TEKST NR 200 1990

LINEÆR ALGEBRA OG ANALYSE

Noter til den naturvidenskabelige basisuddannelse

Mogens Niss

TEKSTER fra

IMFUFA ROSKILDE UNIVERSITETSCENTER
INSTITUT FOR STUDIET AF MATEMATIK OG FYSIK SAMT DERES FUNKTIONER I UNDERVISNING, FORSKNING OG ANVENDELSER

IMFUFA, Roskilde Universitetscenter, Postboks 260, 4000 Roskilde

LINEÆR ALGEBRA OG ANALYSE, Noter til den naturvidenskabelige basisuddannelse.

af: Mogens Niss

IMFUFA tekst nr. 200/90 135 sider ISSN 0106-6242

ABSTRACT

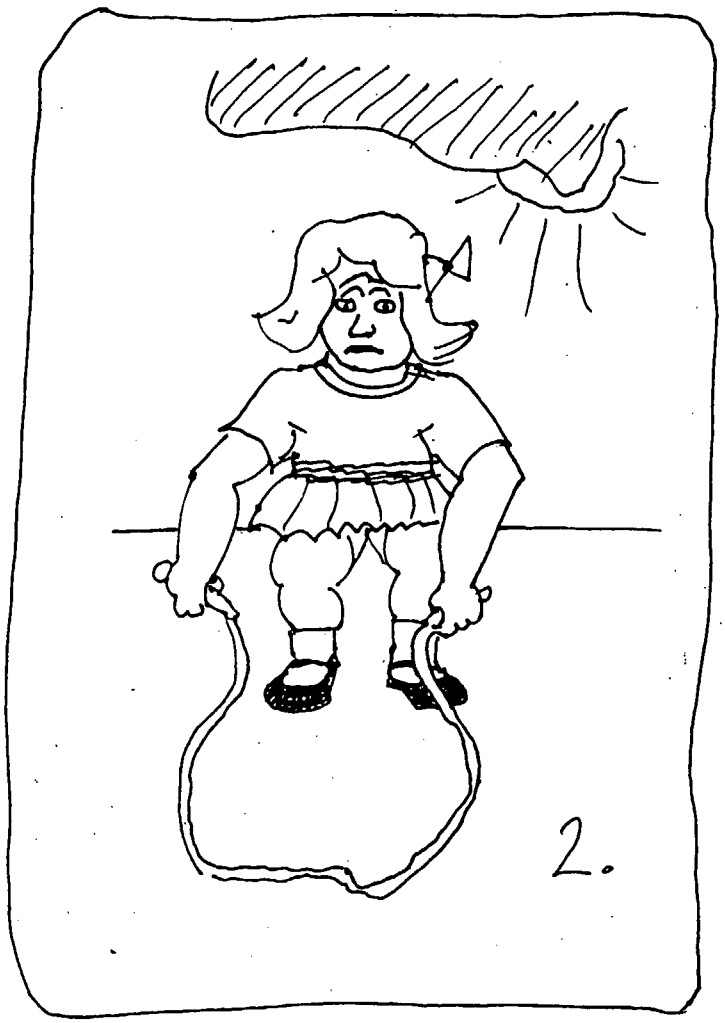
Den foreliggende tekst "Lineær algebra og analyse" er udarbejdet som noter til et pulje A kursus i matematik ved den naturvidenskabelige basisuddannelse. Noterne er blevet til i et antal omgange, idet der tidligere har foreligget foreløbige udgaver. Til noterne knytter sig en opgavesamling som udsendes som et separat hæfte.

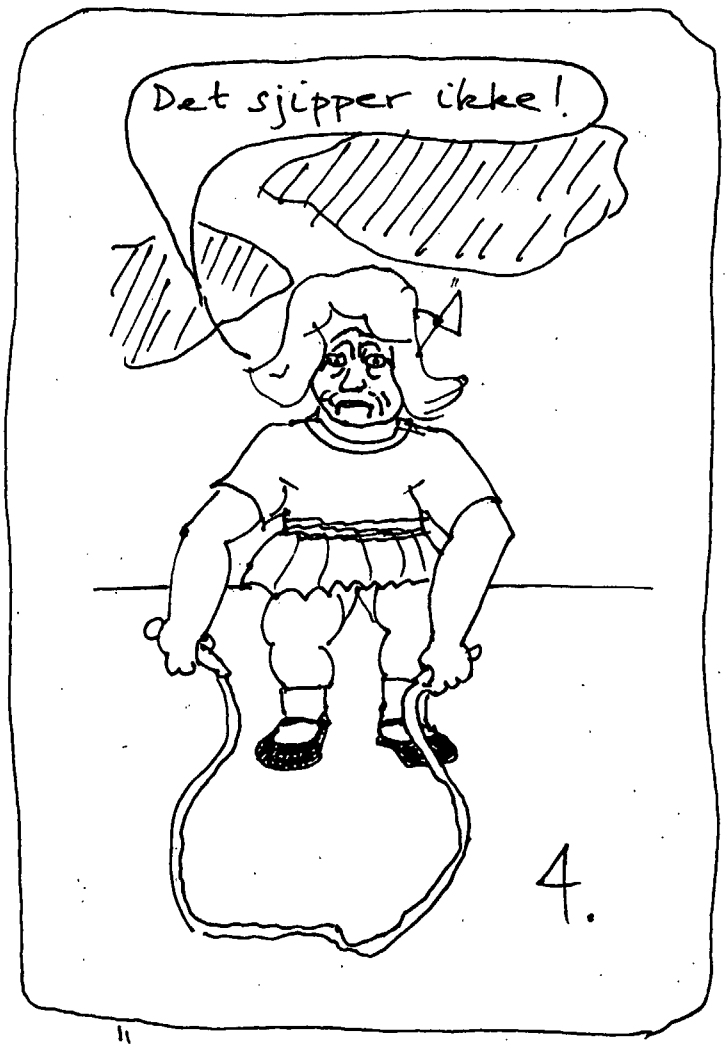
Hensigten med teksten er at give en indføring i grundlæggende dele af den lineære algebra og teorien for lineære differentialligningssystemer, begge dele motiveret i og bragt i spil over for forskellige anvendelsesfelter inden for andre naturvidenskabelige fagområder. Det er samtidig tilstræbt at fremstillingen er matematisk stringent - ikke at forveksle med kedelig - sådan at alle påstande bevises (om end ordet "bevis" ikke er strøet ud over siderne), med mindre dette udtrykkeligt ikke er tilfældet.

Hovedlinjerne i indholdet fremgår af indholdsfortegnelsen. Den lineære algebra udspiller sig i begyndelsen udelukkende i reelle talrum. Siden overføres dele af det opstillede apparat til funktionsrum af hensyn til teorien for lineære differentialligningssystemer, der derefter uden videre kan skumme fløden af den præsenterede lineære algebra.

Mogens Niss
september 1990

INDHOLD





Efter "Radiserne"

Ingers Nis
5.9.1990

MOTIVERENDE INDLEDNING TIL BESKÆFTIGELSEN MED LINEÆR ALGEBRA

Meget kort kan man sige, at den matematiske emnekreds **Lineær Algebra** beskæftiger sig med studiet af fænomenerne/begreberne **proportionalitet** og **additivitet**. Begge dele er grundlæggende både for en kvantitativ (matematisk) behandling af dele af *virkeligheden*, og for opbygningen af en lang række *matematiske discipliner*. **Linearitet**, som er det begreb der sammenfatter proportionalitet og additivitet, er således væsentlig såvel for matematikkens *anvendelse* gennem *matematisk modelbygning* som for *matematikken selv*.

Vi lægger ud med at undersøge nogle simple anvendelseseksempler, som skal give stof og inspiration til den begrebsopbygning vi snart skal begive os ud i.

Eksempel 0: KAGEBAGNING

Lad os allerførst se på en hjemlig situation. Til en fest skal der bages en marmorkage og nogle vaniljekranse. I følge opskriften skal der til marmorkagen bruges 250g smør, 250g mel, 250g sukker, 5 æg (ca. 250g) og 30g kakao. Efter bagningen, der jo giver anledning til et vist svind på grund af fordampning, er der ca. 900g sandkage. Til vaniljekransene benyttes 180g smør, 250g mel, 125g sukker, 1 æg (ca. 50g) samt vaniljekorn (ca. 5g). Heraf kommer ca. 550g færdige småkager.

For at kunne overskue situationen omregner vi råvareforbrugene til 1kg færdig kage af hver slags. Vi kan derefter sammenfatte råvareforbruget til de to slags kager i et skema, hvor alle tal angiver gram:

	marmorkage	vaniljekranse
smør	278	327
mel	278	455
sukker	278	227
æg	278	91
kakao	33	0
vaniljekorn	0	9

Hvis det nu besluttes at lave nogle andre mængder af marmorkage og småkager kan vi gøre forbruget af ingredienser op ved hjælp af skemaet. Skal der bages x_1 kg marmorkage og x_2 kg vaniljekranse, er der brug for (i gram):

smør	$278x_1+327x_2$
mel	$278x_1+455x_2$
sukker	$278x_1+227x_2$
æg	$278x_1+\ 91x_2$

kakao	$33x_1+$	$0x_2$
vaniljekorn	$0x_1+$	$9x_2$

Nu kunne det jo tænkes, at der ikke er tid - eller penge - til at købe ind til kagerne, men at vi må klare os med det vi har i skabet. Der er måske (i gram)

smør:	350
mel:	800
sukker:	500
æg:	300
kakao:	100
vanilje:	10

Spørgsmålet "Kan vi bage marmorkage og småkager, sådan at vi lige netop bruger alle ingredienserne op?"er måske ikke så realistisk i en almindelig husholdning (mere herom lige straks), men matematisk formuleret lyder det: Findes x_1 og x_2, så at

$$\begin{aligned}
350 &= 278x_1+327x_2 \\
800 &= 278x_1+455x_2 \\
500 &= 278x_1+227x_2 \\
300 &= 278x_1+\ 91x_2 \\
100 &= 33x_1+\ \ 0x_2 \\
10 &= 0x_1+\ \ 9x_2?
\end{aligned}$$

Der er altså tale om et system af 6 ligninger med de to ubekendte x_1 og x_2. Nu er svaret på det stillede spørgsmål tydeligvis nej, idet man af de to sidste ligninger straks kan se, at de eneste mulige kandidater er $x_1 = \frac{100}{33}$ og $x_2 = \frac{10}{9}$, som imidlertid ikke passer ind i nogen af de andre ligninger. Det skulle da også være mærkeligt om de kageingredienser man tilfældigvis har på lager netop skulle kunne bruges op i bagningen af marmorkage og vaniljekranse.

For en privat husholdning ville et mere relevant spørgsmål være, hvordan vi med de forhåndenværende ingredienser kan bage marmorkage og vaniljekranse på den "bedste måde". Det er ikke helt klart hvad "bedste måde"skal betyde. Det afhænger af vores interesser i sagen og en række dertil knyttede valg. Én betydning kunne være, at det samlede restlager af ingredienser efter bagningen er så lille som muligt (vi skulle måske ud at rejse). En måske mere nærliggende mulighed var at få bagt så meget som muligt i alt (vi vil gerne hælde så meget kage som vi kan fremskaffe i halsen på gæsterne). Begge problemer kan formuleres matematisk. Vi nøjes med det sidste (forsøg selv med det første, bagefter):

Vi søger den eller de kombinationer af (ikke-negative, naturligvis) x_1 og x_2 (hvis der overhovedet findes nogen), som maksimerer $x_1 + x_2$ under den betingelse at vi ikke

bruger flere ingredienser end vi har, dvs. så at

$$\begin{aligned}
350 &\geq 278x_1+327x_2 \\
800 &\geq 278x_1+455x_2 \\
500 &\geq 278x_1+227x_2 \\
300 &\geq 278x_1+\ 91x_2 \\
100 &\geq \ \ 33x_1+\ \ 0x_2 \\
10 &\geq \ \ \ 0x_1+\ \ 9x_2
\end{aligned}$$

Dette problem er et eksempel på en meget generel klasse af problemer af stor praktisk betydning, de såkaldte **lineære programmeringsproblemer**. De væsentligste metoder til behandlingen af sådanne problemer stammer fra den lineære algebra.

I det foreliggende simple tilfælde er det ikke påkrævet med de store matematiske armsving - men nok nogle små - for at konkludere, at det optimale er at producere 1.009 kg marmorkage og 0.213 kg vaniljekranse, hvis det gælder om at have mest muligt gods til gæsterne. Af skemaet ovenfor kan vi dernæst uden videre beregne forbruget af de enkelte ingredienser.

Eksempel 1: BAGERIET

Vi skal nu løfte blikket noget og undersøge en **generalisering** af det simple eksempel vi netop har behandlet. Tankegangen er faktisk den samme som i Eksempel 0, men fordi problemstillingen formuleres mere generelt kommer der nye momenter ind i sagen.

Et industribageri producerer forskellige slags brød og kager ved anvendelsen af forskellige slags råvarer: diverse melsorter, sukker, smør, margarine, gær, kakao m.m. Lad os sige at der i bageriet benyttes i alt n *forskellige råvarer*, og at der produceres i alt p *forskellige produkter*.

Ser vi først på *produkt nr. 1* kan vi tænke os, at der for at producere 1 kg af varen skal benyttes a_{11} kg af råvare nr. 1, a_{21} kg af råvare nr. 2, og så fremdeles op til a_{n1} kg af råvare nr. n. I tallet a_{k1}, $k = 1, 2, \ldots, n$, står altså det første index, k, for råvarenummeret, mens det andet index, 1, står for produktnummeret. Selve tallet a_{k1} står så for det antal kg af råvare nr. k der indgår i opskriften for 1 kg af produkt nr. 1.

På tilsvarende måde kan vi anskue situationen for de øvrige produkter. Hvis k *betegner råvarenummeret* og m *betegner produktnummeret* skal der *til 1 kg af produkt nr.* m *bruges* a_{km} *kg af råvare nr.* k. Bemærk, at det ikke er givet at summen af råvarerne, $a_{1m}+a_{2m}+\cdots+a_{nm}$, er lig 1, fordi produktionsprocessen (bagningen) ikke nødvendigvis

finder sted under massebevarelse, jfr. Eksempel 0. Bemærk også at i følge sagens natur må alle $a_{km} \geq 0$. Denne restriktion er midlertid ikke væsentlig for det principielle i sagen.

Vi kan sammenfatte situationen ved hjælp af en tabel. Hver **søjle** repræsenterer et bestemt produkt, hver **række** en bestemt råvare. Tallene i en søjle angiver så hvor mange kg af de forskellige råvarer der skal til for at producere 1 kg af produktet med det pågældende søjlenummer:

	1	2	...	p
1	a_{11}	a_{12}	...	a_{1p}
2	a_{21}	a_{22}	...	a_{2p}
$\vdots$	$\vdots$	$\vdots$		$\vdots$
n	a_{n1}	a_{n2}	...	a_{np}

Hvis nu bageriet ønsker at fremstille x_1 kg af produkt nr. 1, x_2 kg af produkt nr. 2 osv. op til x_p kg af produkt nr. p, kan vi bestemme hvor meget der skal benyttes af hver af de n råvarer. Af råvare nr. 1 skal der bruges i alt (kg)

$$y_1 = a_{11}x_1 + a_{12}x_2 + ... + a_{1p}x_p.$$

Af råvare nr. 2:

$$(*) \qquad y_2 = a_{21}x_1 + a_{22}x_2 + ... + a_{2p}x_p,$$

osv. indtil

$$y_n = a_{n1}x_1 + a_{n2}x_2 + ... + a_{np}x_p.$$

Vi kan repræsentere situationen ved en kortfattet og overskuelig skrivemåde, en art stenogram:

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{np} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix}$$

(læg mærke til, at fordi n og p ingenlunde behøver at være ens, er antallet af elementer i x-søjlen ikke nødvendigvis det samme som i y-søjlen, selv om det typografisk kunne se sådan ud).

Diagrammet ovenfor skal fortolkes således: Bageriets *opskrifter* på brød og kager repræsenteres af a-erne i skemaet, hvor hver søjle angår et bestemt produkt. For i bageriet at producere de varemængder som repræsenteres af x-søjlen skal der benyttes de råvaremængder som repræsenteres af y-søjlen. Vi kan altså fortolke a-skemaet som

produktionsgangen, x-søjlen som det ønskede *produktionsresultat*, y-søjlen som det dertil svarende *råvarebehov*.

Hvad nu, hvis bageriet den ene dag ønsker at producere varer svarende til søjlen

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix}, \quad \text{med råvarebehovet} \quad \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

og den anden dag svarende til søjlen

$$\begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_p \end{pmatrix}, \quad \text{med råvarebehovet} \quad \begin{pmatrix} y'_1 \\ y'_2 \\ \vdots \\ y'_n \end{pmatrix}$$

hvad er så det samlede råvarebehov for de to dage? Ja det er jo trivielt. Det samlede råvarebehov er selvfølgelig

$$\begin{pmatrix} y_1 + y'_1 \\ y_2 + y'_2 \\ \vdots \\ y_n + y'_n \end{pmatrix}$$

Med denne indsats af råvarer fremstilles i alt produkterne

$$\begin{pmatrix} x_1 + x'_1 \\ x_2 + x'_2 \\ \vdots \\ x_p + x'_p \end{pmatrix}$$

Det er på den baggrund nærliggende at **aftale en konvention for addition** (en aftale er nødvendig; det står ikke i Grundloven, hvordan man skal lægge søjler sammen):

Søjler af *samme dimension* kan adderes. Det sker pladsvis:

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix} + \begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_p \end{pmatrix} = \begin{pmatrix} x_1 + x'_1 \\ x_2 + x'_2 \\ \vdots \\ x_p + x'_p \end{pmatrix}$$

og tilsvarende for y-søjlerne.

I stedet for at udføre produktionen to forskellige dage, med adskilte sæt råvarer, henholdsvis produkter, kunne den have været udført samme dag (vi går ud fra at bageriets kapacitet - maskiner, ovne, arbejdskraft osv. - tillader det). Opskrifterne, altså a-skemaet, er jo de samme, så forholdene må kunne sammenfattes i skemaet:

$$\begin{pmatrix} y_1 + y_1' \\ y_2 + y_2' \\ \vdots \\ y_n + y_n' \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{np} \end{pmatrix} \begin{pmatrix} x_1 + x_1' \\ x_2 + x_2' \\ \vdots \\ x_p + x_p' \end{pmatrix}$$

Samtidig har vi jo imidlertid efter den netop opstillede konvention, at

$$\begin{aligned} \begin{pmatrix} y_1 + y_1' \\ y_2 + y_2' \\ \vdots \\ y_n + y_n' \end{pmatrix} &= \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} + \begin{pmatrix} y_1' \\ y_2' \\ \vdots \\ y_n' \end{pmatrix} \\ &= \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{np} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix} + \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{np} \end{pmatrix} \begin{pmatrix} x_1' \\ x_2' \\ \vdots \\ x_p' \end{pmatrix} \end{aligned}$$

Ved at sammenholde de to udtryk for $y + y'$-søjlen slutter vi, at

$$\begin{aligned} &\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{np} \end{pmatrix} \begin{pmatrix} x_1 + x_1' \\ x_2 + x_2' \\ \vdots \\ x_p + x_p' \end{pmatrix} = \\ &\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{np} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix} + \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{np} \end{pmatrix} \begin{pmatrix} x_1' \\ x_2' \\ \vdots \\ x_p' \end{pmatrix} \end{aligned}$$

Argumentet for det resultat vi netop er nået frem til trak på vores forestillinger om bagning osv. Men resultatet kan også efterprøves ved direkte udregning, ud fra betydningen af skemaopstillingen, jfr. (*) side 4. Den typiske plads, lad os sige den i'te, i søjlen der kommer ud af venstresiden, er jo efter forskriften side 4

$$\begin{aligned} &a_{i1}(x_1 + x_1') + a_{i2}(x_2 + x_2') + \dots + a_{ip}(x_p + x_p') \\ =&(a_{i1}x_1 + a_{i2}x_2 + \dots + a_{ip}x_p) + (a_{i1}x_1' + a_{i2}x_2' + \dots + a_{ip}x_p'). \end{aligned}$$

På den højreside af lighedstegnet får vi på den i'te plads i søjlen svarende til første klump

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{ip}x_p.$$

Tilsvarende for den anden klump

$$a_{i1}x_1' + a_{i2}x_2' + \cdots + a_{ip}x_p'.$$

I alt bliver den i'te plads i resultatsøjlen for højresiden

$$(a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{ip}x_p) + (a_{i1}x_1' + a_{i2}x_2' + \cdots + a_{ip}x_p'),$$

som netop stemmer overens med resultatet for venstresiden.

Nu har vi allerede indført én operation, **addition på søjler** af reelle tal. Vi har altså udvidet det sædvanlige additionsbegreb for tal til at omfatte en ny slags objekter, søjler af tal. Men vi har også set at vi kunne indføre en slags operation af et rektangulært talskema (a-skemaet) på en talsøjle, idet vi jo ved at foreskrive et bestemt produktionsresultat (x-søjlen) kan beregne det dertil svarende råvarebehov (y-søjlen), givet et sæt opskrifter. Resultatet af operationen af et talskema på en søjle (der skal have lige så mange elementer som talskemaet har søjler) bliver en ny søjle (med lige så mange elementer som talskemaet har rækker). Vi så at denne operation er **additiv**: hvis vi opererer på summen af to søjler, bliver resultatet summen af de resultater der kommer ud af at operere på de to søjler hver for sig. Det som opererer på søjler (altså på sæt af produkter) er talskemaet bestående af a-erne (altså produktionsgangen, opskrifterne).

Også andre reguleringer af bageriets produktion er nærliggende. Det kunne f.eks. tænkes at man af en eller anden grund (f.eks. sæsonsvingninger) ønskede at øge eller mindske hele produktionen med en bestemt faktor a, som vi for fortolkningens skyld må have til at være ikke-negativ. I stedet for at indstille produktionen på at fremstille søjlen

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix}$$

repræsenterende bageriets forskellige varer, vil man producere søjlen

$$\begin{pmatrix} ax_1 \\ ax_2 \\ \vdots \\ ax_p \end{pmatrix}$$

Hele produktionen skal altså være a gange så stor som før (obs! a kan godt være mindre end 1) - den skal finde sted i en ny **skala**. Så ville det være naturligt at aftale

skrivemåden

$$a\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix} \quad \text{for} \quad \begin{pmatrix} ax_1 \\ ax_2 \\ \vdots \\ ax_p \end{pmatrix}$$

Vi siger at den nye søjle er **proportional** med den gamle. Vi har faktisk dermed indført en ny operation på søjler: multiplikation med et reelt tal. Tallet kaldes i sammenhængen en **skalar**, fordi det bestemmer skalaen af produktionen. Operationen kaldes rimeligt nok **skalarmultiplikation**. Til forskel fra sædvanlig multiplikation med tal, der jo ud fra to objekter af samme art (tal) skaffer et tredje objekt af samme art, vedrører skalarmultiplikation to objekter af forskellig art, et tal og en søjle, og resultatet af multiplikationen bliver en søjle.

Vi vil nu undersøge hvilke konsekvenser kravet om en a gange så stor produktion får på råvarebehovet. Overvejer man bageprocesser et øjeblik er det en rimelig antagelse (men ikke nogen naturlov), at hvis produktionen skal skaleres med faktoren a må råvareforbruget også skaleres med faktoren a for alle råvarer

$$a\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} ay_1 \\ ay_2 \\ \vdots \\ ay_n \end{pmatrix}$$

hvoraf - idet vi jo har

$$\begin{pmatrix} ay_1 \\ ay_2 \\ \vdots \\ ay_n \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{np} \end{pmatrix} \begin{pmatrix} ax_1 \\ ax_2 \\ \vdots \\ ax_p \end{pmatrix}$$

og

$$a\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = a\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{np} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix}$$

- det sluttes at

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{np} \end{pmatrix} \begin{pmatrix} ax_1 \\ ax_2 \\ \vdots \\ ax_p \end{pmatrix} = a\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{np} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix}.$$

Også dette resultat kan skaffes ved regulær udregning! Den i'te plads i venstresidens søjle er

$$a_{i1}(ax_1) + a_{i2}(ax_2) + \cdots + a_{ip}(ax_p) = a(a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{ip}x_p).$$

Indholdet i parentesen er netop hvad der står på den i'te plads i søjlen svarende til højresiden før multiplikationen med a.

Dette viser at den indførte operation mellem en søjle og et talskema er en **proportionalitet**, dvs. den fører en skaleret søjle over i den pågældende skalar gange med resultatet af operationen på søjlen før skalering.

Indskud med visse konklusioner på Eksempel 1

Ud fra det foregående kan vi se, at vores betragtninger over hvordan bageriets produktion foregår, fører til at vi må hæfte os ved addition og skalarmultiplikation af talsøjler, såvel som ved en operation mellem et talskema og en talsøjle, som har den egenskab at være additiv og skalarmultiplikativ.

Selv om den valgte skrivemåde med talskemaer og -søjler giver en forenkling i forhold til den oprindelige skrivemåde a la (*), er der dog stadig meget at skrive (det føler man ikke mindst når man selv har tæsket i klaviaturet). En yderligere forenkling ville være ønskelig.

Vi aftaler derfor (noget énvejs, må jeg indrømme), at vi i stedet for x-søjler og y-søjler - som altså stadigvæk kan have forskellig længde - skriver $\underline{x}$ og $\underline{y}$, uanset længden, og i stedet for a-skemaet skriver symbolet A. Så kan vi kort notere det hidtil fundne som

$$\underline{y} = A\underline{x},$$

og, idet vi har indført hvad $\underline{x} + \underline{x}'$ og $a\underline{x}$ (a en skalar) skal betyde:

$$A(\underline{x} + \underline{x}') = A\underline{x} + A\underline{x}'$$
$$A(a\underline{x}) = aA\underline{x}.$$

Analogien med sædvanlig regning med tal og parenteser er tydelig. Der er simplethen tale om en *generalisation* heraf, for hvis A, $\underline{x}$ og $\underline{x}'$ alle kun består af ét tal står der simpelthen tal i formellinjerne ovenfor.

Vi vil kalde et talskema A, bestående af et antal *rækker* (n) og et antal *søjler* (p), for en **matrix**, nærmere bestemt en $n \times p$-**matrix**. Søjler er en speciel slags talskemaer, matricer med kun én søjle. De hidtil betragtede x-søjler er da $p \times 1$-matricer, mens

y-søjlerne er $n \times 1$-matricer. Operationen af talskemaer (matricer) på søjler (en speciel slags matricer) er et specialtilfælde af **matrixmultiplikation**. På dette sted skal vi ikke forfølge dette anliggende nærmere; det vil der blive rigelig anledning til senere i kurset. Vi skal dog gøre én vigtig observation allerede nu: Hvis en matrix skal multipliceres med en søjle (fra højre), må søjlen have præcis så mange rækker (elementer) som matricen har søjler. Resultatet af multiplikationen bliver en søjle med netop så mange rækker som matricen har rækker. I uformelt symbolsprog: "$n \times p$"$\cdot$"$p \times 1$"$=$ "$n \times 1$".

Vi afslutter omtalen af bageri-eksemplet med et par kommentarer. Det vi hidtil har foretaget os i eksemplet kan sammenfattes således: Til et givet input (her kravet til produktionsresultatet, $\underline{x}$) frembringes et output (her råvareforbruget, $\underline{y}$) ved hjælp af en bestemt operation (her oversættelsen af produktionskrav til råvarekrav, repræsenteret af matrixidentiteten $\underline{y} = A\underline{x}$). Imidlertid kan man i industribageriet spørge: Givet en bestemt mængde råvarer, altså en søjle $\underline{y}$, hvilken produktion kan da realiseres med disse råvarer?

Hvis råvarerne skal bruges op inden for den betragtede periode - hvilket ikke er så urealistisk et krav, når man betænker problemerne med friskhed, investeringer, lagerplads osv. - kan spørgsmålet formuleres:

Givet $\underline{y}$, findes der da et $\underline{x}$, så at $\underline{y} = A\underline{x}$? Skriver vi dette spørgsmål ud i detaljer, i overensstemmelse med $(*)$ side 4 lyder det: Givet

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix},$$

findes da

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix}$$

så at

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1p}x_p &= y_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2p}x_p &= y_2 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{np}x_p &= y_n? \end{aligned}$$

Hvis svaret er ja (det er det faktisk ikke altid, som Eksempel 0 viste), hvordan ser da sådanne sæt $(x_1, x_2, \ldots, x_p)$ ud, og hvor mange slags er der? Hvis der er flere sæt,

hvilke(t) er så det "bedste"i en eller anden henseende som må være præciseret nærmere? Vi har med andre ord at gøre med et ligningssystem. På grund af ligningssystemets struktur, jf. behandlingen ovenfor, kaldes det for et **lineært ligningssystem.** Undersøgelserne af sådanne er en væsentlig sag for den lineære algebra.

Er det nu ikke noget krav at råvarerne bruges op inden for den betragtede periode, men blot at der ikke bruges mere af nogen råvare end der er til rådighed, kan man stille spørgsmålet:

Givet $\underline{y}$, findes der da et $\underline{x}$, så at $\underline{y} \geq A\underline{x}$? I detaljer:

$$\begin{aligned}
&a_{11}x_1 + a_{12}x_2 + \ldots + a_{1p}x_p \leq y_1 \\
&a_{21}x_1 + a_{22}x_2 + \ldots + a_{2p}x_p \leq y_2 \\
&\vdots \\
&a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{np}x_p \leq y_n.
\end{aligned}$$

Igen kan man, hvis der er flere løsninger $\underline{x}$ (og det er der ofte), spørge hvilket der i en eller anden henseende er "bedst". I så fald har vi at gøre med et **lineært programmeringsproblem.** Også en sag for den lineære algebra.

Eksempel 2: AFSTEMNING AF KEMISKE LIGNINGER

At lineære ligningssystemer optræder i andre sammenhænge end ved brød- og kagefremstilling, f.eks. inden for kemien, skal belyses i dette eksempel.

Vi betragter en simpel kemisk reaktion, hvis kvalitative indhold antages kendt, men hvor koefficienterne i første omgang er ubestemte:

$$?HgI_4^{--} + ?OH^- + ?NH_3 \rightarrow ?Hg_2NI + ?I^- + ?H_2O.$$

Opgaven er nu at sætte koefficienter ind i denne ligning. I gymnasiet benytter man sædvanligvis en metode baseret på de såkaldte iltnings- eller oxidationstrin. Denne metode er imidlertid ofte lidt uigennemskuelig, fordi den hviler på en række spilleregler der ikke alle har et klart fysisk-kemisk grundlag. Det viser sig nu at man kan afstemme denne og alle andre reaktionsligninger ud fra ét alment princip: Antallet af atomer af hvert stof skal være det samme på begge sider af reaktionspilen, og tilsvarende med antallet af ladninger (regnet med fortegn, naturligvis). Kalder vi de seks ubekendte koefficienter for $x_1, x_2, \ldots, x_6$ kan vi skrive reaktionsligningen

$$x_1HgI_4^{--} + x_2OH^- + x_3NH_3 \rightarrow x_4Hg_2NI + x_5I^- + x_6H_2O.$$

hvor opgaven er at bestemme disse x-er.

Hvis antallet af atomer af et givet grundstof skal være det samme på begge sider, må først for Hg's vedkommende

$$x_1 = 2x_4.$$

Tilsvarende for I:

$$4x_1 = x_4 + x_5.$$

Fortsættes med henholdsvis O, H og N får vi ligningerne

$$x_2 = x_6, \quad x_2 + 3x_3 = 2x_6, \quad x_3 = x_4.$$

Nu resterer der kun at holde regnskab med ladningerne, hvilket giver

$$2x_1 + x_2 = x_5.$$

Fordi dette ligningssystem (tilfældigvis?) er så simpelt, kan det hurtigt reduceres til et ligningssystem i få ubekendte som det er let at finde direkte. Alligvel vil vi for at bane vejen for behandlingen af mere komplicerede situationer skrive systemet alment op, nemlig således:

$$\begin{aligned}
x_1 \qquad\qquad - 2x_4 \qquad\qquad &= 0\\
4x_1 \qquad\qquad - \ x_4 - x_5 \qquad &= 0\\
x_2 \qquad\qquad\qquad - \ x_6 &= 0\\
x_2 + 3x_3 \qquad\qquad - 2x_6 &= 0\\
x_3 - \ x_4 \qquad\qquad &= 0\\
2x_1 + x_2 \qquad\qquad - x_5 \qquad &= 0
\end{aligned}$$

Benytter vi den notation vi udviklede i det foregående eksempel, har vi at gøre med ligningen på matrix-form:

$$\begin{pmatrix} 1 & 0 & 0 & -2 & 0 & 0 \\ 4 & 0 & 0 & -1 & -1 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 \\ 0 & 1 & 3 & 0 & 0 & -2 \\ 0 & 0 & 1 & -1 & 0 & 0 \\ 2 & 1 & 0 & 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

eller kort $A\underline{x} = \underline{0}$, hvor A er den opskrevne 6×6-matrix.

Hvad angår løsningen af ligningssystemet er det uden videre klart, at sættet $(x_1, x_2, \ldots, x_6) = (0, 0, \ldots, 0)$ - vi burde skrive begge sæt som søjler - bestående af lutter nuller er en løsning. Fra et kemisk synspunkt er den ikke videre ophidsende, al den stund ingen af de pågældende stoffer så overhovedet ville foreligge. Det er imidlertid også klart, at hvis $(x_1^0, x_2^0, \ldots, x_6^0)$ er én løsning til systemet, vil også ethvert skalarmultiplum $a(x_1^0, x_2^0, \ldots, x_6^0)$ (hvor a er en skalar) af dette sæt være en løsning. Dette beror selvsagt på at højresiden af ligningssystemet er nulsøjlen. Ellers ville det netop sagte have været forkert.

Ved hjælp af den lineære algebra bliver vi i stand til at afklare løsningsforholdene for lineære ligningssystemer, også det ovenstående. Det vil da vise sig (prøv selv at indse det!), at samtlige løsninger til ligningssystemet har formen

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix} = t \begin{pmatrix} 2 \\ 3 \\ 1 \\ 1 \\ 7 \\ 3 \end{pmatrix},$$

hvor t er et vilkårligt reelt tal.

Nu giver det for en realistisk kemisk fortolkning kun mening at operere med t-er hentet fra de naturlige tal. Og her siger tallet 1 det væsentlige. Alle andre naturlige tal giver blot et multiplum af alle de indgående kemiske stoffer.

I alt får den afstemte ligning skikkelsen

$$2HgI_4^{--} + 3OH^- + NH_3 \rightarrow Hg_2NI + 7I^- + 3H_2O.$$

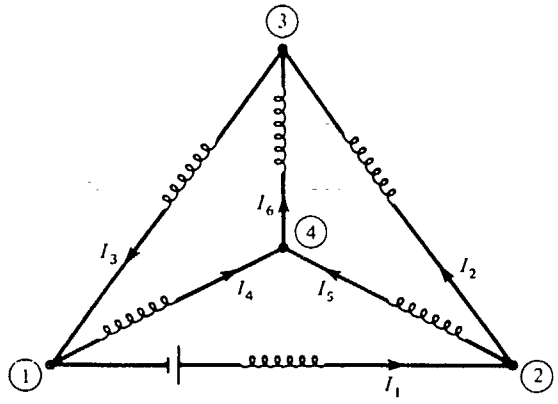
Selvfølgelig ville dette være et noget stort apparat at opstille blot for at afstemme en enkelt simpel kemisk ligning. Men pointen er dels at metoden virker for alle mulige reaktionsligninger, også hvis der forløber flere koblede reaktioner på én gang, men især at metoden ikke hviler på kemisk uigennemskuelige forestillinger eller begreber, men kun på det helt basale: grundstofbevarelse og ladningsbevarelse. For øvrigt giver metoden også mulighed for at klassificere reaktionsligninger på kemisk relevante måder ved hjælp af strukturen i det tilsvarende ligningssystem.

Eksempel 3: KIRCHHOFF's LOVE FOR ELEKTRISKE KREDSLØB

I elektricitetslæren lærer man at for at et jævnstrømskredsløb er i ligevægt må det opfylde bl.a den såkaldte *Kirchhoff's 1. lov*:

Ved hver knude skal summen af de indgående strømme være lig summen af de udgående strømme.

Lad os betragte det kredsløb som er angivet på figuren, hvor de indcirklede tal angiver knuderne.



Kredsløb

I følge Kirchhoff's 1. lov gælder så
knude 1: $I_3 = I_1 + I_4$
knude 2: $I_1 = I_2 + I_5$
knude 3: $I_2 + I_6 = I_3$
knude 4: $I_4 + I_5 = I_6$.

Dette ligningssystem kan fremstilles på matrix-form, idet vi med $\underline{I}$ betegner søjlen

$$\begin{pmatrix} I_1 \\ I_2 \\ I_3 \\ I_4 \\ I_5 \\ I_6 \end{pmatrix},$$

$$\begin{pmatrix} 1 & 0 & -1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} I_1 \\ I_2 \\ I_3 \\ I_4 \\ I_5 \\ I_6 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

eller kort, $M\underline{I} = \underline{0}$. Der er altså tale om et lineært ligningssystem med 4 ligninger og 6 ubekendte.

Eksempel 4: UDVIKLING I MARKEDSANDELE

Lad os forestille os et geografisk område, hvis pilsnerforbrug dækkes af tre bryggerier, X,Y og Z. Bryggeriernes markedsandele, som er foranderlige, gøres op hver måned, og opgaven er at beskrive udviklingen heri gennem tiden. Vi går ud fra en begyndelsessituation i måned nr. 0 med markedsandelene henholdsvis

$$x_0, y_0, z_0 \quad \text{for} \quad X, Y \text{og } Z,$$

hvor x_0, y_0 og z_0 alle er tal i intervallet $[0,1]$. Idet vi går ud fra at de tre bryggerier dækker hele områdets forbrug i den betragtede periode, har vi

$$x_0 + y_0 + z_0 = 1.$$

I måned nr. n betegnes markedsandelene med x_n, y_n og z_n, $n = 1, 2, 3, \ldots$

Vi antager nu at bryggeriet X hver måned er i stand til at fastholde brøkdelen a_{11} $(0 \leq a_{11} \leq 1)$ af sine egne kunder, desuden at vinde brøkdelen a_{12} $(0 \leq a_{12} \leq 1)$ af Y's kunder og brøkdelen a_{13} $(0 \leq a_{13} \leq 1)$ af Z's kunder. Hvis det samlede antal pilsnerkunder i området er konstant N, vil bryggeri X i måned nr. 1 have et antal kunder der er lig

$$a_{11}(x_0 N) + a_{12}(y_0 N) + a_{13}(z_0 N).$$

På den anden side er antallet af X's kunder i måned nr. 1 også lig $x_1 N$, så at

$$x_1 N = a_{11}(x_0 N) + a_{12}(y_0 N) + a_{13}(z_0 N).$$

Ved at forkorte med N finder vi X's markedsandel i måned nr. 1

$$x_1 = a_{11}x_0 + a_{12}y_0 + a_{13}z_0.$$

På tilsvarende måde får vi for de øvrige bryggerier - med de oplagte fortolkninger af a_{ij}'erne: a_{ij} angiver den brøkdel af kunder fra bryggeri nr. j (hvor 1 står for X, 2 for Y og 3 for Z) som bryggeri nr. i vinder/fastholder i løbet af en måned:

$$y_1 = a_{21}x_0 + a_{22}y_0 + a_{23}z_0$$

$$z_1 = a_{31}x_0 + a_{32}y_0 + a_{33}z_0.$$

Med den tidligere indførte matrix-notation får vi

$$\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix}$$

eller kort,

$$\underline{v}_1 = A\underline{v}_0,$$

hvor betydningen af $\underline{v}_0$ og $\underline{v}_1$ er oplagt.

Tallene i søjle nr. j i A angiver de tre brøkdele bryggeri nr. j fastholder og afgiver af kunder i en given måned. Hvis der ikke er andre bryggerier der betjener området, må summen af tallene i hver søjle i A være 1. Søjle i angiver jo hvilken skæbne bryggeri nr. j's kunder undergår ved overgangen til næste måned: De kan forblive hos j eller overtages af ét af de to andre. En matrix med denne egenskab kaldes ofte for en **overgangsmatrix**, fordi den - som det fremgår -kan beskrive overgangen fra én periode til en anden.

Under antagelse af at mønstret fastholdes gennem tiden, kan vi fortsætte processen

$$\begin{aligned} \underline{v}_2 &= A\underline{v}_1 = A(A\underline{v}_0), \\ &\vdots \\ \underline{v}_n &= A(A(\ldots(A\underline{v}_0)\ldots)) \end{aligned}$$

Helt i analogi med vores sædvanlige notation med potenser skriver vi gerne

$$\underline{v}_n = A^n\underline{v}_0.$$

Det er så nærliggende at interessere sig for den langsigtede udvikling af denne proces, dvs. hvad der sker for $n \to \infty$. Også den slags spørgsmål behandles af den lineære algebra, som ved at give kriterier baseret på udseendet af A kan fortælle om udviklingen i sådanne processer. Det viser sig bl.a., at under den forudsætning at alle A's elementer er strengt positive, vil processen i det lange løb nærme sig en situation hvor de tre bryggeriers markedsandele er konstante.

Eksempel 5: ALDERSSTRUKTURERED POPULATIONER

Ved undersøgelser af biologiske populationer er det ofte væsentligt at de enkelte individer i gennem deres livsløb befinder sig i forskellige klasser. Der kan f.eks. være tale om at de gennemløber forskellige biologiske stadier (f.eks. larve, puppe, sommerfugl). Det kan også være at man af udefra kommende grunde kan have interesse i at "anbringe"individerne i klasser (f.eks. fisk efter alder (der jo er forbundet med længde og vægt) af hensyn til bestandens størrelse, fangstregler eller lignende).

Vi betragter nu en sådan aldersstruktureret population. Lad os antage at den er inddelt i $k+1$ klasser med numrene $0, 1, 2, \ldots k$. Den 0'te klasse består af de yngste

individer. Rekrutteringen til den klasse sker sædvanligvis ved fødsel. Den første klasse består af de individer der er én enhed ældre. Der kan være tale om tidsenheder som dage, måneder, år, men også blot om selve klasserne, uanset hvor lang tid kreaturerne befinder sig i hvert af dem. Rekrutteringen til første klasse sker ved at individerne rykker ind i den efter at have gennemløbet 0'te klasse. Og så fremdeles, indtil den sidste klasse. Sædvanligvis består den af de ældste levende individer, og klassen forlades ved død. Men mindre drastiske muligheder kan komme på tale, f.eks. hvis man er interesseret i at følge bestandsudviklingen på forskellige klassetrin i en folkeskole.

Lad os gå ud fra, at vi har at gøre med en population hvis individer kan tælles. Det er egentlig ikke afgørende, men letter forståelsen af det følgende. Vi kunne lige så godt have opereret med f.eks. biomasse. Vi er interesseret i at følge den tidslige udvikling i størrelsen af populationen og dens forskellige klasser. Med $n_0(t)$ betegnes antallet af individer i klasse 0 til tiden t. Med $n_1(t)$ antallet i klasse 1, og så fremdeles indtil $n_k(t)$ i klasse k. Vi tager udgangspunkt i forholdene til begyndelsestidspunktet $t = 0$.

Først vil vi se på rekrutteringen til klasse 0. Den finder sted ved fødsel, dvs. ved at individerne i de øvrige klasser formerer sig. Det er rimeligt at gå ud fra, at hver klasse giver et bidrag til de nyfødte som er proportionalt med antallet af individer i klassen. Proportionalitetsfaktoren vil være forskellig fra klasse til klasse og kan godt være 0. Med $f_0, f_1, \ldots, f_k$ angives de forskellige proportionalitetsfaktorer. Bogstavet f er valgt for at angive fertilitet, altså den gennemsnitlige antal nyfødte pr. individ i den pågældende klasse.

Antallet af individer i klasse 0 til tid 1 vil så være

$$n_0(1) = f_0 n_0(0) + f_1 n_1(0) + \ldots + f_k n_k(0).$$

Rekrutteringen til klasserne $1, \ldots k$ sker simpelthen ved at individer fra den foregående klasse rykker op. Det kan altså ikke i denne model forekomme at nogen bliver i deres klasse ved overgangen til en ny periode. Det antages at der er en fast opryknings(overlevelses)rate fra hver klasse, sådan at en bestemt brøkdel af individerne for klasse j rykker op i klasse $j + 1$, resten dør eller forlader systemet på anden måde. Vi kan antage, at brøkdelen p_0 af individerne fra klasse 0 rykker op i klasse 1, mens brøkdelen p_1 af individerne fra klasse 1 rykker op i klasse 2, osv. De individer der til tid 0 befandt sig i klasse k antages at have forladt populationen til tid 1, jf. bemærkningerne ovenfor. Derved får vi

$$\begin{aligned}
n_1(1) &= p_0 n_0(0), \\
n_2(1) &= p_1 n_1(0), \\
&\vdots \\
n_k(1) &= p_{k-1} n_{k-1}(0).
\end{aligned}$$

De individer der til tid 0 befandt sig i klasse k antages at have forladt populationen til tid 1, jf. bemærkningerne ovenfor.

De opskrevne relationer beskriver altså overgangen fra tid 0 til tid 1. Relationerne kan samles på matrix-form:

$$\begin{pmatrix} n_0(1) \\ n_1(1) \\ \vdots \\ n_k(1) \end{pmatrix} = \begin{pmatrix} f_0 & f_1 & \dots & f_{k-1} & f_k \\ p_0 & 0 & \dots & 0 & 0 \\ 0 & p_1 & \dots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & p_{k-1} & 0 \end{pmatrix} \begin{pmatrix} n_0(0) \\ n_1(0) \\ n_2(0) \\ \vdots \\ n_k(0). \end{pmatrix}$$

En sådan matrix, L, kaldes en **Leslie-matrix** efter den første der (i 40'erne) indførte dem til populationsbeskrivelse.

Forestiller man sig at dette mønster er fast til alle tidspunkter, får vi den generelle overgangsmekanisme beskrevet ved

$$\begin{pmatrix} n_0(t+1) \\ n_1(t+1) \\ \vdots \\ n_k(t+1) \end{pmatrix} = \begin{pmatrix} f_0 & f_1 & \dots & f_{k-1} & f_k \\ p_0 & 0 & \dots & 0 & 0 \\ 0 & p_1 & \dots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & p_{k-1} & 0 \end{pmatrix} \begin{pmatrix} n_0(t) \\ n_1(t) \\ n_2(t) \\ \vdots \\ n_k(t) \end{pmatrix}$$

eller kort

$$\underline{n}(t+1) = L\underline{n}(t), \quad t = 0, 1, 2, \dots.$$

På samme måde som i det foregående eksempel finder vi

$$\underline{n}(t+1) = L^{t+1}\underline{n}(0).$$

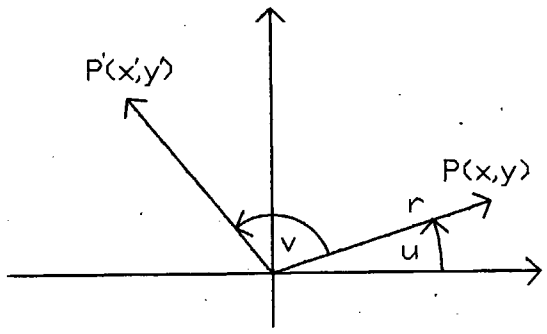
De spørgsmål man her ville stille om den langsigtede udvikling af populationen og dens klasser er de samme som i markedsandelseksemplet. Det matematiske apparat til besvarelse af dem er de samme: lineær algebra.

Eksempel 6: GEOMETRISKE TRANSFORMATIONER

Vi vil slutte indledningen med at undersøge et par geometriske transformationer i planen.

Først **rotation** om et punkt med en bestemt vinkel. Vi forestiller os et koordinatsystem indlagt, så at rotationens centrum er koordinatsystemets begyndelsespunkt. Hvis punktet P har koordinaterne (x, y) kan disse skrives

$$(x, y) = (r \cos u, r \sin u),$$



hvor r er afstanden fra P til begyndelsespunktet O, og u er den vinkel $\underline{OP}$ danner med 1-aksen. Ved rotation med vinklen v føres P over i punktet P' med koordinaterne

$$\begin{aligned}(x', y') &= (r \cos(u + v), r \sin(u + v)) \\ &= (r \cos u \cos v - r \sin u \sin v, r \sin u \cos v + r \cos u \sin v) \\ &= (x \cos v - y \sin v, y \cos v + x \sin v),\end{aligned}$$

(det mellemste lighedstegn følger af additionsformlerne for cos og sin), altså

$$\begin{aligned}x' &= \cos v\, x - \sin v\, y \\ y' &= \sin v\, x + \cos v\, y,\end{aligned}$$

og i matrixnotation

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos v & -\sin v \\ \sin v & \cos v \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

En rotation i planen med vinklen v omkring koordinatsystemets begyndelsespunkt er altså beskrevet ved at punktets nye koordinater fremgår af de gamle ved en bestemt *lineær transformation* af de gamle.

