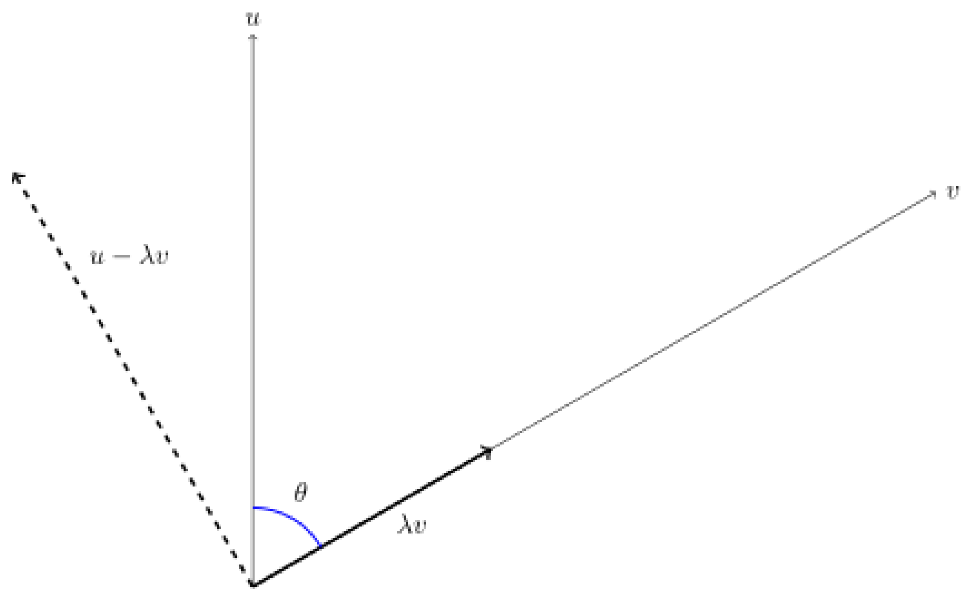
TAL OG LINEÆR ALGEBRA

Niels Lauritzen

Marcel Bökstedt

Pdf version 1.12.2017 af interaktiv bog: https://edtech.dk/LinAlg



Indhold

float table

Kapitel 0

Indledning

Dette materiale er en bearbejdning af noter som oprindeligt er kodet, skrevet og brugt af Niels Lauritzen. Hans stil er meget personlig og engagerende, og jeg har beholdt det meste af det han har forfattet. Niels kan godt lide at fortælle sin personlige mening, og det må jeg hellere lade ham blive ved med. Det vil sige, hvis der senere i denne tekst står “jeg” så er ikke mig men ham, hvis I forstår hvad jeg mener. Når jeg (altså ikke ham, men mig) vil sige noget, så vil jeg omtale mig selv som Marcel, hvilket jo også lyder næsten ens.

Formålet med dette kursus er at lære at bruge lineær algebra. Nu er det sådan, at vi matematiker gerne vil gøre rede for hvorfor de metoder vi anbefaler rent faktisk virker. Det gør vi i beviser, som er matematikers måde at overbevise tvivlere.

0.1 Bevis

Vi mener selv at metoden at bruge beviser på dem er mere humant end alternativet at brænde dem på bål. ♠

Men i dette kursus er vores hovedformål ikke at gøre rede for alle detaljer i argumenterne. Derfor er vi gået med til det kompromis at vi giver de matematisk korrekte argumenter, men vi skjuler dem. Hensigten er at beviserne skal være skrevet ud i alle detaljer. På den ene side betyder det at man bør kunne arbejde sig igennem hvert enkelt argument, og forstå hvorfor de forskellige påstande er sande. På den anden side gør det at nogle af argumenterne fylder meget —man står som en okse foran en mur af tekst. Hvis I ønsker at få den fulde sandhed at vide, må I altså meget gerne klikke på de knappe som er spredt ud over teksten og er mærket “bevis” eller lignende. Da vi I blive oplyst, men at kunne gengive disse beviser er ikke en del af pensum. Der er alligevel tre gode grunde til at studere i det mindste nogle af beviserne. Den første grund er den indlysende at det giver en bedre forståelse af stoffet. Den anden grund er at de fleste af beviserne er gode og ikke altfor vanskelige øvelser i at bruge teorien, så at det faktisk kan være en mindst lige så god øvelse at studere et bevis som at regne en numerisk opgave. Den tredie grund er at beviserne øver i at læse en matematisk tekst. Beviserne i disse noter minder meget om tankegangen i andre matematiske artikler og bøger, men ambitionen er at de skal være mere udførlige end hvad man normalt finder i sådanne tekster.

Et par af beviserne er relativt indviklede, og går udover hvad man forventer af en "øvelse".

0.2 Bevis *

Vi advarer om disse beviser ved at markere dem med en lille stjerne. ♠

Selv om I aldrig læser et eneste bevis, så husk på at det her er ikke et rent regnekursus. I fremtiden er risikoen relativt stor for at I kommer ud for at løse ligninger eller for at på anden måde bruge lineær algebra. Da vil I ikke sidde ned ved et bord og løse ligninger i hånden på et stykke papir i skæret fra et vokslys, men selvfølgelig vil I fodre en computer med ligningerne. Det ville vi også gøre. De regninger vi laver her er primært øvelser for at lære at fortolke de resultater som en maskine giver. Det nytter ikke meget at man ejer en computer som kan beregne egenvektorer som en mis, hvis man ikke er klar over hvad en egenvektor er, og hvad de kan bruges til.

Teorien bygger på en ret abstrakt begrebsdannelse som er vigtig selv om man kun interesserer sig for anvendelser. I har lov til at stole på vores beviser, men I *skal* lære at forstå de begreber vi kommer til at arbejde med. Og erfaringen siger at den bedste måde at lære at forstå disse begreber er ved at bruge dem i praksis, det vil sige ved at selv lave konkrete udregninger.

Noterne er skrevet med henblik på at de skal læses på en computer, tablet eller lignende. Det vil sige, vi gør flittigt brug af html. Vi har lavet en pdf version som kan printes ud, men vi anbefaler at bruge skærmen! Et tip: Hvis det bliver anstrengende for øjnene at læse fra skærmen, kan det være en god idé at invertere farverne (så at skriften bliver hvid på sort baggrund). Mange netlæsere og computers kommer i dag med udvidelser eller indstillinger der gør det muligt.

Hvis der er noget i de her noter som ikke er optimalt, eller måske ren sort snak, så ville det glæde mig (Marcel) meget hvis I bruger annoteringssystemet. Det virker sådan at I kan lave en kommentar som bliver synlig i teksten. Jeg vil skynde mig at svare på sådanne kommentarer. Den nøjagtige fremgangsmåde er såre simpel, og forklaret af Niels i denne video:

VIDEO: https://youtu.be/wACe3pPRFEg

Der findes mange bøger om lineær algebra, og de fleste af dem er udmærkede. I Århus har man på forskellige tidspunkter brugt Nielsen og Salomonsen "Lineær algebra via eksempler", Leon "Linear algebra with applications" og du Plessis "Forelæsningsnoter i lineær algebra". I København bruger man Hesselholt og Wahl "Lineær algebra". Hvis i kigger i nogle af disse eller i andre lignende kilder tager i sikkert ikke varig skade af det.

Kapitel 1

Om tal og vektorer i planen

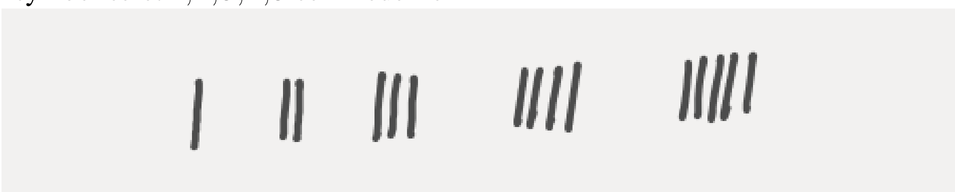
VIDEO: https://youtu.be/80nJknoes7A

I dette kapitel vil vi introducere tal helt fra de naturlige tal $0, 1, 2, \ldots$, komme ind på negative tal og brøker for til sidst at berøre de reelle tal og vektorer i planen. Alt dette for at lede frem til de komplekse tal i næste kapitel. De komplekse tal er fantastiske, men for at komme i dybden med dem kræves at man er fortrolig med lidt trigonometri og vektorer i planen.

