Vektorrom og lineæravbildninger

Harald Hanche-Olsen[1]

Sammendrag. Dette notatet[2] er tenkt som et supplement til Gilbert Strangs bok *Linear algebra and its applications*. Hensikten er å utdype teorien for abstrakte vektorrom, lineæravbildninger og sammenhengen med matriser og basisskifte, som er noe stemoderlig behandlet i boken.

Finner du trykkfeil eller andre feil og uklarheter, vil jeg gjerne høre om det — helst via epost til harald.hanche-olsen@ntnu.no.

Kropper og vektorrom

Til nå har vi sett konkrete eksempler på *reelle* og *komplekse* vektorrom: Underrom av $\mathbb{R}^n$ og $\mathbb{C}^n$. Når vi arbeider med disse vektorrommene er alt basert på *lineærkombinasjoner*: Vi adderer vektorer og multipliserer dem med skalarer (reelle tall om vi arbeider med $\mathbb{R}^n$, komplekse tall om vi arbeider med $\mathbb{C}^n$). Veldig mye av formalismen viser seg å være uavhengig av om vi arbeider med reelle eller komplekse tall. Vi kan også gjennomføre de fleste argumenter uten å arbeide med koordinatene til en vektor. Dette betyr at vi godt kan abstrahere bort mye av det spesifikke ved $\mathbb{R}^n$ og $\mathbb{C}^n$ (det vil si koordinatene!) og fokusere på det vi kan gjøre ved hjelp av lineærkombinasjoner. En aksiomatisk fremgangsmåte lar oss altså fokusere på det essensielle og se bort fra uvesentligheter.

1 Eksempler. Den andre fordelen ved en aksiomatisk metode er at man får mange flere eksempler å arbeide med — her nevner vi bare to eksempler på mer abstrakte vektorrom:

- Mengden av alle tre ganger deriverbare funksjoner av en variabel danner et vektorrom (man kan multiplisere en funksjon $y(x)$ punktvis med en skalar a, det vil si $(ay)(x) = a \cdot y(x)$, og man adderer funksjoner punktvis).

[1]Takk til Idar Hansen og Finn Faye Knudsen for konstruktiv kritikk av manuskriptet.

[2]Notatet stammer fra 1998, men er typesatt på ny fordi originalen bare var tilgjengelig i relativt ukurante formater (dvi og PostScript).

- Ta en lineær homogen differensialligning, for eksempel $y'''(x) + \sin(x)y''(x) - 4\cos(x)y'(x) + x^4y(x) = 0$. Mengden av alle løsninger til denne ligningen er et vektorrom (faktisk et underrom av rommet beskrevet over).

Kropper. Mengden av mulige skalarer har hittil vært enten $\mathbb{R}$ eller $\mathbb{C}$ — og slik skal det fortsatt være, i dette kurset. Men det er nyttig å være klar over at det finnes mange flere eksempler. Alt som kreves av mengden av skalarer er at den danner en *kropp* (på engelsk: field).

En kropp er en mengde $\mathbb{K}$ der det er definert to operasjoner, som vi kaller addisjon og multiplikasjon. Disse operasjonene kombinerer to elementer i $\mathbb{K}$ og produserer et nytt element i $\mathbb{K}$, akkurat slik vi er vant til med reelle eller komplekse tall. Vi skriver også operasjonene på vanlig måte som $a + b$ for summen (resultatet av addisjon) og ab eller $a \cdot b$ for produktet (dvs resultatet av multiplikasjon) av a og b. Disse må oppfylle regnereglene

- $a + 0 = a$
- For hver a finnes b slik at $a + b = 0$
- $a + b = b + a$
- $(a + b) + c = a + (b + c)$
- $a \cdot 1 = a$
- For hver $a \neq 0$ finnes b slik at $ab = 1$
- $ab = ba$
- $(ab)c = a(bc)$
- $a(b + c) = ab + ac$ [3]

Her er a, b og c vilkårlige elementer i $\mathbb{K}$, og 0 og 1 er faste elementer i $\mathbb{K}$. Vi må i tillegg kreve $1 \neq 0$, for ellers ville $\mathbb{K} = \{0\}$ være et eksempel på en kropp, men dette vil vi gjerne unngå.

2 Oppgave. Vis at det, gitt a, kun finnes én b slik at $a + b = 0$. Denne skalaren b kaller vi $-a$.

Vis at det, gitt $a \neq 0$, kun finnes én b slik at $ab = 1$. Denne skalaren b kaller vi a^{-1}. Vi skriver også c/a for $c \cdot a^{-1}$.

[3]Her følger vi den vanlige konvensjonen om at multiplikasjon har høyere presedens enn addisjon, slik at $ab + ac$ betyr $(ab) + (ac)$.

