & -7 & 1 \end{pmatrix}$, $BA = \begin{pmatrix} -5 & 2 & 3 \\ 3 & 1 & -2 \\ 3 & 1 & 1 \end{pmatrix}$

Oppgave 2. $AB = \begin{pmatrix} 5 & 7 \\ 2 & 3 \end{pmatrix}$.

Oppgave 3. $AB = \begin{pmatrix} 8 & -1 & -2 \\ 6 & 3 & 0 \\ -8 & 1 & 2 \end{pmatrix}$.

Oppgave 4. a) AB blir en 8×9 -matrise.

b) B er en 3×5 -matrise

c) B har 7 søyler

Oppgave 5. a) $AB = \begin{pmatrix} 83 & 62 \\ 72 & 48 \\ 50 & 80 \end{pmatrix}$

b) $AC = \begin{pmatrix} 3 & 81 & 41 \\ 0 & 72 & 32 \\ 21 & 36 & 51 \end{pmatrix}$

c) Uttrykket er ikke definert siden matrisene B og C ikke har samme dimensjon (dermed er ikke summen $B + C$ definert).

d) $(BC)^T = \begin{pmatrix} 6 & 27 \\ 72 & 54 \\ 42 & 69 \end{pmatrix}$

e) $B^T C^T$ er ikke definert siden B^T er en 2×2 -matrise og C^T er en 3×2 -matrise, så dimensjonene stemmer ikke overens.

f) $(A + C^T)B = \begin{pmatrix} 89 & 86 \\ 153 & 102 \\ 96 & 144 \end{pmatrix}$

g) $B(A^T - 2C) = \begin{pmatrix} 62 & -80 & -38 \\ 9 & -60 & -51 \end{pmatrix}$

Oppgave 7. Vi finner første rad i DE ved å multiplisere første rad i D med matrisen E :

$$(3, 1, 4) \begin{pmatrix} 2 & 7 & 1 \\ 8 & 2 & 8 \\ 1 & 8 & 2 \end{pmatrix} = (18, 55, 19)$$

Vi finner andre søyle i DE ved å multiplisere matrisen D med andre søyle i E :

$$\begin{pmatrix} 3 & 1 & 4 \\ 1 & 5 & 9 \\ 2 & 6 & 5 \end{pmatrix} \begin{pmatrix} 7 \\ 2 \\ 8 \end{pmatrix} = \begin{pmatrix} 55 \\ 89 \\ 66 \end{pmatrix}$$

Regner du ut hele matriseproduktet på vanlig måte, får du:

$$DE = \begin{pmatrix} 18 & 55 & 19 \\ 51 & 89 & 59 \\ 57 & 66 & 60 \end{pmatrix}$$

Oppgave 8. a) Første og andre søyle i produktmatrisen AB er også like.

b) Andre søyle i produktmatrisen AB består også bare av nuller.

Oppgave 9. 4

Oppgave 10. $A^2 = \begin{pmatrix} 3 & -4 \\ -2 & 3 \end{pmatrix}$, $A^3 = \begin{pmatrix} 7 & -10 \\ -5 & 7 \end{pmatrix}$.

Oppgave 11. a) $\mathbf{y} = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$

b) $C = \begin{pmatrix} 9 & -3 \\ 8 & -3 \end{pmatrix}$

c) $A\mathbf{y} = C\mathbf{x} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$

Oppgave 12. a) $M = \begin{pmatrix} 0.7 & 0.05 & 0.1 \\ 0.15 & 0.75 & 0.15 \\ 0.1 & 0.1 & 0.7 \end{pmatrix}$

$$\begin{aligned} \text{b) } N &= \begin{pmatrix} 1 & 0.2 & 0 \\ 0 & 0.8 & 0.05 \\ 0 & 0 & 0.95 \end{pmatrix} \\ \text{c) } K &= \begin{pmatrix} 0.73 & 0.2 & 0.13 \\ 0.125 & 0.605 & 0.155 \\ 0.095 & 0.095 & 0.665 \end{pmatrix} \end{aligned}$$

Fordeling: A: 258, B: 268.5, C: 313.5

d) Fordeling: A: 282.8, B: 243.3, C: 258.5 (vi tar med flere siffer enn det som er rimelig for å gjøre svaret lett å kontrollere).

Seksjon 1.7

Oppgave 3. Matrisene A , B og D er inverterbare med inverse matriser

$$A^{-1} = \begin{pmatrix} -2 & 7 \\ -1 & 3 \end{pmatrix}, B^{-1} = \begin{pmatrix} 0.2 & 0 \\ 0 & 0.25 \end{pmatrix}, D^{-1} = \begin{pmatrix} -1 & 1 \\ 0.75 & -0.5 \end{pmatrix}$$

Matrisen C er singulær (siden den har en søyle som bare består av nuller).