Noget af det første man lærer som et lille barn er at tælle:

$$1, 2, 3, 4, \ldots$$

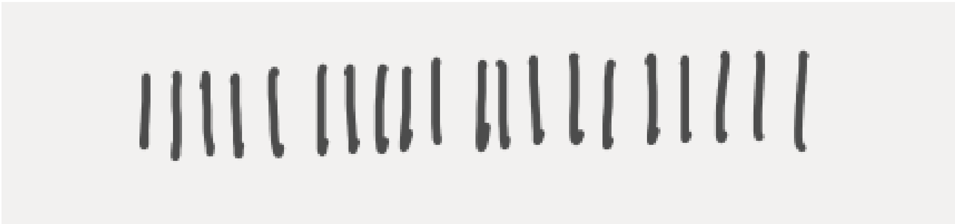
Det at tælle er fundamentalt for mennesket og arkæologiske kilder nævner at mennesket har gjort det i mindst 50.000 år. Man må formode at tællesymbolerne dengang har været primitive uden avancerede symboler som $0, \ldots, 9$. Måske har man symboliseret $1, 2, 3, 4, 5$ som nedenfor



Tallenes historie er fascinerende.

1.1 Det mirakuløse tal nul

Notationen med pinde er ikke specielt økonomisk. Vores titalssystem er i dag så indgroet i vores kultur at vi synes det er spild af blæk at skrive tallet 20 som



Man må ikke glemme at notationen 20 indeholder visdom gemt i årtusinders ophobet menneskelig erfaring. Ved at indføre tallet 0 kan man tælle i grupper af $1, 10, 100, \ldots$. Således dækker notationen 20 over at man tæller 0 grupper af 1 og 2 grupper af 10. I det hele taget er symbolet 0 et mirakel, som har bragt menneskeheden betydeligt videre efter det blev indført at den indiske matematiker Brahmagupta i 628. At have et specielt symbol for ingenting er en smuk abstraktion.

Omkring computerens opfindelse kom der mere fokus på at man nødvendigvis ikke behøver at tælle med hensyn til grupper af størrelser $1, 10, 100, \ldots$. Man kan også tælle binært det vil sige med hensyn til grupper af størrelser $1, 2, 4, 8, 16, \ldots$. I det binære talsystem kan de 20 pinde skrives

$$10100.$$

Måske er dette mere elegant - enten er en gruppe der (1) eller også er den der ikke (0). Læg mærke til at tallet 0 er central lige meget hvilket talsystem man vælger.

1.1 Quiz

Brahmagupta opfandt tallet 0 i år 628. Hvad gælder om tallet 628?

Det skrives 1010001000 i det binære talsystem.

Det skrives 1001110100 i det binære talsystem.

I det oktale talsystem det vil sige med hensyn til grupper af størrelse $1, 8, 64, 512, \ldots$ skrives 628 som 1210.

Det skrives DCXXVIII som romertal. ♠

1.2 De naturlige tal

Rent matematisk har de naturlige tal egentlig ikke noget at gøre med i hvilket talsystem de bliver skrevet op. I den abstrakte matematiske verden giver det god mening at bruge pinde til at repræsentere naturlige tal. Faktisk er det sådan man indfører de naturlige tal med den aksiomatiske metode under navnet Peanos aksiomer.

For nemheds skyld vil vi definere de naturlige tal til at starte med 0 og være

$$0, 1, 2, \dots,$$

hvor vi allerede ved hvordan man adderer og multiplicerer naturlige tal. Mængden af naturlige tal betegnes med $\mathbb{N}$. Årsagen til denne notation skyldes at man i århundreder har brugt tavle og kridt som kommunikationsmiddel. På en tavle er det svært at skrive et boldface N. Det er nemmere at dekorere et N til et blackboard bold N som ovenfor.

1.3 De hele tal

Vi ved godt at der findes negative tal, men hvordan vil vi egentlig forklare dem? Skru tiden nogle hundrede år tilbage og forestil dig hvor svært det har været at komme fra at man har 0 kroner i sin pung og skylder 5 kroner væk til at abstrahere og sige at man har -5 kroner i sin pung. Måske er det også svært at forestille sig i dag.

I matematikkens verden drejer det sig om at kunne løse ligninger. Indenfor de naturlige tal kan man ikke løse de enkleste ligninger, hvor man kun har lov at benytte naturlige tal, x og $+$ indenfor de naturlige tal. For eksempel har ligningen

$$x + 5 = 3 \tag{1.1}$$

ikke løsninger i de naturlige tal. Mere formelt skriver vi at der ikke findes noget $x \in \mathbb{N}$, som opfylder at $x + 5 = 3$. For at kunne løse denne ligning bliver nødt til at udvide de naturlige tal til de hele tal

$$\dots, -3, -2, -1, 0, 1, 2, 3, \dots$$

som betegnes $\mathbb{Z}$ (for Zahlen på tysk). I mængden af heltal har ligningen (1.1) løsningen $x = -2$.

1.4 De rationale tal

Der er stadig ret enkle ligninger som for eksempel

$$2x = 1, \tag{1.2}$$

som vi ikke kan løse med et heltal $x \in \mathbb{Z}$. Det er grunden til at vi indfører mængden af brøker eller rationale tal, som betegnes $\mathbb{Q}$. Brøker er som bekendt tal af formen p/q, hvor $p \in \mathbb{Z}$ og $q \in \mathbb{N}$ med betingelsen at $q \neq 0$. Ligningen (1.2) kan løses med brøken $x = \frac{1}{2}$.

1.2 Quiz

Hvad gælder om ligningen

$$\frac{1}{2}+\frac{x}{3}=1?$$

Den har to løsninger.

Ingen løsninger til ligningen er naturlige tal.

Der er kun en løsning, som er et rationalt tal mellem 1 og 2.

Der er kun en løsning, som er et rationalt tal mellem 0 og 1. ♠

1.4.1 Eksempel

Det er ikke svært at lægge heltal sammen eller gange to brøker sammen, men forestil dig nu at du har glemt hvordan man lægger brøker sammen. Kunne du benytte hjernekraft til at finde ud af det, blot ud fra indfaldsvinklen med at man skal kunne løse ligninger? Lad os tage eksemplet

$$s=\frac{1}{2}+\frac{1}{3}. \tag{1.3}$$

Vi er godt klar over at s bestemt ikke er lig med

$$\frac{1+1}{2+3}=\frac{2}{5}$$

i og med at s må være større end $\frac{1}{2}$. Men hvordan finder vi s som brøk? Her hjælper ligninger os. Vi ved at $x=\frac{1}{2}$ og $y=\frac{1}{3}$ er løsninger til ligningerne

$$\begin{aligned}2x&=1\\3y&=1.\end{aligned}$$

Hvis vi nu ganger første ligning med 3 og anden ligning med 2 får vi ligningerne

$$\begin{aligned}6x&=3\\6y&=2.\end{aligned}$$

Disse to ligninger kan vi nu lægge sammen og få ligningen

$$6x+6y=6(x+y)=5.$$

Derfor er

$$\frac{1}{2}+\frac{1}{3}=x+y=\frac{5}{6}.$$

Et alternativ til ligningerne kunne være at gange ligningen (1.3) igennem med $6 = 2 \cdot 3$ og få

$$6s = 6\left(\frac{1}{2}+\frac{1}{3}\right) = \frac{6}{2}+\frac{6}{3} = 3+2 = 5 \quad \text{detvilsige} \quad s = \frac{5}{6}.$$

Indrømmet, matematisk har vi snydt en smule her. Faktisk har vi også brug for at sige hvornår to brøker er ens, som for eksempel

$$\frac{5}{6} = \frac{10}{12},$$

men det er en anden historie.

1.3 Opgave

Efter præcis samme metode som i eksemplet kan vi finde (med bogstaver) formlen

$$\frac{a}{s}+\frac{b}{t} = \frac{at+bs}{st} \tag{1.4}$$

for addition af de to brøker $\frac{a}{s}$ og $\frac{b}{t}$. Prøv langsomt at gå igennem metoden, som følger: $x = \frac{a}{s}$ og $y = \frac{b}{t}$ er løsninger til ligningerne

$$\begin{aligned} sx &= a \\ ty &= b. \end{aligned}$$

Ved at gange første ligning med t og anden ligning med s og addere ligningerne fremkommer formlen (1.4).