3 Eksempler. Foruten de opplagte eksemplene $\mathbb{R}$, $\mathbb{C}$ og mengden $\mathbb{Q}$ av rasjonale tall kan vi nevne noen mer eksotiske kropper:

- $\mathbb{K} = \{0, 1\}$ med regnereglene $0 + 0 = 1 + 1 = 0$, $0 + 1 = 1 + 0 = 1$, $0 \cdot 0 = 0 \cdot 1 = 1 \cdot 0 = 0$, $1 \cdot 1 = 1$. Denne kroppen kalles også $\mathbb{Z}_2$ eller $\mathbb{GF}_2$ og finner mange anvendelser, blant annet innenfor kodeteori og kryptografi (vektorrom over denne kroppen likeså).
- Mer generelt, om p er et primtall kan $\mathbb{Z}_p = \mathbb{GF}_p$ beskrives som tallene $\{0, 1, \ldots, p-1\}$ hvor regnereglene er addisjon og multiplikasjon modulo p. Disse er også i bruk i kryptografisk sammenheng, der p er et stort primtall (typisk flere hundre siffer).
- Mengden av alle rasjonale funksjoner ($f(x)/g(x)$ der f og g er polynomer) er også en kropp.

Eksemplene er mange fler; teorien for kropper er sentral i Galois-teorien, som handler om løsning av polynomligninger. Teoremet til Nils Henrik Abel om umuligheten av å løse den generelle 5.-gradsligningen ved hjelp av røtter bevises enklest innenfor denne teorien. Men nå er vi langt ute på viddene i forhold til kurset i matrisemetoder!

Vektorrom. Med et *vektorrom* over en kropp $\mathbb{K}$ forstås en ikketom mengde V, som blant annet inneholder et spesielt element O som vi vil kalle nullvektoren, og der det er definert to operasjoner: En addisjon som kombinerer to vektorer (om $x \in V$ og $y \in V$ er $x + y \in V$) og en multiplikasjon som kombinerer en skalar og en vektor (om $a \in \mathbb{K}$ og $x \in V$ er $ax \in V$) slik at følgende aksiomer er oppfylt (der $a, b \in \mathbb{K}$ og $x, y, z \in V$):

- $x + O = x$
- $x + y = y + x$
- $(x + y) + z = x + (y + z)$
- $0x = O$
- $1x = x$
- $(ab)x = a(bx)$
- $(a + b)x = ax + bx$
- $a(x + y) = ax + ay$

4 Oppgave. Vis at $x + (-1)x = O$. (Vi skriver kortere $-x$ for $(-1)x$). Vis også at dersom $a \in \mathbb{K}$, $x \in V$ og $ax = O$ er enten $a = 0$ eller $x = O$.

Ovenfor skrev vi O for nullvektoren i V for ikke å forveksle den med skalaren 0. Fra nå av skriver vi 0 også for nullvektoren, og håper det er klart fra sammenhengen hva vi mener. Egentlig burde vi også brukt et annet symbol enn + for addisjon av vektorer, for ikke å forveksle med addisjon av skalarer. Tilsvarende skulle vi hatt en annen notasjon for multiplikasjon av skalar med vektor. I praksis blir dette for tungvint, og man regner med at sammenhengen gjør det klart hvilken operasjon som brukes.

Et *underrom* av et vektorrom V er en delmengde $U \neq \emptyset$ slik at om $x, y \in U$ og $c \in \mathbb{K}$ er $x + y \in U$ og $cx \in U$.

5 Oppgave. Vis at et underrom av et vektorrom selv er et vektorrom.

Basis og dimensjon. Lineær uavhengighet, underrom utspent av en samling vektorer, basis og dimensjon kan defineres for generelle vektorrom akkurat som for $\mathbb{R}^n$ og underrom av $\mathbb{R}^n$. Vi skal ikke gjenta disse definisjonen her, siden de står hos Strang. Legg bare merke til at disse definisjonene bare bruker vektorromsoperasjonene addisjon og skalar multiplikasjon — vektorkomponenter (koordinater) kommer ikke noe sted inn.

6 Teorem. *Ethvert vektorrom har enten en endelig basis eller en uendelig følge av lineært uavhengige vektorer.*

Vektorrom med en endelig basis kalles *endeligdimensjonale* mens de andre kalles *uendeligdimensjonale.*

Bevis: La V være et vektorrom. Hvis $V = \{0\}$ har V en endelig basis (den tomme basisen, som består av null vektorer). I motsatt fall, la $v_1 \neq 0$ være en vektor i V.

Vi fortsetter ved induksjon. Anta $v_1, \ldots, v_n$ er lineært uavhengige vektorer i V. Vi har to muligheter. Den første er at disse vektorene utspenner hele V, men da er de en basis for V, som altså har en endelig basis. Den andre muligheten er at de ikke utspenner hele V. Da lar vi

v_{n+1} være en vektor som ikke kan skrives som lineærkombinasjon av $v_1, \dots, v_n$. Vi må vise at $v_1, \dots, v_{n+1}$ er lineært uavhengige. Anta derfor

$$\sum_{k=1}^{n+1} c_k v_k = 0.$$

Nå er $c_{n+1} = 0$, for ellers kunne vi skrive v_{n+1} som en lineærkombinasjon av $v_1, \dots, v_n$:

$$v_{n+1} = -\sum_{k=1}^{n} \frac{c_k}{c_{n+1}} v_k.$$

Siden $c_{n+1} = 0$ har vi dermed

$$\sum_{k=1}^{n} c_k v_k = 0,$$

og siden $v_1, \dots, v_n$ er lineært uavhengige (induksjonshypotesen!) er $c_1 = \cdots = c_n = 0$. Dette viser at $v_1, \dots, v_{n+1}$ er lineært uavhengige.

