Lineær Algebra

Lars Hesselholt og Nathalie Wahl

Oktober 2016

Forord

Denne bog er beregnet til et første kursus i lineær algebra, men vi har lagt vægt på at fremstille dette materiale på en sådan måde, at det implicit fremstår i sin naturlige generalitet. Vi har også lagt vægt på ikke-kommutativitet, og vi benytter således ikke, at ab og ba er ens, medmindre det er nødvendigt.

Vores hovedkilde som inspiration har været den uovertrufne præsentation af lineær algebra i N. Bourbaki, Algebra I og Algebra II. Vores gennemgang af determinanten er også stærkt inspireret af H. A. Nielsen, Lineær Algebra, Aarhus Universitet fra 1988. Endvidere er bogens layout og visse figurer overført fra Niels Vigand Pedersen, Lineær Algebra, Københavns Universitet, 2. udgave fra 2009, revideret af Morten Risager.

København, oktober 2016, Lars Hesselholt og Nathalie Wahl

Indhold

0 Skalarer

Hele vejen igennem bogen vil vi arbejde med en grundlæggende mængde af "tal", som vi betegner $\mathbb{F}$. Vi siger, at elementer $a \in \mathbb{F}$ er *skalarer*. De to hovedeksempler, vi har i tankerne, er de reelle tal $\mathbb{F} = \mathbb{R}$ og de komplekse tal $\mathbb{F} = \mathbb{C}$. Vi specificerer nedenfor de aritmetiske strukturer "+" og " · " på mængden af skalarer, som vi skal gøre brug af.

0.1 Legemer

Vi tillader $\mathbb{F}$ at være et vilkårligt legeme, hvilket vil sige en mængde af skalarer udstyret med sum og produkt af disse skalarer, der opfylder de sædvanlige aritmetiske regler. Den præcise definition er som følger.

Definition 0.1.1 Et *legeme* er en triple $(\mathbb{F}, +, \cdot)$, der består af en mængde $\mathbb{F}$ samt to afbildninger $+ : \mathbb{F} \times \mathbb{F} \to \mathbb{F}$ og $\cdot : \mathbb{F} \times \mathbb{F} \to \mathbb{F}$, sådan at der gælder:

(A1) For alle $a, b, c \in \mathbb{F}$ er $(a + b) + c = a + (b + c)$.

(A2) Der findes et element $0 \in \mathbb{F}$, sådan at $a + 0 = a = 0 + a$ for alle $a \in \mathbb{F}$.

(A3) For alle $a \in \mathbb{F}$, findes der et $b \in \mathbb{F}$, sådan at $a + b = 0 = b + a$.

(A4) For alle $a, b \in \mathbb{F}$ er $a + b = b + a$.

(P1) For alle $a, b, c \in \mathbb{F}$ er $(a \cdot b) \cdot c = a \cdot (b \cdot c)$.

(P2) Der findes et element $1 \in \mathbb{F}$, sådan at $a \cdot 1 = a = 1 \cdot a$ for alle $a \in \mathbb{F}$.

(P3) For alle $0 \neq a \in \mathbb{F}$, findes $b \in \mathbb{F}$, sådan at $a \cdot b = 1 = b \cdot a$. Endvidere er $0 \neq 1$.

(P4) For alle $a, b \in \mathbb{F}$ er $a \cdot b = b \cdot a$.

(D1) For alle $a, b, c \in \mathbb{F}$ er $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$.

(D2) For alle $a, b, c \in \mathbb{F}$ er $(a + b) \cdot c = (a \cdot c) + (b \cdot c)$.

Vi forkorter normalt og skriver ab i stedet for $a \cdot b$. Vi vil også normalt misbruge notation og blot skrive $\mathbb{F}$ for legemet $(\mathbb{F}, +, \cdot)$.

Eksempel 0.1.2 Vi giver følgende mere eller mindre velkendte eksempler på legemer.

($\mathbb{R}$) Det bedst kendte eksempel på et legeme er nok legemet af reelle tal $(\mathbb{R},+,\cdot)$, hvor $\mathbb{R}$ er mængden af reelle tal, og hvor "+" og " $\cdot$ " er henholdsvis den sædvanlige sum og det sædvanlige produkt af reelle tal.

($\mathbb{Q}$) Et andet velkendt eksempel er legemet af rationelle tal $(\mathbb{Q},+,\cdot)$, hvor $\mathbb{Q} \subset \mathbb{R}$ er delmængden af rationale tal, og hvor igen "+" og " $\cdot$ " er de sædvanlige sum og produkt operationer.

($\mathbb{C}$) Et meget vigtigt eksempel er legemet af komplekse tal $(\mathbb{C},+,\cdot)$, som består af mængden af komplekse tal

$$\mathbb{C} = \{a+ib \mid a,b \in \mathbb{R}\}$$

med "+" og "$\cdot$" defineret ved henholdsvis

$$\begin{aligned}(a+ib)+(c+id) &= (a+c)+i(b+d)\\ (a+ib)\cdot(c+id) &= (ac-bd)+i(ad+bc).\end{aligned}$$

Det modsatte element af $a+ib$ er $-a+i(-b)$, som vi også skriver $-a-ib$, og, hvis $a,b \neq 0$, da er det multiplikativt inverse element af $a+ib$ givet ved

$$\frac{1}{a+ib} = \frac{a-ib}{(a+ib)(a-ib)} = \frac{a-ib}{a^2+b^2} = \frac{a}{a^2+b^2} - i\frac{b}{a^2+b^2}.$$

($\mathbb{F}_2$) Det mindste legeme er legemet $(\mathbb{F}_2,+,\cdot)$, hvor $\mathbb{F}_2 = \{0,1\}$, og hvor $1+1$ er defineret til at være lig med 0.

Der findes mange andre legemer, endelige som uendelige, og linear algebra fungerer over et vilkårligt legeme. I dette kursus vil vi dog kun bruge $\mathbb{R}$ og $\mathbb{C}$ som eksempler.

De grundlæggende egenskaber (A1)–(A4), (P1)–(P4) og (D1)–(D2), der per definition gælder for sum og produkt i et legeme, har flere konsekvenser, som er velkendte for de reelle og de komplekse tal, men som altid gælder så snart de grundlæggende betingelser er opfyldt. Vi nævner her nogle af de vigtigste:

(1) Elementet $0 \in \mathbb{F}$, der opfylder (A2), er entydigt bestemt. For hvis 0 og $0'$ begge opfylder (A2), da gælder
$$0 = 0 + 0' = 0'.$$
Vi kalder $0 \in \mathbb{F}$ for *nul-elementet* i $\mathbb{F}$.

(2) Givet $a \in \mathbb{F}$, da er elementet $b \in \mathbb{F}$, der opfylder (A3) ligeledes entydigt bestemt. For hvis både b og b' opfylder (A3), da er
$$b = b+0 = b+(a+b') = (b+a)+b' = 0+b' = b'.$$
Vi skriver $-a$ for dette element b og kalder det for det *modsatte element* af a.

(3) Man viser tilsvarende, at elementet $1 \in \mathbb{F}$, der opfylder (P2), er entydigt bestemt; vi kalder dette element for *et-elementet* i $\mathbb{F}$. Og givet $0 \neq a \in \mathbb{F}$, da er elementet $b \in \mathbb{F}$, der opfylder (P3), ligeledes entydigt bestemt; vi kalder dette element for det *multiplikativt inverse element* af a og betegner det a^{-1}.

(4) Hvis $a + c = b + c$, så er $a = b$, idet

$$\begin{aligned} a &\overset{(A2)}{=} a + 0 \overset{(A3)}{=} a + (c + (-c)) \overset{(A1)}{=} (a + c) + (-c) \\ &= (b + c) + (-c) \overset{(A1)}{=} b + (c + (-c)) \overset{(A3)}{=} b + 0 \overset{(A2)}{=} b. \end{aligned}$$

Det vil sige, at vi kan trække et vilkårligt $c \in \mathbb{F}$ fra på begge sider at et lighedstegn. Tilsvarende gælder der for alle $c \in \mathbb{F}$ med $c \neq 0$, at hvis $a \cdot c = b \cdot c$, så er $a = b$, fordi

$$\begin{aligned} a &\overset{(P2)}{=} a \cdot 1 \overset{(P3)}{=} a \cdot (c \cdot c^{-1}) \overset{(P1)}{=} (a \cdot c) \cdot c^{-1} \\ &= (b \cdot c) \cdot c^{-1} \overset{(P1)}{=} b \cdot (c \cdot c^{-1}) \overset{(P3)}{=} b \cdot 1 \overset{(P2)}{=} b. \end{aligned}$$

Det vil sige, at vi kan dividere med $c \in \mathbb{F}$ på begge sider af et lighedstegn, forudsat at c er forskellig fra 0. Vi kan derimod ikke dividere med 0, fordi 0 ikke har et multiplikativt inverst element.

Vi nævner derudover de følgende to konsekvencer af Definition 0.1.1:

Sætning 0.1.3 *Lad $\mathbb{F}$ være et legeme. For alle $a \in \mathbb{F}$ gælder følgende:*

(E1) $a \cdot 0 = 0 = 0 \cdot a$.

(E2) $a \cdot (-1) = -a = (-1) \cdot a$.

Bevis Vi beviser først (E1). Givet $a \in \mathbb{F}$, da er

$$a \cdot 0 \overset{(A2)}{=} a \cdot (0 + 0) \overset{(D1)}{=} (a \cdot 0) + (a \cdot 0),$$

og ved at trække $a \cdot 0$ fra venstre og højre side får vi da, at $0 = (a \cdot 0)$ som ønsket.

Vi viser dernæst (E2). Ifølge bemærkning (2) ovenfor, er $-a \in \mathbb{F}$ det entydigt bestemte element, der opfylder $a + (-a) = 0 = (-a) + a$. Men udregningen

$$\begin{aligned} a + (-1) \cdot a &\overset{(P2)}{=} 1 \cdot a + (-1) \cdot a \overset{(D2)}{=} (1 + (-1)) \cdot a \overset{(A3)}{=} 0 \cdot a \overset{(E1)}{=} 0 \\ (-1) \cdot a + a &\overset{(P2)}{=} (-1) \cdot a + 1 \cdot a \overset{(D1)}{=} a \cdot (1 + (-1)) \overset{(A3)}{=} a \cdot 0 \overset{(E1)}{=} 0 \end{aligned}$$

viser, at elementet $(-1) \cdot a$ også har denne egenskab, hvilket viser (E2). $\square$

Bemærkning 0.1.4 Hvis $(\mathbb{Z},+,\cdot)$ er triplen, hvor $\mathbb{Z}$ er mængden af hele tal, og hvor "+" og " · " er de sædvanlige sum og produkt operationer, så er er alle aksiomerne i Definition 0.1.1 med undtagelse af (P3) opfyldt. For de eneste elementer i $\mathbb{Z}$, der har et multiplikativt inverst element er $+1$ og -1. Dermed er $(\mathbb{Z},+,\cdot)$ altså ikke noget legeme. En triple af denne art kaldes for en *kommutativ ring*. Kommutative ringe forekommer i enorm variation.[1]

0.2 Højre og venstre multiplikation

Hvis $(\mathbb{F},+,\cdot)$ er et legeme, så gælder der ifølge aksiom (P4) at $a\cdot b = b\cdot a$ for alle $a,b\in\mathbb{F}$. Med andre ord, så gør det ingen forskel, om vi ganger a med b fra højre eller fra venstre. Der forekommer dog naturligt også tripler $(\mathbb{F},+,\cdot)$, der opfylder alle aksiomerne for at være et legeme med undtagelse af (P4), og vi siger da, at $(\mathbb{F},+,\cdot)$ er et *skævlegeme*.

Eksempel 0.2.1 Hamilton's *kvaternioner* er skævlegemet $(\mathbb{H},+,\cdot)$, hvor

$$\mathbb{H} = \{a+ib+jc+kd \mid a,b,c,d\in\mathbb{R}\}$$

er mængden af kvaternioner, og hvor "+" er defineret ved

$$\begin{aligned}&(a_1+ib_1+jc_1+kd_1)+(a_2+ib_2+jc_2+kd_2)\\&\quad=(a_1+a_2)+i(b_1+b_2)+j(c_1+c_2)+k(d_1+d_2),\end{aligned}$$

mens " · " er defineret ved

$$\begin{aligned}&(a_1+ib_1+jc_1+kd_1)\cdot(a_2+ib_2+jc_2+kd_2)\\&\quad=(a_1a_2-b_1b_2-c_1c_2-d_1d_2)+i(a_1b_2+b_1a_2+c_1d_2-d_1c_2)\\&\qquad+j(a_1c_2-b_1d_2+c_1a_2+d_1b_2)+k(a_1d_2+b_1c_2-c_1b_2+d_1a_2).\end{aligned}$$

Her er det vigtigt at holde styr på, om vi ganger fra højre eller fra venstre. For eksempel er $i\cdot j = k$, mens $j\cdot i = -k$.

I lineær algebra spiller multiplikation fra højre og venstre forskellige roller, og vi gør det derfor til en regel altid at være forsigtige med, hvilken side vi ganger fra. Den kommutative lov (P4) gælder således *ikke* for multiplikation af matricer, som vi indfører i kapitel 2. Endelig er (P4) slet ikke nødvendig i de fleste af de efterfølgende kapitler, og mange beviser bliver faktisk lettere, når man ikke tillader sig at bruge (P4).

[1] Det mere end 5000 sider lange Stacks Project (stacks.math.columbia.edu) omhandler kommutative ringe og deres egenskaber.

1 Lineære ligningssystemer

I dette kapitel indfører vi lineære ligningssystemer og beskriver en algoritme til at finde deres løsninger. Denne algoritme eller metode, der allerede er beskrevet i den kinesiske tekst "Ni kapitler om den matematiske kunst" skrevet omkring begyndelsen af vor tidsregning, blev langt senere genfundet af først Newton og senere Gauss, og går nu, noget uretfærdigt, under navnet *Gauss elimination*.

Algoritmen består i at omforme ligningssystemer ved hjælp af *rækkeoperationer*, der ikke ændrer løsningsmængden, men gør det muligt umiddelbart at angive denne. Disse rækkeoperationer er et helt grundlæggende værktøj i linear algebra, som benyttes i mange udregninger.

Vi samler al infomation om et lineært ligningssystem i en *matrix*, og vi indfører en speciel form for matricer, som vi kalder matricer på *reduceret echelon form*, for hvilke det tilhørende ligningssystem umiddelbart kan løses. Gauss elimination er da mere præcist en algoritme, der ved hjælp af rækkeoperationer omdanner en given matrix til en entydigt bestemt matrix på reduceret echelon form.

1.1 Lineære ligningssystemer

Vi siger, at et system af ligninger på formen

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

med $a_{ij}, b_i \in \mathbb{F}$ er et lineært ligningssystem med m ligninger i n ubekendte over $\mathbb{F}$. Skalaerne a_{ij} og b_i kaldes henholdsvis ligningssystemets *koefficienter* og *konstanter*. Ved en *løsning* til ligningssystemet forstås en familie af n skalarer

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

der samtidigt tilfredstiller de m ligninger i systemet. Den mængde, der består af alle løsninger til ligningssystemet, kaldes ligningssystemets *løsningsmængde*.

Hvis konstanterne b_i alle er lig med 0 kaldes ligningssystemet *homogent*, og ellers kaldes det *inhomogent*. Et homogent ligningssystem har altid løsningen

$$\boldsymbol{x} = \mathbf{0} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

mens ikke alle inhomogene ligningssystem har løsninger.

Eksempel 1.1.1 (Homogent og inhomogent ligningssystem) Ligningssystemet

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

med 3 ligninger i 3 ubekendte er inhomogent; og ligningssystemet

$$\begin{aligned} 2x_1 - 3x_2 + x_3 &= 0 \\ -x_1 + x_2 - x_3 &= 0 \\ 3x_1 + 2x_2 + 2x_3 &= 0 \end{aligned}$$

er det tilhørende homogene ligningssystem.

Det er nyttigt ikke at skulle skrive de variable x_j hele tiden, og derfor samler vi ligningssystemets koefficienter a_{ij} i en $m \times n$-matrix

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

som vi kalder ligningssystemets *koefficientmatrix*. Vi siger, at matricen har m rækker og n søjler. Hvis vi tilføjer den ekstra søjle

$$\boldsymbol{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix},$$

der består af ligningssystemets konstanter, så får vi en $m \times (n+1)$-matrix

$$(A \mid \boldsymbol{b}) = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array} \right),$$

som vi kalder ligningssystemets *totalmatrix*. Vi har her, som det er sædvane, inkluderet en lodret linie, der adskiller koefficienter og konstanter. Linien er ikke af matematisk betydning og har udelukkende til formål at huske os på, hvilket ligningssystem denne matrix repræsenterer.

Eksempel 1.1.2 (Totalmatrix) Ligningssystemet i eksempel 1.1.1 har totalmatrix

$$\left(\begin{array}{rrr|r} 2 & -3 & 1 & -1 \\ -1 & 1 & -1 & 2 \\ 3 & 2 & 2 & 3 \end{array}\right).$$

Den har 3 rækker og 4 søjler, og er dermed en 3×4-matrix.

1.2 Rækkeoperationer

Lad os som et simpelt eksempel bestemme løsningsmængden for ligningssystemet

$$\begin{aligned} 4x_1 + x_2 &= 9 \\ x_1 - x_2 &= 1. \end{aligned}$$

Vi adderer først den anden ligning til den første og får da ligningssystemet

$$\begin{aligned} 5x_1 \qquad &= 10 \\ x_1 - x_2 &= 1. \end{aligned}$$

Vi ganger dernæst den første ligning med $\frac{1}{5}$ og den anden med -1, hvilket giver

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

Endelig adderer vi den første ligning til den anden, hvilket giver

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

De fire ligningssystemer har alle samme løsningsmængde, da de operationer, vi har udført bevarer løsningsmængden og kan inverteres. Fra det sidste ligningssystem kan vi derfor konkludere, at denne fælles løsningsmængde består af den entydige løsning

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

Vi skal nu vise, at denne løsningsmetode virker for et generelt lineært ligningssystem.

Så vi betragter lineære ligningssystemer

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

bestående af m ligninger i n ubekendte og indfører de følgende tre typer af operationer, der omformer et sådant lineært ligningssystem til et nyt lineært ligningssystem.

Type M: Multplication fra venstre af en ligning med en skalar $c \neq 0$.

Type S: Addition af et venstre multiplum af en ligning til en anden ligning.

Type T: Ombytning af to ligninger.

Her står "M" for *multiplikation*; "S" for *sum*; og "T" for *transposition*. Vi husker på fra afsnit 0.2, at vi skelner mellem venstre og højre multiplikation.

De tre typer af operationer bevarer alle både antallet m af ligninger og antallet n af ubekendte. Endvidere gælder der, at enhver løsning til et givet ligningssystem igen er en løsning til det ligningssystem, der fremkommer ved at udføre en af de tre typer operationer. Desuden har hver af de tre typer operationer en *invers* operation: For hvis vi ganger ligning i med $c \neq 0$ fra venstre, så kan vi gange den nye ligning i med $c^{-1} \neq 0$ fra venstre og derved få den gamle ligning i tilbage. Og hvis vi adderer c gange ligning j til ligning $i \neq j$, så kan vi adderer $-c$ gange ligning j til den nye ligning i og derved få den oprindelige ligning i tilbage. Endelig, hvis vi ombytter ligning i og j, så kan vi ombytte dem igen, og derved komme tilbage til, der hvor vi startede. Derfor er enhver løsning til det nye ligningssystem også en løsning til det oprindelige ligningssystem. Med andre ord har de to ligningssystemer den samme løsningsmængde.

Vi udtrykker nu de tre typer operationer ved ligningssystemernes totalmatricer

$$(A \mid \boldsymbol{b}) = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array}\right),$$

hvor de svarer til de følgende tre typer *rækkeoperationer* på matricer:

Type M: Multplication fra venstre af en række med en skalar $c \neq 0$.

Type S: Addition af et venstre multiplum af en række til en anden række.

Type T: Ombytning af to rækker.

Vi bemærker, at disse rækkeoperationer bevarer antallet af rækker og antallet af søjler.

Rækkeoperioner forstås lettest ved gennemgang af nogle eksempler. Disse eksempler illustrerer samtidigt, hvordan vi angiver rækkeoperationer.

Eksempel 1.2.1 Vi betragter igen ligningssystemet

$$\begin{aligned} 4x_1 + x_2 &= 9 \\ x_1 - x_2 &= 1, \end{aligned}$$

som vi løste i starten af afsnitet. Vi udfører nu de rækkeoperationer på ligningssystemet totalmatrix $(A \mid \boldsymbol{b})$, der svarer til de rækkeoperationer, vi anvendte tidligere.

$$\begin{aligned}
(A \mid \boldsymbol{b}) = & \left(\begin{array}{rr|r} 4 & 1 & 9 \\ 1 & -1 & 1 \end{array}\right) \begin{array}{l} +R_2 \\ \end{array} \\
& \left(\begin{array}{rr|r} 5 & 0 & 10 \\ 1 & -1 & 1 \end{array}\right) \begin{array}{l} \frac{1}{5} \cdot R_1 \\ \end{array} \\
& \left(\begin{array}{rr|r} 1 & 0 & 2 \\ 1 & -1 & 1 \end{array}\right) \begin{array}{l} \\ (-1) \cdot R_2 \end{array} \\
& \left(\begin{array}{rr|r} 1 & 0 & 2 \\ -1 & 1 & -1 \end{array}\right) \begin{array}{l} \\ +R_1 \end{array} \\
(A' \mid \boldsymbol{b}') = & \left(\begin{array}{rr|r} 1 & 0 & 2 \\ 0 & 1 & 1 \end{array}\right)
\end{aligned}$$

Vi bemærker, at de fem matricer ovenfor *ikke* er lig med hinanden, og at det derfor er forkert at skrive lighedstegn mellem dem. Ligningssystemet hørende til $(A' \mid \boldsymbol{b}')$ er

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

hvilket er trivielt at løse.

Svarende til hver af de tre typer af rækkeoperationer "M", "S" og "T", indfører vi nu den tilhørende *inverse* rækkeoperation.

Den inverse af den rækkeoperation, der består i at gange række i med $c \neq 0$ fra venstre, er den rækkeoperation, der består i at gange række i med c^{-1} fra venstre.

Den inverse af den rækkeoperation, der består i at addere c gange række j til række $i \neq j$, er den rækkeoperation, der består i at addere $-c$ gange række j til række i.

Den inverse af den rækkeoperation, der består i at ombytte række i og j, er den samme rækkeoperation.

Hvis en given rækkeoperation omformer matricen A til matricen A', så omformer den inverse rækkeoperation altså matricen A' til matricen A.

Eksempel 1.2.2 Vi betragter igen matricerne $(A \mid \boldsymbol{b})$ og $(A' \mid \boldsymbol{b}')$ fra eksempel 1.2.1. Vi udfører nu de inverse rækkeoperationer af de rækkeoperationer, vi udførte for at omdanne $(A \mid \boldsymbol{b})$ til $(A' \mid \boldsymbol{b}')$, og vi udfører dem endvidere i den omvendte rækkefølge.

$$
\begin{aligned}
(A' \mid \boldsymbol{b}') = &\left(\begin{array}{cc|c} 1 & 0 & 2 \\ 0 & 1 & 1 \end{array}\right) \begin{array}{l} \\ +(-1)\cdot R_1 \end{array} \\
&\left(\begin{array}{cc|c} 1 & 0 & 2 \\ -1 & 1 & -1 \end{array}\right) \begin{array}{l} \\ (-1)\cdot R_2 \end{array} \\
&\left(\begin{array}{cc|c} 1 & 0 & 2 \\ 1 & -1 & 1 \end{array}\right) \begin{array}{l} 5\cdot R_1 \\ \\ \end{array} \\
&\left(\begin{array}{cc|c} 5 & 0 & 10 \\ 1 & -1 & 1 \end{array}\right) \begin{array}{l} +(-1)\cdot R_2 \\ \\ \end{array} \\
(A \mid \boldsymbol{b}) = &\left(\begin{array}{cc|c} 4 & 1 & 9 \\ 1 & -1 & 1 \end{array}\right)
\end{aligned}
$$

De inverse rækkeoperationer, anvendt i omvendt rækkefølge, omdanner altså matricen $(A' \mid \boldsymbol{b}')$ til den oprindelige matrix $(A \mid \boldsymbol{b})$.

Sætning 1.2.3 *Hvis totalmatricen $(A \mid \boldsymbol{b})$ for et lineært ligningssystem kan omdannes til totalmatricen $(A' \mid \boldsymbol{b}')$ for et andet lineært ligningssystem ved at udføre et endeligt antal rækkeoperationer, så har de to ligningssystemer den samme løsningsmængde.*

Bevis Vi beviser sætningen ved induktion på antallet $N \geq 0$ af rækkeoperationer, der benyttes til at omdanne $(A \mid \boldsymbol{b})$ til $(A' \mid \boldsymbol{b}')$. Hvis $N = 0$, så er $(A \mid \boldsymbol{b}) = (A' \mid \boldsymbol{b}')$, og der er derfor ikke noget at vise. Vi antager derfor, at påstanden er bevist for $N = r - 1$, og beviser den for $N = r$. Så lad $(A \mid \boldsymbol{b})$ være totalmatricen for et lineært ligningssystem, og lad $(A' \mid \boldsymbol{b}')$ være en matrix, der er fremkommet fra matricen $(A \mid \boldsymbol{b})$ ved at udføre r rækkeoperationer. Vi skal vise, at ligningssystemerne hørende til $(A \mid \boldsymbol{b})$ og $(A' \mid \boldsymbol{b}')$ har den samme løsningsmængde. Lad $(B \mid \boldsymbol{c})$ være den matrix, der fremkommer fra $(A \mid \boldsymbol{b})$ ved at udføre de $r - 1$ første af de r rækkeoperationer. Per induktion ved vi da, at ligningssystemerne hørende til $(A \mid \boldsymbol{b})$ og $(B \mid \boldsymbol{c})$ har samme løsningsmængde, og det er derfor tilstrækkeligt at vise, at ligningssystemerne hørende til $(B \mid \boldsymbol{c})$ og $(A' \mid \boldsymbol{b}')$ har samme løsningsmængde. Men $(A' \mid \boldsymbol{b}')$ fremkommer fra $(B \mid \boldsymbol{c})$ ved at udføre en enkelt rækkeoperation, og som vi allerede har bemærket har de tilhørende ligningssystemer derfor den samme løsningsmængde. Dette viser induktionsskridtet og dermed sætningen. $\square$

Eksempel 1.2.4 Hvis man er *meget* forsigtig, så kan to eller flere operationer udføres i samme skridt forudsat, at de er *ombyttelige*. De fire operationer

$$
\begin{aligned}
A = &\begin{pmatrix} 0 & 1 & 2 & 1 \\ 1 & 7 & 3 & 1 \\ 0 & 4 & 6 & -6 \end{pmatrix} \begin{array}{l} R_1 \leftrightarrow R_2 \\ \\ \frac{1}{2} \cdot R_3 \end{array} \\
&\begin{pmatrix} 1 & 7 & 3 & 1 \\ 0 & 1 & 2 & 1 \\ 0 & 2 & 3 & -3 \end{pmatrix} \begin{array}{l} +(-1) \cdot R_2 \\ \\ +3 \cdot R_2 \end{array} \\
A' = &\begin{pmatrix} 1 & 6 & 1 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 5 & 9 & 0 \end{pmatrix}
\end{aligned}
$$

omdanner A til A'. De første to operationer er ombyttelige, da de ikke involverer de samme rækker, og det samme gælder for de sidste to operationer.

Vi vil anvende rækkeoperationer til at omdanne totalmatricerne af ligningssystemer, til matricer på en særlig form, som vi kalder *reduceret echelon form*.

Definition 1.2.5 En $m \times n$-matrix $A = (a_{ij})$ siges at være på *reduceret echelon form*, hvis der findes $r \geq 0$ og $1 \leq j_1 < \cdots < j_r \leq n$, sådan at følgende gælder:

(1) For alle $1 \leq s \leq r$ gælder, at

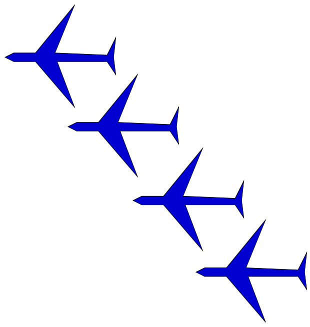
$$a_{ij_s} = \begin{cases} 1 & \text{hvis } i = s, \\ 0 & \text{hvis } i \neq s. \end{cases}$$

(2) For alle $1 \leq s \leq r$ og $1 \leq j < j_s$ gælder, at $a_{sj} = 0$.

(3) For alle $r < i \leq m$ og $1 \leq j \leq n$ gælder, at $a_{ij} = 0$.

Hvis $A = (a_{ij})$ er en $m \times n$-matrix, så siges skalarene a_{ij} at være matricens *indgange*. Betingelsen (1) udtrykker altså, at i den j_s'te søjle, hvor $1 \leq s \leq r$, er indgangen $a_{sj_s} = 1$, mens alle øvrige indgange er lig med 0. Betingelsen (2) udtrykker ligeledes, at i den s'te række, hvor $1 \leq s \leq r$, er indgangen $a_{sj_s} = 1$ den første indgang fra venstre som ikke er lig med 0. Endelig udtrykker betingelsen (3), at indgangene i de nederste $m - r$ rækker alle er lig med 0. Hvis en matrix A er på reduceret echelon form, så kaldes antallet r af ikke-nul rækker for *rangen* af matricen, og den første ikke-nul indgang fra venstre i en række kaldes for rækkens *ledende indgang*. De ledende indgange er altså alle lig med 1, og antallet af ledende indgange er lig med rangen r.

Vi siger endvidere, at en $m \times n$-matrix $A = (a_{ij})$ er på *echelon form*, hvis den opfylder betingelserne (2) og (3) i definition 1.2.5 samt den svagere betingelse



Figur 1.1: Betegnelsen "echelon form" har sin oprindelse i den militære betegnelse "echelon formation". Figuren viser fly, der flyver i echelon formation.

(1') For alle $1 \leq s \leq r$ gælder, at $a_{s,j_s} \neq 0$.

Hvis en matrix A er på echelon form, så kalder vi igen antallet r af ikke-nul rækker for matricens rang, og vi kalder den første ikke-nul indgang fra venstre i en række for rækkens ledende indgang. Der gælder altså igen, at rangen r er lig med antallet af ledende indgange.

Eksempel 1.2.6 De to følgende 4×8-matricer, hvor "$*$" indikerer en vilkårlig skalar, er henholdsvis på reduceret echelon form og på echelon form.

$$\begin{pmatrix} 1 & * & * & 0 & * & 0 & * & * \\ 0 & 0 & 0 & 1 & * & 0 & * & * \\ 0 & 0 & 0 & 0 & 0 & 1 & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 2 & * & * & * & * & * & * \\ 0 & 0 & 4 & * & * & * & * & * \\ 0 & 0 & 0 & 0 & 5 & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 \end{pmatrix}$$

Deres rang er henholdsvis 3 og 4, og de ledende indgange er markeret med blåt. Vi bemærker, at i matricen, der er på reduceret echelon form, er de ledende indgange alle lig med 1, mens alle indgange, der ligger over en ledende indgang er lig med 0.

Vi viser nu den følgende vigtige sætning. Vi skal senere i kapitel 2, hvor vi har lidt flere værktøjer til vores rådighed, vise, at matricen A' i sætningen er entydigt bestemt af den givne matrix A.

Sætning 1.2.7 *Enhver matrix A kan omdannes til en matrix A' på reduceret echelon form ved at udføre et endeligt antal rækkeoperationer.*

Bevis Vi lader A være en $m \times n$-matrix og viser ved induktion på antallet $n \geq 1$ af søjler, at A ved brug af rækkeoperationer kan omdannes til en matrix A' på reduceret echelon form. Vi betragter først det basale tilfælde $n = 1$, hvor

$$A = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}$$

er en enkelt søjle. Hvis $a_{i1} = 0$ for alle i, så er $A' = A$ allerede på reduceret echelon form af rang $r = 0$, og i modsat fald findes $1 \leq i \leq m$ med $a_{i1} \neq 0$. Vi anvender nu først en rækkeoperation af type M til at gange den i'te række med a_{i1}^{-1}, hvorefter vi anvender en rækkeoperation af type T til at ombytte den i'te række med den første række. Herved får vi en matrix B på formen

$$B = \begin{pmatrix} 1 \\ b_{21} \\ \vdots \\ b_{m1} \end{pmatrix}.$$

Endelig anvender vi for alle $2 \leq i \leq m$ en rækkeoperation af type S til at addere $-b_{i1}$ gange den første række til den i'te række, hvorved vi får matricen

$$A' = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

som er på reduceret echelon form af rang $r = 1$. Dette viser påstanden for $n = 1$.

Så vi antager induktivt, at påstanden er vist for $n = p - 1$ og beviser den for $n = p$. Hvis alle indgange i den første søjle er lig med 0, så er

$$A = \left(\begin{array}{c|c} 0 & \\ \vdots & \quad B \quad \\ 0 & \end{array}\right),$$

hvor B er en $m \times (p-1)$–matrix. Den induktive antagelse viser, at der findes en følge af rækkeoperationer, som omdanner B til en matrix B' på reduceret echelon form, og den samme følge af rækkeoperationer omdanner da matricen A til matricen

$$A' = \left(\begin{array}{c|c} 0 & \\ \vdots & \quad B' \quad \\ 0 & \end{array}\right),$$

der som ønsket er på reduceret echelon form. Hvis ikke alle indgange i den første søjle er lig med 0, så findes $1 \leq i \leq m$ med $a_{i1} \neq 0$, og vi kan da som i tilfældet $n = 1$ først anvende en rækkeoperation af type M til at gange den i'te række med a_{i1}^{-1} og derefter anvende en rækkeoperation af type T til at ombytte den i'te række med den første række. Ved disse rækkeoperationer bliver A omdannet til en matrix B på formen

$$B = \left(\begin{array}{c|ccc} 1 & b_{12} & \cdots & b_{1p} \\ \hline 0 & & & \\ \vdots & & C & \\ 0 & & & \end{array}\right),$$

hvor C er en $(m-1)\times(p-1)$–matrix. Den induktive antagelse viser nu, at der findes en følge af rækkeoperationer, der omdanner C til en matrix C' på reduceret echelon form, og da alle indgange i den første søjle i B med undtagelse af $b_{11} = 1$ er lig med 0, så omdanner de tilsvarende rækkeoperationer $m \times p$-matricen B til matricen

$$D = \left(\begin{array}{c|ccc} 1 & b_{12} & \cdots & b_{1p} \\ \hline 0 & & & \\ \vdots & & C' & \\ 0 & & & \end{array}\right).$$

Endelig lader vi $1 \leq j_2 < \cdots < j_r \leq p$ numerere de søjler i D, der indeholder de ledende indgange i C', og anvender for alle $2 \leq s \leq r$ rækkeoperationen af type S, der adderer $-b_{1j_s}$ gange den s'te række i D til den første række i D. Matricen A', der fremkommer herved, er da på reduceret echelon form af rang r. Dette viser induktionsskridtet og dermed sætningen. □

Beviset for sætning 1.2.7 udgør en algoritme, der anvender rækkeoperationer til at omdanne en given matrix til en matrix på reduceret echelon form, og det er denne algoritme, som vi kalder for *Gauss elimination*. Givet et lineært ligningssystem, kan vi specielt anvende denne algoritme på ligningssystemets totalmatrix $(A \mid \boldsymbol{b})$. Vi får derved en matrix $(A' \mid \boldsymbol{b}')$ på reduceret echelon form, og løsningsmængden for det tilhørende ligningssystem kan da ganske umiddelbart angives. Denne løsningsmængde er ifølge sætning 1.2.3 den samme som løsningsmængden for det oprindelige ligningssystem. Vi vil nu illustrere denne metode ved gennemgang af nogle eksempler.

Eksempel 1.2.8 (Entydig løsning) Vi anvender Gauss elimination til at bestemme løsningsmængden til det lineære ligningssystem

$$\begin{array}{rrrcr} 2x_1 & -3x_2 & +x_3 & = & -1 \\ -x_1 & +x_2 & -x_3 & = & 2 \\ 3x_1 & +2x_2 & +2x_3 & = & 3. \end{array}$$

Vi opskriver derfor ligningssystemets totalmatrix $(A \mid \boldsymbol{b})$ og anvender rækkeoperationer til at omdanne denne til en matrix $(A' \mid \boldsymbol{b}')$ på reduceret echelon form.

$$
\begin{aligned}
(A \mid \boldsymbol{b}) = &\left(\begin{array}{rrr|r} 2 & -3 & 1 & -1 \\ -1 & 1 & -1 & 2 \\ 3 & 2 & 2 & 3 \end{array}\right) \begin{array}{l} \\ R_1 \leftrightarrow R_2 \\ \\ \end{array} \\
&\left(\begin{array}{rrr|r} -1 & 1 & -1 & 2 \\ 2 & -3 & 1 & -1 \\ 3 & 2 & 2 & 3 \end{array}\right) \begin{array}{l} \\ +2R_1 \\ +3R_1 \end{array} \\
&\left(\begin{array}{rrr|r} -1 & 1 & -1 & 2 \\ 0 & -1 & -1 & 3 \\ 0 & 5 & -1 & 9 \end{array}\right) \begin{array}{l} +R_2 \\ \\ +5R_2 \end{array} \\
&\left(\begin{array}{rrr|r} -1 & 0 & -2 & 5 \\ 0 & -1 & -1 & 3 \\ 0 & 0 & -6 & 24 \end{array}\right) \begin{array}{l} (-1)\cdot R_1 \\ (-1)\cdot R_2 \\ (-\frac{1}{6})\cdot R_3 \end{array} \\
&\left(\begin{array}{rrr|r} 1 & 0 & 2 & -5 \\ 0 & 1 & 1 & -3 \\ 0 & 0 & 1 & -4 \end{array}\right) \begin{array}{l} +(-2)R_3 \\ +(-1)R_3 \\ \\ \end{array} \\
(A' \mid \boldsymbol{b}') = &\left(\begin{array}{rrr|r} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -4 \end{array}\right)
\end{aligned}
$$

Her har vi markeret de indgange, vi har ønsket at ændre, med rødt; og i den endelige matrix $(A' \mid \boldsymbol{b}')$ har vi markeret de 3 ledende indgange med blåt. Ligningssystemet, der har $(A' \mid \boldsymbol{b}')$ som totalmatrix, er nu

$$
\begin{array}{ccccc} x_1 & & & = & 3 \\ & x_2 & & = & 1 \\ & & x_3 & = & -4. \end{array}
$$

Dette ligningssystem har tydeligvis præcis den ene løsning

$$
\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ -4 \end{pmatrix}.
$$

Ifølge sætning 1.2.3 er denne løsning derfor også den entydige løsning til det oprindelige ligningssystem.

Eksempel 1.2.9 (Løsningsmængde med een parameter) Vi anvender Gauss elimination til at bestemme løsningsmængden til det lineære ligningssystem

$$
\begin{array}{rrrcr} -x_1 & -2x_2 & -5x_3 & = & -3 \\ 2x_1 & +3x_2 & +8x_3 & = & 4 \\ 2x_1 & +6x_2 & +14x_3 & = & 10. \end{array}
$$

Vi opskriver igen totalmatricen for ligningssystemet og omdanner denne til en matrix på reduceret echelon form ved hjælp af rækkeoperationer.

$$
\begin{array}{rl}
(A \mid \boldsymbol{b}) = \left(\begin{array}{rrr|r} -1 & -2 & -5 & -3 \\ 2 & 3 & 8 & 4 \\ 2 & 6 & 14 & 10 \end{array}\right) & \begin{array}{l} \\ +2R_1 \\ +2R_1 \end{array} \\
\left(\begin{array}{rrr|r} -1 & -2 & -5 & -3 \\ 0 & -1 & -2 & -2 \\ 0 & 2 & 4 & 4 \end{array}\right) & \begin{array}{l} +(-2)R_2 \\ \\ +2R_2 \end{array} \\
\left(\begin{array}{rrr|r} -1 & 0 & -1 & 1 \\ 0 & -1 & -2 & -2 \\ 0 & 0 & 0 & 0 \end{array}\right) & \begin{array}{l} (-1)\cdot R_1 \\ (-1)\cdot R_2 \\ \\ \end{array} \\
(A' \mid \boldsymbol{b}') = \left(\begin{array}{rrr|r} 1 & 0 & 1 & -1 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right). &
\end{array}
$$

Vi har igen markeret med rødt de indgange vi ønsker at ændre, og vi har markeret med blåt de 2 ledende indgange. Ligningssystemet, der har $(A' \mid \boldsymbol{b}')$ som totalmatrix, er nu

$$
\begin{array}{rcr}
x_1 \qquad + \ x_3 & = & -1 \\
x_2 + 2x_3 & = & 2 \\
0 & = & 0.
\end{array}
$$

Vi ser heraf, at der for hver værdi t af x_3, findes præcis en værdi af x_1 og af x_2, der giver en løsning til ligningssystemet, nemlig

$$
\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -1-t \\ 2-t \\ t \end{pmatrix}.
$$

Med andre ord er den fælles løsningsmængde for de to ligningssystemer

$$
\left\{ \boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -1-t \\ 2-2t \\ t \end{pmatrix} \;\middle|\; t \in \mathbb{F} \right\}.
$$

Dermed er der uendeligt mange løsninger, som er *parametriseret* ved parameteren $t \in \mathbb{F}$. Vi siger, at variablen x_3 er en *fri variabel*,, da den kan tilordnes en vilkårlig værdi, og bemærker, at den svarer til den søjle i $(A' \mid \boldsymbol{b}')$, der *ikke* indeholder en ledende indgang. Vi siger endvidere, at variablene x_1 og x_2, der svarer til de søjler i $(A' \mid \boldsymbol{b}')$ der indeholder en ledende indgang, er *ledende variable*. Generelt kan de ledende variable på entydigvis udtrykkes ved de frie variable.

Eksempel 1.2.10 (Tom løsningsmængde) Vi anvender igen Gauss elimination til at bestemme løsningsmængden til det lineære ligningssystem

$$\begin{array}{rrrcr} -x_1 & -2x_2 & -5x_3 & = & -3 \\ 2x_1 & +3x_2 & +8x_3 & = & 4 \\ 2x_1 & +6x_2 & +14x_3 & = & 5. \end{array}$$

Dette ligningssystem har samme koefficientmatrix som ligningssystemet i eksempel 1.2.9, men konstantsøjlen er en anden. Vi opskriver totalmatricen for ligningssystemet og omdanner denne til en matrix på reduceret echelon form.

$$\begin{aligned}
(A \mid \boldsymbol{b}) = & \left(\begin{array}{rrr|r} -1 & -2 & -5 & -3 \\ 2 & 3 & 8 & 4 \\ 2 & 6 & 14 & 5 \end{array}\right) \begin{array}{l} \\ +2R_1 \\ +2R_1 \end{array} \\
& \left(\begin{array}{rrr|r} -1 & -2 & -5 & -3 \\ 0 & -1 & -2 & -2 \\ 0 & 2 & 4 & -1 \end{array}\right) \begin{array}{l} +(-2)R_2 \\ \\ +2R_2 \end{array} \\
& \left(\begin{array}{rrr|r} -1 & 0 & -1 & 1 \\ 0 & -1 & -2 & -2 \\ 0 & 0 & 0 & -1 \end{array}\right) \begin{array}{l} (-1)\cdot R_1 \\ (-1)\cdot R_2 \\ (-1)\cdot R_3 \end{array} \\
& \left(\begin{array}{rrr|r} 1 & 0 & 1 & -1 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 0 & 1 \end{array}\right) \begin{array}{l} +R_3 \\ +(-2)R_3 \\ \\ \end{array} \\
(A' \mid \boldsymbol{b}') = & \left(\begin{array}{rrr|r} 1 & 0 & 1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right).
\end{aligned}$$

Vi har igen markeret de indgange, vi ønsker at ændre, med rødt, og vi har markeret de 3 ledende indgange med blåt. Ligningssystemet med $(A' \mid \boldsymbol{b}')$ som totalmatrix er nu

$$\begin{array}{rcr} x_1 \qquad + \ x_3 & = & 0 \\ x_2 + 2x_3 & = & 0 \\ 0 & = & 1. \end{array}$$

Dette ligningssystem har tydeligvis *ingen* løsninger, da den sidste ligning "$0 = 1$" *aldrig* er opfyldt unanset hvilke værdier, vi tilskriver de variable. Ifølge sætning 1.2.3 har det oprindelige ligningssystem derfor hellere ingen løsninger.

Generelt har et ligningssystem med totalmatrix $(A \mid \boldsymbol{b})$ ingen løsninger, hvis og kun hvis konstantsøjlen i den rækkeækvivalente matrix $(A' \mid \boldsymbol{b}')$ på reduceret echelon form indeholder en ledende indgang.

Eksempel 1.2.11 (Løsningsmængde med to parametre) Vi anvender Gauss elimination til at bestemme løsningsmængden til det lineære ligningssystem

$$\begin{array}{rrrrrcr} & & 2x_3 & -x_4 & +8x_5 & = & -13 \\ x_1 & -2x_2 & +3x_3 & +2x_4 & +x_5 & = & 10 \\ 3x_1 & -6x_2 & +10x_3 & +6x_4 & +5x_5 & = & 27 \end{array}$$

Så vi opskriver totalmatricen for ligningssystemet og omdanner denne til en matrix på reduceret echelon form ved hjælp af rækkeoperationer.

$$(A \mid \boldsymbol{b}) = \left(\begin{array}{rrrrr|r} 0 & 0 & 2 & -1 & 8 & -13 \\ 1 & -2 & 3 & 2 & 1 & 10 \\ 3 & -6 & 10 & 6 & 5 & 27 \end{array}\right) \begin{array}{l} R_1 \leftrightarrow R_2 \\ \\ \\ \end{array}$$

$$\left(\begin{array}{rrrrr|r} 1 & -2 & 3 & 2 & 1 & 10 \\ 0 & 0 & 2 & -1 & 8 & -13 \\ \color{red}{3} & -6 & 10 & 6 & 5 & 27 \end{array}\right) \begin{array}{l} \\ \\ +(-3)R_1 \end{array}$$

$$\left(\begin{array}{rrrrr|r} 1 & -2 & \color{red}{3} & 2 & 1 & 10 \\ 0 & 0 & \color{red}{2} & -1 & 8 & -13 \\ 0 & 0 & 1 & 0 & 2 & -3 \end{array}\right) \begin{array}{l} +(-3)R_3 \\ +(-2)R_3 \\ \\ \end{array}$$

$$\left(\begin{array}{rrrrr|r} 1 & -2 & 0 & \color{red}{2} & -5 & 19 \\ 0 & 0 & 0 & -1 & 4 & -7 \\ 0 & 0 & 1 & 0 & 2 & -3 \end{array}\right) \begin{array}{l} +2R_2 \\ \\ \\ \end{array}$$

$$\left(\begin{array}{rrrrr|r} 1 & -2 & 0 & 0 & 3 & 5 \\ 0 & 0 & 0 & \color{red}{-1} & 4 & -7 \\ 0 & 0 & 1 & 0 & 2 & -3 \end{array}\right) \begin{array}{l} \\ (-1)\cdot R_2 \\ \\ \end{array}$$

$$\left(\begin{array}{rrrrr|r} 1 & -2 & 0 & 0 & 3 & 5 \\ 0 & 0 & 0 & 1 & -4 & 7 \\ 0 & 0 & 1 & 0 & 2 & -3 \end{array}\right) \begin{array}{l} \\ R_2 \leftrightarrow R_3 \\ \\ \end{array}$$

$$(A' \mid \boldsymbol{b}') = \left(\begin{array}{rrrrr|r} \color{blue}{1} & -2 & 0 & 0 & 3 & 5 \\ 0 & 0 & \color{blue}{1} & 0 & 2 & -3 \\ 0 & 0 & 0 & \color{blue}{1} & -4 & 7 \end{array}\right)$$

Vi har markeret de indgange, vi ønsker at ændre, med rødt; og vi har markeret de 3 ledende indgange med blåt. Ligningssystemet med $(A' \mid \boldsymbol{b}')$ som totalmatrix er nu

$$\begin{array}{rrrrrcr} x_1 & -2x_2 & & & +3x_5 & = & 5 \\ & & x_3 & & +2x_5 & = & -3 \\ & & & x_4 & -4x_5 & = & 7. \end{array}$$

Vi ser umiddelbart, at for hver værdi $x_2 = t_1$ og $x_5 = t_2$ af de frie variable x_2 og x_5 har ligningssystemet præcis den ene løsning

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 5 & +2t_1 & -3t_2 \\ & t_1 & \\ -3 & & -2t_2 \\ 7 & & +4t_2 \\ & & t_2 \end{pmatrix}.$$

Den fælles løsningsmængde til de to ligningssystemer er derfor uendelig og er parametriseret ved de to parametre $t_1 \in \mathbb{F}$ og $t_2 \in \mathbb{F}$.

Vi angiver nu den generelle algoritme til løsning af lineære ligningssystemer.

Sætning 1.2.12 (Løsning af lineære ligningssystemer ved Gauss elimination)
Løsningsmængden for det lineære ligningssystem

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

bestående af m ligninger i n ubekendte over $\mathbb{F}$ bestemmes ved følgende algoritme.

(1) *Ligningssystemets totalmatrix*

$$(A \mid \boldsymbol{b}) = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array}\right)$$

omdannes ved en følge af rækkeoperationer til en matrix $(A' \mid \boldsymbol{b}')$ på reduceret echelon form af rang $0 \le r \le \min\{m,n\}$.

(2) *Hvis en af de ledende indgange i $(A' \mid \boldsymbol{b}')$ er indeholdt i $\boldsymbol{b}'$, så har ligningssystemet ingen løsninger.*

(3) *Hvis ingen af de ledende indgange i $(A' \mid \boldsymbol{b}')$ er indeholdt i $\boldsymbol{b}'$ og hvis $r = n$, så har ligningssystemet præcis een løsning.*

(4) *Hvis ingen af de ledende indgange i $(A' \mid \boldsymbol{b}')$ er indeholdt i $\boldsymbol{b}'$ og hvis $r < n$, så er løsningsmængden uendelig. Den parametriseres ved at opdele de n variable i de r ledende variable $x_{j_1},\dots,x_{j_r}$ svarende til de søjler i A', der indeholder en ledende indgang, og de resterende $p = n - r$ frie variable; tilordne de frie variable vilkårlige værdier $t_1,\dots,t_p$; og løse ligningssystemet med totalmatrix $(A' \mid \boldsymbol{b}')$ for at bestemme de tilhørende entydige værdier af de ledende variable.*

Bevis Vi har allerede vist i sætning 1.2.7, at der findes en følge af rækkeoperationer, der omdanner totalmatricen $(A \mid \boldsymbol{b})$ til en matrix $(A' \mid \boldsymbol{b}')$ på reduceret echelon form. Med andre ord kan skridtet (1) altid udføres. Desuden har vi vist i sætning 1.2.3, at det givne ligningssystem og ligningssystemet med totalmatrix $(A' \mid \boldsymbol{b}')$ har den samme løsningsmængde. I resten af beviset vil vi derfor udelukkende betragte det sidstnævnte ligningssystem.

Vi viser nu påstanden (2). Så vi antager, at $\boldsymbol{b}'$ indeholder en af de ledende indgange i matricen $(A' \mid \boldsymbol{b}')$. Da denne $m \times (n+1)$-matrix er på reduceret echelon form med $\boldsymbol{b}'$ som sidste søjle, gælder der nødvendigvis, at $j_r = n+1$, og derfor er $a'_{rj} = 0$ for alle $1 \leq j \leq n$, mens $b'_r = 1$. Den r'te ligning i det tilhørende ligningssystem er derfor

$$0 \cdot x_1 + \cdots + 0 \cdot x_n = 1,$$

hvilket som ønsket viser, at ligningssystemet ingen løsninger har.

Vi viser dernæst (3). Da $m \times (n+1)$-matricen $(A' \mid \boldsymbol{b}')$ er på reduceret echelon form og har $r = n$ ledende indgange, og da den sidste søjle $\boldsymbol{b}'$ ikke indeholder nogen ledende indgang, må alle søjler i A' nødvendigvis indeholde en ledende indgang. De første $r = n$ ligninger i det tilhørende ligningssystem er da henholdsvis

$$\begin{array}{ccccc} x_1 & & & = & b'_1 \\ & \ddots & & & \vdots \\ & & x_n & = & b'_n, \end{array}$$

mens de sidste $m - r$ ligninger alle er lig den trivielle ligning "$0 = 0$". Dette viser som ønsket, at ligningssystemet har en entydig løsning.

Vi viser endelig (4). Da $m \times (n+1)$-matricen $(A' \mid \boldsymbol{b}')$ er på reduceret echelon form og har $r < n$ ledende indgange, og da den sidste søjle $\boldsymbol{b}'$ ikke indeholder nogen ledende indgang, konkluderer vi derfor, at r af de n søjler i A' indeholder en ledende indgang, mens de resterende $p = r - n$ søjler ikke indeholder en ledende indgang. Vi indicerer nu med $1 \leq j_1 < \cdots < j_r \leq n$ og $1 \leq k_1 < \cdots < k_p \leq n$ de søjler i A', der henholdsvis indeholder og ikke indeholder en ledende indgang. De første r ligninger i ligningssystemet giver da

$$\begin{aligned} x_{j_1} &= b'_1 - (a'_{1,k_1} x_{k_1} + \cdots + a'_{1,k_p} x_{k_p}) \\ &\vdots \\ x_{j_r} &= b'_r - (a'_{r,k_1} x_{k_1} + \cdots + a'_{r,k_p} x_{k_p}), \end{aligned}$$

idet vi flytter de led, der indeholder en af de fri variable $x_{k_1}, \ldots, x_{k_p}$, til højre side af lighedstegnet. De resterende $m - r$ ligninger er igen alle lig den trivielle ligning "$0 = 0$". Dette viser som ønsket, at for hver værdi $x_{k_1} = t_1, \ldots, x_{k_p} = t_p$ af de frie variable, har ligningssystemet præcis een løsning. □

Bemærkning 1.2.13 Den opmærksomme læser kan måske undre sig over, hvorvidt matricen $(A' \mid \boldsymbol{b}')$ på reduceret echelon form i sætning 1.2.12 (1) og dens rang r afhænger af den anvendte følge af rækkeoperationer. For i sætning 1.2.7 har vi kun vist, at der findes en følge af rækkeoperationer, der omdanner den givne matrix $(A \mid \boldsymbol{b})$ til en matrix $(A' \mid \boldsymbol{b}')$ på reduceret echelon form. Vi viser i sætning 2.5.6 nedenfor, at matricen $(A' \mid \boldsymbol{b}')$ faktisk er entydigt bestemt af $(A \mid \boldsymbol{b})$, og at den ikke afhænger af den valgte følge af rækkeoperationer.

Eksempel 1.2.14 (Ligningssystem med komplekse tal) Vi anvender Gauss elimination til at løse det lineære ligningssystem

$$\begin{array}{rrrcr} x_1 & +(-1+i)x_2 & -x_3 & = & 4-i \\ -x_1 & -2ix_2 & +(2-i)x_3 & = & -2+7i \end{array}$$

med 2 ligninger i 3 variable over de komplekse tal $\mathbb{C}$. Vi opskriver ligningssystemets totalmatrix $(A \mid \boldsymbol{b})$ og omdanner den til en matrix $(A' \mid \boldsymbol{b}')$ på reduceret echolon form ved hjælp af rækkeoperationer.

$$\begin{array}{rl}
(A \mid \boldsymbol{b}) = & \left(\begin{array}{ccc|c} {\color{red}i} & -1+i & -1 & 4-i \\ -1 & -2i & 2-i & -2+7i \end{array}\right) \begin{array}{l} +i\cdot R_2 \\ \end{array} \\
& \left(\begin{array}{ccc|c} 0 & {\color{red}1+i} & 2i & -3-3i \\ -1 & -2i & 2-i & -2+7i \end{array}\right) \begin{array}{l} \frac{1}{2}(1-i)\cdot R_1 \\ \end{array} \\
& \left(\begin{array}{ccc|c} 0 & 1 & 1+i & -3 \\ -1 & {\color{red}-2i} & 2-i & -2+7i \end{array}\right) \begin{array}{l} \\ +2i\cdot R_1 \end{array} \\
& \left(\begin{array}{ccc|c} 0 & 1 & 1+i & -3 \\ {\color{red}-1} & 0 & i & -2+i \end{array}\right) \begin{array}{l} \\ (-1)\cdot R_2 \end{array} \\
& \left(\begin{array}{ccc|c} 0 & 1 & 1+i & -3 \\ 1 & 0 & -i & 2-i \end{array}\right) \; R_1 \leftrightarrow R_2 \\
(A' \mid \boldsymbol{b}') = & \left(\begin{array}{ccc|c} {\color{blue}1} & 0 & -i & 2-i \\ 0 & {\color{blue}1} & 1+i & -3 \end{array}\right)
\end{array}$$

Vi har som tidligere markeret de indgange, vi ønsker at ændre, med rødt og de ledende indgange i matricen $(A' \mid \boldsymbol{b}')$ med blåt. Da $\boldsymbol{b}'$ ikke indeholder ledende indgange, har ligningssystemet løsninger; og da rangen $r = 2$ af $(A' \mid \boldsymbol{b}')$ er mindre end antallet $n = 3$ af variable, så er løsningsmængden uendelig og parametriseret af $p = r - n = 1$ parameter. Endelig aflæser vi umiddelbart fra matricen $(A' \mid \boldsymbol{b}')$, at for hver værdi $t \in \mathbb{C}$ af den frie variabel x_3 er der præcis den ene løsning

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} (2-i)+i\cdot t \\ -3-(1+i)\cdot t \\ t \end{pmatrix}$$

til ligningssystemet. Vi understreger, at $t \in \mathbb{C}$ er en kompleks parameter.

Lemma 1.2.15 *Enhver matrix B på echelon form af rang r kan ved endeligt mange rækkeoperationer omdannes til en matrix B' på reduceret echelon form af samme rang r.*

Bevis Lad B være en $m \times n$-matrix på echelon form, og lad $1 \le j_1 < \cdots < j_r \le n$ indicere de søjler i B, der indeholder en ledende indgang. For alle $1 \le s \le r$ anvender vi først en rækkeoperation af type M til at gange den s'te række med $b_{sj_s}^{-1}$. Herved opnår vi en ny matrix C på echelon form af rang r med ledende indgange

$$c_{sj_s} = b_{sj_s}^{-1} \cdot b_{sj_s} = 1,$$

hvor $1 \le s \le r$. Dernæst anvender vi for alle $2 \le s \le r$ og $1 \le i \le s-1$ den rækkeoperation af type S, der adderer $-c_{ij_s}$ gange den s'te række i C til den i'te række i C. Herved opnår vi en matrix B' på reduceret echelon form af rang r med ledende indgange

$$b'_{sj_s} = c_{sj_s} = 1,$$

hvor $1 \le s \le r$, hvilket viser lemmaet. □

Vi har vist i sætning 1.2.7, at totalmatricen $(A \mid \boldsymbol{b})$ for et givet ligningssystem ved rækkeoperationer kan omdannes til en matrix $(A' \mid \boldsymbol{b}')$ på reduceret echelon form, og vi har givet en fuldstændig beskrivelse af ligningssystemets løsningmængde udtrykt ved matricen $(A' \mid \boldsymbol{b}')$ i sætning 1.2.12. Lemma 1.2.15 viser, at hvis vi kun ønsker at bestemme kvalitative egenskaber ved løsningsmængden såsom, hvorvidt den er tom eller ej; hvorvidt den er endelig eller ej; samt antallet $p = n - r$ af parametre, så kan vi dog nøjes med at omdanne $(A \mid \boldsymbol{b})$ til en matrix $(B \mid \boldsymbol{c})$ på echelon form. Da det normalt kræver færre rækkeoperationer at omdanne $(A \mid \boldsymbol{b})$ til en matrix $(B \mid \boldsymbol{c})$ på echelon form end til en matrix $(A' \mid \boldsymbol{b}')$ på reduceret echelon from, så kan man herved nogen gange spare sig lidt arbejde. Vi bemærker dog, at, modsat $(A' \mid \boldsymbol{b}')$, så er $(B \mid \boldsymbol{c})$ ikke entydigt bestemt ved matricen $(A \mid \boldsymbol{b})$.

Eksempel 1.2.16 (Echelon / reduceret echelon form) Vi betragter ligningssystemet

$$\begin{array}{rrrrcr} 2x_1 & +3x_2 & -4x_3 & -x_4 & = & 12 \\ -2x_1 & & +10x_3 & +3x_4 & = & 4 \\ 2x_1 & +6x_2 & +2x_3 & & = & 2 \end{array}$$

og omdanner først dets totalmatrix $(A \mid \boldsymbol{b})$ til en matrix $(B \mid \boldsymbol{c})$ på echelon form.

$$(A \mid \boldsymbol{b}) = \left(\begin{array}{rrrr|r} 2 & 3 & -4 & -1 & 12 \\ -2 & 0 & 10 & 3 & 4 \\ 2 & 6 & 2 & 0 & 2 \end{array}\right) \begin{array}{l} \\ +R_1 \\ +(-1)\cdot R_1 \end{array}$$

$$\left(\begin{array}{rrrr|r} 2 & 3 & -4 & -1 & 12 \\ 0 & 3 & 6 & 2 & 16 \\ 0 & 3 & 6 & 1 & -10 \end{array}\right) \begin{array}{l} \\ \\ +(-1)\cdot R_2 \end{array}$$

$$(B \mid \boldsymbol{c}) = \left(\begin{array}{rrrr|r} 2 & 3 & -4 & -1 & 12 \\ 0 & 3 & 6 & 2 & 1 \\ 0 & 0 & 0 & -1 & -26 \end{array}\right)$$

Vi har her som tidligere markeret de ledende indgange med blåt. Hvis vi yderligere omdanner $(B \mid \boldsymbol{c})$ til en matrix $(A' \mid \boldsymbol{b}')$ på reduceret echelon form, så vil positionen af de ledende ikke ændres. Derfor kan vi allerede nu konkludere, at ligningssystemet har uendeligt mange løsninger, og at løsningsmængden kan parametriseres ved

$$p = n - r = 4 - 3 = 1$$

parameter. Hvis vi ønsker præcist at angive løsningsmængden, så må vi først bestemme matricen $(A' \mid \boldsymbol{b}')$, hvilket vi i dette tilfælde overlader til læseren.

Vi har hovedsageligt anvendt rækkeoperationer på totalmatricen $(A \mid \boldsymbol{b})$ hørende til et lineært ligningssystem, men vi kan også anvende rækkeoperationer til at omdanne ligningssystemets koefficientmatrix A til en matrix B på echelon form. Vi afslutter dette kapitel med følgende vigtige resultat angående betydningen af rangen af matricen B.

Sætning 1.2.17 *Lad A være en $m \times n$-matrix, der ved hjælp af rækkeoperationer kan omdannes til en $m \times n$-matrix B på echelon form af rang $0 \leq r \leq \min\{m, n\}$. Da gælder:*

(1) *Ligningssystemet med totalmatrix $(A \mid \boldsymbol{b})$ har mindst en løsning for alle valg af konstantsøjle $\boldsymbol{b}$, hvis og kun hvis $r = m$.*

(2) *Ligningssystemet med totalmatrix $(A \mid \boldsymbol{b})$ har højst en løsning for alle valg af konstantsøjle $\boldsymbol{b}$, hvis og kun hvis $r = n$.*

(3) *Ligningssystemet med totalmatrix $(A \mid \boldsymbol{b})$ har netop een løsning for alle valg af konstantsøjle $\boldsymbol{b}$, hvis og kun hvis $r = m = n$.*

Bevis Ifølge lemma 1.2.15 kan vi uden tab af generalitet antage, at matricen $B = A'$ er på reduceret echelon form og af rang r, hvilket vi vil gøre i resten af beviset. Vi vælger også en gang for alle en følge af rækkeoperationer, der omdanner A til A'.

Vi viser først (1). Hvis $r = m$, da indeholder hver række i A' en ledende indgang. Derfor gælder der for alle $\boldsymbol{b}$, at matricen $(A' \mid \boldsymbol{b}')$, der fremkommer fra $(A \mid \boldsymbol{b})$ ved at anvende den valgte følge af rækkeoperationer, er på reduceret echelon form, og at de ledende indgange i $(A' \mid \boldsymbol{b}')$ alle er indholdt i A'. Da $\boldsymbol{b}'$ altså ikke indeholder en ledende indgang, viser sætning 1.2.12 (3)–(4), at ligningssystemet har en løsning. Hvis $r < m$, da er de sidste $m - r$ rækker i A' alle lig med nul-rækken. Lad nu $\boldsymbol{b}' = (b'_i)$, hvor

$$b'_i = \begin{cases} 1 & \text{hvis } i = r + 1, \\ 0 & \text{hvis } i \neq r + 1. \end{cases}$$

Da er matricen $(A' \mid \boldsymbol{b}')$ på reduceret echelon form af rang $r + 1$, og da $\boldsymbol{b}'$ indeholder en ledende indgang, viser sætning 1.2.12 (2), at det tilhørende ligningssystem ingen løsninger har. Ifølge sætning 1.2.3 er det samme derfor tilfældet for ligningssystemet, hvis totalmatrix er den matrix $(A \mid \boldsymbol{b})$, der fås fra $(A' \mid \boldsymbol{b}')$ ved i omvendt rækkefølge at udføre de inverse rækkeoperationer svarende til rækkeoperationerne i den valgte følge af rækkeoperationer. Dette viser (1).

Vi viser dernæst (2). Hvis $r = n$, da indeholder hver søjle i A' en ledende indgang. Lad nu $\boldsymbol{b}$ være et valg af konstantsøjle. Den valgte følge af rækkeoperationer omdanner matricen $(A \mid \boldsymbol{b})$ til en matrix $(A' \mid \boldsymbol{c})$, og vi kan da omdanne $(A' \mid \boldsymbol{c})$ til en matrix $(A' \mid \boldsymbol{b}')$ på reduceret echelon form ved om nødvendigt at anvende yderligere rækkeoperationer. Da hver søjle i A' indeholder en ledende indgang, er rangen r' af $(A' \mid \boldsymbol{b}')$ er enten $r' = r$ eller $r' = r+1$. Hvis $r' = r+1$, da indeholder $\boldsymbol{b}'$ en ledende indgang, hvorfor sætning 1.2.12 (2) viser, at ligningssystemet med totalmatrix $(A \mid \boldsymbol{b})$ ikke har nogen løsninger. Og hvis $r' = r$, da indeholder $\boldsymbol{b}'$ ikke en ledende indgang, hvorfor sætning 1.2.12 (3) viser, at ligningssystemet med totalmatrix $(A \mid \boldsymbol{b})$ præcis har een løsning. Med andre ord gælder der for alle valg af $\boldsymbol{b}$, at ligningssystemet med totalmatrix $(A \mid \boldsymbol{b})$ højst har en løsning. Hvis $r < n$, da gælder der for $\boldsymbol{b} = \boldsymbol{0}$, at ligningssystemet med totalmatrix $(A \mid \boldsymbol{b})$ har uendeligt mange løsninger. For dette følger fra sætning 1.2.12 (4), idet den valgte følge af rækkeoperationer omdanner $(A \mid \boldsymbol{0})$ til matricen $(A' \mid \boldsymbol{0})$, der er på reduceret echelon form af rang $r < n$ og ikke har nogen ledende indgange i konstantsøjlen. Dette viser (2), og endelig følger (3) ved at kombinere (1) og (2). □

Eksempel 1.2.18 Vi betragter 3×3-matricen

$$A = \begin{pmatrix} -1 & -2 & -5 \\ 2 & 3 & 8 \\ 2 & 6 & 14 \end{pmatrix}$$

og undersøger først om, der findes en konstantsøjle $\boldsymbol{b}$ sådan, at ligningssystemet med totalmatrix $(A \mid \boldsymbol{b})$ ikke har nogen løsninger. Vi anvender først rækkeoperationer til at omdanne A til en matrix B på echelon form.

$$\begin{aligned} A = &\begin{pmatrix} -1 & -2 & -5 \\ 2 & 3 & 8 \\ 2 & 6 & 14 \end{pmatrix} \begin{matrix} \\ +2R_1 \\ +2R_1 \end{matrix} \\ &\begin{pmatrix} -1 & -2 & -5 \\ 0 & -1 & -2 \\ 0 & 2 & 4 \end{pmatrix} \begin{matrix} \\ \\ +2R_2 \end{matrix} \\ B = &\begin{pmatrix} -1 & -2 & -5 \\ 0 & -1 & -2 \\ 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

Da B er på echelon form af rang $r = 2 < 3 = m$, konkluderer vi fra Sætning 1.2.17 (1), at en sådan konstantsøjle $\boldsymbol{b}$ findes. Endvidere fortæller beviset for sætningen os, hvordan vi bærer os ad med at finde et sådant $\boldsymbol{b}$:

$$
\begin{aligned}
(B \mid \boldsymbol{c}) = &\left(\begin{array}{rrr|r} -1 & -2 & -5 & 0 \\ 0 & -1 & -2 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right) \begin{array}{l} \\ \\ +(-2)R_2 \end{array} \\
&\left(\begin{array}{rrr|r} -1 & -2 & -5 & 0 \\ 0 & -1 & -2 & 0 \\ 0 & 2 & 4 & 1 \end{array}\right) \begin{array}{l} \\ +(-2)R_1 \\ +(-2)R_1 \end{array} \\
(A \mid \boldsymbol{b}) = &\left(\begin{array}{rrr|r} -1 & -2 & -5 & 0 \\ 2 & 3 & 8 & 0 \\ 2 & 6 & 14 & 1 \end{array}\right)
\end{aligned}
$$

Ligningssystemet med totalmatrix $(B \mid \boldsymbol{c})$ har tydeligvis ingen løsninger, og det samme er derfor tilfældet for ligningssystemet med total matrix $(A \mid \boldsymbol{b})$. Her har vi udført de inverse af de rækkeoperationer, vi benyttede til at omdanne A til B.

2 Matricer og lineære afbildninger

I dette kapitel introducerer og studerer vi reelle og komplekse standard n-dimensionale vektorrum, samt lineære afbildninger imellem disse. I det 1-dimensionale tilfælde er hældningskoefficenten af en lineær afbildning givet ved en skalar, mens den tilsvarende koefficient i højere dimensionale tilfælde i stedet er givet ved en matrix. Vi begynder med en detaljeret introduktion til matricer og deres algebra.

2.1 Matricer

Vi har allerede brugt matricer, så det er nu på høje tid, at vi definerer dem ordentligt. Vi siger synonymt, at en afbildning $x\colon I \to X$ fra en mængde I til en mængde X er en *familie* af elementer i X indiceret ved I, og vi skriver $(x_i)_{i\in I}$, hvor $x_i = x(i)$.

Definition 2.1.1 Lad m og n være naturlige tal. En $m \times n$-matrix med indgange i et legeme $\mathbb{F}$ er en familie $A = (a_{ij})$ af elementer i $\mathbb{F}$ indiceret ved mængden af par (i, j) af naturlige tal $1 \le i \le m$ og $1 \le j \le n$. Vi siger, at elementet a_{ij} er den (i, j)'te indgang i A. Mængden bestående af alle $m \times n$-matricer med indgange i $\mathbb{F}$ betegnes $M_{m,n}(\mathbb{F})$.

Om nødvendigt skriver vi $a_{i,j}$ i stedet for a_{ij}, og vi skriver endvidere

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

for at anskueliggøre matricen $A = (a_{ij}) \in M_{m,n}(\mathbb{F})$. Bemærk, at der ikke er kommaer mellem indgangene i matricen. Givet $A = (a_{ij}) \in M_{m,n}(\mathbb{F})$, kalder vi $1 \times n$-matricen

$$\begin{pmatrix} a_{i1} & a_{i2} & \dots & a_{in} \end{pmatrix} \in M_{1,n}(\mathbb{F})$$

for den i'te *række* i A; og vi kalder $m \times 1$-matricen

$$\begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix} \in M_{m,1}(\mathbb{F})$$

for den j'te *søjle* i A. Hvis A er en $m \times n$-matrix, så siger vi også, at A er en matrix af dimensioner $m \times n$, hvilket læses "m gange n". Vi bemærker, at m og n er antallet af henholdsvis rækker og søjler i A. Hvis $m = n$, så siger vi, at matricen A er *kvadratisk* af orden n, og og vi skriver $M_n(\mathbb{F})$ i stedet for $M_{n,n}(\mathbb{F})$ for mængden kvadratiske matricer af orden n med indgange i $\mathbb{F}$.

Eksempel 2.1.2 Matricerne

$$A = \begin{pmatrix} 2 & -1 & 0 & 3 \\ 3 & 2 & 0 & 0 \end{pmatrix} \in M_{2,4}(\mathbb{R}) \qquad \text{og} \qquad B = \begin{pmatrix} 2i & -1 \\ 0 & 1+i \end{pmatrix} \in M_2(\mathbb{C})$$

er henholdsvis en 2×4-matrix med indgange i $\mathbb{R}$ og en kvadratisk matrix af orden 2 med indgange i $\mathbb{C}$.

Eksempel 2.1.3 (1) *Nulmatricen* af dimensioner $m \times n$ er den matrix

$$O_{m,n} = \begin{pmatrix} 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{pmatrix} \in M_{m,n}(\mathbb{F}),$$

hvis indgange alle er lig med 0. Vi benytter forkortelsen $O = O_{m,n}$, hvis m og n fremgår fra sammenhængen.

(2) Den *modsatte* af en $m \times n$-matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \in M_{m,n}(\mathbb{F})$$

er $m \times n$-matricen

$$-A = \begin{pmatrix} -a_{11} & -a_{12} & \dots & -a_{1n} \\ -a_{21} & -a_{22} & \dots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{m1} & -a_{m2} & \dots & -a_{mn} \end{pmatrix} \in M_{m,n}(\mathbb{F}).$$

(3) *Identitetsmatricen* $I_n \in M_n(\mathbb{F})$ er den kvadratiske matrix af orden n, hvis (i,j)'te indgang er Kronecker's δ_{ij}, som er enten 1 eller 0, eftersom $i = j$ eller $i \neq j$.

$$I_1 = (1), \quad I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \dots \quad I_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$

Vi forkorter ofte $I = I_n$, hvis n fremgår fra sammenhængen.

Vi indfører nu de fundamentale aritmetiske operationer på matricer.