Oppgave 4. $A = \begin{pmatrix} 2 & -5 \\ -1 & 3 \end{pmatrix}$

Oppgave 5. $(AB)^{-1} = B^{-1}A^{-1} = \begin{pmatrix} 14 & 59 \\ 4 & 17 \end{pmatrix}$

10. b) $A^{-1} = \begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix}$

c) Multipliserer vi med A^{-1} på begge sider av matriseligningen, får vi

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 19 \\ 7 \end{pmatrix}$$

Seksjon 1.8

Oppgave 1: a) 14 b) 38 c) 0

Oppgave 2: 11

Oppgave 3: $\frac{23}{2}$

Oppgave 4: 27

Oppgave 5: a) negativt, b) negativt

Oppgave 6: Hint: Tolk determinanten som et areal.

Oppgave 8: Hint: Utrykk arealet som en determinant

Oppgave 9: b) Ligningssystemet har enten ingen eller uendelig mange løsninger avhengig av konstantene c_1 og c_2 . Linjene $a_1x + b_1y = c_1$ og $a_2x + b_2y = c_2$ er nemlig enten parallelle (ingen løsninger) eller sammenfallende (uendelig mange løsninger).

Oppgave 10: a) 84 b) 20 c) 0

Oppgave 11: 7

Oppgave 12: $\frac{5}{3}$

Oppgave 13: Positivt

Oppgave 15: Hint: Tolk determinanten som et volum.

Oppgave 16: -127

Seksjon 2.1

- Oppgave 1:** a) $D_f = \mathbf{R}^2 \setminus \{(0, 0)\}$
 b) $D_f = \{(x, y) \in \mathbf{R}^2 \mid x \neq y \text{ og } x \neq -y\}$
 c) $D_f = \{(x, y) \in \mathbf{R}^2 \mid x + y > 0\}$
 d) $D_f = \{(x, y) \in \mathbf{R}^2 \mid \text{det ikke finnes et heltall } k \text{ slik at } x - y = \frac{\pi}{2} + k\pi\}$
 e) $D_f = \mathbf{R}^3 \setminus \{\text{kuleflaten med sentrum i origo og radius 5}\}$

Seksjon 2.3

Oppgave 1: a) 20

- b) 1
 c) e
 d) 1

Seksjon 2.4

- Oppgave 1:** a) $\frac{\partial f}{\partial x} = 3x^2y + 3y^4$ og $\frac{\partial f}{\partial y} = x^3 + 12xy^3$
 b) $\frac{\partial f}{\partial x} = \frac{2x+3x^2}{y}$ og $\frac{\partial f}{\partial y} = -\frac{x^2+x^3}{y^2}$
 c) $\frac{\partial f}{\partial x} = -\sin(x+y^2)$ og $\frac{\partial f}{\partial y} = -2y\sin(x+y^2)$
 d) $\frac{\partial f}{\partial x} = 2x\ln(xy^2) + x$ og $\frac{\partial f}{\partial y} = \frac{2x^2}{y}$
 e) $\frac{\partial f}{\partial x} = e^{-z}$, $\frac{\partial f}{\partial y} = e^{-z}$ og $\frac{\partial f}{\partial z} = -(x+y)e^{-z}$
 f) $\frac{\partial f}{\partial x} = \frac{z^2}{(1+y^2)\cos^2(x)}$, $\frac{\partial f}{\partial y} = -\frac{2yz^2\tan x}{(1+y^2)^2}$ og $\frac{\partial f}{\partial z} = \frac{2z\tan x}{1+y^2}$
 g) $\frac{\partial f}{\partial x} = \frac{z}{1+(x+y)^2}$, $\frac{\partial f}{\partial y} = \frac{z}{1+(x+y)^2}$ og $\frac{\partial f}{\partial z} = \arctan(x+y)$
 h) $\frac{\partial f}{\partial x} = -(z^2+u)e^{-x+3y}$, $\frac{\partial f}{\partial y} = 3(z^2+u)e^{-x+3y}$, $\frac{\partial f}{\partial z} = 2ze^{-x+3y}$ og $\frac{\partial f}{\partial u} = e^{-x+3y}$