♠

1.4 Opgave

Diofants ungdom varede $1/6$ af hans liv. Han fik skæg efter $1/12$ mere. Efter $1/7$ mere blev han gift. Fem år senere fik han en søn. Sønnen levede halvt så længe som faderen og Diofant døde fire år efter sønnen. Hvor gammel blev Diofant? ♠

1.5 De reelle tal

Vi begyndte med de naturlige tal $\mathbb{N}$ og kunne ikke løse enkle ligninger. Så udbyggede vi til de hele tal $\mathbb{Z}$, men kunne her stadig ikke løse helt simple ligninger som $2x = 1$. Det gjorde at vi "opfandt" brøker eller de rationale tal $\mathbb{Q}$. Her har vi at gøre med tal, hvor man kan addere, subtrahere, multiplicere og dividere (med alle tal undtagen 0). Med symboler har vi lavet kæden

$$\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q}.$$

Til hverdag omgiver vi os praktisk taget kun med rationale tal. Computere kan strengt taget kun håndtere rationale tal. Men rationale tal kan sagtens være overordentligt komplicerede med store tællere og nævnere som for eksempel

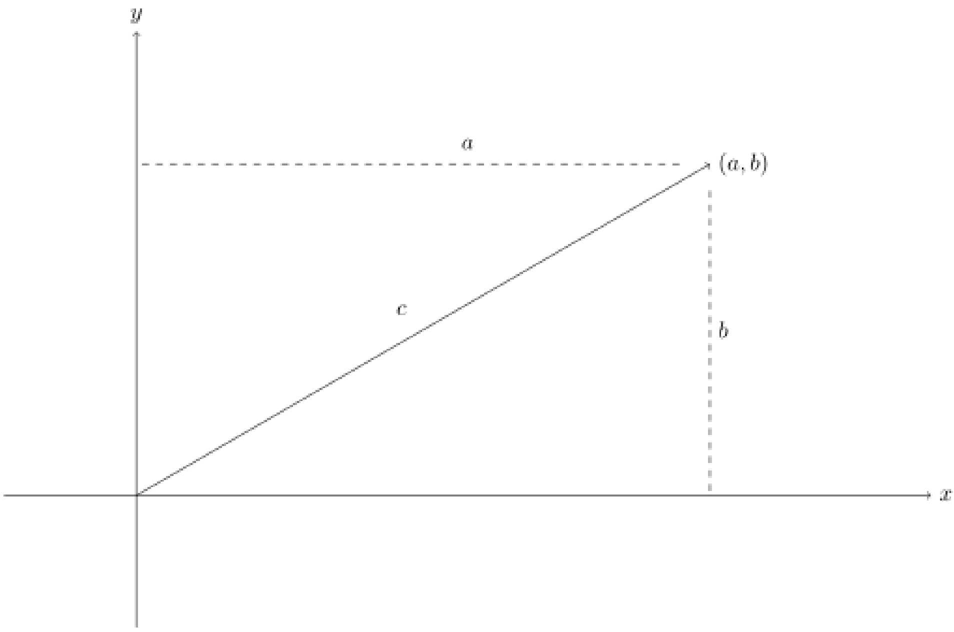
$$\frac{237894619875623478956127341734059237458923745237528903475203857}{534058927340589273450923475029345723049587234095723408957121232}$$

Findes der andre tal end de rationale?

Her støder vi på et af de mest overraskende elementer i matematikkens historie. Svaret er ja og skal findes i Pythagoras' læresætning om længden af hypotenusen i en retvinklet trekant. Som du helt givet husker, siger Pythagoras for en retvinklet trekant med hypotenuselængde c og med katetelængder a og b at

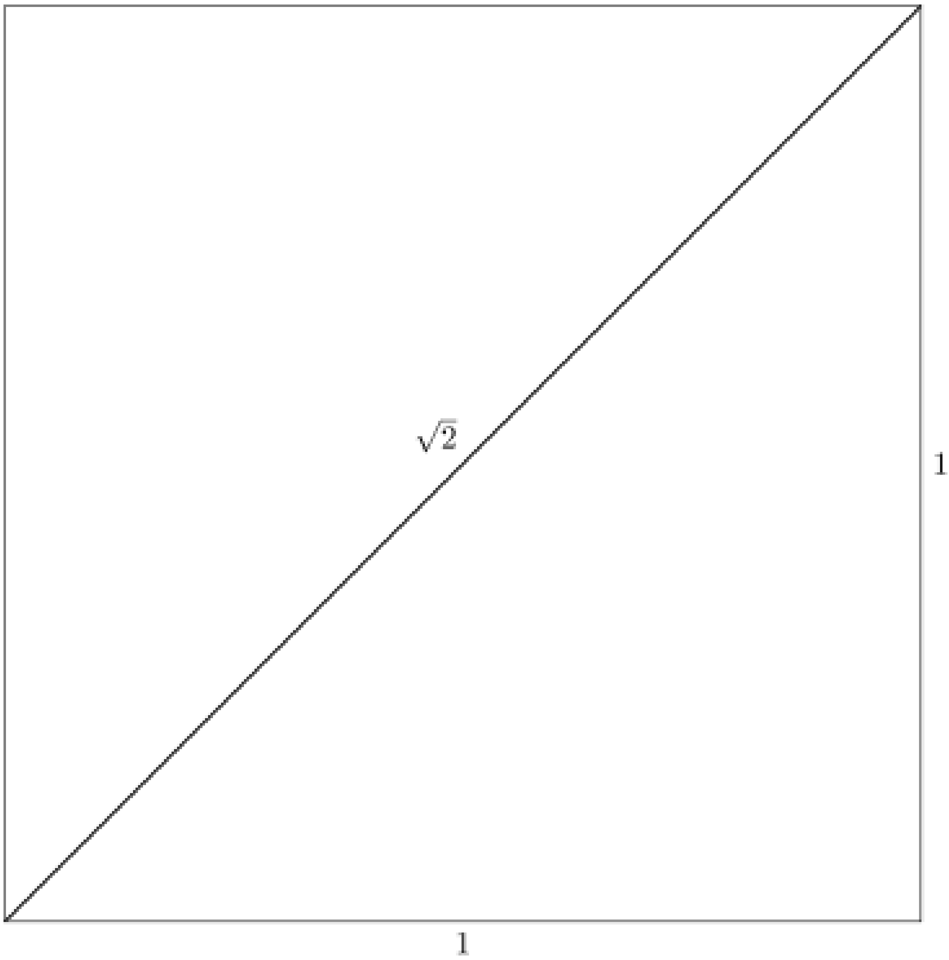
$$a^2 + b^2 = c^2.$$

Vi kan illustrere det med vektorer i et koordinatsystem:



Her siger Pythagoras at længden af vektoren med koordinaterne (a,b) er $\sqrt{a^2+b^2}$ eller i mere dagligdags sprog: Længden af diagonalen i et rektangel med sidelængder a og b er $\sqrt{a^2+b^2}$.

Den totale overraskelse er at længden af diagonalen i et rektangel, hvor begge sidelængder er 1 (det vil sige et kvadrat med sidelængde 1) ikke er et rationalt tal. Tænk lige over det. Noget så naturligt som længden af diagonalen nedenfor er ikke en brøk!



Et af de mest berømte matematiske argumenter, flere tusinde år gammelt, er netop et bevis for at længden af diagonalen ovenfor ikke kan skrives som en brøk. Det er ren matematik, når den er allerbedst.

1.5 Opgave

Hovedingrediensen i det matematiske bevis for at kvadratroden af 2 ikke er et rational tal er følgende udsagn om naturlige tal: Kvadratet at et ulige tal er ulige f.eks., $3^2 = 9, 5^2 = 25, 7^2 = 49$. Kan du lave et bevis for at udsagnet gælder for alle ulige tal? ♠

VIDEO: https://youtu.be/nVFJepH6Q0M

Selvom man til daglig egentlig ikke har brug for irrationale tal, er det i matematikken ekstremt vigtigt at kunne håndtere tal som $\sqrt{2}$. I abstrakt matematik kan man vise at der faktisk er langt flere irrationale tal end rationale. Disse udgør tilsammen de reelle tal, som betegnes $\mathbb{R}$. Det er en anelse teknisk at konstruere de reelle tal matematisk, men vi vil alligevel benytte dem, når vi regner med vektorer i planen.

1.6 Vektorer i planen

En vektor **v** i planen er givet ved dens koordinater $\begin{pmatrix} x \\ y \end{pmatrix}$, som er ordnede par af reelle tal $x, y \in \mathbb{R}$. Af typografiske hensyn skrives vektoren v også som (x, y).

Det er meget naturligt at lægge to vektorer sammen og gange en vektor med et tal på følgende måde: Betragt vektorerne, og bemærk at vektorer er fede!

$$\mathbf{u} = \begin{pmatrix} a \\ b \end{pmatrix} \quad \text{og} \quad \mathbf{v} = \begin{pmatrix} c \\ d \end{pmatrix}$$

samt tallet λ. Så er summen $\mathbf{u} + \mathbf{v}$ lig med

$$\begin{pmatrix} a + c \\ b + d \end{pmatrix}$$

og skalarmultiplikationen $\lambda \mathbf{u}$ lig med

$$\begin{pmatrix} \lambda a \\ \lambda b \end{pmatrix}.$$

Vi betegner mængden af vektorer givet ved deres koordinater som $\mathbb{R}^2$. Fra din baggrund i matematik ved du at prikproduktet mellem **u** og **v** er givet ved formlen

$$\mathbf{u} \cdot \mathbf{v} = ac + bd.$$

Prikproduktet har en masse gode egenskaber, herunder

$$(\lambda \mathbf{u}) \cdot \mathbf{v} = \lambda(\mathbf{u} \cdot \mathbf{v})$$
$$\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w},$$

(1.6) DEFINITION.