Definition 2.1.4 (1) *Summen* af to matricer

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

der begge har dimensioner $m \times n$, er matricen

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

der ligeledes har dimensioner $m \times n$.
(2) *Produktet* af to matricer

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \qquad \text{og} \qquad B = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1p} \\ b_{21} & b_{22} & \dots & b_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \dots & b_{np} \end{pmatrix}$$

af dimensioner henholdsvis $m \times n$ og $n \times p$, er matricen

$$A \cdot B = \begin{pmatrix} c_{11} & c_{12} & \dots & c_{1p} \\ c_{21} & c_{22} & \dots & c_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \dots & c_{mp} \end{pmatrix}$$

hvis (i,k)'te indgang er defineret ved

$$c_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk} \;=\; a_{i1} b_{1k} + a_{i2} b_{2k} + \dots + a_{in} b_{nk}.$$

Produktmatricen $A \cdot B$ har dimensioner $m \times p$.

Vi understreger, at summen af to matricer A og B kun er defineret, hvis A og B har de samme dimensioner; og at produktet af to matricer A og B kun er defineret, hvis antallet af søjler i A er lig antallet af rækker i B. Det har således ikke mening at tale om $A + B$ eller AB, medmindre dimensionerne af A og B er som foreskrevet. Som for multiplikation i $\mathbb{F}$, forkorter vi normalt og skriver AB i stedet for $A \cdot B$.

Eksempel 2.1.5 Summen af matricerne

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

findes, da de to matricer har samme dimensioner 2×3, og er givet ved

$$\begin{aligned} A + B &= \begin{pmatrix} 2 & 0 & 4 \\ 1 & -3 & 6 \end{pmatrix} + \begin{pmatrix} 1 & -7 & 4 \\ 1 & 2 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 2+1 & 0+(-7) & 4+4 \\ 1+1 & -3+2 & 6+0 \end{pmatrix} \\ &= \begin{pmatrix} 3 & -7 & 8 \\ 2 & -1 & 6 \end{pmatrix}. \end{aligned}$$

Her har vi markeret indgangene i A og B med henholdsvis blåt og sort.

Eksempel 2.1.6 Produktet af matricerne

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

findes, da antallet søjler i A er lig med antallet af rækker i B, og er givet ved

$$\begin{aligned} A \cdot B &= \begin{pmatrix} 2 & 7 \\ -1 & 3 \end{pmatrix} \cdot \begin{pmatrix} 2 & -1 & 0 \\ 4 & 1 & 2 \end{pmatrix} \\ &= \begin{pmatrix} 2 \cdot 2 + 7 \cdot 4 & 2 \cdot (-1) + 7 \cdot 1 & 2 \cdot 0 + 7 \cdot 2 \\ (-1) \cdot 2 + 3 \cdot 4 & (-1) \cdot (-1) + 3 \cdot 1 & (-1) \cdot 0 + 3 \cdot 2 \end{pmatrix} \\ &= \begin{pmatrix} 32 & 5 & 14 \\ 10 & 4 & 6 \end{pmatrix}. \end{aligned}$$

Her har vi markeret indgangene i A og B med henholdsvis blåt og sort. Vi bemærker, at produktet $B \cdot A$ ikke er defineret, da antallet af søjler i B er forskelligt fra antallet af rækker i A.

Vi bemærker specielt, at givet en $1 \times n$-matrix og en $n \times 1$-matrix

$$A = \begin{pmatrix} a_1 & a_2 & \dots & a_n \end{pmatrix} \qquad \text{og} \qquad B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix},$$

da er produktet AB defineret og givet ved 1×1-matricen

$$AB = \begin{pmatrix} a_1 b_1 + a_2 b_2 + \cdots + a_n b_n \end{pmatrix}.$$

Generelt, givet en $m \times n$-matrix A og en $n \times p$-matrix B, kan vi da betragte deres produkt som værende den $m \times p$-matricen $AB = C$, hvis indgange er alle de mulige produkter af en række i A med en søjle i B:

$$\begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & & & \vdots \\ \boxed{a_{i1} \; \cdots \; a_{ij} \; \cdots \; a_{in}} & & & & \\ \vdots & & & & \vdots \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} b_{11} & \cdots & b_{1k} & \cdots & b_{1p} \\ \vdots & & \vdots & & \vdots \\ b_{j1} & & b_{jk} & & b_{jp} \\ \vdots & & \vdots & & \vdots \\ b_{n1} & \cdots & b_{nk} & \cdots & b_{np} \end{pmatrix} = \begin{pmatrix} c_{11} & \cdots & c_{1k} & \cdots & c_{1p} \\ \vdots & & \vdots & & \vdots \\ c_{i1} & \cdots & \boxed{c_{ik}} & \cdots & c_{ip} \\ \vdots & & \vdots & & \vdots \\ c_{m1} & \cdots & c_{mk} & \cdots & c_{mp} \end{pmatrix}$$

Da der er m rækker i A og p søjler i B, er der altså $m \times p$ indgange i $AB = C$.

Eksempel 2.1.7 Produktmatricen

$$AB = \begin{pmatrix} \boxed{1 \;\; 0 \;\; 3} \end{pmatrix} \begin{pmatrix} \boxed{\begin{matrix} 2 \\ 4 \\ 1 \end{matrix}} \end{pmatrix} = \begin{pmatrix} 5 \end{pmatrix}$$

er en 1×1-matrix, da der er 1 række i A og 1 søjle i B, mens produktmatricen

$$BA = \begin{pmatrix} \boxed{2} \\ \boxed{4} \\ \boxed{1} \end{pmatrix} \begin{pmatrix} \boxed{1} & \boxed{0} & \boxed{3} \end{pmatrix} = \begin{pmatrix} 2 \cdot 1 & 2 \cdot 0 & 2 \cdot 3 \\ 4 \cdot 1 & 4 \cdot 0 & 4 \cdot 3 \\ 1 \cdot 1 & 1 \cdot 0 & 1 \cdot 3 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 6 \\ 4 & 0 & 12 \\ 1 & 0 & 3 \end{pmatrix}$$

er en 3×3-matrix, da der er 3 rækker i B og 3 søjler i A. Vi bemærker, at det ikke er muligt at sammenligne matricerne AB og BA, da deres dimensioner er forskellige. Så vi kan ikke engang stille spørgsmålet, om hvorvidt $AB = BA$.

Eksempel 2.1.8 I kapitel 1 har vi betragtet lineære ligningssystemer

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

med $a_{ij}, b_i \in \mathbb{F}$. Vi betragter nu matricerne

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \qquad \boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \qquad \text{og} \qquad \boldsymbol{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix},$$

som har dimensioner henholdvis $m \times n$, $n \times 1$ og $m \times 1$. Produktmatricen

$$A\boldsymbol{x} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix}$$

er derfor en $m \times 1$-matrix ligesom $\boldsymbol{b}$. Det oprindelige lineære ligningssystem kan altså ækvivalent skrives som ligning af $m \times 1$-matricer

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

og vi vil i det følgende hovedsageligt benytte denne korte skrivemåde.

Sætning 2.1.9 *For matricer af passende dimensioner gælder følgende:*

(A1) $(A + B) + C = A + (B + C)$

(A2) $A + O = A = O + A$

(A3) $A + (-A) = O = (-A) + A$

(A4) $A + B = B + A$

(P1) $(AB)C = A(BC)$

(P2) $AI_n = A = I_m A$

(D1) $A(B + C) = AB + AC$

(D2) $(A + B)C = AC + BC$

(E1) $AO_{n,n} = O_{m,n} = O_{m,m}A$

(E2) $A(-I_n) = -A = (-I_m)A$

Bevis Vi ved fra definition 0.1.1 og sætning 0.1.3, at de tilsvarende identiteter for addition og multiplikation af skalærer gælder, og vi benytter nu dette til at vise, at identiteterne også gælder for addition og multiplikation af matricer. Identiteterne (A1)–(A4) omhandler alle matricer af samme dimensioner og følger umiddelbart fra de tilsvarende identiteter (A1)–(A4) for skalarer, idet addition for matricer per definition er givet ved addition i $\mathbb{F}$ af de respektive indgange.

Identiteten (P1) omhandler matricer A, B og C af dimensioner $m \times n$, $n \times p$ og $p \times q$. Produkt matricerne $(AB)C$ og $A(BC)$ er begge af dimensioner $m \times q$, og vi ønsker at vise, at deres (i,l)'te indgange er identiske. På den ene side er den (i,l)'te indgang i $(AB)C$ lig med $\sum_{k=1}^{p} e_{ik}c_{kl}$, hvor $e_{ik} = \sum_{j=1}^{n} a_{ij}b_{jk}$ er den (i,k)'te indgang i AB; og på den anden side er den (i,l)'te indgang i $A(BC)$ lig med $\sum_{j=1}^{n} a_{ij}d_{jl}$, hvor $d_{jl} = \sum_{k=1}^{p} b_{jk}c_{kl}$ er den (j,l)te indgang i BC. Vi skal derfor vise, at der gælder følgende identitet af skalarer:

$$\sum_{k=1}^{p}\Big(\sum_{j=1}^{n} a_{ij}b_{jk}\Big)c_{kl} = \sum_{j=1}^{n} a_{ij}\Big(\sum_{k=1}^{p} b_{jk}c_{kl}\Big).$$

For at være helt præcis, så definerer vi her de itererede summer til at være

$$\sum_{j=1}^{n} a_{ij}b_{jk} = (\dots((a_{i1}b_{1k} + a_{i2}b_{2k}) + a_{i3}b_{3k}) + \dots + a_{in}b_{nk}),$$

etc. Det følger dog fra (A1) for skalarer, at ethvert valg af summationsorden vil tilordne den samme værdi til den itererede sum. Identiteten mellem de (i,l)'te indgange i $(AB)C$ og $A(BC)$ fås nu som følger:

$$\begin{aligned}\sum_{k=1}^{p}\Big(\sum_{j=1}^{n} a_{ij}b_{jk}\Big)c_{kl} &= \sum_{k=1}^{p}\sum_{j=1}^{n}(a_{ij}b_{jk})c_{kl} = \sum_{k=1}^{p}\sum_{j=1}^{n} a_{ij}(b_{jk}c_{kl}) \\ &= \sum_{j=1}^{n}\sum_{k=1}^{p} a_{ij}(b_{jk}c_{kl}) = \sum_{j=1}^{n} a_{ij}\Big(\sum_{k=1}^{p} b_{jk}c_{kl}\Big)\end{aligned}$$

Her følger den første identitet fra (D2) for skalarer; den anden identitet fra (P1) for skalarer; og den tredie identitet fra (A1) og (A4) for skalarer; og den sidste identitet fra (D1) for skalarer. Dette beviser, at (P1) gælder for matricer.

Vi beviser dernæst identiteten $AI_n = A$ i (P2); beviset for identiteten $I_mA = A$ er helt tilsvarende. Identiteten $AI_n = A$ omhandler en matrix A af dimensioner $m \times n$ og identitetsmatricen I_n af dimensioner $n \times n$. Vi husker på fra eksempel 2.1.3(3), at den (j,k)'te indgang δ_{jk} i I_n er lig 1, hvis $j = k$, og lig med 0, hvis $j \neq k$. På (i,k)'te indgange svarer identiteten $AI_n = A$ nu til følgende identitet af skalarer:

$$\sum_{j=1}^{n} a_{ij}\delta_{jk} = a_{ik}.$$

så identiteten følger fra (A2), (P2) og (E1) for skalarer. Vi har hermed bevist, at (P2) gælder for matricer.

Vi beviser dernæst identiteten (D1); identiteten (D2) bevises helt tilsvarende. Denne identitet omhandler en $m \times n$-matrix A og $n \times p$-matricer B og C og svarer på (i,k)'te indgange til følgende identitet af skalarer:

$$\sum_{j=1}^{n} a_{ij}(b_{jk} + c_{jk}) = \sum_{j=1}^{n} a_{ij}b_{jk} + \sum_{j=1}^{n} a_{ij}c_{jk}.$$

Denne identitet følger umiddelbart fra (A1), (A4) og (D1) for skalarer, hvilket beviser, at (D1) også gælder for matricer.

Endelig fås identiteterne (E1) og (E2) fra de øvrige identiteter ved at repetere beviset for sætning 0.1.3. Vi har hermed bevist sætningen. □

Eksempel 2.1.10 Vi betragter 2×2-matricerne

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

og udregner de to produkter AB og BA, der begge findes og er af dimensioner 2×2.

$$AB = \begin{pmatrix} \boxed{1 \quad 1} \\ \boxed{0 \quad 1} \end{pmatrix} \begin{pmatrix} \boxed{\begin{matrix} 1 \\ 1 \end{matrix}} & \boxed{\begin{matrix} 0 \\ 1 \end{matrix}} \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$$

$$BA = \begin{pmatrix} \boxed{1 \quad 0} \\ \boxed{1 \quad 1} \end{pmatrix} \begin{pmatrix} \boxed{\begin{matrix} 1 \\ 0 \end{matrix}} & \boxed{\begin{matrix} 1 \\ 1 \end{matrix}} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$$

Da matricerne AB og BA er af samme dimensioner, har det mening at spørge, om de er ens. Men udregningen ovenfor viser, at $AB \neq BA$.

2.2 Vektorrummet $\mathbb{F}^m$

Vi lader nu igen $\mathbb{F}$ være et legeme, for eksempel $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, og definerer

$$\mathbb{F}^m = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \,\middle|\, x_1, \dots, x_m \in \mathbb{F} \right\}$$

til at være mængden $M_{m,1}(\mathbb{F})$ af $m \times 1$-matricer med indgange i $\mathbb{F}$, som vi også kalder for mængden af *søjlevektorer* i $\mathbb{F}$ af dimension m. Vi vil tilsvarende kalde mængden $M_{1,n}(\mathbb{F})$ af $1 \times n$-matricer med indgange i $\mathbb{F}$ for mængden af *rækkevektorer* i $\mathbb{F}$ af dimension n, men vi vil ikke indføre nogen speciel notation for denne mængde.

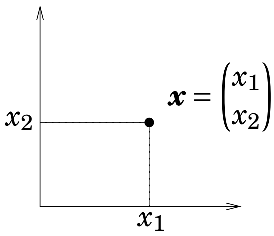
Eksempel 2.2.1 Vi kan visuallisere

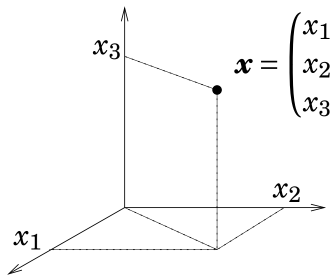
$$\mathbb{R}^2 = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \,\middle|\, x_1, x_2 \in \mathbb{R} \right\} \qquad \text{og} \qquad \mathbb{R}^3 = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \,\middle|\, x_1, x_2, x_3 \in \mathbb{R} \right\}$$

som mængden af punkter i henholdsvis planen og rummet; se figur 2.2.1.

Vi har allerede defineret sum og produkt af matricer, og disse regneoperationer kan specielt anvendes på søjlevektorer som følger. Givet to søjlevektorer

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \qquad \text{og} \qquad \boldsymbol{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$$





Figur 2.1: De reelle vektorrum $\mathbb{R}^2$ og $\mathbb{R}^3$.

i $\mathbb{F}$ af samme dimension m, da er deres sum givet ved

$$\boldsymbol{x} + \boldsymbol{y} = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_m + y_m \end{pmatrix},$$

hvilken igen er en søjlevektor i $\mathbb{F}$ af dimension m.

Vi husker dernæst på, at hvis vi ganger en $m \times 1$-matrix med en 1×1-matrix, så får vi igen en $m \times 1$-matrix. Vi kalder denne operation for *skalar multiplikation*. Som det er sædvane, vil vi også misbruge notation og identificere en 1×1-matrix $(a) \in M_{1,1}(\mathbb{F})$ med dens indgang $a \in \mathbb{F}$. Så det skalare multiplum af en søjlevektor

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \in \mathbb{F}^m$$

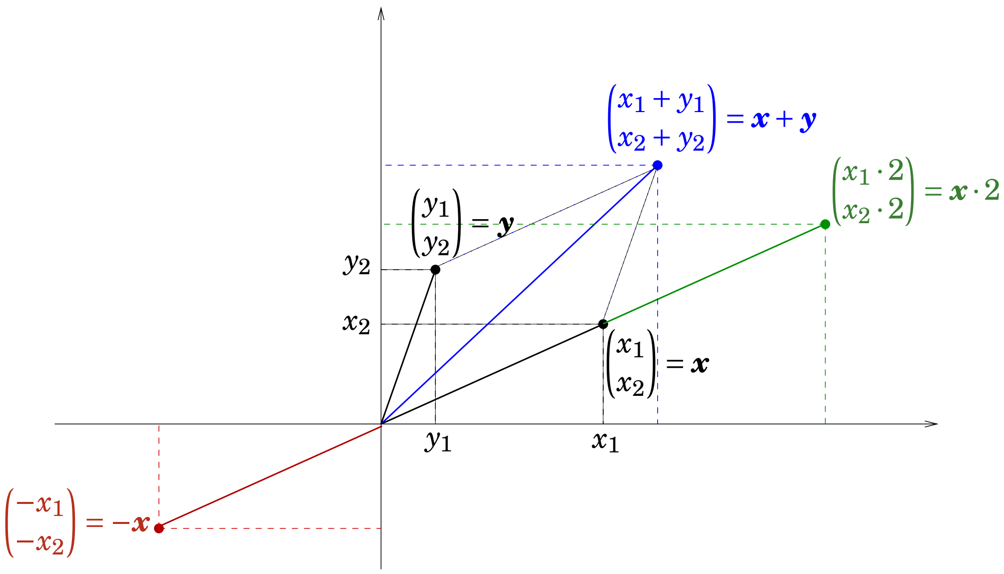
med en skalar $a \in \mathbb{F}$ er søjlevektoreren

$$\boldsymbol{x} \cdot a = \begin{pmatrix} x_1 a \\ \vdots \\ x_m a \end{pmatrix} \in \mathbb{F}^m.$$

Vi forkorter normalt og skriver $\boldsymbol{x}a$ i stedet for $\boldsymbol{x} \cdot a$. Vi understreger, at det omvendte produkt $a \cdot \boldsymbol{x}$ ikke har mening som et matrix produkt, medmindre $m = 1$, da antallet af søjler i a ikke er lig med antallet af rækker i $\boldsymbol{x}$. Endvidere er den *modsatte vektor* af $\boldsymbol{x}$ og nulvektoren $\mathbf{0}$ givet ved henholdsvis

$$-\boldsymbol{x} = \begin{pmatrix} -x_1 \\ \vdots \\ -x_m \end{pmatrix} \qquad \text{og} \qquad \mathbf{0} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Vi henviser til figur 2.2 for en illustration af disse aritmetiske operationer.



Figur 2.2: Addition, skalar multiplikation og modsat vektor i $\mathbb{R}^2$.

Sætning 2.1.9 ovenfor specialiserer nu til den følgende sætning, der udtrykker, at mængden $\mathbb{F}^m$ sammen med operationerne vektorsum og skalarprodukt udgør, hvad vi senere vil kalde et $\mathbb{F}$-vektorrum; se definition 4.1.1.

Sætning 2.2.2 *For søjlevektorer $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathbb{F}^m$ og skalarer $a, b \in \mathbb{F}$ gælder følgende:*

(A1) $(\boldsymbol{x}+\boldsymbol{y})+\boldsymbol{z} = \boldsymbol{x}+(\boldsymbol{y}+\boldsymbol{z})$

(A2) $\boldsymbol{x}+\boldsymbol{0} = \boldsymbol{x} = \boldsymbol{0}+\boldsymbol{x}$

(A3) $\boldsymbol{x}+(-\boldsymbol{x}) = \boldsymbol{0} = (-\boldsymbol{x})+\boldsymbol{x}$

(A4) $\boldsymbol{x}+\boldsymbol{y} = \boldsymbol{y}+\boldsymbol{x}$

(P1) $(\boldsymbol{x} \cdot a) \cdot b = \boldsymbol{x} \cdot (a \cdot b)$

(P2) $\boldsymbol{x} \cdot 1 = \boldsymbol{x}$

(D1) $\boldsymbol{x} \cdot (a+b) = \boldsymbol{x} \cdot a + \boldsymbol{x} \cdot b$

(D2) $(\boldsymbol{x}+\boldsymbol{y}) \cdot a = \boldsymbol{x} \cdot a + \boldsymbol{y} \cdot a$

(E1) $\boldsymbol{x} \cdot 0 = \boldsymbol{0}$

(E2) $\boldsymbol{x} \cdot (-1) = -\boldsymbol{x}$

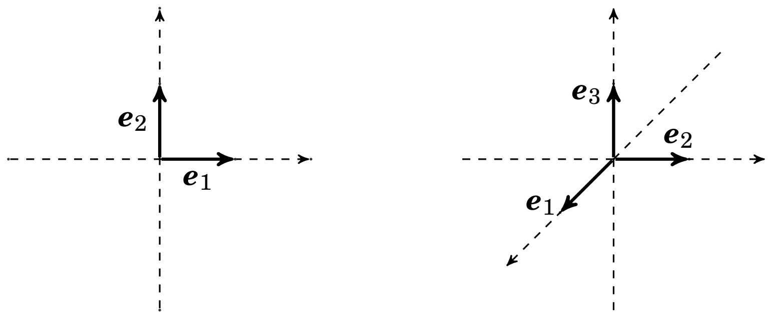
Vi siger, at en familie af elementer i en mængde X indiceret ved $I = \{1, 2, \dots, m\}$ er en *m-tuple* af elementer i X og skriver $(x_1, x_2, \dots, x_m)$ i stedet for $(x_i)_{i \in I}$.

Definition 2.2.3 Hvis $\mathbb{F}$ er et legeme, og hvis m et naturligt tal, da kaldes m-tuplen $(\boldsymbol{e}_1, \dots, \boldsymbol{e}_m)$ af vektorer i $\mathbb{F}^m$, hvor

$$\boldsymbol{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ \boldsymbol{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ \dots, \ \boldsymbol{e}_m = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix},$$

for *standardbasen* for $\mathbb{F}^m$.

Vektorerne $\boldsymbol{e}_1, \dots, \boldsymbol{e}_m$ kaldes for *standardenhedsvektorerne* i $\mathbb{F}^m$. Den følgende figur illustrerer disse vektorer for $\mathbb{F} = \mathbb{R}$.



Figur 2.3: Standardenhedsvektorer i $\mathbb{R}^2$ og $\mathbb{R}^3$.

Lad $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_r)$ være en r-tuple af vektorer i $\mathbb{F}^m$. Vi siger, at en vektor $\boldsymbol{u}$ er en *linearkombination* af $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_r)$, hvis der findes skalarer $a_1, \dots, a_r \in \mathbb{F}$, sådan at

$$\boldsymbol{u} = \boldsymbol{u}_1 a_1 + \boldsymbol{u}_2 a_2 + \dots + \boldsymbol{u}_r a_r.$$

Specielt er nulvektoren $\mathbf{0}$ en linearkombination af 0-tuplen ( ), idet vi vedtager, at en tom sum er lig med $\mathbf{0}$.

Eksempel 2.2.4 Vi betragter de følgende tre vektorer i $\mathbb{F}^3$:

$$\boldsymbol{u} = \begin{pmatrix} 2 \\ 5 \\ 3 \end{pmatrix}, \qquad \boldsymbol{u}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \qquad \text{og} \qquad \boldsymbol{u}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}.$$

Vi ser da, at $\boldsymbol{u}$ er en linearkombination af familien $(\boldsymbol{u}_1, \boldsymbol{u}_2)$, idet

$$\boldsymbol{u} = \begin{pmatrix} 2 \\ 5 \\ 3 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 3 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \cdot 2 + \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \cdot 3 = \boldsymbol{u}_1 \cdot 2 + \boldsymbol{u}_2 \cdot 3.$$

Lemma 2.2.5 *Lad $\mathbb{F}$ være et legeme og lad m være et ikke-negativt heltal. Enhver vektor $\boldsymbol{x} \in \mathbb{F}^m$ kan på entydig vis udtrykkes som en linear kombination*

$$\boldsymbol{x} = \boldsymbol{e}_1 a_1 + \boldsymbol{e}_2 a_2 + \cdots + \boldsymbol{e}_m a_m$$

af standardbasen $(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_m)$.

Bevis Givet en vilkårlig vektor

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} \in \mathbb{F}^m,$$

ønsker vi at undersøge, om der findes $a_1, a_2, \ldots, a_m \in \mathbb{F}$, sådan at

$$\boldsymbol{x} = \boldsymbol{e}_1 a_1 + \boldsymbol{e}_2 a_2 + \cdots + \boldsymbol{e}_m a_m.$$

Vi udregner derfor højresiden, som er lig med

$$\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} a_1 + \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} a_2 + \cdots + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} a_m = \begin{pmatrix} a_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ a_2 \\ \vdots \\ 0 \end{pmatrix} + \cdots + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ a_m \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix},$$

hvoraf vi ser, at $a_1 = x_1$, $a_2 = x_2$, …, $a_m = x_m$ er den entydige løsning til dette problem. □

2.3 Lineære afbildninger

Vi lader igen $\mathbb{F}$ være et legeme, for eksempel $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, og betragter vektorrummet

$$\mathbb{F}^n = M_{n,1}(\mathbb{F})$$

bestående af $n \times 1$-matricer med indgange i $\mathbb{F}$. Hvis nu $A \in M_{m,n}(\mathbb{F})$ er en $m \times n$-matrix med indgange i $\mathbb{F}$, så kan vi for alle $\boldsymbol{x} \in \mathbb{F}^n$ danne produktet $A\boldsymbol{x} \in M_{m,1}(\mathbb{F}) = \mathbb{F}^m$. På denne måde giver $m \times n$-matricen $A \in M_{m,n}(\mathbb{F})$ anledning til en afbildning

$$f : \mathbb{F}^n \to \mathbb{F}^m$$

defineret ved $f(\boldsymbol{x}) = A\boldsymbol{x}$. Matricen $A \in M_{m,n}(\mathbb{F})$ giver altså på denne måde anledning til afbildningen $f\colon \mathbb{F}^n \to \mathbb{F}^m$, der til en søjlevektor af dimension n tilordner en søjlevektor af dimension m. Vi skal i dette afsnit vise, at de afbildninger, der fremkommer således, præcis er de *lineære* afbildninger, som vi nu definerer.

Definition 2.3.1 Lad $\mathbb{F}$ være et legeme. En afbildning $f\colon \mathbb{F}^n \to \mathbb{F}^m$ er lineær, hvis den opfylder følgenge betingelser (L1)–(L2) for alle $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{F}^n$ og $a \in \mathbb{F}$.

(L1) $f(\boldsymbol{x}+\boldsymbol{y}) = f(\boldsymbol{x}) + f(\boldsymbol{y})$.

(L2) $f(\boldsymbol{x}\cdot a) = f(\boldsymbol{x})\cdot a$.

Enhver lineær afbildning $f\colon \mathbb{F}^n \to \mathbb{F}^m$ afbilder nødvendigvis $\mathbf{0} \in \mathbb{F}^n$ til $\mathbf{0} \in \mathbb{F}^m$. For

$$f(\mathbf{0}) = f(\mathbf{0}\cdot 0) = f(\mathbf{0})\cdot 0 = \mathbf{0},$$

hvor den første og sidste identitet fås fra (E1) i sætning 2.2.2, mens den midterste identitet er (L2). Ved at anvende (E2) i sætning 2.2.2 samt (L2), ser vi tilsvarende, at

$$f(-\boldsymbol{x}) = f(\boldsymbol{x}\cdot(-1)) = f(\boldsymbol{x})\cdot(-1) = -f(\boldsymbol{x}).$$

Vi viser nu det følgende mere generelle resultat.

Lemma 2.3.2 *Lad $\mathbb{F}$ være et legeme, og lad $f\colon \mathbb{F}^n \to \mathbb{F}^m$ være en lineær afbildning. Givet vektorer $\boldsymbol{x}_1,\dots,\boldsymbol{x}_k \in \mathbb{F}^n$ og skalarer $a_1,\dots,a_k \in \mathbb{F}$, da gælder, at*

$$f(\boldsymbol{x}_1a_1 + \boldsymbol{x}_2a_2 + \cdots + \boldsymbol{x}_ka_k) = f(\boldsymbol{x}_1)a_1 + f(\boldsymbol{x}_2)a_2 + \cdots + f(\boldsymbol{x}_k)a_k.$$

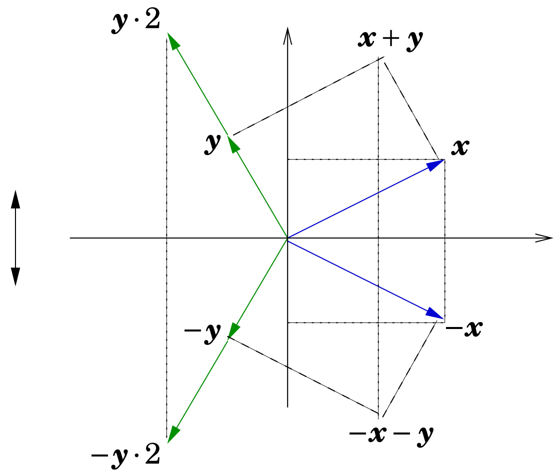
Bevis Vi beviser påstanden ved induktion på $k \geq 0$. I tilfældet $k = 0$, da er påstanden, at $f(\mathbf{0}) = \mathbf{0}$, hvilket vi allerede har vist ovenfor. For den tomme sum er per definition lig med nulvektoren. Så vi antager, at påstanden allerede er bevist for $k = r-1$ og beviser den for $k = r$. Hertil udregner vi

$$\begin{aligned} &f(\boldsymbol{x}_1a_1 + \cdots + \boldsymbol{x}_{r-1}a_{r-1} + \boldsymbol{x}_ra_r) = f(\boldsymbol{x}_1a_1 + \cdots + \boldsymbol{x}_{r-1}a_{r-1}) + f(\boldsymbol{x}_ra_r) \\ &= f(\boldsymbol{x}_1a_1 + \cdots + \boldsymbol{x}_{r-1}a_{r-1}) + f(\boldsymbol{x}_r)a_r = f(\boldsymbol{x}_1)a_1 + \cdots + f(\boldsymbol{x}_{r-1})a_{r-1} + f(\boldsymbol{x}_r)a_r, \end{aligned}$$

hvor de tre identiteter fås fra henholdsvis (L1) med $\boldsymbol{x} = \boldsymbol{x}_1a_1 + \cdots + \boldsymbol{x}_{r-1}a_{r-1}$ og $\boldsymbol{y} = \boldsymbol{x}_ra_r$, fra (L2) med $\boldsymbol{x} = \boldsymbol{x}_r$ og $a = a_r$, og fra den induktive hypotese. Dette viser induktionsskridtet og dermed lemmaet. □

Eksempel 2.3.3 Lad $\mathbb{F} = \mathbb{R}$ og lad $f\colon \mathbb{R} \to \mathbb{R}$ være afbildningen defineret ved $f(x) = 2\cdot x$. Denne afbildning er lineær, idet

$$\begin{aligned} f(x+y) &= 2\cdot(x+y) \overset{(D1)}{=} 2\cdot x + 2\cdot y = f(x) + f(y), \\ f(x\cdot a) &= 2\cdot(x\cdot a) \overset{(P1)}{=} (2\cdot x)\cdot a = f(x)\cdot a, \end{aligned}$$



Figur 2.4: Spejling i x–aksen i $\mathbb{R}^2$ er en lineær afbildning

hvilket viser, at $f:\mathbb{R}\to\mathbb{R}$ opfylder både (L1) og (L2). Vi bemærker, at vi kun har brugt, at $\mathbb{R}$ er et legeme, og at $2\in\mathbb{R}$. Den samme udregning viser mere generelt, at hvis $\mathbb{F}$ er et legeme og $c\in\mathbb{F}$, så er afbildningen $f:\mathbb{F}\to\mathbb{F}$ defineret ved $f(\boldsymbol{x})=c\cdot\boldsymbol{x}$ lineær.

Bemærkning 2.3.4 Afbildningen $g:\mathbb{R}\to\mathbb{R}$ defineret ved $g(x)=2\cdot x+3$ er *ikke* lineær, da for eksempel $g(0)\neq 0$. En afbildning som denne, der er en sum af en lineær afbildning og en konstant afbildning, siges at være en *affin* afbildning. Vi bemærker, at enhver affin afbildning $g:\mathbb{R}\to\mathbb{R}$ er på formen $g(x)=a\cdot x+b$ for passende $a,b\in\mathbb{R}$.

Eksempel 2.3.5 (Spejling) Vi betragter den afbildning $s:\mathbb{R}^2\to\mathbb{R}^2$, der er givet ved spejling i x-aksen i $\mathbb{R}^2$; se figur 2.4. Afbildningen er eksplicit givet ved

$$s\begin{pmatrix}x_1\\x_2\end{pmatrix}=\begin{pmatrix}x_1\\-x_2\end{pmatrix},$$

og den følgende udregning viser, at den opfylder (L1) og (L2):

$$\begin{aligned}
s\left(\begin{pmatrix}x_1\\x_2\end{pmatrix}+\begin{pmatrix}y_1\\y_2\end{pmatrix}\right)&=s\begin{pmatrix}x_1+y_1\\x_2+y_2\end{pmatrix}=\begin{pmatrix}x_1+y_1\\-(x_2+y_2)\end{pmatrix}\\
&=\begin{pmatrix}x_1+y_1\\-x_2-y_2\end{pmatrix}=\begin{pmatrix}x_1\\-x_2\end{pmatrix}+\begin{pmatrix}y_1\\-y_2\end{pmatrix}=s\begin{pmatrix}x_1\\x_2\end{pmatrix}+s\begin{pmatrix}y_1\\y_2\end{pmatrix}\\
s\left(\begin{pmatrix}x_1\\x_2\end{pmatrix}\cdot a\right)&=s\begin{pmatrix}x_1\cdot a\\x_2\cdot a\end{pmatrix}=\begin{pmatrix}x_1\cdot a\\-(x_2\cdot a)\end{pmatrix}=\begin{pmatrix}x_1\cdot a\\(-1)\cdot(x_2\cdot a)\end{pmatrix}\\
&=\begin{pmatrix}x_1\cdot a\\((-1)\cdot x_2)\cdot a\end{pmatrix}=\begin{pmatrix}x_1\\-x_2\end{pmatrix}\cdot a=s\begin{pmatrix}x_1\\x_2\end{pmatrix}\cdot a
\end{aligned}$$

Figur 2.4 illustrerer (L1) og (L2) for spejling i x-aksen.

Vores udregning i eksempel 2.3.3 generaliserer umiddelbart til følgende resultat.

Sætning 2.3.6 *Hvis $A \in M_{m,n}(\mathbb{F})$ er en $m \times n$-matrix med indgange i et legeme $\mathbb{F}$, da er afbildningen $f : \mathbb{F}^n \to \mathbb{F}^m$ defineret ved $f(\boldsymbol{x}) = A\boldsymbol{x}$ en lineær afbildning.*

Bevis Det følger fra sætning 2.1.9, at $f : \mathbb{F}^n \to \mathbb{F}^m$ er lineær. For (D1) viser, at

$$f(\boldsymbol{x}+\boldsymbol{y}) = A(\boldsymbol{x}+\boldsymbol{y}) = A\boldsymbol{x}+A\boldsymbol{y} = f(\boldsymbol{x})+f(\boldsymbol{y}),$$

mens (P1) tilsvarende viser, at

$$f(\boldsymbol{x}a) = A(\boldsymbol{x}a) = (A\boldsymbol{x})a = f(\boldsymbol{x})a,$$

så afbildningen $f : \mathbb{F}^n \to \mathbb{F}^m$ opfylder både (L1) og (L2). $\square$

Eksempel 2.3.7 Multiplikation med 2×2-matricen

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

giver anledning til afbildningen $f : \mathbb{F}^2 \to \mathbb{F}^2$ givet ved

$$f\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \left(\begin{array}{|cc|} \hline 2 & -1 \\ \hline 6 & 3 \\ \hline \end{array}\right)\left(\begin{array}{|c|} \hline x_1 \\ x_2 \\ \hline \end{array}\right) = \begin{pmatrix} 2x_1 - x_2 \\ 6x_1 + 3x_2 \end{pmatrix},$$

som derfor ifølge sætning 2.3.6 er lineær.

Vi skal nu omvendt vise, at *enhver* lineær afbildning fra $\mathbb{F}^n$ til $\mathbb{F}^m$ fremkommer ved multiplikation $m \times n$-matrix. Man kan tænke på denne matrix som den generaliserede hældningskoefficient af den lineære afbildning.

Sætning 2.3.8 *Lad $\mathbb{F}$ være et legeme, lad $f : \mathbb{F}^n \to \mathbb{F}^m$ være en lineær afbildning, og lad*

$$A = \begin{pmatrix} f(\boldsymbol{e}_1) & f(\boldsymbol{e}_2) & \cdots & f(\boldsymbol{e}_n) \end{pmatrix} \in M_{m,n}(\mathbb{F})$$

være den matrix, hvis j'te søjle er søjlevektoren $f(\boldsymbol{e}_j) \in \mathbb{F}^m$. Da gælder, at

$$f(\boldsymbol{x}) = A\boldsymbol{x}$$

for alle $\boldsymbol{x} \in \mathbb{F}^n$.

Bevis Vi skriver

$$f(\boldsymbol{e}_j) = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix} \in \mathbb{F}^m,$$

sådan at

$$A = \begin{pmatrix} f(\boldsymbol{e}_1) & f(\boldsymbol{e}_2) & \cdots & f(\boldsymbol{e}_n) \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \in M_{m,n}(\mathbb{F}).$$

Ifølge lemma 2.2.5 kan vi skrive $\boldsymbol{x} \in \mathbb{F}^n$ entydigt som

$$\boldsymbol{x} = \boldsymbol{e}_1 x_1 + \boldsymbol{e}_2 x_2 + \cdots + \boldsymbol{e}_n x_n = \sum_{j=1}^{n} \boldsymbol{e}_j x_j,$$

og vi anvender da den lineære afbildning $f : \mathbb{F}^n \to \mathbb{F}^m$ på denne ligning og får, at

$$f(\boldsymbol{x}) = f(\sum_{j=1}^{n} \boldsymbol{e}_j x_j) = \sum_{j=1}^{n} f(\boldsymbol{e}_j) x_j$$

ifølge lemma 2.3.2. På samme måde kan $f(\boldsymbol{e}_j) \in \mathbb{F}^m$ skrives entydigt som

$$f(\boldsymbol{e}_j) = \boldsymbol{e}_1 a_{1j} + \boldsymbol{e}_2 a_{2j} + \cdots + \boldsymbol{e}_m a_{mj} = \sum_{i=1}^{m} \boldsymbol{e}_i a_{ij},$$

hvilket vi substituerer i udtrykket for $f(\boldsymbol{x})$ overfor. Herved får vi

$$f(\boldsymbol{x}) = \sum_{j=1}^{n} (\sum_{i=1}^{m} \boldsymbol{e}_i a_{ij}) x_j = \sum_{j=1}^{n} \sum_{i=1}^{m} (\boldsymbol{e}_i a_{ij} x_j) = \sum_{i=1}^{m} \sum_{j=1}^{n} (\boldsymbol{e}_i a_{ij} x_j) = \sum_{i=1}^{m} \boldsymbol{e}_i (\sum_{j=1}^{n} a_{ij} x_j),$$

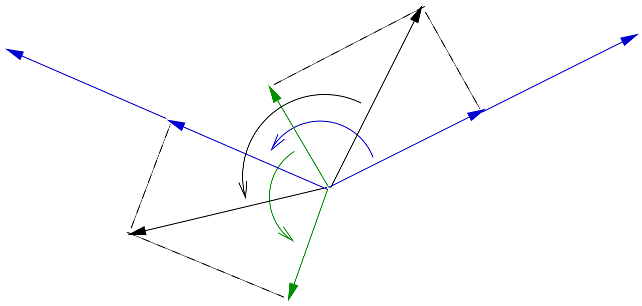
hvilket viser, at

$$f(\boldsymbol{x}) = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = A\boldsymbol{x}$$

som ønsket. □

Eksempel 2.3.9 Vi så i eksempel 2.3.5, at afbildningen $s \colon \mathbb{R}^2 \to \mathbb{R}^2$, der er givet ved spejling i x-aksen, er lineær, og ifølge sætning 2.3.8 er den derfor givet ved venstre multiplikation med en 2×2-matrix A. Sætningen siger endvidere, at

$$A = \begin{pmatrix} s(\boldsymbol{e}_1) & s(\boldsymbol{e}_2) \end{pmatrix} = \begin{pmatrix} s\begin{pmatrix} 1 \\ 0 \end{pmatrix} & s\begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$



Figur 2.5: Rotation med vinkel θ omkring $\mathbf{0}$ i $\mathbb{R}^2$ er en lineær afbildning

hvilket udregningen

$$A\boldsymbol{x} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ -x_2 \end{pmatrix} = s(\boldsymbol{x})$$

da også bekræfter.

Eksempel 2.3.10 Vi betragter afbildningen $r_\theta \colon \mathbb{R}^2 \to \mathbb{R}^2$, der er givet ved rotationen gennem θ radianer imod urets retning omkring $\mathbf{0} \in \mathbb{R}^2$. Figur 2.5 illustrerer, at denne afbildning opfylder $r_\theta(\boldsymbol{x}+\boldsymbol{y}) = r_\theta(\boldsymbol{x}) + r_\theta(\boldsymbol{y})$ og $r_\theta(\boldsymbol{x}a) = r_\theta(\boldsymbol{x})a$ og derfor er lineær, og den er derfor ifølge sætning 2.3.8 givet ved venstre multiplikation 2×2-matricen

$$B = \begin{pmatrix} r_\theta(\boldsymbol{e}_1) & r_\theta(\boldsymbol{e}_2) \end{pmatrix} = \begin{pmatrix} \cos\theta & \cos(\theta + \frac{\pi}{2}) \\ \sin\theta & \sin(\theta + \frac{\pi}{2}) \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Vi vil faktisk definere $r_\theta \colon \mathbb{R}^2 \to \mathbb{R}^2$ til at være afbildningen givet ved $r_\theta(\boldsymbol{x}) = B \cdot \boldsymbol{x}$.

Definition 2.3.11 Lad $\mathbb{F}$ være et legeme og lad $f \colon \mathbb{F}^n \to \mathbb{F}^m$ være en lineær afbildning. Den entydigt bestemte matrix $A \in M_{m,n}(\mathbb{F})$, sådan at der for alle $\boldsymbol{x} \in \mathbb{F}^n$ gælder, at

$$f(\boldsymbol{x}) = A\boldsymbol{x},$$

kaldes for *matricen, der repræsenterer $f \colon \mathbb{F}^n \to \mathbb{F}^m$ med hensyn til standardbaserne* for henholdsvis $\mathbb{F}^n$ og $\mathbb{F}^m$. Matricen har søjler $f(\boldsymbol{e}_1), \dots, f(\boldsymbol{e}_n)$.

Vi skal senere definere og betragte andre baser for $\mathbb{F}^n$ og $\mathbb{F}^m$ end standardbaserne, og vi er derfor forsigtige med at sige, at det her er standardbaserne, vi benytter.

Eksempel 2.3.12 Identitetsafbildningen $\mathrm{id} \colon \mathbb{F}^n \to \mathbb{F}^n$ er defineret ved

$$\mathrm{id}(\boldsymbol{x}) = \boldsymbol{x}.$$

Den er lineær, fordi $\mathrm{id}(\boldsymbol{x}+\boldsymbol{y}) = \boldsymbol{x}+\boldsymbol{y} = \mathrm{id}(\boldsymbol{x}) + \mathrm{id}(\boldsymbol{y})$ og $\mathrm{id}(\boldsymbol{x}a) = \boldsymbol{x}a = \mathrm{id}(\boldsymbol{x})a$, så ifølge sætning 2.3.8 er den givet ved venstre multiplikation med $n \times n$-matricen

$$A = \begin{pmatrix} \mathrm{id}(\boldsymbol{e}_1) & \mathrm{id}(\boldsymbol{e}_2) & \cdots & \mathrm{id}(\boldsymbol{e}_n) \end{pmatrix} = \begin{pmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \cdots & \boldsymbol{e}_n \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} = I_n.$$

Altså er identitetsafbildningen $\mathrm{id}\colon \mathbb{F}^n \to \mathbb{F}^n$ repræsenteret af identitetsmatricen I_n med hensyn til standardbasen $(\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_n)$ for både domænet og codomænet.

Sætning 2.3.13 *Lad $\mathbb{F}$ være et legeme, lad $f\colon \mathbb{F}^n \to \mathbb{F}^m$ og $g\colon \mathbb{F}^p \to \mathbb{F}^n$ være to lineære afbildninger og lad $A \in M_{m,n}(\mathbb{F})$ og $B \in M_{n,p}(\mathbb{F})$ være matricerne, der repræsenterer henholdsvis $f\colon \mathbb{F}^n \to \mathbb{F}^m$ og $g\colon \mathbb{F}^p \to \mathbb{F}^n$ med hensyn til de respektive standardbaser. Den sammensatte afbildning $f \circ g\colon \mathbb{F}^p \to \mathbb{F}^m$ er igen lineær, og matricen $C \in M_{m,p}(\mathbb{F})$, der represæsenterer $f \circ g\colon \mathbb{F}^p \to \mathbb{F}^m$ med hensyn til de respektive standardbaser, er*

$$C = AB.$$

Bevis For alle $\boldsymbol{x} \in \mathbb{F}^p$, gælder der, at

$$(f \circ g)(\boldsymbol{x}) = f(g(\boldsymbol{x})) = A(B\boldsymbol{x}) = (AB)\boldsymbol{x}.$$

Her følger den første lighed fra definitionen af den sammensatte afbildning; den anden fra sætning 2.3.8; og den sidste fra sætning 2.1.9 (P1). Det følger derfor fra sætning 2.3.6, at afbildningen $f \circ g\colon \mathbb{F}^p \to \mathbb{F}^m$ er lineær, og da matricen C, der repræsenterer denne afbildning med hensyn til standardbaserne for $\mathbb{F}^p$ og $\mathbb{F}^m$, er entydigt bestemt, konkluderer vi endvidere, at $C = AB$ som påståét. $\square$

Eksempel 2.3.14 Lad $s\colon \mathbb{R}^2 \to \mathbb{R}^2$ være den lineære afbildning fra eksempel 2.3.5, der er givet ved spejling i x-aksen, og lad $r = r_{\pi/2}\colon \mathbb{R}^2 \to \mathbb{R}^2$ være den lineære afbildning fra eksempel 2.3.10, der er givet ved rotation gennem $\pi/2$ radianer imod urets retning omkring $\mathbf{0} \in \mathbb{R}^2$. Vi ved fra eksempel 2.3.9 og 2.3.10, at $r\colon \mathbb{R}^2 \to \mathbb{R}^2$ og $s\colon \mathbb{R}^2 \to \mathbb{R}^2$ er repræsenteret med hensyn til de respektive standardbaser ved henholdsvis

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

Ifølge sætning 2.3.13 repræsenterer produktmatricerne BA og AB derfor henholdsvis de sammensatte afbildninger $s \circ r\colon \mathbb{R}^2 \to \mathbb{R}^2$ og $r \circ s\colon \mathbb{R}^2 \to \mathbb{R}^2$, hvoraf den første er "først drej så spejl", mens den anden er "først spejl så drej". Vi udregner

$$\begin{aligned} BA &= \left(\begin{array}{|c|c|} \hline 0 & -1 \\ 1 & 0 \\ \hline \end{array}\right)\left(\begin{array}{|cc|} \hline 1 & 0 \\ \hline 0 & -1 \\ \hline \end{array}\right) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \\ AB &= \left(\begin{array}{|cc|} \hline 1 & 0 \\ \hline 0 & -1 \\ \hline \end{array}\right)\left(\begin{array}{|c|c|} \hline 0 & -1 \\ 1 & 0 \\ \hline \end{array}\right) = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}. \end{aligned}$$

Vi bemærker, at de to matricer er forskellige, og konkluderer derfor, at de sammensatte afbildninger $r \circ s$ og $s \circ r$ ligeledes er forskellige. Dette er illustreret i figur 2.6.

Figur 2.6: Man starter med den samme (sort) billede. I den første figure laver man en rotation med vinkel $\frac{\pi}{2}$ (i rød) efterfulgt af en spejling (i blå). I den anden, laver man først samme spejling (i rød) og så rotationen med samme vinkel (i blå). (Se Eksempel 2.3.14 for en forklaring af billedet med matricer.)

Eksempel 2.3.15 En fabrik fremstiller to varer X_1 og X_2 og anvender dertil tre råvarer Y_1, Y_2 og Y_3. Hvis der dagligt fremstilles x_1 enheder af X_1 og x_2 enheder af X_2, så siger vi, at fabrikkens *produktion* er

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Hvis fabrikken dagligt forbruger y_1 enheder af Y_1, y_2 enheder af Y_2 og y_3 enheder af Y_3, så siger vi tilsvarende, at fabrikkens *forbrug* er

$$\boldsymbol{y} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}.$$

Til produktion af 1 enhed af X_1 kræves 3 enheder af Y_1, 2 enheder af Y_2 og 1 enhed af Y_3; og til produktion 1 enhed af X_2 kræves 5 enheder af Y_1, 5 enheder af Y_2 og 3 enheder af Y_3. Der gælder altså følgende sammenhæng mellem forbrug og produktion:

$$\begin{aligned} y_1 &= 3x_1 + 5x_2 \\ y_2 &= 2x_1 + 5x_2 \\ y_3 &= x_1 + 3x_2. \end{aligned}$$

Så hvis vi lader $f : \mathbb{R}^2 \to \mathbb{R}^3$ være den afbildning, der til en ønsket produktion $\boldsymbol{x}$ af varer tilordner det tilsvarende forbrug $\boldsymbol{y}$ af råvarer, da gælder der

$$f\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 3x_1 + 5x_2 \\ 2x_1 + 5x_2 \\ x_1 + 3x_2 \end{pmatrix} = \begin{pmatrix} 3 & 5 \\ 2 & 5 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Dermed er $f : \mathbb{R}^2 \to \mathbb{R}^3$ altså en lineær afbildning.

2.4 Invertible matricer

I dette afsnit undersøger vi de matricer A for hvilke ligningssystemet $A\boldsymbol{x} = \boldsymbol{b}$ præcis har én løsning for alle $\boldsymbol{b}$. Disse matricer, som vi kalder de invertible matricer, repræsenterer således bijektive lineære afbildninger, så vi genkalder først definitionen af en bijektiv afbildninger.

Vi minder om, at en afbildning $f: X \to Y$ mellem to vilkårlige mængder X og Y kaldes for *bijektiv*, hvis der findes en afbildning $g: Y \to X$ i den modsatte retning, sådan at

$$f \circ g = \mathrm{id}_Y \qquad \text{og} \qquad g \circ f = \mathrm{id}_X.$$

Hvis en sådan afbildning $g: Y \to X$ findes, da er den entydigt bestemt. For hvis også afbildningen $g': Y \to X$ opfylder $f \circ g' = \mathrm{id}_X$ og $g' \circ f = \mathrm{id}_Y$, da er

$$g = g \circ \mathrm{id}_Y = g \circ (f \circ g') = (g \circ f) \circ g' = \mathrm{id}_X \circ g' = g'.$$

Vi kalder denne entydigt bestemte afbildning for den *inverse afbildning* af $f: X \to Y$ og betegner den med $f^{-1}: Y \to X$. Vi siger også, at en bijektiv afbildning $f: X \to Y$ er en 1-1 korrespondance fra X til Y, og det følgende lemma forklarer hvorfor.

Lemma 2.4.1 *Lad $f: X \to Y$ være en afbildning. De følgende udsagn er ækvivalente:*

(1) *Afbildningen $f: X \to Y$ er bijektiv.*

(2) *For alle $y \in Y$ findes der præcis et element $x \in X$, sådan at $f(x) = y$.*

Bevis Vi antager først (1) og lader $g: Y \to X$ være den inverse afbildning af $f: X \to Y$. Givet $y \in Y$, da opfylder $x = g(y) \in X$, at

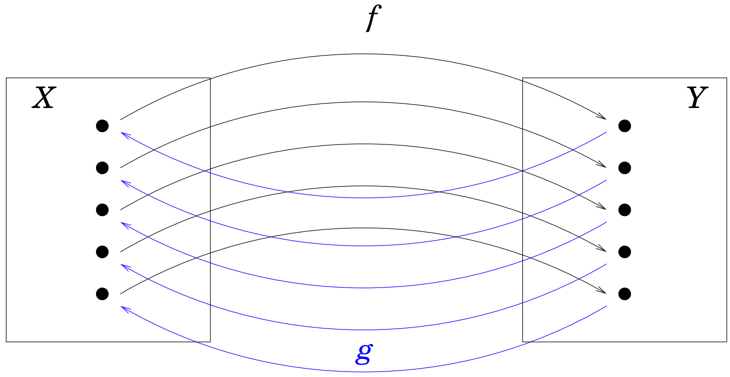
$$f(x) = f(g(y)) = (f \circ g)(y) = \mathrm{id}_Y(y) = y,$$

og hvis også $x' \in X$ opfylder, at $g(x') = y$, da er

$$x' = \mathrm{id}_X(x') = (g \circ f)(x') = g(f(x')) = g(y) = g(f(x)) = (g \circ f)(x) = \mathrm{id}_X(x) = x,$$

hvilket viser (2). Vi antager dernæst (2) og lader $g: Y \to X$ være den afbildning, der til et element $y \in Y$ tilordner det entydigt bestemte element $x \in X$, sådan at $y = f(x)$. Hvis $x \in X$ og $y \in Y$, da er udsagnene "$f(x) = y$" og "$x = g(y)$" derfor ensbetydende. Specielt er udsagnene "$f(g(y)) = y$" og "$g(y) = g(y)$" ensbetydende, hvilket viser, at $f \circ g = \mathrm{id}_Y$. Og tilsvarende er udsagnene "$g(f(x)) = x$" og "$f(x) = f(x)$" ensbetydende, hvilket viser, at $g \circ f = \mathrm{id}_X$. Dette viser (1). $\square$

Eksempel 2.4.2 Vi lader $a \in \mathbb{R}$ og betragter den lineære afbildning $f: \mathbb{R} \to \mathbb{R}$ defineret ved $f(x) = ax$. Hvis $a \neq 0$, da er $f: \mathbb{R} \to \mathbb{R}$ bijektiv, og den inverse afbildning $f^{-1}: \mathbb{R} \to \mathbb{R}$ er givet ved $f^{-1}(y) = a^{-1}y$. Hvis $a = 0$, da er $f(x) = 0$ for alle $x \in \mathbb{R}$, og lemma 2.4.1 viser derfor, at $f: \mathbb{R} \to \mathbb{R}$ ikke er bijektiv.



Figur 2.7: Bijektiv afbildning

Korollar 2.4.3 *Lad $\mathbb{F}$ være et legeme, lad $f\colon \mathbb{F}^n \to \mathbb{F}^m$ være en lineær afbildning, og lad $A \in M_{m,n}(\mathbb{F})$ være matricen, der repræsenterer $f\colon \mathbb{F}^n \to \mathbb{F}^m$ med hensyn til de respektive standardbaser. Da er følgende udsagn ækvivalente:*

(1) *Afbildningen $f\colon \mathbb{F}^n \to \mathbb{F}^m$ er bijektiv.*

(2) *For alle $\boldsymbol{b} \in \mathbb{F}^m$ har det lineære ligningssystem $A\boldsymbol{x} = \boldsymbol{b}$ præcis én løsning $\boldsymbol{x} \in \mathbb{F}^n$.*

Bevis Dette følger umiddelbart fra lemma 2.4.1, idet $f(\boldsymbol{x}) = A\boldsymbol{x}$. □

Sætning 2.4.4 *Hvis $\mathbb{F}$ er et legeme, og hvis $f\colon \mathbb{F}^n \to \mathbb{F}^m$ er en bijektiv lineær afbildning, så gælder der nødvendigvis, at $m = n$.*

Bevis Lad $A \in M_{m,n}(\mathbb{F})$ være den matrix, der repræsenterer $f\colon \mathbb{F}^n \to \mathbb{F}^m$ med hensyn til standardbaserne, og lad B være en matrix på echelon form af rang $0 \leq r \leq \min\{m,n\}$, der er fremkommet fra A ved at udføre en følge af rækkeoperationer. Ifølge korollar 2.4.3 har det lineære ligningssystem $A\boldsymbol{x} = \boldsymbol{b}$ præcis én løsning for alle $\boldsymbol{b} \in \mathbb{F}^m$, og derfor viser sætning 1.2.17, at $r = m = n$. Specielt er $m = n$ som påstået. □

Sætning 2.4.4 viser specielt, at en bijektiv lineær afbildning altid er repræsenteret af en kvadratisk matrix.

Bemærkning 2.4.5 Givet vilkårlige positive heltal m og n, så kan man vise, at der altid findes en bijektiv afbildning $f\colon \mathbb{R}^n \to \mathbb{R}^m$. Sætning 2.4.4 viser imidlertid, at en sådan bijektiv afbildning kun kan være lineær, hvis $m = n$.

Sætning 2.4.6 *Hvis $\mathbb{F}$ er et legeme, og hvis $f\colon \mathbb{F}^n \to \mathbb{F}^n$ er en bijektiv lineær afbildning, så er den inverse afbildning $g\colon \mathbb{F}^n \to \mathbb{F}^n$ også lineær.*

Bevis Vi skal vise, at den inverse afbildning $g\colon \mathbb{F}^n \to \mathbb{F}^n$ opfylder (L1) og (L2). Vi så i beviset for lemma 2.4.1, at "$\boldsymbol{u} = g(\boldsymbol{v})$" og "$f(\boldsymbol{u}) = \boldsymbol{v}$" er ækvivalente udsagn. Specielt er udsagnene "$g(\boldsymbol{x}) + g(\boldsymbol{y}) = g(\boldsymbol{x} + \boldsymbol{y})$" og "$f(g(\boldsymbol{x}) + g(\boldsymbol{y})) = \boldsymbol{x} + \boldsymbol{y}$" derfor ækvivalente, og det sidste udsagn gælder, da $f : \mathbb{F}^n \to \mathbb{F}^n$ opfylder (L1):

$$f(g(\boldsymbol{x}) + g(\boldsymbol{y})) = f(g(\boldsymbol{x})) + f(g(\boldsymbol{y})) = \boldsymbol{x} + \boldsymbol{y}.$$

Dette viser, at $g\colon \mathbb{F}^n \to \mathbb{F}^n$ opfylder (L1). Tilsvarende er "$g(\boldsymbol{x})a = g(\boldsymbol{x}a)$"og "$f(g(\boldsymbol{x})a) = \boldsymbol{x}a$" ækvivalente, og det sidste gælder, da $f : \mathbb{F}^n \to \mathbb{F}^n$ opfylder (L2):

$$f(g(\boldsymbol{x})a) = f(g(\boldsymbol{x}))a = \boldsymbol{x}a.$$

Dette viser, at $g\colon \mathbb{F}^n \to \mathbb{F}^n$ også opfylder (L2). Altså er $g\colon \mathbb{F}^n \to \mathbb{F}^n$ lineær. □

Lad $f : \mathbb{F}^n \to \mathbb{F}^n$ være en bijektiv lineær afbildning, og lad $g\colon \mathbb{F}^n \to \mathbb{F}^n$ være den inverse afbildning, der ifølge sætning 2.4.6 også er lineær. Lad $A \in M_n(\mathbb{F})$ og $B \in M_n(\mathbb{F})$ være de kvadratiske matricer, der repræsenterer henholdsvis $f\colon \mathbb{F}^n \to \mathbb{F}^n$ og $g\colon \mathbb{F}^n \to \mathbb{F}^n$ med hensyn til standardbasen for både domænet og codomænet. Ifølge sætning 2.3.13 og eksempel 2.3.12 er identiteterne $f \circ g = \mathrm{id} = g \circ f$ og matrix identiterne

$$AB = I = BA$$

da ensbetydende. Dette motiverer den følgende definition.

Definition 2.4.7 En kvadratisk matrix A er *invertibel*, hvis der findes en kvadratisk matrix B, sådan at $AB = I = BA$, hvor I er identitetsmatricen.

Hvis A er en kvadratisk matrix, og hvis de to kvadratiske matricer B og B' opfylder, at $AB = I = BA$ og $AB' = I = B'A$, da er $B = B'$. For ifølge sætning 2.1.9 er

$$B = BI = B(AB') = (BA)B' = IB' = B'.$$

Dette følger også fra entydigheden af den inverse afbildning, idet B repræsenterer den lineære inverse afbildning af den lineære afbildning, som A repræsenterer. Vi kalder matricen B for den *inverse matrix* af matricen A og betegner den A^{-1}.

Bemærkning 2.4.8 Lad $\mathbb{F}$ være et legeme, lad $f : \mathbb{F}^n \to \mathbb{F}^n$ være en lineær afbildning, og lad $A \in M_n(\mathbb{F})$ være den kvadratiske matrix, der repræsenterer $f : \mathbb{F}^n \to \mathbb{F}^n$ med hensyn til standardbasen for både domænet og codomænet. Vi har ovenfor vist, at de følgende udsagn (1)–(4) er ækvivalente.

(1) Afbildningen $f : \mathbb{F}^n \to \mathbb{F}^n$ er bijektiv.

(2) Der findes en lineær afbildning $g\colon \mathbb{F}^n \to \mathbb{F}^n$, sådan at $f \circ g = \mathrm{id} = g \circ f$.

(3) Den kvadratiske matrix A er invertibel.

(4) Det lineære ligningssystem $A\boldsymbol{x} = \boldsymbol{b}$ har præcis én løsning for alle $\boldsymbol{b} \in \mathbb{F}^n$.

Hvis en kvadratisk matrix A er invertibel, og hvis dens inverse matrix A^{-1} er kendt, så er den entydige løsning til ligningssystemet $A\boldsymbol{x} = \boldsymbol{b}$ givet ved $\boldsymbol{x} = A^{-1}\boldsymbol{b}$. For

$$A(A^{-1}\boldsymbol{b}) = (AA^{-1})\boldsymbol{b} = I\boldsymbol{b} = \boldsymbol{b}.$$

Vi skal i næste afsnit vise, at opgaven med at bestemme A^{-1} er ækvivalent til opgaven med at bestemme den entydige løsning til $A\boldsymbol{x} = \boldsymbol{e}_j$ for alle j.

Eksempel 2.4.9 Vi betragter matricerne

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

og udregner

$$\begin{aligned} AB &= \left(\begin{array}{|cc|} \hline 1 & 2 \\ \hline -1 & 3 \\ \hline \end{array}\right)\left(\begin{array}{|c|c|} \hline 3/5 & -2/5 \\ 1/5 & 1/5 \\ \hline \end{array}\right) = \begin{pmatrix} 3/5 + 2/5 & -2/5 + 2/5 \\ -3/5 + 3/5 & 2/5 + 3/5 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I \\ BA &= \left(\begin{array}{|cc|} \hline 3/5 & -2/5 \\ \hline 1/5 & 1/5 \\ \hline \end{array}\right)\left(\begin{array}{|c|c|} \hline 1 & 2 \\ -1 & 3 \\ \hline \end{array}\right) = \begin{pmatrix} 3/5 + 2/5 & 6/5 - 6/5 \\ 1/5 - 1/5 & 2/5 + 3/5 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I \end{aligned}$$

hvilket viser, at $B = A^{-1}$ er den inverse matrix af A.

Som eksempel 2.4.9 illustrerer, så er det let at checke om to kvadratiske matricer A og B er hinandens inverse ellers ej. Vi skal i næste afsnit vise, hvordan Gauss elimination kan anvendes til at afgøre om en kvadratisk matrix A er invertibel eller ej og givet fald finde den inverse matrix. Vi skal senere i kapitel 3 se, hvorfor den inverse matrix af en invertibel kvadratisk matrix med heltalsindgange kan forventes at have brøker (med samme nævner!) som indgange.

En kvadratisk matrix på formen

$$\mathrm{diag}(a_1, \dots, a_n) = \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n \end{pmatrix}$$

kaldes for en *diagonalmatrix*. Vi bemærker, at

$$\mathrm{diag}(a_1, \dots, a_n)\,\mathrm{diag}(b_1, \dots, b_n) = \mathrm{diag}(a_1 b_1, \dots, a_n b_n),$$

så produktet af to diagonalmatricer er let at udregne.

Lemma 2.4.10 *Lad $\mathbb{F}$ være et legeme. Diagonalmatricen* $\operatorname{diag}(a_1,\dots,a_n)$ *er invertibel, hvis og kun hvis diagonalindgangene $a_1,\dots,a_n \in \mathbb{F}$ alle er invertible, og i givet fald er*

$$\operatorname{diag}(a_1,\dots,a_n)^{-1} = \operatorname{diag}(a_1^{-1},\dots,a_n^{-1}).$$

Bevis Hvis $a_1,\dots,a_n$ alle er invertible, så er

$$\begin{aligned}\operatorname{diag}(a_1,\dots,a_n)\operatorname{diag}(a_1^{-1},\dots,a_n^{-1}) &= \operatorname{diag}(1,\dots,1) = I,\\ \operatorname{diag}(a_1^{-1},\dots,a_n^{-1})\operatorname{diag}(a_1,\dots,a_n) &= \operatorname{diag}(1,\dots,1) = I,\end{aligned}$$

hvilket viser, at $\operatorname{diag}(a_1,\dots,a_n)$ er invertibel, og at dens inverse matrix er som angivet. Omvendt, hvis $A = \operatorname{diag}(a_1,\dots,a_n)$ er invertibel med invers matrix B, så viser

$$AB = \begin{pmatrix} a_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & a_n \end{pmatrix}\begin{pmatrix} b_{11} & \dots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{n1} & \dots & b_{nn} \end{pmatrix} = \begin{pmatrix} a_1b_{11} & \dots & a_1b_{1n} \\ \vdots & \ddots & \vdots \\ a_nb_{n1} & \dots & a_nb_{nn} \end{pmatrix} = I,$$
$$BA = \begin{pmatrix} b_{11} & \dots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{n1} & \dots & b_{nn} \end{pmatrix}\begin{pmatrix} a_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & a_n \end{pmatrix} = \begin{pmatrix} b_{11}a_1 & \dots & b_{1n}a_n \\ \vdots & \ddots & \vdots \\ b_{n1}a_1 & \dots & b_{nn}a_n \end{pmatrix} = I,$$

at a_i er invertibel med invers $a_i^{-1} = b_{ii}$ for alle $1 \le i \le n$, som ønsket. □

Sætning 2.4.11 *Hvis A og B er invertible matricer af samme dimensioner, da er AB også en invertibel matrix, og dens inverse matrix er givet ved*

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

Bevis Vi sætter $C = B^{-1}A^{-1}$ og udregner

$$\begin{aligned}(AB)C &= (AB)(B^{-1}A^{-1}) = ((AB)B^{-1})A^{-1} = (A(BB^{-1}))A^{-1} = (AI)A^{-1} = AA^{-1} = I,\\ C(AB) &= (B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}(AB)) = B^{-1}((A^{-1}A)B) = B^{-1}(IB) = B^{-1}B = I.\end{aligned}$$

Per entydighed af den inverse matrix konkluderer vi derfor, at $C = (AB)^{-1}$. □

Bemærkning 2.4.12 Hvis $\mathbb{F}$ er et legeme, så skriver vi

$$GL_n(\mathbb{F}) \subset M_n(\mathbb{F})$$

for delmængden bestående af de invertible $n \times n$-matricer. Det følger fra sætning 2.4.11, at matrixproduktet restringerer til et produkt på $GL_n(\mathbb{F})$, der opfylder følgende:

(G1) For alle $A, B, C \in GL_n(\mathbb{F})$ er $(A \cdot B) \cdot C = A \cdot (B \cdot C)$.

(G2) Der findes $I \in GL_n(\mathbb{F})$, sådan at $A \cdot I = A = I \cdot A$ for alle $A \in GL_n(\mathbb{F})$.

(G3) For alle $A \in GL_n(\mathbb{F})$ findes $B \in GL_n(\mathbb{F})$, sådan at $A \cdot B = I = B \cdot A$.

Generelt defineres en *gruppe* til at være et par $(G, \cdot)$ bestående af en mængde G og et produkt "$\cdot$" på G, der opfylder (G1)–(G3). Gruppen $(GL_n(\mathbb{F}), \cdot)$ kaldes den generelle lineære gruppe og er et typisk eksempel på en gruppe. Grupper i almindelighed og den generelle lineære gruppe i særdeleshed optræder mange steder i matematikken.

2.5 Operationsmatricer

I dette afsnit introducerer vi *operationsmatricer*, som er matricer med det egnskab, at venstre multiplikation med disse matricer effektuerer de rækkeoperationer, vi definerede i kapitel 1. Vi beviser også flere sætninger fra kapitel 1 i en mere præcis form. Specielt viser vi, at hvis en matrix A ved rækkeoperationer omdannes til en matrix A' på reduceret echelon form, da afhænger A' kun af A og ikke af den valgte følge af rækkeoperationer. Vi viser også, hvordan Gauss elimination anvendes til at afgøre, om en matrix er invertibel og i givet fald finde den inverse matrix.

Vi indfører nu tre typer af kvadratiske matricer, som vi kalder *operationsmatricer*, og som svarer til hver af de tre typer af rækkeoperationer, som vi definerede i kapitel 1. De tre typer rækkeoperationerne var: Type M, at gange en række med en skalar $c \neq 0$ fra venstre; Type S, at adder c gange en række til en anden række; og Type T, at bytte om på to rækker. Alle tre typer operationsmatricer, som vi nu definerer, er kvadratiske matricer, der kun afviger ganske lidt fra at være diagonalmatricer.

Type M: For alle $1 \leq i \leq m$ og $0 \neq c \in \mathbb{F}$ er $M_i(c)$ den matrix, der fås fra I_m ved at multiplicere den i'te række med c fra venstre.

Type S: For alle $1 \leq i \neq j \leq m$ og $c \in \mathbb{F}$ er $S_{ij}(c)$ den matrix, der fås fra I_m ved at addere c gange den j'te række til den i'te række .

Type T: For alle $1 \leq i < j \leq m$ er T_{ij} den matrix, der fås fra I_m ved ombytning af den i'te række og den j'te række.

Figur 2.8 illustrerer disse matricer.

Vi viser nu, at rækkeoperationer på en $m \times n$-matrix svarer til venstre multiplikation med operationsmatricer af orden m.

$$M_i(c) = \begin{array}{c} \begin{matrix} & i & \end{matrix} \\ \begin{pmatrix} 1 & & \\ & c & \\ & & 1 \end{pmatrix} \end{array} \begin{matrix} \\ i \\ \\ \end{matrix} \qquad S_{ij}(c) = \begin{array}{c} \begin{matrix} & i & j & \end{matrix} \\ \begin{pmatrix} 1 & & & \\ & 1 & c & \\ & & 1 & \\ & & & 1 \end{pmatrix} \end{array} \begin{matrix} \\ i \\ \\ \\ \end{matrix} \qquad T_{ij} = \begin{array}{c} \begin{matrix} & i & j & \end{matrix} \\ \begin{pmatrix} 1 & & & \\ & 0 & 1 & \\ & 1 & 0 & \\ & & & 1 \end{pmatrix} \end{array} \begin{matrix} \\ i \\ j \\ \\ \end{matrix}$$

Figur 2.8: Operationsmatricerne $M_i(c)$, $S_{ij}(c)$ og T_{ij} adskiller sig fra I_m som anført.

Sætning 2.5.1 *Lad $A \in M_{m,n}(\mathbb{F})$ være en $m \times n$-matrix med indgange i et legeme $\mathbb{F}$.*

(1) *For alle $1 \le i \le m$ og $0 \neq c \in \mathbb{F}$ er $M_i(c)A$ den matrix, der fremkommer fra A ved at multiplicere den i'te række med c fra venstre.*

(2) *For alle $1 \le i \neq j \le m$ og $c \in \mathbb{F}$ er $S_{ij}(c)A$ den matrix, der fremkommer fra A ved at addere c gange den j'te række til den i'te række.*

(3) *For alle $1 \le i < j \le m$ er $T_{ij}A$ den matrix, der fremkommer fra A ved at ombytte den i'te række og den j'te række.*

Bevis Lad $\boldsymbol{\delta}_r \in M_{1,m}(\mathbb{F})$ være $1 \times m$-matricen med $(\boldsymbol{\delta}_r)_{1j} = \delta_{rj}$. Matricen $\boldsymbol{\delta}_r$ er altså den r'te række i identitetsmatricen $I_m \in M_{m,m}(\mathbb{F})$, og hvis $B \in M_{m,n}(\mathbb{F})$ er en $m \times n$-matrix, da er $1 \times n$-matricen $\boldsymbol{\delta}_r \cdot B \in M_{1,n}(\mathbb{F})$ den r'te række i B.

Vi viser nu (1). Per definition er

$$\boldsymbol{\delta}_r \cdot M_i(c) = \begin{cases} c \cdot \boldsymbol{\delta}_r & \text{hvis } r = i, \\ \boldsymbol{\delta}_r & \text{hvis } r \neq i, \end{cases}$$

og den r'te række i $M_i(c) \cdot A$ er derfor lig med henholdsvis

$$\boldsymbol{\delta}_r \cdot (M_i(c) \cdot A) = (\boldsymbol{\delta}_r \cdot M_i(c)) \cdot A = (c \cdot \boldsymbol{\delta}_r) \cdot A = c \cdot (\boldsymbol{\delta}_r \cdot A),$$

hvis $r = i$, og

$$\boldsymbol{\delta}_r \cdot (M_i(c) \cdot A) = (\boldsymbol{\delta}_r \cdot M_i(c)) \cdot A = \boldsymbol{\delta}_r \cdot A,$$

hvis $r \neq i$, hvilket netop er påstanden (1).

Vi viser dernæst (2). Operationsmatricen $S_{ij}(c)$ er per definition givet ved

$$\boldsymbol{\delta}_r \cdot S_{ij}(c) = \begin{cases} \boldsymbol{\delta}_r + c \cdot \boldsymbol{\delta}_j & \text{hvis } r = i, \\ \boldsymbol{\delta}_r & \text{hvis } r \neq i, \end{cases}$$

og den r'te række i $S_{ij} \cdot A$ er derfor lig med henholdsvis

$$\begin{aligned}\boldsymbol{\delta}_r \cdot (S_{ij}(c) \cdot A) &= (\boldsymbol{\delta}_r \cdot S_{ij}(c)) \cdot A = (\boldsymbol{\delta}_r + c \cdot \boldsymbol{\delta}_j) \cdot A \\ &= \boldsymbol{\delta}_r \cdot A + (c \cdot \boldsymbol{\delta}_j) \cdot A = \boldsymbol{\delta}_r \cdot A + c \cdot (\boldsymbol{\delta}_j \cdot A),\end{aligned}$$

hvis $r = i$, og

$$\boldsymbol{\delta}_r \cdot (S_{ij}(c) \cdot A) = (\boldsymbol{\delta}_r \cdot S_{ij}(c)) \cdot A = \boldsymbol{\delta}_r \cdot A,$$

hvis $r \neq i$. Dette viser påstanden (2).