Dernæst betragter vi en **spejling** i en linje gennem begyndelsespunktet. Har P igen koordinaterne

$$(x, y) = (r \cos u, r \sin u),$$

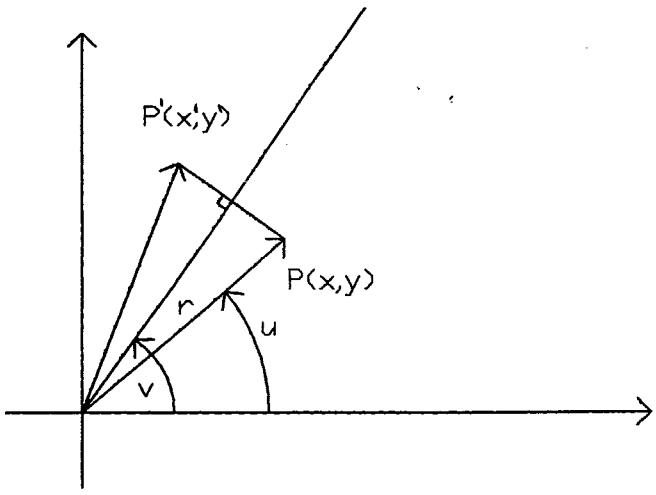
vil det spejlede punkt P' have koordinaterne

$$\begin{aligned}(x', y') &= (r\cos(2v-u), r\sin(2v-u)) \\ &= (r\cos 2v\cos u + r\sin 2v\sin u, r\sin 2v\cos u - r\cos 2v\sin u) \\ &= (x\cos 2v + y\sin 2v, x\sin 2v - y\cos 2v).\end{aligned}$$

Derved får vi

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos 2v & \sin 2v \\ \sin 2v & -\cos 2v \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

der igen er en **lineær transformation** af koordinaterne.



Ser vi sluttelig på en **multiplikation** ud fra koordinatsystemets begyndelsespunkt med en bestemt faktor a, bliver punktet P med koordinater (x, y) overført i punktet P' med koordinater (ax, ay). Også denne transformation er *lineær*, beskrevet ved

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

LINEÆRE LIGNINGSSYSTEMER I

Det er en af den lineære algebras opgaver at levere en teori for løsningsforholdene ved lineære ligningssystemer. Som et første led i opbygningen af en sådan teori vil vi betragte et par eksempler, som dels viser hvorledes man rent praktisk griber løsningen af et lineært ligningssystem an, og som dels viser at der kan optræde forskellige karakteristiske tilfælde.

Eksempel 1.

Vi betragter ligningssystemet

$$
(1) \qquad \begin{aligned}
2\,x_1 + \quad x_2 - 3\,x_3 &= 1 \\
4\,x_1 - 2\,x_2 - 2\,x_3 &= 3 \\
-x_1 + \frac{1}{2}x_2 - \frac{1}{2}x_3 &= \frac{3}{2}
\end{aligned}
$$

Vi er ude på at omforme dette ligningssystem til ét hvor løsningsforholdene er lettere at overskue. Idéen er først at eliminere x_1 fra samtlige ligninger bortset fra den øverste. Derved danner de to sidste ligninger et selvstændigt ligningssystem bestående af kun to ligninger med to ubekendte. På dette mindre ligningssystem kan vi da foretage en tilsvarende omformning ved at fjerne (her) x_2 fra "alle"ligninger på nær den anden. Det bevirker, at x_3 bliver den eneste ubekendte tilbage i den tredje ligning, som derfor bliver umiddelbart løsbar (om overhovedet). Denne skridtvise elimination viser sig at være anvendelig som generel løsningsmetodik ved alverdens lineære ligningssystemer, ikke blot det foreliggende. Metoden bærer navnet **Gauss-elimination** efter én af den nyere tids mest betydningsfulde matematikere, tyskeren Carl Friederich Gauss (1777-1855). Se nu hvordan:

Ligningssystemet er ensbetydende med følgende ligningssystem, som er fremkommet af det oprindelige ved at den anden ligning er erstattet med den selv *plus den første ligning multipliceret* med -2 (den første ligning selv er uberørt):

$$
(2) \qquad \begin{aligned}
2\,x_1 + \quad x_2 - 3\,x_3 &= 1 \\
-4\,x_2 + 4\,x_3 &= 1 \\
-x_1 + \frac{1}{2}x_2 - \frac{1}{2}x_3 &= \frac{3}{2}
\end{aligned}
$$

Fra det nye system kan man komme tilbage til det oprindelige ved at udføre den "modsatte operation": at multiplicere den første ligning med +2 og derefter addere den til den anden. De to systemer er altså virkelig ensbetydende (ækvivalente).

Nu foretages en tilsvarende operation på system (2) med den hensigt at bortskaffe x_1-leddet i den tredje ligning. Den første ligning multipliceres med $\frac{1}{2}$ og adderes til den tredje (den første ligning er stadig uberørt). Derved får vi

$$
\begin{aligned}
2x_1 + \quad x_2 - 3x_3 &= 1 \\
-4x_2 + 4x_3 &= 1 \\
x_2 - 2x_3 &= 2.
\end{aligned}
\tag{3}
$$

Dette ligningssystem er ensbetydende med (2) - med argumenter magen til de ovenstående: vi kan komme tilbage til (2) ved på (3) at udføre den modsatte operation bestående i til den tredje ligning at lægge den første multipliceret med $-\frac{1}{2}$. System (2) er på sin side ensbetydende med (1). Systemet (3) er altså ensbetydende med det oprindelige system.

Nu skaffer vi os af med x_2-leddet i den sidste ligning i (3) ved at multiplicere den anden ligning med $\frac{1}{4}$ og addere den til den tredje. Det giver det ækvivalente ligningssystem

$$
\begin{aligned}
2x_1 + \quad x_2 - 3x_3 &= 1 \\
-4x_2 + 4x_3 &= 1 \\
- \quad x_3 &= \frac{9}{4}.
\end{aligned}
$$

Herved har vi opnået at omforme det oprindelige ligningssystem til et ensbetydende, men meget simplere ligningssystem i **trappeform**. Det trappeformede ligningssystem er simplere fordi det kan løses umiddelbart ved optrævling nedefra:

Af den nederste ligning ses straks at $x_3 = -\frac{9}{4}$. Dette indsættes i den anden ligning, som derefter kun indeholder den ubekendte x_2, der så bestemmes som $x_2 = -\frac{5}{2}$. Sluttelig bestemmes x_1 af den første ligning ved indsættelse af de fundne værdier for x_2 og x_3. Vi finder $x_1 = -\frac{13}{8}$.

Denne analyse viser os at ligningssystemet (1), bestående af tre ligninger med tre ubekendte, har én og kun én løsning (eller for at være lidt mere pedantisk: ét og kun ét løsningssæt), nemlig

$$(x_1, x_2, x_3) = (-\frac{13}{8}, -\frac{5}{2}, -\frac{9}{4}).$$

Det gik jo fint. Går det altid så nemt? Lad os se på nogle flere eksempler.

Eksempel 2.

Ligningssystemet

$$
(1) \qquad \begin{aligned}
2\,x_1 + \quad x_2 - 3\,x_3 &= 1 \\
4\,x_1 - 2\,x_2 + 2\,x_3 &= 3 \\
-x_1 + \frac{1}{2}x_2 - \frac{1}{2}x_3 &= \frac{3}{2}
\end{aligned}
$$

adskiller sig kun fra det foregående ved +'et foran x_3-leddet i den anden ligning. Går vi nu frem på samme måde som før finder vi

$$
(2) \qquad \begin{aligned}
2\,x_1 + \quad x_2 - 3\,x_3 &= 1 \\
-4\,x_2 + 8\,x_3 &= 1 \\
-x_1 + \frac{1}{2}x_2 - \frac{1}{2}x_3 &= \frac{3}{2}.
\end{aligned}
$$

Og videre:

$$
(3) \qquad \begin{aligned}
2x_1 + \quad x_2 - 3x_3 &= 1 \\
-4x_2 + 8x_3 &= 1 \\
x_2 - 2x_3 &= 2.
\end{aligned}
$$

Multiplicerer vi nu den mellemste ligning med $\frac{1}{4}$ og lægger den til den sidste, bliver resultatet

$$
\begin{aligned}
2x_1 + \quad x_2 - 3x_3 &= 1 \\
-4x_2 + 8x_3 &= 1 \\
0 \quad x_3 &= \frac{9}{4}.
\end{aligned}
$$

Den sidste ligning har øjensynlig ingen løsning. Så meget mere har hele ligningssystemet heller ikke en løsning. Metoden virkede udmærket til at afklare løsningsforholdene, der er blot ingen løsninger.

Men hermed er variationerne ikke udtømt. Endnu et eksempel:

Eksempel 3.

Også ligningssystemet

$$
(1) \qquad \begin{aligned} 2\,x_1 + \quad x_2 - 3\,x_3 &= 1 \\ 4\,x_1 + 2\,x_2 - 2\,x_3 &= 3 \\ -x_1 + \frac{1}{2}x_2 - \frac{1}{2}x_3 &= \frac{3}{2} \end{aligned}
$$

adskiller sig kun ubetydeligt (ved +'et foran x_2-leddet i den anden ligning) fra det først undersøgte system. Går vi til værks efter forskrifterne, når vi til

$$
(2) \qquad \begin{aligned} 2\,x_1 + \quad x_2 - 3\,x_3 &= 1 \\ + 4\,x_3 &= 1 \\ -x_1 + \frac{1}{2}x_2 - \frac{1}{2}x_3 &= \frac{3}{2}, \end{aligned}
$$

og derefter til

$$
(3) \qquad \begin{aligned} 2x_1 + x_2 - 3x_3 &= 1 \\ + 4x_3 &= 1 \\ x_2 - 2x_3 &= 2. \end{aligned}
$$

Fulgte vi nu forskriften automatisk - og det skulle "man" (f.eks. et computer-program) jo gøre, hvis man ikke som her opererede på et konkret system med givne tal som koefficienter, men på et generelt system med vilkårlige og derfor principielt ukendte koefficienter - kunne vi ikke komme videre herfra. Efter forskriften skal vi jo skaffe x_2-leddet væk fra den sidste ligning ved at multiplicere den mellemste ligning med en passende konstant og addere resultatet til den sidste ligning. Men eftersom koefficienten til x_2-leddet i den mellemste ligning er 0 mens x_2-koefficienten i den sidste ligning er forskellig fra 0, kan ingen multiplikation i verden klare denne opgave. Hvad gør vi så?

I det foreliggende tilfælde er sagen nemt ordnet. Vi bytter blot om på ligning 2 og ligning 3, hvilket selvsagt ikke ændrer systemet. Det derved ommøblerede system er umiddelbart på trappeform

$$
(4) \qquad \begin{aligned} 2x_1 + x_2 - 3x_3 &= 1 \\ x_2 - 2x_3 &= 2 \\ + 4x_3 &= 1 \end{aligned}
$$

og kan let løses. Det har den entydigt bestemte løsning

$$
(x_1, x_2, x_3) = (-\frac{3}{8}, \frac{5}{2}, \frac{1}{4}).
$$

Dette eksempel viser, at den først beskrevne automatiske forskrift ikke virker i enhver situation. Den må åbenbart suppleres i visse undtagelsessituationer. Vi skal ikke her forfølge sådanne undtagelser nøjere. En total kortlægning af hvad der kan ske er vigtig for teoretiske formål, og som baggrund for at konstruere computer-programmer til løsning af lineære ligningssystemer. Til vores formål rækker det at demonstrere metoden og at antyde et behov for i nogle situationer at supplere den. Ved at gå frem som beskrevet kan man slippe levende igennem alle praktiske ligningsløsningssituationer.

Så meget om metoden. Hvad angår løsningsforholdene har vi oplevet, at fredeligt udseende og næsten ens udseende systemer af tre ligninger med tre ubekendte kan udvise helt forskellige løsningsegenskaber: Der kan være *netop én løsning*, og der kan *ingen løsninger* være. Er det alt? Nej, se selv:

Eksempel 4.

$$
\begin{aligned}
2x_1 + x_2 - 3x_3 &= 1 \\
4x_1 + 2x_2 - 2x_3 &= 3 \\
2x_1 + x_2 + x_3 &= 2.
\end{aligned} \tag{1}
$$

Udfører vi eliminationen af x_1 i de to sidste ligninger på én gang - hvilket øjensynlig er det samme som at foretage den i to skridt - får vi

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

Da de to sidste ligninger er ens, har vi i realiteten kun at gøre med to ligninger med tre ubekendte:

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

Af den sidste ligning fremgår, at $x_3 = -\frac{1}{4}$. Indsættes dette i den første ligning antager systemet skikkelsen

$$
\begin{aligned}
2x_1 + x_2 \quad &= \frac{1}{4} \\
x_3 &= \frac{1}{4}.
\end{aligned} \tag{4}
$$

Her kan vi frit vælge hvad x_1 eller x_2 skal være, idet den anden så kan bestemmes. Sætter vi f.eks. $x_2 = t$, hvor t er et vilkårligt reelt tal, løser $x_1 = -\frac{1}{2}t + \frac{1}{8}$ og $x_2 = t$ den første ligning. Omvendt har enhver løsning denne form. Samlet har vi fundet, at systemet (4) har løsningerne

$$(x_1, x_2, x_3) = (-\frac{1}{2}t + \frac{1}{8}, t, \frac{1}{4}),$$

hvor t er et vilkårligt reelt tal.

Ligningssystemet har altså *uendeligt mange løsninger*.

I behandlingen af tre ligninger med tre ubekendte har vi altså mødt tre forskellige løsningsmuligheder: **netop én løsning; ingen løsninger; uendeligt mange løsninger**. Det viser sig at ét af disse tilfælde altid vil indtræffe. Det er altså ikke muligt at der er f.eks. syv forskellige løsninger. Det viser sig også at situationen er den samme *uanset hvor mange ligninger og hvor mange ubekendte der indgår i ligningssystemet*. For at illustrere dette afsluttes dette afsnit med endnu tre eksempler.

Eksempel 5.

Ligningssystemet

$$(1) \qquad \begin{aligned} -x_1 \qquad + 2x_3 + \ x_4 - \ x_5 &= -1 \\ 2\,x_1 - \ x_2 - \ x_3 + \ x_4 + 3x_5 &= 1 \\ x_1 + 2x_2 - 3x_3 - 2x_4 + \ x_5 &= 0 \end{aligned}$$

består af tre ligninger med fem ubekendte. Løsningen af lignings- systemet gribes an på samme måde som i de foregående eksempler. Bestræbelsen er at nå frem til et ækvivalent ligningssystem, hvor den anden ligning ikke indeholder x_1 og den tredje ikke x_1 og x_2. Vi lægger ud med at skaffe x_1 væk fra de to sidste ligninger. Det sker ved til den anden ligning at lægge 2 gange den første, og ved til den tredje ligning at lægge den første. Vi foretager begge operationer på én gang. Derved når vi frem til det ensbetydende ligningssystem

$$(2) \qquad \begin{aligned} -x_1 \qquad + 2x_3 + \ x_4 - x_5 &= -1 \\ -x_2 + 3x_3 + 3x_4 + x_5 &= -1 \\ 2\,x_2 - \ x_3 - \ x_4 \qquad &= -1. \end{aligned}$$

Nu skaffer vi x_2-leddet væk fra den sidste ligning i (2) ved til denne ligning at lægge 2 gange den mellemste ligning. Resultatet bliver

$$(3) \qquad \begin{aligned} -x_1 \qquad + 2x_3 + \ x_4 - \qquad x_5 &= -1 \\ -x_2 + 3x_3 + 3x_4 + \ x_5 &= -1 \\ 5 \quad x_3 + 5x_4 + 2 \ x_5 &= -3. \end{aligned}$$

Herved er vi nået frem til et ligninssystem på trappeform som er ensbetydende med (1). Vi kan ikke fortsætte forenklingsprocessen længere, fordi en bortskaffelse af x_3 i den sidste ligning ville genindføre x_1 eller x_2 i ligningen. Men der er heller ikke behov for yderligere forenkling; Systemet (3) kan løses uden videre: Uanset hvilke værdier x_4 og x_5 tildeles kan x_3 bestemmes af den sidste ligning, derefter x_2 af den mellemste og endelig x_1 af den første ligning. Det betyder at *for hvilke som helst reelle tal t og u*, vil med

$$
\begin{aligned}
x_4 &= t \text{ og} \\
x_5 &= u, \\
x_3 &= -\frac{3}{5} - t - \frac{2}{5}u \text{ (af sidste ligning)}, \\
x_2 &= 3(-\frac{3}{5} - t - \frac{2}{5}u) + 3t + u + 1 \text{ (af den mellemste ligning samt udtrykket for } x_3) \\
&= -\frac{1}{5}u - \frac{4}{5}, \\
x_1 &= 2(-\frac{3}{5} - t - \frac{2}{5}u) + t - u + 1 \text{ (af den første ligning samt udtrykket for } x_3) \\
&= -t - \frac{9}{5}u - \frac{1}{5}
\end{aligned}
$$

udgøre et løsningssæt til ligninssystemet (3) og dermed til det ensbetydende system (1):

$$(x_1, x_2, x_3, x_4, x_5) = (-t - \frac{9}{5}u - \frac{1}{5}, -\frac{1}{5}u - \frac{4}{5}, -t - \frac{2}{5}u - \frac{3}{5}, t, u).$$

Ethvert sæt af denne form er altså en løsning til (1).

Hvis omvendt $(x_1, x_2, x_3, x_4, x_5)$ er en løsning kan vi altid finde to reelle tal t og u, sådan at løsningen har den opskrevne form. Vi skal jo blot sætte $t = x_4$ og $u = x_5$, så kan x_1, x_2 og x_3 i kraft af (3) ikke slippe for at have den ønskede form. Med udtrykkene ovenfor har vi altså bestemt samtlige løsninger til ligningssystemet. Læg mærke til at både x_4 og x_5 kan vælges frit. Vi siger at der er to **frie variable** i løsningssættene. De øvrige tre er derefter bestemt. Vi siger, at de er **bundne variable** i løsningssættene. Nu er det ikke sådan at det netop *skal* være x_4 og x_5 der er de frie variable, mens resten *skal* være bundne. Havde vi f.eks. skrevet hver lignings venstreside op i modsat orden med x_5 først osf. og derefter benyttet os af trappeomformningen, ville x_2 og x_1 være blevet de frie variable, mens de øvrige var blevet de bundne. Så friheden og bundetheden er ikke knyttet til *bestemte* af de ubekendte. De væsentlige i sagen er *antallet* af frie, henholdsvis bundne, variable. Dette antal er *det samme* uanset den rækkefølge vi skriver leddene op i.

Eksempel 6.

I ligningssystemet

$$
(1) \qquad \begin{aligned}
-2x_1 + \quad x_2 + 2x_3 &= 0 \\
-3x_1 - \frac{1}{2}x_2 - \quad x_3 &= 4 \\
x_1 - \frac{3}{2}x_2 + \quad x_3 &= -2 \\
-\quad x_1 + \frac{5}{2}x_2 - \quad x_3 &= 1
\end{aligned}
$$

er der fire ligninger og tre ubekendte. Også i dette tilfælde er det vores bestræbelse at omforme systemet til et ækvivalent system på trappeform. Vi går frem på den foreskrevne måde ved at først at skaffe x_1 væk i de sidste tre ligninger. Det foregår ved at der til den anden ligning lægges den første gange $-\frac{3}{2}$, til den tredje den første gange $\frac{1}{2}$, og til den sidste den første gange $-\frac{1}{2}$. Derved fremkommer det ensbetydende system

$$
(2) \qquad \begin{aligned}
-2x_1 + \quad x_2 + 2x_3 &= 0 \\
-2x_2 - 4x_3 &= 4 \\
-\quad x_2 + 2x_3 &= -2 \\
2 \quad x_2 - 2x_3 &= 1.
\end{aligned}
$$

Det sidste skridt består i at bortskaffe x_2 i de to sidste ligninger ved til den tredje at lægge $-\frac{1}{2}$ gange den anden, og ved til den fjerde simpelthen at lægge den anden. Resultatet bliver

$$
(3) \qquad \begin{aligned}
-2x_1 + \quad x_2 + 2x_3 &= 0 \\
-2x_2 - 4x_3 &= 4 \\
4 \quad x_3 &= -4 \\
-6x_3 &= 5.
\end{aligned}
$$

Af de to sidste ligninger ses at systemet *ingen løsninger* har, eftersom x_3 ikke kan overkomme på én gang at være -1 (den tredje ligning) og $-\frac{5}{6}$ (den fjerde ligning).

Eksempel 7.

En minimal forandring af ligningssystemet i Eksempel 6 giver en helt anden løsningssi-

tuation. Hvis højresiden i den sidste ligning ændres til 2 har vi at gøre med systemet

$$
(1) \qquad \begin{aligned}
-2x_1 + \quad x_2 + 2x_3 &= 0 \\
-3x_1 - \frac{1}{2}x_2 - \quad x_3 &= 4 \\
x_1 - \frac{3}{2}x_2 + \quad x_3 &= -2 \\
-\quad x_1 + \frac{5}{2}x_2 - \quad x_3 &= 2.
\end{aligned}
$$

De samme operationer som i Eksempel 6 fører til

$$
(2) \qquad \begin{aligned}
-2x_1 + \quad x_2 + 2x_3 &= 0 \\
-2x_2 - 4x_3 &= 4 \\
-\quad x_2 + 2x_3 &= -2 \\
2 \quad x_2 - 2x_3 &= 2,
\end{aligned}
$$

og videre til

$$
(3) \qquad \begin{aligned}
-2x_1 + \quad x_2 + 2x_3 &= 0 \\
-2x_2 - 4x_3 &= 4 \\
4 \quad x_3 &= -4 \\
-6x_3 &= 6.
\end{aligned}
$$

Her er der ingen modsætninger mellem de to sidste ligninger, som begge kræver, at $x_3 = -1$. Derefter finder vi straks $x_2 = 0$ og $x_1 = -1$, sådan at ligningssystemet har *netop én løsning*, nemlig

$$(x_1, x_2, x_3) = (-1, 0, -1).$$

Sagen er her, at ligningssystemet kun *tilsyneladende* består af fire ligninger. I *virkeligheden* er der kun tre uafhængige ligninger. Det forholder sig nemlig sådan, at den sidste ligning i (1) følger af de foregående tre, idet den sidste ligning simpelthen er $\frac{1}{8}$ gange den første $-\frac{1}{4}$ gange den anden $-\frac{3}{2}$ gange den tredje (det forventes du ikke selv at kunne få øje på!).

De tre sidste eksempler viser, at også med ligningssystemer hvor antallet af ligninger ikke er lig antallet af ubekendte bliver løsningsomstændighederne de samme: ingen, netop én, eller uendeligt mange løsninger. I Eksempel 5 så vi desuden at de uendeligt mange løsninger kan være fastlagt ved *flere* frie variable - eller flere *parametre*, som man også siger. Vi benytter også den sprogbrug at løsningsmængden er *fler- dimensional*.

Hermed har vi afsluttet første omgang i behandlingen af lineære ligningssystemer. Flere omgange vil følge. Den indførte løsningsmetode med dens nødvendige modifikationer, som i alle tilfælde har kunnet afklare løsningsforholdene for ligningssystemerne, kan formuleres således : *Ligningssystemet omformes ved hjælp af* **rækkeoperationer**. *En rækkeoperation består i at en given ligning erstattes med : den selv ganget med en skalar* $\neq 0$ *(f. eks. 1), plus én af de andre ligninger ganget med en skalar (som gerne må være 0).* Ved en rækkeoperation har det oprindelige og det omformede ligningssystem de samme løsninger, de er **ækvivalente**. Denne metode har et automatisk præg. Det er faktisk en væsentlig del af pointen med en metode. Det forhindrer imidlertid ikke, at man ved behandlingen af *konkret foreliggende* ligningssystemer ofte kan slippe lettere igennem end ved at bruge automatmetoden. F.eks. er det ikke sikkert at den ligning der tilfældigvis står først er den det er bekvemmest at have på den plads. Ofte vil det være rarest hvis koefficienten til x_1 er 1 eller -1. Det kan også være at man straks kan se at én eller flere af ligningerne på en simpel måde fremkommer af nogle af de andre. Sådanne lettelser skal man naturligvis ikke afstå fra at bruge, når situationen er til det.

TALRUMMENE $\mathbb{R}^n$

Indledning

I alle de foregående afsnit har **sæt af reelle tal** spillet en central rolle. Vi skal nu foretage et systematisk studium af sådanne sæt og lineære operationer på dem.

I alt hvad der følger er n et vilkårligt naturligt tal (dvs. $1, 2, \ldots$). De objekter vi interesserer os for er ordnede sæt bestående af n reelle tal: $(x_1, x_2, \ldots, x_n)$. At sættet er ordnet betyder at rækkefølgen tallene optræder i er væsentlig. Således er (for $n = 4$) sættet $(1, -2, \frac{1}{2}, 0)$ forskelligt fra sættet $(0, 1, \frac{1}{2}, -2)$. Sådanne objekter kalder vi **vektorer**, undertiden n-vektorer hvis vi har brug for at understrege hvilket n der er tale om. Det kan f.eks. være tilfældet hvis vektorer af forskellig længde er i spil. Man kan også støde på betegnelsen "ordnede n-sæt"eller "ordnede n-tupler". De enkelte tal i en vektor, altså x'erne, kaldes ofte vektorens **komponenter**. Det fastsættes at to vektorer er identiske, hvis på hver plads deres komponenter er identiske.

Vi vil i mange sammenhænge benytte notationen $\underline{x}$ som forkortelse af $(x_1, x_2, \ldots, x_n)$, selv om det ikke af denne notation fremgår hvilket n der refereres til. Det må så underforstås af sammenhængen.

Samlingen af samtlige n-vektorer benævnes $\mathbb{R}^n$. Vi har altså per definition

$$\mathbb{R}^n = \{(x_1, x_2, \ldots, x_n) | x_1 \in \mathbb{R}, x_2 \in \mathbb{R}, \ldots, x_n \in \mathbb{R}\}.$$

Vi taler om **talrummet** $\mathbb{R}^n$. Der findes en uendelighed af sådanne talrum, nemlig ét for hvert n. Hvis $n = 1$ har vi øjensynlig blot at gøre med $\mathbb{R}$, de reelle tal selv. Hvis $n = 2$ eller $n = 3$ har vi at gøre med koordinatsættene for punkterne i planen henholdsvis i rummet.

Som nævnt er det væsentligt at holde styr på *rækkefølgen* af tallene i en n-vektor. Derimod er det ikke væsentligt for de begreber vi snart skal indføre om vi opskriver en n-vektor som en række eller en søjle. Sådan var det jo ikke med matricer, som er rektangulære talskemaer hvor både den vandrette og den lodrette orden er væsentlig. **Rækkematricen** $(x_1, x_2, \ldots, x_n)$ er *forskellig fra* **søjlematricen**

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

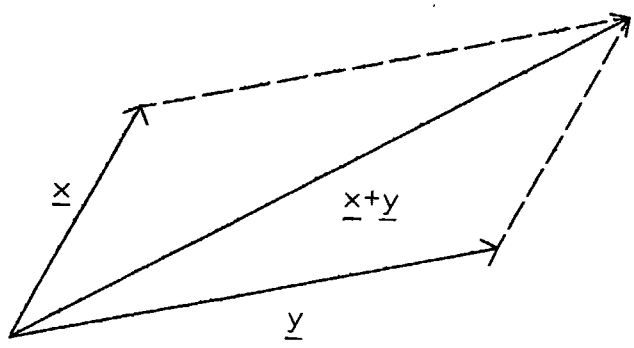
De to forskellige matricer er *forskellige matrix-fremstillinger af én og samme vektor.* Vi vil ofte få brug for at fremstille en given vektor på begge måder, alt efter hvad der er bekvemt i sammenhængen. Men endnu engang: I definitionen af en n-vektor er kun rækkefølgen af tallene væsentlig, ikke om de opskrives vandret eller lodret. I definitionen af matricer er både rækkefølgen og opskrivningen væsentlig.

Addition af vektorer i $\mathbb{R}^n$

Vi vil nu indføre et **additionsbegreb** for vektorer af samme størrelse. Med en nobel sprogbrug hentet fra den abstrakte algebra vil vi indføre en **komposition** $+$ i $\mathbb{R}^n$, dvs. en operation der til to vilkårlige vektorer i $\mathbb{R}^n$ knytter en tredje, deres *sum*. Definitionen er lidet dybsindig, idet vi simpelthen adderer to n-vektorer ved at addere komponenterne pladsvis:

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

Denne måde at indføre addition på er i $\mathbb{R}$ simpelthen den sædvanlige addition af reelle tal, mens den i $\mathbb{R}^2$ og $\mathbb{R}^3$ svarer til addition af geometriske vektorer ("kræfternes parallellogram"), jfr. illustrationen.



Kræfternes parallellogram

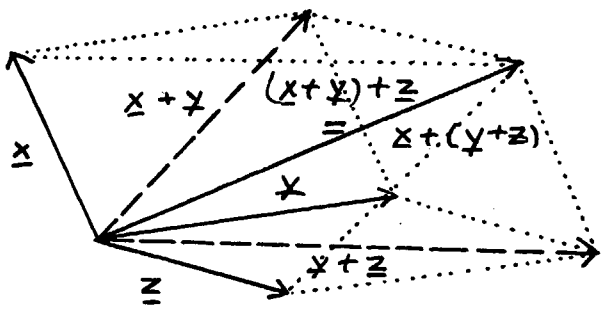
Det er en umiddelbar konsekvens af reglerne for addition af reelle tal, at man kan bytte om på leddene i additionen:

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

- den såkaldte **kommutative lov** - og at man frit kan sætte parenteser således:

$$\begin{aligned}&((x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n)) + (z_1, z_2, \ldots, z_n)\\ &= (x_1, x_2, \ldots, x_n) + ((y_1, y_2, \ldots, y_n) + (z_1, z_2, \ldots, z_n)).\end{aligned}$$

Dette er den såkaldte **associative lov**, hvis indhold i $\mathbb{R}^2$ og $\mathbb{R}^3$ kan fortolkes geometrisk, jf. illustrationen nedenfor. Begge love følger af at addition af vektorer foregår pladsvis.



Der findes netop én vektor som virker **neutralt ved addition**, dvs. opfylder at den adderet til en vilkårlig vektor lader denne vektor uberørt. Vektoren bestående af lutter 0'er bestrider åbenbart dette job, idet

$$(x_1, x_2, \ldots, x_n) + (0, 0, \ldots, 0) = (x_1, x_2, \ldots, x_n)$$

for ethvert $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$. Ingen andre vektorer har denne egenskab, thi hvis $(w_1, w_2, \ldots, w_n)$ opfylder

$$(x_1, x_2, \ldots, x_n) + (w_1, w_2, \ldots, w_n) = (x_1, x_2, \ldots, x_n),$$

må

$$(x_1 + w_1, x_2 + w_2, \ldots, x_n + w_n) = (x_1, x_2, \ldots, x_n).$$

Eftersom to vektorer er identiske netop hvis de har samme komponenter, har vi så

$$\begin{aligned} x_1 + w_1 &= x_1, \\ x_2 + w_2 &= x_2, \\ &\vdots \\ x_n + w_n &= x_n, \end{aligned}$$

hvoraf det følger, at $w_1 = 0, w_2 = 0, \ldots, w_n = 0$. Dette viser at

$$(w_1, w_2, \ldots, w_n) = (0, 0, \ldots, 0).$$

Vektoren hvis komponenter alle er 0 kaldes **nulvektoren** og betegnes $\underline{0}$ (det må fremgå af sammenhængen hvilket n der er tale om).

Det er nu nærliggende at spørge om der til en given vektor $(x_1, x_2, \ldots, x_n)$ findes en anden vektor $(u_1, u_2, \ldots, u_n)$ der neutraliserer den, dvs. opfylder

$$(x_1, x_2, \ldots, x_n) + (u_1, u_2, \ldots, u_n) = (0, 0, \ldots, 0)$$

Dette er ensbetydende med at søge $u_1, u_2, \ldots, u_n$ så at

$$x_1 + u_1 = 0,\ x_2 + u_2 = 0,\ \ldots, x_n + u_n = 0.$$

Disse krav er opfyldt, hvis og kun hvis

$$u_1 = -x_1,\ u_2 = -x_2,\ \ldots, u_n = -x_n.$$

Svaret er altså bekræftende, idet

$$(u_1, u_2, \ldots, u_n) = (-x_1, -x_2, \ldots, -x_n)$$

som den eneste vektor opfylder det stillede krav. Vi kalder også $(-x_1, -x_2, \ldots, -x_n)$ for den **modsatte vektor** til $(x_1, x_2, \ldots, x_n)$, og betegner den

$$-(x_1, x_2, \ldots, x_n).$$

Det er nu muligt at benytte skrivemåden

$$(x_1, x_2, \ldots, x_n) - (y_1, y_2, \ldots, y_n)$$

for

$$(x_1, x_2, \ldots, x_n) + (-(y_1, y_2, \ldots, y_n))\,(= (x_1 - y_1, x_2 - y_2, \ldots, x_n - y_n)).$$

Eksempel 0

I $\mathbb{R}^4$ har vi vektorerne $\underline{x} = (2, -1, 0, 3)$, $\underline{y} = (-5, 2, -1, -2)$ og $\underline{z} = (3, -1, 1, -1)$. Det ses at $\underline{x} + \underline{y} = (-3, 1, -1, 1)$ så at $-(\underline{x} + \underline{y}) = \underline{z}$. *

Omtalen af addition afsluttes ved at samle de fundne resultater i kort notation:

Der indføres en komposition $+$ i $\mathbb{R}^n$, som opfylder

(1) For alle $\underline{x}$ og $\underline{y}$: $\underline{x}+\underline{y}=\underline{y}+\underline{x}$ (den kommutative lov)

(2) For alle $\underline{x}$, $\underline{y}$ og $\underline{z}$: $(\underline{x}+\underline{y})+\underline{z}=\underline{x}+(\underline{y}+\underline{z})$ (den associative lov)

(3) Der findes netop ét neutralt element $\underline{0}$ så at for alle $\underline{x}$ gælder

$$\underline{x}+0=\underline{x}$$

(4) Ethvert $\underline{x}$ har netop ét modsat element, $-\underline{x}$, opfyldende at

$$\underline{x}+(-\underline{x})=\underline{0}.$$

(Disse fire egenskaber udtrykkes ofte samlet sådan, at $(\mathbb{R}^n,+)$ udgør en **abelsk gruppe**).

Alene på grundlag af reglerne 1.-4. kan det sluttes, at *enhver ligning* af formen $\underline{x}+\underline{a}=\underline{b}$ *har netop én løsning*, nemlig $\underline{x}=\underline{b}-\underline{a}$. At $\underline{b}-\underline{a}$ løser ligningen ses ved indsættelse. At der ikke er andre muligheder indses ved at addere $-\underline{a}$ (som eksisterer i følge 4.) til begge sider af ligningen $\underline{x}+\underline{a}=\underline{b}$, hvilket pedantisk opskrevet giver

$$\underline{x}=\underline{x}+(\underline{a}-\underline{a})=(\underline{x}+\underline{a})+(-\underline{a})=\underline{b}+(-\underline{a})=\underline{b}-\underline{a}.$$

Skalarmultiplikation i $\mathbb{R}^n$

Vektorer i $\mathbb{R}^n$ kan skaleres med reelle tal. Det vil sige, at for enhver vektor $\underline{x}=(x_1,x_2,\ldots,x_n)\in\mathbb{R}^n$ og ethvert reelt tal α dannes per definition vektoren

$$\alpha\underline{x}=(\alpha x_1,\alpha x_2,\ldots,\alpha x_n).$$

Denne vektor siges at være fremgået ved **skalarmultiplikation** af vektoren $\underline{x}$ med **skalaren** α. Som det ses er en skalar (her) ikke andet end et reelt tal. Betegnelsen "skalar"i stedet for "tal"tjener til at markere den indbyrdes *rollefordeling* mellem tal og vektorer.

Skalarmultiplikationen er ikke som addition en komposition *inden for* $\mathbb{R}^n$ selv. Det er en operation på et *tal* (fra $\mathbb{R}$) *og en vektor* (fra $\mathbb{R}^n$), med en *vektor* som resultat.

Et par vigtige egenskaber ved skalarmultiplikation følger uden videre af definitionen:

(5) For ethvert $\underline{x}\in\mathbb{R}^n$ er

$$1\underline{x}=\underline{x}.$$

(6) For ethvert $\underline{x}\in\mathbb{R}^n$ og vilkårlige skalarer α og β gælder

$$\alpha(\beta\underline{x})=(\alpha\beta)\underline{x}.$$

Dette ses ved inspektion ud fra definitionen. Ad 5.:

$$1\underline{x} = 1(x_1, x_2, \ldots, x_n) = (x_1, x_2, \ldots, x_n) = \underline{x},$$

og ad 6.:

$$\alpha(\beta\underline{x}) = \alpha(\beta x_1, \beta x_2, \ldots, \beta x_n) = (\alpha\beta x_1, \alpha\beta x_2, \ldots, \alpha\beta x_n) = (\alpha\beta)\underline{x}.$$

Endvidere er det åbenbart at

$$0\underline{x} = \underline{0},$$

idet $0(x_1, x_2, \ldots, x_n) = (0, 0, \ldots, 0) = \underline{0}$.

Vi mangler at undersøge hvordan skalarmultiplikationen spiller sammen med additionen i $\mathbb{R}^n$. Der gælder de **distributive love**

(7) For vilkårlige $\underline{x}$ og $\underline{y}$ i $\mathbb{R}^n$ og enhver skalar α er

$$\alpha(\underline{x} + \underline{y}) = \alpha\underline{x} + \alpha\underline{y}.$$

(8) For ethvert $\underline{x} \in \mathbb{R}^n$ og vilkårlige skalarer α og β er

$$(\alpha + \beta)\underline{x} = \alpha\underline{x} + \beta\underline{x}.$$

Disse love fortæller øjensynlig hvordan man kan "gange ind i" og "sætte uden for parentes" ved skalarmultiplikation der involverer en sum af vektorer eller skalarer.

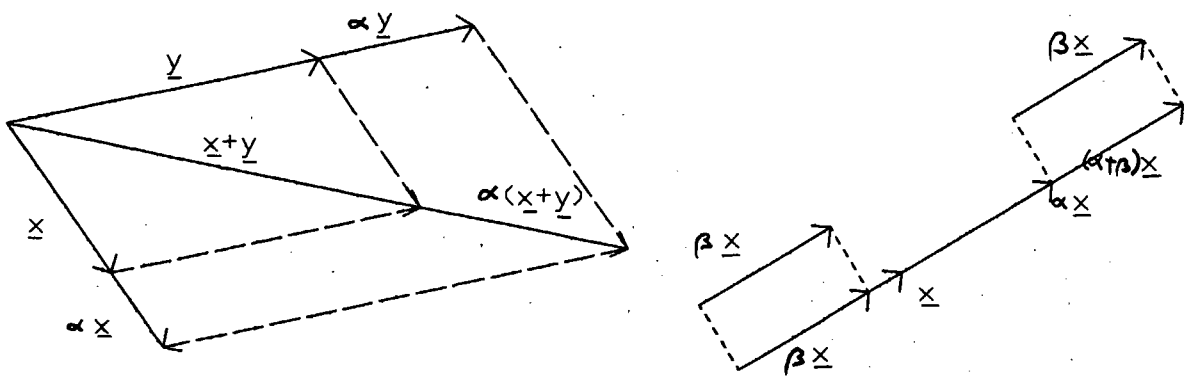
Også her overbeviser man sig om rigtigheden af påstandene ved simpel inspektion hvilende på de indførte definitioner af addition og skalarmultiplikation. Prøv selv at gøre rede for hvilke definitioner eller love der ligger bag de enkelte lighedstegn: Ad 7.:

$$\begin{aligned}
\alpha(\underline{x} + \underline{y}) &= \alpha(x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n) \\
&= (\alpha(x_1 + y_1), \alpha(x_2 + y_2), \ldots, \alpha(x_n + y_n)) \\
&= (\alpha x_1 + \alpha y_1, \alpha x_2 + \alpha y_2, \ldots, \alpha x_n + \alpha y_n) \\
&= (\alpha x_1, \alpha x_2, \ldots, \alpha x_n) + (\alpha y_1, \alpha y_2, \ldots, \alpha y_n) \\
&= \alpha\underline{x} + \alpha\underline{y}.
\end{aligned}$$

Ad 8.:

$$\begin{aligned}
(\alpha + \beta)\underline{x} &= ((\alpha + \beta)x_1, (\alpha + \beta)x_2, \ldots, (\alpha + \beta)x_n) \\
&= (\alpha x_1 + \beta x_1, \alpha x_2 + \beta x_2, \ldots, \alpha x_n + \beta x_n) \\
&= (\alpha x_1, \alpha x_2, \ldots, \alpha x_n) + (\beta x_1, \beta x_2, \ldots, \beta x_n) \\
&= \alpha\underline{x} + \beta\underline{x}.
\end{aligned}$$

Det geometriske indhold i 7. og 8. i $\mathbb{R}^2$ og $\mathbb{R}^3$ kan illustreres således:



Af hensyn til udviklingen af teorien i senere afsnit samler vi her de fundne grundregler for addition og skalarmultiplikation i $\mathbb{R}^n$:

(1) For alle $\underline{x}$ og $\underline{y}$: $\underline{x}+\underline{y}=\underline{y}+\underline{x}$ (den **kommutative** lov)

(2) For alle $\underline{x}$, $\underline{y}$ og $\underline{z}$: $(\underline{x}+\underline{y})+\underline{z}=\underline{x}+(\underline{y}+\underline{z})$ (den **associative** lov)

(3) Der findes netop ét **neutralt** element $\underline{0}$, så at for alle $\underline{x}$

$$\underline{x}+\underline{0}=\underline{x}$$

(4) Ethvert $\underline{x}$ har netop ét **modsat** element, $-\underline{x}$, opfyldende at

$$\underline{x}+(-\underline{x})=\underline{0}.$$

(5) For ethvert $\underline{x}\in\mathbb{R}^n$ er

$$1\underline{x}=\underline{x}.$$

(6) For ethvert $\underline{x}\in\mathbb{R}^n$ og vilkårlige skalarer α og β gælder

$$\alpha(\beta\underline{x})=(\alpha\beta)\underline{x}.$$

(7) For vilkårlige $\underline{x}$ og $\underline{y}$ i $\mathbb{R}^n$ og enhver skalar α er

$$\alpha(\underline{x}+\underline{y})=\alpha\underline{x}+\alpha\underline{y}$$

(den **første distributive** lov)

(8) For ethvert $\underline{x}\in\mathbb{R}^n$ og vilkårlige skalarer α og β er

$$(\alpha+\beta)\underline{x}=\alpha\underline{x}+\beta\underline{x}$$

(den **anden distributive** lov).

Det at $\mathbb{R}^n$ er udstyret med en addition og en skalarmultiplikation der opfylder 1.-8. udtrykkes sædvanligvis ved at sige, at $\mathbb{R}^n$ med disse operationer er organiseret som et **vektorrum**. Undertiden ser man også betegnelsen *lineært rum*. Senere hen vil vi få brug for at studere andre objekt-verdener (end talrummene) der kan organiseres som vektorrum.

De grundliggende lineære egenskaber ved talrummene kan udledes af 1.-8. Når først disse er etableret er det ikke nødvendigt hele tiden at gå tilbage til udgangspunktet i selve definitionerne af addition og skalarmultiplikation. (Men hvis det nu og da viser bekvemt for os at gå tilbage til definitionerne, vil vi dog ikke afholde os fra at gøre det.) Som eksempel på hvordan man kan udlede resultater alene ved hjælp af 1.-8. har vi allerede set resultatet om entydig ligningsløsning. Et par stykker til:

Først ser vi på påstanden: For ethvert $\underline{x} \in \mathbb{R}^n$ gælder

$$0\underline{x} = \underline{0}.$$

Denne påstand har vi ovenfor godtgjort ved netop at benytte definitionerne. Men det er altså ikke nødvendigt, se selv:

$$\underline{x} + 0\underline{x} = 1\underline{x} + 0\underline{x} = (1+0)\underline{x} = 1\underline{x} = \underline{x},$$

hvor vi har benyttet henholdsvis 5., 8., addition i $\mathbb{R}$, samt 5. én gang til. Adderes nu $-\underline{x}$ til begge sider af lighedstegnet, får vi

$$0\underline{x} = -\underline{x} + \underline{x} = \underline{0}.$$

Dernæst kan vi bevise, at for ethvert $\underline{x} \in \mathbb{R}^n$ må

$$(-1)\underline{x} = -\underline{x}.$$

At påstanden skal bevises, hvis vi kun har 1.-8. til rådighed skyldes, at $(-1)\underline{x}$ er fastlagt ved hjælp af skalarmultiplikationen, mens $-\underline{x}$ er fastlagt i forhold til addition. Vi har igen ud fra 5., 8. og addition i $\mathbb{R}$, at

$$\underline{x} + (-1)\underline{x} = 1\underline{x} + (-1)\underline{x} = (1 + (-1))\underline{x} = 0\underline{x} = \underline{0}.$$

Adderes også her $-\underline{x}$ til begge sider af lighedstegnet, finder vi

$$(-1)\underline{x} = -\underline{x}.$$

(Naturligvis kan dette resultat også opnås ved direkte at benytte den måde skalarmultiplikationen indføres på i $\mathbb{R}^n$; men det er altså ikke pointen her.)

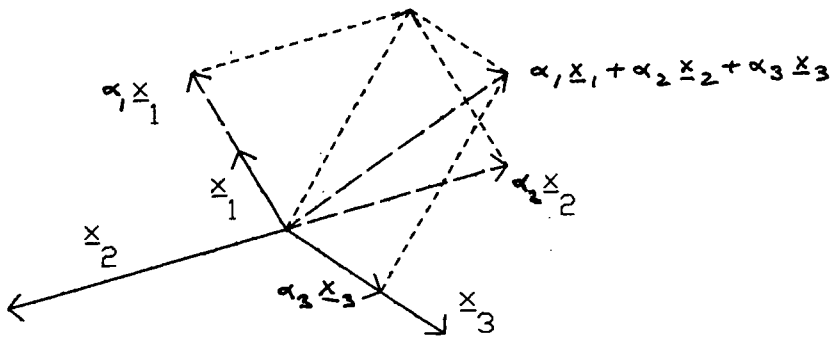
Øvelse. Vis alene på basis af 1.-8., at $\alpha\underline{0} = \underline{0}$ for enhver skalar α. Vis på samme grundlag, at $k\underline{x} = \underline{x} + \underline{x} + \ldots + \underline{x}$ (k led) for ethvert naturligt tal k (som jo også er en skalar) og enhver vektor $\underline{x} \in \mathbb{R}^n$.

Linearkombinationer

Hvis vi har givet et endeligt antal - f.eks. k - vektorer i $\mathbb{R}^n, \underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k$, kan vi danne nye vektorer af formen

$$(*) \qquad \alpha_1 \underline{x}_1 + \alpha_2 \underline{x}_2 + \ldots + \alpha_k \underline{x}_k,$$

hvor $\alpha_1, \alpha_2, \ldots, \alpha_k$ er vilkårlige skalarer. Et sådant udtryk, en sum af skalerede vektorer, kaldes en **linearkombination** af vektorerne $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k$ med **koefficienterne** $\alpha_1, \alpha_2, \ldots, \alpha_k$. Figuren nedenfor illusterer begrebet.



Fastholder vi en bestemt samling af vektorer $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k$, kan vi undersøge mængden af samtlige linearkombinationer $(*)$ af disse vektorer, når koefficienterne varierer frit i $\mathbb{R}$, altså

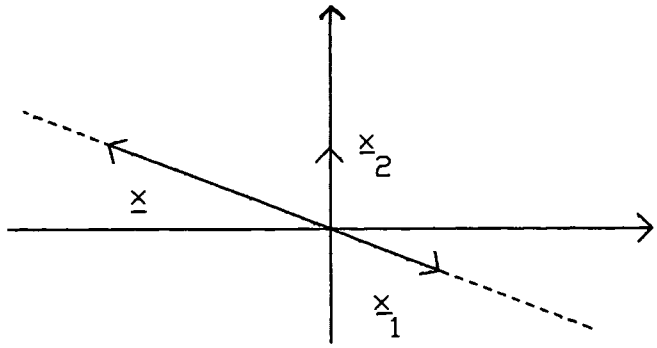
$$\{\alpha_1 \underline{x}_1 + \alpha_2 \underline{x}_2 + \ldots + \alpha_k \underline{x}_k | \alpha_1 \in \mathbb{R}, \alpha_2 \in \mathbb{R}, \ldots, \alpha_k \in \mathbb{R}\}$$

Denne mængde betegnes $span\{\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k\}$ og kaldes det af $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k$ **udspændte underrum** af $\mathbb{R}^n$. Vi siger også at $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k$ **udspænder** $span\{\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k\}$. Nogle gange ser man **frembringe** i stedet for udspænde. Vi skal senere vende tilbage til indholdet og betydningen af begrebet underrum. Foreløbig er det blot et navn.

Nu er det vist tid til et par eksempler.

Eksempel 1

Vi betragter i $\mathbb{R}^2$ vektoren $\underline{x}_1 = (3, -1)$. Det af $\underline{x}_1$ udspændte underrum $span\{\underline{x}_1\}$ består af alle vektorer af formen $\alpha \underline{x}_1$, hvor $\alpha \in \mathbb{R}$. Underrummet udgøres altså af alle talpar på linjen $\{(3\alpha, -\alpha) | \alpha \in \mathbb{R}\}$ gennem $\underline{0}$ (se figur).



Tilføjer vi f.eks. vektoren $\underline{x}_2 = (0,2)$, består det af $\underline{x}_1$, $\underline{x}_2$ udspændte underrum, $span\{\underline{x}_1, \underline{x}_2\}$, af alle vektorer af formen $\alpha_1\underline{x}_1 + \alpha_2\underline{x}_2$, for vilkårlige koefficienter α_1 og α_2. Tydeligvis udgør dette underrum hele $\mathbb{R}^2$: $span\{\underline{x}_1, \underline{x}_2\} = \mathbb{R}^2$. Ethvert givet $\underline{y} = (y_1, y_2)$ kan nemlig fremstilles som linearkombination af $\underline{x}_1$ og $\underline{x}_2$ med koefficienterne $\frac{y_1}{3}$ og $\frac{y_2}{2} + \frac{y_1}{6}$:

$$(y_1, y_2) = \alpha_1\underline{x}_1 + \alpha_2\underline{x}_2 = \alpha_1(3,-1) + \alpha_2(0,2) = (3\alpha_1, -\alpha_1 + 2\alpha_2).$$

Vi får så α_1 og α_2 af ligningssystemet $3\alpha_1 = y_1, -\alpha_1 + 2\alpha_2 = y_2$.