Længden af vektoren **u** *er givet ved formlen*

$$|\mathbf{u}| = \sqrt{\mathbf{u} \cdot \mathbf{u}} = \sqrt{a^2 + b^2} \qquad (\text{NB} : |\mathbf{u}|^2 = \mathbf{u} \cdot \mathbf{u}).$$

En vektor **v** *siges at være en enhedsvektor, hvis den har længde* 1 *det vil sige* $|\mathbf{v}| = 1$.

Vektorerne **u** *og* **v** *siges at være vinkelrette på hinanden hvis* $\mathbf{u} \cdot \mathbf{v} = 0$.

VIDEO: https://youtu.be/Hz7fNEayQSU

1.7 Quiz

Betragt vektorerne

$$\mathbf{u} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad \text{og} \qquad \mathbf{v} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

For hvilket λ er $\mathbf{u} - \lambda \mathbf{v}$ vinkelret på $\mathbf{v}$?

$\lambda = 1.$

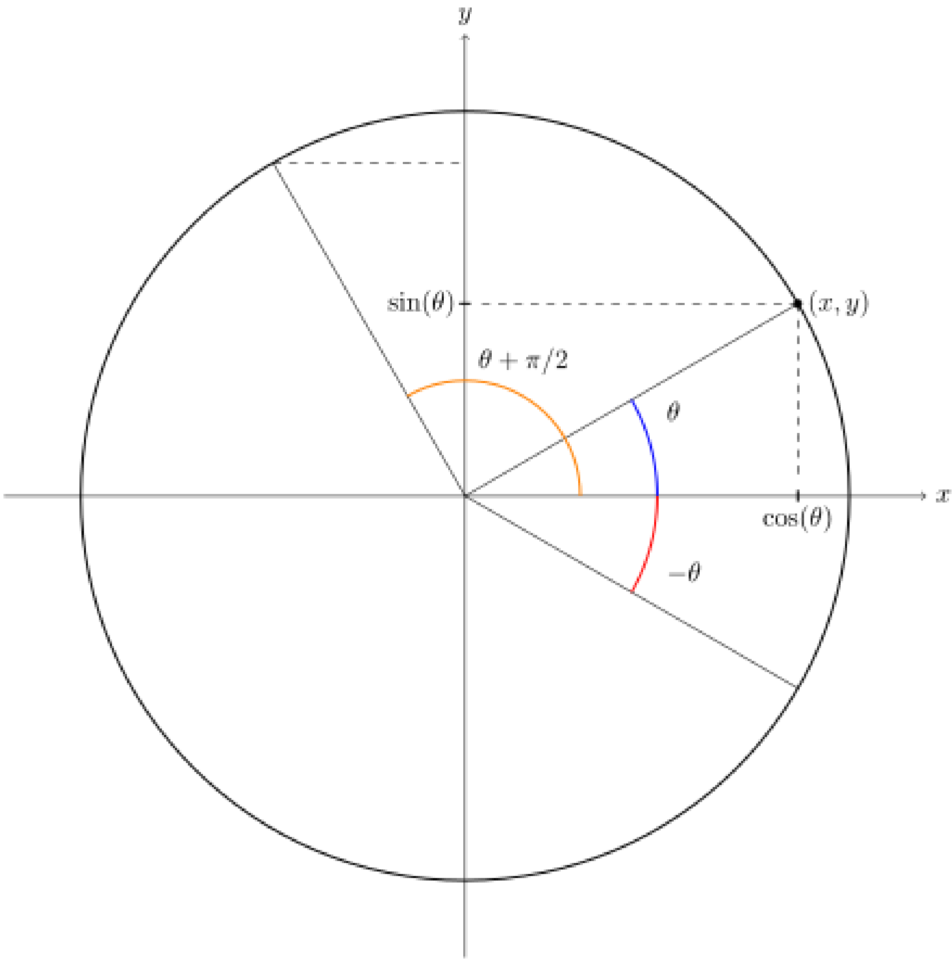
$\lambda = \frac{2}{3}.$

$\lambda = \frac{4}{5}.$

$\lambda = \frac{3}{5}.$ ♠

1.7 Cosinus og sinus til en vinkel

Vi repeterer cosinus og sinus af vinkler. Enhedscirklen nedenfor er netop defineret som mængden af enhedsvektorer det vil sige vektorer med længde 1.



Enhedsvektoren (x,y) på tegningen er entydigt givet ud fra dens vinkel θ med x-aksen. Cosinus, $\cos(\theta)$, til vinklen θ er defineret som x-koordinaten og sinus, $\sin(\theta)$, som y-koordinaten til enhedsvektoren. Denne definition giver omgående den velkendte formel

$$\cos^2(\theta)+\sin^2(\theta)=1.$$

Ud fra tegningen ovenfor kan man også aflæse følgende ligninger:

$$\cos(-\theta)=\cos(\theta) \tag{1.5}$$
$$\sin(-\theta)=-\sin(\theta) \tag{1.6}$$
$$\cos(\theta+\pi/2)=-\sin(\theta) \tag{1.7}$$
$$\sin(\theta+\pi/2)=\cos(\theta) \tag{1.8}$$

VIDEO: https://youtu.be/3XUfu-Lpx1A

1.8 Quiz

Lad x være cosinus til 45 grader (eller $\pi/4$). Hvad gælder om x?

$$x = \sin(\pi/4).$$

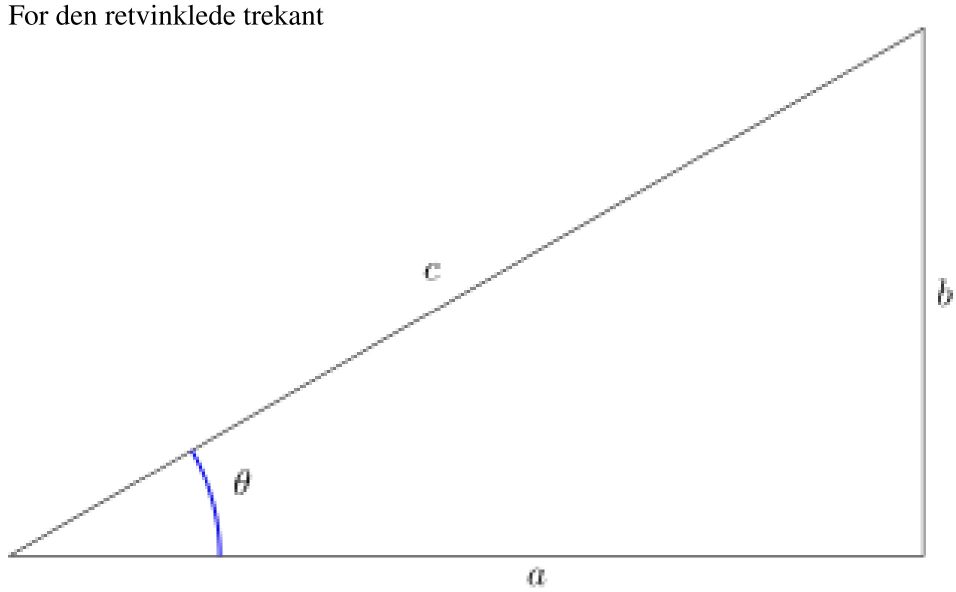
$$2x^2 = 1.$$

$$x = \frac{\sqrt{2}}{2}.$$

$$x = \sin(3\pi/4).$$

♠

For den retvinklede trekant



kan man også via definitionen af cosinus og sinus ud fra enhedscirklen finde frem til formlerne

$$c\cos(\theta) = a$$
$$c\sin(\theta) = b$$

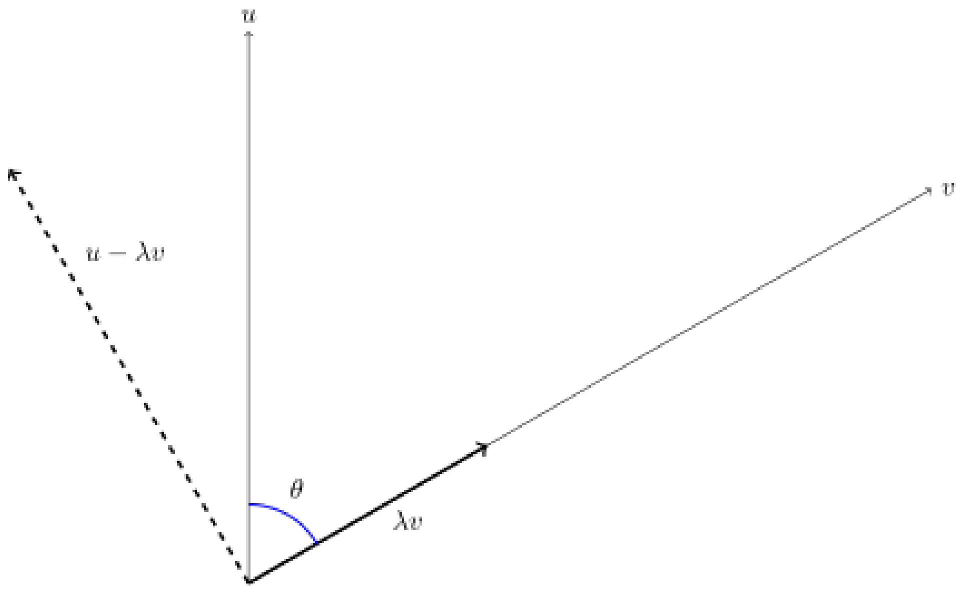
Disse formler er meget nyttige, når man skal regne på vektorer i planen.