Endelig viser vi (3). Operationsmatricen T_{ij} er defineret ved

$$\boldsymbol{\delta}_r \cdot T_{ij} = \begin{cases} \boldsymbol{\delta}_j & \text{hvis } r = i, \\ \boldsymbol{\delta}_i & \text{hvis } r = j, \\ \boldsymbol{\delta}_r & \text{hvis } r \neq i \text{ og } r \neq j, \end{cases}$$

og derfor er den r'te række i $T_{ij} \cdot A$ lig med

$$\boldsymbol{\delta}_r \cdot (T_{ij} \cdot A) = (\boldsymbol{\delta}_r \cdot T_{ij}) \cdot A = \begin{cases} \boldsymbol{\delta}_j \cdot A & \text{hvis } r = i, \\ \boldsymbol{\delta}_i \cdot A & \text{hvis } r = j, \\ \boldsymbol{\delta}_r \cdot A & \text{hvis } r \neq i \text{ og } r \neq j, \end{cases}$$

hvilket viser (3). □

I kapitel 1 indførte vi en invers rækkeoperation hørende til hver rækkeoperation, og oversat til operationsmatricer, giver dette følgende resultat.

Sætning 2.5.2 *Operationsmatricerne $M_i(c)$, $S_{ij}(c)$ og T_{ij} er alle invertible, og deres inverse matricer er givet ved henholdsvis*

$$M_i(c)^{-1} = M_i(c^{-1}), \qquad S_{ij}(c)^{-1} = S_{ij}(-c) \qquad \text{og} \qquad T_{ij}^{-1} = T_{ij}.$$

Bevis Hvis A er lig med $M_i(c)$, $S_{ij}(c)$ eller T_{ij}, så sætter vi B til at være lig med henholdsvis $M_i(c^{-1})$, $S_{ij}(-c)$ eller T_{ij}, og skal da vise, at $AB = I = BA$. Nu er

$$AB = (AB)I = A(BI),$$

og ifølge sætning 2.5.1 er $A(BI)$ den matrix, der fremkommer fra identitetsmatricen ved først at anvende den rækkeoperation, der svarer til B, og derefter den rækkeoperation, der svarer til A. I alle tre tilfælde, er disse rækkeoperationer præcis hinandens inverse rækkeoperationer. Dette viser, at $AB = I$, og $BA = I$ vises tilsvarende. Alternativt kan sætningen også let bevises ved direkte udregning af de relevante matrix produkter. □

Definition 2.5.3 Hvis $A, B \in M_{m,n}(\mathbb{F})$ er $m \times n$-matricer med indgange i et legeme $\mathbb{F}$, da siges matricen A at være *rækkeækvivalent* til matricen B, hvis der findes en endelig følge af operationsmatricer $P_1, \dots, P_k \in M_m(\mathbb{F})$, sådan at $B = P_k \cdots P_2 P_1 A$.

Ifølge sætning 2.5.1 er A altså rækkeækvivalent til B, hvis B fremkommer fra A ved at anvende endeligt mange rækkeoperationer. Vi viser nu, at rækkeækvivalens er en ækvivalensrelation på mængden $M_{m,n}(\mathbb{F})$ af $m \times n$-matricer med indgange i $\mathbb{F}$.

Sætning 2.5.4 *Lad $\mathbb{F}$ være et legeme, og lad m og n være naturlige tal.*

(1) *For alle $A \in M_{m,n}(\mathbb{F})$ gælder, at A er rækkeækvivalent til A.*

(2) *For alle $A, B \in M_{m,n}(\mathbb{F})$ gælder, at A er rækkeækvivalent til B, hvis og kun hvis B er rækkeækvivalent til A.*

(3) *For alle $A, B, C \in M_{m,n}(\mathbb{F})$ gælder, at hvis A er rækkeækvivalent til B, og hvis B er rækkeækvivalent til C, så er A rækkeækvivalent til C.*

Bevis En matrix A fremkommer fra sig selv ved at anvende $k = 0$ rækkeoperationer, hvilket viser (1). Hvis $B = P_k \cdots P_1 A$, da er $A = P_1^{-1} \cdots P_k^{-1} B$, hvilket viser (2). Her har vi anvendt, at operationsmatricer ifølge sætning 2.5.1 er invertible. Hvis $B = P_k \cdots P_1 A$ og $C = Q_l \cdots Q_1 B$, da er $C = Q_l \cdots Q_1 P_k \cdots P_1 A$, hvilket viser (3) og dermed lemmaet. □

Vi omformulerer sætning 1.2.3 og gentager beviset ved hjælp af operationsmatricer.

Sætning 2.5.5 *Hvis totalmatricerne $(A \mid \boldsymbol{b})$ og $(B \mid \boldsymbol{c})$ for ligningssystemerne $A\boldsymbol{x} = \boldsymbol{b}$ og $B\boldsymbol{x} = \boldsymbol{c}$ er rækkeækvivalente, så har de to ligningssystemerne samme løsningsmængde.*

Bevis Hvis $(B \mid \boldsymbol{c}) = P_k \cdots P_1 (A \mid \boldsymbol{b})$, da er specielt $B = P_k \cdots P_1 A$ og $\boldsymbol{c} = P_k \cdots P_1 \boldsymbol{b}$. Hvis derfor $A\boldsymbol{x} = \boldsymbol{b}$, da er også $B\boldsymbol{x} = P_k \cdots P_1 A\boldsymbol{x} = P_k \cdots P_1 \boldsymbol{b} = \boldsymbol{c}$. Omvendt, hvis $B\boldsymbol{x} = \boldsymbol{c}$, da er også $A\boldsymbol{x} = P_1^{-1} \cdots P_k^{-1} B\boldsymbol{x} = P_1^{-1} \cdots P_k^{-1} \boldsymbol{c} = \boldsymbol{b}$. □

Den følgende sætning, hvoraf den første del er ækvivalent med sætning 1.2.7, er vores første hovedsætning.

Sætning 2.5.6 *Lad A være en $m \times n$-matrix med indgange i et legeme $\mathbb{F}$.*

(1) *Der findes en endelig følge af operationsmatricer $P_1, \dots, P_k \in M_m(\mathbb{F})$, sådan at*

$$A' = P_k \dots P_2 P_1 A$$

er på reduceret echelon-form af rang $0 \le r \le \min\{m, n\}$.

(2) *Matricen A' og dens rang r er entydigt bestemte af matricen A og afhænger ikke af den anvendte følge af operationsmatricer.*

Bevis Vi har allerede vist (1) i sætning 1.2.7, og for at vise (2) er det ifølge sætning 2.5.4 tilstrækkelige at vise, at to rækkeækvivalente $m \times n$-matricer A og B, der begge er på reduceret echelon-form, nødvendigvis identiske. Vi beviser denne påstand ved induktion på $n \ge 1$. For $n = 1$ er eneste $m \times 1$-matricer på reduceret echelon-form

$$\mathbf{0} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad \text{og} \quad \boldsymbol{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

og for enhver følge af (operations-)matricer $P_1, \dots, P_k \in M_m(\mathbb{F})$ er $P_k \cdots P_1 \mathbf{0} = \mathbf{0}$, hvilket viser, at $\mathbf{0}$ ikke er rækkeækvivalent til $\boldsymbol{e}_1$. Derfor er to rækkeækvivalente $m \times 1$-matricer A og B nødvendigvis ens, hvilket viser påstanden for $n = 1$.

Vi antager induktivt, at påstanden er vist for $n = p - 1$ og beviser den for $n = p$. Hertil lader vi C og D være de $m \times (p-1)$-matricer, der fremkommer fra henholdsvis A og B ved at fjerne den sidste søjle, sådan at

$$A = \left(\begin{array}{c|c} & a_{1p} \\ C & \vdots \\ & a_{mp} \end{array} \right) \quad \text{og} \quad B = \left(\begin{array}{c|c} & b_{1p} \\ D & \vdots \\ & b_{mp} \end{array} \right).$$

Da A og B er på reduceret echelon form, så er også C og D på reduceret echelon form, og de er endvidere rækkeækvivalente. For hvis $P_1, \dots, P_k$ er operationsmatricer, sådan at $B = P_k \cdots P_1 A$, da er også $C = P_k \cdots P_1 D$. Den induktive antagelse viser derfor, at $m \times (p-1)$-matricerne C og D er ens, hvoraf vi konkluderer, at

$$A - B = \left(\begin{array}{c|c} & a_{1p} - b_{1p} \\ C - D & \vdots \\ & a_{mp} - b_{mp} \end{array} \right) = \left(\begin{array}{c|c} & a_{1p} - b_{1p} \\ O & \vdots \\ & a_{mp} - b_{mp} \end{array} \right).$$

Vi ønsker at vise, at også den sidste søjle er lig med $\mathbf{0}$, sådan at $A = B$, og antager derfor modsætningsvist, at der findes et $1 \le j \le p$, sådan at $a_{jp} \ne b_{jp}$. Da A og B er rækkeækvivalente, så har ligningssystemerne $A\boldsymbol{x} = \mathbf{0}$ og $B\boldsymbol{x} = \mathbf{0}$ ifølge sætning 2.5.5 samme løsningsmængde. Hvis $\boldsymbol{x}$ er en løsning til disse ligningssystemer, da er

$$(A-B)\boldsymbol{x} = A\boldsymbol{x} - B\boldsymbol{x} = \mathbf{0} - \mathbf{0} = \mathbf{0},$$

og $\boldsymbol{x}$ er derfor også en løsning til ligningssystemet $(A-B)\boldsymbol{x} = \mathbf{0}$. Den j'te ligning i dette ligningssystem er "$(a_{jp} - b_{jp})x_p = 0$", og da vi har antaget, at $a_{jp} - b_{jp} \ne 0$, følger det, at $x_p = 0$. Så enhver løsning $\boldsymbol{x}$ til $A\boldsymbol{x} = \mathbf{0}$ og $B\boldsymbol{x} = \mathbf{0}$ har nødvendigvis $x_p = 0$, og derfor er x_p ikke en fri variable i disse ligningssystemer. Da A og B er på reduceret echolon form medfører dette, at deres p'te søjler begge indeholder en ledende indgang. Vi har allerede set, at $m \times (p-1)$-matricerne C og D, der er på reduceret echelon form, er identitiske, og derfor har de specielt samme rang $r-1$. Da $m \times p$-matricerne A og B er på reduceret echelon form, og da deres sidste søjler indeholder en ledende indgang, er disse søjler derfor begge nødvendigvis lig med $\boldsymbol{e}_r$. Dermed er $A = B$, hvilket strider mod vores antagelse, at $A \ne B$, så denne antagelse var altså forkert. Så $A = B$, hvilket viser induktionsskridtet og dermed sætningen. □

Bemærkning 2.5.7 Hvis $\mathbb{F}$ er et legeme, så viser sætning 2.5.4, at rækkeækvivalens er en ækvivalensrelation på mængden $M_{m,n}(\mathbb{F})$, og denne mængde kan derfor opdeles i de disjunkte ækvivalensklasser for denne ækvivalensrelation. Sætning 2.5.6 siger da, at hver rækkeækvivalensklasse af $m \times n$-matricer præcis indeholder én $m \times n$-matrix på reduceret echelon form.

Hvis $A \in M_{m,n}(\mathbb{F})$ er en matrix, så viser sætning 2.5.6, at der er en entydigt bestemt matrix $A' \in M_{m,n}(\mathbb{F})$, der er på reduceret echelon-form og rækkeækvivalent til A. Derfor er specielt rangen af A' entydigt bestemt af A, så følgende definition er meningsfuld. Vi husker på, at rangen af A' er lig med antallet af ledende indgange i A' eller ækvivalent med antallet af ikke-nul rækker i A'.

Definition 2.5.8 Lad A være en $m \times n$-matrix med indgange i et legeme $\mathbb{F}$ og lad A' være den entydigt bestemte matrix på reduceret echelon-form, der er rækkeækvivalent med A. Da er *rangen* af A defineret til at være rangen af A' og betegnes $\operatorname{rank}(A)$.

Vi skal senere vise, at rangen af $A \in M_{m,n}(\mathbb{F})$ er lig med dimensionen af billedet af den lineære afbildning $f\colon \mathbb{F}^n \to \mathbb{F}^m$ givet ved $f(\boldsymbol{x}) = A\boldsymbol{x}$, men det må vente til, vi har defineret vektorrum og deres dimension i kapitel 4. Vi relaterer her rangen af matricen A til injektivitet og surjektivitet af afbildningen $f\colon \mathbb{F}^n \to \mathbb{F}^m$, og genkalder først om definition af disse begreber.

En afbildning $f: X \to Y$ mellem to mængder X og Y kaldes for *injektiv*, hvis der for alle $y \in Y$ findes højst et $x \in X$ med $f(x) = Y$, og den kaldes for *surjektiv*, hvis der for alle $y \in Y$ findes mindst et $x \in X$ med $f(x) = y$. Ifølge lemma 2.4.1 er en afbildning bijektiv, hvis og kun hvis den både er injektiv og surjektiv.

Sætning 2.5.9 *Lad $\mathbb{F}$ være et legeme, lad A være en $m \times n$-matrix med indgange i $\mathbb{F}$, lad $0 \leq r \leq \min\{m,n\}$ være rangen af A og lad $f: \mathbb{F}^n \to \mathbb{F}^m$ være den lineære afbildning defineret ved $f(\boldsymbol{x}) = A\boldsymbol{x}$. Der gælder da følgende:*

(1) *Afbildningen $f: \mathbb{F}^n \to \mathbb{F}^m$ er surjektiv, hvis og kun hvis $r = m$.*

(2) *Afbildningen $f: \mathbb{F}^n \to \mathbb{F}^m$ er injektiv, hvis og kun hvis $r = n$.*

(3) *Afbildningen $f: \mathbb{F}^n \to \mathbb{F}^m$ er bijektiv, hvis og kun hvis $r = m = n$.*

Bevis Per definition er $f: \mathbb{F}^n \to \mathbb{F}^m$ injektiv (resp. surjektiv, resp. bijektiv), hvis og kun hvis ligningssystemet $A\boldsymbol{x} = \boldsymbol{b}$ har højst en løsning (resp. mindst en løsning, resp. præcis én løsning) for alle $\boldsymbol{b} \in \mathbb{F}^m$. Sætning følger derfor fra sætning 1.2.17 og definition 2.5.8. $\square$

Korollar 2.5.10 *Lad $\mathbb{F}$ være et legeme, og lad $f: \mathbb{F}^n \to \mathbb{F}^m$ være en lineær afbildning.*

(1) *Hvis $f: \mathbb{F}^n \to \mathbb{F}^m$ er surjektiv, da er $n \geq m$.*

(2) *Hvis $f: \mathbb{F}^n \to \mathbb{F}^m$ er injektiv, da er $n \leq m$.*

(3) *Hvis $f: \mathbb{F}^n \to \mathbb{F}^m$ er bijektiv, da er $n = m$.*

Bevis Lad $A \in M_{m,n}(\mathbb{F})$ være matricen, der repræsenterer $f: \mathbb{F}^n \to \mathbb{F}^m$ med hensyn til standardbaserne, sådan at $f(\boldsymbol{x}) = A\boldsymbol{x}$. Rangen r af A opfylder da, at $0 \leq r \leq \min\{m,n\}$, og korollaret følger derfor fra sætning 2.5.9. $\square$

Vi bemærker, at blandt kvadratiske matricer af orden n er identitetsmatricen den eneste matrix, der både er på reduceret echelon form og af rang $r = n$.

Sætning 2.5.11 *Lad $\mathbb{F}$ være et legeme. En kvadratisk matrix $A \in M_n(\mathbb{F})$ af orden n er invertibel, hvis og kun hvis den har rang $r = n$. I givet fald findes en endelig følge af operationsmatricer $P_1, \dots, P_k \in M_n(\mathbb{F})$, sådan at $I = A' = P_k \cdots P_1 A$, og dermed*

$$A^{-1} = P_k \cdots P_1.$$

Bevis Vi viser først, at A er invertibel, hvis og kun hvis $r = n$. Lad $A' \in M_n(\mathbb{F})$ være den entydigt bestemte matrix på reduceret echelon from, som er rækkeækvivalent med A, og lad $P_1, \dots, P_k \in M_m(\mathbb{F})$ være operationsmatricer, sådan at $A' = P_k \cdots P_1 A$. Ifølge sætning 2.5.2 er operationsmatricer invertible, og derfor er A altså invertibel, hvis og kun hvis A' er invertibel. Men A' er invertibel, hvis og kun hvis den har rang $r = n$. For hvis $r < n$, da er den n'te række i A' nul, og derfor er $A'\boldsymbol{e}_n = \mathbf{0}$, hvilket viser, at A' ikke er invertibel; og hvis $r = n$, da er $A' = I$ og er derfor specielt invertibel. Dette viser, at A er invertibel, hvis og kun hvis $r = n$. Hvis dette er tilfældet, da er A invertibel og $I = A' = P_k \cdots P_1 A$, og derfor er $A^{-1} = IA^{-1} = P_k \cdots P_1$ som ønsket. □

Korollar 2.5.12 *En $n \times n$-matrix $A \in M_n(\mathbb{F})$ med indgange i et legeme $\mathbb{F}$ er invertibel, hvis og kun hvis $n \times 2n$-matricen $(A \mid I) \in M_{n,2n}(\mathbb{F})$ ved rækkeoperationer kan omdannes til en $n \times 2n$-matrix på formen $(I \mid B) \in M_{n,2n}(\mathbb{F})$, og i givet fald er $A^{-1} = B$.*

Bevis Ifølge sætning 2.5.11 er A invertibel, hvis og kun hvis den ved rækkeoperationer kan omdannes til identitetsmatricen, og dette er tilfældet, hvis og kun hvis $(A \mid I)$ ved rækkeoperationer kan omdannes til en $n \times 2n$-matrix på formen $(I \mid B)$. I givet fald er $(I \mid B)$ på reduceret echelon form og rækkeækvivalent til $(A \mid I)$, og sætning 2.5.6 viser derfor, at den er entydigt bestemt ved $(A \mid I)$. Lad nu $P_1, \dots, P_k \in M_n(\mathbb{F})$ være en følge af operationsmatricer, sådan at $P_k \cdots P_1 A = I$. Da viser sætning 2.5.11, at

$$P_k \cdots P_1 (A \mid I) = (P_k \cdots P_1 A \mid P_k \cdots P_1 I) = (I \mid A^{-1}),$$

hvilket viser, at $B = A^{-1}$ som ønsket. □

Korollar 2.5.12 giver en algoritme til at afgøre om en kvadratisk matrix er invertibel eller ej og i givet fald finde dens inverse. Denne algoritme forstås lettest ved gennemgang af nogle eksempler.

Eksempel 2.5.13 Vi anvender korollar 2.5.12 til at undersøge om 3×3-matricen

$$A = \begin{pmatrix} 8 & 5 & 3 \\ 2 & 2 & 1 \\ 3 & 1 & 1 \end{pmatrix}$$

er invertibel og i givet fald finde dens inverse matrix. Vi omdanner derfor 3×6-matricen

$(A \mid I)$ til en matrix på reduceret echelon-form ved hjælp af rækkeoperationer:

$$
\begin{aligned}
(A \mid I) = &\left(\begin{array}{ccc|ccc} 8 & 5 & 3 & 1 & 0 & 0 \\ 2 & 2 & 1 & 0 & 1 & 0 \\ 3 & 1 & 1 & 0 & 0 & 1 \end{array}\right) \begin{array}{l} +(-3)R_3 \\ +(-1)R_3 \\ \end{array} \\
&\left(\begin{array}{ccc|ccc} -1 & 2 & 0 & 1 & 0 & -3 \\ -1 & 1 & 0 & 0 & 1 & -1 \\ 3 & 1 & 1 & 0 & 0 & 1 \end{array}\right) \begin{array}{l} \\ +(-1)R_1 \\ +3R_1 \end{array} \\
&\left(\begin{array}{ccc|ccc} -1 & 2 & 0 & 1 & 0 & -3 \\ 0 & -1 & 0 & -1 & 1 & 2 \\ 0 & 7 & 1 & 3 & 0 & -8 \end{array}\right) \begin{array}{l} +2R_2 \\ \\ +7R_2 \end{array} \\
&\left(\begin{array}{ccc|ccc} -1 & 0 & 0 & -1 & 2 & 1 \\ 0 & -1 & 0 & -1 & 1 & 2 \\ 0 & 0 & 1 & -4 & 7 & 6 \end{array}\right) \begin{array}{l} (-1)\cdot R_1 \\ (-1)\cdot R_2 \\ \end{array} \\
(I \mid B) = &\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -2 & -1 \\ 0 & 1 & 0 & 1 & -1 & -2 \\ 0 & 0 & 1 & -4 & 7 & 6 \end{array}\right)
\end{aligned}
$$

Her har vi som tidligere indikeret de indgange, vi ønsker at ændre med rødt, og de ledende indgange med blåt. Vi konkluderer nu for det første, at A er invertibel, og for det andet, at

$$A^{-1} = B = \begin{pmatrix} 1 & -2 & -1 \\ 1 & -1 & -2 \\ -4 & 7 & 6 \end{pmatrix}$$

er den inverse matrix.

Eksempel 2.5.14 Vi anvender korollar 2.5.12 til at undersøge om 3×3-matricen

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

er invertibel og i givet fald finde dens inverse matrix. Vi omdanner derfor 3×6-matricen

$(A \mid I)$ til en matrix på reduceret echelon-form ved hjælp af rækkeoperationer:

$$
\begin{aligned}
(A \mid I) = &\left(\begin{array}{rrr|rrr} 2 & 3 & 1 & 1 & 0 & 0 \\ -1 & 0 & -2 & 0 & 1 & 0 \\ 1 & 2 & 0 & 0 & 0 & 1 \end{array}\right) \begin{array}{l} +(-2)R_3 \\ +R_3 \\ \end{array} \\
&\left(\begin{array}{rrr|rrr} 0 & -1 & 1 & 1 & 0 & -2 \\ 0 & 2 & -2 & 0 & 1 & 1 \\ 1 & 2 & 0 & 0 & 0 & 1 \end{array}\right) R_1 \leftrightarrow R_3 \\
&\left(\begin{array}{rrr|rrr} 1 & 0 & 1 & 1 & 0 & -2 \\ 0 & 2 & -2 & 0 & 1 & 1 \\ 0 & -1 & 1 & 1 & 0 & -2 \end{array}\right) +2R_3 \\
&\left(\begin{array}{rrr|rrr} 1 & 0 & 1 & 1 & 0 & -2 \\ 0 & 0 & 0 & -1 & 1 & 2 \\ 0 & -1 & 1 & 1 & 0 & -2 \end{array}\right) R_2 \leftrightarrow R_3 \\
&\left(\begin{array}{rrr|rrr} 1 & 0 & 1 & 1 & 0 & -2 \\ 0 & -1 & 1 & 1 & 0 & -2 \\ 0 & 0 & 0 & -1 & 1 & 2 \end{array}\right)
\end{aligned}
$$

Her har vi som tidligere indikeret de indgange, vi ønsker at ændre med rødt, og de ledende indgange med blåt. Matricen er på echelon form, og vi kan derfor aflæse, at A har rang $r = 2 < 3 = n$, og derfor er A ikke er invertibel.

Hvis A er en $m \times n$-matrix, så påviste vi i sætning 1.2.17 en sammenhæng mellem rangen af A og mængden af løsninger til ligningssystemet $A\boldsymbol{x} = \boldsymbol{b}$, og vi minder om, at denne sammenhæng var opdelt i de tre tilfælde $r = m$, $r = n$, og $r = m = n$. Vi bemærker nu, at hvis A er kvadratisk, sådan at $m = n$, da er disse tre betingelser sammenfaldende og svarer alle til betingelsen, at $r = n$. Vi viser nu, at denne tilsyneladende uskyldige bemærkning har følgende bemærkelsesværdige konsekvenser.

Sætning 2.5.15 *Hvis A og B er $n \times n$-matrix med indgange i et legeme $\mathbb{F}$, da er følgende udsagn ækvivalente:*

(1) $AB = I$.

(2) $BA = I$.

(3) $AB = I$ *og* $BA = I$.

Bevis Det er klart, at (3) medfører både (1) og (2), så vi skal vise, at (1) medfører (3), og at (2) medfører (3). Vi viser nu, at (1) medfører (3), og ved at ombytte A og B viser det samme argument, at også (2) medfører (3). Ifølge sætning 2.5.11 er (3) ækvivalent til udsagnet, at $r = n$, så det er nok at vise, at (1) medfører, at $r = n$. Men $AB = I$ medfører, at ligningssystemet $A\boldsymbol{x} = \boldsymbol{b}$ har mindst en løsning, nemlig $\boldsymbol{x} = B\boldsymbol{b}$, og derfor viser sætning 1.2.17, at $r = n$ som ønsket. Dette beviser sætningen. □

Ligeledes har vi følgende korollar til sætning 1.2.17, eller dens oversættelse, sætning 2.5.9, til lineære abildninger:

Korollar 2.5.16 *Hvis $\mathbb{F}$ er et legeme og $f : \mathbb{F}^n \to \mathbb{F}^n$ en lineær afbildning, da er de følgende udsagn ækvivalente:*

(1) *Afbildningen $f : \mathbb{F}^n \to \mathbb{F}^n$ er surjektiv.*

(2) *Afbildningen $f : \mathbb{F}^n \to \mathbb{F}^n$ er injektiv.*

(3) *Afbildningen $f : \mathbb{F}^n \to \mathbb{F}^n$ er bijektiv.*

Bevis Dette følger umiddelbart fra sætning 2.5.9 idet, ifølge sætningen, så er alle tre betingelse ækvivalente med $r = n$ fordi vi er i det tilfælde hvor $m = n$. □

Bemærkning 2.5.17 Sætning 2.5.15 gælder kun for *kvadratiske* matricer, også selvom matrixprodukterne AB og BA begge findes. Et modeksempel er

$$AB = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \end{pmatrix} = I_1 \qquad \text{og} \qquad BA = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \neq I_2.$$

Bemærkning 2.5.18 Vi skal senere i kapitel 3 også anvende søjleoperationer til at definere og udregne determinanten. Ligesom rækkeoperationer svarer til at gange med operationsmatricer fra venstre, så svarer søjleoperationer til at gange med disse fra højre. Hvis $A \in M_{m,n}(\mathbb{F})$ er en $m \times n$-matrix med indgange i et legeme $\mathbb{F}$, da gælder:

(1) For alle $n \geq j \geq 1$ og $0 \neq c \in \mathbb{F}$ er $AM_j(c)$ den matrix, der fremkommer fra A ved at gange den j'te søjle med c fra højre.

(2) For alle $n \geq k \neq j \geq 1$ og $c \in \mathbb{F}$ er $AS_{kj}(c)$ den matrix, der fremkommer fra A ved at adderer den k'te søjle gange c til den j'te søjle.

(3) For alle $n \geq k > j \geq 1$ er AT_{kj} den matrix, der fremkommer fra A ved at ombytte den j'te søjle og den k'te søjle.

Her er alle operationsmatricerne kvadratiske matricer af orden n. Vi understreger dog, at søjleoperationer *ikke* kan bruges til at løse lineære ligningssystemer, da de modsat rækkeoperationer ændrer løsningsmængden.

2.6 Hermitiske former

På legemet af komplekse tal $\mathbb{C}$ findes der en meget interessant ekstra struktur, nemlig kompleks konjugering, som er defineret ved $(a+ib)^* = a - ib$. Der findes en tilsvarende struktur på skævlegemet af kvaternioner $\mathbb{H}$, nemlig kvaternionisk konjugation, som er defineret ved $(a+ib+jc+kd)^* = a-ib-jc-kd$. Disse er begge eksempler på det følgende begreb, som vi kalder en skævinvolution. En afgørende egenskab ved en skævinvolution er, at den inverterer multiplikationsordenen, og derfor vil vi i dette afsnit afbejde med skævlegemer, hvor den kommutative lov ikke antages at gælde for multiplikationen.

Definition 2.6.1 En *skævinvolution* på et skævlegeme $\mathbb{F}$ er en afbildning

$$(-)^* : \mathbb{F} \to \mathbb{F}$$

med de følgende egenskaber:

(I1) For alle $a, b \in \mathbb{F}$ er $(a+b)^* = a^* + b^*$.

(I2) For alle $a, b \in \mathbb{F}$ er $(ab)^* = b^* a^*$.

(I3) For alle $a \in \mathbb{F}$ er $(a^*)^* = a$.

Vi siger "skæv" og "involution" for at antyde henholdsvis (I2) og (I3). Hvis $\mathbb{F}$ er et legeme, da er identitetsafbildningen en skævinvolution på $\mathbb{F}$. Dette er specielt tilfældet for både $\mathbb{F} = \mathbb{R}$ og $\mathbb{F} = \mathbb{C}$, men for $\mathbb{F} = \mathbb{C}$ vil vi hovedsageligt betragte den skævinvolution, der er givet ved kompleks konjugering. På skævlegemet $\mathbb{H}$ er identitetsafbildningen imidlertid ikke en skævinvolution, da den ikke tilfredsstiller (I2), men som nævnt så er kvaternionisk konjugering er en skævinvolution på $\mathbb{H}$.

Definition 2.6.2 Lad $\mathbb{F}$ være et skævlegeme med skævinvolution $(-)^* : \mathbb{F} \to \mathbb{F}$. Da er den *adjungerede matrix* af en $m \times n$-matrix

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

med indgange i $\mathbb{F}$ den $n \times m$-matrix med indgange i $\mathbb{F}$, der er givet ved

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

En kvadratisk matrix A med indgange i $\mathbb{F}$ er *hermitisk*, hvis $A^* = A$.

Vi bemærker, at antallet af rækker i A^* er lig med antallet af søjler i A og vice versa. Det giver altså kun mening at spørge om $A^* = A$, hvis A er en kvadratisk matrix.

Bemærkning 2.6.3 Hvis $(-)^* : \mathbb{F} \to \mathbb{F}$ er identitetsafbildningen, da benyttes i stedet den følgende terminologi: den adjungerede matrix A^* kaldes for den *transponerede matrix* og skrives A^t; og en hermitisk matrix siges at være en *symmetrisk* matrix.

Eksempel 2.6.4 Vi anskueliggør definitionen af den adjungerede matrix med følgende eksempler på matricer og deres adjungerede matricer.

$$A = \begin{pmatrix} 1 & 0 \\ 2 & -1 \\ 0 & 3 \end{pmatrix}, \qquad A^* = \begin{pmatrix} 1 & 2 & 0 \\ 0 & -1 & 3 \end{pmatrix}.$$

$$B = \begin{pmatrix} 2 & -1 & -3 \\ 0 & 1 & 2 \\ 0 & 0 & -1 \end{pmatrix}, \qquad B^* = \begin{pmatrix} 2 & 0 & 0 \\ -1 & 1 & 0 \\ -3 & 2 & -1 \end{pmatrix}.$$

$$C = \begin{pmatrix} -2 & i \\ -i & -1 \end{pmatrix}, \qquad C^* = \begin{pmatrix} -2 & i \\ -i & -1 \end{pmatrix}.$$

Matricerne B og C er kvadratiske, og det er de adjungerede matricer B^* og C^* dermed også. Matricen C er endvidere hermitisk. Her har vi anvendt kompleks konjugering som skævinvolution på de komplekse tal.

Vi viser nu, at adjungering af matricer har præcis de samme egenskaber som en skævinvolution, hvilket er grunden til, at vi anvender den samme notation for begge.

Sætning 2.6.5 *Lad $\mathbb{F}$ være et skævlegeme og lad $(-)^* : \mathbb{F} \to \mathbb{F}$ være en skævinvolution. For matricer af passende dimensioner med indgange i $\mathbb{F}$ gælder følgende:*

(I1) $(A+B)^* = A^* + B^*$

(I2) $(AB)^* = B^* A^*$

(I3) $(A^*)^* = A$

Bevis For $A = (a_{ij}) \in M_{m,n}(\mathbb{F})$ skriver vi $A^* = (a'_{ij}) \in M_{n,m}(\mathbb{F})$, sådan at $a'_{ij} = (a_{ji})^*$ per definition af den adjungerede matrix. For at vise (I1) lader vi $A, B \in M_{m,n}(\mathbb{F})$ og sætter $C = A + B \in M_{m,n}(\mathbb{F})$. Udregningen

$$c'_{ij} = c^*_{ji} = (a_{ji} + b_{ji})^* = a^*_{ji} + b^*_{ji} = a'_{ij} + b'_{ij}$$

viser da, at $C^* = (A+B)^*$ som ønsket. For at vise (I2) lader vi tilsvarende $A \in M_{m,n}(\mathbb{F})$ og $B \in M_{n,p}(\mathbb{F})$ og sætter $C = AB \in M_{m,p}(\mathbb{F})$, hvorefter udregningen

$$c'_{ik} = c^*_{ki} = (\sum_{j=1}^{n} a_{kj} b_{ji})^* = \sum_{j=1}^{n} (a_{kj} b_{ji})^* = \sum_{j=1}^{n} b^*_{ji} a^*_{kj} = \sum_{j=1}^{n} b'_{ij} a'_{jk}$$

viser, at $C^* = B^* A^*$ som ønsket. For til slut at vise (I3) lader vi $A \in M_{m,n}(\mathbb{F})$ og sætter $C = A^* \in M_{n,m}(\mathbb{F})$. Udregningen

$$c'_{ij} = c^*_{ji} = (a'_{ji})^* = (a^*_{ij})^* = a_{ij}$$

viser, at $C^* = A$, hvilket beviser (I3) og dermed sætningen. □

Definition 2.6.6 Lad $\mathbb{F}$ være et skævlegeme, lad $(-)^* : \mathbb{F} \to \mathbb{F}$ være en skævinvolution, og lad n være et naturligt tal. En *hermitisk form* på $\mathbb{F}^n$ er en afbildning

$$\langle -, - \rangle : \mathbb{F}^n \times \mathbb{F}^n \to \mathbb{F},$$

for hvilken der gælder følgende:

(H1) For alle $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathbb{F}^n$ er $\langle \boldsymbol{x}, \boldsymbol{y} + \boldsymbol{z} \rangle = \langle \boldsymbol{x}, \boldsymbol{y} \rangle + \langle \boldsymbol{x}, \boldsymbol{z} \rangle$.

(H2) For alle $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{F}^n$ og $a \in \mathbb{F}$ er $\langle \boldsymbol{x}, \boldsymbol{y} \cdot a \rangle = \langle \boldsymbol{x}, \boldsymbol{y} \rangle \cdot a$.

(H3) For alle $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathbb{F}^n$ er $\langle \boldsymbol{x} + \boldsymbol{y}, \boldsymbol{z} \rangle = \langle \boldsymbol{x}, \boldsymbol{z} \rangle + \langle \boldsymbol{y}, \boldsymbol{z} \rangle$.

(H4) For alle $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{F}^n$ og $a \in \mathbb{F}$ er $\langle \boldsymbol{x} \cdot a, \boldsymbol{y} \rangle = a^* \cdot \langle \boldsymbol{x}, \boldsymbol{y} \rangle$.

(H5) For alle $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{F}^n$ er $\langle \boldsymbol{y}, \boldsymbol{x} \rangle = \langle \boldsymbol{x}, \boldsymbol{y} \rangle^*$.

Hermitiske former er opkaldt efter Charles Hermite, der også er kendt for at have bevist, at grundtallet e for den naturlige logaritme er transcendentalt.

Vi skal nu vise, at der er en kanonisk 1-1 korrespondance mellem hermitiske former på $\mathbb{F}^n$ og hermitiske matricer $A \in M_n(\mathbb{F})$. Vi bemærker hertil, at hvis $\boldsymbol{x} \in \mathbb{F}^n = M_{n,1}(\mathbb{F})$ er en søjlevektor af dimension n, så er $\boldsymbol{x}^* \in M_{1,n}(\mathbb{F})$ en rækkevektor af dimension n.

Sætning 2.6.7 *Lad $\mathbb{F}$ være et skævlegeme, lad $(-)^* : \mathbb{F} \to \mathbb{F}$ være en skævinvolution, og lad n være et naturligt tal.*

(1) *Hvis $\langle -, - \rangle : \mathbb{F}^n \times \mathbb{F}^n \to \mathbb{F}$ er en hermitisk form, så er matricen $A = (a_{ij}) \in M_n(\mathbb{F})$ med indgange $a_{ij} = \langle \boldsymbol{e}_i, \boldsymbol{e}_j \rangle$ en hermitisk matrix.*

(2) *Hvis $A \in M_n(\mathbb{F})$ er en hermitisk matrix, så er afbildningen $\langle -, - \rangle : \mathbb{F}^n \times \mathbb{F}^n \to \mathbb{F}$ givet ved $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \boldsymbol{x}^* A \boldsymbol{y}$ en hermitisk form.*

Bevis Påstanden (1) følger fra (H5), idet $a_{ij} = \langle \boldsymbol{v}_i, \boldsymbol{v}_j \rangle = \langle \boldsymbol{v}_j, \boldsymbol{v}_i \rangle^* = a_{ji}^*$, og for at bevise påstanden (2), lader vi $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathbb{F}^n$ og $a \in \mathbb{F}$ og udregner

$$\begin{aligned}
\langle \boldsymbol{x}, \boldsymbol{y} + \boldsymbol{z} \rangle &= \boldsymbol{x}^* A(\boldsymbol{y} + \boldsymbol{z}) = \boldsymbol{x}^* A \boldsymbol{y} + \boldsymbol{x}^* A \boldsymbol{z} = \langle \boldsymbol{x}, \boldsymbol{y} \rangle + \langle \boldsymbol{x}, \boldsymbol{z} \rangle \\
\langle \boldsymbol{x}, \boldsymbol{y} \cdot a \rangle &= \boldsymbol{x}^* A \boldsymbol{y} a = \langle \boldsymbol{x}, \boldsymbol{y} \rangle \cdot a \\
\langle \boldsymbol{x} + \boldsymbol{y}, \boldsymbol{z} \rangle &= (\boldsymbol{x} + \boldsymbol{y})^* A \boldsymbol{z} = \boldsymbol{x}^* A \boldsymbol{z} + \boldsymbol{y}^* A \boldsymbol{z} = \langle \boldsymbol{x}, \boldsymbol{z} \rangle + \langle \boldsymbol{y}, \boldsymbol{z} \rangle \\
\langle \boldsymbol{x} \cdot a, \boldsymbol{y} \rangle &= (\boldsymbol{x} \cdot a)^* A \boldsymbol{y} = a^* \cdot \boldsymbol{x}^* A \boldsymbol{y} = a^* \cdot \langle \boldsymbol{x}, \boldsymbol{y} \rangle \\
\langle \boldsymbol{y}, \boldsymbol{x} \rangle &= \boldsymbol{y}^* A \boldsymbol{x} = \boldsymbol{y}^* A^* (\boldsymbol{x}^*)^* = (\boldsymbol{x}^* A \boldsymbol{y})^* = \langle \boldsymbol{x}, \boldsymbol{y} \rangle^*,
\end{aligned}$$

hvilket viser påstanden. Her har vi anvendt sætning 2.6.5. □

Vi kalder matricen $A = (\langle \boldsymbol{e}_i, \boldsymbol{e}_j \rangle)$ i sætning 2.6.7 (1) for *matricen, der repræsenterer den hermitiske form* $\langle -, - \rangle$ *med hensyn til standardbasen* $(\boldsymbol{e}_1, \dots, \boldsymbol{e}_n)$ *for* $\mathbb{F}^n$, og omvendt kalder vi den hermitiske form $\langle -, - \rangle$ i sætning 2.6.7 (2) for den *hermitiske form, der hører til matricen* A *med hensyn til standardbasen* $(\boldsymbol{e}_1, \dots, \boldsymbol{e}_n)$ *for* $\mathbb{F}^n$.

Definition 2.6.8 Hvis $\mathbb{F}$ er et skævlegeme, og $(-)^* \colon \mathbb{F} \to \mathbb{F}$ er en skævinvolution, så siges en hermitisk form $\langle -, - \rangle$ på $\mathbb{F}^n$ at være *ikke-singulær*, hvis den matrix $A \in M_n(\mathbb{F})$, der repræsenter $\langle -, - \rangle$ med hensyn til standardbasen $(\boldsymbol{e}_1, \dots, \boldsymbol{e}_n)$ er invertibel.

Eksempel 2.6.9 (1) Det standard indre produkt på $\mathbb{F}^n$ givet ved

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \boldsymbol{x}^* \boldsymbol{y} = x_1^* y_1 + x_2^* y_2 + \dots + x_n^* y_n$$

er den hermitiske form på $\mathbb{F}^n$, der hører til identitetsmatricen $I_n \in M_n(\mathbb{F})$ med hensyn til standardbasen $(\boldsymbol{e}_1, \dots, \boldsymbol{e}_m)$. Den er ikke-singulær.

(2) Den Minkowski hermitiske form på $\mathbb{F}^n$ givet ved

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \boldsymbol{x}^* \operatorname{diag}(-1, 1, \dots, 1) \boldsymbol{y} = -x_1^* y_1 + x_2^* y_2 + \dots + x_n^* y_n$$

er den hermitiske form på $\mathbb{F}^n$, der hører til diagonalmatricen $\operatorname{diag}(-1, 1, \dots, 1) \in M_n(\mathbb{F})$ med hensyn til standardbasen $(\boldsymbol{e}_1, \dots, \boldsymbol{e}_n)$. Den er ikke-singulær.

(3) Den hyperbolske form på $\mathbb{F}^2$ er den hermitiske form

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \boldsymbol{x}^* \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \boldsymbol{y} = x_1^* y_2 + x_2^* y_1,$$

der hører til den indikerede hermitiske matrix $H \in M_2(\mathbb{F})$. Den er ikke-singulær.

Definition 2.6.10 Lad $\mathbb{F}$ være et skævlegeme, lad $(-)^* : \mathbb{F} \to \mathbb{F}$ være en skævinvolution, og lad $\langle -,- \rangle$ og $\langle -,- \rangle'$ være hermitiske former på henholdsvis $\mathbb{F}^n$ og $\mathbb{F}^m$. En lineær afbildning $f : \mathbb{F}^n \to \mathbb{F}^m$ er en *lineær isometri* med hensyn til $\langle -,- \rangle$ og $\langle -,- \rangle'$, hvis

$$\langle f(\boldsymbol{x}), f(\boldsymbol{y}) \rangle' = \langle \boldsymbol{x}, \boldsymbol{y} \rangle$$

for alle $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{F}^n$, og den er en *isometrisk isomorfi* med hensyn til $\langle -,- \rangle$ og $\langle -,- \rangle'$, hvis den tillige er en isomorfi.

Eksempel 2.6.11 (1) En isometrisk isomorfi af $f : \mathbb{R}^n \to \mathbb{R}^n$ med hensyn til det standard indre produkt $\langle -,- \rangle$ kaldes for en euklidisk transformation.
(2) En isometrisk isomorfi af $f : \mathbb{R}^n \to \mathbb{R}^n$ med hensyn til den Minkowski hermitiske form $\langle -,- \rangle$ kaldes for en Lorentz transformation.

Sætning 2.6.12 *Lad $\mathbb{F}$ være et skævlegeme og lad $(-)^* : \mathbb{F} \to \mathbb{F}$ være en skævinvolution. Lad $\langle -,- \rangle$ og $\langle -,- \rangle'$ være hermitiske former på $\mathbb{F}^n$ og $\mathbb{F}^m$, og lad $A \in M_n(\mathbb{F})$ og $A' \in M_m(\mathbb{F})$ være de matricer, der repræsenterer $\langle -,- \rangle$ og $\langle -,- \rangle'$ med hensyn til standardbaserne. En lineær afbildning $f : \mathbb{F}^n \to \mathbb{F}^m$ er en lineær isometri med hensyn til $\langle -,- \rangle$ og $\langle -,- \rangle'$, hvis og kun hvis den matrix $P \in M_{m,n}(\mathbb{F})$, der repræsenterer denne afbildning med hensyn til de respektive standardbaser, tilfredsstiller, at*

$$P^* A' P = A.$$

Bevis Hvis $P^* A' P = A$, da viser udregningen

$$\langle f(\boldsymbol{x}), f(\boldsymbol{y}) \rangle' = (P\boldsymbol{x})^* A' P \boldsymbol{y} = \boldsymbol{x}^* P^* A' P \boldsymbol{y} = \boldsymbol{x}^* A \boldsymbol{y} = \langle \boldsymbol{x}, \boldsymbol{y} \rangle,$$

hvor $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{F}^n$, at $f : \mathbb{F}^n \to \mathbb{F}^m$ er en lineær isometri. Omvendt, hvis $f : \mathbb{F}^n \to \mathbb{F}^m$ er en lineær isometri, da viser udregningen

$$\boldsymbol{e}_i^* P^* A' P \boldsymbol{e}_j = (P\boldsymbol{e}_i)^* A' P \boldsymbol{e}_j = \langle f(\boldsymbol{e}_i), f(\boldsymbol{e}_j) \rangle' = \langle \boldsymbol{e}_i, \boldsymbol{e}_j \rangle = \boldsymbol{e}_i A \boldsymbol{e}_j,$$

at de (i,j)'te indgange i matricerne $P^* A' P$ og A er identiske, for alle $1 \leq i,j \leq n$, og derfor gælder der, at $P^* A' P = A$ som ønsket. $\square$

Eksempel 2.6.13 Ifølge sætning 2.6.12 viser udregningen

$$P^* A' P = \begin{pmatrix} -1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} -1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} = A,$$

at afbildningen $f : \mathbb{F}^2 \to \mathbb{F}^2$ givet ved

$$f(\boldsymbol{x}) = P\boldsymbol{x} = \begin{pmatrix} -1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \boldsymbol{x}$$

er en isometrisk isomorfi fra $\mathbb{F}^2$ med den Minkowski hermitiske form til $\mathbb{F}^2$ med den hyperbolske form. Her har vi brugt, at $1/\sqrt{2} \in \mathbb{F}$, og i modsat fald findes der ikke nogen sådan isometrisk isomorfi.

3 Determinant

Givet n vektorer i $\mathbb{R}^n$, ønsker vi at definere og udregne det n-dimensionale volume af det parallelepipedum, de n vektorer udspænder. Dette problem fører naturligt til definitionen af determinanten af en $n \times n$-matrix med indgange i $\mathbb{R}$, og denne definition viser sig da at være meningsfuld for en $n \times n$-matrix med indgange i et vilkårligt legeme $\mathbb{F}$. Definitionen nødvendiggør, at multiplikation af skalarer opfylder den kommutative lov, så vi skal derfor i dette kapitel for første gang bruge, at der for skalarer $a, b \in \mathbb{F}$ gælder, at $ab = ba$. En vigtig teoretisk anvendelse af determinanten er, at den afgør, hvorvidt en kvadratisk matrix er invertibel eller ej, og deraf navnet *determinant*, som skyldes Gauss. Determinanten udregnes lettest ved at benytte dens definerende egenskaber. Denne metode kan minde om Gauss elimination med den væsentlige forskel, at nogle operationer ændrer determinanten, men på en kontrolleret måde. Man skal dog være forsigt med ikke at overse denne forskel. Endelig findes der en lukket formel for determinanten af en $n \times n$-matrix, som skyldes Leibniz, men denne formel indeholder $n!$ summander og er mest af teoretisk interesse.

3.1 Determinant af 2×2-matrix

Lad os først betragte det 1-dimensionale tilfælde. Det 1-dimensionale parallelepipedum udspændt af en vektor $\boldsymbol{a}_1 = (a_{11}) \in \mathbb{R}^1$ defineres som intervallet

$$P = \{\boldsymbol{a}_1 c_1 \in \mathbb{R}^1 \mid 0 \leq c_1 \leq 1\} \subset \mathbb{R}^1,$$

og vi definerer det 1-dimensionale volume eller længden af P til at være absolutværdien

$$\operatorname{vol}(P) = |a_{11}|.$$

Tilsvarende er det 2-dimensionale parallelepipedum udspændt af de 2 vektorer

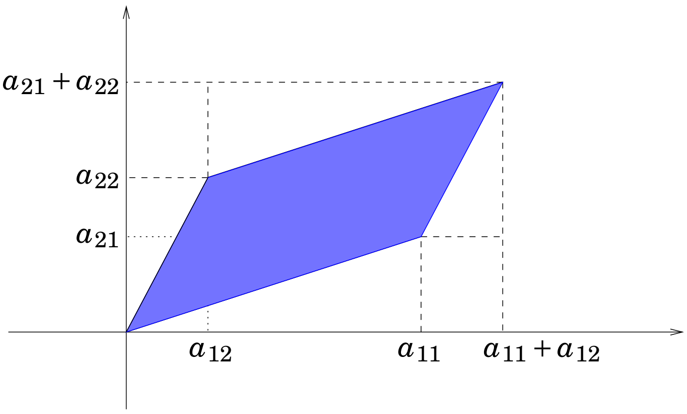
$$\boldsymbol{a}_1 = \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}, \qquad \boldsymbol{a}_2 = \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}$$

defineret til at være parallelogrammet

$$P = \{\boldsymbol{a}_1 c_1 + \boldsymbol{a}_2 c_2 \mid 0 \leqslant c_1, c_2 \leqslant 1\} \subset \mathbb{R}^2;$$

se Figur 3.1 for en illustration af denne delmængde. Som vores notation antyder, vil vi tænke på de to vektorer $\boldsymbol{a}_1, \boldsymbol{a}_2 \in \mathbb{R}^2$ som søjlevektorerne i matricen

$$A = \begin{pmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in M_2(\mathbb{R}).$$

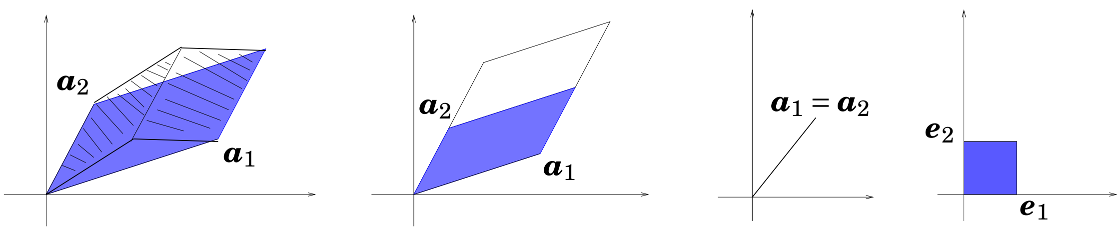


Figur 3.1: Parallelogrammet udspændt af to vektorer

Vi forestiller os, at det 2-dimensionale volumen eller areal af parallelogrammet P tilsvarende skal være givet som absolutværdien

$$\operatorname{vol}(P) = |\det(A)|$$

af en skalar $\det(A)$, der afhænger af 2×2-matricen A. Vi diskuterer nu, hvordan vi vil forvente, at afbildningen $\det\colon M_2(\mathbb{R}^2) \to \mathbb{R}$ skal opføre sig, hvilket figuren illusterer.



Figur 3.2: Determinanten opfylder (1) $\det\begin{pmatrix}\boldsymbol{b}_1+\boldsymbol{c}_1 & \boldsymbol{a}_2\end{pmatrix} = \det\begin{pmatrix}\boldsymbol{b}_1 & \boldsymbol{a}_2\end{pmatrix} + \det\begin{pmatrix}\boldsymbol{c}_1 & \boldsymbol{a}_2\end{pmatrix}$, (2) $\det\begin{pmatrix}\boldsymbol{a}_1 & \boldsymbol{a}_2 c\end{pmatrix} = \det\begin{pmatrix}\boldsymbol{a}_1 & \boldsymbol{a}_2\end{pmatrix}c$, (3) $\det\begin{pmatrix}\boldsymbol{a} & \boldsymbol{a}\end{pmatrix} = 0$ (4) $\det\begin{pmatrix}\boldsymbol{e}_1 & \boldsymbol{e}_2\end{pmatrix} = 1$

Hvis vektoren $\boldsymbol{a}_1 = \boldsymbol{b}_1 + \boldsymbol{c}_1$ er en sum af to vektorer, da forventer vi rimeligvis, at arealet af parallelogrammet udspændt af $\boldsymbol{a}_1$ og $\boldsymbol{a}_2$ er lig med summen af arealerne af parallelogrammerne udspændt af henholdsvis $\boldsymbol{b}_1$ og $\boldsymbol{a}_2$ og af $\boldsymbol{c}_1$ og $\boldsymbol{a}_2$; tilsvarende, hvis $\boldsymbol{a}_2 = \boldsymbol{b}_2 + \boldsymbol{c}_2$ er en sum af to vektorer, da vil vi forvente, at arealet af parallologrammet udspændt af $\boldsymbol{a}_1$ og $\boldsymbol{a}_2$ er lig med summen af arealerne af parallelogrammerne udspændt at henholdsvis $\boldsymbol{a}_1$ og $\boldsymbol{b}_1$ og af $\boldsymbol{a}_1$ og $\boldsymbol{c}_2$. Hvis vi erstatter $\boldsymbol{a}_1$ med $\boldsymbol{a}_1 \cdot c$, men bibeholder $\boldsymbol{a}_2$, da vil vi forvente, at arealet af de tilsvarende parallelogrammer ganges med $|c|$; tilsvarende, hvis vi erstatter $\boldsymbol{a}_2$ med $\boldsymbol{a}_2 \cdot c$, men bibeholder $\boldsymbol{a}_1$, da vil vi igen forvente, at arealet af de tilsvarende parallelogrammer ganges med $|c|$. Hvis $\boldsymbol{a}_1 = \boldsymbol{a}_2$, da forventer vi, at arealet af det parallelogram, de udspænder, er lig med 0. Endelig, vælger vi at normere vores arealmål, sådan at enhedskvadratet har areal 1. Vi ønsker altså, at afbildningen $\det\colon M_2(\mathbb{R}) \to \mathbb{R}$ har de følgende egenskaber:

(D1) Hvis $\boldsymbol{a}_1 = \boldsymbol{b}_1 + \boldsymbol{c}_1$ eller $\boldsymbol{a}_2 = \boldsymbol{b}_2 + \boldsymbol{c}_2$, da gælder

$$\det\begin{pmatrix}\boldsymbol{b}_1 & \boldsymbol{a}_2\end{pmatrix} + \det\begin{pmatrix}\boldsymbol{c}_1 & \boldsymbol{a}_2\end{pmatrix} = \det\begin{pmatrix}\boldsymbol{a}_1 & \boldsymbol{a}_2\end{pmatrix} = \det\begin{pmatrix}\boldsymbol{a}_1 & \boldsymbol{b}_2\end{pmatrix} + \det\begin{pmatrix}\boldsymbol{a}_1 & \boldsymbol{c}_2\end{pmatrix}.$$

(D2) Hvis $\boldsymbol{a}_1$ erstattes af $\boldsymbol{a} \cdot c$ eller $\boldsymbol{a}_2$ erstattes af $\boldsymbol{a}_2 \cdot c$, da gælder

$$\det\begin{pmatrix}\boldsymbol{a}_1 \cdot c & \boldsymbol{a}_2\end{pmatrix} = \det\begin{pmatrix}\boldsymbol{a}_1 & \boldsymbol{a}_1\end{pmatrix} \cdot c = \det\begin{pmatrix}\boldsymbol{a}_1 & \boldsymbol{a}_2 \cdot c\end{pmatrix}.$$

(D3) Hvis $\boldsymbol{a}_1 = \boldsymbol{a}_2$, da er $\det\begin{pmatrix}\boldsymbol{a}_1 & \boldsymbol{a}_1\end{pmatrix} = 0$.

(D4) Hvis $\boldsymbol{a} = \boldsymbol{e}_1$ og $\boldsymbol{a}_2 = \boldsymbol{e}_2$ er standardenhedsvektorerne, da er $\det\begin{pmatrix}\boldsymbol{e}_1 & \boldsymbol{e}_2\end{pmatrix} = 1$.

Det viser sig, at der netop findes een sådan afbildning:

Sætning 3.1.1 *Lad* $\mathbb{F}$ *være et legeme. Afbildningen* $\det\colon M_2(\mathbb{F}) \to \mathbb{F}$ *defineret ved*

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

opfylder (D1)–(D4) *og er entydigt bestemt herved.*

Bemærkning 3.1.2 (1) I mere læsevenlig form siger formlen i sætning 3.1.1, at

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

(2) Man kan udregne arealet af parallelogrammet i figur 3.1 ved at begynde med arealet $(a_{11}+a_{12})(a_{21}+a_{22})$ af det store rektangel og derfra trække arealet af de rektangler, der tilsammen udgør området udenfor parallelogrammet. Denne udregning giver samme formel $a_{11}a_{22} - a_{12}a_{21}$.

Før vi beviser sætning 3.1.1, viser vi først, at egenskaberne (D1)–(D3) medfører de følgende yderligere egenskaber:

(D5) Hvis vektorerne $\boldsymbol{a}_1$ og $\boldsymbol{a}_2$ ombyttes, da gælder

$$\det\begin{pmatrix}\boldsymbol{a}_1 & \boldsymbol{a}_2\end{pmatrix} = -\det\begin{pmatrix}\boldsymbol{a}_2 & \boldsymbol{a}_1\end{pmatrix}.$$

(D6) Hvis et multiplum af en af vektorerne $\boldsymbol{a}_1$ og $\boldsymbol{a}_2$ adderes til den anden, da gælder

$$\det\begin{pmatrix}\boldsymbol{a}_1 + \boldsymbol{a}_2 \cdot c & \boldsymbol{a}_2\end{pmatrix} = \det\begin{pmatrix}\boldsymbol{a}_1 & \boldsymbol{a}_2\end{pmatrix} = \begin{pmatrix}\boldsymbol{a}_1 & \boldsymbol{a}_1 \cdot c + \boldsymbol{a}_2\end{pmatrix}.$$

For (D5) følger fra udregningen

$$\begin{aligned}\det\begin{pmatrix}\boldsymbol{a}_1 & \boldsymbol{a}_2\end{pmatrix}+\det\begin{pmatrix}\boldsymbol{a}_2 & \boldsymbol{a}_1\end{pmatrix} &= \det\begin{pmatrix}\boldsymbol{a}_1 & \boldsymbol{a}_2\end{pmatrix}+\det\begin{pmatrix}\boldsymbol{a}_1 & \boldsymbol{a}_1\end{pmatrix}+\det\begin{pmatrix}\boldsymbol{a}_2 & \boldsymbol{a}_1\end{pmatrix}+\det\begin{pmatrix}\boldsymbol{a}_2 & \boldsymbol{a}_2\end{pmatrix}\\ &= \det\begin{pmatrix}\boldsymbol{a}_1 & \boldsymbol{a}_1+\boldsymbol{a}_2\end{pmatrix}+\det\begin{pmatrix}\boldsymbol{a}_2 & \boldsymbol{a}_1+\boldsymbol{a}_2\end{pmatrix}\\ &= \det\begin{pmatrix}\boldsymbol{a}_1+\boldsymbol{a}_2 & \boldsymbol{a}_1+\boldsymbol{a}_2\end{pmatrix} = 0,\end{aligned}$$

hvor de fire identiteter følger fra henholdsvis (D3), (D1), (D1) og (D3); og på lignende vis følger den første identitet i (D6) fra udregningen

$$\begin{aligned}\det\begin{pmatrix}\boldsymbol{a}_1+\boldsymbol{a}_2\cdot c & \boldsymbol{a}_2\end{pmatrix} &= \det\begin{pmatrix}\boldsymbol{a}_1 & \boldsymbol{a}_2\end{pmatrix}+\det\begin{pmatrix}\boldsymbol{a}_1\cdot c & \boldsymbol{a}_2\end{pmatrix}\\ &= \det\begin{pmatrix}\boldsymbol{a}_1 & \boldsymbol{a}_2\end{pmatrix}+\det\begin{pmatrix}\boldsymbol{a}_2 & \boldsymbol{a}_2\end{pmatrix}\cdot c\\ &= \det\begin{pmatrix}\boldsymbol{a}_1 & \boldsymbol{a}_2\end{pmatrix},\end{aligned}$$

hvor de tre identiteter følger fra (D1), (D2), og (D3); den anden identitet i (D6) vises helt tilsvarende.

Bevis (for Sætning 3.1.1) Man viser let, at den givne afbildning opfylder (D1)–(D4), så vi nøjes med at vise, at disse egenskaber bestemmer afbildningen entydigt. Vi skriver derfor først $\boldsymbol{a}_1 = \boldsymbol{e}_1 a_{11} + \boldsymbol{e}_2 a_{21}$ og får da fra (D1) og (D2), at

$$\det\begin{pmatrix}\boldsymbol{a}_1 & \boldsymbol{a}_2\end{pmatrix} = \det\begin{pmatrix}\boldsymbol{e}_1 a_{11} & \boldsymbol{a}_2\end{pmatrix}+\det\begin{pmatrix}\boldsymbol{e}_2 a_{21} & \boldsymbol{a}_2\end{pmatrix} = \det\begin{pmatrix}\boldsymbol{e}_1 & \boldsymbol{b}\end{pmatrix}a_{11}+\det\begin{pmatrix}\boldsymbol{e}_2 & \boldsymbol{b}\end{pmatrix}a_{21}.$$

Dernæst skriver vi $\boldsymbol{a}_2 = \boldsymbol{e}_1 a_{12} + \boldsymbol{e}_2 a_{22}$ og får fra henholdsvis (D6), (D2) og (D4), at

$$\det\begin{pmatrix}\boldsymbol{e}_1 & \boldsymbol{a}_2\end{pmatrix} = \det\begin{pmatrix}\boldsymbol{e}_1 & \boldsymbol{e}_2 a_{22}\end{pmatrix} = \det\begin{pmatrix}\boldsymbol{e}_1 & \boldsymbol{e}_2\end{pmatrix}a_{22} = a_{22},$$

mens henholdsvis (D6), (D2), (D5) og (D4) viser, at

$$\det\begin{pmatrix}\boldsymbol{e}_2 & \boldsymbol{a}_2\end{pmatrix} = \det\begin{pmatrix}\boldsymbol{e}_2 & \boldsymbol{e}_1 a_{12}\end{pmatrix} = \det\begin{pmatrix}\boldsymbol{e}_2 & \boldsymbol{e}_1\end{pmatrix}a_{12} = -\det\begin{pmatrix}\boldsymbol{e}_1 & \boldsymbol{e}_2\end{pmatrix}a_{12} = -a_{12}.$$

Tilsammen viser disse udregninger altså, at (D1)–(D4) medfører, at

$$\det\begin{pmatrix}\boldsymbol{a}_1 & \boldsymbol{a}_2\end{pmatrix} = a_{22}a_{11} - a_{12}a_{21},$$

som ønsket. □

Eksempel 3.1.3 Vi illustrerer, hvordan (D1)–(D4) og de afledte egenskaber (D5)–(D6) kan anvendes til at udregne determinanten.

$$\det\begin{pmatrix}5 & 3\\ 7 & 4\end{pmatrix} \overset{(D6)}{=} \det\begin{pmatrix}-1 & 3\\ -1 & 4\end{pmatrix} \overset{(D6)}{=} \det\begin{pmatrix}-1 & 0\\ -1 & 1\end{pmatrix} \overset{(D6)}{=} \det\begin{pmatrix}-1 & 0\\ 0 & 1\end{pmatrix} \overset{(D2)}{=} \det\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\cdot(-1) \overset{(D4)}{=} -1.$$

I de første tre ligeheder bruger man (D6) på følgende måde: for den første erstatter vi den første søjle med den første søjle minus 2 gange den anden søjle, for den anden erstatter vi den anden søjle med den anden søjle plus 3 gange den første søjle, og for den tredje erstatter vi den første søjle med den første søjle plus den anden søjle. Den fjerde lighed bruger (D2) på den første søjle og den sidste er givet ved (D4). Formlen "$a_{11}a_{22} - a_{12}a_{21}$" fra sætning 3.1.1 giver naturligvis det samme resultat.

3.2 Determinant for $n \times n$-matricer

Vi definerer nu determinanten af en $n \times n$-matrix med indgange i et legeme $\mathbb{F}$ ved at generalisere den definition for 2×2-matricer, som vi gav i afsnit 3.1. For $\mathbb{F} = \mathbb{R}$ definerer vi endvidere det n-dimensionale volumen af det n-dimensionale parallelepipedum

$$P = \{\boldsymbol{a}_1 c_1 + \cdots + \boldsymbol{a}_n c_n \mid 0 \leq c_1, \ldots, c_n \leq 1\} \subset \mathbb{R}^n$$

udspændt af $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \in \mathbb{R}^n$ til at være absolutværdien

$$\operatorname{vol}(P) = |\det\begin{pmatrix}\boldsymbol{a}_1 & \boldsymbol{a}_2 & \ldots & \boldsymbol{a}_n\end{pmatrix}|$$

af determinanten af den reelle $n \times n$-matrix, der har $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ som søjler.

Sætning 3.2.1 *Lad $\mathbb{F}$ være et legeme og lad n være et naturligt tal. Der findes da en entydigt bestemt afbildning* $\det\colon M_n(\mathbb{F}) \to \mathbb{F}$ *med følgende egenskaber:*

(D1) *For alle $1 \leq k \leq n$ og $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{k-1}, \boldsymbol{b}_k, \boldsymbol{c}_k, \boldsymbol{a}_{k+1}, \ldots, \boldsymbol{a}_n \in \mathbb{F}^n$ gælder*

$$\det\begin{pmatrix}\boldsymbol{a}_1 & \ldots & \boldsymbol{b}_k + \boldsymbol{c}_k & \ldots & \boldsymbol{a}_n\end{pmatrix} = \det\begin{pmatrix}\boldsymbol{a}_1 & \ldots & \boldsymbol{b}_k & \ldots & \boldsymbol{a}_n\end{pmatrix} + \det\begin{pmatrix}\boldsymbol{a}_1 & \ldots & \boldsymbol{c}_k & \ldots & \boldsymbol{a}_n\end{pmatrix}.$$

(D2) *For alle $1 \leq k \leq n$, $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \in \mathbb{F}^m$ og $c \in \mathbb{F}$ gælder*

$$\det\begin{pmatrix}\boldsymbol{a}_1 & \ldots & \boldsymbol{a}_k \cdot c & \ldots & \boldsymbol{a}_n\end{pmatrix} = \det\begin{pmatrix}\boldsymbol{a}_1 & \ldots & \boldsymbol{a}_k & \ldots & \boldsymbol{a}_n\end{pmatrix} \cdot c.$$

(D3) *For alle $1 \leq k < l \leq n$ og $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \in \mathbb{F}^m$ med $\boldsymbol{a}_k = \boldsymbol{a}_l$, gælder*

$$\det\begin{pmatrix}\boldsymbol{a}_1 & \ldots & \boldsymbol{a}_k & \ldots & \boldsymbol{a}_l & \ldots & \boldsymbol{a}_n\end{pmatrix} = 0.$$

(D4) *For standardenhedsvektorerne $\boldsymbol{e}_1, \boldsymbol{e}_2 \ldots, \boldsymbol{e}_n \in \mathbb{F}^n$ gælder*

$$\det\begin{pmatrix}\boldsymbol{e}_1 & \boldsymbol{e}_2 & \ldots & \boldsymbol{e}_n\end{pmatrix} = 1.$$

Bemærkning 3.2.2 Vi kan betragte determinanten som en afbildning

$$\det\colon \mathbb{F}^n \times \cdots \times \mathbb{F}^n \to \mathbb{F},$$

der til en n-tuple $(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n)$ af søjlevektorer tilordner skalaren $\det\begin{pmatrix}\boldsymbol{a}_1 & \ldots & \boldsymbol{a}_n\end{pmatrix}$, og da udtrykker (D1)–(D2), at determinanten opfylder betingelserne (L1)–(L2) for en lineær afbildning i hver faktor $\mathbb{F}^n$. Vi siger, at en sådan afbildning er *multi-lineær*. Endvidere siges en multi-lineær afbildning, der også opfylder (D3), at være *alternerende*.

Således er determinanten den entydige alternerende multi-lineære afbildning, der på standardbasen $(\boldsymbol{e}_1,\dots,\boldsymbol{e}_n)$ tager værdien 1. Denne sidste egenskab (D4) er naturligvis temmelig arbitrær, og beviset for sætning 3.2.1 viser da også mere generelt, at der for ethvert $a \in \mathbb{F}$ findes en entydig multi-lineær alternerende afbildning, der på standardbasen $(\boldsymbol{e}_1,\dots,\boldsymbol{e}_n)$ antager værdien $a \in \mathbb{F}$.

I beviset for sætning 3.2.1 anvender vi undervejs, at de definerede egenskaber ved determinanten har følgende konsekvenser.

Sætning 3.2.3 *Lad $\mathbb{F}$ være et legeme, og lad n være et naturligt tal. Hvis en afbildning* $\det\colon M_n(\mathbb{F}) \to \mathbb{F}$ *tilfredsstiller* (D1)–(D3)*, da gælder endvidere følgende:*

(D5) *Hvis B er matricen, der fremkommer fra A ved ombytning af to søjler, da er*

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

(D6) *Hvis B er matricen, der fremkommer fra A ved at addere et multiplum af en søjle i A til en anden søjle i A, da er*

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

Bevis Vi viser først (D5). Så lad A være en $n \times n$-matrix og lad B være matricen, der fremkommer fra A ved ombytning af den k'te og l'te søjle, hvor $1 \le k < l \le n$.

$$\begin{aligned}
\det(A) + \det(B) &= \det\begin{pmatrix}\dots & \boldsymbol{a}_k & \dots & \boldsymbol{a}_l & \dots\end{pmatrix} + \det\begin{pmatrix}\dots & \boldsymbol{a}_l & \dots & \boldsymbol{a}_k & \dots\end{pmatrix} \\
&\overset{(D3)}{=} \det\begin{pmatrix}\dots & \boldsymbol{a}_k & \dots & \boldsymbol{a}_l & \dots\end{pmatrix} + \det\begin{pmatrix}\dots & \boldsymbol{a}_l & \dots & \boldsymbol{a}_k & \dots\end{pmatrix} \\
&\qquad + \det\begin{pmatrix}\dots & \boldsymbol{a}_k & \dots & \boldsymbol{a}_k & \dots\end{pmatrix} + \det\begin{pmatrix}\dots & \boldsymbol{a}_l & \dots & \boldsymbol{a}_l & \dots\end{pmatrix} \\
&\overset{(D1)}{=} \det\begin{pmatrix}\dots & \boldsymbol{a}_k & \dots & \boldsymbol{a}_k + \boldsymbol{a}_l & \dots\end{pmatrix} + \det\begin{pmatrix}\dots & \boldsymbol{a}_l & \dots & \boldsymbol{a}_k + \boldsymbol{a}_l & \dots\end{pmatrix} \\
&\overset{(D1)}{=} \det\begin{pmatrix}\dots & \boldsymbol{a}_k + \boldsymbol{a}_l & \dots & \boldsymbol{a}_k + \boldsymbol{a}_l & \dots\end{pmatrix} \overset{(D3)}{=} 0,
\end{aligned}$$

hvilket beviser (D5). Tilsvarende, lad B være matricen der fremkommer fra A ved at addere multiplummet $\boldsymbol{a}_l \cdot c$ af den l'te søjle $\boldsymbol{a}_l$ til den k'te søjle $\boldsymbol{a}_k$. Hvis $k < l$, da er

$$\begin{aligned}
\det(B) &= \det\begin{pmatrix}\dots & \boldsymbol{a}_k + \boldsymbol{a}_l \cdot c & \dots & \boldsymbol{a}_l & \dots\end{pmatrix} \\
&\overset{(D1)}{=} \det\begin{pmatrix}\dots & \boldsymbol{a}_k & \dots & \boldsymbol{a}_l & \dots\end{pmatrix} + \det\begin{pmatrix}\dots & \boldsymbol{a}_l \cdot c & \dots & \boldsymbol{a}_l & \dots\end{pmatrix} \\
&\overset{(D2)}{=} \det\begin{pmatrix}\dots & \boldsymbol{a}_k & \dots & \boldsymbol{a}_l & \dots\end{pmatrix} + \det\begin{pmatrix}\dots & \boldsymbol{a}_l & \dots & \boldsymbol{a}_l & \dots\end{pmatrix} \cdot c \\
&\overset{(D3)}{=} \det\begin{pmatrix}\dots & \boldsymbol{a}_k & \dots & \boldsymbol{a}_l & \dots\end{pmatrix} = \det(A),
\end{aligned}$$

hvilket beviser (D6) for $k < l$, og beviset for $k > l$ er ganske tilsvarende. □

For at vise at determinantafbildningen findes, giver vi i beviset for sætning 3.2.1 en rekursiv definition af determinanten, som bruger følgende definition.

Definition 3.2.4 Lad A være en $m \times n$-matrix. Givet $1 \leq i \leq m$ og $1 \leq j \leq n$, er den (i,j)'te *undermatrix* af A den $(m-1)\times(n-1)$-matrix A_{ij}, der fremkommer fra A ved at fjerne den i'te række og den j'te søjle:

$$A_{ij} = \begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & & \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{pmatrix}$$

Hvis A er en kvadratisk matrix af orden n, da er den (i,j)'te undermatrix A_{ij} specielt en kvadratisk matrix af orden $n-1$.

Eksempel 3.2.5 Vi betragter som eksempel 3×3-matricen

$$A = \begin{pmatrix} 5 & -1 & 1 \\ 0 & 2 & -3 \\ 2 & 6 & 4 \end{pmatrix}.$$

Den $(1,1)$'te undermatrix A_{11} og den $(2,3)$'te undermatrix A_{23} er da henholdsvis

$$A_{11} = \begin{pmatrix} 5 & -1 & 1 \\ 0 & 2 & -3 \\ 2 & 6 & 4 \end{pmatrix} = \begin{pmatrix} 2 & -3 \\ 6 & 4 \end{pmatrix} \qquad \text{og} \qquad A_{23} = \begin{pmatrix} 5 & -1 & 1 \\ 0 & 2 & -3 \\ 2 & 6 & 4 \end{pmatrix} = \begin{pmatrix} 5 & -1 \\ 2 & 6 \end{pmatrix},$$

som begge er 2×2-matricer.

