Uddrag af

Lineær Algebra

PEH 1999

KAPITEL 1

VEKTORRUM OG MATRICER

0. Indledning.

Der er mange lighedspunkter mellem plan og rumlig vektoralgebra. Ligheden forstærkes, når man går over til at regne med vektorernes koordinatsæt, altså med vektorer i talrummet $\mathbb{R}^2$, henholdsvis $\mathbb{R}^3$. Det leder til spørgsmålet: hvorfor ikke også indføre regning med talsæt, der indeholder 4 tal? Eller med talsæt, der indeholder 5 tal, osv? Da må man ganske vist give afkald på anskuelige modeller som planen og rummet, hvor man kan "se" vektorerne. Men det betyder ikke, at der savnes muligheder for at anvende en sådan flerdimensional vektoralgebra og den matrixalgebra, der udspringer af den. Tværtimod gælder om de fleste ikke-geometriske anvendelser, at en begrænsning til 2 eller 3 dimensioner virker som en kunstig indsnævring. En økonomisk orienteret model som den, der er omtalt i Eksempel B.2 side MO176 – 179 kommer bedre til sin ret og bliver mere realistisk, hvis man tillader vilkårlige antal råvarer og produkter.

I dette kapitel skal vi generalisere $\mathbb{R}^2$- og $\mathbb{R}^3$-vektoralgebraen til $\mathbb{R}^n$, hvor $n \in \mathbb{N}$ er vilkårlig. (Vi vil endda gå et skridt videre og kort omtale generelle *vektorrum*). Derefter indfører *matricer* med vilkårlige rækkeantal m og søjleantal n. Det anbefales læseren at støtte sig til Kapitel B i MaD-noterne. eller en tilsvarende behandling af de simple tilfælde $m = n = 2$ og $m = n = 3$.

Ikke kun om dette kapitel, men om hele bogen gælder, at stoffet jævnligt belyses med eksempler og øvelser, der skal vise teoriens *anvendelser* inden for områder som økonomi, statistik, biologi og teknik.

1. Talrummet $\mathbb{R}^n$.

Ved det n-dimensionale talrum, betegnet $\mathbb{R}^n$, forstås mængden af alle n-talsæt $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)$. Vi kalder $\boldsymbol{x}$ en n-dimensional *vektor*, og vi siger, at den har tallene x_i $(i = 1, \cdots, n)$ som *koordinater*. Vi vil foreløbig undlade at skelne mellem de to skrivemåder for $\boldsymbol{x}$, hvor koordinatsættet enten "ligger ned", sådan som vi lige skrev det, eller "står på højkant", som i den følgende definition. Den første skrivemåde er pladsbesparende og bruges derfor en del i denne bog, den anden gør det lettere at overskue regning med vektorer og anbefales til læserens egen brug.

Vi regner to vektorer for ens, når de har samme koordinatsæt. Videre indfører vi straks et par regningsarter i $\mathbb{R}^n$, inspireret af reglerne for, hvordan man regner med 2- og 3-dimensionale talvektorer.

DEFINITION 1. Lad $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ og $t \in \mathbb{R}$. Ved summen af vektorerne $\boldsymbol{x}$ og $\boldsymbol{y}$ og produktet af tallet t og vektoren $\boldsymbol{x}$ forstås vektorerne henholdsvis

$$\boldsymbol{x} + \boldsymbol{y} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix} \quad \text{og} \quad t\boldsymbol{x} = \begin{pmatrix} tx_1 \\ tx_2 \\ \vdots \\ tx_n \end{pmatrix}.$$

I $\mathbb{R}^n$ betegner man med $\boldsymbol{o}$ den såkaldte *nulvektor*, $\boldsymbol{o} = (0, 0, \cdots, 0)$. En vektor $\boldsymbol{x}$ kaldes *egentlig*, når den er forskellig fra $\boldsymbol{o}$. Ligeledes betegner man med $-\boldsymbol{x}$

den såkaldte *modsatte vektor* af $\boldsymbol{x}$, altså $-\boldsymbol{x} = (-x_1, -x_2, \cdots, -x_n)$. I stedet for $\boldsymbol{x} + (-\boldsymbol{y})$ skrives som regel $\boldsymbol{x} - \boldsymbol{y}$.

Det viser sig, at man kan regne lige så ubekymret med $\mathbb{R}^n$-vektorer som med $\mathbb{R}^2$- og $\mathbb{R}^3$-vektorer. Det er indholdet af den følgende sætning.

SÆTNING 1. For vilkårlige $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathbb{R}^n$, $s, t \in \mathbb{R}$ gælder

$$\boldsymbol{x} + \boldsymbol{y} = \boldsymbol{y} + \boldsymbol{x} \,, \tag{1}$$
$$(\boldsymbol{x} + \boldsymbol{y}) + \boldsymbol{z} = \boldsymbol{x} + (\boldsymbol{y} + \boldsymbol{z}) \,, \tag{2}$$
$$\boldsymbol{x} + \boldsymbol{o} = \boldsymbol{x} \,, \tag{3}$$
$$t\boldsymbol{x} = \boldsymbol{o} \iff t = 0 \text{ eller } \boldsymbol{x} = \boldsymbol{o} \,, \tag{4}$$
$$s(t\boldsymbol{x}) = (st)\boldsymbol{x} \,, \tag{5}$$
$$(s + t)\boldsymbol{x} = s\boldsymbol{x} + t\boldsymbol{x} \,, \tag{6}$$
$$t(\boldsymbol{x} + \boldsymbol{y}) = t\boldsymbol{x} + t\boldsymbol{y} \,. \tag{7}$$

BEVIS: Hver af de syv formler følger ret umiddelbart af Definition 1, når formlens to sider skrives ud i koordinater. F.eks. bliver (1) til

$$(x_1 + y_1, x_2 + y_2, \cdots, x_n + y_n) = (y_1 + x_1, y_2 + x_2, \cdots, y_n + x_n) \,,$$

og at dette er sandt, følger af, at leddenes orden i en sum af to *tal* er ligegyldig, når denne regel så at sige anvendes på hver koordinat. Beviserne for de seks andre formler går på samme måde, og de overlades til læseren. □

Bemærkning. Vektorregnereglerne (1)-(7) er lettere at bevise, når vi kun tænker på vektorerne som talsæt, end når vi (for $n = 2$ eller $n = 3$) opfatter dem som orienterede liniestykker i plan eller rum.

Hvorfor gør vi da ikke det, at vi også i den indledende fremstilling udskifter de geometriske argumenter med beviser som ovenfor, hvor der kun arbejdes med koordinaterne? Svaret er, at de geometriske beviser indgår i hele opbygningen af vektorbegrebet, herunder når vi indfører vektorkoordinater og udleder udtryk for, hvordan regning med de geometrisk indførte vektorer oversættes til koordinater. At kassere det geometriske udgangspunkt ville derfor svare til at tro, at man kan undvære stueetagen i et hus, når man flytter op på første sal.

EKSEMPEL 1. Lad $\mathbb{R}^4$-vektorerne a, b og c være givet ved $a = (2,0,-1,1)$, $b = (1,3,0,2)$, $c = (5,4,-1,3)$. Hvis man skal udregne et udtryk i sådanne opgivne vektorer, kan det betale sig at reducere udtrykket mest muligt, inden man indsætter vektorernes koordinater. Reduktionen foregår under kombineret brug af reglerne i Sætning 1. Eksempel:

$$\begin{aligned}
d &= 3\cdot(5a-2b-2c) + 8\cdot(-2a+b+c) + 2\cdot(a-c) \\
&= 15a-6b-6c-16a+8b+8c+2a-2c \\
&= (15-16+2)\cdot a + (-6+8)\cdot b + (-6+8-2)\cdot c \\
&= a + 2b \\
&= (2,0,-1,1)+2\cdot(1,3,0,2) \\
&= (4,6,-1,5)\,.
\end{aligned}$$

Vi minder om, at koordinatudregningen til slut er nemmere at overskue, hvis man skriver vektorerne "på højkant", sådan at tilsvarende koordinater kommer til at stå ud for hinanden. Til gengæld fylder regningerne noget mere, især når n er stor.

ØVELSE 1. Lad a, b og c være de tre vektorer i Eksempel 1. Vis, at der findes et tal t, således at der gælder

$$13a+17b+tc = (-7,11,-3,17)\,.$$

Antag, at vektoren på ligningens højre side ændres, idet man vælger de fire koordinater på må og få. Tror du, det da normalt vil være muligt at finde et tal t, således at ligningen er opfyldt? Forsøg at begrunde svaret.

EKSEMPEL 2. **Produktionsplan.** En virksomhed fremstiller 2 produkter A og B ud fra 5 råvarer $R_1, R_2, \cdots, R_5$. Behovet for råvarer ved produktionen beskrives ved følgende skema:

enheder af råvare	R_1	R_2	R_3	R_4	R_5
pr. enhed af produkt A	2	9	8	5	3
pr. enhed af produkt B	3	5	4	6	6

Det er klart, hvordan man nu kan tale om "råvare-forbrugsvektorer", henholdsvis $a = (2,9,8,5,3)$ for produktet A og $b = (3,5,4,6,6)$ for produktet B. Det samlede råvareforbrug ved fremstilling af en enhed af hver af de to produkter er da $a+b =$

$(5, 14, 12, 11, 9)$, og forbruget ved fremstilling af f.eks. 20 enheder A er $20a = (40, 180, 160, 100, 60)$.

En *produktionsplan* er en beslutning om, hvor mange enheder af hvert produkt virksomheden vil fremstille i en given periode. Hvis det f.eks. planlægges at fremstille 140 enheder A og 215 enheder B, kan vi udregne råvareforbruget som

$$140a + 215b = (925,\ 2335,\ 1980,\ 1990,\ 1710).$$

Det skal altså forstås på den måde, at der i alt er behov for 925 enheder af R_1, 2335 enheder af R_2 osv.

Det kan ske, at produktionen er pålagt begrænsninger, fordi lagerbeholdningerne af råvarer ikke kan suppleres i en vis periode. Grunden kan også være, at råvarerne kun kan leveres i begrænsede mængder pr. produktionsperiode. – Antag, at virksomhedens lagre er c_1 enheder R_1, c_2 enheder R_2, $\cdots$, c_5 enheder R_5. Vi sætter $c = (c_1, c_2, \cdots, c_n)$ ("lagervektoren"). Da er kravet til en produktionsplan på x enheder A og y enheder B, at der skal gælde

$$xa + yb \leq c, \tag{8}$$

idet et ulighedstegn mellem vektorer opfattes på den måde, at tegnet skal gælde *i hver enkelt koordinat*. Med givne lagre c_i udmøntes (8) således i fem lineære uligheder i x og y. Mængden af tilladte produktionsplaner svarer til mængden af talpar (x, y), som opfylder de fem uligheder. Det vil typisk være punkterne i et *polygon-område* i XY-systemets første kvadrant, idet produktionsplanen også skal tilfredsstille $x \geq 0$, $y \geq 0$ (man kan ikke fremstille et negativt antal enheder). I visse tilfælde kommer hertil yderligere et krav om, at x og y skal være *hele tal*, fordi de angiver antal af udelelige produktionsenheder.

ØVELSE 2. Betragt virksomheden i Eksempel 2. Antag, at de fem råvarer findes i lagerbeholdninger på henholdsvis 2000, 4500, 3400, 3700 og 3300 enheder. Disse lagre kan ikke suppleres.

(1) Opstil ulighederne (8), og skitsér mængden af tilladte produktionsplaner som et område i XY-systemet.

(2) Vis, at der findes en tilladt produktionsplan, som bevirker, at man netop får opbrugt lagrene af R_3, R_4 og R_5. Tror du, at dét normalt, dvs med tilfældigt valgte a, b og c, vil være muligt? Hvis den førnævnte produktionsplan gennemføres, hvor meget vil der da være tilbage af R_1 og R_2 ?

Regninger af den type, der blev illustreret i Eksempel 2 og Øvelse 2, er almindelige inden for faget *operationsøkonomi*. Dog vil det i realistiske situationer ikke dreje sig om f.eks. 2 produkter og 5 råvarer, men om snesevis, ja måske hundreder af produkter og råvarer. Men principperne er de samme.

Vi vil et par gange vende tilbage til produktionsplanlægning som et vigtigt anvendt eksempel på, at vektorer og matricer kan være effektive redskaber til at koncentrere store talmaterialer og gøre dem lettere at overskue.

2 Lineær afhængighed og uafhængighed i $\mathbb{R}^n$.

Af hensyn til det følgende starter vi med at slå et par ting fast om begrebet *vektorsæt*. Et sådant består af et antal vektorer (som regel fra samme talrum $\mathbb{R}^n$) skrevet op i en bestemt rækkefølge. To sæt regnes kun for ens, hvis de indeholder lige mange vektorer samt stemmer overens vektor for vektor. F.eks. er $(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c})$ og $(\boldsymbol{a}, \boldsymbol{c}, \boldsymbol{b})$ to forskellige sæt, selv om de indeholder de samme tre vektorer. Det kræves ikke, at vektorerne i et sæt er forskellige, dvs $(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{b})$ er også et vektorsæt.

Disse bemærkninger viser, at der er forskel mellem en endelig *mængde* af vektorer og et vektorsæt. Mængden $\{\boldsymbol{a}, \boldsymbol{b}\}$ er den samme som mængden $\{\boldsymbol{b}, \boldsymbol{a}\}$ og ligeledes den samme som $\{\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{b}\}$, mens sættene $(\boldsymbol{a}, \boldsymbol{b})$, $(\boldsymbol{b}, \boldsymbol{a})$ og $(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{b})$ alle er forskellige.

Et *delsæt* af et sæt S fremkommer, når man stryger nogle af vektorerne i S. Vektorerne i delsættet optræder således i samme rækkefølge som i det givne sæt. F.eks. er $(\boldsymbol{a}, \boldsymbol{c})$ og $(\boldsymbol{b}, \boldsymbol{c})$ begge delsæt af $(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c})$, mens $(\boldsymbol{b}, \boldsymbol{a})$ ikke er det.

*

To vektorer i planen eller rummet vil normalt ikke ligge på samme linie, og tre vektorer i rummet vil normalt ikke ligge i samme plan. Kan vi opstille lignende, mere generelle udsagn om vektorer i $\mathbb{R}^n$?

At to vektorer $\boldsymbol{a}$ og $\boldsymbol{b}$ ligger på linie, betyder rent regnemæssigt, at de er proportionale, dvs der findes et tal t, så at f.eks. $\boldsymbol{b} = t\boldsymbol{a}$. Vi vedtager at sige, at to vektorer ligger på linie, også når én af dem er nulvektoren, og er f.eks. $b = o$, kan det udtrykkes ved førnævnte ligning med $t = 0$.

At tre rumlige vektorer a, b og c ligger i samme plan, betyder tilsvarende, at en af dem kan udtrykkes lineært ved de to andre, f.eks. $c = t_1 a + t_2 b$. Specielt ligger vektorerne i samme plan, hvis to af dem er proportionale, svarende til f.eks. $t_1 = 0$ eller $t_2 = 0$ i førnævnte ligning. Vi siger også, at vektorerne ligger i samme plan, hvis én af dem er nulvektoren, og det udtrykkes i ligningen for $t_1 = t_2 = 0$. Det normale er dog, at t_1 og t_2 begge er forskellige fra 0, dvs a, b og c går i forskellige retninger, men hver af dem ligger i planen udspændt af de to andre.

Disse overvejelser af det algebraiske indhold i udtrykkene "ligge på linie", "ligge i samme plan" vil nu blive generaliseret til $\mathbb{R}^n$ under navnet *lineær afhængighed* mellem vektorer. Forinden er det praktisk at indføre en præcis glose for det forhold, at en vektor er udtrykt lineært ved andre vektorer.

DEFINITION 2. **Linearkombination.** Lad $S = (a_1, a_2, \cdots, a_m)$ være et sæt af m vektorer i $\mathbb{R}^n$. Ved en *linearkombination* af vektorerne i S forstås en vektor x af formen

$$x = t_1 a_1 + t_2 a_2 + \cdots + t_m a_m , \tag{1}$$

hvor $t_i \in \mathbb{R}$ for $i = 1, 2, \cdots, m$. Linearkombinationen kaldes *egentlig*, når det vides, at mindst én af t_i'erne er forskellig fra 0.

En linearkombination af en enkelt vektor a er altså et eller andet multiplum af den, $x = ta$, en linearkombination af to vektorer a og b er en vektor af formen $x = t_1 a + t_2 b$ osv.

DEFINITION 3. **Lineær afhængighed og uafhængighed.** Lad $S = (a_1, a_2, \cdots, a_m)$ være et sæt af m vektorer i $\mathbb{R}^n$. Hvis der findes en egentlig linearkombination af vektorerne i S, som er lig med nulvektoren, altså: hvis der findes tal $t_1, t_2, \cdots, t_m$, ikke alle lig med 0, så at

$$x = t_1 a_1 + t_2 a_2 + \cdots + t_m a_m = o , \tag{2}$$

siges vektorerne i S at være *lineært afhængige*. I modsat fald, dvs hvis (2) kun gælder, når alle $t_i = 0$, siges vektorerne at være *lineært uafhængige*.

Definition 3 er drejet en smule i forhold til det, man kunne vente ud fra omtalen af tilfældene $n = 2$ og $n = 3$. Den følgende sætning slår fast, at drejningen er uvæsentlig, idet vi lige så godt kunne have defineret lineær afhængighed og uafhængighed ved et krav om, at mindst én, hhv ingen af vektorerne kan udtrykkes lineært ved de øvrige. Når vi foretrækker Definition 3, er det, fordi vektorerne i den indgår på lige fod.

SÆTNING 2. Et sæt af vektorer i $\mathbb{R}^n$ er lineært afhængige, hvis og kun hvis mindst én af dem kan skrives som en linearkombination af de øvrige.

BEVIS: Antag først, at vektorerne $a_1, a_2, \cdots, a_m$ er lineært afhængige. Der findes da tal $t_1, t_2, \cdots, t_m$, som ikke alle er lig med 0, således at (2) gælder. Lad f.eks. $t_m \neq 0$. Vi kan løse ligningen (2) mht a_m:

$$a_m = -\frac{t_1}{t_m}a_1 - \frac{t_2}{t_m}a_2 - \cdots - \frac{t_{m-1}}{t_m}a_{m-1}.$$

Denne ligning udtrykker netop, at en af vektorerne i S, her a_m, er en linearkombination af de øvrige.

Antag dernæst, at en af vektorerne i S, f.eks. a_m, er en linearkombination af de øvrige, dvs der findes tal $s_1, s_2, \cdots, s_{m-1}$, således at

$$a_m = s_1 a_1 + s_2 a_2 + \cdots + s_{m-1} a_{m-1}.$$

Denne ligning kan omskrives til

$$s_1 a_1 + s_2 a_2 + \cdots + s_{m-1} a_{m-1} + (-1) \cdot a_m = o,$$

som udtrykker, at en linearkombination af vektorerne i S er lig med nulvektoren. Og denne linearkombination er egentlig, da i hvert fald koefficienten til a_m er forskellig fra 0. Dermed er Sætning 2 bevist. □

Bemærk: Sætning 2 hævder ikke, at *enhver* vektor i et lineært afhængigt sæt kan skrives som linearkombination af de øvrige, kun at dette er muligt for *mindst én* af vektorerne. Hvis der gælder en ligning af formen (1), hvor $t_i \neq 0$ for alle i, kan ligningen løses mht enhver af vektorerne. Men hvis f.eks. $t_1 = 0$, er det ikke sikkert, at a_1 kan skrives som en linearkombination af $a_2, \cdots, a_m$. Blandt andet dette illustreres i de to næste eksempler.

EKSEMPEL 3. Lad $\mathbb{R}^4$-vektorerne

$$a_1 = (2,1,1,-3), \quad a_2 = (-1,1,-2,0), \quad a_3 = (1,1,0,-2)$$

være givet. Er de lineært afhængige eller lineært uafhængige?

Vi skriver forsøgsvis

$$a_3 = s_1 a_1 + s_2 a_2$$

og vil nu prøve at finde tal s_1 og s_2, som tilfredsstiller denne ligning. Den er ensbetydende med

$$\begin{pmatrix} 1 \\ 1 \\ 0 \\ -2 \end{pmatrix} = s_1 \begin{pmatrix} 2 \\ 1 \\ 1 \\ -3 \end{pmatrix} + s_2 \begin{pmatrix} -1 \\ 1 \\ -2 \\ 0 \end{pmatrix} = \begin{pmatrix} 2s_1 - s_2 \\ s_1 + s_2 \\ s_1 - 2s_2 \\ -3s_1 \end{pmatrix}.$$

Der er altså tale om et system af fire ligninger (en for hver koordinat) med de to ubekendte s_1 og s_2. Et sådant system vil normalt ikke have løsninger, men i dette tilfælde er tallene valgt, så der *er* en løsning. F.eks. finder man af den sidste ligning, at $s_1 = 2/3$, og indsættes det i den anden ligning, fås videre $s_2 = 1 - s_1 = 1 - 2/3 = 1/3$. Dermed er vi naturligvis ikke færdige med at løse ligningssystemet, vi har kun fundet den eneste *mulige* løsning, men mangler at gøre prøve dels i den første ligning, dels i den tredje ligning. Men som før antydet viser det sig, at prøven stemmer i begge tilfælde (prøv selv efter!), dvs der gælder faktisk

$$a_3 = \frac{2}{3} a_1 + \frac{1}{3} a_2 .$$

Dermed har vi vist, at de tre vektorer er lineært afhængige.

Af nogle af bemærkningerne ovenfor fremgår, at hvis vi skriver tre $\mathbb{R}^4$-vektorer op på må og få, må vi regne med, at de er lineært *u*afhængige. Der opstår nemlig et system af fire ligninger med to ubekendte, og det vil normalt ikke have løsninger. Selv i $\mathbb{R}^3$, hvor vi får tre ligninger med to ubekendte, vil systemet i reglen ikke have nogen løsninger, dvs tre $\mathbb{R}^3$-vektorer vil i reglen være lineært uafhængige. Det stemmer med den anskuelige fornemmelse af, at tre tilfældigt valgte vektorer i rummet normalt ikke vil ligge i samme plan. Tilsvarende vil to tilfældigt valgte vektorer i $\mathbb{R}^2$ normalt være lineært uafhængige, dvs ikke ligge på linie (ikke være proportionale).

EKSEMPEL 4. Lad $\mathbb{R}^5$-vektorerne

$$a_1 = (2,0,-6,0,8), \quad a_2 = (-1,0,3,0,-4), \quad a_3 = (0,1,7,0,5)$$

være givet. Er de lineært afhængige eller lineært uafhængige?

Vi forsøger ligesom i Eksempel 3 at finde to tal s_1 og s_2, således at

$$a_3 = s_1 a_1 + s_2 a_2 \, .$$

Vektorligningen udmønter sig denne gang i et system af fem ligninger med to ubekendte. Ved udskrivning af systemet (gennemfør dette!) viser det sig hurtigt, at det ikke har løsninger. Alene ligningen for vektorernes 2. koordinater, som bliver

$$0 \cdot s_1 + 0 \cdot s_2 = 1 \, ,$$

indeholder en modstrid, dvs den kan ikke være tilfredsstillet af noget talpar (s_1, s_2), og det samme gælder da for hele systemet.

Kan man da slutte, at $S = (a_1, a_2, a_3)$ er et sæt af lineært uafhængige vektorer? Nej, vi kan kun slutte, at a_3 ikke kan skrives som linearkombination af de to andre vektorer i sættet S. Kigger man nærmere på vektorerne a_1 og a_2, viser det sig, at de er proportionale, idet vi f.eks. kan skrive $a_1 = (-2) \cdot a_2$ (eller $a_2 = (-0.5) \cdot a_1$). Dermed kan vi også skrive

$$a_1 = (-2) \cdot a_2 + 0 \cdot a_3 \, ,$$

hvorved en af vektorerne i S er skrevet som linearkombination af de øvrige. Altså er vektorerne a_1, a_2 og a_3 alligevel lineært afhængige. Dette kan også udtrykkes som i Definition 3, f.eks. hvis vi omskriver ligningen ovenfor til

$$1 \cdot a_1 + 2 \cdot a_2 + 0 \cdot a_3 = o \, ,$$

hvor det ses, at udtrykket på venstre side er en egentlig linearkombination af vektorerne i S. De er altså lineært afhængige.

ØVELSE 3. Undersøg i hvert af følgende tilfælde, om vektorerne er lineært afhængige eller uafhængige. (Vink: lineær uafhængighed er ensbetydende med, at ligning (2) side 7 kun har løsningen $t_1 = t_2 = \cdots = t_m = 0$).

(1) $(1,3)$, $(2,-5)$,
(2) $(1,3)$, $(2,-5)$, $(-1,6)$,
(3) $(1,3)$, $(-5,-15)$,
(4) $(1,1,0)$, $(2,0,-1)$, $(3,4,5)$,
(5) $(1,1,0)$, $(2,0,-1)$, $(3,5,1)$,
(6) $(1,1,1,1,)$, $(1,1,1,0)$, $(1,1,0,0)$, $(1,0,0,0)$.

Hvad er karakteristisk for det ligningssystem, der fremkommer under løsningen af spørgsmål (6). – Forsøg at generalisere (6). Det kan gøres på flere måder.

ØVELSE 4. Betragt situationen i Eksempel 2, side 4-5. Antag, at der er givet en lagervektor $\boldsymbol{c}$. Vil lineær afhængighed af vektorerne $\boldsymbol{a}, \boldsymbol{b}$ og $\boldsymbol{c}$ være nok til at sikre, at der kan lægges en produktionsplan, således at alle fem lagre lige netop bruges op?

Vi slutter dette afsnit med at vise nogle simple egenskaber ved sæt af lineært afhængige eller lineært uafhængige vektorer.

SÆTNING 3. (1) Vektorerne i et vilkårligt *delsæt* af et sæt lineært uafhængige vektorer er atter lineært uafhængige.

(2) *Omordning* af vektorerne i et sæt ændrer ikke egenskaberne lineær afhængighed og lineær uafhængighed.

(3) Vektorerne i et sæt, der indeholder *nulvektoren*, er lineært afhængige.

(4) Vektorerne i et sæt, der indeholder *to ens* vektorer, er lineært afhængige. Det samme er tilfældet, hvis blot to af vektorerne i sættet er *proportionale*.

BEVIS: (1) Lad $S = (\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_m)$ være et sæt af m lineært uafhængige $\mathbb{R}^n$-vektorer. Betragt et delsæt S_1 på r af vektorerne i S, hvor $r \leq m$, og lad det for nemheds skyld være de første r vektorer; beviset går på samme måde, når de står på andre pladser. Antag nu, at en linearkombination af S_1-vektorerne er lig med nulvektoren, dvs der gælder en ligning af formen

$$t_1\boldsymbol{a}_1 + \cdots + t_r\boldsymbol{a}_r = \boldsymbol{o}.$$

Dermed har vi også en linearkombination af vektorerne i S, som er lig med nulvektoren, nemlig

$$t_1\boldsymbol{a}_1 + \cdots + t_r\boldsymbol{a}_r + 0 \cdot \boldsymbol{a}_{r+1} + \cdots + 0 \cdot \boldsymbol{a}_m = \boldsymbol{o}.$$

Da vektorerne i S er lineært uafhængige, er alle koefficienter i den sidste ligning lig med nul, dvs vi har $t_1 = \cdots = t_r = 0$. Men dermed er det samtidig vist, at hvis en linearkombination af vektorerne i S_1 er lig med nulvektoren, så er alle koefficienter lig med 0. Altså er vektorerne i S_1 lineært uafhængige.

(2) For ikke at gøre beviset for tungt nøjes vi med at formulere det for et sæt på fire vektorer og en bestemt omordning. Andre tilfælde går på lignende måde. – Lad $S = (\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3, \boldsymbol{a}_4)$, og antag f.eks., at S-vektorerne er lineært afhængige, dvs der findes fire tal t_1, t_2, t_3, t_4, ikke alle lig med 0, således at

$$t_1\boldsymbol{a}_1 + t_2\boldsymbol{a}_2 + t_3\boldsymbol{a}_3 + t_4\boldsymbol{a}_4 = \boldsymbol{o}.$$

Betragt nu det omordnede sæt $S_1 = (a_3, a_1, a_4, a_2)$. For at vise, at der er lineær afhængighed også i dette sæt, er det nok at foretage den tilsvarende omordning af leddene i vektorligningen ovenfor, dvs skrive den på formen

$$t_3 a_3 + t_1 a_1 + t_4 a_4 + t_2 a_2 = o .$$

Da har vi en egentlig linearkombination af vektorerne i S_1, som er lig med o, eller med andre ord: også i S_1 er der lineær afhængighed.

At også lineær uafhængighed i S medfører lineær uafhængighed i S', vises således: Antag uafhængighed i S. Den 'modsatte omordning' af den, der bragte S over i S', vil bringe S' over i S. Iflg. første del af beviset vil lineær afhængighed i S' derfor medføre lineær afhængighed i S, i strid med antagelsen. Med andre ord: der er lineær uafhængighed i S'.

(3) Vi kan i kraft af (2) antage, at det er den første vektor i S, der er lig med nulvektoren, dvs $S = (o, a_2, \cdots, a_m)$. Idet

$$1 \cdot o + 0 \cdot a_2 + \cdots + 0 \cdot a_m = o ,$$

har vi straks angivet en egentlig linearkombination af S-vektorerne, som er lig med o, dvs de er lineært afhængige.

(4) Vi kan i kraft af (2) antage, at det er de to første vektorer i S, der er ens, dvs $S = (a, a, a_3, \cdots, a_m)$. Idet

$$1 \cdot a + (-1) \cdot a + 0 \cdot a_3 + \cdots + 0 \cdot a_m = o ,$$

har vi straks angivet en egentlig linearkombination af S-vektorerne, som er lig med o, dvs de er lineært afhængige. – Hvis de første to vektorer i S i stedet er proportionale, henholdsvis a og ka, skal vi blot ændre den første koefficient til k, så vil vi atter have en (egentlig) linearkombination, som er lig med o, dvs vi har bevist, at vektorerne er lineært afhængige. □

ØVELSE 5. Vis ved et modeksempel, f.eks. med tre vektorer i $\mathbb{R}^2$, at udsagnet i Sætning 3, (1) *ikke* gælder, når udtrykket "lineært uafhængige" erstattes af "lineært afhængige".

Vis endvidere følgende pendant til Sætning (3), (1): Vektorerne i et vilkårligt *udvidet sæt* af et sæt lineært afhængige vektorer (dvs det givne sæt er et delsæt af det udvidede) er atter lineært afhængige.

Gælder der her en tilsvarende regel for lineær uafhængighed?

9. Matricer.

DEFINITION 17. **Matrix.** En $m \times n$ *matrix* er et rektangulært talskema af formen

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

Tallet a_{ij}, der står i matricens i'te *række* og j'te *søjle*, kaldes dens ij'te *element*. Man taler også om matricens *rækkevektorer* $(a_{i1}, a_{i2}, \cdots, a_{in})$, hhv *søjlevektorer* $\boldsymbol{a}_j = (a_{1j}, a_{2j}, \cdots, a_{mj})^{\top}$.

DEFINITION 18. **Matrixregning.** Summen, hhv differensen af to $m \times n$ matricer $\boldsymbol{A} = (a_{ij})$ og $\boldsymbol{B} = (b_{ij})$ defineres ved

$$\boldsymbol{A} + \boldsymbol{B} = (a_{ij} + b_{ij}), \tag{2}$$

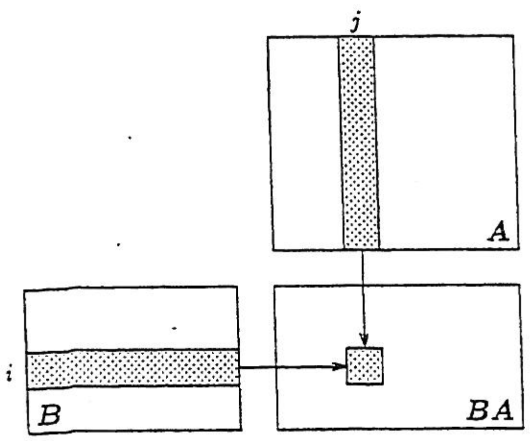
$$\boldsymbol{A} - \boldsymbol{B} = (a_{ij} - b_{ij}). \tag{3}$$

Produktet af et tal $t \in \mathbb{R}$ og en $m \times n$ matrix $\boldsymbol{A} = (a_{ij})$ defineres ved

$$t\boldsymbol{A} = (ta_{ij}). \tag{4}$$

Produktet af en $l \times m$ matrix $\boldsymbol{B} = (b_{ij})$ og en $m \times n$ matrix $\boldsymbol{A} = (a_{ij})$ defineres som $l \times n$ matricen $\boldsymbol{BA} = \boldsymbol{C} = (c_{ij})$, hvor

$$c_{ij} = (\boldsymbol{BA})_{ij} = \sum_{k=1}^{m} b_{ik}a_{kj}. \tag{5}$$



Vi noterer, at $\boldsymbol{A} \pm \boldsymbol{B}$ kun har mening, når $\boldsymbol{A}$ og $\boldsymbol{B}$ har samme antal rækker og samme antal søjler, mens produktet $\boldsymbol{BA}$ kun har mening, når antal søjler i $\boldsymbol{B}$ er lig med antal rækker i $\boldsymbol{A}$. Det ij'te element i $\boldsymbol{BA}$, altså tallet på højre side af (5), er lig med skalarproduktet af den i'te række i $\boldsymbol{B}$ og den j'te søjle i $\boldsymbol{A}$. Det er nyttigt ved udregning "i hånden" af $\boldsymbol{BA}$, jfr. skemaet til venstre.

Regneregler for lineære afbildninger overføres uden videre til matricer:

SÆTNING 20. Mængden af $m \times n$ matricer er et *vektorrum* mht matrix-addition og tal–matrix multiplikation. Specielt optræder *nulmatricen* $\boldsymbol{O}$, der har 0 på alle mn pladser, som "nulelement", mens *den modsatte matrix* til $\boldsymbol{A}$ er $-\boldsymbol{A} = (-a_{ij})$.

Matrixmultiplikation er associativ, den er distributiv (både fra højre og venstre), og en talfaktor i en af faktorerne kan sættes udenfor.

BEVIS: Sætningens første del følger umiddelbart af Sætning 17, side 53, og anden del følger umiddelbart af Sætning 19, side 57. □

Ligesom ved sammensætning af afbildninger er den eneste vigtige regneregel, der *ikke* gælder for matricer, den kommutative lov for matrixmultiplikation. Faktorernes orden i et matrixprodukt er ikke ligegyldig, dvs normalt vil

$$\boldsymbol{AB} \neq \boldsymbol{BA}.$$

Bortset herfra kan man i kraft af Sætning 20 regne ubekymret på matrixudtryk, forudsat at de har mening for de matricer, det drejer sig om.

EKSEMPEL 22. For matricerne $\boldsymbol{A} = \begin{pmatrix} 2 & 0 & 3 \\ -1 & 1 & -1 \end{pmatrix}$, $\boldsymbol{B} = \begin{pmatrix} 1 & 1 & 4 \\ 0 & 3 & -2 \end{pmatrix}$ er

$$\boldsymbol{A}+\boldsymbol{B} = \begin{pmatrix} 2+1 & 0+1 & 3+4 \\ -1+0 & 1+3 & -1-2 \end{pmatrix} = \begin{pmatrix} 3 & 1 & 7 \\ -1 & 4 & -3 \end{pmatrix},$$

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

$$3\boldsymbol{A} = \begin{pmatrix} 6 & 0 & 9 \\ -3 & 3 & -3 \end{pmatrix}.$$

Endvidere finder vi for $\boldsymbol{C} = \begin{pmatrix} 1 & -2 & 3 & 0 \\ 1 & 0 & -1 & 4 \\ 0 & 1 & 2 & 0 \end{pmatrix}$ samt $\boldsymbol{A}$ som ovenfor, at

$$\boldsymbol{AC} = \begin{pmatrix} 2 & -1 & 12 & 0 \\ 0 & 1 & -6 & 4 \end{pmatrix}.$$

Her er f.eks. elementet i 1. række, 3. søjle udregnet som det skalære produkt af 1. række i $\boldsymbol{A}$ og 3. søjle i $\boldsymbol{C}$, dvs som

$$2 \cdot 3 + 0 \cdot (-1) + 3 \cdot 2 = 6 + 0 + 6 = 12.$$

(Det anbefales, at læseren selv checker de øvrige elementer i $\boldsymbol{AC}$).

ØVELSE 27. Udregn med betegnelserne i Eksempel 22

1) $5\boldsymbol{A} - 4\boldsymbol{B}$, 2) $\boldsymbol{BC}$, 3) $(5\boldsymbol{A} - 4\boldsymbol{B})\boldsymbol{C}$.

EKSEMPEL 23. **Nulreglen.** Betegnelsen $\boldsymbol{O}$ for nulmatricen er en smule uklar, idet den forudsætter, at vi arbejder med matricer af en bestemt type $m \times n$. Noget tilsvarende gælder egentlig allerede for nulvektoren $\boldsymbol{o}$. Den simple skrivemåde volder dog som regel ingen problemer, men hvis det i samme regninger er nødvendigt at kunne skelne mellem nulmatricer af forskellig type, kan man f.eks. betegne $m \times n$ nulmatricen som $\boldsymbol{O}_{m\times n}$.

Af definitionen på en nulmatrix er det klart, at der gælder $\boldsymbol{A}\boldsymbol{O}_{m\times n} = \boldsymbol{O}_{l\times n}$ for enhver $l \times m$ matrix $\boldsymbol{A}$, og ligeledes $\boldsymbol{O}_{l\times m}\boldsymbol{B} = \boldsymbol{O}_{l\times n}$ for enhver $m \times n$ matrix $\boldsymbol{B}$. Hvis en af faktorerne i et matrixprodukt er en nulmatrix, er produktet altså ligeledes en nulmatrix. Kan vi også slutte den modsatte vej? Udtrykt på en anden måde, gælder for $l \times n$ matricer $\boldsymbol{A}$ og $m \times n$ matricer $\boldsymbol{B}$, at

$$\boldsymbol{A}\boldsymbol{B} = \boldsymbol{O}_{l\times n} \quad \text{medfører} \quad \boldsymbol{A} = \boldsymbol{O}_{l\times m} \quad \text{eller} \quad \boldsymbol{B} = \boldsymbol{O}_{m\times n}\ ?$$

Svaret er nej, som følgende modeksempel straks viser:

$$\boldsymbol{A} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \boldsymbol{B} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad \boldsymbol{A}\boldsymbol{B} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Der findes også mindre "nul-prægede" modeksempler, f.eks.

$$\boldsymbol{A} = \begin{pmatrix} 1 & -3 \\ 2 & -6 \end{pmatrix}, \quad \boldsymbol{B} = \begin{pmatrix} 12 & -30 \\ 4 & -10 \end{pmatrix}, \quad \boldsymbol{A}\boldsymbol{B} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Nulreglen ville lyde således for matrixmultiplikation: et produkt er en nulmatrix, hvis *og kun hvis* en af faktorerne er en nulmatrix. Men den gælder altså *ikke*. Vi skal dog i Kapitel 2 se, at den "i hovedsagen" gælder, idet faktorerne $\boldsymbol{A}$ og $\boldsymbol{B}$ skal være temmelig specielle, for at deres produkt kan få nuller på alle pladser.

10. Kvadratiske matricer. Transponering.

En matrix med kun én række kaldes en *rækkematrix*, og en matrix med kun én søjle kaldes en *søjlematrix*. I det følgende vil vi ofte opfatte en vektor $\boldsymbol{x} \in \mathbb{R}^n$ som en søjlematrix. Det betyder bl.a., at udtrykket for en lineær afbildning $\boldsymbol{f} : \mathbb{R}^n \to \mathbb{R}^m$ skrevet på den korte og praktiske form

$$\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x}, \tag{1}$$

hvor begge sider er en $m \times 1$ søjlematrix, højresiden fordi den er produkt af en $m \times n$ matrix og en $n \times 1$ matrix. Hvis vi ønsker at opfatte en given vektor $\boldsymbol{x} \in \mathbb{R}^n$ som rækkematrix, skriver vi den $\boldsymbol{x}^\top$.

Ligesom vi ikke behøver at skelne mellem vektoren $\boldsymbol{x}$ og søjlematricen $\boldsymbol{x}$, kan vi stort set sætte lighedstegn mellem en 1×1 matrix (x) og det *tal* x, der er matricens eneste element. Sådan som vi har indført regning med matricer, gælder jo f.eks. $(2)+(5)=(7)$ og $(2)(5)=(10)$, parallelt med $2+5=7$ og $2 \cdot 5 = 10$. Det betyder specielt, at vi kan skrive *det skalære produkt* af to $\mathbb{R}^n$-vektorer $\boldsymbol{x}$ og $\boldsymbol{y}$ som et matrixprodukt, idet der gælder

$$\boldsymbol{x} \cdot \boldsymbol{y} = \boldsymbol{x}^\top \boldsymbol{y} = \boldsymbol{y}^\top \boldsymbol{x}. \tag{2}$$

Når vi nemlig skal udregne f.eks. produktet af $1 \times n$ matrixen $\boldsymbol{x}^\top$ og $n \times 1$ matricen $\boldsymbol{y}$, får vi en 1×1 matrix, hvis eneste element er lig med det skalære produkt af den 1. række i matricen $\boldsymbol{x}^\top$, dvs vektoren $\boldsymbol{x}$, og den 1. søjle i matricen $\boldsymbol{y}$, dvs vektoren $\boldsymbol{y}$.

EKSEMPEL 24. Når $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, kan man danne ikke blot produkterne $\boldsymbol{x}^\top \boldsymbol{y}$ og $\boldsymbol{y}^\top \boldsymbol{x}$, der iflg. (2) begge er lig med tallet $\boldsymbol{x} \cdot \boldsymbol{y}$, men også produkterne $\boldsymbol{x}\boldsymbol{y}^\top$ og $\boldsymbol{y}\boldsymbol{x}^\top$, der til gengæld (for $n > 1$) er mere indviklede størrelser, idet begge er $n \times n$ matricer og i øvrigt normalt forskellige fra hinanden. F.eks. får vi om to $\mathbb{R}^3$-vektorer $\boldsymbol{x} = (x_1, x_2, x_3)^\top$ og $\boldsymbol{y} = (y_1, y_2, y_3)^\top$:

$$\boldsymbol{x}\boldsymbol{y}^\top = \begin{pmatrix} x_1y_1 & x_1y_2 & x_1y_3 \\ x_2y_1 & x_2y_2 & x_2y_3 \\ x_3y_1 & x_3y_2 & x_3y_3 \end{pmatrix}, \quad \boldsymbol{y}\boldsymbol{x}^\top = \begin{pmatrix} y_1x_1 & y_1x_2 & y_1x_3 \\ y_2x_1 & y_2x_2 & y_2x_3 \\ y_3x_1 & y_3x_2 & y_3x_3 \end{pmatrix}.$$

(Det anbefales læseren selv at gå dette efter ved hjælp af skemaet til udregning af et matrixprodukt). Om begge matricer gælder, at de dels har alle tre rækker, dels alle tre søjler proportionale. For andre værdier af n ser regningerne tilsvarende ud. Bemærk i øvrigt, at for $n > 1$ har ingen af matrixprodukterne $\boldsymbol{x}\boldsymbol{y}$, $\boldsymbol{y}\boldsymbol{x}$, $\boldsymbol{x}^\top \boldsymbol{y}^\top$ eller $\boldsymbol{y}^\top \boldsymbol{x}^\top$ mening.

ØVELSE 28. Lad $\boldsymbol{x} = (1,2,3,5)^\top$ og $\boldsymbol{y} = (2,-1,4)^\top$. Opskriv $\boldsymbol{x}^\top$ og $\boldsymbol{y}^\top$. Udregn produkterne $\boldsymbol{x}\boldsymbol{y}^\top$ og $\boldsymbol{y}\boldsymbol{x}^\top$. Var det væsentligt i Eksempel 24, at $\boldsymbol{x}$ og $\boldsymbol{y}$ havde samme dimension? Generalisér resultaterne i eksemplet.

DEFINITION 19. Kvadratisk matrix. En $n \times n$ matrix $\boldsymbol{A}$ siges at være *kvadratisk* af n'te *orden.* Elementerne a_{ii} $(i = 1, \cdots, n)$ udgør den kvadratiske matrix' *diagonal.* Mængden af kvadratiske n'te ordens matricer betegnes $\mathcal{M}_n$.

Glosen "diagonal" om elementerne a_{ii} kan uden problemer bruges også i forbindelse med en ikke-kvadratisk matrix. Hvis $\boldsymbol{A}$ er en $m \times n$ matrix, er længden af diagonalen lig med det mindste af tallene m og n.

DEFINITION 20. Trekantsmatrix og diagonalmatrix. En kvadratisk matrix $\boldsymbol{A}$, hvor alle elementer under diagonalen er lig med 0, dvs $a_{ij} = 0$ for $i > j$, kaldes en *øvre trekantsmatrix.* Tilsvarende defineres *nedre trekantsmatrix.* Hvis alle elementer uden for diagonalen er lig med 0, dvs $a_{ij} = 0$ for $i \neq j$, kaldes $\boldsymbol{A}$ en *diagonalmatrix.* Diagonalmatricen med elementer a_i $(i = 1, \cdots, n)$ i diagonalen betegnes $\mathrm{diag}(a_1, a_2, \cdots, a_n)$.

Eksempler på en øvre trekantsmatrix, hhv en diagonalmatrix (her med $n = 3$):

$$\begin{pmatrix} 2 & 3 & 1 \\ 0 & 5 & 4 \\ 0 & 0 & 7 \end{pmatrix}, \qquad \mathrm{diag}(2,5,7) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 7 \end{pmatrix}.$$

Mængden $\mathcal{M}_n$ er et afsluttet regneområde mht matrixregning. Når man adderer $\mathcal{M}_n$-matricer eller ganger dem med tal, får man atter $\mathcal{M}_n$-matricer, dvs $\mathcal{M}_n$ er et *vektorrum* mht de to regningsarter. Men $\mathcal{M}_n$ er også afsluttet over for matrixmultiplikation, dvs for $\boldsymbol{A}, \boldsymbol{B} \in \mathcal{M}_n$ eksisterer både $\boldsymbol{AB}$ og $\boldsymbol{BA}$, og begge tilhører atter $\mathcal{M}_n$. Stadig vil dog normalt gælde $\boldsymbol{AB} \neq \boldsymbol{BA}$.

Ved multiplikation af $\mathcal{M}_n$-matricer findes der et "neutralelement", dvs en matrix, der opfører sig lidt ligesom tallet 1 ved talregning.

DEFINITION 21. Enhedsmatrix. Den kvadratiske n'te ordens matrix

$$\boldsymbol{E} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}. \tag{3}$$

kaldes *enhedsmatricen* af n'te orden.