1.9 Opgave

Findes en retvinklet trekant med sidelængder 3, 4 og 5? I givet fald, bestem vinklerne i denne trekant. ♠

1.8 Projektionen af en vektor på en anden vektor

Givet to vektorer **u** og **v** som nedenfor, hvor meget (λ) skal vi forkorte eller forlænge **v** med for at afstanden mellem $\lambda\mathbf{v}$ og **u** bliver mindst mulig?



Der er her tale et minimeringsproblem. Vi kender vektorerne **u** og **v** og skal finde tallet λ så længden

$$|\mathbf{u} - \lambda\mathbf{v}|$$

af vektoren $\mathbf{u} - \lambda\mathbf{v}$ bliver minimal. Det er præcis det samme som at finde λ, som minimerer funktionen

$$\begin{aligned} f(\lambda) &= |\mathbf{u} - \lambda\mathbf{v}|^2 = (\mathbf{u} - \lambda\mathbf{v}) \cdot (\mathbf{u} - \lambda\mathbf{v}) = \mathbf{u} \cdot \mathbf{u} - 2\lambda\mathbf{u} \cdot \mathbf{v} + \lambda^2\mathbf{v} \cdot \mathbf{v} \\ &= |\mathbf{v}|^2 \lambda^2 - 2(\mathbf{u} \cdot \mathbf{v})\lambda + |\mathbf{u}|^2. \end{aligned}$$

Faktisk er $f(\lambda)$ en parabel (i λ), som vender benene opad og med bundpunkt for

$$\lambda = -\frac{-2(\mathbf{u} \cdot \mathbf{v})}{2\,|\mathbf{v}|^2} = \frac{\mathbf{u} \cdot \mathbf{v}}{|\mathbf{v}|^2}.$$

Med denne værdi for λ gælder

$$(\mathbf{u} - \lambda\mathbf{v}) \cdot \mathbf{v} = 0$$

det vil sige vektorerne $\mathbf{u} - \lambda \mathbf{v}$ og $\mathbf{v}$ er vinkelrette. Måske ikke så overraskende ud fra tegningen ovenfor. Vektoren $\lambda \mathbf{v}$ kaldes for projektionen af $\mathbf{u}$ på $\mathbf{v}$.

VIDEO: https://youtu.be/th5cvI0NFos

1.10 Quiz

Lad d betegne afstanden fra punktet $(1,1)$ til linjen gennem $(0,0)$ og $(2,1)$. Hvad gælder om d?

$$d = \frac{1}{2}.$$

$$d = 0.447214.$$

$$d = \frac{\sqrt{5}}{5}.$$

$$d = \frac{2}{\sqrt{5}}.$$

♠

Ud fra formlerne for retvinklede trekanter får vi

$$|\mathbf{u}| \cos(\theta) = |\mathbf{v}|\, \lambda$$

og dermed den smukke formel for cosinus til vinklen θ mellem vektorerne $\mathbf{u}$ og $\mathbf{v}$:

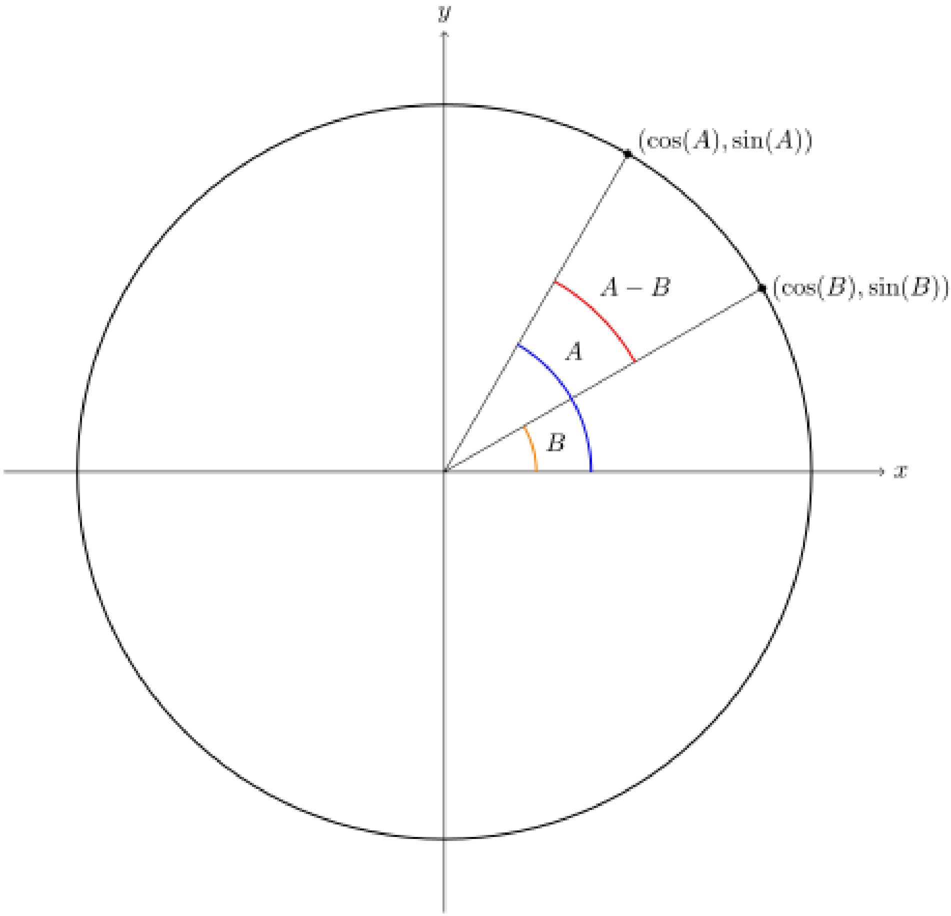
$$\cos(\theta) = \frac{\mathbf{u} \cdot \mathbf{v}}{|\mathbf{u}|\,|\mathbf{v}|}. \tag{1.9}$$

1.9 Cosinus og sinus for summen af to vinkler

Man kan ret nemt overbevise sig om at $\cos(A+B)$ ikke er lig med

$$\cos(A)+\cos(B)$$

for to vinkler A og B. For eksempel er $1=\cos(0)=\cos(0+0)$ ikke lig med $\cos(0)+\cos(0)=2$. Men findes der en formel, som udtrykker $\cos(A+B)$ ved hjælp af cosinus og sinus til A og B?



Ud fra tegningen ovenfor kan vi udlede en formel for cosinus til differencen mellem de to vinkler A og B. Da de to vektorer er enhedsvektorer er deres prikprodukt ud fra den smukke formel (1.9) netop lig med cosinus til forskellen mellem deres vinkler dvs.

$$\cos(A-B)=\begin{pmatrix}\cos(A)\\ \sin(A)\end{pmatrix}\cdot\begin{pmatrix}\cos(B)\\ \sin(B)\end{pmatrix}=\cos(A)\cos(B)+\sin(A)\sin(B). \tag{1.10}$$

VIDEO: https://youtu.be/AzI2QAkqCcs

Ved at lave et mindre hack og udskifte B med $-B$ i formlen (1.10) får vi

$$\cos(A+B) = \cos(A-(-B)) = \cos(A)\cos(-B) + \sin(A)\sin(-B)$$

Ved nu at benytte $\cos(-B) = \cos(B)$ og $\sin(-B) = -\sin(B)$ (se de grundlæggende trigonometriske formler) kommer vi frem til formlen

$$\cos(A+B) = \cos(A)\cos(B) - \sin(A)\sin(B). \tag{1.11}$$

Hvad med additionsformler for sinus? Her benyttes igen de trigonometriske formler til at slutte

$$-\sin(x) = \cos\left(x+\frac{\pi}{2}\right) \quad \text{samt} \quad \cos(x) = \sin\left(x+\frac{\pi}{2}\right)$$

det vil sige

$$\sin(A+B) = -\cos\left(A+B+\frac{\pi}{2}\right) = -\left(\cos(A)\cos\left(B+\frac{\pi}{2}\right) - \sin(A)\sin\left(B+\frac{\pi}{2}\right)\right).$$

og dermed

$$\sin(A+B) = \cos(A)\sin(B) + \sin(A)\cos(B). \tag{1.12}$$

Nu overlades det som en opgave til læseren at overbevise sig om at den sidste formel

$$\sin(A-B) = \sin(A)\cos(B) - \sin(B)\cos(A) \tag{1.13}$$

gælder.