Slik fortsetter vi å plukke lineært uavhengige vektorer i V. Enten kan vi fortsette med dette i det uendelige, eller så stopper prosessen og vi har funnet en basis for V. □

7 Teorem. (Plassmangelteoremet, konkret versjon) *Om n vektorer i $\mathbb{K}^m$ er lineært uavhengige, er $n \le m$.*

Bevis: La de n vektorene være søylene i en matrise A. Da har ligningen $Ax = 0$ bare den trivielle løsningen $x = 0$. Men A har en trappedekomposisjon $PA = LU$, og $Ax = 0$ er ekvivalent med $Ux = 0$. U har høyst m pivot-elementer, så hvis $n > m$ vil ikke alle søyler inneholde et slikt — det vil si at det finnes fri variable. Man kan velge de fri variable så de ikke alle er 0, og løse med hensyn på basisvariablene ved tilbakesubstitusjon, slik at man ender med en ikketriviell løsning av $Ax = 0$. Dette er en motsigelse, og antagelsen $n > m$ må derfor være feil. □

8 Teorem. (Plassmangelteoremet, abstrakt versjon) *Om $v_1, \dots, v_m$ spenner ut et vektorrom V og $w_1, \dots, w_n \in V$ er lineært uavhengige, er $n \le m$.*

Bevis: Siden v–ene utspenner V kan hver w skrives som lineærkombinasjon av v–ene:

$$w_j = \sum_{i=1}^{m} a_{ij} v_i \qquad (j = 1, \dots, n).$$

Vi betrakter en lineærkombinasjon av w–ene:

$$\sum_{j=1}^{n} c_j w_j = \sum_{j=1}^{n} c_j \sum_{i=1}^{m} a_{ij} v_i = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} c_j \right) v_i,$$

men siden $w_1, \dots, w_n$ er lineært uavhengige gjelder implikasjonen

$$\sum_{j=1}^{n} a_{ij} c_j = 0 \text{ for } i = 1, \dots, m \Longrightarrow c_1 = \cdots = c_n = 0.$$

Matrisen (a_{ij}) har altså lineært uavhengige søyler, men disse søylene er n i tallet og ligger i $\mathbb{K}^m$, så $n \le m$ ved det konkrete plassmangelteoremet. □

Ved å anvende dette teoremet i begge retninger på to basiser får vi:

9 Korollar. *To basiser for et endeligdimensjonalt vektorrom har det samme antall elementer.*

Som du sikkert allerede vet, er det antall elementer i en basis som definerer *dimensjonen* til et vektorrom, og korollaret ovenfor sier altså at dimensjonen er veldefinert.

Det kan bemerkes at ovenstående bevis for plassmangelteoremet er nokså uvanlig (det er forresten også *navnet*). Det vanlige beviset er ved induksjon på m. Med litt strev kan du kanskje klare å snekre det sammen selv.

Lineæravbildninger, basiser og koordinater

En *lineæravbildning* fra et vektorrom X til et vektorrom Y er en funksjon $L\colon X \to Y$ som oppfyller kravene[4]

- $L(\xi_1 + \xi_2) = L(\xi_1) + L(\xi_2) \qquad (\xi_1, \xi_2 \in X)$,
- $L(c\xi) = cL(\xi) \qquad (c \in \mathbb{K}, \xi \in X)$.

10 Eksempel. La Y være vektorrommet av alle løsninger y av differensialligningen i eksempel 1 (side 1). Vi kan definere en lineæravbildning $L_x\colon Y \to \mathbb{R}^3$ ved $L_x(y) = [y(x), y'(x), y''(x)]^{\mathrm{T}}$.

11 Oppgave. La $L\colon X \to Y$ være en lineæravbildning. Vis at $\mathscr{N}(L) = \{x \in X \colon L(x) = 0\}$ er et underrom av X (*nullrommet* til L). Vis også at $\mathscr{R}(L) = \{L(x)\colon x \in X\}$ er et underrom av Y (*rekkevidden* til L). Vis at L har en invers avbildning $L^{-1}\colon Y \to X$ hvis og bare hvis $\mathscr{N}(L) = \{0\}$ og $\mathscr{R}(L) = Y$, og at L^{-1} i så fall er lineær.

12 Proposisjon. *En $m \times n$–matrise A med koeffisienter i $\mathbb{K}$ kan brukes til å definere en lineæravbildning $L\colon \mathbb{K}^n \to \mathbb{K}^m$ ved $L(x) = Ax$. Omvendt vil enhver lineæravbildning $L\colon \mathbb{K}^n \to \mathbb{K}^m$ være på denne formen for en entydig gitt matrise A.*

Med andre ord, teorien for matriser er den samme som teorien for lineæravbildninger mellom rom av typen $\mathbb{K}^n$.

