Kapitel 7

Vektorer og matricer

7.1 Vektorer

Som i det forrige kapitel vil vi betegne et tallegeme ved $\mathbb{F}$. Det, vi vil arbejde med nu, virker over ethvert legeme, men læseren kan blot tænke på $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$. Når vi beskriver løsninger til systemer af lineære ligninger, arbejder vi allerede med $\mathbb{F}^n$, mængden af alle n-tupler med værdier i $\mathbb{F}$. Vi har også allerede forklaret, at et sådant n-tupel for nemheds skyld ofte betragtes som en $n \times 1$-matrix. Det betyder blot, at:

$$(v_1, \ldots, v_n) \text{ også kan skrives som } \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}.$$

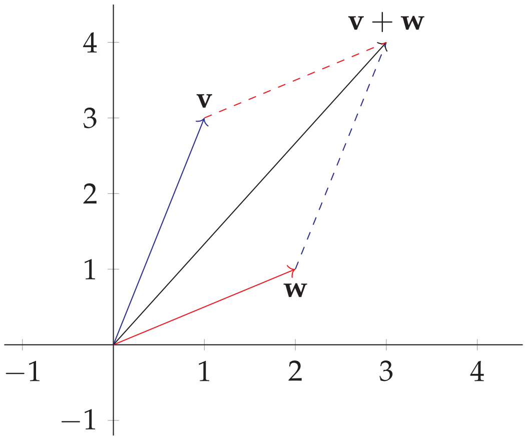
Når et n-tupel skrives som en $n \times 1$-matrix, siger vi, at n-tuplet er skrevet på *vektor*-form. Elementer i $\mathbb{F}^n$ kaldes derfor *vektorer* med n elementer fra $\mathbb{F}$. Hvis alle elementer i en sådan vektor er nul, kalder vi denne vektor *nulvektoren* i $\mathbb{F}^n$.

Bemærkning 7.1.1
Elementer i $\mathbb{F}^{n\times 1}$ kaldes typisk *søjlevektorer*, mens elementer i $\mathbb{F}^{1\times n}$ kaldes *rækkevektorer*.

Vi har i det forrige kapitel allerede benyttet os af, at der findes en naturlig måde at lægge to vektorer $\mathbf{v}$ og $\mathbf{w}$ fra $\mathbb{F}^n$ sammen på, og at man også kan multiplicere en vektor fra $\mathbb{F}^n$ med et element $c \in \mathbb{F}$, ofte kaldet en *skalar* i denne sammenhæng, da multiplikation af en vektor med en konstant kan betragtes som en skalering af vektoren. Mere præcist er addition af vektorer i $\mathbb{F}^n$ defineret som:

$$\begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} + \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} v_1 + w_1 \\ \vdots \\ v_n + w_n \end{bmatrix}. \tag{7.1}$$

For $\mathbb{F} = \mathbb{R}$ og $n = 2$ er addition af to vektorer illustreret i Figur 7.1.

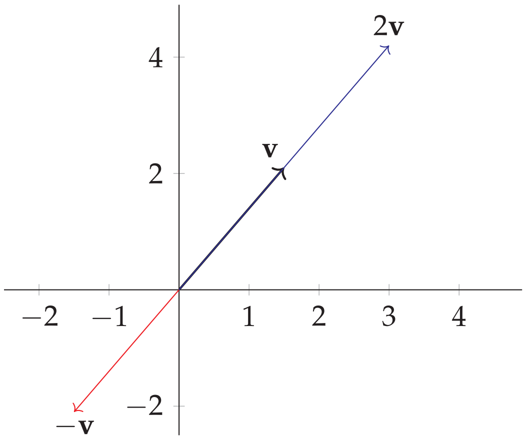


Figur 7.1: Addition af to vektorer $\mathbf{v}$ og $\mathbf{w}$ i $\mathbb{R}^2$.

Produktet af en skalar c fra $\mathbb{F}$ med en vektor i $\mathbb{F}^n$ defineres som:

$$c \cdot \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} c \cdot v_1 \\ \vdots \\ c \cdot v_n \end{bmatrix}. \tag{7.2}$$

I stedet for $(-1) \cdot \mathbf{v}$ kan man også skrive $-\mathbf{v}$. Ligeledes skrives et udtryk som $\mathbf{v} + (-1) \cdot \mathbf{w}$ ofte som $\mathbf{v} - \mathbf{w}$. Man udelader også normalt multiplikationstegnet mellem skalar og vektor. Med andre ord, et udtryk som $c\mathbf{v}$ skal blot læses som $c \cdot \mathbf{v}$. For $\mathbb{F} = \mathbb{R}$ og $n = 2$ er skalering af en vektor illustreret i Figur 7.2.



Figur 7.2: Skalering af en vektor $\mathbf{v} \in \mathbb{R}^2$.

Som med matricer vil vi ofte benytte fed skrift til vektorer og typisk bogstaver som $\mathbf{u}, \mathbf{v}, \mathbf{w}$. Lad os for fremtidig reference opstille følgende samling af egenskaber for vektoraddition og skalarmultiplikation:

Sætning 7.1.1

Lad $\mathbb{F}$ være et legeme, og lad $c, d \in \mathbb{F}$ samt $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{F}^n$. Da gælder:

(i) $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$

(ii) $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$

(iii) $c \cdot (d \cdot \mathbf{u}) = (c \cdot d) \cdot \mathbf{u}$

(iv) $c \cdot (\mathbf{u} + \mathbf{v}) = c \cdot \mathbf{u} + c \cdot \mathbf{v}$

(v) $(c + d) \cdot \mathbf{u} = c \cdot \mathbf{u} + d \cdot \mathbf{u}$

Vi udelader beviset herfor.

Da vi nu har vektorer til rådighed, kan vi se nærmere på flere af deres egenskaber. Vi starter med et eksempel.

Eksempel 7.1.1

Vi er givet vektorerne

$$\mathbf{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \text{ og } \mathbf{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \in \mathbb{R}^2.$$

(a) Udregn $4 \cdot \mathbf{u} + 3 \cdot \mathbf{v}$.

(b) Find c og d, således at $c \cdot \mathbf{u} + d \cdot \mathbf{v} = \mathbf{0}$, hvor $\mathbf{0}$ betegner nulvektoren i $\mathbb{R}^2$.

Svar:

(a) Ved brug af definitionen af skalarmultiplikation og vektoraddition får vi

$$4 \cdot \mathbf{u} + 3 \cdot \mathbf{v} = 4 \cdot \begin{bmatrix} 1 \\ 2 \end{bmatrix} + 3 \cdot \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 8 \end{bmatrix} + \begin{bmatrix} 6 \\ 3 \end{bmatrix} = \begin{bmatrix} 10 \\ 11 \end{bmatrix}.$$

(b) Vi har

$$c \cdot \mathbf{u} + d \cdot \mathbf{v} = c \cdot \begin{bmatrix} 1 \\ 2 \end{bmatrix} + d \cdot \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} c \\ 2c \end{bmatrix} + \begin{bmatrix} 2d \\ d \end{bmatrix} = \begin{bmatrix} c + 2d \\ 2c + d \end{bmatrix}.$$

Hvis vi ønsker, at resultatet skal være nulvektoren, skal vi løse det homogene system af lineære ligninger:

$$\begin{cases} c + 2d = 0 \\ 2c + d = 0 \end{cases}.$$

Ved at trække den første ligning to gange fra den anden ligning, hvilket svarer til at udføre den elementære rækkeoperation $R_2 \leftarrow R_2 - 2 \cdot R_1$, får vi systemet:

$$\begin{cases} c + 2d = 0 \\ 0c - 3d = 0 \end{cases}.$$

Vi kunne fortsætte og bringe systemet på reduceret trappeform, men det er allerede klart nu, at den eneste løsning er $c = d = 0$.