Bevis (for sætning 3.2.1) Vi ønsker at vise eksistens og entydighed af en afbildning $\det\colon M_n(\mathbb{F}) \to \mathbb{F}$, der opfylder (D1)–(D4), og vi begynder med eksistensen. Vi definerer denne afbildning rekursivt for $n \geq 0$ som følger: For $n = 0$ har mængden $M_n(\mathbb{F})$ præcis ét element, nemlig den tomme matrix $(\ )$, og vi definerer da $\det\colon M_0(\mathbb{F}) \to \mathbb{F}$ til at være afbildningen givet ved $\det(\ ) = 1$. For $n \geq 1$ antager vi, at afbildningen er defineret for $n = p-1$ og definerer den for $n = p$ ved formlen

$$\det(A) = \sum_{j=1}^{p} (-1)^{1+j} a_{1j} \det(A_{1j}) \tag{3.2.6}$$

hvor A_{1j} er den $(1,j)$'te undermatrix af A. Vi skal vise at afbildningen defineret på denne måde opfylder (D1)–(D4), hvilket vi gør ved induktion på $n \geq 0$. Hvis $n = 0$, da er (D1)–(D3) trivielt opfyldte, da der ikke findes $1 \leq j \leq n$ eller $1 \leq j < k \leq n$, mens (D4) er opfyldt per definition. Så vi antager induktivt, at afbildningen $\det\colon M_n(\mathbb{F}) \to \mathbb{R}$ opfylder (D1)–(D4) for $n = p-1$, og viser, at den også opfylder (D1)–(D4) for $n = p$. For at vise (D1), lader vi $A, B, C \in M_p(\mathbb{F})$ være kvadratiske matricer af orden p, sådan at

$$a_{ij} = b_{ij} = c_{ij}$$

for alle $1 \le i, j \le p$ med $j \neq k$, mens

$$a_{ik} = b_{ik} + c_{ik},$$

for alle $1 \le i \le p$. Så for alle $1 \le j \le p$ med $j \neq k$ er den j'te søjle i A lig med den j'te søjle i både B og C, mens den k'te søjle i A er lig med summen af den k'te søjle i B og den k'te søjle i C, og vi ønsker at vise, at $\det(A) = \det(B) + \det(C)$. Da (D1) per induktion gælder for kvadratiske matricer af orden $p-1$, gælder der, at

$$\det(A_{1j}) = \det(B_{1j}) + \det(C_{1j})$$

for alle $1 \le j \le p$ med $j \neq k$, mens

$$\det(A_{1k}) = \det(B_{1k}) = \det(C_{1k}),$$

fordi $A_{1k} = B_{1k} = C_{1k}$. Derfor får vi nu ved at indsætte i (3.2.6), at

$$\begin{aligned}
\det(A) &= \sum_{j=1}^{p} (-1)^{1+j} a_{1j} \det(A_{1j}) \\
&= \sum_{j=1}^{k-1} (-1)^{1+j} a_{1j} (\det(B_{1j}) + \det(C_{1j})) + (-1)^{1+k} (b_{1k} + c_{1k}) \det(A_{1k}) \\
&\qquad + \sum_{j=k+1}^{p} (-1)^{1+j} a_{1j} (\det(B_{1j}) + \det(C_{1j})) \\
&= \sum_{j=1}^{p} (-1)^{1+j} b_{1j} \det(B_{1j}) + \sum_{j=1}^{p} (-1)^{1+j} c_{1j} \det(C_{1j}) = \det(B) + \det(C)
\end{aligned}$$

som ønsket. Dette viser, at (D1) gælder for $n = p$, og beviset for (D2) er tilsvarende.

For at vise (D3), betragter vi først tilfældet, hvor $1 \le k < p$ og den k'te og $(k+1)$'te søjle i A er ens. Hvis $j \neq k$ og $j \neq k+1$, da har undermatricen A_{1j} igen to ens søjler, sådan at $\det(A_{1j}) = 0$ per induktion. Desuden er $A_{1k} = A_{1,k+1}$, sådan at

$$\begin{aligned}
\det(A) &= \sum_{j=1}^{p} (-1)^{1+j} a_{1j} \det(A_{1j}) \\
&= (-1)^{1+k} a_{1k} \det(A_{1k}) + (-1)^{1+k+1} a_{1,k+1} \det(A_{1k+1}) = 0,
\end{aligned}$$

hvilket viser, at (D3) gælder i tilfældet, hvor to nabosøjler er ens. Dette medfører ved et argument svarende til beviset for (D5) i sætning 3.2.3, at determinanten af en kvadratisk matrix af orden p skifter fortegn, hvis to nabosøjler ombyttes. Det bruger vi for at vise, at (D3) gælder i det generelle tilfælde. Så antag nu at den j'te og k'te søjle i A er ens, hvor $1 \leq j < k \leq p$. Ved at lave $k - j - 1$ ombytninger af nabosøjler kan vi omdanne A til en matrix B vis j'te og $j+1$'st søjler er ens, og ved at bruge (D5) gentagende gange for nabosøjler, ser vi, at $\det(A) = (-1)^{k-j-1}\det(B)$, og vi har lige vist at $\det(B) = 0$ fordi den har to ens nabosøjler. Så $\det(A) = 0$ og (D3) følger. Endelig følger det fra (3.2.6), at

$$\det\begin{pmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \dots & \boldsymbol{e}_m \end{pmatrix} = 1 \cdot \det\begin{pmatrix} \boldsymbol{e}_1 & \dots & \boldsymbol{e}_{m-1} \end{pmatrix} + 0 + \dots + 0 = \det\begin{pmatrix} \boldsymbol{e}_1 & \dots & \boldsymbol{e}_{m-1} \end{pmatrix},$$

og per induktion er den fælles værdi lig med 1, hvilket viser at også (D4) gælder. Vi har nu vist induktionsskridtet, og vi har således defineret en afbildning

$$\det : M_n(\mathbb{F}) \to \mathbb{F},$$

der opfylder (D1)–(D4).

Vi mangler at vise, at en sådan afbildning er entydigt bestemt. Så vi lader

$$D : M_n(\mathbb{F}) \to \mathbb{F}$$

være en afbildning, der opfylder (D1)–(D4), og skal da vise, at $D = \det$, hvilket vi gør ved induktion på $n \geq 0$. Hvis $n = 0$, da er $D\begin{pmatrix} \end{pmatrix} = 1$ ifølge (D4), hvilket viser, at $D = \det$ i dette tilfælde. Så vi antager, at vi har vist, at $D = \det$ for $n = p - 1$, og viser, at $D = \det$ for $n = p$. Hvis $A = \begin{pmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \dots & \boldsymbol{a}_p \end{pmatrix} \in M_p(\mathbb{F})$, så kan vi anvende lemma 2.2.5 til at skrive den første søjle $\boldsymbol{a}_1$ som en linear kombination

$$\boldsymbol{a}_1 = \boldsymbol{e}_1 a_{11} + \boldsymbol{e}_2 a_{21} + \dots + \boldsymbol{e}_p a_{p1}$$

af standardbasen $(\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_p)$ for $\mathbb{F}^p$ og konkluderer da, at

$$\begin{aligned} D(A) &= D\begin{pmatrix} \boldsymbol{e}_1 a_{11} & \boldsymbol{a}_2 & \dots & \boldsymbol{a}_p \end{pmatrix} + D\begin{pmatrix} \boldsymbol{e}_2 a_{21} & \boldsymbol{a}_2 & \dots & \boldsymbol{a}_p \end{pmatrix} + \dots + D\begin{pmatrix} \boldsymbol{e}_p a_{p1} & \boldsymbol{a}_2 & \dots & \boldsymbol{a}_p \end{pmatrix} \\ &= D\begin{pmatrix} \boldsymbol{e}_1 & \boldsymbol{a}_2 & \dots & \boldsymbol{a}_p \end{pmatrix} a_{11} + D\begin{pmatrix} \boldsymbol{e}_2 & \boldsymbol{a}_2 & \dots & \boldsymbol{a}_p \end{pmatrix} a_{21} + \dots + D\begin{pmatrix} \boldsymbol{e}_p & \boldsymbol{a}_2 & \dots & \boldsymbol{a}_p \end{pmatrix} a_{p1}, \end{aligned}$$

hvor vi anvender henholdsvis (D1) og (D2) for afbildningen D. Vi påstår, at

$$D\begin{pmatrix} \boldsymbol{e}_i & \boldsymbol{a}_2 & \dots & \boldsymbol{a}_p \end{pmatrix} = (-1)^{i-1}\det(A_{i1})$$

for alle $1 \leq i \leq p$, og da det også opfylder (D1)–(D4), så viser denne påstand, at

$$D(A) = \sum_{i=1}^{p} (-1)^{i-1}\det(A_{i1}) a_{i1} = \det(A)$$

som ønsket. For at bevise påstanden, lader vi $1 \leq i \leq p$ være fastholdt og anvender (D6), som er en konsekvens af (D1)–(D3), til at skrive

$$D\begin{pmatrix} \boldsymbol{e}_i & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_p \end{pmatrix} = D\begin{pmatrix} \boldsymbol{e}_i & \boldsymbol{a}_2 - \boldsymbol{e}_i a_{i2} & \cdots & \boldsymbol{a}_p - \boldsymbol{e}_i a_{ip} \end{pmatrix}.$$

Vi bemærker, at matricen til højre kun afhænger af den kvadratiske matrix A_{i1} af orden $p-1$. Mere præcist, hvis $B \in M_{p-1}(\mathbb{F})$ er en vilkårlig kvadratisk matrix af orden $p-1$, da definerer vi $\widetilde{B} \in M_p(\mathbb{F})$ til at være matricen

$$\widetilde{B} = \begin{pmatrix} 0 & b_{1,1} & \dots & b_{1,p-1} \\ \vdots & \vdots & & \vdots \\ 0 & b_{i-1,1} & \dots & b_{i-1,p-1} \\ 1 & 0 & \dots & 0 \\ 0 & b_{i,1} & \dots & b_{i,p-1} \\ \vdots & \vdots & & \vdots \\ 0 & b_{p-1,1} & \dots & b_{p-1,p-1} \end{pmatrix},$$

der er kvadratisk af orden p, og vi har da, at

$$\begin{pmatrix} \boldsymbol{e}_i & \boldsymbol{a}_2 - \boldsymbol{e}_i a_{i2} & \cdots & \boldsymbol{a}_p - \boldsymbol{e}_i a_{ip} \end{pmatrix} = \begin{pmatrix} 0 & a_{1,2} & \dots & a_{1,p} \\ \vdots & \vdots & & \vdots \\ 0 & a_{i-1,2} & \dots & a_{i-1,p} \\ 1 & 0 & \dots & 0 \\ 0 & a_{i+1,2} & \dots & a_{i+1,p} \\ \vdots & \vdots & & \vdots \\ 0 & a_{p,2} & \dots & a_{p,p} \end{pmatrix} = \widetilde{A_{i1}}.$$

Vi bemærker dernæst, at afbildningen $d \colon M_{p-1}(\mathbb{F}) \to \mathbb{F}$ defineret ved

$$d(B) = (-1)^{i-1} D(\tilde{B})$$

opfylder (D1)–(D4). For egenskaberne (D1)–(D3) følger fra de tilsvarende egenskaber for D, mens (D4) følger fra udregningen

$$d\begin{pmatrix} \boldsymbol{e}_1 & \dots & \boldsymbol{e}_{p-1} \end{pmatrix} = (-1)^{i-1} D\begin{pmatrix} \boldsymbol{e}_i & \boldsymbol{e}_1 & \dots & \boldsymbol{e}_{i-1} & \boldsymbol{e}_{i+1} & \dots & \boldsymbol{e}_p \end{pmatrix} = D\begin{pmatrix} \boldsymbol{e}_1 & \dots & \boldsymbol{e}_p \end{pmatrix} = 1,$$

hvor den første identitet er definitionen af d; den anden følger fra (D5) for D, idet vi for at ændre $\begin{pmatrix} \boldsymbol{e}_i & \boldsymbol{e}_1 & \cdots & \boldsymbol{e}_{i-1} & \boldsymbol{e}_{i+1} & \cdots & \boldsymbol{e}_p \end{pmatrix}$ til $\begin{pmatrix} \boldsymbol{e}_1 & \cdots & \boldsymbol{e}_{i-1} & \boldsymbol{e}_i & \boldsymbol{e}_{i+1} & \cdots & \boldsymbol{e}_p \end{pmatrix}$ skal bytte om på to rækker $i-1$ gange; og den tredje følger fra (D4) for D. Det følger derfor fra den induktive hypotese, at $d(B) = \det(B)$ for alle $B \in M_{p-1}(\mathbb{F})$. Specielt er

$$(-1)^{i-1} D(\widetilde{A_{i1}}) = d(\widetilde{A_{i1}}) = \det(A_{i1}),$$

hvilket viser den ønskede påstand, at

$$D\begin{pmatrix} \boldsymbol{e}_i & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_p \end{pmatrix} = D(\widetilde{A_{i1}}) = (-1)^{i-1} \det(A_{i1}).$$

Dette beviser induktionsskridtet og dermed entydigheden af determinanten. □

Det følger fra beviset for sætning 3.2.1, at der gælder følgende rekursive formler for determinanten. Disse formler kaldes for *Laplace udvikling*, da de skyldes Laplace.

Sætning 3.2.7 (Laplace) *Lad A være en $n \times n$-matrix med indgange i et legeme $\mathbb{F}$.*

(D7) *Determinanten af A kan udregnes ved udvikling langs i'te række:*

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}).$$

(D8) *Determinanten af A kan udregnes ved udvikling langs j'te søjle:*

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} \det(A_{ij}) a_{ij}.$$

Bevis Vi beviser først (D7). I eksistensdelen af beviset for sætning 3.2.1 beviste vi, at $\det(A)$ definereret ved udvikling langs første række opfylder (D1)–(D4). Valget af den første række spillede dog ingen rolle, og en mindre modifikation af beviset viser at udvikling af deteminanten langs en vilkårlig række ligeledes opfylder (D1)–(D4). Resultatet følger derfor fra entydigheden af determinanten.

Vi beviser dernæst (D8). Ifølge beviset for entydighedsdelen af sætning 3.2.1, og bemærkningen om at $(-1)^{i-1} = (-1)^{i+1}$, er

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+1} \det(A_{i1}) a_{i1},$$

hvilket viser (D8) for $j = 1$. Vi lader nu B være den matrix, der fremkommer fra A ved at ombytte den første søjle og den j'te søjle, og har da identiteten

$$\det(B) = \sum_{i=1}^{n} (-1)^{i+1} \det(B_{i1}) b_{i1},$$

som vi netop har bemærket. Det følger da fra (D5) er $\det(B) = -\det(A)$, og per definition af matricen B er $b_{i1} = a_{ij}$. Endvidere fremkommer B_{i1} fra $n \times (n-1)$-matricen

$$B_1 = \begin{pmatrix} \boldsymbol{b}_2 & \cdots & \boldsymbol{b}_{j-1} & \boldsymbol{b}_j & \boldsymbol{b}_{j+1} & \cdots & \boldsymbol{b}_n \end{pmatrix} = \begin{pmatrix} \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_{j-1} & \boldsymbol{a}_1 & \boldsymbol{a}_{j+1} & \cdots & \boldsymbol{a}_n \end{pmatrix}$$

ved at fjerne den i'te række, mens A_{ij} fremkommer fra $n \times (n-1)$-matricen

$$A_j = \begin{pmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_{j-1} & \boldsymbol{a}_{j+1} & \cdots & \boldsymbol{a}_n \end{pmatrix}$$

ved at fjerne den i'te række. Vi bemærker nu, at A_j kan omdannes til B_1 ved at foretage $j-2$ ombytninger to søjler. Det følger heraf, at A_{ij} ligeledes kan omdannes til B_{i1} ved at foretage $j-2$ ombytninger af to søjler, og derfor viser (D5), at

$$\det(B_{i1}) = (-1)^{j-2} \det(A_{ij}).$$

Vi har således alt i alt vist, at

$$\begin{aligned}\det(A) &= -\det(B) = -\sum_{i=1}^{n}(-1)^{i+1}\det(B_{i1})b_{i1}\\ &= -\sum_{i=1}^{n}(-1)^{i+1+j-2}\det(A_{ij})a_{ij} = \sum_{i=1}^{n}(-1)^{i+j}\det(A_{ij})a_{ij}\end{aligned}$$

som ønsket. Dette viser (D8) og dermed sætningen. □

Eksempel 3.2.8 Vi udregner determinanten af matricen

$$A = \begin{pmatrix} 5 & -1 & 1 \\ 0 & 2 & -3 \\ 2 & 6 & 4 \end{pmatrix}$$

fra eksempel 3.2.5. Laplace udvikling langs første søjle giver

$$\det(A) = \det(A_{11})\cdot 5 - \det(A_{21})\cdot 0 + \det(A_{31})\cdot 2,$$

hvor

$$A_{11} = \begin{pmatrix} 5 & -1 & 1 \\ 0 & 2 & -3 \\ 2 & 6 & 4 \end{pmatrix} = \begin{pmatrix} 2 & -3 \\ 6 & 4 \end{pmatrix} \quad \text{og} \quad A_{31} = \begin{pmatrix} 5 & -1 & 1 \\ 0 & 2 & -3 \\ 2 & 6 & 4 \end{pmatrix} = \begin{pmatrix} -1 & 1 \\ 2 & -3 \end{pmatrix}.$$

Vi behøver ikke at udregne A_{21} og dens determinant, da $a_{21} = 0$. Fra sætning 3.1.1 har vi nu, at $\det(A_{11}) = 2\cdot 4 - (-3)\cdot 6 = 26$ og $\det(A_{31}) = (-1)\cdot(-3) - 1\cdot 2 = 1$, så

$$\det(A) = 26\cdot 5 + 1\cdot 2 = 132.$$

Vi bemærker, at det altid er en fordel at udvikle en determinant langs en række eller søjle, der indeholder mange nuller—jo flere desto bedre. For dermed behøver man ikke udregne de tilsvarende summander i Laplace udviklingen. I dette eksempel kunne vi lige så godt have valgt at udvikle determinanten langs den anden række.

Bemærkning 3.2.9 Det er naturligvis vigtigt at huske fortegnene i udviklingen af determinant efter en række eller en søjle. De minder om et skakbræt

$$(+), \quad \begin{pmatrix} + & - \\ - & + \end{pmatrix}, \quad \begin{pmatrix} + & - & + \\ - & + & - \\ + & - & + \end{pmatrix}, \quad \begin{pmatrix} + & - & + & - \\ - & + & - & + \\ + & - & + & - \\ - & + & - & + \end{pmatrix}, \quad \dots \quad ,$$

hvor man begynder med "+" i det øverste venstre hjørne.

Sætning 3.2.10 *Lad $\mathbb{F}$ være et legeme og lad $(-)^* : \mathbb{F} \to \mathbb{F}$ være en skævinvolution. Hvis A er en $n \times n$-matrix med indgange i $\mathbb{F}$, og hvis A^* er den adjungerede matrix, da er*

$$\det(A^*) = \det(A)^*.$$

Bevis Beviset er igen ved induktion på $n \geq 0$. For $n = 0$ er sætningen triviel, så vi antager, at sætningen allerede er bevist for $n = p-1$ og beviser den for $n = p$. Lad os skrive $a'_{ij} = a^*_{ji}$ for den (i,j)'te indgang i A^*. Da endvidere $(A^*)_{ij} = (A_{ji})^*$, gælder

$$\begin{aligned}\det(A^*) &= \sum_{i=1}^{p} (-1)^{i+1} \det((A^*)_{i1}) a'_{i1} = \sum_{j=1}^{p} (-1)^{1+j} \det((A_{1j})^*) a^*_{1j} \\ &= \sum_{j=1}^{p} (-1)^{1+j} \det(A_{1j})^* a^*_{1j} = (\sum_{j=1}^{p} (-1)^{1+j} a_{1j} \det(A_{1j}))^* = \det(A)^*,\end{aligned}$$

hvor den første lighed følger fra (D8); den anden fra vores indledende bemærninger; den tredje fra den induktive hypotese, at sætningen gælder for $(p-1) \times (p-1)$-matricer; den fjerde fra definitionen af en antiinvolution; og den femte fra (D7). Vi har nu vist induktionsskridtet og dermed sætningen. □

Korollar 3.2.11 *Lad A være en $n \times n$-matrix med indgange i et legeme $\mathbb{F}$, og lad A^t være den transponerede $n \times n$-matrix, da gælder*

$$\det(A^t) = \det(A).$$

Bevis Dette er netop udsagnet i sætning 3.2.10 i det tilfælde, hvor $(-)^* = \mathrm{id} : \mathbb{F} \to \mathbb{F}$. □

Korollar 3.2.12 *Hvis A er en hermitisk kompleks matrix, da er* $\det(A)$ *et reelt tal.*

Bevis Vi lader $\mathbb{F} = \mathbb{C}$ og lader $\sigma : \mathbb{C} \to \mathbb{C}$ være kompleks konjugation. Vi minder om, at en kvadratisk matrix A med indgange i $\mathbb{C}$ per definition er hermitisk, hvis $A = A^*$. Vi konkluderer derfor fra sætning 3.2.10, at

$$\det(A) = \det(A^*) = \det(A)^*,$$

hvilket som ønsket viser, at $\det(A)$ er et reelt tal. □

Den følgende sætning giver en version af (D1)–(D3) og (D5)–(D6) for rækkevektorer i stedet for søjlevektorer.

Sætning 3.2.13 *Lad $\mathbb{F}$ være et legeme og n et naturligt tal. Der gælder følgende:*

(D1') *For alle $1 \le j \le n$ og $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{j-1}, \boldsymbol{b}_j, \boldsymbol{c}_j, \boldsymbol{a}_{j+1}, \ldots, \boldsymbol{a}_n \in M_{1,n}(\mathbb{F})$ er*

$$\det\begin{pmatrix} \boldsymbol{a}_1 \\ \vdots \\ \boldsymbol{b}_j + \boldsymbol{c}_j \\ \vdots \\ \boldsymbol{a}_n \end{pmatrix} = \det\begin{pmatrix} \boldsymbol{a}_1 \\ \vdots \\ \boldsymbol{b}_j \\ \vdots \\ \boldsymbol{a}_n \end{pmatrix} + \det\begin{pmatrix} \boldsymbol{a}_1 \\ \vdots \\ \boldsymbol{c}_j \\ \vdots \\ \boldsymbol{a}_n \end{pmatrix}.$$

(D2') *For alle $1 \le j \le n$, $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \in M_{1,n}(\mathbb{F})$ og $c \in \mathbb{F}$ gælder*

$$\det\begin{pmatrix} \boldsymbol{a}_1 \\ \vdots \\ c \cdot \boldsymbol{a}_j \\ \vdots \\ \boldsymbol{a}_n \end{pmatrix} = c \cdot \det\begin{pmatrix} \boldsymbol{a}_1 \\ \vdots \\ \boldsymbol{a}_j \\ \vdots \\ \boldsymbol{a}_n \end{pmatrix}.$$

(D3') *Hvis $A \in M_{n,n}(\mathbb{F})$ har to ens rækker, da er* $\det(A) = 0$.

(D5') *Hvis B fremkommer ved ombytning af to rækker i A, da er* $\det(B) = -\det(A)$.

(D6') *Hvis B fremkommer fra A ved at addere et multiplum af en række i A til en anden række i A, da er* $\det(B) = \det(A)$.

Bevis Hvis $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \in M_{1,n}(\mathbb{F})$ er rækkevektorer, da er $\boldsymbol{a}_1^t, \ldots, \boldsymbol{a}_n^t \in M_{n,1}(\mathbb{F})$ søjlevektorer, og ifølge korollar 3.2.12 har $n \times n$-matricen

$$A = \begin{pmatrix} \boldsymbol{a}_1 \\ \vdots \\ \boldsymbol{a}_n \end{pmatrix} \in M_n(\mathbb{F})$$

den samme determinant som dens transponerede $n \times n$-matrix

$$A^t = \begin{pmatrix} \boldsymbol{a}_1 \\ \vdots \\ \boldsymbol{a}_n \end{pmatrix}^t = \begin{pmatrix} \boldsymbol{a}_1^t & \cdots & \boldsymbol{a}_n^t \end{pmatrix} \in M_n(\mathbb{F}).$$

De ønskede identiteter (D1')–(D3') og (D5')–(D6') angående rækkevektorer følger nu fra de tilsvarende identiteter (D1)–(D3) og (D5)–(D6) angående søjlevektorer. □

Vi gennemgår i næste afsnit, hvordan rækkeoperationer og søjleoperationer anvendes til at udregne determinanter. Vi viser her den følgende vigtige egenskab.

Sætning 3.2.14 *Hvis A og B er $n \times n$-matricer med indgange i et legeme $\mathbb{F}$, da gælder*

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

Bevis Givet $a \in \mathbb{F}$, betragter vi afbildningen $\det_a : M_n(\mathbb{F}) \to \mathbb{F}$ defineret ved

$$\det_a(B) = a \cdot \det(B).$$

Ved en mindre modifikation af beviset for entydighedsdelen af sætning 3.2.1 viser vi, at denne afbildning er entydigt bestemt ved, at den opfylder (D1)–(D3) i sætning 3.2.1 samt den nye betingelse $(\mathrm{D4})_a$, at $\det_a(I_n) = a$. Det er derfor tilstrækkeligt at vise, at afbildningen $d : M_n(\mathbb{F}) \to \mathbb{F}$ defineret ved

$$d(B) = \det(AB)$$

opfylder (D1)–(D3) samt $(\mathrm{D4})_a$, hvor $a = \det(A)$. For da giver entydighedsudsagnet, at $d(B) = \det_a(B)$, hvilket præcis giver $\det(AB) = \det(A)\det(B)$ som ønsket. Vi viser nu, at afbildningen $d : M_n(\mathbb{F}) \to \mathbb{F}$ opfylder (D1). Hvis $B = \begin{pmatrix} \boldsymbol{b}_1 & \cdots & \boldsymbol{b}_k & \cdots & \boldsymbol{b}_n \end{pmatrix}$, da er

$$AB = \begin{pmatrix} \boxed{a_{11} \quad a_{12} \quad \cdots \quad a_{1n}} \\ \vdots \quad\quad \vdots \quad\quad\quad\quad \vdots \\ \boxed{a_{n1} \quad a_{n2} \quad \cdots \quad a_{nn}} \end{pmatrix} \begin{pmatrix} \begin{array}{|c|c|c|c|c|} b_{11} & \cdots & b_{1k} & \cdots & b_{1n} \\ b_{21} & \cdots & b_{2k} & \cdots & b_{2n} \\ \vdots & & \vdots & & \vdots \\ b_{n1} & \cdots & b_{nk} & \cdots & b_{nn} \end{array} \end{pmatrix} = \begin{pmatrix} A\boldsymbol{b}_1 & \cdots & A\boldsymbol{b}_k & \cdots & A\boldsymbol{b}_n \end{pmatrix}$$

ifølge definitionen af matrix produktet. Hvis derfor $\boldsymbol{b}_k = \boldsymbol{c}_k + \boldsymbol{d}_k$, da er

$$\begin{aligned} d\begin{pmatrix} \boldsymbol{b}_1 & \cdots & \boldsymbol{b}_k & \cdots & \boldsymbol{b}_n \end{pmatrix} &= \det\begin{pmatrix} A\boldsymbol{b}_1 & \cdots & A(\boldsymbol{c}_k + \boldsymbol{d}_k) & \cdots & A\boldsymbol{b}_n \end{pmatrix} \\ &= \det\begin{pmatrix} A\boldsymbol{b}_1 & \cdots & A\boldsymbol{c}_k + A\boldsymbol{d}_k & \cdots & A\boldsymbol{b}_n \end{pmatrix} \\ &= \det\begin{pmatrix} A\boldsymbol{b}_1 & \cdots & A\boldsymbol{c}_k & \cdots & A\boldsymbol{b}_n \end{pmatrix} + \det\begin{pmatrix} A\boldsymbol{b}_1 & \cdots & A\boldsymbol{d}_k & \cdots & A\boldsymbol{b}_n \end{pmatrix} \\ &= d\begin{pmatrix} \boldsymbol{b}_1 & \cdots & \boldsymbol{c}_k & \cdots & \boldsymbol{b}_n \end{pmatrix} + d\begin{pmatrix} \boldsymbol{b}_1 & \cdots & \boldsymbol{d}_k & \cdots & \boldsymbol{b}_n \end{pmatrix}, \end{aligned}$$

hvilket viser, at $d : M_n(\mathbb{F}) \to \mathbb{F}$ opfylder (D1). Her følger den første lighed og fjerde lighed fra definitionen af matrix produktet, den anden lighed følger fra den distributive lov for matrix multiplikation, mens den tredje lighed følger fra (D1) for determinanten. Beviset for (D2) og (D3) er tilsvarende, mens $(\mathrm{D4})_a$ fås fra $d(I) = \det(AI) = \det(A) = a$. □

Korollar 3.2.15 *Hvis en kvadratisk matrix A er invertibel, så er* $\det(A)$ *invertibel, og*

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

Specielt er $\det(A) \neq 0$.

Bevis Vi har at $AA^{-1} = I = A^{-1}A$. Fra Sætning 3.2.14 og (D4) har vi derfor

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

som viser, at $\det(A)$ er invertibel med invers $\det(A^{-1})$. □

Bemærkning 3.2.16 Hvis A og B er $n\times n$-matricer, da medfører den kommutative lov for multiplikation af skalarer samt sætning 3.2.14, at der altid gælder

$$\det(AB) = \det(A)\det(B) = \det(B)\det(A) = \det(BA),$$

også selvom matricerne AB og BA sædvanligvis ikke er ens. Vi bemærker også, at

$$\det(A+B) \neq \det(A) + \det(B),$$

undtaget i trivielle tilfælde.

Eksempel 3.2.17 Vi betragter 2×2-matricerne

$$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{og} \quad B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{med} \quad AB = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} \quad \text{og} \quad BA = \begin{pmatrix} 1 & -1 \\ 0 & -1 \end{pmatrix}.$$

Selvom $AB \neq BA$, så er $\det(AB) = -1 = \det(BA)$.

Vi giver endelig en lukket formel for determinanten af en $n\times n$-matrix. En bijektiv afbildning $\sigma\colon \{1,2,\dots,n\} \to \{1,2,\dots,n\}$ kaldes for en *permutation af n bogstaver*. Der er

$$n! = n\cdot(n-1)\cdot(n-2)\cdot\ \dots\ \cdot 1$$

forskellige permutationer af n bogstaver. En permutation af n bogstaver kaldes for en *transposition*, hvis der findes $1 \le k < l \le n$, sådan at

$$\sigma(i) = \begin{cases} l & \text{hvis } i = k, \\ k & \text{hvis } i = l, \\ i & \text{ellers,} \end{cases}$$

og man kan vise, at enhver permutation kan udtrykkes som en sammensætning af et antal transpositioner. En sådan opskrivning er dog ikke entydig. Givet en permutation af n bogstaver $\sigma\colon \{1,2,\dots,n\} \to \{1,2,\dots,n\}$, tilordner vi $n\times n$-matricen

$$P(\sigma) = \begin{pmatrix} \boldsymbol{e}_{\sigma(1)} & \boldsymbol{e}_{\sigma(2)} & \cdots & \boldsymbol{e}_{\sigma(n)} \end{pmatrix} \in M_n(\mathbb{F}),$$

som vi kalder for den tilhørende *permutationsmatrix*. Hvis σ og τ er to permutation af n bogstaver, så er deres sammensætning $\sigma\circ\tau$ igen en permutation af n bogstaver, og

$$P(\sigma\circ\tau) = P(\sigma)\cdot P(\tau).$$

Vi definerer *fortegnet* af en permutationen σ af n bogstaver til at være determinanten

$$\operatorname{sign}(\sigma) = \det(P(\sigma)) = \det\begin{pmatrix} \boldsymbol{e}_{\sigma(1)} & \boldsymbol{e}_{\sigma(2)} & \cdots & \boldsymbol{e}_{\sigma(n)} \end{pmatrix} \in \mathbb{F}.$$

Det følger fra sætning 3.2.14, at hvis σ og τ er to permutationer af n bogstaver, da er

$$\operatorname{sign}(\sigma \circ \tau) = \operatorname{sign}(\sigma)\operatorname{sign}(\tau),$$

og det følger endvidere fra (D5) at fortegnet af en transposition er -1. Generelt gælder der altså, at hvis σ er en permutation af n bogstaver, og hvis σ er en sammensætning af r tranpositioner, da er $\operatorname{sign}(\sigma) = (-1)^r \in \mathbb{F}$.

Sætning 3.2.18 (Leibniz) *Lad $A = (a_{ij})$ være en kvadratisk matrix af orden n med indgange i et legeme $\mathbb{F}$. Da er*

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

hvor summen løber over de $n!$ mulige permutationer af n bogstaver.

Bevis For alle $1 \leq j \leq n$ skriver vi den j'te søjle $\boldsymbol{a}_j$ i A som en linear kombination

$$\boldsymbol{a}_j = \sum_{i=1}^{n} \boldsymbol{e}_i a_{ij}$$

af standardbasen $(\boldsymbol{e}_1, \dots, \boldsymbol{e}_n)$ for $\mathbb{F}^n$. Vi får da ved gentagen brug af (D1) og (D2), at

$$\begin{aligned} \det(A) &= \det\begin{pmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_n \end{pmatrix} \\ &= \sum_{\sigma} \det\begin{pmatrix} \boldsymbol{e}_{\sigma(1)} & \boldsymbol{e}_{\sigma(2)} & \cdots & \boldsymbol{e}_{\sigma(n)} \end{pmatrix} a_{\sigma(1),1} a_{\sigma(2),2} \cdots a_{\sigma(n),n}, \end{aligned}$$

hvor summen løber over de n^n mulige afbildninger $\sigma \colon \{1, 2, \dots, n\} \to \{1, 2, \dots, n\}$. Hvis en sådan afbildning σ ikke er injektiv, da har matricen $\begin{pmatrix} \boldsymbol{e}_{\sigma(1)} & \boldsymbol{e}_{\sigma(2)} & \cdots & \boldsymbol{e}_{\sigma(n)} \end{pmatrix}$ to eller flere søjler, der er ens, og derfor viser (D3), at dens determinant lig med nul. Det er således kun summander, der er indicerede af injektive, eller ækvivalent, bijektive afbildninger $\sigma \colon \{1, 2, \dots, n\} \to \{1, 2, \dots, n\}$, som bidrager til summen. Dette viser sætningen. $\square$

Eksempel 3.2.19 For en 2×2-matrix specialiserer sætning 3.2.18 til den formel

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

som vi gav i sætning 3.1.1; og for en 3×3-matrix, får vi formlen

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

med $3! = 6$ summander. Vi advarer dog mod at bruge denne formel, da man meget let kommer til at lave fejl.

3.3 Triangulære matricer

Der er to hovedmetoder til udregning af determinanter: (1) at bruge Laplace udvikling langs en række eller en søjle samt (2) at anvende rækkeoperationer og søjleoperationer til at omdanne matricen til en triangulær matrix, som vi nu definerer. Man skal dog her være forsigtig med at huske på, at en rækkeoperation af type $M_i(c)$ og en søjleoperation af type $M_j(c)$ begge ændrer determinantens værdi med faktoren c.

Definition 3.3.1 En $n \times n$-matrix $B = (b_{ij})$ kaldes for *nedre triangulær*, hvis $b_{ij} = 0$ for alle $1 \leq i < j \leq m$; *øvre triangulær*, hvis $b_{ij} = 0$ for alle $1 \leq j < i \leq m$; og *triangulær*, hvis den er enten nedre triangulær eller øvre triangulær.

Specielt er en diagonal matrix en triangulær matrix. Generelt er diagonalindgangene i en triangulær matrix af særlig betydning. For eksempel er en øvre triangulær matrix på echelon form, hvis og kun hvis dens diagonalindgange alle er forskellige fra 0.

Eksempel 3.3.2 Blandt 3×3-matricerne

$$A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 8 & 0 \\ 8 & 7 & 2 \end{pmatrix}, \qquad B = \begin{pmatrix} 2 & 6 & 3 \\ 0 & 4 & -6 \\ 0 & 0 & 12 \end{pmatrix} \qquad \text{og} \qquad C = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

er A nedre triangulær og B øvre triangulær, mens C er både øvre og nedre trianguær og derfor en diagonal matrix. Vi har markeret diagonalindgangene med blåt.

Sætning 3.3.3 *Determinanten af en triangulær $n \times n$-matrix B er lig med produktet*

$$\det(B) = b_{11} b_{22} \cdots b_{nn}$$

af diagonalindgangene.

Bevis Vi beviser udsagnet for B nedre triangulær; beviset for B øvre triangulær er ganske tilsvarende. Beviset er ved induktion på $n \geq 0$, og tilfældet $n = 0$ gælder trivielt. Vi antager derfor, at udsagnet er vist for $n = p - 1$ og viser det for $n = p$. Så vi lader

$$B = \begin{pmatrix} b_{11} & 0 & 0 & \cdots & 0 \\ b_{21} & b_{22} & 0 & \cdots & 0 \\ b_{31} & b_{32} & b_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b_{p1} & b_{p2} & b_{p3} & \cdots & b_{pp} \end{pmatrix}$$

og får ved Laplace udvikling langs første række, at

$$\det(B) = b_{11} \cdot \det(B_{11}) + 0 \cdot \det(B_{12}) + \cdots + 0 \cdot \det(B_{1p}) = b_{11} \cdot \det(B_{11}).$$

Endvidere er undermatricen B_{11} en nedre triangulær $(p-1) \times (p-1)$-matrix, og dens diagonalindgange er $b_{22}, \ldots, b_{pp}$, så ifølge den induktive antagelse er

$$\det(B_{11}) = b_{22} \cdots b_{pp}.$$

Dette viser induktionsskridtet og dermed sætningen. □

Det er altså meget let at bestemme determinanten af en triangulær matrix. Generelt for en kvadratisk matrix A består den mest effektive strategi til bestemmelse af $\det(A)$ i at anvende række- og søjleoperationer til at omdanne A til en triangulær matrix B, hvorved $\det(A)$ udtrykkes ved $\det(B)$ ved hjælp af sætning 3.2.1, 3.2.3 og 3.2.13. Da $\det(B)$ er givet ved sætning 3.3.3, bestemmer dette $\det(A)$. Vi vil nu illustrerer denne fremgang ved følgende eksempel.

Eksempel 3.3.4 Vi betragter 4×4-matricen

$$A = \begin{pmatrix} 1 & 1 & 3 & 5 \\ 0 & 0 & 2 & 3 \\ 1 & 2 & 1 & 3 \\ 3 & 1 & 5 & 9 \end{pmatrix},$$

og anvender rækkeoperationer til at omdanne den til en triangulær matrix B, idet vi er omhyggelige med at notere, hvorledes disse operationer påvirker determinanten.

$$\det(A) \overset{(D5')}{=} -\det\begin{pmatrix} 1 & 1 & 3 & 5 \\ 1 & 2 & 1 & 3 \\ 0 & 0 & 2 & 3 \\ 3 & 1 & 5 & 9 \end{pmatrix} \overset{(D1')}{=} -\det\begin{pmatrix} 1 & 1 & 3 & 5 \\ 0 & 1 & -2 & -2 \\ 0 & 0 & 2 & 3 \\ 0 & -2 & -4 & -6 \end{pmatrix}$$
$$\overset{(D1')}{=} -\det\begin{pmatrix} 1 & 1 & 3 & 5 \\ 0 & 1 & -2 & -2 \\ 0 & 0 & 2 & 3 \\ 0 & 0 & -8 & -10 \end{pmatrix} \overset{(D1')}{=} -\det\begin{pmatrix} 1 & 1 & 3 & 5 \\ 0 & 1 & -2 & -2 \\ 0 & 0 & 2 & 3 \\ 0 & 0 & 0 & 2 \end{pmatrix} = -4.$$

Her har vi først anvendt (D5') til at ombytte den anden og tredje række, hvilket ændrer determinantens fortegn. Derefter har vi anvendt (D1') til at addere -1 gange 1. række til 2. række og 3 gange 1. række til 4. række. Vi har så igen anvendt (D1') til at addere 2 gange 2. række til 4. række, og endelig har vi anvendt (D1') til at addere 4 gange 3. række til 4. række. Da anvendelser af (D1') ikke ændrer determinanten, har vi ved brug af rækkeoperationer identificeret $\det(A) = -\det(B)$, hvor B er triangulær. Endelig viser sætning 3.3.3, at $\det(B) = 4$, og vi konkluderer derfor, at $\det(A) = -4$.

Eksempel 3.3.5 Til sammenligning bestemmer vi determinanten af 4×4-matricen A i eksempel 3.3.4 ved Laplace udvikling langs 2. række.

$$
\begin{aligned}
\det\begin{pmatrix} 1 & 1 & 3 & 5 \\ 0 & 0 & 2 & 3 \\ 1 & 2 & 1 & 3 \\ 3 & 1 & 5 & 9 \end{pmatrix} &= -0 + 0 - 2 \cdot \det\begin{pmatrix} 1 & 1 & 3 & 5 \\ 0 & 0 & 2 & 3 \\ 1 & 2 & 1 & 3 \\ 3 & 1 & 5 & 9 \end{pmatrix} + 3 \cdot \det\begin{pmatrix} 1 & 1 & 3 & 5 \\ 0 & 0 & 2 & 3 \\ 1 & 2 & 1 & 3 \\ 3 & 1 & 5 & 9 \end{pmatrix} \\
&= -2 \cdot \det\begin{pmatrix} 1 & 1 & 5 \\ 1 & 2 & 3 \\ 3 & 1 & 9 \end{pmatrix} + 3 \cdot \det\begin{pmatrix} 1 & 1 & 3 \\ 1 & 2 & 1 \\ 3 & 1 & 5 \end{pmatrix} \\
&= -2 \cdot \left(\begin{pmatrix} 1 & 1 & 5 \\ 1 & 2 & 3 \\ 3 & 1 & 9 \end{pmatrix} \cdot 5 - \det\begin{pmatrix} 1 & 1 & 5 \\ 1 & 2 & 3 \\ 3 & 1 & 9 \end{pmatrix} \cdot 3 + \det\begin{pmatrix} 1 & 1 & 5 \\ 1 & 2 & 3 \\ 3 & 1 & 9 \end{pmatrix} \cdot 9 \right) \\
&\quad + 3 \cdot \left(3 \cdot \det\begin{pmatrix} 1 & 1 & 3 \\ 1 & 2 & 1 \\ 3 & 1 & 5 \end{pmatrix} - 1 \cdot \begin{pmatrix} 1 & 1 & 3 \\ 1 & 2 & 1 \\ 3 & 1 & 5 \end{pmatrix} + 5 \cdot \det\begin{pmatrix} 1 & 1 & 3 \\ 1 & 2 & 1 \\ 3 & 1 & 5 \end{pmatrix} \right) \\
&= -2 \cdot \left(\det\begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix} \cdot 5 - \det\begin{pmatrix} 1 & 1 \\ 3 & 1 \end{pmatrix} \cdot 3 + \det\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \cdot 9 \right) \\
&\quad + 3 \cdot \left(3 \cdot \det\begin{pmatrix} 1 & 3 \\ 2 & 1 \end{pmatrix} - 1 \cdot \det\begin{pmatrix} 1 & 3 \\ 1 & 1 \end{pmatrix} + 5 \cdot \det\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \right) \\
&= -2 \cdot \big((-5) \cdot 5 - (-2) \cdot 3 + 1 \cdot 9 \big) \\
&\quad + 3 \cdot \big(3 \cdot (-5) - 1 \cdot (-2) + 5 \cdot 1 \big) \\
&= -2 \cdot (-10) + 3 \cdot (-8) = 20 - 24 = -4.
\end{aligned}
$$

Vi har valgt at udvikle determinanten langs 2. række, fordi den har mange nuller. Vi har derefter udviklet den første og anden 3×3-matrix efter henholdsvis 3. søjle og 2. række, markeret med rødt. Og endelig har vi brugt formlen i sætning 3.1.1 til at udregne de seks 2×2-determinanter.

Laplace udvikling er mest anvendelig til bestemmelse af determinanten af matricer med (rigtigt) mange nuller. Det er dog lettere at lave fejl og sværere at opdage dem ved Laplace udvikling end ved udregning ved hjælp af række- og søjleoperationer.

Eksempel 3.3.6 Vi udregner determinanten af matricen

$$
A = \begin{pmatrix} 2 & 16 & 3 \\ 4 & 8 & -6 \\ 8 & 16 & 12 \end{pmatrix}
$$

ved at anvende søjleoperationer til at omdanne A til den triangulære matrix B og holde styr på, hvordan dette påvirker determinanten.

$$\det(A) = \det\begin{pmatrix} 2 & 16 & 3 \\ 4 & 8 & -6 \\ 8 & 16 & 12 \end{pmatrix} \overset{(D2)}{=} \det\begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & -2 \\ 4 & 2 & 4 \end{pmatrix} \cdot 2^4 \cdot 3 \overset{(D6)}{=} \det\begin{pmatrix} 1 & 0 & 0 \\ 2 & -3 & -4 \\ 4 & -6 & 0 \end{pmatrix} \cdot 2^4 \cdot 3$$

$$\overset{(D2)}{=} \det\begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 1 \\ 4 & 2 & 0 \end{pmatrix} \cdot 2^6 \cdot 3^2 \overset{(D6)}{=} \det\begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 4 & 2 & -2 \end{pmatrix} \cdot 2^6 \cdot 3^2 = -2^7 3^2 = -1152.$$

Vi har først anvendt (D2) til at flytte den fælles faktor 2 fra første søjle, den fælles faktor 8 fra anden søjle og den fælles faktor 3 fra tredje søjle udenfor determinanten; dernæst har vi anvendt (D6) og har adderet første søjle gange -2 til anden søjle og første søjle gange -1 til tredje søjle; vi har så anvendt (D2) til at flytte den fælles faktor -3 i anden søjle og den fælles faktor -4 i tredje søjle udenfor determinanten; og endelig har vi anvendt (D6) og har adderet anden søjle gange -1 til tredje søjle. Sætning 3.3.3 viser, at determinanten af den nedre triangulære matrix er lig med -2.

Eksempel 3.3.7 Man kan ofte med fordel anvende både (D1)–(D6) og de tilsvarende rækkevarianter (D1')–(D6') til at udregne determinanter. Som eksempel udregner vi følgende determinant ved at ombytte rækker og søjler, idet vi husker, at en ombytning af to rækker eller søjler skifter fortegn på determinanten.

$$\det\begin{pmatrix} 0 & 0 & 0 & 0 & 3 \\ 0 & 2 & 0 & 0 & 5 \\ 0 & 3 & -2 & 0 & -4 \\ 0 & -6 & 1 & 2 & 2 \\ 8 & 1 & 2 & 3 & 4 \end{pmatrix} \overset{(D5)}{=} -\det\begin{pmatrix} 8 & 1 & 2 & 3 & 4 \\ 0 & 2 & 0 & 0 & 5 \\ 0 & 3 & -2 & 0 & -4 \\ 0 & -6 & 1 & 2 & 2 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix}$$

$$\overset{(D5)}{=} +\det\begin{pmatrix} 8 & 1 & 2 & 3 & 4 \\ 0 & -6 & 1 & 2 & 2 \\ 0 & 3 & -2 & 0 & -4 \\ 0 & 2 & 0 & 0 & 5 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix} \overset{(D5')}{=} -\det\begin{pmatrix} 8 & 3 & 2 & 1 & 4 \\ 0 & 2 & 1 & -6 & 2 \\ 0 & 0 & -2 & 3 & -4 \\ 0 & 0 & 0 & 2 & 5 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix} = 2^6 \cdot 3 = 192.$$

Her har vi markeret med rødt de rækker og søjler, vi ombytter, og i den øvre triangulære matrix til sidst har vi markeret diagonalindgangene med blåt. Til sidst har vi benyttet sætning 3.3.3 til at udregne determinanten af den triangulære matrix.

3.4 Determinant og invers matrix

Vi har allerede set, at determinanten af en invertibel matrix er invertibel. Vi skal nu vise det omvendte udsagn, idet vi giver en formel for den inverse matrix udtrykt ved determinanten. Vi husker på, at et element i et legeme er invertibelt, hvis og kun hvis det er forskellig fra nul.

Sætning 3.4.1 *Lad $\mathbb{F}$ være et legeme, lad $A \in M_n(\mathbb{F})$, og antag, at $\det(A) \in \mathbb{F}$ er invertibel. Da er matricen A invertibel og matricen $C \in M_n(\mathbb{F})$, hvis (i,j)'te indgang er*

$$c_{ij} = (-1)^{i+j}\det(A_{ji})\det(A)^{-1},$$

er den inverse matrix af A. Her er A_{ji} den (j,i)'te undermatrix af A.

Bevis Vi udregner den (i,k)'te indgang

$$d_{ik} = \sum_{j=1}^{n} a_{ij}c_{jk} = \sum_{j=1}^{n}(-1)^{j+k}a_{ij}\det(A_{kj})\det(A)^{-1}$$

i produkt matricen $D = AC$, og betragter $i = k$ og $i \neq k$ særskilt. Hvis $i = k$, da er

$$d_{ii} = \Big(\sum_{j=1}^{n}(-1)^{j+i}a_{ij}\det(A_{ij})\Big)\det(A)^{-1} = \det(A)\det(A)^{-1} = 1,$$

hvor den midderste identitet er Laplace udvikling af $\det(A)$ langs i'te række. Hvis $i \neq k$, da betragter vi matricen B, der fremkommer fra A ved at erstatte den k'te række med den i'te række. På den ene side viser (D3'), at $\det(B) = 0$, da den i'te og k'te række i B er ens. Og på den anden side giver Laplace udvikling langs den k'te række, at

$$\det(B) = \sum_{j=1}^{n}(-1)^{k+j}b_{kj}\det(B_{kj}) = \sum_{j=1}^{n}(-1)^{k+j}a_{ij}\det(A_{kj}),$$

hvor den anden lighed følger fra definitionen af B. Derfor er

$$d_{ik} = (-1)^{i+k}\det(B)\det(A)^{-1} = 0.$$

Dermed har vi vist at $AC = D = I$. Vi ser ligeledes, at $CA = I$, ved enten at foretage en lignende udregning eller ved at bruge sætning 2.5.15, så $C = A^{-1}$. □

Bemærkning 3.4.2 Vi lader $A \in M_n(\mathbb{R})$ være en invertibel matrix og antager, at alle indgange i A er heltal. Da er $\det(A)$ og $\det(A_{ji})$ også heltal, og sætning 3.4.1 viser derfor, at indgangene i $C = A^{-1}$ er rationale tal med samme nævner $\det(A)$. Dette forklarer at matricen A^{-1} i eksempel 2.4.9 havde indgangene af form $m/5$ med $m \in \mathbb{Z}$, idet $\det(A) = 5$ i dette eksempel.

Eksempel 3.4.3 For 2×2-matricer viser sætning 3.4.1, at

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

forudsat at determinanten $d = a_{11}a_{22} - a_{21}a_{12}$ er invertibel.

Bemærkning 3.4.4 Hvis $\mathbb{F}$ er et legeme, så skriver vi

$$SL_n(\mathbb{F}) \subset GL_n(\mathbb{F})$$

for delmængden af de invertible $n \times n$-matricer, der består af de $n \times n$-matricer, der har determinant lig med 1. At sådanne matricer er invertible følger fra sætning 3.4.1, idet $1 \in \mathbb{F}$ er invertibel. Det følger endvidere fra korollar 3.2.14, at hvis $A, B \in SL_n(\mathbb{F})$, da er også $AB \in SL_n(\mathbb{F})$. Dermed er $(SL_n(\mathbb{F}), \cdot)$ en gruppe, som defineret i bemærkning 2.4.12. Vi kalder denne gruppe for den specielle lineære gruppe.

Hvis A er invertibel, så har ligningssystemet $A\boldsymbol{x} = \boldsymbol{b}$ den entydige løsning $\boldsymbol{x} = A^{-1}\boldsymbol{b}$. Denne løsning kan også udtrykkes ganske elegant ved hjælp af determinanter.

Sætning 3.4.5 (Cramers regel) *Lad A være en invertibel $n \times n$-matrix med indgange i et legeme $\mathbb{F}$, lad $\boldsymbol{a}_1, \dots, \boldsymbol{a}_n \in \mathbb{F}^n$ være søjlevektorerne i A og lad $b \in \mathbb{F}^n$ være en vilkårlig søjlevektor. Da er den entydige løsning $\boldsymbol{x} = A^{-1}\boldsymbol{b}$ til ligningen $A\boldsymbol{x} = \boldsymbol{b}$ givet ved*

$$x_i = \frac{\det\begin{pmatrix}\boldsymbol{a}_1 & \dots & \boldsymbol{a}_{i-1} & \boldsymbol{b} & \boldsymbol{a}_{i+1} & \dots & \boldsymbol{a}_n\end{pmatrix}}{\det\begin{pmatrix}\boldsymbol{a}_1 & \dots & \boldsymbol{a}_{i-1} & \boldsymbol{a}_i & \boldsymbol{a}_{i+1} & \dots & \boldsymbol{a}_n\end{pmatrix}},$$

hvor x_i er den i'te koordinat i vektoren $\boldsymbol{x}$.

Bevis Formlen for $A^{-1} = C = (c_{ij})$ fra sætning 3.4.5 viser, at

$$x_i = \begin{pmatrix}c_{i1} & c_{i2} & \cdots & c_{in}\end{pmatrix}\boldsymbol{b} = \sum_{j=1}^{n} c_{ij}b_j = \sum_{j=1}^{m} (-1)^{i+j}\det(A_{ji})\det(A)^{-1}b_j,$$

og vi skal derfor vise, at

$$\sum_{j=1}^{m} (-1)^{i+j} b_j \det(A_{ji}) = \det\begin{pmatrix}\boldsymbol{a}_1 & \dots & \boldsymbol{a}_{i-1} & \boldsymbol{b} & \boldsymbol{a}_{i+1} & \dots & \boldsymbol{a}_n\end{pmatrix}.$$

Men summen på venstre side er præcis udviklingen langs i'te søjle af determinanten på højre side, så den ønskede identitet følger fra sætning 3.2.7. $\square$

3.5 Polynomier

Vi har allerede bemærket, at for at kunne definere determinanten er det nødvendigt, at den kommutative lov gælder for multiplikation af skalarer. For (D2) medfører, at

$$ab = \det\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}ab = \det\begin{pmatrix}a & 0\\ 0 & 1\end{pmatrix}b = \det\begin{pmatrix}a & 0\\ 0 & b\end{pmatrix} = \det\begin{pmatrix}1 & 0\\ 0 & b\end{pmatrix}a = \det\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}ba = ba.$$

Vi har dog intetsteds i dette kapitel anvendt, at ethvert element $a \in \mathbb{F}$ med $a \neq 0$ er invertibelt. Vi har således kun brugt, at $\mathbb{F}$ er en kommutativ ring, som vi nu definerer. Specielt gælder alle resultater i dette kapitel også, hvis vi erstatter $\mathbb{F}$ med den kommutative ring $\mathbb{F}[t]$ af polynomier med koefficienter i $\mathbb{F}$, som vi indfører i dette afsnit.

Definition 3.5.1 En *kommutativ ring* er en triple $(R, +, \cdot)$ af en mængde R og to afbildninger $+\colon R \times R \to R$ og $\cdot\colon R \times R \to R$, sådan at:

(A1) For alle $a, b, c \in R$ er $(a + b) + c = a + (b + c)$.

(A2) Der findes et element $0 \in R$, sådan at $a + 0 = a = 0 + a$ for alle $a \in R$.

(A3) For alle $a \in R$, findes $b \in R$, sådan at $a + b = 0 = b + a$.

(A4) For alle $a, b \in R$ er $a + b = b + a$.

(P1) For alle $a, b, c \in R$ er $(a \cdot b) \cdot c = a \cdot (b \cdot c)$.

(P2) Der findes et element $1 \in R$, sådan at $a \cdot 1 = a = 1 \cdot a$ for alle $a \in R$.

(P4) For alle $a, b \in R$ er $a \cdot b = b \cdot a$.

(D1) For alle $a, b, c \in R$ er $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$

(D2) For alle $a, b, c \in R$ er $(a + b) \cdot c = (a \cdot c) + (b \cdot c)$.

Vi misbruger ofte notation og skiver blot R for den kommutative ring $(R, +, \cdot)$. Et element $a \in R$ er *invertibelt*, hvis der findes $b \in R$, sådan at $a \cdot b = 1 = b \cdot a$. Vi bemærker, at et *legeme* per definition er en kommutativ ring, hvori ethvert element, der ikke er lig med 0, er invertibelt.

Eksempel 3.5.2 (1) Mængden $\mathbb{Z}$ af hele tal sammen med afbildningerne $+\colon \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$ og $\cdot\colon \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$ givet ved henholdsvis sum og produkt af hele tal, udgør en kommutativ ring. Denne kommutative ring er ikke et legeme, da der for eksempel ikke findes noget helt tal b, sådan at $2 \cdot b = 1 = b \cdot 2$.
(2) Mængden $C^0(\mathbb{R})$ af kontinuerte funktioner $f\colon \mathbb{R} \to \mathbb{R}$ sammen med afbildningerne $+\colon C^0(\mathbb{R}) \times C^0(\mathbb{R}) \to C^0(\mathbb{R})$ og $\cdot\colon C^0(\mathbb{R}) \times C^0(\mathbb{R}) \to C^0(\mathbb{R})$ defineret ved

$$(f + g)(x) = f(x) + g(x) \qquad \text{og} \qquad (f \cdot g)(x) = f(x) \cdot g(x)$$

udgør en kommutativ ring. Elementerne "0" og "1" i denne kommutative ring er de konstante funktioner med værdi henholdsvis 0 og 1. Denne kommutative ring er heller ikke et legeme, da der for eksempel ikke findes nogen kontinuert funktion $f\colon \mathbb{R} \to \mathbb{R}$ sådan at $\mathrm{id} \cdot f = 1 = f \cdot \mathrm{id}$. For $(\mathrm{id} \cdot f)(0) = \mathrm{id}(0) \cdot f(0) = 0 \cdot f(0) = 0 \neq 1 = 1(0)$.

Vi definerer et *polynomium* med koefficienter i en kommutativ ring R til at være en følge $(a_0, a_1, a_2, \dots)$ af elementer i en kommutativ ring R, sådan at $a_i \neq 0$ for højst endeligt mange i. Vi vil dog altid indføre en variabel "t" og skrive

$$\sum_{i \geq 0} a_i t^i$$

i stedet for $(a_0, a_1, a_2, \dots)$, og vi siger da, at a_i er koefficienten til t^i. Vi definerer nu henholdsvis *summen* og *produktet* af to polynomier ved

$$\begin{aligned}
(\sum_{i \geq 0} a_i t^i) + (\sum_{i \geq 0} b_i t^i) &= \sum_{i \geq 0} (a_i + b_i) t^i \\
(\sum_{i \geq 0} a_i t^i) \cdot (\sum_{j \geq 0} b_j t^j) &= \sum_{k \geq 0} (\sum_{i+j=k} a_i \cdot b_j) t^k.
\end{aligned}$$

Vi bemærker, at sum og produkt af polynomier er defineret ved, at vi lader som om, den variable "t" var et element i R. Normalt undlader vi at skrive de koefficienter a_i, der er lig med 0. Så en typisk sum og et typisk produkt af to polynomier er for eksempel

$$\begin{aligned}
(2t^3 - t^2 + 3t + 1) + (t^2 + 2t + 1) &= 2t^3 + 5t + 2, \\
(2t^3 - t^2 + 3t + 1) \cdot (t^2 + 2t + 1) &= 2t^5 + 3t^4 + 3t^3 + 5t^2 + 5t + 1.
\end{aligned}$$

Specielt skriver vi 0 for det polynomium, hvis koefficienter alle er 0, og vi skriver 1 for det polynomium, hvis koefficient til t^i er 1 for $i = 0$ og 0 for $i > 0$.

Sætning 3.5.3 *Hvis R er en kommutativ ring, da udgør mængden $R[t]$ af polynomier med koefficienter i R sammen med sum og produkt af polynomier en kommutativ ring.*

Bevis Vi viser (P1) og (D1); (D2) vises ligesom (D1), mens de øvrige egenskaber følger umiddelbart fra de tilsvarende egenskaber for R. Ifølge definitionen af produktet af polynomier samt henholdsvis (D2) og (D1) for skalarer gælder der, at

$$\begin{aligned}
(\sum_{i \geq 0} a_i t^i \cdot \sum_{j \geq 0} b_j t^j) \cdot \sum_{k \geq 0} c_k t^k &= \sum_{l \geq 0} (\sum_{i+j+k=l} (a_i b_j) c_k) t^l, \\
\sum_{i \geq 0} a_i t^i \cdot (\sum_{j \geq 0} b_j t^j \cdot \sum_{k \geq 0} c_k t^k) &= \sum_{l \geq 0} (\sum_{i+j+k=l} a_i (b_j c_k)) t^l,
\end{aligned}$$

og dermed følger (P1) for polynomier fra (P1) for skalarer.

$$\begin{aligned}
\sum_{i \geq 0} a_i t^i \cdot (\sum_{j \geq 0} b_j t^j + \sum_{j \geq 0} c_j t^j) &= \sum_{k \geq 0} (\sum_{i+j=k} a_i (b_j + c_k)) t^k, \\
(\sum_{i \geq 0} a_i t^i \cdot \sum_{j \geq 0} b_j t^j) + (\sum_{i \geq 0} a_i t^i \cdot \sum_{j \geq 0} c_j t^j) &= \sum_{k \geq 0} (\sum_{i+j=k} (a_i b_j + a_i c_j)) t^k,
\end{aligned}$$

således at (D1) for polynomier følger fra (D1) for skalarer. □

Hvis d er et helt tal, så siger vi, at et polynomium

$$p = p(t) = \sum_{i \geq 0} a_i t^i \in R[t]$$

har *grad højst d* og skriver $\deg(p) \leq d$, hvis $a_i = 0$ for alle $i > d$. Vi skriver også

$$R[t]_{\leq n} = \{p \in R[t] \mid \deg(p) \leq d\} \subset R[t]$$

for delmængden af polynomier af grad højst d. Specielt er nulpolynomiet det eneste polynium, der har grad højst d for alle d, mens de konstante polynomier præcis er de polynomier, der har grad højst 0.

Lemma 3.5.4 *Lad $p, q \in R[t]$ være polynomier med koeffcienter i en kommutativ ring R. Da gælder følgende:*

(1) *Hvis* $\deg(p) \leq d$ *og* $\deg(q) \leq e$, *da er* $\deg(p+q) \leq \max\{d, e\}$.

(2) *Hvis* $\deg(p) \leq d$ *og* $\deg(q) \leq e$, *da er* $\deg(p \cdot q) \leq de$.

Bevis Både (1) og (2) følger direkte fra definitionerne. □

Vi betragter nu determinantafbildningen

$$\det\colon M_n(R[t]) \to R[t]$$

for $n \times n$-matricer med indgange i $R[t]$. Denne findes ifølge sætning 3.2.1, der som nævnt gælder for enhver kommutativ ring.

Sætning 3.5.5 *Lad R være en kommutativ ring, lad $R[t]$ være den kommutative ring af polynomier med koefficienter i R, og lad $P = (p_{ij}) \in M_n(R[t])$ være en matrix med indgange i $R[t]$. Hvis $\deg(p_{ij}) \leq d$ for alle $1 \leq i, j \leq n$, da gælder*

$$\deg(\det(P)) \leq dn.$$

Bevis Beviset er igen ved induktion på $n \geq 0$; tilfældet $n = 0$ er trivielt. Så vi antager, at sætningen gælder for $n = r - 1$ og viser, at den gælder for $n = r$. Ved Laplace udvikling langs første række får vi, at

$$\det(P) = \sum_{j=1}^{r} (-1)^{1+j} p_{1j} \det(P_{1j}).$$

Her er $\deg(p_{1j}) \leq d$ per antagelse, og $\deg(\det(P_{1j})) \leq d(n-1)$ per induktion, og derfor viser lemma 3.5.4, at $\deg(\det(P)) \leq d + d(n-1) = dn$ som ønsket. □

Vi forklarer nu, hvordan den variable "t" kan tilordnes en værdi, dvs. formelt hvordan man evaluerer et polynomium på et element. En afbildning

$$f : R \to S$$

mellem to kommutative ringe siges at være en *ring homomorfi*, hvis den opfylder:

(R1) For alle $a, b \in R$ er $f(a+b) = f(a) + f(b)$.

(R2) For alle $a, b \in R$ er $f(a \cdot b) = f(a) \cdot f(b)$.

(R3) $f(1_R) = 1_S$.

En ring homomorfi $\eta \colon R \to S$ og et element $s \in S$ giver anledning til en ring homomorfi

$$\mathrm{ev}_{\eta,s} : R[t] \to S,$$

der er defineret ved formlen

$$\mathrm{ev}_{\eta,s}(\sum_{i \geq 0} a_i t^i) = \sum_{i \geq 0} \eta(a_i) s^i.$$

Bemærk, at sum og produkt af polynomier netop er defineret, sådan at denne afbildning er en ring homomorfi. Hvis $\eta \colon R \to S$ er underforstået, så skriver vi også $p(s)$ i stedet for $\mathrm{ev}_{\eta,s}(p)$ og siger, at $p(s)$ fremkommer fra $p(t)$ ved at *substituere $s \in S$ for t*.

Eksempel 3.5.6 (1) Lad $R = \mathbb{R}$ og $S = \mathbb{R}$ med $\eta = \mathrm{id} \colon \mathbb{R} \to \mathbb{R}$. Vi kan da substituere $a \in \mathbb{R}$ for t, hvilket giver $p(a) = \mathrm{ev}_a(p) \in \mathbb{R}$.
(2) Vi betragter også $R = \mathbb{R}$ og $S = \mathbb{C}$ og lader $\eta \colon \mathbb{R} \to \mathbb{C}$ være inklusionen af de reelle tal i de komplekse tal. Vi kan da substituere $a \in \mathbb{C}$ for t, hvilket giver $p(a) = \mathrm{ev}_a(p) \in \mathbb{C}$.
(3) Lad $R = \mathbb{R}$ og lad $S = C^0(\mathbb{R})$ være den kommutative ring af kontinuerte funktioner fra eksempel 3.5.2. Vi lader $\eta \colon \mathbb{R} \to C^0(\mathbb{R})$ være afbildningen, der til $a \in \mathbb{R}$ tilordner den konstante funktion $\eta(a) \colon \mathbb{R} \to \mathbb{R}$ med værdi a, og bemærker, at identitetsafbildningen $\mathrm{id} \colon \mathbb{R} \to \mathbb{R}$ et element i $S = C^0(\mathbb{R})$. Vi skriver "x" for dette element, hvilket giver

$$\mathrm{ev}_{\eta,x} : \mathbb{R}[t] \to C^0(\mathbb{R}),$$

der er en ring homomorfi og afbilder $p(t)$ til den kontinuerte funktion $p(x)$, der fås ved at substituere $x \in S$ for t.

Vi siger, at $s \in S$ er en *rod* i polynomiet $p(t) \in R[t]$, hvis $p(s) = 0$. For eksempel har $p(t) = t^2 + 1 \in \mathbb{R}[t]$ ingen rødder i $\mathbb{R}$, mens $\pm i$ er rødder i $\mathbb{C}$. Den følgende vigtige sætning kaldes for algebraens fundamentalsætning.

Sætning 3.5.7 (Gauss-Argand) *Ethvert ikke-konstant polynomium med koefficienter i $\mathbb{C}$ har mindst en rod i $\mathbb{C}$.*

Bevis Vi henviser til en af de følgende lærebøger for et bevis.

(1) David S. Dummit og Richard M. Foote. Abstract algebra. Third Edition. John Wiley & Sons, Inc., Hoboken, NJ, 2004.

(2) Serge Lang. Algebra. Revised third edition. Graduate Texts in Mathematics, 211. Springer-Verlag, New York, 2002.

De to bøger præsenterer det samme bevis. I den først bog, er beviset givet i Theorem 35 på side 615–617; og i den anden bog, er beviset givet i Example 5 på side 272–273. □

Man viser da ved polynomiumsdivision, at ethvert polynomium $p(t) \in \mathbb{C}[t]$ af grad præcis $d \geq 0$ har præcis d rødder i $\mathbb{C}$, talt med multiplicitet.

Sætning 3.5.8 *Lad $p(t) = a_d t^d + a_{d-1} t^{d-1} + \cdots + a_1 t + a_0 \in \mathbb{C}[t]$ være et polynomium med koefficienter i $\mathbb{C}$, sådan at $d \geq 0$ og $a_d \neq 0$. Da findes $z_1, \dots, z_d \in \mathbb{C}$, sådan at*

$$p(t) = a_d(t - z_1) \cdot \cdots \cdot (t - z_d).$$

Bevis Påstanden følger fra sætning 3.5.7 ved brug af den euklidiske algoritme. Denne er bevist i Theorem 1.1 i Lang's bog (2) på side 173–174. □

Vi bemærker, at beviset for algebraens fundamentalsætning ikke er konstruktivt, så sætning siger altså blot, at der en findes en rod, men siger ikke noget om, hvordan man bærer sig ad med at finde en rod. For et generelt polynomium $p(t) \in \mathbb{C}[t]$, kan man kun vise, at $z \in \mathbb{C}$ tilnærmelsesvis opfylder $p(z) = 0$. For ligegyldigt, hvor mange cifre af $p(z)$, man udregner, så kan man generelt ikke vide, at der ikke vil dukke ikke-nul cifre op i $p(z)$, hvis man regner længere. Man skal derfor være påpasselig med også at forstå, hvordan denne unøjagtighed influerer de konklusioner, man træffer ved at antage, at ligheden $p(z) = 0$ gælder eksakt.

4 Vektorrum

Hvis A er en $m \times n$-matrix med indgange i et legeme $\mathbb{F}$, så har vi i kapitel 1 beskrevet en algoritme, der producerer en parametrisering af løsningsmængden

$$N_A = \{\boldsymbol{x} \in \mathbb{F}^n \mid A\boldsymbol{x} = \boldsymbol{0}\} \subset \mathbb{F}^n$$

til det homogene lineære ligningssystem $A\boldsymbol{x} = \boldsymbol{0}$; se sætning 1.2.12. Denne algoritme afhænger dog af nogle temmeligt arbitrære valg såsom definitionen af en matrix på reduceret echelon form. I modsætning hertil afhænger løsningmængden N_A selv ikke af nogle valg, og vi ønsker derfor at udvikle vores begrebsverden, sådan at vi direkte kan udtrykke, hvad denne løsningsmængde "er" for en størrelse. Hertil indfører vi begrebet *vektorrum*, som vi allerede har stiftet bekendtskab med i afsnit 2.2.

4.1 Vektorrum og lineære afbildninger

Vi lader $\mathbb{F}$ være et legeme, og som før tænker vi på $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$. Hele dette kapitel virker dog ligeså godt for skævlegemer, så $\mathbb{F} = \mathbb{H}$ er også tilladt. Et (abstrakt) vektorrum består af en mængde, hvis elementer vi kalder vektorer, udstyret med to operationer "+" og "$\cdot$", som vi kalder vektorsum og skalar multiplikation.

Definition 4.1.1 Lad $\mathbb{F}$ være et legeme. Et *højre $\mathbb{F}$-vektorrum* er en triple $(V, +, \cdot)$, der består af en mængde V og to afbildninger $+ : V \times V \to V$ og $\cdot : V \times \mathbb{F} \to V$, sådan at:

(A1) For alle $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in V$ er $(\boldsymbol{x} + \boldsymbol{y}) + \boldsymbol{z} = \boldsymbol{x} + (\boldsymbol{y} + \boldsymbol{z})$.

(A2) Der findes et element $\boldsymbol{0} \in V$, sådan at $\boldsymbol{x} + \boldsymbol{0} = \boldsymbol{x} = \boldsymbol{0} + \boldsymbol{x}$ for alle $\boldsymbol{x} \in V$.

(A3) For alle $\boldsymbol{x} \in V$, findes $\boldsymbol{y} \in V$, sådan at $\boldsymbol{x} + \boldsymbol{y} = \boldsymbol{0} = \boldsymbol{y} + \boldsymbol{x}$.

(A4) For alle $\boldsymbol{x}, \boldsymbol{y} \in V$ er $\boldsymbol{x} + \boldsymbol{y} = \boldsymbol{y} + \boldsymbol{x}$.

(V1) For alle $\boldsymbol{x} \in V$ og $a, b \in \mathbb{F}$ er $(\boldsymbol{x} \cdot a) \cdot b = \boldsymbol{x} \cdot (ab)$.

(V2) For alle $\boldsymbol{x}, \boldsymbol{y} \in V$ og $a \in \mathbb{F}$ er $(\boldsymbol{x} + \boldsymbol{y}) \cdot a = (\boldsymbol{x} \cdot a) + (\boldsymbol{y} \cdot a)$.

(V3) For alle $\boldsymbol{x} \in V$ og $a, b \in \mathbb{F}$ er $\boldsymbol{x} \cdot (a + b) = (\boldsymbol{x} \cdot a) + (\boldsymbol{x} \cdot b)$.

(V4) For alle $\boldsymbol{x} \in V$ er $\boldsymbol{x} \cdot 1 = x$.

Man definerer *venstre* $\mathbb{F}$-*vektorrum* tilsvarende; den eneste forskel er, at vektorerne kan multipliceres med skalarer fra venstre i stedet for fra højre.

Bemærkning 4.1.2 Lad $\mathbb{F}$ være et legeme og lad $(V,+,\cdot)$ være et $\mathbb{F}$-vektorrum.
(1) Vektoren $\mathbf{0} \in V$ i (A2) er entydigt bestemt. For hvis $\mathbf{0}$ og $\mathbf{0}'$ begge opfylder (A2), da viser (A2) for henholdsvis $\mathbf{0}$ og $\mathbf{0}'$, at $\mathbf{0}' = \mathbf{0} + \mathbf{0}' = \mathbf{0}$. Vi kalder $\mathbf{0} \in V$ for *nulvektoren*.
(2) Givet $\boldsymbol{x} \in V$, da er vektoren $\boldsymbol{y} \in V$ i (A3) entydigt bestemt. For hvis $\boldsymbol{y}$ og $\boldsymbol{y}'$ begge opfylder (A3), så viser (A3) for henholdsvis $\boldsymbol{y}$ og $\boldsymbol{y}'$ samt (A1) og (A2), at

$$\boldsymbol{y} = \boldsymbol{y} + \mathbf{0} = \boldsymbol{y} + (\boldsymbol{x} + \boldsymbol{y}') = (\boldsymbol{y} + \boldsymbol{x}) + \boldsymbol{y}' = \mathbf{0} + \boldsymbol{y}' = \boldsymbol{y}'.$$

Vi skriver $-\boldsymbol{x}$ for vektoren $\boldsymbol{y}$ og kalder den for den *modsatte vektor* af $\boldsymbol{x}$.
(3) Vi definerer *differencen* af to vektorer $\boldsymbol{x}$ og $\boldsymbol{y}$ til at være vektoren $\boldsymbol{x} - \boldsymbol{y} = \boldsymbol{x} + (-\boldsymbol{y})$.
(4) Man viser som i sætning 0.1.3, at $\boldsymbol{x} \cdot 0 = \mathbf{0}$ og $\boldsymbol{x} \cdot (-1) = -\boldsymbol{x}$.