1.10 Opgaver

1.10.1

1.11 Quizopgave

Lad d betegne afstanden fra punktet $(2,1)$ til linjen gennem $(0,0)$ og $(1,3)$. Hvad gælder om d?

$$d > \frac{3}{2}.$$

$$1 < d < \frac{3}{2}$$

$$2d^2 = 5.$$

$$3d^2 = 5.$$

♠

1.10.2

Giv et præcist argument for at

$$\sin(2\theta) = 2\sin(\theta)\cos(\theta).$$

1.10.3

Lommeregneren siger at $\sin(\pi/3)$ cirka er 0.866. Giv et geometrisk argument for at

$$\sin(\pi/3) = \frac{\sqrt{3}}{2}$$

ved hjælp af en retvinklet trekant, hvor de to ikke rette vinkler er $\theta = \pi/3$ og $\theta/2 = \pi/6$ (det vil sige henholdsvis 60 og 30 grader).

Kapitel 2

De komplekse tal

VIDEO: https://youtu.be/6WHkB$_n2mos$

Lad os vende tilbage til historien om ligninger og tal. Det vil føre for vidt at forklare, hvordan man systematisk indfører tal, som ikke behøver være brøker. Kort sagt laver man dem som uendelige decimaltal, men detaljerne er ret kedelige og lidt besværlige. Så vi vil antage at vi ved hvad (reelle) tal er. Man kan ikke finde et reelt tal, som løser ligningen

$$x^2 + 1 = 0,$$

—intet reelt tal ganget med sig selv giver -1. Enhver enkel lommerregner vil blinke en fejlmeddelelse, hvis du forsøger at tage kvadratroden til -1. Derimod vil et moderne computer algebra system som Maple eller Mathematica formentlig give dig symbolet I som output. Hvad er dette I?

Lad os først komme ind på hvor anderledes (og mere spændende) de komplekse tal er med et eksempel fra computergrafik.

2.1 Mandelbrotmængden

Geometrisk opholder de komplekse tal sig i to dimensioner, mens de almindelige tal kun bevæger sig på en linje i en dimension.

Man kan benytte de komplekse tal til at generere såkaldte fraktaler som Mandelbrotmængden nedenfor. Mandelbrotmængden er blevet kaldt et af de smukkeste og mest komplicerede objekter i moderne matematik. I næste afsnit ser vi nærmere på addition og multiplikation af komplekse tal. Mandelbrotmængden er frembragt ene og alene ved at iterere multiplikationer og additioner med komplekse tal.

VIDEO: https://youtu.be/I9$_zAXotgbg$

2.2 Regneregler

Mandelbrotmængden fremkommer ved mange iterationer bestående af multiplikationer og additioner med komplekse tal. Vi har brug for at vide hvad komplekse tal er og hvordan man regner med dem.

En specielt vigtig regneregel for reelle tal hedder den distributive lov. Den siger at

$$x(y+z) = xy + xz$$

for $x, y, z \in \mathbb{R}$. En anden velkendt regel siger at faktorernes orden er ligegyldig eller at multiplikation er kommutativ det vil sige $xy = yx$.

2.1 Opgave

Forklar så detaljeret som muligt, hvorfor

$$(x+y)(z+w) = xz + yz + xw + yw$$

ved at bruge den distributive lov og den kommutative lov for multiplikation. Skal vi bruge den kommutative lov? Hvorfor? ♠

Lad os antage at vi oven i de reelle tal kaster et opdigtet eller *imaginært* tal i ind, som har egenskaben at

$$i^2 = -1.$$

Hvis a, b, c, d er reelle tal og vi antager at i adlyder de almindelige regneregler får vi følgende udregning:

$$\begin{aligned}
(a+bi)(c+di) &= (a+bi)c + (a+bi)di \\
&= ac + bic + adi + bidi \\
&= ac + bci + adi + bdi^2 \\
&= ac + bci + adi - bd \\
&= (ac - bd) + (ad + bc)i
\end{aligned}$$

Det vil sige to tal på formen $x + yi$, hvor x og y er reelle tal, ganger sammen til et tal af samme *standard form*.

VIDEO: https://youtu.be/$_3V2woRQfSo$

Faktisk kan man vise at, når man definerer multiplikation som ovenfor og addition som

$$(a+bi) + (c+di) = (a+c) + (b+d)i,$$

så får man en mængde af nye tal, som opfylder alle velkendte regneregler for reelle tal, som f.eks.

$$
\begin{aligned}
x(yz) &= (xy)z \\
x(y+z) &= xy+xz \\
xy &= yx \\
1x &= x
\end{aligned}
$$

2.2 Opgave

Den ukronede konge blandt regnereglerne er reglen $x(yz) = (xy)z$, som kaldes den associative lov. Uformelt siger den at vi selv kan vælge hvordan vi udregner et produkt xyz: Det gør ikke nogen forskel om vi først ganger x sammen med y og så ganger z på eller om vi først ganger y sammen med z og så ganger x på. Det giver det samme resultat. Hvis vi ikke havde den associative lov ville (den aritmetiske) verden være kaotisk.

Faktisk kan vi allerede nu se at den associative lov gælder for vores nye tal af formen $a+bi$: Lad $x = a+bi, y = c+di$ og $z = e+fi$, hvor a,b,c,d,e,f er reelle tal.

1. Udregn $u = yz$ ved at sætte ind i formlen.
2. Udregn $v = xy$ ved at sætte ind i formlen.
3. Udregn xu og vz ved at sætte ind i formlen.
4. Overbevis dig nu om at $x(yz) = (xy)z$ det vil sige at $xu = vz$.

♠

Hvorfor benyttes notationen i for $\sqrt{-1}$? Begrundelsen er historisk og daterer sig tilbage til 1500-tallet, hvor den italienske matematiker Cardano havde behov for at regne med opdigtede eller *imaginære* tal for at finde en formel til løsning af tredjegradsligninger. Det komplekse tal i kaldes for *den imaginære enhed*. Man skal dog ikke forledes til at tro at de komplekse tal kun er et påfund opdigtet for fem lange sekler siden af en tosset matematiker. De dukker hele tiden op i anvendelser. Den ene af de to mest fundamentale fysiske teorier, kvantefysikken, er formuleret i termer af komplekse tal og vektorrum.

Hvad er så den anden mest fundamentale teori?

Einsteins generelle relativitetsteori. Den bruger også vektorrum, men ikke så meget de komplekse tal. ♠

(2.3) DEFINITION.

De komplekse tal er mængden

$$\mathbb{C} = \{x + yi \mid x, y \in \mathbb{R}\},$$

hvor multiplikation *er givet som*

$$(a+bi)(c+di) = (ac-bd) + (ad+bc)i$$

og addition *som*

$$(a+bi) + (c+di) = (a+c) + (b+d)i.$$

For et komplekst tal $z = x + iy$ defineres

1. Realdelen *som*

$$\mathrm{Re}(z) = x$$

2. Imaginærdelen *som*

$$\mathrm{Im}(z) = y.$$

3. Det komplekst konjugerede tal

$$\bar{z} = x - iy$$

To komplekse tal er identiske hvis og kun hvis deres real- og imaginærdele er ens.

Ved hjælp af formlen for multiplikation af komplekse tal ses for $z = c + di$, at

$$z\bar{z} = (c+di)(c-di) = c^2 + d^2$$

ved indsættelse af $a = c$, $b = d$, $c = c$ og $d = -d$. Denne observation gør at vi kan udlede følgende divisionsformel:

$$\begin{aligned} \frac{a+bi}{c+di} &= \frac{(c-di)(a+bi)}{(c-di)(c+di)} \\ &= \frac{(ac+bd)+(bc-ad)i}{c^2+d^2} \\ &= \frac{ac+bd}{c^2+d^2} + \frac{bc-ad}{c^2+d^2}i. \end{aligned}$$

Divisionen af to komplekse tal giver altså igen et komplekst tal. Vi antager at tælleren $c + di$ ikke er lig 0 det vil sige c og d ikke begge er 0 eller ækvivalent hermed at $c^2 + d^2 \neq 0$.