Et udtryk som $4 \cdot \mathbf{u} + 3 \cdot \mathbf{v}$ kaldes en *linearkombination* af vektorerne $\mathbf{u}$ og $\mathbf{v}$. Mere generelt, givet vektorer $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{F}^m$ og skalarer $c_1, \ldots, c_n \in \mathbb{F}$ er et udtryk på formen

$$c_1 \cdot \mathbf{v}_1 + \cdots + c_n \cdot \mathbf{v}_n$$

en linearkombination af vektorerne $\mathbf{v}_1, \ldots, \mathbf{v}_n$. Anden del af eksemplet viser, at den eneste linearkombination af vektorerne $\mathbf{u}$ og $\mathbf{v}$, der er lig med nulvektoren, er linearkombinationen $0 \cdot \mathbf{u} + 0 \cdot \mathbf{v}$. En sekvens af vektorer kan have denne egenskab generelt, hvilket beskrives i følgende.

Definition 7.1.1

En sekvens af vektorer $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{F}^m$ kaldes *lineært uafhængig*, hvis og kun hvis ligningen $c_1 \cdot \mathbf{v}_1 + \cdots c_n \cdot \mathbf{v}_n = \mathbf{0}$ kun kan være sand for $c_1 = \cdots = c_n = 0$.

Hvis sekvensen af vektorer $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{F}^m$ ikke er lineært uafhængig, siges den at være *lineært afhængig*.

Med andre ord er en sekvens af vektorer $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{F}^m$ lineært uafhængig, hvis og kun hvis den eneste linearkombination af vektorerne, der er lig med nulvektoren, opnås for $c_1 = \cdots = c_n = 0$. Ved hjælp af nogle logiske udtryk kan lineær uafhængighed af en sekvens af vektorer $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{F}^m$ formuleres som følger:

$$\text{for alle } c_1, \ldots, c_n \in \mathbb{F} \text{ gælder: } c_1 \cdot \mathbf{v}_1 + \cdots + c_n \cdot \mathbf{v}_n = \mathbf{0} \Rightarrow c_1 = \cdots = c_n = 0. \tag{7.3}$$

På samme måde kan lineær afhængighed af sekvensen af vektorer $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{F}^m$ formuleres på følgende måde:

$$\text{der eksisterer } c_1, \ldots, c_n \in \mathbb{F}, \text{ således at: } c_1 \cdot \mathbf{v}_1 + \cdots c_n \cdot \mathbf{v}_n = \mathbf{0} \ \wedge \ \text{ ikke alle } c_i \text{ er nul.} \tag{7.4}$$

I stedet for at sige, at en sekvens af vektorer $\mathbf{v}_1, \ldots, \mathbf{v}_n$ er lineært (u)afhængig, er det også ganske almindeligt blot at sige, at vektorerne $\mathbf{v}_1, \ldots, \mathbf{v}_n$ er lineært (u)afhængige. Vi vil ofte bruge denne måde at formulere tingene på.

Eksempel 7.1.2

Sekvensen af vektorer bestående af

$$\mathbf{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \text{ og } \mathbf{v} = \begin{bmatrix} 2 \\ 4 \end{bmatrix} \in \mathbb{R}^2$$

er lineært afhængig. Da $\mathbf{v} = 2 \cdot \mathbf{u}$, ser vi nemlig, at $(-2) \cdot \mathbf{u} + \mathbf{v} = \mathbf{0}$.

Dette eksempel illustrerer et mere generelt princip: To vektorer $\mathbf{u}$ og $\mathbf{v}$ er lineært afhængige, hvis og kun hvis den ene er en skalering af den anden. Hvis for eksempel $\mathbf{u} = c \cdot \mathbf{v}$, så er $1 \cdot \mathbf{u} + (-c) \cdot \mathbf{v} = \mathbf{0}$, hvilket viser, at vektorerne er lineært afhængige. På samme måde, hvis

$\mathbf{v} = c \cdot \mathbf{u}$, så er $(-c) \cdot \mathbf{u} + 1 \cdot \mathbf{v} = \mathbf{0}$, hvilket igen viser, at vektorerne er lineært afhængige. Omvendt, hvis vektorerne er lineært afhængige, findes der $c, d \in \mathbb{F}$, som ikke begge er nul, således at $c \cdot \mathbf{u} + d \cdot \mathbf{v} = \mathbf{0}$. Hvis $c \neq 0$, så får vi $\mathbf{u} = (-d/c) \cdot \mathbf{v}$, således at $\mathbf{v}$ er en skalering af $\mathbf{u}$. Hvis $d \neq 0$, opnår vi på samme måde, at $\mathbf{v} = (-c/d) \cdot \mathbf{u}$, hvilket viser at $\mathbf{u}$ i dette tilfælde er en skalering af $\mathbf{v}$. Intuitivt kan man derfor sige, at to vektorer $\mathbf{u}$ og $\mathbf{v}$ er lineært afhængige, hvis og kun hvis der findes en linje gennem origo, som indeholder både $\mathbf{u}$ og $\mathbf{v}$.

Eksempel 7.1.3

Sekvensen af vektorer bestående af

$$\mathbf{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \text{ og } \mathbf{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \in \mathbb{R}^2$$

er lineært uafhængig. Vi har nemlig set i Eksempel 7.1.1, at ligningen $c \cdot \mathbf{u} + d \cdot \mathbf{v} = \mathbf{0}$ medfører, at $c = d = 0$.

Dette eksempel illustrerer, at den lineære uafhængighed af en sekvens af vektorer kan undersøges ved brug af teorien om systemer af lineære ligninger. Dette er faktisk tilfældet generelt, og det generelle resultat er følgende:

Lemma 7.1.2

Lad vektorer $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{F}^m$ være givet, og lad $\mathbf{A} \in \mathbb{F}^{m \times n}$ være den $m \times n$-matrix, hvis søjler er $\mathbf{v}_1, \ldots, \mathbf{v}_n$, det vil sige

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

Sekvensen af vektorer $\mathbf{v}_1, \ldots, \mathbf{v}_n$ er lineært uafhængig, hvis og kun hvis det homogene system af lineære ligninger med koefficientmatrix $\mathbf{A}$ kun har nulvektoren $\mathbf{0} \in \mathbb{F}^n$ som løsning.

Bevis. Antag først, at sekvensen af vektorer $\mathbf{v}_1, \ldots, \mathbf{v}_n$ er lineært uafhængig, og lad $(c_1, \ldots, c_n) \in \mathbb{F}^n$ være en løsning til det homogene system af lineære ligninger med koefficientmatrix $\mathbf{A}$. Dette system kan direkte omskrives til ligningen $c_1 \cdot \mathbf{v}_1 + \cdots + c_n \cdot \mathbf{v}_n = \mathbf{0}$. Ved at bruge vores antagelse om, at sekvensen af vektorer $\mathbf{v}_1, \ldots, \mathbf{v}_n$ er lineært uafhængig, ser vi, at $(c_1, \ldots, c_n) = (0, \ldots, 0)$.

Antag nu omvendt, at det homogene system af lineære ligninger med koefficientmatrix $\mathbf{A}$ kun har nulvektoren $\mathbf{0} \in \mathbb{F}^n$ som løsning. Hvis $(c_1, \ldots, c_n) \in \mathbb{F}^n$ opfylder $c_1 \cdot \mathbf{v}_1 + \cdots + c_n \cdot \mathbf{v}_n = \mathbf{0}$, kan vi straks konkludere, at $(c_1, \ldots, c_n)$ også er en løsning til det homogene system af lineære ligninger med koefficientmatrix $\mathbf{A}$. Men så kan vi konkludere, at $(c_1, \ldots, c_n) = (0, \ldots, 0)$. $\square$

Dette lemma fører til en enkel karakterisering af lineær uafhængighed:

Sætning 7.1.3

Lad $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{F}^m$ være givet, og lad $\mathbf{A} \in \mathbb{F}^{m \times n}$ være matricen med søjlerne $\mathbf{v}_1, \ldots, \mathbf{v}_n$. Sekvensen af vektorer $\mathbf{v}_1, \ldots, \mathbf{v}_n$ er lineært uafhængig, hvis og kun hvis matricen $\mathbf{A}$ har rang n.