Der findes rigtig mange vektorrum inden for matematik. Vi giver nogle eksempler.

Eksempel 4.1.3 (1) Mængden $M_{m,1}(\mathbb{F})$ af $m \times 1$-matricer med indgange i $\mathbb{F}$ sammen med afbildningerne $+ : M_{m,1}(\mathbb{F}) \times M_{m,1}(\mathbb{F}) \to M_{m,1}(\mathbb{F})$ og $\cdot : M_{m,1}(\mathbb{F}) \times M_{1,1}(\mathbb{F}) \to \mathbb{M}_{m,1}(\mathbb{F})$, der er givet ved henholdsvis matrixsum og matrixprodukt, udgør et højre $\mathbb{F}$-vektorrum, idet vi identificerer $M_{1,1}(\mathbb{F})$ med $\mathbb{F}$. Vi skriver $(\mathbb{F}^m,+,\cdot)$ for dette højre $\mathbb{F}$-vektorrum og kalder det for $\mathbb{F}$-vektorrummet af søjlevektorer af dimension m. Vi har allerede betragtet dette vektorrum i afsnit 2.2, og vi skal senere se, at det faktisk har dimension m.

(2) Mængden $M_{1,n}(\mathbb{F})$ af $1 \times n$-matricer med indgange i $\mathbb{F}$ sammen med afbildningerne $+ : M_{1,n}(\mathbb{F}) \times M_{1,n}(\mathbb{F}) \to M_{1,n}(\mathbb{F})$ og $\cdot : M_{1,1}(\mathbb{F}) \times M_{1,n}(\mathbb{F}) \to \mathbb{M}_{1,n}(\mathbb{F})$ givet ved henholdsvis matrixsum og matrixprodukt udgør et venstre $\mathbb{F}$-vektorrum, hvor vi identificerer $M_{1,1}(\mathbb{F})$ med $\mathbb{F}$. Vi indfører ikke anden notation for dette vektorrum, der har dimension n.

(3) Hvis triplen $(\mathbb{F},+,\cdot)$ er et legeme, da er den samme triple også både et højre og et venstre vektorrum. For aksiomerne for et legeme i Definition 0.1.1 medfører aksiomerne for et vektorrum i Definition 4.1.1. Disse $\mathbb{F}$-vektorrum har begge dimension 1.

(4) Vi har i (3) givet mængden $\mathbb{C}$ af komplekse tal en struktur af højre $\mathbb{C}$-vektorrum. Vi kan give denne en struktur af højre $\mathbb{R}$-vektorrum med vektorrum $+ : \mathbb{C} \times \mathbb{C} \to \mathbb{C}$ og skalar multiplikation og $\cdot : \mathbb{C} \times \mathbb{R} \to \mathbb{C}$ defineret ved henholdsvis

$$(x_1 + ix_2) + (y_1 + iy_2) = (x_1 + y_1) + i(x_2 + y_2),$$
$$(x_1 + ix_2) \cdot a = x_1 a + ix_2 a.$$

Vi skal senere vise, at dette højre $\mathbb{R}$-vektorrum har dimension 2.

(5) Mængden $\mathbb{C}$ af komplekse tal kan ligeledes gives en struktur af højre $\mathbb{Q}$-vektorrum med vektorrum $+ : \mathbb{C} \times \mathbb{C} \to \mathbb{C}$ og skalar multiplikation $\cdot : \mathbb{C} \times \mathbb{Q} \to \mathbb{C}$ defineret ved

$$(x_1 + ix_2) + (y_1 + iy_2) = (x_1 + y_1) + i(x_2 + y_2),$$
$$(x_1 + ix_2) \cdot a = x_1 a + ix_2 a.$$

Dette højre $\mathbb{Q}$-vektorrum er har uendelig dimension $2^{\aleph_0}$.

(6) Mængden $\mathbb{F}[t]_{\leq d}$ af polynomier med koefficienter i $\mathbb{F}$ og af grad højst d sammen med afbildningerne $+\colon \mathbb{F}[t]_{\leq d} \times \mathbb{F}[t]_{\leq d} \to \mathbb{F}[t]_{\leq d}$ og $\cdot\colon \mathbb{F}[t]_{\leq d} \times \mathbb{F}[t]_{\leq 0} \to \mathbb{F}[t]_{\leq d}$ defineret til at være henholdsvis sum og produkt af polynomier udgør et højre $\mathbb{F}$-vektorrum, idet vi identificerer $\mathbb{F}[t]_{\leq 0}$ med $\mathbb{F}$. Det har dimension $d+1$.

(7) Mængden $C^0(\mathbb{R})$ af kontinuerte funktioner $f\colon \mathbb{R} \to \mathbb{R}$ sammen med afbildningerne $+\colon C^0(\mathbb{R}) \times C^0(\mathbb{R}) \to C^0(\mathbb{R})$ og $\cdot\colon C^0(\mathbb{R}) \times \mathbb{R} \to C^0(\mathbb{R})$ defineret ved

$$(f+g)(x) = f(x) + g(x) \qquad \text{og} \qquad (f \cdot a)(x) = f(x) \cdot a$$

udgør et højre $\mathbb{R}$-vektorrum. Det har uendelig dimension $2^{2^{\aleph_0}}$.

Vi vil i denne bog altid betragte højre $\mathbb{F}$-vektorrum, og vi vil derfor forkorte og sige $\mathbb{F}$-vektorrum i stedet for højre $\mathbb{F}$-vektorrum. Hvis $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, så siger vi også, at et $\mathbb{F}$-vektorrum er henholdsvis et reelt vektorrum og et komplekst vektorrum. Hvis $\mathbb{F}$ er underforstået, så vil vi sommetider sige vektorrum i stedet for $\mathbb{F}$-vektorrum. Desuden vil vi, som det er sædvane, misbruge notation og skrive V for vektorrummet $(V, +, \cdot)$.

Sætning 4.1.4 *Lad $\mathbb{F}$ være et legeme, lad $(V, +, \cdot)$ være et $\mathbb{F}$-vektorrum, og lad $V' \subset V$ være en delmængde, sådan at følgende gælder:*

(1) *Nulvektoren $\mathbf{0} \in V$ tilhører $V' \subset V$.*

(2) *Hvis $\boldsymbol{x}, \boldsymbol{y} \in V' \subset V$, så er også $\boldsymbol{x} + \boldsymbol{y} \in V' \subset V$.*

(3) *Hvis $\boldsymbol{x} \in V' \subset V$ og $a \in \mathbb{F}$, så er også $\boldsymbol{x} \cdot a \in V' \subset V$.*

Lad afbildningerne $+'\colon V' \times V' \to V'$ og $\cdot'\colon V' \times \mathbb{F} \to V'$ være defineret ved henholdsvis $\boldsymbol{x} +' \boldsymbol{y} = \boldsymbol{x} + \boldsymbol{y}$ og $\boldsymbol{x} \cdot' a = \boldsymbol{x} \cdot a$, dvs. med sum og skalarmultiplikation fra V. Da er også $(V', +', \cdot')$ et $\mathbb{F}$-vektorrum.

Bevis Det følger fra henholdsvis (2) og (3), at afbildningerne $+'$ og $\cdot'$ er veldefinerede, og vi ønsker at vise, at de opfylder (A1)–(A4) og (V1)–(V4). Da afbildningerne $+$ og $\cdot$ opfylder (A1)–(A4) og (V1)–(V4), følger det umiddelbart, at også afbildningerne $+'$ og $\cdot'$ opfylder (A1), (A4) og (V1)–(V4). Vi viser nu, at $(V', +', \cdot')$ også opfylder (A2) og (A3). Så lad $\mathbf{0} \in V$ være nulvektoren. Per antagelse (1) er $\mathbf{0} \in V'$, og desuden er

$$\boldsymbol{x} +' \mathbf{0} = \boldsymbol{x} + \mathbf{0} = \boldsymbol{x} = \mathbf{0} + \boldsymbol{x} = \mathbf{0} +' \boldsymbol{x}$$

fordi (A2) gælder i V. Dette viser, at også $(V', +', \cdot)$ opfylder (A2). Hvis endelig $\boldsymbol{x} \in V'$, så viser (3), at også $\boldsymbol{x} \cdot (-1) \in V'$, og udregningen

$$\boldsymbol{x} +' (\boldsymbol{x} \cdot (-1)) = \boldsymbol{x} + (\boldsymbol{x} \cdot (-1)) = \boldsymbol{x} + (-\boldsymbol{x}) = \mathbf{0} = (-\boldsymbol{x}) + \boldsymbol{x} = (\boldsymbol{x} \cdot (-1)) + \boldsymbol{x} = (\boldsymbol{x} \cdot (-1)) +' \boldsymbol{x}$$

viser derfor, at $(V', +', \cdot')$ opfylder (A3). $\square$

Vi kalder $\mathbb{F}$-vektorrummet $(V', +', \cdot')$ defineret i sætning 4.1.4 for et *underrum* af det givne $\mathbb{F}$-vektorrum $(V, +, \cdot)$. Vi vil dog normalt blot sige, at V' er et underrum af V. Vi vil ligeledes misbruge notation og skrive $+$ og $\cdot$ i stedet for henholdsvis $+'$ og $\cdot'$.

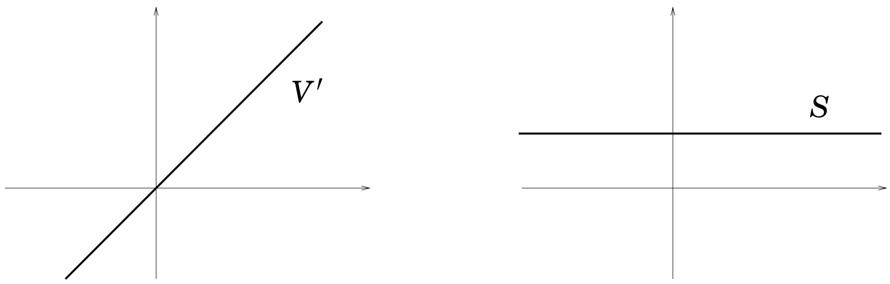
Eksempel 4.1.5 (1) Vi påstår, at delmængden

$$V' = \left\{ \boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 \;\middle|\; x_1 = x_2 \right\} \subset V = \mathbb{R}^2$$

opfylder (1)–(3) i sætning 4.1.4. For $\boldsymbol{0} \in V'$, da $0 = 0$, så (1) gælder; hvis $\boldsymbol{x}, \boldsymbol{y} \in V'$, så er også $\boldsymbol{x} + \boldsymbol{y} \in V'$, fordi "$x_1 = x_2$ og $y_1 = y_2$" medfører "$x_1 + y_1 = x_2 + y_2$", så (2) gælder; og hvis $\boldsymbol{x} \in V'$ og $a \in \mathbb{R}$, da er $\boldsymbol{x} \cdot a \in V'$, idet "$x_1 = x_2$" medfører "$x_1 a = x_2 a$", så (3) gælder.
(2) Modsat påstår vi, at delmængden

$$S = \left\{ \boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 \;\middle|\; x_2 = 1 \right\} \subset V = \mathbb{R}^2$$

ikke opfylder (1)–(3) i sætning 4.1.4. Den opfylder faktisk *ingen* af (1)–(3) i sætning 4.1.4. For $\boldsymbol{0} \notin S$, da $0 \neq 1$; hvis $\boldsymbol{x}, \boldsymbol{y} \in S$, da er $\boldsymbol{x} + \boldsymbol{y} \notin S$, da $1 + 1 \neq 1$; og hvis $\boldsymbol{x} \in S$, da er $b x \cdot a \notin S$, da $1 \cdot a \neq 1$, medmindre $a = 1$.



Figur 4.1: $V' \subset \mathbb{R}^2$ er et underrum, mens $S \subset \mathbb{R}^2$ ikke er det; se Eksempel 4.1.5.

Lemma 4.1.6 *Hvis A er en $m \times n$-matrix med indgange i et legeme $\mathbb{F}$, da er delmængden*

$$N_A = \{\boldsymbol{x} \in \mathbb{F}^n \mid A\boldsymbol{x} = \boldsymbol{0}\} \subset \mathbb{F}^n$$

af løsninger til det homogene ligningssystem $A\boldsymbol{x} = \boldsymbol{0}$ et underrum af $\mathbb{F}^n$.

Bevis Vi viser, at (1)–(3) i sætning 4.1.4 er opfyldt. Da $A\boldsymbol{0} = \boldsymbol{0}$, er (1) opfyldt; hvis $A\boldsymbol{x} = \boldsymbol{0}$ og $A\boldsymbol{y} = \boldsymbol{0}$, da er $A(\boldsymbol{x} + \boldsymbol{y}) = A\boldsymbol{x} + A\boldsymbol{y} = \boldsymbol{0}$, så (2) er opfyldt; og hvis $A\boldsymbol{x} = \boldsymbol{0}$, da er også $A(\boldsymbol{x}a) = (A\boldsymbol{x})a = \boldsymbol{0}a = \boldsymbol{0}$, så (3) er opfyldt. □

Vi kalder underrummet $N_A \subset \mathbb{F}^n$ for *nulrummet* af matricen $A \in M_{m,n}(\mathbb{F})$. Dette består per definition af mængden af løsninger til den homogene ligning $A\boldsymbol{x} = \boldsymbol{0}$. Vi bemærker, at mængden af løsninger til den inhomogene ligning $A\boldsymbol{x} = \boldsymbol{b}$ ikke er et underrum af vektorrummet $\mathbb{F}^n$, hvis $\boldsymbol{b} \neq \boldsymbol{0}$. For da er $A\boldsymbol{0} \neq \boldsymbol{b}$, så (1) i sætning 4.1.4 er ikke opfyldt.

Bemærkning 4.1.7 Eksempel 4.1.5 er et specialtilfælde af lemma 4.1.6. For hvis A er 1×2-matricen $\begin{pmatrix} 1 & -1 \end{pmatrix}$, så er nulrummet

$$N_A = \{\boldsymbol{x} \in \mathbb{R}^2 \mid A\boldsymbol{x} = \boldsymbol{0}\} = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 \mid \begin{pmatrix} 1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0 \right\} \subset \mathbb{R}^2$$

præcis underrummet $V' \subset \mathbb{R}^2$ i eksempel 4.1.5. Tilsvarende hvis $B = \begin{pmatrix} 0 & 1 \end{pmatrix}$ og $\boldsymbol{b} = 1$, så er løsningsmængden til det uhomogene ligningssystem $B\boldsymbol{x} = \boldsymbol{b}$ lig med

$$\{\boldsymbol{x} \in \mathbb{R}^2 \mid B\boldsymbol{x} = \boldsymbol{b}\} = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 \mid \begin{pmatrix} 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 1 \right\} \subset \mathbb{R}^2$$

netop delmængden $S \subset \mathbb{R}^2$ i eksempel 4.1.5.

Vi har allerede defineret og studeret lineære afbildninger $f : \mathbb{F}^n \to \mathbb{F}^m$ i sektion 2.3, og vi udvider nu dette begreb til afbildninger mellem generelle vektorrum.

Definition 4.1.8 Lad $\mathbb{F}$ være et legeme og lad V og W være $\mathbb{F}$-vektorrum. En afbildning $f : V \to W$ er *lineær*, hvis den opfylder følgende betingelser:

(L1) For alle $\boldsymbol{x}, \boldsymbol{y} \in V$ gælder $f(\boldsymbol{x} + \boldsymbol{y}) = f(\boldsymbol{x}) + f(\boldsymbol{y})$.

(L2) For alle $\boldsymbol{x} \in V$ og $a \in \mathbb{F}$ gælder $f(\boldsymbol{x} \cdot a) = f(\boldsymbol{x}) \cdot a$.

Vi understreger, at en lineær afbildning $f : V \to W$ mellem to vektorrum altid afbilder nulvektoren i V til nulvektoren i W. For det følger fra (L1), at

$$f(\boldsymbol{0}) + f(\boldsymbol{0}) = f(\boldsymbol{0} + \boldsymbol{0}) = f(\boldsymbol{0}),$$

hvorfra udsagnet følger ved at trække $f(\boldsymbol{0})$ fra på begge sider af lighedstegnet. Vi vil også sige, at en afbildning $g : V \to W$ er *affin*, hvis afbildningen $f : V \to W$ defineret ved $f(\boldsymbol{x}) = g(\boldsymbol{x}) - g(\boldsymbol{0})$ er lineær. For eksempel er identitetsafbildningen $\mathrm{id} : V \to V$ og nulafbildningen $\boldsymbol{0} : V \to W$ begge linære. Den konstante afbildning $\boldsymbol{b} : V \to W$ med værdi $\boldsymbol{b} \in W$ er altid affin, og den er lineær hvis og kun hvis $\boldsymbol{b} = \boldsymbol{0}$.

Eksempel 4.1.9 Lad $\mathbb{F}$ være et legeme og lad $\mathbb{F}[t]$ være $\mathbb{F}$-vektorrummet af polynomier med koefficienter i $\mathbb{F}$. Da er afbildningen $D : \mathbb{F}[t] \to \mathbb{F}[t]$, der til et polynomium

$$p(t) = \sum_{n \geq 0} a_n t^n$$

tilordner det formelt afledte polynomium

$$p'(t) = \sum_{n \geq 1} n a_n t^{n-1},$$

en lineær afbildning.

Den følgende resultat giver en måde at associere en lineær afbildning til en tuple af vektorer i et vektorrum. Denne konstruktion kommer til at spille en særlig rolle i resten af kapitlet.

Lemma 4.1.10 *Lad V være et $\mathbb{F}$-vektorrum og lad $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ være en n-tuple af vektorer i V. Da er afbildningen $h : \mathbb{F}^n \to V$ defineret ved*

$$h(\boldsymbol{x}) = \boldsymbol{v}_1 x_1 + \boldsymbol{v}_2 x_2 + \cdots + \boldsymbol{v}_n x_n$$

en lineær afbildning.

Bevis Vi viser at h opfylder (L1) og (L2). For alle $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{F}^n$ har vi, at

$$\begin{aligned} h(\boldsymbol{x} + \boldsymbol{y}) &= \boldsymbol{v}_1(x_1 + y_1) + \cdots + \boldsymbol{v}_n(x_n + y_n) \\ &= (\boldsymbol{v}_1 x_1 + \cdots + \boldsymbol{v}_n x_n) + (\boldsymbol{v}_1 y_1 + \cdots + \boldsymbol{v}_n y_n) = h(\boldsymbol{x}) + h(\boldsymbol{y}), \end{aligned}$$

hvilket viser, at h opfylder (L1). Her følger den anden lighed ved gentagen anvendelse af (V3). Og for alle $\boldsymbol{x} \in \mathbb{F}^n$ og $a \in \mathbb{F}$ har vi, at

$$h(\boldsymbol{x}a) = \boldsymbol{v}_1(x_1 a) + \cdots + \boldsymbol{v}_n(x_n a) = (\boldsymbol{v}_1 x_1 + \cdots + \boldsymbol{v}_n x_n)a = h(\boldsymbol{x})a,$$

hvilket viser, at h opfylder (L2). Her følger den anden lighed ved gentagen anvendelse af (V1). □

Eksempel 4.1.11 Lad $V = C^0([0, 2\pi])$ være vektorrummet af kontinuere afbildninger $f : [0, 2\pi] \to \mathbb{R}$ fra eksempel 4.1.3(7). Da er $(\cos(x), \sin(x))$ et par af elementer i V, og lemma 4.1.10 viser da, at afbildningen $h : \mathbb{R}^2 \to C^0([0, 2\pi])$ defineret ved

$$h\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \cos(x)a_1 + \sin(x)a_2$$

er lineær.

Lemma 4.1.12 *Lad $\mathbb{F}$ være et legeme, lad U, V og W være $\mathbb{F}$-vektorrum, og lad $f : V \to W$ og $g : U \to V$ være to sammensættelige lineære afbildninger. Da er også den sammensatte afbildning $f \circ g : U \to W$ lineær.*

Bevis Vi viser, at $f \circ g : U \to W$ opfylder (L1) og (L2). For alle $\boldsymbol{x}, \boldsymbol{y} \in U$, gælder

$$(f \circ g)(\boldsymbol{x} + \boldsymbol{y}) = f(g(\boldsymbol{x} + \boldsymbol{y})) = f(g(\boldsymbol{x}) + g(\boldsymbol{y})) = f(g(\boldsymbol{x})) + f(g(\boldsymbol{y})) = (f \circ g)(\boldsymbol{x}) + (f \circ g)(\boldsymbol{y}),$$

hvilket viser, at $f \circ g : U \to W$ opfylder (L1). Tilsvarende gælger for alle $\boldsymbol{x} \in U$ og $a \in \mathbb{F}$, at

$$(f \circ g)(\boldsymbol{x} \cdot a) = f(g(\boldsymbol{x} \cdot a)) = f(g(\boldsymbol{x}) \cdot a) = f(g(\boldsymbol{x})) \cdot a,$$

hvilket viser, at $f \circ g : U \to W$ opfylder (L2). □

Definition 4.1.13 Lad $\mathbb{F}$ være et legeme. En lineær afbildning $f: V \to W$ mellem to $\mathbb{F}$-vektorrum er en *isomorfi*, hvis der findes en lineær afbildning $g: W \to V$, sådan at $f \circ g = \mathrm{id}_W$ og $g \circ f = \mathrm{id}_V$.

Det følgende resultat viser, at en lineær afbildning er en isomorfi, hvis og kun hvis den er bijektiv.

Sætning 4.1.14 *Lad $\mathbb{F}$ være et legeme og lad $f: V \to W$ være en lineær afbildningen mellem to $\mathbb{F}$-vektorrum. Da er følgende udsagn ækvivalente:*

(1) *Afbildningen $f: V \to W$ er en bijektion.*

(2) *Der findes en lineær afbildning $g: W \to V$, sådan at $f \circ g = \mathrm{id}_W$ og $g \circ f = \mathrm{id}_V$.*

Bevis Vi antager først (2) og viser (1). Da identitetsafbildningen er bijektiv, medfører identiteten $f \circ g = \mathrm{id}_W$, at f er surjektiv, mens identiteten $g \circ f = \mathrm{id}_V$ medfører, at f er injektiv. Altså er f bijektiv, hvilket viser (1). Vi antager dernæst (1) og viser (2). Da f er bijektiv, kan vi definere en afbildning $g: W \to V$ ved at fastsætte, at $g(\boldsymbol{y}) = \boldsymbol{x}$ hvis og kun hvis $\boldsymbol{y} = f(\boldsymbol{x})$. Per definition gælder da, at $f \circ g = \mathrm{id}_W$ og $g \circ f = \mathrm{id}_V$, men vi skal vise, at $g: W \to V$ er lineær. Da $f: V \to W$ er lineær, gælder for alle $\boldsymbol{y}, \boldsymbol{z} \in W$ og $a \in \mathbb{F}$, at

$$\begin{aligned} \boldsymbol{y} + \boldsymbol{z} &= f(g(\boldsymbol{y})) + f(g(\boldsymbol{z})) = f(g(\boldsymbol{y}) + g(\boldsymbol{z})), \\ \boldsymbol{y} \cdot a &= f(g(\boldsymbol{y})) \cdot a = f(g(\boldsymbol{y}) \cdot a), \end{aligned}$$

hvilket viser henholdsvis, at $g(\boldsymbol{y} + \boldsymbol{z}) = g(\boldsymbol{y}) + g(\boldsymbol{z})$ og $g(\boldsymbol{y} \cdot a) = g(\boldsymbol{y}) \cdot a$. Altså er g lineær, hvilket afslutter beviset for (2). □

Vi viser nu det følgende meget nyttige resultat.

Lemma 4.1.15 *Hvis $\mathbb{F}$ er et legeme, og hvis $f: V \to W$ er en lineær afbildning mellem to $\mathbb{F}$-vektorrum, da er de følgende udsagn ækvivalente:*

(1) *Afbildningen $f: V \to W$ er injektiv.*

(2) *Hvis $\boldsymbol{v} \in V$ og $f(\boldsymbol{v}) = \boldsymbol{0}$, så er $\boldsymbol{v} = \boldsymbol{0}$.*

Bevis Vi minder om, at $f: V \to W$ er injektiv, hvis der for alle $\boldsymbol{w} \in W$, at der højst findes et $\boldsymbol{v} \in V$ med $f(\boldsymbol{v}) = \boldsymbol{w}$. Så (1) medfører (2) per definition. Vi antager nu (2) og viser (1). Vi skal vise, at hvis $\boldsymbol{v}_1, \boldsymbol{v}_2 \in V$ og $f(\boldsymbol{v}_1) = f(\boldsymbol{v}_2)$, da er $\boldsymbol{v}_1 = \boldsymbol{v}_2$. Så lad $\boldsymbol{v} = \boldsymbol{v}_1 - \boldsymbol{v}_2$. Da er

$$\begin{aligned} f(\boldsymbol{v}) = f(\boldsymbol{v}_1 - \boldsymbol{v}_2) = f(\boldsymbol{v}_1 + \boldsymbol{v}_2 \cdot (-1)) &\overset{(L1)}{=} f(\boldsymbol{v}_1) + f(\boldsymbol{v}_2 \cdot (-1)) \\ &\overset{(L2)}{=} f(\boldsymbol{v}_1) + f(\boldsymbol{v}_2) \cdot (-1) = f(\boldsymbol{v}_1) - f(\boldsymbol{v}_2) = \boldsymbol{0}, \end{aligned}$$

og (2) medfører derfor, at $\boldsymbol{v} = \boldsymbol{0}$ og dermed $\boldsymbol{v}_1 = \boldsymbol{v}_2$ som ønsket. □

Eksempel 4.1.16 Vi lader $h\colon \mathbb{R}^2 \to C^0([0,2\pi])$ være den lineære afbildning fra eksempel 4.1.11, som er defineret ved $h(\boldsymbol{a}) = \cos(x)a_1 + \sin(x)a_2$. Vi ønsker at vise, at denne afbildning er injektiv, og ifølge lemma 4.1.15 er det tilstrækkeligt at vise, at hvis $\boldsymbol{a} \in \mathbb{R}^2$ og $h(\boldsymbol{a}) = \boldsymbol{0} \in C^0([0,2\pi])$, så er $\boldsymbol{a} = \boldsymbol{0}$. Vi husker på, at $\boldsymbol{0} \in C^0([0,2\pi])$ er nulafbildningen $\boldsymbol{0}\colon [0,2\pi] \to \mathbb{R}$, som er defineret ved $\boldsymbol{0}(x) = 0$ for alle $x \in [0,2\pi]$. Vi skal derfor vise, at hvis

$$\cos(x)a_1 + \sin(x)a_2 = 0,$$

for alle $x \in [0,2\pi]$, så er både $a_1 = 0$ og $a_2 = 0$. Vi kan betragte denne ligning som et lineært ligningssystem i de to variable a_1 og a_2 med uendeligt mange ligninger indiceret ved $x \in [0,2\pi]$. For $x = 0$ har vi ligningen

$$a_1 = \cos(0)a_1 + \sin(0)a_2 = 0,$$

hvilket viser, at $a_1 = 0$. Og for $x = \frac{\pi}{2}$ har vi ligningen

$$a_2 = \cos(\tfrac{\pi}{2})a_1 + \sin(\tfrac{\pi}{2})a_2 = 0,$$

hvilket viser, at $a_2 = 0$.

4.2 Basis for et vektorrum

Vi skal nu indføre tre helt centrale begreber for vektorrum. Disse begreber knytter sig ikke til enkelte vektorer i et vektorrum, men derimod til *familier* af vektorer, og de kræver derfor nogen tilvænning. De tre centrale egenskaber af en familie af vektorer i et vektorrum V, som vi nu vil definere, er, at den kan være *lineært uafhængig*, den kan *frembringe* V, eller den kan være en *basis for* V. For letheds skyld vil vi for det meste kun betragte *endelige* familier; vi betragter uendelige familier sidst i afsnittet.

Vi siger synonymt, at en afbildning $x\colon I \to X$ fra en endelig mængde I til en mængde X er en *endelig familie* af elementer i X indiceret ved I, og vi skriver $(x_i)_{i\in I}$, hvor $x_i = x(i)$. Hvis $I = \{1,2,\dots,n\}$, så skriver vi sommetider $(x_1,x_2,\dots,x_n)$ i stedet for $(x_i)_{i\in I}$ og siger, at denne familie er en n-tuple af elementer i X. For eksempel er standardbasen $(\boldsymbol{e}_1,\dots,\boldsymbol{e}_n)$ fra definition 2.2.3 en n-tuple af vektorer i $\mathbb{F}^n$. Og for enhver mængde X, findes der præcis én 0-tuple af elementer i X, nemlig den tomme familie $(\ )$. Vi understreger, at en familie af elementer ikke er det samme som en mængde af elementer. For eksempel er 1-tuplen (1) og 2-tuplen $(1,1)$ forskellige, mens mængderne $\{1\}$ og $\{1,1\}$ er ens. Vi betegner altid familier med runde paranteser, mens mængder betegnes med Tuborg paranteser.

Lad nu $(\boldsymbol{v}_i)_{i\in I}$ være en endelig familie af vektorer i et $\mathbb{F}$-vektorrum V. Hvis $(a_i)_{i\in I}$ er en familie af skalarer i $\mathbb{F}$ indiceret ved den samme mængde I, da kan vi danne summen

$$\boldsymbol{v} = \sum_{i\in I} \boldsymbol{v}_i a_i \in V,$$

og vi siger, at vektoren $\boldsymbol{v}$ er en *linearkombination* af familien $(\boldsymbol{v}_i)_{i\in I}$.

Eksempel 4.2.1 (1) Vi betragter de to vektorer

$$\boldsymbol{v}_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \qquad \text{og} \qquad \boldsymbol{v}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

i $\mathbb{R}^2$. En vektor $\boldsymbol{v} \in \mathbb{R}^2$ kan skrives som en linearkombination af familien $(\boldsymbol{v}_1)$ hvis og kun hvis, der findes $a \in \mathbb{R}$, sådan at

$$\boldsymbol{v} = \begin{pmatrix} 2a \\ a \end{pmatrix}.$$

I modsætning hertil, kan enhver vektor $\boldsymbol{v} \in \mathbb{R}^2$ skrives som en linearkombination af familien $(\boldsymbol{v}_1, \boldsymbol{v}_2)$, idet

$$\boldsymbol{v} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}(x_1 - x_2) + \begin{pmatrix} 1 \\ 1 \end{pmatrix}(-x_1 + 2x_2).$$

Denne måde at skrive $\boldsymbol{v}$ som en linearkombination af $(\boldsymbol{v}_1, \boldsymbol{v}_2)$ er endvidere entydig.

(2) Lad $(\boldsymbol{a}_1, \dots, \boldsymbol{a}_n)$ være familien af vektorer i $\mathbb{F}^m$, der består af søjlerne i $m \times n$-matricen $A \in M_{m,n}(\mathbb{F})$. En vektor $\boldsymbol{b} \in \mathbb{F}^m$ er da en linear kombination

$$\boldsymbol{b} = \boldsymbol{a}_1 x_1 + \boldsymbol{a}_2 x_2 + \cdots + \boldsymbol{a}_n x_n$$

af denne familie, hvis og kun hvis ligningssystemet $A\boldsymbol{x} = \boldsymbol{b}$ har en løsning $\boldsymbol{x} \in \mathbb{F}^n$.

(3) Lad $\mathbb{F}[t]_{\leq d}$ være vektorrummet af polynomier med koeffcienter i $\mathbb{F}$ af grad højst d fra eksempel 4.1.3 (6). Enhver vektor $p(t)$ i dette vektorrum er en linearkombination af familien af vektorer $(t^0, t^1, \dots, t^d)$. For hvis $p(t) = \sum_{0 \leq n \leq d} a_n t^n$, da er

$$p(t) = \sum_{0 \leq n \leq d} t^n \cdot a_n.$$

Vi bemærker, at familien af skalarer $(a_0, a_1, \dots, a_d)$ er entydigt bestemt af $p(t)$.

(4) Lad $C^0([0, 2\pi], \mathbb{C})$ være $\mathbb{C}$-vektorrummet af kontinuerte funktioner $f\colon [0, 2\pi] \to \mathbb{C}$. Vektoren $\exp(ix)$ i dette vektorrum er en linearkombination

$$\exp(ix) = \cos(x) \cdot 1 + \sin(x) \cdot i$$

af familien af vektorer $(\cos(x), \sin(x))$.

Lemma 4.2.2 *Lad $\mathbb{F}$ være et legeme, lad V være et $\mathbb{F}$-vektorum, og lad $(\boldsymbol{v}_i)_{i \in I}$ være en endelig familie af vektorer i V. Delmængden $W = \{\sum_{i \in I} \boldsymbol{v}_i a_i \mid a_i \in \mathbb{F}\} \subset V$, der består af alle linearkombinationer af familien $(\boldsymbol{v}_i)_{i \in I}$, er et underrum af V.*

Bevis Vi viser, at $W \subset V$ opfylder (1)–(3) i sætning 4.1.4. Vi bemærker, at $\boldsymbol{0} = \sum_{i \in I} \boldsymbol{v}_i \cdot 0$ er i W, hvilket viser (1). Og vi bemærker, at hvis $\boldsymbol{v} = \sum_{i \in I} \boldsymbol{v}_i a_i$ og $\boldsymbol{w} = \sum_{i \in I} \boldsymbol{v}_i b_i$ er i W, da er $\boldsymbol{v} + \boldsymbol{w} = \sum_{i \in I} \boldsymbol{v}_i a_i + \sum_{i \in I} \boldsymbol{v}_i b_i = \sum_{i \in I} \boldsymbol{v}_i (a_i + b_i)$ også i W, hvilket viser (2). Endelig, hvis $\boldsymbol{v} = \sum_{i \in I} \boldsymbol{v}_i a_i$ er i W og $a \in \mathbb{F}$, så er $\boldsymbol{x}a = (\sum_{i \in I} \boldsymbol{v}_i a_i) a = \sum_{i \in I} \boldsymbol{v}_i (a_i a)$ også i W, hvilket viser (3). $\square$

Underrummet $W \subset V$ i lemma 4.2.2 kaldes også for *spannet* af familien $(\boldsymbol{v}_i)_{i \in I}$ og betegnes $\operatorname{span}((bv_i)_{i \in I})$.

De følgende begreber er nogle af de vigtigste matematiske begreber overhovedet.

Definition 4.2.3 Lad $\mathbb{F}$ være et legeme, lad V være et $\mathbb{F}$-vektorrum og lad $(\boldsymbol{v}_i)_{i\in I}$ være en endelig familie af vektorer i V.

(1) Familien af vektorer $(\boldsymbol{v}_i)_{i\in I}$ siges at være *lineært uafhængig*, hvis den eneste familie af skalarer $(a_i)_{i\in I}$, for hvilken der gælder, at

$$\sum_{i\in I} \boldsymbol{v}_i a_i = \boldsymbol{0},$$

er nulfamilien $(a_i)_{i\in I}$, hvor $a_i = 0$ for alle $i \in I$. Ellers siges familien $(\boldsymbol{v}_i)_{i\in I}$ at være *lineært afhængig*.

(2) Familien af vektorer $(\boldsymbol{v}_i)_{i\in I}$ siges at *frembringe* eller udspænde V, hvis enhver vektor $\boldsymbol{v} \in V$ kan skrives som en linearkombination

$$\boldsymbol{v} = \sum_{i\in I} \boldsymbol{v}_i a_i$$

af familien $(\boldsymbol{v}_i)_{i\in I}$.

(3) Familien af vektorer $(\boldsymbol{v}_i)_{i\in I}$ siges at være en *basis* for V, hvis den både er lineært uafhængig og frembringer V.

Eksempel 4.2.4 (1) Vi betragter de tre vektorer

$$\boldsymbol{v}_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \qquad \boldsymbol{v}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad \text{og} \qquad \boldsymbol{v}_3 = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$$

i $\mathbb{R}^2$; vi har betragtet de første to i eksempel 4.2.1. Familien $(\boldsymbol{v}_1)$ er lineært uafhængig, fordi $\boldsymbol{v}_1 \cdot a_1 = \boldsymbol{0}$ medfører, at $a_1 = 0$. Den frembringer imidlertid ikke $\mathbb{R}^2$, da vektoren $\boldsymbol{v}_2$ for eksempel ikke kan skrives som en linearkombination af $(\boldsymbol{v}_1)$. Familien $(\boldsymbol{v}_1, \boldsymbol{v}_2)$ er også lineært uafhængig, idet $\boldsymbol{v}_1 a_1 + \boldsymbol{v}_2 a_2 = \boldsymbol{0}$, hvis og kun hvis

$$\begin{aligned} 2a_1 + a_2 &= 0 \\ a_1 + a_2 &= 0, \end{aligned}$$

og dette ligningssystem har kun den løsning, hvor både $a_1 = 0$ og $a_2 = 0$. Vi så allerede i eksempel 4.2.1, at $(\boldsymbol{v}_1, \boldsymbol{v}_2)$ frembringer $\mathbb{R}^2$, så dermed er $(\boldsymbol{v}_1, \boldsymbol{v}_2)$ en basis for $\mathbb{R}^2$. Familien $(\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3)$ frembringer også $\mathbb{R}^2$, fordi delfamilien $(\boldsymbol{v}_1, \boldsymbol{v}_2)$ gør det, men den er ikke lineært uafhængig. For ligningssystemet $\boldsymbol{v}_1 a_1 + \boldsymbol{v}_2 a_2 + \boldsymbol{v}_3 a_3 = \boldsymbol{0}$ har løsninger, hvor a_1, a_2 og a_3 ikke alle er nul, for eksempel $a_1 = 2$, $a_2 = -5$ og $a_3 = 1$.

(2) I ethvert vektorrum V er den tomme familie ( ) lineært uafhængig, fordi betingelsen i definition 4.2.3 (1) trivielt er opfyldt. Den tomme familie er en basis, hvis og kun hvis $V = \{\mathbf{0}\}$. For den tomme sum er per definition lig med $\mathbf{0}$.
(3) Standardbasen $(\boldsymbol{e}_1, \dots, \boldsymbol{e}_m)$ fra definition 2.2.3 er en basis for $\mathbb{F}^m$.
(4) En endelig familie af vektorer $(\boldsymbol{v}_i)_{i \in I}$, for hvilken der findes $h \in I$, sådan at $\boldsymbol{v}_h = 0$, er lineært afhængig. For lad $(a_i)_{i \in I}$ være familien af skalarer, hvor $a_h = 1$ og $a_i = 0$ for $i \neq h$. Da er $\sum_{i \in I} \boldsymbol{v}_i a_i = \mathbf{0}$, selvom $(a_i)_{i \in I}$ ikke er nulfamilien.

Eksempel 4.2.5 Vi betragter det homogene ligningssystem $A\boldsymbol{x} = \mathbf{0}$, hvor

$$A = \begin{pmatrix} 0 & 0 & 2 & -1 & 8 \\ 1 & -2 & 3 & 2 & 1 \\ 3 & -6 & 10 & 6 & 5 \end{pmatrix}.$$

Vi har tidligere i eksempel 1.2.11 udregnet, at matricen

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

er den entydigt bestemte matrix på reduceret echelon form, der er rækkeækvivalent med A. Løsningsmængden til ligningsystemet $A\boldsymbol{x} = \mathbf{0}$ er dermed

$$N_A = \left\{ \begin{pmatrix} 2t_1 & -3t_2 \\ t_1 & \\ & -2t_2 \\ & 4t_2 \\ & t_2 \end{pmatrix} \;\middle|\; t_1, t_2 \in \mathbb{R} \right\}.$$

Med andre ord kan enhver løsning til ligningen $A\boldsymbol{x} = \mathbf{0}$ skrives entydigt på formen

$$\boldsymbol{x} = \boldsymbol{v}_1 t_1 + \boldsymbol{v}_2 t_2,$$

hvor $t_1, t_2 \in \mathbb{F}$ og hvor $\boldsymbol{v}_1, \boldsymbol{v}_2 \in N_A \subset \mathbb{F}^5$ er vektorerne

$$\boldsymbol{v}_1 = \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \qquad \boldsymbol{v}_2 = \begin{pmatrix} -3 \\ 0 \\ -2 \\ 4 \\ 1 \end{pmatrix}.$$

Vi har således udregnet, at familien $(\boldsymbol{v}_1, \boldsymbol{v}_2)$ udgør en basis for nulrummet $N_A \subset \mathbb{F}^5$.

Bemærkning 4.2.6 Lad $(\boldsymbol{a}_1, \dots, \boldsymbol{a}_n)$ være familien af vektorer i $\mathbb{F}^m$, der består af søjlerne i en $m \times n$-matrix $A \in M_{m,n}(\mathbb{F})$.

(1) Familien $(\boldsymbol{a}_1,\dots,\boldsymbol{a}_n)$ er lineært uafhængig, hvis og kun hvis $\boldsymbol{x}=\boldsymbol{0}$ er den eneste løsning til ligningssystemet $A\boldsymbol{x}=\boldsymbol{0}$, idet enhver sådan løsning netop udtrykker $\boldsymbol{0}$ som en linearkombination $\boldsymbol{a}_1x_1+\cdots+\boldsymbol{a}_nx_n=\boldsymbol{0}$ af $(\boldsymbol{a}_1,\dots,\boldsymbol{a}_n)$.
(2) Familien $(\boldsymbol{a}_1,\dots,\boldsymbol{a}_n)$ frembringer $\mathbb{F}^m$, hvis og kun $A\boldsymbol{x}=\boldsymbol{b}$ har mindst en løsning for alle $\boldsymbol{b}\in\mathbb{F}^m$, igen fordi en sådan løsning præcis udtrykker $\boldsymbol{b}$ som en linearkombination $\boldsymbol{a}_1x_1+\cdots+\boldsymbol{a}_nx_n=\boldsymbol{b}$ af $(\boldsymbol{a}_1,\dots,\boldsymbol{a}_n)$.
(3) Familien $(\boldsymbol{a}_1,\dots,\boldsymbol{a}_n)$ er en basis for $\mathbb{F}^m$, hvis og kun hvis $A\boldsymbol{x}=\boldsymbol{b}$ har præcis én løsning for alle $\boldsymbol{b}\in\mathbb{F}^m$, og dette holder hvis og kun hvis A er invertibel.

Eksempel 4.2.7 Vi udregende i eksempel 3.2.8, at matricen

$$A=\begin{pmatrix}5&-1&1\\0&2&-3\\2&6&4\end{pmatrix}\in M_3(\mathbb{R})$$

har determinant $\det(A)=132\neq 0$. Da $\det(A)$ er invertibel i $\mathbb{R}$ er A derfor invertibel ifølge sætning 3.4.1. Derfor viser bemærkning 4.2.6 altså, at familien

$$\left(\boldsymbol{a}_1=\begin{pmatrix}5\\0\\2\end{pmatrix},\ \boldsymbol{a}_2=\begin{pmatrix}-1\\2\\6\end{pmatrix},\ \boldsymbol{a}_3=\begin{pmatrix}1\\-3\\4\end{pmatrix}\right)$$

er en basis for $\mathbb{R}^3$.

Vi viste i lemma 4.1.10, at en n-tuple $(\boldsymbol{v}_1,\dots,\boldsymbol{v}_n)$ af vektorer i et $\mathbb{F}$-vektorrum V giver anledning til en lineær afbildning $h\colon\mathbb{F}^n\to V$. Vi viser nu, at denne afbildning præcis er injektiv, surjektiv eller bijektiv eftersom n-tuplen $(\boldsymbol{v}_1,\dots,\boldsymbol{v}_n)$ er lineært uafhængig, frembringer V eller er en basis for V.

Sætning 4.2.8 *Lad $\mathbb{F}$ være et legeme, og lad V være et $\mathbb{F}$-vektorrum. Lad $(\boldsymbol{v}_1,\dots,\boldsymbol{v}_n)$ være en n-tuple af vektorer i V, og lad $h\colon\mathbb{F}^n\to V$ være afbildningen defineret ved*

$$h(\boldsymbol{x})=\boldsymbol{v}_1x_1+\cdots+\boldsymbol{v}_nx_n.$$

Da gælder følgende:

(1) *Familien $(\boldsymbol{v}_1,\dots,\boldsymbol{v}_n)$ er lineært uafhængig, hvis og kun hvis $h\colon\mathbb{F}^n\to V$ er injektiv.*

(2) *Familien $(\boldsymbol{v}_1,\dots,\boldsymbol{v}_n)$ frembringer V, hvis og kun hvis $h\colon\mathbb{F}^n\to V$ er surjektiv.*

(3) *Familien $(\boldsymbol{v}_1,\dots,\boldsymbol{v}_n)$ er en basis for V, hvis og kun hvis $h\colon\mathbb{F}^n\to V$ er bijektiv.*

Bevis Per definition er familien $(\boldsymbol{v}_1,\dots,\boldsymbol{v}_n)$ linært uafhængig, hvis og kun hvis der for alle $\boldsymbol{x} \in \mathbb{F}^n$ gælder, at hvis $h(\boldsymbol{x}) = \boldsymbol{0}$, så er $\boldsymbol{x} = \boldsymbol{0}$. Og ifølge lemma 4.1.15 er dette tilfældet, hvis og kun hvis $h\colon \mathbb{F}^n \to V$ er injektiv. Dette viser (1).

Familien $(\boldsymbol{v}_1,\dots,\boldsymbol{v}_n)$ frembringer per definition V, hvis og kun hvis alle $\boldsymbol{v} \in V$ kan skrives som en linearkombination $\boldsymbol{v} = \boldsymbol{v}_1x_1 + \dots + \boldsymbol{v}_nx_n$. Dette betyder præcis, at $h\colon \mathbb{F}^n \to V$ er surjektiv, hvorfor (2) følger.

Endelig er $(\boldsymbol{v}_1,\dots,\boldsymbol{v}_n)$ en basis, hvis og kun hvis den både er lineært uafhængig og frembringer V, og vi har netop vist, at dette er tilfældet, hvis og kun hvis afbildningen $h\colon \mathbb{F}^n \to V$ både er injektiv og surjektiv. Men en afbildning er både injektiv og surjektiv, hvis og kun hvis den er bijektiv. Dette viser (3). $\square$

Hvis $(\boldsymbol{v}_1,\dots,\boldsymbol{v}_n)$ er en basis for V, da er afbildningen $h\colon \mathbb{F}^n \to V$ i sætning 4.2.8 altså en lineær isomorfi. Vi skal i næste afsnit se, at dette tillader os, at oversætte alle lineær algebra spørgsmål, der angår V til tilsvarende spørgsmål, der angår $\mathbb{F}^n$. Det er dog vigtigt at huske på, at denne oversættelse *afhænger af valget af basis* for V.

I det tilfældet hvor $V = \mathbb{F}^m$ giver sætningen følgende resultat, som også kan ses som en omformulering af bemærkning 4.2.6.

Korollar 4.2.9 *Lad A være en $m \times n$-matrix af rang $0 \le r \le \min\{m,n\}$ med indgange i et legeme $\mathbb{F}$, og lad $(\boldsymbol{a}_1,\dots,\boldsymbol{a}_n)$ være familien af søjlevektorer i A.*

(1) *Familien $(\boldsymbol{a}_1,\dots,\boldsymbol{a}_n)$ er lineært uafhængig, hvis og kun hvis $r = n$.*

(2) *Familien $(\boldsymbol{a}_1,\dots,\boldsymbol{a}_n)$ frembringer $\mathbb{F}^m$, hvis og kun hvis $r = m$.*

(3) *Familien $(\boldsymbol{a}_1,\dots,\boldsymbol{a}_n)$ er en basis for $\mathbb{F}^m$, hvis og kun hvis $r = m = n$.*

Specielt er familien $(\boldsymbol{a}_1,\dots,\boldsymbol{a}_n)$ en basis hvis og kun hvis A er invertibel.

Bevis Vi bemærker, at afbildningen $h\colon \mathbb{F}^n \to \mathbb{F}^m$ i sætning 4.2.8, der hører til familien $(\boldsymbol{a}_1,\dots,\boldsymbol{a}_n)$ af vektorer i $\mathbb{F}^n$, er givet ved

$$h(\boldsymbol{x}) = \boldsymbol{a}_1x_1 + \dots + \boldsymbol{a}_nx_n = A\boldsymbol{x}.$$

Påstanden følger derfor fra sætning 4.2.8 og sætning 2.5.9. $\square$

Eksempel 4.2.10 Vi betragter igen matricen

$$A = \begin{pmatrix} 0 & 0 & 2 & -1 & 8 \\ 1 & -2 & 3 & 2 & 1 \\ 3 & -6 & 10 & 6 & 5 \end{pmatrix}$$

fra eksempel 4.2.5. Den entydigt bestemte matrix A', der er rækkeævivalent til A og på reduceret echelon form, udregnes til at være

$$A' = \begin{pmatrix} 1 & -2 & 0 & 0 & 3 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & -4 \end{pmatrix}.$$

Matricen A' har tre ledene indgange, og derfor er $\operatorname{rank}(A) = 3$. Matricen A er en 3×5-matrix, så $m = 3$ og $n = 5$ i dette eksempel. Idet $\operatorname{rank}(A) = 3 = m$, så siger korollar 4.2.9 at søjlerne i A frembringer $\mathbb{R}^3$.

Vi skal nu vise, at et vektorrum, der er frembragt af en endelig familie af vektorer, har en basis. Vi viser et mere præcist resultat, som er mere anvendeligt. Generelt, hvis $(x_i)_{i \in I}$ er en familie af elementer i en mængde X, og hvis $J \subset I$ er en delmængde, så siger vi, at familien $(x_i)_{i \in J}$ er en *delfamilie* af familien $(x_i)_{i \in I}$. Specielt er alle delfamilier af en n-tuple $(x_1, \dots, x_n)$ af formen $(x_{i_1}, \dots, x_{i_p})$, hvor $0 \le p \le n$ og $1 \le i_1 < \cdots < i_p \le n$. Den følgende sætning er den lineære algebras hovedsætning.

Sætning 4.2.11 *Lad $\mathbb{F}$ være et legeme og lad V være et $\mathbb{F}$-vektorrum. Hvis $(\boldsymbol{v}_i)_{i \in I}$ er en endelig familie af vektorer i V, som frembringer V, og hvis $(\boldsymbol{v}_i)_{i \in K}$ er en lineært uafhængig delfamilie deraf, da findes en mængde J med $K \subset J \subset I$, sådan at delfamilien $(\boldsymbol{v}_i)_{i \in J}$ er en basis for V.*

Bevis Vi ønsker at finde $K \subset J \subset I$, sådan at $(\boldsymbol{v}_i)_{i \in J}$ er en basis for V. Vi ved, at der findes $K \subset J \subset I$, sådan at $(\boldsymbol{v}_i)_{i \in J}$ er lineært uafhængig, for det er rigtigt for $J = K$. Vi vælger derfor et $K \subset J \subset I$, sådan at $(\boldsymbol{v}_i)_{i \in J}$ er lineært uafhængig, og sådan at J er maksimal i den forstand, at hvis $J \subset S \subset I$ og $(\boldsymbol{v}_i)_{i \in S}$, da er $S = J$. Vi vil nu vise, at familien $(\boldsymbol{v}_i)_{i \in J}$ også frembringer V.

Så vi antager modsætningsvis, at dette ikke er tilfældet og viser, at det strider mod maksimaliteten af J. Ifølge vores antagelse findes der en vektor $\boldsymbol{v} \in V$, som ikke er en linearkombination af $(\boldsymbol{v}_i)_{i \in J}$, og da familien $(\boldsymbol{v}_i)_{i \in I}$ frembringer V, kan vi yderligere antage, at $\boldsymbol{v} = \boldsymbol{v}_h$, hvor $h \in I$ og $h \notin J$. Vi lader $S = J \cup \{h\} \subset I$ og viser, at familien $(\boldsymbol{v}_i)_{i \in S}$ er lineært uafhængig. Så lad

$$\sum_{i \in S} \boldsymbol{v}_i a_i = \boldsymbol{0}$$

være en linearkombination, der er lig med nulvektoren. Hvis $a_h \neq 0$, da er

$$\boldsymbol{v}_h = \sum_{i \in J} \boldsymbol{v}_i \cdot (-a_i a_h^{-1})$$

en linearkombination af $(\boldsymbol{v}_i)_{i \in J}$. Men da vi har netop antaget, at $\boldsymbol{v}_h$ ikke er en linearkombination af $(\boldsymbol{v}_i)_{i \in J}$, konkluderer vi, at $a_h = 0$. Deraf følger, at

$$\sum_{i \in J} \boldsymbol{v}_i a_i = \sum_{i \in S} \boldsymbol{v}_i a_i = \boldsymbol{0},$$

og da $(\boldsymbol{v}_i)_{i \in J}$ er lineært uafhængig, følger det, at $a_i = 0$ for alle $i \in J$. Vi har altså hermed vist, at $a_i = 0$ for alle $i \in S$, hvilket viser, at $(\boldsymbol{v}_i)_{i \in J'}$ er lineært uafhængig. Men $S \neq J$, så dette strider mod maksimaliteten af J, og vi konkluderer derfor, at vores antagelse, at familien $(\boldsymbol{v}_i)_{i \in J}$ ikke frembringer V er forkert. Så familien $(\boldsymbol{v}_i)_{i \in J}$ frembringer V og er samtidigt lineært uafhængig og derfor en basis for V. Dette viser sætningen. $\square$

Vi siger, at et vektorrum V er *endeligt frembragt*, hvis der findes en endelig familie af vektorer i V, som frembringer V. Sætning 4.2.11 viser specielt, at ethvert endeligt frembragt vektorrum V har en endelig basis. For lad $(\boldsymbol{v}_i)_{i\in I}$ være en endelig familie af vektorer i V, der frembringer V. Da den tomme delfamilie $(\ )$ er lineært uafhængig, viser sætning 4.2.11, at der findes $\emptyset \subset J \subset I$, sådan at $(\boldsymbol{v}_i)_{i\in J}$ er en basis for V.

Eksempel 4.2.12 Vi betragter igen familien $(\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3)$ af vektorer i $\mathbb{R}^2$, hvor

$$\boldsymbol{v}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \boldsymbol{v}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad \text{og} \qquad \boldsymbol{v}_3 = \begin{pmatrix} 1 \\ 3 \end{pmatrix}.$$

Som vi bemærkede i eksempel 4.2.4 (1), frembringer denne familie $\mathbb{R}^2$, men den er ikke lineært uafhængig. De tre delfamilier $(\boldsymbol{v}_1, \boldsymbol{v}_2)$, $(\boldsymbol{v}_1, \boldsymbol{v}_3)$ og $(\boldsymbol{v}_2, \boldsymbol{v}_3)$ er alle lineært uafhængige og frembringer $\mathbb{R}^2$, og de udgør dermed alle baser for $\mathbb{R}^2$. De første to af disse indeholder den lineært uafhængige familie $(\boldsymbol{v}_1)$ som en delfamilie, mens dette ikke er tilfældet for den tredie.

Sætning 4.2.11 siger specielt, at enhver endelig familie af vektorer, der frembringer et vektorrum, har en delfamilie, der er en basis for dette vektorrum. Sætningen siger dog *ikke*, at en sådan basis er entydig, og eksempel 4.2.12 viser da også, at dette ikke er tilfældet. Sætningen og dens bevis siger heller ikke noget om, *hvordan* man bærer sig ad med, at finde en sådan basis. Det vender vi tilbage til i afsnit 4.4.

Sætning 4.2.13 *Lad $\mathbb{F}$ være et legeme, og lad V være et $\mathbb{F}$-vektorrum, der er frembragt af en familie $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_m)$ af m vektorer. Hvis $(\boldsymbol{w}_1, \dots, \boldsymbol{w}_n)$ er en lineært uafhængig familie af n vektorer i V, så er $n \leq m$.*

Bevis Ifølge sætning 4.2.11 findes der en delfamilie $(\boldsymbol{v}_{i_1}, \dots, \boldsymbol{v}_{i_k})$ af $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_m)$, som er en basis for V. Derfor viser sætning 4.2.8, at afbildningen $h \colon \mathbb{F}^k \to V$ defineret ved

$$h(\boldsymbol{x}) = \boldsymbol{v}_{i_i} x_1 + \dots + \boldsymbol{v}_{i_k} x_k$$

er bijektiv. Samme sætning viser, at afbildningen $g \colon \mathbb{F}^n \to V$ defineret ved

$$g(\boldsymbol{y}) = \boldsymbol{w}_1 y_1 + \dots + \boldsymbol{w}_n y_n$$

er injektiv. Dermed er den sammensatte afbildning $h^{-1} \circ g \colon \mathbb{F}^n \to \mathbb{F}^k$ injektiv. Da denne afbildning også er lineær, viser korollar 2.5.10, at $n \leq k$. Da også $k \leq m$, følger det, at $n \leq m$ som ønsket. □

Eksempel 4.2.14 I eksempel 4.2.4 (1) viste vi, at familien $(\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3)$ af vektorer i $\mathbb{R}^2$ ikke er lineært uafhængig. Men dette følger også fra sætning 4.2.13, som viser, at en lineært uafhængig familie af vektorer $\mathbb{R}^2$ aldrig kan bestå af mere end 2 vektorer. For familien $(\boldsymbol{v}_1, \boldsymbol{v}_2)$ består af 2 vektorer og frembringer $\mathbb{R}^2$.

Sætning 4.2.15 *Lad $\mathbb{F}$ være et legeme og lad V være et endeligt frembragt $\mathbb{F}$-vektorrum. Hvis både $(\boldsymbol{v}_1,\ldots,\boldsymbol{v}_m)$ og $(\boldsymbol{w}_1,\ldots,\boldsymbol{w}_n)$ er baser for V, så er $m = n$.*

Bevis Da familien $(\boldsymbol{v}_1,\ldots,\boldsymbol{v}_m)$ frembringer V, og da familien $(\boldsymbol{w}_1,\ldots,\boldsymbol{w}_n)$ er lineært uafhængig, viser sætning 4.2.13, at $m \geq n$, og uligheden $m \leq n$ vises tilsvarende. □

Sætning 4.2.15 viser, at det følgende dimensionsbegreb er veldefineret.

Definition 4.2.16 Lad V være et endeligt frembragt vektorrum over et legeme $\mathbb{F}$, og lad $(\boldsymbol{v}_1,\ldots,\boldsymbol{v}_d)$ være en basis for V. Da kaldes det naturlige tal

$$\dim_{\mathbb{F}}(V) = d$$

for *dimensionen* af V.

Hvis $\mathbb{F}$ er underforstået, så skriver vi $\dim(V)$ i stedet for $\dim_{\mathbb{F}}(V)$.

Eksempel 4.2.17 (1) Hvis $\mathbb{F}$ er et vilkårligt legeme, da er

$$\dim_{\mathbb{F}}(\mathbb{F}^m) = m.$$

For standardbasen $(\boldsymbol{e}_1,\boldsymbol{e}_2,\ldots,\boldsymbol{e}_m)$ er en basis for $\mathbb{F}^m$.

(2) Underrummet $N_A \subset \mathbb{F}^5$, som vi betragtede i eksempel 4.2.5, har dimension

$$\dim_{\mathbb{F}}(N_A) = 2.$$

For familien $(\boldsymbol{v}_1,\boldsymbol{v}_2)$ fra eksempel 4.2.5 er en basis.

(3) Det reelle vektorrum $\mathbb{C}$, som vi betragtede i eksempel 4.1.3 (3), har dimension

$$\dim_{\mathbb{R}}(\mathbb{C}) = 2.$$

For $(1,i)$ er en basis, da ethvert $z \in \mathbb{C}$ kan skrives entydigt som $z = 1\cdot a + i\cdot b$ med $a,b \in \mathbb{R}$.

(4) Lad $\mathbb{F}[t]_{\leq d}$ være $\mathbb{F}$-vektorrummet af polynomier med koefficienter i $\mathbb{F}$ og af grad højst d, som vi betragtede i eksempel 4.1.3 (5). Det har dimension

$$\dim_{\mathbb{F}}(\mathbb{F}[t]_{\leq d}) = d+1.$$

For familien $(t^0,t^1,\ldots,t^d)$ er en basis, da ethvert polynomium $p(t)$ af grad højst d kan skrives entydigt som en linear kombination $p(t) = t^0\cdot a_0 + t^1\cdot a_1 + \cdots + t^d\cdot a_d$, og index mængden $\{0,1,\ldots,d\}$, som har $d+1$ elementer.

Vi ønsker nu at vise, at hvis V er et endeligt frembragt vektorrum, og hvis $W \subset V$ er et underrum, så er $\dim(W) \leq \dim(V)$. Det er dog ikke trivielt at vise, at vektorrummet W er endeligt frembragt, så vi viser først, at dette faktisk er tilfældet.

Sætning 4.2.18 *Lad $\mathbb{F}$ være et legeme, lad V være et endeligt frembragt $\mathbb{F}$-vektorrum, og lad $W \subset V$ være et underrum. Da er W endeligt frembragt, og* $\dim(W) \leq \dim(V)$.

Bevis Vi viser først, at W er endeligt frembragt. Så vi lader $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_m)$ være en familie af vektorer, der frembringer V, og antage, at W ikke er endeligt frembragt. Vi viser da, at der for alle $n \geq 0$ findes en lineært uafhængig familie $(\boldsymbol{w}_1, \dots, \boldsymbol{w}_n)$ af vektorer i W, hvilket giver en modstrid med sætning 4.2.13. Beviset for denne påstand er ved induktion på $n \geq 0$, og tilfældet $n = 0$ gælder, da den tomme familie er lineært uafhængig. Så vi antager, at påstanden er bevist for $n = p-1$ og viser den for $n = p$. Per induktion findes der en lineært uafhængige familie $(\boldsymbol{w}_1, \dots, \boldsymbol{w}_{p-1})$ af $p-1$ vektorer i W, og da W per antagelse ikke er endeligt frembragt, er underrummet $W' \subset W$ frembragt af denne familie ikke W. Vi kan derfor vælge en vektor $\boldsymbol{w}_p \in W \smallsetminus W'$, og vi viser nu, at familien $(\boldsymbol{w}_1, \dots, \boldsymbol{w}_{p-1}, \boldsymbol{w}_p)$ af p vektorer i W er lineært uafhængig. Så vi lader

$$\boldsymbol{w}_1 a_1 + \dots + \boldsymbol{w}_{p-1} a_{p-1} + \boldsymbol{w}_p a_p = \mathbf{0}$$

være en linearkombination af $(\boldsymbol{w}_1, \dots, \boldsymbol{w}_p)$, som er lig med nulvektoren, og skal da vise, at $a_i = 0$ for alle $1 \leq i \leq p$. Hvis $a_p \neq 0$, da er

$$\boldsymbol{w}_p = -(\boldsymbol{w}_1 a_1 + \dots + \boldsymbol{w}_{p-1} a_{p-1}) a_p^{-1},$$

hvilket strider mod, at $\boldsymbol{w}_p \notin W'$. Så $a_p = 0$, og derfor er

$$\boldsymbol{w}_1 a_1 + \dots + \boldsymbol{w}_{p-1} a_{p-1} = \mathbf{0},$$

hvorfra følger, at $a_i = 0$ for $1 \leq i \leq p-1$, da $(\boldsymbol{w}_1, \dots, \boldsymbol{w}_{p-1})$ er lineært uafhængig. Dette viser induktionsskridtet og dermed påstanden, at der for alle $n \geq 0$ findes en lineært uafhængig familie $(\boldsymbol{w}_1, \dots, \boldsymbol{w}_n)$ af vektorer i W. Denne familie er da også en lineært uafhængig familie af vektorer i V, hvilket er i modstrid med sætning 4.2.13. Vi slutter derfor, at vores antagelse, at W ikke er endeligt frembragt var forkert, hvilket viser, at W er endeligt frembragt som ønsket. Vi anvender endelig sætning 4.2.11 til at vælge baser $(\boldsymbol{w}_1, \dots, \boldsymbol{w}_n)$ for W og $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_m)$ for V. Familien $(\boldsymbol{w}_1, \dots, \boldsymbol{w}_n)$ er da specielt en lineær uafhængig familie af vektorer i V, mens familien $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_m)$ af vektorer i V frembringer V. Derfor viser sætning 4.2.13, at

$$\dim(W) = n \leq m = \dim(V)$$

som ønsket. $\square$

Sætning 4.2.19 *Lad $\mathbb{F}$ være et legeme, lad V være et endeligt frembragt $\mathbb{F}$-vektorrum, og lad $d = \dim_{\mathbb{F}}(V)$. Da gælder følgende:*

(1) *Enhver lineært uafhængig familie $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_m)$ af vektorer i V er en delfamilie af en basis $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_m, \boldsymbol{v}_{m+1}, \dots, \boldsymbol{v}_d)$ for V.*

(2) *Enhver familie $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ af vektorer i V, der frembringer V, har en delfamilie $(\boldsymbol{v}_{i_1}, \dots, \boldsymbol{v}_{i_d})$, der er en basis for V.*

Bevis For at bevise (1) lader vi $(\boldsymbol{v}_{m+1}, \dots, \boldsymbol{v}_p)$ være en endelig familie af vektorer i V, der frembringer V. Da frembringer familien $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_m, \boldsymbol{v}_{m+1}, \dots, \boldsymbol{v}_p)$ ligeledes V, og den indeholder den lineært uafhængige familie $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_m)$ som en delfamilie. Sætning 4.2.11 viser derfor, at der findes en delmængde

$$\{1,2,\dots,m\} \subset J \subset \{1,2,\dots,m,m+1,\dots,p\},$$

sådan at delfamilien $(\boldsymbol{v}_i)_{i \in J}$ er en basis for V, og ved om nødvendigt at reindicere, kan vi antage, at $J = \{1,2,\dots,m,m+1,\dots,d\}$. Dette viser (1).

Endelig følger (2) umiddelbart fra sætning 4.2.11, som viser, at der findes

$$\emptyset \subset \{i_1,\dots,i_d\} \subset \{1,2,\dots,n\},$$

sådan at delfamilien $(\boldsymbol{v}_{i_1}, \dots, \boldsymbol{v}_{i_d})$ er en basis for V. □

Bemærkning 4.2.20 Vi skitserer her, hvordan begreber og sætninger i dette afsnit udvides til vektorrum, der ikke antages at være endeligt frembragte.

Vi definerer en linearkombination af en familie af vektorer i et vektorrum som følger. Hvis $a = (a_i)_{i \in I}$ er en familie af skalarer, da definerer vi dens *support* til at være

$$\operatorname{supp}(a) = \{i \in I \mid a_i \neq 0\} \subset I,$$

og vi siger, at $a = (a_i)_{i \in I}$ har *endelig support*, hvis $\operatorname{supp}(a) \subset I$ er en endelig delmængde. Vi lader nu V være et vektorrum og lader $(\boldsymbol{v}_i)_{i \in I}$ være en familie af vektorer i V. Hvis $(a_i)_{i \in I}$ er en familie af skalarer, der er indiceret ved den samme mængde I og som har endelig support, da definerer vi

$$\sum_{i \in I} \boldsymbol{v}_i a_i = \sum_{i \in \operatorname{supp}(a)} \boldsymbol{v}_i a_i$$

og kalder denne sum for en *linearkombination* af familien $(\boldsymbol{v}_i)_{i \in I}$. Vi bemærker, at mens højresiden er en endelig sum og derfor meningsfuld, er venstresiden blot et symbol, der ved denne ligning defineres til at være synonymt med højresiden. Med denne definition af linearkombination kan definition 4.2.3 nu udvides til vilkårlige familier mutatis mutandis.

Den lineære algebras hovedsætning gælder generelt og siger, at hvis $(\boldsymbol{v}_i)_{i\in I}$ er en familie af vektorer i V, der frembringer V, og hvis $(\boldsymbol{v}_i)_{i\in K}$ er en lineært uafhængig delfamilie, da findes $K \subset J \subset I$, sådan at delfamilien $(\boldsymbol{v}_i)_{i\in J}$ er en basis for V. Man kan dog ikke bevise denne sætning uden at anvende Zorn's lemma.[1] Et vektorrum V har altid en familie af vektorer, der frembringer V, nemlig identitetsfamilien $(\boldsymbol{v})_{\boldsymbol{v}\in V}$, og den tomme familie er en lineært uafhængig delfamilie heraf, så dermed har V en basis, som kan være (vilkårligt) uendelig.

Man kan også vise, at to baser for det samme vektorrum nødvendigvis består af det samme "antal" vektorer. Mere præcist, hvis $(\boldsymbol{v}_i)_{i\in I}$ og $(\boldsymbol{w}_j)_{j\in J}$ er to baser for V, da har mængderne I og J samme kardinalitet[2] α, og man definerer da dimensionen af V til at være $\dim_{\mathbb{F}}(V) = \alpha$.

4.3 Matrix repræsentation af lineære afbildninger

Som Descartes har lært os, er det lettest at lave udregninger ved at bruge koordinater. Vi indfører nu koordinater for endelige frembragte vektorrum og matrix repræsentation for lineær afbildninger mellem endelige frembragte vektorrum. Det er dog meget vigtigt at huske, at begge begreber afhænger af valg af baser.

Til enhver n-tuple $(\boldsymbol{v}_1,\dots,\boldsymbol{v}_n)$ af vektorer i V har vi i lemma 4.1.10 tilordnet en lineær afbildning $h\colon \mathbb{F}^n \to V$, og vi har vist i sætning 4.2.8, at denne afbildning en isomorfi, hvis og kun hvis $(\boldsymbol{v}_1,\dots,\boldsymbol{v}_n)$ er en basis for V.

Definition 4.3.1 Lad V være et endeligt frembragt vektorrum over et legeme $\mathbb{F}$. Lad $(\boldsymbol{v}_1,\dots,\boldsymbol{v}_n)$ være en basis for V, og lad $h\colon \mathbb{F}^n \to V$ være afbildningen defineret ved

$$h(\boldsymbol{x}) = \boldsymbol{v}_1 x_1 + \boldsymbol{v}_2 x_2 + \cdots + \boldsymbol{v}_n x_n.$$

Givet $\boldsymbol{v} \in V$, da kaldes den entydigt bestemte søjlevektor $\boldsymbol{x} \in \mathbb{F}^n$, sådan at $h(\boldsymbol{x}) = \boldsymbol{v}$, for *koordinaterne af $\boldsymbol{v} \in V$ med hensyn til basen $(\boldsymbol{v}_1,\dots,\boldsymbol{v}_n)$.*

Eksempel 4.3.2 (1) Vi betragter igen basen

$$\left(\boldsymbol{v}_1 = \begin{pmatrix}2\\1\end{pmatrix},\ \boldsymbol{v}_2 = \begin{pmatrix}1\\1\end{pmatrix}\right)$$

[1] A. Blaas. *Existence of bases implies the axiom of choice*. Axiomatic Set Theory (Boulder, CO, 1983), 31–33, Contemp. Math., 31, Amer. Math. Soc., Providence, RI, 1984.

[2] Kenneth Kunen. Set Theory. Studies in Logic (London), 34. College Publications, London, 2011. Kardinaliteten af en mængde er defineret i Definition I.10.11.

for $V = \mathbb{R}^2$ fra eksempel 4.2.4 (1). Afbildningen $h\colon \mathbb{R}^2 \to V$ er i det tilfælde givet ved

$$h(\boldsymbol{x}) = \boldsymbol{v}_1 x_1 + \boldsymbol{v}_2 x_2 = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

så per definition er koordinaterne af vektoren

$$\boldsymbol{y} = \begin{pmatrix} 5 \\ 3 \end{pmatrix} \in V$$

med hensyn til basen $(\boldsymbol{v}_1, \boldsymbol{v}_2)$ givet ved den entydige løsning til ligningssystemet

$$\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 5 \\ 3 \end{pmatrix},$$

nemlig

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 5 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 5 \\ 3 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

(2) Standardbasen $(\boldsymbol{e}_1, \dots, \boldsymbol{e}_n)$ er en basis for $V = \mathbb{F}^n$. Afbildingen

$$h\colon \mathbb{F}^n \to V$$

er i dette tilfælde identitetsafbildningen, og koordinaterne af $\boldsymbol{x} \in \mathbb{F}^n$ med hensyn til standard basen er dermed $\boldsymbol{x}$ selv.

(3) Vi betragter vektorrummet $V = \mathbb{F}[t]_{\leq d}$ af polynomier med koefficienter i $\mathbb{F}$ og af grad højst d fra eksempel 4.1.3. Familien $(t^0, t^1, \dots, t^d)$ er en basis for V og den tilhørende afbildningen $h\colon \mathbb{F}^{d+1} \to V$ er givet ved

$$h(\boldsymbol{x}) = t^0 \cdot x_1 + t^1 \cdot x_2 + \cdots + t^d \cdot x_{d+1} = x_1 + x_2 t + \cdots + x_{d+1} t^d.$$

Dermed er koordinaterne for vektoren $p(t) = a_0 + a_1 t + \cdots + a_d t^d \in V$ med hensyn til basen $(t^0, t^1, \dots, t^d)$ for V givet ved

$$\boldsymbol{x} = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_d \end{pmatrix} \in \mathbb{F}^{d+1}.$$

Vi bemærker det forvirrende indexskift, som skyldes, at vi et dette afsnit kun tillader baser at være indiceret ved $I = \{1, 2, \dots, n\}$ og ikke ved for eksempel $I = \{0, 1, \dots, d\}$.

Vi skal nu vise, at givet to endeligt frembragte vektorrum samt baser for disse, da kan enhver lineær afbildning mellem disse vektorrum på entydig vis repræsenteres ved en matrix. Det er vigtigt at huske denne sætning helt præcist, da den danner grundlaget for alle beregninger, der har med lineære afbildninger at gøre.

Sætning 4.3.3 *Lad $\mathbb{F}$ være et legeme, og lad $f: V \to W$ være en lineær afbildning fra et n-dimensionalt $\mathbb{F}$-vektorrum til et m-dimensionalt $\mathbb{F}$-vektorrum. Lad $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ og $(\boldsymbol{w}_1, \dots, \boldsymbol{w}_m)$ være baser for henholdsvis V og W, og lad*

$$A = (a_{ij}) \in M_{m,n}(\mathbb{F})$$

være matricen, hvis j'te søjle er koordinaterne for $f(\boldsymbol{v}_j) \in W$ med hensyn til $(\boldsymbol{w}_1, \dots, \boldsymbol{w}_m)$. Hvis $\boldsymbol{v} \in V$ og $\boldsymbol{w} = f(\boldsymbol{v}) \in W$, og hvis $\boldsymbol{x} \in \mathbb{F}^n$ og $\boldsymbol{y} \in \mathbb{F}^m$ er koordinaterne for henholdsvis $\boldsymbol{v} \in V$ med hensyn til $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ og for $\boldsymbol{w} \in W$ med hensyn til $(\boldsymbol{w}_1, \dots, \boldsymbol{w}_m)$, da er

$$\boldsymbol{y} = A\boldsymbol{x}.$$

Matricen $A \in M_{m,n}(\mathbb{F})$ er endvidere entydigt bestemt ved denne egenskab.