SÆTNING 21. Enhedsmatricen $\boldsymbol{E}$ af n'te orden opfylder

$$\boldsymbol{AE} = \boldsymbol{A}, \qquad \boldsymbol{EB} = \boldsymbol{B}, \tag{4}$$

hvor $\boldsymbol{A}$ er en vilkårlig $m \times n$ matrix, og hvor $\boldsymbol{B}$ er en vilkårlig $n \times p$ matrix.

BEVIS: Først bemærkes, at række- og søjleantal på venstre og højre side stemmer i begge formlerne i (4). Dernæst finder vi det ij'te element i $\boldsymbol{AE}$, som er lig med den i'te række i $\boldsymbol{A}$ skalært multipliceret med den j'te søjle i $\boldsymbol{E}$, dvs med e_j. Elementet er altså

$$(a_{i1}, \cdots, a_{ij}, \cdots, a_{in}) \cdot (0, \cdots, 1, \cdots, 0) = 0 + \cdots + a_{ij} + \cdots + 0 = a_{ij}.$$

Men når det gælder på alle pladser, er $\boldsymbol{AE}$ lig med $\boldsymbol{A}$, som hævdet i den første formel. Den anden formel, $\boldsymbol{EB} = \boldsymbol{B}$, vises på samme måde. □

Formel (4) kan ses fra et andet synspunkt: Den *lineære afbildning*, der hører til en kvadratisk matrix af n'te orden, afbilder $\mathbb{R}^n$ ind i sig selv; specielt vil $e(\boldsymbol{x}) = \boldsymbol{Ex} = \boldsymbol{x}$ være den *identiske afbildning* i $\mathbb{R}^n$, og (4) udtrykker, at e er "neutral" ved sammensætning med vilkårlige lineære afbildninger, idet

$$\boldsymbol{f} \circ \mathrm{e} = \boldsymbol{f}, \qquad \mathrm{e} \circ \boldsymbol{g} = \boldsymbol{g},$$

som endda gælder, også når $\boldsymbol{f}$, hhv $\boldsymbol{g}$ er ikke-lineære.

I resten af denne bog har de kvadratiske matricer en fremtrædende rolle. Lad os derfor nævne, at Sætning 21 ofte kun skal bruges i det specialtilfælde, hvor $\boldsymbol{A}$ og $\boldsymbol{B}$ begge tilhører $\mathcal{M}_n$, og hvor der ingen grund er til at skelne mellem dem. Formel (4) antager da skikkelsen

$$\boldsymbol{AE} = \boldsymbol{EA} = \boldsymbol{A}. \tag{5}$$

SÆTNING 22. Mængden af øvre $n \times n$ trekantsmatricer er et afsluttet regneområde inden for $\mathcal{M}_n$. Det samme gælder om mængden af nedre $n \times n$ trekantsmatricer.

BEVIS: Lad $\boldsymbol{A}$ og $\boldsymbol{B}$ være to øvre $n \times n$ trekantsmatricer, og lad $t \in \mathbb{R}$. Det er klart, at både $\boldsymbol{A} + \boldsymbol{B}$ og $t\boldsymbol{A}$ også er øvre $n \times n$ trekantsmatricer. (Mængden af sådanne er altså et *underrum* af $\mathcal{M}_n$.) Vi mangler at vise, at også produktet $\boldsymbol{BA}$ er en øvre $n \times n$ trekantsmatrix.

Betragt et element under diagonalen i $\boldsymbol{BA}$, f.eks. det ij'te, hvor altså $i > j$. Dette element er lig med skalarproduktet af den i'te række i $\boldsymbol{B}$, som har 0 på de første $i-1$ pladser, og den j'te søjle i $\boldsymbol{A}$, som er forskellig fra 0 kun på de første j pladser. Da $i-1 \geq j$, vil hvert led i skalarproduktet indeholde mindst én nulfaktor og derfor være lig med 0. Da det gælder overalt under diagonalen, er $\boldsymbol{BA}$ en øvre trekantsmatrix. - Beviset for nedre trekantsmatricer går på tilsvarende måde. □

SÆTNING 23. Mængden af diagonalmatricer af n'te orden er et afsluttet regneområde inden for $\mathcal{M}_n$, idet der gælder: for $\boldsymbol{A} = \text{diag}(a_1, \cdots, a_n)$, $\boldsymbol{B} = \text{diag}(b_1, \cdots, b_n)$ og $t \in \mathbb{R}$ er

$$\boldsymbol{A} + \boldsymbol{B} = \text{diag}(a_1 + b_1, a_2 + b_2, \cdots, a_n + b_n), \tag{6}$$
$$t\boldsymbol{A} = \text{diag}(ta_1, ta_2, \cdots, ta_n), \tag{7}$$
$$\boldsymbol{AB} = \boldsymbol{BA} = \text{diag}(a_1b_1, a_2b_2, \cdots, a_nb_n). \tag{8}$$

BEVIS: (6) og (7) indses umiddelbart. For at vise (8) bemærker vi, at det ij'te element i $\boldsymbol{AB}$ er lig med skalarproduktet af den i'te række i $\boldsymbol{A}$, som er $(0, 0, \cdots, a_i, \cdots, 0)$ med a_i på plads i, og den j'te søjle i $\boldsymbol{B}$, som er $(0, 0, \cdots, b_j, \cdots, 0)$ med b_j på plads j. Dette skalarprodukt er 0, når $i \neq j$, og a_ib_i, når $i = j$. Følgelig er $\boldsymbol{AB}$ lig med højre side af (8). Da denne ikke kan "mærke", at der byttes om på $\boldsymbol{A}$ og $\boldsymbol{B}$, er også $\boldsymbol{BA}$ lig med $\text{diag}(a_1b_1, a_2b_2, \cdots, a_nb_n)$. Dermed er (8) vist. □

Som et supplement til vektorregningsarterne skal vi endnu indføre følgende nyttige skrivemåde:

DEFINITION 22. **Transponering.** Ved den *transponerede* af $m \times n$ matricen $\boldsymbol{A}$, betegnet $\boldsymbol{A}^\top$, forstås den $n \times m$ matrix, der fås af $\boldsymbol{A}$ ved spejling om diagonalen, idet rækker og søjler i $\boldsymbol{A}$ bytter roller.

Definitionen på transponering kan også udtrykkes i formlen

$$A = (a_{ij}) \qquad \text{medfører} \qquad (A^{\top})_{ij} = a_{ji}. \tag{9}$$

Eksempel på transponering:

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & -1 \\ 0 & -7 & -8 \\ 11 & 2 & -3 \end{pmatrix}, \qquad A^{\top} = \begin{pmatrix} 1 & 4 & 0 & 11 \\ 2 & 5 & -7 & 2 \\ 3 & -1 & -8 & -3 \end{pmatrix}.$$

I forlængelse af Sætning 22 og Sætning 23 bemærkes: Når A er en øvre trekantsmatrix, er $A^{\top}$ en nedre, og omvendt. Når A er en n'te ordens diagonalmatrix, er simpelthen $A^{\top} = A$. Bemærk også, at den brug, vi tidligere har indført af tegnet $^{\top}$, nemlig til at skifte mellem række- og søjleskrivemåden for en $\mathbb{R}^n$-vektor, harmonerer med brugen af tegnet til transponering.

SÆTNING 24. For vilkårlige $m \times n$ matricer A og B samt $t \in \mathbb{R}$ gælder

$$(A + B)^{\top} = A^{\top} + B^{\top}, \tag{10}$$

$$(tA)^{\top} = tA^{\top}, \tag{11}$$

$$(A^{\top})^{\top} = A. \tag{12}$$

For vilkårlige matricer A og B, hhv $m \times n$ og $n \times p$, gælder

$$(AB)^{\top} = B^{\top} A^{\top}. \tag{13}$$

BEVIS: (10), (11) og (12) følger umiddelbart af Definition 22. For at vise (13) bemærker vi først, at række- og søjleantal på formlens to sider stemmer, da venstre side er transponeret af en $m \times p$ matrix, dvs en $p \times m$ matrix, mens højre side er produkt af en $p \times n$ matrix og en $n \times m$ matrix. Endvidere er det ij'te element på venstre side lig med det ji'te element af AB, dvs j'te række i A skalært multipliceret med i'te søjle i B. Samtidig er det ij'te element på højre side af (13) lig med række i i $B^{\top}$, dvs søjle i i B, skalært multipliceret med søjle j i $A^{\top}$, dvs række j i A. Da det vilkårlige element på hver side af (13) således er samme tal, er formlen sand. □

De to første formler i Sætning 24 udtrykker tilsammen, at transponering er en *lineær* proces. Ved kombineret brug af formlerne følger, at også transponering af et mere indviklet udtryk, hvor matricer adderes, subtraheres og ganges med tal, svarer til transponering af hver af de optrædende matricer. Eksempel: $(3\boldsymbol{A} - 5\boldsymbol{B} + 4\boldsymbol{C})^{\top} = 3\boldsymbol{A}^{\top} - 5\boldsymbol{B}^{\top} + 4\boldsymbol{C}^{\top}$.

Formel (12) udtrykker, at transponering er en *involutorisk* proces, dvs "når den udføres to gange, kommer man hjem igen". Dette princip møder man mange forskellige steder i matematikken, f.eks. ved dannelse af modsat tal, reciprok, modsat vektor, komplementærmængde samt ved diverse former for spejling i plan og rum.

Formel (13) er knapt så oplagt som de tre andre. En intuitiv fornemmelse af (13) kan man få ved at tænke sig foretaget en "spejling om diagonalen" af skemaet til udregning af $\boldsymbol{AB}$, se side 61 (hvor dog $\boldsymbol{A}$ og $\boldsymbol{B}$ skal bytte rolle). Derved bliver det til et skema til udregning af $\boldsymbol{B}^{\top}\boldsymbol{A}^{\top}$, som således er lig med det transponerede af resultatet fra før, dvs lig med $(\boldsymbol{AB})^{\top}$.

EKSEMPEL 25. Formel (13) kan udtrykkes i ord således: den transponerede af et produkt af to matricer er lig med produktet af deres transponerede *i modsat rækkefølge*. Reglen kan med uændret formulering generaliseres til et produkt af tre faktorer:

$$(\boldsymbol{ABC})^{\top} = \boldsymbol{C}^{\top}\boldsymbol{B}^{\top}\boldsymbol{A}^{\top},$$

hvor det blot forudsættes, at produktet på venstre side har mening. Formlen vises ved at benytte (13) to gange:

$$\begin{aligned}(\boldsymbol{ABC})^{\top} &= (\boldsymbol{A}(\boldsymbol{BC}))^{\top}\\ &= (\boldsymbol{BC})^{\top}\boldsymbol{A}^{\top}\\ &= \boldsymbol{C}^{\top}\boldsymbol{B}^{\top}\boldsymbol{A}^{\top}.\end{aligned}$$

Med tilsvarende bevis kan formlen generaliseres videre til et produkt af fire faktorer, fem faktorer osv.

ØVELSE 29. Lad $\boldsymbol{A} = \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}$, $\boldsymbol{B} = \begin{pmatrix} -1 & 2 & 1 \\ 0 & 1 & 3 \end{pmatrix}$. Udregn dels $\boldsymbol{AB}$, dels $\boldsymbol{B}^{\top}\boldsymbol{A}^{\top}$. Er (13) opfyldt?

DEFINITION 23. **Symmetri og skævsymmetri.** En kvadratisk matrix $\boldsymbol{A}$ kaldes *symmetrisk*, når $\boldsymbol{A}^{\top} = \boldsymbol{A}$, dvs når $a_{ji} = a_{ij}$ for alle i, j. En kvadratisk matrix $\boldsymbol{B}$ kaldes *skævsymmetrisk* (eller: *antisymmetrisk*), når $\boldsymbol{B}^{\top} = -\boldsymbol{B}$, dvs når $b_{ji} = -b_{ij}$ for alle i, j.

I en skævsymmetrisk matrix er alle diagonalelementer lig med 0. For $i = j$ får vi nemlig $b_{ii} = -b_{ii}$, som ved overflytning og division med 2 giver $b_{ii} = 0$.

Eksempler på hhv en symmetrisk og en skævsymmetrisk matrix:

$$\boldsymbol{A} = \begin{pmatrix} 1 & 5 & -7 \\ 5 & 3 & 2 \\ -7 & 2 & -3 \end{pmatrix}, \qquad \boldsymbol{B} = \begin{pmatrix} 0 & -1 & 4 \\ 1 & 0 & 9 \\ -4 & -9 & 0 \end{pmatrix}.$$

EKSEMPEL 26. **Symmetri og skævsymmetri.** *Nulmatricen er den eneste $n \times n$ matrix, der er både symmetrisk og skævsymmetrisk.* Antag nemlig, at der gælder både $\boldsymbol{A}^{\top} = \boldsymbol{A}$ og $\boldsymbol{A}^{\top} = -\boldsymbol{A}$. Ved at addere disse to formler og derefter dividere med 2 får vi $\boldsymbol{A}^{\top} = \boldsymbol{O}$ og dermed $\boldsymbol{A} = \boldsymbol{O}$.

Symmetri og skævsymmetri er lineære egenskaber i den forstand, at hvis f.eks. $\boldsymbol{A}_1$ og $\boldsymbol{A}_2$ begge er symmetriske, så er dels $\boldsymbol{A}_1 + \boldsymbol{A}_2$, dels $t\boldsymbol{A}_1$ (hvor $t \in \mathbb{R}$) også symmetriske. Det indses således:

$$(\boldsymbol{A}_1 + \boldsymbol{A}_2)^{\top} = \boldsymbol{A}_1^{\top} + \boldsymbol{A}_2^{\top} = \boldsymbol{A}_1 + \boldsymbol{A}_2, \qquad (t\boldsymbol{A}_1)^{\top} = t\boldsymbol{A}_1^{\top} = t\boldsymbol{A}_1.$$

Tilsvarende regler gælder for skævsymmetri, med analogt bevis. Mængden af symmetriske og mængden af skævsymmetriske $n \times n$ matricer er altså begge *underrum* af $\mathcal{M}_n$.

En vilkårlig kvadratisk matrix $\boldsymbol{A}$ kan på netop én måde skrives som sum af en symmetrisk og en skævsymmetrisk matrix. Antag nemlig, at vi har opløsningen

$$\boldsymbol{A} = \boldsymbol{S} + \boldsymbol{T}, \tag{14}$$

hvor $\boldsymbol{S}$ er symmetrisk og $\boldsymbol{T}$ skævsymmetrisk. Ved at transponere begge sider af denne formel får vi under brug af forudsætningerne

$$\boldsymbol{A}^{\top} = \boldsymbol{S} - \boldsymbol{T}.$$

Videre kan vi dels addere disse to formler og dividere med 2, dels trække den anden fra den første og dividere med 2. Derved får vi hhv

$$\boldsymbol{S} = \frac{1}{2}(\boldsymbol{A} + \boldsymbol{A}^{\top}), \qquad \boldsymbol{T} = \frac{1}{2}(\boldsymbol{A} - \boldsymbol{A}^{\top}). \tag{15}$$

Formel (15) angiverer altså de eneste *mulige* $\boldsymbol{S}$ og $\boldsymbol{T}$ i opløsningen (14). Omvendt kan man nemt gå efter, at matricerne $\boldsymbol{S}$ og $\boldsymbol{T}$ i (15) virkelig *er* hhv symmetrisk og skævsymmetrisk, samt at de har summen $\boldsymbol{A}$. Dermed er påstanden vist.

ØVELSE 30. Gennemfør opspaltningen (14) i Eksempel 26 for 3×3 matricen

$$A = \begin{pmatrix} 4 & 8 & -7 \\ -1 & 3 & 5 \\ 2 & 9 & 11 \end{pmatrix},$$

dvs skriv $\boldsymbol{A}$ som sum af en symmetrisk matrix $\boldsymbol{S}$ og en skævsymmetrisk matrix $\boldsymbol{T}$.

11. Affin afbildning.

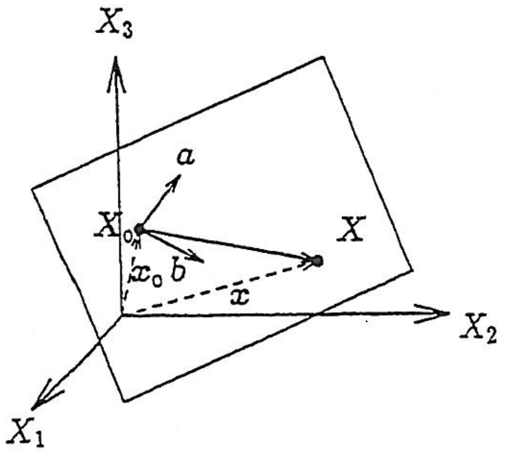
> DEFINITION 24. **Affin afbildning.** En afbildning $f : \mathbb{R}^n \curvearrowright \mathbb{R}^m$ kaldes *affin*, hvis den er af formen
>
> $$f(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}, \tag{1}$$
>
> hvor $\boldsymbol{A}$ er en $m \times n$ matrix og $\boldsymbol{b}$ en $\mathbb{R}^m$-vektor.

Man kan opfatte den affine afbildning (1) som sammensat af den lineære afbildning, der har $\boldsymbol{A}$ som matrix, og den "parallelforskydning" i $\mathbb{R}^m$, der bestemmes ved vektoren $\boldsymbol{b}$. Når specielt $\boldsymbol{b} = \boldsymbol{o}$, er f lineær.

Mens man ved en lineær afbildning $f(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x}$ i reglen tænker på $\boldsymbol{x}$ og $\boldsymbol{A}\boldsymbol{x}$ som *vektorer* i hhv $\mathbb{R}^n$ og $\mathbb{R}^m$, så er det ved en affin afbildning ofte mere naturligt at opfatte dem som *punkter* (evt. som stedvektorer til punkter), smlgn. side 27-28 og side 41-45. Derved kan man f.eks. tale om, at en parameterfremstilling for en linie i $\mathbb{R}^n$, se side 41-42, er det samme som et regneudtryk for en affin afbildning fra $\mathbb{R}$ til $\mathbb{R}^n$, og billedmængden er selve linien. En beslægtet situation møder vi i det følgende eksempel.

EKSEMPEL 27. **Parameterfremstilling for plan.** En plan i $\mathbb{R}^3$ kan være givet ved, at den indeholder punktet X_0 og er parallel med hver af vektorerne $\boldsymbol{a}$ og $\boldsymbol{b}$ (som antages lineært uafhængige, dvs ikke-parallelle).

Et vilkårligt punkt $X \in \mathbb{R}^3$ ligger altså i planen, hvis og kun hvis $\overrightarrow{X_0X} = \boldsymbol{x} - \boldsymbol{x}_0$ kan skrives som en linearkombination af $\boldsymbol{a}$ og $\boldsymbol{b}$, dvs hvis og kun hvis der findes tal



s og t, således at stedvektoren $\boldsymbol{x}$ er af formen

$$\boldsymbol{x} = s\boldsymbol{a} + t\boldsymbol{b} + \boldsymbol{x}_0 \,. \tag{2}$$

(Vi flyttede leddet $\boldsymbol{x}_0$ over på højre side.) Udtrykket (2) kaldes en *parameterfremstilling for planen.* Værdierne af de to parametre s og t kan vælges frit og uafhængigt af hinanden, og når $(s,\ t)$ gennemløber $\mathbb{R}^2$, vil X gennemløbe den plan, der er tale om.

Ved udskrivning i koordinater af (2) fås et udtryk af formen

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} sa_1 + tb_1 + x_{01} \\ sa_2 + tb_2 + x_{02} \\ sa_3 + tb_3 + x_{03} \end{pmatrix} = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \\ a_3 & b_3 \end{pmatrix} \begin{pmatrix} s \\ t \end{pmatrix} + \begin{pmatrix} x_{01} \\ x_{02} \\ x_{03} \end{pmatrix},$$

dvs parameterfremstillingen er regneudtryk for en affin afbildning fra $\mathbb{R}^2$ ("parameterplanen") til $\mathbb{R}^3$, og billedmængden ved denne afbildning er selve planen.

Begrebet "parameterfremstilling for en plan" er, ligesom for en linie (i plan eller rum) stærkt flertydig, idet $\boldsymbol{x}_0$, $\boldsymbol{a}$ og $\boldsymbol{b}$ kan vælges på mangfoldige måder. Mens en linie i $\mathbb{R}^3$ beskrives simplest ved en parameterfremstilling, må man nok sige, at en plan i $\mathbb{R}^3$ beskrives simplest ved en ligning. Men der er situationer, hvor en parameterfremstilling vil være at foretrække.

ØVELSE 31. Bestem en parameterfremstilling for planen i $\mathbb{R}^3$ gennem punkterne $(1,0,2)$, $(3,4,5)$ og $(4,2,-1)$. Bestem også en ligning for planen.

12. Eksempler på anvendelser.

Som afslutning på kapitlet skal vi se nogle eksempler på problemer i andre fag, hvor matrixregning kan være til nytte. Et par af eksemplerne tages op igen senere i bogen, når vi har fået flere metoder til rådighed. I opgavesektionen bag i bogen suppleres med en række andre anvendelser, hvoraf nogle, men ikke alle, ligger i forlængelse af stoffet nedenfor.

EKSEMPEL 28. **Lineær programmering.** (Økonomi). En virksomhed fremstiller 5 produkter P_1, P_2, P_3, P_4, P_5 ud fra 3 råvarer R_1, R_2, R_3. Vi betragter en periode, hvor der fremstilles x_j enheder P_j $(j = 1,2,3,4,5)$ og forbruges y_i enheder R_i $(i = 1,2,3)$. Antag endvidere, at der bruges a_{ij} enheder R_i til 1 enhed P_j $(i = 1,2,3,\ j = 1,2,3,4,5)$. Da er det samlede forbrug af R_i givet ved ligningen

$$y_i = a_{i1}x_1 + a_{i2}x_2 + a_{i3}x_3 + a_{i4}x_4 + a_{i5}x_5$$

$(i = 1,2,3)$. De tre ligninger kan skrives under ét på vektor-matrixform som

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}, \tag{1}$$

hvor $\boldsymbol{x} = (x_1, x_2, x_3, x_4, x_5)^\top$ er *produktionsvektoren*, $\boldsymbol{y} = (y_1, y_2, y_3)^\top$ er *råvareforbrugsvektoren*, og matricen $\boldsymbol{A} = (a_{ij})$ rummer de såkaldte *tekniske koefficienter*.

Med (1) er der konstrueret en temmelig generel model for produktionsplanlægning, men den forudsætter dog, at råvareforbruget ved fremstilling af det enkelte produkt er proportionalt med produktionens størrelse. Det er ikke altid opfyldt, specielt ikke for råstoffet arbejdstid, hvor en øget produktion i kraft af rationalisering ikke behøver at medføre et tilsvarende øget forbrug af arbejdstimer.

I praksis kan man naturligvis ikke vælge $\boldsymbol{x}$ frit i (1). For det første skal der gælde $x_j \geq 0$ for alle j, for det andet vil x_j'erne af praktiske grunde være pålagt visse begrænsninger opadtil. En vigtig type begrænsning består i, at råvarerne for den periode, det handler om, kun foreligger (eller kun kan skaffes) i visse bestemte *lagerkapaciteter* b_i, dvs der skal gælde $y_1 \leq b_1$, $y_2 \leq b_2$ og $y_3 \leq b_3$. Idet vi vedtager at samle disse tre uligheder under ét på formen $\boldsymbol{y} \leq \boldsymbol{b}$ (hvor $\boldsymbol{b} = (b_1, b_2, b_3)^\top$), skal x_j'erne altså i kraft af (1) opfylde

$$\boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}. \tag{2}$$

At finde mængden af såkaldte *tilladte punkter* $\boldsymbol{x}$, der tilfredsstiller disse uligheder (foruden $\boldsymbol{x} \geq \boldsymbol{o}$), svarer til at finde de mulige produktionsplaner for virksomheden. Blandt disse planer vil man ofte være interesseret i at finde den, der fører til størst mulig indtjening for virksomheden. Hvis vi forenkler problemet ved at holde udgifter til råvarer udenfor, og hvis vi antager, at produktet P_j kan sælges til en pris (eller: med en fortjeneste) af størrelsen p_j pr. enhed ($j = 1,2,3,4,5$), bliver den samlede indtjening ("dækningsbidraget") lig med

$$Q(\boldsymbol{x}) = p_1x_1 + p_2x_2 + \cdots + p_5x_5 = \boldsymbol{p}^\top \boldsymbol{x}.$$

Problemet kan dermed under ét formuleres på følgende måde:

$$\text{Maksimér } Q(\boldsymbol{x}) = \boldsymbol{p}^\top \boldsymbol{x} \text{ under bibetingelserne } \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b},\ \boldsymbol{x} \geq \boldsymbol{o}. \tag{3}$$

Dette er en af prototyperne på problemstillinger, som man møder inden for *lineær programmering*, der er en underdisciplin af teorien for matematisk *optimering*. I vektor-matrix formuleringen (3) kan vi generelt tænke på en virksomhed, der fremstiller n produkter ud fra m råvarer, hvorved produktionsvektoren $\boldsymbol{x}$ og prisvektoren $\boldsymbol{p}$ er $\mathbb{R}^n$-vektorer, råvareforbrugs-vektoren $\boldsymbol{y}$ og lagerkapacitets-vektoren $\boldsymbol{b}$ er $\mathbb{R}^m$-vektorer, og de tekniske koefficienter udgør en $m \times n$ matrix $\boldsymbol{A}$.

Beslægtet med (3) er situationen, hvor en lineær "objektfunktion" $Q(\boldsymbol{x})$ ikke skal maksimeres, men *minimeres*, og hvor ulighederne typisk "vender den anden vej", dvs de er af formen $\boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{b}$ (men stadig $\boldsymbol{x} \geq \boldsymbol{o}$). En sådan problemstilling kan f.eks. opstå i forbindelse med *foderplanlægning* ved husdyrproduktion, hvor man vil minimere udgifterne til foder, men samtidig stiller visse minimumskrav til foderets indhold af en række næringsstoffer, mineraler m.v. Der er også problemer, hvor nogle af bibetingelserne (evt. alle) er *ligninger* og ikke uligheder.

Lineær programmering med diverse specialtilfælde, varianter og generalisationer har fået stor betydning i moderne teoretisk økonomi, først og fremmest i det område, der betegnes *operationsøkonomi*.

EKSEMPEL 29. **Deformation.** (Fysik). Et kasseformet elastisk legeme, f.eks. en gummiklods, påvirkes af en deformerende kraft med komponenter k_1, k_2 og k_3 i tre retninger vinkelret på hinanden, dvs i det tilsvarende koordinatsystem er kraften $k = (k_1, k_2, k_3)^\top$. Eksperimenter har vist, at når kun k_1 virker, er den *relative deformation* i de tre akseretninger hhv $0.13\,k_1$, $-0.04\,k_2$ og $-0.04\,k_3$, dvs for $k_1 = 1$ bliver legemet 13% længere i kraftens retning og 4% kortere i de to retninger på tværs heraf. (I praksis gælder udtrykkene dog normalt kun med rimelig nøjagtighed for *små* værdier af k_1.) Tilsvarende udtryk gælder for hver af de to andre kraftkomponenter betragtet isoleret, og hvis legemet er *isotropt*, dvs hvis det "opfører sig ens i alle retninger" mht deformation, så er koefficienterne de samme. Antages yderligere, at virkningerne af de tre komponenter kan adderes, får vi følgende udtryk for de relative deformationer, altså for "længdeudvidelsesforholdene" λ_1, λ_2 og λ_3 i de tre akseretninger:

$$\begin{aligned}\lambda_1 &= 0.13\,k_1 - 0.04\,k_2 - 0.04\,k_3\,,\\ \lambda_2 &= -0.04\,k_1 + 0.13\,k_2 - 0.04\,k_3\,,\\ \lambda_3 &= -0.04\,k_1 - 0.04\,k_2 + 0.13\,k_3\,,\end{aligned}$$

eller på vektor-matrix form:

$$\boldsymbol{l} = \boldsymbol{A}\boldsymbol{k}, \qquad \text{hvor } \boldsymbol{l} = \begin{pmatrix}\lambda_1\\ \lambda_2\\ \lambda_3\end{pmatrix}, \quad \boldsymbol{A} = \begin{pmatrix}0.13 & -0.04 & -0.04\\ -0.04 & 0.13 & -0.04\\ -0.04 & -0.04 & 0.13\end{pmatrix}.$$

Vi kan også opstille modellen med en lidt mere generel deformationsmatrix

$$\boldsymbol{A} = \begin{pmatrix}a & -b & -b\\ -b & a & -b\\ -b & -b & a\end{pmatrix}, \tag{4}$$

hvor $a, b \in \mathbb{R}_+$. Bemærk, at $\boldsymbol{A}$ er *symmetrisk*.

ØVELSE 32. Betragt en akseparallel enhedsterning i $X_1X_2X_3$-systemet, der underkastes deformation iflg. modellen i det foregående eksempel, med koefficientmatrix $\boldsymbol{A}$ givet ved (4), hvor talstørrelserne a og b antages væsentlig mindre end 1. Antag, at der virker en enhedskraft i X_1-aksens retning, dvs vi har $\boldsymbol{k} = (1, 0, 0)^\top$. Hvad er rumfanget af den kasse, som terningen deformeres til? Vis, at rumfanget er uændret (*inkompressibelt* materiale), hvis og kun hvis $a = 2b$.

EKSEMPEL 30. **Epidemi.** (Biologi). I en given population betragtes en epidemi af en smitsom sygdom, som kan medføre døden. Forløbet af epidemien kan forenklet beskrives ved, at man løbende angiver, hvor mange individer der befinder sig i hver af følgende grupper:

1° raske (dvs endnu ikke angrebne),
2° syge,
3° døde,
4° helbredte (og, antager vi, dermed immune over for sygdommen).

Vi betegner de fire antal til tidspunkt t med hhv r_t, s_t, d_t og h_t $\quad (t = 0, 1, 2, 3, \cdots)$, og vi sætter $\boldsymbol{x}_t = (r_t, s_t, d_t, h_t)^\top$. Det anvendte tidsstep er 1 uge, dvs optællingstidspunktet t kan fortolkes som "uge t".

Om den sygdom og den population, det her drejer sig om, vil vi gøre følgende mere præcise antagelser: Enhver rask har på ethvert tidspunkt 20% risiko for at blive ramt af sygdommen inden for den følgende uge (vi forudsætter altså, næppe særlig realistisk, at smittefaren er uafhængig af antal syge), og enhver syg har på ethvert tidspunkt 10% risiko for at dø samt 20% chance for at blive helbredt inden for den følgende uge. Epidemien antages i øvrigt at strække sig over så kort tid, at man kan se bort fra fødsler samt fra andre dødsårsager end sygdommen.

Ved sammenligning af uge t og uge $t+1$ udmøntes forudsætningerne således:

1°. Af de r_t raske er 80% stadig raske, dvs

$$r_{t+1} = 0.8\, r_t\,.$$

2°. Af de r_t raske er 20% nu syge, og af de s_t syge er $100 - 10 - 20 = 70\%$ stadig syge, dvs

$$s_{t+1} = 0.2\, r_t + 0.7\, s_t\,.$$

3°. Antallet af døde er øget med 10% af de syge, dvs

$$d_{t+1} = 0.1\, s_t + d_t\,.$$

4°. Antallet af helbredte er øget med 20% af de syge, dvs

$$h_{t+1} = 0.2\, s_t + h_t\,.$$

De fire udledte ligninger kan samles på vektor-matrix form:

$$\begin{pmatrix} r_{t+1} \\ s_{t+1} \\ d_{t+1} \\ h_{t+1} \end{pmatrix} = \begin{pmatrix} 0.8 & 0 & 0 & 0 \\ 0.2 & 0.7 & 0 & 0 \\ 0 & 0.1 & 1 & 0 \\ 0 & 0.2 & 0 & 1 \end{pmatrix} \begin{pmatrix} r_t \\ s_t \\ d_t \\ h_t \end{pmatrix}$$

eller

$$\boldsymbol{x}_{t+1} = \boldsymbol{A}\boldsymbol{x}_t\,, \tag{5}$$

hvor $\boldsymbol{A}$ er koefficientmatricen.

Hvis epidemien bryder ud i en population bestående af 100 raske individer, altså

$$\boldsymbol{x}_0 = (100,\, 0,\, 0,\, 0)^{\mathsf{T}},$$

så får vi ved indsættelse i (5)

$$\boldsymbol{x}_1 = (80,\, 20,\, 0,\, 0)^{\mathsf{T}},$$

og efter samme princip i de følgende uger (der rundes af til heltal):

$$\begin{aligned}
\boldsymbol{x}_2 &= (64,\ 30,\ 2,\ 4)^{\mathsf{T}}, \\
\boldsymbol{x}_3 &= (51,\ 34,\ 5,\ 10)^{\mathsf{T}}, \\
\boldsymbol{x}_4 &= (41,\ 34,\ 8,\ 17)^{\mathsf{T}}, \\
\boldsymbol{x}_5 &= (33,\ 32,\ 11,\ 24)^{\mathsf{T}}.
\end{aligned}$$

Det ses, at epidemien kulminerer efter tre til fire ugers forløb. Man kan i øvrigt vise, at $\boldsymbol{x}_t$ for $t \to \infty$ går imod $(0,\ 0,\ 33,\ 67)^{\mathsf{T}}$, svarende til, at når der er gået lang tid, vil alle have haft sygdommen, og en trediedel af dem er døde, mens resten er helbredte og immune. Vi vender tilbage til epidemi-eksemplet i Kapitel 4, hvor vi vil være bedre i stand til at undersøge modellens opførsel.

EKSEMPEL 31. **Varians-kovarians matricen.** (Statistik.) En vektor, hvis komponenter (koordinater) er stokastiske variable, kaldes en *stokastisk vektor*. Dens forventning defineres som vektoren, der fremkommer, når hver komponent erstattes af sin forventning. (På tilsvarende måde defineres *stokastisk matrix* samt forventningen af en sådan. I dette eksempel møder vi dog kun stokastiske vektorer, mens de optrædende matricer har "rigtige tal" som elementer).

Betragt n stokastiske variable X_i med forventning $EX_i = \mu_i$ $(i = 1, 2, \cdots, n)$. Vi sætter $\boldsymbol{X} = (X_1, X_2, \cdots, X_n)^{\mathsf{T}}$, således at $E\boldsymbol{X} = (\mu_1, \mu_2, \cdots, \mu_n)^{\mathsf{T}}$. Sæt endvidere

$$\begin{aligned}
& \sigma_i^2 = \operatorname{Var} X_i && = E\{(X_i - \mu_i)^2\} \\
\text{og for } i \neq j: \quad & \sigma_{ij} = \operatorname{Cov}(X_i, X_j) && = E\{(X_i - \mu_i)(X_j - \mu_j)\}.
\end{aligned}$$

Ved *varians-kovarians matricen* for $\boldsymbol{X}$ forstås den kvadratiske n'te ordens matrix

$$\boldsymbol{V} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2n} \\ \vdots & & & \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_n^2 \end{pmatrix}.$$

Ud fra definitionen af σ_{ij} er det klart, at der gælder $\sigma_{ji} = \sigma_{ij}$, dvs matricen $\boldsymbol{V}$ er *symmetrisk*. Den udgør ikke blot en overskuelig måde at opskrive de talstørrelser på, der tilsammen beskriver X_i'ernes variation, men viser sig også at kunne udnyttes i regneudtryk for variationen af mere komplicerede størrelser, som afhænger af X_i'erne. Derved undgår man, at disse regneudtryk svulmer op og bliver uhåndterlige. Vi omtaler kort det grundlæggende eksempel på en sådan situation.

Antag, at vi sammen med X_i'erne har brug for nogle andre stokastiske variable $Y_1, Y_2, \cdots, Y_m$, som er udtrykt lineært ved X_i'erne:

$$Y_i = a_{i1}X_1 + a_{i2}X_2 + \cdots + a_{in}X_n \qquad (i = 1, 2, \cdots, m), \tag{6}$$

eller på vektor-matrix form

$$\boldsymbol{Y} = \boldsymbol{A}\boldsymbol{X}, \tag{7}$$

hvor $\boldsymbol{A}$ er den $m \times n$ matrix, der dannes af koefficienterne a_{ij} i ligningerne (6).

Vi vil nu udregne dels forventningen, dels varians-kovarians matricen for $\boldsymbol{Y}$, udtrykt ved de tilsvarende størrelser for $\boldsymbol{X}$. Ved at tage forventningen af hver side af (6) og sammenstille de fremkomne ligninger får vi

$$E\boldsymbol{Y} = \boldsymbol{A}\,E\boldsymbol{X}.$$

Denne ligning kan opfattes som en generalisering af den regel for stokastiske variable, der siger, at "en konstant faktor bibeholdes, når man danner forventning", dvs $Y = aX$ medfører $EY = a\,EX$. Reglen gælder så at sige også for en konstant matrix-faktor mellem stokastiske vektorer, som ikke behøver at have samme dimension. Reglen gælder i øvrigt også for en matrix-faktor "efter" den stokastiske vektor, dvs for en sammenhæng af typen $\boldsymbol{Z}^\top = \boldsymbol{X}^\top\boldsymbol{B}$. Det får vi brug for om lidt.

En anden regel for stokastiske variable siger, at "en konstant faktor kvadreres, når man danner varians", dvs $Y = aX$ medfører $\operatorname{Var} Y = a^2 \operatorname{Var} X$. Det er vort mål at generalisere også denne regel til vektortilfældet $\boldsymbol{Y} = \boldsymbol{A}\boldsymbol{X}$. Først bemærkes, at $\boldsymbol{V}$ iflg. Eksempel 24, side 64 kan skrives på formen

$$\boldsymbol{V} = E\{(\boldsymbol{X} - E\boldsymbol{X})(\boldsymbol{X} - E\boldsymbol{X})^\top\}$$

og tilsvarende for $\boldsymbol{Y}$'s varians-kovarians matrix, som vi betegner $\boldsymbol{W}$:

$$\begin{aligned}
\boldsymbol{W} &= E\{(\boldsymbol{Y} - E\boldsymbol{Y})(\boldsymbol{Y} - E\boldsymbol{Y})^\top\} \\
&= E\{(\boldsymbol{A}\boldsymbol{X} - \boldsymbol{A}\,E\boldsymbol{X})(\boldsymbol{A}\boldsymbol{X} - \boldsymbol{A}\,E\boldsymbol{X})^\top\} \\
&= E\{\boldsymbol{A}(\boldsymbol{X} - E\boldsymbol{X})(\boldsymbol{X}^\top\boldsymbol{A}^\top - E\boldsymbol{X}^\top\boldsymbol{A}^\top)\} \\
&= \boldsymbol{A}\,E\{(\boldsymbol{X} - E\boldsymbol{X})(\boldsymbol{X} - E\boldsymbol{X})^\top\}\boldsymbol{A}^\top,
\end{aligned}$$

hvor vi satte $\boldsymbol{A}$ og $\boldsymbol{A}^\top$ uden for $E\{\ \}$ iflg. reglen ovenfor. Altså:

$$\boldsymbol{W} = \boldsymbol{A}\boldsymbol{V}\boldsymbol{A}^\top. \tag{8}$$

Ved (8) er sammenhængen mellem varians-kovarians matricerne for hhv $\boldsymbol{X}$ og $\boldsymbol{Y}$ givet. I det specielle tilfælde, hvor $m = n = 1$, forenkles de to stokastiske vektorer til stokastiske variable X og Y, $\boldsymbol{A}$ er en 1×1 matrix (a), og varians-kovarians matricerne er ligeledes begge 1×1 matricer, med hhv $\operatorname{Var} X$ og $\operatorname{Var} Y$ som eneste element. Vi kan dermed skrive (8) som en talligning:

$$\operatorname{Var} Y = a \cdot \operatorname{Var} X \cdot a = a^2 \operatorname{Var} X.$$

Det fremgår, at (8) virkelig er en generalisation af førnævnte regel for variansdannelse.

ØVELSE 33. Fire stokastiske variable X_1, X_2, Y_1, Y_2 er forbundet ved

$$\begin{aligned} Y_1 &= 2X_1 + 3X_2, \\ Y_2 &= X_1 + 4X_2. \end{aligned}$$

Det oplyses, at X_1 og X_2 er uafhængige og har samme varians, som er 1. Find korrelationskoefficienten mellem Y_1 og Y_2, dvs tallet

$$\rho = \frac{\operatorname{Cov}(Y_1, Y_2)}{\sqrt{\operatorname{Var} X_1}\sqrt{\operatorname{Var} X_2}}.$$

[Vink: Find først varians-kovarians matricen $\boldsymbol{W}$ for $(Y_1, Y_2)^{\mathrm{T}}$ ved brug af (8), hvor $\boldsymbol{V}$ iflg. det forudsatte er lig med 2×2 enhedsmatricen.]

EKSEMPEL 32. **Aldersfordelt populationsvækst.** (Biologi). Betragt en dyrepopulation, hvis størrelse beskrives ved antal hunner x_t ved en optælling i år t. (Af hensyn til forholdene ved formering nøjes vi med at tælle hunnerne.) Det antages, at hunnerne i gennemsnit har 70% chance for at overleve mindst et år, og at de i gennemsnit føder 0.6 hununger, som er i live et år senere. Antal hunner i år $t+1$ findes som summen af de overlevende og de nyfødte:

$$x_{t+1} = \frac{70}{100} x_t + 0.6\, x_t = 1.3\, x_t. \tag{9}$$

Ved successiv anvendelse af (9) fås

$$x_t = 1.3\, x_{t-1} = 1.3^2 x_{t-2} = \cdots = 1.3^t x_0.$$

Man siger, at populationen *vokser eksponentielt*, idet tidsvariablen t optræder som eksponent. Hvis populationen starter til tidspunkt 0 med f.eks. $x_0 = 100$ hunner, vil den i løbet af 20 år vokse til $100 \cdot 1.3^{20} \approx 19\,000$ hunner. En sådan vækst er ikke mulig på længere sigt. Modellens antagelser gælder kun, så længe populationens størrelse ikke er begyndt at hæmme væksten pga mangel på føde og plads. Når det sker, aftager både fødsels- og overlevelsesrate, og væksten flader ud. Under i øvrigt stabile forhold vil x_t da nærme sig en (nogenlunde) konstant værdi, *bærekapaciteten* mht den pågældende art af det afgrænsede område, der betragtes.

Men denne fase med *begrænset vækst* vil vi ikke gå nærmere ind på her. Vi vil som i starten af eksemplet antage, at formerings- og overlevelsesmønstret er konstant, men til gengæld forfiner vi beskrivelsen på en anden led, idet vi tager hensyn til, at hunnerne i populationen ikke er ens.

Den mest markante forskel mellem dyrene er, at de er af forskellig *alder*. For atter at have et konkret udgangspunkt vil vi betragte en art med maksimal levealder på 5 år. Størrelse og fordeling på aldersgrupper af en given population kan da beskrives ved vektoren

$$\boldsymbol{x}_t = (x_{t0}, x_{t1}, x_{t2}, x_{t3}, x_{t4})^{\mathsf{T}},$$

hvis koordinater er antallene af hhv 0-, 1-, 2-, 3- og 4-årige hunner. I de fem aldersgrupper antager vi, at der er følgende *aldersspecifikke* overlevelses- og fødselsrater:

alder	gnsn. overlevelseschance	gnsn. antal hununger
0	70 %	0.8
1	80 %	0.9
2	60 %	0.7
3	40 %	0.5
4	(0 %)	0.4

Populationen $\boldsymbol{x}_t$ giver anledning til følgende antal nyfødte hunner 1 år senere:

$$x_{t+1,0} = 0.8\,x_{t0} + 0.9\,x_{t1} + 0.7\,x_{t2} + 0.5\,x_{t3} + 0.4\,x_{t4}.$$

De øvrige komponenter i $\boldsymbol{x}_{t+1}$ findes som brøkdelen af overlevende fra den foregående komponent i $\boldsymbol{x}_t$. Altså får vi ved brug af skemaet ovenfor:

$$\begin{aligned} x_{t+1,1} &= 0.7\,x_{t0}, \\ x_{t+1,2} &= 0.8\,x_{t1}, \\ x_{t+1,3} &= 0.6\,x_{t2}, \\ x_{t+1,4} &= 0.4\,x_{t3}. \end{aligned}$$

De fem ligninger kan samles på vektor-matrix form til

$$\boldsymbol{x}_{t+1} = \boldsymbol{A}\boldsymbol{x}_t, \qquad \text{hvor} \quad \boldsymbol{A} = \begin{pmatrix} 0.8 & 0.9 & 0.7 & 0.5 & 0.4 \\ 0.7 & 0 & 0 & 0 & 0 \\ 0 & 0.8 & 0 & 0 & 0 \\ 0 & 0 & 0.6 & 0 & 0 \\ 0 & 0 & 0 & 0.4 & 0 \end{pmatrix}. \tag{10}$$

Ganske som i det ikke-aldersopdelte tilfælde, vi så på i starten af eksemplet, kan vi *iterere* (10) og derved udlede følgende direkte udtryk for $\boldsymbol{x}_t$:

$$\boldsymbol{x}_t = \boldsymbol{A}\boldsymbol{x}_{t-1} = \boldsymbol{A}^2\boldsymbol{x}_{t-2} = \cdots = \boldsymbol{A}^t\boldsymbol{x}_0. \tag{11}$$

For en forelagt startpopulation $\boldsymbol{x}_0$ kan man ved brug af (10)/(11) så at sige udregne populationens udvikling i tidens løb. Der er atter tale om en slags "eksponentiel vækst", men med den væsentlige komplikation i forhold til det ikke-aldersopdelte tilfælde, at det er en *matrix*, ikke et tal, vi opløfter til potens. Også denne model vender vi tilbage til i Kapitel 4, hvor vi vil være bedre rustet til at udtale os om, hvordan en matrix opfører sig, når den opløftes til voksende potens. Det viser sig, at modellen på langt sigt forudsiger "næsten-eksponentiel vækst". Noget andet er så, at inden denne stabilisering når at indtræde, har populationens vækst gjort modellen urealistisk, jfr. betragtningerne ovenfor.

En model af typen $\boldsymbol{x}_{t+1} = \boldsymbol{A}\boldsymbol{x}_t$, som vi har mødt dels i Eksempel 30 (epidemien), dels i (10) ovenfor, kaldes en *matrixprojektion*. For en art med en anden maksimalalder og andre overlevelses- og formeringsrater kan man under lignende forudsætninger som ovenfor opstille en model analog med (10). Karakteristisk for "projektionsmatricen" $\boldsymbol{A}$ vil være, at den er kvadratisk, med fødselsraterne som elementer i øverste række og overlevelsesraterne i "subdiagonalen", dvs den skrå række lige under diagonalen.

ØVELSE 34. Betragt en art, hvor alle hunner overlever til en alder af præcis 2 år. Det første år får de i gennemsnit 1 hununge, det andet år 3 hununger. Opstil en model af formen (10) for en population af arten. Sæt $\boldsymbol{x}_0 = (100, 0)^\mathsf{T}$, og udregn populationsvektorer for et antal af de følgende år. Har du indtryk af, at fordelingen mellem "unge" og "gamle", dvs forholdet x_{t0}/x_{t1}, stabiliserer sig for voksende t? Idet vi definerer den årlige *vækstfaktor* som forholdet

$$\frac{x_{t+1,0} + x_{t+1,1}}{x_{t0} + x_{t1}}$$

mellem de samlede antal hunner i hhv år $t+1$ og år t, skal du også undersøge, om denne størrelse synes at stabilisere sig for voksende t.