Bevis. Dette følger af Korollar 6.4.5 og Lemma 7.1.2. □

Eksempel 7.1.4

Betragt de følgende tre vektorer i $\mathbb{C}^3$:

$$\mathbf{u} = \begin{bmatrix} 1 \\ 0 \\ 1+i \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 0 \\ 1+i \\ 0 \end{bmatrix} \text{ og } \mathbf{w} = \begin{bmatrix} 1+i \\ -1+5i \\ 2i \end{bmatrix}.$$

(a) Er vektorerne $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ lineært uafhængige?

(b) Er vektorerne $\mathbf{u}$, $\mathbf{v}$ lineært uafhængige?

(c) Er vektoren $\mathbf{u}$ lineært uafhængig?

Svar: Den generelle strategi for denne type spørgsmål er at bruge Sætning 7.1.3. Husk, at for at beregne rangen af en matrix er det ifølge Definition 6.3.2, definitionen på rangen af en matrix, tilstrækkeligt at bestemme dens reducerede trappeform. Lad os nu besvare de tre spørgsmål ét ad gangen.

(a) Sætning 7.1.3 medfører, at vi for at finde svaret skal bestemme rangen af matricen

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 1+i \\ 0 & 1+i & -1+5i \\ 1+i & 0 & 2i \end{bmatrix}.$$

Vi har

$$\begin{bmatrix} 1 & 0 & 1+i \\ 0 & 1+i & -1+5i \\ 1+i & 0 & 2i \end{bmatrix} \underset{R_3 \leftarrow R_3 - (1+i) \cdot R_1}{\longrightarrow} \begin{bmatrix} 1 & 0 & 1+i \\ 0 & 1+i & -1+5i \\ 0 & 0 & 0 \end{bmatrix}$$
$$\underset{R_2 \leftarrow (1+i)^{-1} \cdot R_2}{\longrightarrow} \begin{bmatrix} 1 & 0 & 1+i \\ 0 & 1 & 2+3i \\ 0 & 0 & 0 \end{bmatrix}.$$

Vi kan konkludere, at $\rho(\mathbf{A}) = 2$, hvilket er mindre end tre, altså antallet af vektorer vi arbejder med. Derfor er vektorerne $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ lineært afhængige.

(b) I dette tilfælde skal vi beregne rangen af matricen

$$\mathbf{B} = \begin{bmatrix} 1 & 0 \\ 0 & 1+i \\ 1+i & 0 \end{bmatrix}.$$

Anvendes de samme elementære rækkeoperationer som ved løsning af første spørgsmål, får vi, at den reducerede trappeform af **B** er matricen

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

Vi ser, at $\rho(\mathbf{B}) = 2$, hvilket er lig med antallet af vektorer, vi arbejder med. Derfor er vektorerne $\mathbf{u}$, $\mathbf{v}$ lineært uafhængige.

(c) Hvis vi kun overvejer vektoren $\mathbf{u}$, skal vi bestemme rangen af matricen

$$\mathbf{C} = \begin{bmatrix} 1 \\ 0 \\ 1+i \end{bmatrix}.$$

Denne matrix har rang ét, da den ene søjle, som denne matrix består af, ikke er en nulsøjle. Vi kan konkludere, at sekvensen bestående af vektoren $\mathbf{u}$ er lineært uafhængig. Generelt er en sekvens bestående af kun én vektor $\mathbf{u} \in \mathbb{F}^m$ lineært uafhængig, hvis og kun hvis $\mathbf{u} \neq \mathbf{0}$.

7.2 Matricer og vektorer

Da vi studerede systemer af lineære ligninger, introducerede vi begrebet matrix. En matrix $\mathbf{A} \in \mathbb{F}^{m \times n}$ blev introduceret som et rektangulært skema indeholdende $m \times n$ elementer fra et givet legeme $\mathbb{F}$:

$$\mathbf{A} = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}.$$

Nogle gange skriver man blot $\mathbf{A} = [a_{ij}]_{1 \le i \le m, 1 \le j \le m}$ for at holde det kort. Når en matrix gives på denne form, er elementet a_{ij}, som nogle gange også skrives som $a_{i,j}$, det element, der befinder sig i række i og søjle j i matricen $\mathbf{A}$. Det er også normal notation at betegne dette element ved $\mathbf{A}_{ij}$ eller $\mathbf{A}_{i,j}$. Matricen $\mathbf{A}$, der er givet ovenfor, har m rækker: $[a_{i1} \ \dots \ a_{in}]$ for

$i = 1, \dots, m$ og n søjler:

$$\begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix} \text{ for } j = 1, \dots, n.$$

Vi vil kalde rækker i en matrix for *rækkevektorer* og på tilsvarende vis søjler i en matrix for *søjlevektorer*.

Det viser sig at være ganske nyttigt at kunne multiplicere en matrix med en vektor. Vi definerer følgende:

Definition 7.2.1

Lad $\mathbf{A} = (a_{ij})_{1 \le i \le m, 1 \le j \le m} \in \mathbb{F}^{m \times n}$ være en matrix og $\mathbf{v} = (v_1, \dots, v_n) \in \mathbb{F}^n$ en vektor. Da definerer vi $\mathbf{A} \cdot \mathbf{v} \in \mathbb{F}^m$ som følger:

$$\begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \cdot \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} a_{11} \cdot v_1 + \cdots a_{1n} \cdot v_n \\ \vdots \\ a_{m1} \cdot v_1 + \cdots a_{mn} \cdot v_n \end{bmatrix}$$

Bemærk, at vi ikke kan multiplicere en hvilken som helst matrix med en hvilken som helst vektor. Deres størrelser skal "passe" sammen: antallet af søjler i matricen skal være lige med antallet af elementer i vektoren. Hvis dette ikke er tilfældet, er den tilsvarende matrix-vektor-multiplikation ikke defineret.

Eksempel 7.2.1

Lad

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \text{ og } \mathbf{v} = \begin{bmatrix} 7 \\ -1 \\ -2 \end{bmatrix}.$$

Udregn $\mathbf{A} \cdot \mathbf{v}$.

Svar: Ved at benytte Definition 7.2.1 får vi, at:

$$\mathbf{A} \cdot \mathbf{v} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \cdot \begin{bmatrix} 7 \\ -1 \\ -2 \end{bmatrix} = \begin{bmatrix} 1 \cdot 7 + 2 \cdot (-1) + 3 \cdot (-2) \\ 4 \cdot 7 + 5 \cdot (-1) + 6 \cdot (-2) \end{bmatrix} = \begin{bmatrix} -1 \\ 11 \end{bmatrix}.$$

Bemærk, at matrix-vektor-produktet opstår meget naturligt, når man betragter et system af lineære ligninger. Et system af lineære ligninger

$$\begin{cases} a_{11} \cdot x_1 & + & \cdots & + & a_{1n} \cdot x_n & = & b_1 \\ & & \vdots & & & & \vdots \\ a_{m1} \cdot x_1 & + & \cdots & + & a_{mn} \cdot x_n & = & b_m \end{cases}$$

kan udtrykkes som

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}. \tag{7.5}$$

Nu hvor vi har defineret et matrix-vektor-produkt, kan man stille spørgsmålet, om man mere generelt også kan multiplicere matricer med hinanden. Svaret viser sig at være ja, forudsat at deres størrelser passer sammen. Mere præcist kan vi gøre følgende:

Definition 7.2.2

Lad $\mathbf{A} \in \mathbb{F}^{m\times n}$ og $\mathbf{B} \in \mathbb{F}^{n\times \ell}$. Antag, at søjlerne i $\mathbf{B}$ er givet ved $\mathbf{b}_1, \ldots, \mathbf{b}_\ell \in \mathbb{F}^n$, altså antag, at

$$\mathbf{B} = \begin{bmatrix} | & & | \\ \mathbf{b}_1 & \ldots & \mathbf{b}_\ell \\ | & & | \end{bmatrix}.$$

Så definerer vi

$$\mathbf{A} \cdot \mathbf{B} = \begin{bmatrix} | & & | \\ \mathbf{A} \cdot \mathbf{b}_1 & \ldots & \mathbf{A} \cdot \mathbf{b}_\ell \\ | & & | \end{bmatrix}.$$

Bemærk, at matrixproduktet $\mathbf{A} \cdot \mathbf{B}$ kun er defineret, hvis antallet af søjler i $\mathbf{A}$ er lig med antallet af rækker i $\mathbf{B}$. Hvis disse tal stemmer overens, er $\mathbf{A} \cdot \mathbf{B}$ en matrix med m rækker og ℓ søjler. Med andre ord, hvis $\mathbf{A} \in \mathbb{F}^{m\times n}$, og $\mathbf{B} \in \mathbb{F}^{n\times \ell}$, så er $\mathbf{A} \cdot \mathbf{B} \in \mathbb{F}^{m\times \ell}$.