Bevis Ifølge definition 4.3.1 kan vi skrive $\boldsymbol{v} \in V$ entydigt som

$$\boldsymbol{v} = \sum_{j=1}^{n} \boldsymbol{v}_j x_j,$$

hvor $\boldsymbol{x} \in \mathbb{F}^n$ er koordinaterne af $\boldsymbol{v}$ med hensyn til basen $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$. Vi anvender f på denne ligning og udnytter henholdsvis (L1) og (L2) til at omskrive højresiden:

$$f(\boldsymbol{v}) = f(\sum_{j=1}^{n} \boldsymbol{v}_j x_j) = \sum_{j=1}^{n} f(\boldsymbol{v}_j x_j) = \sum_{j=1}^{n} f(\boldsymbol{v}_j) x_j.$$

Ifølge definition 4.3.1, kan vi ligeledes skrive vektoren $f(\boldsymbol{v}_j) \in W$ entydigt som

$$f(\boldsymbol{v}_j) = \sum_{i=1}^{m} \boldsymbol{w}_i a_{ij},$$

hvor $\boldsymbol{a}_j \in \mathbb{F}^m$ er koordinaterne af $f(\boldsymbol{v}_j)$ med hensyn til basen $(\boldsymbol{w}_1, \dots, \boldsymbol{w}_m)$. Vi substituerer disse sidste ligninger i udtrykket for $\boldsymbol{w} = f(\boldsymbol{v})$ overfor, hvilket giver

$$\boldsymbol{w} = \sum_{j=1}^{n} (\sum_{i=1}^{m} \boldsymbol{w}_i a_{ij}) x_j = \sum_{i=1}^{m} \boldsymbol{w}_i (\sum_{j=1}^{n} a_{ij} x_j).$$

Da koordinaterne $\boldsymbol{y} \in \mathbb{F}^m$ af vektoren $\boldsymbol{w}$ med hensyn til basen $(\boldsymbol{w}_1, \dots, \boldsymbol{w}_m)$ er entydigt bestemt, aflæser vi derfor af denne ligning, at

$$y_i = \sum_{j=1}^{n} a_{ij} x_j$$

for alle $1 \leq i \leq m$. Dermed har vi vist, at $\boldsymbol{y} = A\boldsymbol{x}$ som ønsket. Endelig er matricen A entydigt bestemt, idet koordinaterne $\boldsymbol{a}_j \in \mathbb{F}^m$ for vektoren $f(\boldsymbol{v}_j) \in W$ med hensyn til basen $(\boldsymbol{w}_1, \dots, \boldsymbol{w}_m)$ er entydigt bestemte. $\square$

Sætning 4.3.3 giver en 1-1 korrespondance mellem lineære afbildninger $f\colon V \to W$ og matricer $A \in M_{m,n}(\mathbb{F})$. Denne 1-1 korrespondance afhænger af valget af baser $(\boldsymbol{v}_1,\dots,\boldsymbol{v}_n)$ for V og $(\boldsymbol{w}_1,\dots,\boldsymbol{w}_m)$ for W, og vi indfører derfor følgende terminologi.

Definition 4.3.4 Lad $\mathbb{F}$ være et legeme, og lad $f\colon V \to W$ være en lineær afbildning fra et n-dimensionalt $\mathbb{F}$-vektorrum til et m-dimensionalt $\mathbb{F}$-vektorrum. Lad $(\boldsymbol{v}_1,\dots,\boldsymbol{v}_n)$ og $(\boldsymbol{w}_1,\dots,\boldsymbol{w}_m)$ være baser for henholdsvis V og W. Da kaldes matricen

$$A = (a_{ij}) \in M_{m,n}(\mathbb{F}),$$

hvis j'te søjle er koordinaterne for $f(\boldsymbol{v}_j) \in W$ med hensyn til basen $(\boldsymbol{w}_1,\dots,\boldsymbol{w}_m)$ for W, for *matricen, der repræsenterer $f\colon V \to W$ med hensyn til baserne $(\boldsymbol{v}_1,\dots,\boldsymbol{v}_n)$ for V og $(\boldsymbol{w}_1,\dots,\boldsymbol{w}_m)$ for W.*

Eksempel 4.3.5 (1) Lad $\mathbb{F}$ være et legeme, og betragt vektorerne $\boldsymbol{v}_1, \boldsymbol{v}_2 \in \mathbb{F}^2$, hvor

$$\boldsymbol{v}_1 = \begin{pmatrix} 7 \\ 2 \end{pmatrix}, \qquad \boldsymbol{v}_2 = \begin{pmatrix} 3 \\ 1 \end{pmatrix}.$$

Familien $(\boldsymbol{v}_1, \boldsymbol{v}_2)$ er en basis for $\mathbb{F}^2$, da matricen $A = \begin{pmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 \end{pmatrix}$ er invertibel. (Det er den for eksempel, fordi $\det(A) = 1$.) Per definition er matricen $P \in M_2(\mathbb{F})$, der repræsenterer identitetsafbildningen $\mathrm{id}\colon \mathbb{F}^2 \to \mathbb{F}^2$ med hensyn til basen $(\boldsymbol{v}_1, \boldsymbol{v}_2)$ for domænet og standardbasen $(\boldsymbol{e}_1, \boldsymbol{e}_2)$ for codomænet givet ved

$$P = \begin{pmatrix} 7 & 3 \\ 2 & 1 \end{pmatrix}.$$

For søjlene i P er koordinaterne af henholdsvis $\boldsymbol{v}_1 = \mathrm{id}(\boldsymbol{v}_1)$ og $\boldsymbol{v}_2 = \mathrm{id}(\boldsymbol{v}_2)$ med hensyn til standardbasen $(\boldsymbol{e}_1, \boldsymbol{e}_2)$ for $\mathbb{F}^2$. Denne matrix er altså *ikke* identitetsmatricen.
(2) Vi betragter den lineære afbildning $D\colon \mathbb{F}[t]_{\leq 2} \to \mathbb{F}[t]_{\leq 1}$, der til et polynomium

$$p(t) = a_0 + a_1 t + a_2 t^2$$

tilordner det formelt afledte polynomium

$$p'(t) = a_1 + 2a_2 t.$$

Matricen $A \in M_{2,3}(\mathbb{F})$, der repræsenterer D med hensyn til basen (t^0, t^1, t^2) for domænet og (t^0, t^1) for codomænet, er da

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

idet

$$\begin{aligned} D(t^0) &= 0 = t^0 \cdot 0 + t^1 \cdot 0 \\ D(t^1) &= 1 = t^0 \cdot 1 + t^1 \cdot 0 \\ D(t^2) &= 2t = t^0 \cdot 0 + t^1 \cdot 2. \end{aligned}$$

Sætning 4.3.6 *Lad $\mathbb{F}$ være et legeme, lad U, V og W være $\mathbb{F}$-vektorrum af dimension henholdsvis p, n og m, og lad $(\boldsymbol{u}_1,\dots,\boldsymbol{u}_p)$, $(\boldsymbol{v}_1,\dots,\boldsymbol{v}_n)$ og $(\boldsymbol{w}_1,\dots,\boldsymbol{w}_m)$ være baser for henholdsvis U, V og W. Lad $f\colon V \to W$ og $g\colon U \to V$ være lineære afbildninger, og lad $A \in M_{m,n}(\mathbb{F})$, $B \in M_{n,p}(\mathbb{F})$ og $C \in M_{m,p}(\mathbb{F})$ være matricerne, der repræsenterer henholdsvis $f\colon V \to W$, $g\colon U \to V$ og $f \circ g\colon U \to W$ med hensyn til de givne baser. Da gælder*

$$C = AB.$$

Bevis For $\boldsymbol{u} \in U$ lader vi $\boldsymbol{x} \in \mathbb{F}^p$, $\boldsymbol{y} \in \mathbb{F}^n$ og $\boldsymbol{z} \in \mathbb{F}^m$ være koordinaterne for henholdsvis $\boldsymbol{u} \in U$ med hensyn til basen $(\boldsymbol{u}_1,\dots,\boldsymbol{u}_p)$, $\boldsymbol{v} = g(\boldsymbol{u}) \in V$ med hensyn til basen $(\boldsymbol{v}_1,\dots,\boldsymbol{v}_n)$, og $\boldsymbol{w} = f(g(\boldsymbol{u})) \in W$ med hensyn til basen $(\boldsymbol{w}_1,\dots,\boldsymbol{w}_m)$. Ifølge sætning 4.3.3 gælder da, at

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

og da matricen C, der repræsenterer $f \circ g\colon U \to W$ med hensyn til baserne $(\boldsymbol{u}_1,\dots,\boldsymbol{u}_p)$ for U og $(\boldsymbol{w}_1,\dots,\boldsymbol{w}_m)$ for W er entydigt bestemt, konkluderer vi, at $C = AB$ som ønsket. □

Bemærkning 4.3.7 Den følgende figur illustrerer sætning 4.3.5.

$$\begin{tikzcd} U \arrow[r, "g"] \arrow[rr, bend left, "f \circ g"] & V \arrow[r, "f"] & W \end{tikzcd}$$

$$\begin{tikzcd} (u_1,\dots,u_p) \arrow[r, dotted, "B"] \arrow[rr, dotted, bend right, "AB"'] & (v_1,\dots,v_n) \arrow[r, dotted, "A"] & (w_1,\dots,w_m) \end{tikzcd}$$

Vi understreger, at den rækkefølge, hvori de to afbildninger f og g sammensættes, er den samme som den rækkefølge, hvori de to matricer A og B, der repræsenterer dem med hensyn til de valgte baser, multipliceres.

Korollar 4.3.8 *Lad $\mathbb{F}$ være et legeme, lad $f\colon V \to W$ være en lineær afbildning fra et n-dimensionalt $\mathbb{F}$-vektorrum V til et m-dimensionalt $\mathbb{F}$-vektorrum W, og lad $A \in M_{m,n}(\mathbb{F})$ være matricen, der repræsenterer $f\colon V \to W$ med hensyn til baser $(\boldsymbol{v}_1,\dots,\boldsymbol{v}_n)$ for V og $(\boldsymbol{w}_1,\dots,\boldsymbol{w}_m)$ for W. Da er de følgende udsagn ækvivalente:*

(1) *Der findes en lineær afbildning $g\colon W \to V$, sådan at $f \circ g = \mathrm{id}_W$ og $g \circ f = \mathrm{id}_V$.*

(2) *Der findes en matrix $B \in M_{n,m}(\mathbb{F})$, sådan at $AB = I_m$ og $BA = I_n$.*

I givet fald repræsenterer matricen $B \in M_{n,m}(\mathbb{F})$ afbildningen $g\colon W \to V$ med hensyn til basen $(\boldsymbol{w}_1,\dots,\boldsymbol{w}_m)$ for W og $(\boldsymbol{v}_1,\dots,\boldsymbol{v}_n)$ for V.

Bevis Vi antager først (1) og viser (2). Hertil lader vi $B \in M_{n,m}(\mathbb{F})$ være matricen, der repræsenterer $g\colon W \to V$ med hensyn til baserne $(\boldsymbol{w}_1,\dots,\boldsymbol{w}_m)$ for W og $(\boldsymbol{v}_1,\dots,\boldsymbol{v}_n)$ for V. Sætning 4.3.6 viser da, at $AB \in M_m(\mathbb{F})$ repræsenterer $\mathrm{id}_W\colon W \to W$ med hensyn til den *samme* basis $(\boldsymbol{w}_1,\dots,\boldsymbol{w}_m)$ for både domænet og codomænet, hvilket viser, at $AB = I_m$. Vi ser tilsvarende, at $BA \in M_n(\mathbb{F})$ repræsenterer $\mathrm{id}_V\colon V \to V$ med hensyn til den *samme* basis $(\boldsymbol{v}_1,\dots,\boldsymbol{v}_n)$ for både domænet og codomænet, og derfor er også $BA = I_n$, hvilket viser (2).

Omvendt, hvis (2) gælder, da definerer vi $g\colon W \to V$ til at være afbildningen, der til $\boldsymbol{w} = \boldsymbol{w}_1 y_1 + \cdots + \boldsymbol{w}_m y_m \in W$ tilordner $\boldsymbol{v} = \boldsymbol{v}_1 x_1 + \cdots + \boldsymbol{v}_n x_n \in V$, hvor

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1m} \\ b_{21} & b_{22} & \cdots & b_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nm} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}$$

med $B = (b_{ij}) \in M_{n,m}(\mathbb{F})$ den givne matrix. Afbildningen $g\colon W \to V$ er da lineær, og ifølge sætning 4.3.6 er $f \circ g = \mathrm{id}_V$ og $g \circ f = \mathrm{id}_W$, hvilket viser (1). □

Bemærkning 4.3.9 Hvis de ækvivalente udsagn (1) og (2) i korollar 4.3.8 gælder, da er nødvendigvis $m = n$. Lad os vise det mere præcise udsagn, at hvis $(\boldsymbol{v}_1,\dots,\boldsymbol{v}_n)$ er en basis for V, da er $(f(\boldsymbol{v}_1),\dots,f(\boldsymbol{v}_n))$ en basis for W. Givet $\boldsymbol{w} \in W$, kan vi skrive $\boldsymbol{v} = g(\boldsymbol{w}) \in V$ som en linearkombination $\boldsymbol{v} = \boldsymbol{v}_1 a_1 + \cdots + \boldsymbol{v}_n a_n$, så

$$\boldsymbol{w} = f(\boldsymbol{v}) = f(\boldsymbol{v}_1 a_1 + \cdots + \boldsymbol{v}_n a_n) = f(\boldsymbol{v}_1)a_1 + \cdots + f(\boldsymbol{v}_n)a_n,$$

hvilket viser, at familien $(f(\boldsymbol{v}_1),\dots,f(\boldsymbol{v}_n))$ frembringer W. Og hvis en linearkombination $f(\boldsymbol{v}_1)a_1 + \cdots + f(\boldsymbol{v}_n)a_n$ er lig med nulvektoren, da er også

$$\boldsymbol{v}_1 a_1 + \cdots + \boldsymbol{v}_n a_n = g(f(\boldsymbol{v}_1))a_1 + g(f(\boldsymbol{v}_n))a_n = g(f(\boldsymbol{v}_1)a_1 + \cdots + f(\boldsymbol{v}_n)a_n)$$

lig med nulvektoren, og da $(\boldsymbol{v}_1,\dots,\boldsymbol{v}_n)$ er lineært uafhængig, er $(f(\boldsymbol{v}_1),\dots,f(\boldsymbol{v}_n))$ også lineært uafhængig. Specielt har V og W derfor samme dimension n.

Eksempel 4.3.10 Vi vender tilbage til eksempel 4.3.5, men vi ønsker nu at bestemme matricen $B \in M_2(\mathbb{F})$, der repræsenterer $\mathrm{id}\colon \mathbb{F}^2 \to \mathbb{F}^2$ med hensyn til standardbasen $(\boldsymbol{e}_1,\boldsymbol{e}_2)$ for domænet og basen $(\boldsymbol{v}_1,\boldsymbol{v}_2)$ for codomænet. Da identitetsafbildningen er sin egen inverse afbildning, viser korollar 4.3.8, at

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

hvor vi har brugt eksempel 3.4.3 til at udregne den inverse matrix.

Vi skal nu se, at matricen, der repræsenterer en lineær afbildning med hensyn til et valg af baser for dens domæne og codomæne, bestemmer matricen, der repræsenterer afbildningen med hensyn til et andet valg af baser for dens domæne og codomæne.

Sætning 4.3.11 *Lad $\mathbb{F}$ være det legeme, og lad $f\colon V \to W$ være en lineær afbildning fra et n-dimensionalt $\mathbb{F}$-vektorrum til et m-dimensionalt $\mathbb{F}$-vektorrum. Lad $A \in M_{m,n}(\mathbb{F})$ være matricen, der repræsenterer $f\colon V \to W$ med hensyn til baser $(\boldsymbol{v}_1,\dots,\boldsymbol{v}_n)$ for V og $(\boldsymbol{w}_1,\dots,\boldsymbol{w}_m)$ for W, og lad $B \in M_{m,n}(\mathbb{F})$ være matricen, der repræsenterer $f\colon V \to W$ med hensyn til baser $(\boldsymbol{v}'_1,\dots,\boldsymbol{v}'_n)$ for V og $(\boldsymbol{w}'_1,\dots,\boldsymbol{w}'_m)$ for W. Lad endvidere $P \in M_n(\mathbb{F})$ være matricen, der repræsenterer $\mathrm{id}_V\colon V \to V$ med hensyn til basen $(\boldsymbol{v}'_1,\dots,\boldsymbol{v}'_n)$ for domænet og basen $(\boldsymbol{v}_1,\dots,\boldsymbol{v}_n)$ for codomænet, og lad $Q \in M_m(\mathbb{F})$ være matricen, der repræsenterer $\mathrm{id}_W\colon W \to W$ med hensyn til basen $(\boldsymbol{w}'_1,\dots,\boldsymbol{w}'_m)$ for domænet og basen $(\boldsymbol{w}_1,\dots,\boldsymbol{w}_m)$ for codomænet. Da er*

$$B = Q^{-1}AP.$$

Bevis Vi lader C være matricen for $\mathrm{id}_W \circ f = f \circ \mathrm{id}_V\colon V \to W$ med hensyn til baserne $(\boldsymbol{v}'_1,\dots,\boldsymbol{v}'_n)$ for V og $(\boldsymbol{w}_1,\dots,\boldsymbol{w}_m)$ for W. Det følger da fra sætning 4.3.6, at

$$C = QB = AP,$$

med Q, B, A og P som i sætningen. Da identitetsafbildning $\mathrm{id}_W\colon W \to W$ er invertibel, viser korollar 4.3.8 at matricen Q er invertibel, så sætningen følger. $\square$

Bemærkning 4.3.12 Den følgende figur illustrerer sætning 4.3.11.

$$\begin{tikzcd}
(\boldsymbol{v}_1,\dots,\boldsymbol{v}_n) \arrow[rrr, dotted, bend left, "A"] & & & (\boldsymbol{w}_1,\dots,\boldsymbol{w}_m) \\
& V \arrow[r, "f"] & W & \\
& V \arrow[u, "\mathrm{id}_V"] \arrow[r, "f"] & W \arrow[u, "\mathrm{id}_W"] & \\
(\boldsymbol{v}'_1,\dots,\boldsymbol{v}'_n) \arrow[uuu, dotted, bend left, "P"] \arrow[rrr, dotted, bend right, "B"] & & & (\boldsymbol{w}'_1,\dots,\boldsymbol{w}'_m) \arrow[uuu, dotted, bend right, "Q"']
\end{tikzcd}$$

Vi ønsker at skrive den prikkede pil "B" som en sammensætning af de øvrige prikkede pile og bemærker, at dette kræver, at pilen "Q" vendes om. Pilen i den modsatte retning er da "Q^{-1}", og den findes, fordi Q er invertibel.

Eksempel 4.3.13 Lad $f\colon \mathbb{F}^3 \to \mathbb{F}^2$ være den lineære afbildning givet ved $f(\boldsymbol{x}) = A\boldsymbol{x}$ med

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

Matricen $A \in M_{2,3}(\mathbb{F})$ repræsenterer altså $f\colon \mathbb{F}^3 \to \mathbb{F}^2$ med hensyn til standardbaserne $(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$ for domænet og $(\boldsymbol{e}_1, \boldsymbol{e}_2)$ for codomænet. Vi ønsker nu at bestemme matricen $B \in M_{2,3}(\mathbb{F})$, der repræsenterer $f\colon \mathbb{F}^3 \to \mathbb{F}^2$ med hensyn til de nye baser

$$\left(\boldsymbol{v}_1' = \begin{pmatrix} 2 \\ 0 \\ 3 \end{pmatrix}, \boldsymbol{v}_2' = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \boldsymbol{v}_3' = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}\right) \qquad \text{og} \qquad \left(\boldsymbol{w}_1' = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \boldsymbol{w}_2' = \begin{pmatrix} 3 \\ 2 \end{pmatrix}\right)$$

for henholdsvis $\mathbb{F}^3$ og $\mathbb{F}^2$. Vi aflæser direkte fra definitionen af vektorerne i disse nye baser, at matricen

$$P = \begin{pmatrix} 2 & 0 & 1 \\ 0 & 1 & 0 \\ 3 & 1 & 1 \end{pmatrix} \in M_3(\mathbb{F})$$

repræsenterer identitetsafbildningen $\mathrm{id}_{\mathbb{F}^3}\colon \mathbb{F}^3 \to \mathbb{F}^3$ med hensyn til basen $(\boldsymbol{v}_1', \boldsymbol{v}_2', \boldsymbol{v}_3')$ for domænet og standardbasen $(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$ for codomænet; og at matricen

$$Q = \begin{pmatrix} 1 & 3 \\ 1 & 2 \end{pmatrix} \in M_2(\mathbb{F})$$

repræsenterer identitetsafbildningen $\mathrm{id}_{\mathbb{F}^2}\colon \mathbb{F}^2 \to \mathbb{F}^2$ med hensyn til basen $(\boldsymbol{w}_1', \boldsymbol{w}_2')$ for domænet og standardbasen $(\boldsymbol{e}_1, \boldsymbol{e}_2)$ for codomænet. Vi illusterer den information, vi har samlet ovenfor, i den nedenstående figur. Vi konkluderer derfor fra sætning 4.3.11, at matricen B er givet ved

$$B = Q^{-1}AP = \begin{pmatrix} -2 & 3 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 2 & 4 & 1 \\ 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} 2 & 0 & 1 \\ 0 & 1 & 0 \\ 3 & 1 & 1 \end{pmatrix} = \begin{pmatrix} -8 & -13 & -3 \\ 5 & 6 & 2 \end{pmatrix},$$

hvor vi igen benytter eksempel 3.4.3 til at udregne Q^{-1}.

$$\begin{tikzcd}
(\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3) \arrow[rrr, dotted, bend left, "{\begin{pmatrix}2&4&1\\1&-1&0\end{pmatrix}}"] & & & (\boldsymbol{e}_1,\boldsymbol{e}_2) \\
& \mathbb{F}^3 \arrow[r, "f"] & \mathbb{F}^2 & \\
& \mathbb{F}^3 \arrow[r, "f"] \arrow[u, "\mathrm{id}_{\mathbb{F}^3}"'] & \mathbb{F}^2 \arrow[u, "\mathrm{id}_{\mathbb{F}^2}"'] & \\
(\boldsymbol{v}'_1,\boldsymbol{v}'_2,\boldsymbol{v}'_3) \arrow[uuu, dotted, bend left, "{\begin{pmatrix}2&0&1\\0&1&0\\3&1&1\end{pmatrix}}"] \arrow[rrr, dotted, bend right, "B"'] & & & (\boldsymbol{w}'_1,\boldsymbol{w}'_2) \arrow[uuu, dotted, bend right, "{\begin{pmatrix}1&3\\1&2\end{pmatrix}}"']
\end{tikzcd}$$

Bemærkning 4.3.14 Vi opfordrer læseren til at huske det lette bevis for sætning 4.3.11 snarere end det komplicerede udsagn. For at kunne skifte koordinater og lave lignende udregninger, er det kun nødvendigt at huske følgende:

(a) Definitionen af matricen, der repræsenterer en lineær afbildning med hensyn til givne baser for dens domæne og codomæne.

(b) At hvis to linæere afbildning sammensættes, da multiplicereres de matricer, der repræsenterer dem, i samme rækkefølge.

Koordinatskift svarer til at anvende disse principper på identitetsafbildningen. Det er en god idé at tegne figurer, som vi har gjort ovenfor, for ikke at blive forvirret over, hvilken vej afbildninger går, og hvilken rækkefølge, de sammensættes i.

Eksempel 4.3.15 Vi betragter en vilkårlig vektor

$$\boldsymbol{v} = \boldsymbol{e}_1 x_1 + \boldsymbol{e}_2 x_2 + \boldsymbol{e}_3 x_3 \in \mathbb{F}^3$$

og ønsker at bestemme koordinaterne

$$\boldsymbol{v} = \boldsymbol{v}'_1 y_1 + \boldsymbol{v}'_2 y_2 + \boldsymbol{v}'_3 y_3 \in \mathbb{F}^3$$

for denne vektor med hensyn til basen $(\boldsymbol{v}'_1,\boldsymbol{v}'_2,\boldsymbol{v}'_3)$ for $\mathbb{F}^3$ fra eksempel 4.3.13. Vi anvender derfor sætning 4.3.3 på identitetsafbildningen $\mathrm{id}\colon \mathbb{F}^3 \to \mathbb{F}^3$. Sætningen siger, at hvis A er matricen, der repræsenterer $\mathrm{id}\colon \mathbb{F}^3 \to \mathbb{F}^3$ med hensyn til standard basen $(\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3)$ for domænet og basen $(\boldsymbol{v}'_1,\boldsymbol{v}'_2,\boldsymbol{v}'_3)$ for codomænet, da er

$$\boldsymbol{y} = A\boldsymbol{x}.$$

Vi har i eksempel 4.3.13 bestemt matricen P, der repræsenterer $\mathrm{id}\colon \mathbb{F}^3 \to \mathbb{F}^3$ med hensyn til basen $(\boldsymbol{v}_1', \boldsymbol{v}_2', \boldsymbol{v}_3')$ for domænet og standardbasen $(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$ for codomænet, hvilket er let at gøre. Korollar 4.3.8 viser, at $A = P^{-1}$, og vi udregner nu den inverse matrix ved hjælp af rækkeoperationer.

$$
\begin{aligned}
(P \mid I) &= \left(\begin{array}{ccc|ccc} 2 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 3 & 1 & 1 & 0 & 0 & 1 \end{array}\right) \begin{array}{l} \\ \\ +(-1)R_1 \end{array} \\
& \left(\begin{array}{ccc|ccc} 2 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & -1 & 0 & 1 \end{array}\right) \begin{array}{l} +(-2)R_3 \\ \\ \\ \end{array} \\
& \left(\begin{array}{ccc|ccc} 0 & -2 & 1 & 3 & 0 & -2 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & -1 & 0 & 1 \end{array}\right) \begin{array}{l} +2R_2 \\ \\ +(-1)R_2 \end{array} \\
& \left(\begin{array}{ccc|ccc} 0 & 0 & 1 & 3 & 2 & -2 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & -1 & -1 & 1 \end{array}\right) \begin{array}{l} \\ R_1 \leftrightarrow R_3 \\ \\ \end{array} \\
(I \mid P^{-1}) &= \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & -1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 3 & 2 & -2 \end{array}\right).
\end{aligned}
$$

Vi har her markeret de indgange, vi har ønsket at ændre undervejs, med rødt. Altså er koordinaterne af $\boldsymbol{y} \in \mathbb{F}^3$ med hensyn til basen $(\boldsymbol{v}_1', \boldsymbol{v}_2', \boldsymbol{v}_3')$ for $\mathbb{F}^3$ givet ved

$$
\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} -1 & -1 & 0 \\ 0 & 1 & 0 \\ 3 & 2 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -x_1 - x_2 \\ x_2 \\ 3x_1 + 2x_2 - 2x_3 \end{pmatrix}.
$$

Vi bemærker, at mens det er let at finde koordinaterne med hensyn til standardbasen for $\mathbb{F}^n$ fra kooridinaterne med hensyn til en given basis $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ for $\mathbb{F}^n$, så kræver det omvendt en udregning at finde koordinaterne med hensyn til basen $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ for $\mathbb{F}^n$ fra koordinaterne med hensyn til standardbasen for $\mathbb{F}^n$. Man skal nemlig typisk invertere en matrix.

Bemærkning 4.3.16 Lineære afbildninger mellem uendeligt dimensionale vektorrum forekommer naturligt og er vigtige at studere. For eksempel har mængden $C^0([a,b])$ af kontinuerte reelle funktioner $f\colon [a,b] \to \mathbb{R}$ en reel vektorrumsstruktur med vektorsum og skalar multiplikation defineret ved $(f+g)(x) = f(x) + g(x)$ og $(f \cdot a)(x) = f(x) \cdot a$, og afbildningen $I\colon C^0([a,b]) \to \mathbb{R}$ defineret ved

$$
I(f) = \int_a^b f(x)dx
$$

er da en lineær afbildning. Dimensionen af $C^0([a,b])$ er imidlertid $2^{2^{\aleph_0}}$, så det er ikke muligt at studere dette vektorrum med de algebraiske metoder, vi har udviklet her. I stedet betragter man $C^0([a,b])$ som et *topologisk* vektorrum, hvilket gør det muligt at tale om (visse) uendelige summer, og som sådant udgør den tællelige familie $(x^n)_{n\in\mathbb{N}}$ ifølge Weierstrass' approksimationssætning en slags basis.

4.4 Kerne og billede

Vi definerer kernen og billedet af en lineær afbildning og viser, hvordan disse begreber giver anledning til en langt bedre forståelse af lineære ligningssystemer.

Definition 4.4.1 Lad $\mathbb{F}$ være et legeme, Hvis $f\colon V\to W$ er en lineær afbildning mellem to $\mathbb{F}$-vektorrum, da kaldes underrummene

$$\ker(f)=\{\boldsymbol{v}\in V\mid f(\boldsymbol{v})=\boldsymbol{0}\}\subset V \qquad \text{og} \qquad \operatorname{im}(f)=\{f(\boldsymbol{v})\in W\mid \boldsymbol{v}\in V\}\subset W$$

henholdsvis *kernen* af $f\colon V\to W$ og *billedet* af $f\colon V\to W$.

Vi bemærker, at $\ker(f)\subset V$ er et underrum. For $f(\boldsymbol{0})=\boldsymbol{0}$; hvis $f(\boldsymbol{u})=\boldsymbol{0}$ og $f(\boldsymbol{v})=\boldsymbol{0}$, da er $f(\boldsymbol{u}+\boldsymbol{v})=f(\boldsymbol{u})+f(\boldsymbol{v})=\boldsymbol{0}+\boldsymbol{0}=\boldsymbol{0}$; og hvis $f(\boldsymbol{v})=\boldsymbol{0}$, da er $f(\boldsymbol{v}\cdot a)=f(\boldsymbol{v})\cdot a=\boldsymbol{0}\cdot a=\boldsymbol{0}$. Tilsvarende er også $\operatorname{im}(f)\subset W$ et underrum. For $\boldsymbol{0}=f(\boldsymbol{0})$; hvis $\boldsymbol{w}=f(\boldsymbol{u})$ og $\boldsymbol{z}=f(\boldsymbol{v})$, da er $\boldsymbol{w}+\boldsymbol{z}=f(\boldsymbol{u})+f(\boldsymbol{v})=f(\boldsymbol{u}+\boldsymbol{v})$; og hvis $\boldsymbol{w}=f(\boldsymbol{v})$, så er $\boldsymbol{w}\cdot a=f(\boldsymbol{v})\cdot a=f(\boldsymbol{v}\cdot a)$.

Eksempel 4.4.2 (1) For nulafbildningen $\boldsymbol{0}\colon V\to W$ gælder, at $\ker(\boldsymbol{0})=V$ og $\operatorname{im}(\boldsymbol{0})=\{\boldsymbol{0}\}$, og for identitetsafbildningen $\mathrm{id}_V\colon V\to V$ er $\ker(\mathrm{id}_V)=\{\boldsymbol{0}\}$ og $\operatorname{im}(\mathrm{id}_V)=V$.
(2) Hvis $f\colon\mathbb{F}^n\to\mathbb{F}^m$ repræsenteret af matricen A i standardbasen, så er $\ker(f)=N_A$ lig med nulrummet for A og $\operatorname{im}(f)$ er søjlerummet, det underrum af $\mathbb{F}^m$ frembragt af søjlene i matricen A.

Vi ser af eksempel 4.4.2, at der for både nulafbildningen og identitetsafbildningen gæder, at summen af dimensionerne af kernen og billedet er lig med dimensionen af domænet. Vi skal nu vise, at denne identitet, der kaldes Grassmann's dimensionsformel, gælder helt generelt.

Sætning 4.4.3 *Lad $\mathbb{F}$ være et legeme. Hvis $f\colon V\to W$ er en lineær afbildning mellem endeligt dimensionale $\mathbb{F}$-vektorrum, da gælder*

$$\dim(\ker(f))+\dim(\operatorname{im}(f))=\dim(V).$$

Bevis Vi vælger baser $(\boldsymbol{v}_1,\dots,\boldsymbol{v}_p)$ og $(\boldsymbol{w}_1,\dots,\boldsymbol{w}_r)$ for henholdsvis $\ker(f)$ og $\operatorname{im}(f)$, hvilket er muligt ifølge sætning 4.2.11, og vælger yderligere en familie $(\boldsymbol{v}_{p+1},\dots,\boldsymbol{v}_{p+r})$ af vektorer i V, sådan at $f(\boldsymbol{v}_{p+i})=\boldsymbol{w}_i$, for alle $1\le i\le r$. Vi påstår, at $(\boldsymbol{v}_1,\dots,\boldsymbol{v}_p,\boldsymbol{v}_{p+1},\dots,\boldsymbol{v}_{p+r})$ er en basis for V, hvorfra sætningen umiddelbart følger. For at vise påstanden, viser vi først, at familien $(\boldsymbol{v}_1,\dots,\boldsymbol{v}_p,\boldsymbol{v}_{p+1},\dots,\boldsymbol{v}_{p+r})$ er lineært uafhængig. Så vi lader

$$\boldsymbol{v}_1a_1+\cdots+\boldsymbol{v}_pa_p+\boldsymbol{v}_{p+1}a_{p+1}+\cdots+\boldsymbol{v}_{p+r}a_{p+r}=\boldsymbol{0}$$

være en linearkombination, der er lig med $\boldsymbol{0}$, og skal vise at $a_i=0$ for alle $1\le i\le p+r$. Vi anvender den lineære afbildning $f\colon V\to W$ på begge sider af ligningen og får

$$\boldsymbol{w}_1a_{p+1}+\cdots+\boldsymbol{w}_ra_{p+r}=\boldsymbol{0},$$

idet $f(\boldsymbol{v}_i)=\boldsymbol{0}$ for $1\le i\le p$ og $f(\boldsymbol{v}_{p+i})=\boldsymbol{w}_i$ for $1\le i\le r$. Da familien $(\boldsymbol{w}_1,\dots,\boldsymbol{w}_r)$ er lineært uafhængig, konkluderer vi, at $a_{p+i}=0$ for $1\le i\le r$. Den oprindelige ligning er altså

$$\boldsymbol{v}_1a_1+\cdots+\boldsymbol{v}_pa_p=\boldsymbol{0},$$

og da også $(\boldsymbol{v}_1,\dots,\boldsymbol{v}_p)$ er lineært uafhængig, følger det, at $a_i=0$ for $1\le i\le p$. Dette viser, at familien $(\boldsymbol{v}_1,\dots,\boldsymbol{v}_p,\boldsymbol{v}_{p+1},\dots,\boldsymbol{v}_{p+r})$ er lineært uafhængig som ønsket. Vi mangler at vise, at denne familie også frembringer V. Så lad $\boldsymbol{v}\in V$ være en vilkårlig vektor, og lad $\boldsymbol{w}=f(\boldsymbol{v})\in W$. Da $(\boldsymbol{w}_1,\dots,\boldsymbol{w}_r)$ frembringer W, kan vi skrive $\boldsymbol{w}=\boldsymbol{w}_1b_1+\cdots+\boldsymbol{w}_rb_r$ som en linearkombination heraf. Men da er

$$f(\boldsymbol{v}-(\boldsymbol{v}_{p+1}b_1+\cdots+\boldsymbol{v}_{p+r}b_r))=\boldsymbol{w}-(\boldsymbol{w}_1b_1+\cdots+\boldsymbol{w}_rb_r)=\boldsymbol{0},$$

hvilket viser, at $\boldsymbol{v}-(\boldsymbol{v}_{p+1}b_1+\cdots+\boldsymbol{v}_{p+r}b_r)\in\ker(f)$, og vi kan derfor skrive denne vektor som en linearkombination

$$\boldsymbol{v}-(\boldsymbol{v}_{p+1}b_1+\cdots+\boldsymbol{v}_{p+r}b_r)=\boldsymbol{v}_1a_1+\cdots+\boldsymbol{v}_pa_p$$

af familien $(\boldsymbol{v}_1,\dots,\boldsymbol{v}_p)$, der per antagelse frembringer $\ker(f)$. Men dermed er

$$\boldsymbol{v}=\boldsymbol{v}_1a_1+\cdots+\boldsymbol{v}_pa_p+\boldsymbol{v}_{p+1}b_1+\cdots+\boldsymbol{v}_{p+r}b_r$$

en linearkombination af $(\boldsymbol{v}_1,\dots,\boldsymbol{v}_p,\boldsymbol{v}_{p+1},\dots,\boldsymbol{v}_{p+r})$ som ønsket. Dette viser påstanden og dermed sætningen. □

Eksempel 4.4.4 (1) For nulafbildningen $\boldsymbol{0}\colon V\to W$ har kernen samme dimension som V, mens billedet har dimension 0. For identitetsafbildningen $\mathrm{id}_V\colon V\to V$ har kernen dimension 0, mens billedet har samme dimension som V.
(2) Den lineære afbildning $f\colon\mathbb{F}^4\to\mathbb{F}^3$ defineret ved

$$f(\boldsymbol{x})=\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&0&0\end{pmatrix}\boldsymbol{x}$$

har $\ker(f) = \{\boldsymbol{x} \in \mathbb{F}^4 \mid x_1 = 0 \text{ og } x_2 = 0\}$ og $\operatorname{im}(f) = \{\boldsymbol{y} \in \mathbb{F}^3 \mid y_3 = 0\}$. Så $\ker(f)$ har basis $(\boldsymbol{e}_3, \boldsymbol{e}_3)$ og dimension 2, mens $\operatorname{im}(f)$ har basis $(\boldsymbol{e}_1, \boldsymbol{e}_2)$ og dimension 2. Dermed er

$$\dim(\ker(f)) + \dim(\operatorname{im}(f)) = 2 + 2 = 4 = \dim(\mathbb{F}^4),$$

som det sig bør ifølge sætning 4.4.3.

Definition 4.4.5 Lad $\mathbb{F}$ være et legeme og lad $f : V \to W$ være en lineær afbildning mellem endeligt dimensionale $\mathbb{F}$-vektorrum. Da kaldes dimensionen

$$\operatorname{rank}(f) = \dim(\operatorname{im}(f))$$

af billedet af $f : V \to W$ for *rangen* af $f : V \to W$.

Lad A være en $m \times n$-matrix med indgange i et legeme $\mathbb{F}$. Vi betragter den lineære afbildning $f : \mathbb{F}^n \to \mathbb{F}^m$ defineret ved $f(\boldsymbol{x}) = A\boldsymbol{x}$. Dens kerne

$$\ker(f) = N_A = \{\boldsymbol{x} \in \mathbb{F}^n \mid A\boldsymbol{x} = \boldsymbol{0}\} \subset \mathbb{F}^n$$

er lig med nulrummet for A, mens dens billede

$$\operatorname{im}(f) = R_A = \{A\boldsymbol{x} \in \mathbb{F}^m \mid \boldsymbol{x} \in \mathbb{F}^n\} \subset \mathbb{F}^m$$

er lig med søjlerummet for A. Vi viser nu, hvordan Gauss-elimination giver algoritmer til bestemmelse af baser for begge af disse underrum.

Sætning 4.4.6 *Lad A være en $m \times n$-matrix med indgange i et legeme $\mathbb{F}$ og lad A' være den entydigt bestemte matrix på reduceret echelon form, der er rækkeækvivalent med A. Hvis $1 \le j_1 < \cdots < j_r \le n$ indicerer de søjler i A', der indeholder en ledende indgang, da er familien $(\boldsymbol{a}_{j_1}, \dots, \boldsymbol{a}_{j_r})$ bestående af de tilsvarende søjler i A en basis for $R_A \subset \mathbb{F}^m$.*

Bevis Vi vælger en invertibel matrix $P \in M_m(\mathbb{F})$, sådan at $A' = PA$, hvilket er muligt ifølge sætning 2.5.6. Vi minder om, at for en vilkårlig $m \times n$-matrix C er dens j'te søjle givet ved $\boldsymbol{c}_j = C\boldsymbol{e}_j$. Specielt gælder der for søjlerne $\boldsymbol{a}_j$ og $\boldsymbol{a}'_j$ i A og A', at

$$P\boldsymbol{a}_j = PA\boldsymbol{e}_j = A'\boldsymbol{e}_j = \boldsymbol{a}'_j$$

Vi påstår, at familien $(\boldsymbol{e}_1, \dots, \boldsymbol{e}_r)$ er en basis for $R_{A'}$. For per definitionen af en matrix på reduceret echelon form er $\boldsymbol{a}'_{j_s} = \boldsymbol{e}_s$ for alle $1 \le s \le r$, hvilket viser, at den lineært uafhængige familie $(\boldsymbol{e}_1,\dots,\boldsymbol{e}_r)$ er en familie af vektorer i $R_{A'}$, og desuden er $a'_{ij} = 0$ for alle $r < i \leq n$ og $1 \leq j \leq n$, hvilket viser, at familien $(\boldsymbol{e}_1,\dots,\boldsymbol{e}_r)$ frembringer $R_{A'}$. Dette viser påstanden. Endelig gælder der for alle $1 \leq s \leq r$, at

$$\boldsymbol{a}_{j_s} = P^{-1}\boldsymbol{a}'_{j_s} = P^{-1}\boldsymbol{e}_s,$$

og vi slutter derfor fra bemærkning 4.3.9, at familien $(\boldsymbol{a}_{j_1},\dots,\boldsymbol{a}_{j_r})$ er en basis for R_A. □

Sætningen giver specielt en algoritme, der til en given n-tuple af vektorer i $\mathbb{F}^m$, der frembringer $\mathbb{F}^m$, giver en delfamilie, der er en basis for $\mathbb{F}^m$. Vi illustrerer nu dette.

Eksempel 4.4.7 Vi betragter 5-tuplen af vektorer i $\mathbb{F}^3$ givet ved

$$\left(\boldsymbol{a}_1 = \begin{pmatrix}1\\3\\2\end{pmatrix},\ \boldsymbol{a}_2 = \begin{pmatrix}2\\4\\3\end{pmatrix},\ \boldsymbol{a}_3 = \begin{pmatrix}1\\9\\5\end{pmatrix},\ \boldsymbol{a}_4 = \begin{pmatrix}3\\0\\1\end{pmatrix},\ \boldsymbol{a}_5 = \begin{pmatrix}2\\7\\8\end{pmatrix}\right),$$

og vi anvender da sætning 4.4.6 til at finde en delfamilie heraf, der en basis for $\mathbb{F}^3$, hvis dette er muligt. Vi omdanner derfor matricen

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

hvis j'te søjle er $\boldsymbol{a}_j$, til en matrix A' på reduceret echelon form ved hjælp af rækkeoperationer, hvilket givet

$$A' = \begin{pmatrix}1 & 0 & 7 & 0 & -39\\0 & 1 & -3 & 0 & 31\\0 & 0 & 0 & 1 & -7\end{pmatrix}.$$

der har ledene indgange i søjlerne 1, 2 og 4. Ifølge sætning 4.4.6 er $(\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_4)$ derfor en basis $\mathbb{F}^3$. Vi bemærker, at grunden til, at en sådan basis fandtes i dette tilfælde, er, at $\operatorname{rank}(A) = 3 = m$. Hvis dette ikke havde været tilfædet, så ville den oprindelige tuple ikke frembringe hele $\mathbb{F}^3$, men et underrum deraf, og vi ville da i stedet have fundet en basis for dette underrum.

Korollar 4.4.8 *Lad $\mathbb{F}$ være et legeme, lad A være en $m \times n$-matrix med indgange i $\mathbb{F}$, og lad $f\colon \mathbb{F}^n \to \mathbb{F}^m$ være den lineære afbildning defineret ved $f(\boldsymbol{x}) = A\boldsymbol{x}$. Da gælder*

$$\operatorname{rank}(f) = \operatorname{rank}(A).$$

Bevis Sætning 4.4.6 viser, at dimensionen af søjlerummet R_A er lig med antallet r af ledende indgang i den entydigt bestemte matrix B, der er på reduceret echelon form og rækkeækvivalent med A. Der gælder altså, at

$$\operatorname{rank}(f) = \dim(\ker(f)) = \dim(R_A) = r = \operatorname{rank}(A)$$

som ønsket. □

Sætning 4.4.9 *Lad A være en $m \times n$-matrix med indgange i et legeme $\mathbb{F}$ og lad A' være den entydigt bestemte matrix på reduceret echelon form, der er rækkeækvivalent med A. Lad $1 \le j_1 < \cdots < j_r \le n$ og $1 \le k_1 < \cdots < k_p \le n$ indicere de søjler i A', der henholdsvis indeholder og ikke indeholder ledende indgange, og for $1 \le i \le p$ lad*

$$\boldsymbol{c}_i = \boldsymbol{e}_{k_i} - \sum_{s=1}^{r} \boldsymbol{e}_{j_s} a'_{s,k_i}.$$

Da udgør familien $(\boldsymbol{c}_1, \ldots, \boldsymbol{c}_p)$ en basis for nulrummet $N_A = N_{A'} \subset \mathbb{F}^n$.

Bevis Vi benytter igen, at $A' \cdot \boldsymbol{e}_j = \boldsymbol{a}'_j$ er den j'te søjle i A', og at A' er på reduceret echelon form. Vi udregner ved brug heraf, at

$$A' \cdot \boldsymbol{c}_i = A' \cdot (\boldsymbol{e}_{k_i} - \sum_{s=1}^{r} \boldsymbol{e}_{j_s} a'_{s,k_i}) = \boldsymbol{a}'_{k_i} - \sum_{s=1}^{r} \boldsymbol{e}_s a'_{s,k_i} = \mathbf{0}.$$

For den j_s'te søjle i A' er $\boldsymbol{e}_s$, og da $a'_{s,k_i} = 0$ for $r < s \le m$. Dermed er $\boldsymbol{c}_i \in N_{A'}$ for alle $1 \le i \le p$. Vi vil nu vise, at familien $(\boldsymbol{c}_1, \ldots, \boldsymbol{c}_p)$ er lineært uafhængig. Så lad

$$\boldsymbol{c}_1 a_1 + \boldsymbol{c}_2 a_2 + \cdots + \boldsymbol{c}_p a_p = \mathbf{0}$$

være en linearkombination, der er lig med nulvektoren. For $1 \le i \le p$ udtrykker den k_i'te ligning i dette ligningssystem, at $a_i = 0$, da $\boldsymbol{e}_{k_i}$ kun optræder i vektoren $\boldsymbol{c}_{k_i}$, hvilket viser, at $(\boldsymbol{c}_1, \ldots, \boldsymbol{c}_p)$ er lineært uafhængig som ønsket. Endelig viser sætning 4.4.3 og korollar 4.4.8, at $\dim(N_{A'}) = n - r = p$. Vi har dermed vist, at familien $(\boldsymbol{c}_1, \ldots, \boldsymbol{c}_p)$ er en lineært uafhængig p-tuple af vektorer i det p-dimensionale vektorrum $N_{A'}$, og den er derfor en basis ifølge sætning 4.2.19. $\square$

Vi illustrerer endelig med et eksempel algoritmerne i sætning 4.4.9 og sætning 4.4.6, der producerer baser for henholdsvis kernen og billedet af en lineær afbildning $f \colon \mathbb{F}^n \to \mathbb{F}^m$.

Eksempel 4.4.10 Vi betragter afbildningen $f \colon \mathbb{F}^4 \to \mathbb{F}^3$ givet ved $f(\boldsymbol{x}) = A\boldsymbol{x}$, hvor

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

og ønsker at bestemme en basis for $\ker(f)$ og $\operatorname{im}(f)$. Vi omdanner derfor først A til en matrix A' på reduceret echelon form ved hjælp af rækkeoperationer.

$$
\begin{aligned}
A = &\begin{pmatrix} 2 & 4 & 3 & 7 \\ 0 & 0 & 1 & 1 \\ 1 & 2 & 1 & 3 \end{pmatrix} \begin{matrix} +(-2)R_3 \\ \\ \\ \end{matrix} \\
&\begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 2 & 1 & 3 \end{pmatrix} \begin{matrix} +(-1)R_2 \\ \\ +(-1)R_2 \end{matrix} \\
&\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 2 & 0 & 2 \end{pmatrix} \; R_1 \leftrightarrow R_3 \\
A' = &\begin{pmatrix} 1 & 2 & 0 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}
\end{aligned}
$$

Her har vi igen markeret de indgange, vi ønsker at ændre, med rødt, og vi har markeret de $r = 2$ ledende indgange i $A' = B$ med blåt. Så $\operatorname{im}(f)$ har dimension $r = 2$, og Grassmann's dimensionsformel viser, at $\ker(f)$ har dimension $n - r = 4 - 2 = 2$. Ifølge sætning 4.4.6 er

$$
\left(\boldsymbol{a}_1 = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix},\ \boldsymbol{a}_3 = \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} \right)
$$

en basis for $\operatorname{im}(f)$; og ifølge sætning 4.4.9 er $(\boldsymbol{c}_1, \boldsymbol{c}_2)$ en basis for $\ker(f)$, hvor

$$
\boldsymbol{c}_1 = \boldsymbol{e}_2 - (\boldsymbol{e}_{j_1} b_{12} + \boldsymbol{e}_{j_2} b_{22}) = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} 2 - \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} 0 = \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix}
$$

og

$$
\boldsymbol{c}_2 = \boldsymbol{e}_4 - (\boldsymbol{e}_{j_1} b_{14} + \boldsymbol{e}_{j_2} b_{24}) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} 2 - \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} 1 = \begin{pmatrix} -2 \\ 0 \\ -1 \\ 1 \end{pmatrix}.
$$

5 Egenværdier og egenrum

I dette kapitel introducerer vi de vigtige begreber egenværdi og egenrum af en lineær afbildning. Disse begreber er dels kraftfulde redskaber til at forstå lineære afbildninger og er dels vigtige begreber i sig selv. I kvantemekanik er de observable således lineære afbildninger, mens de mulige værdier, som disse observable kan antage ved måling, netop er de lineære afbildningers egenværdier. Så selvom definitionen af egenværdier og egenrum er ganske simpel, så er det en meget righoldig teori, vi herved udvikler, og en teori der på mange måder udfordrer vores intuition. Vi skal i dette kapitel igen bruge, at den kommutative lov gælder for multplikation af skalarer.

5.1 Egenværdier og egenrum for kvadratiske matricer

Givet to kvadratiske matricer $A = (a_{ij}) \in M_n(\mathbb{F})$ og $B = (b_{ij}) \in M_n(\mathbb{F})$ med indgange i et legeme $\mathbb{F}$, da definerer vi $A + Bt \in M_n(\mathbb{F}[t])$ til at være den kvadratiske matrix med indgange i polynomiumsringen $\mathbb{F}[t]$ givet ved

$$A + Bt = (a_{ij} + b_{ij}t) \in M_n(\mathbb{F}[t]).$$

Vi kan da betragte dens determinant

$$\det(A + Bt) \in \mathbb{F}[t],$$

som ifølge sætning 3.5.5 er et polynomium af grad højst n. Vi skal her kun betragte tilfældet, hvor $B = -I$ er den modsatte matrix af identitetsmatricen.

Definition 5.1.1 Hvis $\mathbb{F}$ er et legeme og $A \in M_n(\mathbb{F})$, da kaldes polynomiet

$$\chi_A(t) = \det(A - It) = \det \begin{pmatrix} a_{11} - t & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - t & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} - t \end{pmatrix}$$

for det *karakteristike polynomium* af A.

For eksempel er det karakteristiske polynomium af en 2×2-matrix er givet ved

$$\chi_A(t) = \det\begin{pmatrix} a_{11} - t & a_{12} \\ a_{21} & a_{22} - t \end{pmatrix} = t^2 - (a_{11} + a_{22})t + (a_{11}a_{22} - a_{12}a_{21}),$$

hvilket vi kan skrive mere kompakt som $\chi_A(t) = t^2 - \mathrm{tr}(A)t + \det(A)$, hvor $\mathrm{tr}(A) = a_{11} + a_{22}$ er summen af diagonal elementerne.

Eksempel 5.1.2 Betragt matricen

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 3 & 5 \\ 0 & 2 & 4 \end{pmatrix}.$$

Da er det karakteristiske polynomium af A lig med

$$\begin{aligned} \chi_A(t) = \det(A - It) = \det\begin{pmatrix} 1-t & 2 & 0 \\ 0 & 3-t & 5 \\ 0 & 2 & 4-t \end{pmatrix} &= (1-t)\det\begin{pmatrix} 3-t & 5 \\ 2 & 4-t \end{pmatrix} \\ &= (1-t)\big((3-t)(4-t) - 10\big) = -t^3 + 15t^2 - 9t + 2 \end{aligned}$$

Sætning 5.1.3 *Lad $A \in M_n(\mathbb{F})$ være en kvadratisk matrix med indgange i et legeme $\mathbb{F}$. For enhver invertibel matrix $P \in M_n(\mathbb{F})$ gælder, at*

$$\chi_{P^{-1}AP}(t) = \chi_A(t).$$

Bevis Givet $Q \in M_n(\mathbb{F}[t])$, da er den (i,k)'te indgang i $Q \cdot (It)$ lig med

$$\sum_{j=1}^{n} q_{ij} \cdot \delta_{jk} t = q_{ik} \cdot t,$$

mens den (i,k)'te indgang i $(It) \cdot Q$ er lig med

$$\sum_{j=1}^{n} \delta_{ij} t \cdot q_{jk} = t \cdot q_{ik}.$$

Da den kommutative lov gælder for multiplikation af polynomier, konkluderer vi derfor, at $Q \cdot (It) = (It) \cdot Q$. Specielt er

$$P^{-1}(A - It)P = P^{-1}AP - P^{-1}(It)P = P^{-1}AP - P^{-1}P(It) = P^{-1}AP - It,$$

og derfor viser sætning 3.2.14, at

$$\begin{aligned} \chi_A(t) &= \det(A - It) = \det(P)^{-1}\det(A - It)\det(P) = \det(P^{-1}(A - It)P) \\ &= \det(P^{-1}AP - It) = \chi_{P^{-1}AP}(t) \end{aligned}$$

som ønsket. □

Givet en kvadratisk matrix $A \in M_n(\mathbb{F})$, da kan vi substituere skalaren $\lambda \in \mathbb{F}$ for t i det karakteristiske polynomium $\chi_A(t) \in \mathbb{F}[t]$, hvorved vi får skalaren

$$\chi_A(\lambda) = \det(A - I\lambda) = \det\begin{pmatrix} a_{11}-\lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22}-\lambda & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{m2} & \dots & a_{nn}-\lambda \end{pmatrix} \in \mathbb{F}.$$

Vi vil derfor også skrive $I\lambda$ for diagonalmatricen $\operatorname{diag}(\lambda, \dots, \lambda)$.

Lemma 5.1.4 *Hvis $\mathbb{F}$ er et legeme, og hvis $A \in M_n(\mathbb{F})$ er en kvadratisk matrix, da er nulrummet*

$$N_{A-I\lambda} = \{\boldsymbol{v} \in \mathbb{F}^n \mid A\boldsymbol{v} = \boldsymbol{v}\lambda\} \subset \mathbb{F}^n$$

for alle $\lambda \in \mathbb{F}$.

Bevis Per definition er $\boldsymbol{v} \in N_{A-I\lambda}$, hvis og kun hvis $(A - I\lambda)\boldsymbol{v} = \boldsymbol{0}$, og dette gælder, hvis og kun hvis $A\boldsymbol{v} = (I\lambda)\boldsymbol{v}$. Bemærk nu at for alle $\boldsymbol{v} \in \mathbb{F}^n$, så er

$$(I\lambda)\boldsymbol{v} = \begin{pmatrix} \lambda v_1 \\ \vdots \\ \lambda v_n \end{pmatrix} \quad \text{mens} \quad \boldsymbol{v}\lambda = \begin{pmatrix} v_1\lambda \\ \vdots \\ v_n\lambda \end{pmatrix}.$$

Da vi har antaget, at den kommutative lov gælder for multiplikation af skalarer, konkluderer vi derfor, at $(I\lambda)\boldsymbol{v} = \boldsymbol{v}\lambda$, hvilket beviser lemmaet. □

Vi udtrykker lemma 5.1.4 ved at sige, at på underrummet $N_{A-I\lambda} \subset \mathbb{F}^n$ er venstre multiplikation med A givet ved højre skalering med vægt λ.

Definition 5.1.5 Lad A være en $n \times n$-matrix med indgange i et legeme $\mathbb{F}$.

(1) En skalar $\lambda \in \mathbb{F}$ er en *egenværdi* for A, hvis $N_{A-I\lambda} \neq \{\boldsymbol{0}\}$.

(2) Hvis $\lambda \in \mathbb{F}$ er en egenværdi for A, så kaldes nulrummet $N_{A-I\lambda}$ for *egenrummet* hørende til egenværdien $\lambda \in \mathbb{F}$.

(3) Hvis $\lambda \in \mathbb{F}$ er en egenværdi for A, så kaldes enhver vektor $\boldsymbol{v} \in N_{A-I\lambda}$, der ikke er nulvektoren, for en *egenvektor* for A hørende til egenværdien $\lambda \in \mathbb{F}$.

Vi bemærker, at mens egenrummet $N_{A-I\lambda}$ hørende til egenværdi $\lambda \in \mathbb{F}$ er entydigt bestemt, så er der typisk (uendeligt) mange egenvektorer $\boldsymbol{v}$ hørende til egenværdi λ.

Sætning 5.1.6 *Egenværdierne af en kvadratisk matrix $A \in M_n(\mathbb{F})$ med indgange i et legeme $\mathbb{F}$ er præcis rødderne i $\mathbb{F}$ af det karakteristiske polynomium $\chi_A(t) \in \mathbb{F}[t]$.*

Bevis Per definition er $\lambda \in \mathbb{F}$ en egenværdi for A, hvis den homogene ligning

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

har en ikke-triviel løsning $\boldsymbol{x} = \boldsymbol{v} \neq \boldsymbol{0}$, og da matricen $A - I\lambda$ er kvadratisk, er dette tilfældet, hvis og kun hvis matricen $A - I\lambda$ ikke er invertibel. Ifølge sætning 3.4.1 er matricen $A - I\lambda$ invertibel, hvis og kun hvis dens determinant $\chi_A(\lambda) = \det(A - I\lambda)$ er invertibel, og da $\mathbb{F}$ er et legeme, er $\chi_A(\lambda)$ invertibel, hvis og kun hvis $\chi_A(\lambda) \neq 0$. Vi har hermed vist, at $\lambda \in \mathbb{F}$ er en egenværdi for A, hvis og kun hvis $\chi_A(\lambda) = 0$ som ønsket. □

Givet en egenværdi $\lambda \in \mathbb{F}$ for en kvadratisk matrix $A \in M_n(\mathbb{F})$, så kaldes dimensionen af egenrummet $N_{A-I\lambda}$ hørende til λ for den *geometriske multiplicitet* af λ, mens det største hele tal $k \geq 1$, for hvilket der gælder $\chi_A(t) = (\lambda - t)^k \cdot q(t) \in \mathbb{F}[t]$, kaldes for den *algebraiske multiplicitet* af λ. Vi beviser i sætning 5.3.7, at den algebraiske multiplicitet af en egenværdi altid er større end eller lig med den geometriske multiplicitet.

Eksempel 5.1.7 (1) Identitetsmatricen $I \in M_n(\mathbb{F})$ har karakteristisk polynomium

$$\chi_I(t) = \det(I - It) = (1 - t)^n.$$

Sætning 5.1.6 viser derfor, at $\lambda = 1$ er eneste egenværdi. Den algebraisk multiplicitet er lig med n, og da $N_{I-I1} = N_O = \mathbb{F}^n$, er også den geometrisk multiplicitet lig med n.
(2) Nulmatricen $O \in M_n(\mathbb{F})$ har karakteristisk polynomium

$$\chi_O(t) = \det(O - It) = \det(-It) = (-t)^n.$$

Så $\lambda = 0$ er eneste egenværdi med algebraisk multiplicitet n, og da $N_{O-I0} = N_O = \mathbb{F}^n$, er den geometriske multiplicitet også lig med n.
(3) Det karakteristiske polynomium af matricen

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in M_2(\mathbb{F})$$

udregnes til at være

$$\chi_A(t) = \det\begin{pmatrix} -t & -1 \\ 1 & -t \end{pmatrix} = t^2 + 1.$$

Hvis $\mathbb{F} = \mathbb{R}$, så har matricen A altså ingen egenværdier, men hvis $\mathbb{F} = \mathbb{C}$, så er

$$\chi_A(t) = (t - i)(t + i),$$

hvilket viser, at A har egenværdierne $\lambda_1 = +i$ og $\lambda_2 = -i$, og at begge har algebraisk multiplicitet 1. For at udregne egenrummet hørende til $\lambda_1 = +i$, skal vi løse ligningssystemet $(A - Ii)\boldsymbol{x} = \mathbf{0}$. Vi omdanner matricen $B = A - Ii$ ved hjælp af rækkeoperationer til en matrix B' på reduceret echelon-form.

$$
\begin{aligned}
B = &\begin{pmatrix} -i & -1 \\ 1 & -i \end{pmatrix} \quad i \cdot R_1 \\
&\begin{pmatrix} 1 & -i \\ 1 & -i \end{pmatrix} \quad +(-1)R_1 \\
B' = &\begin{pmatrix} 1 & -i \\ 0 & 0 \end{pmatrix}
\end{aligned}
$$

Egenrummet for $A \in M_2(\mathbb{C})$ hørende til $\lambda_1 = +i$ er altså givet ved

$$N_{A-Ii} = \left\{ \begin{pmatrix} i \\ 1 \end{pmatrix} \cdot t \mid t \in \mathbb{C} \right\} \subset \mathbb{C}^2.$$

Tilsvarende, for at finde egenrummet hørende til $\lambda = -1$ skal vi løse ligningssystemet $(A - I(-i))\boldsymbol{x} = \mathbf{0}$. Vi omdanner matricen $C = A - I(-i)$ ved hjælp af rækkeoperationer til en matrix C' på reduceret echelon-form.

$$
\begin{aligned}
C = &\begin{pmatrix} i & -1 \\ 1 & i \end{pmatrix} \quad (-i) \cdot R_1 \\
&\begin{pmatrix} 1 & i \\ 1 & i \end{pmatrix} \quad +(-1)R_1 \\
C' = &\begin{pmatrix} 1 & i \\ 0 & 0 \end{pmatrix}
\end{aligned}
$$

Heraf ser vi, at egenrummet for $A \in M_2(\mathbb{C})$ hørende til egenværdien $\lambda_2 = -i$ er givet ved

$$N_{A+Ii} = \left\{ \begin{pmatrix} -i \\ 1 \end{pmatrix} \cdot t \;\middle|\; t \in \mathbb{C} \right\} \subset \mathbb{C}^2.$$

Egenværdierne $\lambda_1 = +i$ og $\lambda_2 = -i$ har dermed også begge geometrisk multiplicitet 1.
(4) Det karakteristiske polynomium af matricen

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in M_2(\mathbb{F})$$

er $\chi_A(t) = (1-t)^2$, hvilket viser, at A har $\lambda = 1$ som eneste egenværdi med algebraisk multiplicitet 2. Det tilhørende egenrum

$$N_{A-I1} = \left\{ \boldsymbol{x} \in \mathbb{F}^2 \;\middle|\; \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \boldsymbol{x} = \mathbf{0} \right\} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix} \cdot t \;\middle|\; t \in \mathbb{F} \right\} \subset \mathbb{F}^2$$

har imidlertid dimension 1, så egenværdien $\lambda = 1$ har altså geometriske multiplicitet 1, hvilket er mindre end den algebraisk multiplicitet 2.

Sætning 5.1.8 *Lad $\mathbb{F}$ være et legeme, og lad A være en $n \times n$-matrix med indgange i $\mathbb{F}$. Lad $(\lambda_1, \dots, \lambda_p)$ være en familie af parvis forskellige egenværdier for A, og lad $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_p)$ være en familie af vektorer i $\mathbb{F}^m$, sådan at $\boldsymbol{v}_i$ er en egenvektor hørende til λ_i for alle $1 \le i \le p$. Da er familien $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_p)$ lineært uafhængig. Specielt er $p \le n$.*

Bevis Vi beviser udsagnet ved induktion på $p \ge 0$. Hvis $p = 0$, så gælder udsagnet, da den tomme familie altid er lineært uafhængig. Vi antager derfor, at udsagnet allerede er bevist for $p = r - 1$ og beviser, at det gælder for $p = r$. Så vi lader

$$\boldsymbol{v}_1 c_1 + \cdots + \boldsymbol{v}_r c_r = \mathbf{0}$$

være en linear kombination af $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_r)$, som er lig med nulvektoren, og skal vise, at $c_i = 0$ for alle $1 \le i \le r$. Hertil bemærker vi, at

$$(A - I\lambda_i)(\boldsymbol{v}_j) = \boldsymbol{v}_j \cdot (\lambda_j - \lambda_i),$$

hvilket viser, at

$$(A - I\lambda_1) \dots (A - I\lambda_{r-1})(\boldsymbol{v}_1 c_1 + \cdots + \boldsymbol{v}_r c_r) = \boldsymbol{v}_r c_r (\lambda_r - \lambda_1) \cdots (\lambda_r - \lambda_{r-1}).$$

Da venstresiden per antagelse er lig med $\mathbf{0}$, konkluderer vi, at $c_r = 0$. Dermed er

$$\boldsymbol{v}_1 c_1 + \cdots + \boldsymbol{v}_{r-1} c_{r-1} = \mathbf{0},$$

og den induktive hypotese viser derfor, at $c_1 = \cdots = c_{r-1} = 0$. Altså er $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_r)$ lineært uafhængig, hvilket viser induktionsskridtet. Endelig følger den sidste påstand, at der nødvendigvis gælder $p \le n$, fra sætning 4.2.13. □

Korollar 5.1.9 *Lad $\mathbb{F}$ være et legeme, og lad $A \in M_n(\mathbb{F})$ være en kvadratisk matrix med indgange i $\mathbb{F}$. Hvis $(\lambda_1, \dots, \lambda_n)$ er en familie af n parvis forskellige egenværdier for A, da er enhver familie $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ af tilhørende egenvektorer en basis for $\mathbb{F}^n$.*

Bevis Familien af egenvektorer $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ er lineært uafhængig ifølge sætning 5.1.8, og da den består af $n = \dim_{\mathbb{F}}(\mathbb{F}^n)$ vektorer, viser sætning 4.2.19, at den er en basis. □

Eksempel 5.1.10 Vi så i eksempel 5.1.7, at matricen

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in M_2(\mathbb{C})$$

har egenværdier $\lambda_1 = +i$ og $\lambda_2 = -i$, og at vektorerne

$$\boldsymbol{v}_1 = \begin{pmatrix} i \\ 1 \end{pmatrix} \qquad \text{og} \qquad \boldsymbol{v}_2 = \begin{pmatrix} -i \\ 1 \end{pmatrix},$$

er egenvektorer for henholdsvis λ_1 og λ_2. Dermed er $(\boldsymbol{v}_1, \boldsymbol{v}_2)$ en familie af egenvektorer hørende til $n = 2$ forskellige egenværdier for A, og derfor viser korollar 5.1.9, at $(\boldsymbol{v}_1, \boldsymbol{v}_2)$ er en basis for $\mathbb{C}^2$. (Dette følger også fra $\det\begin{pmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 \end{pmatrix} = 2i \neq 0$.)

5.2 Diagonaliserbare matricer

En kvadratisk matrix $A \in M_n(\mathbb{F})$ kaldes *diagonaliserbar*, hvis der findes en invertibel matrix $P \in M_n(\mathbb{F})$, sådan at $P^{-1}AP$ er en diagonalmatrix D.

En diagonalmatrix D svarer geometrisk til en vægtet skalering $f\colon \mathbb{F}^n \to \mathbb{F}^n$ langs hver af vektorerne i standardbasen $(\boldsymbol{e}_1,\ldots,\boldsymbol{e}_n)$ for $\mathbb{F}^n$. Tilsvarende svarer en diagonaliserbar matrix A til en vægtet skalering $f\colon \mathbb{F}^n \to \mathbb{F}^n$ langs hver af vektorerne i generel basis $(\boldsymbol{v}_1,\ldots,\boldsymbol{v}_n)$ for $\mathbb{F}^n$.

Sætning 5.2.1 *Lad $\mathbb{F}$ være et legeme, lad $A,P \in M_n(\mathbb{F})$ være kvadratiske matricer, og lad $(\boldsymbol{v}_1,\ldots,\boldsymbol{v}_n)$ være familien af søjlevektorer i P. Da er følgende udsagn ækvivalente:*

(1) *Matricen P er invertibel og $P^{-1}AP$ er en diagonal matrix.*

(2) *Familien $(\boldsymbol{v}_1,\ldots,\boldsymbol{v}_n)$ er en basis for $\mathbb{F}^n$ og består af egenvektorer for A.*

Hvis dette er tilfældet og hvis $P^{-1}AP = \operatorname{diag}(\lambda_1,\ldots,\lambda_n)$, så gælder endvidere, at $\boldsymbol{v}_i$ er en egenvektor for A hørende til egenværdien λ_i for alle $1 \le i \le n$.

Bevis Husk fra bemærkning 4.2.6 at matricen $P \in M_n(\mathbb{F})$ er invertibel, hvis og kun hvis hvis familien

$$(\boldsymbol{v}_1,\ldots,\boldsymbol{v}_n) = (P\boldsymbol{e}_1,\ldots,P\boldsymbol{e}_n)$$

er en basis for $\mathbb{F}^n$. Der gælder endvidere, at $P^{-1}AP = \operatorname{diag}(\lambda_1,\ldots,\lambda_n)$, hvis og kun hvis $P^{-1}AP\boldsymbol{e}_i = \boldsymbol{e}_i\lambda_i$ for alle $1 \le i \le n$, hvilket gælder, hvis og kun hvis $AP\boldsymbol{e}_i = P\boldsymbol{e}_i\lambda_i$ for alle $1 \le i \le n$. Idet $P\boldsymbol{e}_i = \boldsymbol{v}_i$, gælder dette hvis og kun hvis $A\boldsymbol{v}_i = \boldsymbol{v}_i\lambda_i$ for alle $1 \le i \le n$, som præcis siger at basen $(\boldsymbol{v}_1,\ldots,\boldsymbol{v}_n)$ består af egenvektorer. $\square$

Eksempel 5.2.2 Vi undersøger om matricen

$$A = \begin{pmatrix} -4 & -6 \\ 1 & 1 \end{pmatrix}$$

er diagonaliserbar. Dette er ifølge sætning 5.2.1 tilfældet, hvis og kun hvis der findes en basis $(\boldsymbol{v}_1,\boldsymbol{v}_2)$ for $\mathbb{F}^2$, der består af egenvektorer for A, og vi begynder derfor med at udregne det karakteristiske polynomium.

$$\begin{aligned}\chi_A(t) &= \det\begin{pmatrix} -4-t & -6 \\ 1 & 1-t \end{pmatrix} \overset{(D6')}{=} \det\begin{pmatrix} -2-t & -4-2t \\ 1 & 1-t \end{pmatrix} \overset{(D2')}{=} (2+t)\det\begin{pmatrix} -1 & -2 \\ 1 & 1-t \end{pmatrix} \\ &\overset{(D6')}{=} (2+t)\det\begin{pmatrix} -1 & -2 \\ 0 & -1-t \end{pmatrix} \overset{(D2')}{=} (2+t)(1+t)\det\begin{pmatrix} -1 & -2 \\ 0 & -1 \end{pmatrix} = (2+t)(1+t).\end{aligned}$$

Vi ser, at A har to forskellige egenværdier $\lambda_1 = -2$ og $\lambda_2 = -1$, så ifølge korollar 5.1.9 har $\mathbb{F}^2$ derfor en basis $(\boldsymbol{v}_1, \boldsymbol{v}_2)$ bestående af egenvektorer. Vi bestemmer egenrummet hørende til $\lambda_1 = -2$ og omdanner derfor matricen $B = A - I(-2)$ til en matrix B' på reduceret echelon-form ved hjælp af rækkeoperationer.

$$
\begin{aligned}
B = &\begin{pmatrix} -2 & -6 \\ 1 & 3 \end{pmatrix} \quad (-\tfrac{1}{2})\cdot R_1 \\
&\begin{pmatrix} 1 & 3 \\ 1 & 3 \end{pmatrix} \quad +(-1)R_1 \\
B' = &\begin{pmatrix} 1 & 3 \\ 0 & 0 \end{pmatrix}
\end{aligned}
$$

Heraf ser vi, at egenrummet hørende til $\lambda_1 = -2$ er givet ved

$$N_{A-I(-2)} = \left\{ \begin{pmatrix} -3 \\ 1 \end{pmatrix} \cdot t \;\middle|\; t \in \mathbb{F} \right\},$$

og en tilsvarende udregning viser, at egenrummet hørende til $\lambda_2 = -1$ er givet ved

$$N_{A-I(-2)} = \left\{ \begin{pmatrix} -2 \\ 1 \end{pmatrix} \cdot t \;\middle|\; t \in \mathbb{F} \right\}.$$

Sætning 5.2.1 viser derfor, at matricen

$$P = \begin{pmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 \end{pmatrix} = \begin{pmatrix} -3 & -2 \\ 1 & 1 \end{pmatrix}$$

er invertibel, og at

$$P^{-1}AP = \begin{pmatrix} -2 & 0 \\ 0 & -1 \end{pmatrix} = \operatorname{diag}(-2, -1),$$

hvilket også er let at eftervise.

Bemærkning 5.2.3 En anden fordel ved diagonalmatricer er, at det er let at udregne deres potenser. For hvis $D = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$, da er $D^k = \operatorname{diag}(\lambda_1^k, \dots, \lambda_n^k)$. Hvis A er diagonaliserbar, så kan man tilsvarende udregne A^k mere økonomisk som følger. Vi potensopløfter diagonal matricen $D = P^{-1}AP$ som før, og bemærker, at

$$D^k = (P^{-1}AP)^k = P^{-1}APP^{-1}AP \dots P^{-1}AP = P^{-1}A^kP,$$

idet vi fjerner de $k-1$ forekomster af $PP^{-1} = I$. Derfor er

$$A^k = PD^kP^{-1},$$

så op til et koordinatskift er det ligeså let at udregne potenser af en diagonaliserbar matrix som af en diagonalmatrix. Vi bemærker også, at A^k igen er diagonaliserbar.