For $\alpha_1 = \frac{y_1}{3}$ og $\alpha_2 = \frac{1}{2}(y_2 + \alpha_1) = \frac{y_2}{2} + \frac{y_1}{6}$ har vi så

$$(y_1, y_2) = (\frac{y_1}{3})(3,-1) + (\frac{y_2}{2} + \frac{y_1}{6})(0,2) = (\frac{y_1}{3})\underline{x}_1 + (\frac{y_2}{2} + \frac{y_1}{6})\underline{x}_2.$$

Det er imidlertid ikke sådan at hvilke som helst to vektorer i $\mathbb{R}^2$ udspænder hele planen. F.eks. vil vi ved at tilføje $\underline{x} = (-6,2)$ til $\underline{x}_1$ ikke opnå at øge det udspændte underrum, eftersom

$$\alpha_1(3,-1) + \alpha_2(-6,2) = (3\alpha_1 - 6\alpha_2, -\alpha_1 + 2\alpha_2) = (\alpha_1 - 2\alpha_2)(3,-1)$$

stadig blot gennemløber den samme rette linje gennem $\underline{0}$ som før. Der gælder altså her, at

$$span\{\underline{x}_1, \underline{x}\} = span\{\underline{x}_1\}.$$

Skulle vi sluttelig få den idé at tilføje flere vektorer, f.eks. $\underline{x}_3, \ldots, \underline{x}_k$ til de hidtil betragtede $\underline{x}_1$ og $\underline{x}_2$, vil dette naturligvis ikke øge $span\{\underline{x}_1, \underline{x}_2\}$, der jo allerede udgør hele $\mathbb{R}^2$. Vi har altså

$$span\{\underline{x}_1, \underline{x}_2\} = span\{\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k\}.$$

Det er således ikke alene vektorernes antal der afgør hvor stor en delmængde af talrummet de udspænder. *

Eksempel 2

Hvis vi i $\mathbb{R}^3$ sætter $\underline{x}_1 = (3,1,2)$, udspænder $\underline{x}_1$ en ret linje i rummet gennem $\underline{0}$:

$$span\{\underline{x}_1\} = \{\alpha(3,1,2) | \alpha \in \mathbb{R}\} = \{(3\alpha, \alpha, 2\alpha) | \alpha \in \mathbb{R}\}.$$

Inddrager vi $\underline{x}_2 = (-1,-1,4)$, vil $\underline{x}_1$ og $\underline{x}_2$ til sammen udspænde

$$span\{\underline{x}_1, \underline{x}_2\} = \{\alpha(3,1,2) + \beta(-1,-1,4) | \alpha \in \mathbb{R}, \beta \in \mathbb{R}\},$$

der udgør en plan i $\mathbb{R}^3$ gennem $\underline{0}$. (Igen kunne et andet valg af $\underline{x}_2$, feks. $\left(-\frac{3}{10}, -\frac{1}{10}, -\frac{2}{10}\right)$, have bevirket at det af $\underline{x}_1$ og $\underline{x}_2$ udspændte underrum var det samme som det $\underline{x}_1$ alene kunne klare at udspænde.)

Hvad nu hvis der optræder en tredje vektor, $\underline{x}_3$? Der er to muligheder. Enten udspænder $\underline{x}_1$, $\underline{x}_2$ og $\underline{x}_3$ den samme plan som før, eller de udspænder hele $\mathbb{R}^3$. Det første er tilfældet, hvis $\underline{x}_3$ selv er en linearkombination af $\underline{x}_1$ og $\underline{x}_2$ (eks.: $\underline{x}_3 = (1,-1,10)$). Så vil jo enhver linearkombination af $\underline{x}_1$, $\underline{x}_2$ og $\underline{x}_3$ kunne fremstilles som en linearkombination af $\underline{x}_1$ og $\underline{x}_2$ alene.

Den anden mulighed indtræffer f.eks. med $\underline{x}_3 = (0,2,1)$. Enhver vektor $\underline{y} = (y_1, y_2, y_3)$ i $\mathbb{R}^3$ kan nemlig fremstilles som en linearkombination af $\underline{x}_1$, $\underline{x}_2$ og $\underline{x}_3$

$$(y_1, y_2, y_3) = \alpha(3,1,2) + \beta(-1,-1,4) + \gamma(0,2,1) = (3\alpha - \beta, \alpha - \beta + 2\gamma, 2\alpha + 4\beta + \gamma)$$

med passende valg af koefficienter α, β og γ. At bestemme sådanne koefficienter kommer jo ud på at løse ligningssystemet

$$\begin{aligned} 3\alpha - \beta &= y_1 \\ \alpha - \beta + 2\gamma &= y_2 \\ 2\alpha + 4\beta + \gamma &= y_3 \end{aligned}$$

i α, β og γ. Ved anvendelse af de metoder der blev indført i afsnittet om lineære ligningssystemer, finder vi at dette system (med forbehold for regnefejl) har løsningerne

$$(\alpha, \beta, \gamma) = \frac{1}{30}(9y_1 - y_2 + 2y_3, -3y_1 - 3y_2 + 6y_3, -6y_1 + 14y_2 + 2y_3).$$

Med dette valg af koefficienter kan $\underline{y}$ altså fremstilles som linearkombination af $\underline{x}_1$, $\underline{x}_2$ og $\underline{x}_3$. Tilføjer vil til disse vektorer en hvilken som helst endelig samling af vektorer, $\underline{x}_4, \ldots, \underline{x}_k$ udspændes stadig ikke mere end $\mathbb{R}^3$. Hver af de nye vektorer er selv linearkombinationer af $\underline{x}_1$, $\underline{x}_2$, og $\underline{x}_3$, så en linearkombination af samtlige $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k$ vil kunne fremstilles som en linearkombination af $\underline{x}_1$, $\underline{x}_2$, og $\underline{x}_3$ alene. *

Med inspiration fra disse eksempler vil vi notere et par generelle observationer - faktisk små sætninger - gyldige i ethvert $\mathbb{R}^n$.

* Det gælder altid at $span\{\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k\}$ vil indeholde $\underline{0}$, idet $\underline{0}$ fremkommer ved at benytte koefficienterne $\alpha_1 = \alpha_2 = \ldots = \alpha_k = 0$.

* Hvis der til $\mathbf{U} = span\{\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k\}$ føjes vektorer (de kunne hedde $\underline{y}_1, \ldots, \underline{y}_p$) der allerede ligger i $\mathbf{U}$, dvs. selv er linearkombinationer af $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k$, øges $span\{\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k\}$ ikke, dvs.

$$span\{\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k, \underline{y}_1, \ldots, \underline{y}_p\} = span\{\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k\}.$$

Et argument der formaliserer det ovenstående: Når $\underline{y}$'erne er linearkombinationer af $\underline{x}$'erne findes passende koefficienter (betegnet $a_{ij}, i = 1, 2, \ldots, p, j = 1, 2, \ldots, k$), så at

$$\begin{aligned}
\underline{y}_1 &= a_{11}\underline{x}_1 + a_{12}\underline{x}_2 + \ldots + a_{1k}\underline{x}_k \\
\underline{y}_2 &= a_{21}\underline{x}_1 + a_{22}\underline{x}_2 + \ldots + a_{2k}\underline{x}_k \\
&\vdots \\
\underline{y}_p &= a_{p1}\underline{x}_1 + a_{p2}\underline{x}_2 + \ldots + a_{pk}\underline{x}_k.
\end{aligned}$$

En vilkårlig linearkombination af $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k, \underline{y}_1, \ldots, \underline{y}_p$ har skikkelsen

$$\begin{aligned}
&\alpha_1\underline{x}_1 + \alpha_2\underline{x}_2 + \ldots + \alpha_k\underline{x}_k + \beta_1\underline{y}_1 + \ldots + \beta_p\underline{y}_p \\
&= \alpha_1\underline{x}_1 + \alpha_2\underline{x}_2 +, \ldots, +\alpha_k\underline{x}_k \\
&\quad + \beta 1(a_{11}\underline{x}_1 + a_{12}\underline{x}_2 + \ldots + a_{1k}\underline{x}_k) \\
&\quad + \beta 2(a_{21}\underline{x}_1 + a_{22}\underline{x}_2 + \ldots + a_{2k}\underline{x}_k) \\
&\quad \vdots \\
&\quad + \beta p(a_{p1}\underline{x}_1 + a_{p2}\underline{x}_2 + \ldots + a_{pk}\underline{x}_k) \\
&= \text{linearkombination af } \underline{x}\text{'erne (skriv selv koefficienterne op, jeg gider ikke).}
\end{aligned}$$

* Tilføjes $\underline{0}$ til $\{\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k\}$ ændres $span\{\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k\}$ ikke. Dette er et specialtilfælde af det foregående, idet $\underline{0}$ jo er en linearkombination af $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k$ med alle koefficienterne lig 0.

* Hvis $\{\underline{u}_1, \underline{u}_2, \ldots, \underline{u}_p\}$ er en delmængde af $\{\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_q\}$, er $span\{\underline{u}_1, \underline{u}_2, \ldots, \underline{u}_p\}$ en delmængde af $span\{\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_q\}$. Enhver linearkombination af $\underline{u}_1, \underline{u}_2, \ldots, \underline{u}_p$ er nemlig også en linearkombination af $\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_q$, idet vi foran de vektorer som ikke er i $\{\underline{u}_1, \underline{u}_2, \ldots, \underline{u}_p\}$ blot anbringer koefficienten 0. Så bliver resultatet en linearkombination af $\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_q$.

* For enhver endelig mængde af vektorer, $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k$, vil underrummet $\mathbf{U} = span\{\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k\}$ være *lukket over for addition og skalarmultiplikation.* Det betyder dels at hvis $\underline{x}$ og $\underline{y}$ begge ligger i $\mathbf{U}$, vil $\underline{x} + \underline{y}$ også ligge i $\mathbf{U}$, og dels at $\alpha\underline{x}$ også

ligger i **U** når $\underline{x}$ gør det. Argumentet for det er selvsagt, at hvis $\underline{x}$ og $\underline{y}$ begge er linearkombinationer af $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k$, gælder det samme for $\underline{x} + \underline{y}$ og for $\alpha\underline{x}$. Desuden *må* **U** *for egen regning opfylde kravene 1.- 8.* Det er ikke nogen overraskelse, at det opfylder kravene 1., 2., og 5.-8. De gælder jo for alle vektorer i $\mathbb{R}^n$, og derfor selvfølgelig også for dem der befinder sig i **U**. Men også 3. og 4. er opfyldt af **U** for sig. Dvs. (3.) at den neutrale vektor $\underline{0}$ befinder sig i **U** - det så vi ovenfor at den altid gør - og at (4.) den modsatte vektor $-\underline{x}$ til en vektor $\underline{x} \in$ **U** også befinder sig i **U**. At det sidste er sandt skyldes, at hvis $\underline{x}$ er en linearkombination af $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k$, så er $-\underline{x}$ det også, blot med de modsatte koefficienter.

$*$ $\mathbb{R}^n$ er selv et underrum i $\mathbb{R}^n$! Det er nemlig altid muligt at angive et endeligt antal vektorer, der frembringer hele $\mathbb{R}^n$. *OBS! Dette er meget vigtigt.* Sætter vi nemlig

$$\begin{aligned}
\underline{e}_1 &= (1,0,0,\ldots,0)\\
\underline{e}_2 &= (0,1,0,\ldots,0)\\
\underline{e}_3 &= (0,0,1,\ldots,0)\\
&\vdots\\
\underline{e}_n &= (0,0,0,\ldots,1)
\end{aligned}$$

gælder åbenbart for et vilkårligt $\underline{x} = (x_1, x_2, \ldots, x_n)$, at

$$\begin{aligned}
\underline{x} &= (x_1, x_2, \ldots, x_n)\\
&= (x_1, 0, \ldots, 0) + (0, x_2, \ldots, 0) + \ldots + (0, 0, \ldots, x_n)\\
&= x_1(1, 0, \ldots, 0) + x_2(0, 1, \ldots, 0) + \ldots + x_n(0, 0, \ldots, 1)\\
&= x_1\underline{e}_1 + x_2\underline{e}_2 + \ldots + x_n\underline{e}_n.
\end{aligned}$$

Dette viser, at $\underline{x}$ er en linearkombination af $\underline{e}_1, \underline{e}_2, \ldots, \underline{e}_n$. Enhver vektor i $\mathbb{R}^n$ er altså en linearkombination af disse vektorer, som kaldes (de sædvanlige) *grundvektorer*. Vi kan skrive

$$\mathbb{R}^n = span\{\underline{e}_1, \underline{e}_2, \ldots, \underline{e}_n\}.$$

Disse grundvektorer er ikke de eneste der fremstiller alle vektorer i $\mathbb{R}^n$. Der er uendeligt mange andre valg.

$*$ I ethvert $\mathbb{R}^n$ vil mængden $\{\underline{0}\}$ alene bestående af nulvektoren i $\mathbb{R}^n$ være et underrum, idet jo $\{\underline{0}\} = span\{\underline{0}\}$. Eftersom vi ovenfor indså at ethvert underrum vil indeholde $\underline{0}$, er $\{\underline{0}\}$ *det mindste underrum* i $\mathbb{R}^n$, sådan at forstå at det er en delmængde af ethvert andet.

$*$ Vi har hidtil kun beskæftiget med underrum udspændt af endeligt mange vektorer. Der er imidlertid ingen problemer med også at tale om underrum udspændt af

en vilkårlig mængde af vektorer, endelig eller uendelig. Hvis nemlig M er en vilkårlig delmængde af $\mathbb{R}^n$, kan vi danne samtlige linearkombinationer af vilkårlige sæt bestående af *endeligt* mange vektorer fra M. Denne samling linearkombinationer betegnes **span**(M). Vi har altså defineret

$$span(\mathrm{M}) = \{\alpha_1\underline{x}_1 + \alpha_2\underline{x}_2 + \ldots + \alpha_j\underline{x}_j | j \in \mathrm{N}, \underline{x}_1, \underline{x}_2 \ldots, \underline{x}_j \in \mathrm{M}, \alpha_1, \alpha_2, \ldots, \alpha_j \in \mathbb{R}\}$$

Læg mærke til at j ikke - som før k - er et fast tal. Hvis M er en endelig mængde bestående af k vektorer er vi tilbage i det tidligere behandlede tilfælde. Så kan j antage alle værdier i $\{1, 2, \ldots, k\}$, men vi kan få samtlige linearkombinationer frem ved kun at betragte $j = k$, idet linearkombinationerne svarende til mindre værdier af j fås ved at have nogle af koefficienterne til at være 0. Hvis derimod M er en uendelig mængde, er der ingen øvre grænse for hvor store sæt af endeligt mange vektorer vi kan plukke ud af M. Derfor kan j antage alle værdier i N. Det er vigtigt at være klar over at vi kun tillader linearkombinationer af *endeligt mange* vektorer. Lineær algebra er ikke en sportsgren der dyrker uendelige summer.

Lineær uafhængighed

I eksemplerne i foregående afsnit så vi at ét og samme underrum i $\mathbb{R}^n$ kan udspændes af mange forskellige sæt vektorer. Bortset fra i det særlige tilfælde hvor underrummet kun består af nulvektoren, $\underline{0}$, er der ikke nogen øvre grænse for hvor mange vektorer der kan udspænde et givet underrum. Vi kan nemlig blot til et udspændende sæt $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k$ tilføje vilkårlige vektorer, som er linearkombinationer af disse $\underline{x}$'er. Derimod må der altid være et mindste antal vektorer der kan udspænde et givet underrum. Der kan jo ikke være færre end én vektor i et udspændende sæt. Det er derfor nærliggende nærmere at undersøge mulighederne for at bestemme et mindste sæt udspændende vektorer for et givet underrum.

Vi vil nu påbegynde denne undersøgelse. Til den ende er begrebet **lineær uafhængighed** centralt.

Et sæt af endeligt mange vektorer, $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k$ i $\mathbb{R}^n$ siges at være **lineært uafhængigt**, hvis nulvektoren $\underline{0}$ kun kan fremstilles som linearkombination af $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k$ ved at alle koefficienterne er 0. Altså lidt mere formelt: $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k$ er per definition **lineært uafhængige** hvis og kun hvis det gælder, at man af

$$\alpha_1\underline{x}_1 + \alpha_2\underline{x}_2 + \ldots + \alpha_k\underline{x}_k = \underline{0}$$

kan slutte at $\alpha_1 = \alpha_2 = \ldots = \alpha_k = 0$.

Eftersom det altid er muligt at fremstille $\underline{0}$ som linearkombination af et hvilket som helst sæt af endeligt mange vektorer ved at lade alle koefficienterne være 0, er det

nærliggende at kalde en sådan linearkombination **triviel** (dette ord bruges ofte i matematisk jargon som karakterisering af det bundløst banale). Derefter er det forståeligt, hvorfor man ofte kalder en linearkombination der fremstiller $\underline{0}$ *uden* at alle koefficienterne er 0, for **ikke-triviel**. Vi bruger også vendingen "en ikke-triviel fremstilling af $\underline{0}$". På den måde kan man omformulere definitionen af lineær uafhængighed:

Et sæt af endeligt mange vektorer er lineært uafhængigt, netop hvis det er umuligt at give en ikke-triviel fremstilling af $\underline{0}$ ved hjælp af disse vektorer.

Eksempel 3

I $\mathbb{R}^2$ betragter vi vektorerne $\underline{x}_1 = (3,-1), \underline{x}_2 = (0,2), \underline{x}_3 = (-6,2)$ som vi også behandlede i Eksempel 1. Her er sættet bestående af $\underline{x}_1$ og $\underline{x}_2$ lineært uafhængigt, thi hvis

$$\alpha_1\underline{x}_1 + \alpha_2\underline{x}_2 = \underline{0},$$

altså hvis $\alpha_1(3,-1) + \alpha_2(0,2) = (0,0)$, har vi at α_1 og α_2 må opfylde ligningssystemet

$$\begin{aligned} 3\alpha_1 &= 0 \\ -\alpha_1 + 2\alpha_2 &= 0. \end{aligned}$$

Dette ligningssystemt har åbenbart ikke andre løsninger end

$$\alpha_1 = \alpha_2 = 0.$$

Men så er $\underline{x}_1$ og $\underline{x}_2$ lineært uafhængige. Læg mærke til at vi i Eksempel 1 fandt at disse to vektorer udspænder hele $\mathbb{R}^2$.

Derimod er vektorerne $\underline{x}_1$ og $\underline{x}_3$ ikke lineært uafhængige. Thi da $\underline{x}_3$ er proportional med $\underline{x}_1$, idet vi har $\underline{x}_3 = -2\underline{x}_1$, kan disse vektorer ikke være lineært uafhængige. For idet

$$\underline{x}_1 + \frac{1}{2}\underline{x}_3 = \underline{x}_1 + \frac{1}{2}(-2\underline{x}_1) = \underline{0},$$

har vi en fremstilling af $\underline{0}$ som linearkombination af $\underline{x}_1$ og $\underline{x}_3$ med andre koefficienter end lutter 0'er. I Eksempel 1 så vi, at $\underline{x}_1$ og $\underline{x}_3$ ikke udspændte hele $\mathbb{R}^2$.

Når $\underline{x}_1$ og $\underline{x}_3$ ikke er lineært uafhængige, er hele sættet $\underline{x}_1, \underline{x}_2, \underline{x}_3$ "endnu mindre" lineært uafhængigt. Det er jo let at fremstille $\underline{0}$ som linearkombination af disse tre vektorer uden at **alle** koefficienterne er 0:

$$\underline{x}_1 + 0\underline{x}_2 + \frac{1}{2}\underline{x}_3 = \underline{0}.$$

Vi har blot proppet $0\underline{x}_2$ ind i den tidligere linearkombination af $\underline{x}_1$ og $\underline{x}_3$, der allerede selv kunne klare at levere en ikke-triviel fremstilling af $\underline{0}$. *

Hvis et sæt af endeligt mange vektorer i $\mathbb{R}^n$ ikke er lineært uafhængigt, siges det at være **lineært afhængigt**. Med et lineært afhængigt sæt *findes* altså en ikke-triviel fremstilling af $\underline{0}$.

En række enkle, men vigtige observationer byder sig umiddelbart til:

* *Enhver enkelt vektor $\underline{x}$ som ikke er $\underline{0}$, udgør et lineært uafhængigt sæt.* Thi hvis $\alpha\underline{x} = \underline{0}$ var en ikke-triviel fremstilling af $\underline{0}$, altså med α forskellig fra 0, ville $\underline{x} = (1/\alpha)\alpha\underline{x} = (1/\alpha)\underline{0} = \underline{0}$. Dette er i strid med forudsætningen om at $\underline{x}$ ikke er $\underline{0}$.

* *Intet sæt af vektorer i $\mathbb{R}^n$ der indeholder $\underline{0}$ kan være lineært uafhængigt.* Med andre ord, ethvert sådant sæt er lineært afhængigt. En linearkombination af vektorerne hvor der står en (vilkårlig) koefficient forskellig fra 0 foran $\underline{0}$, og koefficienten 0 foran resten af vektorerne i sættet, giver jo en ikke-triviel fremstilling (ikke alle koefficienter er 0) af $\underline{0}$.

Det er en konsekvens heraf, at hvis der til et lineært uafhængigt sæt af vektorer føjes $\underline{0}$ bliver det udvidede sæt lineært afhængigt.

* *Hvis vi har at gøre med et lineært uafhængigt sæt af vektorer i $\mathbb{R}^n$, er ethvert delsæt også lineært uafhængigt.* Thi hvis der fandtes en ikke-triviel fremstilling af $\underline{0}$ ved hjælp af delsættet, kunne man skaffe en ikke-triviel fremstilling af $\underline{0}$ ud fra det oprindelige sæt ved simpelthen at anbringe koefficienten 0 foran de vektorer der ikke indgår i delsættet. Men da det oprindelige sæt er lineært uafhængigt er det umuligt at skaffe en ikke-triviel fremstilling af $\underline{0}$ med dette sæt.

Det følger af dette resultat, at *hvis man til et lineært afhængigt sæt af vektorer føjer et hvilket som helst antal vektorer, er det resulterende sæt også lineært afhængigt.* Var det nemlig lineært uafhængigt, måtte - efter det ovenstående - det samme gælde for ethvert delsæt, herunder det oprindelige, som imidlertid var forudsat at være lineært afhængigt.

* *Hvis i et sæt af vektorer i $\mathbb{R}^n$ én (eller flere) af vektorerne i sættet er en linearkombination af nogle af de øvrige, må sættet være lineært afhængigt.* Lad os nemlig antage, at $\underline{x}$ og $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k$ indgår i sættet sådan at

$$\underline{x} = \alpha_1\underline{x}_1 + \alpha_2\underline{x}_2 + \ldots + \alpha_k\underline{x}_k$$

for passende koefficienter $\alpha_1, \alpha_2, \ldots, \alpha_k$. Så er

$$1\underline{x} - \alpha_1\underline{x}_1 - \alpha_2\underline{x}_2 - \ldots - \alpha_k\underline{x}_k = \underline{0}$$

en ikke-triviel fremstilling af nulvektoren, da koefficienten til $\underline{x}$ ikke er 0. Det bevirker, at $\underline{x}$ og $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k$ udgør et lineært afhængigt sæt. Efter det foregående er så også det sæt de er plukket ud af lineært afhængigt.

Et specialtilfælde af den betragtede situation har vi, hvis et sæt indeholder én eller flere identiske vektorer. Et sådant sæt er altså nødvendigvis lineært afhængigt.

Det næste resultat er så vigtigt at det ophøjes til en sætning med et pænt bevis.

SÆTNING. *Et sæt af endeligt mange vektorer i $\mathbb{R}^n$ er lineært afhængigt hvis og kun hvis (mindst) én af vektorerne er en linearkombination af de øvrige.*

BEVIS: Der er to påstande at vise:

(1) Hvis $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k$ er et lineært afhængigt sæt af vektorer, er mindst én af dem en linearkombination af de øvrige. Dette indses således: Da sættet er lineært afhængigt, findes der en ikke-triviel linearkombination af dem som fremstiller nulvektoren:

$$\alpha_1 \underline{x}_1 + \alpha_2 \underline{x}_2 + \ldots + \alpha_k \underline{x}_k = \underline{0},$$

hvor altså mindst ét af α'erne ikke er 0. Lad os antage at α_j ikke er 0. Så kan vi isolere $\alpha_j \underline{x}_j$ på den ene side af lighedstegnet, og derefter multiplicere ligningen med $\frac{1}{\alpha_j}$.(Hele denne operation er lovlig i kraft af egenskaberne 1.-8.) Dette giver en fremstilling af $\underline{x}_j$ som linearkombination af de øvrige. Formelt skrevet op bliver resultatet

$$\underline{x}_j = -(\frac{\alpha_1}{\alpha_j})\underline{x}_1 - \ldots - (\frac{\alpha_{j-1}}{\alpha_j})\underline{x}_{j-1} - (\frac{\alpha_{j+1}}{\alpha_j})\underline{x}_{j+1} - \ldots - (\frac{\alpha_k}{\alpha_j})\underline{x}_k,$$

hvor α_j eventuelt kan være α_1 eller α_k. Hermed har vi vist "kun hvis"-delen af sætningen.

(2) Hvis én af vektorerne er en linearkombination af de øvrige, er sættet lineært afhængigt. Denne påstand har vi godtgjort under ét af observationspunkterne ovenfor. Dermed er også "hvis"-delen klaret. *Q.E.D.*

BASIS OG DIMENSION

Basisbegrebet

Vi skal nu nærme os en besvarelse af det spørgsmål som blev stillet i begyndelsen af det foregående afsnit, "hvor få vektorer kan vi nøjes med til at udspænde et givet underrum af $\mathbb{R}^n$?". I besvarelsen er begrebet **basis** centralt.

Et endeligt sæt af vektorer $\underline{b}_1, \underline{b}_2, \ldots, \underline{b}_k$ i et underrum $\mathbf{U}$ af $\mathbb{R}^n$ kaldes en **basis** for $\mathbf{U}$, hvis følgende *to krav* er opfyldt:

(1) Enhver vektor i $\mathbf{U}$ kan fremstilles som en linearkombination af $\underline{b}_1, \underline{b}_2, \ldots, \underline{b}_k$.

Eftersom enhver linearkombination af vektorer i $\mathbf{U}$ selv ligger i $\mathbf{U}$ (et underrum er lukket over for addition og skalarmultiplikation, jf. tidligere bemærkning) kan dette også udtrykkes:

$$\mathbf{U} = span\{\underline{b}_1, \underline{b}_2, \ldots, \underline{b}_k\}.$$

(2) Sættet $\underline{b}_1, \underline{b}_2, \ldots, \underline{b}_k$ er lineært uafhængigt.

Foreløbig ved vi ikke hvilke underrum der overhovedet besidder en basis.

Underrummet $\{\underline{0}\}$ *har ingen basis*, da $\underline{0}$, som er alene om at udfylde $\{\underline{0}\}$, ikke er et lineært uafhængigt sæt.

Men derudover kan vi konstatere at $\mathbb{R}^n$, der jo er et underrum i sig selv, har en basis. Vi har nemlig tidligere indset, at vektorerne $\underline{e}_1, \underline{e}_2, \ldots, \underline{e}_n$, bestemt ved

$$\begin{aligned}
\underline{e}_1 &= (1, 0, 0, \ldots, 0) \\
\underline{e}_2 &= (0, 1, 0, \ldots, 0) \\
\underline{e}_3 &= (0, 0, 1, \ldots, 0) \\
&\vdots \\
\underline{e}_n &= (0, 0, 0, \ldots, 1)
\end{aligned}$$

udspænder hele $\mathbb{R}^n$. Der resterer derfor blot at vise at de er lineært uafhængige. Men en linearkombination af $\underline{e}$'erne der fremstiller nulvektoren,

$$\alpha_1 \underline{e}_1 + \alpha_2 \underline{e}_2 + \ldots + \alpha_n \underline{e}_n = \underline{0},$$

må opfylde $(\alpha_1, \alpha_2, \ldots, \alpha_n) = \underline{0} = (0, 0, \ldots, 0)$ (udfør selv "udregningen"). Med andre ord må samtlige α'er være 0. Fremstillingen er derfor nødt til at være triviel, hvilket viser at $\underline{e}$'erne er lineært uafhængige. De udgør således en basis for $\mathbb{R}^n$.

SÆTNING. *Ethvert underrum* $\mathbf{U} = span\{\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k\}$, *som ikke udelukkende består af* $\underline{0}$, *har en basis.*

BEVIS: Hvis sættet $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k$ er lineært uafhængigt er påstanden åbenbart sand, al den stund $\underline{x}$'erne per definition udspænder $\mathbf{U}$. Hvis sættet derimod er lineært afhængigt, er det muligt at $k-1$ af vektorerne er lineært uafhængige. Hvis heller ikke det er tilfældet, er det muligt at $k-2$ af dem er det, og så fremdeles. På et tidspunkt må vi nå frem til det største antal $\underline{x}$'er der danner et lineært uafhængigt sæt. Om ikke andet består dette sæt kun af én vektor. Mindst én af vektorerne $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k$ er jo ikke $\underline{0}$, og en enkelt vektor forskellig fra $\underline{0}$ danner, i følge den første observation i det foregående afsnit, et lineært uafhængigt sæt.

Konklusionen af disse betragtninger er, at der findes et største antal, det kunne hedde j, lineært uafhængige vektorer blandt $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k$. Det er meget muligt at der findes flere forskellige måder at vælge disse j lineært uafhængige vektorer på. Vi gør bare brug af én af måderne. Der er ikke noget i vejen for at antage at de første j $\underline{x}$'er, $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_j$, udgør et maksimalt antal lineært uafhængige vektorer blandt $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k$. Ellers kunne vi bare have omnummereret dem.

Godt. Nu udgør altså $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_j$ et maksimalt lineært uafhængigt sæt blandt $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k$. Det indebærer at uanset hvilken af de resterende vektorer der føjes til $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_j$, bliver det udvidede sæt lineært afhængigt (ellers var j jo ikke det største antal lineært uafhængige). Føjes f.eks. $\underline{x}_{j+1}$ til $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_j$, er sættet $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_j, \underline{x}_{j+1}$ lineært afhængigt. Der må derfor findes en ikke-triviel fremstilling af $\underline{0}$ ved hjælp af disse vektorer

$$(*) \qquad \alpha_1 \underline{x}_1 + \ldots + \alpha_j \underline{x}_j + \alpha_{j+1} \underline{x}_{j+1} = \underline{0},$$

hvor mindst ét af α'erne er forskelligt fra 0. Nu kan det ikke tænkes at $\alpha_{j+1} = 0$. For i så fald ville vi jo have

$$\alpha_1 \underline{x}_1 + \ldots + \alpha_j \underline{x}_j = \underline{0},$$

hvor mindst ét blandt $\alpha_1, \ldots, \alpha_j$ må være forskelligt fra 0 (når nu α_{j+1} ikke er det). Men så ville der jo foreligge en ikke-triviel fremstilling af $\underline{0}$ ved hjælp af det lineært uafhængige sæt $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_j$; altså en modstrid med definitionen af lineær uafhængighed. Når α_{j+1} ikke er 0, kan $\underline{x}_{j+1}$ isoleres i $(*)$, dvs. fremstilles som linearkombination af $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_j$.

Med helt parallelle argumenter kan vi indse, at også de resterende $\underline{x}_{j+2}, \ldots, \underline{x}_k$ kan fremstilles som linearkombinationer af $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_j$.

Nu er ræsonnementet næsten fuldført. Alle vektorerne i $\mathbf{U}$ er linearkombinationer af $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k$. Men da $\underline{x}_{j+1}, \underline{x}_{j+2}, \ldots, \underline{x}_k$ er linearkombinationer af $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_j$, kan enhver linearkombination af $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k$ også skrives som en linearkombination af $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_j$. Dette sæt af vektorer udspænder altså hele $\mathbf{U}$. Da det samtidig er

lineært uafhængigt, opfylder det begge krav til en basis. Underrummet **U** besidder altså en basis. *Q.E.D.*

Eksempel 4

Gangen i beviset kan illustreres således: I $\mathbb{R}^3$ har — i kraft af sætningen — $\mathbf{U} = span\{(1,0,2),(-1,-1,5),(2,1,-3)\}$ en basis. Da $(2,1,-3) = (1,0,2) - (-1,-1,5)$ er de tre vektorer ikke lineært uafhængige. Det er derimod $(1,0,2)$ og $(-1,-1,5)$ (check selv!). Eftersom derved enhver linearkombination af de tre vektorer

$$\begin{aligned}
&\alpha_1(1,0,2) + \alpha_2(-1,-1,5) + \alpha_3(2,1,3) \\
&= \alpha_1(1,0,2) + \alpha_2(-1,-1,5) + \alpha_3\left((1,0,2) - (-1,-1,5)\right) \\
&= (\alpha_1 + \alpha_3)(1,0,2) + (\alpha_2 - \alpha_3)(-1,-1,5)
\end{aligned}$$

en linearkombination af de to første, udgør disse en basis for **U**. $*$

Dimension

Den ovenfor fundne basis for **U** blevet konstrueret ud fra det sæt $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k$ som udspændte **U**. Men der vil være uendeligt mange andre sæt som udspænder **U**. Måske kunne det tænkes, at der af dem kunne konstrueres baser for **U** med et andet antal elementer end j. Det viser sig, at dette ikke er muligt. Også dette resultat er så centralt, at det må mejsles ud i en sætning:

SÆTNING. *Enhver basis for et underrum* $\mathbf{U} = span\{\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k\}$ *af* $\mathbb{R}^n$ *har det samme antal elementer. Da* $\mathbb{R}^n$ *selv er et underrum gælder påstanden også for* $\mathbb{R}^n$.

BEVIS: Lad os antage et både $\underline{b}_1, \underline{b}_2, \ldots, \underline{b}_j$ og $\underline{c}_1, \underline{c}_2, \ldots, \underline{c}_i$ er baser for **U**. Vi skal vise at $i = j$. Hvis dette ikke var tilfældet, er enten $j < i$ eller $i < j$.

Vi ser først hvad der sker, hvis $j < i$.

Enhver af vektorerne $\underline{c}_1, \underline{c}_2, \ldots, \underline{c}_i$ er en linearkombination af $\underline{b}_1, \underline{b}_2, \ldots, \underline{b}_j$, som jo udspænder **U**. Specielt:

$$\underline{c}_1 = \alpha_1\underline{b}_1 + \alpha_2\underline{b}_2 + \ldots + \alpha_j\underline{b}_j.$$

Her kan ikke alle α'erne være 0, da så $\underline{c}_1$ ville være $\underline{0}$, i strid med at $\underline{c}_1$ indgår i et lineært uafhængigt sæt. Vi kan f.eks. antage at α_1 er forskellig fra 0. Så kan $\underline{b}_1$ udtrykkes som en linearkombination af $\underline{c}_1$ og $\underline{b}_2, \ldots, \underline{b}_j$. Det bevirker, at enhver linearkombination

af $\underline{b}_1, \underline{b}_2, \ldots, \underline{b}_j$, dvs. enhver vektor i $\mathbf{U}$, også kan fremstilles som en linearkombination af $\underline{c}_1, \underline{b}_2, \ldots, \underline{b}_j$. Det gælder specielt $\underline{c}_2$.

Vektoren $\underline{c}_2$ er altså en linearkombination af $\underline{c}_1, \underline{b}_2, \ldots, \underline{b}_j$. I denne linearkombination kan ikke alle $\underline{b}$-koefficienterne være 0, da i så fald $\underline{c}_2$ ville være proportional med $\underline{c}_1$, i strid med at de er lineært uafhængige. Vi kan gå ud fra at koefficienten til $\underline{b}_2$ ikke er 0, hvilket bevirker at $\underline{b}_2$ er en linearkombination af $\underline{c}_1, \underline{c}_2, \underline{b}_3, \ldots, \underline{b}_j$. Enhver linearkombination af $\underline{c}_1, \underline{b}_2, \ldots, \underline{b}_j$ - dvs. enhver vektor i $\mathbf{U}$ - kan dermed fremstilles som en linearkombination af $\underline{c}_1, \underline{c}_2, \underline{b}_3, \ldots, \underline{b}_j$. Dette sæt udspænder altså hele $\mathbf{U}$.

På denne måde kan vi blive ved med at udskifte et $\underline{b}$ ad gangen med et $\underline{c}$, og stadig have at gøre med et sæt der udspænder hele $\mathbf{U}$. Eftersom $i > j$ er der $\underline{c}$'er nok til at skifte alle $\underline{b}$'erne ud. Dermed er altså $\underline{c}_1, \underline{c}_2, \ldots, \underline{c}_j$ et sæt af vektorer der udspænder hele $\mathbf{U}$. Specielt må sættet udspænde de resterende $\underline{c}$'er (hvoraf der er mindst ét): $\underline{c}_{j+1}, \ldots, \underline{c}_i$. Disse $\underline{c}$'er er altså linearkombinationer af $\underline{c}_1, \underline{c}_2, \ldots, \underline{c}_j$. Men så fortæller den foregående sætning, at hele sættet $\underline{c}_1, \underline{c}_2, \ldots, \underline{c}_i$ er lineært afhængigt, i strid med at det udgør en basis. Vi føres altså til at konkludere, at forudsætningen $i > j$ ikke kan opretholdes.

Muligheden $i < j$ kan imidlertid heller ikke opretholdes. For hele det foregående ræsonnement kan jo gentages med $\underline{b}$'erne og $\underline{c}$'erne i ombyttede roller. Tilbage er kun muligheden $i = j$, som dermed er bevist. *Q.E.D.*

Det fælles antal vektorer i en basis for et givet underrum $\mathbf{U} = span\{\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k\}$ af $\mathbb{R}^n$ er et karakteristisk tal for underrummet. Tallet kaldes underrummets **dimension** og betegnes gerne $\mathbf{dim}(\mathbf{U})$. For talrummet $\mathbb{R}^n$ fandt vi at de n grundvektorer udgør en basis. Vi har derfor

SÆTNING. $\mathbb{R}^n$ *har dimensionen* n.

Da underrummet $\mathbf{U} = \{\underline{0}\}$ ikke har nogen basis giver det god mening at tillægge det dimensionen 0 : $dim\{\underline{0}\} = 0$.

I de følgende punkter er det hele tiden underforstået, at de underrum $\mathbf{U}$ der betragtes er af skikkelsen $span\{\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k\}$.

$*$ *For ethvert tal* $k = 1, 2, \ldots, n$ *findes et underrum* $\mathbf{U}$ *af* $\mathbb{R}^n$ *som har dimensionen* k. Vi kan nemlig blot vælge de k første af de n grundvektorer. De er lineært uafhængige, da de er et delsæt af et lineært uafhængigt sæt. De er derfor en basis for det underrum de udspænder. Underrummet har altså dimensionen k. De underrum der har dimensionen $n - 1$ kaldes **hyperplaner** i $\mathbb{R}^n$. Dette begreb er åbenbart en slags generalisation af begrebet plan i $\mathbb{R}^3$. Hyperplanerne i $\mathbb{R}^3$ er netop planerne gennem $\underline{0}$. Det er også i klar

analogi med situationen i $\mathbb{R}^2$ og $\mathbb{R}^3$, at vi kalder de 1-dimensionale underrum i $\mathbb{R}^n$ for rette linjer gennem 0. De har jo formen

$$\{\alpha \underline{a} \mid \alpha \in \mathbb{R}\},$$

hvor $\underline{a}$ ikke er $\underline{0}$. Denne mængde kaldes **den rette linje gennem $\underline{0}$ med retningen $\underline{a}$.**

Det er så også nærliggende at kalde en mængde af formen

$$\{\alpha \underline{a} + \underline{b} \mid \alpha \in \mathbb{R}\},$$

hvor stadig $\underline{a}$ er forskellig fra $\underline{0}$, og hvor $\underline{b}$ er en vilkårlig vektor i $\mathbb{R}^n$, for **den rette linje gennem $\underline{b}$ med retningen $\underline{a}$.** Obs! Denne mængde er *ikke* et underrum i $\mathbb{R}^n$ (hvorfor ikke?).

$*$ *I ethvert underrum* **U** *af dimension k vil et vilkårligt sæt af k lineært uafhængige vektorer fra* **U** *udgøre en basis.* Lad nemlig dette sæt hedde $\underline{c}_1, \underline{c}_2, \ldots, \underline{c}_k$. Desuden findes en basis for **U**, $\underline{b}_1, \underline{b}_2, \ldots, \underline{b}_k$. Nu kan vi med nøjagtigt det samme ræsonnement som i beviset for den næstsidste sætning skifte alle $\underline{b}$'erne ud med $\underline{c}$'er og opnå, at det resulterende sæt, altså $\underline{c}_1, \underline{c}_2, \ldots, \underline{c}_k$, frembringer **U**. Da det samtidig er lineært uafhængigt udgør det en basis for **U**.

Påstanden gælder også for $\mathbf{U} = \mathbb{R}^n$.

$*$ Som en pendant til det foregående punkt gælder: *I ethvert underrum U af dimension k vil et vilkårligt sæt af k vektorer som frembringer U udgøre en basis.* For at indse det lader vi igen sættet hedde $\underline{c}_1, \underline{c}_2, \ldots, \underline{c}_k$. Påstanden er godtgjort hvis det kan bevises at $\underline{c}$'erne er lineært uafhængige.

Var de i stedet lineært afhængige, måtte der være et største lineært uafhængigt delsæt (ikke alle $\underline{c}$'erne kan jo være $\underline{0}$). Dette delsæt måtte have færre end k elementer. Tilføjes et vilkårligt af de øvrige $\underline{c}$'er til dette maksimale lineært uafhængige delsæt, opstår et lineært afhængigt sæt. Derfor må den tilføjede vektor være en linearkombination af vektorerne i delsættet. Det ville føre til at enhver vektor i **U** - der jo er en linearkombination af $\underline{c}_1, \underline{c}_2, \ldots, \underline{c}_k$ - kunne fremstilles som en linearkombination af det lineært uafhængige delsæt af $\underline{c}_1, \underline{c}_2, \ldots, \underline{c}_k$ med færre elementer end k. Men så ville delsættet være en basis for **U** med færre end k elementer, i strid med at **U** har dimensionen k.

$*$ *I et underrum* **U** *af dimension k vil et vilkårligt sæt af $k+1$ eller flere vektorer være lineært afhængigt.* I følge det næstsidste punkt ville jo i modsat fald k af dem udgøre en basis for **U**. Derfor ville disse k frembringe hele **U**, specielt de resterende vektorer i sættet, i strid med antagelsen om at sættet var lineært uafhængigt. *Dimensionen af underrummet angiver altså det største antal lineært uafhængige vektorer i* **U**.

Specielt vil et vilkårligt sæt af $n+1$ eller flere vektorer i $\mathbb{R}^n$ være lineært afhængigt.

Mere om underrum

$*$ De ovenstående betragtninger er gennemført for underrum af formen $\mathbf{U} = span\{\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k\}$. Hvad nu hvis $\mathbf{M}$ er en vilkårlig delmængde af $\mathbb{R}^n$, og $\mathbf{U} = span(\mathbf{M})$? *Det gør ingen forskel!* Ethvert sådant underrum er nemlig nødvendigvis af formen $span\{\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k\}$ for passende valg af $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k$. Det overbeviser vi os om på følgende måde:

Der findes i $\mathbf{M}$ et største antal lineært uafhængige vektorer, jf. det foregående punkt. Dette antal kunne være k, og $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k$ et dertil svarende lineært uafhængigt sæt fra $\mathbf{M}$. Det kan godt være at $k = n$, men det spiller i øvrigt ingen rolle. Tilføjer vi til $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k$ en hvilken som helst vektor $\underline{x}$ fra $\mathbf{M}$ som ikke selv ligger i dette sæt, bliver det udvidede sæt lineært afhængigt, og med det sædvanlige argument ses, at $\underline{x}$ er en linearkombination af $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k$. Heraf følger at enhver vektor i $\mathbf{M}$ er en linearkombination af dette sæt.

Da $\{\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k\}$ er en delmængde af $\mathbf{M}$, er også $span\{\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k\}$ en delmængde af $span(\mathbf{M}) = \mathbf{U}$. Men omvendt må det også gælde at enhver vektor i $\mathbf{U}$ faktisk ligger i $span\{\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k\}$. Thi enhver vektor $\underline{y}$ i $\mathbf{U}$ er en linearkombination af vektorer fra $\mathbf{M}$, og hver af disse er en linearkombination af $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k$. Derved bliver $\underline{y}$ i alt en linearkombination af $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k$, altså et element i $span\{\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k\}$, hvilket var påstanden. Summa summarum er $span\{\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k\}$ og $\mathbf{U}$ delmængder af hinanden. Men så er de identiske:

$$\mathbf{U} = span\{\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k\}.$$

$*$ Vi kan nu give en meget nyttig karakterisering af underrummene $\mathbf{U} = span(\mathbf{M})$ i $\mathbb{R}^n$. Det er uden videre klart, at hvis $\underline{x}$ og $\underline{y}$ er to vektorer fra $\mathbf{U}$, vil også (1) $\underline{x} + \underline{y}$ og (2) $\alpha\underline{x}$ ligge i $\mathbf{U}$, hvor α er en vilkårlig skalar. Det skyldes simpelthen at både $\underline{x}$ og $\underline{y}$ er linearkombinationer af vektorer fra $\mathbf{M}$, hvorved også $\underline{x} + \underline{y}$ og $\alpha\underline{x}$ er linearkombinationer af vektorer fra $\mathbf{M}$, hvilket netop er kriteriet for at de befinder sig i $\mathbf{U}$.

Hvis nu $\mathbf{V}$ er en delmængde af $\mathbb{R}^n$, om hvilken vi kun ved at den opfylder (1) og (2), må $\mathbf{V}$ være et underrum. Nærmere bestemt er $\mathbf{V} = span(\mathbf{V})$. Thi på den ene side er det klart, at enhver vektor $\underline{x}$ i $\mathbf{V}$ er en linearkombination af vektorer i $\mathbf{V}$, nemlig $\underline{x} = 1\underline{x}$! Altså er $\mathbf{V}$ en delmængde af $span(\mathbf{V})$. Omvendt vil enhver linearkombination af vektorer fra $\mathbf{V}$ selv tilhøre $\mathbf{V}$ i kraft af (gentagen anvendelse) af (1) og (2) (check selv detaljerne). Det bevirker at $span(\mathbf{V})$ er en delmængde af $\mathbf{V}$. Men så er de to mængder ens.

De fundne resultater fortjener at blive samlet i en sætning, som altså er blevet bevist:

SÆTNING.

(a) *Ethvert underrum i* $\mathbb{R}^n$, $U = span(M)$, *hvor* $\mathbf{M}$ *er en vilkårlig delmængde af* $\mathbb{R}^n$, *er af formen* $\mathbf{U} = span\{\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k\}$ *for et passende sæt* $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k$ *fra* $\mathbf{M}$.

(b) *En delmængde* $\mathbf{V}$ *af* $\mathbb{R}^n$ *er et underrum i* $\mathbb{R}^n$, *hvis og kun hvis* $\mathbf{V}$ *opfylder*

(1) *Hvis* $\underline{x}$ *og* $\underline{y}$ *tilhører* $\mathbf{V}$, *vil også* $\underline{x} + \underline{y}$ *tilhøre* $\mathbf{V}$

(2) *Hvis* $\underline{x}$ *tilhører* $\mathbf{V}$ *og* α *er en skalar, vil* $\alpha\underline{x}$ *også tilhøre* $\mathbf{V}$.

Det er altså ikke andre underrum i $\mathbb{R}^n$ end dem af formen $span\{\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k\}$. De kan karakteriseres ved hjælp af (1) og (2).

Vi kan samtidig slå fast, at de delmængder af $\mathbb{R}^n$, der opfylder kravene 1.-8. - dvs. de samme delmængder som opfylder (1) og (2) - altså netop er underrummene i $\mathbb{R}^n$.

Koordinatfremstilling

Hvis $\mathbf{U}$ er et k-dimensionalt underrum i $\mathbb{R}^n$, og $\underline{b}_1, \underline{b}_2, \ldots, \underline{b}_k$ er en basis for $\mathbf{U}$, har enhver vektor $\underline{x}$ i $\mathbf{U}$ en fremstilling som linearkombination af $\underline{b}_1, \underline{b}_2, \ldots, \underline{b}_k$:

$$\underline{x} = \alpha_1\underline{b}_1 + \alpha_2\underline{b}_2 + \ldots + \alpha_k\underline{b}_k,$$

med passende koefficienter $\alpha_1, \alpha_2, \ldots, \alpha_k$. Principielt kunne det tænkes at $\underline{x}$ havde en anden fremstilling som linearkombination af $\underline{b}$'erne med nogle andre koefficienter $\beta_1, \beta_2, \ldots, \beta_k$:

$$\underline{x} = \beta_1\underline{b}_1 + \beta_2\underline{b}_2 + \ldots + \beta_k\underline{b}_k.$$

Dette er imidlertid ikke muligt, α-sættet og β-sættet må være identiske. Thi trækker vi den sidste fremstilling fra den første, får vi en fremstilling af $\underline{0}$ som linearkombination af $\underline{b}_1, \underline{b}_2, \ldots, \underline{b}_k$:

$$\underline{0} = (\alpha_1 - \beta_1)\underline{b}_1 + (\alpha_2 - \beta_2)\underline{b}_2 + \ldots + (\alpha k - \beta_k)\underline{b}_k.$$

Da $\underline{b}$'erne udgør en basis, er de lineært uafhængige, hvorved de kun tillader en triviel fremstilling af $\underline{0}$. Samtlige koefficienter i denne fremstilling må altså være 0. Med andre ord må α'erne og β'erne være identiske.

Der findes altså én og kun én fremstilling af $\underline{x}$ *som linearkombination af vektorerne i en given basis.* De entydigt bestemte koefficienter i denne fremstilling kaldes $\underline{x}$'s **koordinater** i forhold til basen. I forhold til en anden basis er $\underline{x}$'s koordinater selvsagt nogle andre, jf. Eksempel 5 nedenfor.