Bevis: Den første delen — at multiplikasjon med en matrise er lineær — er enkel nok, og vi overlater beviset til leseren. Vi ser på omvendingen.

Vi kan få innsikt ved å betrakte den naturlige basisen i $\mathbb{K}^n$: Den består av vektorene $e_1, \ldots, e_n$ der e_j er null på alle plasser med unntak av en ener i posisjon j. Dersom $L(x) = Ax$ får vi spesielt $L(e_j) = Ae_j$, men Ae_j er søyle j i A. Dette viser at alle søylene i A, og dermed A selv, er entydig gitt av L. Det viser også hvordan vi må konstruere A

[4] Jeg vil i hovedsak bruke greske bokstaver for elementer i et generelt vektorrom, mens jeg bruker latinske bokstaver for vektorer i rom av formen $\mathbb{K}^n$. Siden $\mathbb{K}^n$ selv er et vektorrom, kan jeg selvsagt ikke følge denne regelen helt konsistent.

dersom vi bare er gitt L: Søyle j i A må være $L(e_j)$. Formelt: $A = [L(e_1), \ldots, L(e_n)]$.

Med denne definisjonen av A må vi vise $L(x) = Ax$. Men en generell vektor $x \in \mathbb{K}^n$ kan skrives $x = \sum_{j=1}^n x_j e_j$, og dermed blir

$$L(x) = L\left(\sum_{j=1}^n x_j e_j\right) = \sum_{j=1}^n x_j L(e_j)$$

men dette siste er akkurat det samme uttrykket vi får for Ax når vi vet at $L(e_j)$ er j'te søyle i A. □

Dersom vi har gitt en endelig følge $\alpha = (\alpha_1, \ldots, \alpha_n)$ av vektorer i et vektorrom X kan vi på en naturlig måte danne en lineæravbildning $\mathscr{L}_\alpha\colon \mathbb{K}^n \to X$ gitt ved

$$\mathscr{L}_\alpha(c) = \sum_{j=1}^n c_j \alpha_j.$$

Den virker altså simpelthen ved å bruke elementene i c som koeffisienter i en lineærkombinasjon av basiselementene.

13 Oppgave. Vis at $\mathscr{L}_\alpha$ har en invers avbildning $X \to \mathbb{K}^n$ dersom vektorene i α danner en basis for X.

Vi skriver $\mathscr{M}_\alpha$ for den inverse avbildningen til $\mathscr{L}_\alpha$, og kaller den *koordinatavbildningen* til basisen α fordi $x = \mathscr{M}_\alpha(\xi)$ er koordinatene til ξ i basisen: Med andre ord, x_j er de entydig gitte tallene som gir oss $\xi = \sum_{j=1}^n x_j \alpha_j$.

14 Proposisjon. *Anta X er gitt en basis $\alpha_1, \ldots, \alpha_n$, og vi er gitt vektorer $\beta_1, \ldots, \beta_n \in Y$. Da finnes en, og kun en, lineæravbildning $L\colon X \to Y$ slik at $L(\alpha_j) = \beta_j$ for $j = 1, \ldots, n$.*

Med andre ord, vi kan foreskrive verdiene av en lineæravbildning på elementene i en basis etter forgodtbefinnende, og valget bestemmer lineæravbildningen entydig. $\beta_1, \ldots, \beta_n \in Y$ trenger ikke utspenne Y eller være lineært uavhengige.

Bevis: Lineariteten viser at vi må ha

$$L\left(\sum_{j=1}^{n} x_j\alpha_j\right) = \sum_{j=1}^{n} x_j\beta_j,$$

som viser entydigheten, fordi enhver vektor i X kan skrives på formen $\sum_{j=1}^{n} x_j\alpha_j$. Men formelen kan og brukes til å *definere* L, slik at eksistensen er bevist. Mer elegant kan vi skrive $L(\xi) = \mathscr{L}_\beta(\mathscr{M}_\alpha(\xi))$, noe som umiddelbart gir lineariteten av L siden den er en sammensetning av lineære avbildninger. $\square$

15 Oppgave. Fyll inn detaljene i beviset ovenfor.

I følge proposisjon 12 er det en 1–1 korrespondanse mellom matriser og lineæravbildninger mellom rom av typen $\mathbb{K}^n$. Dette resultatet lar seg generalisere til generelle endeligdimensjonale vektorrom, men det krever at vi *velger* en basis i hvert av de involverte rommene. Det presise resultatet er som følger:

16 Teorem. *Anta $L\colon X \to Y$ er en lineæravbildning, og at vektorrommene X og Y er utstyrt med hver sin basis: henholdsvis $\alpha = (\alpha_1,\ldots,\alpha_n)$ og $\beta = (\beta_1,\ldots,\beta_m)$. Da finnes en entydig matrise $\mathscr{M}_{\beta\leftarrow\alpha}(L)$ slik at*

$$\mathscr{M}_\beta(L(\xi)) = \mathscr{M}_{\beta\leftarrow\alpha}(L)\cdot\mathscr{M}_\alpha(\xi) \qquad (\xi \in X).$$

Dette gir en en-til-en korrespondanse mellom mengden av lineæravbildninger $X \to Y$ og alle $m \times n$–matriser.