Sæt i stedet $\boldsymbol{x} = (50, 50)^\mathsf{T}$. Hvordan udvikler populationen sig med denne startvektor? Hvordan stemmer det med, hvad du fandt for vækstfaktoren ovenfor?

Forelæsning F3

Matematik og modeller

Blok 3 – Forår 2006

EKSEMPEL 2 – Side LA133

```
Vi løser ligningen Ax=b, hvor A er:          og b er:
> A   [,1] [,2] [,3] [,4] [,5]
R1     1    1    0    2    1                  b=(2,7,2,17)
R2     2    3    1    5    2
R3    -1    0    1    1   -2
R4     3    6    3   13    1

Vi bruger totalmatricen: Ab= [A,b]
```


Række operationer

```
> Ab<-matrix(c(1,2,-1,3,1,3,0,6,0,1,1,3,2,5,1,13,1,2,-2,1,2,7,2,17),4)

> Ab
   [,1] [,2] [,3] [,4] [,5] [,6]
R1    1    1    0    2    1    2        Vi foretager den rækkeoperation,
R2    2    3    1    5    2    7        der giver 0 på pladsen (2,1)
R3   -1    0    1    1   -2    2        vedhjælp af R1:
R4    3    6    3   13    1   17
> A1<-Ab; A1<-roperation2(A1,2,1,1) ; A1    R2 → R2 - 2*R1:
   [,1] [,2] [,3] [,4] [,5] [,6]
R1    1    1    0    2    1    2
R2    0    1    1    1    0    3
R3   -1    0    1    1   -2    2
R4    3    6    3   13    1   17
> A1<-roperation2(A1,3,1,1); A1             R3 → R3 + R1:
   [,1] [,2] [,3] [,4] [,5] [,6]
R1    1    1    0    2    1    2
R2    0    1    1    1    0    3
R3    0    1    1    3   -1    4
R4    3    6    3   13    1   17
> A1<-roperation2(A1,4,1,1); A1             R4 → R4 - 3*R1:
   [,1] [,2] [,3] [,4] [,5] [,6]
R1    1    1    0    2    1    2
R2    0    1    1    1    0    3
R3    0    1    1    3   -1    4
R4    0    3    3    7   -2   11
```


LA 82

Vi samler nu rækkeoperationer, der skaffer nuller under elementet (1,1) vha. R1, etc

```
> A1<-Ab
> A1
     [,1] [,2] [,3] [,4] [,5] [,6]
[1,]    1    1    0    2    1    2
[2,]    2    3    1    5    2    7
[3,]   -1    0    1    1   -2    2
[4,]    3    6    3   13    1   17
> A1<-roperation3(A1,1,1)
> A1
     [,1] [,2] [,3] [,4] [,5] [,6]
[1,]    1    1    0    2    1    2
[2,]    0    1    1    1    0    3
[3,]    0    1    1    3   -1    4
[4,]    0    3    3    7   -2   11
```

```
> A1<-roperation3(A1,2,2)
> A1
     [,1] [,2] [,3] [,4] [,5] [,6]
[1,]    1    1    0    2    1    2
[2,]    0    1    1    1    0    3
[3,]    0    0    0    2   -1    1
[4,]    0    0    0    4   -2    2
> A1<-roperation3(A1,3,4)
> A1
     [,1] [,2] [,3] [,4] [,5] [,6]
[1,]    1    1    0    2    1    2
[2,]    0    1    1    1    0    3
[3,]    0    0    0    2   -1    1
[4,]    0    0    0    0    0    0
```

```
Vi kan nu løse ligningerne
Løsningsrummet har 2 parametre!

rang(A)=3, rang(Ab)=3
```


Vi ændrer A-matricen og gentager spøgen:

```
> A[4,6]<-8; A1<-A; A1

     [,1] [,2] [,3] [,4] [,5] [,6]
[1,]    1    1    0    2    1    2
[2,]    2    3    1    5    2    7
[3,]   -1    0    1    1   -2    2
[4,]    3    6    3   13    1    8
> A1<-roperation3(A1,1,1)
> A1
     [,1] [,2] [,3] [,4] [,5] [,6]
[1,]    1    1    0    2    1    2
[2,]    0    1    1    1    0    3
[3,]    0    1    1    3   -1    4
[4,]    0    3    3    7   -2    2
> A1<-roperation3(A1,2,2)
> A1
     [,1] [,2] [,3] [,4] [,5] [,6]
[1,]    1    1    0    2    1    2
[2,]    0    1    1    1    0    3
[3,]    0    0    0    2   -1    1
[4,]    0    0    0    4   -2   -7
> A1<-roperation3(A1,3,4)
> A1
     [,1] [,2] [,3] [,4] [,5] [,6]
[1,]    1    1    0    2    1    2
[2,]    0    1    1    1    0    3
[3,]    0    0    0    2   -1    1
[4,]    0    0    0    0    0   -9
```

Der er ingen løsninger: `rang(A)=3`, `rang(Ab)=4`. Definitionen af `rang(A)` følger:

Rangen af en n x m - matrtix

Definition: Rangen af en m x n matrix `A` `rang(A)` er det maksimale antal lin. uafhængige søjler. eller $\varrho(A) = \text{rang}(A)$

Sætning: `rang(A)` ændres ikke ved rækkeoperationer og søjleombytninger,

$\text{rang}(A) = \text{rang}(A^T)$,

`rang(A)` er mindre en `min(m,n)`.

Beregning af `rang(A)`: Foretag række-operationer, og evt. række- og søjle-ombytninger, så `A` bliver en trappe-matrix, med 0'er under "trappen":

`rang(A)` = # trin på "trappen".

Sætning: Ligningssystemet `Ax=b` har løsninger hvis og kun hvis

`rang(A)=rang([A,b])`

Normalt:

`m=n:` éntydig løsning

`m>n:` ingen løsning

`m<n:` uendelig mange løsninger (n-m parametre)

LA 84

KAPITEL 3

LINEÆRE LIGNINGSSYSTEMER

0. Indledning.

Temaet i dette kapitel er grundlæggende i lineær algebra, og det dukker i forskellige forklædninger op i de fleste af dens anvendelser. Det drejer sig om at "løse m lineære ligninger med n ubekendte". Med den sprogbrug, vi indførte i Kapitel 1, kan problemet formuleres således: Løs matrixligningen $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$, hvor $\boldsymbol{A}$ er en given $m \times n$ matrix og $\boldsymbol{b} \in \mathbb{R}^m$ en given vektor.

Lad os begynde med at se på de tilfælde, vi umiddelbart kan overskue, fordi ligningen $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ og dens løsningsmængde kan fortolkes på en tallinie, i planen eller i rummet. Via taleksempler giver vi en oversigt over de ni tilfælde, der svarer til $m = 1, 2, 3$ kombineret med $n = 1, 2, 3$.

EKSEMPEL 1. *Tilfældet* $m = 1$, $n = 1$. Ligningen $3x = 12$ kan karakteriseres som "1 lineær ligning med 1 ubekendt". Den har netop én løsning, nemlig $x = \frac{1}{3} \cdot 12 = 4$. Generelt vil ligningen $ax = b$ normalt have den *entydige* løsning $x = \frac{b}{a}$. Men ordet normalt skal tages bogstaveligt: det kunne jo tænkes, at $a = 0$, og da har ligningen *ingen* løsninger. Med mindre der samtidig gælder $b = 0$, for da har ligningen *uendelig mange* løsninger, idet $0 \cdot x = 0$ er sandt for alle $x \in \mathbb{R}$.

Allerede disse i sig selv trivielle overvejelser rummer et par af de pointer, der viser sig at gælde alment. For det første: Der kan være 1) ingen løsninger, 2) en entydig løsning eller 3) uendelig mange løsninger. For det andet: Ud fra antal ubekendte og antal ligninger forventer man en bestemt af disse tre muligheder, men man kan ikke være sikker på, at den "normale" situation foreligger - det kommer an på systemet. Vi belyser disse bemærkninger med flere af de anskuelige tilfælde.

Tilfældet $m = 1, n = 2$. Ligningen $2x_1 + 3x_2 = 7$ kan karakteriseres som "1 lineær ligning med 2 ubekendte". Den har uendelig mange løsninger, idet vi kan vælge f.eks. x_2 frit, hvorefter x_1 kan findes af ligningen, som så vil være tilfredsstillet af det konstruerede talpar $\boldsymbol{x} = (x_1, x_2)$. Sættes $x_2 = t$, får vi $x_1 = \frac{1}{2}(7-3t) = \frac{7}{2} - \frac{3}{2}t$, dvs løsningen kan skrives samlet på formen

$$\begin{aligned} x_1 &= \tfrac{7}{2} - \tfrac{3}{2}t, \\ x_2 &= t \end{aligned} \qquad \text{eller} \qquad \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \frac{7}{2} \\ 0 \end{pmatrix} + t \begin{pmatrix} -\frac{3}{2} \\ 1 \end{pmatrix},$$

hvor $t \in \mathbb{R}$. Dette er en parameterfremstilling for løsningsmængden, der kan fortolkes som en ret linie i X_1X_2-systemet.

Generelt har ligningen $a_1x_1 + a_2x_2 = b$ uendelig mange løsninger, idet den fremstiller en ret linie i X_1X_2-systemet. Det forudsætter dog, at mindst et af tallene a_1 og a_2 er $\neq 0$. Når $a_1 = a_2 = 0$, $b \neq 0$, er der ingen løsninger, og når $a_1 = a_2 = b = 0$, er der uendelig mange løsninger, nemlig alle $\boldsymbol{x} \in \mathbb{R}^2$.

Tilfældet $m = 1, n = 3$. Ligningen $2x_1 + 3x_2 + 5x_3 = 7$ kan karakteriseres som "1 lineær ligning med 3 ubekendte". Den har uendelig mange løsninger, idet vi kan vælge både f.eks. x_2 og x_3 frit, hvorefter x_1 kan findes af ligningen, som så vil være tilfredsstillet af det konstruerede talsæt $\boldsymbol{x} = (x_1, x_2, x_3)$. Med $x_2 = t_1, x_3 = t_2$ får vi $x_1 = \frac{1}{2}(7 - 3t_1 - 5t_2) = \frac{7}{2} - \frac{3}{2}t_1 - \frac{5}{2}t_2$, dvs løsningen kan skrives samlet på formen

$$\begin{aligned} x_1 &= \tfrac{7}{2} - \tfrac{3}{2}t_1 - \tfrac{5}{2}t_2, \\ x_2 &= \quad\; t_1 \quad\quad\;\; , \\ x_3 &= \qquad\qquad t_2 \end{aligned} \qquad \text{eller} \qquad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \frac{7}{2} \\ 0 \\ 0 \end{pmatrix} + t_1 \begin{pmatrix} -\frac{3}{2} \\ 1 \\ 0 \end{pmatrix} + t_2 \begin{pmatrix} -\frac{5}{2} \\ 0 \\ 1 \end{pmatrix} \tag{1}$$

hvor $t_1, t_2 \in \mathbb{R}$. Iflg. Eksempel 27, side 71-72 er (1) parameterfremstilling for en *plan* i $X_1X_2X_3$-systemet. Det er så lidt af en smagssag, om man i grunden opnår en forenkling ved at omskrive "mængden af $\boldsymbol{x} \in \mathbb{R}^3$, der tilfredsstiller ligningen $2x_1 + 3x_2 + 5x_3 = 7$" til "mængden af $\boldsymbol{x} \in \mathbb{R}^3$, der kan skrives på formen (1)."

Generelt har ligningen $a_1x_1 + a_2x_2 + a_3x_3 = b$ uendelig mange løsninger, idet den fremstiller en plan i $X_1X_2X_3$-systemet. Det forudsætter dog, at mindst en af a_i'erne er $\neq 0$. Når $a_1 = a_2 = a_3 = 0$, $b \neq 0$, er der ingen løsninger, og når $a_1 = a_2 = a_3 = b = 0$, er der uendelig mange løsninger, nemlig alle $\boldsymbol{x} \in \mathbb{R}^3$.

Tilfældet $m = 2, n = 1$. Ligningerne $3x = 12$, $5x = -3$ udgør tilsammen et system på "2 lineære ligninger med 1 ubekendt". Systemet har ingen løsninger. Den første ligning er nemlig kun opfyldt af $x = 4$, og denne x-værdi stemmer ikke i den anden ligning (der kun er opfyldt af $x = -\frac{3}{5}$.)

Generelt har ligningerne $a_1 x = b_1$, $a_2 x = b_2$ ingen fælles løsning. Som ved de foregående systemer kan det dog for specielle værdier af konstanterne ske, at der er en entydig løsning, evt. (når alle fire koefficienter er 0) at der er uendelig mange.

Tilfældet $m = 2$, $n = 2$. Ligningerne $2x_1 + 3x_2 = 7$, $5x_1 + 2x_2 = 1$ udgør tilsammen et system på "2 lineære ligninger med 2 ubekendte." De løses på velkendt måde, f.eks. ved lige store koefficienters metode. Det viser sig, at $x_1 = -1$, $x_2 = 3$.

Generelt vil 2 lineære ligninger med 2 ubekendte normalt have en entydig løsning, svarende til, at 2 linier i planen (X_1X_2-systemet) skærer hinanden i ét punkt. For specielle værdier af koefficienterne kan det dog ske, enten at der ingen løsninger er (typisk svarende til, at de to linier er parallelle), eller at der er uendelig mange løsninger (de to linier falder helt sammen).

Tilfældet $m = 2$, $n = 3$. Ligningerne $2x_1 + 3x_2 + 5x_3 = 7$, $x_1 + x_2 + 2x_3 = 3$ udgør tilsammen et system på "2 lineære ligninger med 3 ubekendte". Systemet har uendelig mange løsninger, idet vi kan vælge x_3 frit, hvorefter x_1 og x_2 kan findes af de to ligninger. Mere præcist: Sætter vi $x_3 = t$, antager systemet formen $2x_1 + 3x_2 = 7 - 5t$, $x_1 + x_2 = 3 - 2t$. Hvis vi ganger ligningerne med hhv -1 og 3 og derefter lægger dem sammen, elimineres x_2, og vi får $-2x_1 + 3x_1 = -7 + 5t + 9 - 6t$ eller $x_1 = 2 - t$. Indsættes dette i den anden af ligningerne, får vi $x_2 = 3 - 2t - x_1 = 3 - 2t - (2 - t) = 1 - t$. Løsningsmængden kan dermed skrives samlet på formen

$$\begin{aligned} x_1 &= 2 - t, \\ x_2 &= 1 - t, \\ x_3 &= t \end{aligned} \qquad \text{eller} \qquad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}.$$

Dette er en *parameterfremstilling* for løsningsmængden, der kan fortolkes som en ret linie i $X_1X_2X_3$-systemet.

Generelt vil 2 lineære ligninger med 3 ubekendte normalt have uendelig mange løsninger, svarende til, at 2 planer i rummet ($X_1X_2X_3$-systemet) skærer hinanden i en linie. For specielle værdier af koefficienterne kan det dog også ske, at der ingen løsninger er (typisk svarende til, at de to planer er parallelle).

Det skulle nu være muligt at se en tendens i de seks behandlede tilfælde, og vi omtaler derfor de sidste tre (med $m = 3$) lidt mere summarisk.

Tilfældet $m = 3$, $n = 1$. Et system på "3 lineære ligninger med 1 ubekendt" har normalt ingen løsninger, af lignende grunde som i tilfældet $m = 2$, $n = 1$. For specielle værdier af konstanterne kan det dog ske, at der er en entydig løsning, evt. (når alle seks koefficienter er =0) at der er uendelig mange.

Tilfældet $m = 3, n = 2$. Et system på "3 lineære ligninger med 2 ubekendte" har normalt ingen løsninger. Løser man nemlig delsystemet af f.eks. de to første ligninger, giver det i reglen en entydig løsning, der imidlertid kun undtagelsesvis også stemmer i den tredje ligning. Geometrisk svarer hertil, at tre linier i X_1X_2-systemet normalt ikke går gennem samme punkt. Det kan dog ske, at de gør det, og på linie hermed kan systemet for specielle værdier af de optrædende koefficienter have en entydig løsning eller evt. uendelig mange løsninger.

Tilfældet $m = 3, n = 3$. Et system på "3 lineære ligninger med 3 ubekendte" har normalt en entydig løsning, der kan findes på velkendt måde ved successiv elimination. Stereometrisk svarer hertil, at de tre planer, der fremstilles i $X_1X_2X_3$-systemet af de enkelte ligninger, normalt skærer hinanden i ét punkt. Det kan dog også ske, at de ikke gør det, og svarende hertil kan systemet for specielle værdier af de optrædende koefficienter enten være uden løsninger eller have uendelig mange.

Om de ni tilfælde, der omtales i Eksempel 1, gælder sammenfattende:

Når $\underline{m > n}$, er der normalt *ingen* løsninger.
Når $\underline{m = n}$, er der normalt en *entydig* løsning.
Når $\underline{m < n}$, er der normalt *uendelig mange* løsninger. Specielt er der for $n - m = 1$ typisk en "enkelt uendelighed" af løsninger (linie), for $n - m = 2$ en "dobbelt uendelighed" af løsninger (plan).

I næste afsnit skal vi se, at de hovedlinier, der her er antydet, holder stik for matrixligningen $\boldsymbol{Ax} = \boldsymbol{b}$ generelt.

1. Lineært ligningsystem.

Ved et *lineært ligningssystem* forstås et system af typen

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

hvor *koefficienterne* a_{ij} og *højresiderne* b_i er givne tal. Man søger løsningsmængden bestående af de $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)^\top \in \mathbb{R}^n$, der tilfredsstiller alle m ligninger. Ligningerne kan samles under ét på matrixform som

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}\,. \tag{2}$$

Her indfører vi et par vigtige betegnelser.

DEFINITION 1. **Koefficientmatrix og totalmatrix.** Ved *koefficientmatricen* for systemet (1) forstås den $m \times n$ matrix $\boldsymbol{A} = (a_{ij})$, der dannes af koefficienterne til de ubekendte på ligningernes venstresider. Ved *totalmatricen* for (1) forstås den $m \times (n+1)$ matrix $(\boldsymbol{A}\,|\,\boldsymbol{b})$, der fås af koefficientmatricen ved at tilføje systemets højresider som en ekstra søjle.

Endnu en måde at skrive systemet (1) på er

$$x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \cdots + x_n\boldsymbol{a}_n = \boldsymbol{b}\,, \tag{3}$$

hvor $\boldsymbol{a}_j$'erne som sædvanlig betegner søjlevektorerne i $\boldsymbol{A}$. Formulering (3) baner vejen til følgende karakterisering af den situation, at systemets løsningsmængde ikke er tom, altså af, at der er mindst én løsning.

SÆTNING 1. Løsningsmængden til systemet (1) er ikke-tom, hvis og kun hvis dets koefficientmatrix og dets totalmatrix har samme rang, eller udtrykt i en formel: hvis og kun hvis $\rho(\boldsymbol{A}) = \rho(\boldsymbol{A}\,|\,\boldsymbol{b})$.

$\rho(A) = \text{rang}(A)$!

BEVIS: At der er løsninger til systemet, betyder iflg. (3), at $\boldsymbol{b}$ kan skrives på mindst én måde som linearkombination af søjlerne i $\boldsymbol{A}$. Med andre ord, at $\boldsymbol{b}$ tilhører det underrum af $\mathbb{R}^m$, der udspændes af disse søjler. Men det er igen det samme som at sige, at søjle-maksimalgraden af $\boldsymbol{A}$ ikke øges, når vi tilføjer $\boldsymbol{b}$ som ekstra søjle. Da rangen af en matrix er lig med dens søjle-maksimalgrad, følger sætningen. □

ØVELSE 1. Løs hvert af systemerne 1) $\begin{pmatrix} 5 & 2 \\ 5 & 3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 7 \\ 17 \end{pmatrix}$,

2) $\begin{pmatrix} 2 & 1 \\ 6 & 3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 7 \\ 17 \end{pmatrix}$, 3) $\begin{pmatrix} 2 & 1 \\ 6 & 3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 7 \\ 21 \end{pmatrix}$.

Udregn i hvert af de tre tilfælde dels $\rho(\boldsymbol{A})$, dels $\rho(\boldsymbol{A}\,|\,\boldsymbol{b})$. Er Sætning 1 opfyldt?

Nu til vort hovedproblem, systemet (1). Hvordan kan vi ændre og helst forenkle dette system uden at ændre løsningsmængden? Derom handler

SÆTNING 2. Løsningsmængden for systemet (1) er uændret ved hver af følgende operationer udført på systemets totalmatrix $(\boldsymbol{A}\,|\,\boldsymbol{b})$:

1° en rækkeoperation,
2° en rækkeombytning,
3° multiplikation af en række med et tal $t \neq 0$.

Løsningsmængden for (1) er i det væsentlige uændret ved

4° en søjleombytning i koefficientmatricen $\boldsymbol{A}$,

idet denne blot bevirker, at de to tilsvarende ubekendte (koordinater) skal ombyttes i enhver løsning $\boldsymbol{x}$.

BEVIS: 1° At den i'te række i $(\boldsymbol{A}\,|\,\boldsymbol{b})$ ganges med t og lægges til den j'te række, svarer til den ændring af systemet (1), at den i'te ligning ganges med t og lægges til den j'te ligning. Et talsæt $\boldsymbol{x} \in \mathbb{R}^n$, der tilfredsstiller (1), vil også tilfredsstille det ændrede system. Da man kan regne tilbage fra det ændrede system til (1), idet den i'te ligning ganges med $-t$ og lægges til den j'te, er en løsning til det ændrede system også en løsning til (1). Med andre ord, de to systemer har samme løsningsmængde.

2° At den i'te række i $(\boldsymbol{A}\,|\,\boldsymbol{b})$ ombyttes med den j'te række, svarer til, at den i'te ligning i systemet (1) ombyttes med den j'te ligning. Da rækkefølgen af ligningerne i (1) ikke spiller nogen rolle for, hvilke talsæt der tilfredsstiller hele systemet, har de to systemer samme løsningsmængde.

3° At den i'te række i $(\boldsymbol{A}\,|\,\boldsymbol{b})$ ganges med $t \neq 0$, svarer til den ændring af systemet (1), at der ganges igennem med t på begge sider i den i'te ligning.

Hvis talsættet $\boldsymbol{x}$ tilfredsstiller (1), vil det også tilfredsstille det ændrede system. Da man kan regne tilbage fra dette til (1) ved at gange igennem med $1/t$ i den i'te ligning, er en løsning til det ændrede system også en løsning til (1). De to systemer har således samme løsningsmængde.

4° At den i'te søjle i $\boldsymbol{A}$ ombyttes med den j'te søjle, svarer til, at x_i-leddene i hver enkelt ligning i systemet ombyttes med x_j-leddene. Det ændrer ikke ligningerne og dermed heller ikke systemets løsningsmængde. Blot må man huske, at i et talsæt, der er frembragt som løsning ud fra det ændrede system, står x_i-værdien på den j'te plads og omvendt, mens øvrige ubekendte er på de rigtige pladser (med mindre der er lavet flere søjleombytninger). □

Det er ikke noget tilfælde, at Sætning 2 ikke omtaler søjleoperationer på $(\boldsymbol{A}\,|\,\boldsymbol{b})$ eller $\boldsymbol{A}$. Dem skal man holde sig fra, da de ændrer løsningsmængden radikalt. Husk også, at højresiden $\boldsymbol{b}$ *ikke* må indgå i søjleombytninger. Med disse begrænsninger er Sætning 2 et nyttigt redskab, når man vil "regne ensbetydende" på et lineært ligningssystem og derved successivt forenkle det. Sætningen koncentrerer regningerne til det væsentlige, nemlig koefficienter og højresider i systemet. På dette grundlag vil vi nu skitsere en *standardmetode* til at løse (1), dvs til at frembringe en parameterfremstilling for løsningsmængden.

Først opskrives totalmatricen $(\boldsymbol{A}\,|\,\boldsymbol{b})$. Ved rækkeoperationer, rækkeombytninger og tal-gange-række samt om nødvendigt søjleombytninger (NB! kun i $\boldsymbol{A}$) skaffer vi nuller under diagonalen og tal $\neq 0$ så langt ned gennem diagonalen som muligt. Herved når vi til en matrix af formen

$$\begin{pmatrix} \hat{a}_{11} & \hat{a}_{12} & \dots & \hat{a}_{1r} & \hat{a}_{1,r+1} & \dots & \hat{a}_{1n} & \hat{b}_1 \\ 0 & \hat{a}_{22} & \dots & \hat{a}_{2r} & \hat{a}_{2,r+1} & \dots & \hat{a}_{2n} & \hat{b}_2 \\ & \vdots & & & & & & \\ 0 & 0 & \dots & \hat{a}_{rr} & \hat{a}_{r,r+1} & \dots & \hat{a}_{rn} & \hat{b}_r \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 & \hat{b}_{r+1} \\ & \vdots & & & & & & \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 & \hat{b}_m \end{pmatrix} = (\hat{\boldsymbol{A}}\,|\,\hat{\boldsymbol{b}}) \ ,$$

hvor $\hat{a}_{ii} \neq 0$ for $i = 1, 2, \cdots, r$. Processen stopper, når vi er nået ned i nederste række ($r = m$), eller når der er et antal nulrækker tilbage i koefficientmatrixdelen. Så længe denne indeholder bare ét element $\neq 0$, kan man ved række/-søjleombytninger indrangere dette element som "næste" diagonalelement.

I matricen $(\hat{A} \mid \hat{b})$ kommer det eklatant til udtryk, om løsningsmængden er tom eller ej. Hvis der nemlig i den nederste hale af højresider, dvs tallene $\hat{b}_i$ $(i = r+1, \cdots, n)$, findes bare ét, som er forskelligt fra 0, så gælder

$$\rho(\hat{A}) = r < \rho(\hat{A} \mid \hat{b}) = r + 1, \tag{4}$$

og da er der iflg. Sætning 1 ingen løsninger. (Det ses også direkte af, at den tilsvarende ligning i systemet $\hat{A}x = \hat{b}$ er af formen $0 = \hat{b}_i$.) Hvis omvendt alle disse $\hat{b}_i$ er $= 0$, er $\rho(\hat{A}) = \rho(\hat{A} \mid \hat{b}) = r$, dvs da *er* der løsninger. I dette tilfælde kan de $m - r$ nederste rækker i $(\hat{A} \mid \hat{b})$ simpelthen bortkastes, da de svarer til ligninger, som alle er af formen $0 = 0$ og derfor er opfyldt uanset x. Tilbage er et system på r ligninger med n ubekendte, hvis totalmatrix er den øverste $r \times (n+1)$ blok i $(\hat{A} \mid \hat{b})$. Strukturen af matricen $(\hat{A} \mid \hat{b})$ medfører, at der gælder $r \leq n$.

For $m \leq n$ vil man normalt, men naturligvis ikke altid, finde $r = m$, dvs ingen nulrækker forneden. For $m > n$ vil der til gengæld blive frembragt mindst $m - n$ nulrækker i koefficient-delen, og en eller flere af de tilsvarende højresider $\hat{b}_{r+1}, \cdots, \hat{b}_m$ vil normalt, men naturligvis ikke altid, være forskellig fra 0. I så fald kan vi slutte, at løsningsmængden er tom.

Antag, at der er løsninger, dvs eventuelle nulrækker forneden i koefficient-delen af $(\hat{A} \mid \hat{b})$ har også 0 på højre side og kan bortkastes. Vi skelner mellem to tilfælde, $r = n$ og $r < n$.

For $r = n$ kan ligningerne løses *entydigt* nedefra, idet vi af den nederste ligning finder x_n, indsætter dette i den næstnederste og finder x_{n-1}, osv.

For $r < n$ bliver der *uendelig mange løsninger*. Vi kan nemlig vælge $x_{r+1}, \cdots, x_n$ frit og uafhængigt af hinanden og derpå finde $x_1, \cdots, x_r$, sådan som det lige blev beskrevet i tilfældet $r = n$. Man siger, at der er en "$(n - r)$-dobbelt uendelighed" af løsninger. En parameterfremstilling for løsningsmængden kan man få ved at regne videre på totalmatricen for de r ligninger, der er tilbage. Denne matrix omdannes ved nye rækkeoperationer "nedefra" samt nye række-gange-tal operationer, indtil den er af formen

$$\begin{pmatrix} 1 & 0 & \cdots & 0 & -b_{1,r+1} & \cdots & -b_{1n} & c_1 \\ 0 & 1 & \cdots & 0 & -b_{2,r+1} & \cdots & -b_{2n} & c_2 \\ & \vdots & & & & & & \\ 0 & 0 & \cdots & 1 & -b_{r,r+1} & \cdots & -b_{rn} & c_r \end{pmatrix} = (E_r \mid -B \mid c),$$

hvor $\boldsymbol{B}$ er en $r \times (n-r)$ matrix og $\boldsymbol{c} \in \mathbb{R}^r$, mens $\boldsymbol{E}_r$ betegner $\mathcal{M}_r$-enhedsmatricen. At denne omformning er mulig, følger af, at diagonalkoefficienterne $\hat{a}_{ii}$ $(i = 1, \cdots, r)$ alle er $\neq 0$. I øvrigt kan række-gange-tal operationer også inddrages tidligere, hvis det forekommer bekvemt.

Ligningssystemet med $(\boldsymbol{E}_r \mid -\boldsymbol{B} \mid \boldsymbol{c})$ som totalmatrix er $(\boldsymbol{E}_r \mid -\boldsymbol{B})\boldsymbol{x} = \boldsymbol{c}$. Idet vi sætter $\boldsymbol{x} = (\boldsymbol{x}', \boldsymbol{x}'')^{\mathsf{T}}$, hvor

$$\boldsymbol{x}' = (x_1, x_2, \cdots, x_r)^{\mathsf{T}}, \qquad \boldsymbol{x}'' = (x_{r+1}, \cdots, x_n)^{\mathsf{T}},$$

bliver førnævnte system til $\boldsymbol{E}_r\boldsymbol{x}' - \boldsymbol{B}\boldsymbol{x}'' = \boldsymbol{c}$, eller via en lille omskrivning:

$$\boldsymbol{x}' = \boldsymbol{B}\boldsymbol{x}'' + \boldsymbol{c}. \tag{5}$$

Det blev nævnt, at når $n > r$, kan man vælge $x_{r+1}, \cdots, x_n$ frit og derefter løse ligningerne mht $x_1, \cdots, x_r$. Det er netop, hvad ligning (5) udtrykker i vektor-matrix form: Man kan vælge $\boldsymbol{x}''$ som en vilkårlig vektor $\boldsymbol{t} \in \mathbb{R}^{n-r}$ og derefter udregne $\boldsymbol{x}'$ som $\boldsymbol{x}' = \boldsymbol{B}\boldsymbol{t} + \boldsymbol{c}$. Resultatet er en *parameterfremstilling for løsningsmængden.* Opskrevet udførligt ser den således ud:

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_r \\ x_{r+1} \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_{1,r+1}t_1 + \ldots + b_{1n}t_{n-r} + c_1 \\ b_{2,r+1}t_1 + \ldots + b_{2n}t_{n-r} + c_2 \\ \vdots \\ b_{r,r+1}t_1 + \ldots + b_{rn}t_{n-r} + c_r \\ t_1 \\ \ddots \\ t_{n-r} \end{pmatrix}, \tag{6}$$

hvor $t_1, \cdots, t_{n-r} \in \mathbb{R}$. Her er dog forudsat, at det ikke var nødvendigt at ombytte søjler på vejen fra $(\boldsymbol{A} \mid \boldsymbol{b})$ til $(\hat{\boldsymbol{A}} \mid \hat{\boldsymbol{b}})$. Hvis det var det, må de tilsvarende ubekendte x_i og x_j bytte plads i udtrykket (6), jfr. Sætning 2, 4°.

Tre bemærkninger. 1) Den fremgangsmåde i tilfældet $r < n$, der er beskrevet ovenfor, fungerer faktisk også for $r = n$. Da kommer der ingen parametre t_j ind i billedet, vi får blot omskrevet systemet $\hat{\boldsymbol{A}}\boldsymbol{x} = \hat{\boldsymbol{b}}$ til $\boldsymbol{x} = \boldsymbol{c} = (c_1, \ldots, c_n)^{\mathsf{T}}$, som er systemets entydige løsning.

2) Det understreges, at (6) ikke er den eneste måde at opskrive løsningsmængden til (1) på. Der vil tværtimod være talrige muligheder. F.eks. kan man i reglen frit udvælge $n - r$ af de ubekendte og "bruge dem som parametre", dvs

det skal ikke nødvendigvis være netop de $n-r$ sidste ubekendte. Og parametrene behøver ikke engang at være identiske med nogle af de ubekendte. Den beskrevne metode, der fører frem til (6), er blot mest praktisk, når man én gang er nået til mellemstadiet $(\hat{\boldsymbol{A}} \,|\, \hat{\boldsymbol{b}})$.

3) I vektor-matrix koncentreret form kan (6) skrives

$$\boldsymbol{x} = \begin{pmatrix} \boldsymbol{B} \\ \boldsymbol{E}_{n-r} \end{pmatrix} \boldsymbol{t} + \begin{pmatrix} \boldsymbol{c} \\ \boldsymbol{o}_{n-r} \end{pmatrix} \qquad (\boldsymbol{t} \in \mathbb{R}^{n-r}),$$

hvor $\boldsymbol{E}_{n-r}$ og $\boldsymbol{o}_{n-r}$ betegner hhv $\mathcal{M}_{n-r}$-enhedsmatricen og $\mathbb{R}^{n-r}$-nulvektoren. Denne formel kan ses som et elegant modstykke til formlen $(\boldsymbol{E}_r \,|\, -\boldsymbol{B})\boldsymbol{x} = \boldsymbol{c}$. I praksis er den udførlige version (6) dog nok nemmest at bruge.

Resumé af standard løsningsproceduren. Vi afrunder beskrivelsen af standardmetoden til løsning af systemet (1) med at opskrive den i skematisk oversigtsform. De fire matrixomdannelser, der kan komme på tale, er

1° rækkeoperationer i totalmatricen,
2° rækkeombytninger i totalmatricen,
3° række-gange-tal operationer (med en faktor $\neq 0$) i totalmatricen,
4° søjleombytninger i koefficientmatricen, om nødvendigt.

Løsning af systemet $\boldsymbol{Ax} = \boldsymbol{b}$ udføres i følgende skridt:

1. Opskriv totalmatricen $(\boldsymbol{A} \,|\, \boldsymbol{b})$.

2. Omdan den vha operationerne 1°, 2°, 3° samt om nødvendigt 4° til formen $(\hat{\boldsymbol{A}} \,|\, \hat{\boldsymbol{b}})$.

3. Har eventuelle nulrækker i $\hat{\boldsymbol{A}}$ alle 0 også i $\hat{\boldsymbol{b}}$-søjlen? Hvis nej, er man færdig, og facit er: *der er ingen løsninger.* Hvis ja (herunder hvis der ingen nulrækker er), fortsættes.

4. Bortkast eventuelle nulrækker i totalmatricen. Omdan derefter vha 1°, 2° og 3° den resterende $r \times (n+1)$ totalmatrix til formen $\boldsymbol{E}_r \,|\, -\boldsymbol{B} \,|\, \boldsymbol{c})$.

5. Opskriv "næsten-løsningen" (6).

6. Blev der udført søjleombytninger (iv) i starten? Hvis nej, angiver (6) løsningen, og man er færdig. Hvis ja, skal man først foretage de tilsvarende ombytninger (i modsat rækkefølge af før, hvis det drejer sig om flere) af de ubekendte i udtrykket (6). Efter denne ændring angiver (6) løsningen.

EKSEMPEL 2. Løs ligningssystemet

$$\begin{aligned}
x_1 + x_2 \qquad\quad + 2x_4 + x_5 &= 2,\\
2x_1 + 3x_2 + x_3 + 5x_4 + 2x_5 &= 7,\\
-x_1 \qquad + x_3 + x_4 - 2x_5 &= 2,\\
3x_1 + 6x_2 + 3x_3 + 13x_4 + x_5 &= 17.
\end{aligned}$$

Vi opskriver totalmatricen og omdanner den som beskrevet ovenfor. (Eksemplet er konstrueret, så der også bliver brug for en søjleombytning.)

$$\begin{pmatrix} 1 & 1 & 0 & 2 & 1 & 2 \\ 2 & 3 & 1 & 5 & 2 & 7 \\ -1 & 0 & 1 & 1 & -2 & 2 \\ 3 & 6 & 3 & 13 & 1 & 17 \end{pmatrix} \qquad \begin{aligned} &rk.2 \to rk.2 - 2 \cdot rk.1 \\ &rk.3 \to rk.3 + 1 \cdot rk.1 \\ &rk.4 \to rk.4 - 3 \cdot rk.1 \end{aligned}$$

$$\curvearrowright \begin{pmatrix} 1 & 1 & 0 & 2 & 1 & 2 \\ 0 & 1 & 1 & 1 & 0 & 3 \\ 0 & 1 & 1 & 3 & -1 & 4 \\ 0 & 3 & 3 & 7 & -2 & 11 \end{pmatrix} \qquad \begin{aligned} &rk.3 \to rk.3 - 1 \cdot rk.2 \\ &rk.4 \to rk.4 - 3 \cdot rk.2 \end{aligned}$$

$$\curvearrowright \begin{pmatrix} 1 & 1 & 0 & 2 & 1 & 2 \\ 0 & 1 & 1 & 1 & 0 & 3 \\ 0 & 0 & 0 & 2 & -1 & 1 \\ 0 & 0 & 0 & 4 & -2 & 2 \end{pmatrix} \qquad rk.4 \to rk.4 - 2 \cdot rk.3$$

$$\curvearrowright \begin{pmatrix} 1 & 1 & 0 & 2 & 1 & 2 \\ 0 & 1 & 1 & 1 & 0 & 3 \\ 0 & 0 & 0 & 2 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \qquad \begin{aligned} &\text{(der er løsninger !)} \\ &\text{nulrækken fjernes} \end{aligned}$$

$$\curvearrowright \begin{pmatrix} 1 & 1 & 0 & 2 & 1 & 2 \\ 0 & 1 & 1 & 1 & 0 & 3 \\ 0 & 0 & 0 & 2 & -1 & 1 \end{pmatrix} \qquad sø.3 \leftrightarrow sø.5$$

$$\curvearrowright \begin{pmatrix} 1 & 1 & 1 & 2 & 0 & 2 \\ 0 & 1 & 0 & 1 & 1 & 3 \\ 0 & 0 & -1 & 2 & 0 & 1 \end{pmatrix} \qquad rk.1 \to rk.1 + 1 \cdot rk.3$$

$$\curvearrowright \begin{pmatrix} 1 & 1 & 0 & 4 & 0 & 3 \\ 0 & 1 & 0 & 1 & 1 & 3 \\ 0 & 0 & -1 & 2 & 0 & 1 \end{pmatrix} \qquad \begin{aligned} &rk.1 \to rk.1 - 1 \cdot rk.2 \\ &rk.3 \to (-1) \cdot rk.3 \end{aligned}$$

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

Dermed kan løsningsmængden opskrives som

$$\begin{pmatrix} x_1 \\ x_2 \\ x_5 \\ x_4 \\ x_3 \end{pmatrix} = \begin{pmatrix} -3t_1 + t_2 \\ -t_1 - t_2 + 3 \\ 2t_1 \quad -1 \\ t_1 \\ t_2 \end{pmatrix} \qquad (t_1, t_2 \in \mathbb{R}).$$

Bemærk, at der er taget højde for søjleombytningen. Undervejs til løsningen tillod vi os i øvrigt at fjerne en 0-række, *inden* vi var nået til formen $(\hat{\boldsymbol{A}} \,|\, \hat{\boldsymbol{b}})$, hvilket naturligvis ikke spiller nogen rolle.

ØVELSE 2. Vis på lignende måde som i eksempel 2, at ligningssystemet

$$\begin{aligned} x_1 + x_2 \qquad\quad &= 2\,, \\ 2x_1 + 3x_2 + x_3 &= 7\,, \\ -x_1 \qquad + x_3 &= 2\,, \\ 3x_1 + 6x_2 + 3x_3 &= 17 \end{aligned}$$

ikke har nogen løsninger.

ØVELSE 3. Vis, at ligningssystemet

$$\begin{aligned} x_1 \qquad + x_3 &= 2\,, \\ 2x_1 + x_2 + 2x_3 &= 7\,, \\ -x_1 + x_2 - 2x_3 &= 2\,, \\ 3x_1 + 3x_2 + x_3 &= 17 \end{aligned}$$

har den entydige løsning $\boldsymbol{x} = (3, 3, -1)^{\mathsf{T}}$.

ØVELSE 4. Løs ligningssystemet

$$\begin{aligned} x_1 + x_2 + x_3 + x_4 &= 1\,, \\ x_1 + 2x_2 + 3x_3 + 4x_4 &= 5\,, \\ x_1 + 4x_2 + 9x_3 + 16x_4 &= 25\,. \end{aligned}$$

EKSEMPEL 3. Løs for enhver værdi af $a \in \mathbb{R}$ ligningssystemet

$$\begin{pmatrix} 1 & a & 1 \\ 2a & 2 & a+1 \\ a & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3 \\ 2a+4 \\ 3a \end{pmatrix}.$$

Som i Eksempel 2, med regninger gældende for alle $a \in \mathbb{R}$, finder vi:

$$\begin{pmatrix} 1 & a & 1 & 3 \\ 2a & 2 & a+1 & 2a+4 \\ a & 1 & 1 & 3a \end{pmatrix} \qquad \begin{matrix} rk.2 \to rk.2 - 2 \cdot rk.3 \\ rk.2 \leftrightarrow rk.3 \end{matrix}$$

$$\curvearrowright \begin{pmatrix} 1 & a & 1 & 3 \\ a & 1 & 1 & 3a \\ 0 & 0 & a-1 & -4a+4 \end{pmatrix} \qquad rk.2 \to rk.2 - a \cdot rk.1$$

$$\curvearrowright \begin{pmatrix} 1 & a & 1 & 3 \\ 0 & 1-a^2 & 1-a & 0 \\ 0 & 0 & a-1 & -4a+4 \end{pmatrix}.$$

Vi er nu nået til formen $(\hat{A} \,|\, \hat{b})$, dog forudsat at alle tre diagonalelementer er $\neq 0$, hvilket de er for $a \neq \pm 1$. Vi deler undersøgelsen op i tre tilfælde.

$\underline{a = 1}$: Den totalmatrix, vi nåede frem til ovenfor, bliver lig med

$$\begin{pmatrix} 1 & 1 & 1 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Løsningsmængden er ikke tom. Den består af alle $x \in \mathbb{R}^3$, der opfylder

$$x_1 + x_2 + x_3 = 3.$$

Hvis man ikke er tilfreds med denne beskrivelse af løsningsmængden (som er en plan), kan man udlede en parameterfremstilling for den, f.eks. ved at sætte $x_2 = s$, $x_3 = t$, hvorved ligningen giver $x_1 = 3 - s - t$. Det samlede udtryk bliver dermed

$$x = (3 - s - t,\ s,\ t)^{\mathsf{T}} \qquad (s, t \in \mathbb{R}).$$

$\underline{a = -1}$: Totalmatricen bliver

$$\begin{pmatrix} 1 & -1 & 1 & 3 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & -2 & 8 \end{pmatrix}.$$

Her vil en enkelt rækkeoperation, hvor $rk.2$ lægges til $rk.3$, bringe denne på formen $0\ 0\ 0\ 8$. Altså: *ingen løsninger*. Det ses også af, at 2. og 3. ligning i det ændrede system giver hhv $x_3 = 0$ og $x_3 = -4$, der ikke kan gælde samtidig.

$\underline{a \neq \pm 1}$: Da totalmatricen fra før er regulær, er der en entydig løsning. For at finde den udfører vi to række-gange-tal operationer, idet vi dividerer $rk.2$ og $rk.3$ igennem med hhv $1 - a$ og $a - 1$. Derved forenkles totalmatricen til

$$\begin{pmatrix} 1 & a & 1 & 3 \\ 0 & a+1 & 1 & 0 \\ 0 & 0 & 1 & -4 \end{pmatrix}.$$

Af nederste ligning i det tilsvarende system fås $x_3 = -4$. Ved indsættelse heraf i den midterste ligning fås $x_2 = -x_3/(a+1) = 4/(a+1)$. Endelig kan vi indsætte de to fundne ubekendte i den øverste ligning og derved finde den tredje, som bliver $x_1 = 3 - ax_2 - x_3 = 7 - 4a/(a+1)$. Samlet er den entydige løsning

$$\boldsymbol{x} = \left(7 - \frac{4a}{a+1},\ \frac{4}{a+1},\ -4\right)^{\mathsf{T}}.$$

Vi slutter dette afsnit med, i forlængelse af side 126, at opsummere nogle af resultaterne om ligningssystemet $\boldsymbol{Ax} = \boldsymbol{b}$.

Løsningsmængden til systemet afhænger først og fremmest af forholdet mellem antal ligninger m og antal ubekendte n. Desuden afhænger den af de aktuelle koefficienter a_{ij} og højresider b_i.

$\underline{m > n}$: Normalt *ingen* løsninger, da man må forvente, at $\rho(\boldsymbol{A}) = n$ og $\rho(\boldsymbol{A}\,|\,\boldsymbol{b}) = n + 1$.

$\underline{m = n}$: Normalt *netop én* løsning, da man må forvente, at $\boldsymbol{A}$ er regulær, og i så fald er $\boldsymbol{Ax} = \boldsymbol{b}$ ensbetydende med $\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b}$.

$\underline{m < n}$: Normalt en "$(n-m)$*-dobbelt uendelighed* af løsninger, da man må forvente, at $\rho(\boldsymbol{A}) = \rho(\boldsymbol{A}\,|\,\boldsymbol{b}) = m$. Ligningerne kan da løses med hensyn til m af de variable, der derved udtrykkes lineært ved de øvrige $n - m$ variable. Disse kan vælges frit som uafhængige "parametre".

Det skal endnu en gang fremhæves, at udtryk som "normalt", "må forvente" o.lign. dækker over mangfoldige undtagelser. Det gælder også ved diverse anvendelser. Sagen er, at når man opskriver et system med *tilfældigt valgte koefficienter*, er det mest sandsynligt, at man ender i en af de opridsede normalsituationer. Men når tallene foreligger på forhånd, har man ingen garanti for, at de er "tilfældige". Heller ikke, når de stammer fra et anvendelsesområde og måske endda er "grimme tal". Typisk kan systemet indeholde en eller flere koefficienter, der bør behandles som arbitrære konstanter (se Eksempel 3 ovenfor), og for specielle værdier af disse koefficienter kan det da ske, at f.eks. $\rho(\boldsymbol{A})$ bliver mindre end "normalt". Derfor er det af værdi at udvikle en generel løsningsprocedure, sådan som vi har gjort.