Eksempel 5

I Eksempel 2 betragtede vi i $\mathbb{R}^3$ vektorerne $\underline{x}_1 = (3,1,2), \underline{x}_2 = (-1,-1,4)$ og $\underline{x}_3 = (0,2,1)$, og indså at enhver vektor i $\mathbb{R}^3$ kan fremstilles som linearkombination af disse tre. Da dimensionen af $\mathbb{R}^3$ er 3, udgør disse vektorer en basis for $\mathbb{R}^3$.

I eksemplet fandt vi endvidere, at vektoren $\underline{y}$ kan skrives som

$$\underline{y} = \alpha(3,1,2) + \beta(-1,-1,4) + \gamma(0,2,1),$$

med $\alpha = \frac{1}{30}(9y_1 - y_2 + 2y_3), \beta = \frac{1}{30}(-3y_1 - 3y_2 + 6y_3)$ og $\gamma = \frac{1}{30}(-6y_1 + 14y_2 + 2y_3)$.

Sættet (α, β, γ) er således $\underline{y}$'s koordinater i basen $\underline{x}_1, \underline{x}_2, \underline{x}_3$. F.eks. har $\underline{e}_1 = (1,0,0)$ i *denne basis* koordinaterne $(\frac{9}{30}, -\frac{3}{30}, -\frac{6}{30})$. For $\underline{e}_2$ og $\underline{e}_3$ er koordinaterne henholdvis $(-\frac{1}{30}, -\frac{3}{30}, \frac{14}{30})$ og $(\frac{2}{30}, \frac{6}{30}, \frac{2}{30})$.

Vektoren $\underline{y}$'s koordinater i grundbasen $\{\underline{e}_1, \underline{e}_2, \underline{e}_3\}$ er (y_1, y_2, y_3), som er forskellig fra $\underline{y}$'s koordinater i basen $\{\underline{x}_1, \underline{x}_2, \underline{x}_3\}$. *

Det er altså vigtigt at være opmærksom på forskellen mellem en vektor - som selv er et talsæt - og dens koordinater i en given basis, som *også* er et talsæt, men ikke det samme, med mindre basen er grundbasen.

Eksempel 6

I $\mathbb{R}^7$ betragtes underummet $\mathbf{U}$ udspændt af vektorerne $\underline{x}_1 = (0,2,1,0,-3,1,4)$, $\underline{x}_2 = (-1,1,-2,0,0,3,0)$ og $\underline{x}_3 = (-3,-5,-10,0,12,5,-16)$. Her er $\underline{x}_3 = -4\underline{x}_1 + 3\underline{x}_2$, mens $\underline{x}_1$ og $\underline{x}_2$ øjensynlig er lineært uafhængige. Derfor er $\mathbf{U}$ et 2- dimensionalt underrum i $\mathbb{R}^7$, med $\{\underline{x}_1, \underline{x}_2\}$ som basis. Vektoren $\underline{x}_3$'s koordinater i denne basis er $(-4,3)$.

LINEÆRE AFBILDNINGER MELLEM TALRUM

Indledning

I ved godt hvor væsentlige reelle funktioner af reelle variable er, både i anvendelsesforbindelser og inden døre i matematikken selv. De funktioner man beskæftiger sig med i den gymnasiale matematikundervisning afhænger praktisk taget alle af én variabel med reelle tal som funktionsværdier. I har også - men måske ikke i matematikundervisningen - mødt eksempler på reelle funktioner af to variable - det kunne f.eks. være et geografisk steds højde over havoverfladen som funktion af stedets koordinater; eller spændingsforskellen U som funktion af resistans R og strømstyrke I i elektriske kredse (Ohms lov: $U = RI$); eller en indespærret luftarts temperatur T som funktion af dens tryk P og rumfang V ($T = \frac{PV}{nR}$); eller den samfundsmæssige produktion Y som funktion af arbejdskraften L og kapitalapparatet I ($Y = cL^{\alpha}I^{1-\alpha}$) etc.

Der er imidlertid ingen grund til at begrænse sig til sådanne funktioner af én eller to variable. Både fra et anvendelsessynspunkt og fra et matematisk synspunkt er der særdeles god mening i at betragte funktioner hvis variable hentes fra ét talrum $\mathbb{R}^n$, og hvis værdier befinder sig i et andet talrum $\mathbb{R}^p$, hvor både n og p kan være vilkårlige naturlige tal.

Vi kan f.eks. forestille os et vi i ethvert punkt på jordoverfladen - fastlagt ved (x_1, x_2) = (længdegrad,breddegrad) - ønskede at repræsentere et sæt meteorologiske data, f.eks. stedets lufttryk, lufttemperatur, luftfugtighed, vindhastighed, regnmålerstand, ved en vektor $(y_1, y_2, y_3, y_4, y_5)$ i $\mathbb{R}^5$. Hvis vi for ethvert punkt på Jorden havde en sådan repræsentation ville vi i realiteten stå med en **afbildning** F - en generaliseret funktion; man bruger ordet afbildning, hvis værdierne ikke ligger i $\mathbb{R}$ - der til ethvert punkt i et område $\mathbf{A}$ af $\mathbb{R}^2$ havde knyttet et punkt i $\mathbb{R}^5$:

$$F : \mathbf{A} \to \mathbb{R}^5,$$

hvor altså $\mathbf{A}$ er en delmængde af $\mathbb{R}^2$. Vi vil nemlig ikke bruge hele $\mathbb{R}^2$, fordi kun punkter svarende til sædvanlige længdegrader og breddegrader skal indgå i foretagendet. Det betyder at længdegraden kan antage en vinkelværdi mellem $-\pi$ og π (svarende til mellem 180 vestlig og 180 østlig længde), mens breddegraden kan antage en vinkelværdi mellem $-\frac{\pi}{2}$ (Sydpolen) og $\frac{\pi}{2}$ (Nordpolen). Som A vælger vi derfor $A = \{(x_1, x_2) | x_1 \in [-\pi, \pi[,\ x_2 \in [-\frac{\pi}{2}, \frac{\pi}{2}]\}$.

I denne opstilling forestiller vi os de meteorologiske data indsamlet på ét og kun ét tidspunkt. Tiden indgår ikke i modellen. Der er imidlertid intet i vejen for at tænke sig de samme slags data på et bestemt sted givet til forskellige tidspunkter. Det kunne foregå ved at vi ikke blot så på datasættet som funktion af stedets koordinater,

men også af tidspunktet. Den uafhængige variable skulle altså ikke længere være en vektor $(x_1, x_2) \in \mathbb{R}^2$, men en vektor $(x_1, x_2, x_3) \in \mathbb{R}^3$, hvor den nye komponent x_3 repræsenterer tiden. Så ville vi have at gøre med en ny afbildning

$$G : \mathbf{B} \to \mathbb{R}^5,$$

hvor $\mathbf{B}$ er en delmængde af $\mathbb{R}^3$ (nærmere bestemt $\mathbf{B} = \mathbf{A} \times \mathbb{R}$).

Der kan måske være grund til at gøre opmærksom på, at ingen af de to ovenstående "meteorologiske modeller" befatter sig med *indsamlingen* af data. Det er naturligvis ikke muligt at indsamle data i ethvert punkt i et uendeligt område af et talrum. Modellerne skal derfor anskues som *principmodeller*: I princippet eksisterer de anførte data for hvert eneste punkt på Jorden.

Dette var blot et enkelt eksempel blandt uendeligt mange til at illustrere fornuften i at se på generelle afbildninger mellem højere-dimensionale talrum.

Ligesom der findes uendeligt mange typer af "gymnasiefunktioner" (reelle funktioner af én reel variabel), findes der naturligvis uendeligt mange slags afbildninger mellem talrum. Vi skal i denne forbindelse kun betragte én eneste slags - men hvilken slags! - de **lineære afbildninger**. Skønt de lineære afbildninger kun udgør en lille, om end uendelig, skare i den store fauna af afbildninger, er det en meget vigtig skare. Dels fordi den i sig selv indfanger nogle centrale egenskaber ved verden og ved matematikken, og dels fordi den ved at bestå af særligt "tamme" afbildninger er hovedhjælpemidlet til at analysere og forstå mere "vilde" eksemplarer i afbildningsfaunaen.

Ingen af de ovenfor omtalte afbildninger er faktisk lineære. Imidlertid er vi stødt på lineære afbildninger i den "motiverende indledning", selv om vi ikke kaldte dem sådan dér, og selv om vi stadig ikke har defineret begrebet. Således betragtede vi i eksemplet *Bageriet* en situation at der til et givet produktkrav $\underline{x}$ (en vektor i $\mathbb{R}^n$ skrevet som $n \times 1$-søjle) svarede et råvarebehov $\underline{y}$ (en vektor i $\mathbb{R}^p$ skrevet som $p \times 1$-søjle), bestemt af $\underline{y} = A\underline{x}$, hvor A var den $p \times n$-matrix der beskriver produktionsgangen. Det betyder i virkeligheden, at der til ethvert $\underline{x} \in \mathbb{R}^n$ (hvoraf dog kun nogle er bagningsmæssigt realistiske, hvilket vi ser bort fra her) svarer et $\underline{y} \in \mathbb{R}^p$ (hvoraf igen kun nogle er realistiske). Dermed har vi at gøre med en afbildning

$$F : \mathbb{R}^n \to \mathbb{R}^p,$$

givet ved $\underline{y} = F(\underline{x}) = A\underline{x}$.

I indledningsafsnittet indførtes i eksemplet *Markedsandele* en afbildning af $\mathbb{R}^3$ ind i $\mathbb{R}^3$, givet ved $F(\underline{v}) \doteq A\underline{v}$, hvor $\underline{v}$ angiver de tre bryggeriers markedsandele i en bestemt måned, og $F(\underline{v})$ markedsandelene i den følgende måned. I eksemplet *Populationer* havde vi at gøre med afbildningen $F : \mathbb{R}^k \to \mathbb{R}^k$ bestemt ved

$$\underline{n}(1) = F(\underline{n}(0)) = L\underline{n}(0),$$

hvor L er en Leslie-matrix, og $\underline{n}(0)$ og $\underline{n}(1)$ er k- vektorer skrevet som søjler, der repræsenterer populationssammensætningen til perioderne 0 og 1.

Endelig betragtedes *Geometriske Transformationer*, som rotationer, drejninger og punktmultiplikation ud fra $\underline{0}$, der alle er afbildninger af $\mathbb{R}^2$ ind i $\mathbb{R}^2$ af formen $\underline{y} = A\underline{x}$.

Læg mærke til, at i alle disse eksempler var afbildningerne givet ved forskrifter af formen $\underline{y} = A\underline{x}$, hvor $\underline{x}$ og $\underline{y}$ er vektorer i passende talrum og A en dertil svarende matrix. Det vil lidt senere vise sig at dette netop er karakteriserende for lineære afbildninger. Men det er ikke sådan de defineres. Det sker således:

Begrebet lineær afbildning

Vi har at gøre med to talrum $\mathbb{R}^n$ og $\mathbb{R}^p$, hvor n og p kan være hvilke som helst naturlige tal. En afbildning

$$F : \mathbb{R}^n \to \mathbb{R}^p$$

siges at være **lineær**, hvis F tilfredsstiller følgende to krav:

(1) For vilkårlige vektorer $\underline{x}$ og $\underline{y}$ i $\mathbb{R}^n$ er:

$$F(\underline{x} + \underline{y}) = F(\underline{x}) + F(\underline{y})$$

(F er **additiv**)

og

(2) For enhver vektor $\underline{x}$ i $\mathbb{R}^n$ og enhver skalar α er:

$$F(\alpha\underline{x}) = \alpha F(\underline{x})$$

(F **respekterer skalering**).

De to krav (1) og (2) kan sammenfattes til ét:

L For vilkårlige vektorer $\underline{x}$ og $\underline{y} \in \mathbb{R}^n$ og vilkårlige skalarer α og β er:

$$F(\alpha\underline{x} + \beta\underline{y}) = \alpha F(\underline{x}) + \beta F(\underline{y})$$

(F **respekterer linearkombinering**).

Gør selv rede for dette som en **øvelse**.

Hvis $n = p$, d. v. s. $F : \mathbb{R}^n \to \mathbb{R}^p$, bruges ofte betegnelsen en **lineær transformation** om F.

$*$ Lineære afbildninger bærer deres navn med rette. *De afbilder en ret linje i en ret linje (eller i et punkt).* Den typiske rette linje i $\mathbb{R}^n$ er nemlig bestemt ved at gå gennem punktet $\underline{b}$ med retningen $\underline{a}$, hvor $\underline{a}$ ikke er $\underline{0}$. Den er derfor givet ved

$$\{t\underline{a} + \underline{b} \mid t \in \mathbb{R}\}.$$

For ethvert punkt på denne linje har vi, da F respekterer linearkombinering, at

$$F(t\underline{a} + \underline{b}) = tF(\underline{a}) + F(\underline{b}).$$

Det viser, at ethvert punkt på linjen i $\mathbb{R}^n$ givet ved $\underline{a}$ og $\underline{b}$ afbildes i et punkt på linjen i $\mathbb{R}^p$ gennem punktet $F(\underline{b})$ med retningen $F(\underline{a})$. Dette forudsætter dog, at $F(\underline{a})$ er forskellig fra $\underline{0}$, da der ellers ikke er tale om nogen ret linje, men kun om punktet $\underline{b}$. I det tilfælde hvor der tale om en ret linje, kommer alle punkter på linjen med som billede ved F af et punkt på den oprindelige linje. Thi punktet givet ved $tF(\underline{a}) + F(\underline{b})$, dvs. ved parameterværdien t, er jo billede af punktet $t\underline{a} + \underline{b}$ svarende til den samme parameterværdi.

Eksempel 1

De lineære afbildninger fra $\mathbb{R}$ ind i $\mathbb{R}$ - svarende til at $\mathbb{R}^n$ og $\mathbb{R}^p$ begge er $\mathbb{R}$ - må specielt opfylde $F(\alpha x) = \alpha x$ for alle $x \in \mathbb{R}$ og alle $\alpha \in \mathbb{R}$. Det er der meget få slags funktioner der kan stå model til. Vi har nemlig for ethvert $x \in \mathbb{R}$, at

$$F(x) = F(x \cdot 1) = xF(1),$$

idet x som reelt tal både kan opfattes som en vektor i $\mathbb{R}$ (det sker i $F(x)$) og som en skalar (det sker i $F(x \cdot 1)$ og i $xF(x)$). Kalder vi konstanten $F(1)$ for c, har vi altså at F må være givet ved

$$F(x) = cx, \text{ for } x \in \mathbb{R}.$$

Det er bemærkelsesværdigt, at dette er en følge alene af at F skal respektere skalering. Additivitetskravet har ikke været inddraget i ræsonnementet.

En sådan funktion kaldes af oplagte grunde for en *proportionalitet*. Enhver proportionalitet opfylder åbenbart fordringerne til en lineær afbildning, og altså også additivitetskravet (check selv!). *Af lineære funktioner af $\mathbb{R}$ i $\mathbb{R}$ findes der altså ikke andre end proportionaliteterne.*

Så fredelige funktioner som

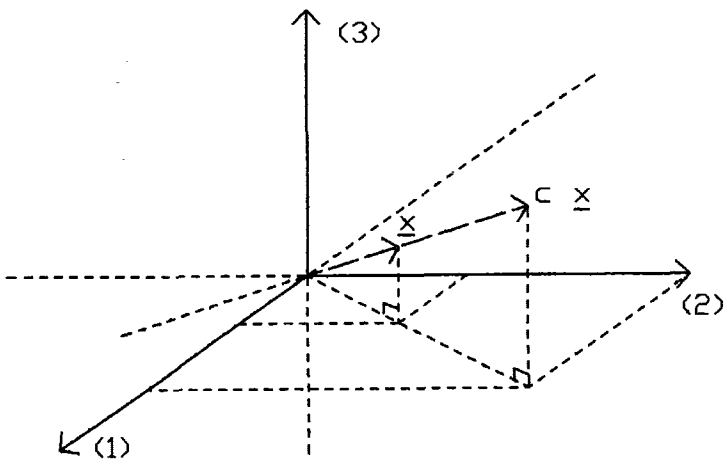
$$F(x) = cx + d,$$

med d forskellig fra 0, er altså ikke lineære. Vi har da også f.eks. $F(2) = 2c + d$, der ikke er lig $(c + d) + (c + d) = F(1) + F(1)$.

Eksempel 2

Også i $\mathbb{R}^n$ er **proportionaliteterne**, dvs. afbildninger af $\mathbb{R}^n$ ind i $\mathbb{R}^n$ givet ved

$$F(\underline{x}) = c\underline{x},$$

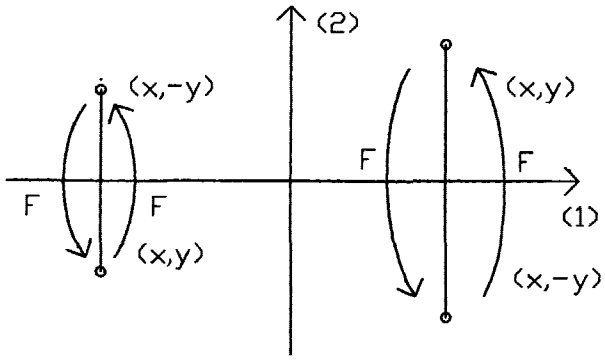


interessante lineære afbildninger (kontrollér selv at linearitetskravene er opfyldt).

Der er imidlertid mange andre lineære afbildninger. F.eks. er i $\mathbb{R}^2$, afbildningen F defineret ved (spejling i x-aksen)

$$F(x, y) = (x, -y), \text{for } (x, y) \in \mathbb{R}^2$$

lineær.



Thi for vilkårlige (u,v), (x,y) i $\mathbb{R}^2$ og vilkårlige α og β i $\mathbb{R}$, har vi

$$\begin{aligned} F(\alpha(u,v)+\beta(x,y)) &= F(\alpha u+\beta x, \alpha v+\beta y) \\ &= (\alpha u+\beta x, -\alpha v-\beta y) \\ &= (\alpha u, -\alpha v)+(\beta x, -\beta y) \\ &= \alpha(u,-v)+\beta(x,-y) \\ &= \alpha F(u,v)+\beta F(x,y), \end{aligned}$$

hvilket viser at F respekterer linearkombinering. $*$

Eksempel 3

Et andet vigtigt eksempel på lineære afbildninger, denne gang af $\mathbb{R}^n$ ind i $\mathbb{R}$, har vi i F'er af typen

$$F(\underline{x}) = F(x_1, x_2, \ldots, x_n) = a_1x_1 + a_2x_2 + \ldots + a_nx_n, \text{for } \underline{x} \in \mathbb{R}^n$$

hvor $a_1, a_2, \ldots, a_n$ er et vilkårligt, men fast sæt af reelle tal. At F virkelig er en lineær afbildning ses af, at

$$\begin{aligned} F(\underline{x}+\underline{y}) &= a_1(x_1+y_1)+a_2(x_2+y_2)+\ldots+a_n(x_n+y_n) \\ &= a_1x_1+a_2x_2+\ldots+a_nx_n+a_1y_1+a_2y_2+\ldots+a_ny_n \\ &= F(\underline{x})+F(\underline{y}) \end{aligned}$$

og

$$\begin{aligned} F(\alpha\underline{x}) &= a_1(\alpha x_1)+\alpha_2(\alpha x_2)+\ldots+a_n(\alpha x_n) \\ &= \alpha(a_1x_1+a_2x_2+\ldots+a_nx_n) \\ &= \alpha F(\underline{x}).* \end{aligned}$$

Også dette afsnit afsluttes med et par almene observationer:

$*$ *For enhver lineær afbildning* $F:\mathbb{R}^n \to \mathbb{R}^p$ *gælder, at* $F(\underline{0}) = \underline{0}$ (hvor vi tillader os at benytte betegnelsen $\underline{0}$ for nulvektoren både i $\mathbb{R}^n$ og i $\mathbb{R}^p$). Det følger f.eks. af at F respekterer skalering: $F(\underline{0}) = F(0\cdot\underline{0}) = 0F(\underline{0}) = \underline{0}$. (Vis, som en øvelse, at det også følger af at F er additiv.)

$*$ Afbildningen $F:\mathbb{R}^n \to \mathbb{R}^p$, som er *konstant* $\underline{0}$, dvs. er defineret ved, at

$$F(\underline{x}) = \underline{0}, \text{ for alle } \underline{x} \in \mathbb{R}^n,$$

er tydeligvis *en lineær afbildning*. Ingen andre konstante funktioner er lineære. (Check selv begge dele.)

$*$ *OBS! Vigtigt!* Hvis $\underline{x} \in \mathbb{R}^n$ skrives som $n \times 1$- søjle og A er en $p \times n$ matrix, bestemmes ved

$$F(\underline{x}) = A\underline{x},$$

hvor $F(\underline{x})$ skrives som en $p \times 1$-søjle, en afbildning $F : \mathbb{R}^n \to \mathbb{R}^p$. *Enhver sådan afbildning er lineær*, thi vi har i indledningen set, at

$$A(\underline{x} + \underline{x}') = A\underline{x} + A\underline{x}',$$

dvs.

$$F(\underline{x} + \underline{y}) = F(\underline{x}) + F(\underline{y}),$$

og

$$A(\alpha\underline{x}) = \alpha A\underline{x},$$

dvs.

$$F(\alpha\underline{x}) = \alpha F(\underline{x}),$$

for alle $\underline{x}$, $\underline{x}'$ i $\mathbb{R}^n$, og alle $\alpha \in \mathbb{R}$.

Vi skal i det næste afsnit indse, at enhver lineær afbildning mellem to talrum har en sådan matrixfremstilling, når der er valgt baser i de to talrum.

Det næste punkt ophøjes til en sætning:

SÆTNING. *En lineær afbildning afbilder et underrum i et underrum.*

BEVIS: Hvis den lineære afbildning hedder $F : \mathbb{R}^n \to \mathbb{R}^p$, og $\mathbf{U}$ er et underrum i $\mathbb{R}^n$, er påstanden i sætningen altså, at billedmængden $F(\mathbf{U})$ (som per definition er den mængde i $\mathbb{R}^p$ som er alle billeder af vektorer i $\mathbf{U}$) er et underrum i $\mathbb{R}^p$. Det kan vi bevise ved at godtgøre, at $F(\mathbf{U})$ opfylder underrumskravene (1) og (2).

Først (1): Hvis $\underline{y}$ og $\underline{y}'$ ligger i $F(\mathbf{U})$, har de formen

$$\underline{y} = F(\underline{x}),$$

og

$$\underline{y}' = F(\underline{x}'),$$

hvor $\underline{x}, \underline{x}' \in \mathbf{U}$. Da $\mathbf{U}$ er et underrum vil $\underline{x} + \underline{x}'$ tilhøre $\mathbf{U}$. Da der desuden - på grund af F's linearitet - gælder, at

$$\underline{y} + \underline{y}' = F(\underline{x}) + F(\underline{x}') = F(\underline{x} + \underline{x}'),$$

ser vi at $\underline{y} + \underline{y}'$ er billedet ved F af en vektor - $\underline{x} + \underline{x}'$ - fra $\mathbf{U}$. Det betyder, at $\underline{y} + \underline{y}'$ befinder sig i $F(\mathbf{U})$.

Dernæst (2): Hvis $\underline{y} \in F(\mathbf{U})$, og $\alpha \in \mathbb{R}$, gælder at

$$\underline{y} = F(\underline{x})$$

og

$$\alpha\underline{y} = F(\alpha\underline{x}),$$

for et $\underline{x} \in \mathbf{U}$. Da $\alpha\underline{x} \in \mathbf{U}$, fordi $\mathbf{U}$ er et underrum, er $\alpha\underline{y} = F(\alpha\underline{x})$ billedet ved F af en vektor fra $\mathbf{U}$. Det viser, at $\alpha\underline{y}$ befinder sig i $F(\mathbf{U})$.

I alt opfylder $F(\mathbf{U})$ begge underrumsbetingelserne. *Q.E.D.*

$*$ Eftersom $\mathbb{R}^n$ er et specielt underrum i sig selv, følger af denne sætning at $F(\mathbb{R}^n)$ er et underrum i $\mathbb{R}^p$. Dette underrum - der ingenlunde behøver at udfylde hele $\mathbb{R}^p$ - kaldes **billedrummet** ved F.

$*$ Mængden af vektorer i $\mathbb{R}^n$, der ved en lineær afbildning

$$F : \mathbb{R}^n \to \mathbb{R}^p$$

afbildes i $\underline{0}$ i $\mathbb{R}^p$, kaldes **nulrummet** for F. Det betegnes ofte $N(F)$:

$$N(F) = \{\underline{x} \in \mathbb{R}^n | F(\underline{x}) = \underline{0}\}.$$

Nulrummet er et underrum i $\mathbb{R}^n$. Thi lad $\underline{x}$ og $\underline{x}'$ begge tilhøre $N(F)$, dvs. $F(\underline{x}) = \underline{0}$ og $F(\underline{x}') = \underline{0}$, og lad α være en vilkårlig skalar. Så vil $F(\underline{x} + \underline{x}') = F(\underline{x}) + F(\underline{x}') = \underline{0}$, dvs. $\underline{x} + \underline{x}' \in N(F)$, og $F(\alpha\underline{x}) = \alpha F(\underline{x}) = \underline{0}$, dvs. $\alpha\underline{x} \in N(F)$. Dette viser, at $N(F)$ opfylder de to underrumsbetingelser.

Matrixfremstillinger af lineære afbildninger

Hvilke slags lineære afbildninger findes der mellem to talrum? I dette afsnit vil dette spørgsmål blive behandlet.

Lad os forestille os de involverede talrum $\mathbb{R}^n$ og $\mathbb{R}^p$ forsynet med baser, henholdsvis $\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_n$ for $\mathbb{R}^n$ og $\underline{b}_1, \underline{b}_2, \ldots, \underline{b}_p$ for $\mathbb{R}^p$. Disse baser kan være vilkårlige, og behøver således ikke at bestå af grundvektorerne i de to rum; selv om denne mulighed også er dækket af de følgende betragtninger. Vi interesserer os for en vilkårlig lineær afbildning F fra $\mathbb{R}^n$ til $\mathbb{R}^p$.

For at få et overblik over hvordan F virker, fremstiller vi alle vektorer i $\mathbb{R}^n$, respektive $\mathbb{R}^p$, ved hjælp af deres koordinater i de valgte baser:

$$\underline{x} = \alpha_1 \underline{a}_1 + \alpha_2 \underline{a}_2 + \ldots + \alpha_n \underline{a}_n$$

og

$$\underline{y} = \beta_1 \underline{b}_1 + \beta_2 \underline{b}_2 + \ldots + \beta_p \underline{b}_p.$$

Da F er lineær, gælder så at

$$\begin{aligned} F(\underline{x}) &= F(\alpha_1 \underline{a}_1 + \alpha_2 \underline{a}_2 + \ldots + \alpha_n \underline{a}_n) \\ &= \alpha_1 F(\underline{a}_1) + \alpha_2 F(\underline{a}_2) + \ldots + \alpha_n F(\underline{a}_n). \end{aligned}$$

Nu har alle vektorerne $F(\underline{a}_1), F(\underline{a}_2), \ldots, F(\underline{a}_n)$ hver deres koordinatfremstilling i $\underline{b}$-basen:

$$\begin{aligned} F(\underline{a}_1) &= c_{11}\underline{b}_1 + c_{21}\underline{b}_2 + \ldots + c_{p1}\underline{b}_p \\ F(\underline{a}_2) &= c_{12}\underline{b}_1 + c_{22}\underline{b}_2 + \ldots + c_{p2}\underline{b}_p \\ &\vdots \\ F(\underline{a}_n) &= c_{1n}\underline{b}_1 + c_{2n}\underline{b}_2 + \ldots + c_{pn}\underline{b}_p, \end{aligned}$$

hvor i c_{ij} det andet index refererer til $\underline{a}$-nummeret j, mens det første refererer til den $\underline{b}$-vektor c_{ij} er koefficient til. Indsættes dette i fremstillingen for $F(\underline{x})$ ovenfor, finder vi

$$\begin{aligned} F(\underline{x}) = &\alpha_1(c_{11}\underline{b}_1 + c_{21}\underline{b}_2 + \ldots + c_{p1}\underline{b}_p) + \\ &\alpha_2(c_{12}\underline{b}_1 + c_{22}\underline{b}_2 + \ldots + c_{p2}\underline{b}_p) + \\ &\vdots \\ &\alpha_n(c_{1n}\underline{b}_1 + c_{2n}\underline{b}_2 + \ldots + c_{pn}\underline{b}_p), \end{aligned}$$

der ved ordning efter $\underline{b}$'erne bliver til

$$\begin{aligned} &(c_{11}\alpha_1 + c_{12}\alpha_2 + \ldots + c_{1n}\alpha_n)\underline{b}_1 + \\ &(c_{21}\alpha_1 + c_{22}\alpha_2 + \ldots + c_{2n}\alpha_n)\underline{b}_2 + \\ &\vdots \\ &(c_{p1}\alpha_1 + c_{p2}\alpha_2 + \ldots + c_{pn}\alpha_n)\underline{b}_p. \end{aligned}$$

Heraf aflæses uden videre, at $F(\underline{x})$'s koordinater i $\mathbb{R}^p$'s basis er givet ved

$$\begin{aligned} \beta_1 &= c_{11}\alpha_1 + c_{12}\alpha_2 + \ldots + c_{1n}\alpha_n \\ \beta_2 &= c_{21}\alpha_1 + c_{22}\alpha_2 + \ldots + c_{2n}\alpha_n \\ &\vdots \\ \beta_p &= c_{p1}\alpha_1 + c_{p2}\alpha_2 + \ldots + c_{pn}\alpha_n. \end{aligned}$$

Dette kan også skrives på matrixform:

$$\begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ c_{21} & c_{22} & \dots & c_{2n} \\ \vdots & \vdots & & \vdots \\ c_{p1} & c_{p2} & \dots & c_{pn} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix},$$

eller kort

$$\underline{\beta} = A\underline{\alpha},$$

hvor α-søjlen er $\underline{x}$'s koordinater i $\underline{a}$-basen i $\mathbb{R}^n$, og β-søjlen er $F(\underline{x})$'s koordinater i $\underline{b}$-basen i $\mathbb{R}^p$, og hvor A er en $p \times n$ matrix der er karakteriseret ved at *dens j'te søjle udgøres af koordinaterne til billedet af $\underline{a}_j$, $F(\underline{a}_j)$, i $\underline{b}$-basen.*

Dette viser, at *enhver lineær afbildning $F : \mathbb{R}^n \to \mathbb{R}^p$ kan beskrives ved hjælp af en matrixfremstilling $\underline{\beta} = A\underline{\alpha}$, hvor $\underline{\alpha}$ og $\underline{\beta}$ betegner koordinaterne for $\underline{x}$, henholdsvis $F(\underline{x})$, i forhold til vilkårligt valgte baser i henholdsvis $\mathbb{R}^n$ og $\mathbb{R}^p$.* Fremstillingen afhænger naturligvis af de valgte baser. Læg specielt mærke til, at det ikke er den samme matrix A der optræder i enhver situation. Vælges andre baser i $\mathbb{R}^n$ eller $\mathbb{R}^p$ fås et andet A.

Resultatet gælder specielt, hvis vi fra begyndelsen har valgt *grundvektorerne som baser i begge rum.* I så fald er koordinatsættet for $\underline{x}$ lig med $\underline{x}$ selv, og tilsvarende er koordinatersættet for $F(\underline{x})$ lig med $F(\underline{x})$ selv. I denne situation forbinder matrixfremstillingen altså direkte vektoren $\underline{x}$ selv med dens billede $F(\underline{x})$:

$$F(\underline{x}) = A\underline{x}.$$

Sammenholdes dette resultat med en tidligere bemærkning om at *enhver* afbildning defineret ved en forskrift af typen $(*)$ er lineær, har vi indset at *de lineære afbildninger mellem to talrum netop er dem som er bestemt ved forskriften $F(\underline{x}) = A\underline{x}$ når grundvektorerne i de to rum er valgt som baser.*

LINEÆRE AFBILDNINGER OG LINEÆRE LIGNINGSSYSTEMER

Ved hjælp af matrixfremstillingen i det foregående afsnit kan vi beskrive F's billedrum og nulrum på en simpel måde.

At $\underline{y} \in \mathbb{R}^p$ tilhører F's billedrum $F(\mathbb{R}^n)$ betyder, at der findes en vektor $\underline{x} \in \mathbb{R}^n$, så at $\underline{y} = A\underline{x}$, eller med andre ord:

Vektoren $\underline{y} \in \mathbb{R}^n$ tilhører billedrummet for F hvis og kun hvis ligningssystemet $\underline{y} = A\underline{x}$ har en løsning $\underline{x} \in \mathbb{R}^n$.

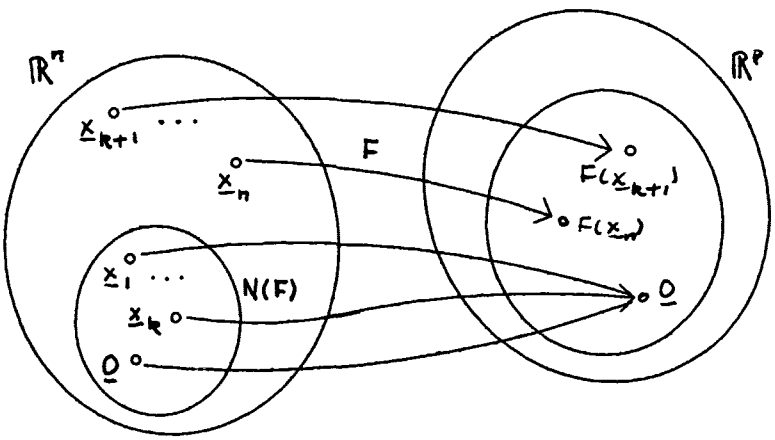
Ligeledes har vi, at $\underline{x} \in \mathbb{R}^n$ *tilhører nulrummet for F*, hvis og kun hvis $A\underline{x} = \underline{0}$, dvs. *hvis og kun hvis $\underline{x}$ er en løsning til ligningssystemet $A\underline{x} = \underline{0}$.*

Om sammenhængen mellem dimensionerne af nulrummet og billedrummet for en lineær afbildning $F : \mathbb{R}^n \to \mathbb{R}^p$ gælder følgende vigtige resultat. Før vi kan formulere det, minder vi om, at *hvis et underrum* U *af et talrum udelukkende består af* $\underline{0}$, *tillægger vi det dimensionen 0.*

Dimensionssætningen

SÆTNING (DIMENSIONSSÆTNINGEN). $dimN(F) + dimF(\mathbb{R}^n) = dim\mathbb{R}^n = n.$

Inden beviset bringes ser vi på hvad sætningen egentlig udsiger. Nulrummet og billedrummet for F skal altså dele dimensionen n (som er et fast tal) mellem sig. Sætningen udtrykker således i skarp form den lidt løse intuition, at jo større nulrummet for F er, dvs. jo flere vektorer i $\mathbb{R}^n$ der ved F sendes over i $\underline{0}$ i $\mathbb{R}^p$, jo "færre forskellige værdier" kan F antage, dvs. jo mindre fylder billedrummet $F(\mathbb{R}^n)$ i $\mathbb{R}^p$. Hvis nulrummet alene består af nulvektoren, har billedrummet samme dimension som definitionsrummet for F. I alle andre tilfælde er billedrummets dimension mindre end definitionsrummets. En lineær afbildning kan altså aldrig "blæse sit definitionsrum op".



BEVIS: (Konsulter figuren ovenfor)

(1) Først betragtes den situation, hvor $N(F)$ ikke kun består af $\underline{0}$. Lad $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k$ være en basis for $N(F)$.

I den ekstreme situation hvor $k = n$, udgør $N(F)$ hele $\mathbb{R}^n$, og så vil $F(\mathbb{R}^n)$ alene bestå af $\underline{0}$ og dermed have dimensionen 0. I så fald passer påstanden i sætningen åbenbart.

Hvis $k < n$, må der findes en vektor $\underline{x}_{k+1}$ i $\mathbb{R}^n$, der sammen med $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k$ danner et lineært uafhængigt sæt. I modsat fald ville jo enhver vektor i $\mathbb{R}^n$ være en linearkombination af $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k$ i strid med at dimensionen af $\mathbb{R}^n$ er $n > k$. Hvis $k + 1 = n$ udgør det lineært uafhængige sæt $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k, \underline{x}_{k+1}$ en basis for $\mathbb{R}^n$. Ellers er $k + 1 < n$, og der må på samme måde som før findes en vektor $\underline{x}_{k+2}$, der sammen med $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k, \underline{x}_{k+1}$ danner et lineært uafhængigt sæt. Således kan vi om fornødent fortsætte indtil vi har fundet supplerende vektorer $\underline{x}_{k+1}, \underline{x}_{k+2}, \ldots, \underline{x}_n$, så at det samlede sæt $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k, \underline{x}_{k+1}, \underline{x}_{k+2}, \ldots, \underline{x}_n$ er lineært uafhængigt, og dermed - da det består af n vektorer - en basis for $\mathbb{R}^n$.

Nu påstår jeg, at billederne af de tilføjede vektorer:

$$F(\underline{x}_{k+1}), F(\underline{x}_{k+2}), \ldots, F(\underline{x}_n)$$

udgør en basis for billedrummet $F(\mathbb{R}^n)$. Hvis det er rigtigt er sætningen bevist i de betragtede tilfælde. Der findes jo så $n-k = dim(\mathbb{R}^n) - dim N(F)$ basisvektorer for $F(\mathbb{R}^n)$.

Først indser vi, at $F(\underline{x}_{k+1}), F(\underline{x}_{k+2}), \ldots, F(\underline{x}_n)$ er lineært uafhængige i $\mathbb{R}^p$. Lad til den ende

$$(*) \qquad \alpha_{k+1} F(\underline{x}_{k+1}) + \alpha_{k+2} F(\underline{x}_{k+2}) + \ldots + \alpha_n F(\underline{x}_n) = \underline{0}$$

være en fremstilling af nulvektoren i $\mathbb{R}^p$. På grund af F's linearitet, gælder så, at

$$F(\alpha_{k+1}\underline{x}_{k+1} + \alpha_{k+2}\underline{x}_{k+2} + \ldots + \alpha_n \underline{x}_n) = \underline{0},$$

hvilket viser, at vektoren $\alpha_{k+1}\underline{x}_{k+1} + \alpha_{k+2}\underline{x}_{k+2} + \ldots + \alpha_n \underline{x}_n$ ligger i nulrummet for F. Den må derfor være en linearkombination af $N(F)$- basen $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k$. Nu må alle α'erne være 0. Hvis ikke, var ét af dem forskelligt fra 0, f.eks. α_{k+j}, sådan at $\underline{x}_{k+j}$ ville være en linearkombination af de øvrige $\underline{x}$'er (altså $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_{k+j-1}, \underline{x}_{k+j+1}, \ldots, \underline{x}_n$), i strid med at $\underline{x}$'erne er lineært uafhængige. Men når alle α'erne er 0 er fremstillingen $(*)$ triviel. Da dette gælder enhver sådan fremstilling er $F(\underline{x}_{k+1}), F(\underline{x}_{k+2}), \ldots, F(\underline{x}_n)$ lineært uafhængige, som påstået.

Dernæst skal det godtgøres, at $F(\underline{x}_{k+1}), F(\underline{x}_{k+2}), \ldots, F(\underline{x}_n)$ udspænder billedrummet for F. Det følger af at hvis $\underline{y} \in F(\mathbb{R}^n)$, eksisterer et $\underline{x} \in \mathbb{R}^n$, så at $\underline{y} = F(\underline{x})$. Da $\underline{x}$'erne udgør en basis for $\mathbb{R}^n$ har $\underline{x}$ en fremstilling

$$\underline{x} = \alpha_1\underline{x}_1 + \alpha_2\underline{x}_2 + \ldots + \alpha_k\underline{x}_k + \alpha_{k+1}\underline{x}_{k+1} + \alpha_{k+2}\underline{x}_{k+2} + \ldots + \alpha_n\underline{x}_n$$

med passende α'er som koordinater. Da F er lineær gælder, at

$$\begin{aligned}
\underline{y} &= F(\underline{x}) \\
&= F(\alpha_1\underline{x}_1 + \alpha_2\underline{x}_2 + \ldots + \alpha_k\underline{x}_k + \alpha_{k+1}\underline{x}_{k+1} + \alpha_{k+2}\underline{x}_{k+2} + \ldots + \alpha_n\underline{x}_n) \\
&= \alpha_1 F(\underline{x}_1) + \alpha_2 F(\underline{x}_2) + \ldots + \alpha_k F(\underline{x}_k) \\
&+ \alpha_{k+1}F(\underline{x}_{k+1}) + \alpha_{k+2}F(\underline{x}_{k+2}) + \ldots + \alpha_n F(\underline{x}_n).
\end{aligned}$$

Men her ligger $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k$ i $N(F)$, så at F afbilder dem alle i $\underline{0}$. Derved bliver

$$\underline{y} = \alpha_{k+1}F(\underline{x}_{k+1}) + \alpha_{k+2}F(\underline{x}_{k+2}) + \ldots + \alpha_n F(\underline{x}_n).$$

Men så er $\underline{y}$ jo også frembragt af $F(\underline{x}_{k+1}), F(\underline{x}_{k+2}), \ldots, F(\underline{x}_n)$, som påstået. Altså udspænder disse vektorer billedrummet for F.

(2) Vi mangler blot at behandle den situation hvor $N(F)$ kun omfatter $\underline{0}$. I den situation skal vil blot indse at $dim F(\mathbb{R}^n) = n$. Men her kan vi blot vælge en basis $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_n$ for $\mathbb{R}^n$. Dens billeder vil med nøjagtigt de samme argumenter som netop er præsenteret udspænde $F(\mathbb{R}^n)$.

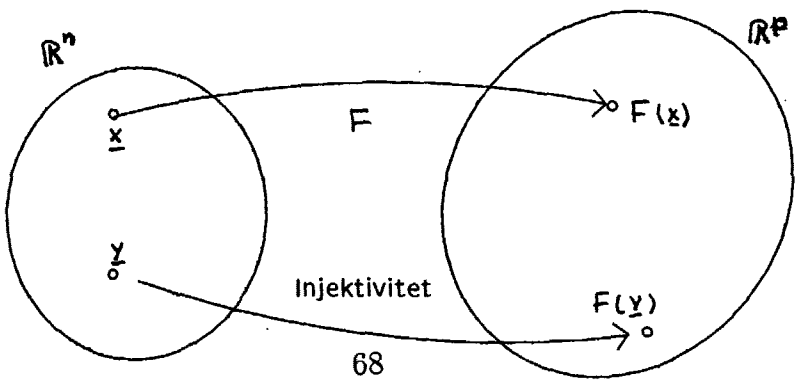
Hermed er sætningen bevist. *Q.E.D.*

Kender vi dimensionen af nulrummet for en lineær afbildning F, fortæller dimensionssætningen noget om hvor meget F "klapper $\mathbb{R}^n$ sammen" under afbildningen.

Hvis $N(F) = \{\underline{0}\}$ klapper F overhovedet ikke $\mathbb{R}^n$ sammen. Det leder os til en vigtig definition.

En lineær afbildning kaldes **injektiv**, hvis to forskellige vektorer i $\mathbb{R}^n$ altid har forskellige billeder i $\mathbb{R}^p$, eller for at være præcis, hvis og kun hvis F opfylder:

(I) For alle $\underline{x}, \underline{y} \in \mathbb{R}^n$ vil $F(\underline{x}) = F(\underline{y})$ medføre, at $\underline{x} = \underline{y}$.



Det viser sig ikke at være nødvendigt at undersøge kravet i (I) for alle par $\underline{x}$ og $\underline{y}$. Man kan nøjes med at undersøge den tilsyneladende svagere betingelse om andre vektorer end $\underline{0}$ ved F afbildes i $\underline{0}$. Der gælder nemlig

$*$ *F er injektiv hvis og kun hvis F opfylder:*

(I$*$) *For alle* $\underline{x} \in \mathbb{R}^n$ *vil* $F(\underline{x}) = \underline{0}$ *medføre, at* $\underline{x} = \underline{0}$.

At denne påstand er sand, ser vi således: For det første er det klart, at (I) medfører (I$*$). (I$*$) er nemlig et specialtilfælde af (I) med parret $\underline{x} = \underline{x}$ og $\underline{y} = \underline{0}$ (hvor vi selvsagt benytter, at $F(\underline{0}) = \underline{0}$). Men (I$*$) medfører på sin side (I). For hvis $\underline{x}$ og $\underline{y}$ er vilkårlige vektorer i $\mathbb{R}^n$, sådan at $F(\underline{x}) = F(\underline{y})$, er $F(\underline{x} - \underline{y}) = \underline{0}$, efterdi F er lineær. Så aktiverer vi (I$*$) på vektoren $\underline{x} - \underline{y}$, og slutter deraf at $\underline{x} - \underline{y} = \underline{0}$, eller anderledes sagt, at $\underline{x} = \underline{y}$. Men det var et bevis for at (I) gælder.

Da betingelsen (I$*$) kommer ud på, at nulrummet for F alene består af $\underline{0}$, kan vi levere en ækvivalent formulering af det netop fundne:

F er injektiv netop hvis nulrummet for F, N(F), kun består af $\underline{0}$.

Ved at sammenholde dette med dimensionssætningen, kan vi skaffe endnu en ækvivalent formulering:

F er injektiv netop hvis dimensionen af F's billedrum $F(\mathbb{R}^n)$ *er lig n, dimensionen af definitionsrummet for F.*

Og dermed er vi tilbage ved udgangspunktet: F er injektiv netop hvis F ikke klapper definitionsrummet sammen under afbildningen.

En sidste ækvivalent formulering kan vi få ved at inddrage en matrixfremstilling for $F(\underline{x}) = A\underline{x}$:

F er injektiv hvis og kun hvis ligningssystemet $A\underline{x} = \underline{0}$ *kun har den trivielle løsning* $\underline{x} = \underline{0}$.

En enkelt vigtig observation:

$*$ *Hvis* $F : \mathbb{R}^n \to \mathbb{R}^p$ *er injektiv, afbilder F enhver basis for* $\mathbb{R}^n$ *på en basis for* $F(\mathbb{R}^n)$. Lad nemlig $\underline{b}_1, \underline{b}_2, \ldots, \underline{b}_n$ være en basis for $\mathbb{R}^n$. Så er $F(\underline{b}_1), F(\underline{b}_2), \ldots, F(\underline{b}_n)$ en basis for $\mathbf{U} = F(\mathbb{R}^n)$, påstår jeg.

For det første udspænder $F(\underline{b}_1), F(\underline{b}_2), \ldots, F(\underline{b}_n)$ $\mathbf{U}$. Thi hvis $\underline{y}$ tilhører $\mathbf{U}$, eksisterer per definition af $\mathbf{U}$ et $\underline{x} \in \mathbb{R}^n$ så at $F(\underline{x}) = \underline{y}$. Men når $\underline{b}$'erne er en basis for $\mathbb{R}^n$, må

$\underline{x} = \alpha_1\underline{b}_1 + \alpha_2\underline{b}_2 + \ldots + \alpha_n\underline{b}_n$ med passende koefficienter $\alpha_1, \alpha_2, \ldots, \alpha_n$. Så vil på grund af F's linearitet

$$F(\underline{x}) = \alpha_1 F(\underline{b}_1) + \alpha_2 F(\underline{b}_2) + \ldots + \alpha_n F(\underline{b}_n),$$

hvilket demonstrerer at $\underline{y} = F(\underline{x})$ er udspændt af $F(\underline{b}_1), F(\underline{b}_2), \ldots, F(\underline{b}_n)$. Læg mærke til at denne del af ræsonnementet ikke benytter, at F er injektiv!

Dernæst er $F(\underline{b}_1), F(\underline{b}_2), \ldots, F(\underline{b}_n)$ lineært uafhængige. For hvis

$$\alpha_1 F(\underline{b}_1) + \alpha_2 F(\underline{b}_2) + \ldots + \alpha_n F(\underline{b}_n) = \underline{0},$$

er også

$$F(\alpha_1\underline{b}_1 + \alpha_2\underline{b}_2 + \ldots + \alpha_n\underline{b}_n) = \underline{0},$$

hvoraf, da F er injektiv (der var den!),

$$\alpha_1\underline{b}_1 + \alpha_2\underline{b}_2 + \ldots + \alpha_n\underline{b}_n = \underline{0}.$$

Men da $\underline{b}$'erne udgør en basis er de lineært uafhængige, hvorfor samtlige α'er er 0. Men så er vi færdige (med beviset, altså).

Dimensionssætningen fortæller noget om hvor stor en del F's billedrum faktisk udgør af $\mathbb{R}^p$ - det rum F afbilder over i. F.eks. fortæller den os at billedrummets dimension højst er n. Hvis derfor p er større end n kan billedrummet ikke udfylde hele $\mathbb{R}^p$. Men det er ikke sådan, at bare p <n er $F(\mathbb{R}^n) = \mathbb{R}^p$, f.eks. kan F's billedmængde bestå af $\underline{0}$ alene. Den situation, hvor F's billedrum faktisk udfylder hele $\mathbb{R}^p$ interesserer os så meget, at vi giver den et særligt navn:

En lineær afbildning $F : \mathbb{R}^n \to \mathbb{R}^p$ kaldes **surjektiv**, hvis F's billedrum udfylder hele $\mathbb{R}^p$, dvs. hvis og kun hvis $F(\mathbb{R}^n) = \mathbb{R}^p$. Eller anderledes sagt, hvis og kun hvis der for ethvert $\underline{y} \in \mathbb{R}^p$ findes et $\underline{x} \in \mathbb{R}^n$, så at $F(\underline{x}) = \underline{y}$. Vi benytter den malende sprogbrug, at F afbilder $\mathbb{R}^n$ **på** $\mathbb{R}^p$.