VIDEO: https://youtu.be/2WCToS0llzQ

En meget vigtig egenskab er at hvis $a+bi \neq 0$, så findes $c+di$ så $(a+bi)(c+di)=1$. Vi ved godt denne regel er korrekt for de reelle tal det vil sige, hvis $x \neq 0$ så findes et reelt tal y så $xy=1$. Her kan vi bare sætte $y=1/x$. På samme måde kan vi sætte $c+di=1/(a+bi)$.

2.4 Opgave

Find en formel for $1/(a+bi)$, hvor a og b ikke begge er 0. ♠

2.5 Opgave

1. Find $z \in \mathbb{C}$, som opfylder at $z(1+i)=(1+2i)$.
2. Find $z_1, z_2 \in \mathbb{C}$, som opfylder

$$\begin{aligned}(1-2i)z_1+z_2 &= 8+4i\\ z_1+iz_2 &= -3+5i\end{aligned}$$

♠

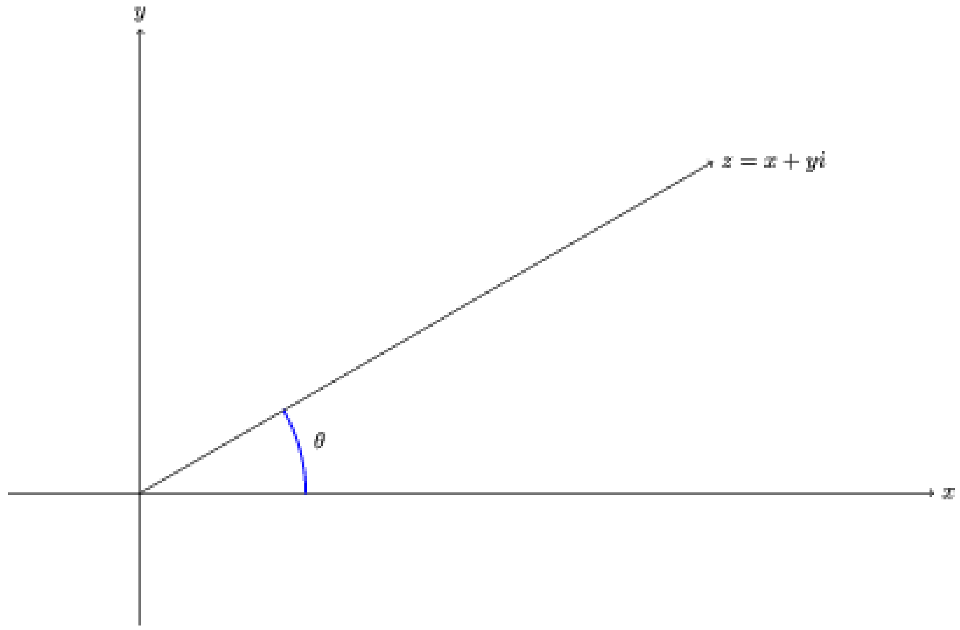
2.6 Lidt vanskelig Opgave

Vis at hvis $y_1y_2 \in \mathbb{C}$ og $y_1y_2=0$ så er enten $y_1=0$ eller $y_2=0$. Vis derefter at hvis $y_1,y_2,x \in \mathbb{C}$ og $y_1x=y_2x$, så er enten $x=0$ eller $y_1=y_2$. Det vil sige, der er ingen seriøse konkurrenter til titlen $1/x$. ♠

2.3 Geometrisk fortolkning og polær form

Den geometriske fortolkning af de komplekse tal blev først introduceret af den danske matematiker Caspar Wessel (1745-1818) i 1797.

Det forekommer ret naturligt at opfatte et komplekst tal $z=x+iy$, som punktet (x,y) i et koordinatsystem.



Med dette geometriske billede ligger det lige for at indføre følgende definitioner.

(2.7) DEFINITION.

Lad $z = x + iy$ være et komplekst tal. Længden af vektoren (x,y) kaldes modulus for z og betegnes $|z|$. Vinklen som (x,y) danner med x-aksen kaldes argumentet for z og defineres som en vinkel i intervallet $[0,2\pi)$. For et reelt tal $x \in \mathbb{R}$ definerer vi det komplekse tal

$$e^{ix} = \cos(x) + i\sin(x). \tag{2.1}$$

Læg mærke til at hvis θ er argumentet for det komplekse tal z, så er

$$\frac{z}{|z|} = e^{i\theta}. \tag{2.2}$$

Det er fordi at $z/|z|$ præcis svarer til vektoren (x,y) ganget med dens reciprokke længde. Denne vektor er en enhedsvektor, som danner vinklen θ med x-aksen. Det vil sige den svarer præcis til det komplekse tal $e^{i\theta}$.

Ligningen (2.2) giver den smukke geometriske repræsentation

$$z = |z|\,e^{i\theta} \tag{2.3}$$

af det komplekse tal z. Fremstillingen (2.3) kaldes for den polære form af z.

Denne repræsentation fortjener at blive kaldt smuk, fordi den afspejler sig mirakuløst i multiplikationen af komplekse tal: Lad z_1 og z_2 være to komplekse tal med argumenter henholdsvis θ_1 og θ_2. Så er

$$z_1 z_2 = |z_1|\,e^{i\theta_1}\,|z_2|\,e^{i\theta_2} = |z_1|\,|z_2|\,e^{i\theta_1}e^{i\theta_2} = |z_1|\,|z_2|\,e^{i(\theta_1+\theta_2)}. \tag{2.4}$$

Med ord har vi gjort rede for at

> Man multiplicerer to komplekse tal ved at multiplicere deres længder og addere deres argumenter.

Det sidste lighedstegn i (2.4) har vi rent faktisk ikke vist og det er da også et af hovedresultaterne:

(2.8) SÆTNING.

For to tal $x, y \in \mathbb{R}$ gælder formlen

$$e^{ix}e^{iy} = e^{i(x+y)}.$$

Bevis

Bevis. Vi bruger multiplikation af komplekse tal og ganger $e^{ix} = \cos(x) + i\sin(x)$ og $e^{iy} = \cos(y) + i\sin(y)$ sammen:

$$\begin{aligned}
&e^{ix}e^{iy} = (\cos(x) + i\sin(x))(\cos(y) + i\sin(y)) = \\
&\cos(x)\cos(y) - \sin(x)\sin(y) + i(\cos(x)\sin(y) + \sin(x)\cos(y)) = \\
&\cos(x+y) + i\sin(x+y) = e^{i(x+y)}.
\end{aligned}$$

Her forekommer miraklet i næstsidste og sidste linje ovenfor. Multiplikationen af komplekse tal indeholder additionsformlerne:

$$\begin{aligned}
\cos(x+y) &= \cos(x)\cos(y) - \sin(x)\sin(y) \\
\sin(x+y) &= \cos(x)\sin(y) + \sin(x)\cos(y)
\end{aligned}$$

for cosinus og sinus. □

♠

Læg i øvrigt mærke til verdens smukkeste formel

$$e^{i\pi} = -1.$$

Formlen kombinerer på minimal vis fire af de allervigtigste konstanter i matematikken: e, i, π og 1 og følger af definitionen i (2.1). XKCD har også sin mening om den sag.

> VIDEO: https://youtu.be/6HTzvbPhtE0

2.3.1 De Moivres formel

Abraham de Moivre var en fransk matematiker, som udover at beskæftige sig med sandsynlighedsteori også fik sit navn udødeliggjort gennem *De Moivres formel*. Denne formel siger i al sin enkelhed at der for et naturligt tal n og et reelt tal x gælder

$$(\cos(x) + i\sin(x))^n = \cos(nx) + i\sin(nx). \tag{2.5}$$

Det er ikke svært at bevise formlen via Sætning 8 ovenfor, som medfører at

$$(\cos(x) + i\sin(x))^n = (e^{ix})^n = e^{inx} = \cos(nx) + i\sin(nx).$$

Ikke desto mindre er (2.5) et mirakel, som markerer den stærke forbindelse mellem komplekse tal og trigonometriske funktioner. For eksempel kan man benytte De Moivres formel til at udlede formler som

$$\cos(3x) = 4\cos(x)^3 - 3\cos(x). \tag{2.6}$$

VIDEO: https://youtu.be/$_V p_{fJ} Hz2g$

2.9 Quiz

Hvad er $\cos(2x)$?

$2\sin(x)\cos(x) + 1$

$2\cos(x)^2 - 1$

Ingen af de foregående svarmuligheder. ♠

2.10 Opgave

Find en stamfunktion til $\cos(x)^3$. ♠

2.4 Andengradsligningen og højeregradsligninger

Lad os rette opmærksomheden mod ligningen

$$z^n = 1, \tag{2.7}$$

hvor n er et naturligt tal. Hvis vi kun begrænser os til de reelle tal, har (2.7) højst to løsninger (for eksempel for $n = 2$) og nogle gange kun en (for eksempel for $n = 3$). I de komplekse tals domæne har vi to dimensioner og kan boltre os både lodret og vandret. En løsning z til (2.7) bliver nødt til at have modulus 1 det vil sige $z = e^{i\varphi}$ med $\varphi \in [0, 2\pi)$. Da

$$z^n = e^{in\varphi} = \cos(n\varphi) + i\sin(n\varphi) = 1$$

har vi altså $m2\pi = n\varphi$ for et helt tal $m \in \mathbb{Z}$. Hvis φ er argumentet for z har vi altså kun mulighederne $m = 0, 1, \ldots, n-1$.