2. Homogent og inhomogent ligningssystem.

Vi har tidligere brugt gloserne 'homogen' og 'inhomogen' om en lineær ligning, hvis højreside er hhv $=0$ og $\neq 0$. Mere alment opstilles

> DEFINITION 2. **Homogent og inhomogent ligningssystem.**
> For $\boldsymbol{b} \neq \boldsymbol{o}$ kaldes systemet $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ *inhomogent*, mens systemet $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{o}$ kaldes *homogent*. For et givet inhomogent system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ taler man om det *tilhørende* homogene system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{o}$.

Når løsningen til et inhomogent system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ er fundet på formen (6), side 131, kan man nemt opstille et tilsvarende udtryk for løsningen til det tilhørende homogene system: Hvis blot konstantleddene c_i i de r øverste ligninger erstattes med 0, har man løsningen til $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{o}$. Det indses således: Når standardmetoden anvendes på det homogene system, vil de samme regninger, der fører fra $(\boldsymbol{A} \,|\, \boldsymbol{b})$ til $(\hat{\boldsymbol{A}} \,|\, \hat{\boldsymbol{b}})$ side 129 og videre til $(\boldsymbol{E}_r \,|\, -\boldsymbol{B} \,|\, \boldsymbol{c})$ side 130, føre fra $(\boldsymbol{A} \,|\, \boldsymbol{o})$ til $(\hat{\boldsymbol{A}} \,|\, \boldsymbol{o})$ og videre til $(\boldsymbol{E}_r \,|\, -\boldsymbol{B} \,|\, \boldsymbol{o})$. Nullerne på ligningernes højresider følger så at sige med hele vejen.

Af (6), side 131 med alle $c_i = 0$ fremgår endvidere, at løsningsmængden til det homogene system er et *underrum* V af $\mathbb{R}^n$, udspændt af vektorerne

$$\begin{gathered}(b_{1,r+1}, \cdots, b_{r,r+1}, 1, 0, \cdots, 0)^{\mathsf{T}}, \\ (b_{1,r+2}, \cdots, b_{r,r+2}, 0, 1, \cdots, 0)^{\mathsf{T}}, \\ \vdots \\ (b_{1n}, \cdots, b_{rn}, 0, 0, \cdots, 1)^{\mathsf{T}}.\end{gathered}$$

Strukturen af de sidste $n-r$ koordinater i disse vektorer viser, at de er lineært uafhængige og dermed udgør en *basis* for V. Heraf følger også, at $\dim(V) = n - r = n - \rho(\boldsymbol{A})$. Specielt har en *hyperplan gennem nulpunktet*, defineret som løsning til én homogen lineær ligning, dimensionen $n-1$, som nævnt uden bevis side 44. Samme sted viste vi vha *linearitetsbetingelserne* side 31, at en hyperplan gennem nulpunktet er et underrum af $\mathbb{R}^n$. Beviset kan ved små ændringer overføres på løsningsmængden til $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{o}$, men så får vi ikke resultatet om underrummets dimension med.

Vi samler resultaterne i en sætning, der også rummer en vigtig karakterisering af sammenhængen mellem løsningen til et lineært ligningssystem og det tilhørende homogene system.

SÆTNING 3. (1) Løsningsmængden til det homogene system $\boldsymbol{Ax} = \boldsymbol{o}$ er et underrum af $\mathbb{R}^n$ med dimension $n - r$, hvor r er rangen af $\boldsymbol{A}$.

(2) Når man til en enkelt løsning til systemet $\boldsymbol{Ax} = \boldsymbol{b}$ adderer samtlige løsninger til det tilhørende homogene system $\boldsymbol{Ax} = \boldsymbol{o}$, får man samtlige løsninger til det inhomogene system.

BEVIS: (1) er vist ovenfor. For at vise (2) går vi frem som i Eksempel 16, side 44-45 i specialtilfældet $m = 1$. Lad L og V være løsningsmængderne til hhv $\boldsymbol{Ax} = \boldsymbol{b}$ og $\boldsymbol{Ax} = \boldsymbol{o}$, og antag $\boldsymbol{x}_0 \in L$, dvs der gælder $\boldsymbol{Ax}_0 = \boldsymbol{b}$. Vi skal vise

$$L = \{\boldsymbol{x}' + \boldsymbol{x}_0 \,|\, \boldsymbol{x}' \in V\}.$$

Antag først $\boldsymbol{x} = \boldsymbol{x}' + \boldsymbol{x}_0$, hvor $\boldsymbol{x}' \in V$. Da er

$$\boldsymbol{Ax} = \boldsymbol{Ax}' + \boldsymbol{Ax}_0 = \boldsymbol{o} + \boldsymbol{b} = \boldsymbol{b},$$

som viser, at $\boldsymbol{x} \in L$. Antag dernæst $\boldsymbol{x} \in L$, dvs $\boldsymbol{Ax} = \boldsymbol{b}$. Vi kan skrive $\boldsymbol{x}$ på formen $(\boldsymbol{x} - \boldsymbol{x}_0) + \boldsymbol{x}_0$, hvor første led tilhører V, da

$$\boldsymbol{A}(\boldsymbol{x} - \boldsymbol{x}_0) = \boldsymbol{Ax} - \boldsymbol{Ax}_0 = \boldsymbol{b} - \boldsymbol{b} = \boldsymbol{o}.$$

Dermed er $\boldsymbol{x}$ af formen $\boldsymbol{x}' + \boldsymbol{x}_0$, hvor $\boldsymbol{x}' \in V$, og beviset er færdigt. □

Udtrykket (6), side 131 er af formen $\boldsymbol{x}' + \mathrm{c}_0$, hvor $\mathrm{c}_0 = (c_1, \cdots, c_r, 0, \cdots, 0)^{\mathsf{T}}$ er en løsning til det inhomogene system, idet den fremkommer for $\boldsymbol{t} = \boldsymbol{o}$. Men (6) er ikke i sig selv et *bevis* for (2) i Sætning 3, da c_0 blot er én af de (normalt) uendelig mange muligheder for at vælge $\boldsymbol{x}_0$.

EKSEMPEL 4. Løs ligningssystemet

$$\begin{aligned}
x_1 + x_2 \phantom{{}+x_3} + 2x_4 + x_5 &= 0,\\
2x_1 + 3x_2 + x_3 + 5x_4 + 2x_5 &= 0,\\
-x_1 \phantom{{}+x_2} + x_3 + x_4 - 2x_5 &= 0,\\
3x_1 + 6x_2 + 3x_3 + 13x_4 + x_5 &= 0.
\end{aligned}$$

Der er tale om det homogene system hørende til systemet i Eksempel 2, side 133-134. Iflg. udtrykket øverst side 134 kan vi straks opskrive det tilsvarende løsningsudtryk for systemet ovenfor, som efter tilbagebytning af x_3 og x_5 ser således ud:

$$x = \begin{pmatrix} -3t_1 + t_2 \\ -t_1 - t_2 \\ t_2 \\ t_1 \\ 2t_1 \end{pmatrix} \qquad (t_1, t_2 \in \mathbb{R}).$$

Løsningsmængden er et todimensionalt $\mathbb{R}^5$-underrum udspændt af $(-3,-1,0,1,2)^\mathsf{T}$ og $(1,-1,1,0,0)^\mathsf{T}$.

ØVELSE 5. Betragt den lineære afbildning $f : \mathbb{R}^5 \curvearrowright \mathbb{R}^3$ givet ved $f(x) = Ax$, hvor

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

(1) Find mængden af $x \in \mathbb{R}^5$, der opfylder $f(x) = o$.

(2) Find mængden af $x \in \mathbb{R}^5$, der opfylder $f(x) = f(x_0)$, hvor vektoren x_0 er opgivet som $(7,3,-5,-2,4)^\mathsf{T}$. [Advarsel: Pas på, at du ikke kommer til at udføre overflødigt arbejde! Der er stort set ikke brug for nye regninger i (2).]

EKSEMPEL 5. **Kerne.** I videregående lineær algebra har man ofte brug for underrummet af løsninger til det homogene system $Ax = o$. Man kalder dette underrum for *kernen* af A, $\ker(A)$. Når A opfattes som matrix for en lineær afbildning $f : \mathbb{R}^n \curvearrowright \mathbb{R}^m$, taler man sommetider i stedet om kernen af f.

Nogle af de resultater, vi har fundet i det foregående, kan dermed gives følgende formulering: Når $m \geq n$, er $\ker(A)$ normalt det trivielle underrum $\{o\}$, men er der lineær afhængighed mellem søjlerne i A, er kernen ikke-triviel. Når $m < n$, er dimensionen af kernen mindst $n - m$. Kun for $A = O$ er kernen "trivielt stor", dvs lig med hele $\mathbb{R}^n$. - Alment gælder $\dim(\ker(A)) = n - \rho(A)$.

Vi vil nu indføre en matrix $\mathbf{E}$, der opfører sig særlig simpelt mht. multiplikation, og som derfor vil vise sig at være nyttig.

Definition B.2.2 (Enhedsmatricen) Matricen

$$\mathbf{E} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

kaldes *enhedsmatricen.*

Den følgende sætning viser, at en matrix ikke ændrer sig, når den bliver ganget med enhedsmatricen.

Sætning B.2.1 (Multiplikation med enhedsmatricen) For enhver matrix $\mathbf{M}$ gælder der

$$\mathbf{ME} = \mathbf{EM} = \mathbf{M}.$$

Bevis for Sætning B.2.1. Med

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

har vi

$$\mathbf{ME} = \begin{pmatrix} a & c \\ b & d \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a\cdot 1 + c\cdot 0 & a\cdot 0 + c\cdot 1 \\ b\cdot 1 + d\cdot 0 & b\cdot 0 + d\cdot 1 \end{pmatrix} = \begin{pmatrix} a & c \\ b & d \end{pmatrix} = \mathbf{M}$$

og tilsvarende

$$\mathbf{EM} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} a & c \\ b & d \end{pmatrix} = \begin{pmatrix} 1\cdot a + 0\cdot b & 1\cdot c + 0\cdot d \\ 0\cdot a + 1\cdot b & 0\cdot c + 1\cdot d \end{pmatrix} = \begin{pmatrix} a & c \\ b & d \end{pmatrix} = \mathbf{M}.$$

□

Repetition 2x2-matricer!

B.3 Invers matrix og determinant

En ligning af formen $ax = b$ (hvor a og b er tal og hvor $a \neq 0$) løses ved at dividere med a på begge sider af lighedstegnet, hvilket giver $x = \frac{b}{a}$. Sagt på en anden måde kan man gange med a^{-1} på begge sider af lighedstegnet og få $x = a^{-1}b$. I mange sammenhænge (se f.eks. Anvendelseseksempel B.1) har man brug for at løse tilsvarende ligninger

$$\mathbf{Mv} = \mathbf{p},$$

LA 140

hvor $\mathbf{M}$ er en given matrix og $\mathbf{p}$ en given vektor. I lighed med løsningen af ligningen $ax = b$ kunne man gange på begge sider med $\mathbf{M}^{-1}$, og derved opnå

$$\mathbf{v} = \mathbf{M}^{-1}\mathbf{p}.$$

Spørgsmålet er så blot, hvad vi mener med den inverse matrix "$\mathbf{M}^{-1}$".

For $a \neq 0$ kan tallet $a^{-1} = \frac{1}{a}$ defineres som løsningen x til ligningen $ax = 1$. Tilsvarende kan man anskue den inverse matrix $\mathbf{M}^{-1}$ som en matrix $\mathbf{X}$, der opfylder begge ligningerne $\mathbf{MX} = \mathbf{E}$ og $\mathbf{XM} = \mathbf{E}$ (husk at faktorernes rækkefølge ikke er underordnet), hvor $\mathbf{E}$ er enhedsmatricen (se Definition B.2.2). Dette gør vi nu til den formelle definition af den inverse matrix.

Definition B.3.1 (Invers matrix) Matricen $\mathbf{M}$ har en invers matrix, hvis der findes en matrix $\mathbf{X}$ sådan at

$$\mathbf{MX} = \mathbf{XM} = \mathbf{E}.$$

I så fald kaldes $\mathbf{X}$ den *inverse matrix* til $\mathbf{M}$ og betegnes $\mathbf{M}^{-1}$.

Bemærkning B.3.1 Definition B.3.1 siger indirekte, at hvis $\mathbf{M}$ har en invers matrix, så er der kun én invers matrix. Dette kan ses på følgende måde: Hvis $\mathbf{X}_1$ og $\mathbf{X}_2$ begge er inverse matricer til $\mathbf{M}$, har vi

$$\mathbf{X}_1 = \mathbf{X}_1\mathbf{E} = \mathbf{X}_1(\mathbf{MX}_2) = (\mathbf{X}_1\mathbf{M})\mathbf{X}_2 = \mathbf{EX}_2 = \mathbf{X}_2,$$

dvs. $\mathbf{X}_1 = \mathbf{X}_2$.

Som definitionen antyder, er det ikke alle matricer, der har en invers matrix. Vi ønsker nu at besvare følgende spørgsmål vedrørende "invers matrix".

- Hvornår har en matrix en invers matrix?
- Hvordan bestemmer man den inverse matrix?
- Hvordan kan vi bruge den inverse matrix til at løse ligninger med?

Det viser sig, at den såkaldte determinant af en matrix spiller en vigtig rolle i forbindelse med den inverse matrix.

Definition B.3.2 (Determinant) For en matrix

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

defineres *determinanten*, det M, ved udtrykket

$$\det M = ad - bc.$$

Nogle gange angiver man også determinanten ved at ændre de store parenteser om matricen til store numerisk-tegn, dvs.

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

Bemærkning B.3.2 I nogle sammenhænge (f.eks. i andre lærebøger) kalder man en matrix M *invertibel*, hvis den har en invers matrix. Tilsvarende siges M at være *regulær*, hvis $\det M \neq 0$ og at være *singulær*, hvis $\det M = 0$.

Eksempel B.3.1 For matricen

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

har vi

$$\det M = \begin{vmatrix} 4 & 2 \\ -3 & 1 \end{vmatrix} = 4 \cdot 1 - (-3) \cdot 2 = 10.$$

Bemærk at matricen

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

skrives med store parenteser, mens determinanten

$$\det M = \begin{vmatrix} 4 & 2 \\ -3 & 1 \end{vmatrix},$$

som jo er et tal, skrives med store numerisk-tegn. Der er altså stor forskel på at benytte store parenteser og store numerisk-tegn.

LA142

Før vi går videre med den inverse matrix, viser vi nogle regneregler for determinanter.

Sætning B.3.1 (Regneregler for determinanter) Lad $\mathbf{M}$ og $\mathbf{N}$ være to matricer.

(a) "Determinanten af produktet er produktet af determinanterne", dvs.

$$\det(\mathbf{NM}) = \det \mathbf{N} \cdot \det \mathbf{M}.$$

(b) Hvis $\mathbf{M}$ har en invers matrix, så er $\det \mathbf{M} \neq 0$, og der gælder

$$\det(\mathbf{M}^{-1}) = \frac{1}{\det \mathbf{M}}.$$

(c) For et tal $k \in \mathbb{R}$ gælder der

$$\det(k\mathbf{M}) = k^2 \cdot \det \mathbf{M}.$$

Bevis for Sætning B.3.1.

(a) Med

$$\mathbf{M} = \begin{pmatrix} a_1 & c_1 \\ b_1 & d_1 \end{pmatrix} \quad \text{og} \quad \mathbf{N} = \begin{pmatrix} a_2 & c_2 \\ b_2 & d_2 \end{pmatrix}$$

har vi ifølge Definition B.2.1(c)

$$\mathbf{NM} = \begin{pmatrix} a_2a_1 + c_2b_1 & a_2c_1 + c_2d_1 \\ b_2a_1 + d_2b_1 & b_2c_1 + d_2d_1 \end{pmatrix}$$

og dermed

$$\begin{aligned} \det(\mathbf{NM}) &= \begin{vmatrix} a_2a_1 + c_2b_1 & a_2c_1 + c_2d_1 \\ b_2a_1 + d_2b_1 & b_2c_1 + d_2d_1 \end{vmatrix} \\ &= (a_2a_1 + c_2b_1)(b_2c_1 + d_2d_1) - (b_2a_1 + d_2b_1)(a_2c_1 + c_2d_1) \\ &= a_2a_1b_2c_1 + a_2a_1d_2d_1 + c_2b_1b_2c_1 + c_2b_1d_2d_1 \\ &\quad - b_2a_1a_2c_1 - b_2a_1c_2d_1 - d_2b_1a_2c_1 - d_2b_1c_2d_1 \\ &= a_2a_1d_2d_1 + c_2b_1b_2c_1 - b_2a_1c_2d_1 - d_2b_1a_2c_1. \end{aligned}$$

På den anden side er

$$\begin{aligned} \det \mathbf{N} \cdot \det \mathbf{M} &= (a_1d_1 - b_1c_1)(a_2d_2 - b_2c_2) \\ &= a_1d_1a_2d_2 - a_1d_1b_2c_2 - b_1c_1a_2d_2 + b_1c_1b_2c_2, \end{aligned}$$

hvoraf det ses, at der gælder $\det(\mathbf{NM}) = \det \mathbf{N} \cdot \det \mathbf{M}$.

(b) Når (a) anvendes på formlen $\mathbf{MM}^{-1} = \mathbf{E}$, fås

$$\det \mathbf{M} \cdot \det(\mathbf{M}^{-1}) = \det \mathbf{E} = 1,$$

hvoraf (b) følger.

LA 143

(c) Med

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

har vi

$$\det(k\mathbf{M}) = \begin{vmatrix} ka & kc \\ kb & kd \end{vmatrix} = ka \cdot kd - kb \cdot kc = k^2(ad - bc) = k^2 \cdot \det \mathbf{M}.$$

□

I Bemærkning B.6.2 vil vi give et mere intuitivt argument for Sætning B.3.1(a).

Eksempel B.3.2 For

$$\mathbf{M} = \begin{pmatrix} 4 & 2 \\ -3 & 1 \end{pmatrix} \quad \text{og} \quad \mathbf{N} = \begin{pmatrix} 2 & -1 \\ 0 & 6 \end{pmatrix}$$

har vi $\det \mathbf{M} = 4 \cdot 1 - (-3) \cdot 2 = 10$ (som i Eksempel B.3.1) og $\det \mathbf{N} = 2 \cdot 6 - 0 \cdot (-1) = 12$. Endvidere udregnede vi i Eksempel B.2.1

$$\mathbf{NM} = \begin{pmatrix} 11 & 3 \\ -18 & 6 \end{pmatrix},$$

så $\det \mathbf{NM} = 11 \cdot 6 - (-18) \cdot 3 = 120 = 12 \cdot 10$. Dermed har vi eftervist formlen $\det(\mathbf{NM}) = \det \mathbf{N} \cdot \det \mathbf{M}$ i dette eksempel.

Bemærkning B.3.3

(a) Der findes *ingen* simple regler for $\det(\mathbf{M} + \mathbf{N})$ og $\det(\mathbf{M} - \mathbf{N})$.

(b) Idet $\det \mathbf{N}$ og $\det \mathbf{M}$ er tal, følger det af Sætning B.3.1(a), at

$$\det(\mathbf{NM}) = \det \mathbf{N} \cdot \det \mathbf{M} = \det \mathbf{M} \cdot \det \mathbf{N} = \det(\mathbf{MN}).$$

Selvom de to matricer $\mathbf{NM}$ og $\mathbf{MN}$ normalt er forskellige (se bl.a. Eksempel B.2.1), har de altså den samme determinant.

Den følgende sætning viser, hvordan man afgør, om en given matrix har en invers matrix, og hvordan man i så fald bestemmer den inverse matrix.

Sætning B.3.2 (Bestemmelse af den inverse matrix) Matricen

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

har en invers matrix, netop hvis

$$\det \mathbf{M} = ad - bc \neq 0.$$

I så fald er den inverse matrix givet ved formlen

$$\mathbf{M}^{-1} = \frac{1}{\det \mathbf{M}} \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}$$

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

Det er forholdsvist let at huske formlen for $\mathbf{M}^{-1}$ for en 2×2 matrix $\mathbf{M}$:

- byt om på de to diagonalelementer, a og d,
- skift fortegn på de to andre elementer, b og c,
- dividér med determinanten på alle pladser.

Bevis for Sætning B.3.2. Vi har allerede i Sætning B.3.1 set, at hvis matricen $\mathbf{M}$ har en invers matrix, så må der gælde $\det \mathbf{M} \neq 0$. Derfor skal vi vise følgende: hvis $\det \mathbf{M} \neq 0$, da har $\mathbf{M}$ en invers matrix og den inverse er givet ved formlen i Sætning B.3.2. Vi lader derfor

$$\mathbf{X} = \frac{1}{\det \mathbf{M}} \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}.$$

Ved udregning finder vi dermed

$$\begin{aligned} \mathbf{XM} &= \frac{1}{\det \mathbf{M}} \begin{pmatrix} d & -c \\ -b & a \end{pmatrix} \begin{pmatrix} a & c \\ b & d \end{pmatrix} \\ &= \frac{1}{\det \mathbf{M}} \begin{pmatrix} da - cb & dc - cd \\ -ba + ab & -bc + ad \end{pmatrix} \\ &= \frac{1}{\det \mathbf{M}} \begin{pmatrix} \det \mathbf{M} & 0 \\ 0 & \det \mathbf{M} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathbf{E}. \end{aligned}$$

På tilsvarende vis får vi $\mathbf{MX} = \mathbf{E}$. Dette viser, at $\mathbf{M}^{-1}$ har den ønskede form. □

LA 145

I Bemærkning B.3.5 vil vi komme ind på, hvordan man er kommet frem til udtrykket for $\mathbf{M}^{-1}$ fra Sætning B.3.2.

Eksempel B.3.3 For matricen

$$\mathbf{M} = \begin{pmatrix} 4 & 2 \\ -3 & 1 \end{pmatrix}$$

fandt vi i Eksempel B.3.1, at $\det \mathbf{M} = 10 \neq 0$, så $\mathbf{M}$ har en invers matrix givet ved

$$\mathbf{M}^{-1} = \frac{1}{10}\begin{pmatrix} 1 & -2 \\ -(-3) & 4 \end{pmatrix} = \begin{pmatrix} 0.1 & -0.2 \\ 0.3 & 0.4 \end{pmatrix}.$$

Som kontrol kan vi ved udregning eftervise, at der faktisk gælder

$$\begin{aligned} \mathbf{M}\mathbf{M}^{-1} &= \begin{pmatrix} 4 & 2 \\ -3 & 1 \end{pmatrix}\begin{pmatrix} 0.1 & -0.2 \\ 0.3 & 0.4 \end{pmatrix} \\ &= \begin{pmatrix} 4\cdot 0.1 + 2\cdot 0.3 & 4\cdot(-0.2) + 2\cdot 0.4 \\ -3\cdot 0.1 + 1\cdot 0.3 & -3\cdot(-0.2) + 1\cdot 0.4 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathbf{E} \end{aligned}$$

og på tilsvarende vis $\mathbf{M}^{-1}\mathbf{M} = \mathbf{E}$.

For tal $a \neq 0$ og $b \neq 0$ gælder der $\frac{1}{ab} = \frac{1}{a} \cdot \frac{1}{b}$, dvs. $(ab)^{-1} = a^{-1}b^{-1}$. Vi vil nu se, at det samme gælder for matricer dog med en væsentlig ændring.

Sætning B.3.3 (Den inverse af et produkt) Når matricerne $\mathbf{M}$ og $\mathbf{N}$ begge har en invers matrix, så har produktet $\mathbf{NM}$ ligeledes en invers matrix givet ved

$$(\mathbf{NM})^{-1} = \mathbf{M}^{-1}\mathbf{N}^{-1}.$$

Bemærkning B.3.4 Bemærk at rækkefølgen af $\mathbf{M}$ og $\mathbf{N}$ bliver byttet om, når man tager invers matrix. En måde (for nogle) at huske dette på er, at man først tager sokker *på* og derefter sko *på*, men at man omvendt først tager sko *af* og derefter sokker *af*. Rækkefølgen af sko og sokker bliver altså byttet om, når man tager dem af i stedet for på.

Bevis for Sætning B.3.3. Lad $\mathbf{X} = \mathbf{M}^{-1}\mathbf{N}^{-1}$. Så har vi

$$(\mathbf{NM})\mathbf{X} = (\mathbf{NM})(\mathbf{M}^{-1}\mathbf{N}^{-1}) = \mathbf{N}(\mathbf{M}\mathbf{M}^{-1})\mathbf{N}^{-1} = \mathbf{N}\mathbf{E}\mathbf{N}^{-1} = \mathbf{N}\mathbf{N}^{-1} = \mathbf{E}.$$

På tilsvarende vis får man $\mathbf{X}(\mathbf{NM}) = \mathbf{E}$, hvoraf vi slutter, at $\mathbf{X} = \mathbf{M}^{-1}\mathbf{N}^{-1}$ faktisk er den inverse matrix til $\mathbf{NM}$. Der gælder altså $(\mathbf{NM})^{-1} = \mathbf{M}^{-1}\mathbf{N}^{-1}$. □

LA 146

Som vi var inde på i begyndelsen af afsnittet, kan matrixinvertering benyttes til at løse ligninger. Mere præcist har vi følgende sætning.

Sætning B.3.4 (Løsning af ligninger ved matrixinvertering) Vi betragter et ligningssystem

$$\mathbf{M}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r \\ s \end{pmatrix},$$

hvor $\mathbf{M}$ er en given matrix og $\binom{r}{s}$ en given vektor. Hvis $\mathbf{M}$ har en invers matrix, så er der netop én løsning $\binom{x}{y}$ til ligningssystemet, nemlig

$$\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{M}^{-1}\begin{pmatrix} r \\ s \end{pmatrix}.$$

Med vektornotation kan dette formuleres ved, at ligningssystemet $\mathbf{Mv} = \mathbf{p}$ netop har løsningen

$$\mathbf{v} = \mathbf{M}^{-1}\mathbf{p},$$

når $\mathbf{M}$ har en invers matrix.

Bevis for Sætning B.3.4. Når $\mathbf{M}$ har en invers matrix, så kan vi gange med $\mathbf{M}^{-1}$ på begge sider af lighedstegnet i ligningen $\mathbf{Mv} = \mathbf{p}$, hvilket giver

$$\mathbf{Mv} = \mathbf{p} \quad \Leftrightarrow \quad \mathbf{M}^{-1}\mathbf{Mv} = \mathbf{M}^{-1}\mathbf{p} \quad \Leftrightarrow \quad \mathbf{v} = \mathbf{M}^{-1}\mathbf{p},$$

som ønsket. □

Eksempel B.3.4 I Eksempel B.1.1 løste vi ligningerne

$$\begin{cases} 4x + 2y = 3 \\ -3x + y = -2 \end{cases}$$

vha. lige store koefficienters metode. Ligningerne kan også skrives

$$\begin{pmatrix} 4 & 2 \\ -3 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3 \\ -2 \end{pmatrix}.$$

Ved matrixinvertering (se Eksempel B.3.3) finder vi

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 4 & 2 \\ -3 & 1 \end{pmatrix}^{-1}\begin{pmatrix} 3 \\ -2 \end{pmatrix} = \begin{pmatrix} 0.1 & -0.2 \\ 0.3 & 0.4 \end{pmatrix}\begin{pmatrix} 3 \\ -2 \end{pmatrix} = \begin{pmatrix} 0.7 \\ 0.1 \end{pmatrix},$$

som vi også fandt i Eksempel B.1.1.

I Eksempel B.1.1 krævede det nogle flere mellemregninger at løse ligningerne

$$\begin{cases} 4.3\,x + 2.1\,y = 3 \\ -3.2\,x + 1.1\,y = -2 \end{cases}$$

LA 147

Determinant af en n x n - martrix

- Determinant kendes for 2 x 2 og 3 x 3 (Se MD131 og MD164)
- Determinanten `det(D)` af en trekants-matrix `D` er lig produktet af diagonalelementerne:
  $det(D)=d_1*d_2* \dots *d_n$
- Determinanten ændres ikke ved rækkeoperationer
- Determinanten skifter fortegn ved række (eller søjle) ombytninger
- $det(A) = det(A^T)$
- $det(AB) = det(A)det(B)$

7

Lad `A` være en `n x n` - matrix:

`A` kaldes *regulær*, såfremt `det(A)` er forskellig fra `0`.

Sætning:

`det(A)=0` hvis og kun hvis `rang(A) < n`.

Beregning af determinant:

- Bring `A` på trekant-form `T` vha. rækkeoperationer og række- og søjleombytninger:
  $det(A) = (-1)^{(\#ombytninger)} det(T)$
- Udvikling efter søjle

8

LA 148

Determinant beregnet ved udvikling efter sidste søjle

Lad `A` være en `n x n` matrix og `A[k,l]` et element i matricen. Definer *komplementet* `K[k,l]` som $(-1)^{(k+l)}\det(A_{kl})$, hvor A_{kl} er `(n-1)x(n-1)` matricen, der fåes fra `A` ved at fjerne `k`'te række og `l`'te søjle. Der gælder:

`det(A) = Sum(k=1:n,A[k,n]*K[k,n])`

Eksempel:

```
> B  [,1] [,2] [,3] [,4] [,5]
0år   0.8  0.9  0.7  0.5  0.4
1år   0.7  0.0  0.0  0.0  0.0
2år   0.0  0.8  0.0  0.0  0.0
3år   0.0  0.0  0.6  0.0  0.0
4år   0.0  0.0  0.0  0.4  0.0
```

```
> B15 [,1] [,2] [,3] [,4]
[1,]   0.7  0.0  0.0  0.0
[2,]   0.0  0.8  0.0  0.0
[3,]   0.0  0.0  0.6  0.0
[4,]   0.0  0.0  0.0  0.4
```

```
> det(B)
[1] 0.05376

> B[1,5]*(-1)^(1+5)*0.7*0.8*0.6*0.4
[1] 0.05376
```

9

Invers til `n x n` - matrix

Hvis `A` er regulær (`det(A)` ikke `0`), så har `A` en invers matrix A^{-1}, dvs:

$AA^{-1} = A^{-1}A = E$, og

Ligningen `Ax=b` har den éntydige løsning: $x = A^{-1}b$

Beregning af A^{-1}:

- `2 x 2` se MD132
- `n > 2` brug R , lommeregner eller dominometoden

10

LA 149

Dominometoden:

- Dan `n x 2n` - matricen `[A,E]`
- Foretag rækkeoperationer i `[A,E]` til `A` er blevet en diagonalmatrix
- Divider hver række med diagonal-elementet
- Den fremkomne matrix er $[E,A^{-1}]$

Eksempel fra LAOpg98:

```
> AE [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8]
[1,]    1    1    0    0    1    0    0    0
[2,]    1    1    1    0    0    1    0    0
[3,]    0    1    1    1    0    0    1    0
[4,]    0    0    1    1    0    0    0    1
```

Bliver ved rækkeoperationer til:

```
> EAinv[,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8]
[1,]    1    0    0    0    1    0   -1    1
[2,]    0    1    0    0    0    0    1   -1
[3,]    0    0    1    0   -1    1    0    0
[4,]    0    0    0    1    1   -1    0    1
```

Prøv at gøre det?

Det rækkeoperationerne gør her er at de tilsammen svarer til at gange med A^{-1}:

$$A^{-1}(A,E) = (A^{-1}A, A^{-1}E) = (E, A^{-1})$$

LA 150

KAPITEL 4

SPEKTRALTEORI.

0. Indledning.

I dette kapitel vil vi først udlede sammenhængen mellem vektorers opløsninger efter forskellige baser i samme vektorrum. Derved generaliseres begrebet *koordinat* til rum af højere dimension .

vi opbygge et n-dimensionalt modstykke til en 'stiv drejning' af koordinatsystemet, dvs koordinatskifte mellem to sædvanlige retvinklede koordinatsystemer i $\mathbb{R}^2$ og $\mathbb{R}^3$. Det leder til indførelse af begrebet *ortogonal matrix.*

Kapitlets hovedemne er følgende problem: Givet en lineær afbildning f : $\mathbb{R}^n \curvearrowright \mathbb{R}^n$, hvilke vektorer $x \in \mathbb{R}^n$ afbildes da ved f i et multiplum af sig selv, dvs hvilke vektorer opfylder $f(x) = \lambda x$, hvor λ er et tal? At finde sådanne *egenvektorer* x og tilhørende *egenværdier* λ har vist sig at være af betydning også i de forskellige anvendelser af matrixalgebraen. Vi skal give en indføring i *spektralteori*, der undersøger dette spørgsmål. Herved spiller *symmetriske matricer* en særlig rolle, ikke mindst fordi de knytter sig til en vigtig type ikke-lineære funktioner af n variable, såkaldte *kvadratiske former*, som bruges flittigt inden for økonomisk teori.

Kapitlet slutter med en omtale af spektralegenskaber for *ikke-negative matricer.* De har interesse set fra et anvendelsessynspunkt, især fordi de optræder i mange af den type diskrete matematiske modeller, vi gav et par eksempler på allerede i slutningen af Kapitel 1.

1. Basisskift.

Hvad vi nedenfor udleder om basisskift, skal i det store og hele kun bruges for talrummene $\mathbb{R}^n$.

Vi betragter et $\mathbb{R}^n$ som et *vektorrum* V. V har en *basis* $E = (E_1, E_2, \cdots, E_n)$, således at en vilkårlig vektor $X \in V$ kan skrives entydigt på formen

$$X = x_1 E_1 + x_2 E_2 + \cdots + x_n E_n \,. \tag{1}$$

> DEFINITION 1. **Koordinatsæt mht basis.** Sættet $x = (x_1, \cdots, x_n)^\top \in \mathbb{R}^n$ fastlagt ved (1) kaldes X's *koordinatsæt mht basen* E.

Som nævnt vil regning i V med vektorerne X betyde tilsvarende regninger i $\mathbb{R}^n$ med koordinatsættene x. I øvrigt: Når talsættene x nedenfor indgår i matrixregneudtryk, skal de som sædvanlig opfattes på søjleform, men hvor der kun er tale om $\mathbb{R}^n$-vektorregning, vil vi ikke tage det så nøje med, om vi skriver f.eks. $(3,5)$, $(3,5)^\top$ eller $\binom{3}{5}$, dvs vil vi udelade "$\top$", når det ikke ser ud til at kunne give anledning til misforståelser.

EKSEMPEL 1. Det mest nærliggende eksempel på situationen beskrevet ovenfor går ud på, at V er $\mathbb{R}^n$, og at vi benytter den naturlige basis $E = (e_1, e_2, \cdots, e_n)$. Derved får enhver $\mathbb{R}^n$-vektor x så at sige sig selv som koordinatsæt, idet

$$x = x_1 e_1 + x_2 e_2 + \cdots + x_n e_n \,.$$

En anden mulighed er, at V stadig er $\mathbb{R}^n$, men at vi som basis benytter et andet sæt lineært uafhængige $\mathbb{R}^n$-vektorer, $Q = (q_1, q_2, \cdots, q_n)$. Når vi refererer til en sådan basis, forskellig fra den naturlige, må vi passe på at skelne mellem, om et $\mathbb{R}^n$-sæt skal opfattes som koordinatsæt eller som "sig selv", dvs som $\mathbb{R}^n$-vektor.

Antag f.eks. $V = \mathbb{R}^2$ og $Q = (q_1, q_2)$, hvor $q_1 = (3,1)$, $q_2 = (-1,2)$. Den vektor $\boldsymbol{x}$, der har koordinatsættet $(4,5)$ mht basen Q, er da

$$\boldsymbol{x} = 4\begin{pmatrix}3\\1\end{pmatrix} + 5\begin{pmatrix}-1\\2\end{pmatrix} = \begin{pmatrix}7\\14\end{pmatrix}.$$

På tilsvarende måde kan vi finde koordinatsættet $\boldsymbol{y} = (y_1, y_2)$ for vektoren $(2,9)$ mht basen Q ved at løse vektorligningen

$$\begin{pmatrix}2\\9\end{pmatrix} = y_1\begin{pmatrix}3\\1\end{pmatrix} + y_2\begin{pmatrix}-1\\2\end{pmatrix};$$

der igen er ensbetydende med systemet

$$3y_1 - y_2 = 2, \qquad y_1 + 2y_2 = 9.$$

Ved at løse disse to lineære ligninger i to ubekendte på sædvanlig vis finder man det søgte koordinatsæt, som viser sig at være

$$\boldsymbol{y} = (y_1, y_2) = \left(\frac{13}{7}, \frac{25}{7}\right).$$

Koordinater mht en valgt basis kan også benyttes, når V er et n-dimensionalt *underrum* af et talrum $\mathbb{R}^m$ af højere dimension, altså med $m > n$, og vi vælger en eller anden basis for V.

Lad f.eks. V være den plan gennem nulpunktet i $\mathbb{R}^3$, der har ligning $x_1 - x_2 + 2x_3 = 0$. Der gælder: $n = \dim V = 2$, $m = 3$. Som en basis for V kan vi benytte et vilkårligt sæt af to ikke-proportionale vektorer, som opfylder V's ligning. Vi vælger $Q = (\boldsymbol{a}, \boldsymbol{b})$, hvor $\boldsymbol{a} = (1,1,0)$ og $\boldsymbol{b} = (0,2,1)$. Der er derefter ikke noget i vejen for f.eks. at sige, at en $\mathbb{R}^3$-vektor $\boldsymbol{x} \in V$ "har koordinaterne $(3,5)$ mht Q". Det betyder, at

$$\boldsymbol{x} = 3\boldsymbol{a} + 5\boldsymbol{b} = 3\begin{pmatrix}1\\1\\0\end{pmatrix} + 5\begin{pmatrix}0\\2\\1\end{pmatrix} = \begin{pmatrix}3\\13\\5\end{pmatrix}.$$

(Bemærk, at $\boldsymbol{x}$ opfylder planens ligning, idet $3 - 13 + 2 \cdot 5 = 0$.)

For endelig - som en undtagelse - at illustrere, at V ikke nødvendigvis skal være en mængde, hvis elementer er talsæt, nævner vi det følgende eksempel: V er vektorrummet $\mathcal{P}_2$ af polynomier af højst anden grad, og E er sættet (p_0, p_1, p_2), hvor $p_0(x) = 1$, $p_1(x) = x$, $p_2(x) = x^2$. At $p \in \mathcal{P}_2$ har koordinatsættet $(2,-1,3)$ mht basen E, betyder da, at p er polynomiet $p(x) = 2 - x + 3x^2$.

ØVELSE 1. (1) Find den vektor i $\mathbb{R}^2$, der har koordinatsættet $(5,-1)$ mht basen $Q = ((1,2),\ (2,5))$. Find koordinatsættet mht samme basis for vektoren $(35,-47)$. (For at undgå at blande vektorerne "selv" og deres koordinatsæt sammen kunne man sætte koordinater i [ ], f.eks. $[5,-1]$. Det gøres dog ikke i denne bog.)

(2) Find den vektor i $\mathbb{R}^3$, der har koordinatsættet $(1,3,4)$ mht basen $Q = ((1,-1,0),$ $(0,2,1),\ (0,-1,3))$. Find koordinatsættet mht Q for vektoren $(14,7,0)$.

Betragt atter et vilkårligt n-dimensionalt vektorrum V, og antag nu, at der er givet to forskellige baser $E = (\boldsymbol{E}_1,\cdots,\boldsymbol{E}_n)$ og $E^* = (\boldsymbol{E}_1^*,\cdots,\boldsymbol{E}_n^*)$ i V. En vilkårlig vektor $\boldsymbol{X} \in V$ har koordinatsæt mht begge baser. Lad $\boldsymbol{x}$ være koordinatsættet for $\boldsymbol{X}$ mht E, og lad $\boldsymbol{y}$ være koordinatsættet for $\boldsymbol{X}$ mht E^*. Hvad er sammenhængen mellem $\boldsymbol{x}$ og $\boldsymbol{y}$?

For at kunne besvare det spørgsmål må vi vide, hvordan vektorerne i den ene basis udtrykkes ved vektorerne i den anden. Da situationen ofte er den, at den ene basis E foreligger, hvorefter man går over til en anden basis E^*, kan vi lette sprogbrugen ved at tale om hhv 'den gamle' og 'den nye basis'. Vi antager, at sammenhængen mellem de to er givet på den måde, at vi kender *de nye basisvektorers gamle koordinater.* Lad $\boldsymbol{q}_j = (q_{1j},\cdots,q_{nj})^{\mathsf{T}}$ være koordinatsættet for $\boldsymbol{E}_j^*$ mht E, dvs der gælder

$$\boldsymbol{E}_j^* = \sum_{i=1}^{n} q_{ij}\boldsymbol{E}_i = q_{1j}\boldsymbol{E}_1 + \cdots + q_{nj}\boldsymbol{E}_n \qquad (j = 1,\cdots,n)\,.$$

Om en vilkårlig vektor $\boldsymbol{X} \in V$ har vi da:

$$\boldsymbol{X} = \sum_{i=1}^{n} x_i\boldsymbol{E}_i = \sum_{j=1}^{n} y_j\boldsymbol{E}_j^* = \sum_{j=1}^{n}\sum_{i=1}^{n} y_j q_{ij}\boldsymbol{E}_i = \sum_{i=1}^{n}\left(\sum_{j=1}^{n} q_{ij}y_j\right)\boldsymbol{E}_i\,.$$

Da koordinaterne x_i mht E er *entydige*, følger heraf:

$$x_i = \sum_{j=1}^{n} q_{ij}y_j = q_{i1}y_1 + \cdots + q_{in}y_n \qquad (i = 1,\cdots,n)\,. \tag{2}$$

Vi vil samle - og let udbygge - disse overvejelser i en sætning:

SÆTNING 1. Lad E og E^* være to baser for et givet n-dimensionalt vektorrum V, og antag, at E^*-basisvektorernes koordinatsæt mht E er hhv $q_1, \cdots, q_n$. En vilkårlig vektors koordinatsæt $\boldsymbol{x}$ mht E og $\boldsymbol{y}$ mht E^* er da forbundet ved formlen

$$\boldsymbol{x} = \boldsymbol{Q}\boldsymbol{y}, \tag{3}$$

hvor *basisskiftmatricen* $\boldsymbol{Q} \in \mathcal{M}_n$ har q_j'erne som søjler. Matricen $\boldsymbol{Q}$ er regulær, og sammenhængen mellem $\boldsymbol{x}$ og $\boldsymbol{y}$ den modsatte vej er

$$\boldsymbol{y} = \boldsymbol{Q}^{-1}\boldsymbol{x}\ . \tag{4}$$

BEVIS: Formel (3) er identisk med (2) samlet på vektor-matrix form. Hvis $\boldsymbol{Q}$ var singulær, kunne vi finde en linearkombination af E^*-vektorerne, som var nulvektoren, i strid med, at E^* er en basis for V. Altså er $\boldsymbol{Q}$ regulær, og når vi ganger med $\boldsymbol{Q}^{-1}$ fra venstre i (3), går den over i (4). □

I de fleste tilfælde vil vi være i den situation, at V er $\mathbb{R}^n$, og E er den naturlige basis $(e_1, \cdots, e_n)$, mens den nye basis er et vektorsæt $Q = (q_1, \cdots, q_n)$. De nye basisvektorer q_j har da sig selv som koordinatsæt mht E, og basisskiftmatricen $\boldsymbol{Q}$ har q_j'erne som søjler.

EKSEMPEL 2. Betragt igen $\mathbb{R}^2$-situationen i Eksempel 1. Den kan opfattes på den måde, at vi sammen med den naturlige basis E også benytter basen $Q = (q_1, q_2)$, hvor $q_1 = (3,1)$ og $q_2 = (-1,2)$. Basisskiftmatricen fra E til Q er da $\boldsymbol{Q} = \begin{pmatrix} 3 & -1 \\ 1 & 2 \end{pmatrix}$. Hvis vi om en vektor $\boldsymbol{x}$ ved, at dens koordinatsæt mht Q er $(4,5)$, kan vi finde dens koordinatsæt mht E, dvs "vektoren selv", af formel (3):

$$\boldsymbol{x} = \boldsymbol{Q}\begin{pmatrix} 4 \\ 5 \end{pmatrix} = \begin{pmatrix} 3 & -1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 4 \\ 5 \end{pmatrix} = \begin{pmatrix} 7 \\ 14 \end{pmatrix}.$$

Ligeledes kan vi af (4) finde koordinatsættet $\boldsymbol{y}$ mht Q for vektoren $\boldsymbol{x} = (2,9)$:

$$\boldsymbol{y} = \boldsymbol{Q}^{-1}\boldsymbol{x} = \frac{1}{\det \boldsymbol{Q}}\begin{pmatrix} 2 & 1 \\ -1 & 3 \end{pmatrix}\begin{pmatrix} 2 \\ 9 \end{pmatrix} = \begin{pmatrix} \frac{13}{7} \\ \frac{25}{7} \end{pmatrix}.$$

Altså samme regninger som i Eksempel 1, men sat bedre i system via Sætning 1.

ØVELSE 2. (1) Find i $\mathbb{R}^2$ basisskiftmatricen (jfr. Øvelse 1)

1° fra den naturlige basis til basen $((1,2),\ (2,5))$,

2° fra basen $((1,2),\ (2,5))$ til den naturlige basis.

Afprøv med dette basisskift formlerne (3) og (4) på vektorerne i Øvelse 1, første del.

(2) Regn på samme måde anden del af Øvelse 1 ved matrixregning.

EKSEMPEL 3. Lad E, E^* og E^{**} betegne tre baser for et n-dimensionalt vektorrum V. Lad endvidere $\boldsymbol{Q}$ og $\boldsymbol{Q}^*$ være basisskiftmatricen fra E til E^*, hhv basisskiftmatricen fra E^* til E^{**}. Om en vilkårlig V-vektors koordinatsæt $\boldsymbol{x}$, $\boldsymbol{y}$ og $\boldsymbol{z}$ mht de tre baser gælder da

$$\boldsymbol{x} = \boldsymbol{Q}\boldsymbol{y}, \qquad \boldsymbol{y} = \boldsymbol{Q}^*\boldsymbol{z}.$$

Ved kombination af de to ligninger følger

$$\boldsymbol{x} = \boldsymbol{Q}\boldsymbol{Q}^*\boldsymbol{z}.$$

Dermed har vi vist, at *basisskiftmatricen fra basis E til basis E^{**} er $\boldsymbol{Q}\boldsymbol{Q}^*$.*

En lignende regel kan udledes for et basisskift i $\mathbb{R}^n$ mellem to vilkårlige baser, hvoraf ingen behøver at være den naturlige. Lad P og Q være to baser for $\mathbb{R}^n$, og lad $\boldsymbol{P}$ og $\boldsymbol{Q}$ være de tilhørende $\mathcal{M}_n$-matricer, dvs $\boldsymbol{P}$ er så at sige basisskiftmatricen fra den naturlige basis til basen P, med P-vektorerne som søjler, og tilsvarende er $\boldsymbol{Q}$ basisskiftmatricen fra den naturlige basis til basen Q, med Q-vektorerne som søjler.