Bevis: Beviset lar seg lettest forklare ved diagrammet

$$\begin{tikzcd} X \arrow[r, "L"] \arrow[d, "\mathscr{M}_\alpha"'] & Y \arrow[d, "\mathscr{M}_\beta"] \\ \mathbb{K}^n \arrow[r, "M"'] & \mathbb{K}^m \end{tikzcd}$$

der pilene står for lineæravbildninger. Øverst har vi den gitte lineæravbildningen $L\colon X \to Y$, og de to vertikale pilene står for koordinatavbildningene gitt ved de to basisene. Siden $\mathscr{M}_\alpha$ er invertibel kan vi

fylle inn den nederste lineæravbildningen $M\colon \mathbb{K}^n \to \mathbb{K}^m$ slik at diagrammet *kommuterer*, det vil si at om vi følger pilene to veier gjennom diagrammet fra ett sted til et annet er den tilhørende sammensetningen av avbildninger den samme begge veier. I vårt tilfelle betyr dette

$$(*) \qquad \mathscr{M}_\beta(L(\xi)) = M(\mathscr{M}_\alpha(\xi)) \qquad (\xi \in X)$$

og alt vi trenger gjøre for å definere M slik at dette holder er å starte i nedre venstre hjørne i diagrammet, gå mot pilen opp til X, deretter mot høyre til Y og så ned til $\mathbb{K}^m$ — slik (les uttrykket på høyresiden nedenfor fra høyre mot venstre for å se sammenhengen med forklaringen foran):

$$M(x) = (\mathscr{M}_\beta(L(\mathscr{L}_\alpha(x)))) \qquad (x \in \mathbb{K}^n).$$

Nå gjenstår det bare å henvise til proposisjon 12, som fastslår at det finnes en entydig matrise $\mathscr{M}_{\beta\leftarrow\alpha}$ som implementerer lineæravbildningen M, det vil si $M(x) = \mathscr{M}_{\beta\leftarrow\alpha}x$ for $x \in \mathbb{K}^n$. Men setter vi dette uttrykket for M inn i $(*)$ får vi akkurat den formelen vi skulle bevise. $\square$

Matrisen $\mathscr{M}_{\beta\leftarrow\alpha}(L)$ er i prinsipp enkel å beregne. For vi skal jo ha

$$\mathscr{M}_\beta(L(\alpha_j)) = \mathscr{M}_{\beta\leftarrow\alpha}(L)\mathscr{M}_\alpha(\alpha_j) = \mathscr{M}_{\beta\leftarrow\alpha}(L)e_j$$

som er j'te søyle i $\mathscr{M}_{\beta\leftarrow\alpha}(L)$. Om vi kjenner $L(\alpha_j)$ er det altså bare å regne ut koordinatene til denne vektoren i basisen β og sette svaret inn som j'te søyle i $\mathscr{M}_{\beta\leftarrow\alpha}(L)$.

17 Eksempel. Vi lar $\lambda > 0$ være et fast tall og lar X være rommet av løsninger til $\ddot{x}(t) + \lambda x(t) = 0$. Dette rommet har en basis bestående av de to funksjonene $\alpha_1(t) = \cos\sqrt{\lambda}t$ og $\alpha_2(t) = \sin\sqrt{\lambda}t$.

Definer $L\colon X \to \mathbb{R}^2$ ved $L(x) = [x(0), x(\pi)]^{\mathrm{T}}$. Vi bruker den naturlige basisen e i $\mathbb{R}^2$, og merker oss at $\mathscr{M}_e(u) = u$ når $u \in \mathbb{R}^2$ (u er sin egen koordinatvektor — det er derfor vi kaller basisen naturlig).

Dermed er utregningen av $\mathscr{M}_{e\leftarrow\alpha}(L)$ ganske likefrem:

$$\mathscr{M}_e(L(\alpha_1)) = L(\alpha_1) = \begin{bmatrix} 1 \\ \cos\sqrt{\lambda}\pi \end{bmatrix},$$

$$\mathscr{M}_e(L(\alpha_2)) = L(\alpha_2) = \begin{bmatrix} 0 \\ \sin\sqrt{\lambda}\pi \end{bmatrix};$$

$$\mathscr{M}_{e\leftarrow\alpha}(L) = \begin{bmatrix} 1 & 0 \\ \cos\sqrt{\lambda}\pi & \sin\sqrt{\lambda}\pi \end{bmatrix}.$$

18 Oppgave. La P_n være rommet av polynomer av grad høyst n i en variabel x. Finn matrisen til derivasjonsoperatoren $D\colon P_n \to P_{n-1}$ der vi bruker basisen $1, x, x^2, \ldots, x^n$ for P_n (og tilsvarende for P_{n-1}).