I matrixformulering (stadig med $F(\underline{x}) = A\underline{x}$) udtrykkes surjektivitet således:

$*$ *F er surjektiv hvis og kun hvis for ethvert $\underline{y} \in \mathbb{R}^p$ ligningssystemet $A\underline{x} = \underline{y}$ har en løsning $\underline{x} \in \mathbb{R}^n$.*

$*$ Af dimensionssætningen slutter vi, at *hvis F er surjektiv, er* $\dim N(F) = n - p$. Hvis derfor specielt $n = p$, er $\dim N(F) = 0$, således at F også er injektiv. Dvs. *en surjektiv afbildning af et talrum* **på sig selv** *er også injektiv*. Dimensionssætningen viser også at, hvis $F : \mathbb{R}^n \to \mathbb{R}^n$ (altså $n = p$) *er injektiv, er den også surjektiv*. Thi $dimF(\mathbb{R}^n) = n - dimN(F) = n - 0 = n$. Sådan forholder det sig ikke i almindelighed. Injektivitet og surjektivitet er uafhængige egenskaber, sådan at forstå, at en afbildning kan være surjektiv uden at være injektiv, og injektiv uden at være surjektiv. Se selv:

Eksempel 1

Den lineære afbildning $F : \mathbb{R}^2 \to \mathbb{R}^4$ givet ved

$$F(x,y) = (x+y,\, x-y,\, 2x+3y,\, \frac{1}{2}x - 4y)$$

(hvorfor er den forresten lineær?), er injektiv. Thi

$$F(x,y) = \underline{0}$$

netop hvis $x+y=0$, $x-y=0$, $2x+3y=0$, $\frac{1}{2}x-4y=0$, hvilket er ensbetydende med at $x=0$ og $y=0$, sådan at nulrummet for F alene består af $\underline{0}$. Derimod er F ikke surjektiv, eftersom dimensionen af billedrummet er lig 2 (dimensionssætningen slår til igen), mens $\mathbb{R}^4$ er 4-dimensional.

Den lineære afbildning $G : \mathbb{R}^3 \to \mathbb{R}^2$ defineret ved

$$G(x,y,z) = (x,y)$$

er øjensynlig surjektiv, da enhver vektor i $\mathbb{R}^2$, (x,y), er billede ved G af en vektor i $\mathbb{R}^3$, ja faktisk af uendeligt mange, nemlig af alle de vektorer der har formen $(x,y,$hvad som helst$)$. Det er i samme åndedrag klart, at G ikke er injektiv, da f.eks. både $(x,y,0)$ og $(x,y,1)$ afbildes i (x,y). $*$

De lineære afbildninger der er *både injektive og surjektive* er af stor vigtighed, hvorfor de tildeles et særligt navn: De kaldes **bijektive**. Ofte kaldes en bijektiv afbildning også en **énentydig korrespondance**.

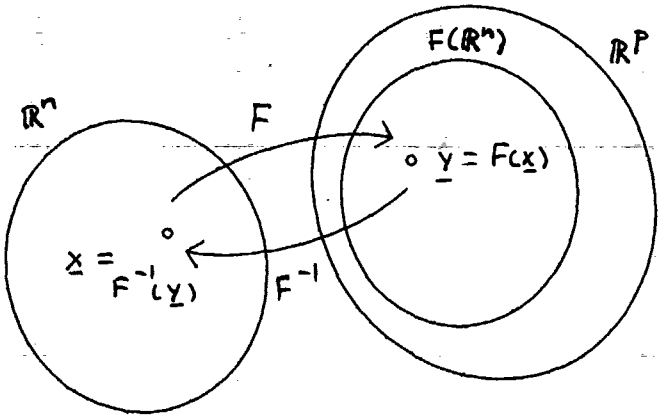
$*$ At en lineær afbildning $F : \mathbb{R}^n \to \mathbb{R}^p$ er *bijektiv* kan udtrykkes: *enhver vektor i* $\mathbb{R}^p$ *er billede ved* F *af netop én vektor i* $\mathbb{R}^n$. I kraft af vores undersøgelser kan dette kun finde sted, hvis $n = p$.

Øvelse: Opskriv en række ækvivalente formuleringer af bijektivitetsegenskaben ved at koble de forskellige formuleringer af injektivitet med de forskellige formuleringer af surjektivitet.

$*$ Selv om en injektiv afbildning ikke nødvendigvis er surjektiv på *hele* $\mathbb{R}^p$, vil den naturligvis være surjektiv på *billedrummet* $F(\mathbb{R}^n)$, idet jo enhver vektor i billedrummet selvfølgelig er billede af en vektor i $\mathbb{R}^n$. Den er altså en bijektiv afbildning fra $\mathbb{R}^n$ til $F(\mathbb{R}^n)$. Vi siger at $\mathbb{R}^n$ og $F(\mathbb{R}^n)$ er i *énentydig korrespondance*. Enhver vektor $\underline{y}$ i billedrummet er billede af netop én vektor $\underline{x}$ i $\mathbb{R}^n$. Denne vektor betegnes $F^{-1}(\underline{y})$. Derved fastlægges en afbildning

$$F^{-1} : F(\mathbb{R}^n) \to \mathbb{R}^n.$$

Afbildningen kaldes **den omvendte afbildning til** F. *Enhver injektiv afbildning har altså en omvendt, defineret på dens billedrum.*



Omvendt afbildning

Sidste om lineære ligningsystemer

Vi afrunder dette afsnit med analysere *løsningsforholdene for lineære ligningssystemer* ved hjælp af det apparat vi nu har opstillet.

Lad os betragte et vilkårligt lineært ligningssystem

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

hvor A er en $p \times n$-matrix, $\underline{x}$ en vektor i $\mathbb{R}^n$ opskrevet som søjle, og $\underline{b}$ en vektor i $\mathbb{R}^p$ opskrevet som søjle. Hvis $\underline{b} = \underline{0}$ kaldes ligningssystemet **homogent**. hvis $\underline{b} \neq \underline{0}$ kaldes det **inhomogent**. Mængden af løsningerne til (*) kaldes **løsningsrummet**.

Først betragter vi det *homogene system*, $A\underline{x} = \underline{0}$. Løsningsrummet for dette system udgør netop *nulrummet* for den lineære afbildning $F : \mathbb{R}^n \to \mathbb{R}^p$, hvis matrix er A (når grundvektorerne er valgt som baser i begge rum). Løsningsrummet for det homogene system er altså et *underrum* i $\mathbb{R}^n$. Et homogent ligningssystem har *altid løsninger*, om ikke andet så kun $\underline{0}$.

Spørgsmålet er nu: For hvilke andre $\underline{b} \in \mathbb{R}^p$ har ligningssystemet overhovedet løsninger? Svaret er, at *ligningssystemet har løsninger netop hvis* $\underline{b}$ *tilhører billedrummet for* F. Eftersom

$$A\underline{x} = x_1\underline{a}_1 + x_2\underline{a}_2 + \ldots + x_n\underline{a}_n$$

- hvor vi med $\underline{a}$'er har betegnet søjlerne $p \times 1$- søjlerne i A - er billedrummet for F *udspændt af søjlerne i* A. Billedrummets dimension er altså lig *antallet af lineært uafhængige søjler i* A.

Hvis billedrummet ikke udgør hele $\mathbb{R}^p$, har ligningssystemet ikke løsninger for ethvert $\underline{b}$. Det kan altså sagtens hænde, at der ingen løsninger er.

Man kan vise, men det vil vi afstå fra her, at *antallet af lineært uafhængige søjler i en vilkårlig matrix er det samme som antallet af lineært uafhængige rækker*. Det fælles tal kaldes **matricens rang**, og betegnes ofte med r.

$*$ Med andre ord er *billedrummets dimension lig med matricens rang r.*

Ved hjælp af dimensionssætningen slutter vi, at løsningsrummet til den homogene ligning (der jo er lig nulrummet for F) har en dimension der er lig

$$n - \dim F(\mathbb{R}^n) = n - r.$$

Dette resultat fortjener at blive opstillet som

SÆTNING. *Løsningsrummet L_h til et* **homogent** *ligningssystem $A\underline{x} = \underline{0}$, hvor A er en $p \times n$- matrix, er et underrum af $\mathbb{R}^n$. Dets dimension er lig $n - r$, hvor n altså er antallet af søjler i matricen, og r er matricens rang.*

Hvis et inhomogent ligningssystem $A\underline{x} = \underline{b}$, $\underline{b} \neq \underline{0}$, overhovedet har løsninger, kan løsningsrummet karakteriseres ved hjælp af løsningsrummet for **det tilsvarende homogene** ligningssystem, som det kaldes. Lad nemlig $\underline{x}_0$ være en enkelt løsning - vilkårlig, men fastholdt under betragtningerne - til det inhomogene system. En sådan løsning kaldes ofte en **partikulær løsning**. Så vil det om en eventuel anden løsning, $\underline{x}$, til det inhomogene system gælde, at

$$\underline{x} - \underline{x}_0 \text{ er en løsning til det tilsvarende homogene system,}$$

thi $A(\underline{x} - \underline{x}_0) = A\underline{x} - A\underline{x}_0 = \underline{b} - \underline{b} = \underline{0}$. Derved tilhører $\underline{x} - \underline{x}_0$ løsningsrummet L_h for det homogene system, eller anderledes sagt,

$$\underline{x} = \underline{x}_0 + \text{en løsning til det homogene system.}$$

Bruger vi skrivemåden $\underline{x}_0 + L_h$ for mængden

$$\{\underline{x} \in \mathbb{R}^n \mid \underline{x} = \underline{x}_0 + \underline{z}, \quad \text{hvor } \underline{z} \in L_h\},$$

og betegner L_{ih} løsningsrummet for det inhomogene system, har vi således

$$L_{ih} = \underline{x}_0 + L_h.$$

En sådan mængde, hvor en fast vektor uden for et givet underrum **U** er adderet til samtlige vektorer i **U**, kaldes et **sideunderrum** til **U**. Mængden er ikke selv et underrum, da den ikke indeholder $\underline{0}$. Selv om et sideunderrum ikke selv er et underrum, giver det ikke anledning til misforståelser at tillægge det samme *dimension* som det tilhørende underrum.

Også her er det på sin plads at opsamle resultaterne i en sætning:

SÆTNING. *Løsningsrummet for et* **inhomogent** *ligningssystem* $A\underline{x} = \underline{b}$, *hvor* A *er en* $p \times n$*- matrix og* $\underline{b} \in \mathbb{R}^p$, *er* **enten** *tomt, nemlig hvis* $\underline{b}$ *ikke tilhører underrummet i* $\mathbb{R}^p$ *udspændt af søjlerne i* A, **eller** *et sideunderrum til løsningsrummet for det tilsvarende homogene system. Dets dimensionen er* $n-r$, *hvor* n *altså er antallet af søjler i matricen og* r *er matricens rang.*

Med disse sætninger samt de øvrige resultater i hånden, kan vi gøre et par observationer:

* Hvis et *homogent* ligningssystem har **fuld rang**, dvs. samtlige søjler er lineært uafhængige så at $r = n$, har systemet *ikke andre løsninger end* $\underline{0}$, thi løsningsrummets dimension er så 0. Hvis et *inhomogent* ligningssystem har fuld rang, har det *enten ingen eller netop én løsning*.

* Hvis et *homogent* ligningssystem *ikke har fuld rang*, er $r < n$. Så er der *uendeligt mange* løsninger, idet løsningsrummet mindst har dimensionen 1. Hvis et *inhomogent* ligningssystem ikke har fuld rang, har det *enten ingen løsninger, eller uendeligt mange*, idet løsningsmængden er et sideunderrum af dimension mindst 1.

Denne situation foreligger, *hvis et ligningssystem har flere ubekendte end ligninger*. For hvis $p < n$, kan systemet ikke have fuld rang, eftersom antallet af lineært uafhængige søjler er lig antallet af lineært uafhængige rækker.

SAMMENSÆTNING AF LINEÆRE AFBILDNINGER

Hidtil har vi kun beskæftiget os med en enkelt lineær afbildning ad gangen. Nu vil vi undersøge hvordan *to lineære afbildninger* kan bringes i samspil. For at få en fornemmelse af hvad der er på færde, lægger vi ud med at betragte et par eksempler.

Eksempel 1

Vi betragter to lineære afbildninger, $F : \mathbb{R}^3 \to \mathbb{R}^4$ og $G : \mathbb{R}^4 \to \mathbb{R}^2$, defineret ved forskrifterne $F(\underline{x}) = A\underline{x}$, hvor A er en 4×3-matrix, og $G(\underline{y}) = B\underline{y}$, hvor B er en 2×4-matrix. Lad os se på en situation med konkrete matricer:

$$F(\underline{x}) = \begin{pmatrix} -1 & 2 & 0 \\ 2 & 1 & 3 \\ 2 & 0 & 1 \\ 0 & -3 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \qquad G(\underline{y}) = \begin{pmatrix} 0 & 1 & -1 & 2 \\ 2 & -3 & 3 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix}$$

Tager vi nu et $\underline{x} \in \mathbb{R}^3$, vil det ved F afbildes i en vektor $F(\underline{x})$ i $\mathbb{R}^4$, altså i det rum G virker på. Derfor kan vi danne billedet ved G af $F(\underline{x})$, det hedder $G(F(\underline{x}))$ og befinder sig i $\mathbb{R}^2$. Alt i alt har vi dannet en afbildning fra $\mathbb{R}^3$ til $\mathbb{R}^2$, defineret ved

$$\underline{x} \to G(F(\underline{x})).$$

Vi vil nu bestemme en forskrift for denne *sammensatte* afbildning af F og G, der også benævnes $G \circ F$.

Først har vi

$$F(x_1, x_2, x_3) = \begin{pmatrix} -1 & 2 & 0 \\ 2 & 1 & 3 \\ 2 & 0 & 1 \\ 0 & -3 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -x_1 + 2x_2 \\ 2x_1 + x_2 + 3x_3 \\ 2x_1 + x_3 \\ -3x_2 + 2x_3 \end{pmatrix}.$$

Sendes dette ind i forskriften for G, opstår

$$\begin{aligned} G(F(x_1, x_2, x_3)) &= \begin{pmatrix} 0 & 1 & -1 & 2 \\ 2 & -3 & 3 & 0 \end{pmatrix} \begin{pmatrix} -x_1 + 2x_2 \\ 2x_1 + x_2 + 3x_3 \\ 2x_1 + x_3 \\ -3x_2 + 2x_3 \end{pmatrix} \\ &= \begin{pmatrix} 2x_1 + x_2 + 3x_3 - (2x_1 + x_3) + 2(-3x_2 + 2x_3) \\ 2(-x_1 + 2x_2) - 3(2x_1 + x_2 + 3x_3) + 3(2x_1 + x_3) \end{pmatrix} \\ &= \begin{pmatrix} -5x_2 + 6x_3 \\ -2x_1 + x_2 - 6x_3 \end{pmatrix} = \begin{pmatrix} 0 & -5 & 6 \\ -2 & 1 & -6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}. \end{aligned}$$

Nettoudbyttet af denne operation er altså en forskrift for den sammensatte afbildning $G \circ F$ af F og G (læg mærke til at der i $G \circ F$ er byttet om på rækkefølgen, fordi F benyttes først - dvs. tættest på den variable $\underline{x}$ - mens G benyttes sidst):

$$G(F(x_1, x_2, x_3)) = \begin{pmatrix} 0 & -5 & 6 \\ -2 & 1 & -6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

Heraf ser vi bl.a., at $G \circ F$ er en lineær afbildning fra $\mathbb{R}^3$ til $\mathbb{R}^2$. Nu er det klart, at matricen for $G \circ F$ er skabt ved hjælp af operationer på elementerne i matricerne A og B, men hvordan det nærmere er foregået fortaber sig i de konkrete talregnerier. Nedenfor vil vi trænge ind i mønstret bag regnerierne. *

Men først endnu et eksempel.

Eksempel 2

I det foregående eksempel var F og G defineret direkte ved deres matrixfremstillinger i forhold til baser bestående af grundvektorerne i alle rummene $\mathbb{R}^2$, $\mathbb{R}^3$ og $\mathbb{R}^4$. I det her foreliggende eksempel er der tale om, at der i de involverede rum er valgt nogle andre baser end grundbaserne.

Afbildningen $F : \mathbb{R}^2 \to \mathbb{R}^3$ er givet i forhold til, at der i $\mathbb{R}^2$ er valgt basen $\underline{u}_1 = (1, 2)$, $\underline{u}_2 = (0, -1)$, mens der i $\mathbb{R}^3$ er valgt basen $\underline{v}_1 = (1, 1, 0)$, $\underline{v}_2 = (1, 0, 1)$, $\underline{v}_3 = (0, 0, -1)$. F.eks. har vektoren $\underline{x} = (x_1, x_2) = (-5, 9)$ koordinaterne $\underline{\xi} = (-5, -19)$ i $\underline{u}$- basen.

I disse baser er $F : \mathbb{R}^2 \to \mathbb{R}^3$ givet ved matricen

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

Det vil sige, at hvis $\underline{x}$'s koordinater i $\underline{u}$-basen for $\mathbb{R}^2$ er $\underline{\xi} = (\xi_1, \xi_2)$, bliver koordinaterne for $F(\underline{x})$ i $\underline{v}$-basen for $\mathbb{R}^3$ lig

$$\begin{pmatrix} 1 & 1 \\ 0 & -2 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}.$$

På tilsvarende måde antager vi $G : \mathbb{R}^3 \to \mathbb{R}^2$ givet ved matricen

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

Vi fandt, at $F(\underline{x})$'s koordinater i $\underline{v}$-basen var givet ved

$$\begin{pmatrix} 1 & 1 \\ 0 & -2 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = \begin{pmatrix} \xi_1 + \xi_2 \\ -2\xi_2 \\ -\xi_1 + 3\xi_2 \end{pmatrix}.$$

Så bliver koordinaterne for $G(F(\underline{x}))$ - ikke at forveksle med $G(F(\underline{x}))$ selv, da vi ikke arbejder i grundbasen - bestemt som

$$\begin{aligned} \begin{pmatrix} 2 & -4 & 0 \\ 3 & 1 & -2 \end{pmatrix} \begin{pmatrix} \xi_1 + \xi_2 \\ -2\xi_2 \\ -\xi_1 + 3\xi_2 \end{pmatrix} &= \begin{pmatrix} 2(\xi_1 + \xi_2) - 4(-2\xi_2) \\ 3(\xi_1 + \xi_2) + (-2\xi_2) - 2(-\xi_1 + 3\xi_2) \end{pmatrix} \\ &= \begin{pmatrix} 2\xi_1 + 10\xi_2 \\ 5\xi_1 - 5\xi_2 \end{pmatrix} \\ &= \begin{pmatrix} 2 & 10 \\ 5 & -5 \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}, \end{aligned}$$

hvilket viser at også denne gang er $G \circ F$ en lineær afbildning. $*$

Nu er vi rustet til at behandle den generelle situation.

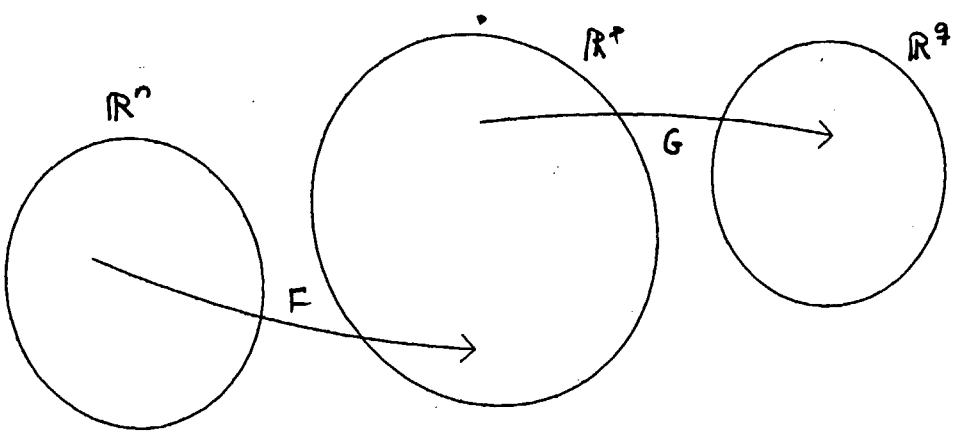
Vi forestiller os derfor givet to lineære afbildninger

$$F : \mathbb{R}^n \to \mathbb{R}^p$$

og

$$G : \mathbb{R}^p \to \mathbb{R}^q$$

jfr. vedstående figur.



Betragter vi et $\underline{x} \in \mathbb{R}^n$ vil det ved F afbildes i $F(\underline{x}) \in \mathbb{R}^p$, hvor G kan få fingre i det: $G(F(\underline{x})) \in \mathbb{R}^q$. Derved har vi i realiteten at gøre med en afbildning fra $\mathbb{R}^n$ ind i $\mathbb{R}^q$. Denne afbildning kaldes den **sammensatte** afbildning af F og G, og benævnes $G \circ F$. Den er altså defineret ved

$$(G \circ F)(\underline{x}) = G(F(\underline{x})), \quad \underline{x} \in \mathbb{R}^n$$

$$G \circ F : \mathbb{R}^n \to \mathbb{R}^q.$$

Der gælder nu, at også *den sammensatte afbildning $G \circ F$ er lineær*. Dette godtgøres ved at vi indser, at $G \circ F$ respekterer linearkombinering:

$$\begin{aligned}(G \circ F)(\alpha\underline{x} + \beta\underline{y}) &= G(F(\alpha\underline{x} + \beta\underline{y})) \\ &= G(\alpha F(\underline{x}) + \beta F(\underline{y})) \\ &= \alpha G(F(\underline{x})) + \beta G(F\underline{y})) \\ &= \alpha(G \circ F)(\underline{x}) + \beta(G \circ F)(\underline{y}),\end{aligned}$$

hvor lighedstegn nr. 2 skyldes at F er lineær, lighedstegn nr. 3 skyldes at G er lineær, mens det første og det sidste lighedstegn kommer af definitionen på $G \circ F$.

Eftersom alle lineære afbildninger har en matrixfremstilling i forhold til givne baser i de involverede rum, gælder det samme for $G \circ F$. Lad der derfor i $\mathbb{R}^n$ være valgt basen $\underline{u}_1, \underline{u}_2, \ldots, \underline{u}_n$, og i $\mathbb{R}^q$ - som $G \circ F$'s billeder jo ligger i - basen $\underline{w}_1, \underline{w}_2, \ldots, \underline{w}_q$. Der findes så en matrix $q \times n$-matrix C, så at

$$C\underline{\xi}$$

er koordinaterne for $(G \circ F)(\underline{x})$ i w-basen for $\mathbb{R}^q$, hvis $\underline{\xi} = (\xi_1, \xi_2, \ldots, \xi_n)$ er koordinaterne for $\underline{x}$ i u-basen for $\mathbb{R}^n$. **Obs!** Det er vigtigt at huske, at $\underline{\xi}$ *ikke er identisk* med $\underline{x}$ og $C\underline{\xi}$ ikke med $(G \circ F)(\underline{x})$, hvis ikke de valgte baser i de to rum består af grundvektorerne.

Vi vil nu finde ud af hvordan matricen C konkret ser ud. Det vil vi gøre ved at efterforske hvordan $G \circ F$ virker på et $\underline{x} \in \mathbb{R}^n$ ud fra viden om hvordan F og G virker. Denne viden er lagret i matrixfremstillinger af henholdsvis F og G, når der i $\mathbb{R}^n$ og $\mathbb{R}^q$ er valgt de nævnte baser, og i $\mathbb{R}^p$ også er valgt en basis, $\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_p$.

Vi har altså, at $F(\underline{x})$'s koordinater i $\underline{v}$-basen for $\mathbb{R}^p$ er givet ved

$$A\underline{\xi},$$

hvor A er en $p \times n$-matrix.

På tilsvarende måde er koordinaterne for $G(\underline{y})$ givet ved

$$B\underline{\tau},$$

hvor $\underline{\tau} = (\tau_1, \tau_2, \ldots, \tau_p)$ er $\underline{y}$'s koordinater i $\underline{v}$-basen for $\mathbb{R}^p$, og hvor B er en $q \times p$-matrix.

Men eftersom $\underline{\tau} = A\underline{\xi}$, har vi at $G(F(\underline{x})$'s koordinater - dvs. $(G \circ F)(\underline{x})$'s koordinater - er

$$C\underline{\xi} = B\underline{\tau} = B(A\underline{\xi})$$

For at udregne eksplicit hvad dette indebærer, tænker vi os at matricerne A og B er møbleret således:

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{p1} & a_{p2} & \ldots & a_{pn} \end{pmatrix}$$

og

$$B = \begin{pmatrix} b_{11} & b_{12} & \ldots & b_{1p} \\ b_{21} & b_{22} & \ldots & b_{2p} \\ \vdots & \vdots & & \vdots \\ b_{q1} & b_{q2} & \ldots & b_{qp} \end{pmatrix}.$$

Så får vi, idet $\underline{\xi} = (\xi_1, \xi_2, \ldots, \xi_n)$, at $\underline{\tau}$ skrevet som søjle er givet ved

$$\begin{pmatrix} \tau_1 \\ \tau_2 \\ \vdots \\ \tau_p \end{pmatrix} = \begin{pmatrix} a_{11}\xi_1 + a_{12}\xi_2 + \cdots + a_{1n}\xi_n \\ a_{21}\xi_1 + a_{22}\xi_2 + \cdots + a_{2n}\xi_n \\ \vdots \\ a_{p1}\xi_1 + a_{p2}\xi_2 + \cdots + a_{pn}\xi_n \end{pmatrix}$$

Denne søjle er nu foderet for G's matrix, B, så at

$$B \begin{pmatrix} \tau_1 \\ \tau_2 \\ \vdots \\ \tau_p \end{pmatrix} = B \begin{pmatrix} a_{11}\xi_1 + a_{12}\xi_2 + \cdots + a_{1n}\xi_n \\ a_{21}\xi_1 + a_{22}\xi_2 + \cdots + a_{2n}\xi_n \\ \vdots \\ a_{p1}\xi_1 + a_{p2}\xi_2 + \cdots + a_{pn}\xi_n \end{pmatrix}.$$

Nu indsættes B. Resultatet af operationen bliver en $q \times 1$- søjle. Udtrykt ved τ-søjlen får vi ved anvendelse af reglerne for multiplikation af en matrix med en (acceptabel) søjle

$$\begin{pmatrix} b_{11}\tau_1 + b_{12}\tau_2 + \cdots + b_{1p}\tau_p \\ b_{21}\tau_1 + b_{22}\tau_2 + \cdots + b_{2p}\tau_p \\ \vdots \\ b_{q1}\tau_1 + b_{q2}\tau_2 + \cdots + b_{qp}\tau_p \end{pmatrix}.$$

Når vi indsætter hvad τ'erne er -jf. søjlen ovenfor - opstår et langt og derfor uoverskueligt, men principielt letforståeligt udtryk, ét for hvert element i $q \times 1$-søjlen. Dens *første* element er

$$b_{11}(a_{11}\xi_1 + a_{12}\xi_2 + \cdots + a_{1n}\xi_n) + b_{12}(a_{21}\xi_1 + a_{22}\xi_2 + \cdots + a_{2n}\xi_n) +$$

$$\cdots + b_{1p}(a_{p1}\xi_1 + a_{p2}\xi_2 + \cdots + a_{pn}\xi_n),$$

som vi ved at samle efter ξ'erne kan omordne til

$$(b_{11}a_{11} + b_{12}a_{21} + \cdots + b_{1p}a_{p1})\xi_1 + (b_{11}a_{12} + b_{12}a_{22} + \cdots + b_{1p}a_{p2})\xi_2 +$$

$$\cdots + (b_{11}a_{1n} + b_{12}a_{2n} + \cdots + b_{1p}a_{pn})\xi_n.$$

Inden vi angiver de øvrige elementer i søjlen, er det nok værd at se nøjere på, hvad det egentlig er der står i udtrykket lige ovenfor. Parentesen foran ξ_1 består af *den første række i B- matricen (matrix)multipliceret med den første søjle i A- matricen.* Parentesen foran ξ_2 består tilsvarende af den første række i B-matricen multipliceret med den anden søjle i A–matricen, og så fremdeles.

Det *andet* element i $q \times 1$-søjlen er skabt efter samme princip som det første. Den *andet* element er

$$b_{21}(a_{11}\xi_1 + a_{12}\xi_2 + \cdots + a_{1n}\xi_n) + b_{22}(a_{21}\xi_1 + a_{22}\xi_2 + \cdots + a_{2n}\xi_n) +$$

$$\cdots + b_{2p}(a_{p1}\xi_1 + a_{p2}\xi_2 + \cdots + a_{pn}\xi_n),$$

som vi ved atter at samle efter ξ'erne kan omordne til

$$(b_{21}a_{11} + b_{22}a_{21} + \cdots + b_{2p}a_{p1})\xi_1 + (b_{21}a_{12} + b_{22}a_{22} + \cdots + b_{2p}a_{p2})\xi_2 +$$

$$\cdots + (b_{21}a_{1n} + b_{22}a_{2n} + \cdots + b_{2p}a_{pn})\xi_n.$$

Her er parentesen foran ξ_j, hvor j = 1,2,...,n, lig med den anden række i B-matricen multipliceret med den j'te søjle i A-matricen.

Og sådan fortsætter det, indtil det sidste - det *q'te* - element i søjlen, som er

$$b_{q1}(a_{11}\xi_1 + a_{12}\xi_2 + \cdots + a_{1n}\xi_n) + b_{q2}(a_{21}\xi_1 + a_{22}\xi_2 + \cdots + a_{2n}\xi_n) +$$

$$\cdots + b_{qp}(a_{p1}\xi_1 + a_{p2}\xi_2 + \cdots + a_{pn}\xi_n),$$

som ved ordning efter ξ'erne bliver til

$$(b_{q1}a_{11} + b_{q2}a_{21} + \cdots + b_{qp}a_{p1})\xi_1 + (b_{q1}a_{12} + b_{q2}a_{22} + \cdots + b_{qp}a_{p2})\xi_2 +$$

$$\cdots + (b_{q1}a_{1n} + b_{q2}a_{2n} + \cdots + b_{qp}a_{pn})\xi_n.$$

Hermed har vi udtrykt samtlige elementer i koordinatsættet for $G(F(\underline{x}))$ ved hjælp af ξ'erne og elementerne i matricerne A og B. Lad os samle hele molevitten i én søjle:

$$\begin{pmatrix} (b_{11}a_{11} + b_{12}a_{21} + \cdots + b_{1p}a_{p1})\xi_1 + & \cdots & + (b_{11}a_{1n} + b_{12}a_{2n} + \cdots + b_{1p}a_{pn})\xi_n \\ (b_{21}a_{11} + b_{22}a_{21} + \cdots + b_{2p}a_{p1})\xi_1 + & \cdots & + (b_{21}a_{1n} + b_{22}a_{2n} + \cdots + b_{2p}a_{pn})\xi_n \\ & \vdots & \\ (b_{q1}a_{11} + b_{q2}a_{21} + \cdots + b_{qp}a_{p1})\xi_1 + & \cdots & + (b_{q1}a_{1n} + b_{q2}a_{2n} + \cdots + b_{qp}a_{pn})\xi_n \end{pmatrix}$$

Hvis vi anstrenger os lidt for at se igennem tågerne i denne formelophobning, opdager vi at søjlens første element er resultatet af en række tal (b, a-parenteserne) multipliceret med ξ-søjlen. Det forholder sig lige sådan med de øvrige elementer i søjlen. Det viser, at vi har at gøre med en $q \times p$-matrix multipliceret med ξ-søjlen:

$$\begin{pmatrix} (b_{11}a_{11} + b_{12}a_{21} + \cdots + b_{1p}a_{p1}) & \dots & (b_{11}a_{1n} + b_{12}a_{2n} + \cdots + b_{1p}a_{pn}) \\ (b_{21}a_{11} + b_{22}a_{21} + \cdots + b_{2p}a_{p1}) & \dots & (b_{21}a_{1n} + b_{22}a_{2n} + \cdots + b_{2p}a_{pn}) \\ & \vdots & \\ (b_{q1}a_{11} + b_{q2}a_{21} + \cdots + b_{qp}a_{p1}) & \dots & (b_{q1}a_{1n} + b_{q2}a_{2n} + \cdots + b_{qp}a_{pn}) \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{pmatrix}.$$

Det vi her har fundet er $q \times p$-matricen C for den sammensatte afbildning $G \circ F$. Denne matrix er skabt ved manøvrer med elementerne i A og B. Imidlertid så vi jo, at koordinaterne for $(G \circ F)(\underline{x}) = G(F(\underline{x}))$ er givet ved $C\underline{\xi} = B(A\underline{\xi})$. Den fundne matrix C benævnes derfor **matrixproduktet** BA af matricerne B og A og noteres:

$$C = BA.$$

Læg mærke til at F's matrix A står til højre og G's matrix B står til venstre i dette matrixprodukt svarende til rækkefølgen af F og G i $G \circ F$.

Matrixproduktets egenskaber

Allerførst ser vi på, hvordan matrixproduktet BA er fremkommet af A og B. Ved direkte beskuelse af matrixproduktet som det er skrevet op ovenfor, ser vi at $C = BA$ består af q **rækker og** n **søjler**. Dette skal sammenholdes med at B består af q rækker og p søjler, mens A består af p rækker og n søjler. En simpel måde at huske dette på er:

$$q \times n = (q \times p)(p \times n),$$

hvor altså antallet - p - af søjler i B lig antallet af rækker i A, "går ud" af regnestykket.

I samme åndedrag konstaterer vi, at vi *kun kan multiplicere en matrix B med en matrix A (fra højre), hvis antallet af søjler i B er lig antallet af rækker i A.* Det er da uden videre klart, at vi ikke kan vente i almindelighed at have lov til at multiplicere B med A *fra venstre*. Dette er simpelthen kun muligt, hvis antallet af søjler i A er lig antallet af rækker i B, dvs. hvis B er en $q \times n$-matrix samtidig med at A er en $n \times q$-matrix. Allerede af den grund er *matrixproduktet ikke kommutativt*. Men selv hvis både BA og AB kan dannes, er de sjældent ens. BA er jo en $q \times n$-matrix, mens AB er en $n \times q$-matrix.

Nå, men hvad står der nu i $C = BA$? Jo, ved fortsat ransagning af udtrykket for C ser vi, at elementet anbragt på plads $(1, 1)$, dvs. i den første række og den første søjle er dannet ved at den første række i B er (matrix)multipliceret med den første søjle i

A. De øvrige elementer i den første række af C er dannet ved at den første række i B er multipliceret med de øvrige søjler i A, én ad gangen. I den anden række i C er forholdene tilsvarende, bortset fra at det nu er den anden række i B, der multipliceres med søjlerne fra A for at danne elementerne i den anden række i C. Således fortsætter det. Den almene regel, som vi vil nedfælde som en *definition* er altså, at

elementet på den (i,j)'te plads i produktmatricen $C = BA$ er skabt ved at den i'te række i B (matrix)multipliceres med den j'te søjle i A.

For en umiddelbar betragtning kunne det se ud som om, der i selve definitionen af matrixmultiplikation allerede indgår et begreb om matrixmultiplikation. Sådan står det jo udtrykkeligt i formuleringen, som derved ser cirkulær ud. Det er den imidlertid ikke, idet der i definitionen kun benyttes multiplikation mellem en række og en *søjle* (fra højre), som vi tidligere har indført.

Lad os inden vi kaster os over et par eksempler gøre endnu en observation:

$*$ Den enkelte søjle i C, lad os sige nr. j, består af rækkerne i B multipliceret med søjle nr. j fra A, med andre ord af matrix-søjleproduktet $B(j$'te søjle i $A)$ (som tidligere indført). Derfor kan vi *anskue matrixproduktet BA som en sammenstilling af de søjler der fremkommer ved at multiplicere B med søjlerne i A, én efter én.*

Nu til et par eksempler.

Eksempel 3

Matrixproduktet $C = BA$ af 2×4 og 4×3 matricerne fra Eksempel 1:

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

er 2×3-matricen

$$C = \begin{pmatrix} (0+2-2+0) & (0+1+0-6) & (0+3-1+4) \\ (-2-6+6+0) & (4-3+0+0) & (0-9+3+0) \end{pmatrix} = \begin{pmatrix} 0 & -5 & 6 \\ -2 & 1 & -6 \end{pmatrix},$$

sådan som vi også fandt det, men på en mindre gennemskuelig måde, i Eksempel 1. I dette tilfælde er det ikke muligt at danne produktet AB, da antallet af søjler i A ikke er lig antallet af rækker i B.

Eksempel 4

Her ser vi på matricerne fra Eksempel 2:

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

Produktmatricen $C = BA$ fås så som 2×2-matricen

$$C = \begin{pmatrix} (2+0+0) & (2+8+0) \\ (3+0+2) & (3-2-6) \end{pmatrix} = \begin{pmatrix} 2 & 8 \\ 5 & -5 \end{pmatrix},$$

som vi også fandt i Eksempel 2. I dette eksempel er det muligt at danne AB, som bliver en 3×3-matrix:

$$\begin{pmatrix} (2+3) & (-4+1) & (0-2) \\ (0-6) & (0-2) & (0+4) \\ (-2+9) & (4+3) & (0-6) \end{pmatrix} = \begin{pmatrix} 5 & -3 & -2 \\ -6 & -2 & 4 \\ 7 & 7 & -6 \end{pmatrix}.$$

Det fremgår, at BA og AB ikke ligner hinanden særlig meget.

Et par observationer i tilknytning til matrixmultiplikation byder sig umiddelbart til.

$*$ Hvis en vilkårlig matrix A multipliceres fra højre eller venstre med en **nulmatrix**, dvs. matrix bestående af lutter 0'er - og med et "lovligt" antal rækker, respektive søjler - bliver resultatet en nulmatrix. Dette følger åbenbart af, at alle de produktsummer der danner elementerne i produktmatricen er 0.

$*$ Vi har tidligere set, at hvis en $1 \times n$-matrix ganges (fra højre) med en $n \times 1$-matrix, bliver resultatet et *tal*.

$*$ Hvis en $n \times 1$-matrix A ganges (fra højre) med en $1 \times n$-matrix B bliver resultatet en $n \times n$-*matrix*. På den (i,j)'te plads i produktmatricen står jo den i'te række i A (der jo her blot består af elementet a_{i1}) ganget med den j'te søjle i B (der her blot består af elementet b_{1j}), med tallet $a_{i1}b_{1j}$ som resultat.

Én slags matricer er af stor betydning i matrixregningen. Det er de såkaldte **enhedsmatricer**. Enhver enhedsmatrix er en *kvadratisk* matrix, dvs. en matrix med lige mange rækker og søjler. Den n'te enhedsmatrix har 1-taller i diagonalen, altså på pladserne (i,i) med samme række- søjlenummer, mens alle andre pladser er besat med 0'er:

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

hvor der er n rækker og søjler.

Det særlige ved enhedsmatricerne - eller, hvis det af sammenhængen klart fremgår at der kun er tale om én, så enhedsmatricen - er, at de er **neutrale** ved multiplikation fra både højre og venstre. Det vil sige, at hvis A er en vilkårlig $p \times n$-matrix er både

$$E_p A = A, \quad \text{og} \quad AE_n = A.$$

Den første af disse påstande indses således: Vi skal bestemme det (i,j)'te element i $E_p A$. Det gøres ved at multiplicere den i'te række i E_p med den j'te søjle i A, dvs.

$$\begin{pmatrix} 0 & 0 & \dots & 1 & \dots & 0 \end{pmatrix} \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{pj} \end{pmatrix} = a_{ij},$$

idet 1-tallet i den i'te række i E_p står på den i'te plads. Da dette er situationen på alle pladser (i,j), stemmer $E_p A$ overens med A. På tilsvarende måde finder vi det (i,j)'te element i AE_n:

$$\begin{pmatrix} a_{i1} & a_{i2} & \dots & a_{in} \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = a_{ij},$$

hvor 1-tallet i den j'te søjle af E_n står på den j'te plads. Det konkluderes altså igen, at AE_n stemmer overens med A.

For lidt siden blev det demonstreret at *matrixmultiplikation ikke er kommutativ*. Det er til gengæld det eneste der er "galt" med matrixmultiplikation. Ellers opfører den sig pænt. Der gælder nemlig

SÆTNING. *Matrixmultiplikationen opfylder egenskaberne (for de matricer, hvor produkterne er defineret):*

$$\begin{aligned} A(BC) &= (AB)C \text{ associativitet} \\ A(B+C) &= AB + AC \text{ og } (F+G)H = FH + GH \text{ distributivitet} \\ A(\lambda B) &= (\lambda A)B = \lambda AB. \end{aligned}$$

Her skal vi huske fra indledningsafsnittet, at addition af matricer af samme dimension sker pladsvis. Påstanden i sætningen godtgøres uden vanskeligheder ved udregning efter

definitionen. Hvad associativiteten angår, er der dog tale om en besværlig udregning, som passende kan udføres som en øvelse. For i øvrigt at belyse hvad der er på færde, vil vi bevise den første distributive lov.

Lad til den ende A være en $p \times n$-matrix, mens både B og C er $n \times m$-matricer (det vigtige er n'et). For at udregne det (i,j)'te element i $A(B+C)$ skal den i'te række i A multipliceres med den j'te søjle i $B+C$:

$$
\begin{aligned}
&\begin{pmatrix} a_{i1} & a_{i2} & \dots & a_{in} \end{pmatrix} \begin{pmatrix} b_{1j}+c_{1j} \\ b_{2j}+c_{2j} \\ \vdots \\ b_{nj}+c_{nj} \end{pmatrix} \\
&= a_{i1}(b_{1j}+c_{1j}) + a_{i2}(b_{2j}+c_{2j}) + \quad \dots \quad + a_{in}(b_{nj}+c_{nj}) \\
&= a_{i1}b_{1j} + a_{i1}c_{1j} + a_{i2}b_{2j} + a_{i2}c_{2j} + \quad \dots \quad + a_{in}b_{nj} + a_{in}c_{nj} \\
&= (a_{i1}b_{1j} + a_{i2}b_{2j} + \dots + a_{in}b_{nj}) + (a_{i1}c_{1j} + a_{i2}c_{2j} + \dots + a_{in}c_{nj}),
\end{aligned}
$$

der netop er summen af det (i,j)'te led i AB og det (i,j)'te led i AC, dvs. i alt det (i,j)'te led i $AB+AC$.

Bijektive afbildninger og invertible matricer

Vi skal nu se lidt nærmere på matricerne for *bijektive* lineære afbildninger. I et tidligere afsnit konstaterede vi ved hjælp af dimensionssætningen, at hvis $F : \mathbb{R}^n \to \mathbb{R}^p$ er bijektiv, må $n = p$. Dvs. en bijektiv lineær afbildning er en afbildning af et talrum $\mathbb{R}^n$ på sig selv:

$$F : \mathbb{R}^n \to \mathbb{R}^n.$$

I en hvilken som helst matrixfremstilling for en bijektiv lineær afbildning, må matricen derfor være *kvadratisk*.

Nu har vi også tidligere indset, at en injektiv lineær afbildning F har en *omvendt* afbildning

$$F^{-1} : \mathbb{R}^n \to \mathbb{R}^n.$$

Denne afbildning er defineret ved at hvis $\underline{y} \in \mathbb{R}^n$ er $F^{-1}(\underline{y})$ = det éntydigt bestemte $\underline{x} \in \mathbb{R}^n$, som opfylder, at $F(\underline{x}) = \underline{y}$. Nu er også F^{-1} *lineær*. Selv om det næppe er særligt overraskende, har vi faktisk ikke godtgjort det. Det ordner vi lige i farten:

Vi skal indse: hvis $\underline{y}_1$ og $\underline{y}_2$ ligger i $\mathbb{R}^n$, og α_1 og α_2 er vilkårlige skalarer, er

$$F^{-1}(\alpha_1 \underline{y}_1 + \alpha_2 \underline{y}_2) = \alpha_1 F^{-1}(\underline{y}_1) + \alpha_2 F^{-1}(\underline{y}_2)$$

Ved at benytte F på begge sider, ser vi at denne identitet er ensbetydende med identiteten

$$F(F^{-1}(\alpha_1 \underline{y}_1 + \alpha_2 \underline{y}_2)) = F(\alpha_1 F^{-1}(\underline{y}_1) + \alpha_2 F^{-1}(\underline{y}_2)).$$

Her er venstresiden (da $F \circ F^{-1} =$ den identiske afbildning) lig med

$$\alpha_1 \underline{y}_1 + \alpha_2 \underline{y}_2,$$

mens højresiden i kraft af F's linearitet er lig

$$\alpha_1 F(F^{-1}(\underline{y}_1)) + \alpha_2 F(F^{-1}(\underline{y}_2)) = \alpha_1 \underline{y}_1 + \alpha_2 \underline{y}_2.$$

Da altså venstresiden og højresiden efter brug af F tydeligvis er identiske, er de oprindelige venstre- og højresider identiske. Derfor er F^{-1} lineær.

Eksempel 5

Hvis $F : \mathbb{R}^3 \to \mathbb{R}^3$ er defineret ved

$$\underline{y} = F(\underline{x}) = A\underline{x},$$

hvor

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

er F bijektiv, da A er kvadratisk og dens søjler lineært uafhængige (kontrollér selv det). Så bestemmes $F^{-1}(\underline{x})$, lig det $\underline{x} = (x_1, x_2, x_3)$ for hvilket $\underline{y} = A\underline{x}$, som den entydigt bestemte løsning til ligningssystemet

$$\begin{pmatrix} 2 & 0 & 1 \\ -1 & 1 & 0 \\ 0 & 3 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}.$$

Ved udførelse af de sædvanlige gymnastiske øvelser finder vi, at

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 & 3 & -1 \\ 2 & 4 & -1 \\ -3 & -6 & 2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}.$$

Altså er

$$F^{-1}(\underline{y}) = \begin{pmatrix} 2 & 3 & -1 \\ 2 & 4 & -1 \\ -3 & -6 & 2 \end{pmatrix} \underline{y}.$$

Vi slutter eksemplet med at multiplicere matricen for F og matricen F^{-1} både fra højre og venstre:

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

$$\begin{pmatrix} 2 & 0 & 1 \\ -1 & 1 & 0 \\ 0 & 3 & 2 \end{pmatrix} \begin{pmatrix} 2 & 3 & -1 \\ 2 & 4 & -1 \\ -3 & -6 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

At dette resultat ikke er nogen tilfældighed vil blive åbenbart nedenfor. $*$

Vi vil nu studere *situationen i almindelighed.*

Først minder vi om, at hvis $F : \mathbb{R}^n \to \mathbb{R}^n$ er *injektiv, så er den også bijektiv.* Thi da $dimF(\mathbb{R}^n) = n - dimN(F) = n$ (fordi F er injektiv) er F surjektiv.

Hvis derfor $F : \mathbb{R}^n \to \mathbb{R}^n$ er injektiv eksisterer også $F^{-1} : \mathbb{R}^n \to \mathbb{R}^n$. Den er i følge det foregående lineær, hvorfor den har en matrixfremstilling i forhold til en given basis for $\mathbb{R}^n$. Har F i denne basis matricen A kan vi kalde matricen for F^{-1} for B. Hvad kan der nu siges om B?

Jo, matricen for $F^{-1} \circ F$ er efter det foregående afsnit lig BA. På den anden side er $F^{-1} \circ F = I_n$, hvor I_n er den **identiske afbildning** af $\mathbb{R}^n$ ind i sig selv, der lader alle vektorer uberørt:

$$I_n(\underline{x}) = \underline{x}.$$

Den identiske afbildning har åbenbart i enhver basis enhedsmatricen E_n som matrix. Når derfor $F^{-1} \circ F = I_n$ er så $BA = E_n$.

Nu gælder der imidlertid også, at $F \circ F^{-1} = I_n$. Eftersom matricen for $F \circ F^{-1}$ er AB, har vi så at $AB = E_n$.

Summa summarum har vi om matricen B for F^{-1}, at

$$AB = BA = E_n.$$

Det vil sige, at B er en slags *reciprok* til A. Derfor, og inspireret af at B er matricen for F^{-1}, betegnes B med

$$B = A^{-1},$$

og kaldes den **inverse matrix** - somme tider også den *omvendte* - for A.

Da F og F^{-1} er hinandens omvendte, må vi have

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

Vi har altså indset følgende:

Hvis A er en $(n \times n)$-matrix for en *injektiv* (dermed her en *bijektiv*) afbildning $F : \mathbb{R}^n \to \mathbb{R}^n$, findes en invers matrix A^{-1} til A. Vi siger, at A er **invertibel**. Den inverse matrix A^{-1} er matrix for F^{-1}, den omvendte afbildning til F.

Er nu omvendt A en $n \times n$-matrix for hvilken der findes en $n \times n$-matrix B, så at

$$AB = BA = E_n,$$

er A matrix for en lineær afbildning $F : \mathbb{R}^n \to \mathbb{R}^n$ defineret ved $F(\underline{x}) = A\underline{x}$. Denne afbildning er da injektiv. For hvis $F(\underline{x}) = F(\underline{x}')$, er $A\underline{x} = A\underline{x}'$, hvoraf:

$$\underline{x} = E_n\underline{x} = (BA)\underline{x} = B(A\underline{x}) =$$

$$B(A\underline{x}') = (BA)\underline{x}' = E_n\underline{x}' = \underline{x}',$$

hvilket viser, at $\underline{x} = \underline{x}'$. F er altså injektiv (dvs. bijektiv). Så findes A^{-1}, der jo også opfylder

$$AA^{-1} = A^{-1}A = E_n.$$

Nu må $B = A^{-1}$, thi

$$B = BE_n = B(AA^{-1}) = (BA)A^{-1} = E_nA^{-1} = A^{-1}.$$

Konklusionen på disse betragtninger er: Hvis der til $n \times n$- matricen A findes en $n \times n$-matrix B, så at

$$AB = BA = E_n,$$

er A invertibel og $A^{-1} = B$.