Betragt nu en vilkårlig vektor $\boldsymbol{x} \in \mathbb{R}^n$. Dens koordinatsæt $\boldsymbol{y}$ mht P og $\boldsymbol{z}$ mht Q er forbundet med $\boldsymbol{x}$ ved

$$\boldsymbol{x} = \boldsymbol{P}\boldsymbol{y} = \boldsymbol{Q}\boldsymbol{z}.$$

Ved at gange fra venstre med $\boldsymbol{P}^{-1}$ får vi heraf

$$\boldsymbol{y} = \boldsymbol{P}^{-1}\boldsymbol{Q}\boldsymbol{z}.$$

Dermed har vi vist, at *basisskiftmatricen fra basis P til basis Q er $\boldsymbol{P}^{-1}\boldsymbol{Q}$.*

ØVELSE 3. Find ved brug af den sidste regel i Eksempel 3 basisskiftmatricen fra $\mathbb{R}^2$-basen $((1,2),\ (2,5))$ til $\mathbb{R}^2$-basen $((2,3),\ (-1,1))$.

ØVELSE 4. (1) Givet $\mathbb{R}^4$-vektorerne

$$a_1 = (1,1,1,0), \qquad a_2 = (0,1,1,1), \qquad b_1 = (2,5,5,3), \qquad b_2 = (1,0,0,-1).$$

Vis, at $Q_1 = (a_1, a_2)$ og $Q_2 = (b_1, b_2)$ er to forskellige baser for *samme* 2-dimensionale underrum U af $\mathbb{R}^4$. Find basisskiftmatricen fra Q_1 til Q_2. [Vink: Der er tale om en 2×2 matrix, ikke en 4×4 matrix.]

(2) Vis, at vektoren $(9, 12, 12, 3)$ tilhører underrummet U, og find dens koordinatsæt mht hver af de to baser.

(3) Vis, at U også kan beskrives som løsningsmængde til de to homogene lineære ligninger $x_1 + x_4 = x_2$, $x_2 = x_3$.

Lad os gentage, at vi i det store og hele vil holde os til tilfældet $V = \mathbb{R}^n$. De fleste resultater kan dog ret nemt overføres til generelle vektorrum, herunder til underrum af rum af type $\mathbb{R}^n$. Bemærk i øvrigt, at basisskiftudtrykkene (3) og (4) svarer til udtrykkene for koordinatskifte i $\mathbb{R}^2$, se Grundbog i Matematik side 89. Her i $\mathbb{R}^n$ nøjes vi dog med det "nulpunktsbevarende tilfælde".

2. Ortogonal matrix. (§2 overspringes !)

DEFINITION 2. **Ortogonal basis. Ortonormeret basis.** En basis $Q = (q_1, \cdots, q_n)$ for $\mathbb{R}^n$ kaldes *ortogonal*, når basisvektorerne er indbyrdes ortogonale, dvs når $q_i \cdot q_j = 0$ for $i \neq j$. Når basisvektorerne desuden alle er enhedsvektorer, dvs når

$$q_i \cdot q_j = \begin{cases} 1 & \text{for } i = j\,, \\ 0 & \text{for } i \neq j\,, \end{cases} \tag{1}$$

siger man, at basen Q er *ortonormeret.*

Den naturlige basis i $\mathbb{R}^n$ er ortonormeret. Som et andet eksempel på en ortonormeret basis kan i $\mathbb{R}^2$ nævnes $(q_1, q_2) = ((\frac{4}{5}, -\frac{3}{5}), (\frac{3}{5}, \frac{4}{5}))$. Det er nemt at gå efter, at $|q_1| = |q_2| = 1$, samt at $q_1 \cdot q_2 = 0$.

Vi bemærker, at hvis et vektorsæt $Q = (q_1, \cdots, q_n)$ opfylder (1), er det automatisk en basis for $\mathbb{R}^n$. Hvis vi nemlig i ligningen

$$t_1 q_1 + \cdots + t_i q_i + \cdots + t_n q_n = o$$

ganger skalært med q_i på begge sider, får vi $t_i = 0 \quad (i = 1, \cdots, n)$. Der er altså lineær uafhængighed i Q, som følgelig er en basis for $\mathbb{R}^n$.

DEFINITION 3. **Ortogonal matrix.** En $n \times n$ matrix Q kaldes *ortogonal*, når dens søjler er en ortonormeret basis for $\mathbb{R}^n$.

Af bemærkningerne lige efter Definition 2 følger, at

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} \frac{4}{5} & -\frac{3}{5} \\ \frac{3}{5} & \frac{4}{5} \end{pmatrix}$$

er eksempler på ortogonale matricer.

Betegnelsen 'ortogonal matrix' kan synes inkonsekvent: Hvorfor ikke nøjes med at kræve, at søjlerne er ortogonale, og så bruge f.eks. betegnelsen 'ortonormal' om en matrix, hvor søjlerne yderligere vides at være enhedsvektorer? Det er der faktisk også nogle der gør! Også sprog*brugen* er altså en smule inkonsekvent. At Definition 3 alligevel er den, de fleste foretrækker, forklares af den følgende sætning. Hovedpointen i den er, at når *søjlerne* i Q er enhedsvektorer, som er indbyrdes ortogonale, så gælder det samme automatisk for *rækkerne*. Det er langt fra indlysende, men alligevel er beviset kort. Det skyldes, at vejen er banet gennem begreberne invers matrix og komplement i Kapitel 2.

SÆTNING 2. Følgende udsagn om en $\mathcal{M}_n$-matrix Q er ensbetydende:

1° Q er ortogonal.
2° $Q^\top Q = E$.
3° $Q^{-1} = Q^\top$.
4° $Q^\top$ er ortogonal.

BEVIS: Formlen
$$Q^\top Q = E$$

udtrykker præcis det samme som formel (1) i Definition 2, for det ij'te element i produktet på venstre side er lig med det skalære produkt af den i'te række i $Q^\top$, som er q_i (!), og den j'te søjle i Q, som er q_j. Dermed er det vist, at 1° og 2° er ensbetydende. Samtidig udtrykker 2°, at $Q^\top$ er invers til Q, dvs 2° og 3° er ensbetydende.

Antag, at 1° gælder, og dermed også 2° og 3°. Da har vi

$$(Q^\top)^\top Q^\top = QQ^\top = QQ^{-1} = E. \tag{2}$$

I (2) spiller $Q^\top$ så at sige rollen som Q i 2°, dvs (2) udtrykker, at $Q^\top$ er ortogonal. Hvis omvendt $Q^\top$ er ortogonal, kan man heraf slutte, at $(Q^\top)^\top = Q$ er ortogonal. Dermed er det også godtgjort, at 2° og 4° er ensbetydende, og vi er færdige med beviset. □

SÆTNING 3. Om en ortogonal $\mathcal{M}_n$-matrix Q gælder

$$\det Q = \pm 1. \tag{3}$$

BEVIS: Af 2° i Sætning 2 fås

$$\det(Q^\top Q) = \det Q^\top \det Q = (\det Q)^2 = \det E = 1,$$

som straks leder til (3). □

DEFINITION 4. **Egentligt og uegentligt ortogonal matrix.** En ortogonal matrix Q kaldes *egentligt ortogonal*, hvis $\det Q = 1$, og *uegentligt ortogonal*, hvis $\det Q = -1$.

Hvis man tænker på den ortogonale matrix Q som matrix for en lineær afbildning i $\mathbb{R}^n$, er den et n-dimensionalt modstykke til en nulpunktsbevarende *stiv bevægelse*, dvs en drejning om nulpunktet i $\mathbb{R}^2$ eller $\mathbb{R}^3$, evt. sammensat med en spejling om en linie, hhv en plan. Ved en sådan stiv bevægelse i $\mathbb{R}^2$ går enhedskvadratet over i et dermed kongruent, evt. spejlvendt kongruent kvadrat med siden 1, dvs med areal $\det Q = \pm 1$. Og ved den tilsvarende afbildning i $\mathbb{R}^3$ går enhedsterningen over i en dermed kongruent, evt. spejlvendt kongruent terning med kantlængde 1, dvs med rumfang $\det Q = \pm 1$.

Tænker man på Q som matrix for et basisskift i $\mathbb{R}^n$, er den et n-dimensionalt modstykke til overgang fra ét sædvanligt retvinklet koordinatsystem til et andet i $\mathbb{R}^2$ eller $\mathbb{R}^3$, altså en drejning af et sådant system om nulpunktet, evt. kombineret med en ombytning af to af akserne, svarende til $\det Q = -1$.

Både når Q fortolkes i forbindelse med en lineær afbildning og et basisskift, kan ortogonalitet altså opfattes som en slags *formbevarende* egenskab, mens egentlig og uegentlig ortogonalitet er n-dimensionale modstykker til begreberne 'retvendt' og 'spejlvendt' i plan eller rum.

ØVELSE 5. (1) Bestem tallene a og b, således at matricen $\begin{pmatrix} 3/5 & a \\ 4/5 & b \end{pmatrix}$ er egentligt ortogonal. Bestem derefter a og b, således at matricen er uegentligt ortogonal.

(2) Vis, at enhver ortogonal 2×2 matrix enten er af formen $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ eller af formen $\begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}$. Antag, at matricen er af den første af de to typer. Hvad er den geometriske betydning af θ, når matricen opfattes som hørende til en lineær afbildning i $\mathbb{R}^2$? Og når den opfattes som hørende til et basisskift i $\mathbb{R}^2$?

ØVELSE 6. Vis, at 3×3 matricen $Q = \begin{pmatrix} \frac{2}{3} & -\frac{2}{3} & -\frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \end{pmatrix}$ er ortogonal. Undersøg, om Q er egentligt eller uegentligt ortogonal.

Sammensættes stive bevægelser i plan eller rum, fås en stiv bevægelse, og successiv overgang mellem sædvanlige retvinklede systemer i plan/rum er lig med en enkelt overgang af samme art. Tilsvarende gælder:

SÆTNING 4. Hvis $\mathcal{M}_n$-matricerne $\boldsymbol{P}$ og $\boldsymbol{Q}$ er ortogonale, er også produktmatricen $\boldsymbol{PQ}$ ortogonal.

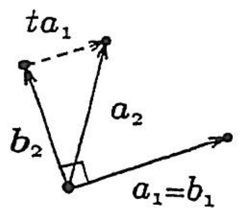
BEVIS: Vi benytter Sætning 2, 2°: Af $\boldsymbol{P}^{\top}\boldsymbol{P} = \boldsymbol{E}$ og $\boldsymbol{Q}^{\top}\boldsymbol{Q} = \boldsymbol{E}$ følger

$$(\boldsymbol{PQ})^{\top}(\boldsymbol{PQ}) = \boldsymbol{Q}^{\top}\boldsymbol{P}^{\top}\boldsymbol{PQ} = \boldsymbol{Q}^{\top}\boldsymbol{Q} = \boldsymbol{E},$$

og dermed er påstanden vist. □

EKSEMPEL 4. **Ortogonalisering.**

Betragt et ægte underrum U af $\mathbb{R}^n$, givet ved en basis $Q = (a_1, a_2, \cdots, a_r)$, dvs $\dim U = r$. Vi vil beskrive, hvordan man ud fra den givne basis Q kan konstruere en *ortogonal* basis $Q^* = (b_1, b_2, \cdots, b_r)$ for U.



Som første vektor vælger vi $b_1 = a_1$. Dernæst søger vi $t \in \mathbb{R}$, således at $b_2 = a_2 - tb_1$ (der tilhører U uanset t) er ortogonal på b_1:

$$b_1 \cdot b_2 = b_1 \cdot a_2 - t(b_1)^2 = 0 \quad \text{eller} \quad t = \frac{b_1 \cdot a_2}{(b_1)^2},$$

dvs for denne værdi af t er b_1 og b_2 indbyrdes ortogonale. Videre søger vi $t_1, t_2 \in \mathbb{R}$, således at $b_3 = a_3 - t_1 b_1 - t_2 b_2$ er ortogonal på såvel b_1 som b_2:

$$b_1 \cdot b_3 = b_1 \cdot a_3 - t_1(b_1)^2 - t_2(b_1 \cdot b_2) = 0 \qquad \text{eller} \qquad t_1 = \frac{b_1 \cdot a_3}{(b_1)^2},$$

$$b_2 \cdot b_3 = b_2 \cdot a_3 - t_1(b_2 \cdot b_1) - t_2(b_2)^2 = 0 \qquad \text{eller} \qquad t_2 = \frac{b_2 \cdot a_3}{(b_2)^2}.$$

(Vi benyttede undervejs, at $b_1 \cdot b_2 = 0$). For disse værdier af t_1 og t_2 er b_1, b_2 og b_3 altså indbyrdes ortogonale.

Det skulle nu være klart, hvordan man fortsætter processen og til sidst når frem til et sæt $Q^* = (b_1, \cdots, b_r)$ af indbyrdes ortogonale U-vektorer. Pga den lineære uafhængighed af a_i'erne kan ingen af b_i'erne blive lig med o (i udtrykket for b_i har a_i jo koefficienten 1), og da b_i'erne er indbyrdes ortogonale, ses ligesom på side 162, at de er lineært uafhængige. De udgør med andre ord en basis for U, og denne basis er ortogonal.

Metoden, der betegnes *ortogonalisering* (evt.: Gram-Schmidt ortogonalisering), kan ved en lille udvidelse bruges til at konstruere en *ortonormeret* basis for U. Man skal blot afslutte med at normere hver enkelt basisvektor, dvs erstatte b_i med $\beta_i = b_i/|b_i|$ $(i = 1, \cdots, r)$. Endvidere kan metoden generaliseres til et vilkårligt n-dimensionalt vektorrum med indre produkt.

ØVELSE 7. I $\mathbb{R}^4$ betragtes det 3-dimensionale underrum $U = \mathrm{sp}(a_1, a_2, a_3)$, hvor

$$a_1 = (1,0,1,0)^\top, \qquad a_2 = (4,1,2,1)^\top, \qquad a_3 = (0,0,1,-1)^\top.$$

Benyt ortogonalisering, der er beskrevet i det foregående eksempel, til at konstruere en ortogonal basis for U. Find også en ortonormeret basis for U.

3. Lineær afbildning og basisskift.

Til en lineær afbildning $f : \mathbb{R}^n \curvearrowright \mathbb{R}^n$ har vi hidtil knyttet én bestemt matrix A, entydigt fastlagt ved, at dens j'te søjlevektor a_j er $f(e_j)$, altså billedet ved f af den j'te naturlige basisvektor e_j. Eller mere udførligt: det vilkårlige element a_{ij} i A er lig med i'te koordinat af $f(e_j)$ mht den naturlige basis E. Hvis $x \in \mathbb{R}^n$ er vilkårlig, og x' er dens billede ved f, $x' = f(x)$, så vil der, som omtalt side 63-64 og benyttet talrige gange i det foregående, gælde

$$x' = Ax \, . \tag{1}$$

Det viser sig, at indførelsen af koordinater mht en vilkårlig basis leder til en kraftig udvidelse af begrebet 'matrix hørende til en lineær afbildning'. For en given afbildning f kan man *mht en vilkårlig basis* $Q = (q_1, \cdots, q_n)$ knytte en matrix A til f. Denne matrix er defineret ved, at dens j'te søjlevektor a_j er koordinatsættet mht Q til $f(q_j)$, altså til billedet ved f for den j'te vektor i basen Q. Eller mere udførligt: a_{ij} er lig med i'te koordinat mht Q af $f(q_j)$. Hvis vi nu lader x og x' betegne koordinatsættene mht Q for hhv en vilkårlig $\mathbb{R}^n$-vektor og dens billede ved f, så vil der, som et ræsonnement helt mage til dét på side 63-64 viser, gælde

$$x' = Ax \, . \tag{2}$$

Denne formel er formelt den samme som (1), men reelt en stærk udvidelse af den, idet der for samme f kan optræde mange forskellige matricer A, alt afhængigt af, hvilken basis vi har valgt at benytte. Formel (2) indeholder (1) som specialtilfælde, nemlig når den naturlige basis E er valgt.

Hvordan er sammenhængen mellem matricerne for *samme* afbildning mht *forskellige* baser ? For at finde svaret på dette spørgsmål betragter vi en lineær afbildning $f : \mathbb{R}^n \curvearrowright \mathbb{R}^n$ og to baser E^* og E^{**}. Lad A være matricen for f mht E^*, og lad tilsvarende B være matricen for f mht E^{**}. Lad endvidere $\xi \in \mathbb{R}^n$ være en vilkårlig vektor, lad x og x' være koordinatsættene for hhv ξ og $f(\xi)$ mht basen E^*, og lad tilsvarende y og y' være koordinatsættene for hhv ξ og $f(\xi)$ mht basen E^{**}. Lad endelig Q være basisskiftmatricen fra basis E^* til basis E^{**}.

Vi kan nu skrive en række sammenhænge mellem de fire koordinatsæt op. Af definitionen på A og B følger, at

$$x' = Ax, \qquad y' = By. \tag{3}$$

Da x og y er koordinatsæt for samme vektor mht hver sin af de to baser, og da det samme gælder x' og y', er endvidere

$$x = Qy, \qquad x' = Qy'. \tag{4}$$

Ved kombination af (3) og (4) samt brug af, at Q er regulær, fås

$$y' = Q^{-1}x' = Q^{-1}Ax = Q^{-1}AQy.$$

Når denne formel sammenholdes med højre del af (3), og når vi desuden benytter, at f's matrix mht en given basis (her: basen E^{**}) er entydigt bestemt, får vi den ønskede sammenhæng mellem A og B. Vi formulerer dette resultat som en sætning, hvor kun de væsentlige størrelser indgår.

SÆTNING 5. Lad A og B være matricerne for en lineær afbildning f i $\mathbb{R}^n$ mht to givne baser, og lad Q være basisskiftmatricen fra den første til den anden af disse baser. Da gælder

$$B = Q^{-1}AQ. \tag{5}$$

Hvis specielt Q er ortogonal, antager (5) formen

$$B = Q^{\top}AQ. \tag{6}$$

BEVIS: (Fremgår af ovenstående samt af Sætning 2, 3°.) □

Det skal straks fremhæves, at selv om vi nu har indført en flertydighed i begrebet 'matrix for en lineær afbildning', vil vi stadigvæk normalt hæfte os mest ved dens matrix mht *den naturlige basis*, altså den, vi hidtil har opfattet som 'afbildningens matrix', og vi vil fortsætte med at bruge dette udtryk og forstå det som hidtil, når intet andet siges.

Hvad skal vi da overhovedet med den udvidelse, vi har gennemført? Svaret er, at hvis man over for en given afbildning f er i stand til at foretage det rigtige basisskift, kan man ofte opnå et simplere udtryk for afbildningen end det, der forelå i første omgang. Når f og A er givet, kan Sætning 5 ses som udgangspunkt for at finde en "bedre" basis, dvs bestemme et basisskift Q, således at $\boldsymbol{B} = \boldsymbol{Q}^{-1}\boldsymbol{AQ}$ er "så pæn som muligt". *Spektralteorien*, som vi kommer til i næste afsnit, er velegnet til at angribe dette problem.

Lad os desuden bemærke, at situationen normalt er knapt så indviklet, som det kunne se ud i overvejelserne forud for Sætning 5. I de fleste tilfælde vil 'den første basis' være den naturlige, dvs $E^* = E$ og dermed $\boldsymbol{\xi} = \boldsymbol{x}$. Dermed vil den nye basis som omtalt i kapitlets første afsnit være tæt knyttet til basisskiftmatricen $\boldsymbol{Q}$, idet den vil være $E^{**} = Q = (\boldsymbol{q}_1, \cdots, \boldsymbol{q}_n)$, sættet af $\boldsymbol{Q}$'s søjlevektorer. Men vi valgte at gøre sætningen mere generel - med den omkostning, at argumentationen blev lidt mere indviklet.

EKSEMPEL 5. En lineær afbildning $\boldsymbol{f} : \mathbb{R}^2 \curvearrowright \mathbb{R}^2$ er givet ved, at den har matricen $\boldsymbol{A} = \begin{pmatrix} 3 & 2 \\ 2 & 3 \end{pmatrix}$ mht den naturlige basis. Udtrykket for afbildningen er altså:

$$\boldsymbol{f}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1' \\ x_2' \end{pmatrix} = \begin{pmatrix} 3x_1 + 2x_2 \\ 2x_1 + 3x_2 \end{pmatrix}.$$

Vi vil finde udtrykket for samme afbildning mht en anden basis svarende til, at X_1X_2-systemet drejes $-45°$ om nulpunktet. Da går $\boldsymbol{e}_1 = (1,0)$ over i $\boldsymbol{q}_1 = (\cos(-45°), \sin(-45°)) = (\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}})$, og $\boldsymbol{e}_2 = (0,1)$ går over i $\boldsymbol{q}_2 = (\cos 45°, \sin 45°) = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$. Basisskiftmatricen er altså

$$\boldsymbol{Q} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}.$$

Bemærk, at $\boldsymbol{Q}$ er ortogonal. Det stemmer med, at basisskiftet er knyttet til en stiv bevægelse (drejning) af det givne system, der svarer til den naturlige basis.

Matricen $\boldsymbol{B}$ for $\boldsymbol{f}$ i det drejede system, Y_1Y_2-systemet, finder vi af Sætning 5:

$$\begin{aligned}\boldsymbol{B} = \boldsymbol{Q}^{\mathsf{T}}\boldsymbol{AQ} &= \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 3 & 2 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \\ &= \left(\frac{1}{\sqrt{2}}\right)^2 \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 3 & 2 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \\ &= \frac{1}{2}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 5 \\ -1 & 5 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 2 & 0 \\ 0 & 10 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 5 \end{pmatrix}.\end{aligned}$$

Udtrykket for $\boldsymbol{f}$ i Y_1Y_2-koordinater er dermed

$$\begin{pmatrix} y_1' \\ y_2' \end{pmatrix} = \boldsymbol{By} = \begin{pmatrix} y_1 \\ 5y_2 \end{pmatrix},$$

hvilket viser, at afbildningen er *en ret affinitet omkring Y_1-aksen i forholdet* 5. Dét var ikke så nemt at se af det oprindelige udtryk for $\boldsymbol{f}$. Her har vi altså et eksempel på, at et basisskift kan føre til et mere gennemsigtigt udtryk for en given lineær afbildning. Det fremgår dog ikke, hvordan man finder på at vælge lige netop det basisskift, der svarer til drejningen på $-45°$. Baggrunden herfor skulle gerne blive gjort klarere på de følgende sider.

ØVELSE 8. Tegn en figur, der illustrerer afbildningen $\boldsymbol{f}$ i det foregående eksempel. Udregn nogle billeder af simple vektorer i $\mathbb{R}^2$, og afsæt både vektorerne og deres billeder på figuren. Ser det ud til, at afbildningen virkelig, som vi udregnede i eksemplet, er en ret affinitet i forhold 5 omkring linien $x_2 = -x_1$?

ØVELSE 9. Vis, at når en lineær afbildning i $\mathbb{R}^n$ har en *symmetrisk* matrix mht den naturlige basis, og man derefter foretager et basisskift med en *ortogonal* basisskiftmatrix, da får afbildningen også en symmetrisk matrix mht den nye basis. Udtrykt på en anden og mere formel måde: Vis, at når $\boldsymbol{A}$ og $\boldsymbol{Q}$ i Sætning 5 er hhv symmetrisk og ortogonal, da er $\boldsymbol{B}$ også symmetrisk. - Afbildningen i Eksempel 5 illustrerer denne regel.

ØVELSE 10. En lineær afbildning $\boldsymbol{f} : \mathbb{R}^2 \curvearrowright \mathbb{R}^2$ har matricen $\boldsymbol{A} = \begin{pmatrix} 66 & -12 \\ -12 & 59 \end{pmatrix}$ mht den naturlige basis. Koordinatsystemet drejes den spidse vinkel θ, hvor $\cos\theta = \frac{3}{5}$. Find matricen for $\boldsymbol{f}$ mht den nye basis.

4. Egenværdi og egenvektor.

Betragt atter en lineær afbildning $\boldsymbol{f} : \mathbb{R}^n \curvearrowright \mathbb{R}^n$. Man kan stille spørgsmålet: Hvilke vektorer ændres på en særlig pæn eller simpel måde ved afbildningen ? F.eks., hvilke vektorer (ud over $\boldsymbol{o}$) går ved $\boldsymbol{f}$ over i sig selv, dvs opfylder $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{x}$? Hvis $\boldsymbol{f}$ er afbildningen i $\mathbb{R}^2$ med matrix $\boldsymbol{A} = \begin{pmatrix} 2 & 3 \\ 3 & 10 \end{pmatrix}$, så har vektoren $\binom{3}{-1}$ denne egenskab, idet

$$\begin{pmatrix} 2 & 3 \\ 3 & 10 \end{pmatrix}\begin{pmatrix} 3 \\ -1 \end{pmatrix} = \begin{pmatrix} 3 \\ -1 \end{pmatrix}.$$

Ved nærmere eftersyn viser det sig dog, at det kun er de færreste lineære afbildninger, der har sådanne *fixelementer*, bortset naturligvis fra o.

Et knap så stærkt krav er, at billedvektoren $f(x)$ skal være *proportional* med den givne vektor x. At spørge efter sådanne vektorer har vist sig at være en frugtbar problemstilling, både matematisk og i forbindelse med en række af den lineære algebras anvendelser.

DEFINITION 5. **Egenvektor og egenværdi.** Lad f være en lineær afbildning af $\mathbb{R}^n$ ind i sig selv. En egentlig $\mathbb{R}^n$-vektor x siges at være *egenvektor* for f, hvis der findes et tal $\lambda \in \mathbb{R}$, således at

$$f(x) = \lambda x, \tag{1}$$

og λ kaldes en *egenværdi* for f. Man siger også, at x er en egenvektor *hørende til* egenværdien λ. - Lad A være f's matrix mht den naturlige basis i $\mathbb{R}^n$, således at (1) er ensbetydende med

$$Ax = \lambda x. \tag{2}$$

Da betegnes λ og x også som hhv egenværdi og egenvektor for A.

EKSEMPEL 6. Taleksemplet i teksten ovenfor kan med de nye betegnelser formuleres således: $\mathbb{R}^2$-afbildningen f med matrix $A = \begin{pmatrix} 2 & 3 \\ 3 & 10 \end{pmatrix}$ (eller blot: matricen A) har tallet 1 som egenværdi og $\binom{3}{-1}$ som en tilhørende egenvektor. En anden egenvektor for f hørende til samme egenværdi er $\binom{15}{-5}$, idet der gælder

$$\begin{pmatrix} 2 & 3 \\ 3 & 10 \end{pmatrix}\begin{pmatrix} 15 \\ -5 \end{pmatrix} = \begin{pmatrix} 15 \\ -5 \end{pmatrix}.$$

Endvidere har f tallet 11 som egenværdi og $\binom{1}{3}$ som en tilhørende egenvektor, idet

$$\begin{pmatrix} 2 & 3 \\ 3 & 10 \end{pmatrix}\begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 11 \\ 33 \end{pmatrix}.$$

ØVELSE 11. Vis følgende regel: f har 0 som egenværdi, hvis og kun hvis dens matrix er singulær, dvs hvis og kun hvis f ikke er enentydig.

Ligning (2) kan skrives $\boldsymbol{Ax} = \lambda \boldsymbol{x} = \lambda \boldsymbol{Ex}$, eller ved overflytning

$$(\boldsymbol{A} - \lambda \boldsymbol{E})\boldsymbol{x} = \boldsymbol{o}\,. \tag{3}$$

Iflg. reglen side 153 har (3) egentlige løsninger, hvis og kun hvis

$$\det(\boldsymbol{A} - \lambda \boldsymbol{E}) = 0\,. \tag{4}$$

DEFINITION 6. **Karakteristisk ligning. Reduktionsdeterminant.** Ligning (4) kaldes *den karakteristiske ligning* for afbildningen f (eller: for matricen $\boldsymbol{A}$), og determinanten på venstre side, altså

$$\begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ & \vdots & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{vmatrix}, \tag{5}$$

kaldes *reduktionsdeterminanten* for $\boldsymbol{A}$.

SÆTNING 6. 1° En lineær afbildning $f : \mathbb{R}^n \curvearrowright \mathbb{R}^n$, dens matrix $\boldsymbol{A}$ samt enhver anden matrix $\boldsymbol{B} \in \mathcal{M}_n$, der kan være matrix for f mht en eller anden basis, har de samme egenværdier, som er rødderne i ligning (4). Der er *højst* n sådanne egenværdier.

2° Egenvektorerne hørende til en given egenværdi λ_0 for f findes ved at indsætte $\lambda = \lambda_0$ i (3) og løse dette system. Med tilføjelse af $\boldsymbol{o}$ udgør mængden af egenvektorer hørende til λ_0 et *underrum* af $\mathbb{R}^n$.

3° Egenvektorer hørende til *forskellige* egenværdier for f er lineært uafhængige.

BEVIS: 1° Ved udregning af (5) fås et polynomium af n'te grad i λ, og et sådant har højst n rødder. De er iflg. Definition 5 fælles egenværdier for afbildningen f og matricen $\boldsymbol{A}$. At enhver anden $\boldsymbol{B}$, der er matrix for f mht en basis Q, har de samme egenværdier, er ligeledes klart ud fra Definition 5, idet egenværdierne knytter sig til afbildningen, og (1) medfører $\boldsymbol{By} = \lambda \boldsymbol{y}$, hvor $\boldsymbol{y}$ er $\boldsymbol{x}$'s koordinatsæt mht Q. Man kan dog også føre et mere formelt bevis for, at A og B har samme reduktionsdeterminant og dermed samme egenværdier. Det foregår således: Iflg. Sætning 5 er $B = Q^{-1}AQ$, hvoraf følger

$$\begin{aligned}\det(B - \lambda E) &= \det(Q^{-1}AQ - \lambda E) \\ &= \det(Q^{-1}(A - \lambda E)Q) \\ &= \frac{1}{\det Q} \cdot \det(A - \lambda E) \cdot \det Q = \det(A - \lambda E).\end{aligned}$$

2° Ved at indsætte $\lambda = \lambda_0$ i (3) får vi et homogent lineært ligningssystem med n ligninger og n ubekendte. Da systemets koefficientmatrix er singulær (sådan er λ_0 jo bestemt!), er *mindst én af ligningerne en linearkombination af de øvrige* og kan dermed undværes, dvs der er med sikkerhed andre løsninger end o. Løsningsmængden er iflg. Sætning 3, side 138 et underrum af $\mathbb{R}^n$.

3° Antag lineær afhængighed mellem nogle f-egenvektorer q_i hørende til *forskellige* egenværdier λ_i. Lad $m(> 1)$ være det *mindst mulige* antal q_i'er, der tilfredsstiller en ligning af formen

$$t_1q_1 + \cdots + t_mq_m = o \qquad (\text{alle } t_i \neq 0)\,. \tag{6}$$

Ved at tage f-værdien på begge sider af denne ligning bliver den til

$$t_1\lambda_1q_1 + \cdots + t_m\lambda_mq_m = o\,. \tag{7}$$

Nu ganger vi (6) med $-\lambda_1$ og lægger den til (7):

$$t_2(\lambda_2 - \lambda_1)q_2 + \cdots + t_m(\lambda_m - \lambda_1)q_m = o\,.$$

Der er mindst én koefficient $\lambda_i - \lambda_1 (\neq 0)$ på venstre side, dvs vi har en egentlig linearkombination af $m-1$ af q_i'erne, som er $= o$, i modstrid med definitionen af m. Lineær afhængighed mellem q_i'erne er således ikke mulig. □

Bemærkninger. (1) Det underrum af $\mathbb{R}^n$, der iflg. 2° i sætningen udgøres af egenvektorerne hørende til egenværdien λ_0, har dimension $n - r$, hvor $r = \rho(A - \lambda_0E)$. Som regel er dimensionen dog kun 1, dvs egenvektorerne hørende til en bestemt egenværdi udgør som regel en linie gennem nulpunktet. Har man en enkelt λ_0-egenvektor q, har man for $r = n - 1$ stort set dem alle, idet mængden af dem da er $\{tq \mid t \in \mathbb{R}\}$.

(2) Antallet af rødder i en ligning af n'te grad er højst n, selv når en eventuel dobbeltrod tælles med 2 gange, en tredobbelt rod med 3 gange etc. Når det i Sætning 6, 1° blev vist, at f højst har n egenværdier, gælder dette resultat derfor *også*, selv om vi tæller en rod i (4), λ_0, med dens *multiplicitet*, dvs lige så mange gange, faktoren $\lambda - \lambda_0$ går op i reduktionsdeterminanten.

(3) Mængden af egenværdier for f kaldes f's *spektrum*, og tilsvarende for en kvadratisk matrix $\boldsymbol{A}$. Når $\boldsymbol{Q}$ er regulær, har $\boldsymbol{A}$ og $\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}$ iflg. Sætning 6, 1° samme spektrum. *Spektralteori* er læren om egenværdi- og egenvektor egenskaber for kvadratiske matricer - og for andre størrelser, som teorien kan generaliseres til.

Nedenfor gives nogle konkrete eksempler på, hvordan man bestemmer egenværdier og tilhørende egenvektorer for en forelagt matrix. De rækkeoperationer i koefficientmatricen, der benyttes undervejs ved løsning af simple homogene ligningssystemer, er ikke forklaret i detaljer. Specielt om tilfældet $n = 3$ bemærkes, at for at finde egenvektorerne hørende til en given egenværdi skal man normalt løse to ikke-proportionale homogene ligninger

$$b_1x_1 + b_2x_2 + b_3x_3 = 0\,,$$
$$c_1x_1 + c_2x_2 + c_3x_3 = 0\,.$$

Det kan man f.eks. gøre ved at danne *krydsproduktet* af de to koefficientvektorer, $\boldsymbol{v} = \boldsymbol{b} \times \boldsymbol{c}$, hvorefter løsningen er $\boldsymbol{x} = t\boldsymbol{v}$ $(t \in \mathbb{R})$. Man kan også sætte en af de tre variable lig med en parameter t og løse ligningerne mht de to andre. (Da er det altså i realiteten standardmetoden, man bruger). Under alle omstændigheder skal løsningen naturligvis være den samme, uanset hvilken metode man finder den ved.

EKSEMPEL 7. Matricen $\begin{pmatrix} 2 & 3 \\ 3 & 10 \end{pmatrix}$ har den karakteristiske ligning

$$\begin{vmatrix} 2-\lambda & 3 \\ 3 & 10-\lambda \end{vmatrix} = 0\,,$$

der ved løsning giver $\lambda = 1$ og $\lambda = 11$. De tilhørende egenvektorer findes således:

$$\underline{\lambda = 1}: \begin{pmatrix} 1 & 3 \\ 3 & 9 \end{pmatrix} \boldsymbol{x} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \Leftrightarrow \quad x_1 + 3x_2 = 0 \quad \Leftrightarrow \quad \underline{\boldsymbol{x} = t\begin{pmatrix} 3 \\ -1 \end{pmatrix}} \; (t \in \mathbb{R})\,,$$

$$\underline{\lambda = 11}: \begin{pmatrix} -9 & 3 \\ 3 & -1 \end{pmatrix} \boldsymbol{x} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \Leftrightarrow \quad 3x_1 - x_2 = 0 \quad \Leftrightarrow \quad \underline{\boldsymbol{x} = t\begin{pmatrix} 1 \\ 3 \end{pmatrix}} \; (t \in \mathbb{R})\,.$$

EKSEMPEL 8. Matricen $\begin{pmatrix} -2 & 1 & -2 \\ 2 & -3 & 4 \\ 2 & -5 & 6 \end{pmatrix}$ har den karakteristiske ligning

$$\begin{vmatrix} -2-\lambda & 1 & -2 \\ 2 & -3-\lambda & 4 \\ 2 & -5 & 6-\lambda \end{vmatrix} = 0\,,$$

der ved udregning giver $-\lambda^3 + \lambda^2 + 2\lambda = 0$. Heraf finder man egenværdierne til $\lambda = -1$, $\lambda = 0$ og $\lambda = 2$. Egenvektorer hørende til $\lambda = -1$ bestemmes således:

$$\underline{\lambda = -1}: \begin{pmatrix} -1 & 1 & -2 \\ 2 & -2 & 4 \\ 2 & -5 & 7 \end{pmatrix} x = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \Leftrightarrow \quad \begin{pmatrix} -1 & 1 & -2 \\ 0 & 0 & 0 \\ 0 & -3 & 3 \end{pmatrix} x = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\Leftrightarrow \quad \begin{pmatrix} 1 & -1 & 2 \\ 0 & 1 & -1 \end{pmatrix} x = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \text{der giver} \quad \underline{x = t\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}} \quad (t \in \mathbb{R}).$$

På lignende måde findes:

$$\underline{\lambda = 0}: \quad \underline{x = t\begin{pmatrix} -1 \\ 2 \\ 2 \end{pmatrix}} \quad (t \in \mathbb{R}), \qquad \underline{\lambda = 2}: \quad \underline{x = t\begin{pmatrix} -1 \\ 2 \\ 3 \end{pmatrix}} \quad (t \in \mathbb{R}).$$

EKSEMPEL 9. Matricen $\begin{pmatrix} 5 & 4 & 0 \\ 4 & 20 & 0 \\ 0 & 0 & 4 \end{pmatrix}$ har den karakteristiske ligning

$$\begin{vmatrix} 5-\lambda & 4 & 0 \\ 4 & 20-\lambda & 0 \\ 0 & 0 & 4-\lambda \end{vmatrix} = 0\,,$$

der ved løsning giver $\lambda = 4$ (dobbeltrod) samt $\lambda = 21$, og tilhørende egenvektorer:

$$\underline{\lambda = 4}: \begin{pmatrix} 1 & 4 & 0 \\ 4 & 16 & 0 \\ 0 & 0 & 0 \end{pmatrix} x = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \Leftrightarrow \quad x_1 + 4x_2 = 0$$

$$\Leftrightarrow \quad \underline{x = t_1\begin{pmatrix} 4 \\ -1 \\ 0 \end{pmatrix} + t_2\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}} \quad (t_1, t_2 \in \mathbb{R})\,,$$

$$\underline{\lambda = 21}: \begin{pmatrix} -16 & 4 & 0 \\ 4 & -1 & 0 \\ 0 & 0 & -17 \end{pmatrix} x = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \Leftrightarrow \quad 4x_1 - x_2 = 0 \quad \text{og} \quad x_3 = 0,$$

hvor vi sætter $x_1 = t$ og derefter finder løsningen $\underline{x = t \begin{pmatrix} 1 \\ 4 \\ 0 \end{pmatrix}}$ $(t \in \mathbb{R})$.

EKSEMPEL 10. Matricen $\begin{pmatrix} 1 & -4 \\ 4 & 1 \end{pmatrix}$ har den karakteristiske ligning

$$\begin{vmatrix} 1-\lambda & -4 \\ 4 & 1-\lambda \end{vmatrix} = 0 \qquad \text{eller} \qquad \lambda^2 - 2\lambda + 17 = 0,$$

der ikke har nogen rødder, dvs matricen (og enhver lineær afbildning i $\mathbb{R}^2$, som den kan være matrix for) har ingen egenværdier og dermed heller ingen egenvektorer.

5. Diagonalisering.

Vi vil nu nærme os det spørgsmål, der blev rejst side 167-168: Givet en lineær afbildning $\boldsymbol{f}$ med matrix $\boldsymbol{A}$, hvordan finder vi et basisskift, således at f's matrix $\boldsymbol{B}$ mht den nye basis er så simpel som muligt? Spørgsmålet viser sig tæt forbundet med dét at finde egenværdier og egenvektorer for $\boldsymbol{A}$.

Først vil vi afgrænse, hvad vi mener med, at $\boldsymbol{f}$'s matrix er "så simpel som muligt". Det skal betyde, at den er en *diagonalmatrix*. Sådanne har iflg. Sætning 23, side 67 en simpel og afrundet algebra, hvortil kommer følgende yderst enkle spektralegenskaber:

> SÆTNING 7. $\mathcal{M}_n$-diagonalmatricen $\boldsymbol{D} = \operatorname{diag}(\lambda_1, \cdots, \lambda_n)$ har sine diagonalelementer λ_i som egenværdier og den naturlige basisvektor $\boldsymbol{e}_i$ som egenvektor hørende til λ_i $(i = 1, \cdots, n)$.

BEVIS: Ved indsættelse af $\boldsymbol{D}$ på $\boldsymbol{A}$'s plads i (5), side 171 fås straks, at den karakteristiske ligning for $\boldsymbol{D}$ er

$$(\lambda_1 - \lambda)(\lambda_2 - \lambda) \cdots (\lambda_n - \lambda) = 0,$$

der har rødderne $\lambda_i \quad (i = 1, 2, \cdots, n)$. Endvidere ses, at $\boldsymbol{D}\boldsymbol{e}_i = \lambda_i \boldsymbol{e}_i$, hvilket viser sætningens anden påstand. □

Hvis f.eks. $\boldsymbol{D} = \text{diag}(3, 3, 5)$, er egenvektorrummet hørende til egenværdien 3 lig med det 2-dimensionale underrum udspændt af e_1 og e_2, altså X_1X_2-planen. Mere generelt: et diagonalelement, der optræder r gange, er r-dobbelt egenværdi, og det tilhørende egenvektorrum er r-dimensionalt.

Hvad er karakteristisk for en lineær afbildning $\boldsymbol{f}$, hvis matrix mht den naturlige basis er af diagonalform? F.eks. kan afbildningen i $\mathbb{R}^3$ med matrix $\text{diag}(2, 3, \frac{1}{2})$ opfattes som *sammensat af rette affiniteter i de tre akseretninger* i forhold hhv 2, 3, og $\frac{1}{2}$, og altså med koordinatplanerne som affinitetsplaner. Opfattelsen kan generaliseres til $\mathbb{R}^n$. På lignende måde kan en afbildning, hvis matrix mht en *ortonormeret* basis er af diagonalform, opfattes som sammensat af rette affiniteter i basisvektorernes retninger.

DEFINITION 7. **Diagonalisering.** En $\mathcal{M}_n$-matrix $\boldsymbol{A}$ kaldes *diagonaliserbar*, hvis der findes en regulær $\mathcal{M}_n$-matrix $\boldsymbol{Q}$, således at $\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}$ er en diagonalmatrix $\boldsymbol{D}$. Man siger da, at $\boldsymbol{Q}$ *diagonaliserer* $\boldsymbol{A}$.

SÆTNING 8. 1° $\mathcal{M}_n$-matricen $\boldsymbol{A}$ kan diagonaliseres, hvis og kun hvis den har n *lineært uafhængige egenvektorer.* I så fald vil en matrix $\boldsymbol{Q}$ med n lineært uafhængige egenvektorer $q_1, \cdots, q_n$ som søjler diagonalisere $\boldsymbol{A}$, idet der gælder $\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q} = \boldsymbol{D}$, hvor $\boldsymbol{D} = \text{diag}(\lambda_1, \cdots, \lambda_n)$ har egenværdien λ_i hørende til q_i som sit i'te diagonalelement $(i = 1, \cdots, n)$.

2° Når $\boldsymbol{f}$ er en lineær afbildning i $\mathbb{R}^n$ med matrix $\boldsymbol{A}$, og denne diagonaliseres af matricen $\boldsymbol{Q}$, idet $\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q} = \boldsymbol{D} = \text{diag}(\lambda_1, \cdots, \lambda_n)$, så vil $\boldsymbol{f}$ have matricen $\boldsymbol{D}$ mht den basis Q, der dannes af søjlerne i $\boldsymbol{Q}$.

3° Når $\boldsymbol{A} \in \mathcal{M}_n$ kan diagonaliseres, har den n egenværdier, forudsat at hver egenværdi tælles med lige så mange gange, som den er rod i $\boldsymbol{A}$'s karakteristiske ligning.

BEVIS: 1° Antag først, at $\boldsymbol{A}$ kan diagonaliseres, og lad $\boldsymbol{Q}$ være en matrix, der gør det:

$$\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q} = \boldsymbol{D}\,, \qquad \boldsymbol{D} = \operatorname{diag}(\lambda_1, \cdots, \lambda_n)\,. \tag{1}$$

Når (1) ganges fra venstre med $\boldsymbol{Q}$ på begge sider, bliver den til

$$\boldsymbol{A}\boldsymbol{Q} = \boldsymbol{Q}\boldsymbol{D}\,. \tag{2}$$

Den j'te søjle på venstre side af denne formel er $\boldsymbol{A}\boldsymbol{q}_j$, og iflg. Sætning 8, side 105 er den j'te søjle på højre side $\lambda_j \boldsymbol{q}_j$, dvs (2) er ensbetydende med formlerne

$$\boldsymbol{A}\boldsymbol{q}_j = \lambda_j \boldsymbol{q}_j \qquad (j = 1, \cdots, n). \tag{3}$$

Disse formler udtrykker, at $\boldsymbol{q}_1, \cdots \boldsymbol{q}_n$ er egenvektorer for $\boldsymbol{A}$, og da de er søjler i den regulære matrix $\boldsymbol{Q}$, er de lineært uafhængige.

Antag dernæst, at $\boldsymbol{A}$ har n lineært uafhængige egenvektorer, dvs der gælder et sæt formler af typen (3). Ved sammenstilling bliver de til (2), hvor $\boldsymbol{Q}$ er matricen med $\boldsymbol{q}_j$'erne som søjler. Da de er lineært uafhængige, er $\boldsymbol{Q}$ regulær, og vi kan gange med $\boldsymbol{Q}^{-1}$ fra venstre på begge sider af (2), hvorved den går over i (1), som udtrykker, at $\boldsymbol{A}$ diagonaliseres af $\boldsymbol{Q}$. Dermed er 1° bevist.

2° Følger direkte af Sætning 5, side 167.

3° Antag, at $\boldsymbol{A}$ kan diagonaliseres, idet $\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q} = \boldsymbol{D} = \operatorname{diag}(\lambda_1, \cdots, \lambda_n)$. Da har $\boldsymbol{A}$ og $\boldsymbol{D}$ samme egenværdier iflg. Sætning 6, 1°, og den sidste påstand følger af Sætning 7. □

EKSEMPEL 11. I Eksempel 7 betragtede vi matricen $\boldsymbol{A} = \begin{pmatrix} 2 & 3 \\ 3 & 10 \end{pmatrix}$, der viste sig at have egenværdierne $\lambda_1 = 1$ og $\lambda_2 = 11$. Vi fandt også $\boldsymbol{A}$'s egenvektorer. Som en egenvektor hørende til egenværdien 1 vælger vi $\boldsymbol{q}_1 = \begin{pmatrix} 3 \\ -1 \end{pmatrix}$ og som en egenvektor hørende til egenværdien 11 vælger vi $\boldsymbol{q}_2 = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$. Det medfører, at matricen $\boldsymbol{Q} = \begin{pmatrix} 3 & 1 \\ -1 & 3 \end{pmatrix}$ diagonaliserer $\boldsymbol{A}$, idet

$$\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q} = \operatorname{diag}(1, 11) = \begin{pmatrix} 1 & 0 \\ 0 & 11 \end{pmatrix}.$$

Bemærk, at hver søjle i Q er "bestemt på nær en faktor". F.eks. kan vi i stedet for q_1 benytte $q_1^* = 2q_1 = (6, -2)^\top$ og i stedet for q_2 benytte $q_2^* = -3q_1 = (-3, -9)^\top$. Matricen $Q^* = \begin{pmatrix} 6 & -3 \\ -2 & -9 \end{pmatrix}$ vil med andre ord også diagonalisere A til $\operatorname{diag}(1, 11)$. Ligeledes kan man ombytte søjlerne i en matrix, der diagonaliserer A til $\operatorname{diag}(1, 11)$, men da bliver resultatet $\operatorname{diag}(11, 1)$.