Oppgave 2: a) $\nabla f = (2xy, x^2)$

- b) $\nabla f = (\cos(xy^2z) - xy^2z\sin(xy^2z), -2x^2yz\sin(xy^2z), -x^2y^2\sin(xy^2z))$
 c) $\nabla f = (w\cos ve^{u\cos v}, -wue^{u\cos v}\sin v, e^{u\cos v})$
 d) $\nabla f = (\frac{z_2z_3}{1+z_1^2z_2^2}, \frac{z_1z_3}{1+z_1^2z_2^2}, \arctan(z_1z_2) + e^{z_3})$

Oppgave 3: a) $f'((1, 2); (3, -1)) = 11$

- b) $f'((1, 0); (-1, 1)) = -1$
 c) $f'((1, 0, 1); (1, 1, -1)) = -1$
 d) $f'((\frac{\pi}{2}, 1, 0); (2, 0, -1)) = -1$

Oppgave 4: a) I retningen gitt av vektoren $(24, 173)$

- b) I retningen gitt av vektoren $(1, 1, 0)$
 c) I retningen gitt av vektoren $(3, 0, -3, 1)$

Oppgave 5: Hint: Bruk at $\Delta V \approx V'((r, h); (\Delta r, \Delta h))$ Anslått usikkerhet: $\Delta V \approx 3.8m^3$

Oppgave 6: Tommelfingerregelen gjelder når høyden i centimeter er det dobbelte av vekten i kilo (f.eks høyde 180 cm og vekt 90 kilo). Er man lettere enn dette, er regelen “for snill” i den forstand at BMI går opp hvis man legger på en kilo og en centimeter.

Seksjon 2.5

- Oppgave 1:** a) $\frac{\partial^2 f}{\partial x^2} = 6y$, $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} = 6x + 4y$, $\frac{\partial^2 f}{\partial y^2} = 4x$
 b) $\frac{\partial^2 f}{\partial x^2} = 0$, $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} = \cos y$, $\frac{\partial^2 f}{\partial y^2} = -x \sin y$
 c) $\frac{\partial^2 f}{\partial x^2} = (x^2 + 4x + 2)e^{x-y}$, $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} = -(x^2 + 2x)e^{x-y}$, $\frac{\partial^2 f}{\partial y^2} = x^2 e^{x-y}$
 d) $\frac{\partial^2 f}{\partial x^2} = 2z$, $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} = 0$, $\frac{\partial^2 f}{\partial x \partial z} = \frac{\partial^2 f}{\partial z \partial x} = 2x$, $\frac{\partial^2 f}{\partial y^2} = -2z^2$,
 $\frac{\partial^2 f}{\partial y \partial z} = \frac{\partial^2 f}{\partial z \partial y} = -4yz$, $\frac{\partial^2 f}{\partial z^2} = -2y^2$

- Oppgave 2:** a) $(6xy + 6x^2yz + x^3yz^2)e^{xz}$.
 b) $-20x^3y^4 \cos(xyz) + 10x^4y^5z \sin(xyz) + x^5y^6z^2 \cos(xyz)$

Seksjon 2.6

- Oppgave 1:** a) $\mathbf{F}'(x, y) = \begin{pmatrix} 2xy & x^2 \\ 1 & 2y \end{pmatrix}$
 b) $\mathbf{F}'(x, y) = \begin{pmatrix} 2xye^{x^2y+z} & x^2e^{x^2y+z} & e^{x^2y+z} \\ yz^2 & xz^2 & 2xyz \end{pmatrix}$
 c) $\mathbf{F}'(x, y) = \begin{pmatrix} \arctan(xy) + \frac{xy}{1+x^2y^2} & \frac{x^2}{1+x^2y^2} \\ \ln y & \frac{x}{y} \\ y \cos y^2 & x \cos y^2 - 2xy^2 \sin y^2 \end{pmatrix}$
 d) $\mathbf{F}'(x, y, z, u) = \begin{pmatrix} y \sin(xu^2) + xyu^2 \cos(xu^2) & x \sin(xu^2) & 0 & 2x^2yu \cos(xu^2) \\ 0 & 0 & 2zu & z^2 \end{pmatrix}$

Seksjon 2.7

- Oppgave 1:** $\frac{\partial k}{\partial x} = 8xy^2 + 1$, $\frac{\partial k}{\partial y} = 8x^2y + 2y$

- Oppgave 2:** $\frac{\partial k}{\partial x} = 2ye^{-2y(x+z)} - (4xy^2 + 2yz)e^{-2y(x+z)}$,
 $\frac{\partial k}{\partial y} = 2xe^{-2y(x+z)} - 2(2xy + z)(x + z)e^{-2y(x+z)}$,
 $\frac{\partial k}{\partial z} = e^{-2y(x+z)} - 2(2xy^2 + yz)e^{-2y(x+z)}$