A er matrix for en injektiv (bijektiv) afbildning præcis hvis alle søjlerne i A er lineært uafhængige (dimensionssætningen nok engang). Eftersom A er kvadratisk er dette ensbetydende med at også rækkerne er lineært uafhængige.

Nu har vi sparet sammen til en sætning:

SÆTNING. *For en $n \times n$-matrix A er følgende udsagn ensbetydende:*

(1) *A er invertibel, dvs. matrix for en injektiv afbildning $F : \mathbb{R}^n \to \mathbb{R}^n$*
(2) *Der findes netop én matrix B så at $AB = BA = E_n$ (Denne matrix er $B = A^{-1}$, matricen for F's omvendte afbildning F^{-1})*
(3) *Søjlerne i A er lineært uafhængige*
(4) *Rækkerne i A er lineært uafhængige*

Eksempel 6

Matricen

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

er invertibel, da dens søjler er lineært uafhængige (godtgør selv det). Dens inverse er

$$A^{-1} = \begin{pmatrix} 5 & -2 & 2 & -1 \\ -9 & \frac{7}{2} & -3 & 2 \\ \frac{9}{2} & -\frac{3}{2} & \frac{3}{2} & -1 \\ -3 & 1 & -1 & 1 \end{pmatrix}.$$

Kontrollér selv denne påstand ved udregning.

$*$ Ovenfor så vi, at $(A^{-1})^{-1} = A$. En anden observation er, at

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

Der gælder nemlig

$$(AB)(B^{-1}A^{-1}) = A[B(B^{-1}A^{-1})] = A[(BB^{-1})A^{-1}] =$$

$$A(E_nA^{-1}) = AA^{-1} = E_n$$

og

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

$$(B^{-1}E_n)B = B^{-1}B = E_n,$$

altså i alt

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

Dette viser, at $(B^{-1}A^{-1})$ er AB's inverse.

De kriterier for invertibilitet af en matrix som er indeholdt i sætningen ovenfor er aldeles udmærkede, men de er ikke særligt *operationelle*, sådan at forstå at der ikke uden videre er nogen automatisk metode til at afgøre om de er opfyldt. I betragtning af at invertibilitet er en så vigtig egenskab ved matricer for både teoretiske og praktiske formål, er der behov for et mere operationelt kriterium. Derom handler næste afsnit.

Determinanter

Et kriterium af den art vi efterspørger har at gøre med begrebet *determinant*. Enhver *kvadratisk* matrix får (om lidt, hav venligst tålmodighed) knyttet et reelt tal, matricens determinant, til sig. Dette tal er en karakteristisk størrelse for matricen på samme måde som vægten af en person er en karakteristtisk størrelse for personen. Determinanten fortæller en del vigtige ting om "sin" matrix, bl.a. om den er invertibel eller ej. Lidt senere skal vi også bruge den til andre formål.

Teorien for determinanter er omfattende og indviklet. Det er derfor kun muligt at behandle dem nødtørftigt her.

Først tager vi et meget lille skridt: Vi definerer determinanten af en 1×1-matrix (a). Determinanten skal simpelthen være tallet i matricen selv:

$$det(a) = a.$$

Dernæst indfører vi determinanter for 2×2-matricer. Hvis A er en 2×2-matrix:

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix},$$

sættes

$$detA = a_{11}a_{22} - a_{21}a_{12}.$$

Udtrykket ser måske bekendt ud for nogen af jer, enten fra en tidligere beskæftigelse med 2×2-matricer, eller fra løsningen af to ligninger med to ubekendte.

Nu tager vi en elevatortur. Vi tænker os, at vi allerede har tillagt alle $(n-1) \times (n-1)$-matricer en determinant (én hver, altså). Så vil vi definere determinanten af en vilkårlig $n \times n$-matrix. Lad os betragte en sådan:

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

Nu tager vi fat i den første søjle i A. Forestiller vi os dannet en matrix af A ved at slette den første række og den første søjle i A har vi at gøre med en $(n-1) \times (n-1)$- matrix. Den betegnes A_{11} og kaldes *komplementet* til elementet a_{11} i A. Som $(n-1) \times (n-1)$-matrix har den, efter forudsætningen, en determinant.

Dernæst går vi til det næste plads $(2,1)$ i den første søjle. Igen sletter vi den pågældende række - her den anden - og søjlen fra A. Herved opstår en ny $(n-1) \times (n-1)$-matrix, A_{21}, komplementet til a_{21}. Også A_{21} har en determinant.

Således fortsættes ned i gennem første søjle for A. Derved har vi høstet n stks. $(n-1)\times(n-1)$-matricer, A_{11}, A_{21},..., A_{n1}, der alle har determinanter. Nu er vi klar til at definere determinanten for A, men inden vil jeg give en advarsel: Prøv ikke at forstå, hvorfor i alverden definitionen ser sådan ud. Det er ikke muligt på det grundlag vi kan skabe i dette kursus. Tag den i stedet til efterretning, og lær at omgås den gennem eksempler. Så skal definitionen ikke udsættes længere:

$$detA = (-1)^{1+1}a_{11}detA_{11} + (-1)^{2+1}a_{21}detA_{21} + \quad \ldots \quad + (-1)^{n+1}a_{n1}detA_{n1}.$$

Det er vist påkrævet straks at prøve definitionen af på konkrete eksempler. Allerførst er der grund til at undersøge om den passer for 2×2-matricer, dvs. om denne måde at danne determinanten af en 2×2-matrix på, ud fra determinanten af 1×1-matricer, stemmer overens med den tidligere definition. Til den ende betragtes igen

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

Vi skal opsøge den første søjle og dens komplementer. Komplementet til a_{11} fås ved at slette den første række og den første søjle i A. Derved opstår $A_{11} = (a_{22})$, hvis determinant er a_{22}. Komplementet til a_{21} er $A_{21} = (a_{12})$, hvis determinant er a_{12}. Dermed bliver efter den generelle definition

$$detA = (-1)^{1+1}a_{11}detA_{11} + (-1)^{2+1}a_{21}detA_{21} = a_{11}a_{22} - a_{21}a_{12}.$$

Minsandten! Det blev det rigtige.

Når vi har defineret determinanten af en 2×2-matrix, kan vi definere determinanten af enhver 3×3-matrix ved at benytte den generelle motor. Vi ser på en vilkårlig 3×3-matrix

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

Igen går vi til den første søjle i A og opsøger komplementerne. Komplementet til a_{11} er 2×2-matricen

$$A_{11} = \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix}.$$

På tilsvarende måde finder vi komplementerne til a_{21} og a_{31}:

$$A_{21} = \begin{pmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{pmatrix} \quad \text{og} \quad A_{31} = \begin{pmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{pmatrix}.$$

Det er ingredienserne i $detA$:

$$detA = (-1)^{1+1}a_{11}detA_{11} + (-1)^{2+1}a_{21}detA_{21} + (-1)^{3+1}a_{31}detA_{31} =$$

$$a_{11}(a_{22}a_{33} - a_{32}a_{23}) - a_{21}(a_{12}a_{33} - a_{32}a_{13}) + a_{31}(a_{12}a_{23} - a_{22}a_{13}).$$

Ved at omordne efter fortegn m.m. bliver resultatet

$$detA = (a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{21}a_{32}a_{13}) - (a_{13}a_{22}a_{31} + a_{12}a_{21}a_{33} + a_{23}a_{32}a_{11}).$$

Denne opskrivningsrækkefølge er valgt fordi den gør det muligt at visualisere beregningen. De positive led i determinanten er de tre "højreskrå" produkter af elementerne i A, mens de negative er de tre "venstreskrå":

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

Nu har vi defineret determinanten af alle 3×3-matricer. Så kan vi gå videre til 4×4-matricer ... Nej, I skal nok blive skånet for det generelle. Et eksempel må række.

Eksempel 7

Matricen fra Eksempel 6 så sådan ud

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

Dens determinant bestemmes ved

$$\begin{aligned} detA &= (-1)^{1+1}(-1)detA_{11} + (-1)^{2+1}0detA_{21} + (-1)^{3+1}3detA_{31} + (-1)^{4+1}0detA_{41} \\ &= -detA_{11} + 3detA_{31}, \end{aligned}$$

hvor

$$A_{11} = \begin{pmatrix} 2 & 4 & 0 \\ 2 & 0 & -1 \\ 0 & 2 & 3 \end{pmatrix} \qquad \text{og} \qquad A_{31} = \begin{pmatrix} 0 & 2 & 1 \\ 2 & 4 & 0 \\ 0 & 2 & 3 \end{pmatrix}.$$

Heraf finder vi

$$detA_{11} = (2\cdot 0\cdot 3 + 4\cdot(-1)\cdot 0 + 0\cdot 2\cdot 2) - (0\cdot 0\cdot 0\cdot + 4\cdot 2\cdot 3 + 2\cdot(-1)\cdot 2) = -(24-4) = -20$$

og

$$detA_{31} = (0\cdot 4\cdot 3 + 2\cdot 0\cdot 0 + 1\cdot 2\cdot 2) - (1\cdot 4\cdot 0 + 2\cdot 2\cdot 3 + 0\cdot 2\cdot 0) = 4 - 12 = -8,$$

som ved indsættelse giver

$$detA = 20 + 3 \cdot (-8) = -4.*$$

På denne måde kan man altså **rekursivt**, som det hedder, tillægge enhver kvadratisk matrix en determinant ved at bygge på determinanter for matricerne ét nummer mindre.

Den ovenstående definition af determinanter tog udgangspunkt i den første søjle, et for så vidt vilkårligt valg. Imidlertid gælder det - men det må vi afstå fra at bevise - at vi lige så godt kunne have benyttet en hvilken som helst af de andre søjler, *resultatet var blevet det samme*. Generelt betegner vi med A_{ij} den $(n-1) \times (n-1)$-matrix der fremkommer af A, hvis vi sletter den i'te række og den j'te søjle - den kaldes naturligt nok **komplementet** til elementet a_{ij}. Så får vi ved at tage udgangspunkt i den k'te søjle, hvor k er ethvert af tallene $1, 2, \ldots, n$, at

$$detA = (-1)^{1+k} a_{1k} detA_{1k} + (-1)^{2+k} a_{2k} detA_{2k} + \quad \ldots \quad + (-1)^{n+k} a_{nk} detA_{nk}.$$

Man siger at determinanten er beregnet ved **udvikling** efter den k'te søjle.

Men vi kunne også have benyttet os af en vilkårlig *række*. Der gælder nemlig også for den k'te række, at

$$detA = (-1)^{k+1} a_{k1} detA_{k1} + (-1)^{k+2} a_{k2} detA_{k2} + \quad \ldots \quad + (-1)^{k+n} a_{kn} detA_{kn}.$$

Man siger, at determinanten er beregnet ved udvikling efter den k'te række.

Når det gælder beregning af determinanten for en konkret matrix, er det bekvemt at have den frihed i valget af beregningsudgangspunkt som fremgår af de to sidste resultater. Regnerierne vil blive lettest hvis man vælger en række/søjle med mange 0'er.

En væsentlig motivation for at vi overhovedet interesserer os (moderat) for determinanter, var deres evne til at karakterisere de invertible matricer. Den kommer til udtryk i

SÆTNING. *En kvadratisk matrix A er invertibel hvis og kun hvis dens determinant er forskellig fra 0: $detA \neq 0$,*

som vi ikke giver noget bevis for.

Afsnittet afsluttes med et par nyttige egenskaber ved determinanter.

$*$ Hvis én af rækkerne eller én af søjlerne i A består af lutter 0'er, er $detA = 0$. Det følger umiddelbart ved udvikling efter den pågældende række/søjle.

$*$ Hvis A og B er $n \times n$-matricer gælder $det(AB) = detA \cdot detB$. Dette bevises ikke.

$*$ Med A^T betegnes den **transponerede** for matricen A, dvs. den matrix der fremgår af A ved at opskrive dens søjler som rækker: Elementet b_{ij} på plads (i,j) i A^T er givet ved a_{ji}. Hvis A er kvadratisk er også A^T kvadratisk:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \quad \text{og} \quad A^T = \begin{pmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{pmatrix}.$$

Der gælder nu, at $detA^T = detA$, som vi dog ikke vil bevise.

$*$ Hvis A er en **øvre trekantsmatrix**, dvs. A har lutter 0'er under diagonalen:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix},$$

kan determinanten beregnes på en meget simpel måde, idet den simpelthen består af produktet af diagonalelementerne:

$$detA = a_{11}a_{22}\dots a_{nn}.$$

Dette følger faktisk let af definitionen. Udvikler vi nemlig efter første søjle, er der kun ét led, nemlig $a_{11}detA_{11}$, alle de andre bliver 0. Imidlertid er også A_{11} en øvre trekantsmatrix. Udvikler vi den efter dens første søjle, får vi dens determinant til at være a_{22} multipliceret med determinanten af den matrix der fremgår af A ved at slette dens to første søjler og to første rækker. Det er på ny en øvre trekantsmatrix, hvor der i øverste venstre hjørne står a_{33}. På denne måde gnaver vi os ned igennem A, indtil vi har konsumeret den sidste bid, determinanten til (a_{nn}), altså tallet a_{nn}.

$*$ Hvis A er en **nedre trekantsmatrix**, dvs. en matrix med lutter 0'er over diagonalen, er dens determinant igen lig produktet af diagonalelementerne. Det kan indses enten ved at ræsonnement helt magen til det foregående, eller ved at lægge mærke til at A^T er en øvre trekantsmatrix, netop når A er en nedre trekantsmatrix, og udnytte at $detA^T = detA$.

$*$ Enhedsmatricen er både en øvre og en nedre trekantsmmatrix. Dens determinant er derfor lig produktet af diagonalelementerne, altså 1: $detE_n = 1$.

BASISSKIFT

Flere gange i det foregående har vi strejfet det forhold at en lineær afbildning har forskellige matrixfremstillinger i forskellige baser. I dette afsnit vil vi foretage en systematisk undersøgelse af spørgsmålet. Vores interesse for sagen skyldes, at der i forbindelse med en lineær afbildning ofte er "naturlige" baser som er forskellige fra baserne bestående af grundvektorerne. For at få en fornemmelse af hvad der er på færde lægger vi ud med et eksempel, før vi går til den generelle situation.

Eksempel 1

I $\mathbb{R}^2$ forestiller vi os givet basen $\underline{b}_1 = (1,2)$, $\underline{b}_2 = (-1,1)$, og i $\mathbb{R}^3$ basen $\underline{c}_1 = (2,0,-1)$, $\underline{c}_2 = (1,1,1)$, $\underline{c}_3 = (0,-2,2)$ (check selv at der virkelig er tale om baser i de to rum). En lineær afbildning $F : \mathbb{R}^2 \to \mathbb{R}^3$ er givet ved dens billeder af den valgte basis i $\mathbb{R}^2$:

$$F(\underline{b}_1) = 3\underline{c}_1 + 2\underline{c}_3, \quad F(\underline{b}_2) = -\underline{c}_1 + 2\underline{c}_2 - 3\underline{c}_3.$$

Det betyder at F's matrixfremstilling i disse baser er

$$\underline{\tau} = \begin{pmatrix} 3 & -1 \\ 0 & 2 \\ 2 & -3 \end{pmatrix} \underline{\xi},$$

hvor $\underline{\xi}$ er $\underline{x}$'s koordinater i $\underline{b}$-basen for $\mathbb{R}^2$, mens $\underline{\tau}$ er billedvektorens, $F(\underline{x})$'s, koordinater i $\underline{c}$-basen i $\mathbb{R}^3$. Så vidt så godt.

Nu skiftes baser i både $\mathbb{R}^2$ og $\mathbb{R}^3$. Vi forestiller os at vi går over til den nye basis $\underline{b}'_1 = (2,0)$, $\underline{b}'_2 = (-3,2)$, i $\mathbb{R}^2$, mens vi skifter til basen $\underline{c}'_1 = (1,2,0)$, $\underline{c}'_2 = (3,0,1)$, $\underline{c}'_3 = (0,0,-2)$ i $\mathbb{R}^3$.

For at undersøge hvordan basisskiftet foregår i hvert af de to rum ser vi først på $\mathbb{R}^2$. Med $\underline{x}$ betegnes en vilkårlig vektor i $\mathbb{R}^2$. I den oprindelige basis har den koordinaterne (ξ_1, ξ_2):

$$\underline{x} = \xi_1 \underline{b}_1 + \xi_2 \underline{b}_2,$$

mens den i den nye basis har koordinaterne ξ'_1 og ξ'_2:

$$\underline{x} = \xi'_1 \underline{b}'_1 + \xi'_2 \underline{b}'_2.$$

Hvordan hænger nu ξ-erne og ξ'-erne sammen? Ja, det kan vi besvare ved at udnytte, at de nye basisvektorer jo kan fremstilles i den gamle basis:

$$\underline{b}'_1 = \frac{2}{3}\underline{b}_1 - \frac{4}{3}\underline{b}_2, \quad \text{og} \quad \underline{b}'_2 = -\frac{1}{3}\underline{b}_1 + \frac{8}{3}\underline{b}_2.$$

(Koefficienterne heri findes ved at løse ligningerne

$$\underline{b}_1' = \alpha_1 \underline{b}_1 + \beta_1 \underline{b}_2$$

og

$$\underline{b}_2' = \alpha_2 \underline{b}_1 + \beta_2 \underline{b}_2$$

i α_1 og α_2.) Ved at indsætte det i fremstillingen af $\underline{x}$ i $\underline{b}'$-basen får vi

$$\underline{x} = \xi_1'(\frac{2}{3}\underline{b}_1 - \frac{4}{3}\underline{b}_2) + \xi_2'(-\frac{1}{3}\underline{b}_1 + \frac{8}{3}\underline{b}_2) = (\frac{2}{3}\xi_1' - \frac{1}{3}\xi_2')\underline{b}_1 + (-\frac{4}{3}\xi_1' + \frac{8}{3}\xi_2')\underline{b}_2.$$

Dette er faktisk en fremstilling af $\underline{x}$ i $\underline{b}$-basen. Sammenholdes den med den oprindelige fremstilling af $\underline{x}$ i denne basis, må der gælde - da der kun findes én sådan fremstilling - at

$$\xi_1 = \frac{2}{3}\xi_1' - \frac{1}{3}\xi_2', \quad \xi_2 = -\frac{4}{3}\xi_1' + \frac{8}{3}\xi_2',$$

eller på matrixform

$$\begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = \begin{pmatrix} \frac{2}{3} & -\frac{1}{3} \\ -\frac{4}{3} & \frac{8}{3} \end{pmatrix} \begin{pmatrix} \xi_1' \\ \xi_2' \end{pmatrix} = S \begin{pmatrix} \xi_1' \\ \xi_2' \end{pmatrix},$$

hvor S blot er en kort betegnelse for matricen. Bogstavet S står for "skift". Denne matrix, der er invertibel (dens determinant er $4 \neq 0$) angiver hvordan overgangen til en ny basis forplanter sig til et skift i koordinater for en given vektor:

$$\begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = S \begin{pmatrix} \xi_1' \\ \xi_2' \end{pmatrix}.$$

I matrixligningen udtrykkes de *gamle* koordinater ved hjælp af de *nye*. Da S er invertibel kan de nye koordinater lige så godt udtrykkes ved hjælp af de gamle:

$$\begin{pmatrix} \xi_1' \\ \xi_2' \end{pmatrix} = S^{-1} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}.$$

Derefter går vi frem på helt analog måde i $\mathbb{R}^3$. Vi ser på $\underline{y}$, en vilkårlig vektor i $\mathbb{R}^3$, med koordinaterne (τ_1, τ_2, τ_3) i den gamle basis, $\underline{c}$-basen, og med koordinaterne $(\tau_1', \tau_2', \tau_3')$ i den nye basis, $\underline{c}'$-basen:

$$\underline{y} = \tau_1 \underline{c}_1 + \tau_2 \underline{c}_2 + \tau_3 \underline{c}_3, \text{ og } \underline{y} = \tau_1' \underline{c}_1' + \tau_2' \underline{c}_2' + \tau_3' \underline{c}_3'.$$

Eftersom

$$\underline{c}_1' = 0\underline{c}_1 + 1\underline{c}_2 - \frac{1}{2}\underline{c}_3, \quad \underline{c}_2' = 1\underline{c}_1 + 1\underline{c}_2 + \frac{1}{2}\underline{c}_3, \quad \underline{c}_3' = \frac{2}{5}\underline{c}_1 - \frac{4}{5}\underline{c}_2 - \frac{2}{5}\underline{c}_3,$$

(prøv selv efter!), er

$$\underline{y} = \tau_1'(0\underline{c}_1 + 1\underline{c}_2 - \frac{1}{2}\underline{c}_3) + \tau_2'(1\underline{c}_1 + 1\underline{c}_2 + \frac{1}{2}\underline{c}_3) + \tau_3'(\frac{2}{5}\underline{c}_1 - \frac{4}{5}\underline{c}_2 - \frac{2}{5}\underline{c}_3) =$$

$$(0\tau_1' + 1\tau_2' + \frac{2}{5}\tau_3')\underline{c}_1 + (1\tau_1' + 1\tau_2' - \frac{4}{5}\tau_3')\underline{c}_2 + (-\frac{1}{2}\tau_1' + \frac{1}{2}\tau_2' - \frac{4}{5}\tau_3')\underline{c}_3.$$

Ved at sammenholde dette med $\underline{y}$'s oprindelige fremstilling i $\underline{c}$-basen finder vi

$$\begin{pmatrix} \tau_1 \\ \tau_2 \\ \tau_3 \end{pmatrix} = \begin{pmatrix} 0 & 1 & \frac{2}{5} \\ 1 & 1 & -\frac{4}{5} \\ -\frac{1}{2} & \frac{1}{2} & -\frac{4}{5} \end{pmatrix} \begin{pmatrix} \tau_1' \\ \tau_2' \\ \tau_3' \end{pmatrix} = T \begin{pmatrix} \tau_1' \\ \tau_2' \\ \tau_3' \end{pmatrix}.$$

Igen var resultatet af anstrengelserne en oversættelse af basisskiftet i $\mathbb{R}^3$ til en angivelse af, at de *gamle* koordinater er lig T gange de *nye* koordinater, altså en sammenhæng af samme art som i $\mathbb{R}^2$. Atter er koordinatskiftsmatricen, her T, *invertibel* (ses f.eks. af at dens determinant er $\frac{4}{5} \neq 0$). De nye koordinater er udtrykt ved de gamle: $\underline{\tau}' = T^{-1}\underline{\tau}$.

Ved hjælp af koordinatskiftsmatricerne i de to rum kan vi nu angive F's matrixfremstilling i de nye baser. Vi havde jo i de oprindelige baser

$$\underline{\tau} = \begin{pmatrix} 3 & -1 \\ 0 & 2 \\ 2 & -3 \end{pmatrix} \underline{\xi}.$$

Heri skal vi blot indsætte

$$\underline{\tau} = T\underline{\tau}' \quad \text{og} \quad \underline{\xi} = S \begin{pmatrix} \xi_1' \\ \xi_2' \end{pmatrix},$$

hvilket giver

$$T\underline{\tau}' = \begin{pmatrix} 3 & -1 \\ 0 & 2 \\ 2 & -3 \end{pmatrix} S\underline{\xi}'.$$

Multipliceres denne identitet på begge sider fra venstre med T^{-1}, fremkommer F's matrixfremstilling i de nye baser:

$$\underline{\tau}' = T^{-1} \begin{pmatrix} 3 & -1 \\ 0 & 2 \\ 2 & -3 \end{pmatrix} S\underline{\xi}',$$

der ved indsættelse af T^{-1} og S giver

$$\underline{\tau}' = \begin{pmatrix} -\frac{1}{4} & \frac{5}{8} & -\frac{3}{4} \\ \frac{3}{4} & \frac{1}{8} & \frac{1}{4} \\ \frac{5}{8} & -\frac{5}{16} & -\frac{5}{8} \end{pmatrix} \begin{pmatrix} 3 & -1 \\ 0 & 2 \\ 2 & -3 \end{pmatrix} \begin{pmatrix} \frac{2}{3} & -\frac{1}{3} \\ -\frac{4}{3} & \frac{8}{3} \end{pmatrix} \underline{\xi}' =$$

$$\begin{pmatrix} -\frac{20}{3} & \frac{65}{6} \\ \frac{7}{2} & -\frac{17}{4} \\ -\frac{5}{12} & \frac{55}{24} \end{pmatrix} \underline{\xi}'.*$$

I virkeligheden er det mere besværligt at gennemføre betragtninger som dem i eksemplet med konkrete tal (I skal ikke tro at jeg har nydt udregningerne !), fordi de bærende idéer camoufleres af taljongleringen. Her skal I se, hvordan man griber sagen an i den generelle situation.

Lad os gå ud fra et der i $\mathbb{R}^n$ er givet en basis $\underline{b}_1, \underline{b}_2, \ldots, \underline{b}_n$. Den kalder vi den "gamle" basis. Enhver vektor $\underline{x}$ i $\mathbb{R}^n$ har så en entydigt bestemt koordinatfremstilling i denne basis:

$$\underline{x} = \xi_1 \underline{b}_1 + \xi_2 \underline{b}_2 + \ldots + \xi_n \underline{b}_n$$

Derefter forestiller vi os indført en "ny" basis i $\mathbb{R}^n$, $\underline{b}'_1, \underline{b}'_2, \ldots, \underline{b}'_n$. I forhold til denne har $\underline{x}$ koordinatfremstillingen

$$\underline{x} = \xi'_1 \underline{b}'_1 + \xi'_2 \underline{b}'_2 + \ldots + \xi'_n \underline{b}'_n.$$

De nye basisvektorer er alle linearkombinationer af de gamle:

$$\begin{aligned} \underline{b}'_1 &= s_{11}\underline{b}_1 + s_{21}\underline{b}_2 + \ldots + s_{n1}\underline{b}_n, \\ \underline{b}'_2 &= s_{12}\underline{b}_1 + s_{22}\underline{b}_2 + \ldots + s_{n2}\underline{b}_n, \\ &\vdots \\ \underline{b}'_n &= s_{1n}\underline{b}_1 + s_{2n}\underline{b}_2 + \ldots + s_{nn}\underline{b}_n. \end{aligned}$$

De indsættes nu i $\underline{x}$'s b'-fremstilling:

$$\begin{aligned} \underline{x} = &\xi'_1(s_{11}\underline{b}_1 + s_{21}\underline{b}_2 + \ldots + s_{n1}\underline{b}_n) \\ &+\xi'_2(s_{12}\underline{b}_1 + s_{22}\underline{b}_2 + \ldots + s_{n2}\underline{b}_n) \\ &\vdots \\ &+\xi'_n(s_{1n}\underline{b}_1 + s_{2n}\underline{b}_2 + \ldots + s_{nn}\underline{b}_n), \end{aligned}$$

som ved omordning efter $\underline{b}$-erne bliver til

$$\begin{aligned} &(s_{11}\xi'_1 + s_{12}\xi'_2 + \ldots + s_{1n}\xi'_n)\underline{b}_1 \\ &+(s_{21}\xi'_1 + s_{22}\xi'_2 + \ldots + s_{2n}\xi'_n)\underline{b}_2 \\ &\vdots \\ &+(s_{n1}\xi'_1 + s_{n2}\xi'_2 + \ldots + s_{nn}\xi'_n)\underline{b}_n. \end{aligned}$$

Sammenholdes dette med $\underline{x}$'s oprindelige fremstilling i $\underline{b}$-basen slutter vi at

$$\begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{pmatrix} = \begin{pmatrix} s_{11}\xi_1' + s_{12}\xi_2' + \ldots + s_{1n}\xi_n' \\ s_{21}\xi_1' + s_{22}\xi_2' + \ldots + s_{2n}\xi_n' \\ \vdots \\ s_{n1}\xi_1' + s_{n2}\xi_2' + \ldots + s_{nn}\xi_n' \end{pmatrix} = \begin{pmatrix} s_{11} & s_{12} & \ldots & s_{1n} \\ s_{21} & s_{22} & \ldots & s_{2n} \\ \vdots & \vdots & \vdots & \\ s_{n1} & s_{n2} & \ldots & s_{nn} \end{pmatrix} \begin{pmatrix} \xi_1' \\ \xi_2' \\ \vdots \\ \xi_n' \end{pmatrix}.$$

Hermed har vi opnået at forbinde de *gamle* og *nye* koordinater i $\mathbb{R}^n$ ved hjælp af **koordinatskiftsmatricen** S:

$$\underline{\xi} = S\underline{\xi}'.$$

Matricen S er matrix for en lineær afbildning af $\mathbb{R}^n$ ind i sig selv. Sådan som S er dannet, består *den j'te søjle i S af koordinaterne til den j'te nye basisvektor, $\underline{b}_j'$, i den gamle basis.* Det betyder, at S er matricen for den lineære afbildning af $\mathbb{R}^n$ ind i sig selv - forsynet med den oprindelige basis - der afbilder den gamle basis på den nye.

Matricen S *er invertibel.* Thi hvis $S\underline{\xi}' = \underline{0}$, er $\underline{\xi} = \underline{0}$ koordinaterne i $\underline{b}$-basen for $\underline{0}$. Sættet $\underline{\xi}'$ er så koordinaterne for $\underline{0}$ i den nye basis. Men da nulvektoren har de samme koordinater i alle baser, er $\underline{\xi}' = 0$. Dette viser, at S er matrix for en injektiv afbildning af $\mathbb{R}^n$ ind i sig selv. Denne afbildning er så bijektiv, hvorfor S er invertibel. *Ethvert basisskift er altså en bijektiv lineær afbildning af talrummet på sig selv.*

Enhver invertibel $n \times n$-matrix S er matrix for et basisskift i $\mathbb{R}^n$. Hvis nemlig $\underline{b}_1, \underline{b}_2, \ldots, \underline{b}_n$ er en given basis for $\mathbb{R}^n$, beskriver søjlerne i S koordinaterne i denne basis for n vektorer $\underline{b}_1', \underline{b}_2', \ldots, \underline{b}_n'$. Disse vektorer er lineært uafhængige - og dermed en basis for $\mathbb{R}^n$. Thi er

$$\begin{aligned} \underline{0} =& \alpha_1\underline{b}_1' + \alpha_2\underline{b}_2' + \ldots + \alpha_n\underline{b}_n' \\ =& \alpha_1(s_{11}\underline{b}_1 + s_{21}\underline{b}_2 + \ldots + s_{n1}\underline{b}_n) \\ &+\alpha_2(s_{12}\underline{b}_1 + s_{22}\underline{b}_2 + \ldots + s_{n2}\underline{b}_n) \\ &\vdots \\ &+\alpha_n(s_{1n}\underline{b}_1 + s_{2n}\underline{b}_2 + \ldots + s_{nn}\underline{b}_n), \end{aligned}$$

der ved omordning efter $\underline{b}$-erne giver

$$\begin{aligned} &(s_{11}\alpha_1 + s_{12}\alpha_2 + \ldots + s_{1n}\alpha_n)\underline{b}_1 \\ +&(s_{21}\alpha_1 + s_{22}\alpha_2 + \ldots + s_{2n}\alpha_n)\underline{b}_2 \\ &\vdots \\ +&(s_{n1}\alpha_1 + s_{n2}\alpha_2 + \ldots + s_{nn}\alpha_n)\underline{b}_n. \end{aligned}$$

Dette viser, at

$$\underline{0} = S\underline{\alpha},$$

og da søjlerne i S er lineært uafhængige fordi S er forudsat invertibel, må samtlige α'er være 0. Dette viser, at $\underline{b}'$erne er lineært uafhængige.

Ud fra ræsonnementerne ovenfor kan vi altså konkludere, at *de $n \times n$-matricer S der beskriver et basisskift i $\mathbb{R}^n$ er netop de invertible matricer.*

Nu kan vi undersøge hvad der sker med en lineær afbildnings matrixfremstilling, hvis der skiftes baser i begge de rum den arbejder i.

Vi forestiller os givet $F : \mathbb{R}^n \to \mathbb{R}^p$, hvor $\mathbb{R}^n$ og $\mathbb{R}^p$ er forsynet med givne baser. I forhold til disse baser har F en matrixfremstilling bestemt ved $p \times n$-matricen A:

$$\underline{\tau} = A\underline{\xi}.$$

Nu skiftes baser i både $\mathbb{R}^n$ og $\mathbb{R}^p$. Overgangen til den nye basis i $\mathbb{R}^n$ er beskrevet ved $\underline{\xi} = S\underline{\xi}'$, hvor som før $\underline{\xi}'$ er koordinaterne i den nye basis. På tilsvarende måde er overgangen til den nye basis beskrevet ved $\underline{\tau} = T\underline{\tau}'$. Matrixfremstillingen for F i de nye baser fremkommer så ved i den oprindelige fremstilling at indsætte $\underline{\tau}$ udtrykt ved $\underline{\tau}'$, og $\underline{\xi}$ ved $\underline{\xi}'$

$$T\underline{\tau}' = A(S\underline{\xi}'),$$

hvoraf den nye matrixfremstillingen skaffes ved multiplikation med T^{-1} fra venstre:

$$\underline{\tau}' = T^{-1}AS\underline{\xi}'.$$

Dette viser at F's matrix i de nye baser er $A' = T^{-1}AS$.

I det specialtilfælde som egentlig interesserer os mest her, er $n = p$, dvs. F afbilder $\mathbb{R}^n$ ind i sig selv, og det er ét og samme basiskift der foregår i "begge" rum. Det bevirker, at $T = S$, så at vi i dette tilfælde får $A' = S^{-1}AS$. Desuden er både A og A' kvadratiske matricer.

$*$ I denne situation har både A og A' determinanter, og der gælder

$$\begin{aligned} detA' &= det(S^{-1}AS) = detS^{-1}detA\,detS \\ &= detS^{-1}detS\,detA = det(S^{-1}\cdot S)detA = detE_n\,detA \\ &= detA. \end{aligned}$$

Vi ser altså at enhver matrix for F har samme determinant. Dette viser at determinanten er knyttet til den lineære afbildning selv, ikke blot til den "tilfældige" matrixfremstilling afbildningen måtte være iklædt.

Egenværdier og egenvektorer

I et tidligere afsnit omtaltes en særligt venligsindet slags lineære afbildninger af et talrum ind i sig selv, de såkaldte proportionaliteter. De er givet ved $F(\underline{x}) = \alpha\underline{x}$ for alle $\underline{x} \in \mathbb{R}^n$. Som vi ved er ikke alle lineære afbildninger af et talrum ind i sig selv af denne type. Alligevel kan det godt være at en afbildning der ikke er en proportionalitet i *hele* rummet kan være det i *dele* af rummet. Denne mulighed vil blive nærmere undersøgt i dette afsnit. Alle de afbildninger der optræder er lineære afbildninger $F : \mathbb{R}^n \to \mathbb{R}^n$.

Først et par begreber. Vi siger at en *egentlig* vektor $\underline{x}$, altså en vektor $\neq 0$, er en **egenvektor** for F, hvis der findes en skalar λ, så at

$$F(\underline{x}) = \lambda\underline{x}.$$

Skalaren λ kaldes den til $\underline{x}$ hørende **egenværdi**.

Hvis $\underline{x}$ er en egenvektor er ethvert multiplum $\alpha\underline{x}$ også en egenvektor med den samme egenværdi: $F(\alpha\underline{x}) = \alpha F(\underline{x}) = \alpha(\lambda\underline{x}) = \lambda(\alpha\underline{x})$. Det betyder altså at F virker som en proportionalitet på underrummet udspændt af $\underline{x}$. Det forhindrer ikke at F kan have andre egenvektorer svarende til samme egenværdi, som ikke er et multiplum af $\underline{x}$. Mængden af alle egenvektorer hørende til en given egenværdi λ

$$E_\lambda = \{\underline{x} | F(\underline{x}) = \lambda\underline{x}\}$$

kaldes det til λ hørende **egenrum**. Berettigelsen af betegnelsen "rum" ligger i at *ethvert egenrum er et underrum i* $\mathbb{R}^n$. Thi hvis $\underline{x} \in E_\lambda$ og $\underline{x}' \in E_\lambda$ gælder det samme for $\alpha\underline{x} + \alpha'\underline{x}'$, idet

$$\begin{aligned} F(\alpha\underline{x} + \alpha'\underline{x}') &= \alpha F(\underline{x}) + \alpha' F(\underline{x}') = \alpha(\lambda\underline{x}) + \alpha'(\lambda\underline{x}') \\ &= \lambda(\alpha\underline{x} + \alpha'\underline{x}'). \end{aligned}$$

Dimensionen af egenrummet hørende til egenværdien λ kaldes λ's **geometriske multiplicitet**.

Som det fremgår opererer vi kun med egentlige egenvektorer. Der er jo ikke noget (nyt) interessant i at $F(\underline{0}) = \lambda\underline{0}$ for alle λ. Derimod kan 0 godt være en egenværdi, nemlig hvis der findes en egentlig vektor $\underline{x}$ så at $F(\underline{x}) = \underline{0}$. Dette er tydeligvis ensbetydende med at nulrummet for F består af mere end blot nulvektoren, hvilket på sin side er ensbetydende med at F ikke er injektiv.

SÆTNING. *Egenvektorer fra forskellige egenrum, dvs. svarende til forskellige egenværdier, er lineært uafhængige.*

BEVIS: Denne påstand kræver et lille bevis. Lad $\lambda_1, \lambda_2, \ldots, \lambda_k$ være forskellige egenværdier for F. Er så $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k$ egenvektorer svarende til hver af disse egenvektorer,

er $\underline{x}$'erne lineært uafhængige. Ellers findes et største antal j, $j < k$, lineært uafhængige blandt dem. Vi kan antage nummereringen foretaget sådan at de j første, $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_j$, er lineært uafhængige. Så vil $\underline{x}_{j+1}$ tilhøre $span\{\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_j\}$:

$$\underline{x}_{j+1} = \alpha_1 \underline{x}_1 + \alpha_2 \underline{x}_2 + \ldots + \alpha_k \underline{x}_j,$$

hvor ikke alle α'erne er 0, da egenvektorer er egentlige vektorer. Så benyttes F på begge sider:

$$\begin{aligned}\lambda_{j+1}\underline{x}_{j+1} &= F(\underline{x}_{j+1}) = F(\alpha_1 \underline{x}_1 + \alpha_2 \underline{x}_2 + \ldots + \alpha_k \underline{x}_j) \\ &= \alpha_1 F(\underline{x}_1) + \alpha_2 F(\underline{x}_2) + \ldots + \alpha_k F(\underline{x}_j) \\ &= \alpha_1 \lambda_1 \underline{x}_1 + \alpha_2 \lambda_2 \underline{x}_2 + \ldots + \alpha_j \lambda_j \underline{x}_j.\end{aligned}$$

Hvis $\lambda_{j+1} = 0$ har vi at gøre med en ikke-triviel fremstilling af $\underline{0}$ (ikke alle $\alpha_i \lambda_i$ kan være 0, da λ'erne er forskellige og ikke alle α er 0) som linearkombination af de lineært uafhængige $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_j$. Dette er ikke muligt. Hvis på den anden side $\lambda_{j+1} \neq 0$, ville vi ved division opnå

$$\underline{x}_{j+1} = \alpha_1 \frac{\lambda_1}{\lambda_{j+1}} \underline{x}_1 + \alpha_2 \frac{\lambda_2}{\lambda_{j+1}} \underline{x}_2 + \ldots + \alpha_j \frac{\lambda_j}{\lambda_{j+1}} \underline{x}_j,$$

der giver en fremstilling af $\underline{x}_{j+1}$ som linearkombination af $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_j$. Denne fremstilling er forskellig fra den første. For λ'erne er indbyrdes forskellige, sådan at brøkerne $\frac{\lambda_1}{\lambda_{j+1}}, \frac{\lambda_2}{\lambda_{j+1}}, \ldots, \frac{\lambda_j}{\lambda_{j+1}}$ alle er forskellige fra 1, samtidig med at ikke alle α er 0. Vi har altså to forskellige fremstillinger af vektoren $\underline{x}_{j+1}$ i en basis for $span\{\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_j\}$. Heller ikke det er muligt. Modstriderne kom af antagelsen om at ikke hele sættet af de betragtede egenvektorer var lineært uafhængige. Det må de derfor være.

Hvordan afgør vi nu om en lineær afbildning $F : \mathbb{R}^n \to \mathbb{R}^n$ har nogen egenvektorer? Hvis A er F's matrix i *grundbasen* i $\mathbb{R}^n$, er F givet ved $F(\underline{x}) = A\underline{x}$. At F har en egenvektor kommer så ud på at der findes et tal λ så at ligningssystemet $A\underline{x} = \lambda\underline{x}$ har en ikke-triviel løsning $\underline{x}$. Dette er det samme som at det homogene ligningssystem $(A - \lambda E)\underline{x} = \underline{0}$ har en ikke triviel-løsning. Men det er videre ensbetydende med at matricen $A - \lambda E$ ikke er invertibel, hvilket slutteligt er ækvivalent med at $det(A - \lambda E) = 0$. Summa summarum: *λ er en egenværdi for A hvis og kun hvis*

$$det(A - \lambda E) = 0.$$

Lad os lige se lidt nærmere på $det(A - \lambda E)$. Vi har

$$det(A - \lambda E) = det \begin{pmatrix} a_{11} - \lambda & a_{12} & a_{13} & \ldots & a_{1n} \\ a_{21} & a_{22} - \lambda & a_{23} & \ldots & a_{2n} \\ a_{31} & a_{32} & a_{33} - \lambda & \ldots & a_{3n} \\ \vdots & \vdots & \vdots & & \\ a_{n1} & a_{n2} & a_{n3} & \ldots & a_{nn} - \lambda \end{pmatrix}.$$

Benytter man definitionen på en determinant af en $n \times n$- matrix, kan man indse at $det(A - \lambda E)$ er et *n'te grads polynomium i* λ. Vi afstår for at bevise dette her (det kan gøres ved induktion). Dette polynomium kaldes **det karakteristiske polynomium** for F, eller blot for A. Vi har altså at F har egenvektorer hvis og kun hvis det karakteristiske polynomium for F har reelle rødder. Og λ er en egenværdi for F netop hvis λ er rod i det karakteristiske polynomium. Antallet af gange λ er rod i det karakteristiske polynomium kaldes λ's **algebraiske multiplicitet**. Eftersom et n'te grads polynomium højst har n reelle rødder, har F *højst n egenværdier*. Dette stemmer også overens med at egenvektorerne hørende til forskellige egenværdier er lineært uafhængige, sammenholdt med at $dim\mathbb{R}^n = n$.

Hvis λ er en egenværdi for F er de tilhørende egenvektorer bestemt som løsningerne til ligningssystemet

$$(A - \lambda E)\underline{x} = \underline{0}.$$

Eksempel 1

Lad $F : \mathbb{R}^2 \to \mathbb{R}^2$ være givet ved

$$F(\underline{x}) = \begin{pmatrix} 2 & -2 \\ -3 & 1 \end{pmatrix} \underline{x}.$$

Så er det karakteristiske polynomium

$$det \begin{pmatrix} 2-\lambda & -2 \\ -3 & 1-\lambda \end{pmatrix} = (2-\lambda)(1-\lambda) - 6 = \lambda^2 - 3\lambda - 4.$$

Det har rødderne $\lambda = 4$, og $\lambda = -1$, så at F har egenværdierne 4 og -1.

For at finde de tilhørende egenvektorer skal vi løse ligningssystemerne

$$\begin{pmatrix} 2-4 & -2 \\ -3 & 1-4 \end{pmatrix} \underline{x} = \underline{0} \quad \text{og} \quad \begin{pmatrix} 2+1 & -2 \\ -3 & 1+1 \end{pmatrix} \underline{x} = \underline{0},$$

der har løsningerne henholdsvis

$$t \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad \text{og} \quad u \begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

for reelle t og u. Det betyder at $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ er en basis for egenrummet E_4, mens $\begin{pmatrix} 2 \\ 3 \end{pmatrix}$ er en basis for egenrummet E_{-1}. I de to egenrum virker F som en proportionalitet, i E_4 med proportionalitetsfaktoren 4, i E_{-1} med proportionalitetsfaktoren -1. Egenvektorer fra de to rum er, i følge de generelle overvejelser ovenfor, lineært uafhængige. Vælger

vi $\underline{b}_1' = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ og $\underline{b}_2' = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$ udgør de en ny basis for $\mathbb{R}^2$. Eftersom $F(\alpha_1 \underline{b}_1' + \alpha_2 b_2') = \alpha_1 4 \underline{b}_1 + \alpha_2(-1)\underline{b}_2$, har F i denne basis matricen

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

Det fremgår altså at ved at overgå til en "naturlig" basis for F, nemlig dens proportionalitetsretninger, får F en meget mere håndterlig matrixfremstilling.

Eksempel 2

Afbildningen $F : \mathbb{R}^2 \to \mathbb{R}^2$, defineret ved

$$F(\underline{x}) = \begin{pmatrix} 2 & -1 \\ 3 & 1 \end{pmatrix},$$

har det karakteristiske polynomium $(2-\lambda)(1-\lambda)+3 = \lambda^2 - 3\lambda + 5$, der har diskriminanten -11 og derfor ingen reelle rødder. F har derfor ingen reelle egenværdier og ingen egenvektorer. Der er altså ingen retninger hvor F virker som en proportionalitet. $*$

Dette var altså et eksempel på en lineær afbildning der *ikke har nogen egenvektorer.* I modsætning hertil står proportionaliteterne, $F(\underline{x}) = \lambda \underline{x}$, der har *alle vektorer som egenvektorer* svarende til én og samme egenværdi, λ. Det næste eksempel giver et nyt indblik i hvad der kan ske.

Eksempel 3

Afbildningen $F : \mathbb{R}^3 \to \mathbb{R}^3$ givet ved

$$F(\underline{x}) = A\underline{x} = \begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{pmatrix} \underline{x}$$

har det karakteristiske polynomium

$$det \begin{pmatrix} 1-\lambda & 2 & 2 \\ 2 & 1-\lambda & 2 \\ 2 & 2 & 1-\lambda \end{pmatrix} = -\lambda^3 + 3\lambda^2 + 9\lambda + 5$$

(udfør selv mellemregningerne). For at gøre det lettere at finde rødderne gør vi den observation, at enhver vektor af formen $\underline{x} = (x, x, x)$ er egenvektor til F, idet $F(x,x,x) = 5(x,x,x)$. Den tilsvarende egenværdi er altså $\lambda = 5$, som altså er rod i det karakteristiske

polynomium. Det betyder at faktoren $(\lambda-5)$ kan "sættes uden for" i det karakteristiske polynomium ved hjælp af polynomiers division. Derved finder vi $-\lambda^3+3\lambda^2+9\lambda+5 = (\lambda-5)(-\lambda^2-2\lambda-1) = -(\lambda-5)(\lambda+1)^2$, sådan at F også har egenværdien $\lambda=-1$ med algebraisk multiplicitet 2. Nu skal vi finde egenvektorerne svarende til de to egenværdier (vi har allerede fundet nogle til $\lambda=5$, men måske kunne der være flere?). Vi skal løse ligningssystemerne

$$\begin{pmatrix} -4 & 2 & 2 \\ 2 & -4 & 2 \\ 2 & 2 & -4 \end{pmatrix} \underline{x} = \underline{0}, \quad \text{og} \quad \begin{pmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{pmatrix} \underline{x} = \underline{0}.$$

Ved løsning af disse systemer viser det sig at egenrummet $E_5 = span\{(1,1,1)\}$, mens $E_{-1} = span\{(-1,1,0);(-1,0,1)\}$. Det sidste viser, at den geometriske multiplicitet af $\lambda=-1$ er 2, lige som den algebraiske. Kalder vi $\underline{b}_1' = (1,1,1)$, $\underline{b}_2' = (-1,1,0)$ og $\underline{b}_3' = (-1,0,1)$ har vi øjensynlig at gøre med en *basis af egenvektorer* for $\mathbb{R}^3$. Har $\underline{x}$ i denne basis koordinaterne (x_1', x_2', x_3'), altså $\underline{x} = x_1'\underline{b}_1' + x_2'\underline{b}_2' + x_3'\underline{b}_3'$, er

$$\begin{aligned} F(\underline{x}) &= F(x_1'\underline{b}_1' + x_2'\underline{b}_2' + x_3'\underline{b}_3') \\ &= x_1'F(\underline{b}_1') + x_2'F(\underline{b}_2') + x_3'F(\underline{b}_3') \\ &= x_1'5\underline{b}_1' + x_2'(-1)\underline{b}_2' + x_3'(-1)\underline{b}_3', \end{aligned}$$

hvilket viser at F i egenvektorbasen har matricen

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

Overgangen $\underline{x} = S\underline{x}'$ fra den oprindelige basis af grundvektorer til den nye basis af egenvektorer sker ved hjælp af koordinatskiftsmatricen

$$S = \begin{pmatrix} 1 & -1 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

Læg mærke til at den første søjle i S er en egenvektor for egenværdien $\lambda=5$, mens de to næste søjler er lineært uafhængige egenvektorer for egenværdien $\lambda=-1$. Dette er ikke nogen tilfældighed, jf. nedenfor. $*$

Øvelse Undersøg hvad der sker med A' hvis i eksemplet S møbleres af egenvektorerne i en anden rækkefølge, eller af andre egenvektorer for de samme egenværdier.

I betragtningerne om egenværdier og egenvektorer er vi gået ud fra en lineær afbildning F med en given matrix A i en given basis. Hvad nu hvis vi i stedet havde F givet ved dens matrixfremstilling i en anden basis, sådan at matricen havde været en

anden, A'? Finder vi så andre egenværdier og egenvektorer? På forhånd må svaret naturligvis være nej. Egenværdier og egenvektorer knytter sig til selve afbildningen og ikke til den "tilfældige" basis den måtte være givet i. Men så må dette forhold afspejles i matrixregneriet. Det gør det da også. Hvis nemlig T er et vilkårligt basisskift i $\mathbb{R}^n$ (ikke nødvendigvis ét der har med egenvektorer at gøre) gælder om det karakteristiske polynomium for A':

$$\begin{aligned} det(A' - \lambda E) &= det(T^{-1}AT - \lambda E) = det(T^{-1}AT - \lambda T^{-1}T) \\ &= det(T^{-1}AT - T^{-1}(\lambda E)T) = det(T^{-1}(A - \lambda E)T) \\ &= detT^{-1}det(A - \lambda E)detT = det(A - \lambda E)detT^{-1}detT \\ &= det(A - \lambda E). \end{aligned}$$

(Hvis der er et af flere af lighedstegnene du er i tvivl om, så regn fra højre mod venstre.) Resultatet viser at F har *det samme karakteristiske polynomium i enhver basis.* Egenværdierne bestemmes altså lige så vel af den ene matrixfremstilling for F som den anden.