Eksempel 5.2.4 Vi anvender bemærkning 5.2.3 på den diagonaliserbare matrix A i eksempel 5.2.2 og får, at

$$A^k = P\operatorname{diag}((-2)^k,(-1)^k)P^{-1} = \begin{pmatrix} -3 & -2 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} (-2)^k & 0 \\ 0 & (-1)^k \end{pmatrix}\begin{pmatrix} -1 & -2 \\ 1 & 3 \end{pmatrix}$$
$$\begin{pmatrix} -3(-2)^k & -2(-1)^k \\ (-2)^k & (-1)^k \end{pmatrix}\begin{pmatrix} -1 & -2 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 3(-2)^k-2(-1)^k & 6(-2)^k-6(-1)^k \\ -(-2)^k+(-1)^k & -2(-2)^k+3(-1)^k \end{pmatrix}.$$

Her har vi anvendt eksempel 3.4.3 til at udregne P^{-1}.

Sætning 5.2.5 *Lad A og B være to $n \times n$-matricer med indgange i et legeme $\mathbb{F}$. Hvis der findes en basis $(\boldsymbol{v}_1,\dots,\boldsymbol{v}_n)$ for $\mathbb{F}^n$ bestående af vektorer, der både er egenvektorer for A og B, så gælder der nødvendigvis, at $AB = BA$.*

Bevis Ifølge sætning 5.2.1 er matricen $P = \begin{pmatrix} \boldsymbol{v}_1 & \dots & \boldsymbol{v}_n \end{pmatrix}$ invertibel, og matricerne

$$\begin{aligned} P^{-1}AP &= \operatorname{diag}(\lambda_1,\dots,\lambda_n), \\ P^{-1}BP &= \operatorname{diag}(\mu_1,\dots,\mu_n) \end{aligned}$$

er diagonalmatricer. Vi konkluderer derfor, at

$$AB = P\operatorname{diag}(\lambda_1,\dots,\lambda_n)P^{-1}P\operatorname{diag}(\mu_1,\dots,\mu_n)P^{-1} = P\operatorname{diag}(\lambda_1\mu_1,\dots,\lambda_n\mu_n)P^{-1}$$

og

$$BA = P\operatorname{diag}(\mu_1,\dots,\mu_n)P^{-1}P\operatorname{diag}(\lambda_1,\dots,\lambda_n)P^{-1} = P\operatorname{diag}(\mu_1\lambda_1,\dots,\mu_n\lambda_n)P^{-1}.$$

Heraf følger som ønsket, at $AB = BA$, hvor vi anvender, at den kommutative lov gælder for multiplikation af skalarer. □

Bemærkning 5.2.6 Til trods for, at sætning 5.2.5 er ganske let at vise, så har den enorme konsekvenser, idet den er grundlaget for Heisenberg's ubestemthedsrelationer. I kvantemekanik svarer observable og tilstande henholdsvis til matricer og vektorer, og en observabel A har en veldefineret værdi λ i tilstanden $\boldsymbol{v}$, hvis og kun hvis $\boldsymbol{v}$ er en egenvektor for A med egenværdi λ. Sætning 5.2.5 siger derfor, at hvis $AB \neq BA$, så er værdierne af de observable A og B ikke samtidigt veldefinerede.

Hvis matricen A er diagonaliserbar og $P^{-1}AP = \operatorname{diag}(\lambda_1,\dots,\lambda_n)$, så følger det fra sætning 5.3.6, at det karakteristiske polynomium faktoriserer som et produkt

$$\chi_A(t) = \chi_{P^{-1}AP}(t) = \det\big(\operatorname{diag}(\lambda_1 - t,\dots,\lambda_n - t)\big) = (\lambda_1 - t)\cdot\ldots\cdot(\lambda_n - t)$$

af førstegradspolynomier. Det omvendte udsagn gælder imidlertid ikke, medmindre egenværdierne $\lambda_1,\dots,\lambda_n$ er parvis forskellige. Vi har i stedet følgende resultat.

Sætning 5.2.7 *Lad $A \in M_n(\mathbb{F})$ være en kvadratisk matrix med indgange i et legeme $\mathbb{F}$. Da er følgende udsagn ækvivalente:*

(1) *Det karakteristiske polynomium er et produkt af førstegradspolynomier,*

$$\chi_A(t) = (\lambda_1 - t) \cdot \ldots \cdot (\lambda_n - t).$$

(2) *Der findes en invertibel matrix $P \in M_n(\mathbb{F})$, sådan at matricen $P^{-1}AP$ er en øvre triangulær matrix.*

Bevis Vi antager først (1) og beviser (2). Beviset er ved induktion på $n \geq 0$. Tilfældet $n = 0$ gælder, da den tomme matrix både er invertibel og øvre triangulær. Så vi antager, at (2) gælder for $n = r - 1$ og viser, at (2) gælder for $n = r$. Vi vælger en egenvektor $\boldsymbol{v}_1$ hørende til λ_1 og supplerer ved hjælp af sætning 4.2.19 til en basis $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ for $\mathbb{F}^n$. Matricen $Q = \begin{pmatrix} \boldsymbol{v}_1 & \ldots & \boldsymbol{v}_n \end{pmatrix}$ er da invertibel (bemærkning 4.2.6), og

$$B = Q^{-1}AQ = \left(\begin{array}{c|ccc} \lambda_1 & b_{12} & \cdots & b_{1p} \\ \hline 0 & & & \\ \vdots & & B_{11} & \\ 0 & & & \end{array}\right)$$

fordi $Q^{-1}AQ\boldsymbol{e}_1 = Q^{-1}A\boldsymbol{v}_1 = Q^{-1}\boldsymbol{v}_1\lambda_1 = \boldsymbol{e}_1\lambda_1$. Ifølge sætning 5.3.6 er $\chi_B(t) = \chi_A(t)$, og ved Laplace udvikling af determinanten efter første søjle får vi endvidere, at

$$\chi_B(t) = \det(B - I_r t) = (\lambda_1 - t)\det(B_{11} - I_{r-1}t) = (\lambda_1 - t)\chi_{B_{11}}(t).$$

Per entydighed af polynomiumsdivision[1] følger derfor, at

$$\chi_{B_{11}}(t) = (\lambda_2 - t) \cdot \ldots \cdot (\lambda_r - t).$$

Den induktive antagelse, at (2) gælder for $n = r - 1$, viser nu, at der findes en invertibel matrix $R_{11} \in M_{r-1}(\mathbb{F})$, sådan at matricen $C_{11} = R_{11}^{-1}B_{11}R_{11} \in M_{r-1}(\mathbb{F})$ er øvre triangulær. Vi lader da

$$R = \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & R_{11} & \\ 0 & & & \end{array}\right) \in M_r(\mathbb{F})$$

[1] Dette følger fra Euklid's algoritme. Et bevis kan findes på side 173–174 i Serge Lang, Algebra. Revised third edition. Graduate Texts in Mathematics, 211. Springer-Verlag, New York, 2002.

og $P = QR \in M_r(\mathbb{F})$. Da er matricen

$$C = P^{-1}AP = R^{-1}Q^{-1}AQR = R^{-1}BR = \left(\begin{array}{c|ccc} \lambda_1 & c_{12} & \cdots & c_{1r} \\ \hline 0 & & & \\ \vdots & & C_{11} & \\ 0 & & & \end{array}\right) \in M_r(\mathbb{F})$$

øvre triangulær, hvilket viser induktionsskridtet og derfor (2). Endelig følger den omvendte implikation, at (2) medfører (1), fra sætning 5.3.6 og sætning 3.3.3. □

Vi bemærker, at algebraens fundamentalsætning (sætning 3.5.7) medfører, at den første betingelse (1) i sætning 5.2.7 altid er opfyldt for $\mathbb{F} = \mathbb{C}$. Dette gælder dog ikke nødvendigs for $\mathbb{F} = \mathbb{R}$. Forskellen er afspejlet i det følgende resultat.

Sætning 5.2.8 (1) *Hvis $A \in M_n(\mathbb{R})$ er en symmetrisk reel kvadratisk matrix, så findes der en invertibel reel matrix $P \in M_n(\mathbb{R})$, sådan at $P^{-1}AP \in M_n(\mathbb{R})$ er øvre triangulær.*
(2) *Hvis $A \in M_n(\mathbb{C})$ er en vilkårlig kompleks kvadratisk matrix, så findes der en invertibel komplex matrix $P \in M_n(\mathbb{C})$, sådan at $P^{-1}AP \in M_n(\mathbb{C})$ er øvre triangulær.*

Bevis Vi begynder med at vise (2). Ifølge korollar 3.5.8 er ethvert polynomium med komplekse koefficienter et produkt af førstegradspolynomier. Dette gælder specielt for det karakteristiske polynomium $\chi_A(t)$, og udsagnet (2) følger derfor fra sætning 5.2.7.

Vi viser dernæst (1). Vi lader $B \in M_n(\mathbb{C})$ være matricen $A \in M_n(\mathbb{R})$ betragtet som en kompleks matrix. Det følger som før fra korollar 3.5.8, at

$$\chi_B(t) = (\lambda_1 - t) \cdot \ldots \cdot (\lambda_n - t) \in \mathbb{C}[t],$$

hvor $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ er egenværdierne af B. Vi påstår, at disse egenværdier alle er reelle tal. Da den reelle matrix A er symmetrisk, er den komplekse matrix B hermitisk, sådan at $B = B^*$. Lad nu $\lambda \in \mathbb{C}$ være en egenværdi for B, og lad $\mathbf{0} \neq \boldsymbol{z} \in \mathbb{C}^n$ være en dertil hørende egenvektor. Da er $B\boldsymbol{z} = \boldsymbol{z}\lambda$, og derfor er $\boldsymbol{z}^*B^* = \lambda^*\boldsymbol{z}^*$, hvilket medfører, at

$$\lambda^*\boldsymbol{z}^*\boldsymbol{z} = \boldsymbol{z}^*B^*\boldsymbol{z} = \boldsymbol{z}^*B\boldsymbol{z} = \boldsymbol{z}^*\boldsymbol{z}\lambda,$$

idet $B^* = B$. Men $\boldsymbol{z} \neq \mathbf{0}$, så $\boldsymbol{z}^*\boldsymbol{z} \neq 0$, og vi konkluderer derfor, at $\lambda^* = \lambda$, hvilket viser påstanden. Polynomiumsidentiteten ovenfor viser derfor, at

$$\chi_A(t) = (\lambda_1 - t) \cdot \ldots \cdot (\lambda_n - t) \in \mathbb{R}[t]$$

med $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$, og derfor følger (1) ved at anvende sætning 5.2.7 som før. □

5.3 Egenværdier og egenrum for lineære endomorfier

Lad $\mathbb{F}$ være et legeme. Vi siger, at en lineær afbildning $f: V \to V$, hvis domæne og codomæne er det samme $\mathbb{F}$-vektorrum V, er en *lineær endomorfi* af $\mathbb{F}$-vektorrummet V. Hvis $\mathbb{F}$-vektorrummet V er endeligt dimensionalt med basis $(\boldsymbol{v}_1,\dots,\boldsymbol{v}_n)$, da er enhver lineær endomorfi $f: V \to V$ entydigt bestemt ved den kvadratiske matrix $A \in M_n(\mathbb{F})$, der repræsenterer $f: V \to V$ med hensyn til basen $(\boldsymbol{v}_1,\dots,\boldsymbol{v}_n)$ for både domænet og codomænet; se definition 4.3.4. Hvis $(\boldsymbol{w}_1,\dots,\boldsymbol{w}_n)$ ligeledes er en basis for V, og hvis matricen $B \in M_n(\mathbb{F})$ repræsenterer $f: V \to V$ med hensyn til basen $(\boldsymbol{w}_1,\dots,\boldsymbol{w}_n)$ for både domænet og codomænet, da gælder ifølge sætning 4.3.11, at

$$B = P^{-1}AP,$$

hvor P er matricen, der repræsenterer identitetsafbildningen $\mathrm{id}_V: V \to V$ med hensyn til baserne $(\boldsymbol{w}_1,\dots,\boldsymbol{w}_n)$ for domænet og $(\boldsymbol{v}_1,\dots,\boldsymbol{v}_n)$ for codomænet. Den følgende figur illustrerer koordinatskiftet mellem de to baser.

$$\begin{tikzcd}
(\boldsymbol{v}_1,\dots,\boldsymbol{v}_n) \arrow[rrr, dotted, bend left, "A"] & & & (\boldsymbol{v}_1,\dots,\boldsymbol{v}_n) \\
& V \arrow[r, "f"] & V & \\
& V \arrow[r, "f"] \arrow[u, "\mathrm{id}_V"'] & V \arrow[u, "\mathrm{id}_V"'] & \\
(\boldsymbol{w}_1,\dots,\boldsymbol{w}_n) \arrow[uuu, dotted, bend left, "P"] \arrow[rrr, dotted, bend right, "B"'] & & & (\boldsymbol{w}_1,\dots,\boldsymbol{w}_n) \arrow[uuu, dotted, bend right, "P"']
\end{tikzcd}$$

Ifølge sætning 5.3.6 er de karakteristiske polynomier $\chi_A(t)$ og $\chi_B(t)$ derfor identitiske, hvilket viser, at den følgende definition er meningsfuld.

Definition 5.3.1 Lad $\mathbb{F}$ være et legeme, og lad V være et $\mathbb{F}$-vektorrum V af endelig dimension n. Det *karakteristiske polynomium* af en lineær endomorfi $f: V \to V$ er da det karakteristiske polynomium

$$\chi_f(t) = \chi_A(t) = \det(A - It) \in \mathbb{F}[t]$$

af den kvadratiske matrix $A \in M_n(\mathbb{F})$, der repræsenterer $f: V \to V$ med hensyn til en fælles basis $(\boldsymbol{v}_1,\dots,\boldsymbol{v}_n)$ for domænet og codomænet.

Det er afgørende i definition 5.3.1, at vi anvender den *samme* basis $(\boldsymbol{v}_1,\dots,\boldsymbol{v}_n)$ for både domænet og codomænet af endomorfien $f\colon V\to V$, men derudover gør det ingen forskel, hvilken basis vi anvender.

Lad nu V være et vilkårligt $\mathbb{F}$-vektorrum. Hvis $\lambda\in\mathbb{F}$, da definerer formlen

$$(\mathrm{id}_V\,\lambda)(\boldsymbol{v})=\boldsymbol{v}\lambda$$

en lineær endomorfi $\mathrm{id}_V\,\lambda\colon V\to V$. For hvis $\boldsymbol{v},\boldsymbol{w}\in V$ og $\mu\in\mathbb{F}$, da gælder

$$\begin{aligned}(\mathrm{id}_V\,\lambda)(\boldsymbol{v}+\boldsymbol{w})&=(\boldsymbol{v}+\boldsymbol{w})\lambda=\boldsymbol{v}\lambda+\boldsymbol{w}\lambda=(\mathrm{id}_V\,\lambda)(\boldsymbol{v})+(\mathrm{id}_V\,\lambda)(\boldsymbol{w})\\(\mathrm{id}_V\,\lambda)(\boldsymbol{v}\mu)&=(\boldsymbol{v}\mu)\lambda=\boldsymbol{v}(\mu\lambda)=\boldsymbol{v}(\lambda\mu)=(\boldsymbol{v}\lambda)\mu=(\mathrm{id}_V\,\lambda)(\boldsymbol{v})\mu,\end{aligned}$$

hvor vi har brugt den kommutative lov for multiplikation af skalarer. Specielt gælder der, at for enhver lineær endomorfi $f\colon V\to V$ og $\lambda\in\mathbb{F}$, er

$$V_\lambda=\{\boldsymbol{v}\in V\mid f(\boldsymbol{v})=\boldsymbol{v}\lambda\}=\ker(f-\mathrm{id}_V\,\lambda)\subset V$$

et underrum af vektorrummet V.

Definition 5.3.2 Lad V være et vektorrum over et legeme $\mathbb{F}$, og lad $f\colon V\to V$ være en lineær endomorfi.

(1) En skalar $\lambda\in\mathbb{F}$ er en *egenværdi* for $f\colon V\to V$, hvis $V_\lambda\neq\{\mathbf{0}\}$.

(2) Hvis $\lambda\in\mathbb{F}$ er en egenværdi for $f\colon V\to V$, så kaldes underrummet $V_\lambda\subset V$ for *egenrummet* for $f\colon V\to V$ hørende til egenværdien $\lambda\in\mathbb{F}$.

(3) Hvis $\lambda\in\mathbb{F}$ er en egenværdi for $f\colon V\to V$, så kaldes enhver vektor $\boldsymbol{v}\in V_\lambda$, der ikke er nulvektoren, for en *egenvektor* for $f\colon V\to V$ hørende til egenværdien $\lambda\in\mathbb{F}$.

Eksempel 5.3.3 (1) Nulafbildningen $\mathbf{0}\colon V\to V$ har $\lambda=0$ som eneste egenværdi, og

$$V_0=\ker(\mathbf{0}-\mathrm{id}_V\,0)=\ker(\mathbf{0})=V.$$

(2) Identitetsafbildningen $\mathrm{id}_V\colon V\to V$ har $\lambda=1$ som eneste egenværdi, og

$$V_1=\ker(\mathrm{id}_V-\mathrm{id}_V\,1)=\ker(\mathbf{0})=V.$$

(3) Mængden $V=C^\infty(\mathbb{R})$ af funktioner $f\colon\mathbb{R}\to\mathbb{R}$, der er vilkårligt ofte differentiable, har en reel vektorrumsstruktur, hvor vektorsum og skalarmultiplikation er givet ved henholdsvis $(f+g)(x)=f(x)+g(x)$ og $(f\cdot\lambda)(x)=f(x)\cdot\lambda$, og afbildningen $D\colon V\to V$, der til $f\in V$ tilordner den afledede $f'\in V$, er en lineær endomorfi. Ethvert $\lambda\in\mathbb{R}$ er en egenværdi for $D\colon V\to V$, og det tilhørende egenrum er det 1-dimensionale underrum

$$V_\lambda=\{e^{\lambda x}\cdot c\mid c\in\mathbb{R}\}\subset V,$$

da funktionerne $f(x) = e^{\lambda x} \cdot c$ udgør alle løsninger til differentialligningen $Df = f\lambda$. Så endomorfien $D\colon V \to V$ har specielt uendeligt mange egenværdier. Sætning 5.1.8 viser, at dette kun er muligt, fordi vektorrummet V er uendeligt dimensionalt.

Vi skal vise, at egenværdierne af en lineær endomorfi af et endeligt frembragt $\mathbb{F}$-vektorrum præcis er rødderne i $\mathbb{F}$ af dets karakteristiske polynomium. Vi viser først et helt generelt resultat.

Lemma 5.3.4 *Lad $\mathbb{F}$ være et legeme, lad V være et endeligt dimensionalt $\mathbb{F}$-vektorrum, og lad $g\colon V \to V$ være en lineære endomorfi. Lad $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ være en basis for V, lad $h\colon \mathbb{F}^n \to V$ være den lineære isomorfi givet ved $h(\boldsymbol{x}) = \boldsymbol{v}_1 x_1 + \dots + \boldsymbol{v}_n x_n$, og lad $B \in M_n(\mathbb{F})$ være matricen, der repræsenterer $g\colon V \to V$ med hensyn til denne basis for både domænet og codomænet. Da gælder, at $h(N_B) = \ker(g)$.*

Bevis Vi bemærker, at sætning 4.3.3 udtrykker, at der for alle $\boldsymbol{x} \in \mathbb{F}^n$ gælder, at

$$g(h(\boldsymbol{x})) = h(B\boldsymbol{x}).$$

Dette medfører, at $h(N_B) \subset \ker(g)$. For hvis nu $\boldsymbol{x} \in N_B$, da er $B\boldsymbol{x} = \boldsymbol{0}$, og derfor viser denne identitet, at også $g(h(\boldsymbol{x})) = h(B\boldsymbol{x}) = h(\boldsymbol{0}) = \boldsymbol{0}$. For at vise, at også $\ker(g) \subset h(N_B)$, lader vi $\boldsymbol{v} \in \ker(g)$ og skriver $\boldsymbol{v} = h(\boldsymbol{x})$ for et entydigt $\boldsymbol{x} \in \mathbb{F}^n$. Da er $h(B\boldsymbol{x}) = g(h(\boldsymbol{x})) = g(\boldsymbol{v}) = \boldsymbol{0}$, og da $h\colon \mathbb{F}^n \to V$ er injektiv, følger det, at $\boldsymbol{x} = \boldsymbol{0}$ som ønkset. □

Korollar 5.3.5 *Lad $\mathbb{F}$ være et legeme, lad V være et endeligt dimensionalt $\mathbb{F}$-vektorrum, og lad $f\colon V \to V$ være en lineære endomorfi. Lad $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ være en basis for V, lad $h\colon \mathbb{F}^n \to V$ være den lineære isomorfi givet ved $h(\boldsymbol{x}) = \boldsymbol{v}_1 x_1 + \dots + \boldsymbol{v}_n x_n$, og lad $A \in M_n(\mathbb{F})$ være matricen, der repræsenterer $f\colon V \to V$ med hensyn til denne basis for både domænet og codomænet. Da gælder for alle $\lambda \in \mathbb{F}$, at*

$$h(N_{A-I\lambda}) = V_\lambda.$$

Specielt er λ er en egenværdi for $f\colon V \to V$, hvis og kun hvis λ er en egenværdi for A.

Bevis Identiteten $h(N_{A-I\lambda}) = V_\lambda$ følger fra lemma 5.3.4 med $g = f - \mathrm{id}_V \lambda\colon V \to V$. For matricen B, der repræsenterer $g\colon V \to V$ med hensyn til den fælles basis $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ for domænet og codomænet, er $B = A - I\lambda$, så $N_B = N_{A-I\lambda}$, og $\ker(g) = \ker(f - \mathrm{id}_V \lambda) = V_\lambda$.

Endelig er λ per definition en egenværdi for $f\colon V \to V$, hvis og kun hvis $V_\lambda \neq \{\boldsymbol{0}\}$, og da $h\colon \mathbb{F}^n \to V$ er en isomorfi, så viser identiteten ovenfor, at dette er tilfældet, hvis og kun hvis $N_{A-I\lambda} \neq \{\boldsymbol{0}\}$, hvilket er definitionen på, at λ er en egenværdi for A. □

Sætning 5.3.6 *Lad $\mathbb{F}$ være et legeme og V et endeligt dimensionalt $\mathbb{F}$-vektorrum. Da er egenværdierne af en lineær endomorfi $f\colon V \to V$ netop rødderne i $\mathbb{F}$ af det karakteristiske polynomium $\chi_f(t) \in \mathbb{F}[t]$.*

Bevis Lad $A \in M_n(\mathbb{F})$ være matricen, der repræsenterer $f: V \to V$ med hensyn til en fælles basis $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ for domænet og codomænet. Ifølge korollar 5.3.5 er $\lambda \in \mathbb{F}$ en egenværdi for $f: V \to V$, hvis og kun hvis $\lambda \in \mathbb{F}$ er en egenværdi for A, og ifølge sætning 5.1.6 er dette tilfældet, hvis og kun hvis $\lambda \in \mathbb{F}$ er en rod i $\chi_A(t) = \chi_f(t)$. □

Hvis $\lambda \in \mathbb{F}$ er en egenværdi for den lineære endomorfi $f: V \to V$, så definerer vi igen den *geometriske multiplicitet* til at være dimensionen af egenrummet V_λ hørende til λ; og vi definerer den *algebraiske multiplicitet* til at være det maksimale hele tal $k \geq 0$, sådan at der gælder $\chi_f(t) = (\lambda - t)^k \cdot q(t) \in \mathbb{F}[t]$ for et $q(t) \in \mathbb{F}[t]$.

Sætning 5.3.7 *Lad $\mathbb{F}$ være et legeme og V et endeligt dimensionalt $\mathbb{F}$-vektorrum. Givet en egenværdi $\lambda \in \mathbb{F}$ for en lineær endomorfi $f: V \to V$, da er den algebraiske multiplicitet af λ større end eller lig med den geometriske multiplicitet af λ.*

Bevis Vi vælger først en basis $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ for egenrummet V_λ og supplerer dernæst denne til en basis $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n, \boldsymbol{w}_1, \dots, \boldsymbol{w}_p)$ for V. Da vektorerne $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$ er egenvektorer for $f: V \to V$ hørende til λ, er matricen $A \in M_{n+p}(\mathbb{F})$, der repræsenterer $f: V \to V$ med hensyn til basen $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n, \boldsymbol{w}_1, \dots, \boldsymbol{w}_p)$ for både domænet og codomænet, på formen

$$A = \left(\begin{array}{c|c} I_n\lambda & B \\ \hline O_{p,n} & C \end{array}\right),$$

hvor $I_n \in M_{n,n}(\mathbb{F})$ er identitetsmatricen og $O_{p,n} \in M_{p,n}(\mathbb{F})$ er nulmatricen, og hvor de to matricer $B \in M_{n,p}(\mathbb{F})$ og $C \in M_{p,p}(\mathbb{F})$ kan være vilkårlige. Ved at gentage Laplace udvikling af determinanten efter første søjle n gange får vi da, at

$$\chi_f(t) = \det(A - I_{n+p}t) = \left(\begin{array}{c|c} I_n\lambda - I_n t & B \\ \hline O_{p,n} & C - I_p t \end{array}\right) = (\lambda - t)^n \det(C - I_p t).$$

Dette viser, at den algebraiske multiplicitet af λ større end eller lig med den geometriske multiplicitet n som ønsket. □

Eksempel 5.3.8 Vi udregner egenværdierne og de tilhørende egenrum for den lineære endomorfi $f: \mathbb{F}^3 \to \mathbb{F}^3$ defineret ved $f(\boldsymbol{x}) = A\boldsymbol{x}$, hvor

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

Vi begynder med at udregne det karakteristiske polynomium.

$$
\begin{aligned}
\chi_f(t) &= \det\begin{pmatrix} -1-t & 0 & -2 \\ 3 & 2-t & 2 \\ 1 & -1 & 3-t \end{pmatrix} \overset{(D6')}{=} \det\begin{pmatrix} 2-t & 2-t & 0 \\ 3 & 2-t & 2 \\ 1 & -1 & 3-t \end{pmatrix} \\
&\overset{(D2')}{=} (2-t)\det\begin{pmatrix} 1 & 1 & 0 \\ 3 & 2-t & 2 \\ 1 & -1 & 3-t \end{pmatrix} \overset{(D6')}{=} (2-t)\det\begin{pmatrix} 1 & 1 & 0 \\ 0 & -1-t & 2 \\ 0 & -2 & 3-t \end{pmatrix} \\
&\overset{(D6)}{=} (2-t)\det\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1-t & 2 \\ 0 & -2 & 3-t \end{pmatrix} \overset{(D6)}{=} (2-t)\det\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1-t & 2 \\ 0 & 1-t & 3-t \end{pmatrix} \\
&\overset{(D2)}{=} (2-t)(1-t)\det\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 1 & 3-t \end{pmatrix} \overset{(D6)}{=} (2-t)(1-t)\det\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1-t \end{pmatrix} \\
&= (2-t)(1-t)^2.
\end{aligned}
$$

Vi ser herfra, at $\lambda_1 = 1$ og $\lambda_2 = 2$ er de eneste egenværdier af A, og at deres algebraiske multipliciteter er henholdsvis 2 og 1.

For at udregne egenrummet N_{A-I} hørende til $\lambda_1 = 1$ anvender vi rækkeoperationer til at omdanne $B = A - I$ til en matrix B' på reduceret echelon-form.

$$
\begin{aligned}
B = &\begin{pmatrix} -2 & 0 & -2 \\ 3 & 1 & 2 \\ 1 & -1 & 2 \end{pmatrix} \quad \begin{matrix} +2\cdot R_3 \\ +(-3)\cdot R_3 \\ \end{matrix} \\
&\begin{pmatrix} 0 & -2 & 2 \\ 0 & 4 & -4 \\ 1 & -1 & 2 \end{pmatrix} \quad R_1 \leftrightarrow R_3 \\
&\begin{pmatrix} 1 & -1 & 2 \\ 0 & 4 & -4 \\ 0 & -2 & 2 \end{pmatrix} \quad \tfrac{1}{4}\cdot R_2 \\
&\begin{pmatrix} 1 & -1 & 2 \\ 0 & 1 & -1 \\ 0 & -2 & 2 \end{pmatrix} \quad \begin{matrix} +R_2 \\ \\ +3R_2 \end{matrix} \\
B' = &\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}
\end{aligned}
$$

Vi ser heraf, at egenrummet hørende til $\lambda_1 = 1$ er givet ved

$$
N_{A-I} = \left\{ \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} t \;\middle|\; t \in \mathbb{F} \right\} \subset \mathbb{F}^3,
$$

og vi ser tilsvarende, at egenrummet N_{A-I2} hørende til $\lambda_2 = 2$ er givet ved

$$N_{A-I2} = \left\{ \begin{pmatrix} -2 \\ 1 \\ 3 \end{pmatrix} t \;\middle|\; t \in \mathbb{F} \right\} \subset \mathbb{F}^3.$$

Den geometriske multiplicitet af både $\lambda = 1$ og $\lambda = 2$ altså lig med 1.

6 Vektorrum med indre produkt

Dette kapitel omhandler geometriske strukturer på vektorrum såsom længden af en vektor og vinklen mellem to vektorer. Disse strukturer er imidlertid ikke lineære men derimod kvadratiske af natur. De er derfor ikke udtrykt ved vektorrummet alene, men afhænger af et valg af en yderligere struktur, der kaldes for et *indre produkt*, som vi definerer først. I dette kapitel antager vi, at $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, men i de tre første afsnit gælder alle udsagn og definitioner også for $\mathbb{F} = \mathbb{H}$ med kvaternionisk konjugation.

6.1 Indre produkt

Vi betragter vektorrum over $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, og for $a \in \mathbb{F}$ definerer vi

$$a^* = \begin{cases} a, & \text{hvis } \mathbb{F} = \mathbb{R}, \\ \bar{a}, & \text{hvis } \mathbb{F} = \mathbb{C}. \end{cases}$$

Vi bemærker, at hvis $a \in \mathbb{F}$ opfylder, at $a = a^*$, så er $a \in \mathbb{R}$, og i dette tilfælde er udsagnet "$a > 0$" derfor meningsfuldt.

Definition 6.1.1 Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$. Et *indre produkt* på et $\mathbb{F}$-vektorrum V er en afbildning $\langle -,- \rangle \colon V \times V \to \mathbb{F}$, der opfylder følgende:

(H1) For alle $\boldsymbol{u},\boldsymbol{v},\boldsymbol{w} \in V$ er $\langle \boldsymbol{u}, \boldsymbol{v}+\boldsymbol{w} \rangle = \langle \boldsymbol{u},\boldsymbol{v} \rangle + \langle \boldsymbol{u},\boldsymbol{w} \rangle$.

(H2) For alle $\boldsymbol{u},\boldsymbol{v} \in V$ og $a \in \mathbb{F}$ er $\langle \boldsymbol{u}, \boldsymbol{v}a \rangle = \langle \boldsymbol{u},\boldsymbol{v} \rangle a$.

(H3) For alle $\boldsymbol{u},\boldsymbol{v},\boldsymbol{w} \in V$ og $a \in \mathbb{F}$ er $\langle \boldsymbol{u}+\boldsymbol{v}, \boldsymbol{w} \rangle = \langle \boldsymbol{u},\boldsymbol{w} \rangle + \langle \boldsymbol{v},\boldsymbol{w} \rangle$.

(H4) For alle $\boldsymbol{u},\boldsymbol{v} \in V$ og $a \in \mathbb{F}$ er $\langle \boldsymbol{u}a, \boldsymbol{v} \rangle = a^* \langle \boldsymbol{u},\boldsymbol{v} \rangle$.

(H5) For alle $\boldsymbol{u},\boldsymbol{v} \in V$ er $\langle \boldsymbol{v}, \boldsymbol{u} \rangle = \langle \boldsymbol{u},\boldsymbol{v} \rangle^*$.

(P) For alle $\boldsymbol{0} \neq \boldsymbol{u} \in V$ er $\langle \boldsymbol{u}, \boldsymbol{u} \rangle > 0$.

Et $\mathbb{F}$-*vektorrum med indre produkt* er et par $(V, \langle -,- \rangle)$ bestående af et $\mathbb{F}$-vektorrum og et indre produkt $\langle -,- \rangle \colon V \times V \to \mathbb{F}$.

Vi minder om, at vi i afsnit 2.6 har defineret en hermitisk form på V til at være en afbildning $\langle -,- \rangle : V \times V \to \mathbb{F}$, der opfylder (H1)–(H5), idet vi dog kun betragtede $V = \mathbb{F}^n$. Mens hermitiske former er defineret generelt for $\mathbb{F}$ et skævlegeme med skævinvolution, så giver positivitetsbetingelsen (P) i definitionen af et indre produkt kun mening, hvis $\mathbb{F}$ er $\mathbb{R}$, $\mathbb{C}$ eller $\mathbb{H}$ med de respektive konjugeringsafbildninger som skævinvolution.

Eksempel 6.1.2 (1) Standard-indreproduktet $\langle -,- \rangle : \mathbb{F}^n \times \mathbb{F}^n \to \mathbb{F}$ defineret ved

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \boldsymbol{x}^* \boldsymbol{y} = x_1^* y_1 + \cdots + x_n^* y_n$$

er et indre produkt på $\mathbb{F}^n$.
(2) Vi lader $A = A^* \in M_n(\mathbb{F})$ være er en hermitisk matrix og husker fra sætning 2.6.7, at afbildningen $\langle -,- \rangle : \mathbb{F}^n \times \mathbb{F}^n \to \mathbb{F}$ defineret ved

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \boldsymbol{x}^* A \boldsymbol{y}$$

er en hermitisk form, hvilket vil sige, at den opfylder (H1)–(H5). Hvis denne afbildning også opfylder (P), så siger vi, at den hermitiske matrix $A \in M_m(\mathbb{F})$ er *positiv definit*. En diagonal matrix $A = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$ er således hermitisk, hvis og kun hvis $\lambda_1, \dots, \lambda_n$ alle er reelle tal, og den er hermitisk og positiv definit, hvis og kun hvis $\lambda_1, \dots, \lambda_n$ alle er positive reelle tal. For eksempel er $A = \operatorname{diag}(-1, 1, \dots, 1)$, der svarer til den Minkowski hermitiske form, ikke positiv definit.
(3) Lad $V = C^0([a,b])$ være det uendeligtdimensionale reelle vektorrum af kontinuerte funktioner $f : [a,b] \to \mathbb{R}$ fra bemærkning 4.3.16. Da definerer formlen

$$\langle f, g \rangle = \int_a^b f(x) g(x) dx$$

et indre produkt $\langle -,- \rangle : V \times V \to \mathbb{R}$.

Definition 6.1.3 Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, og lad V være et $\mathbb{F}$-vektorrum. Givet et indre produkt $\langle -,- \rangle : V \times V \to \mathbb{F}$, da kaldes afbildningen $\|-\| : V \to \mathbb{R}$ defineret ved

$$\|\boldsymbol{v}\| = \sqrt{\langle \boldsymbol{v}, \boldsymbol{v} \rangle}$$

for *normen* hørende til $\langle -,- \rangle$.

Vi kalder også $\|\boldsymbol{v}\|$ for *længden* af $\boldsymbol{v}$ med hensyn til $\langle -,- \rangle$. Vi bemærker, at normen af $\boldsymbol{v}$ er veldefineret, fordi $\langle \boldsymbol{v}, \boldsymbol{v} \rangle$ er et ikke-negativt reelt tal, og at normen af $\boldsymbol{v}$ selv er et ikke-negativt reelt tal. Der gælder endvidere, at

$$\|\boldsymbol{v} a\| = \|\boldsymbol{v}\| |a|$$

for alle $\boldsymbol{v} \in V$ og $a \in \mathbb{F}$. Her er $|a| = \sqrt{a^* a} \in \mathbb{R}$ absolutværdien af $a \in \mathbb{F}$.

Vi viser nu, at absolutværdien af det indre produkt af to vektorer altid er mindre eller lig med produktet af deres norm. Denne ulighed, der kendes som *Cauchy-Schwarz' ulighed* og skyldes Cauchy, Schwarz og Bunyakovsky, gælder i ethvert vektorrum med indre produkt. Det er en enormt anvendelig ulighed, hvilket vi illustrerer med beviset for trekantsuligheden nedenfor.

Sætning 6.1.4 (Cauchy-Schwarz) *Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, og lad $(V, \langle -,- \rangle)$ være et $\mathbb{F}$-vektorrum med indre produkt. For alle $\boldsymbol{v}, \boldsymbol{w} \in V$ gælder, at*

$$|\langle \boldsymbol{v}, \boldsymbol{w} \rangle| \leq \|\boldsymbol{v}\| \|\boldsymbol{w}\|.$$

Bevis Hvis $\boldsymbol{v} = 0$, er påstanden triviel, så vi antager, at $\boldsymbol{v} \neq \boldsymbol{0}$. Beviset i dette tilfælde beror på følgende snedige trik: Vi betragter vektoren

$$\boldsymbol{z} = \mathrm{pr}_{\boldsymbol{v}}(\boldsymbol{w}) = \boldsymbol{v} \cdot \frac{\langle \boldsymbol{v}, \boldsymbol{w} \rangle}{\langle \boldsymbol{v}, \boldsymbol{v} \rangle},$$

som vi kalder for projektionen af $\boldsymbol{w}$ på $\boldsymbol{v}$. (Se også definition 6.2.7.) For udregningen

$$\begin{aligned} \langle \boldsymbol{z}, \boldsymbol{w} - \boldsymbol{z} \rangle &= \left\langle \boldsymbol{v} \cdot \frac{\langle \boldsymbol{v}, \boldsymbol{w} \rangle}{\langle \boldsymbol{v}, \boldsymbol{v} \rangle},\ \boldsymbol{w} - \boldsymbol{v} \cdot \frac{\langle \boldsymbol{v}, \boldsymbol{w} \rangle}{\langle \boldsymbol{v}, \boldsymbol{v} \rangle} \right\rangle \\ &= \left(\frac{\langle \boldsymbol{v}, \boldsymbol{w} \rangle}{\langle \boldsymbol{v}, \boldsymbol{v} \rangle} \right)^* \cdot \langle \boldsymbol{v}, \boldsymbol{w} \rangle \ - \ \left(\frac{\langle \boldsymbol{v}, \boldsymbol{w} \rangle}{\langle \boldsymbol{v}, \boldsymbol{v} \rangle} \right)^* \cdot \langle \boldsymbol{v}, \boldsymbol{v} \rangle \cdot \frac{\langle \boldsymbol{v}, \boldsymbol{w} \rangle}{\langle \boldsymbol{v}, \boldsymbol{v} \rangle} = 0 \end{aligned}$$

viser, at $\boldsymbol{z}$ og $\boldsymbol{w} - \boldsymbol{z}$ er ortogonale; se definition 6.2.1 nedenfor. Vi skriver nu $\boldsymbol{w}$ som $\boldsymbol{z} + (\boldsymbol{w} - \boldsymbol{z})$ og udregner

$$\langle \boldsymbol{w}, \boldsymbol{w} \rangle = \big\langle \boldsymbol{z} + (\boldsymbol{w} - \boldsymbol{z}), \boldsymbol{z} + (\boldsymbol{w} - \boldsymbol{z}) \big\rangle = \langle \boldsymbol{z}, \boldsymbol{z} \rangle + \langle \boldsymbol{w} - \boldsymbol{z}, \boldsymbol{w} - \boldsymbol{z} \rangle \geq \langle \boldsymbol{z}, \boldsymbol{z} \rangle,$$

hvor uligheden til højre gælder fordi $\langle \boldsymbol{w} - \boldsymbol{z}, \boldsymbol{w} - \boldsymbol{z} \rangle \geq 0$. Vi bemærker, at uligheden giver mening, fordi begge sider af uligheden er reelle tal. Vi udregner endelig, at

$$\langle \boldsymbol{z}, \boldsymbol{z} \rangle = \left\langle \boldsymbol{v} \cdot \frac{\langle \boldsymbol{v}, \boldsymbol{w} \rangle}{\langle \boldsymbol{v}, \boldsymbol{v} \rangle}, \boldsymbol{v} \cdot \frac{\langle \boldsymbol{v}, \boldsymbol{w} \rangle}{\langle \boldsymbol{v}, \boldsymbol{v} \rangle} \right\rangle = \left(\frac{\langle \boldsymbol{v}, \boldsymbol{w} \rangle}{\langle \boldsymbol{v}, \boldsymbol{v} \rangle} \right)^* \cdot \langle \boldsymbol{v}, \boldsymbol{v} \rangle \cdot \frac{\langle \boldsymbol{v}, \boldsymbol{w} \rangle}{\langle \boldsymbol{v}, \boldsymbol{v} \rangle} = \frac{\langle \boldsymbol{v}, \boldsymbol{w} \rangle^*}{\langle \boldsymbol{v}, \boldsymbol{v} \rangle} \cdot \langle \boldsymbol{v}, \boldsymbol{w} \rangle = \frac{|\langle \boldsymbol{v}, \boldsymbol{w} \rangle|^2}{\langle \boldsymbol{v}, \boldsymbol{v} \rangle},$$

idet $\langle \boldsymbol{v}, \boldsymbol{v} \rangle^* = \langle \boldsymbol{v}, \boldsymbol{v} \rangle$. Så uligheden ovenfor er derfor ækvivalent med uligheden

$$\langle \boldsymbol{v}, \boldsymbol{v} \rangle \cdot \langle \boldsymbol{w}, \boldsymbol{w} \rangle \geq |\langle \boldsymbol{v}, \boldsymbol{w} \rangle|^2,$$

fordi $\langle \boldsymbol{v}, \boldsymbol{v} \rangle > 0$. Sætningen følger nu ved at uddrage kvadratroden på begge sider. □

Lad V være et $\mathbb{F}$-vektorrum med indre produkt $\langle -,- \rangle$ og tilhørende norm $\|-\|$. Givet to vektorer $\boldsymbol{v}$ og $\boldsymbol{w}$ i V, kan vi danne trekanten med hjørner $\boldsymbol{0}$, $\boldsymbol{v}$ og $\boldsymbol{v}+\boldsymbol{w}$, og vi kan betragte $\|\boldsymbol{v}\|$, $\|\boldsymbol{w}\|$ og $\|\boldsymbol{v}+\boldsymbol{w}\|$ som længderne af de tre sider i denne trekant. Trekantsudligheden, som vi nu viser, siger altså at længden af én side i en trekant altid er mindre end summen af længderne af de to andre sider.

Sætning 6.1.5 (Trekantsuligheden) *Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, og lad $(V, \langle -,- \rangle)$ være et $\mathbb{F}$-vektorrum med indre produkt. Da gælder for alle $\boldsymbol{v}, \boldsymbol{w} \in V$, at*

$$\|\boldsymbol{v} + \boldsymbol{w}\| \leq \|\boldsymbol{v}\| + \|\boldsymbol{w}\|.$$

Bevis Vi bemærker først, at der for alle $a = x + iy \in \mathbb{C}$ gælder, at

$$a + a^* = (x + iy) + (x - iy) = 2x \leq 2\sqrt{x^2 + y^2} = 2|a|,$$

og at samme ulighed gælder trivielt også for alle $a \in \mathbb{R}$. Vi udregner nu

$$\begin{aligned}\|\boldsymbol{v} + \boldsymbol{w}\|^2 &= \langle \boldsymbol{v} + \boldsymbol{w}, \boldsymbol{v} + \boldsymbol{w} \rangle = \langle \boldsymbol{v}, \boldsymbol{v} \rangle + \langle \boldsymbol{v}, \boldsymbol{w} \rangle + \langle \boldsymbol{w}, \boldsymbol{v} \rangle + \langle \boldsymbol{w}, \boldsymbol{w} \rangle \\ &= \|\boldsymbol{v}\|^2 + \langle \boldsymbol{v}, \boldsymbol{w} \rangle + \langle \boldsymbol{v}, \boldsymbol{w} \rangle^* + \|\boldsymbol{w}\|^2 \leq \|\boldsymbol{v}\|^2 + 2|\langle \boldsymbol{v}, \boldsymbol{w} \rangle| + \|\boldsymbol{w}\|^2 \\ &\leq \|\boldsymbol{v}\|^2 + 2\|\boldsymbol{v}\|\|\boldsymbol{w}\| + \|\boldsymbol{w}\|^2 = (\|\boldsymbol{v}\| + \|\boldsymbol{w}\|)^2,\end{aligned}$$

hvor den første af de to uligheder følger af den ulighed, vi viste i begyndelsen af beviset, og den anden er Cauchy-Schwarz' ulighed. Den ønskede ulighed følger ved at uddrage kvadratroden. □

For *reelle* vektorrum med indre produkt er Cauchy-Schwarz' ulighed ækvivalent med udsagnet, at der for alle vektorer $\boldsymbol{v}$ og $\boldsymbol{w}$ i vektorrummet gælder, at

$$-\|\boldsymbol{v}\|\|\boldsymbol{w}\| \leq \langle \boldsymbol{v}, \boldsymbol{w} \rangle \leq \|\boldsymbol{v}\|\|\boldsymbol{w}\|.$$

Hvis $\boldsymbol{v} \neq \boldsymbol{0}$ og $\boldsymbol{w} \neq \boldsymbol{0}$, så kan vi ækvivalent skrive denne ulighed som

$$-1 \leq \frac{\langle \boldsymbol{v}, \boldsymbol{w} \rangle}{\|\boldsymbol{v}\|\|\boldsymbol{w}\|} \leq 1,$$

fordi $\|\boldsymbol{v}\| > 0$ og $\|\boldsymbol{w}\| > 0$. Vi definerer nu *vinklen* mellem $\boldsymbol{v}$ og $\boldsymbol{w}$ til at være det entydigt bestemte reelle tal $0 \leq \theta \leq \pi$, sådan at

$$\cos\theta = \frac{\langle \boldsymbol{v}, \boldsymbol{w} \rangle}{\|\boldsymbol{v}\|\|\boldsymbol{w}\|}.$$

Vi har altså den sædvanlige sammenhæng mellem vinkler og indre produkt, men vi benytter nu det indre produkt til at definere vinkler og ikke omvendt.

Eksempel 6.1.6 (1) Vi betragter $\mathbb{R}^3$ med standard-indreproduktet og udregner vinklen mellem de to vektorer

$$\boldsymbol{x} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \qquad \text{og} \qquad \boldsymbol{y} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Per definition er denne det entydigt bestemte reelle tal $0 \leq \theta \leq \pi$, sådan at

$$\cos\theta = \frac{\langle \boldsymbol{x}, \boldsymbol{y}\rangle}{\sqrt{\langle \boldsymbol{x}, \boldsymbol{x}\rangle}\sqrt{\langle \boldsymbol{y}, \boldsymbol{y}\rangle}} = \frac{2}{\sqrt{2}\sqrt{3}} = \frac{\sqrt{6}}{2},$$

hvilket approksimativt er $\theta \doteq 35,26°$.

(2) Vi lader $(C^0([0,1]), \langle -,-\rangle)$ være det reelle vektorrum med indre produkt fra eksempel 6.1.2 (3), og lader $f,g\colon [0,1] \to \mathbb{R}$ være de to vektorer heri defineret ved henholdsvis $f(x) = x$ og $g(x) = x^2$. For at bestemme vinklen θ mellem f og g udregner vi:

$$\begin{aligned}
\langle f, f\rangle &= \int_0^1 f(x)f(x)dx = \int_0^1 x^2 dx = \frac{1}{3} \\
\langle f, g\rangle &= \int_0^1 f(x)g(x)dx = \int_0^1 x^3 dx = \frac{1}{4} \\
\langle g, g\rangle &= \int_0^1 g(x)g(x)dx = \int_0^1 x^4 dx = \frac{1}{5}
\end{aligned}$$

Der gælder altså, at

$$\cos\theta = \frac{\langle f, g\rangle}{\sqrt{\langle f, f\rangle}\cdot\sqrt{\langle g, g\rangle}} = \frac{\sqrt{3}\cdot\sqrt{5}}{4} = \frac{\sqrt{15}}{4}$$

og dermed, at $\theta \doteq 14,48°$.

6.2 Ortogonalitet

Mens vinkelbegrebet kun giver mening for reelle vektorrum med indre produkt, så giver ortogonalitet mening for vilkårlige $\mathbb{F}$-vektorrum med indre produkt.

Definition 6.2.1 Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, og lad $(V, \langle -,-\rangle)$ være et $\mathbb{F}$-vektorrum med indre produkt.

(1) En vektor $\boldsymbol{u} \in V$ er en *enhedsvektor* med hensyn til $\langle -,-\rangle$, hvis $\|\boldsymbol{u}\| = 1$.

(2) To vektorer $\boldsymbol{v}, \boldsymbol{w} \in V$ er *ortogonale* med hensyn til $\langle -,-\rangle$, hvis $\langle \boldsymbol{v}, \boldsymbol{w}\rangle = 0$.

(3) En familie $(\boldsymbol{v}_i)_{i\in I}$ af vektorer i V er *ortogonal* med hensyn til $\langle -,-\rangle$, hvis den består af vektorer, der er parvis ortogonale med hensyn til $\langle -,-\rangle$.

(4) En familie $(\boldsymbol{u}_i)_{i\in I}$ af vektorer i V er *ortonormal* med hensyn til $\langle -,-\rangle$, hvis den både er ortogonal med hensyn til $\langle -,-\rangle$ og består af enhedsvektorer med hensyn til $\langle -,-\rangle$.

Vi bemærker, at $\|\boldsymbol{u}\| = 1$, hvis og kun hvis $\langle \boldsymbol{u}, \boldsymbol{u}\rangle = 1$.

Eksempel 6.2.2 (1) Standardbasen $(\boldsymbol{e}_1, \dots, \boldsymbol{e}_m)$ for $\mathbb{F}^m$ er ortonormal med hensyn til standard-indreproduktet, fordi $\langle \boldsymbol{e}_i, \boldsymbol{e}_j \rangle = \delta_{ij}$.
(2) Den følgende familie $(\boldsymbol{v}_1, \boldsymbol{v}_2)$ i $\mathbb{F}^2$ er ortogonal, men ikke ortonormal, med hensyn til standard-indreproduktet.

$$\Big(\boldsymbol{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \ \boldsymbol{v}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}\Big)$$

For selvom $\langle \boldsymbol{v}_1, \boldsymbol{v}_2 \rangle = 0$, så er $\langle \boldsymbol{v}_1, \boldsymbol{v}_1 \rangle = 2 \neq 1$ og $\langle \boldsymbol{v}_2, \boldsymbol{v}_2 \rangle = 2 \neq 1$.
(3) Ved en udregning af integraler ses, at i vektorrummet $C^0([0,2\pi])$ af kontinuerte funktioner $f : [0,2\pi] \to \mathbb{R}$ er familien

$$\Big(\frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}}\cos(x), \frac{1}{\sqrt{\pi}}\sin(x), \frac{1}{\sqrt{\pi}}\cos(2x), \frac{1}{\sqrt{\pi}}\sin(2x), \frac{1}{\sqrt{\pi}}\cos(3x), \frac{1}{\sqrt{\pi}}\sin(3x), \ \dots\Big)$$

ortonormal med hensyn til det indre produkt, som vi definerede i eksempel 6.1.2 (3). Det underrum af $C^0([0,2\pi])$, der er frembragt af denne familie, kaldes for vektorrummet af *trigonometriske polynomier*. Trigonometriske polynomer bruges for eksempel til at modelere signaler.

Lemma 6.2.3 *Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, og lad $(V, \langle -,- \rangle)$ være et $\mathbb{F}$-vektorrum med indre produkt. Enhver ortonormal familie af vektorer i V er lineært uafhængig.*

Bevis Lad $(\boldsymbol{v}_i)_{i\in I}$ være en ortonormal familie af vektorer i V og lad

$$\sum_{i\in I} \boldsymbol{v}_i a_i = \boldsymbol{0}$$

være en linear kombination af $(\boldsymbol{v}_i)_{i\in I}$, der er lig med nulvektoren. (Hvis familien ikke er endelig, så anvendes definitionen givet i bemærkning 4.2.20.) Vi skal da vise, at $a_i = 0$ for alle $i \in I$. Men ligningen ovenfor viser, at

$$\langle \boldsymbol{v}_i, \sum_{j\in I} \boldsymbol{v}_j a_j \rangle = \sum_{j\in I} \langle \boldsymbol{v}_i, \boldsymbol{v}_j \rangle a_j = \sum_{j\in I} \delta_{ij} a_j = a_i$$

er lig med $\langle \boldsymbol{v}_i, \boldsymbol{0} \rangle = 0$ for alle $i \in I$ som ønsket. □

Det følgende simple men enormt anvendelige resultat viser, at hvis en basis for et vektorrum er ortonormal med hensyn til et givet indre produkt, så kan koordinaterne af en vektor med hensyn til denne basis findes ved udregning af indre produkter.

Sætning 6.2.4 *Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, lad $(V, \langle -,- \rangle)$ være et $\mathbb{F}$-vektorrum med indre produkt, og lad $(\boldsymbol{u}_i)_{i\in I}$ være en ortonormal basis for V. Da er*

$$\boldsymbol{w} = \sum_{i\in I} \boldsymbol{u}_i \cdot \langle \boldsymbol{u}_i, \boldsymbol{w} \rangle$$

for alle $\boldsymbol{w} \in V$.

Bevis Da enhver vektor $\boldsymbol{w} \in V$ kan skrives entydigt som en linear kombination

$$\boldsymbol{w} = \sum_{j \in I} \boldsymbol{u}_j \cdot a_j$$

af basen $(\boldsymbol{u}_j)_{j \in I}$, følger lemmaet af udregningen

$$\langle \boldsymbol{u}_i, \boldsymbol{w} \rangle = \langle \boldsymbol{u}_i, \sum_{j \in I} \boldsymbol{u}_j \cdot a_j \rangle = \sum_{j \in I} \langle \boldsymbol{u}_i, \boldsymbol{u}_j \cdot a_j \rangle = \sum_{j \in I} \langle \boldsymbol{u}_i, \boldsymbol{u}_j \rangle \cdot a_j = a_i.$$

Her følger den anden og tredje identitet fra henholdsvis (H1) og (H2), mens den sidste identitet følger fra $\langle \boldsymbol{u}_i, \boldsymbol{u}_j \rangle = \delta_{ij}$. □

Eksempel 6.2.5 Ved at normere vektorerne i den ortogonale basis

$$\left(\boldsymbol{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix},\ \boldsymbol{v}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}\right)$$

fra eksempel 6.2.2 får vi den ortonormale basis

$$\left(\boldsymbol{u}_1 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix},\ \boldsymbol{u}_2 = \begin{pmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}\right).$$

Vi ser nu ved at anvende sætning 6.2.4, at vektoren

$$\boldsymbol{w} = \begin{pmatrix} 3 \\ 5 \end{pmatrix} \in \mathbb{F}^2$$

har koordinater

$$\boldsymbol{x} = \begin{pmatrix} \langle \boldsymbol{u}_1, \boldsymbol{w} \rangle \\ \langle \boldsymbol{u}_2, \boldsymbol{w} \rangle \end{pmatrix} = \begin{pmatrix} 8/\sqrt{2} \\ 2/\sqrt{2} \end{pmatrix} = \begin{pmatrix} 4 \\ 1 \end{pmatrix} \cdot \sqrt{2}$$

med hensyn til den ortonormale basis $(\boldsymbol{u}_1, \boldsymbol{u}_2)$ for $\mathbb{F}^2$.

Vi viser omvendt, at hvis koordinaterne for to vektorer med hensyn til en ortonormal basis for et vektorrum med indre produkt er kendte, så er det let at udregne det indre produkt af de to vektorer: det indre produkt i koordinater med hensyn til en ortonormale basis er altid givet ved standard-indreproduktet på $\mathbb{F}^n$.

Sætning 6.2.6 *Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, lad $(V, \langle -, - \rangle)$ være et endeligdimensionalt $\mathbb{F}$-vektorrum med indre produkt, og lad $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$ være en orthonormal basis for V. Lad $\boldsymbol{v}, \boldsymbol{w} \in V$, og lad $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{F}^n$ være koordinaterne for henholdsvis $\boldsymbol{v}$ og $\boldsymbol{w}$ med hensyn til basen $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$. Da er*

$$\langle \boldsymbol{v}, \boldsymbol{w} \rangle \ = \ \boldsymbol{x}^* \boldsymbol{y} \ = \ x_1^* y_1 + \cdots + x_n^* y_n.$$

Bevis Den ønskede formel følger fra udregningen

$$\begin{aligned}\langle \boldsymbol{v},\boldsymbol{w}\rangle &= \langle \sum_{i\in I} \boldsymbol{u}_i \cdot x_i, \sum_{j\in J} \boldsymbol{u}_j \cdot y_j\rangle = \sum_{i\in I} \langle \boldsymbol{u}_i \cdot x_i, \sum_{j\in J} \boldsymbol{u}_j \cdot y_j\rangle = \sum_{i\in I}\sum_{j\in I} \langle \boldsymbol{u}_i \cdot x_i, \boldsymbol{u}_j \cdot y_j\rangle \\ &= \sum_{i\in I}\sum_{j\in I} x_i^* \cdot \langle \boldsymbol{u}_i, \boldsymbol{u}_j \cdot y_j\rangle = \sum_{i\in I}\sum_{j\in I} x_i^* \cdot \langle \boldsymbol{u}_i, \boldsymbol{u}_j\rangle \cdot y_j = \sum_{i\in I} x_i^* y_i,\end{aligned}$$

hvor anden, tredje, fjerde og femte identitet fås fra egenskaberne ved et indre produkt, mens den sidste identitet fås fra $\langle \boldsymbol{u}_i, \boldsymbol{u}_j\rangle = \delta_{ij}$. □

Definition 6.2.7 Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, lad $(V, \langle -,-\rangle)$ være et $\mathbb{F}$-vektorrum med indre produkt. Givet vektorer $\boldsymbol{v},\boldsymbol{w} \in V$, da kaldes vektoren

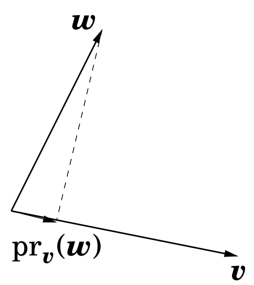
$$\mathrm{pr}_{\boldsymbol{v}}(\boldsymbol{w}) = \boldsymbol{v} \cdot \frac{\langle \boldsymbol{v},\boldsymbol{w}\rangle}{\langle \boldsymbol{v},\boldsymbol{v}\rangle}$$

for den *ortogonale projektion* af $\boldsymbol{w}$ på $\boldsymbol{v}$ med hensyn til $\langle -,-\rangle$.

Vi bemærker, at $\boldsymbol{w} - \mathrm{pr}_{\boldsymbol{v}}(\boldsymbol{w})$ er ortogonal på $\boldsymbol{v}$, idet

$$\langle \boldsymbol{v}, \boldsymbol{w} - \mathrm{pr}_{\boldsymbol{v}}(\boldsymbol{w})\rangle = \Big\langle \boldsymbol{v}, \boldsymbol{w} - \boldsymbol{v} \cdot \frac{\langle \boldsymbol{v},\boldsymbol{w}\rangle}{\langle \boldsymbol{v},\boldsymbol{v}\rangle}\Big\rangle = \langle \boldsymbol{v},\boldsymbol{w}\rangle - \langle \boldsymbol{v},\boldsymbol{v}\rangle \cdot \frac{\langle \boldsymbol{v},\boldsymbol{w}\rangle}{\langle \boldsymbol{v},\boldsymbol{v}\rangle} = 0,$$

samt at $\boldsymbol{w} = \mathrm{pr}_{\boldsymbol{v}}(\boldsymbol{w}) + (\boldsymbol{w} - \mathrm{pr}_{\boldsymbol{v}}(\boldsymbol{w}))$, hvilket er trivielt men ofte ganske brugbart.



Figur 6.1: Ortogonal projektion $\mathrm{pr}_{\boldsymbol{v}}(\boldsymbol{w})$ af vektoren $\boldsymbol{w}$ på vektoren $\boldsymbol{v}$

Eksempel 6.2.8 Figur 6.1 illustrerer det følgende eksempel på ortogonal projektion i $\mathbb{R}^2$ med hensyn til standard-indreproduktet:

$$\boldsymbol{v} = \begin{pmatrix} 5 \\ -1 \end{pmatrix}, \qquad \boldsymbol{w} = \begin{pmatrix} 2 \\ 4 \end{pmatrix} \qquad \text{og} \qquad \mathrm{pr}_{\boldsymbol{v}}(\boldsymbol{w}) = \begin{pmatrix} 5 \\ -1 \end{pmatrix} \cdot \frac{10-4}{25+1} = \begin{pmatrix} 5 \\ -1 \end{pmatrix} \cdot \frac{3}{13}.$$

Vi skal nu vise, at ethvert endeligdimensionalt vektorrum med indre produkt har en ortonormal basis. Mere præcist angiver vi en algoritme, der til en given basis tilordner en ny basis, der er ortonormal med hensyn til det indre produkt. Algoritmen kaldes for Gram-Schmidt ortogonalisering.

Sætning 6.2.9 *Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, lad $(V, \langle -,- \rangle)$ være et $\mathbb{F}$-vektorrum med indre produkt, og lad $(\boldsymbol{w}_1, \dots, \boldsymbol{w}_n)$ være en lineært uafhængig familie af vektorer i V. Lad $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ være familien af vektorer i V, hvori $\boldsymbol{v}_j$ er defineret rekursivt ved*

$$\boldsymbol{v}_j = \boldsymbol{w}_j - \sum_{1 \le k < j} \operatorname{pr}_{\boldsymbol{v}_k}(\boldsymbol{w}_j)$$

for alle $1 \le j \le n$; og lad endvidere $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$ være familien af vektorer i V, hvor

$$\boldsymbol{u}_j = \boldsymbol{v}_j \cdot \|\boldsymbol{v}_j\|^{-1}$$

for alle $1 \le j \le n$. Da gælder følgende:

(1) *Familierne $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$, $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ og $(\boldsymbol{w}_1, \dots, \boldsymbol{w}_n)$ frembringer samme underrum af V.*

(2) *Familien $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ er ortogonal med hensyn til $\langle -,- \rangle$.*

(3) *Familien $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$ er ortonormal med hensyn til $\langle -,- \rangle$.*

Specielt, hvis $(\boldsymbol{w}_1, \dots, \boldsymbol{w}_n)$ er en basis for V, da er $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ og $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$ baser for V, der er henholdvis ortogonal og ortonormal med hensyn til $\langle -,- \rangle$.

Bevis Vi beviser påstanden ved induktion på $n \ge 0$. Tilfældet $n = 0$ er trivielt, så vi antager, at påstanden er vist for $n = r - 1$, og viser den for $n = r$. Ifølge den induktive antagelse frembringer $(r-1)$-tuplerne $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_{r-1})$ og $(\boldsymbol{w}_1, \dots, \boldsymbol{w}_{r-1})$ samme underrum af V, og derfor frembringer r-tuplerne $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_{r-1}, \boldsymbol{w}_r)$ og $(\boldsymbol{w}_1, \dots, \boldsymbol{w}_{r-1}, \boldsymbol{w}_r)$ også samme underrum W af V. Fra definitionen af $\boldsymbol{v}_r$ har vi endvidere, at

$$\boldsymbol{w}_r = \boldsymbol{v}_r + \sum_{1 \le k < r} \boldsymbol{v}_k \cdot \frac{\langle \boldsymbol{v}_k, \boldsymbol{w}_r \rangle}{\langle \boldsymbol{v}_k, \boldsymbol{v}_k \rangle},$$

hvilket viser, at r-tuplerne $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_{r-1}, \boldsymbol{v}_r)$ og $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_{r-1}, \boldsymbol{w}_r)$ begge frembringer W, og vi har dermed vist, at $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_r)$ og $(\boldsymbol{w}_1, \dots, \boldsymbol{w}_r)$ begge frembringer W. Da r-tuplen $(\boldsymbol{w}_1, \dots, \boldsymbol{w}_r)$ er lineært uafhængig, udgør denne en basis for W, og sætning 4.2.19 viser da, at r-tuplen $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_r)$ ligeledes er en basis for W. Specielt er vektorerne $\boldsymbol{v}_1, \dots, \boldsymbol{v}_r$ alle forskellige fra $\boldsymbol{0}$, hvilket viser, at skalarene $\|\boldsymbol{v}_1\|, \dots, \|\boldsymbol{v}_r\|$ alle er forskellige fra 0, således at vektorerne $\boldsymbol{u}_1, \dots, \boldsymbol{u}_r$ er veldefinerede. Endvidere viser

$$\boldsymbol{v}_1 \cdot a_1 + \dots + \boldsymbol{v}_r \cdot a_r = \boldsymbol{u}_1 \cdot (\|\boldsymbol{v}_1\| \cdot a_1) + \dots + \boldsymbol{u}_r \cdot (\|\boldsymbol{v}_r\| \cdot a_r),$$

at også $(\boldsymbol{u}_1,\ldots,\boldsymbol{u}_r)$ frembringer W. Dette viser induktionsskridtet for (1).

Ifølge den induktive antagelse er $(\boldsymbol{v}_1,\ldots,\boldsymbol{v}_{r-1})$ endvidere ortogonal med hensyn til $\langle -,-\rangle$. For at vise, at også $(\boldsymbol{v}_1,\ldots,\boldsymbol{v}_r)$ er ortogonal med hensyn til $\langle -,-\rangle$, skal vi vise derfor blot vise, at $\langle \boldsymbol{v}_i,\boldsymbol{v}_r\rangle = 0$ for alle $1 \le i \le r-1$. Så vi lader $1 \le i \le r-1$ være fastholdt og udregner, at

$$\begin{aligned}
\langle \boldsymbol{v}_i,\boldsymbol{v}_r\rangle &= \Big\langle \boldsymbol{v}_i,\boldsymbol{w}_r - \sum_{1\le k<r} \boldsymbol{v}_k \cdot \frac{\langle \boldsymbol{v}_k,\boldsymbol{w}_r\rangle}{\langle \boldsymbol{v}_k,\boldsymbol{v}_k\rangle}\Big\rangle \\
&\overset{(H2)}{=} \langle \boldsymbol{v}_i,\boldsymbol{w}_r\rangle - \sum_{1\le k<r} \langle \boldsymbol{v}_i,\boldsymbol{v}_k\rangle \cdot \frac{\langle \boldsymbol{v}_k,\boldsymbol{w}_r\rangle}{\langle \boldsymbol{v}_k,\boldsymbol{v}_k\rangle} \\
&= \langle \boldsymbol{v}_i,\boldsymbol{w}_r\rangle - \langle \boldsymbol{v}_i,\boldsymbol{v}_i\rangle \cdot \frac{\langle \boldsymbol{v}_i,\boldsymbol{w}_r\rangle}{\langle \boldsymbol{v}_i,\boldsymbol{v}_i\rangle} \;=\; 0,
\end{aligned}$$

hvor vi har brugt, at $\langle \boldsymbol{v}_i,\boldsymbol{v}_k\rangle = 0$ for alle $1 \le k < r$ med $i \ne k$. Dermed er $(\boldsymbol{v}_1,\ldots,\boldsymbol{v}_r)$ ortogonal med hensyn til $\langle -,-\rangle$, hvilket viser induktionsskridtet for (2).

Endelig viser udregningen

$$\langle \boldsymbol{u}_i,\boldsymbol{u}_j\rangle = \langle \boldsymbol{v}_i \cdot \|\boldsymbol{v}_i\|^{-1}, \boldsymbol{v}_j \cdot \|\boldsymbol{v}_j\|^{-1}\rangle = \|\boldsymbol{v}_i\|^{-1} \cdot \langle \boldsymbol{v}_i,\boldsymbol{v}_j\rangle \cdot \|\boldsymbol{v}_j\|^{-1} = \begin{cases} 1 & \text{hvis } i = j, \\ 0 & \text{hvis } i \ne j, \end{cases}$$

at $(\boldsymbol{u}_1,\ldots,\boldsymbol{u}_r)$ er ortonormal med hensyn til $\langle -,-\rangle$. Dette viser induktionsssskridtet for (3) og dermed sætningen. Den sidste bemærkning følger fra (1)–(3) og sætning 4.2.19. □

Korollar 6.2.10 *Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$. Ethvert endeligt frembragt $\mathbb{F}$-vektorrum med indre produkt $(V, \langle -,-\rangle)$ har en basis, der er ortonormal med hensyn til $\langle -,-\rangle$.*

Bevis Ifølge sætning 4.2.11 findes der en endelig basis for V. Vi vælger derfor en sådan basis $(\boldsymbol{w}_1,\ldots,\boldsymbol{w}_n)$, og anvender Gram-Schmidt ortogonalisering til at omdanne den til en basis $(\boldsymbol{u}_1,\ldots,\boldsymbol{u}_n)$, der er ortonormal med hensyn til $\langle -,-\rangle$. □

Eksempel 6.2.11 Vi anvender Gram-Schmidt ortogonalisering til at omdanne basen

$$\left(\boldsymbol{w}_1 = \begin{pmatrix}1\\1\\1\end{pmatrix},\ \boldsymbol{w}_2 = \begin{pmatrix}1\\0\\2\end{pmatrix},\ \boldsymbol{w}_3 = \begin{pmatrix}0\\1\\1\end{pmatrix}\right)$$

for $\mathbb{F}^3$ til en basis $(\boldsymbol{v}_1,\boldsymbol{v}_2\boldsymbol{v}_3)$, der er ortogonal med hensyn til standard-indreproduktet

$\langle -,- \rangle$, og til en basis $(\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3)$, der er ortonormal med hensyn til $\langle -,- \rangle$.

$$\begin{aligned}
\boldsymbol{v}_1 &= \boldsymbol{w}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \\
\boldsymbol{v}_2 &= \boldsymbol{w}_2 - \boldsymbol{v}_1 \cdot \frac{\langle \boldsymbol{v}_1, \boldsymbol{w}_2 \rangle}{\langle \boldsymbol{v}_1, \boldsymbol{v}_1 \rangle} = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \cdot \frac{1+0+2}{1+1+1} = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \\
\boldsymbol{v}_3 &= \boldsymbol{w}_3 - \boldsymbol{v}_1 \cdot \frac{\langle \boldsymbol{v}_1, \boldsymbol{w}_3 \rangle}{\langle \boldsymbol{v}_1, \boldsymbol{v}_1 \rangle} - \boldsymbol{v}_2 \cdot \frac{\langle \boldsymbol{v}_2, \boldsymbol{w}_3 \rangle}{\langle \boldsymbol{v}_2, \boldsymbol{v}_2 \rangle} \\
&= \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \cdot \frac{0+1+1}{1+1+1} - \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \cdot \frac{0-1+1}{0+1+1} = \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} \cdot \frac{1}{3}.
\end{aligned}$$

Vi har hermed fundet den nye basis

$$\left(\boldsymbol{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \ \boldsymbol{v}_2 = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}, \ \boldsymbol{v}_3 = \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} \cdot \frac{1}{3} \right),$$

der er ortogonal med hensyn til $\langle -,- \rangle$, og vi mangler nu blot at normere vektorerne i denne for at finde basen $(\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3)$, der er ortonormal med hensyn til $\langle -,- \rangle$. Vi kan lige så godt normere vektorerne i basen

$$\left(\boldsymbol{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \ \boldsymbol{v}_2 = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}, \ \boldsymbol{v}_3' = \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} \right),$$

der også er ortogonal med hensyn til $\langle -,- \rangle$, sådan at vi slipper for den ubehagelige konstant $\frac{1}{3}$ i vores udregninger. Vi finder dermed, at basen

$$\left(\boldsymbol{u}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \cdot \frac{1}{\sqrt{3}}, \ \boldsymbol{u}_2 = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \cdot \frac{1}{\sqrt{2}}, \ \boldsymbol{u}_3 = \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} \cdot \frac{1}{\sqrt{6}} \right)$$

er ortonormal med hensyn til $\langle -,- \rangle$.

Definition 6.2.12 Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$ og lad $(V, \langle -,- \rangle)$ være et $\mathbb{F}$-vektorrum med indre produkt. Hvis $U \subset V$ er et underrum, så kaldes underrummet

$$U^{\perp} = \{ \boldsymbol{v} \in V \mid \langle \boldsymbol{u}, \boldsymbol{v} \rangle = 0 \text{ for alle } \boldsymbol{u} \in U \} \subset V$$

for det *ortogonale komplement* af $U \subset V$ med hensyn til $\langle -,- \rangle$.

For at eftervise, at $U^\perp \subset V$ er et underrum, tjekker vi, at (1)–(3) i sætning 4.1.4 er opfyldte. For alle $\boldsymbol{u} \in U$, er $\langle \boldsymbol{u}, \boldsymbol{0} \rangle = 0$, så (1) er opfyldt. Hvis $\boldsymbol{v}_1, \boldsymbol{v}_2 \in U^\perp$, så gælder for alle $\boldsymbol{u} \in U$, at $\langle \boldsymbol{u}, \boldsymbol{v}_1 + \boldsymbol{v}_2 \rangle = \langle \boldsymbol{u}, \boldsymbol{v}_1 \rangle + \langle \boldsymbol{u}, \boldsymbol{v}_2 \rangle = 0 + 0 = 0$, så (2) er opfyldt. Hvis $\boldsymbol{v} \in U^\perp$ og $a \in \mathbb{F}$, så gælder for alle $\boldsymbol{u} \in U$, at $\langle \boldsymbol{u}, \boldsymbol{v} \cdot a \rangle = \langle \boldsymbol{u}, \boldsymbol{v} \rangle \cdot a = 0 \cdot a = 0$, så også (3) er opfyldt. Dermed er $U^\perp \subset V$ et underrum. Vi bemærker også, at

$$U \cap U^\perp = \{\boldsymbol{0}\}$$

på grund af positivitetsbetingelsen (P) i definition af et indre produkt.