- Oppgave 3:** $\frac{\partial h}{\partial x_2} = 54x_1^5x_2^2x_3^2 \sin x_3$, $\frac{\partial h}{\partial x_3} = 36x_1^5x_2^3x_3 \sin x_3 + 18x_1^5x_2^3x_3^2 \cos x_3$

- Oppgave 4:** $\frac{\partial h}{\partial v} = e^{3uv^3}u \cos(uv) \left(9v^2 \cos(uv) - 2 \sin(uv) \right)$

- Oppgave 5:** $\mathbf{H}'(1, -2) = \begin{pmatrix} 13 & -7 \\ 10 & 0 \end{pmatrix}$

- Oppgave 6:** $\mathbf{H}'(-1, -2, 1) = \begin{pmatrix} 5 & -13 & 3 \\ 2 & 6 & -2 \end{pmatrix}$

Seksjon 2.8

- Oppgave 1:** $\begin{pmatrix} 2 & -1 & 1 \\ -1 & 1 & -3 \end{pmatrix}$

Oppgave 2: $\begin{pmatrix} -1 & 0 \\ 2 & -2 \\ -3 & 4 \\ 4 & 7 \end{pmatrix}$

Oppgave 3: $-\begin{pmatrix} 6 \\ 3 \end{pmatrix}$

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

Oppgave 5: $\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$

Oppgave 6: $\begin{pmatrix} 2 \cos \theta & -2 \sin \theta \\ 2 \sin \theta & 2 \cos \theta \end{pmatrix}$

Oppgave 7: $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$

Oppgave 8: $\begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix}$

Oppgave 10: $\begin{pmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{pmatrix}$

Oppgave 11: $x = -\frac{3}{7}, y = \frac{1}{7}, z = \frac{1}{7}, u = \frac{2}{7}$

b) $\mathbf{T}(\mathbf{e}_1) = \begin{pmatrix} -\frac{2}{7} \\ -\frac{4}{7} \end{pmatrix}, \mathbf{T}(\mathbf{e}_2) = \begin{pmatrix} \frac{3}{7} \\ -\frac{1}{7} \end{pmatrix}$

c) $\begin{pmatrix} -\frac{2}{7} & \frac{3}{7} \\ -\frac{4}{7} & -\frac{1}{7} \end{pmatrix}$

Oppgave 12: a) $x = \frac{1}{4}, y = \frac{1}{4}. \mathbf{T}^4(\mathbf{e}_1) = \begin{pmatrix} 353 \\ -136 \end{pmatrix}$

b) $u = -\frac{1}{2}, v = \frac{1}{2}. \mathbf{T}^4(\mathbf{e}_2) = \begin{pmatrix} -544 \\ 353 \end{pmatrix}$

c) $\begin{pmatrix} 353 & -544 \\ -136 & 353 \end{pmatrix}$

Oppgave 14: c) $x = 1, y = 2. A^{10}\mathbf{a} = \begin{pmatrix} 3^{10} + 2 \\ 3^{10} - 2 \end{pmatrix} = \begin{pmatrix} 59051 \\ 59047 \end{pmatrix}$

Seksjon 2.9

Oppgave 1: $A = \begin{pmatrix} 2 & -3 & 1 \\ -1 & 0 & 1 \end{pmatrix}, \mathbf{c} = -\begin{pmatrix} 7 \\ 2 \end{pmatrix}$

Oppgave 2: $\mathbf{r}(t) = \begin{pmatrix} 5t + 11 \\ -4t - 10 \end{pmatrix}$

Oppgave 3: $T_{\mathbf{a}}\mathbf{F} = \begin{pmatrix} -4x + 4y - 10 \\ 2x - 2y + 2 \end{pmatrix}$

Oppgave 4: $T_{\mathbf{a}}\mathbf{F} = \begin{pmatrix} 4y \\ x + 2y \\ 24x + y - 32 \end{pmatrix}$

Oppgave 5: Matrise $\begin{pmatrix} 1 & -2 \\ 4 & 1 \end{pmatrix}$, konstantledd $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$

Oppgave 6: Matrise $\begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix}$, konstantledd $\begin{pmatrix} 3 \\ -1 \end{pmatrix}$

Oppgave 7: a) Matrise $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$, konstantledd $\begin{pmatrix} 6 \\ 0 \end{pmatrix}$

b) Matrise $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, konstantledd $\begin{pmatrix} 0 \\ -4 \end{pmatrix}$

Oppgave 8: a) Matrise $\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$, konstantledd $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$