Gennem eksemplerne ovenfor har vi set at det nu og da er muligt at skifte til en basis af egenvektorer for en givet lineær afbildning F. Derved fik F en særligt simpel matrixfremstilling, nemlig en diagonalmatrix, hvor diagonalen er møbleret af egenværdierne for afbildningen. I andre tilfælde havde den betragtede afbildning ingen egenvektorer og egenværdier, og ingen diagonalmatrixfremstilling bød sig til. Vi vil nu afrunde afsnittet med en kort omtale af dette spørgsmål i almindelighed.

Vi siger, at en lineær afbildning $F : \mathbb{R}^n \to \mathbb{R}^n$ kan **diagonaliseres**, hvis der findes en basis for $\mathbb{R}^n$ hvori F kan fremstilles ved en diagonalmatrix. Først og fremmest konstaterer vi, at *hvis F kan diagonaliseres findes en basis for $\mathbb{R}^n$ bestående af egenvektorer.* Lad nemlig F have matricen

$$\begin{pmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ 0 & 0 & \lambda_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & 0 \\ 0 & 0 & 0 & \dots & \lambda_n \end{pmatrix},$$

i basen $\underline{b}'_1, \underline{b}'_2, \dots, \underline{b}'_n$. Så er billedet af $\underline{b}'_1$ - der jo har koordinaterne $(1, 0, \dots, 0)$ i denne basis - lig $\lambda_1 \underline{b}'_1$, dvs. en egenvektor hørende til egenværdien λ_1. På tilsvarende måde er $\underline{b}'_2, \dots, \underline{b}'_2$ egenvektorer for henholdsvis $\lambda_2, \dots \lambda_n$. Det ses altså, at $\underline{b}'$-erne simpelthen er en basis af egenvektorer.

Hvis omvendt $F : \mathbb{R}^n \to \mathbb{R}^n$ har et sæt egenvektorer der udgør en basis for $\mathbb{R}^n$, kan F diagonaliseres. Thi lad $\underline{b}'_1, \underline{b}'_2, \dots, \underline{b}'_n$ være en basis for $\mathbb{R}^n$ af egenvektorer for F, hørende til egenværdierne $\lambda_1, \lambda_2, \dots, \lambda_n$, hvor λ-erne kan være forskellige eller nogle

af dem ens, som det nu falder. Har vektoren $\underline{x}$ i denne basis koordinatfremstillingen $\underline{x} = x_1'\underline{b}_1' + x_2'\underline{b}_2' + \ldots + x_n'\underline{b}_n'$, gælder

$$\begin{aligned} F(\underline{x}) &= F(x_1'\underline{b}_1' + x_2'\underline{b}_2' + \ldots + x_n'\underline{b}_n') = x_1'F(\underline{b}_1') + x_2'F(\underline{b}_2') + \ldots + x_n'F(\underline{b}_n') \\ &= x_1'\lambda_1\underline{b}_1' + x_2'\lambda_2\underline{b}_2' + \ldots + x_n'\lambda_n\underline{b}_n'. \end{aligned}$$

Kaldes koordinaterne for $F(\underline{x})$ for $(y_1', y_2', \ldots, y_n')$ fremgår det, at F's matrixfremstilling i $\underline{b}'$-basen er

$$\begin{pmatrix} y_1' \\ y_2' \\ \vdots \\ y_n' \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 & 0 & \ldots & 0 \\ 0 & \lambda_2 & 0 & \ldots & 0 \\ 0 & 0 & \lambda_3 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ldots & 0 \\ 0 & 0 & 0 & \ldots & \lambda_n \end{pmatrix} \begin{pmatrix} x_1' \\ x_2' \\ \vdots \\ x_n' \end{pmatrix},$$

hvilket netop er en diagonalfremstilling af F.

Vi har dermed overbevist os om at F kan diagonaliseres hvis og kun hvis $\mathbb{R}^n$ har *en basis af egenvektorer for F. Dette er ensbetydende med at der findes n lineært uafhængige egenvektorer for F.*

Hvis F kan diagonaliseres, men ikke i den oprindelige basis har en diagonalmatrix, sker overgangen til den nye basis som sædvanlig ved

$$\underline{x} = S\underline{x}',$$

hvor søjlerne i S er koordinaterne for de nye basisvektorer, *altså egenvektorerne*, i den gamle basis. Rækkefølgen af vektorerne i S bestemmer rækkefølgen af egenværdier i den resulterende diagonalmatrix.

Det vil føre for vidt i denne fremstilling at give en udtømmende diskussion af hvilke lineære afbildninger, eller anderledes sagt: hvilke kvadratiske matricer, der kan diagonaliseres. Et par resultater følger dog af den foranstående udvikling af teorien:

∗ *Hvis F opfylder at summen af de geometriske multipliciteter for dens forskellige egenværdier (altså summen af egenrummenes dimensioner) er lig n, kan F diagonaliseres.* Da egenvektorer hørende til forskellige egenværdier er lineært uafhængige, findes nemlig i denne situation n lineært uafhængige egenvektorer for F. I kraft af deres antal udgør de en basis for $\mathbb{R}^n$, og F kan diagonaliseres.

Et vigtigt specialtilfælde heraf har vi i

∗ *Hvis F har n forskellige egenværdier kan F diagonaliseres.* Thi egenvektorerne hørende til de n forskellige egenværdier er lineært uafhængige og dermed en basis for $\mathbb{R}^n$, sådan at F kan diagonaliseres.

$*$ En **symmetrisk** matrix er en kvadratisk matrix der er lig sin transponerede, dvs. elementet på enhver plads (i,j) er lig elementet på plads (j,i). *Enhver symmetrisk matrix kan diagonaliseres.* Beviset for dette udelades.

Vi slutter diskussionen med et eksempel der viser eksistensen af afbildninger der nok har egenvektorer, men alligevel ikke kan diagonaliseres.

Eksempel 4

Hvis $F : \mathbb{R}^2 \to \mathbb{R}^2$ i grundbasen er givet ved en matrix af formen

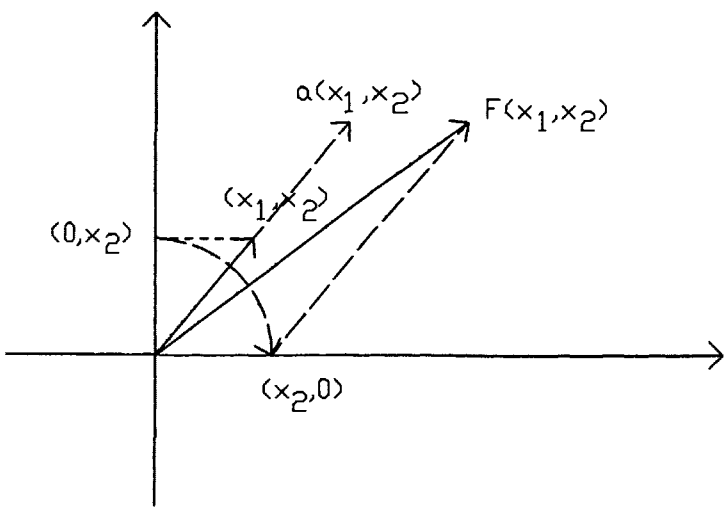
$$\begin{pmatrix} a & 1 \\ 0 & a \end{pmatrix},$$

er det klart at det karakteristiske polynomium er $(a-\lambda)^2$, der har $\lambda = a$ som dobbeltrod. F har altså kun den ene egenværdi a. De tilhørende egenvektorer - F's eneste - er løsningerne til ligningssystemet

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \underline{x} = \underline{0},$$

der øjensynlig er $\underline{x} = t\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\quad t \in \mathbb{R}$. Det er altså ikke muligt at skaffe en basis for $\mathbb{R}^2$ bestående af egenvektorer for F.

Vi kan forstå F's adfærd geometrisk. Der gælder jo, at $F(x_1, x_2) = (ax_1 + x_2, ax_2) = a(x_1, x_2) + (x_2, 0)$. Det viser at F afbilder en vektor på et multiplum af den *plus* vektoren $(x_2, 0)$, der ødelægger proportionalitetsbilledet. Kun for vektorer på x_1-aksen bliver denne vektor $\underline{0}$. $*$



LINEÆRE DIFFERENTIALLIGNINGER

Indledning

En stor del af kapitlerne om lineær algebra har direkte eller indirekte beskæftiget sig med lineære ligningssystemer. I et sådant system er de ubekendte et sæt af *reelle tal*, som skal tilfredsstille de lineære ligningsbånd systemet består af.

Vi skal nu behandle ligninger af en anden type, de såkaldte **differentialligninger** som udgør et af de matematiske emner som har størst betydning for matematikkens anvendelse i videnskabelige og praktiske forbindelser. I en *enkelt* differentialligning er den *ubekendte en funktion*. En sådan er altid defineret på en delmængde af $\mathbb{R}$, evt. på hele $\mathbb{R}$.

Der findes mange slags ligninger med funktioner som ubekendte uden at de af den grund er differentialligninger, f.eks. ligningen $f^2 - 2f + 1 = 0$ (hvor f er den ubekendte funktion). Vi har først at gøre med en differentialligning, hvis der er tale om *bånd der forbinder en ubekendt funktion med en eller flere af dens afledede*. Hvis der kun indgår afledede af første orden kaldes differentialligningen, lidet overraskende, for en **1. ordens differentialligning**. Hvis den højeste orden af afledede der indgår er n, kaldes differentialligningen for en **n'te ordens differentialligning**. Generelt er en differentialligning af n'te orden et bånd af typen

$$F(t, f, f', \ldots, f^{(n)}) = 0,$$

hvor t er den uafhængige variable i den søgte funktion f, og hvor som sædvanlig $f', f'', \ldots$ f^n er de afledede af f til og med orden n. Båndet er opskrevet som en funktion F af hele molevitten. Et helt vildt, tilfældigt og uvigtigt eksempel:

$$t^3 f(t) f''(t) - sin\frac{2t+1}{f'''(t)} - 1989 = 0.$$

Nogle gange ser man i stedet for $f, f', \ldots, f^n$ betegnelser som $x(t)$, $\frac{dx}{dt}$, $\frac{d^2x}{dt^2}, \ldots, \frac{d^nx}{dt^n}$, eller $x(t), x'(t), \ldots, x^{(n)}(t)$. Læg i de sidste tilfælde mærke til at x-et står for *en funktion* og ikke for den uafhængige variable, som altså er t. Det sidste valg skyldes at mange, men dog ikke alle, differentialligninger beskriver fænomener der udvikler sig i tiden, som så betragtes som den uafhængige variable. I denne fremstilling vil vi fortrinsvis benytte den notation hvor de ubekendte funktioner betegnes med x, y, z, x_1, x_2, osv., mens den uafhængige variable som nævnt betegnes med t.

Ved en **løsning** til en differentialligning forstås ganske enkelt en funktion med tilhørende definitionsmængde, som passer ind i differentialligningen. Nærmere bestemt skal

de løsningsfunktioner vi interesserer os for skal være defineret på et *åbent* interval i $\mathbb{R}$. Vi medregner både intervaller af formen $]a, \infty[$, $]-\infty, a[$ samt $\mathbb{R}$ selv. Når intervallet skal være åbent, er det fordi vi skal kunne differentiere funktionen (det relevante antal gange) i ethvert punkt i definitionsmængden, uden at tænke på om der mon kun kan gås til grænsen fra højre eller fra venstre. (Egentlig kan vi tænke på en differentialligning som et sæt af uendeligt mange ligninger med uendeligt mange ubekendte. Differentialligningen skal jo være opfyldt for alle t i et vist interval, det giver én ligning for hvert punkt t som den ubekendte funktion skal opfylde, og den ubekendte funktion skal have uendeligt mange værdier, én for hvert t).

Hidtil har vi kun omtalt en enkelt differentialligning. Men lige som vi i beskæftigelsen med sædvanlige ligninger, med tal som ubekendte, havde god grund til at studere ligning*ssystemer* hvor flere ubekendte bindes sammen i flere "samtidige" ligninger, har vi brug for at studere **differentialligningssystemer**. Et differentialligningssystem består af flere sammenhørende differentialligninger i flere ubekendte funktioner. En løsning er så et *sæt* af funktioner, hver defineret på et åbent interval, som til sammen passer ind i ligningerne. Et eksempel: Differentialligningssystemet af 1. orden,

$$x'(t) = ax(t)(1 - \frac{x(t) + y(t)}{K})$$
$$y'(t) = by(t)(1 - \frac{x(t) + y(t)}{L})$$

giver en generel model af dynamikken i to konkurrerende biologiske populationer, der til tid t har størrelserne henholdsvis $x(t)$ og $y(t)$. Tallene a, b, K, L er konstanter.*

Et andet eksempel på et differentialligningssystem har vi i:

$$C_v'(t) = -\alpha C_v(t) + \beta C_s(t) + \gamma$$
$$C_s'(t) = \delta C_v(t) - \eta C_s(t).$$

Dette ligningssystem danner den såkaldte *Lorenzen-model* for fosforomsætningen i en sø. $C_v(t)$ angiver gennemsnitskoncentrationen af fosfor i søvandet til tiden t, mens $C_s(t)$ angiver gennemsnitskoncentrationen af fosfor i sediment. Størrelserne α, β, γ, δ og η er positive konstanter.

Modellens *første* ligning udtrykker den antagelse, at ændringen af fosforkoncentrationen i søvandet er sammensat af tre delændringer: (a) en vis brøkdel af søvandets fosfor afgives til sedimentet, (b) en vis brøkdel af sedimentets fosfor frigøres til søvandet; endelig, (c), tilføres søvandet fra omgivelserne en konstant mængde fosfor pr. tidsenhed. Modellens *anden* ligning forestiller sig, at ændringen i sedimentkoncentrationen er sammensat af (a) en positiv tilvækst, hidrørende fra at en brøkdel af vandets fosfor overgår til sedimentet på genopløselig form, og en anden brøkdel bindes permanent i sedimentet, og (b) en negativ tilvækst der skyldes at en vis brøkdel af sedimentets fosfor frigøres

til vandet. Jeg er blevet opmærksom på modellen, fordi den er blevet behandlet i et 2. semesters projekt i mit basishus (02.1, 1990).*

Desværre er det kun en forsvindende brøkdel af verdens differentialligninger man kan løse. Det er dog muligt ad teoretisk vej at sige noget principielt og alment om løsningsforholdene for en større klasse af differentialligninger, end dem man kan løse eksplicit. Den samlede differentialligningsteori er uhyre kompliceret, mangesidet og righoldig på detailstudier. På grund af emnets vigtighed også for anvendelsesformål er teorien enorm. I dette kursus kan vi kun berøre et lille, om end meget centralt, udsnit af emnet. Dette udsnit falder inden for klassen af de såkaldte **lineære differentialligningssystemer**, hvor vi skal behandle den grundlæggende teori, der helt hviler på de resultater vi har opnået i lineær algebra. Lorenzen-modellens differentialligningssystem er et eksempel på et lineært system. Nå, men nu er det vist bedst at gå til håndgribeligheder.

Elementære grundtrin. Funktionsrum

Verdens simpleste differentialligning er $x'(t) = 0$. Den er et eksempel på en lineær differentialligning - et begreb vi endnu ikke kender betydningen af. Ligningen kaldes **homogen**, fordi højresiden er nulfunktionen. Dens løsninger består af alle konstante funktioner defineret på hele den reelle akse: $\{x : \mathbb{R} \to \mathbb{R} \mid x(t) = c,\ t \in \mathbb{R}, \text{ hvor } c \in \mathbb{R}\}$. En lidt mere kompliceret differentialligning har vi i $x'(t) = a(t)$, hvor a er en given reel funktion defineret på et åbent interval I. Lad os for nemheds skyld antage at a er kontinuert. Også denne differentialligning vil blive kaldt lineær, og **inhomogen** hvis a ikke er nulfunktionen. Ligningens løsninger består af samtlige stamfunktioner til a (de findes fordi a er kontinuert): $\{x : \mathbb{R} \to \mathbb{R} \mid x(t) = \int_{t_0}^{t} a(u)\,du + c,\ t \in \mathbb{R}, \text{ hvor } c \in \mathbb{R}\}$. Her er $\int_{t_0}^{t} a(u)\,du$ den stamfunktion til a der i t_0 har værdien 0. Et specialtilfælde har vi, hvis a er en konstant funktion. I så fald bliver løsningsmængden $\{x : \mathbb{R} \to \mathbb{R} \mid x(t) = at + c,\ t \in \mathbb{R}, \quad \text{hvor}, c \in \mathbb{R}\}$, idet vi har valgt $t_0 = 0$.

Inden vi går over til at behandle nogle lidt mere udfordrende problemstillinger vil vi gøre en observation der er væsentlig for det følgende. Observationen har allerede betydning i forhold til de ovenstående simple eksempler.

Vi betragter mængden af *samtlige* reelle funktioner defineret på et givet interval I i $\mathbb{R}$ (I må gerne være $\mathbb{R}$)

$$\mathbf{F}(I) = \{f \mid f : I \to \mathbb{R}\}$$

Sådanne funktioner kan adderes på sædvanlig måde, *punktvis* som vi siger: Summen af to funktioner på I, f og g, er defineret ved $(f + g)(t) = f(t) + g(t)$, $t \in I$. Nulfunktionen, altså den funktion der er konstant 0 på I, er øjensynlig neutral ved addition. Additionen er kommutativ og associativ, og enhver funktion f på I har en modsat funk-

tion nemlig $-f$. Dette viser at *addition af funktioner* på et interval opfylder præcis de samme betingelser som addition i talrummene $\mathbb{R}^n$.

Vi kan også multiplicere funktionerne på I med *skalarer*, dvs. reelle tal, på den sædvanlige måde: Hvis igen f er en funktion på I defineres λf ved $(\lambda f)(t) = \lambda f(t)$. Skalaren 1 er åbenbart neutral ved denne skalarmultiplikation. I øvrigt opfylder den de samme regneregler som skalarmultiplikation i talrummene. Den virker distributivt over for addition: $(\lambda + \mu)f = \lambda f + \mu f$ og $\lambda(f + g) = \lambda f + \lambda g$. Og $(\lambda)(\mu f) = (\lambda\mu)f$.

Alt i alt opfylder additionen og skalarmultiplikationen i $\mathbf{F}(I)$ de tidligere oplistede regler 1.-8. gældende for talrummene. Derfor er $\mathbf{F}(I)$ *et vektorrum*. Elementerne (vektorerne) i $\mathbf{F}(I)$ er altså funktioner. Vektorrum bestående af funktioner kaldes **funktionsrum**. Begreberne **linearkombination, udspænding, lineær uafhængighed** (samt naturligvis *afhængighed*), **basis** og **dimension** lader sig uden videre overføre til denne situation:

Hvis $\mathbf{M}$ er en mængde af funktioner på et åbent interval I (dvs. $\mathbf{M}$ er en delmængde af $\mathbf{F}(I)$), forstår vi ved $span(M)$ mængden af alle linearkombinationer som kan dannes ved hjælp af endeligt mange funktioner fra $\mathbf{M}$:

$$span(M) = \{\lambda_1 f_1 + \lambda_2 f_2 + \ldots + \lambda_k f_k \mid k \in \mathbf{N}, \quad f_1, f_2, \ldots, f_k \in M, \lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{R}\}.$$

Læg mærke til at antallet, k, af elementer i disse linearkombinationer er vilkårligt. Et element i $span(M)$ er altså en funktion g, givet ved $g(t) = \lambda_1 f_1(t) + \lambda_2 f_2(t) + \ldots + \lambda_k f_k(t)$ for alle $t \in I$. I overensstemmelse med vores tidligere sprogbrug omtaler vi $span(M)$ som **underrummet udspændt af M**. Det ses at underrummene i $\mathbf{F}(I)$ netop er de delmængder af $\mathbf{F}(I)$ som er stabile, dvs. lukkede, over for linearkombinering.

Et sæt af funktioner $f_1, f_2, \ldots, f_k$ fra $\mathbf{F}(I)$ på I er **lineært uafhængigt**, netop hvis en linearkombination af f'erne kun fremstille nulfunktionen på I hvis linearkombinationen er triviel. Præcist sagt: Hvis

$$\lambda_1 f_1(t) + \lambda_2 f_2(t) + \ldots + \lambda_k f_k(t) = 0, \quad \text{for alle}\, t \in I,$$

må alle λ'erne være 0.

Det er væsentligt at gøre sig klart, at både når det gælder linearkombinering, udspænding, lineær uafhængighed osv., angår lighedstegnene hele *funktioner*, altså samtlige funktionsværdier på det relevante interval I, og ikke kun værdierne i et enkelt eller nogle få punkter.

En **(endelig) basis** for et underrum $\mathbf{U}$ i $\mathbf{F}(I)$ er et *endeligt* sæt af funktioner, der (1) udspænder $\mathbf{U}$, og (2) er lineært uafhængigt. Antallet af funktioner i en sådan basis kaldes underrummets **dimension**. Underrummet alene bestående af nulfunktionen har ikke nogen basis. Det tillægges dimensionen 0.

I talrummene fandt vi at alle underrum (inklusive talrummene selv) besidder (endelige) baser. Sådan er det langtfra i funktionsrummene. Der findes f.eks. *ingen endelig basis* for $\mathbf{F}(I)$ selv. Ingen nok så stor samling af endeligt mange funktioner fra $\mathbf{F}(I)$ er righoldig nok til at frembringe enhver anden funktion fra $\mathbf{F}(I)$ som linearkombination af funktioner fra samlingen. De underrum af $\mathbf{F}(I)$ som ikke har en endelig basis kaldes **uendeligdimensionale**.

Der findes en masse vigtige underrum i $\mathbf{F}(I)$. Et vigtigt underrum udgøres af de *differentiable funktioner* på I, benævnt $\mathbf{D}(I)$. At disse funktioner virkelig udgør et underrum i $\mathbf{F}(I)$ følger simpelthen af, at en linearkombination af differentiable funktioner igen er en differentiabel funktion. Et endnu vigtigere underrum udgøres af den mængde af funktioner på I, der ud over at være differentiable også opfylder at deres afledede funktion er kontinuert. Igen er det klart at denne delmængde er stabil over for linearkombinering. Funktioner af denne type kaldes kort **kontinuert differentiable**. De betegnes med $\mathbf{C}(I)$. Der gælder åbenbart at $\mathbf{C}(I)$ er et underrum i $\mathbf{D}(I)$, der på jo på sin side er et underrum i $\mathbf{F}(I)$. Heller ingen af disse underrum har endelige baser.

Ser vi nu på løsningsrummet $\{x : \mathbb{R} \to \mathbb{R} \,|\, x(t) = c,\, t \in \mathbb{R}, \quad \text{hvor}\, c \in \mathbb{R}\}$ for den homogene differentialligning $x'(t) = 0$ er det øjensynlig et underrum i $\mathbf{C}(\mathbb{R})$. Hvis vi med 1 betegner funktionen det er konstant lig 1 på $\mathbb{R}$ har vi simpelthen $\{x : \mathbb{R} \to \mathbb{R} \,|\, x(t) = c,\, t \in \mathbb{R}, \quad \text{hvor}\, c \in \mathbb{R}\} = span\{1\}$. Da elementet 1 desuden danner et lineært uafhængigt sæt (funktionen er jo ikke nulfunktionen), har vi fundet at 1 udgør en basis for underrummet. Dette er således 1-dimensionalt. Løsningsrummet $\{x : \mathbb{R} \to \mathbb{R} \,|\, x(t) = \int_{t_0}^{t} a(u)\,du + c,\, t \in \mathbb{R}, \quad \text{hvor}\, c \in \mathbb{R}\}$ til den inhomogene differentialligning $x'(t) = a(t)$ er det *sideunderrum* til $span\{1\}$ i $\mathbf{C}(I)$ der "går gennem" elementet x_0, givet ved $x_0(t) = \int_{t_0}^{t} a(u)\,du$, hvor $t \in \mathbb{R}$.

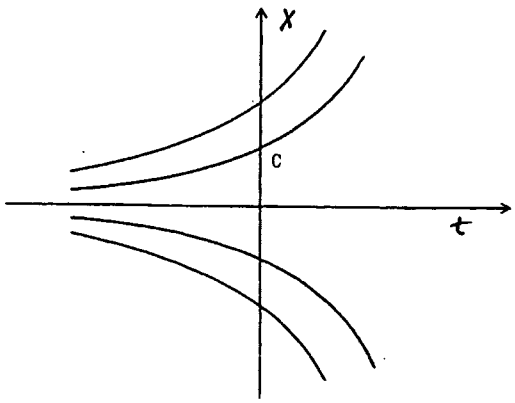
De første ikke-trivielle differentialligninger man støder på, generaliserer de ovennævnte. Ligningen $x'(t) = ax(t)$, hvor a er et vilkårligt reelt tal, har $x'(t) = 0$ som specialtilfælde (når $a = 0$). Som bekendt har løsningsmængden til $x'(t) = ax(t)$ formen $\{x : \mathbb{R} \to \mathbb{R} \,|\, x(t) = ce^{at}, t \in \mathbb{R} \quad \text{hvor}\, c \in \mathbb{R}\}$.

(Hvis du ikke kan huske hvorfor, så se her: Allerførst konstaterer vi at nulfunktionen på $\mathbb{R}$ er en løsning. Dernæst søger vi de eventuelle positive løsninger x. For dem er differentialligningen ensbetydende med ligningen $\frac{x'(t)}{x(t)} = a$. Men da $\frac{x'(t)}{x(t)} = (lnx(t))'$ kommer dette ud på at $(lnx(t))' = a$. Dette er ensbetydende med at $lnx(t)$ er en stamfunktion til a, dvs. $lnx(t) = at + b$, $t \in \mathbb{R}$, hvor b er en vilkårlig reel konstant. Men det viser, at $x(t) = e^{at+b} = e^b e^{at}, t \in \mathbb{R}$, og da e^b gennemløber de positive reelle tal når b gennemløber $\mathbb{R}$ kan vi lige så godt skrive c i stedet for e^b. De positive løsninger har altså formen $ce^{at}, c \in \mathbb{R}_+$. De negative løsninger er øjensynlig $ce^{at}, c \in \mathbb{R}_-$. Bortset fra disse - og fra nulløsningen - findes der ikke andre løsninger på $\mathbb{R}$. Thi hvis x er en

løsning på $\mathbb{R}$, må der om funktionen w, givet ved $w(t) = e^{-at}x(t)$, $t \in \mathbb{R}$, gælde at

$$w'(t) = (e^{-at}x(t))' = -ae^{-at}x(t) + e^{-at}x'(t) = -ae^{-at}x(t) + e^{-at}ax(t) = 0,$$

hvor det næstsidste lighedstegn skyldes at x er en løsning til differentialligningen. Vores udregning viser at funktionen w er en stamfunktion til 0. Der gælder derfor at $e^{-at}x(t) = w(t) = c$, hvor c er en vilkårlig reel konstant. Men så har x den påståede form. Faktisk kunne vi, bagklogt, have nøjedes med de sidste betragtninger, der jo udpeger samtlige løsninger. Men idéen til at inddrage w fik vi først efter at have fundet at x_0, givet ved $x(t) = e^{at}$, $t \in \mathbb{R}$, er en nulpunktsfri løsning til differentialligningen.)



Løsningsrummet til differentialligningen $x'(t) = ax(t)$ er altså

$$\{x : \mathbb{R} \to \mathbb{R} \mid c \in \mathbb{R}, x(t) = ce^{at}, \quad \text{hvor } t \in \mathbb{R}\} = span\{x_0\},$$

hvoraf det fremgår at løsningsrummet er et **1-dimensionalt underrum i F(ℝ)**. Vi kan også skrive ligningen som $x'(t) - ax(t) = 0$, hvilket begrunder at vi kalder den **en homogen differentialligning af 1. orden med konstante koefficienter**.

Hvis der er givet et vilkårligt "tidspunkt" t_0 og en vilkårlig funktionsværdi x_0, findes der netop én løsning x til differentialligningen, som til tidspunktet t_0 har værdien x_0. Vi skal nemlig blot vælge parameteren c, så at $ce^{at_0} = x_0$, dvs. $c = x_0e^{-at_0}$. Det problem at efterspørge, blandt løsningerne til en differentialligning, den/de løsninger som til et givet tidspunkt t_0 har en given værdi x_0 kaldes et **begyndelsesværdiproblem**. Vi har just indset, at for en homogen 1. ordens differentialligning har ethvert begyndelsesproblem en entydigt bestemt løsning.

En inhomogen differentialligning af 1. orden med konstante koefficienter er en ligning af formen $x'(t) - ax(t) = b$, hvor både a og b er konstanter. For at ligningen virkelig skal være inhomogen, må $b \neq 0$. Lad os minde om dens løsninger. Øjensynlig

er den konstante funktion $t \to -\frac{b}{a}, t \in \mathbb{R}$ én løsning på $\mathbb{R}$. Hvis x er en anden løsning til ligningen, defineret på $\mathbb{R}$, ser vi at $x + \frac{b}{a}$ er en løsning til den tilsvarende homogene ligning, idet $(x+\frac{b}{a})'(t) = x'(t) = ax(t)+b = a(x(t)+\frac{b}{a}) = a(x+\frac{b}{a})(t)$. Dette indebærer i følge det foregående, at $x(t) + \frac{b}{a} = ce^{at}$, $t \in \mathbb{R}$, eller med andre ord at

$$x(t) = ce^{at} - \frac{b}{a}, \; t \in \mathbb{R},$$

der er et sideunderrum til løsningsrummet for den homogene, "gående igennem" funktionen $t \to -\frac{b}{a}$.

Også her har ethvert begyndelsesværdiproblem netop én løsning. Er nemlig t_0 og x_0 givne, vil den løsning der er givet ved at $c = e^{-at_0}(x_0 + \frac{b}{a})$ løse begyndelsesværdiproblemet.

I behandlingen senere i dette kapitel har vi brug for at kunne løse *en inhomogen 1. ordens differentialligning* som *ikke* har konstante koefficienter, nemlig en ligning af formen

$$x'(t) - kx(t) = g(t),$$

hvor k godt nok er en reel konstant men hvor g er en kontinuert funktion givet på et interval I. For at løse denne ligning benytter vi os af et lille trick. Enhver funktion x på et interval kan skrives som $x(t) = y(t)e^{kt}$ for en passende funktion y, idet vi jo simpelthen kan sætte $y(t) = x(t)e^{-kt}$, $t \in I$. Vi undersøger nu betingelsen for at $x(t)\,(= y(t)e^{kt})$ er en løsning til den betragtede differentialligning på I. Det kommer jo pr. definition ud på at

$$(y(t)e^{kt})' = ky(t)e^{kt} + g(t), \text{ d.v.s. } \quad y'(t)e^{kt} + ky(t)e^{kt} = ky(t)e^{kt} + g(t), \; t \in I.$$

Dette er åbenbart det samme som at $y'(t)e^{kt} = g(t)$, $t \in I$, hvilket er ensbetydende med at $y'(t) = g(t)e^{-kt}$, $t \in I$. Men dette sidste er et rent stamfunktionsproblem, hvis løsning er

$$y(t) = \int_{t_0}^{t} g(t)e^{-kt}dt + \beta, \; t \in I, \; \beta \in \mathbb{R}.$$

Her er t_0 et vilkårligt, men fast tal i I.

Vi har altså at løsningerne til den inhomogene differentialligning er givet ved

$$x(t) = e^{kt}y(t) = e^{kt}(\int_{t_0}^{t} g(t)e^{-kt}dt + \beta), \; t \in I, \; \beta \in \mathbb{R}.$$

Begyndelsesværdiproblemet svarende til denne situation har en entydigt bestemt løsning for $t_0 \in I$ og x_0 vilkårlig i $\mathbb{R}$. Den fås ved at opsøge den blandt stamfunktionerne der til "tid" t_0 er 0, og sætte $\beta = x_0e^{-kt_0}$.

Lineære differentialligningssystemer af 1.orden med konstante koefficienter

Vi vil først se på et overmåde simpelt *homogent system* af lineære 1. ordens differentialligninger med konstante koefficienter, nemlig

$$\begin{aligned} x_1'(t) &= a_1 x_1(t) \\ x_2'(t) &= a_2 x_2(t) \\ &\vdots \\ x_n'(t) &= a_n x_n(t), \end{aligned}$$

hvor $a_1, a_2, \ldots, a_n$ er reelle konstanter. Som det fremgår har ligningerne i virkeligheden ingen ting med hinanden at gøre, de er helt **afkoblede**. Hver ligning kan løses for sig efter de tidligere beskrevne retningslinjer. Således finder vi

$$x_1(t) = c_1 e^{a_1 t},\ x_2(t) = c_2 e^{a_2 t},\ \ldots x_n(t) = c_n e^{a_n t},$$

hvor $c_1, c_2, \ldots, c_n$ er vilkårlige reelle konstanter, og hvor alle funktionerne er defineret for $t \in \mathbb{R}$.

I det følgende vil vi gøre rede for at løsningen af en vigtig klasse af komplicerede og stærkt forbundne lineære 1. ordens differentialligninger kan føres tilbage et system af den netop behandlede art, ved anvendelsen af, ja hvad mon, lineær algebra!. Først vil vi dog skrive det ovenstående system på matrixform:

$$\begin{pmatrix} x_1'(t) \\ x_2'(t) \\ \vdots \\ x_n'(t) \end{pmatrix} = \begin{pmatrix} a_1 & 0 & 0 & \ldots & 0 \\ 0 & a_2 & 0 & \ldots & 0 \\ 0 & 0 & a_3 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & 0 & \ldots & a_n \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{pmatrix}.$$

Dette fortjener et par kommentarer. Tidligere har vi kun arbejdet med rækker eller søjler hentet fra talrummene $\mathbb{R}^n$. Tilsvarende har vi kun opereret med produkter af matricer og søjler, hvor søjlerne var dannet af tal. Der er selvfølgelig intet som helst i vejen for som specialtilfælde heraf at betragte søjler som er talsæt af funktionsværdier, én søjle for hvert t i det relevante interval. Matrixoperationerne og reglerne for dem virker selv sagt lige så godt hvis tallene i søjlerne tilfældigvis er funktionsværdier. Hvis vi for *ethvert* t betragter vektoren $(x_1(t), x_2(t), \ldots, x_n(t))$, skrevet som række eller søjle som det nu passer os, har vi fastlagt en afbildning fra $\mathbb{R}$ ind i $\mathbb{R}^n$. Denne afbildning betegnes $(x_1, x_2, \ldots, x_n)$ og kaldes gerne en **vektorfunktion**. Undertiden vil vi bruge den korte notation $\mathbf{x}$ for vektorfunktionen $(x_1, x_2, \ldots, x_n)$ (i analogi med at vi tidligere har indført den korte betegnelse $\underline{x}$ for *talrums*vektoren $(x_1, x_2, \ldots, x_n)$, når x-erne står for tal). Det er endvidere nærliggende at benytte notationen $\mathbf{x}'$ for vektorfunktionen

af de afledede: $(x_1', x_2', \ldots, x_n')$. På denne måde kan vi sammenfatte det ovenstående således

$$\mathbf{x}' = A\mathbf{x},$$

hvor A er den opskrevne diagonalmatrix. (Pas på ikke at forveksle dette med basisskift, bare fordi der står ' i ligningen.) Løsningssættet kan opskrives som vektorfunktionen bestemt ved

$$\mathbf{x}(t) = \begin{pmatrix} c_1 e^{a_1 t} \\ c_2 e^{a_2 t} \\ \vdots \\ c_n e^{a_n t} \end{pmatrix}, \quad t \in \mathbb{R}$$

hvor $c_1, c_2, \ldots, c_n$ er vilkårlige reelle tal.

Også her kan vi formulere - og besvare - et til situationen hørende **begyndelsesværdiproblem**: Lad der være givet et $t_0 \in \mathbb{R}$ og en vektor $\underline{x}_0 = (x_1, x_2, \ldots, x_n)$ i $\mathbb{R}^n$. Bestem den/de eventuelle løsningsvektorfunktioner $\mathbf{x}$ defineret på $\mathbb{R}$, som opfylder, at $\mathbf{x}(t_0) = \underline{x}_0$. Det ses at ethvert sådant problem har netop én løsning, nemlig den som er fastlagt ved at $c_1 = x_1 e^{-a_1 t_0}, c_2 = x_2 e^{-a_2 t_0}, \ldots, c_n = x_n e^{-a_n t_0}$.

Den generelle løsning til det oprindelige problem kan skrives som

$$\mathbf{x}(t) = c_1 \begin{pmatrix} e^{a_1 t} \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} 0 \\ e^{a_2 t} \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \ldots + c_n \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ e^{a_n t} \end{pmatrix}.$$

Det fremgår altså, at enhver løsningsvektorfunktion $\mathbf{x}$ til differentialligningen $\mathbf{x}' = A\mathbf{x}$ er en linearkombination af løsningsvektorfunktionerne $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$, givet ved

$$\mathbf{x}_1(t) = \begin{pmatrix} e^{a_1 t} \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \mathbf{x}_2(t) = \begin{pmatrix} 0 \\ e^{a_2 t} \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots \text{og} \quad \mathbf{x}_n(t) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ e^{a_n t} \end{pmatrix}.$$

Dette lægger op til et lille

Indskud: Ligesom vi i det foregående afsnit har organiseret de reelle funktioner på et givet interval I som et vektorrum ved (punktvis) addition og skalarmultiplikation, kan vi organisere vektorfunktionerne - med n komponenter - på I som et vektorrum ved punktvis addition og skalarmultiplikation på hver af de n pladser. Dette er aldeles analogt til den måde hvorpå vi organiserede sæt af reelle tal, ved at addere og skalarmultiplicere

tallene pladsvis. Benytter vi betegnelsen $\mathbf{F}_n(I)$ for mængden af vektorfunktioner med n komponenter på intervallet I, er $\mathbf{F}_n(I)$ altså et vektorrum med den netop indførte addition og skalarmultiplikation.

Begreber som "linearkombination", "udspænding", "linear (u)afhængighed", "basis" og "dimension" m.m. overføres uden videre til den nye situation. Vi bemærkede tidligere at $\mathbf{F}(I)$, der for resten er lig $\mathbf{F}_1(I)$, ikke har nogen endelig basis. Derfor har $\mathbf{F}_n(I)$ selvfølgelig heller ikke nogen endelig basis.

I fortsættelse af betegnelserne for de forskellige rum af reelle funktioner på et interval I, betegner vi med $\mathbf{D}_n(I)$ underrummet af $\mathbf{F}_n(I)$ bestående af de vektorfunktioner, hvis komponentfunktioner er differentiable på I. (At der virkelig er tale om et underrum skyldes at en linearkombination af sådanne vektorfunktioner har komponenter der er en linearkombination af differentiable funktioner.) Ligeledes betegner vi med $\mathbf{C}_n(I)$ det underrum af $\mathbf{D}_n(I)$ (og dermed af $\mathbf{F}_n(I)$) der udgøres af de vektorfunktioner hvis komponentfunktioner alle er kontinuert differentiable. *Indskud slut.*

De n løsninger $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ er lineært uafhængige. Thi lad os betragte en fremstilling af *nulvektorfunktionen* 0 som linearkombination af disse vektorfunktioner

$$0 = \alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2 + \ldots + \alpha_n \mathbf{x}_n$$

Dette er ensbetydende med at der for ethvert $t \in \mathbb{R}$ gælder

$$\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \alpha_1 \begin{pmatrix} e^{a_1 t} \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} 0 \\ e^{a_2 t} \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \ldots + \alpha_n \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ e^{a_n t} \end{pmatrix} = \begin{pmatrix} \alpha_1 e^{a_1 t} \\ \alpha_2 e^{a_2 t} \\ \vdots \\ \alpha_n e^{a_n t} \end{pmatrix}.$$

Eftersom de indgående eksponentialfunktioner ikke ligefrem er konstant 0 (de antager som bekendt overhovedet ikke værdien 0, så chancerne er små), må α-erne alle være 0. Men så er $\mathbf{x}$-erne *lineært uafhængige*. Samtidig fandt vi ovenfor at $\mathbf{x}$-erne *udspænder* mængden af løsninger til differentialligningen. Dette viser at løsningsmængden er et *underrum* (i $\mathbf{C}_n(\mathbb{R})$, da alle komponentfunktionerne åbenbart er kontinuert differentiable) med $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ som basis. Løsningsrummet, som vi nu er berettiget til at kalde det, har således *dimensionen* n.

Nu er vi rede til at behandle et generelt **homogent** system af 1. ordens lineære

differentialligninger med konstante koefficienter. Derved forstår vi et system af formen

$$\begin{aligned}
x_1'(t) &= a_{11}x_1(t) + a_{12}x_2(t) + \ldots + a_{n1}x_n(t) \\
x_2'(t) &= a_{21}x_1(t) + a_{22}x_2(t) + \ldots + a_{n2}x_n(t) \\
&\vdots \\
x_n'(t) &= a_{n1}x_1(t) + a_{n2}x_2(t) + \ldots + a_{nn}x_n(t),
\end{aligned}$$

hvor alle a'erne er konstante, og hvor $t \in \mathbb{R}$. Efter helt samme retningslinjer som før kan vi skrive dette system på matrixform

$$\mathbf{x}' = A\mathbf{x},$$

hvor A - der gerne kaldes **koefficientmatricen** - her står for matricen

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

Lidt senere skal vi se nogle eksempler på at sådanne systemer kan opstå som modeller for fænomener og problemstillinger uden for matematikken selv. Men inden da er det bekvemt at have onduleret den matematiske behandling af systemet.

Idéen i behandlingen er at foretage - hvis det er muligt - et basisskift i $\mathbb{R}^n$ så at systemet ovenfor bringes på diagonalform. Dermed menes at det oprindelige system erstattes med ét på diagonalform - som så kan løses som ovenfor - der har en brugbar forbindelse til det oprindelige system. Vi går til værks sådan her:

Vi ved at matricen A kan diagonaliseres netop hvis der findes en basis for $\mathbb{R}^n$ bestående af egenvektorer for A, hvilket jo ikke altid er tilfældet. Diagonaliseringen udføres ved anvendelse af den matrix S, hvis søjler er egenvektorerne for A udtrykt i den oprindelige basis - og her vælger vi som denne basis simpelthen grundvektorerne i $\mathbb{R}^n$. I følge den tidligere udviklede teori opnår vi at matricen D, givet ved

$$D = S^{-1}AS,$$

er en diagonalmatrix, hvor der i diagonalen står egenværdierne for A i en rækkefølge der svarer til egenvektorernes rækkefølge i S. Så kan vi udtrykke det oprindelige differentialligningssystem ved hjælp af D og S, idet

$$\mathbf{x}' = A\mathbf{x} = (SDS^{-1})\mathbf{x},$$

hvoraf

$$S^{-1}\mathbf{x}' = DS^{-1}\mathbf{x}.$$

Før vi går videre, må vi sikre at der er mening i galskaben. Hvad er der egentlig på færde med vektorfunktionen $S^{-1}\mathbf{x}$ osv? Jo, hvis vi har givet en vektorfunktion $\mathbf{x}$ defineret på et interval I, og en kvadratisk matrix M, er $M\mathbf{x}$ en vektorfunktion, bestemt ved at

$$M\mathbf{x}(t) = \begin{pmatrix} m_{11}x_1(t) + m_{12}x_2(t) + \ldots + m_{1n}x_n(t) \\ m_{21}x_1(t) + m_{22}x_2(t) + \ldots + m_{2n}x_n(t) \\ \vdots \\ m_{n1}x_1(t) + m_{n2}x_2(t) + \ldots + m_{nn}x_n(t) \end{pmatrix}$$

for $t \in \mathbb{R}$. Her ser vi, idet differentiation foregår komponentvis, at

$$(M\mathbf{x})'(t) = \begin{pmatrix} m_{11}x_1'(t) + m_{12}x_2'(t) + \ldots + m_{1n}x_n'(t) \\ m_{21}x_1'(t) + m_{22}x_2'(t) + \ldots + m_{2n}x_n'(t) \\ \vdots \\ m_{n1}x_1'(t) + m_{n2}x_2'(t) + \ldots + m_{nn}x_n'(t) \end{pmatrix} = M\mathbf{x}'(t).$$

Konklusionen er, at $(M\mathbf{x})' = M\mathbf{x}'$. Benyttes dette med S^{-1} i M's rolle, har vi altså $(S^{-1}\mathbf{x})' = S^{-1}\mathbf{x}'$.

Resultatet fra før $S^{-1}\mathbf{x}' = DS^{-1}\mathbf{x}$ kan altså udtrykkes $(S^{-1}\mathbf{x})' = DS^{-1}\mathbf{x}$. Men det viser, at vektorfunktionen $S^{-1}\mathbf{x}$ er en løsning til et differentialligningssystem på *diagonalform*: $\mathbf{y}' = D\mathbf{y}$, hvis løsninger vi allerede kender. Hvis diagonalelementerne i D, dvs. egenværdierne i A, er $\lambda_1, \lambda_2, \ldots, \lambda_n$ (som ikke nødvendigvis er forskellige) gælder

$$S^{-1}\mathbf{x}(t) = \begin{pmatrix} \alpha_1 e^{\lambda_1 t} \\ \alpha_2 e^{\lambda_2 t} \\ \vdots \\ \alpha_n e^{\lambda_n t} \end{pmatrix},$$

hvor α-erne er vilkårlige reelle tal. Multiplikation med S fra venstre giver os *løsningerne til det oprindelige ligningssystem*:

$$\mathbf{x}(t) = S \begin{pmatrix} \alpha_1 e^{\lambda_1 t} \\ \alpha_2 e^{\lambda_2 t} \\ \vdots \\ \alpha_n e^{\lambda_n t} \end{pmatrix},$$

hvor α-erne er vilkårlige reelle tal og $t \in \mathbb{R}$. Det ses at alle løsningsvektorfunktioner har kontinuerte afledede, dvs. $\mathbf{x} \in \mathbf{C}_n(I)$.

Også i denne situation har begyndelsesværdiproblemet en entydigt bestemt løsning. Er nemlig $t_0 \in \mathbb{R}$ og $\underline{0} \in \mathbb{R}^n$ givet, kan vi bestemme α-er så at

$$\mathbf{x}_0 = \mathbf{x}(t_0) = S \begin{pmatrix} \alpha_1 e^{\lambda_1 t_0} \\ \alpha_2 e^{\lambda_2 t_0} \\ \vdots \\ \alpha_n e^{\lambda_n t_0} \end{pmatrix},$$

idet

$$\begin{pmatrix} \alpha_1 e^{\lambda_1 t_0} \\ \alpha_2 e^{\lambda_2 t_0} \\ \vdots \\ \alpha_n e^{\lambda_n t_0} \end{pmatrix} = S^{-1}\mathbf{x}_0,$$

uden videre bestemmer α-erne pladsvis.

Dette væsentlige resultat er først og fremmest frembragt ved hjælp af metoderne fra lineær algebra. Men bagved det hele ligger det afgørende moment at *differentiation foregår lineært*, dvs. den afledede af en linearkombination af funktioner er lig den samme linearkombination af funktionernes afledede.

Lad det endnu engang være understreget, at forudsætningen for at denne behandling er mulig er at A kan diagonaliseres. Det fordrer dels at det karakteristiske polynomium har "tilstrækkeligt mange" reelle rødder - faktisk skal summen af deres geometriske multipliciteter være lig n - og dels at der findes en basis for $\mathbb{R}^n$ af egenvektorer for A. Imidlertid er det også muligt at løse visse klasser af 1. ordens lineære differentialligninger med konstante koefficienter, hvor koefficientmatricen ikke opfylder disse forudsætninger. Vi må afstå fra at give en systematisk behandling af disse tilfælde her.

Vi foretager en omskrivning af løsningsforskriften som givet ovenfor:

$$\mathbf{x}(t) = S(\alpha_1 \begin{pmatrix} e^{\lambda_1 t} \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} 0 \\ e^{\lambda_2 t} \\ \vdots \\ 0 \end{pmatrix} + \alpha_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ e^{\lambda_n t} \end{pmatrix}), \; t \in \mathbb{R}$$

der ved videre anvendelse af reglerne for matrixregning giver

$$\mathbf{x}(t) = \alpha_1 S \begin{pmatrix} e^{\lambda_1 t} \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \alpha_2 S \begin{pmatrix} 0 \\ e^{\lambda_2 t} \\ \vdots \\ 0 \end{pmatrix} + \alpha_n S \begin{pmatrix} 0 \\ 0 \\ \vdots \\ e^{\lambda_n t} \end{pmatrix}, \; t \in \mathbb{R}.$$

Dette viser, at enhver løsning $\mathbf{x}$ til differentialligningen er en linearkombination af løsningerne (skrevet lidt uformelt):

$$S \begin{pmatrix} e^{\lambda_1 t} \\ 0 \\ \vdots \\ 0 \end{pmatrix}, S \begin{pmatrix} 0 \\ e^{\lambda_2 t} \\ \vdots \\ 0 \end{pmatrix}, \ldots \text{og} \quad S \begin{pmatrix} 0 \\ 0 \\ \vdots \\ e^{\lambda_n t} \end{pmatrix},$$

som dermed *udspænder løsningsmængden. Disse løsninger er også lineært uafhængige.* Thi antag, at nulvektorfunktionen 0 er fremstillet som linearkombination af dem, så at

for ethvert $t \in \mathbb{R}$

$$0 = \alpha_1 S \begin{pmatrix} e^{\lambda_1 t} \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \alpha_2 S \begin{pmatrix} 0 \\ e^{\lambda_2 t} \\ \vdots \\ 0 \end{pmatrix} + \alpha_n S \begin{pmatrix} 0 \\ 0 \\ \vdots \\ e^{\lambda_n t} \end{pmatrix}.$$

Ved at sætte S uden for parentes og atter samle de linearkombinerede søjler til én, finder vi at for ethvert reelt t er

$$0 = S \begin{pmatrix} \alpha_1 e^{\lambda_1 t} \\ \alpha_2 e^{\lambda_2 t} \\ \vdots \\ \alpha_n e^{\lambda_n t} \end{pmatrix}.$$

Da S som koordinatsskiftsmatrix er invertibel kan vi multiplicere med S^{-1} fra venstre på begge sider, hvorved

$$0 = S^{-1}0 = \begin{pmatrix} \alpha_1 e^{\lambda_1 t} \\ \alpha_2 e^{\lambda_2 t} \\ \vdots \\ \alpha_n e^{\lambda_n t} \end{pmatrix}, t \in \mathbb{R}.$$

Da dette som før kun er opfyldt hvis alle α-erne er 0, er de betragtede løsninger lineært uafhængige. De udgør altså en *basis* for løsningsrummet. Også i dette tilfælde har vi derfor indset at *løsningerne til det homogene differentialligningssystem* $\mathbf{x}' = A\mathbf{x}$ *er et n-dimensionalt underrum i* $\mathbf{C}_n(R)$.