Om $L\colon X \to Y$ og $M\colon Y \to Z$ er to lineæravbildninger skriver vi $M \circ L$ for *sammensetningen* $X \to Y$, altså

$$M \circ L(\xi) = M(L(\xi))$$

Sammensetning av lineæravbildninger svarer til multiplikasjon av matriser:

19 Teorem. *Anta $L\colon X \to Y$ og $M\colon Y \to Z$ er to lineæravbildninger. Anta X er utstyrt med en basis α, Y er utstyrt med basisen β og Z har basisen γ. Da gjelder*[5]

$$\mathscr{M}_{\gamma\leftarrow\alpha}(M \circ L) = \mathscr{M}_{\gamma\leftarrow\beta}(M)\mathscr{M}_{\beta\leftarrow\alpha}(L).$$

20 Oppgave. Bevis teorem 19 ved å betrakte diagrammet

$$\begin{tikzcd} X \arrow[r, "L"] \arrow[d, "\mathscr{M}_\alpha"'] & Y \arrow[r, "M"] \arrow[d, "\mathscr{M}_\beta"] & Z \arrow[d, "\mathscr{M}_\gamma"] \\ \mathbb{K}^n \arrow[r, "\mathscr{M}_{\beta\leftarrow\alpha}(L)"'] & \mathbb{K}^m \arrow[r, "\mathscr{M}_{\gamma\leftarrow\beta}(M)"'] & \mathbb{K}^l \end{tikzcd}$$

[5]Det er på grunn av denne formelen jeg har valgt notasjonen slik jeg gjorde: Legg merke til hvordan pilene naturlig føyer seg sammen — fra α via β til γ.

Skifte av basis. Et nyttig, men kanskje uventet, spesialtilfelle av teorem 16 er tilfellet $X = Y$ og $L = I$ (identitetsavbildningen): Vi tenker oss altså at ett vektorrom X er utstyrt med to basiser α og β. Teorem 16 sier da at det finnes en matrise $\mathscr{M}_{\beta\leftarrow\alpha}(I)$ med egenskapen

$$\mathscr{M}_\beta(\xi) = \mathscr{M}_{\beta\leftarrow\alpha}(I) \cdot \mathscr{M}_\alpha(\xi)$$

Nå trengte vi knapt den fulle kraften av teoremet for å innse det: $\mathscr{M}_{\beta\leftarrow\alpha}(I)$ er bare matrisen som implementerer lineæravbildningen $\mathscr{M}_\beta \circ \mathscr{L}_\alpha\colon \mathbb{K}^n \to \mathbb{K}^n$.

Siden multiplikasjon med $\mathscr{M}_{\beta\leftarrow\alpha}(I)$ konverterer koordinatene til en vektor ξ fra basisen α til β kalles matrisen en *basisskiftematrise*.

I mange fysikk- og ingeniøranvendelser er det vanlig å snakke om en basisskiftematrise, eller avbildningen den implementerer, som en *passiv transformasjon* (det er typisk en dreining av koordinatsystemet), mens en avbildning $X \to X$ kalles en *aktiv transformasjon* (den kan for eksempel representere en fysisk forflytning av et stivt legeme).

Basisskiftematriser er mye brukt. Det er bare å tenke på alle situasjoner der mer enn ett koordinatsystem er naturlig: Om man for eksempel modellerer en robotarm er det naturlig at hver bevegelig del av armen er utstyrt med sitt eget koordinatsystem. I tillegg har man et koordinatsystem som ligger fast i omgivelsene. Tilsvarende betraktning gjelder ved styring av en satelitt eller romferge, som har sitt eget koordinatsystem som kan roteres i forhold til et koordinatsystem som ligger fast i rommet.

Basisskiftematrisen kan brukes ikke bare til å konvertere koordinater for vektorer, men også for å oversette matrisen til en lineæravbildning fra ett koordinatsystem til et annet. To gangers bruk av teorem 19 gir nemlig resultatet

$$\mathscr{M}_{\beta\leftarrow\beta}(L) = \mathscr{M}_{\beta\leftarrow\alpha}(I)\mathscr{M}_{\alpha\leftarrow\alpha}(L)\mathscr{M}_{\alpha\leftarrow\beta}(I)$$

der vi forutsetter at $L\colon X \to X$ er lineær og X har de to basisene α og β.

21 Oppgave. Tegn et diagram som forklarer resultatet ovenfor. Vis også at

$$\mathscr{M}_{\alpha\leftarrow\beta}(I) = \mathscr{M}_{\beta\leftarrow\alpha}(I)^{-1}$$

og konkluder at $\mathscr{M}_{\beta\leftarrow\beta}(L)$ og $\mathscr{M}_{\alpha\leftarrow\alpha}(L)$ er similære.

Resultatet er viktig, blant annet fordi det tillater oss å definere *determinanten* til en lineæravbildning av et endeligdimensjonalt vektorrom inn i seg selv: Bare ta determinanten av matrisen til avbildningen i en eller annen basis. Siden similære matriser har samme determinant, er valget av basis irrelevant for den resulterende determinanten. Det samme gjelder andre størrelser som similære matriser har felles, som trasen og alle koeffisientene i det karakteristiske polynomet.

22 Oppgave. La X være et endeligdimensjonalt vektorrom. Vis at en lineæravbildning $L\colon X \to X$ er invertibel hvis og bare hvis $\det L \neq 0$. (Hint: Bruk det tilsvarende resultatet for matriser.)