I Eksempel 8 fandt vi for matricen $A = \begin{pmatrix} -2 & 1 & -2 \\ 2 & -3 & 4 \\ 2 & -5 & 6 \end{pmatrix}$, at den har egenværdier $\lambda_1 = -1$, $\lambda_2 = 0$ og $\lambda_3 = 2$, og at en tilhørende egenvektor kan vælges som hhv $q_1 = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$, $q_2 = \begin{pmatrix} -1 \\ 2 \\ 2 \end{pmatrix}$ og $q_3 = \begin{pmatrix} -1 \\ 2 \\ 3 \end{pmatrix}$. Det medfører, at matricen $Q = \begin{pmatrix} -1 & -1 & -1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{pmatrix}$ diagonaliserer A, idet

$$Q^{-1}AQ = \operatorname{diag}(-1, 0, 2). \tag{4}$$

Også her kan man gange søjlerne i Q med vilkårlige faktorer $\neq 0$, uden at egenskaben (4) ændres. Og bytter man rundt på søjlerne i en matrix Q, der opfylder (4), bevirker det blot den tilsvarende ombytning af diagonalelementerne på højre side.

I Eksempel 9 fandt vi for matricen $A = \begin{pmatrix} 5 & 4 & 0 \\ 4 & 20 & 0 \\ 0 & 0 & 4 \end{pmatrix}$, at den har egenværdierne $\lambda_1 = 4$ og $\lambda_3 = 21$. Hørende til λ_1 fandt vi endvidere et 2-dimensionalt egenvektorrum, udspændt af f.eks. $q_1 = (4, -1, 0)^\top$ og $q_2 = (0, 0, 1)^\top$, mens λ_3's egenvektorrum er 1-dimensionalt og udspændt af $q_3 = (1, 4, 0)^\top$. Matricen Q med de tre nævnte egenvektorer som søjler vil da diagonalisere A, idet

$$Q^{-1}AQ = \operatorname{diag}(4, 4, 21). \tag{5}$$

Her vil der være endnu større frihed end ovenfor til at vælge Q, idet de to første søjler kan erstattes med vilkårlige *linearkombinationer* af q_1 og q_2, blot skal søjlerne være ikke-proportionale.

I Eksempel 10 fandt vi for matricen $A = \begin{pmatrix} 1 & -4 \\ 4 & 1 \end{pmatrix}$, at den ikke har nogen egenværdier og egenvektorer. Den kan derfor *ikke* diagonaliseres.

ØVELSE 12. (Fortsættelse af Eksempel 11). Betragt $\mathcal{M}_n$-matricen $\boldsymbol{A}$ fra Eksempel 7. Kontrollér ved direkte udregning, at der med den matrix $\boldsymbol{Q}$, vi fandt i Eksempel 11, gælder $\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q} = \text{diag}(1, 11)$. Kontrollér det samme med $\boldsymbol{Q}$ erstattet af $\boldsymbol{Q}^*$. Prøv også at udregne $\boldsymbol{R}^{-1}\boldsymbol{A}\boldsymbol{R}$, hvor $\boldsymbol{R}$ er den matrix, der fås ved ombytning af søjlerne i $\boldsymbol{Q}$.

Et vigtigt tilfælde, hvor diagonalisering altid er mulig, er følgende, der illustreres af de to første matricer behandlet i Eksempel 11:

SÆTNING 9. Når $\boldsymbol{A} \in \mathcal{M}_n$ har n *forskellige* egenværdier $\lambda_1, \cdots, \lambda_n$, kan $\boldsymbol{A}$ diagonaliseres.

BEVIS: Lad egenværdierne være $\lambda_1, \cdots, \lambda_n$, og vælg en egenvektor for hver af dem, hhv $\boldsymbol{q}_1, \cdots, \boldsymbol{q}_n$. Disse er iflg. Sætning 6, 3° lineært uafhængige, og dermed følger påstanden af Sætning 8, 1°. □

Kan man slutte modsat vej i Sætning 9, dvs kan en matrix *kun* diagonaliseres, hvis den har n forskellige egenværdier? Svaret er nej, som det ses af den matrix, der blev behandlet i Eksempel 8 og i tredje del af Eksempel 11: Den har kun to egenværdier, men kan alligevel diagonaliseres, idet den har tre lineært uafhængige egenvektorer. Det samme gælder diagonalmatricen diag(2,2,5), som kun har to egenværdier, men den er jo på forhånd diagonaliseret! Vi nævner uden bevis følgende regel: Hvis $\det(\boldsymbol{A} - \lambda\boldsymbol{E})$ kan opspaltes i n førstegradsfaktorer, og egenvektorrummet for hver egenværdi λ_i har dimension r, hvor r er multipliciteten af roden λ_i, så kan $\boldsymbol{A}$ diagonaliseres.

Kan man slutte modsat vej i Sætning 8, 3°, dvs kan enhver matrix, hvis reduktionsdeterminant kan opløses fuldstændigt i førstegradsfaktorer, diagonaliseres? Også her er svaret nej, som det følgende eksempel viser.

EKSEMPEL 12. For matricen $\boldsymbol{A} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ finder vi $\det(\boldsymbol{A} - \lambda\boldsymbol{E}) = (1 - \lambda)^2$, dvs $\lambda_1 = 1$ er den eneste egenværdi. De tilhørende egenvektorer fås af $(\boldsymbol{A} - \lambda\boldsymbol{E})\boldsymbol{x} = \boldsymbol{o}$, der ved udskrivning giver $x_2 = 0$ og $0 = 0$. Egenvektorerne er altså alle vektorer på X_1-aksen: $\boldsymbol{x} = t\boldsymbol{e}_1 = \binom{t}{0}$ $(t \in \mathbb{R})$. Heraf følger, at der *ikke* findes to lineært uafhængige egenvektorer, og $\boldsymbol{A}$ kan derfor ikke diagonaliseres, til trods for, at den har "to" (sammenfaldende) egenværdier.

ØVELSE 13. (1) Gør rede for, at matricen $\begin{pmatrix} 0 & c^2 \\ -1 & 2c \end{pmatrix}$ $(c \in \mathbb{R})$ *ikke* kan diagonaliseres for nogen værdi af c.

(2) Undersøg for hver af matricerne

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

om den kan diagonaliseres. [Vink: De to matricer har samme reduktionsdeterminant, som nemmest udregnes ved opløsning efter 3. række.]

EKSEMPEL 13. **Ækvivalente matricer.** To $\mathcal{M}_n$-matricer $\boldsymbol{A}$ og $\boldsymbol{B}$ kaldes *ækvivalente* (eller 'similære', engelsk: 'similar'), og man skriver $\boldsymbol{A} \sim \boldsymbol{B}$, hvis der findes en regulær $\mathcal{M}_n$-matrix $\boldsymbol{Q}$, således at $\boldsymbol{B} = \boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}$.

Relationen $\sim$ har følgende tre egenskaber:

1° $\boldsymbol{A} \sim \boldsymbol{A}$,

2° $\boldsymbol{A} \sim \boldsymbol{B}$ medfører $\boldsymbol{B} \sim \boldsymbol{A}$,

3° $\boldsymbol{A} \sim \boldsymbol{B}$ og $\boldsymbol{B} \sim \boldsymbol{C}$ medfører $\boldsymbol{A} \sim \boldsymbol{C}$.

For at vise 1° behøver vi blot bemærke, at $\boldsymbol{A} = \boldsymbol{E}^{-1}\boldsymbol{A}\boldsymbol{E}$. For at vise 2° antager vi $\boldsymbol{A} \sim \boldsymbol{B}$, dvs $\boldsymbol{B} = \boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}$, der ved multiplikation med $\boldsymbol{Q}$ fra venstre og $\boldsymbol{Q}^{-1}$ fra højre giver $\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{B}\boldsymbol{Q}^{-1} = (\boldsymbol{Q}^{-1})^{-1}\boldsymbol{B}\boldsymbol{Q}^{-1}$, altså $\boldsymbol{B} \sim \boldsymbol{A}$. For at vise 3° antager vi $\boldsymbol{A} \sim \boldsymbol{B}$ og $\boldsymbol{B} \sim \boldsymbol{C}$, dvs $\boldsymbol{B} = \boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}$ og $\boldsymbol{C} = \boldsymbol{R}^{-1}\boldsymbol{B}\boldsymbol{R}$, hvoraf

$$\boldsymbol{C} = \boldsymbol{R}^{-1}\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}\boldsymbol{R} = (\boldsymbol{Q}\boldsymbol{R})^{-1}\boldsymbol{A}(\boldsymbol{Q}\boldsymbol{R}),$$

og denne formel er udtryk for, at $\boldsymbol{A} \sim \boldsymbol{C}$.

Egenskaberne 1° – 3° er de sædvanlige matematiske krav til en ækvivalensrelation. Man kan vise, at når de er opfyldt, giver det mening at tale om de såkaldte 'ækvivalensklasser', dvs $\mathcal{M}_n$-delmængder af indbyrdes ækvivalente matricer. Af resultaterne ovenfor fremgår, at en ækvivalensklasse kan opfattes som mængden af matricer, der kan være knyttet til en bestemt lineær afbildning, idet alle mulige baser $\boldsymbol{Q}$ tages i betragtning. Nogle klasser indeholder diagonalmatricer og består derfor af matricer, der kan diagonaliseres, andre gør ikke. F.eks. vil den ækvivalensklasse i $\mathcal{M}_3$, der indeholder $\operatorname{diag}(3, 5, 7)$, indeholde dels fem andre diagonalmatricer, svarende til de fem andre permutationer af tallene 3, 5 og 7 ned gennem diagonalen, dels mange andre $\mathcal{M}_3$-matricer, nemlig alle med egenværdier 3, 5 og 7. Det er knapt så enkelt at karakterisere klasser, der indeholder diagonalmatricer som f.eks. $\operatorname{diag}(3, 3, 7)$, og det samme gælder klasser helt uden diagonalmatricer. Vi vil i øvrigt ikke dyrke ækvivalens-synspunktet i denne fremstilling.

SÆTNING 13. Hvis $A \in \mathcal{M}_n$ har n indbyrdes ortogonale egenvektorer, er den symmetrisk.

BEVIS: Lad A have n indbyrdes ortogonale egenvektorer. Disse kan vælges som enhedsvektorer, og matricen Q med dem som søjler er da en ortogonal matrix, der diagonaliserer A:

$$D = Q^\top AQ \, ,$$

hvor D er en diagonalmatrix. Ganges på begge sider af denne ligning fra højre med $Q^\top$ $(= Q^{-1})$ og fra venstre med Q, fås $A = QDQ^\top$. Heraf følger

$$A^\top = (QDQ^\top)^\top = (Q^\top)^\top D^\top Q^\top = QDQ^\top = A \, ,$$

som viser, at A er symmetrisk. □

EKSEMPEL 15. Nogle af matricerne i eksemplerne 7-12 illustrerer forholdene ved diagonalisering af symmetriske matricer. F.eks. betragtede vi i Eksempel 7 den symmetriske matrix $A = \begin{pmatrix} 2 & 3 \\ 3 & 10 \end{pmatrix}$. Iflg. Eksempel 11 diagonaliseres den til diag(1, 11) af enhver matrix, hvis søjler er proportionale med hhv $(3,-1)^\top$ og $(1,3)^\top$. Disse to vektorer ses at være indbyrdes ortogonale, i overensstemmelse med Sætning 10. En ortogonal matrix, som diagonaliserer A, får man da ved at normere de to søjler, hvilket fører til

$$Q^{**} = \begin{pmatrix} \frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \\ -\frac{1}{\sqrt{10}} & \frac{3}{\sqrt{10}} \end{pmatrix}, \qquad \text{som opfylder} \qquad (Q^{**})^\top AQ^{**} = \operatorname{diag}(1,11) \, .$$

Ombytter man de to søjler i Q^{**}, får man en matrix R, der diagonaliserer A til diag(11, 1).

Matricen i Eksempel 8 er ikke symmetrisk. Dens egenvektorer er heller ikke indbyrdes ortogonale. Havde de været det, ville det være i modstrid med Sætning 13!

Matricen i Eksempel 9, $A = \begin{pmatrix} 5 & 4 & 0 \\ 4 & 20 & 0 \\ 0 & 0 & 4 \end{pmatrix}$, diagonaliseres iflg. Eksempel 11 til diag(4, 4, 21) af enhver matrix, hvis to første søjler er ikke-proportionale egenvektorer hørende til egenværdien 4, mens tredje søjle skal være en egenvektor hørende

8. Dominerende egenværdi.

DEFINITION 10. Antag, at $\boldsymbol{A} \in \mathcal{M}_n$ har n egenværdier (hvoraf nogle evt. falder sammen), og at en af dem er numerisk større end de øvrige. Den betegnes da som (strengt) *dominerende egenværdi* for $\boldsymbol{A}$.

Vi kan antage, at egenværdierne er $\lambda_1, \cdots, \lambda_n \in \mathbb{R}$, og at det er λ_1, der er dominerende. Den opfylder altså

$$|\lambda_1| > |\lambda_i| \qquad (i = 2, 3, \cdots, n) .$$

Hvis der er flere egenværdier, der er numerisk større end de øvrige, kaldes de *svagt* dominerende. Hvis f.eks. en $\mathcal{M}_5$-matrix har egenværdierne $7, -7, 5, 2$ og -1, er altså både 7 og -7 svagt dominerende.

I slutningen af Kapitel 1 omtalte vi såkaldte *matrixmodeller* af typen

$$\boldsymbol{x}_{t+1} = \boldsymbol{A}\boldsymbol{x}_t , \tag{1}$$

hvor $\boldsymbol{x}_t \in \mathbb{R}^n$ for $t = 0, 1, 2, 3, \cdots$, og hvor $\boldsymbol{A} \in \mathcal{M}_n$ er en given matrix. Ved at *iterere* (1), altså anvende den successivt, fås

$$\boldsymbol{x}_t = \boldsymbol{A}^t \boldsymbol{x}_0 . \tag{2}$$

Det har derfor interesse at kunne udtale sig om, hvordan *potenserne* $\boldsymbol{A}^t$ af en given $\mathcal{M}_n$-matrix $\boldsymbol{A}$ varierer for voksende t. Det viser sig, at begrebet 'dominerende egenværdi' er nyttigt i den sammenhæng.

SÆTNING 19. Antag, at $\boldsymbol{A} \in \mathcal{M}_n$ har egenværdierne $\lambda_1, \cdots, \lambda_n$, og at λ_1 er dominerende egenværdi, med tilhørende egenvektor $\boldsymbol{q}_1$. Der findes en matrix $\boldsymbol{C} \in \mathcal{M}_n$, således at

$$\lim_{t \to \infty} \left(\frac{1}{\lambda_1} \boldsymbol{A} \right)^t = \boldsymbol{C} . \tag{3}$$

Endvidere har $\boldsymbol{C}$ rang 1, idet dens søjler alle er multipla af $\boldsymbol{q}_1$.

BEVIS: Vi viser kun sætningen i det tilfælde, hvor A kan diagonaliseres til $\operatorname{diag}(\lambda_1, \cdots, \lambda_n) = D$, dvs. der findes en regulær $\mathcal{M}_n$-matrix Q (som vælges med q_1 i 1. søjle), således at $D = Q^{-1}AQ$. Ganger vi her fra venstre med Q og fra højre med Q^{-1}, får vi

$$A = QDQ^{-1}.$$

Sæt $D' = \frac{1}{\lambda_1}D = \operatorname{diag}(1, \frac{\lambda_2}{\lambda_1}, \cdots, \frac{\lambda_n}{\lambda_1})$. Da er $\frac{1}{\lambda_1}A = QD'Q^{-1}$, dvs

$$\begin{aligned}
\left(\frac{1}{\lambda_1}A\right)^t &= (QD'Q^{-1})(QD'Q^{-1})\cdots(QD'Q^{-1}) \qquad (t \text{ faktorer}) \\
&= Q\,(D')^t\,Q^{-1} \\
&= Q \operatorname{diag}\left(1, \left(\frac{\lambda_2}{\lambda_1}\right)^t, \cdots, \left(\frac{\lambda_n}{\lambda_1}\right)^t\right) Q^{-1}.
\end{aligned}$$

Da λ_1 er dominerende egenværdi, gælder $|\frac{\lambda_i}{\lambda_1}| < 1$ og følgelig $\left(\frac{\lambda_i}{\lambda_1}\right)^t \to 0$ for $t \to \infty$ $(i = 2, \cdots, n)$. Altså vil $(D')^t$ gå imod $D'' = \operatorname{diag}(1, 0, \cdots, 0)$ for $t \to \infty$, hvilket igen medfører, at $\left(\frac{1}{\lambda_1}A\right)^t$ går imod $QD''Q^{-1}$.

Lad 1. række i Q^{-1} være $r_1 = (r_{11}, r_{12}, \cdots, r_{1n})$. Tallene r_{1i} er ikke alle lig med 0, da det ville være i modstrid med, at Q er regulær. Vi minder også om, at Q's 1. søjle er $q_1 = (q_{11}, q_{21}, \cdots, q_{n1})^\mathsf{T}$. Det viser sig, at vi ikke får brug for de øvrige elementer i Q og Q^{-1}. Udregning giver:

$$\begin{aligned}
\lim_{t\to\infty}\left(\frac{1}{\lambda_1}A\right)^t &= \begin{pmatrix} q_{11} & * & \cdot\cdot & * \\ q_{21} & * & \cdot\cdot & * \\ & \vdots & & \\ q_{n1} & * & \cdot\cdot & * \end{pmatrix} \begin{pmatrix} 1 & 0 & \cdot\cdot & 0 \\ 0 & 0 & \cdot\cdot & 0 \\ & \vdots & & \\ 0 & 0 & \cdot\cdot & 0 \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} & \cdot\cdot & r_{1n} \\ * & * & \cdot\cdot & * \\ & \vdots & & \\ * & * & \cdot\cdot & * \end{pmatrix} \\
&= \begin{pmatrix} q_{11} & * & \cdot\cdot & * \\ q_{21} & * & \cdot\cdot & * \\ & \vdots & & \\ q_{n1} & * & \cdot\cdot & * \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} & \cdot\cdot & r_{1n} \\ 0 & 0 & \cdot\cdot & 0 \\ & \vdots & & \\ 0 & 0 & \cdot\cdot & 0 \end{pmatrix} \\
&= \begin{pmatrix} r_{11}q_{11} & r_{12}q_{11} & \cdots & r_{1n}q_{11} \\ r_{11}q_{21} & r_{12}q_{21} & \cdots & r_{1n}q_{21} \\ & \vdots & & \\ r_{11}q_{n1} & r_{12}q_{n1} & \cdots & r_{1n}q_{n1} \end{pmatrix} = C.
\end{aligned}$$

Af udtrykket for C fremgår, at den har den struktur, sætningen hævder: Alle søjler er multipla (og ikke alle med faktoren 0 !) af q_1, der er egenvektor for A hørende til λ_1. □

Sætning 19 har følgende vigtige konsekvens for matrixmodeller:

SÆTNING 20. Betragt modellen $x_{t+1} = Ax_t$, hvor A er diagonaliserbar, idet $Q^{-1}AQ = D = \operatorname{diag}(\lambda_1, \cdots, \lambda_n)$, med λ_1 som *dominerende* egenværdi. Lad q_1 være 1. søjle i Q og dermed en egenvektor for A hørende til λ_1. Lad endelig r_1 være 1. række i Q^{-1}. Da gælder

$$\lim_{t\to\infty} \frac{1}{\lambda_1^t} x_t = c\, q_1 , \qquad \text{hvor} \quad c = r_1 \cdot x_0 . \tag{4}$$

BEVIS: Ved brug af bl.a. formel (3) i Sætning 19 finder vi:

$$\lim_{t\to\infty} \frac{1}{\lambda_1^t} x_t = \lim_{t\to\infty} \left(\frac{1}{\lambda_1} A\right)^t x_0 = \begin{pmatrix} r_{11}q_{11} & r_{12}q_{11} & \cdots & r_{1n}q_{11} \\ r_{11}q_{21} & r_{12}q_{21} & \cdots & r_{1n}q_{21} \\ & & \vdots & \\ r_{11}q_{n1} & r_{12}q_{n1} & \cdots & r_{1n}q_{n1} \end{pmatrix} x_0$$

$$= x_{01}r_{11}q_1 + \cdots + x_{0n}r_{1n}q_1 = c\, q_1 ,$$

hvor c har den værdi, der angives i sætningen. □

Sætning 20 er lidt indviklet i sin formulering, og man foretrækker derfor somme tider at give resultatet på en mindre eksakt måde, der til gengæld skulle forenkle indholdet og gøre klart, hvad det betyder for anvendelserne.

Betragt modellen $x_{t+1} = Ax_t$, hvor A har λ_1 som dominerende egenværdi og q_1 som tilhørende egenvektor. Da gælder for store t:

$$x_t \approx c\, \lambda_1^t\, q_1 . \tag{5}$$

Det vil sige, at x_t i det lange løb bliver "næsten proportional" med q_1 og i hver koordinat vokser "næsten eksponentielt" med faktor λ_1.

ØVELSE 24. **Ikke fuldt antal egenværdier.**

Betragt matricen $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & a \\ 0 & -a & 0 \end{pmatrix}$, hvor $a \in \mathbb{R} \setminus \{0\}$.

(1) Vis, at A har $\lambda_1 = 1$ som eneste egenværdi. Betingelserne i Sætning 19-20 er altså *ikke* opfyldt.

(2) Udregn potenserne A^2, A^3 og A^4. Vis, at $\lim_{t \to \infty} A^t = \text{diag}(1,0,0)$, når $|a| < 1$, mens A^t *divergerer* for $t \to \infty$, når $|a| \geqq 1$. Heri ligger, at det afhænger af størrelsen af a, om A opfører sig som i Sætning 19-20 eller ikke.

ØVELSE 25. **Svag dominans.** Betragt matricen $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, der iflg. Eksempel 12, side 179 har $\lambda_1 = 1$ som eneste egenværdi.

(1) Vis ved induktion, at $A^t = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \quad (t = 1, 2, 3, \cdots)$.

Det betyder, at $(\frac{1}{\lambda_1} A)^t$ divergerer for $t \to \infty$.

(2) Undersøg ligeledes spørgsmålet om konvergens eller divergens af A^t, når 1) $A = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}$, 2) $A = E = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

I kapitlets sidste afsnit skal vi se på en type matricer A, som er aktuelle ved anvendelserne, og hvor konklusionerne i Sætning 19-20 holder, uanset om der er "for få" egenværdier.

9. Ikke-negative matricer. Modeleksempler.

I den videregående spektralteori spiller matricer med *ikke-negative* elementer en fremtrædende rolle, også fordi de optræder i forbindelse med mange matrixmodeller og andre anvendelser. Pionerarbejde inden for området af O. Perron (1907) og G. Frobenius (1908-12) har gjort, at man taler om *Perron-Frobenius teori.* Vi anfører et par af hovedresultaterne i tillempet form.

DEFINITION 11. **Ikke-negativ matrix. Positiv matrix.**
En matrix $A(\neq O)$ kaldes *ikke-negativ*, hvis dens elementer alle opfylder $a_{ij} \geq 0$, og man skriver da $A > O$. En matrix A kaldes *positiv*, hvis dens elementer alle opfylder $a_{ij} > 0$, og man skriver da $A \gg O$. Betegnelserne ikke-negativ og positiv bruges også om vektorer.

SÆTNING 21. **Perron-Frobenius sætninger.**

PF 1. Hvis $A \in \mathcal{M}_n$ er positiv, har den en positiv egenvektor q_1 og en tilhørende positiv egenværdi λ_1, som er strengt "virtuelt dominerende" i følgende betydning: Ikke blot vil enhver eventuel anden egenværdi λ' opfylde $|\lambda'| < \lambda_1$, men konklusionen i Sætning 19 holder, dvs der gælder $\lim_{t\to\infty}(\frac{1}{\lambda_1}A)^t = C$, hvor C's søjler alle er multipla af q_1.

PF 2. Hvis $A \in \mathcal{M}_n$ er ikke-negativ, har den en ikke-negativ egenvektor q_1 og en tilhørende ikke-negativ egenværdi λ_1, som er svagt "virtuelt dominerende", dvs enhver eventuel anden egenværdi λ' opfylder $|\lambda'| \leq \lambda_1$.

PF 3. Hvis $A \in \mathcal{M}_n$ er ikke-negativ, og hvis der findes et naturligt tal m således, at matricen A^m er positiv, holder de samme konklusioner som for en positiv matrix, jfr. PF 1.

BEVIS: UDELADES. □

På lignende måde som i foregående afsnit ses, at når A er positiv, eller blot "quasi-positiv" som i PF 3, så gælder for modellen $x_{t+1} = Ax_t$:

$$x_t \approx c\,\lambda_1^t\, q_1 \qquad \text{for store } t\,. \tag{1}$$

EKSEMPEL 24. For matricen $A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$ fås de første potenser til

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

$$A^4 = \begin{pmatrix} 1 & 2 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \qquad A^5 = \begin{pmatrix} 2 & 2 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix}.$$

Det ses, at $\boldsymbol{A}$ opfylder kravet i PF 3 med $m = 5$, dvs $\boldsymbol{A}$ har en "virtuelt dominerende egenværdi" λ_1. Reduktionsdeterminanten findes til

$$\det(\boldsymbol{A} - \lambda \boldsymbol{E}) = \begin{vmatrix} -\lambda & 1 & 1 \\ 1 & -\lambda & 0 \\ 0 & 1 & -\lambda \end{vmatrix} = -\lambda^3 + \lambda + 1 = R(\lambda) \, .$$

Af $R(1) = 1$, $R(2) = -5$ ses, at R har et nulpunkt i intervallet $1 < \lambda < 2$. Da $R'(\lambda) = -3\lambda^2 + 1$ er negativ for $\lambda > 1$, aftager $R(\lambda)$ i dette område og har altså kun ét nulpunkt der, dvs førnævnte nulpunkt er det λ_1, som PF 1 taler om. Tilnærmet bestemmelse, f.eks. ved Newton's metode, giver $\underline{\lambda_1 = 1.3247}$. (I øvrigt har R ikke andre reelle rødder). En tilhørende egenvektor $\boldsymbol{q}_1$ bestemmes derefter som sædvanlig ved at løse $\boldsymbol{A}\boldsymbol{x} = 1.3247\,\boldsymbol{x}$, der skrives ud i koordinater som

$$-1.3247\,x_1 + x_2 + x_3 = 0 \, , \qquad x_1 - 1.3247\,x_2 = 0 \, , \qquad x_2 - 1.3247\,x_3 = 0 \, .$$

Her kan vi med sikkerhed stryge en ligning(!), og vi vælger at stryge den første. Endvidere vælger vi $x_3 = 1$ (løsningsvektoren er jo bestemt på nær en konstant faktor). Indsat i 3. ligning giver det $x_2 = 1.3247$, der videre i 2. ligning giver $x_1 = 1.3247^2 = 1.7548$. Vi har dermed fundet egenvektoren $\underline{\boldsymbol{q}_1 = (1.7548\,,\ 1.3247\,,\ 1)^{\top}}$.

ØVELSE 26. Lad $\boldsymbol{A}$ være matricen i det foregående eksempel. Udregn potenserne $\boldsymbol{A}^{10}$ og $\boldsymbol{A}^{20}$. [Vink: Kvadrér $\boldsymbol{A}^5$ og dernæst $\boldsymbol{A}^{10}$!] Synes du, at søjlerne i disse potenser begynder at udvise proportionalitet med $\boldsymbol{q}_1$, sådan som de skal iflg. PF 3 ?

EKSEMPEL 25. For matricen $\boldsymbol{A} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$ fås de første potenser til

$$\boldsymbol{A}^2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} = \boldsymbol{A}^{\top}, \qquad \boldsymbol{A}^3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \boldsymbol{E},$$
$$\boldsymbol{A}^4 = \boldsymbol{A}^3\boldsymbol{A} = \boldsymbol{A}\, , \qquad \boldsymbol{A}^5 = \boldsymbol{A}^3\boldsymbol{A}^2 = \boldsymbol{A}^2 = \boldsymbol{A}^{\top}.$$

Potenserne danner altså den divergente *cykliske følge*

$$\boldsymbol{A},\ \boldsymbol{A}^{\top},\ \boldsymbol{E},\ \boldsymbol{A},\ \boldsymbol{A}^{\top},\ \boldsymbol{E},\ \cdots \, . \tag{2}$$

Hvad udtaler Sætning 21 om $\boldsymbol{A}$? Da $\boldsymbol{A} \geq \boldsymbol{O}$, er vi i tilfælde PF 2, men (2) viser, at vi *ikke* er i tilfælde PF 3. Egenværdierne findes:

$$\det(\boldsymbol{A} - \lambda \boldsymbol{E}) = \begin{vmatrix} -\lambda & 1 & 0 \\ 0 & -\lambda & 1 \\ 1 & 0 & -\lambda \end{vmatrix} = 1 - \lambda^3 = (1 - \lambda)(\lambda^2 + \lambda + 1) \, .$$

Her er sidste faktor et andengradspolynomium uden nulpunkter, dvs der er kun én egenværdi, $\lambda_1 = 1$. En tilhørende egenvektor viser sig i øvrigt at være $q_1 = (1,1,1)^{\mathsf{T}}$. Men det er klart af (2), at potenserne af A *ikke* nærmer sig søjlevis proportionalitet med q_1. Eksemplet viser dermed, at der ikke uden videre kan slækkes på forudsætningen i PF 3.

EKSEMPEL 26. For matricen $A = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ fås $A^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ og derefter $A^3 = O = A^4 = A^5 = \cdots$. Endvidere er $\det(A - \lambda E) = -\lambda^3$, dvs der er tre sammenfaldende egenværdier $\lambda_1 = \lambda_2 = \lambda_3 = 0$, selv om $A \neq O$. Udtrykket "ikke-negativ egenværdi" i PF 2 kan altså heller ikke uden videre skærpes. En tilhørende egenvektor viser sig i øvrigt at være $q_1 = e_1 = (1,0,0)^{\mathsf{T}}$. Men der er ikke andre egenvektorer end den og dens multipla, dvs A kan ikke diagonaliseres.

ØVELSE 27. (1) Kontrollér de egenvektorer q_1, der er anført for matricerne i de to foregående eksempler.

(2) Matricen A i Eksempel 25 kan opfattes som en *permutationsmatrix*, jfr. Eksempel 14, side 181, idet den svarer til den *cykliske* permutation $(1,2,3) \to (2,3,1)$, der netop er brugt som illustration side 181. Brug det til at forklare det resultat om følgen af A's potenser, som vi fandt i Eksempel 25.

Vi skal til slut omtale et par eksempler på matrixmodeller af interesse for bl.a. biologiske anvendelser. Der er, navnlig i det andet tilfælde, tale om generelle versioner af to eksempler i sidste afsnit af Kapitel 1.

EKSEMPEL 27. **Overgangsmatrix.** (Transitionsmatrix.)

Til tidspunkterne $t = 0,1,2,3,\cdots$ betragtes en mængde, hvis elementer ved hvert tidspunkt er i én af n tilstande $T_1, T_2, \cdots, T_n$. Antag, at sandsynligheden for overgang fra tilstand T_j til tilstand T_i i løbet af en tidsenhed er p_{ij}. Det forudsættes, at denne sandsynlighed for alle i, j er uforandret i tiden og uafhængig af elementets 'forhistorie'. Ingen elementer forsvinder fra mængden, og ingen nye kommer til.

Lad x_{ti} være frekvensen (den relative hyppighed, "brøkdelen") af elementer i tilstand T_i til tidspunkt t $(i = 1, \cdots, n)$. Vektoren $x_t = (x_{t1}, \cdots, x_{tn})^{\mathsf{T}}$ har således koordinatsummen 1. Idet antallet af elementer i mængden antages at være meget stort, vil frekvensen af elementer, som mellem tidspunkter t og $t+1$ går fra tilstand T_j til tilstand T_i, være $p_{ij}x_{tj}$. Summeres disse frekvenser over j, fås frekvensen af *samtlige* elementer, der går over til den i'te tilstand, altså $x_{t+1,i}$. Dermed har vi udledt

$$x_{t+1,i} = p_{i1}x_{t1} + \cdots + p_{in}x_{tn} = \sum_{j=1}^{n} p_{ij}x_{tj} \tag{3}$$

$(i = 1, \cdots, n)$. Disse n ligninger kan sammenstilles til matrixligningen

$$\boldsymbol{x}_{t+1} = \boldsymbol{P}\boldsymbol{x}_t\,, \tag{4}$$

hvor $\boldsymbol{P} = (p_{ij})$ betegnes som processens *overgangsmatrix* eller *transitionsmatrix*. Den er *ikke-negativ* og har dén egenskab, at *elementsummen i hver søjle er* 1:

$$\sum_{i=1}^{n} p_{ij} = 1 \qquad (j = 1, \cdots, n)\,, \tag{5}$$

fordi denne sum er sandsynligheden for, at et element i tilstand T_j går til tilstand T_1 *eller* til tilstand T_2 *eller* $\cdots$ til tilstand T_n.

Vha (5) kan vi bevise, at $\boldsymbol{P}$ har egenværdien $\lambda = 1$. Adderer vi nemlig *rækkevektorerne* i matricen $\boldsymbol{P} - 1 \cdot \boldsymbol{E}$, får vi

$$\left(\sum_{i=1}^{n} p_{i1} - 1\,, \quad \sum_{i=1}^{n} p_{i2} - 1\,, \quad ,\cdots, \quad \sum_{i=1}^{n} p_{in} - 1 \right) = \boldsymbol{o}\ (!)\,,$$

og heraf følger, at rækkerne i samme matrix er lineært afhængige, med andre ord: $\det(\boldsymbol{P} - 1 \cdot \boldsymbol{E}) = 0$. Det kan vises, at 1 er dén - strengt eller svagt - dominerende egenværdi, som $\boldsymbol{P}$ skal have iflg. Sætning 21.

Hvorfor har egenværdien $\lambda_1 = 1$ interesse for evt. brug af denne model? Dette har at gøre med begrebet *ligevægt* for det system, som modellen beskriver. Systemet siges at være i en ligevægtstilstand, når frekvenserne for de forskellige tilstande holder sig uforandrede, dvs når $\boldsymbol{x}_{t+1} = \boldsymbol{x}_t$. Da må $\boldsymbol{x}_t$ have en sådan værdi $\boldsymbol{x}^*$, at

$$\boldsymbol{P}\,\boldsymbol{x}^* = \boldsymbol{x}^*\,, \tag{6}$$

hvilket netop udtrykker, at $\boldsymbol{x}^*$ er en egenvektor hørende til egenværdien 1.

Ligevægten $\boldsymbol{x}^*$ skal ud over (6) opfylde $\boldsymbol{x}_1^* + \boldsymbol{x}_2^* + \cdots + \boldsymbol{x}_n^* = 1$, idet dens koordinater også er frekvenser, der skal summere til 1. I almindelighed afhænger det af $\boldsymbol{P}$'s udseende, om $\boldsymbol{x}^*$ derved er entydigt bestemt, eller om der er flere mulige ligevægtstilstande. Den simple epidemimodel i Eksempel 30, side 75-77 er et eksempel på det sidste, idet enhver kombination af døde og helbredte (men ingen raske eller syge) ses at være en ligevægt i den forstand, der blev defineret ovenfor.

Den nærmere undersøgelse af overgangsmatricer, deres ligevægtsforhold og andre egenskaber tilhører den teoretiske statistik og sandsynlighedsregning, hvor man dog i reglen fortolker $\boldsymbol{x}_t$ på en lidt anden måde, nemlig som en vektor af sandsynligheder for det enkelte objekts tilstand, ikke af relative hyppigheder. Processen af overgange, ofte generaliseret til dét tilfælde, at der kan optræde uendelig mange tilstande, kaldes en *Markov-kæde* eller *Markov-proces*. Det kan i øvrigt nævnes, at man i teorien for Markov-kæder traditionelt plejer at opskrive vektor/matrix ligningerne *transponeret* i forhold til det, vi har gjort ovenfor, dvs man skriver

$$\hat{\boldsymbol{x}}_{t+1} = \hat{\boldsymbol{x}}_t \, \hat{\boldsymbol{P}} \ ,$$

hvor $\hat{\boldsymbol{x}}_t$ er en *række*vektor, og $\hat{\boldsymbol{P}}$'s ij'te element angiver sandsynligheden for overgang fra T_i til T_j, dvs i forhold til udgaven ovenfor er $\hat{\boldsymbol{x}}_t = \boldsymbol{x}_t^{\mathsf{T}}$ og $\hat{\boldsymbol{P}} = \boldsymbol{P}^{\mathsf{T}}$.

ØVELSE 28. Hvis alle elementer i overgangsmatricen $\boldsymbol{P}$ (se Eksempel 27) er *positive*, har systemet iflg. Sætning 21, PF 1 en entydig, *positiv* ligevægt $\boldsymbol{x}^*$. Fortolk i lys heraf edb-udskriften

```
P = ((0.81, 0.19), (0.58, 0.42))
P**T,  T = 1,2,3,4,5,6,10,20
···((0.81, 0.19), (0.58, 0.42))
···((0.7663, 0.2337), (0.7134, 0.2866))
···((0.7562, 0.2438), (0.7441, 0.2559))
···((0.7539, 0.2461), (0.7511, 0.2489))
···((0.7534, 0.2466), (0.7528, 0.2472))
···((0.7533, 0.2467), (0.7531, 0.2469))
···((0.7532, 0.2468), (0.7532, 0.2468))
···((0.7532, 0.2468), (0.7532, 0.2468))
```

som en fremskrivning af en proces med to tilstande og med overgangsmatrix

$$\boldsymbol{P} = \begin{pmatrix} 0.81 & 0.58 \\ 0.19 & 0.42 \end{pmatrix} .$$

Matrixpotensernes søjler, der kan aflæses i edb-udskriften som hhv 1. og 2. sæt i de beregnede potenser $\boldsymbol{P}^t$, angiver frekvensfordelingen efter t overgange, når alle elementer starter i første, hhv. i anden tilstand. Hvad sker i begge tilfælde med frekvensfordelingen for $t \to \infty$? Stemmer det med Eksempel 27?

Bestem den egenvektor $\boldsymbol{x}^*$ for $\boldsymbol{P}$, der hører til egenværdien 1 og desuden opfylder $x_1^* + x_2^* = 1$. Stemmer den med beregningerne ovenfor?

forvirrende overskrift

Ligevægt og stabilitet

Definition B.9.1 (n × n matrix)

(a) En $n \times n$ *matrix* er et talskema af formen

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

med n rækker og n søjler. Tallet a_{ij}, der står i den i'te række og j'te søjle, kaldes det *ij'te element* i matricen.

(b) Matricen $\mathbf{M}$ ganges med en vektor $\mathbf{v} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ i $\mathbb{R}^n$ på følgende måde:

$$\mathbf{Mv} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n \end{pmatrix}.$$

Faktisk er det denne simple definition, som tager en stor del af æren for nytten af matricer, idet den gør det muligt at opskrive komplicerede sammenhænge på kort og overskuelig form (se anvendelseseksemplerne).

Eksempel B.9.1 Matricen

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

er en 3×3 matrix. Produktet af $\mathbf{M}$ og vektoren $\mathbf{v} = \begin{pmatrix} 5 \\ 6 \\ 7 \end{pmatrix}$ er

$$\mathbf{Mv} = \begin{pmatrix} 2 & 0 & -1 \\ 1 & 1 & 0 \\ 1 & -3 & 1 \end{pmatrix} \begin{pmatrix} 5 \\ 6 \\ 7 \end{pmatrix} = \begin{pmatrix} 2\cdot 5 + 0\cdot 6 + (-1)\cdot 7 \\ 1\cdot 5 + 1\cdot 6 + 0\cdot 7 \\ 1\cdot 5 + (-3)\cdot 6 + 1\cdot 7 \end{pmatrix} = \begin{pmatrix} 3 \\ 11 \\ -6 \end{pmatrix}.$$

Bemærkning B.9.1 Fremskrivning af en begyndelsesvektor $\mathbf{v}_0$ i $\mathbb{R}^n$ vha. en $n \times n$ matrix $\mathbf{M}$ foregår som for 2×2 matricer (se Bemærkning B.2.2): Vektorerne $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \ldots$ defineres ved

$$\mathbf{v}_1 = \mathbf{Mv}_0, \qquad \mathbf{v}_2 = \mathbf{Mv}_1, \qquad \mathbf{v}_3 = \mathbf{Mv}_2, \qquad \ldots$$

dvs. $\mathbf{v}_{t+1} = \mathbf{Mv}_t$. Vektoren $\mathbf{v}_t$ kan også udtrykkes som $\mathbf{v}_t = \mathbf{M}^t\mathbf{v}_0$.

LA 207

Den følgende definition viser, at vi regner med $n \times n$ matricer stort set som med 2×2 matricer. Vi skriver ikke formlen for produktet ud i detaljer, dels fordi den bliver ret kompliceret, og dels fordi man i konkrete tilfælde som regel vil enten udføre beregningerne vha. computer eller benytte sig af fremgangsmåden med skemaet

	M
N	NM

(se Eksempel B.9.2).

Definition B.9.2 (Regning med n × n matricer) Lad M og N være to $n \times n$ matricer

$$\mathbf{M} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \quad \text{og} \quad \mathbf{N} = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{pmatrix}$$

(a) Summen $\mathbf{M} + \mathbf{N}$ defineres ved at lægge sammen plads for plads:

$$\mathbf{M} + \mathbf{N} = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \cdots & a_{2n} + b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} + b_{n1} & a_{n2} + b_{n2} & \cdots & a_{nn} + b_{nn} \end{pmatrix}.$$

Tilsvarende defineres differensen $\mathbf{M} - \mathbf{N}$.

(b) For et tal k defineres $k\mathbf{M}$ ved at gange k ind på alle pladser:

$$k\mathbf{M} = \begin{pmatrix} ka_{11} & ka_{12} & \cdots & ka_{1n} \\ ka_{21} & ka_{22} & \cdots & ka_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ ka_{n1} & ka_{n2} & \cdots & ka_{nn} \end{pmatrix}$$

(c) *Matrixproduktet* NM defineres vha. skemaet

	M
N	NM

Eksempel B.9.2 Med

$$\mathbf{M} = \begin{pmatrix} 2 & 0 & -1 \\ 1 & 1 & 0 \\ 1 & -3 & 1 \end{pmatrix} \quad \text{og} \quad \mathbf{N} = \begin{pmatrix} 0 & 5 & 1 \\ 2 & -2 & 3 \\ 7 & 0 & -1 \end{pmatrix}$$

har vi

$$\mathbf{M} + \mathbf{N} = \begin{pmatrix} 2 & 5 & 0 \\ 3 & -1 & 3 \\ 8 & -3 & 0 \end{pmatrix}$$

LA 208

og

$$7\mathbf{M} = \begin{pmatrix} 14 & 0 & -7 \\ 7 & 7 & 0 \\ 7 & -21 & 7 \end{pmatrix}.$$

For at udregne produktet **NM** opstilles skemaet

	M
N	**NM**

dvs.

			2	0	-1
			1	1	0
			1	-3	1
0	5	1	$0 \cdot 2 + 5 \cdot 1 + 1 \cdot 1$	$0 \cdot 0 + 5 \cdot 1 + 1 \cdot (-3)$	$0 \cdot (-1) + 5 \cdot 0 + 1 \cdot 1$
2	-2	3	$2 \cdot 2 + (-2) \cdot 1 + 3 \cdot 1$	$2 \cdot 0 + (-2) \cdot 1 + 3 \cdot (-3)$	$2 \cdot (-1) + (-2) \cdot 0 + 3 \cdot 1$
7	0	-1	$7 \cdot 2 + 0 \cdot 1 + (-1) \cdot 1$	$7 \cdot 0 + 0 \cdot 1 + (-1) \cdot (-3)$	$7 \cdot (-1) + 0 \cdot 0 + (-1) \cdot 1$

Heraf ses det, at

$$\mathbf{NM} = \begin{pmatrix} 6 & 2 & 1 \\ 5 & -11 & 1 \\ 13 & 3 & -8 \end{pmatrix}.$$

Vi generaliserer nu begreberne enhedsmatrix og invers matrix, som vi kender fra 2×2 matricer.

Definition B.9.3 (Enhedsmatricen) $n \times n$ matricen

$$\mathbf{E} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

kaldes $n \times n$ *enhedsmatricen.*

Sætning B.9.1 (Multiplikation med enhedsmatricen) For enhver $n \times n$ matrix **M** gælder der

$$\mathbf{ME} = \mathbf{EM} = \mathbf{M}.$$

Definition B.9.4 (Invers matrix) $n \times n$ matricen **M** har en invers matrix, hvis der findes en $n \times n$ matrix **X** sådan at

$$\mathbf{MX} = \mathbf{XM} = \mathbf{E}.$$

I så fald kaldes **X** den *inverse matrix* til **M** og betegnes $\mathbf{M}^{-1}$.

For 2×2 matricer har vi set, dels hvordan man udregner determinanten, og dels at hvis determinanten er forskellig fra nul, så har vi en formel for den inverse matrix. Det er faktisk muligt at generalisere dette til $n \times n$ matricer: Man kan definere determinanten af en $n \times n$ matrix og finde en formel for den inverse (når determinanten er forskellig fra nul). I disse noter vil vi hverken give den generelle definition af determinanten eller formlen for den inverse, men vil nøjes med definitionen på determinanten af en 3×3 matrix.

Definition B.9.5 (Determinant af 3×3 matrix) For en 3×3 matrix

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

defineres *determinanten* $\det \mathbf{M}$ ved udtrykket

$$\begin{aligned} \det \mathbf{M} &= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \\ &= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{31}a_{22}a_{13} - a_{32}a_{23}a_{11} - a_{33}a_{21}a_{12} \\ &= \left| \begin{array}{ccc|cc} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{array} \right. = \searrow + \searrow + \searrow - \nearrow - \nearrow - \nearrow, \end{aligned}$$

hvor den sidste ligning er en praktisk huskeregel: Man tilføjer de to første søjler en gang til og danner de seks "skrå" produkter, hvoraf de tre med $\searrow$ regnes med "+" og de tre med $\nearrow$ regnes med "−".