Vi er er nu i stand til at løse en omfattende klasse af 1. ordens lineære homogene differentialligninssystemer. Hvad kan vi stille op med de tilsvarende **inhomogene** systemer af formen

$$\mathbf{x}' = A\mathbf{x} + \underline{b},$$

hvor $\underline{b}$ er en ikke-triviel konstant, dvs. en vektor $\neq 0$ i $\mathbb{R}^n$? Jo, først og fremmest kan vi konstatere, at to vektorer $\mathbf{x}$ og $\mathbf{x}_0$ er løsninger til dette inhomogene system, hvis og kun hvis deres differens er en løsning til det tilsvarende homogene system. Thi

$$A(\mathbf{x} - \mathbf{x}_0)' = A\mathbf{x}' - A\mathbf{x}_0' = (A\mathbf{x} + \underline{b}) - (A\mathbf{x}_0 + \underline{b}) = A\mathbf{x} - A\mathbf{x}_0 = A(\mathbf{x} - \mathbf{x}_0).$$

Betegner vi løsningsrummet for det homogene system med L_h og for det inhomogene med L_{ih}, har vi altså fundet: Hvis $\mathbf{x}_0 \in L_{ih}$ gælder, at $\mathbf{x} \in L_{ih}$ hvis og kun hvis $\mathbf{x} - \mathbf{x}_0 \in L_h$.

Det indebærer at hvis vi blot kender én enkelt løsning - en **partikulær** løsning - til det inhomogene system, f.eks. $\mathbf{x}_0$, fås samtlige andre løsninger $\mathbf{x}$ ved til den partikulære løsning at lægge løsningsrummet for det tilsvarende homogene. Vi kan formulere dette som en

SÆTNING. *Løsningsmængden L_{ih} for det inhomogene differentialligningssystem $\mathbf{x}' = A\mathbf{x} + \underline{b}$ er sideunderrummet til L_h gennem en vilkårlig partikulær løsning $\mathbf{x}_0$ til det inhomogene system:*

$$L_{ih} = L_h + \mathbf{x}_0$$

Såfremt der findes en konstant vektor $\underline{c} \in \mathbb{R}$ så at $A\underline{c} = -\underline{b}$ (det drejer sig altså om at løse et ligningssystem), vil den konstante vektorfunktion $\mathbf{x}$, givet ved $\mathbf{x}(t) = \underline{c}$ for $t \in \mathbb{R}$, være en partikulær løsning til det inhomogene system. Indsættes dette i det inhomogene system får vi jo, at

$$\mathbf{x}'(t) = (\underline{c})' = 0 = -\underline{b} + \underline{b} = A\underline{c} + \underline{b} = A\mathbf{x} + \underline{b}.$$

Denne mulighed indtræffer altid hvis A er invertibel. I så fald kan vi ved at vælge $\underline{c} = -A^{-1}\underline{b}$ angive en konstant, partikulær løsning til det inhomogene system, nemlig vektorfunktionen givet ved $\mathbf{x}(t) = -A^{-1}\underline{b}$ for alle $t \in \mathbb{R}$. Dette kan kombineres med resultatet for løsning af homogene systemer, hvor koefficientmatricen kan diagonaliseres. Derved opstår resultatet:

SÆTNING. *Hvis matricen A er diagonaliserbar, og der desuden findes en vektor $\underline{c}$ så at $A\underline{c} = -\underline{b}$ (det er f.eks. tilfældet hvis A er invertibel), er alle løsningerne til det lineære 1. ordens differentialligningssystem $\mathbf{x}' = A\mathbf{x} + \underline{b}$ givet på formen*

$$\mathbf{x}(t) = S \begin{pmatrix} \alpha_1 e^{\lambda_1 t} \\ \alpha_2 e^{\lambda_2 t} \\ \vdots \\ \alpha_n e^{\lambda_n t} \end{pmatrix} + \underline{c},$$

hvor α-erne er vilkårlige reelle tal, og $t \in \mathbb{R}$.

Dette resultat dækker også det homogene tilfælde, idet vi for $\underline{b} = 0$ genfinder det tidligere opnåede.

En konstant løsning til et differentialligningssystem kaldes ofte en **ligevægtsløsning**.

Eksempel

Lad os betragte det homogene differentialligningssystem

$$\begin{pmatrix} x_1'(t) \\ x_2'(t) \\ x_3'(t) \end{pmatrix} = \begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix}.$$

I et tidligere eksempel fandt vi egenværdierne for koefficientmatricen til $\lambda = 5$ og $\lambda = -1$ (dobbeltrod). Vi fandt også at den kan diagonaliseres ved anvendelse af koordinatsskiftsmatricen

$$S = \begin{pmatrix} 1 & -1 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

Efter de opnåede resultater har differentialligningssystemet så løsningerne

$$\begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix} = S \begin{pmatrix} \alpha_1 e^{5t} \\ \alpha_2 e^{-t} \\ \alpha_3 e^{-t} \end{pmatrix} = \begin{pmatrix} \alpha_1 e^{5t} - \alpha_2 e^{-t} - \alpha_3 e^{-t} \\ \alpha_1 e^{5t} + \alpha_2 e^{-t} \\ \alpha_1 e^{5t} + \alpha_3 e^{-t} \end{pmatrix},$$

hvor $t \in \mathbb{R}$. (Kontrollér selv ved differentiation at resultatet passer).

Dernæst kigger vi på et tilknyttet inhomogent system

$$\begin{pmatrix} x_1'(t) \\ x_2'(t) \\ x_3'(t) \end{pmatrix} = \begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix} + \begin{pmatrix} 5 \\ 5 \\ 5 \end{pmatrix}.$$

Eftersom A er invertibel, med

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

kan vi udnytte sætningen ovenfor. Vi finder samtlige løsninger til det inhomogene system:

$$\begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix} = S \begin{pmatrix} \alpha_1 e^{5t} \\ \alpha_2 e^{-t} \\ \alpha_3 e^{-t} \end{pmatrix} - A^{-1} \begin{pmatrix} 5 \\ 5 \\ 5 \end{pmatrix},$$

$t \in \mathbb{R}, \alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$. Ved indsættelse af S og A opnås

$$= \begin{pmatrix} \alpha_1 e^{5t} - \alpha_2 e^{-t} - \alpha_3 e^{-t} \\ \alpha_1 e^{5t} + \alpha_2 e^{-t} \\ \alpha_1 e^{5t} + \alpha_3 e^{-t} \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix},$$

hvor $t \in \mathbb{R}$ og $\alpha_1, \alpha_2, \alpha_3$ er vilkårlige reelle tal.

Par af 1.ordens differentialligninger med anvendelser

Efter nu at have behandlet hovedsporet i den generelle teori for lineære differentialligningssystemer af første orden, er det tid at betragte et par anvendelseseksempler. Ud over at illustrere hvad apparatet kan bruges til, lægger eksemplerne op til formuleringen og løsningen af problemer som ikke endnu er blevet behandlet i hovedsporet.

Eksempel 1

Lad os forestille os en population der er opdelt i to grupper, børn og voksne. Hvis antallet af børn til tid t er $u(t)$ og antallet af voksne er $v(t)$ tillader vi os - med en kraftig idealisering - at betragte u og v som differentiable funktioner af tiden. Tilvæksten (regnet med fortegn) pr. tidsenhed til voksenpopulationen forudsættes at bestå i en positiv tilvækst hidrørende fra at nogle af børnene bliver voksne, og en negativ tilvækst hidrørende fra at nogle af de voksne dør eller udvandrer. Gør vi den antagelse at den positive tilvækst er proportional med den tilstedeværende børnepopulation, mens den negative tilvækst er proportional med voksenbestanden, får vi differentialligningen

$$v'(t) = au(t) - bv(t),$$

hvor a og b er ikke-negative konstanter. Gør vi tilsvarende den antagelse at den positive tilvækst i børnepopulationen hidrører fra at voksenpopulationen får et afkom proportionalt med dens egen størrelse, og at den negative tilvækst skyldes at en vis brøkdel af børnene bliver voksne, dør eller udvandrer, får vi yderligere differentialligningen

$$u'(t) = cv(t) - du(t),$$

med c og d som ikke-negative konstanter. Disse to differentialligninger udgør et lineært, homogent system af to 1. ordens differentialligninger:

$$\begin{pmatrix} v'(t) \\ u'(t) \end{pmatrix} = \begin{pmatrix} -b & a \\ c & -d \end{pmatrix} \begin{pmatrix} v(t) \\ u(t) \end{pmatrix}.$$

Koefficientmatricen har det karakteristiske polynomium $\lambda^2 + (b+d)\lambda + (bd - ac)$, hvis diskriminant - med en lille smule udregning - er $D = (b-d)^2 + 4ac$. I alle "typiske" tilfælde er D positiv - skal den være 0 må $b = d$ samtidig med at a eller c er lig 0, hvor det sidste svarer til at ingen børn bliver voksne eller ingen voksne får børn, hvilket giver en lovlig ucharmerende population. I de typiske tilfælde er der to forskellige egenværdier (den ene negativ og den anden positiv), og vi kan fare frem i overensstemmelse med den ovenfor udviklede teori. Eftersom vi lidt senere skal betragte to-dimensionale systemer generelt, skal vi afstå fra at gå videre for øjeblikket.

Eksempel 2

I den mekaniske fysik forestiller man sig et lod af massen M ophængt i en vægtløs fjeder, der enten er sammentrykket eller udtrukket. Slipper man fjederen vil loddet udføre en bevægelse langs sin lodrette akse. Bevægelsen er underkastet Newton's 2. lov: den resulterende kraft er lig masse gange acceleration. Hvis $x(t)$ beskriver loddets afstand til et givet måleniveau til tid t, er dets acceleration (som er rettet lodret, opad eller nedad) $x''(t)$, hvorfor den resulterende kraft er $Mx''(t)$. Denne er imidlertid lig med

summen af tyngdekraften på loddet, altså Mg, hvor g er tyngdeaccelerationen, og den såkaldte fjederkraft, idet vi i første omgang ser bort fra alle former for friktion. I følge Hooke's lov er fjederkraften proportional med loddets flytning fra udgangsstilliingen, med en positiv proportionalitetsfaktor k. Vi har altså

$$Mx''(t) = Mg - kx(t),$$

eller m.a.o.

$$x''(t) = g - \frac{k}{M}x(t).$$

Denne 2.ordens differentialligning er *ækvivalent med et system af to 1. ordens differentialligninger*. Sætter vi nemlig

$$x_1(t) = x'(t)$$

må (idet $x'' = x_1'$) x og x_1 jo tilfredsstille systemet

$$\begin{aligned} x'(t) &= x_1(t) \\ x_1'(t) &= g - \frac{k}{M}x(t). \end{aligned}$$

Hvis omvendt x og x_1 tilfredsstiller dette system, må åbenbart $x''(t) = x_1'(t) = g - \frac{k}{M}x(t)$, sådan at x tilfredsstiller den oprindelige ligning. Dette viser ækvivalensen mellem 2.ordens ligningen og 1.ordens ligningssystemet. Denne ækvivalens er ikke en tilfældighed knyttet til netop dette eksempel, men af en almen natur. *Enhver n'te-ordens differentialligning er ækvivalent med et system bestående af n 1. ordens differentialligninger.*

Dette system kan nu skrives på matrixform

$$\begin{pmatrix} x'(t) \\ x_1'(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -\frac{k}{M} & 0 \end{pmatrix} \begin{pmatrix} x(t) \\ x_1(t) \end{pmatrix} + \begin{pmatrix} 0 \\ g \end{pmatrix},$$

altså et lineært, inhomogent system. Eftersom koefficientmatricen øjensynlig er invertibel (determinanten er jo $\frac{k}{M} \neq 0$) med den inverse

$$\begin{pmatrix} 0 & -\frac{M}{k} \\ 1 & 0 \end{pmatrix},$$

har vi i følge den opstillede teori en partikulær løsning til den inhomogene i den konstante vektorfunktion

$$-\begin{pmatrix} 0 & -\frac{M}{k} \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ g \end{pmatrix} = \begin{pmatrix} \frac{Mg}{k} \\ 0 \end{pmatrix}.$$

Hvad nu med den homogene ligning? Ja, det karakteristiske polynomium er jo $\lambda^2 + \frac{k}{M}$. Det har jo desværre ingen reelle løsninger, så her kommer den hidtil udviklede teori til kort. Lidt senere skal vi se, hvordan den sag gribes an.

Inden da vil vi se på den variation af den ovenfor beskrevne situation hvor der tages hensyn til friktion. Ved små hastigheder for loddet får man en god tilnærmelse til virkeligheden ved at tilføje en friktionskraft der er proportional med (og modsat rettet) loddets hastighedsvektor. Derved får vi at gøre med differentialligningen

$$Mx''(t) = Mg - fx'(t) - kx(t),$$

dvs.

$$x''(t) = g - \frac{f}{M}x'(t) - \frac{k}{M}x(t),$$

hvor friktionskoefficienten f er en positiv konstant. På tilsvarende måde som før omsættes denne 2.ordens differentialligning til et lineært 1. ordens system ved at vi sætter $x_1 = x'$:

$$\begin{aligned} x'(t) &= x_1(t) \\ x_1'(t) &= g - \frac{f}{M}x_1(t) - \frac{k}{M}x(t), \end{aligned}$$

eller på matrixform

$$\begin{pmatrix} x'(t) \\ x_1'(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -\frac{k}{M} & -\frac{f}{M} \end{pmatrix} \begin{pmatrix} x(t) \\ x_1(t) \end{pmatrix} + \begin{pmatrix} 0 \\ g \end{pmatrix}.$$

Også dette system har en invertibel koefficientmatrix (determinanten er stadig lig $\frac{k}{M}$), nemlig

$$\begin{pmatrix} -\frac{f}{k} & -\frac{M}{k} \\ 1 & 0 \end{pmatrix},$$

hvilket giver den samme partikulære løsning som før

$$\begin{pmatrix} \frac{Mg}{k} \\ 0 \end{pmatrix}.$$

Hvad den homogene ligning angår, har koefficientmatricen her det karakteristiske polynomium $\lambda^2 + \frac{f}{M}\lambda + \frac{k}{M}$, hvis diskriminant er $D = \frac{f^2-4kM}{M^2}$. Hvis $D > 0$ har systemet to forskellige reelle egenværdier λ_1 og λ_2, som begge er negative ($\sqrt{(f^2 - 4kM)} < f$, da $k > 0$). I den situation duer den hidtil udviklede teori. Koordinatskiftsmatricen bliver

$$\begin{pmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{pmatrix},$$

så at løsningen til systemet bliver

$$\begin{pmatrix} x \\ x_1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{pmatrix} \begin{pmatrix} \alpha_1 e^{\lambda_1 t} \\ \alpha_2 e^{\lambda_2 t} \end{pmatrix} + \begin{pmatrix} \frac{Mg}{k} \\ 0 \end{pmatrix}.$$

Heraf ses at x (som jo var den størrelse vi er interesseret i), x_1 er blot en hjælpestørrelse, bestemmes som

$$x(t) = \alpha_1 e^{\lambda_1 t} + \alpha_2 e^{\lambda_2 t} + \frac{Mg}{k}, t \in \mathbb{R}.$$

Da λ-erne begge er negative ser vi, at udsvinget x går mod konstanten $\frac{Mg}{k}$. Man siger at der foreligger *eksponentiel dæmpning mod ligevægt.*

Hvis $D = 0$ er $\lambda = -\frac{f}{2M}$ reel dobbeltrod, og der er ikke en basis for $\mathbb{R}^2$ af egenvektorer for koefficientmatricen. Er endelig $D < 0$ er der ingen reelle rødder. I de to sidste tilfælde har vi (endnu) ingen redskaber til at løse differentialligningssystemet. Men nu har vi efterhånden opsparet et behov for at komme videre med teoriudviklingen.

Et generelt system af to lineære 1. ordens differentialligninger er givet ved formen

$$\begin{pmatrix} x_1' \\ x_2' \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

hvor a, b, c, d er vilkårlige reelle konstanter. Koefficientmatricen for systemet betegnes kort A. Det karakteristiske polynomium for dette system er $(a - \lambda)(d - \lambda) - bc = \lambda^2 - (a+d)\lambda + (ad - bc)$, hvis diskriminant er $D = (a+d)^2 - 4(ad - bc) = (a-d)^2 + 4bc$. Løsningsforholdene falder i tre forskellige klasser, afhængigt af D's humør. Vi behandler én ad gangen.

I. D > 0. I dette tilfælde er der to forskellige reelle egenværdier, med tilhørende lineært uafhængige egenvektorer, hvorfor tilfældet falder ind under den allerede behandlede generelle teori. (Lorenzen-modellens ligninger falder ind under dette tilfælde.)

II. D = 0. At $D = 0$ vil sige at $(a - d)^2 + 4bc = 0$, eller m.a.o. at $bc = -\frac{1}{4}(a - d)^2$. I denne situation er der en reel dobbeltrod $\lambda = \frac{a+d}{2}$. Egenvektorerne for A hørende til denne egenværdi findes som løsningerne til ligningssystemet

$$\begin{pmatrix} a - \frac{a+d}{2} & b \\ c & d - \frac{a+d}{2} \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} \frac{a-d}{2} & b \\ c & -\frac{a-d}{2} \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

dvs.

$$\frac{a-d}{2} u + bv = 0$$
$$cu - \frac{a-d}{2} v = 0.$$

Her optræder nu flere undertilfælde.

Hvis $\mathbf{c = 0}$, er - i kraft af identiteten i begyndelsen af afsnittet - $a = d$. Er nu også $b = 0$, er A faktisk en diagonalmatrix fra begyndelsen og vi er tilbage i grundtilfældet.

Hvis $b \neq 0$, har ligningssystemet løsningerne $v = 0$, u fri i $\mathbb{R}$. Egenrummet er altså éndimensionalt, og vi kan ikke finde to lineært uafhængige egenvektorer for A. *Men* vi klarer os endda. Differentialligningssystemet har i denne situation skikkelsen

$$\begin{pmatrix} x_1' \\ x_2' \end{pmatrix} = \begin{pmatrix} a & b \\ 0 & a \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

eller skrevet ud:

$$\begin{aligned} x_1'(t) &= ax_1(t) + bx_2(t) \\ x_2'(t) &= ax_2(t). \end{aligned}$$

Her kan vi først løse den sidste ligning, som jo alene omhandler x_2. Vi får $x_2(t) = \alpha_1 e^{at}$, $t \in \mathbb{R}$, hvor α_1 er et vilkårligt reelt tal. Dette kan nu indsættes i den første differentialligning, som derved alene omhandler x_1:

$$x_1'(t) = ax_1(t) + b\alpha_1 e^{at}.$$

Men denne type differentialligninger har vi løst i et tidligere afsnit. Løsningen er

$$x_1(t) = e^{at}\left(\int_{t_0}^{t} b\alpha_1 e^{at} e^{-at} dt + \alpha_2\right) = e^{at}(\alpha_1 bt + \alpha_2),\ t \in \mathbb{R}.$$

Løsningen til differentialligningssystemet er altså, når $c = 0$, $b \neq 0$, givet ved

$$\begin{aligned} x_1(t) &= e^{at}(\alpha_1 bt + \alpha_2) \\ x_2(t) &= \alpha_1 e^{at}, \end{aligned}$$

$t \in \mathbb{R}$, α_1, $\alpha_2 \in \mathbb{R}$.

Hvis $\mathbf{c} \neq \mathbf{0}$, finder vi igen at egenrummet for den enlige egenværdi $\lambda = \frac{a+d}{2}$ er éndimensionalt. Det er udspændt af vektoren

$$\begin{pmatrix} \frac{a-d}{2c} \\ 1 \end{pmatrix}$$

(check selv det ved at løse det opstillede ligningssystem i u og v). Atter engang kan vi ikke diagonalisere A. Men erfaringerne ovenfor viser at vi kan klare os med mindre. Kan vi nemlig finde en koordinatskiftsmatrix S, så at

$$S^{-1}AS = \begin{pmatrix} \frac{a+d}{2} & 1 \\ 0 & \frac{a+d}{2} \end{pmatrix} = J,$$

kan vi komme igennem med den samme teknik som i det umiddelbart foregående undertilfælde. Men en sådan matrix S findes faktisk. Regnearbejdet er ikke større end at

man kan prøve sig frem (men der findes nu en - i sin generelle form frygteligt indviklet - teori for sagen). Vi finder

$$S = \begin{pmatrix} \frac{a-d}{2c} & \frac{1}{c} \\ 1 & 0 \end{pmatrix}.$$

(Kontroller selv ved udregning at $AS = SJ$. Heraf følger det ønskede, idet S er invertibel fordi dens determinant er lig $-\frac{1}{c}$.)

Benytter vi nu den fundne matrix S til at skrive $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = S \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$, får vi at differentialligningssystemet

$$\begin{pmatrix} x_1' \\ x_2' \end{pmatrix} = A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

er ensbetydende med det transformerede differentialligningssystem

$$S \begin{pmatrix} y_1' \\ y_2' \end{pmatrix} = AS \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.$$

Men dette system kan skrives

$$\begin{pmatrix} y_1' \\ y_2' \end{pmatrix} = S^{-1}AS \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \frac{a+d}{2} & 1 \\ 0 & \frac{a+d}{2} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.$$

Dette system har vi faktisk løst i undertilfældet ovenfor, blot med $\frac{a+d}{2}$ i a's rolle og 1 i b's. Vi opnår derfor

$$\begin{aligned} y_1(t) &= e^{\frac{a+d}{2}t}(\alpha_1 t + \alpha_2) \\ y_2(t) &= \alpha_1 e^{\frac{a+d}{2}t}, \end{aligned}$$

$t \in \mathbb{R}$, $\alpha_1, \alpha_2 \in \mathbb{R}$. Løsningen til det oprindelige system fås så ved at transformere y-erne med S:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = S \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \frac{a-d}{2c} & \frac{1}{c} \\ 1 & 0 \end{pmatrix} \begin{pmatrix} e^{\frac{a+d}{2}t}(\alpha_1 t + \alpha_2) \\ \alpha_1 e^{\frac{a+d}{2}t} \end{pmatrix}.$$

Ved udførelse af den sidste matrixmultiplikation slutter vi, at hvis $c \neq 0$ er løsningerne til det betragtede system

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} e^{\frac{a+d}{2}t}(\alpha_1(\frac{a-d}{2}t + \frac{1}{c}) + \alpha_2 \frac{a-d}{2}) \\ e^{\frac{a+d}{2}t}(\alpha_1 t + \alpha_2) \end{pmatrix},$$

hvor $t \in \mathbb{R}$ og α_1, α_2 er vilkårlige reelle konstanter.

III. D < 0. I dette tilfælde kan vi ikke gøre rede for løsningsovervejelserne uden at inddrage **komplekse tal**, og det er der ikke tid til i denne forbindelse. Vi må slå os til tåls med resultatet. Man kan vise at løsningerne til differentialligningssystemet i dette tilfælde har skikkelsen

$$\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = \begin{pmatrix} e^{\frac{a+d}{2}t}(\alpha_1 cos\frac{\sqrt{-D}}{2}t + \frac{(a-d)\alpha_1+2b\alpha_2}{\sqrt{-D}} sin\frac{\sqrt{-D}}{2}t) \\ e^{\frac{a+d}{2}t}(\alpha_2 cos\frac{\sqrt{-D}}{2}t + \frac{2c\alpha_1+(d-a)\alpha_2}{\sqrt{-D}} sin\frac{\sqrt{-D}}{2}t) \end{pmatrix},$$

hvor α_1, α_2 er vilkårlige reelle konstanter. Ved at ommøblere dette udtryk kan man fremstille resultatet som en linearkombination af to vektorfunktioner med koefficienterne α_1 og α_2. Idet det ikke er vanskeligt at indse at de to vektorfunktioner er lineært uafhængige, viser dette at løsningsrummet også i dette tilfælde er todimensionalt.

Vi har nu erhvervet tilstrækkelig baggrund for at behandle de situationer i Eksempel 2 som vi før måtte lade ligge.

Eksempel 2, fortsat

Først behandles tilfældet uden friktion. Dér fik vi det karakteristiske polynomium $\lambda^2 + \frac{k}{M}$, hvis diskriminant er $D = -\frac{4k}{M} < 0$. Situationen falder derfor ind under det netop behandlede tilfælde. For at finde løsningerne til det homogene system,kan vi indsætte direkte i den opnåede formel, idet $a = d = 0$, $b = 1$, $c = -\frac{k}{M}$ og $-D = \frac{4k}{M}$:

$$x_(t) = \alpha_1 cos\sqrt{\frac{k}{M}}t + \alpha_2 \frac{1}{\sqrt{\frac{k}{M}}} sin\sqrt{\frac{k}{M}}t$$

$$x_1(t) = \alpha_2 cos\sqrt{\frac{k}{M}}t - \alpha_1\sqrt{\frac{k}{M}} sin\sqrt{\frac{k}{M}}t.$$

Løsningerne til det inhomogene system er derfor

$$\begin{pmatrix} x_(t) \\ x_1(t) \end{pmatrix} = \begin{pmatrix} \alpha_1 cos\sqrt{\frac{k}{M}}t + \alpha_2 \frac{1}{\sqrt{\frac{k}{M}}} sin\sqrt{\frac{k}{M}}t \\ \alpha_2 cos\sqrt{\frac{k}{M}}t - \alpha_1\sqrt{\frac{k}{M}} sin\sqrt{\frac{k}{M}}t \end{pmatrix} + \begin{pmatrix} \frac{Mg}{k} \\ 0 \end{pmatrix},$$

hvor det i virkeligheden kun er udtrykket for x vi efterspørger, idet x_1 jo simpelthen er x's afledede. Det fremgår at x som linearkombination af trigonometriske funktioner udviser periodiske svingninger.

I den situation hvor der indgår friktion skylder vi at gøre rede for tilfældene $D = 0$ og $D < 0$.

$D = 0$: Her er $f = 2\sqrt{kM}$, $a = 0$, $c = -\frac{k}{M}$, $d = -\frac{f}{M} = -2\sqrt{\frac{k}{M}}$. Det homogene system får efter den generelle analyse af dobbeltrodssituationen løsningerne

$$\begin{pmatrix} x(t) \\ x_1(t) \end{pmatrix} = \begin{pmatrix} e^{-\sqrt{\frac{k}{M}}t}(\alpha_1(\sqrt{\frac{k}{M}}t - \frac{M}{k}) + \alpha_2\sqrt{\frac{k}{M}}) \\ e^{-\sqrt{\frac{k}{M}}t}(\alpha_1 t + \alpha_2) \end{pmatrix},$$

hvorefter det inhomogene system får løsningerne

$$\begin{pmatrix} x(t) \\ x_1(t) \end{pmatrix} = \begin{pmatrix} e^{-\sqrt{\frac{k}{M}}t}(\alpha_1(\sqrt{\frac{k}{M}}t - \frac{M}{k}) + \alpha_2\sqrt{\frac{k}{M}}) \\ e^{-\sqrt{\frac{k}{M}}t}(\alpha_1 t + \alpha_2) \end{pmatrix} + \begin{pmatrix} \frac{Mg}{k} \end{pmatrix}.$$

Læg mærke til at der gælder $e^{-\sqrt{\frac{k}{M}}t}t \to 0$ for $t \to \infty$ sådan at $x(t) \to \frac{Mg}{k}$ for $t \to \infty$.

$D < 0$: I dette tilfælde er $\sqrt{-D} = \frac{\sqrt{4kM-f^2}}{M}$, mens $a = 0$, $b = 1$, $c = -\frac{k}{M}$, $d = -\frac{f}{M}$. Løsningen til det homogene system er så

$$\begin{pmatrix} x(t) \\ x_1(t) \end{pmatrix} = \begin{pmatrix} e^{\frac{-f}{2M}t}(\alpha_1 cos\frac{\sqrt{4kM-f^2}}{2M}t + \frac{f\alpha_1+2M\alpha_2}{\sqrt{4kM-f^2}}sin\frac{\sqrt{4kM-f^2}}{2M}t) \\ e^{\frac{-f}{2M}t}(\alpha_2 cos\frac{\sqrt{4kM-f^2}}{2M}t - \frac{2k\alpha_1+f\alpha_2}{\sqrt{4kM-f^2}}sin\frac{\sqrt{4kM-f^2}}{2M}t) \end{pmatrix},$$

og til det inhomogene

$$\begin{pmatrix} x(t) \\ x_1(t) \end{pmatrix} = \begin{pmatrix} e^{\frac{-f}{2M}t}(\alpha_1 cos\frac{\sqrt{4kM-f^2}}{2M}t + \frac{f\alpha_1+2M\alpha_2}{\sqrt{4kM-f^2}}sin\frac{\sqrt{4kM-f^2}}{2M}t) \\ e^{\frac{-f}{2M}t}(\alpha_2 cos\frac{\sqrt{4kM-f^2}}{2M}t - \frac{2k\alpha_1+f\alpha_2}{\sqrt{4kM-f^2}}sin\frac{\sqrt{4kM-f^2}}{2M}t) \end{pmatrix} + \begin{pmatrix} \frac{Mg}{k} \\ 0 \end{pmatrix}.$$

Også her er det egentlig kun x der interesserer os:

$$x(t) = e^{\frac{-f}{2M}t}(\alpha_1 cos\frac{\sqrt{4kM-f^2}}{2M}t + \frac{f\alpha_1+2M\alpha_2}{\sqrt{4kM-f^2}}sin\frac{\sqrt{4kM-f^2}}{2M}t) + \frac{Mg}{k},$$

hvoraf man bl.a. ser at eksponentialfaktoren $e^{-\frac{f}{2M}t}$ for $t \to \infty$ leverer en dæmpning af de svingninger der hidrører fra $sin-$ og $cos-$leddene.

Stikordsregister

Liste over tidligere udsendte tekster kan ses på IMFUFA's hjemmeside: http://mmf.ruc.dk eller rekvireres på sekretariatet, tlf. 46 74 22 63 eller e-mail: imfufa@ruc.dk.

332/97 ANOMAL SWELLING AF LIPIDE DOBBELTLAG
Specialerappoprt af: Stine Korremann
Vejleder: Dorthe Posselt

333/97 Biodiversity Matters
an extension of methods found in the literature on monetisation of biodiversity
by: Bernd Kuemmel

334/97 LIFE-CYCLE ANALYSIS OF THE TOTAL DANISH ENERGY SYSTEM
by: Bernd Kuemmel and Bent Sørensen

335/97 Dynamics of Amorphous Solids and Viscous Liquids
by: Jeppe C. Dyre

336/97 Problem-orientated Group Project Work at Roskilde University
by: Kathrine Legge

337/97 Verdensbankens globale befolkningsprognose
- et projekt om matematisk modellering
af: Jørn Chr. Bendtsen, Kurt Jensen, Per Pauli Petersen

338/97 Kvantisering af nanolederes elektriske ledningsevne
Første modul fysikprojekt
af: Søren Dam, Esben Danielsen, Martin Niss,
Esben Friis Pedersen, Frederik Resen Steenstrup
Vejleder: Tage Christensen

339/97 Defining Discipline
by: Wolfgang Coy

340/97 Prime ends revisited - a geometric point of view -
by: Carsten Lunde Petersen

341/97 Two chapters on the teaching, learning and assessment of geometry
by: Mogens Niss

342/97 A global clean fossil scenario DISCUSSION PAPER prepared by Bernd Kuemmel for the project LONG-TERM SCENARIOS FOR GLOBAL ENERGY DEMAND AND SUPPLY

343/97 IMPORT/EKSPORT-POLITIK SOM REDSKAB TIL OPTIMERET UDNYTTELSE AF EL PRODUCERET PÅ VE-ANLÆG
af: Peter Meibom, Torben Svendsen, Bent Sørensen

344/97 Puzzles and Siegel disks
by: Carsten Lunde-Petersen

345/98 Modeling the Arterial System with Reference to an Anesthesia Simulator
Ph.D. Thesis
by: Mette Sofie Olufsen

346/98 Klyngedannelse i en hulkatode-forstøvningsproces
af: Sebastian Horst
Vejledere: Jørn Borggren, NBI, Niels Boye Olsen

347/98 Verificering af Matematiske Modeller
- en analyse af Den Danske Eulerske Model
af: Jonas Blomqvist, Tom Pedersen, Karen Timmermann, Lisbet Øhlenschlæger
Vejleder: Bernhelm Booss-Bavnbek

348/98 Case study of the environmental permission procedure and the environmental impact assessment for power plants in Denmark
by: Stefan Krüger Nielsen
project leader: Bent Sørensen

349/98 Tre rapporter fra FAGMAT - et projekt om tal og faglig matematik i arbejdsmarkedsuddannelserne
af: Lena Lindenskov og Tine Wedege

350/98 OPGAVESAMLING - Bredde-Kursus i Fysik 1976 - 1998
Erstatter teksterne 3/78, 261/93 og 322/96

351/98 Aspects of the Nature and State of Research in Mathematics Education
by: Mogens Niss

352/98 The Herman-Swiatec Theorem with applications
by: Carsten Lunde Petersen

353/98 Problemløsning og modellering i en almendannende matematikundervisning
Specialerapport af: Per Gregersen og Tomas Højgaard Jensen

354/98 A Global Renewable Energy Scenario
by: Bent Sørensen and Peter Meibom

355/98 Convergence of rational rays in parameter spaces
by: Carsten Lunde Petersen and Gustav Ryd

356/98 Terrænmodellering
Analyse af en matematisk model til konstruktion af digitale terrænmodeller
Modelprojekt af: Thomas Frommelt, Hans Ravnkjær Larsen og Arnold Skimminge
Vejleder: Johnny Ottesen

357/98 Cayleys Problem
En historisk analyse af arbejdet med Cayleys problem fra 1870 til 1918
Et matematisk videnskabsfagsprojekt af: Rikke Degn, Bo Jakobsen, Bjarke K.W. Hansen, Jesper S. Hansen, Jesper Udesen, Peter C. Wulff
Vejleder: Jesper Larsen

358/98 Modeling of Feedback Mechanisms which Control the Heart Function in a View to an Implementation in Cardiovascular Models
Ph.D. Thesis by: Michael Danielsen

359/99 Long-Term Scenarios for Global Energy Demand and Supply
Four Global Greenhouse Mitigation Scenarios
by: Bent Sørensen (with contribution from Bernd Kuemmel and Peter Meibom)

360/99 SYMMETRI I FYSIK
En Meta-projektrapport af: Martin Niss, Bo Jakobsen & Tune Bjarke Bonné
Vejleder: Peder Voetmann Christiansen

361/99 Symplectic Functional Analysis and Spectral Invariants
by: Bernhelm Booß-Bavnbek, Kenro Furutani

362/99 Er matematik en naturvidenskab? - en udspænding af diskussionen
En videnskabsfagsprojekt-rapport af: Martin Niss
Vejleder: Mogens Nørgaard Olesen

363/99 EMERGENCE AND DOWNWARD CAUSATION
by: Donald T. Campell, Mark H. Bickhard, and Peder V. Christiansen

364/99 Illustrationens kraft - Visuel formidling af fysik
Integreret speciale i fysik og kommunikation
af Sebastian Horst
Vejledere: Karin Beyer, Søren Kjørup

365/99 To know - or not to know - mathematics, that is a question of context
by: Tine Wedege

366/99 LATEX FOR FORFATTERE - En introduktion til LATEX og IMFUFA-LATEX
af: Jørgen Larsen

367/99 Boundary Reduction of Spectral Invariants and Unique Continuation Property
by: Bernhelm Booss-Bavnbek

368/99 Kvartvejsraport for projektet SCENARIER FOR SAMLET UDNYTTELSE AF BRINT SOM ENERGIBÆRER I DANMARKS FREMTIDIGE ENERGISYSTEM
Projektleder: Bent Sørensen

369/99 Dynamics of Complex Quadratic Correspondences
by: Jacob S. Jalving
Supervisor: Carsten Lunde Petersen

370/99 OPGAVESAMLING - Bredde-Kursus i Fysik 1976 - 1999
Eksamensopgaver fra perioden 1976 - 1999. Denne tekst erstatter tekst nr. 350/98

371/99 Bevisets stilling - beviser og bevisførelse i en gymnasial matematik undervisning
Et matematikspeciale af: Maria Hermannsson
Vejleder: Mogens Niss

372/99 En kontekstualiseret matematikhistorisk analyse af ikke-lineær programmering: Udviklingshistorie og multipel opdagelse
Ph.d.-afhandling af Tinne Hoff Kjeldsen

373/99 Criss-Cross Reduction of the Maslov Index and a Proof of the Yoshida-Nicolaescu Theorem
by: Bernhelm Booss-Bavnbek, Kenro Furutani and Nobukazu Otsuki

374/99 Det hydrauliske spring - Et eksperimentelt studie af polygoner og hastighedsprofiler
Specialeafhandling af: Anders Marcussen
Vejledere: Tomas Bohr, Clive Ellegaard, Bent C. Jørgensen

375/99 Begrundelser for Matematikundervisningen i den lærde skole hhv. gymnasiet 1884-1914
Historiespeciale af Henrik Andreassen, cand.mag. i Historie og Matematik

376/99 Universality of AC conduction in disordered solids
by: Jeppe C. Dyre, Thomas B. Schrøder

377/99 The Kuhn-Tucker Theorem in Nonlinear Programming: A Multiple Discovery?
by: Tinne Hoff Kjeldsen

378/00 Solar energy preprints:
1. Renewable energy sources and termal energy storage
2. Integration of photovoltaic cells into the global energy system
by: Bent Sørensen

379/00 EULERS DIFFERENTIALREGNING
Eulers indførelse af differentialregningen stillet over for den moderne
En tredjesemesters projektrapport på den naturvidenskabelige basisuddannelse
af: Uffe Thomas Volmer Jankvist, Rie Rose Møller Pedersen, Maja Bagge Pedersen
Vejleder: Jørgen Larsen

380/00 MATEMATISK MODELLERING AF HJERTEFUNKTIONEN
Isovolumetrisk ventrikulær kontraktion og udpumpning til det cardiovaskulære system
af: Gitte Andersen (3.moduls-rapport), Jakob Hilmer og Stine Weisbjerg (speciale)
Vejleder: Johnny Ottesen

381/00 Matematikviden og teknologiske kompetencer hos kortuddannede voksne
- Rekognosceringer og konstruktioner i grænselandet mellem matematikkens didaktik og forskning i voksenuddannelse
Ph. d.-afhandling af Tine Wedege

382/00 Den selvundvigende vandring
Et matematisk professionsprojekt
af: Martin Niss, Arnold Skimminge
Vejledere: Viggo Andreasen, John Villumsen

383/00 Beviser i matematik
af: Anne K.S.Jensen, Gitte M. Jensen, Jesper Thrane, Karen L.A.W. Wille, Peter Wulff
Vejleder: Mogens Niss

384/00 Hopping in Disordered Media: A Model Glass Former and A Hopping Model
Ph.D. thesis by: Thomas B. Schrøder
Supervisor: Jeppe C. Dyre

385/00 The Geometry of Cauchy Data Spaces
This report is dedicated to the memory of Jean Leray (1906-1998)
by: B. Booss-Bavnbek, K. Furutani, K. P. Wojciechowski

386/00 Neutrale mandatfordelingsmetoder – en illusion?
af: Hans Henrik Brok-Kristensen, Knud Dyrberg, Tove Oxager, Jens Sveistrup
Vejleder: Bernhelm Booss-Bavnbek

387/00 A History of the Minimax Theorem: von Neumann's Conception of the Minimax Theorem - - a Journey Through Different Mathematical Contexts
by: Tinne Hoff Kjeldsen

388/00 Behandling af impuls ved kilder og dræn i C. S. Peskins 2D-hjertemodel
et 2. moduls matematik modelprojekt
af: Bo Jakobsen, Kristine Niss
Vejleder: Jesper Larsen

389/00 University mathematics based on problemoriented student projects: 25 years of experience with the Roskilde model
By: Mogens Niss
Do not ask what mathematics can do for modelling. Ask what modelling can do for mathematics!
by: Johnny Ottesen

390/01 SCENARIER FOR SAMLET UDNYTTELSE AF BRINT SOM ENERGIBÆRER I DANMARKS FREMTIDIGE ENERGISYSTEM Slutrapport, april 2001
Projektleder: Bent Sørensen
Projektdeltagere: DONG: Aksel Hauge Petersen, Celia Juhl, Elkraft System[#]: Thomas Engberg Pedersen[##], Hans Ravn, Charlotte Søndergren, Energi 2[#]: Peter Simonsen, RISØ Systemanalyseafd.: Kaj Jørgensen[*], Lars Henrik Nielsen, Helge V. Larsen, Poul Erik Morthorst, Lotte Schleisner, RUC: Finn Sørensen[**], Bent Sørensen
[#]Indtil 1/1-2000 Elkraft, [##] fra 1/5-2000 Cowi Consult
[*] Indtil 15/6-1999 DTU Bygninger & Energi, [**] fra 1/1-2001 Polypeptide Labs.
Projekt 1763/99-0001 under Energistyrelsens Brintprogram

391/01 Matematisk modelleringskompetence – et undervisningsforløb i gymnasiet
3. semesters Nat.Bas. projekt af: Jess Tolstrup Boye, Morten Bjørn-Mortensen, Sofie Inari Castella, Jan Lauridsen, Maria Gøtzsche, Ditte Mandøe Andreasen
Vejleder: Johnny Ottesen

392/01 "PHYSICS REVEALED" THE METHODS AND SUBJECT MATTER OF PHYSICS
an introduction to pedestrians (but not excluding cyclists)
PART III: PHYSICS IN PHILOSOPHICAL CONTEXT
by: Bent Sørensen.

393/01 Hilberts matematikfilosofi
Specialerapport af: Jesper Hasmark Andersen
Vejleder: Stig Andur Pedersen

394/01 "PHYSICS REVEALED" THE METHODS AND SUBJECT MATTER OF PHYSICS
an introduction to pedestrians (but not excluding cyclists)
PART II: PHYSICS PROPER
by: Bent Sørensen.

395/01 Menneskers forhold til matematik. Det har sine årsager!
Specialeafhandling af: Anita Stark, Agnete K. Ravnborg
Vejleder: Tine Wedege

396/01 2 bilag til tekst nr. 395: Menneskers forhold til matematik. Det har sine årsager!
Specialeafhandling af: Anita Stark, Agnete K. Ravnborg
Vejleder: Tine Wedege

397/01 En undersøgelse af solvents og kædelængdes betydning for anomal swelling i phospholipiddobbeltlag
2. modul fysikrapport af: Kristine Niss, Arnold Skimminge, Esben Thormann, Stine Timmermann
Vejleder: Dorthe Posselt

398/01 Kursusmateriale til "Lineære strukturer fra algebra og analyse" (E1)
Af: Mogens Brun Heefelt

399/01 Undergraduate Learning Difficulties and Mathematical Reasoning
Ph.D Thesis by: Johan Lithner
Supervisor: Mogens Niss

400/01 On Holomorphic Critical quasi circle maps
By: Carsten Lunde Petersen

401/01 Finite Type Arithmetic
Computable Existence Analysed by Modified Realisability and Functional Interpretation
Master's Thesis by: Klaus Frovin Jørgensen
Supervisors: Ulrich Kohlenbach, Stig Andur Pedersen and Anders Madsen

402/01 Matematisk modellering ved den naturvidenskabelige basisuddannelse
- udvikling af et kursus
Af: Morten Blomhøj, Tomas Højgaard Jensen, Tinne Hoff Kjeldsen og Johnny Ottesen

403/01 Generaliseringer i integralteorien
- En undersøgelse af Lebesgue-integralet, Radon-integralet og Perron-integralet
Et 2. modul matematikprojekt udarbejdet af: Stine Timmermann og Eva Uhre
Vejledere: Bernhelm Booss-Bavnbek og Tinne Hoff Kjeldsen

404/01 "Mere spredt fægtning"
Af: Jens Højgaard Jensen

405/01 Real life routing
- en strategi for et virkeligt vrp
Et matematisk modelprojekt af: David Heiberg Backchi, Rasmus Brauner Godiksen, Uffe Thomas Volmer Jankvist, Jogvan Martin Poulsen og Neslihan Saglanmak
Vejleder: Jørgen Larsen

406/01 Opgavesamling til dybdekursus i fysik
Eksamensopgaver stillet i perioden juni 1976 til juni 2001
Denne tekst erstatter tekst nr. 25/1980 + efterfølgende tillæg

407/01 Unbounded Fredholm Operators and Spectral Flow
By: Bernhelm Booss-Bavnbek, Matthias Lesch, John Phillips

408/02 Weak UCP and Perturbed Monopole Equations
By: Bernhelm Booss-Bavnbek, Matilde Marcolli, Bai-Ling Wang

409/02 Algebraisk ligningsløsning fra Cardano til Cauchy
- et studie af kombinationers, permutationers samt invariansbegrebets betydning for den algebraiske ligningsløsning *før* Gauss, Abel og Galois
Videnskabsfagsprojekt af: David Heiberg Backchi, Uffe thomas Volmer Jankvist, Neslihan Saglanmak
Vejleder: Bernhelm Booss-Bavnbek

410/02 2 projekter om modellering af influenzaepidemier
Influenzaepidemier- et matematisk modelleringsprojekt
Af: Claus Jørgensen, Christina Lohfert, Martin Mikkelsen, Anne-Louise H. Nielsen
Vejleder: Morten Blomhøj
Influenza A: Den tilbagevendende plage – et modelleringsprojekt
Af: Beth Paludan Carlsen, Christian Dahmcke, Lena Petersen, Michael Wagner
Vejleder: Morten Blomhøj

411/02 Polygonformede hydrauliske spring
Et modelleringsprojekt af: Kåre Stokvad Hansen, Ditte Jørgensen, Johan Rønby Pedersen, Bjørn Toldbod
Vejleder: Jesper Larsen

412/02 Hopfbifurkation og topologi i væskestrømning – en generel analyse samt en behandling af strømningen bag en cylinder
Et matematisk modul III professionsprojekt af. Kristine Niss, Bo Jakobsen
Vejledere: Morten Brøns, Johnny Ottesen

413/03 "Elevernes stemmer" Fysikfaget, undervisningen og lærerroller, som eleverne opfatter det i det almene gymnasium i Danmark
Af: Carl Angell, Albert Chr. Paulsen

414/03 Feltliniediagrammer En vej til forståelse?
Et 1. modul fysikprojekt af: Ditte Gundermann, Kåre Stokvad Hansen, Ulf Rørbæk Pedersen
Vejleder: Tage Emil Christensen

415/03 FYSIKFAGET I FORANDRING Læring og undervisning i fysik i gymnasiet med fokus på dialogiske processer, autenticitet og kompetenceudvikling
Ph.d.-afhandling i fysikdidaktik af: Jens Dolin

416/03 Fourier og Funktionsbegrebet
- Overgangen fra Eulers til Dirichlets funktionsbegreb
Projektrapport af: Rasmus Brauner Godiksen, Claus Jørgensen, Tony Moyer Hanberg, Bjørn Toldbod
Vejleder: Erik von Essen

417/03 The Semiotic Flora of Elementary particles
By: Peder Voetmann Christiansen

418/03 Militærmatematik set med kompetencebriller
3. modul projektrapport af: Gitte Jensen og specialerapport af: Jesper Thrane
Vejleder: Tine Wedege

419/03 Energy Bond Graphs – a semiotic formalization of modern physics
By: Peder Voetmann Christiansen

420/03 Stemning og Musikalsk Konsonans
Et matematisk modelleringsprojekt af: Claus Jørgensen
Vejleder: Johnny Ottesen

421/03 OPGAVESAMLING
Bredde-kursus i fysik 1976 – 2003.
Denne tekst erstatter tekst nr. 370/99

422/03 Vurdering af dynamisk blodstrømningsmodel
- ved numerisk simulering med FEMLAB
Et 2. modul matematikprojekt af: Sofie Inari Castella, Ingunn Gunnarsdóttir og Jacob Kirkensgaard Hansen
Vejleder: Johnny Ottesen

423/03 Fysikkens historie i en almendannende fysikundervisning
- Eksemplificeret med Millikan Ehrenhaft kontroversen
Specialerapport af: Marianne Wilcken Bjerregaard
Vejleder: Albert Chr. Paulsen

424/03 Dielectric and Shear Mechanical Relaxation in Glass Forming Liquids
- A thorough analysis and experimental test of the DiMarzio-Bishop model
Master thesis in physics by: Kristine Niss and Bo Jakobsen
Supervised by: Niels Boye Olsen

425/03 Fysiske forklaringer i undervisning
Specialerapport af: Kirsten Ringgard Jensen
Vejleder: Jens Højgaard Jensen