23 Eksempel. (For de som kan litt kvantemekanikk). Den stasjonære Schrödingerligningen for en partikkel med energi E i et potensial V i én romdimensjon kan skrives

$$(*) \qquad -\psi''(x) + V(x)\psi(x) = E\psi(X)$$

(i dimensjonsløse variable). Anta at $V(x) = 0$ for x utenfor et begrenset intervall $[a,b]$. Utenfor dette intervallet kan enhver løsning av ligningen skrives som en lineærkombinasjon av de to løsningene $e^{\pm ikx}$ der $k = \sqrt{E}$. Mer presist kan vi produsere løsninger α_1, α_2, β_1 og β_2 gyldige for alle x slik at

$$\begin{aligned} \alpha_1(x) &= e^{ikx}, \quad \alpha_2(x) = e^{-ikx} \qquad (x < a) \\ \beta_1(x) &= e^{ikx}, \quad \beta_2(x) = e^{-ikx} \qquad (x > b) \end{aligned}$$

Teorien for ordinære differensialligninger garanterer at α er en basis for rommet av løsninger for $(*)$, mens β er en annen slik basis. Basisskiftematrisene $S = \mathscr{M}_{\alpha\leftarrow\beta}(I)$ og $S^{-1} = \mathscr{M}_{\beta\leftarrow\alpha}(I)$ kalles gjerne *spredningsmatrisene* for potensialet V. Løsningen $\beta_1 = S_{11}\alpha_1 + S_{21}\alpha_2$ tolkes som resultatet av et eksperiment der en partikkel skytes inn fra venstre (intensitet $|S_{11}|^2$) og delvis reflekteres (intensitet $|S_{21}|^2$), delvis slippes gjennom (intensitet 1). S er et eksempel på en basisskiftematrise som kan være vanskelig å regne ut, fordi beregningen krever løsning av en differensialligning med variable koeffisienter.

Egenverdier. Anta $L\colon X \to X$ er en lineæravbildning. Akkurat som for matriser definerer vi egenverdier og egenvektorer ved ligningen

$$L(\xi) = \lambda\xi.$$

λ kalles en *egenverdi* med tilhørende *egenvektor* ξ dersom ligningen holder med $\xi \neq 0$. Akkurat som for matriser ser vi at, dersom X er endeligdimensjonal, er λ en egenverdi for L hvis og bare hvis $\det(L - \lambda I) \neq 0$.[6] Dersom X har en basis $\alpha = (\alpha_1, \dots, \alpha_n)$ av egenvektorer for L, ser vi at

$$\mathscr{M}_{\alpha\leftarrow\alpha}(L) = \Lambda$$

der Λ er diagonalmatrisen med $\lambda_1, \dots, \lambda_n$ på diagonalen (for $L(\alpha_j) = \lambda_j\alpha_j$ har koordinatvektor $\lambda_j e_j$ i basisen α, og $[\lambda_1 e_1, \dots, \lambda_n e_n] = \Lambda$.) Man finner

$$L(\mathscr{L}_\alpha(x)) = \mathscr{L}_\alpha(\Lambda x) \qquad (x \in \mathbb{K}^n).$$

24 Oppgave. Dersom ikke alt var klart i avsnittet ovenfor, så fyll inn detaljene.

25 Oppgave. Anta du er gitt en $n \times n$–matrise A. Den definerer en lineæravbildning $L\colon \mathbb{K}^n \to \mathbb{K}^n$ ved $L(x) = Ax$. Dersom $\alpha = (\alpha_1, \dots, \alpha_n)$ er en basis av egenvektorer for A (og dermed for L), kan du danne en matrise $S = [\alpha_1, \dots, \alpha_n]$ med disse egenvektorene som søyler. I følge det ovenstående definerer S også en lineæravbildning $\mathscr{L}_\alpha\colon \mathbb{K}^n \to \mathbb{K}^n$. Vis at $\mathscr{L}_\alpha(x) = Sx$ for alle $x \in \mathbb{K}^n$, og konkluder at teorien foran gir $AS = S\Lambda$.

Indreprodukt

Et *indreprodukt* på et reelt vektorrom V er en operasjon som til to vektorer ξ og η tilordner en skalar (ξ,η)[7] og som oppfyller disse kravene:

- $(\xi, a\eta + b\zeta) = a\,(\xi,\eta) + b\,(\xi,\zeta)$ når $\xi,\eta,\zeta \in V$ og $a,b \in \mathbb{R}$,
- $(\xi,\eta) = (\eta,\xi)$ for alle $\xi,\eta \in V$,
- $(\xi,\xi) > 0$ dersom $\xi \neq 0$.

Det første kravet er altså at indreproduktet er lineært i annen faktor alene, mens det andre kravet er at det er symmetrisk (uendret når

[6] I er identitetsavbildningen: $I(\xi) = \xi$ for alle ξ.

[7] Andre mye brukte skrivemåter for et generelt indreprodukt er $\langle\xi|\eta\rangle$, $\langle\xi,\eta\rangle$ og $(\xi|\eta)$.

faktorene bytter plass). Det følger at det er lineært også i første faktor, og det kalles derfor *bilineært*.