En anden måde at tilgå definitionen af matrixproduktet på er at opstille en formel for elementerne i produktet $\mathbf{A} \cdot \mathbf{B}$ én ad gangen. Lad os sige, at vi ønsker at finde en formel for det (i, j)'te element i produktet, $(\mathbf{A} \cdot \mathbf{B})_{i,j}$, altså elementet i række i og søjle j. Dette svarer til at bestemme det i'te element i produktet $\mathbf{A} \cdot \mathbf{b}_j$, hvor $\mathbf{b}_j$ er den j'te søjle i $\mathbf{B}$. Dette er i sig selv præcis det samme som resultatet af at gange den i'te række i matricen $\mathbf{A}$ med den j'te søjle i matricen $\mathbf{B}$. Da den i'te række i $\mathbf{A}$ kan skrives som $[a_{i1} \ \ldots \ a_{in}]$ og den j'te søjle i $\mathbf{B}$ som

$$\mathbf{b}_j = \begin{bmatrix} b_{1j} \\ \vdots \\ b_{mj} \end{bmatrix},$$

ser vi, at

$$(\mathbf{A} \cdot \mathbf{B})_{i,j} = [a_{i1} \ \ldots \ a_{in}] \cdot \begin{bmatrix} b_{1j} \\ \vdots \\ b_{nj} \end{bmatrix} = a_{i1} \cdot b_{1j} + \cdots + a_{in} \cdot b_{nj}.$$

Ved at bruge summationssymbolet fra Afsnit 5.3 kan vi omskrive denne formel som følger:

$$(\mathbf{A} \cdot \mathbf{B})_{i,j} = \sum_{r=1}^{n} a_{ir} \cdot b_{rj}. \tag{7.6}$$

Eksempel 7.2.2

I dette eksempel sætter vi $\mathbb{F} = \mathbb{R}$, og vi har givet

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

Udregn, hvis det er muligt, matrixprodukterne $\mathbf{A} \cdot \mathbf{B}$ og $\mathbf{B} \cdot \mathbf{A}$.

Svar: Først betragtes matrixproduktet $\mathbf{A} \cdot \mathbf{B}$. Da $\mathbf{A} \in \mathbb{R}^{2\times 3}$, og $\mathbf{B} \in \mathbb{R}^{3\times 3}$, er produktet $\mathbf{A} \cdot \mathbf{B}$ defineret. Vi har allerede udregnet produktet af $\mathbf{A}$ med den første søjle i $\mathbf{B}$ i Eksempel 7.2.1, så vi vil ikke gentage de beregninger her. Tages det i betragtning, får vi:

$$\begin{aligned} \mathbf{A} \cdot \mathbf{B} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \cdot \begin{bmatrix} 7 & 0 & 0 \\ -1 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix} &= \begin{bmatrix} -1 & 1\cdot 0 + 2 \cdot 1 + 3 \cdot 0 & 1 \cdot 0 + 2 \cdot 0 + 3 \cdot 1 \\ 11 & 4 \cdot 0 + 5 \cdot 1 + 6 \cdot 0 & 4 \cdot 0 + 5 \cdot 0 + 6 \cdot 1 \end{bmatrix} \\ &= \begin{bmatrix} -1 & 2 & 3 \\ 11 & 5 & 6 \end{bmatrix}. \end{aligned}$$

Nu ser vi på matrixproduktet $\mathbf{B} \cdot \mathbf{A}$. Da $\mathbf{B}$ har tre søjler, og $\mathbf{A}$ har to rækker, er matrixproduktet $\mathbf{B} \cdot \mathbf{A}$ ikke defineret.

Dette eksempel viser, at der gælder generelt, at $\mathbf{A} \cdot \mathbf{B} \neq \mathbf{B} \cdot \mathbf{A}$. Med andre ord er matrixmultiplikation ikke kommutativ. Faktisk kan det endda ske, som vi lige har set, at et af produkterne ikke er defineret. Selv hvis begge produkter er defineret, betyder rækkefølgen af matricerne stadig noget, og $\mathbf{A} \cdot \mathbf{B} \neq \mathbf{B} \cdot \mathbf{A}$ generelt. Tag for eksempel et kig på $\mathbf{A} = [1\,0]$ og $\mathbf{B} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Så

$$\mathbf{A} \cdot \mathbf{B} = [1\,0] \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 1 \cdot 0 + 0 \cdot 1 = 0, \text{ og } \mathbf{B} \cdot \mathbf{A} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \cdot [1\,0] = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

Lad os også definere addition af matricer:

Definition 7.2.3

Lad $\mathbf{A}, \mathbf{A}' \in \mathbb{F}^{m\times n}$ være givet, altså

$$\mathbf{A} = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}, \text{ og } \mathbf{A}' = \begin{bmatrix} a'_{11} & \dots & a'_{1n} \\ \vdots & & \vdots \\ a'_{m1} & \dots & a'_{mn} \end{bmatrix}.$$

Så definerer vi $\mathbf{A} + \mathbf{A}'$ som følger:

$$\begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} + \begin{bmatrix} a'_{11} & \dots & a'_{1n} \\ \vdots & & \vdots \\ a'_{m1} & \dots & a'_{mn} \end{bmatrix} = \begin{bmatrix} a_{11} + a'_{11} & \dots & a_{1n} + a'_{1n} \\ \vdots & & \vdots \\ a_{m1} + a'_{m1} & \dots & a_{mn} + a'_{mn} \end{bmatrix}.$$

Addition af matricer er kun defineret, hvis de har samme størrelse. På elementniveauet ser vi, at $(\mathbf{A} + \mathbf{A}')_{i,j} = a_{ij} + a'_{ij}$. Addition og multiplikation af matricer opfylder mange af de samme regler som addition og multiplikation af reelle eller komplekse tal. Vi samler nogle af disse i følgende sætning. Den primære forskel, som allerede blev nævnt tidligere, er, at matrixmultiplikation generelt ikke er kommutativ.

Sætning 7.2.1

Lad $\mathbb{F}$ være et legeme. Da gælder

(i) $\mathbf{A} + \mathbf{A}' = \mathbf{A}' + \mathbf{A}$ for alle $\mathbf{A}, \mathbf{A}' \in \mathbb{F}^{m \times n}$.

(ii) $(\mathbf{A} + \mathbf{A}') + \mathbf{A}'' = \mathbf{A} + (\mathbf{A}' + \mathbf{A}'')$ for alle $\mathbf{A}, \mathbf{A}', \mathbf{A}'' \in \mathbb{F}^{m \times n}$.

(iii) $\mathbf{A} \cdot (\mathbf{B} \cdot \mathbf{C}) = (\mathbf{A} \cdot \mathbf{B}) \cdot \mathbf{C}$ for alle $\mathbf{A} \in \mathbb{F}^{m \times n}$, $\mathbf{B} \in \mathbb{F}^{n \times \ell}$ og $\mathbf{C} \in \mathbb{F}^{\ell \times k}$.

(iv) $\mathbf{A} \cdot (\mathbf{B} + \mathbf{B}') = \mathbf{A} \cdot \mathbf{B} + \mathbf{A} \cdot \mathbf{B}'$ for alle $\mathbf{A} \in \mathbb{F}^{m \times n}$ og $\mathbf{B}, \mathbf{B}' \in \mathbb{F}^{n \times \ell}$.