Eksempel 6.2.13 Vi betragter $\mathbb{R}^2$ med standard-indreproduktet. Underrummet

$$U = \{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 \mid x_1 = x_2\} \subset \mathbb{R}^2$$

fra eksempel 4.1.5 har da ortogonalt komplement

$$U^\perp = \{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 \mid x_1 = -x_2\} \subset \mathbb{R}^2.$$

Sætning 6.2.14 *Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, lad $(V, \langle -,- \rangle)$ være et $\mathbb{F}$-vektorrum med indre produkt, lad $U \subset V$ være et endeligdimensionalt underrum, og lad $U^\perp \subset V$ være dets ortogonale komplement. Enhver vektor $\boldsymbol{w} \in V$ kan da skrives entydigt som*

$$\boldsymbol{w} = \mathrm{pr}_U(\boldsymbol{w}) + (\boldsymbol{w} - \mathrm{pr}_U(\boldsymbol{w}))$$

med $\mathrm{pr}_U(\boldsymbol{w}) \in U$ og $\boldsymbol{w} - \mathrm{pr}_U(\boldsymbol{w}) \in U^\perp$. Hvis endvidere $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ er en basis for U, der er ortogonal med hensyn til $\langle -,- \rangle$, så er

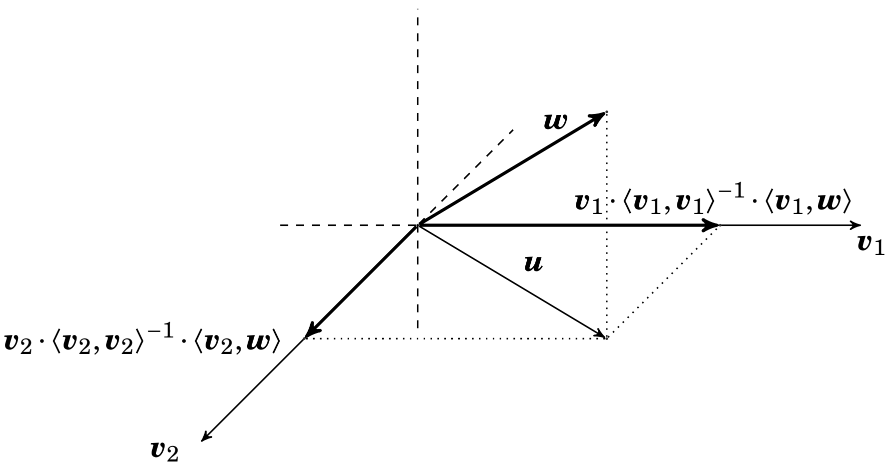
$$\mathrm{pr}_U(\boldsymbol{w}) = \boldsymbol{v}_1 \cdot \frac{\langle \boldsymbol{v}_1, \boldsymbol{w} \rangle}{\langle \boldsymbol{v}_1, \boldsymbol{v}_1 \rangle} + \cdots + \boldsymbol{v}_n \cdot \frac{\langle \boldsymbol{v}_n, \boldsymbol{w} \rangle}{\langle \boldsymbol{v}_n, \boldsymbol{v}_n \rangle}.$$

Bevis Vi bemærker først, at hvis $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$ er en vilkårlig basis for U, og hvis $\boldsymbol{v} \in V$ opfylder, at $\langle \boldsymbol{u}_i, \boldsymbol{v} \rangle = 0$ for alle $1 \le i \le n$, så er $\boldsymbol{v} \in U^\perp$. For vi kan skrive en vilkårlig vektor $\boldsymbol{u} \in U$ som en linear kombination $\boldsymbol{u} = \boldsymbol{u}_1 a_1 + \cdots + \boldsymbol{u}_n a_n$, og

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle = \langle \boldsymbol{u}_1 a_1 + \cdots + \boldsymbol{u}_n a_n, \boldsymbol{v} \rangle = a_1^* \langle \boldsymbol{u}_1, \boldsymbol{v} \rangle + \cdots + a_n^* \langle \boldsymbol{u}_n, \boldsymbol{v} \rangle = 0.$$

Vi lader nu $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$ være en basis for U, der er ortonormal med hensyn til $\langle -,- \rangle$. En sådan findes ifølge korollar 6.2.10. Hvis $\boldsymbol{w} = \boldsymbol{u} + \boldsymbol{v}$ med $\boldsymbol{u} \in U$ og $\boldsymbol{v} \in U^\perp$, da er

$$\langle \boldsymbol{u}_i, \boldsymbol{w} \rangle = \langle \boldsymbol{u}_i, \boldsymbol{u} + \boldsymbol{v} \rangle = \langle \boldsymbol{u}_i, \boldsymbol{u} \rangle + \langle \boldsymbol{u}_i, \boldsymbol{v} \rangle = \langle \boldsymbol{u}_i, \boldsymbol{u} \rangle.$$



Figur 6.2: Den ortogonale projektion $\boldsymbol{u} = \mathrm{pr}_U(\boldsymbol{w})$ af $\boldsymbol{w}$ på underrummet $U \subset \mathbb{R}^3$ med den ortogonale basis $(\boldsymbol{v}_1, \boldsymbol{v}_2)$.

Derfor er vektoren $\boldsymbol{u} \in U$ nødvendigvis givet ved

$$\boldsymbol{u} = \boldsymbol{u}_1 \cdot \langle \boldsymbol{u}_1, \boldsymbol{w} \rangle + \cdots + \boldsymbol{u}_n \cdot \langle \boldsymbol{u}_n, \boldsymbol{w} \rangle.$$

Vi definerer nu $\boldsymbol{u} \in U$ til at være denne vektor og definerer $\boldsymbol{v} = \boldsymbol{w} - \boldsymbol{u}$. Vi mangler at vise, at $\boldsymbol{v} \in U^{\perp}$. Da $\langle \boldsymbol{u}_i, \boldsymbol{u}_j \rangle = \delta_{ij}$, så gælder for alle $1 \leq i \leq n$, at

$$\langle \boldsymbol{u}_i, \boldsymbol{v} \rangle = \langle \boldsymbol{u}_i, \boldsymbol{w} - (\boldsymbol{u}_1 \cdot \langle \boldsymbol{u}_1, \boldsymbol{w} \rangle + \cdots + \boldsymbol{u}_n \cdot \langle \boldsymbol{u}_n, \boldsymbol{w} \rangle) \rangle = \langle \boldsymbol{u}_i, \boldsymbol{w} \rangle - \langle \boldsymbol{u}_i, \boldsymbol{w} \rangle = 0.$$

Altså er $\boldsymbol{v} \in U^{\perp}$ som ønsket. □

Definition 6.2.15 Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, lad $(V, \langle -, - \rangle)$ være et $\mathbb{F}$-vektorrum med indre produkt, og lad $U \subset V$ være et endeligdimensionalt underrum. Afbildningen

$$\mathrm{pr}_U \colon V \to V$$

kaldes for den *ortogonale projektion* på $U \subset V$ med hensyn til $\langle -, - \rangle$.

Den formel for $\mathrm{pr}_U(\boldsymbol{w})$, vi angav i sætning 6.2.14, samt egenskaberne (H1)–(H2) af et indre produkt viser, at den ortogonal projektion $\mathrm{pr}_U \colon V \to V$ er en lineær afbildning.

Eksempel 6.2.16 Vi ønsker at bestemme den ortogonale projektion $\mathrm{pr}_U \colon \mathbb{R}^3 \to \mathbb{R}^3$ på underrummet $U \subset \mathbb{R}^3$ frembragt af familien

$$\left(\boldsymbol{w}_1 = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}, \ \boldsymbol{w}_2 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} \right)$$

med hensyn til standard-indreproduktet på $\mathbb{R}^3$. Vi anvender derfor Gram-Schmidt algoritmen til at omdanne basen $(\boldsymbol{w}_1, \boldsymbol{w}_2)$ til en ny basis $(\boldsymbol{v}_1, \boldsymbol{v}_2)$, der er ortogonal med hensyn til standard-indreproduktet.

$$\begin{aligned} \boldsymbol{v}_1 &= \boldsymbol{w}_1 \\ \boldsymbol{v}_2 &= \boldsymbol{w}_2 - \boldsymbol{v}_1 \cdot \frac{\langle \boldsymbol{v}_1, \boldsymbol{w}_2 \rangle}{\langle \boldsymbol{v}_1, \boldsymbol{v}_1 \rangle} = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} - \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \cdot \frac{0-2+0}{0+1+1} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \end{aligned}$$

Dermed er den ortogonale projektion $\mathrm{pr}_U \colon \mathbb{R}^3 \to \mathbb{R}^3$ givet ved

$$\mathrm{pr}_U(\boldsymbol{x}) = \boldsymbol{v}_1 \cdot \frac{\langle \boldsymbol{v}_1, \boldsymbol{x} \rangle}{\langle \boldsymbol{v}_1, \boldsymbol{v}_1 \rangle} + \boldsymbol{v}_2 \cdot \frac{\langle \boldsymbol{v}_2, \boldsymbol{x} \rangle}{\langle \boldsymbol{v}_2, \boldsymbol{v}_2 \rangle} = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \cdot \frac{0 - x_2 + x_3}{0+1+1} + \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \cdot \frac{x_1 + x_2 + x_3}{1+1+1},$$

hvoraf vi aflæser, at $\mathrm{pr}_U(\boldsymbol{x}) = A\boldsymbol{x}$, hvor

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

Med andre ord er $A \in M_3(\mathbb{R})$ den matrix, der repræsenterer $\mathrm{pr}_U \colon \mathbb{R}^3 \to \mathbb{R}^3$ med hensyn til standardbasen $(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$ for både domænet og codomænet.

6.3 Lineære isometrier

Vi betragter lineære afbildninger mellem vektorrum med indre produkt, der bevarer det indre produkt. Disse afbildninger bevarer da også de geometriske strukturer, som det indre produkt repræsenterer, såsom længder og vinkler. Vi viser specielt, at en lineær isomorfi $f \colon V \to V$ af et endeligt dimensionalt vektorrum bevarer et indre produkt $\langle -, - \rangle$ på V, hvis og kun hvis matricen A, der repræsenterer f med hensyn til en fælles basis for både domæne og codomæne, der er *ortonormal* med hensyn til $\langle -, - \rangle$, opfylder, at $A^* = A^{-1}$. Her er A^* den adjungerede matrix defineret i afsnit 2.6.

Definition 6.3.1 Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, og lad $(V, \langle -, - \rangle_V)$ og $(W, \langle -, - \rangle_W)$ være $\mathbb{F}$-vektorrum med indre produkt. En lineær afbildning $f \colon W \to V$, der opfylder

$$\langle f(\boldsymbol{w}), f(\boldsymbol{w}') \rangle_V = \langle \boldsymbol{w}, \boldsymbol{w}' \rangle_W$$

for alle $\boldsymbol{w}, \boldsymbol{w}' \in W$, kaldes for en *lineær isometri* med hensyn til $\langle -, - \rangle_W$ og $\langle -, - \rangle_V$.

Eksempel 6.3.2 (1) Vi betragter igen rotationen $r_\theta\colon \mathbb{R}^2 \to \mathbb{R}^2$ gennem θ radianer imod urets retning fra eksempel 2.3.10, der er givet ved $r_\theta(\boldsymbol{x}) = B_\theta \boldsymbol{x}$ med

$$B_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}.$$

Udregningen

$$\begin{aligned} \langle r_\theta(\boldsymbol{x}), r_\theta(\boldsymbol{y})\rangle &= \Big\langle \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \boldsymbol{x}, \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \boldsymbol{y} \Big\rangle \\ &= \cos(\theta)x_1 - \sin(\theta)x_2)(\cos(\theta)y_1 - \sin(\theta)y_2) \\ &\quad + (\sin(\theta)x_1 + \cos(\theta)x_2)(\sin(\theta)y_1 + \cos(\theta)y_2) \\ &= (\cos(\theta)^2 + \sin(\theta)^2)x_1y_1 + (\cos(\theta)^2 + \sin(\theta)^2)x_2y_2 \\ &= x_1y_1 + x_2y_2 = \langle \boldsymbol{x}, \boldsymbol{y}\rangle \end{aligned}$$

viser da, at $r_\theta\colon \mathbb{R}^2 \to \mathbb{R}^2$ er en lineær isometri med hensyn til standard-indreproduktet.
(2) Vi betragter også den vægtede skaleringsafbilding $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ givet ved

$$f(\boldsymbol{x}) = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \boldsymbol{x}.$$

Udregningen

$$\langle f(\boldsymbol{x}), f(\boldsymbol{y})\rangle = \Big\langle \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \boldsymbol{x}, \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \boldsymbol{y} \Big\rangle = \lambda_1^2 x_1y_1 + \lambda_2^2 x_2y_2$$

viser da, at $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ er en lineær isometri med hensyn til standard-indreproduktet på $\mathbb{R}^2$, hvis og kun hvis $|\lambda_1| = 1$ og $|\lambda_2| = 1$.

Navnet "isometri" er noget forvirrende, idet det kunne antyde, at en lineær isometri altid er en isomorfi, hvilket *ikke* er tilfældet. Der gælder derimod følgende.

Lemma 6.3.3 *Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, og lad $(V, \langle -,-\rangle_V)$ og $(W, \langle -,-\rangle_W)$ være to $\mathbb{F}$-vektorrum med indre produkt.*

(1) *En lineær isometri $f\colon W \to V$ er nødvendigvis injektiv.*

(2) *Hvis en lineær isometri $f\colon W \to V$ er en isomorfi, da er også den inverse afbildning $f^{-1}\colon V \to W$ en lineær isometri.*

Bevis For at vise (1), lader vi $\boldsymbol{w} \in W$ og antager, at $f(\boldsymbol{w}) = \boldsymbol{0}$. Da er

$$\langle \boldsymbol{w}, \boldsymbol{w}\rangle_W = \langle f(\boldsymbol{w}), f(\boldsymbol{w})\rangle_V = \langle \boldsymbol{0}, \boldsymbol{0}\rangle_V = 0,$$

og positivitetsegenskaben (P) af $\langle -,-\rangle_W$ medfører derfor, at $\boldsymbol{w} = \boldsymbol{0}$. Da afbildningen $f\colon W \to V$ er lineær, konkluderer vi derfor, at den er injektiv.

For at vise (2), skal vi vise, at der for alle $\boldsymbol{v}, \boldsymbol{v}' \in V$ gælder, at

$$\langle \boldsymbol{v}, \boldsymbol{v}' \rangle_V = \langle f^{-1}(\boldsymbol{v}), f^{-1}(\boldsymbol{v}') \rangle_W.$$

Da $f : W \to V$ er en bijektion, så er $f^{-1}(\boldsymbol{v}) = \boldsymbol{w}$, hvis og kun hvis $\boldsymbol{v} = f(\boldsymbol{w})$, og tilsvarende er $f^{-1}(\boldsymbol{v}') = \boldsymbol{w}'$, hvis og kun hvis $\boldsymbol{v}' = f(\boldsymbol{w}')$. Derfor er det udsagn, vi ønsker at vise, ækvivalent til udsagnet, at der for alle $\boldsymbol{w}, \boldsymbol{w}' \in W$ gælder, at

$$\langle f(\boldsymbol{w}), f(\boldsymbol{w}') \rangle_V = \langle \boldsymbol{w}, \boldsymbol{w}' \rangle_W,$$

og dette sidste udsagn gælder, da $f : W \to V$ er en lineær isometri. □

Eksempel 6.3.4 Lad $\boldsymbol{u} \in \mathbb{R}^n$ være en enhedsvektor med hensyn til standard-indreproduktet. Da er afbildningen $\sigma_{\boldsymbol{u}} : \mathbb{F} \to \mathbb{F}^n$ defineret ved $\sigma_{\boldsymbol{u}}(a) = \boldsymbol{u}a$ en lineær isometri, men den er kun en isomorfi, hvis $n = 1$.

Vi indfører nu begrebet *adjunktion* for lineære afbildninger mellem vektorrum med indre produkt. Vi anvender senere dette begreb til bedre at forstå lineære isometrier.

Definition 6.3.5 Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, og lad $(V, \langle -,- \rangle_V)$ og $(W, \langle -,- \rangle_W)$ være $\mathbb{F}$-vektorrum med indre produkt. To lineære afbildninger $f : W \to V$ og $g : V \to W$ siges at være *adjungerede* med hensyn til $\langle -,- \rangle_V$ og $\langle -,- \rangle_W$, hvis

$$\langle \boldsymbol{v}, f(\boldsymbol{w}) \rangle_V = \langle g(\boldsymbol{w}), \boldsymbol{v} \rangle_W$$

for alle $\boldsymbol{v} \in V$ og $\boldsymbol{w} \in W$.

Vi skal vise, at enhver lineær afbildning mellem to endeligt dimensionale vektorrum med indre produkt har en entydigt bestemt adjungeret afbildning. Vi begynder med at vise entydighedsudsagnet, som gælder helt generelt.

Lemma 6.3.6 *Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, og lad $(V, \langle -,- \rangle_V)$ og $(W, \langle -,- \rangle_W)$ være to $\mathbb{F}$-vektorrum med indre produkt, og lad $f : W \to V$ være en lineære afbildning. Hvis $g, h : V \to W$ er to lineære afbildninger, der er begge adjungerede til f med hensyn til $\langle -,- \rangle_V$ og $\langle -,- \rangle_W$, så er $g = h$.*

Bevis Vi antager, at både $g : V \to W$ og $h : V \to W$ er adjungerede til $f : W \to V$ med hensyn til $\langle -,- \rangle_V$ og $\langle -,- \rangle_W$. For alle $\boldsymbol{v} \in V$ og $\boldsymbol{w} \in W$, gælder da, at

$$\langle g(\boldsymbol{w}) - h(\boldsymbol{w}), \boldsymbol{v} \rangle_W = \langle g(\boldsymbol{w}), \boldsymbol{v} \rangle_W - \langle h(\boldsymbol{w}), \boldsymbol{v} \rangle_W = \langle \boldsymbol{w}, f(\boldsymbol{v})_V - \langle \boldsymbol{w}, f(\boldsymbol{v})_V = 0.$$

Dette gælder specielt for $\boldsymbol{v} = g(\boldsymbol{w}) - h(\boldsymbol{w})$, og positivitetsegenskaben (P) ved et indre produkt medfører derfor, at $g(\boldsymbol{w}) - h(\boldsymbol{w}) = \boldsymbol{0}$ som ønsket. □

Sætning 6.3.7 *Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, lad $(V, \langle -,- \rangle_V)$ og $(W, \langle -,- \rangle_W)$ være to endeligt dimensionale $\mathbb{F}$-vektorrum med indre produkt, og lad $f\colon W \to V$ være en lineær afbildning. Lad $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_m)$ og $(\boldsymbol{w}_1, \dots, \boldsymbol{w}_n)$ være baser for V og W, der er ortonormale med hensyn til henholdsvis $\langle -,- \rangle_V$ og $\langle -,- \rangle_W$.*

(1) *Der findes en entydigt bestemt lineær afbildning $g\colon V \to W$, sådan at $f\colon W \to V$ og $g\colon V \to W$ er adjungerede med hensyn til $\langle -,- \rangle_V$ og $\langle -,- \rangle_W$.*

(2) *Hvis matricen $A \in M_{m,n}(\mathbb{F})$ repræsenterer afbildningen $f\colon W \to V$ med hensyn de givne ortonormale baser, så repræsenterer den adjungerede matrix $A^* \in M_{n,m}(\mathbb{F})$ den adjungerede afbildning $g\colon V \to W$ med hensyn til disse baser.*

Bevis Vi påstår, at hvis to lineære afbildning $f\colon W \to V$ og $g\colon W \to V$ er repræsenterede ved adjungerede matricer $A \in M_{m,n}(\mathbb{F})$ og $A^* \in M_{n,m}(\mathbb{F})$ med hensyn til ortonormale baser $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_m)$ for V og $(\boldsymbol{w}_1, \dots, \boldsymbol{w}_n)$ for W, så er de adjungerede lineære afbildninger. Dette viser både (1) og (2), da entydighedsudsagnet i (1) allerede er vist i lemma 6.3.6. Så lad $\boldsymbol{v} \in V$ og $\boldsymbol{w} \in W$, og lad $\boldsymbol{x} \in \mathbb{F}^m$ og $\boldsymbol{y} \in \mathbb{F}^n$ være koordinaterne for $\boldsymbol{v}$ og $\boldsymbol{w}$ med hensyn til de givne baser. Fra sætning 6.2.6 har vi

$$\langle \boldsymbol{v}, f(\boldsymbol{w}) \rangle_V = \boldsymbol{x}^*(A\boldsymbol{y}) \quad \text{og} \quad \langle g(\boldsymbol{v}), \boldsymbol{w} \rangle_W = (A^*\boldsymbol{x})^*\boldsymbol{y},$$

mens sætning 2.6.5 giver, at $\boldsymbol{x}^*A\boldsymbol{y} = (A^*\boldsymbol{x})^*\boldsymbol{y}$, hvilket viser sætningen. □

Eksempel 6.3.8 Vi lader $W = \mathbb{R}^2$ og $V = \mathbb{R}^3$ med standard-indreprodukterne og betragter den lineære afbildning $f\colon V \to W$ defineret ved $f(\boldsymbol{x}) = A\boldsymbol{x}$, hvor

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

Matricen A repræsenterer altså $f\colon V \to W$ med hensyn til standardbaserne $(\boldsymbol{e}_1, \boldsymbol{e}_2)$ og $(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$, og da disse er ortogonale med hensyn til standard-indreprodukterne, så repræsenterer den adjungerede matrix

$$A^* = A^t = \begin{pmatrix} 2 & 0 & 1 \\ 3 & 1 & -1 \end{pmatrix},$$

ifølge sætning 6.3.7 den adjungerede afbildning $g\colon V \to W$. Udregningen

$$\begin{aligned} \langle \boldsymbol{x}, f(\boldsymbol{y}) \rangle &= x_1(2y_1 + 3y_2) + x_2y_2 + x_3(y_1 - y_2) \\ &= (2x_1 + x_3)y_1 + (3x_1 + x_2 - x_3)y_2 = \langle g(\boldsymbol{x}), \boldsymbol{y} \rangle \end{aligned}$$

illustrerer dette forhold.

Udsagnet, at en lineær afbildning er en isometri, kan udtrykkes ved hjælp af dens adjungerede lineære afbildning på følgende vis.

Lemma 6.3.9 *Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, og lad $(V, \langle -,- \rangle_V)$ og $(W, \langle -,- \rangle_W)$ være to $\mathbb{F}$-vektorrum med indre produkt. Hvis to lineære afbildninger $f \colon W \to V$ og $g \colon W \to V$ er adjungerede med hensyn til $\langle -,- \rangle_V$ og $\langle -,- \rangle_W$, da er $f \colon W \to V$ en lineær isometri med hensyn til $\langle -,- \rangle_W$ og $\langle -,- \rangle_V$, hvis og kun hvis $g \circ f = \mathrm{id}_W \colon W \to W$.*

Bevis Da f og g er adjungerede, så gælder der for alle $\boldsymbol{w}, \boldsymbol{w}' \in W$, at

$$\langle f(\boldsymbol{w}), f(\boldsymbol{w}') \rangle_V = \langle (g \circ f)(\boldsymbol{w}), \boldsymbol{w}' \rangle_W.$$

Derfor er $f \colon W \to V$ er en lineær isometri, hvis og kun hvis

$$\langle (g \circ f)(\boldsymbol{w}), \boldsymbol{w}' \rangle_W = \langle \boldsymbol{w}, \boldsymbol{w}' \rangle_W,$$

for alle $\boldsymbol{w}, \boldsymbol{w}' \in W$. Hvis dette gælder, så får vi for $\boldsymbol{w}' = (g \circ f)(\boldsymbol{w}) - \boldsymbol{w}$, at

$$\langle \boldsymbol{w}', \boldsymbol{w}' \rangle_W = \langle (g \circ f)(\boldsymbol{w}) - \boldsymbol{w}, \boldsymbol{w}' \rangle_W = \langle (g \circ f)(\boldsymbol{w}), \boldsymbol{w}' \rangle_W - \langle \boldsymbol{w}, \boldsymbol{w}' \rangle_W = 0,$$

hvorfra vi konkluderer, at $\boldsymbol{w}' = \boldsymbol{0}$ og dermed, at $g \circ f = \mathrm{id}_W$. Hvis omvendt $g \circ f = \mathrm{id}_W$, så er betingelsen ovenfor trivielt opfyldt. □

Sætning 6.3.10 *Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, og lad $(V, \langle -,- \rangle_V)$ og $(W, \langle -,- \rangle_W)$ være to endeligt dimensionale $\mathbb{F}$-vektorrum med indre produkt. Lad $f \colon W \to V$ være en lineær afbildning, og lad $A \in M_{m,n}(\mathbb{F})$ være matricen, der repræsenterer $f \colon W \to V$ med hensyn til ortonormale baser $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_m)$ og $(\boldsymbol{w}_1, \dots, \boldsymbol{w}_n)$ for henholdsvis V og W.*

(1) *Afbildningen $f \colon W \to V$ er en lineær isometri, hvis og kun hvis $A^*A = I_n$.*

(2) *Afbildningen $f \colon W \to V$ er en isometrisk isomorfi, hvis og kun hvis A er invertibel og $A^* = A^{-1}$.*

Bevis Ifølge sætning 6.3.7 er den adjungerede afbildning $g \colon V \to W$ repræsenteret af den adjungerede matrix $A^* \in M_{n,m}(\mathbb{F})$ med hensyn til de givne baser. Derfor er (1) blot en oversættelse af lemma 6.3.9. For at vise (2) antager vi først, at $f \colon W \to V$ er en isometrisk isomorfi, hvilket vil sige, at $f \colon W \to V$ både er en isometri og en isomorfi. Da $f \colon W \to V$ er en isometri, følger det fra (1), at $A^*A = I$, og da $f \colon W \to V$ er en isomorfi, er A invertibel, og vi har endvidere, at

$$A^* = A^*(AA^{-1}) = (A^*A)A^{-1} = A^{-1}$$

som ønsket. Omvendt, hvis A er invertibel og $A^* = A^{-1}$, da er $f \colon W \to V$ en isomorfi, og at $f \colon W \to V$ tillige er en isometri følger fra (1), da $A^*A = A^{-1}A = I_n$. □

Eksempel 6.3.11 (1) Afbildningen $f \colon \mathbb{R}^2 \to \mathbb{R}^3$ fra eksempel 6.3.8 er ikke en isometri med hensyn til standard-indreprodukterne. For udregningen

$$A^*A = \begin{pmatrix} 2 & 0 & 1 \\ 3 & 1 & -1 \end{pmatrix} \begin{pmatrix} 2 & 3 \\ 0 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 5 & 5 \\ 5 & 11 \end{pmatrix}$$

viser, at den påkrævede identitet $A^*A = I_2$ ikke gælder.
(2) Vi har allerede vist, at rotationsafbildningen $r_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ fra eksempel 6.3.2 er en isometrisk isomorfi med hensyn til standard-indreprodukterne, men udregningen

$$B_\theta^* = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix} = B_\theta^{-1}$$

bekræfter, at dette er tilfældet.

Bemærkning 6.3.12 For en kvadratisk matrix A er betingelsen, at $A^*A = I_n$, ensbetydende med, at familien $(\boldsymbol{a}_1, \dots, \boldsymbol{a}_n)$ af søjler i A er ortonormal med hensyn til standard-indreproduktet på $\mathbb{F}^n$. For den (i,j)'te indgang i A^*A er produktet af den i'te række $\boldsymbol{a}_i^*$ i A^* og den j'te søjle $\boldsymbol{a}_j$ i A, hvilket er netop er $\langle \boldsymbol{a}_i, \boldsymbol{a}_j \rangle = \boldsymbol{a}_i^* \boldsymbol{a}_j$. Derfor gælder der, altså at $A^*A = I$, hvis og kun hvis $\langle \boldsymbol{a}_i, \boldsymbol{a}_j \rangle = \delta_{ij}$ for alle $1 \leq i, j \leq n$.
(2) Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, og lad $(V, \langle -,- \rangle_V)$ være et $\mathbb{F}$-vektorrum med indre produkt af endelig dimension n. Vi husker fra sætning 4.2.8, at en n-tuple $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ af vektorer i V er en basis, hvis og kun hvis den lineær afbildning $h \colon \mathbb{F}^n \to V$ defineret ved $h(\boldsymbol{x}) = \boldsymbol{v}_1 x_1 + \dots + \boldsymbol{v}_n x_n$ er en isomorfi. I givet fald er denne afbildning en isometri med hensyn til standard-indreproduktet $\langle -,- \rangle$ på $\mathbb{F}^n$ og det givne indre produkt $\langle -,- \rangle_V$ på V, hvis og kun hvis basen $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ er ortonormal med hensyn til $\langle -,- \rangle_V$.

Vi vil i resten af dette afsnit udelukkende betragte lineære endomorfier, hvilket er linære afbildninger, hvis domæne og codomæne er identitiske.

Definition 6.3.13 Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$ og lad $(V, \langle -,- \rangle)$ være et endeligdimensionalt $\mathbb{F}$-vektorrum med indre produkt. Lad $f \colon V \to V$ være en lineær endomorfi, og lad $g \colon V \to V$ være den adjungerede lineære endomorfi med hensyn til $\langle -,- \rangle$.

(1) Endomorfien $f \colon V \to V$ er *selvadjungeret* med hensyn til $\langle -,- \rangle$, hvis $f = g$.

(2) Endomorfien $f \colon V \to V$ er *normal* med hensyn til $\langle -,- \rangle$, hvis $f \circ g = g \circ f$.

For lineære endomorfier specialiserer resultater ovenfor til følgende resultat.

Sætning 6.3.14 *Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, lad $(V, \langle -, - \rangle)$ være et endeligdimensionalt $\mathbb{F}$-vektorrum med indre produkt, og lad $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ være en basis for V, der er ortonormal med hensyn til $\langle -, - \rangle$. Lad $f \colon V \to V$ være en lineær endomorfi, og lad $A \in M_n(\mathbb{F})$ være matricen, der repræsenterer $f \colon V \to V$ med hensyn den givne ortonormale basis.*

(1) *Endomorfien $f \colon V \to V$ er en isometri med hensyn til $\langle -, - \rangle$, hvis og kun hvis*

$$A^* A = I.$$

Dette er endvidere tilfældet, hvis og kun hvis $A^* = A^{-1}$.

(2) *Endomorfien $f \colon V \to V$ er selvadjungeret med hensyn til $\langle -, - \rangle$, hvis og kun hvis*

$$A^* = A.$$

(3) *Endomorfien $f \colon V \to V$ er normal med hensyn til $\langle -, - \rangle$, hvis og kun hvis*

$$A^* A = A\, A^*.$$

Bevis Påstanden (1) følger fra sætning 6.3.7 og lemma 6.3.9, samt sætning 2.5.15 idet A er kvadratisk. Påstande (2) og (3) følger umiddelbart fra sætning 6.3.7 og definition 6.3.13. □

De klassiske betegnelser for de kvadratiske matricer $A \in M_n(\mathbb{F})$, der opfylder enten $A^* = A^{-1}$ eller $A^* = A$, afhænger desværre af $\mathbb{F}$. Det er der ikke noget at gøre ved, så vi anfører disse klassiske betegnelser i den følgende definition, hvor vi også genkalder nogle definitioner fra afsnit 2.6.

Definition 6.3.15 Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$.

(1) En kvadratisk matrix $Q \in M_m(\mathbb{R})$, der opfylder $Q^* = Q^{-1}$, kaldes *ortogonal*; og en kvadratisk matrix $U \in M_m(\mathbb{C})$, der opfylder $U^* = U^{-1}$, kaldes *unitær*.

(2) En kvadratisk matrix $A \in M_m(\mathbb{R})$, der opfylder $A^* = A$, kaldes *symmetrisk*; og en kvadratisk matrix $A \in M_m(\mathbb{C})$, der opfylder $A^* = A$, kaldes *hermitisk*.

(3) En kvadratisk matrix $A \in M_m(\mathbb{F})$, der opfylder $A^* A = AA^*$, kaldes *normal*.

Vi bemærker, at i det reelle tilfælde er $P^* = P^t$ lig med den transponerede matrix, og at i begge tilfælde er $P^* = P^{-1}$, hvis og kun hvis familien $(\boldsymbol{p}_1, \dots, \boldsymbol{p}_m)$ af søjler i P er ortonormal med hensyn til standard-indreproduktet på $\mathbb{F}^m$.

Eksempel 6.3.16 Enhver symmetrisk reel matrix er normal, men det omvendte er ikke tilfældet. For eksempel er matricen

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

er normal, idet

$$A^*A = \begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix} = AA^*,$$

men A er ikke symmetrisk. Bemærk også, at hvis B er en reel kvadratisk matrix, da er matricen $A = B^*B$ altid symmetrisk. Det tilsvarende gælder i det komplekse tilfælge.

Sætning 6.3.17 *Lad enten* $\mathbb{F} = \mathbb{R}$ *eller* $\mathbb{F} = \mathbb{C}$.

(1) *Hvis både* $A, B \in M_m(\mathbb{F})$ *opfylder* $A^* = A^{-1}$ *og* $B^* = B^{-1}$, *da er også* $(AB)^* = (AB)^{-1}$.

(2) *Hvis* $A \in M_m(\mathbb{F})$ *opfylder* $A^* = A^{-1}$, *da er også* $(A^{-1})^* = (A^{-1})^{-1}$.

Bevis De to påstande følger fra de to udregninger

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

hvor vi anvender sætning 2.6.5. □

Bemærkning 6.3.18 En kvadratisk matrix $S \in M_m(\mathbb{H})$, der opfylder $S^* = S^{-1}$ kaldes en *unitær symplektisk* matrix, og sætning 6.3.17 gælder ligeledes i dette tilfælde. Man anvender klassisk betegnelserne

$$O(m) \subset GL_m(\mathbb{R}), \qquad U(m) \subset GL_m(\mathbb{C}) \qquad \text{og} \qquad \mathrm{Sp}(m) \subset GL_m(\mathbb{H})$$

for delmængden af de invertible $m \times m$-matricer, der består af de $m \times m$-matricer A, der opfylder $A^* = A^{-1}$. Sætning 6.3.17 viser, at $(O(m), \cdot)$, $(U(m), \cdot)$ og $(\mathrm{Sp}(m), \cdot)$ er grupper, og de kaldes den *ortogonale gruppe* den *unitære gruppe* og den *symplektiske gruppe*.

I modsætning hertil så behøver en kvadratisk matrix A, der opfylder $A^* = A$, ikke at være invertibel; og hvis både $A^* = A$ og $B^* = B$, så er $(AB)^* = B^*A^* = BA$, hvilket generelt heller ikke er lig med AB. Så selvadjungerede matricer udgør ikke en gruppe.

6.4 Spektralsætningen

I dette afsnit viser vi de såkaldte spektralsætninger for lineære endomorfier af endeligt dimensionale reelle og komplekse vektorrum med indre produkt. Sætningen siger, at der findes en ortonormal basis for vektorrummet, der består af egenvektorer for den lineære endomorfi, hvis og kun hvis denne er selvadjungeret i det reelle tilfælde og normal i det komplekse tilfælde. Beviset anvender algebraens fundamentalsætning til at faktorisere det karakteristike polynomium hørende til den lineære endomorfi som et produkt af førstegradspolynomier, og denne del af beviset er derfor ikke konstruktivt. Men givet denne faktorisering, er resten af beviset algoritmisk, hvilket vi illustrerer med nogle passende eksempler.

Sætning 6.4.1 *Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$. En matrix $A \in M_n(\mathbb{F})$ er både triangulær og normal, hvis og kun hvis den er en diagonal matrix.*

Bevis Hvis $A = \operatorname{diag}(a_1, \dots, a_n)$ er en diagonal matrix, så er A specielt en triangulær matrix, og udregningen

$$\begin{aligned} A^*A &= \operatorname{diag}(a_1,\dots,a_n)^* \operatorname{diag}(a_1,\dots,a_n) = \operatorname{diag}(a_1^*a_1,\dots,a_n^*a_n), \\ AA^* &= \operatorname{diag}(a_1,\dots,a_n)\operatorname{diag}(a_1,\dots,a_n)^* = \operatorname{diag}(a_1a_1^*,\dots,a_na_n^*) \end{aligned}$$

viser, at A er normal, idet den kommutativ lov for multiplikation gælder i $\mathbb{F}$. Så vi antager omvendt, at $A = (a_{ij}) \in M_n(\mathbb{F})$ er en normal øvre triangulær matrix og viser, at A er en diagonal matrix; beviset i tilfældet, hvor A er en normal nedre triangulær matrix er tilsvarende. Da A er normal, gælder der for alle $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{F}^n$, at

$$\langle A\boldsymbol{x}, A\boldsymbol{y}\rangle = \langle A^*A\boldsymbol{x}, \boldsymbol{y}\rangle = \langle AA^*\boldsymbol{x}, \boldsymbol{y}\rangle = \langle (A^*)^*A^*\boldsymbol{x}, \boldsymbol{y}\rangle = \langle A^*\boldsymbol{x}, A^*\boldsymbol{y}\rangle,$$

og ved at sætte $\boldsymbol{x} = \boldsymbol{y} = \boldsymbol{e}_k$, konkluderer vi derfor, at

$$|a_{1k}|^2 + \dots + |a_{nk}|^2 = \langle A\boldsymbol{e}_k, A\boldsymbol{e}_k\rangle = \langle A^*\boldsymbol{e}_k, A^*\boldsymbol{e}_k\rangle = |a_{k1}|^2 + \dots + |a_{kn}|^2$$

for alle $1 \le k \le n$. Da A endvidere er øvre triangulær, er $a_{ij} = 0$ for $i > j$, og det følger derfor fra den identitet, vi netop har fundet, at

$$|a_{1k}|^2 + \dots + |a_{kk}|^2 = |a_{kk}|^2 + \dots + |a_{kn}|^2$$

for alle $1 \le k \le n$. Vi viser ved induktion på $1 \le i \le n$, at $a_{ij} = 0$ for $i < j$, og begynder med tilfældet $i = 1$. Den ovenstående identitet for $k = 1$ viser, at

$$|a_{11}|^2 = |a_{11}|^2 + |a_{12}|^2 + \dots + |a_{1n}|^2,$$

hvorfor $a_{12} = \cdots = a_{1n} = 0$ som ønsket. Så vi antager, at påstanden er vist for $i = r-1$, og viser den for $i = r$. Der gælder da, at

$$|a_{rr}|^2 = |a_{1r}|^2 + \cdots + |a_{rr}|^2 = |a_{rr}|^2 + \cdots + |a_{rn}|^2,$$

idet den første identitet følger fra den induktive hypotese, mens den anden identitet er den identitet, vi viste ovenfor, for $k = r$. Derfor er $a_{r,r+1} = \cdots = a_{rn} = 0$, hvilket viser induktionsskridtet og dermed påstanden. Der gælder altså, at $a_{ij} = 0$ for $i \neq j$, hvilket som ønsket viser, at A er en diagonal matrix. □

Den følgende sætning kaldes for spektralsætningen for symmetriske matricer.

Sætning 6.4.2 *For en kvadratisk matrix $A \in M_n(\mathbb{R})$ er følgende ækvivalent:*

(1) *Matricen A er symmetrisk.*

(2) *Der findes en ortogonal matrix $Q \in M_n(\mathbb{R})$, sådan at $Q^{-1}AQ$ er en diagonal matrix.*

Bevis Vi antager først (1) og viser (2). Ifølge sætning 5.2.8 findes der en invertibel matrix $P \in M_n(\mathbb{R})$, sådan at matricen $P^{-1}AP \in M_n(\mathbb{R})$ er øvre triangulær. Endvidere producerer Gram-Schmidt algoritmen en øvre triangulær matrix $C \in M_n(\mathbb{R})$, sådan at matricen $Q = PC \in M_n(\mathbb{R})$ er ortogonal. For hvis $(\boldsymbol{p}_1, \dots, \boldsymbol{p}_n)$ er basen, der består af søjlerne i P, så giver Gram-Schmidt algoritmen en basis $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$, der er ortonormal med hensyn til standard-indreproduktet på $\mathbb{R}^n$, hvor

$$\begin{aligned}
\boldsymbol{u}_1 &= \boldsymbol{p}_1 c_{11} \\
\boldsymbol{u}_2 &= \boldsymbol{p}_1 c_{12} + \boldsymbol{p}_2 c_{22} \\
&\ \vdots \\
\boldsymbol{u}_n &= \boldsymbol{p}_1 c_{1n} + \boldsymbol{p}_2 c_{2n} + \cdots + \boldsymbol{p}_n c_{nn},
\end{aligned}$$

og Q er da den ortogonale matrix med søjler $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n$, mens C er den øvre triangulære matrix med indgange c_{ij}. Vi påstår, at $Q^{-1}AQ$ er både normal og øvre triangulær. For

$$\begin{aligned}
(Q^{-1}AQ)^*Q^{-1}AQ &= (Q^*AQ)^*Q^*AQ = Q^*A^*QQ^*AQ = Q^*A^*AQ \\
Q^{-1}AQ(Q^{-1}AQ)^* &= Q^*AQ(Q^*AQ)^* = Q^*AQQ^*A^*Q = Q^*AA^*Q,
\end{aligned}$$

hvor vi bruger at $Q^* = Q^{-1}$. Da A er symmetrisk, er den specielt normal, så vi har nu vist, at $Q^{-1}AQ$ er normal. Endvidere er

$$Q^{-1}AQ = (PC)^{-1}APC = C^{-1}P^{-1}APC,$$

hvor $P^{-1}AP$ og C er øvre triangulære matricer. Men produktet af to øvre triangulære matricer er igen en øvre triangulær matrix, og sætning 3.4.1 viser, at også den inverse matrix af en invertibel øvre triangulær matrix igen er øvre triangulær. Dette viser vores påstand, og sætning 6.4.1 viser, at denne medfører (2).

Vi antager dernæst (2) og viser (1). Så lad $Q \in M_n(\mathbb{R})$ være en ortogonal matrix, sådan at $Q^{-1}AQ = D$ er en diagonal matrix. Da D er symmetrisk, så viser udregningen

$$A^* = (QDQ^{-1})^* = (QDQ^*)^* = QD^*Q^* = QDQ^* = A,$$

at også A er symmetrisk. Dette viser (1) og dermed sætningen. $\square$

Den næste sætning kaldes for spektralsætningen for normale komplekse matricer.

Sætning 6.4.3 *For en kompleks kvadratisk matrix $A \in M_n(\mathbb{C})$ er følgende ækvivalent:*

(1) *Matricen A er normal.*

(2) *Der findes en unitær matrix $U \in M_n(\mathbb{C})$, sådan at $U^{-1}AU$ er en diagonal matrix.*

Bevis Beviset for, at (1) medfører (2) er helt som beviset for den tilsvarende implikation i sætning 6.4.2. Den eneste forskel er, at "$\mathbb{R}$" og "ortogonal" erstattes med henholdsvis "$\mathbb{C}$" og "unitær". Vi antager derfor (2) og viser (1). Så lad $U \in M_n(\mathbb{C})$ være en unitær matrix, sådan at $U^{-1}AU = D$ er en diagonal matrix. Da er $A = UDU^{-1} = UDU^*$, og da D er normal, så viser udregningen

$$\begin{aligned} A^*A &= (UDU^*)^*UDU^* = UD^*U^*UDU^* = UD^*DU^*, \\ AA^* &= UDU^*(UDU^*)^* = UDU^*UD^*U^* = UDD^*U^* \end{aligned}$$

at også A er normal. Dette viser (1) og dermed sætningen. $\square$

Bemærkning 6.4.4 Vi anvender sætning 6.4.3 til at vise, at en normal kompleks matrix $A \in M_n(\mathbb{C})$ er hermitisk, hvis og kun hvis alle dens egenværdier er reelle. Da A er normal, så findes der ifølge sætningen en unitær matrix $U \in M_n(\mathbb{C})$, sådan at

$$U^*AU = \operatorname{diag}(\lambda_1, \dots, \lambda_n) \in M_n(\mathbb{C})$$

er en diagonal matrix D. Hvis nu $A = A^*$ er hermitisk, så er

$$D^* = (U^*AU)^* = U^*A^*U = U^*AU = D,$$

hvilket viser, at D også er hermitisk, og derfor er egenværdierne $\lambda_1, \dots, \lambda_n$ alle er reelle. Omvendt, hvis egenværdierne $\lambda_1, \dots, \lambda_n$ alle er reelle, da er $D = D^*$ hermitisk, og den tilsvarende udregning

$$A^* = (UDU^*) = UD^*U^* = UDU^* = A$$

viser da, at A er hermitisk, som ønsket.

Vi oversætter nu sætning 6.4.2 og sætning 6.4.3 til udsagn om lineære endomorfier af henholdsvis reelle og komplekse vektorrum med indre produkt.

Sætning 6.4.5 *Lad $(V, \langle -,- \rangle)$ være et endeligdimensionalt reelt vektorrum med indre produkt. For en lineær endomorfi $f \colon V \to V$ er følgende ækvivalent:*

(1) *Endomorfien $f \colon V \to V$ er selvadjungeret.*

(2) *Der findes en basis for V, der både er ortonormal med hensyn til $\langle -,- \rangle$ og består af egenvektorer for $f \colon V \to V$.*

Bevis Vi antager først (1) og viser (2). Vi vælger en basis $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ for V, som er ortonormal med hensyn til $\langle -,- \rangle$, hvilket er muligt ifølge korollar 6.2.10. Da den lineære endomorfi $f \colon V \to V$ er selvadjungeret med hensyn til $\langle -,- \rangle$, viser sætning 6.3.14, at matricen $A \in M_n(\mathbb{R})$, der repræsenterer $f \colon V \to V$ med hensyn til basen $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ for både domænet og codomænet, er symmetrisk. Derfor findes der ifølge sætning 6.4.2 en ortogonal matrix $Q = (q_{ij}) \in M_n(\mathbb{R})$, sådan at

$$Q^{-1}AQ = \operatorname{diag}(a_1, \dots, a_n)$$

er en diagonal matrix. Da $Q \in M_n(\mathbb{R})$ er ortogonal, er familien $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$, hvor

$$\boldsymbol{u}_j = \boldsymbol{v}_1 q_{1j} + \boldsymbol{v}_2 q_{2j} + \dots + \boldsymbol{v}_n q_{nj},$$

en basis for V, der er ortonormal med hensyn til $\langle -,- \rangle$. For

$$\begin{aligned} \langle \boldsymbol{u}_i, \boldsymbol{u}_j \rangle &= \langle \boldsymbol{v}_1 q_{1i} + \boldsymbol{v}_2 q_{2i} + \dots + \boldsymbol{v}_n q_{ni}, \boldsymbol{v}_1 q_{1j} + \boldsymbol{v}_2 q_{2j} + \dots + \boldsymbol{v}_n q_{nj} \rangle \\ &= q_{1i} q_{1j} + q_{2i} q_{2j} + \dots + q_{ni} q_{nj} = \boldsymbol{q}_i^* \boldsymbol{q}_j = \delta_{ij}. \end{aligned}$$

Desuden er Q den matrix, der repræsenterer identitetsafbildningen $\mathrm{id}_V \colon V \to V$ med hensyn til basen $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$ for domænet og basen $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ for codomænet. Derfor viser sætning 4.3.11 at matricen, der repræsenterer $f \colon V \to V$ med hensyn til basen $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$ for både domænet og codomænet, er diagonalmatricen

$$Q^{-1}AQ = \operatorname{diag}(\lambda_1, \dots, \lambda_n) \in M_n(\mathbb{R}).$$

Med andre ord er $\boldsymbol{u}_j$ en egenvektor for $f \colon V \to V$ hørende til egenværdien λ_j, hvilket viser (2). Den følgende figur illustrerer koordinatskiftet.

$$\begin{tikzcd}
(\boldsymbol{v}_1,\dots,\boldsymbol{v}_n) \arrow[rrr, bend left, dotted, "A"] & & & (\boldsymbol{v}_1,\dots,\boldsymbol{v}_n) \\
& V \arrow[r, "f"] & V & \\
& V \arrow[r, "f"] \arrow[u, "\mathrm{id}_V"'] & V \arrow[u, "\mathrm{id}_V"'] & \\
(\boldsymbol{u}_1,\dots,\boldsymbol{u}_n) \arrow[uuu, bend left, dotted, "Q"] \arrow[rrr, bend right, dotted, "Q^{-1}AQ"] & & & (\boldsymbol{u}_1,\dots,\boldsymbol{u}_n) \arrow[uuu, bend right, dotted, "Q"']
\end{tikzcd}$$

Vi antager omvendt (2) og viser (1). Så vi lader $(\boldsymbol{u}_1,\dots,\boldsymbol{u}_n)$ være en basis for V, der er ortonormal med hensyn til $\langle -,- \rangle$ og består af egenvektorer for $f\colon V \to V$. Lad $\lambda_j \in \mathbb{R}$ være egenværdien, sådan at $f(\boldsymbol{u}_j) = \boldsymbol{u}_j\lambda_j$. Da er $D = \mathrm{diag}(\lambda_1,\dots,\lambda_n) \in M_n(\mathbb{R})$ matricen, der repræsenterer $f\colon V \to V$ med hensyn til den valgte basis for både domænet og codomænet. Da D er symmetrisk, og da $(\boldsymbol{u}_1,\dots,\boldsymbol{u}_n)$ er ortonormal med hensyn til $\langle -,- \rangle$, viser sætning 6.3.7 derfor, at $f\colon V \to V$ er selvadjungeret med hensyn til $\langle -,- \rangle$. Dette viser (1) og dermed sætningen. □

Sætning 6.4.6 *Lad $(V, \langle -,- \rangle)$ være et endeligdimensionalt komplekst vektorrum med indre produkt. For en lineær endomorfi $f\colon V \to V$ er følgende ækvivalent:*

(1) *Endomorfien $f\colon V \to V$ er normal.*

(2) *Der findes en basis for V, der både er ortonormal med hensyn til $\langle -,- \rangle$ og består af egenvektorer for $f\colon V \to V$.*

Bevis Beviset er tilsvarende beviset for sætning 6.4.5, idet vi anvender sætning 6.4.3 i stedet for sætning 6.4.2 og erstatter "$\mathbb{R}$" med "$\mathbb{C}$", "symmetrisk" og "selvadjungeret" med "normal", og "ortogonal" med "unitær". □

Vi bemærker, at spektralsætningerne og deres beviser kun siger, at en ortonormal basis af egenvektorer findes, men de siger ikke hvordan man bærer sig ad med at finde en sådan basis. Problemet er at finde rødder i det karateristiske polynomium. Hvis det er gjort, så findes der en algoritme, som vi illustrerer ved nogle eksempler nedenfor, der angiver den ønskede basis. Undervejs kan man spare sig en del arbejde ved at bemærke sig følgende.

Korollar 6.4.7 *Lad $A \in M_n(\mathbb{F})$ være enten en symmetrisk reel matrix eller en normal kompleks matrix. Egenvektorer $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{F}^n$ der hører til forskellige egenværdier $\lambda, \mu \in \mathbb{F}$, er nødvendigvis ortogonale med hensyn til standard-indreproduktet på $\mathbb{F}^n$.*

Bevis Vi viser det komplekse tilfælde; det reelle tilfælde er helt tilsvarende. Vi lader $\lambda, \mu \in \mathbb{F}$ være to forskellige egenværdier for A og angiver baser for egenrummene $N_{A-I\lambda}$ og $N_{A-I\mu}$ som følger. Ifølge sætning 6.4.3 findes der en unitær matrix $U \in M_n(\mathbb{C})$, sådan at $U^*AU = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$ er en diagonalmatrix. Så hvis $I = \{i \in \{1, \dots, n\} \mid \lambda_i = \lambda\}$ og $J = \{j \in \{1, \dots, n\} \mid \lambda_j = \mu\}$, da er $(U\boldsymbol{e}_i \mid i \in I)$ og $(U\boldsymbol{e}_j \mid j \in J)$ baser for henholdsvis $N_{A-I\lambda}$ og $N_{A-I\mu}$. Hvis nu $\boldsymbol{x} = \sum_{i \in I} U\boldsymbol{e}_i a_i \in N_{A-I\lambda}$ og $\boldsymbol{y} = \sum_{j \in J} U\boldsymbol{e}_j b_j \in N_{A-I\mu}$, da er

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \sum_{i \in I, j \in J} a_i^* \langle U\boldsymbol{e}_i, U\boldsymbol{e}_j \rangle b_j = \sum_{i \in I, j \in J} a_i^* \langle \boldsymbol{e}_i, \boldsymbol{e}_j \rangle b_j = 0$$

som ønsket, idet $I \cap J = \emptyset$. Se også sætning 6.3.14. □

Eksempel 6.4.8 Vi ønsker at finde en basis for $\mathbb{R}^3$, der er ortonormal med hensyn til standard-indreproduktet, og som endvidere består af egenvektorer for den lineære endomorfi $f\colon \mathbb{R}^3 \to \mathbb{R}^3$ givet ved $f(\boldsymbol{x}) = A\boldsymbol{x}$, hvor $A \in M_3(\mathbb{R})$ er den symmetriske matrix

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

Vi ved fra Sætning 6.4.2 at dette er muligt. For at bestemme egenværdierne af A udregner vi først det karakteristiske polynomium.

$$\begin{aligned}
\chi_A(t) = \det(A - It) &= \det \begin{pmatrix} 1-t & 2 & -1 \\ 2 & -2-t & 2 \\ -1 & 2 & 1-t \end{pmatrix} \\
&\overset{(D6')}{=} \det \begin{pmatrix} 0 & 2(2-t) & -t(2-t) \\ 0 & 2-t & 2(2-t) \\ -1 & 2 & 1-t \end{pmatrix} \overset{(D2')}{=} (2-t)^2 \det \begin{pmatrix} 0 & 2 & -t \\ 0 & 1 & 2 \\ -1 & 2 & 1-t \end{pmatrix} \\
&\overset{(D6')}{=} (2-t)^2 \det \begin{pmatrix} 0 & 0 & -4-t \\ 0 & 1 & 2 \\ -1 & 2 & 1-t \end{pmatrix} \overset{(D2')}{=} (2-t)^2(-4-t) \det \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 2 \\ -1 & 2 & 1-t \end{pmatrix} \\
&\overset{(D5')}{=} -(2-t)^2(-4-t) \det \begin{pmatrix} -1 & 2 & 1-t \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} = (2-t)^2(-4-t)
\end{aligned}$$

Egenværdierne for A er altså $\lambda_1 = 2$ og $\lambda_2 = -4$. Her har vi benyttet sætning 3.3.3 til at udregne determinanten af den triangulære matrix.

Vi anvender dernæst Gauss elimination til at finde baser for egenrummene hørende til $\lambda_1 = 2$ og $\lambda_2 = -4$. Så vi omdanner matricen $B = A - I \cdot 2$ til en rækkeækvivalent matrix B' på reduceret echelon-form:

$$B = \begin{pmatrix} -1 & 2 & -1 \\ 2 & -4 & 2 \\ -1 & 2 & -1 \end{pmatrix} \begin{matrix} \\ +2 \times R_1 \\ +(-1) \times R_1 \end{matrix}$$

$$\begin{pmatrix} -1 & 2 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{matrix} (-1) \times R_1 \\ \\ \\ \end{matrix}$$

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

Egenrummet $N_{A-I\cdot 2}$ hørende til $\lambda_1 = 2$, som er løsningsmængden til ligningssystemet $B\boldsymbol{x} = \boldsymbol{0}$, kan derfor aflæses fra den rækkeækvivalent matrix B' til at have basis

$$\left(\boldsymbol{w}_1 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \quad \boldsymbol{w}_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}\right).$$

Tilsvarende omdanner vi matricen $C = A + I \cdot 4$ til en rækkeækvivalent matrix C' på reduceret echelon-form:

$$C = \begin{pmatrix} 5 & 2 & -1 \\ 2 & 2 & 2 \\ -1 & 2 & 5 \end{pmatrix} \begin{matrix} +5 \times R_3 \\ +2 \times R_3 \\ \\ \end{matrix}$$

$$\begin{pmatrix} 0 & 12 & 24 \\ 0 & 6 & 12 \\ -1 & 2 & 5 \end{pmatrix} \begin{matrix} (1/12) \times R_1 \\ (1/6) \times R_2 \\ (-1) \times R_3 \end{matrix}$$

$$\begin{pmatrix} 0 & 1 & 2 \\ 0 & 1 & 2 \\ 1 & -2 & -5 \end{pmatrix} \begin{matrix} +(-1) \times R_2 \\ \\ +2 \times R_2 \end{matrix}$$

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 2 \\ 1 & 0 & -1 \end{pmatrix} \begin{matrix} \\ R_1 \leftrightarrow R_3 \\ \\ \end{matrix}$$

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

Egenrummet $N_{A+I\cdot 4}$ hørende til $\lambda_2 = -4$, som er løsningsmængden til ligningssystemet

$C\boldsymbol{x} = \boldsymbol{0}$, kan derfor aflæses fra C' til at have basis

$$\left(\boldsymbol{w}_3 = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}\right)$$

idet x_3 er eneste frie variabel.

Vi har nu fundet en basis $(\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3)$ for $\mathbb{R}^3$, bestående af egenvektorer for A. Denne basis er dog ikke ortonormal med hensyn til standard-indreproduktet, og vi anvender derfor Gram-Schmidt algoritmen til at omdanne den til en ortonormal basis $(\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3)$. Bemærk dog, at vi allerede ved fra korollar 6.4.7, at $\boldsymbol{w}_3$ er ortogonal til både $\boldsymbol{w}_1$ og $\boldsymbol{w}_2$, idet de tilhører forskellige egenværdier for den symmetriske matrix A. Vi finder først den ortogonale basis $(\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3)$, hvor

$$\begin{aligned}
\boldsymbol{v}_1 &= \boldsymbol{w}_1 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} \\
\boldsymbol{v}_2 &= \boldsymbol{w}_2 - \boldsymbol{v}_1 \cdot \frac{\langle \boldsymbol{v}_1, \boldsymbol{w}_2 \rangle}{\langle \boldsymbol{v}_1, \boldsymbol{v}_1 \rangle} = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} - \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} \cdot \frac{-2+0+0}{4+1+0} = \begin{pmatrix} -1 \\ 2 \\ 5 \end{pmatrix} \cdot \frac{1}{5} \\
\boldsymbol{v}_3 &= \boldsymbol{w}_3 = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}.
\end{aligned}$$

Endelig får vi ved normalisering af vektorerne i denne basis den ortonormale basis

$$\left(\boldsymbol{u}_1 = \begin{pmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \\ 0 \end{pmatrix},\quad \boldsymbol{u}_2 = \begin{pmatrix} -1/\sqrt{30} \\ 2/\sqrt{30} \\ 5/\sqrt{30} \end{pmatrix},\quad \boldsymbol{u}_3 = \begin{pmatrix} 1/\sqrt{6} \\ -2/\sqrt{6} \\ 1/\sqrt{6} \end{pmatrix}\right),$$

der består af egenvektorer for A. Den ortogonale matrix $Q \in M_3(\mathbb{R})$, der repræsenterer identitetsafbildningen $\mathrm{id}\colon \mathbb{R}^3 \to \mathbb{R}^3$ med hensyn til den nye basis $(\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3)$ for domænet og standardbasen $(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$ for codomænet, er da givet ved

$$Q = \begin{pmatrix} 2/\sqrt{5} & -1/\sqrt{30} & 1/\sqrt{6} \\ 1/\sqrt{5} & 2/\sqrt{30} & -2/\sqrt{6} \\ 0 & 5/\sqrt{30} & 1/\sqrt{6} \end{pmatrix} \in M_3(\mathbb{R}),$$

og ifølge sætning 4.3.11 gælder der endvidere, at

$$Q^{-1}AQ = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -4 \end{pmatrix} \in M_3(\mathbb{R}),$$

idet $\boldsymbol{u}_1$, $\boldsymbol{u}_2$ og $\boldsymbol{u}_3$ er egenvektorer for $f\colon \mathbb{R}^3 \to \mathbb{R}^3$ hørende til egenværdierne 2, 2 og -4.

Lad os endelig sige lidt om, hvordan man opdager mulige fejl i en udregning som denne. Den største risiko for fejl er i udregningen af det karakteristiske polynomium, og en fejl i denne udregning vil gøre alle følgende udregninger forkerte. En sådan fejl er dog let at opdage. For hvis λ *ikke* er en egenværdi for A, så er $B = A - I\lambda$ en invertibel matrix, og ved Gauss elimination vil vi derfor nødvendigvis få, at den tilhørende rækkeækvivalente matrix på reduceret echelon-form er $B' = I$. Så hvis Gauss elimination af $B = A - I\lambda$ giver $B' = I$, så ved vi altså, at λ *ikke* er en egenværdi.

Eksempel 6.4.9 Vi ønsker at finde en basis for $\mathbb{C}^2$, der er ortonormal med hensyn til standard-indreproduktet og består af egenvektorer for den lineære endomorfi $f : \mathbb{C}^2 \to \mathbb{C}^2$ givet ved $f(\boldsymbol{x}) = A\boldsymbol{x}$, hvor $A \in M_2(\mathbb{C})$ er den normale matrix

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

fra Eksempel 6.3.16. Det karakteristiske polynomium er givet ved

$$\chi_A(t) = \det(A - It) = \det\begin{pmatrix} 2-t & -1 \\ 1 & 2-t \end{pmatrix} = (2-t)^2 + 1 = t^2 - 4t + 5,$$

og det har discriminant $\Delta = (-4)^2 - 4 \cdot 1 \cdot 5 = -4$ og rødder

$$\lambda_1 = \frac{-(-4) + \sqrt{-4}}{2 \cdot 1} = 2 + i \qquad \text{og} \qquad \lambda_2 = \frac{-(-4) - \sqrt{-4}}{2 \cdot 1} = 2 - i.$$

Vi anvender nu Gauss elimination til at bestemme baser for de to egenrum, der begge nødvendigvis er 1-dimensionale. Vi omdanner først $B = A - I\lambda_1$ til en rækkeækvivalent matrix B' på reduceret echelon-form.

$$\begin{aligned} B = &\begin{pmatrix} -i & -1 \\ 1 & -i \end{pmatrix} \quad +i \times R_2 \\ &\begin{pmatrix} 0 & 0 \\ 1 & -i \end{pmatrix} \quad R_1 \leftrightarrow R_2 \\ B' = &\begin{pmatrix} 1 & -i \\ 0 & 0 \end{pmatrix} \end{aligned}$$

Herved finder vi basen

$$(\boldsymbol{v}_1 = \begin{pmatrix} i \\ 1 \end{pmatrix})$$

for egenrummet $N_{A - I\lambda_1}$ hørende til $\lambda_1 = 2 + i$. Vi omdanner ligeledes $C = A - I\lambda_2$ til en matrix C' på reduceret echelon-form.

$$\begin{aligned} C = &\begin{pmatrix} i & -1 \\ 1 & i \end{pmatrix} \quad +(-i) \times R_2 \\ &\begin{pmatrix} 0 & 0 \\ 1 & i \end{pmatrix} \quad R_1 \leftrightarrow R_2 \\ C' = &\begin{pmatrix} 1 & i \\ 0 & 0 \end{pmatrix} \end{aligned}$$

Herved finder vi basen

$$\left(\boldsymbol{v}_2 = \begin{pmatrix} -i \\ 1 \end{pmatrix}\right)$$

for egenrummet $N_{A-I\lambda_2}$ hørende til $\lambda_2 = 2-i$. Basen $(\boldsymbol{v}_1, \boldsymbol{v}_2)$ for $\mathbb{C}^2$ er allerede at være ortogonal, som den skal være ifølge korollar 6.4.7, så vi behøver ikke at anvende Gram-Schmidt. Endelig får vi ved normalisering direkte den ortonormale basis

$$\left(\boldsymbol{u}_1 = \begin{pmatrix} -i/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix},\quad \boldsymbol{u}_2 = \begin{pmatrix} i/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}\right),$$

for $\mathbb{C}^2$, der består af egenvektorer for $f\colon \mathbb{C}^2 \to \mathbb{C}^2$. Den unitære matrix $U \in M_2(\mathbb{C})$, der repræsenterer identitetsafbildningen $\mathrm{id}\colon \mathbb{C}^2 \to \mathbb{C}^2$ med hensyn til den nye basis $(\boldsymbol{u}_1, \boldsymbol{u}_2)$ for domænet og standardbasen $(\boldsymbol{e}_1, \boldsymbol{e}_2)$ for codomænet, er da givet ved

$$U = \begin{pmatrix} i/\sqrt{2} & -i/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \in M_2(\mathbb{C}),$$

og sætning 4.3.11 viser derfor, at

$$U^{-1}AU = \begin{pmatrix} 2+i & 0 \\ 0 & 2-i \end{pmatrix} \in M_2(\mathbb{C}),$$

idet $\boldsymbol{u}_1$ og $\boldsymbol{u}_2$ er egenvektorer for $f\colon \mathbb{C}^2 \to \mathbb{C}^2$ med egenværdier henholdsvis $2+i$ og $2-i$.

6.5 Klassifikation af hermitiske former

Vi lader fortsat $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$ og lader $(-)^*\colon \mathbb{F} \to \mathbb{F}$ være skævinvolutionen givet ved

$$a^* = \begin{cases} a & \text{hvis } \mathbb{F} = \mathbb{R}, \\ \bar{a} & \text{hvis } \mathbb{F} = \mathbb{C}. \end{cases}$$

Vi betragter hermitiske former $\langle -,- \rangle\colon V \times V \to \mathbb{F}$ på $\mathbb{F}$-vektorrum V og bemærker, at valget af skævinvolution medfører, at $\langle \boldsymbol{v}, \boldsymbol{v} \rangle$ er et reelt tal for alle $\boldsymbol{v} \in V$. Dermed er den følgende definition meningsfuld.

Definition 6.5.1 Lad $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, og lad $\langle -,- \rangle\colon V \times V \to \mathbb{F}$ være en hermitisk form på et $\mathbb{F}$-vektorrum V.

(1) Et underrum $W \subset V$ er *positiv definit med hensyn til* $\langle -,- \rangle$, hvis $\langle \boldsymbol{w}, \boldsymbol{w} \rangle > 0$ for alle $\boldsymbol{0} \neq \boldsymbol{w} \in W$.

(2) Et underrum $W \subset V$ er *negativ definit med hensyn til* $\langle -,- \rangle$, hvis $\langle \boldsymbol{w}, \boldsymbol{w} \rangle < 0$ for alle $\boldsymbol{0} \neq \boldsymbol{w} \in W$.

Eksempel 6.5.2 (1) En hermitisk form $\langle -,- \rangle \colon V \times V \to \mathbb{F}$ er et indre produkt, hvis og kun hvis ethvert underrum $W \subset V$ positiv definit med hensyn til $\langle -,- \rangle$, hvis og kun hvis V er positiv definit med hensyn til $\langle -,- \rangle$.
(2) I tilfældet $V = \mathbb{R}^n$ med den Minkowski hermitiske form

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = -x_1 y_1 + x_2 y_2 + \dots + x_n y_n$$

siger vi synonymt, at et underrum $W \subset V$, der er positiv definit med hensyn til $\langle -,- \rangle$, er "spacelike", og vi siger ligeledes, at et underrum $W \subset V$, der er negativ definit med hensyn til $\langle -,- \rangle$, er "timelike".

Lad n være et naturligt tal, og lad p, q og r være naturlige tal, sådan at $p+q+r=n$. Vi definerer da $D(p,q,r) \in M_n(\mathbb{F})$ til at være diagonalmatricen

$$D(p,q,r) = \operatorname{diag}(-1,\dots,-1,0,\dots,0,1,\dots 1),$$

hvor de første p diagonalindgange er lig med -1; de næste q diagonalindgange er lig med 0; og de sidste r diagonalindgange er lig med 1. Vi skal nu vise følgende sætning, der kaldes Sylvester's inerti sætning.

Sætning 6.5.3 (Sylvester) *Lad enten $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$, og lad $\langle -,- \rangle \colon V \times V \to \mathbb{F}$ være en hermitisk form på et n-dimensionalt $\mathbb{F}$-vektorrum.*

(1) *Der findes en basis $(\boldsymbol{v}_1,\dots,\boldsymbol{v}_n)$ for V og naturlige tal p, q og r, sådan at der for alle $\boldsymbol{v},\boldsymbol{w} \in V$ med koordinater $\boldsymbol{x},\boldsymbol{y} \in \mathbb{F}^n$ med hensyn til basen $(\boldsymbol{v}_1,\dots,\boldsymbol{v}_n)$ gælder, at*

$$\langle \boldsymbol{v}, \boldsymbol{w} \rangle = \boldsymbol{x}^* D(p,q,r) \boldsymbol{y}.$$

(2) *De naturlige tal p, q og r afhænger ikke af valget af basis og er karakteriseret som følger: p er den maksimale dimension af et underrum $W \subset V$, der er negativ definit med hensyn til $\langle -,- \rangle$; r er den maksimale dimension af et underrum $W \subset V$, der er positiv definit med hensyn til $\langle -,- \rangle$; og $q = n-p-r$.*

Bevis Vi viser først (1) for $\mathbb{F} = \mathbb{R}$, og vælger hertil en vilkårlig basis $(\boldsymbol{w}_1,\dots,\boldsymbol{w}_n)$ for V. Beviset for sætning 2.6.7 viser da, at matricen

$$A = (\langle \boldsymbol{w}_i, \boldsymbol{w}_j \rangle) \in M_n(\mathbb{R})$$

er symmetrisk, og at $\langle \boldsymbol{v}, \boldsymbol{w} \rangle = \boldsymbol{x}^* A \boldsymbol{y}$ for alle $\boldsymbol{v},\boldsymbol{w} \in V$ med tilhørende koordinater $\boldsymbol{x},\boldsymbol{y} \in \mathbb{R}^n$ med hensyn til den valgte basis. Ifølge spektralsætningen (sætning 6.4.2) findes der en ortogonal matrix $Q \in M_n(\mathbb{R})$, sådan at

$$Q^* A Q = \operatorname{diag}(\lambda_1,\dots,\lambda_n) \in M_n(\mathbb{R})$$

er en diagonalmatrix. Vi påstår, at vi kan vælge $Q \in M_n(\mathbb{R})$, sådan at

$$\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n.$$

For ellers kan vi vælge en permutation af n bogstaver $\sigma\colon \{1,\dots,n\} \to \{1,\dots,n\}$, sådan at der gælder $\lambda_{\sigma(1)} \leq \lambda_{\sigma(2)} \leq \cdots \leq \lambda_{\sigma(n)}$, og permutationsmatricen $P(\sigma) \in M_n(\mathbb{F})$ er da ortogonal, idet dens søjler er standardenhedsvektorer, og

$$P(\sigma)^* Q^* A Q P(\sigma) = \operatorname{diag}(\lambda_{\sigma(1)},\dots,\lambda_{\sigma(n)}) \in M_n(\mathbb{R}).$$

Derfor kan og vil vi antage, at $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$, idet vi ellers erstatter Q med $QP(\sigma)$.

Vi lader nu p, q og r være de naturlige tal, sådan at $\lambda_i < 0$ for $1 \leq i \leq p$; $\lambda_i = 0$ for $p+1 \leq i \leq p+q$; og $\lambda_i > 0$ for $p+q+1 \leq i \leq n$. Vi lader dernæst $R = \operatorname{diag}(\rho_1,\dots,\rho_n) \in M_n(\mathbb{R})$ være diagonalmatricen med diagonalindgange

$$\rho_i = \begin{cases} |\lambda_i|^{-\frac{1}{2}} & \text{hvis } \lambda_i \neq 0, \\ 1 & \text{hvis } \lambda_i = 0. \end{cases}$$

Matricen R er symmetrisk, og den er valgt, sådan at der gælder

$$P^* A P = R^* Q^* A Q R = D(p,q,r),$$

hvor vi skriver $P = (p_{ij}) = QR \in M_n(\mathbb{R})$. Endelig definerer vi

$$\boldsymbol{v}_j = \boldsymbol{w}_1 p_{1j} + \boldsymbol{w}_2 p_{2j} + \cdots + \boldsymbol{w}_n p_{nj}$$

for alle $1 \leq j \leq n$. Da er $(\boldsymbol{v}_1,\dots,\boldsymbol{v}_n)$ en basis for V, og hvis $\boldsymbol{x},\boldsymbol{y} \in \mathbb{F}^n$ er koordinaterne for $\boldsymbol{v},\boldsymbol{w} \in V$ med hensyn til basen $(\boldsymbol{v}_1,\dots,\boldsymbol{v}_n)$, da er $P\boldsymbol{x}, P\boldsymbol{y} \in \mathbb{F}^n$ koordinaterne for $\boldsymbol{v} \in V$ med hensyn til basen $(\boldsymbol{w}_1,\dots,\boldsymbol{w}_n)$. Derfor er

$$\langle \boldsymbol{v},\boldsymbol{w}\rangle = (P\boldsymbol{x})^* A (P\boldsymbol{y}) = \boldsymbol{x}^* P^* A P \boldsymbol{y} = \boldsymbol{x}^* D(p,q,r)\boldsymbol{y}$$

som ønsket. Dette viser (1) for $\mathbb{F} = \mathbb{R}$; beviset for $\mathbb{F} = \mathbb{C}$ er helt analogt, idet man blot erstatter "symmetrisk" med "hermitisk", "ortogonal" med "unitær" og sætning 6.4.2 med sætning 6.4.3.

Vi mangler at vise (2). Vi betragter igen $\mathbb{F} = \mathbb{R}$; tilfældet $\mathbb{F} = \mathbb{C}$ er helt analogt. Det er ikke klart, at de naturlige tal p, q og r i beviset for (1) er uafhængige af vores valg af den ortogonal matrix $Q \in M_n(\mathbb{R})$. Vi træffer nu et valg af en sådan matrix $Q \in M_n(\mathbb{R})$ og finder derefter en basis $(\boldsymbol{v}_1,\dots,\boldsymbol{v}_n)$ for V som i beviset for (1). Vi lader $V_{<0} \subset V$ være underrummet frembragt af $(\boldsymbol{v}_1,\dots,\boldsymbol{v}_p)$ og lader $V_{\geq 0} \subset V$ være underrummet frembragt af $(\boldsymbol{v}_{p+1},\dots,\boldsymbol{v}_n)$. Underrummet $V_{<0} \subset V$ har dimension p og er negativ definit med hensyn til $\langle -,-\rangle$. Vi ønsker at vise, at hvis $W \subset V$ er et underrum, der er negativ definit med hensyn til $\langle -,-\rangle$, så er dimensionen af W højst p. Men $W \cap V_{\geq 0} = \{\boldsymbol{0}\}$, og derfor er

$$\dim(W) + \dim(V_{\geq 0}) \leq \dim(V),$$

hvilket viser, at $\dim(W) \leq p$ som ønsket. Hermed har vi vist, at p er den maksimale dimension af et underrum $W \subset V$, der er negativ definit med hensyn til $\langle -,- \rangle$. Beviset for, at r er den maksimale dimension af et underrum $W \subset V$, der er positiv definit med hensyn til $\langle -,- \rangle$ er helt tilsvarende. Dette afslutter beviset. $\square$

Bemærkning 6.5.4 Lad $(V, \langle -,- \rangle)$ være som i sætning 6.5.3. Vi understreger, at et maksimalt underrum $W \subset V$, der er enten negativ definit eller positiv definit med hensyn til den hermitiske form $\langle -,- \rangle$, *ikke* er entydigt. Det er kun de respektive dimensioner p og r af sådanne underrum, der er entydigt bestemte.

Indeks