26 Eksempler. Tre mye brukte eksempler på reelle indreproduktrom:

- $\mathbb{R}^n$ med indreproduktet $(x, y) = x^{\mathrm{T}}y$.
- Dersom W er en positivt definit $n \times n$–matrise, $\mathbb{R}^n$ med indreproduktet $(x, y) = x^{\mathrm{T}}Wy$.
- La w være en rell funksjon på et intervall J, og anta $w > 0$ i J. Mengden av alle reelle funksjoner f på J som oppfyller $\int_J f(x)^2 w(x)\,\mathrm{d}x < \infty$ kan utstyres med indreproduktet $(f, g) = \int_J f(x)g(x)w(x)\,\mathrm{d}x$.

Vi kan også definere indreprodukt på komplekse vektorrom. Men vi kan ikke kreve at indreproduktet skal være bilineært, for da blir jo $(i\xi, i\xi) = -(\xi, \xi)$ og det blir håpløst å kreve at $(\xi, \xi) > 0$ for alle vektorer $\xi \neq 0$. Istedet endrer vi på symmetrien, slik at kravene blir

- $(\xi, a\eta + b\zeta) = a\,(\xi, \eta) + b\,(\xi, \zeta)$ når $\xi, \eta, \zeta \in V$ og $a, b \in \mathbb{R}$,
- $(\xi, \eta) = \overline{(\eta, \xi)}$ for alle $\xi, \eta \in V$,
- $(\xi, \xi) > 0$ dersom $\xi \neq 0$.

Man kan legge merke til at indreproduktet nå blir lineært bare i annen faktor, mens det i den første faktor oppfyller $(a\xi, \eta) = \overline{a}\,(\xi, \eta)$ (*vis dette*). Vi sier gjerne at indreproduktet er *antilineært* i første faktor, og man kaller det komplekse indreproduktet *sesquilineært*.[8]

27 Eksempler. Tre mye brukte eksempler på komplekse indreproduktrom:

- $\mathbb{C}^n$ med indreproduktet $(x, y) = x^{\mathrm{H}}y$.

[8]Forstavelsen *sesqui* kommer fra latin for «en og en halv». Indreproduktet er liksom halvannen ganger lineært. For øvrig er det mer vanlig i matematisk literatur å la indreproduktet være lineært i første faktor og antilineært i andre faktor. Konvensjonen som er brukt her er mer vanlig blant fysikere, men har også vunnet en viss utbredelse blant anvendte matematikere.

- Dersom W er en positivt definit $n \times n$–matrise, $\mathbb{C}^n$ med indreproduktet $(x, y) = x^{\mathrm{H}}Wy$.
- La w være en reell funksjon på et intervall J, og anta $w > 0$ i J. Mengden av alle komplekse funksjoner f på J som oppfyller $\int_J |f(x)|^2\, w(x)\,\mathrm{d}x < \infty$ kan utstyres med indreproduktet $(f, g) = \int_J \overline{f(x)}g(x)w(x)\,\mathrm{d}x$.

Vi definerer *normen* eller *lengden* av en vektor ξ som $\|\xi\| = \sqrt{(\xi, \xi)}$, og kaller vektorer ξ og η *ortogonale* dersom $(\xi, \eta) = 0$.

28 Proposisjon. *Ethvert endeligdimensjonalt reelt eller komplekst vektorrom har en ortonormal basis.*

Bevis: Gram-Schmidt-ortogonalisering fungerer akkurat på samme måte som i $\mathbb{R}^n$ (eller $\mathbb{C}^n$), med det samme beviset. Du må bare skrive om $x^{\mathrm{T}}y$ til (x, y) gjennom hele beviset. □

En lineæravbildning $L\colon V \to V$ kalles *symmetrisk* (for reelle vektorrom V) eller *hermitisk* (for komplekse vektorrom) dersom den oppfyller $(L\xi, \eta) = (\xi, L\eta)$ for alle $\xi, \eta \in V$.

29 Oppgave. La A være en $n \times n$–matrise og la $L\colon \mathbb{R}^n \to \mathbb{R}^n$ (eller $L\colon \mathbb{C}^n \to \mathbb{C}^n$) være multiplikasjon med A. Vis at L er symmetrisk (hermitisk) hvis og bare hvis A er symmetrisk (eller hermitisk). (*Hint.* Hva er (Le_j, e_i)?)

Generaliser: Anta $L\colon V \to V$ er en symmetrisk avbildning der det reelle vektorrommet V er gitt et indreprodukt (eller at L er hermitisk på et komplekst vektorrom). La α være en ortonormal basis for V. Vis at L er symmetrisk (hermitisk) hvis og bare hvis $\mathscr{M}_{\alpha\leftarrow\alpha}(L)$ er symmetrisk (hermitisk).

Vis til slutt at man kan velge en ortonormal basis β slik at $\mathscr{M}_{\beta\leftarrow\beta}(L)$ er en diagonalmatrise. (*Hint.* Diagonalisér $\mathscr{M}_{\alpha\leftarrow\alpha}(L)$.)

Register