(v) $(\mathbf{A} + \mathbf{A}') \cdot \mathbf{B} = \mathbf{A} \cdot \mathbf{B} + \mathbf{A}' \cdot \mathbf{B}$ for alle $\mathbf{A}, \mathbf{A}' \in \mathbb{F}^{m \times n}$ og $\mathbf{B} \in \mathbb{F}^{n \times \ell}$.

Bevis. Vi vil bevise det tredje punkt og overlade de andre dele til læseren. Anvendes Ligning (7.6) på produktet $(\mathbf{B} \cdot \mathbf{C})$ får vi, at $(\mathbf{B} \cdot \mathbf{C})_{s,j} = \sum_{r=1}^{\ell} b_{sr} \cdot c_{rj}$. Ved at benytte dette samt Ligning (7.6) på produktet $\mathbf{A} \cdot (\mathbf{B} \cdot \mathbf{C})$ opnår vi efter en smule omskrivning, at:

$$\begin{aligned}
(\mathbf{A} \cdot (\mathbf{B} \cdot \mathbf{C}))_{i,j} &= \sum_{s=1}^{n} a_{is} \cdot (\mathbf{B} \cdot \mathbf{C})_{s,j} \\
&= \sum_{s=1}^{n} a_{is} \cdot \sum_{r=1}^{\ell} b_{sr} \cdot c_{rj} \\
&= \sum_{s=1}^{n} \sum_{r=1}^{\ell} a_{is} \cdot (b_{sr} \cdot c_{rj}) \\
&= \sum_{s=1}^{n} \sum_{r=1}^{\ell} (a_{is} \cdot b_{sr}) \cdot c_{rj} \\
&= \sum_{r=1}^{\ell} \sum_{s=1}^{n} (a_{is} \cdot b_{sr}) \cdot c_{rj} \\
&= \sum_{r=1}^{\ell} \left(\sum_{s=1}^{n} a_{is} \cdot b_{sr} \right) \cdot c_{rj}
\end{aligned}$$

$$\begin{aligned} &= \sum_{r=1}^{\ell} (\mathbf{B} \cdot \mathbf{C})_{i,r} \cdot c_{rj} \\ &= ((\mathbf{A} \cdot \mathbf{B}) \cdot \mathbf{C})_{i,j}. \end{aligned}$$

□

Vi afslutter dette afsnit med at forklare to yderligere matrixoperationer. Vi har allerede set, at vektorer kan multipliceres med en skalar. Generaliseringen til matricer er umiddelbar: for $c \in \mathbb{F}$ og

$$\mathbf{A} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \in \mathbb{F}^{m \times n} \text{ defineres } c \cdot \mathbf{A} = \begin{bmatrix} c \cdot a_{11} & \cdots & c \cdot a_{1n} \\ \vdots & & \vdots \\ c \cdot a_{m1} & \cdots & c \cdot a_{mn} \end{bmatrix}. \tag{7.7}$$

Afslutningsvis findes der en måde at ombytte rollerne af rækker og søjler i en matrix $\mathbf{A}$ på. Dette kaldes *transponering* af matricen og kan betegnes $\mathbf{A}^T$. Mere præcist, hvis vi har givet

$$\mathbf{A} = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \in \mathbb{F}^{m \times n}, \text{ så defineres } \mathbf{A}^T = \begin{bmatrix} a_{11} & \dots & a_{m1} \\ \vdots & & \vdots \\ a_{1n} & \dots & a_{mn} \end{bmatrix}. \tag{7.8}$$

Eksempel 7.2.3

Lad matricen

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \in \mathbb{R}^{2 \times 3}$$

være givet. Bestem $\mathbf{A}^T$.

Svar:

Vi har

$$\mathbf{A}^T = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}.$$

Bemærk, at hvis $\mathbf{A} \in \mathbb{F}^{m \times n}$, så er $\mathbf{A}^T \in \mathbb{F}^{n \times m}$. På elementniveauet har vi ganske enkelt, at det (i, j)'te element i $\mathbf{A}^T$ er lig med det (j, i)'te element i $\mathbf{A}$.

Transponeringen opfører sig fint i forbindelse med matrixaddition og matrixprodukter. Dette præciseres i følgende sætning.

Sætning 7.2.2

Lad $\mathbb{F}$ være et legeme. Da gælder:

(i) $(\mathbf{A}^T)^T = \mathbf{A}$ for alle $\mathbf{A} \in \mathbb{F}^{m \times n}$.

(ii) $(\mathbf{A} + \mathbf{A}')^T = \mathbf{A}^T + (\mathbf{A}')^T$ for alle $\mathbf{A}, \mathbf{A}' \in \mathbb{F}^{m \times n}$.

(iii) $(\mathbf{A} \cdot \mathbf{B})^T = \mathbf{B}^T \cdot \mathbf{A}^T$ for alle $\mathbf{A} \in \mathbb{F}^{m \times n}$ og $\mathbf{B} \in \mathbb{F}^{n \times \ell}$.

Bevis. Vi beviser kun den første påstand. Generelt er det (i,j)'te element i $\mathbf{B}^T$ lig med det (j,i)'te element i $\mathbf{B}$ for enhver matrix $\mathbf{B}$. Anvendes dette først på matricen $\mathbf{A}^T$ og derefter på matricen $\mathbf{A}$, får vi, at $((\mathbf{A}^T)^T)_{i,j} = (\mathbf{A}^T)_{j,i} = (\mathbf{A})_{i,j}$. Dette viser, at matricerne $(\mathbf{A}^T)^T$ og $\mathbf{A}$ har præcis de samme elementer, og dermed at de er ens. □

Det er vigtigt at bemærke sig rækkefølgen af multiplikationen i punkt 3 før og efter transponering. Transponering så at sige vender rækkefølgen af elementer i et produkt om. Der er en god grund til dette. Givet matricer $\mathbf{A} \in \mathbb{F}^{m \times n}$ og $\mathbf{B} \in \mathbb{F}^{n \times \ell}$, så er produktet $\mathbf{A}^T \cdot \mathbf{B}^T$ generelt set ikke engang defineret! Antallet af søjler i $\mathbf{A}^T$ er m, mens antallet af rækker i $\mathbf{B}^T$ er ℓ. Dog er der ingen problemer med produktet $\mathbf{B}^T \cdot \mathbf{A}^T$, da antallet af søjler i $\mathbf{B}^T$ er n, hvilket er det samme som antallet af rækker i $\mathbf{A}^T$. Selvom disse observationer ikke direkte beviser punkt tre fra Sætning 7.2.2, så forklarer de, hvorfor det er ganske naturligt, at multiplikationsrækkefølgen er som angivet.

7.3 Kvadratiske matricer

Hvis antallet af rækker og af søjler i en matrix er ens, kaldes den en *kvadratisk* matrix. Med andre ord, en matrix $\mathbf{A}$ er en kvadratisk matrix, hvis $\mathbf{A} \in \mathbb{F}^{n \times n}$ for et positivt heltal n. De elementer, der befinder sig på de (i,i)'te positioner i en kvadratisk matrix, kaldes matricens *diagonalelementer*. For eksempel er diagonalelementerne i matricen

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

tallene 1, 5 og 9. Samlet udgør diagonalelementerne, hvad der kaldes *diagonalen* i en kvadratisk matrix.

For et givet n kaldes den $n \times n$-matrix, der har 1'ere i sin diagonal og 0'er overalt derudover, for *identitetsmatricen*, og den betegnes $\mathbf{I}_n$. For $n = 4$ som et eksempel har vi dermed

$$\mathbf{I}_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Denne matrix kaldes identitetsmatricen fordi den ikke har nogen effekt på en vektor, når den multipliceres med vektoren. En direkte udregning viser, at $\mathbf{I}_n \cdot \mathbf{v} = \mathbf{v}$ for alle $\mathbf{v} \in \mathbb{F}^n$. Derfor er funktionen $L : \mathbb{F}^n \to \mathbb{F}^n$ defineret ved $L(\mathbf{v}) = \mathbf{I}_n \cdot \mathbf{v}$ blot identitetsfunktionen. Med denne matrix på plads er følgende definition mulig:

Definition 7.3.1

En kvadratisk matrix $\mathbf{A} \in \mathbb{F}^{n \times n}$ kaldes *invertibel*, hvis der findes en matrix $\mathbf{B} \in \mathbb{F}^{n \times n}$, således at

$$\mathbf{A} \cdot \mathbf{B} = \mathbf{B} \cdot \mathbf{A} = \mathbf{I}_n.$$

Matricen $\mathbf{B}$, hvis den eksisterer, kaldes den inverse matrix af $\mathbf{A}$ og betegnes ved $\mathbf{A}^{-1}$.