Eksempel B.9.3 Med

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

er

$$\begin{aligned} \det \mathbf{M} &= \begin{vmatrix} 2 & -1 & 0 \\ 3 & 0 & 2 \\ -1 & 1 & 2 \end{vmatrix} \\ &= \left| \begin{array}{ccc|cc} 2 & -1 & 0 & 2 & -1 \\ 3 & 0 & 2 & 3 & 0 \\ -1 & 1 & 2 & -1 & 1 \end{array} \right. \\ &= 2 \cdot 0 \cdot 2 + (-1) \cdot 2 \cdot (-1) + 0 \cdot 3 \cdot 1 - (-1) \cdot 0 \cdot 0 - 1 \cdot 2 \cdot 2 - 2 \cdot 3 \cdot (-1) \\ &= 4. \end{aligned}$$

Vi har set, at udtrykket for en 2×2-determinant består af to led (ét med "+" og ét med "−") samt at udtrykket for en 3×3-determinant består af seks led (tre med "+" og tre med "−"). Generelt kan man vise, at udtrykket for en $n \times n$-determinant består af $n!$ led (halvdelen med "+" og halvdelen med "−"). Der er derfor næsten altid et uoverskueligt regnearbejde forbundet med at skulle finde determinanten af, for ikke at tale om den inverse til 4×4 og større matricer. Det kan heldigvis gøres ved hjælp af computerprogrammer som f.eks. R.

Vi vil nu se, at Sætning B.3.3 om den inverse af et produkt også gælder for $n \times n$ matricer. (Beviset for denne sætning er i øvrigt præcist det samme som for 2×2 matricer.)

Sætning B.9.2 (Den inverse af et produkt) Når $n \times n$ matricerne $\mathbf{M}$ og $\mathbf{N}$ begge har en invers matrix, så har produktet $\mathbf{NM}$ ligeledes en invers matrix givet ved

$$(\mathbf{NM})^{-1} = \mathbf{M}^{-1}\mathbf{N}^{-1}.$$

I Afsnit B.1 angav vi tre metoder til løsning af to ligninger med to ubekendte: substitutionsmetoden, lige store koefficienters metode og matrixinvertering. De to første metoder bliver ret besværlige at anvende, hvis man benytter dem til at løse n ligninger med n ubekendte (omend der findes en "smart" udgave af lige store koefficienters metode). Løsning af ligninger ved matrixinvertering fungerer i princippet på samme måde som for to ligninger med to ubekendte:

Sætning B.9.3 (Løsning af ligninger ved matrixinvertering) Vi betragter et ligningssystem med n ligninger og n ubekendte på formen

$$\mathbf{Mv} = \mathbf{p},$$

hvor $\mathbf{M}$ er en given $n \times n$ matrix og $\mathbf{p}$ er en given vektor i $\mathbb{R}^n$. Hvis matricen $\mathbf{M}$ har en invers matrix, så kan løsningen $\mathbf{v}$ til ligningssystemet findes som

$$\mathbf{v} = \mathbf{M}^{-1}\mathbf{p}.$$

Vi har selvfølgelig stadigvæk problemet med at bestemme den inverse matrix $\mathbf{M}^{-1}$, men hvis dette f.eks. gøres vha. R, så følger løsningen til ligningssystemet let af Sætning B.9.3.

Ligevægte spiller en væsentlig rolle for mange anvendelser af $n \times n$ matricer (se f.eks. Anvendelseseksempel B.6), så begrebet er mindst lige så relevant for $n \times n$ matricer som for 2×2 matricer.

Definition B.9.6 (Ligevægt) Givet en $n \times n$ matrix $\mathbf{M}$. En vektor $\mathbf{v}^*$ i $\mathbb{R}^n$ siges at være en *ligevægt* for matricen $\mathbf{M}$, hvis der gælder

$$\mathbf{Mv}^* = \mathbf{v}^*.$$

Bemærkning B.9.2 Idet

$$\mathbf{o} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

er *nulvektoren* i $\mathbb{R}^n$, gælder der (som for 2×2 matricer) for enhver $n \times n$ matrix $\mathbf{M}$, at

$$\mathbf{M}\mathbf{o} = \mathbf{o},$$

så nul-vektoren er altid en ligevægt. Endvidere kan man som i Bemærkning B.4.1 vise, at $\mathbf{M}$ ikke har andre ligevægte, hvis matricen $\mathbf{M} - \mathbf{E}$ har en invers matrix.

Vi henviser til Anvendelseseksempel B.6 for et konkret eksempel på bestemmelse af ligevægte.

Som for 2×2 matricer viser det sig, at $n \times n$ overgangsmatricer har ligevægte $\neq \mathbf{o}$.

Definition B.9.7 (Overgangsmatrix) En $n \times n$ matrix

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

kaldes en *overgangsmatrix,* hvis $a_{ij} \geq 0$ for alle i, j, og der endvidere gælder, at "hver søjle lagt sammen giver 1", dvs. hvis

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

LA 212

Bemærkning B.9.3 Med

$$\mathbf{v}_t = \begin{pmatrix} x_{1t} \\ x_{2t} \\ \vdots \\ x_{nt} \end{pmatrix}$$

får vi

$$\mathbf{v}_{t+1} = \mathbf{M}\mathbf{v}_t = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_{1t} \\ x_{2t} \\ \vdots \\ x_{nt} \end{pmatrix}$$

$$= \begin{pmatrix} a_{11}x_{1t} + a_{12}x_{2t} + \cdots + a_{1n}x_{nt} \\ a_{21}x_{1t} + a_{22}x_{2t} + \cdots + a_{2n}x_{nt} \\ \vdots \\ a_{n1}x_{1t} + a_{n2}x_{2t} + \cdots + a_{nn}x_{nt} \end{pmatrix}.$$

Hvis $\mathbf{v}_t$ (som i Anvendelseseksempel B.6) angiver, hvordan en "population" er fordelt i n grupper, kan a_{ij} fortolkes som sandsynligheden for, at et individ i gruppe j overgår til gruppe i. For fastholdt j skal summen af disse sandsynligheder være 1 (individet er jo at finde i netop én af grupperne efter overgangen), dvs. $a_{1j} + a_{2j} + \cdots + a_{nj} = 1$, hvilket netop er kravet i Definition B.9.7.

Sætning B.9.4 (Ligevægt for overgangsmatrix) Enhver $n \times n$ overgangsmatrix har en ligevægt $\neq \mathbf{o}$.

Bevis for Sætning B.9.4. Udelades. Det er en del sværere end det tilsvarende bevis for 2×2 matricer (Sætning B.4.1). □

Vi henviser til Anvendelseseksempel B.6 for et konkret eksempel på bestemmelse af ligevægte.

Bemærkning B.9.4 I Afsnit B.6 knyttede vi til en given 2×2 matrix

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

en lineær afbildning f fra $\mathbb{R}^2$ ind i $\mathbb{R}^2$ ved udtrykket

$$f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax + cy \\ bx + dy \end{pmatrix}.$$

LA 213

Tilsvarende kan man til en given $n \times n$ matrix $\mathbf{M}$ knytte en lineær afbildning f fra $\mathbb{R}^n$ ind i $\mathbb{R}^n$ ved udtrykket

$$f(\mathbf{v}) = \mathbf{M}\mathbf{v}.$$

I tilfældet $n = 3$ kan man vise et resultat, der svarer til Sætning B.6.1 om arealforhold: En mængde af punkter, F, i $\mathbf{R}^3$ (som f.eks. punkterne i en terning), bliver ved f afbildet over i en ny mængde af punkter, $f(F)$, i $\mathbf{R}^3$ bestående af punkterne $f\binom{x}{y}$, hvor $\binom{x}{y}$ ligger i F. Hvis volumenet af mængden F er V og volumenet af billedmængden $f(F)$ er V', kan man vise, at der gælder

$$V' = |\det \mathbf{M}| \cdot V.$$

Vi vil nu generalisere begreberne affin afbildning samt ligevægt for en sådan til også at gælde for $n \times n$ matricer.

Definition B.9.8 (Affin afbildning) For en given $n \times n$ matrix $\mathbf{M}$ og en given vektor $\mathbf{q}$ i $\mathbb{R}^n$ definerer vi den til $\mathbf{M}$ og $\mathbf{q}$ hørende *affine afbildning* f fra $\mathbb{R}^n$ ind i $\mathbb{R}^n$ ved udtrykket

$$f(\mathbf{v}) = \mathbf{M}\mathbf{v} + \mathbf{q}.$$

Definition B.9.9 (Ligevægt for affin afbildning) Lad f være den affine afbildning

$$f(\mathbf{v}) = \mathbf{M}\mathbf{v} + \mathbf{q}$$

givet ved en $n \times n$ matrix $\mathbf{M}$ og en vektor $\mathbf{q}$ i $\mathbb{R}^n$. En vektor $\mathbf{v}^*$ i $\mathbb{R}^n$ siges at være en *ligevægt* for f, hvis der gælder

$$f(\mathbf{v}^*) = \mathbf{v}^* \qquad \text{dvs.} \qquad \mathbf{M}\mathbf{v}^* + \mathbf{q} = \mathbf{v}^*.$$

Sætning B.7.1 om eksistens og bestemmelse af ligevægte for en affin afbildning kan direkte overføres til at gælde for $n \times n$ matricer.

Sætning B.9.5 (Ligevægt for affin afbildning) Lad f være den affine afbildning

$$f(\mathbf{v}) = \mathbf{M}\mathbf{v} + \mathbf{q}.$$

Hvis matricen $\mathbf{M} - \mathbf{E}$ har en invers matrix, så har f netop én ligevægt $\mathbf{v}^*$ givet ved

$$\mathbf{v}^* = -(\mathbf{M} - \mathbf{E})^{-1}\mathbf{q}.$$

LA 214

I Afsnit B.8 så vi, at egenværdier og -vektorer er nyttige, når man f.eks. ønsker at fremskrive en population. Vi vi nu generalisere disse begreber til $n \times n$ matricer.

Definition B.9.10 (Egenvektorer og -værdier) Lad $\mathbf{M}$ være en $n \times n$ matrix. Hvis der om en vektor $\mathbf{v} \neq \mathbf{o}$ i $\mathbb{R}^n$ og et tal λ gælder

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

så siges $\mathbf{v}$ at være en *egenvektor* for $\mathbf{M}$ med tilhørende *egenværdi* λ.

Bemærkning B.9.5 Som for 2×2 matricer (Sætning B.8.1) kan egenværdierne λ for en $n \times n$ matrix $\mathbf{M}$ *i princippet* bestemmes som løsningerne til ligningen

$$\det(\mathbf{M} - \lambda \mathbf{E}) = 0.$$

I praksis er det dog besværligt at opskrive $\det(\mathbf{M} - \lambda\mathbf{E})$ udtrykt ved λ (det bliver et n'te-gradspolynomium i variablen λ), og derefter bestemme egenværdierne for $\mathbf{M}$ som de (højst n) rødder i dette polynomium. I disse noter vil vi derfor ikke komme yderligere ind på bestemmelse af egenværdier for $n \times n$ matricer vha. ligningen $\det(\mathbf{M} - \lambda\mathbf{E}) = 0$.

For en $n \times n$ matrix $\mathbf{M}$ og en begyndelssvektor $\mathbf{v}_0$ i $\mathbb{R}^n$ definerer vi fremskrivningerne $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \ldots$ ved

$$\mathbf{v}_1 = \mathbf{M}\mathbf{v}_0, \qquad \mathbf{v}_2 = \mathbf{M}\mathbf{v}_1, \qquad \mathbf{v}_3 = \mathbf{M}\mathbf{v}_2, \qquad \ldots$$

dvs. $\mathbf{v}_{t+1} = \mathbf{M}\mathbf{v}_t$. (Se også Bemærkning B.9.1.) Vektorerne $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \ldots$ kan også beregnes som

$$\mathbf{v}_t = \mathbf{M}^t \mathbf{v}_0.$$

Den tilnærmelsesvise opførsel af $\mathbf{v}_t$ for store værdier af t beskrives af følgende generalisation af Sætning B.8.2.

Sætning B.9.6 (Fremskrivning med matrix) Lad $\mathbf{M}$ være en $n \times n$ matrix med n forskellige egenværdier $\lambda_1, \lambda_2, \ldots, \lambda_n$ og antag at $|\lambda_1| > |\lambda_k|$ for $k = 2, \ldots, n$. Lad $\mathbf{q}_1$ være en egenvektor for $\mathbf{M}$ hørende til egenværdien λ_1. For en vikårlig vektor $\mathbf{v}_0$ i $\mathbb{R}^n$ sætter vi $\mathbf{v}_t = \mathbf{M}^t\mathbf{v}_0$. Hvis $\mathbf{v}_0$ ikke er en sum af egenvektorer hørende til egenværdierne $\lambda_2, \lambda_3, \ldots, \lambda_n$ (dvs. hvis $\mathbf{v}_0$ ikke er af formen $\mathbf{v}_0 = \mathbf{q}_2 + \mathbf{q}_3 + \ldots + \mathbf{q}_n$, hvor $\mathbf{q}_k$ er en egenvektor hørende til egenværdien λ_k), så findes et tal c, således at

$$\frac{\mathbf{v}_t}{\lambda_1^t} = \frac{\mathbf{M}^t\mathbf{v}_0}{\lambda_1^t} \to c\,\mathbf{q}_1 \qquad \text{når } t \to \infty. \tag{B.7}$$

Bevis for Sætning B.9.6. Udelades. □

Bemærkning B.9.6 Når $|\lambda_1| > |\lambda_k|$ for $k = 2, \ldots, n$ som i Sætning B.9.6, så kaldes λ_1 den *dominerende egenværdi* for matricen $\mathbf{M}$.

LA 215

I Sætning B.9.4 så vi, at en overgangsmatrix altid har en ligevægt $\mathbf{v}^* \neq \mathbf{o}$. Vi vil nu benytte Sætning B.9.6 til at undersøge denne ligevægt nærmere. Vi har ikke den nødvendige teori til rådighed til at formulere sætningen helt præcist, men vi vil kun betragte overgangsmatricer, for hvilke sætningen gælder.

Sætning B.9.7 (Fremskrivning med overgangsmatrix) Gældende for de fleste $n \times n$ overgangsmatricer $\mathbf{M}$. Lad $\mathbf{v}^* \neq \mathbf{o}$ være en ligevægt for $\mathbf{M}$. For en vikårlig vektor $\mathbf{v}_0$ sætter vi $\mathbf{v}_t = \mathbf{M}^t\mathbf{v}_0$. Så gælder følgende:

(a) $\mathbf{v}^*$ er en egenvektor for $\mathbf{M}$ med tilhørende egenværdi 1.

(b) 1 er den dominerende egenværdi for matricen $\mathbf{M}$.

(c) Der findes en konstant c, således at $\mathbf{v}_t \to c\mathbf{v}^*$ når $t \to \infty$.

Bevis for Sætning B.9.7. Udelades. □

Fremskrivning af en begyndelsesvektor $\mathbf{v}_0$ i $\mathbb{R}^n$ vha. en affin afbildning $f(\mathbf{v}) = \mathbf{M}\mathbf{v} + \mathbf{q}$ (hvor $\mathbf{M}$ er en $n \times n$ matrix og $\mathbf{q}$ en vektor i $\mathbb{R}^n$) foregår som for 2×2 matricer; se (B.6) på s. 159: Vektorerne $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \ldots$ defineres ved

$$\mathbf{v}_1 = f(\mathbf{v}_0), \qquad \mathbf{v}_2 = f(\mathbf{v}_1), \qquad \mathbf{v}_3 = f(\mathbf{v}_2), \qquad \ldots$$

dvs. $\mathbf{v}_{t+1} = f(\mathbf{v}_t)$. Vi slutter af med at generalisere Sætning B.8.4 til $n \times n$ matricer.

Sætning B.9.8 (Fremskrivning med affin afbildning) Lad f være den affine afbildning

$$f(\mathbf{v}) = \mathbf{M}\mathbf{v} + \mathbf{q}$$

givet ved en $n \times n$ matrix $\mathbf{M}$ og en vektor $\mathbf{q}$ i $\mathbb{R}^n$. Antag at $\mathbf{M} - \mathbf{E}$ har en invers matrix og lad $\mathbf{v}^*$ være ligevægten for f. Antag endvidere, at $\mathbf{M}$ har n forskellige egenværdier $\lambda_1, \lambda_2, \ldots, \lambda_n$, som opfylder $|\lambda_k| < 1$ for $k = 1, \ldots, n$. For en vikårlig vektor $\mathbf{v}_0$ i $\mathbb{R}^n$ sætter vi

$$\mathbf{v}_1 = f(\mathbf{v}_0), \qquad \mathbf{v}_2 = f(\mathbf{v}_1), \qquad \mathbf{v}_3 = f(\mathbf{v}_2) \qquad \text{osv.}$$

Så gælder der

$$\mathbf{v}_t \to \mathbf{v}^* \qquad \text{når } t \to \infty.$$

Bevis for Sætning B.9.8. Udelades. □

Bemærkning B.9.7 Ligevægten $\mathbf{v}^*$ for en affin afbildning fra $\mathbb{R}^n$ ind i $\mathbf{R}^n$ kaldes ofte *stabil*, hvis $\mathbf{v}_t \to \mathbf{v}^*$ når $t \to \infty$ uanset hvilken begyndelsesvektor $\mathbf{v}_0$, man vælger. Sætning B.9.8 siger dermed, at ligevægten er stabil, hvis $|\lambda_k| < 1$ for $k = 1, \ldots, n$.

Anvendelseseksempler

Anvendelseseksempel B.1 Kaninpopulation

Vi betragter en population af kaniner. Vi siger at en kanin er ung i dens første leveår og derefter, at den er gammel. Man har observeret, at unge hunkaniner i snit har 70% chance for at overleve til året efter, mens chancen for at en gammel hunkanin overlever til året efter er 40%. Unge hunkaniner får i gennemsnit 2.0 hununger om året, mens gamle hunkaniner i gennemsnit får 1.5 hununger om året.

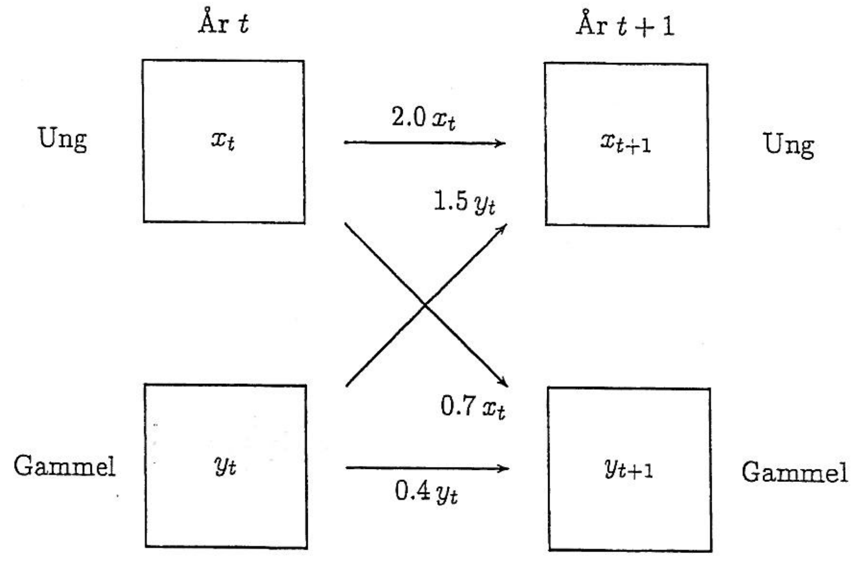
(a) [Matematisk modellering]
For at undersøge hvordan kaninpopulationen udvikler sig med tiden, vil vi foretage en matematisk modellering af situationen. Først giver vi antallene af hunkaniner "matematiske navne", og lader derfor x_t og y_t betegne antallet af hhv. unge og gamle hunkaniner i år t. Hvis vi kender antallene af hunkaniner i år t, så kan antallene af hunkaniner i år $t+1$ beregnes på følgende måde: Først beregnes antallet af unge hunkaniner i år $t+1$. De x_t unge hunkaniner i år t føder tilsammen $2.0 \cdot x_t$ hunkaniner, og disse bliver til unge hunkaniner i år $t+1$. Tilsvarende føder de y_t gamle hunkaniner i år t tilsammen $1.5 \cdot y_t$ hunkaniner, og disse bliver ligeledes til unge hunkaniner i år $t+1$. Det samlede antal unge hunkaniner i år $t+1$ er derfor $2.0 \cdot x_t + 1.5 \cdot y_t$, dvs.

$$x_{t+1} = 2.0 \cdot x_t + 1.5 \cdot y_t. \tag{B.8}$$

Dernæst beregnes antallet af gamle hunkaniner i år $t+1$. Af de x_t unge hunkaniner i år t overlever $0.7 \cdot x_t$ hunkaniner, og disse bliver til gamle hunkaniner i år $t+1$. Tilsvarende overlever $0.4 \cdot y_t$ af de y_t gamle hunkaniner i år t og disse vil fortsat være gamle hunkaniner i år $t+1$. Det samlede antal gamle hunkaniner i år $t+1$ er derfor $0.7 \cdot x_t + 0.4 \cdot y_t$, dvs.

$$y_{t+1} = 0.7 \cdot x_t + 0.4 \cdot y_t. \tag{B.9}$$

Disse sammenhænge mellem kaninpopulationerne i år t og $t+1$ kan angives skematisk i et såkaldt *kompartmentdiagram* (se Figur B.10):



Figur B.10: Kompartmentdiagram til beskrivelse af udviklingen i populationerne

LA 217

(b) Ligningerne (B.8) og (B.9) kan sammen angives på formen

$$\begin{pmatrix} x_{t+1} \\ y_{t+1} \end{pmatrix} = \begin{pmatrix} 2.0\,x_t + 1.5\,y_t \\ 0.7\,x_t + 0.4\,y_t \end{pmatrix}, \tag{B.10}$$

og hvis vi indfører vektor- og matrixnotation

$$\mathbf{v}_t = \begin{pmatrix} x_t \\ y_t \end{pmatrix} \qquad \text{og} \qquad \mathbf{M} = \begin{pmatrix} 2.0 & 1.5 \\ 0.7 & 0.4 \end{pmatrix}$$

følger det af (B.10) at

$$\mathbf{v}_{t+1} = \mathbf{M}\mathbf{v}_t. \tag{B.11}$$

(c) Hvis vi antager, at der i år 0 er 100 unge og ingen gamle hunkaniner, dvs. $\mathbf{v}_0 = \binom{100}{0}$, så kan vi succesivt benytte (B.11) til at beregne populationerne i de følgende år:

$$\mathbf{v}_1 = \begin{pmatrix} 2.0 & 1.5 \\ 0.7 & 0.4 \end{pmatrix} \begin{pmatrix} 100 \\ 0 \end{pmatrix} = \begin{pmatrix} 200 \\ 70 \end{pmatrix},$$

$$\mathbf{v}_2 = \begin{pmatrix} 2.0 & 1.5 \\ 0.7 & 0.4 \end{pmatrix} \begin{pmatrix} 200 \\ 70 \end{pmatrix} = \begin{pmatrix} 505 \\ 168 \end{pmatrix},$$

$$\mathbf{v}_3 = \begin{pmatrix} 2.0 & 1.5 \\ 0.7 & 0.4 \end{pmatrix} \begin{pmatrix} 505 \\ 168 \end{pmatrix} = \begin{pmatrix} 1262 \\ 420.7 \end{pmatrix} \quad \text{osv.}$$

Da der er tale om antal kaniner, er det mest rimeligt at runde af til hele tal og angive $\mathbf{v}_3$ som $\binom{1262}{421}$.

Hvis vi i stedet for antager, at der i år 0 er 100 gamle og ingen unge hunkaniner, dvs. $\mathbf{v}_0 = \binom{0}{100}$, så fås

$$\mathbf{v}_1 = \begin{pmatrix} 2.0 & 1.5 \\ 0.7 & 0.4 \end{pmatrix} \begin{pmatrix} 0 \\ 100 \end{pmatrix} = \begin{pmatrix} 150 \\ 40 \end{pmatrix},$$

$$\mathbf{v}_2 = \begin{pmatrix} 2.0 & 1.5 \\ 0.7 & 0.4 \end{pmatrix} \begin{pmatrix} 150 \\ 40 \end{pmatrix} = \begin{pmatrix} 360 \\ 121 \end{pmatrix},$$

$$\mathbf{v}_3 = \begin{pmatrix} 2.0 & 1.5 \\ 0.7 & 0.4 \end{pmatrix} \begin{pmatrix} 360 \\ 121 \end{pmatrix} = \begin{pmatrix} 901.5 \\ 300.4 \end{pmatrix} \quad \text{osv.}$$

Igen afrundes $\mathbf{v}_3$ og vi får $\mathbf{v}_3 = \binom{902}{300}$.

(d) [Matematisk problembehandling]
Det oplyses at der i år 3 var 160 unge og 50 gamle hunkaniner. Vi ønsker at bestemme antallene af unge og gamle hunkaniner året før, dvs. i år 2. Vi ved altså at $x_3 = 160$ og $y_3 = 50$, og vil vide hvad x_2 og y_2 er. Vi vil angive to forskellige måder at gøre dette på.

(i) Ved lige store koefficienters metode: Ved indsættelse i (B.8) og (B.9) fås ligningerne

$$\begin{cases} 160 = 2.0\,x_2 + 1.5\,y_2 \\ \;\;50 = 0.7\,x_2 + 0.4\,y_2. \end{cases}$$

Ved at gange første ligning med 0.4 og anden ligning med -1.5 fås lige store koefficienter (med modsat fortegn) på y_2 i de 2 ligninger:

$$\begin{cases} \;\;64 = \;\;0.8\;\;x_2 + 0.6\,y_2 \\ -75 = -1.05\,x_2 - 0.6\,y_2. \end{cases}$$

LA 218

Ved at lægge disse 2 ligninger sammen fås

$$-11 = -0.25\,x_2 \qquad \text{dvs.} \qquad x_2 = 44.$$

Til sidst indsættes $x_2 = 44$ i ligningen $160 = 2.0\,x_2 + 1.5\,y_2$, hvilket giver

$$160 = 88 + 1.5\,y_2 \qquad \text{dvs.} \qquad y_2 = 48.$$

I år 2 var der altså 44 unge og 48 gamle hunkaniner.

(ii) Ved matrixinvertering: Vi bemærker først at

$$\det \mathbf{M} = 2.0 \cdot 0.4 - 0.7 \cdot 1.5 = -0.25 \neq 0,$$

så $\mathbf{M}$ har den inverse matrix

$$\mathbf{M}^{-1} = \frac{1}{-0.25}\begin{pmatrix} 0.4 & -1.5 \\ -0.7 & 2.0 \end{pmatrix} = \begin{pmatrix} -1.6 & 6.0 \\ 2.8 & -8.0 \end{pmatrix}.$$

Af formlen $\mathbf{v}_3 = \mathbf{M}\mathbf{v}_2$ fås ved at gange med den inverse matrix $\mathbf{M}^{-1}$ på begge sider, at $\mathbf{v}_2 = \mathbf{M}^{-1}\mathbf{v}_3$, dvs.

$$\mathbf{v}_2 = \begin{pmatrix} -1.6 & 6.0 \\ 2.8 & -8.0 \end{pmatrix}\begin{pmatrix} 160 \\ 50 \end{pmatrix} = \begin{pmatrix} 44 \\ 48 \end{pmatrix}.$$

En fordel ved denne metode er, at det nu er let at beregne $\mathbf{v}_2$ for andre værdier af $\mathbf{v}_3$. Hvis f.eks. $\mathbf{v}_3 = \binom{155}{50}$, så fås

$$\mathbf{v}_2 = \begin{pmatrix} -1.6 & 6.0 \\ 2.8 & -8.0 \end{pmatrix}\begin{pmatrix} 155 \\ 50 \end{pmatrix} = \begin{pmatrix} 52 \\ 34 \end{pmatrix}.$$

(e) I begge tilfælde i (c) ovenfor ser det ud som om, der efterhånden bliver 3 gange så mange unge som gamle kaniner, samt at antallet af såvel unge som gamle kaniner omtrent vokser med en faktor 2.5 hvert år. Vi betragter derfor den situation, hvor der i startpopulationen er præcis 3 gange så mange unge som gamle kaniner, dvs. $x_0 = 3y_0$. Så er $\mathbf{v}_0 = \binom{3y_0}{y_0}$ og dermed

$$\mathbf{v}_1 = \begin{pmatrix} 2.0 & 1.5 \\ 0.7 & 0.4 \end{pmatrix}\begin{pmatrix} 3y_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} 7.5\,y_0 \\ 2.5\,y_0 \end{pmatrix} = 2.5\,\mathbf{v}_0,$$

så efter et år er antallet af såvel unge som gamle kaniner 2.5 gange så stort, og der er stadig 3 gange så mange unge som gamle kaniner. Ligningen ovenfor kan skrives som

$$\mathbf{M}\mathbf{v}_0 = 2.5\,\mathbf{v}_0,$$

hvor altså $\mathbf{v}_0 = \binom{3y_0}{y_0}$. Dette betyder iflg. Definition B.8.1, at $\mathbf{v}_0$ er en egenvektor (når $y_0 \neq 0$) for $\mathbf{M}$ med tilhørende egenværdi 2.5.

(f) Vi afslutter dette eksempel med at advare om, at den opstillede model i en vis forstand er meget simplificeret. For eksempel har vi tolket oplysningen om, at unge hunkaniner i gennemsnit får 2.0 hununger om året, som at de x_t unge hunkaniner i år t tilsammen føder *præcis* $2.0 \cdot x_t$ hunkaniner. Dette svarer lidt til at sige, at hvis man slår plat og krone 6 gange, så vil man få plat *præcis* 3 gange. Man vil i *gennemsnit* få plat 3 gange, men der er jo også en vis sandsynlighed for at få plat f.eks. 1 eller 2 gange. Tages dette i betragtning kommer der altså sandsynlighedsregning ind i billedet.

I vores kanin-eksempel er der tilsvarende for en ung hunkanin i et givet år en vis sandsynlighed for at hun får 0 hununger, en vis sandsynlighed for at hun får 1 hununge, en vis sandsynlighed for at hun får 2 hununger, en vis sandsynlighed for at hun får 3 hununger osv. Det eneste vi benytter i vores opstilling af problemet i (a) er, at hun *i gennemsnit* får 2 hununger. Vi har valgt udelukkende at regne på gennemsnittene, så vi benytter kun de gennemsnitlige antal unger og de gennemsnitlige overlevelseschancer. Det er faktisk muligt at opstille en model (en såkaldt *stokastisk model*), der tager hensyn til de tilfældige variationer i såvel antallet af unger som overlevelseschancerne. (Vi vil dog ikke komme ind på sådanne modeller i disse noter.)

(I Anvendelseseksempel B.7 og Anvendelseseksempel B.8 vil vi se nogle generalisationer af modellen i dette anvendelseseksempel.)

Anvendelseseksempel B.2 Produktionsplan

(a) En foderstoffabrik fremstiller to kraftfodertyper, F_1 og F_2, ud fra to råvarer, R_1 og R_2. Til fremstilling af 1 tons F_1 benyttes 200 kg R_1 og 600 kg R_2 (samt 200 kg fyldstof, som vi her vil se bort fra). Til 1 tons F_2 benyttes 150 kg R_1, 800 kg R_2 (og 50 kg fyldstof). Vi ønsker at opstille en model for beregning af fabrikkens råvareforbrug.

[Matematisk modellering]
Vi betegner den daglige produktion af F_1 og F_2 med hhv. x og y, begge opgivet i tons, og lader x' og y' betegne det daglige forbrug af hhv. R_1 og R_2, ligeledes i tons. Af oplysningerne følger det, at

$$\begin{cases} x' = 0.2\,x + 0.15\,y \\ y' = 0.6\,x + 0.8\ \ y. \end{cases} \tag{B.12}$$

Eksempelvis vil en daglig produktion af 3 tons F_1 og 2 tons F_2 (dvs. $x = 3$ og $y = 2$) give et dagligt råvareforbrug på

$$x' = 0.2 \cdot 3 + 0.15 \cdot 2 = 0.9 \qquad \text{og} \qquad y' = 0.6 \cdot 3 + 0.8 \cdot 2 = 3.4,$$

dvs. 0.9 tons R_1 og 3.4 tons R_2.

Udtrykket (B.12) for råvareforbruget kan også skrives

$$\mathbf{v}' = \mathbf{M}\mathbf{v},$$

hvor $\mathbf{v} = \binom{x}{y}$ er produktionsvektoren, $\mathbf{v}' = \binom{x'}{y'}$ er råvareforbrugsvektoren, og hvor

$$\mathbf{M} = \begin{pmatrix} 0.2 & 0.15 \\ 0.6 & 0.8 \end{pmatrix}.$$

Med matrix- og vektornotation kan råvareforbruget hørende til en produktion på $\mathbf{v} = \binom{3}{2}$ bestemmes ved

$$\mathbf{v}' = \begin{pmatrix} 0.2 & 0.15 \\ 0.6 & 0.8 \end{pmatrix} \begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 0.9 \\ 3.4 \end{pmatrix},$$

som vi også kom frem til ovenfor.

LA 220

(b) Råvarerne indeholder to sporstoffer S_1 og S_2, som man ønsker at holde kontrol med. Det har vist sig, at 1 tons af R_1 indeholder 3 enheder S_1 og 5 enheder S_2, mens 1 tons af R_2 indeholder 2 enheder S_1 og 7 enheder S_2.

[Matematisk modellering]
I et parti råvarer på x' tons R_1 og y' tons R_2 betegner vi det samlede indhold af sporstofferne S_1 og S_2 med hhv. x'' og y'', begge opgivet i antal enheder. Af oplysningerne finder vi da sammenhængen

$$\begin{cases} x'' = 3x' + 2y' \\ y'' = 5x' + 7y'. \end{cases} \tag{B.13}$$

Dette kan også skrives på formen

$$\mathbf{v}'' = \mathbf{N}\mathbf{v}',$$

hvor $\mathbf{v}' = \binom{x'}{y'}$ er råvarevektoren, $\mathbf{v}'' = \binom{x''}{y''}$ er sporstofvektoren, og hvor

$$\mathbf{N} = \begin{pmatrix} 3 & 2 \\ 5 & 7 \end{pmatrix}.$$

Eksempelvis giver $\mathbf{v}' = \binom{x'}{y'} = \binom{0.9}{3.4}$ et indhold af sporstoffer på

$$\mathbf{v}'' = \begin{pmatrix} 3 & 2 \\ 5 & 7 \end{pmatrix} \begin{pmatrix} 0.9 \\ 3.4 \end{pmatrix} = \begin{pmatrix} 9.5 \\ 28.3 \end{pmatrix}.$$

(c) I (a) undersøgte vi sammenhængen mellem den daglige produktion og det tilsvarende råvareforbrug, og i (b) undersøgte vi sammenhængen mellem råvareforbruget og det tilsvarende indhold af sporstoffer. Vi vil nu kombinere disse to sammenhænge, og derved finde sammenhængen mellem den daglige produktion og det tilsvarende indhold af sporstoffer. Ved indsættelse af (B.12) i (B.13) får vi

$$\begin{cases} x'' = 3x' + 2y' = 3(0.2\,x + 0.15\,y) + 2(0.6\,x + 0.8\,y) = 1.8\,x + 2.05\,y, \\ y'' = 5x' + 7y' = 5(0.2\,x + 0.15\,y) + 7(0.6\,x + 0.8\,y) = 5.2\,x + 6.35\,y. \end{cases} \tag{B.14}$$

Det er faktisk lettere at bestemme denne sammenhæng mellem produktionen og det tilsvarende indhold af sporstofferne vha. matricerne $\mathbf{M}$ og $\mathbf{N}$. Ved indsættelse af $\mathbf{v}' = \mathbf{M}\mathbf{v}$ i $\mathbf{v}'' = \mathbf{N}\mathbf{v}'$ fås

$$\mathbf{v}'' = \mathbf{N}(\mathbf{M}\mathbf{v}) = (\mathbf{N}\mathbf{M})\mathbf{v},$$

hvor matricen $\mathbf{NM}$ bestemmes ved matrixmultiplikation som i Definition B.2.1:

		0.2	0.15
		0.6	0.8
3	2	$3 \cdot 0.2 + 2 \cdot 0.6$	$3 \cdot 0.15 + 2 \cdot 0.8$
5	7	$5 \cdot 0.2 + 7 \cdot 0.6$	$5 \cdot 0.15 + 7 \cdot 0.8$

Heraf ses det, at

$$\mathbf{NM} = \begin{pmatrix} 3 & 2 \\ 5 & 7 \end{pmatrix} \begin{pmatrix} 0.2 & 0.15 \\ 0.6 & 0.8 \end{pmatrix} = \begin{pmatrix} 1.8 & 2.05 \\ 5.2 & 6.35 \end{pmatrix},$$

hvilket stemmer med udregningen i (B.14).

Vi har tidligere set, at en daglig produktion på $\mathbf{v} = \binom{3}{2}$ fører til et råvareforbrug på $\mathbf{v}' = \binom{0.9}{3.4}$, samt at indholdet af sporstoffer i denne mængde råvarer er $\mathbf{v}'' = \binom{9.5}{28.3}$. Vha. udtrykket $\mathbf{v}'' = (\mathbf{NM})\mathbf{v}$ kan man springe mellemregningen af råvareforbruget $\mathbf{v}'$ over og direkte bestemme

$$\mathbf{v}'' = \begin{pmatrix} 1.8 & 2.05 \\ 5.2 & 6.35 \end{pmatrix} \begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 9.5 \\ 28.3 \end{pmatrix}.$$

(Dette resultat kan selvfølgelig også fås ved indsættelse af $x = 3$ og $y = 2$ i (B.14).)

(d) Foderstoffabrikken kan forbedre kvaliteten af de to typer kraftfoder ved også at benytte en tredje råvare R_3. Til fremstilling af 1 tons F_1 vil der nu blive brugt 160 kg R_1, 300 kg R_2 og 340 kg R_3 (samt stadigvæk 200 kg fyldstof), mens der til 1 tons F_2 vil blive brugt 120 kg R_1, 400 kg R_2 og 430 kg R_3 (samt 50 kg fyldstof).

[Matematisk modellering]
For at opstille en model for fabrikkens nye råvareforbrug betegner vi som hidtil den daglige produktion af F_1 og F_2 med hhv. x og y (begge opgivet i tons), men lader nu x', y' og z' betegne det daglige forbrug af hhv. R_1, R_2 og R_3 (ligeledes i tons). Vi får da følgende ligninger

$$\begin{cases} x' = 0.16x + 0.12y \\ y' = 0.3\ \ x + 0.4\ \ y \\ z' = 0.34x + 0.43y. \end{cases}$$

Disse ligningerne kan som i (a) angives på matrixform, men her er der brug for en matrix med 3 rækker og 2 søjler (en 3×2 matrix). Vi får da

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 0.16x + 0.12y \\ 0.3\ \ x + 0.4\ \ y \\ 0.34x + 0.43y \end{pmatrix} = \begin{pmatrix} 0.16 & 0.12 \\ 0.3 & 0.4 \\ 0.34 & 0.43 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}. \tag{B.15}$$

Vi ændrer nu betegnelserne brugt i (a), (b) og (c) og lader

$$\mathbf{v} = \begin{pmatrix} x \\ y \end{pmatrix}, \quad \mathbf{v}' = \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} \quad \text{og} \quad \mathbf{M} = \begin{pmatrix} 0.16 & 0.12 \\ 0.3 & 0.4 \\ 0.34 & 0.43 \end{pmatrix}.$$

Dermed kan (B.15) skrives på formen

$$\mathbf{v}' = \mathbf{M}\mathbf{v}.$$

Eksempelvis giver en produktion på $\mathbf{v} = \binom{3}{2}$ et råvareforbrug på

$$\mathbf{v}' = \begin{pmatrix} 0.16 & 0.12 \\ 0.3 & 0.4 \\ 0.34 & 0.43 \end{pmatrix} \begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 0.72 \\ 1.7 \\ 1.88 \end{pmatrix}.$$

(e) Den nye råvare R_3 indeholder 2 enheder af sporstoffet S_2 pr. tons, men intet af sporstoffet S_1. Idet x', y' og z' angiver mængderne af råvarerne R_1, R_2 og R_3 (i tons), kan det samlede indhold x'' og y'' af sporstofferne S_1 og S_2 beregnes af formlerne

$$\begin{cases} x'' = 3x' + 2y' \\ y'' = 5x' + 7y' + 2z' \end{cases}$$

LA 222

Ved brug af en 2×3 matrix kan dette som i (b) skrives på formen

$$\begin{pmatrix} x'' \\ y'' \end{pmatrix} = \begin{pmatrix} 3 & 2 & 0 \\ 5 & 7 & 2 \end{pmatrix} \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}. \tag{B.16}$$

Dette kan også skrives på matrixformen $\mathbf{v}'' = \mathbf{N}\mathbf{v}'$, hvor vi har sat

$$\mathbf{v}'' = \begin{pmatrix} x'' \\ y'' \end{pmatrix} \quad \text{og} \quad \mathbf{N} = \begin{pmatrix} 3 & 2 & 0 \\ 5 & 7 & 2 \end{pmatrix}$$

(matricen $\mathbf{N}$ er altså en anden end i (b) og (c)). Eksempelvis vil indholdet af sporstoffer $\mathbf{v}'' = \binom{x''}{y''}$ i et råvareforbrug på

$$\mathbf{v}' = \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 0.72 \\ 1.7 \\ 1.88 \end{pmatrix}$$

være

$$\mathbf{v}'' = \begin{pmatrix} 3 & 2 & 0 \\ 5 & 7 & 2 \end{pmatrix} \begin{pmatrix} 0.72 \\ 1.7 \\ 1.88 \end{pmatrix} = \begin{pmatrix} 5.56 \\ 19.26 \end{pmatrix}.$$

(f) Endelig kan vi i lighed med (c) benytte (B.15) og (B.16) til at bestemme det samlede indhold x'' og y'' af sporstofferne i et parti foderstoffer på x tons F_1 og y tons F_2:

$$\begin{aligned} \begin{pmatrix} x'' \\ y'' \end{pmatrix} &= \begin{pmatrix} 3 & 2 & 0 \\ 5 & 7 & 2 \end{pmatrix} \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} \\ &= \begin{pmatrix} 3 & 2 & 0 \\ 5 & 7 & 2 \end{pmatrix} \begin{pmatrix} 0.16 & 0.12 \\ 0.3 & 0.4 \\ 0.34 & 0.43 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \\ &= \begin{pmatrix} 1.08 & 1.16 \\ 3.58 & 4.26 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}. \end{aligned} \tag{B.17}$$

Bemærk at dette involverede produktet $\mathbf{NM}$ af 2×3 matricen $\mathbf{N}$ med 3×2 matricen $\mathbf{M}$. Ved udregningen af dette produkt kan man f.eks. benytte skemaet

	M
N	**NM**

dvs.

			0.16	0.12
			0.3	0.4
			0.34	0.43
3	2	0	$3 \cdot 0.16 + 2 \cdot 0.3 + 0 \cdot 0.34$	$3 \cdot 0.12 + 2 \cdot 0.4 + 0 \cdot 0.43$
5	7	2	$5 \cdot 0.16 + 7 \cdot 0.3 + 2 \cdot 0.34$	$5 \cdot 0.12 + 7 \cdot 0.4 + 2 \cdot 0.43$

Med matrixnotation kan udregningen (B.17) skrives

$$\mathbf{v}'' = \mathbf{N}\mathbf{v}' = \mathbf{N}(\mathbf{M}\mathbf{v}) = \mathbf{N}\mathbf{M}\mathbf{v}.$$

Anvendelseseksempel B.3 Svampesygdom i plantage
I en stor granplantage rammes træerne til tider af svampesygdom. Træerne kan groft inddeles i to grupper: raske og syge. Man har konstateret, at et raskt træ med 6% sandsynlighed vil være blevet sygt året efter. Ligeledes vil et sygt træ med 72% sandsynlighed været blevet raskt året efter. Vi antager endvidere, at ingen træer dør af sygdommen eller af andre årsager, og at der heller ikke plantes nye træer. Der er i alt 26000 træer i plantagen.

Ud fra disse oplysninger kan vi opstille følgende model: vi sætter r_t hhv. s_t til antallet af raske træer hhv. antallet af syge træer i år t. De raske træer året efter, r_{t+1}, udgøres af 94% af de raske træer i år, dvs. $0.94\,r_t$, samt af 72% af de syge træer i år, dvs. $0.72\,s_t$. Dermed finder vi

$$r_{t+1} = 0.94\,r_t + 0.72\,s_t,$$

og tilsvarende $s_{t+1} = 0.06r_t + 0.28s_t$. På matrixform har vi derfor

$$\begin{pmatrix} r_{t+1} \\ s_{t+1} \end{pmatrix} = \begin{pmatrix} 0.94 & 0.72 \\ 0.06 & 0.28 \end{pmatrix} \begin{pmatrix} r_t \\ s_t \end{pmatrix}.$$

Af denne matrixmodel fremgår ikke, at der hele tiden er 26000 træer i plantagen, og det er vi således nødt til at formulere særskilt:

$$r_t + s_t = 26000$$

for alle t. Vi bemærker endvidere, at matricen

$$\mathbf{M} = \begin{pmatrix} 0.94 & 0.72 \\ 0.06 & 0.28 \end{pmatrix}$$

er en overgangsmatrix ifølge Definition B.4.2.

Vi vil undersøge fordelingen mellem raske og syge træer, og se om denne fordeling stabiliserer sig over en lang periode. Vi foretager først nogle eksperimenter. Lad os tænke os, at der til at begynde med er 13000 raske og 13000 syge træer; hvordan udvikler fordelingen mellem raske og syge træer sig så? Vi finder, med hjælp fra computer, at

$$\begin{pmatrix} r_4 \\ s_4 \end{pmatrix} = \begin{pmatrix} 0.94 & 0.72 \\ 0.06 & 0.28 \end{pmatrix}^4 \begin{pmatrix} 13000 \\ 13000 \end{pmatrix} \simeq \begin{pmatrix} 23974 \\ 2026 \end{pmatrix},$$

$$\begin{pmatrix} r_8 \\ s_8 \end{pmatrix} = \begin{pmatrix} 0.94 & 0.72 \\ 0.06 & 0.28 \end{pmatrix}^8 \begin{pmatrix} 13000 \\ 13000 \end{pmatrix} \simeq \begin{pmatrix} 24000 \\ 2000 \end{pmatrix}.$$

(Her har vi bl.a. udeladt udregningen af $\binom{r_1}{s_1}$, $\binom{r_2}{s_2}$ og $\binom{r_3}{s_3}$, og bruger i stedet for potensopløftning af matricen for at udregne $\binom{r_4}{s_4}$.) På samme måde går det, hvis alle træerne til at begynde med er syge:

$$\begin{pmatrix} r_4 \\ s_4 \end{pmatrix} = \begin{pmatrix} 0.94 & 0.72 \\ 0.06 & 0.28 \end{pmatrix}^4 \begin{pmatrix} 0 \\ 26000 \end{pmatrix} \simeq \begin{pmatrix} 23944 \\ 2056 \end{pmatrix},$$

$$\begin{pmatrix} r_8 \\ s_8 \end{pmatrix} = \begin{pmatrix} 0.94 & 0.72 \\ 0.06 & 0.28 \end{pmatrix}^8 \begin{pmatrix} 0 \\ 26000 \end{pmatrix} \simeq \begin{pmatrix} 24000 \\ 2000 \end{pmatrix}.$$

Det ser altså stærkt ud til, at fordelingen med 24000 raske træer og 2000 syge træer er en ligevægt. Vi kan faktisk direkte bestemme ligevægten, og vise, at der ikke er andre mulige ligevægte. Ifølge Sætning B.4.1 er

$$\mathbf{v}^* = \begin{pmatrix} r^* \\ s^* \end{pmatrix} = k \begin{pmatrix} 0.72 \\ 0.06 \end{pmatrix}$$

LA 224

en ligevægt for matricen **M** for et vilkårligt tal k. Tallet k bestemmes ud fra oplysningen om, at der er 26000 træer i plantagen, idet der også skal gælde

$$26000 = r^* + s^* = 0.72\,k + 0.06\,k = 0.78\,k,$$

dvs. $k = 26000/0.78 \simeq 33333.33$. Videre finder vi

$$r^* = 0.72\,k = 24000 \qquad \text{og} \qquad s^* = 0.06\,k = 2000.$$

Vi har altså vist, at der kun er en ligevægt med $r^* + s^* = 26000$, nemlig $\binom{r^*}{s^*} = \binom{24000}{2000}$.