Invertible matricer vil dukke op i mange situationer senere, men allerede når man løser visse systemer af lineære ligninger, kan de være nyttige. Antag for eksempel, at man ønsker at løse systemet af lineære ligninger $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$, hvor vi har en kvadratisk koefficientmatrix $\mathbf{A}$, en vektor indeholdende de variable $\mathbf{x} = (x_1, \ldots, x_n)$ samt en højreside $\mathbf{b} = (b_1, \ldots, b_n)$. Hvis koefficientmatricen $\mathbf{A}$ har en invers, kan vi multiplicere fra venstre med $\mathbf{A}^{-1}$ og reducere: $\mathbf{A}^{-1} \cdot (\mathbf{A} \cdot \mathbf{x}) = (\mathbf{A}^{-1} \cdot \mathbf{A}) \cdot \mathbf{x} = \mathbf{I}_n \cdot \mathbf{x}$. Dette betyder, at ligningen $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$ medfører, at $\mathbf{x} = \mathbf{A}^{-1} \cdot \mathbf{b}$. Omvendt, hvis $\mathbf{x} = \mathbf{A}^{-1} \cdot \mathbf{b}$, så vil en multiplikation med $\mathbf{A}$ fra venstre give $\mathbf{A} \cdot \mathbf{x} = \mathbf{A} \cdot (\mathbf{A}^{-1} \cdot \mathbf{b}) = (\mathbf{A} \cdot \mathbf{A}^{-1}) \cdot \mathbf{b} = \mathbf{I}_n \cdot \mathbf{b} = \mathbf{b}$. Dermed har vi vist, at:

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b}, \text{ hvis og kun hvis } \mathbf{x} = \mathbf{A}^{-1} \cdot \mathbf{b}, \text{ forudsat at } \mathbf{A}^{-1} \text{ eksisterer.} \tag{7.9}$$

Denne observation leder til en fin konsekvens for rangen af invertible matricer:

Lemma 7.3.1

Lad $\mathbf{A} \in \mathbb{F}^{n \times n}$ være givet, og antag, at dens inverse matrix eksisterer. Da er $\rho(\mathbf{A}) = n$.

Bevis. Ligning (7.9) medfører, at det homogene system af lineære ligninger $\mathbf{A} \cdot \mathbf{x} = \mathbf{0}$ kun har løsningen $\mathbf{x} = \mathbf{A}^{-1} \cdot \mathbf{0} = \mathbf{0}$. Men ifølge Korollar 6.4.5 vil rangen af $\mathbf{A}$ da være lig med n. □

Vi kunne nævne endnu flere sammenhænge, men det vil vi vende tilbage til senere. Spørgsmålet er nu, hvordan man finder ud af, hvornår en matrix har en invers, og hvis den gør, hvordan man bestemmer den. Vi vil først finde et algoritmisk svar og derefter beskrive en teoretisk karakterisering af invertible matricer.

Det første vi vil gøre, er at finde en algoritme, der for en given $n \times n$-matrix $\mathbf{A}$ bestemmer en $n \times n$-matrix $\mathbf{B}$, således at $\mathbf{A} \cdot \mathbf{B} = \mathbf{I}_n$, hvis en sådan eksisterer. Derfor vil outputtet fra algoritmen enten være, at en sådan $\mathbf{B}$ ikke eksisterer, eller den vil returnere en sådan $\mathbf{B}$. Bemærk, at ifølge Definition 7.3.1 skal den inverse af $\mathbf{A}$, her betegnet $\mathbf{B}$, opfylde $\mathbf{A} \cdot \mathbf{B} = \mathbf{I}_n$ *og* $\mathbf{B} \cdot \mathbf{A} = \mathbf{I}_n$. Heldigvis viser det sig, at $\mathbf{A} \cdot \mathbf{B} = \mathbf{I}_n$ *medfører* $\mathbf{B} \cdot \mathbf{A} = \mathbf{I}_n$, således at den algoritme, vi er ved at beskrive, rent faktisk vil bestemme den inverse matrix $\mathbf{B} = \mathbf{A}^{-1}$, forudsat at den eksisterer.

Lad os betegne den i'te søjle i enhedsmatricen $\mathbf{I}_n$ ved $\mathbf{e}_i$ for $i = 1, \ldots, n$. Da har vi for eksempel for $n = 4$, at

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \text{ og } \mathbf{e}_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Idéen med algoritmen, der finder inverse matricer, er følgende: Vi forsøger at finde en matrix $\mathbf{B} \in \mathbb{F}^{n\times n}$, således at $\mathbf{A} \cdot \mathbf{B} = \mathbf{I}_n$ for en given $\mathbf{A} \in \mathbb{F}^{n\times n}$. Lad os nu betegne søjlerne i $\mathbf{B}$ ved $\mathbf{b}_1, \ldots, \mathbf{b}_n$. Den i'te søjle i $\mathbf{A} \cdot \mathbf{B}$ er jævnfør definitionen af matrixproduktet lig med $\mathbf{A} \cdot \mathbf{b}_i$, mens den i'te søjle i $\mathbf{I}_n$ er lig med $\mathbf{e}_i$. Derfor er $\mathbf{A} \cdot \mathbf{b}_i = \mathbf{e}_i$ for alle i mellem 1 og n. Omvendt, hvis $\mathbf{A} \cdot \mathbf{b}_i = \mathbf{e}_i$ for alle i mellem 1 og n, så har matricerne $\mathbf{A} \cdot \mathbf{B}$ og $\mathbf{I}_n$ de samme søjler, og derfor gælder $\mathbf{A} \cdot \mathbf{B} = \mathbf{I}_n$. Dermed ser vi, at

$$\mathbf{A} \cdot \mathbf{B} = \mathbf{I}_n, \text{ hvis og kun hvis } \mathbf{A} \cdot \mathbf{b}_i = \mathbf{e}_i \text{ for alle } i \text{ mellem 1 og } n.$$

Derfor kan vi bestemme $\mathbf{b}_i$ ved at løse det inhomogene system af lineære ligninger $\mathbf{A} \cdot \mathbf{x} = \mathbf{e}_i$.

Fra teorien fra det foregående kapitel ser vi, at for at finde ud af, om ligningssystemet $\mathbf{A} \cdot \mathbf{x} = \mathbf{e}_i$ har en løsning, er det tilstrækkeligt at bestemme den reducerede trappeform af totalmatricen $[\mathbf{A}|\mathbf{e}_i]$. Hvis $\rho(\mathbf{A}) = \rho([\mathbf{A}|\mathbf{e}_i])$, så findes der ifølge Korollar 6.4.3 en løsning, og ellers ikke. Hvis der for alle i mellem 1 og n gælder, at $\rho(\mathbf{A}) = \rho([\mathbf{A}|\mathbf{e}_i])$, vil vi derfor være i stand til at bestemme en matrix $\mathbf{B} \in \mathbb{F}^{n\times n}$, således at $\mathbf{A} \cdot \mathbf{B} = \mathbf{I}_n$.

Man kunne nu vælge at arbejde sig igennem ligningssystemet $\mathbf{A} \cdot \mathbf{x} = \mathbf{e}_i$ ét i ad gangen og dermed beregne én søjle i matricen $\mathbf{B}$ ad gangen, hvis den eksisterer. Men da første del af de tilsvarende totalmatricer altid er ens, nemlig $\mathbf{A}$, vil det være langt hurtigere at håndtere alle n systemer på én gang ved at bestemme den reducerede trappeform af matricen $[\mathbf{A}|\mathbf{e}_1|\mathbf{e}_2|\ldots|\mathbf{e}_n] = [\mathbf{A}|\mathbf{I}_n]$.

Derfor er en algoritme, der kan afgøre, om en kvadratisk matrix $\mathbf{A} \in \mathbb{F}^{n\times n}$ har en invers og i bekræftende fald bestemme den, følgende:

(i) Bestem den reducerede trappeform af $n \times 2n$-matricen $[\mathbf{A}|\mathbf{I}_n]$. Dette kan gøres ved hjælp af elementære rækkeoperationer, ligesom vi gjorde i Afsnit 6.3.

(ii) Hvis den resulterende reducerede trappeform ikke er på formen $[\mathbf{I}_n|\mathbf{B}]$, konkludér da, at $\mathbf{A}$ ikke har en invers.

(iii) Hvis den er på formen $[\mathbf{I}_n|\mathbf{B}]$, konkludér da, at $\mathbf{A}$ har en invers, nemlig $\mathbf{A}^{-1} = \mathbf{B}$.

For at se hvordan dette virker i praksis, vil vi gennemgå to eksempler.

Eksempel 7.3.1

Lad $\mathbb{F} = \mathbb{R}$ og

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}.$$

Afgør, om denne matrix har en invers, og bestem den i så fald.

Svar: Først bestemmer vi den reducerede trappeform af matricen $[\mathbf{A}|\mathbf{I}_2]$. Vi får:

$$\begin{aligned} [\mathbf{A}|\mathbf{I}_2] &= \begin{bmatrix} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \end{bmatrix} \underset{R_2 \leftarrow R_2 - 3\cdot R_1}{\longrightarrow} \begin{bmatrix} 1 & 2 & 1 & 0 \\ 0 & -2 & -3 & 1 \end{bmatrix} \\ &\underset{R_2 \leftarrow (-1/2)\cdot R_2}{\longrightarrow} \begin{bmatrix} 1 & 2 & 1 & 0 \\ 0 & 1 & 3/2 & -1/2 \end{bmatrix} \underset{R_1 \leftarrow R_1 - 2\cdot R_2}{\longrightarrow} \begin{bmatrix} 1 & 0 & -2 & 1 \\ 0 & 1 & 3/2 & -1/2 \end{bmatrix}. \end{aligned}$$

Derfor konkluderer vi, at $\mathbf{A}$ har en invers, nemlig

$$\mathbf{A}^{-1} = \begin{bmatrix} -2 & 1 \\ 3/2 & -1/2 \end{bmatrix}.$$

Eksempel 7.3.2

Lad $\mathbb{F} = \mathbb{R}$ og

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 5 & 7 & 9 \end{bmatrix}.$$

Afgør, om denne matrix har en invers, og bestem den i så fald.

Svar: Vi starter med at bestemme den reducerede trappeform af matricen $[\mathbf{A}|\mathbf{I}_3]$. Vi får:

$$\begin{aligned} [\mathbf{A}|\mathbf{I}_3] &= \begin{bmatrix} 1 & 2 & 3 & 1 & 0 & 0 \\ 4 & 5 & 6 & 0 & 1 & 0 \\ 5 & 7 & 9 & 0 & 0 & 1 \end{bmatrix} \\ &\underset{R_2 \leftarrow R_2 - 4\cdot R_1}{\longrightarrow} \begin{bmatrix} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -3 & -6 & -4 & 1 & 0 \\ 5 & 7 & 9 & 0 & 0 & 1 \end{bmatrix} \\ &\underset{R_3 \leftarrow R_3 - 5\cdot R_1}{\longrightarrow} \begin{bmatrix} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -3 & -6 & -4 & 1 & 0 \\ 0 & -3 & -6 & -5 & 0 & 1 \end{bmatrix} \\ &\underset{R_3 \leftarrow R_3 - R_2}{\longrightarrow} \begin{bmatrix} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -3 & -6 & -4 & 1 & 0 \\ 0 & 0 & 0 & -1 & -1 & 1 \end{bmatrix}. \end{aligned}$$

Selvom vi endnu ikke er færdige med bestemmelsen af den reducerede trappeform af $[\mathbf{A}|\mathbf{I}_3]$, har vi allerede her fundet en trappeform af den. Pivotelementerne kan aflæses på dette stadie, og vi ser, at de er placeret i matricens første, anden og fjerde søjle. Fortsætter vi reduktionen til den reducerede trappeform, vil de første tre elementer i tredje række forblive nul. Læseren opfordres til at udregne den reducerede trappeform for at sikre sig, at dette faktisk er korrekt. Dermed vil den reducerede trappeform af $[\mathbf{A}|\mathbf{I}_3]$ ikke være på formen $[\mathbf{I}_3|\mathbf{B}]$. Vi konkluderer, at matricen $\mathbf{A}$ ikke har en invers.

I princippet har vi nu en algoritme, der kan afgøre, om en kvadratisk matrix har en invers, og som i bekræftende fald kan bestemme den. Vi har dog ikke vist, at algoritmen er korrekt. Med andre ord, hvis vi følger trinnene i algoritmen, vil resultatet så altid være, som det skal være? Først og fremmest skal vi sikre os, at hvis den reducerede trappeform af $n \times 2n$-matricen $[\mathbf{A}|\mathbf{I}_n]$ ikke er på formen $[\mathbf{I}_n|\mathbf{B}]$, så har $\mathbf{A}$ ikke en invers. Og for det andet skal vi sikre os, at hvis en matrix $\mathbf{B} \in \mathbb{F}^{n\times n}$ opfylder $\mathbf{A} \cdot \mathbf{B} = \mathbf{I}_n$, så gælder der også, at $\mathbf{B} \cdot \mathbf{A} = \mathbf{I}_n$, således at vi kan konkludere, at $\mathbf{B}$ er den inverse af $\mathbf{A}$. Vi vil adressere disse spørgsmål i resten af dette afsnit. Det vil vise sig, at alt er, som det skal være, og man kan vise, at:

$$\begin{aligned} \mathbf{A}^{-1} \text{ eksisterer} \quad &\Leftrightarrow \quad \text{den reducerede trappeform af } [\mathbf{A}|\mathbf{I}_n] \text{ er på formen } [\mathbf{I}_n|\mathbf{B}] \\ &\Leftrightarrow \quad \rho(\mathbf{A}) = n \text{ (dvs. rangen af } \mathbf{A} \text{ er } n). \end{aligned} \tag{7.10}$$

En læser, der er villig til at acceptere dette uden bevis, kan springe resten af dette afsnit over, men for andre læsere vil vi give et bevis nedenfor.

Sætning 7.3.2

Lad $\mathbf{A} \in \mathbb{F}^{n\times n}$ være en kvadratisk matrix. Følgende udsagn er logisk ækvivalente:

(i) Den reducerede trappeform af $n \times 2n$-matricen $[\mathbf{A}|\mathbf{I}_n]$ er på formen $[\mathbf{I}_n|\mathbf{B}]$ for en kvadratisk matrix $\mathbf{B} \in \mathbb{F}^{n\times n}$.

(ii) Der eksisterer en kvadratisk matrix $\mathbf{B} \in \mathbb{F}^{n\times n}$, således at $\mathbf{A} \cdot \mathbf{B} = \mathbf{I}_n$.