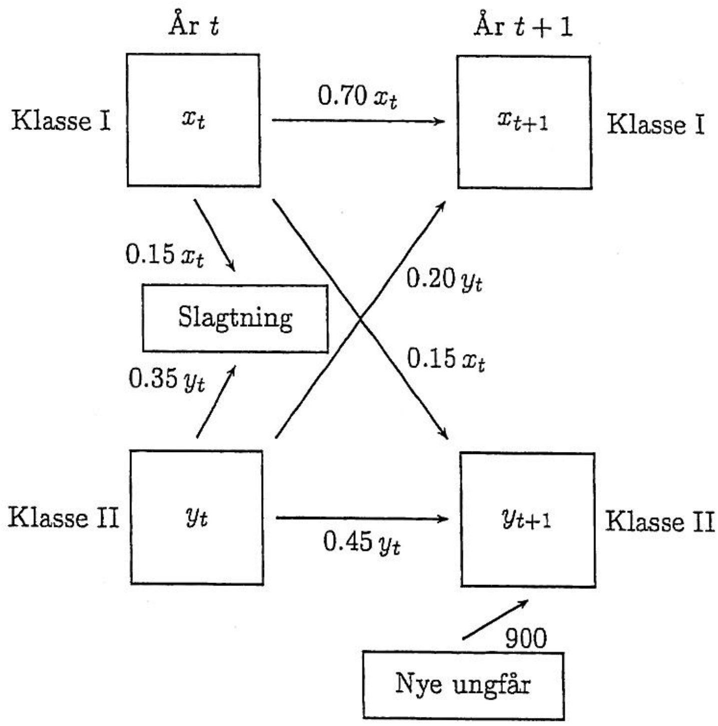
Anvendelseseksempel B.4 Fårebestand

En fåreavler opdeler hvert år sin bestand i to klasser,

klasse I:	får med høj uldproduktion,
klasse II:	får med normal uldproduktion.

Det har vist sig, at der hvert år sker dels en oprykning af 20% af klasse II fårene til klasse I, dels en nedrykning af 15% af klasse I fårene til klasse II. Endvidere udtages der årligt 15% af klasse I fårene og 35% af klasse II fårene til slagtning. Endelig tillægges der hvert år 900 ungfår, som alle starter i klasse II.

(a) [Matematisk modellering]
Vi ønsker en kort, præcis beskrivelse af bevægelserne i de to klasser fra år til år. Til dette formål lader vi x_t og y_t betegne antallene af får i hhv. klasse I og klasse II i år t ($t = 1, 2, 3, \ldots$), og angiver skematisk oplysningerne ovenfor i et kompartmentdiagram (som i Anvendelseseksempel B.1):



Figur B.11: Kompartmentdiagram til beskrivelse af udviklingen i klasserne

LA 225

Nogle enkelte af tallene fremgår ikke direkte af oplysningerne, men er fremkommet ved mellemregninger:

- Af klasse I fårene nedrykkes 15% til klasse II og 15% udtages til slagtning, så der er $(100-15-15)\% = 70\%$ af klasse I fårene, der bliver ved med at være klasse I får.
- Af klasse II fårene oprykkes 20% til klasse I og 35% udtages til slagtning, så der er $(100-20-35)\% = 45\%$ af klasse II fårene, der bliver ved med at være klasse II får.

Af diagrammet ses det, at vi kan udtrykke x_{t+1} og y_{t+1} ved x_t og y_t på følgende vis:

$$\begin{cases} x_{t+1} = 0.70\,x_t + 0.20\,y_t \\ y_{t+1} = 0.15\,x_t + 0.45\,y_t + 900. \end{cases}$$

(b) Ligningerne ovenfor kan skrives på formen

$$\begin{pmatrix} x_{t+1} \\ y_{t+1} \end{pmatrix} = \begin{pmatrix} 0.70 & 0.20 \\ 0.15 & 0.45 \end{pmatrix} \begin{pmatrix} x_t \\ y_t \end{pmatrix} + \begin{pmatrix} 0 \\ 900 \end{pmatrix},$$

dvs.

$$\mathbf{v}_{t+1} = \mathbf{M}\mathbf{v}_t + \mathbf{q},$$

hvor

$$\mathbf{v}_t = \begin{pmatrix} x_t \\ y_t \end{pmatrix}, \qquad \mathbf{M} = \begin{pmatrix} 0.70 & 0.20 \\ 0.15 & 0.45 \end{pmatrix} \qquad \text{og} \qquad \mathbf{q} = \begin{pmatrix} 0 \\ 900 \end{pmatrix}.$$

Dette viser, at sammenhængen mellem $\mathbf{v}_t = \binom{x_t}{y_t}$ og $\mathbf{v}_{t+1} = \binom{x_{t+1}}{y_{t+1}}$ kan opfattes som en affin afbildning:

$$\mathbf{v}_{t+1} = f(\mathbf{v}_t), \qquad \text{hvor} \qquad f(\mathbf{v}) = \mathbf{M}\mathbf{v} + \mathbf{q}.$$

Antag, at der f.eks. i 1992 var 500 får i klasse I og 1400 får i klasse II, dvs. at $\mathbf{v}_{1992} = \binom{500}{1400}$. Vi finder da:

$$v_{1993} = \mathbf{M}\mathbf{v}_{1992} + \mathbf{q} = \begin{pmatrix} 0.70 & 0.20 \\ 0.15 & 0.45 \end{pmatrix} \begin{pmatrix} 500 \\ 1400 \end{pmatrix} + \begin{pmatrix} 0 \\ 900 \end{pmatrix} = \begin{pmatrix} 630 \\ 1605 \end{pmatrix}$$

altså 630 får i klasse I og 1605 får i klasse II.

(c) I nogle sammenhænge kan det være af interesse at kunne foretage en "tilbageskrivning" af bestanden, dvs. udtrykke fårenes fordeling på de to klasser et givet år ved fordelingen det følgende år. Af udtrykket $\mathbf{v}_{t+1} = \mathbf{M}\mathbf{v}_t + \mathbf{q}$ får vi

$$\mathbf{M}\mathbf{v}_t = \mathbf{v}_{t+1} - \mathbf{q}.$$

Idet $\det \mathbf{M} = 0.70 \cdot 0.45 - 0.15 \cdot 0.20 = 0.285 \neq 0$ har $\mathbf{M}$ en invers matrix og ved at gange med denne på begge sider af lighedstegnet fås

$$\mathbf{v}_t = \mathbf{M}^{-1}(\mathbf{v}_{t+1} - \mathbf{q}).$$

Vides det f.eks., at $\mathbf{v}_{1995} = \binom{1000}{1500}$, så fås

$$\begin{aligned} \mathbf{v}_{1994} &= \mathbf{M}^{-1}(\mathbf{v}_{1995} - \mathbf{q}) \\ &= \frac{1}{0.285}\begin{pmatrix} 0.45 & -0.20 \\ -0.15 & 0.70 \end{pmatrix}\left(\begin{pmatrix} 1000 \\ 1500 \end{pmatrix} - \begin{pmatrix} 0 \\ 900 \end{pmatrix}\right) \\ &= \frac{1}{0.285}\begin{pmatrix} 0.45 & -0.20 \\ -0.15 & 0.70 \end{pmatrix}\begin{pmatrix} 1000 \\ 600 \end{pmatrix} \\ &= \frac{1}{0.285}\begin{pmatrix} 330 \\ 270 \end{pmatrix} \simeq \begin{pmatrix} 1158 \\ 947 \end{pmatrix}, \end{aligned}$$

dvs. i 1994 bestod bestanden af 1158 klasse I og 947 klasse II får.

LA 22c

(d) Vi vil nu bestemme ligevægten for vores model

$$\begin{pmatrix} x_{t+1} \\ y_{t+1} \end{pmatrix} = \begin{pmatrix} 0.70 & 0.20 \\ 0.15 & 0.45 \end{pmatrix} \begin{pmatrix} x_t \\ y_t \end{pmatrix} + \begin{pmatrix} 0 \\ 900 \end{pmatrix}.$$

Da

$$\mathbf{M} - \mathbf{E} = \begin{pmatrix} -0.30 & 0.20 \\ 0.15 & -0.55 \end{pmatrix},$$

har vi $\det(\mathbf{M} - \mathbf{E}) = -0.30 \cdot (-0.55) - 0.15 \cdot 0.20 = 0.135 \neq 0$, så $\mathbf{M} - \mathbf{E}$ har en invers matrix. Sætning B.7.1 giver derfor, at modellen har netop én ligevægt $\mathbf{v}^*$ givet ved

$$\mathbf{v}^* = -(\mathbf{M} - \mathbf{E})^{-1}\mathbf{q} = -\frac{1}{0.135} \begin{pmatrix} -0.55 & -0.20 \\ -0.15 & -0.55 \end{pmatrix} \begin{pmatrix} 0 \\ 900 \end{pmatrix} = -\frac{1}{0.135} \begin{pmatrix} -180 \\ -495 \end{pmatrix} \simeq \begin{pmatrix} 1333 \\ 2000 \end{pmatrix}.$$

Fortolkningen af denne ligevægt er, at hvis der et år er 1333 får af klasse I og 2000 får af klasse II, så vil der året efter ligeledes være 1333 får af klasse I og 2000 får af klasse II. I dette tilfælde vil bestanden altså ikke ændre sig fra år til år.

(e) Ved *iteration* (dvs. gentagen brug) af udtrykket $\mathbf{v}_{t+1} = \mathbf{M}\mathbf{v}_t + \mathbf{q}$ kan man fremskrive bestandens udvikling fra år til år. Vi finder,

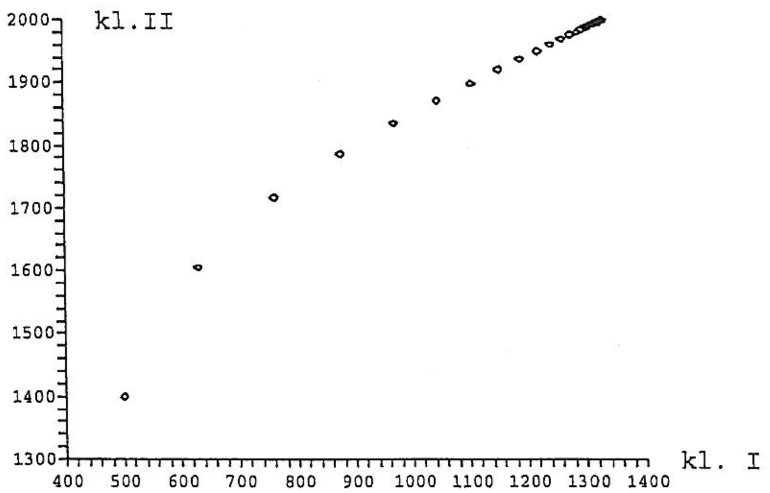
$$v_{1994} = \mathbf{M}\mathbf{v}_{1993} + \mathbf{q} = \begin{pmatrix} 0.70 & 0.20 \\ 0.15 & 0.45 \end{pmatrix} \begin{pmatrix} 5630 \\ 1605 \end{pmatrix} + \begin{pmatrix} 0 \\ 900 \end{pmatrix} = \begin{pmatrix} 762 \\ 1717 \end{pmatrix},$$

og ved at fortsætte på denne måde får man:

år	antal kl. I	antal kl. II
1992	500	1400
1993	630	1605
1994	762	1717
1995	877	1787
1996	971	1837
1997	1047	1872
⋮	⋮	⋮
2021	1332	2000
2022	1333	2000
2023	1333	2000

Der sker altså en tilnærmelse mod modellens ligevægt. Dette kan også illustreres grafisk som på Figur B.12.

LA 22D



Figur B.12: Fårebestandens udvikling over mange år

(f) Egenværdierne for matricen $\mathbf{M}$ kan bestemmes som løsningerne til andengradsligningen

$$\begin{vmatrix} 0.70-\lambda & 0.20 \\ 0.15 & 0.45-\lambda \end{vmatrix} = \lambda^2 - 1.15\,\lambda + 0.285 = 0.$$

Vi får heraf, at egenværdierne er 0.36 og 0.79 (afrundet). Da begge disse er numerisk mindre end 1, kan vi slutte af Sætning B.8.4, at der gælder $\mathbf{v}_t \to \mathbf{v}^* = \binom{1333}{2000}$ for $t \to \infty$, dvs. der er konvergens mod ligevægten, og det endda uanset hvilken startfordeling vi valgte.

Anvendelseseksempel B.5 Nationaløkonomisk model

I følgende nationaløkonomiske model betragtes sammenhængen mellem nationalproduktet Y_t, forbruget C_t og investeringerne I_t i år t.

(a) Der gøres følgende økonomiske antagelser:

- Nationalproduktet er summen af forbruget og investeringerne, dvs.

$$Y_t = C_t + I_t. \tag{B.18}$$

- Forbruget i år $t+1$ består dels af 80% af nationalproduktet i år t og dels af en konstant del på 4, dvs.

$$C_{t+1} = 0.8\,Y_t + 4. \tag{B.19}$$

- Investeringerne i år $t+1$ er halvdelen af forbrugs*stigningen* fra år t til år $t+1$, dvs.

$$I_{t+1} = \tfrac{1}{2}(C_{t+1} - C_t). \tag{B.20}$$

(b) I modellen ovenfor optræder de 3 variable Y_t, C_t og I_t, men ved at indsætte (B.18) i (B.19) kan vi "slippe af med" Y_t og derved få en (matematisk) simplere model med kun de 2 variable C_t og I_t:

$$C_{t+1} = 0.8\,C_t + 0.8\,I_t + 4. \tag{B.21}$$

LA 228

Vi ønsker endvidere at udtrykke I_{t+1} udelukkende ved størrelser fra år t. I (B.20) er dette ikke opfyldt, idet I_{t+1} også afhænger af C_{t+1}, men dette kan afhjælpes ved at indsætte (B.21) i (B.20):

$$I_{t+1} = \tfrac{1}{2}(0.8\,C_t + 0.8\,I_t + 4 - C_t) = -0.1\,C_t + 0.4\,I_t + 2. \tag{B.22}$$

(c) De 2 ligninger (B.21) og (B.22) kan sammenfattes i den affine matrixmodel

$$\begin{pmatrix} C_{t+1} \\ I_{t+1} \end{pmatrix} = \begin{pmatrix} 0.8 & 0.8 \\ -0.1 & 0.4 \end{pmatrix} \begin{pmatrix} C_t \\ I_t \end{pmatrix} + \begin{pmatrix} 4 \\ 2 \end{pmatrix}. \tag{B.23}$$

(d) Hvis vi i år 0 har et forbrug på 15 og ingen investering, dvs. $(C_0, I_0) = (15, 0)$, så kan forbrug og investeringer i de følgende år udregnes succesivt vha. (B.23) (angivet med 2 decimaler):

t	0	1	2	3	4	5	6
C_t	15.00	16.00	17.20	18.24	19.01	19.51	19.81
I_t	0.00	0.50	0.60	0.52	0.38	0.25	0.15

Det ser ud som om forbruget nærmer sig 20, mens investeringerne igen falder til 0. Dette antyder, at $\binom{C^*}{I^*} = \binom{20}{0}$ er en ligevægt for modellen. Ved at indsætte $\binom{C_t}{I_t} = \binom{20}{0}$ i (B.23) fås $\binom{C_{t+1}}{I_{t+1}} = \binom{20}{0}$, så $\binom{C^*}{I^*} = \binom{20}{0}$ er virkelig en ligevægt. Bemærk endvidere, at forbruget til at starte med var lavere end ligevægtsforbruget $C^* = 20$, hvilket førte til positiv investering.

Anvendelseseksempel B.6 Børn på feriekoloni

Vi betragter en feriekoloni, hvor 700 børn tilbringer sommeren. Børnene er dag for dag opdelt i tre hold med hver sin opgave:

HOLD 1	laver mad
HOLD 2	gør rent
HOLD 3	arrangerer leg, spil og sport.

Hver morgen afgøres det ved et terningkast for hver enkelt, om hun/han skal blive på sit hold eller skifte til et af de to andre. Reglerne er:

- Hvis et barn på HOLD 1 slår en sekser, skifter det til HOLD 3; hvis det slår en firer eller femmer, skifter det til HOLD 2; ellers bliver det på HOLD 1.
- Hvis et barn på HOLD 2 slår en sekser, skifter det til HOLD 3; hvis det slår en firer eller femmer, skifter det til HOLD 1; ellers bliver det på HOLD 2.
- Hvis et barn på HOLD 3 slår en sekser skifter det til HOLD 2; ellers skifter det til HOLD 1. (Man kan altså ikke blive på HOLD 3 to dage i træk.)

Vi vil beskrive udviklingen i antal børn på hver af holdene fra dag til dag. Lad x_{1t}, x_{2t} og x_{3t} betegne antal børn på dag t på hhv. HOLD 1, HOLD 2 og HOLD 3, og sæt

$$\mathbf{x}_t = \begin{pmatrix} x_{1t} \\ x_{2t} \\ x_{3t} \end{pmatrix}.$$

LA 229

Vi kan da beskrive sammenhængen mellem holdstørrelserne på dag $t+1$ og dag t på følgende vis:

$$\mathbf{x}_{t+1} = \mathbf{P}\mathbf{x}_t \qquad \text{hvor} \qquad \mathbf{P} = \begin{pmatrix} \frac{3}{6} & \frac{2}{6} & \frac{5}{6} \\ \frac{2}{6} & \frac{3}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} & 0 \end{pmatrix}.$$

I 1. søjle i matricen **P** optræder sandsynlighederne $\frac{3}{6}$, $\frac{2}{6}$ og $\frac{1}{6}$ for at overgå fra HOLD 1 til hhv. HOLD 1, HOLD 2 og HOLD 3. Tilsvarende for 2. og 3. søjle i **P**. Matricen **P** er altså en overgangsmatrix ifølge Definition B.9.7.

Ved ankomsten betragtes alle som værende på køkkenholdet, dvs. $\mathbf{x}_0 = \left(\begin{smallmatrix}700\\0\\0\end{smallmatrix}\right)$. Vi kan da succesivt beregne holdstørrelserne de følgende dage:

$$\mathbf{x}_1 = \begin{pmatrix} \frac{3}{6} & \frac{2}{6} & \frac{5}{6} \\ \frac{2}{6} & \frac{3}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} & 0 \end{pmatrix} \begin{pmatrix} 700 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 350 \\ 233 \\ 117 \end{pmatrix},$$

$$\mathbf{x}_2 = \begin{pmatrix} \frac{3}{6} & \frac{2}{6} & \frac{5}{6} \\ \frac{2}{6} & \frac{3}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} & 0 \end{pmatrix} \begin{pmatrix} 350 \\ 233 \\ 117 \end{pmatrix} = \begin{pmatrix} 350 \\ 253 \\ 97 \end{pmatrix},$$

$$\mathbf{x}_3 = \begin{pmatrix} \frac{3}{6} & \frac{2}{6} & \frac{5}{6} \\ \frac{2}{6} & \frac{3}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} & 0 \end{pmatrix} \begin{pmatrix} 350 \\ 253 \\ 97 \end{pmatrix} = \begin{pmatrix} 340 \\ 259 \\ 101 \end{pmatrix},$$

$$\mathbf{x}_4 = \begin{pmatrix} \frac{3}{6} & \frac{2}{6} & \frac{5}{6} \\ \frac{2}{6} & \frac{3}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} & 0 \end{pmatrix} \begin{pmatrix} 340 \\ 259 \\ 101 \end{pmatrix} = \begin{pmatrix} 340 \\ 260 \\ 100 \end{pmatrix},$$

$$\mathbf{x}_5 = \begin{pmatrix} \frac{3}{6} & \frac{2}{6} & \frac{5}{6} \\ \frac{2}{6} & \frac{3}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} & 0 \end{pmatrix} \begin{pmatrix} 340 \\ 260 \\ 100 \end{pmatrix} = \begin{pmatrix} 340 \\ 260 \\ 100 \end{pmatrix},$$

hvor vi (efter alle udregninger) har afrundet til hele tal ("Hele børn leger bedst"). Det ser ud til, at $\mathbf{x}_t$ nærmer sig en ligevægt $\mathbf{x}^*$ givet ved

$$\mathbf{x}^* = \begin{pmatrix} 340 \\ 260 \\ 100 \end{pmatrix}.$$

Vi kontrollerer, at denne vektor $\mathbf{x}^*$ faktisk *er* en ligevægt for modellen ved at foretage udregningen

$$\begin{pmatrix} \frac{3}{6} & \frac{2}{6} & \frac{5}{6} \\ \frac{2}{6} & \frac{3}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} & 0 \end{pmatrix} \begin{pmatrix} 340 \\ 260 \\ 100 \end{pmatrix} = \begin{pmatrix} 340 \\ 260 \\ 100 \end{pmatrix}.$$

Vi slutter af med at opsummere, hvad vi har foretaget os: Vi fandt frem til modellens ligevægt ved ud fra en given startfordeling at udregne holdstørrelserne de følgende dage. Herved blev vi i stand til at gætte på, at en vis vektor er en ligevægt, hvilket vi derefter kontrollerede.

LA 230

Anvendelseseksempel B.7 Aldersopdelt populationsvækst

(a) Når man ønsker at beskrive størrelsen af en dyrepopulation, er man ofte ikke kun interesseret i populationens samlede størrelse, men også i at kende antallet af individer i hver aldersklasse. Vi deler populationen op i 0-årige, 1-årige, 2-årige osv. og lader den sidste alderklasse bestå af dyr med en alder på mindst n år. (Nogle gange vil det dog være mere hensigtsmæssigt at regne med f.eks. uger eller timer i stedet for år.) Det er lettest kun at regne med antallene af hunner. Vi starter med at betragte populationen til tidspunktet $t = 0$ og beskriver populationen ved vektoren

$$\mathbf{v}_0 = \begin{pmatrix} x_{00} \\ x_{10} \\ \vdots \\ x_{n0} \end{pmatrix},$$

hvor x_{00} er antallet af 0-årige hunner i år 0, x_{10} er antallet af 1-årige hunner i år 0 osv. Mere generelt beskrives populationen i år t ved vektoren

$$\mathbf{v}_t = \begin{pmatrix} x_{0t} \\ x_{1t} \\ \vdots \\ x_{nt} \end{pmatrix},$$

hvor

- x_{0t} er antallet af 0-årige hunner i år t,
- x_{1t} er antallet af 1-årige hunner i år t,
- osv.
- x_{nt} er antallet af hunner med en alder på mindst n år i år t.

For at kunne beregne populationsvektoren $\mathbf{v}_{t+1}$ i år $t+1$ ud fra populationsvektoren $\mathbf{v}_t$ i år t har man brug for at kende fødsels- og overlevelsesraterne. Antag, at

- 0-årige hunner i gennemsnit får b_0 hununger om året (som overlever til året efter),
- 1-årige hunner i gennemsnit får b_1 hununger om året (som overlever til året efter),
- osv.
- hunner med en alder på mindst n år i gennemsnit får b_n hunner om året (som overlever til året efter).

Tallene $b_0, b_1, \ldots, b_n$ kaldes de *alderssspecifikke fødselsrater*. Da antallet $x_{0,t+1}$ af 0-årige hunner i år $t+1$ er lig med antallet af hununger født i år t, har vi

$$x_{0,t+1} = b_0 x_{0t} + b_1 x_{1t} + b_2 x_{2t} + \ldots + b_n x_{nt}.$$

Antag endvidere, at

- 0-årige hunner har sandsynligheden p_0 for at overleve til året efter,
- 1-årige hunner har sandsynligheden p_1 for at overleve til året efter,
- osv.
- hunner med en alder på mindst n år har sandsynligheden p_n for at overleve til året efter. (Hvis der ikke er nogen hunner, der bliver $n+1$ år, så vil $p_n = 0$.)

LA 231

Tallene $p_0, p_1, \ldots, p_n$ kaldes de *aldersspecifikke overlevelsesrater.* Da antallet $x_{1,t+1}$ af 1-årige hunner i år $t+1$ er lig med antallet af overlevende 0-årige hununger fra år t, har vi

$$x_{1,t+1} = p_0 x_{0t}.$$

Tilsvarende har vi

$$x_{2,t+1} = p_1 x_{1t}, \qquad x_{3,t+1} = p_2 x_{2t} \qquad \text{osv.},$$

mens

$$x_{n,t+1} = p_{n-1} x_{n-1,t} + p_n x_{nt},$$

idet hunnerne med en alder på mindst n år i år $t+1$ er de overlevende hununger, der i år t havde en alder på *mindst* $n-1$ år. Ovenstående ligninger kan sammenfattes til

$$\begin{aligned}
x_{0,t+1} &= b_0 x_{0t} + b_1 x_{1t} + \ldots + b_{n-1} x_{n-1,t} + b_n x_{nt}\\
x_{1,t+1} &= p_0 x_{0t}\\
x_{2,t+1} &= \qquad p_1 x_{1t}\\
&\vdots\\
x_{n,t+1} &= \qquad\qquad p_{n-1} x_{n-1,t} + p_n x_{nt}.
\end{aligned}$$

På matrixform kan dette skrives

$$\mathbf{v}_{t+1} = \mathbf{M}\mathbf{v}_t, \qquad \text{hvor} \qquad \mathbf{M} = \begin{pmatrix} b_0 & b_1 & \cdots & b_{n-1} & b_n \\ p_0 & 0 & \cdots & 0 & 0 \\ 0 & p_1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & p_{n-1} & p_n \end{pmatrix}.$$

For $n = 1$, 2 og 3 ser $\mathbf{M}$ derfor ud som følger:

$$\begin{pmatrix} b_0 & b_1 \\ p_0 & p_1 \end{pmatrix}, \qquad \begin{pmatrix} b_0 & b_1 & b_2 \\ p_0 & 0 & 0 \\ 0 & p_1 & p_2 \end{pmatrix} \qquad \text{og} \qquad \begin{pmatrix} b_0 & b_1 & b_2 & b_3 \\ p_0 & 0 & 0 & 0 \\ 0 & p_1 & 0 & 0 \\ 0 & 0 & p_2 & p_3 \end{pmatrix}.$$

Modellen $\mathbf{v}_{t+1} = \mathbf{M}\mathbf{v}_t$ kaldes en matrixmodel for aldersopdelt populationsvækst og matricen $\mathbf{M}$ kaldes ofte en *Lesliematrix.* I (c) vil vi se et mere konkret eksempel på denne model. Se endvidere Anvendelseseksempel B.1 og Anvendelseseksempel B.8.

(b) Modellen i (a) er beskrevet vha. de aldersspecifikke fødselsrater $b_0, b_1, \ldots, b_n$ og de aldersspecifikke overlevelsesrater $p_0, p_1, \ldots, p_n$. I økologi beskriver man sommetider aldersopdelt populationsvækst på en lidt anden måde: De aldersspecifikke fødselsrater benyttes stadigvæk, men i stedet for overlevelsesraterne benyttes tallene $l_0, l_1, \ldots, l_n$, hvor l_k er sandsynligheden på fødselstidspunktet for, at en hun bliver mindst k år (altså er i live, når den er k år gammel), dvs.

- $l_0 = 1$,
- l_1 er sandsynligheden på fødselstidspunktet for, at en hun bliver mindst 1 år,
- osv.
- l_n er sandsynligheden på fødselstidspunktet for, at en hun bliver mindst n år.

Vi vil her antage, at ingen dyr bliver mere end n år gamle, dvs. at $l_{n+1} = 0$, hvilket også kan udtrykkes ved, at overlevelsesraten p_n er lig med 0. Ofte angives værdierne af b_k og l_k i et skema, der kaldes en livstabel ("life-table" på engelsk) for populationen.

Sammenhængen mellem de aldersspecifikke overlevelsesrater $p_0, p_1, \ldots, p_n$ og tallene $l_0, l_1, \ldots, l_n$ er ganske simpel: Hvis en hun bliver mindst k år, betyder det, at den har overlevet fra år 0 til år 1, fra år 1 til år 2 osv. til og med at den har overlevet fra år $k-1$ til år k. Heraf følger det, at

$$l_k = p_0 \cdot p_1 \cdots p_{k-1} \qquad \text{for } k = 1, \ldots, n+1.$$

Denne formel udtrykker l_k ved overlevelsesraterne. Hvis man ønsker at gå den anden vej, kan man benytte formlen

$$p_k = \frac{p_0 \cdot p_1 \cdots p_k}{p_0 \cdot p_1 \cdots p_{k-1}} = \frac{l_{k+1}}{l_k} \qquad \text{for } k = 0, \ldots, n.$$

Endelig nævner vi, at man nogle gange i økologi også benytter de *aldersspecifikke mortalitetsrater* $q_0, q_1, \ldots, q_n$, hvor q_k angiver sandsynligheden for at en k-årig hun dør, inden hun bliver $k+1$ år. Mortalitetsraterne hænger simpelt sammen med overlevelsesraterne, idet der gælder $q_k = 1 - p_k$ (sandsynligheden for at dø er 1 minus sandsynligheden for at overleve).

(c) Vi vil nu illustrere (a) og (b) med et konkret eksempel. For en population af egern oplyses det, at

- den maksimale levealder er 4 år,
- de aldersspecifikke fødselsrater er $b_0 = 0$, $b_1 = 1.7$, $b_2 = 2.6$, $b_3 = 2.1$ og $b_4 = 0$,
- der gælder $l_0 = 1$, $l_1 = 0.4$, $l_2 = 0.24$, $l_3 = 0.12$ og $l_4 = 0.024$.

Livstabellen for populationen ser altså i dette tilfælde ud som følger:

k	b_k	l_k
0	0	1
1	1.7	0.4
2	2.6	0.24
3	2.1	0.12
4	0	0.024

Heraf fås

$$p_0 = \frac{l_1}{l_0} = \frac{0.4}{1} = 0.4, \qquad p_1 = \frac{l_2}{l_1} = \frac{0.24}{0.4} = 0.6,$$

$$p_2 = \frac{l_3}{l_2} = \frac{0.12}{0.24} = 0.5 \qquad \text{og} \qquad p_3 = \frac{l_4}{l_3} = \frac{0.024}{0.12} = 0.2.$$

Samtidig har vi $p_4 = 0$, da den maksimale levealder er 4 år. Ud fra disse oplysninger kan vi nu opskrive matrixmodellen

$$\begin{pmatrix} x_{0,t+1} \\ x_{1,t+1} \\ x_{2,t+1} \\ x_{3,t+1} \\ x_{4,t+1} \end{pmatrix} = \begin{pmatrix} 0 & 1.7 & 2.6 & 2.1 & 0 \\ 0.4 & 0 & 0 & 0 & 0 \\ 0 & 0.6 & 0 & 0 & 0 \\ 0 & 0 & 0.5 & 0 & 0 \\ 0 & 0 & 0 & 0.2 & 0 \end{pmatrix} \begin{pmatrix} x_{0t} \\ x_{1t} \\ x_{2t} \\ x_{3t} \\ x_{4t} \end{pmatrix},$$

hvor x_{kt} er antallet af k-årige hunner i år t for $k = 0, 1, 2, 3, 4$.

LA 233

Anvendelseseksempel B.8 Kaninpopulation opdelt i tre aldersklasser

(a) I Anvendelseseksempel B.1 opstillede vi modellen

$$\begin{pmatrix} x_{t+1} \\ y_{t+1} \end{pmatrix} = \begin{pmatrix} 2.0 & 1.5 \\ 0.7 & 0.4 \end{pmatrix} \begin{pmatrix} x_t \\ y_t \end{pmatrix},$$

hvor x_t er antallet af 0-årige hunkaniner og y_t er antallet af hunkaniner med en alder på mindst 1 år i år t.

Vi vil modificere modellen fra Anvendelseseksempel B.1 ved at dele kaninerne op i tre aldersklasser. Det viser sig, at 1-årige hunkaniner har såvel en højere overlevelsesrate som en højere fødselsrate end hunkaniner med en alder på mindst 2 år. Vi lader

- x_{0t} være antallet af 0-årige hunkaniner i år t,
- x_{1t} være antallet af 1-årige hunkaniner i år t,
- x_{2t} være antallet af hunkaniner med en alder på mindst 2 år i år t,

og lader

$$\mathbf{v}_t = \begin{pmatrix} x_{0t} \\ x_{1t} \\ x_{2t} \end{pmatrix}$$

være populationsvektoren i år t. Som i Anvendelseseksempel B.7 lader vi b_0, b_1 og b_2 være de aldersspecifikke fødselsrater og lader p_0, p_1 og p_2 være de aldersspecifikke overlevelsesrater. Som i Anvendelseseksempel B.1 har vi $b_0 = 2.0$ og $p_0 = 0.7$, og derudover har man observeret, at $b_1 = 1.6$ og $b_2 = 1.3$ samt $p_1 = 0.6$ og $p_2 = 0.3$. Modellen for aldersopdelt populationsvækst fra Anvendelseseksempel B.7 bliver derfor i dette tilfælde til

$$\begin{pmatrix} x_{0,t+1} \\ x_{1,t+1} \\ x_{2,t+1} \end{pmatrix} = \begin{pmatrix} 2.0 & 1.6 & 1.3 \\ 0.7 & 0 & 0 \\ 0 & 0.6 & 0.3 \end{pmatrix} \begin{pmatrix} x_{0t} \\ x_{1t} \\ x_{2t} \end{pmatrix}.$$

(b) Ifølge Sætning B.9.6 findes der en konstant c, således at

$$\frac{\mathbf{v}_t}{\lambda_1^t} \to c\,\mathbf{q}_1 \qquad \text{når } t \to \infty,$$

hvor λ_1 er den dominerende egenværdi og $\mathbf{q}_1$ er en egenvektor for $\mathbf{M}$ hørende til egenværdien λ_1. Endvidere nævnte vi efter beviset for Sætning B.8.2, at dette betyder, at der gælder

$$\mathbf{v}_{t+1} \approx \lambda_1 \mathbf{v}_t,$$

dvs.

$$x_{1,t+1} \approx \lambda_1 x_{1t}, \qquad x_{2,t+1} \approx \lambda_1 x_{2t} \qquad \text{og} \qquad x_{3,t+1} \approx \lambda_1 x_{3t}$$

for store værdier af t. Anderledes udtrykt gælder der altså

$$\frac{x_{1,t+1}}{x_{1t}} \simeq \lambda_1, \qquad \frac{x_{2,t+1}}{x_{2t}} \simeq \lambda_1 \qquad \text{og} \qquad \frac{x_{3,t+1}}{x_{3t}} \simeq \lambda_1$$

for store værdier af t, så vi kan bestemme tilnærmede værdier af λ_1 ved at udregne forholdene

$$\frac{x_{1,t+1}}{x_{1t}}, \qquad \frac{x_{2,t+1}}{x_{2t}} \qquad \text{og} \qquad \frac{x_{3,t+1}}{x_{3t}}$$

LA 234

for store værdier af t. Som startpopulation vælger vi

$$\mathbf{v}_0 = \begin{pmatrix} 100 \\ 0 \\ 0 \end{pmatrix}$$

(forsøg gerne selv vha. R med andre startpopulationer) og får da

$$\mathbf{v}_1 = \begin{pmatrix} 2.0 & 1.6 & 1.3 \\ 0.7 & 0 & 0 \\ 0 & 0.6 & 0.3 \end{pmatrix} \begin{pmatrix} 100 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 200 \\ 70 \\ 0 \end{pmatrix}$$

$$\mathbf{v}_2 = \begin{pmatrix} 2.0 & 1.6 & 1.3 \\ 0.7 & 0 & 0 \\ 0 & 0.6 & 0.3 \end{pmatrix} \begin{pmatrix} 200 \\ 70 \\ 0 \end{pmatrix} = \begin{pmatrix} 512 \\ 140 \\ 42 \end{pmatrix}$$

$$\mathbf{v}_3 = \begin{pmatrix} 2.0 & 1.6 & 1.3 \\ 0.7 & 0 & 0 \\ 0 & 0.6 & 0.3 \end{pmatrix} \begin{pmatrix} 512 \\ 140 \\ 42 \end{pmatrix} = \begin{pmatrix} 1303 \\ 358 \\ 97 \end{pmatrix}$$

$$\mathbf{v}_4 = \begin{pmatrix} 2.0 & 1.6 & 1.3 \\ 0.7 & 0 & 0 \\ 0 & 0.6 & 0.3 \end{pmatrix} \begin{pmatrix} 1303 \\ 358 \\ 97 \end{pmatrix} = \begin{pmatrix} 3304 \\ 912 \\ 244 \end{pmatrix}$$

$$\mathbf{v}_5 = \begin{pmatrix} 2.0 & 1.6 & 1.3 \\ 0.7 & 0 & 0 \\ 0 & 0.6 & 0.3 \end{pmatrix} \begin{pmatrix} 3304 \\ 912 \\ 244 \end{pmatrix} = \begin{pmatrix} 8385 \\ 2313 \\ 620 \end{pmatrix}$$

$$\mathbf{v}_6 = \begin{pmatrix} 2.0 & 1.6 & 1.3 \\ 0.7 & 0 & 0 \\ 0 & 0.6 & 0.3 \end{pmatrix} \begin{pmatrix} 8385 \\ 2313 \\ 620 \end{pmatrix} = \begin{pmatrix} 21276 \\ 5869 \\ 1574 \end{pmatrix} \quad \text{osv.}$$

(hvor vi har afrundet til hele tal, efter at udregningerne er udført). Vi udregner så forholdene

$$\frac{x_{1,t+1}}{x_{1t}} \qquad \text{for } t = 0, 1, 2, 3, 4, 5,$$

dvs. (med tre decimaler)

$$\frac{x_{11}}{x_{10}} = \frac{200}{100} = 2.000$$

$$\frac{x_{12}}{x_{11}} = \frac{512}{200} = 2.560$$

$$\frac{x_{13}}{x_{12}} = \frac{1303}{512} = 2.544$$

$$\frac{x_{14}}{x_{13}} = \frac{3304}{1303} = 2.537$$

$$\frac{x_{15}}{x_{14}} = \frac{8385}{3304} = 2.538$$

$$\frac{x_{16}}{x_{15}} = \frac{21276}{8385} = 2.538.$$

LA 235

Det ser altså ud til, at forholdene $\frac{x_{1,t+1}}{x_{1t}}$ stabiliserer sig omkring 2.538. Vi udregner dernæst de tilsvarende forhold

$$\frac{x_{2,t+1}}{x_{2t}} \quad \text{og} \quad \frac{x_{3,t+1}}{x_{3t}},$$

dvs.

$$\frac{x_{22}}{x_{21}} = \frac{140}{70} = 2.000$$
$$\frac{x_{23}}{x_{22}} = \frac{358}{140} = 2.560$$
$$\frac{x_{24}}{x_{23}} = \frac{912}{358} = 2.544$$
$$\frac{x_{25}}{x_{24}} = \frac{2313}{912} = 2.537$$
$$\frac{x_{26}}{x_{25}} = \frac{5869}{2313} = 2.538$$

($\frac{x_{21}}{x_{20}}$ er ikke defineret, idet $x_{20} = 0$) og

$$\frac{x_{33}}{x_{32}} = \frac{97}{42} = 2.300$$
$$\frac{x_{34}}{x_{33}} = \frac{244}{97} = 2.526$$
$$\frac{x_{35}}{x_{34}} = \frac{620}{244} = 2.542$$
$$\frac{x_{36}}{x_{35}} = \frac{1574}{620} = 2.537$$

($\frac{x_{31}}{x_{30}}$ og $\frac{x_{32}}{x_{31}}$ er ikke definerede, idet $x_{30} = x_{31} = 0$). Igen ses det, at forholdene stabiliserer sig omkring 2.538. Ud fra alle disse beregninger slutter vi, at λ_1 kan tilnærmes ved

$$\lambda_1 \simeq 2.538.$$

(Man kan vise, at der med 5 decimaler gælder $\lambda_1 = 2.53754$.) Antallet af hunkaniner i hver af de tre aldersklasser vokser altså omtrent med en faktor 2.538 hvert år.

(c) Når Sætning B.9.6 formuleres som

$$\mathbf{v}_t \approx c\lambda_1^t \mathbf{q}_1, \qquad \text{hvor} \qquad \mathbf{q}_1 = \begin{pmatrix} q_{11} \\ q_{21} \\ q_{31} \end{pmatrix},$$

dvs.

$$x_{1t} \approx c\lambda_1^t q_{11}, \qquad x_{2t} \approx c\lambda_1^t q_{21} \qquad \text{og} \qquad x_{3t} \approx c\lambda_1^t q_{31},$$

så følger det, at vektoren $\mathbf{v}_t$ er tilnærmelsesvist proportional med $\mathbf{q}_1$ for store værdier af t. (Proportionalitetsfaktoren er $c\lambda_1^t$.) Specielt gælder der

$$\frac{x_{1t}}{x_{3t}} \simeq \frac{q_{11}}{q_{31}} \qquad \text{og} \qquad \frac{x_{2t}}{x_{3t}} \simeq \frac{q_{21}}{q_{31}},$$

så vi kan bestemme tilnærmede værdier af

$$\frac{q_{11}}{q_{31}} \quad \text{og} \quad \frac{q_{21}}{q_{31}}$$

ved at udregne forholdene

$$\frac{x_{1t}}{x_{3t}} \quad \text{og} \quad \frac{x_{2t}}{x_{3t}}$$

for store værdier af t. For $t = 2, 3, 4, 5, 6$ (forholdene er ikke definerede for $t = 0$ og $t = 1$, idet $x_{30} = x_{31} = 0$) finder vi (med to decimaler)

$$\frac{x_{12}}{x_{32}} = \frac{512}{42} = 12.19, \qquad \frac{x_{22}}{x_{32}} = \frac{140}{42} = 3.33$$

$$\frac{x_{13}}{x_{33}} = \frac{1303}{97} = 13.48, \qquad \frac{x_{23}}{x_{33}} = \frac{358}{97} = 3.71$$

$$\frac{x_{14}}{x_{34}} = \frac{3304}{244} = 13.54, \qquad \frac{x_{24}}{x_{34}} = \frac{912}{244} = 3.74$$

$$\frac{x_{15}}{x_{35}} = \frac{8385}{620} = 13.52, \qquad \frac{x_{25}}{x_{35}} = \frac{2313}{620} = 3.73$$

$$\frac{x_{16}}{x_{36}} = \frac{21276}{1574} = 13.52, \qquad \frac{x_{26}}{x_{36}} = \frac{5869}{1574} = 3.73.$$

Det ses, at forholdene $\frac{x_{1t}}{x_{3t}}$ stabiliserer sig omkring 13.52, mens forholdene $\frac{x_{2t}}{x_{3t}}$ stabiliserer sig omkring 3.73, hvoraf vi slutter, at $\frac{q_{11}}{q_{31}}$ og $\frac{q_{21}}{q_{31}}$ kan tilnærmes ved

$$\frac{q_{11}}{q_{31}} \simeq 13.52 \quad \text{og} \quad \frac{q_{21}}{q_{31}} \simeq 3.73.$$

Sætter vi $q_{31} = 1$ fås dermed egenvektoren

$$\mathbf{q}_1 = \begin{pmatrix} 13.52 \\ 3.73 \\ 1 \end{pmatrix}.$$

Hvis vi dividerer $\mathbf{q}_1$ med $13.52 + 3.72 + 1 = 18.25$, fås egenvektoren

$$\begin{pmatrix} 0.74 \\ 0.20 \\ 0.06 \end{pmatrix},$$

hvor summen af de tre koordinater er lig med 1. Dette betyder, at der i det lange løb opstår en stabil fordeling mellem de tre aldersklasser med hhv. 74%, 20% og 6% af hunkaninerne i de tre klasser.

LA 23D

Opgaver

Opgave B.1

Lad

$$\mathbf{M} = \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix} \quad \text{og} \quad \mathbf{N} = \begin{pmatrix} 0 & 1 \\ -2 & 3 \end{pmatrix}.$$

(a) Udregn $\mathbf{M} + 2\mathbf{N}$ og $\mathbf{NM}$.

(b) Udregn determinanterne af $\mathbf{M}, \mathbf{N}$ og $\mathbf{NM}$. Kontrollér, at der gælder

$$\det(\mathbf{NM}) = \det \mathbf{N} \cdot \det \mathbf{M}.$$

(c) Bestem $\det(\mathbf{M}^2)$ og $\det(\mathbf{M}^5)$.
[Vink: Det er *ikke* nødvendigt at udregne $\mathbf{M}^2$ og $\mathbf{M}^5$.]

Opgave B.2

Udregn matricen $(\mathbf{M} + 2\mathbf{N})(\mathbf{M} + \mathbf{N}) - (\mathbf{M} + \mathbf{N})(\mathbf{M} + 2\mathbf{N})$, hvor

$$\mathbf{M} = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix} \quad \text{og} \quad \mathbf{N} = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}.$$

[Vink: Det kan betale sig at starte med at reducere udtrykket.]

Opgave B.3

(a) Løs ligningssystemet

$$\begin{cases} 2x - y = 5 \\ x + 3y = -1 \end{cases}$$

ved lige store koefficienters metode.

(b) Opskriv ligningssystemet på formen

$$\mathbf{M}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} p \\ q \end{pmatrix},$$

hvor $\mathbf{M}$ er en 2×2 matrix og $\binom{p}{q}$ en vektor.

(c) Løs ligningssystemet ved matrixinvertering.

Opgave B.4

Bestem den inverse til hver af matricerne

$$\begin{pmatrix} 2 & 3 \\ 5 & 7 \end{pmatrix}, \quad \begin{pmatrix} 8 & 6 \\ -6 & 8 \end{pmatrix} \quad \text{og} \quad \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix},$$

hvor a er et tal.