Bevis. Antag, at den reducerede trappeform af matricen $[\mathbf{A}|\mathbf{I}_n]$ er på formen $[\mathbf{I}_n|\mathbf{B}]$ for en kvadratisk matrix $\mathbf{B} \in \mathbb{F}^{n\times n}$. Lad $\mathbf{b}_i$ være den i'te søjle i matricen $\mathbf{B}$. Så kan de samme elementære rækkeoperationer, der anvendes til at transformere matricen $[\mathbf{A}|\mathbf{I}_n]$ til formen $[\mathbf{I}_n|\mathbf{B}]$, også anvendes til at transformere matricen $[\mathbf{A}|\mathbf{e}_i]$ til $[\mathbf{I}_n|\mathbf{b}_i]$. Da $[\mathbf{I}_n|\mathbf{b}_i]$ er på reduceret trappeform, kan vi konkludere, at den reducerede trappeform af matricen $[\mathbf{A}|\mathbf{e}_i]$ er lig med $[\mathbf{I}_n|\mathbf{b}_i]$. Dette medfører, at $\mathbf{b}_i$ er en løsning til systemet af lineære ligninger $\mathbf{A} \cdot \mathbf{x} = \mathbf{e}_i$. Derfor gælder $\mathbf{A} \cdot \mathbf{B} = \mathbf{I}_n$. Specifikt gælder $\mathbf{A} \cdot \mathbf{B} = \mathbf{I}_n$ for en kvadratisk matrix $\mathbf{B} \in \mathbb{F}^{n\times n}$, nemlig matricen i højre del af den reducerede trappeform af $[\mathbf{A}|\mathbf{I}_n]$.

Antag omvendt, at $\mathbf{A} \cdot \mathbf{B} = \mathbf{I}_n$ for en kvadratisk matrix $\mathbf{B} \in \mathbb{F}^{n\times n}$. Da har systemet af lineære ligninger $\mathbf{A} \cdot \mathbf{x} = \mathbf{e}_i$ for alle i fra 1 til n en løsning, nemlig den i'te søjle i matricen $\mathbf{B}$. Vi påstår, at den reducerede trappeform af $[\mathbf{A}|\mathbf{I}_n]$ kun indeholder pivotelementer i sine første n søjler. Vi vil bevise denne påstand ved et modstridbevis. Antag derfor, at den reducerede trappeform af

$[\mathbf{A}|\mathbf{I}_n]$ har en pivot i en søjle med indeks $n+i$ for et $i > 0$. Da vil den reducerede trappeform af matricen $[\mathbf{A}|\mathbf{e}_i]$ indeholde en pivot i sin $(n+1)$'te søjle. Specifikt vil $\mathbf{A}$ og $[\mathbf{A}|\mathbf{e}_i]$ ikke have samme rang. Men så har systemet $\mathbf{A}\cdot\mathbf{x} = \mathbf{e}_i$ ifølge Korollar 6.4.3 ingen løsning. Da vi allerede har observeret, at det har en løsning, opnår vi her en modstrid. Dette beviser påstanden om, at den reducerede trappeform af $[\mathbf{A}|\mathbf{I}_n]$ kun indeholder pivotelementer i sine første n søjler. Næste påstand er, at rangen af $[\mathbf{A}|\mathbf{I}_n]$ er lig med n. For at opnå en modstrid vil vi antage, at den reducerede trappeform af $[\mathbf{A}|\mathbf{I}_n]$ indeholder en nulrække. Ved at se på anden del af matricen, $\mathbf{I}_n$, kan vi konkludere, at der findes en sekvens af elementære rækkeoperationer, der kan transformere $\mathbf{I}_n$ om til en matrix med en nulrække. Men mens $\mathbf{I}_n$ er en matrix med rang n, kan en $n\times n$-matrix med en nulrække højst have rang $n-1$. Dette beviser den anden påstand. Ved at kombinere de to påstande konkluderer vi, at den reducerede trappeform af $[\mathbf{A}|\mathbf{I}_n]$ indeholder en pivot i hver af sine første n søjler. Så er den på formen $[\mathbf{I}_n|\mathbf{C}]$, hvor højre del er en kvadratisk matrix $\mathbf{C} \in \mathbb{F}^{n\times n}$. □

Korollar 7.3.3

Lad $\mathbf{A} \in \mathbb{F}^{n\times n}$ være givet. Da eksisterer der en matrix $\mathbf{B} \in \mathbb{F}^{n\times n}$, således at $\mathbf{A}\cdot\mathbf{B} = \mathbf{I}_n$, hvis og kun hvis $\rho(\mathbf{A}) = n$.

Bevis. Hvis $\mathbf{A}\cdot\mathbf{B} = \mathbf{I}_n$ for en $\mathbf{B} \in \mathbb{F}^{n\times n}$, så er den reducerede trappeform af $n\times 2n$-matricen $[\mathbf{A}|\mathbf{I}_n]$ ifølge Sætning 7.3.2 på formen $[\mathbf{I}_n|\mathbf{C}]$ for en $\mathbf{C} \in \mathbb{F}^{n\times n}$. Men så er den reducerede trappeform af $\mathbf{A}$ selv $\mathbf{I}_n$, hvilket medfører, at $\rho(\mathbf{A}) = n$.

Omvendt, hvis $\rho(\mathbf{A}) = n$, er den reducerede trappeform af $\mathbf{A}$ lig med $\mathbf{I}_n$. Men så er den reducerede trappeform af $[\mathbf{A}|\mathbf{I}_n]$ på formen $[\mathbf{I}_n|\mathbf{C}]$ for en kvadratisk matrix $\mathbf{C} \in \mathbb{F}^{n\times n}$. Ifølge Sætning 7.3.2 kan vi konkludere, at der eksisterer en $\mathbf{B} \in \mathbb{F}^{n\times n}$, således at $\mathbf{A}\cdot\mathbf{B} = \mathbf{I}_n$. □

Korollar 7.3.4

Lad $\mathbf{A} \in \mathbb{F}^{n\times n}$ være en kvadratisk matrix, og antag, at der eksisterer en kvadratisk matrix $\mathbf{B} \in \mathbb{F}^{n\times n}$, således at $\mathbf{A}\cdot\mathbf{B} = \mathbf{I}_n$. Da gælder $\mathbf{B}\cdot\mathbf{A} = \mathbf{I}_n$, og derfor er $\mathbf{B} = \mathbf{A}^{-1}$ den inverse af $\mathbf{A}$.

Bevis. For at konkludere, at $\mathbf{B}$ er den inverse af $\mathbf{A}$, skal vi vise, at $\mathbf{A}\cdot\mathbf{B} = \mathbf{B}\cdot\mathbf{A} = \mathbf{I}_n$. Da det er givet, at $\mathbf{A}\cdot\mathbf{B} = \mathbf{I}_n$, har vi tilbage at vise, at $\mathbf{B}\cdot\mathbf{A} = \mathbf{I}_n$.

Bemærk nu, at $\mathbf{A}\cdot(\mathbf{B}\cdot\mathbf{A}) = (\mathbf{A}\cdot\mathbf{B})\cdot\mathbf{A} = \mathbf{I}_n\cdot\mathbf{A} = \mathbf{A} = \mathbf{A}\cdot\mathbf{I}_n$. Derfor gælder $\mathbf{A}\cdot(\mathbf{B}\cdot\mathbf{A}-\mathbf{I}_n) = \mathbf{A}\cdot(\mathbf{B}\cdot\mathbf{A}) - \mathbf{A}\cdot\mathbf{I}_n = \mathbf{0}$, hvor $\mathbf{0}$ her betegner $n\times n$-nulmatricen.

Bemærk, at den forrige ligning medfører, at enhver søjle i $\mathbf{B}\cdot\mathbf{A}-\mathbf{I}_n$ er en løsning til det homogene system af ligninger $\mathbf{A}\cdot\mathbf{x} = \mathbf{0}$. Ifølge det forrige korollar har matricen $\mathbf{A}$ dog rang n. Derfor ved vi fra Korollar 6.4.5, at systemet $\mathbf{A}\cdot\mathbf{x} = \mathbf{0}$ kun har løsningen $\mathbf{x} = \mathbf{0}$. Derfor er alle søjler i $\mathbf{B}\cdot\mathbf{A}-\mathbf{I}_n$ nulsøjler, hvilket medfører, at $\mathbf{B}\cdot\mathbf{A} = \mathbf{I}_n$. Dette er netop, hvad vi skulle vise. □

Korollar 7.3.5

Lad $\mathbf{A} \in \mathbb{F}^{n \times n}$ være givet. Da eksisterer den inverse matrix, hvis og kun hvis $\rho(\mathbf{A}) = n$.

Bevis. Dette følger af de to tidligere korollarer. $\square$