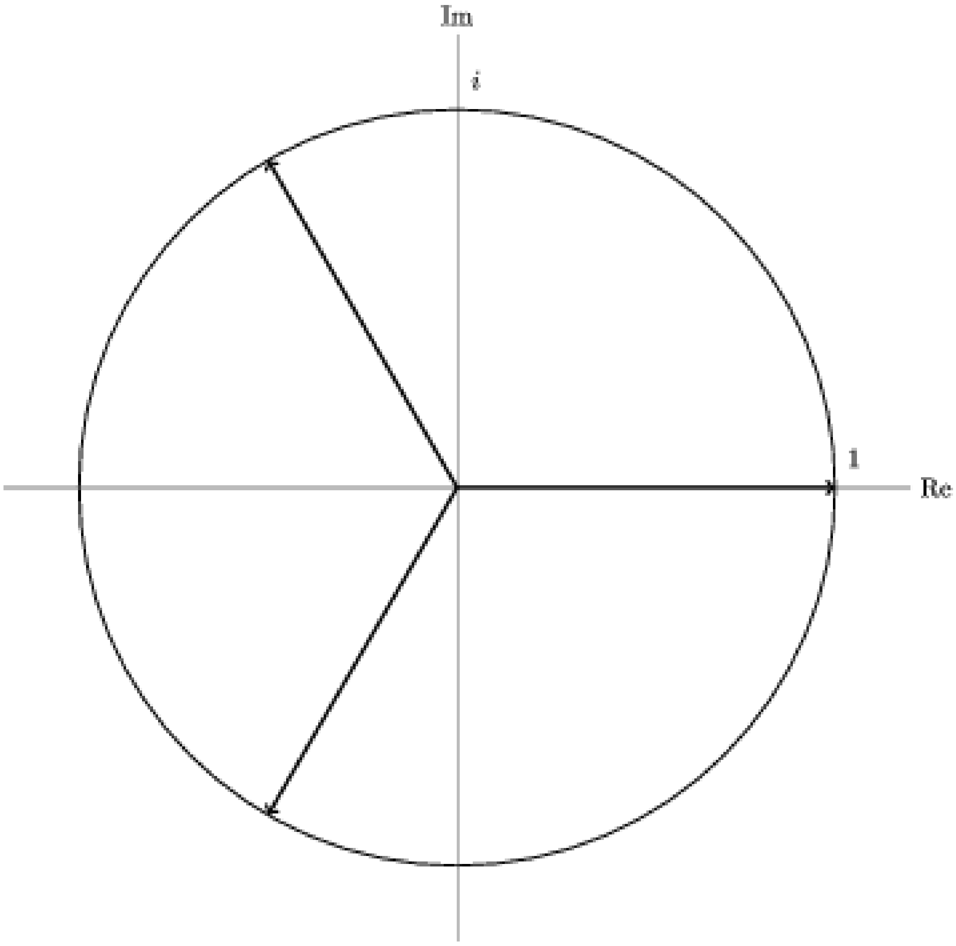
Dermed kan alle løsninger til (2.7) skrives som passende potenser af $\varepsilon_n = e^{i2\pi/n}$:

$$\varepsilon_n^m = e^{m\frac{2\pi i}{n}},$$

hvor $m = 0, 1, \ldots, n-1$. Vi har faktisk bevist at (2.7) altid har n forskellige løsninger over de komplekse tal.

2.11 Eksempel

Lad os som eksempel tage ligningen $z^3 = 1$. Den har løsningerne, som fremkommer ved at tredele enhedscirklen;



i $\mathbb{C}$ det vil sige

$$\begin{aligned}
z_0 &= 1 \\
z_1 &= e^{i\frac{2\pi}{3}} = -\frac{1}{2} + \frac{\sqrt{3}}{2}i \\
z_2 &= e^{i\frac{4\pi}{3}} = -\frac{1}{2} - \frac{\sqrt{3}}{2}i
\end{aligned}$$

♠

2.12 Quiz

Lad z være en løsning til $z^8 = 1$. Hvilke muligheder er der for z?

$|z| = 2$

$z = i$

Argumentet for z er $\pi/8$.

$$z = \frac{\sqrt{2}}{2} + i\frac{\sqrt{2}}{2}.$$

♠

(2.13) SÆTNING.

Lad n være et naturligt tal og $a = re^{i\theta}$ et komplekst tal med modulus $r \neq 0$ og argument θ. Så har ligningen

$$z^n = a \tag{2.8}$$

løsningen

$$z_0 = \sqrt[n]{r}e^{i\frac{\theta}{n}}.$$

Ligningen (2.8) *har n forskellige løsninger og de er*

$$z_0, z_0\varepsilon_n, \ldots, z_0\varepsilon_n^{n-1},$$

hvor $\varepsilon_n = e^{i\frac{2\pi}{n}}$.

Bevis

At opløfte et komplekst tal til n-te potens svarer til at opløfte dets modulus til n-te og gange dets argument med n. Derfor er $z_0^n = re^{i\theta}$ og dermed en løsning til $z^n = a$. Antag nu at $u^n = a$. Så vil

$$\frac{u^n}{z_0^n} = \left(\frac{u}{z_0}\right)^n = 1.$$

Dermed vil $\frac{u}{z_0}$ være en løsning til $z^n = 1$, hvorfor

$$u = z_0\varepsilon_n^m$$

for et eller andet m blandt $0, \ldots, n-1$. ♠

2.4.1 Den gode gamle andengradsligning

Her støder vi på det verdensberømte trick (completing the square):

$$az^2 + bz + c = a\left(z + \frac{b}{2a}\right)^2 - \frac{b^2}{4a} + c.$$

Ved en lettere omskrivning ses at løsninger til andengradsligningen $az^2 + bz + c = 0$ opfylder

$$\left(z + \frac{b}{2a}\right)^2 = \frac{b^2 - 4ac}{4a^2}.$$

Det giver så den klassiske formel

$$z = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a},$$

som giver rigtig god mening også for komplekse tal a, b, c. Vi har nemlig set i Sætning 13 at ligningen

$$w^2 = b^2 - 4ac \tag{2.9}$$

altid kan løses det vil sige andengradsligninger over de komplekse tal har altid løsninger! Det gør ikke nogen forskel hvilken af de to modsat rettede løsninger w vi vælger i (2.9) som $\sqrt{b^2 - 4ac}$. Den klassiske formel gælder stadig på grund af $\pm$.

<u>**2.14 Opgave**</u>

Vis at $\varepsilon_3 = e^{i\frac{2\pi}{3}}$ er en løsning til andengradsligningen

$$z^2 + z + 1 = 0.$$

Benyt dette til at finde et udtryk for sinus til 120 grader ($2\pi/3$) ved hjælp af løsningsformlen for andengradsligningen. ♠

2.4.2 Algebraens fundamentalsætning

Vi så ovenfor at andengradsligninger altid har løsninger i de komplekse tal. Det helt enestående er at dette resultat også gælder for n-te grads ligninger det vil sige ligninger af formen

$$a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0 = 0,$$

hvor $a_n, \dots, a_0 \in \mathbb{C}$ og $a_n \neq 0$. Vi har blot vist det for $n = 2$, men det gælder for alle $n = 1, 2, 3, \dots$!

Denne perle kaldes for algebraens fundamentalsætning. Det er spændende at læse om historien bag denne sætning.

Der findes ikke et algebraisk bevis for sætningen a la det vi lavede for $n = 2$.

2.5 Om komplekse tal og periodiske fænomener

Cosinus og sinus er rasende interessante funktioner. De er matematikkens fremmeste våben i beskrivelsen af periodiske fænomener som for eksempel planetbaner og bølgebevægelser. De bliver endnu mere anvendelige, når man betragter dem ved hjælp af den komplekse eksponentialfunktion e^{ix}.

En periodisk funktion $f(t)$ er en funktion, som gentager sig selv efter et bestemt tidsrum T det vil sige $f(t) = f(t+T)$. For eksempel er både sinus og cosinus periodiske funktioner med periode $T = 2\pi$.

Uden at afsløre den fulde sandhed

Spoiler

Den fulde sandhed vil blive helt og totalt afsløret i et kursus i Fourieranalyse.

♠ kan jeg her skrive at man normalt kigger på cosinus og sinus funktioner på

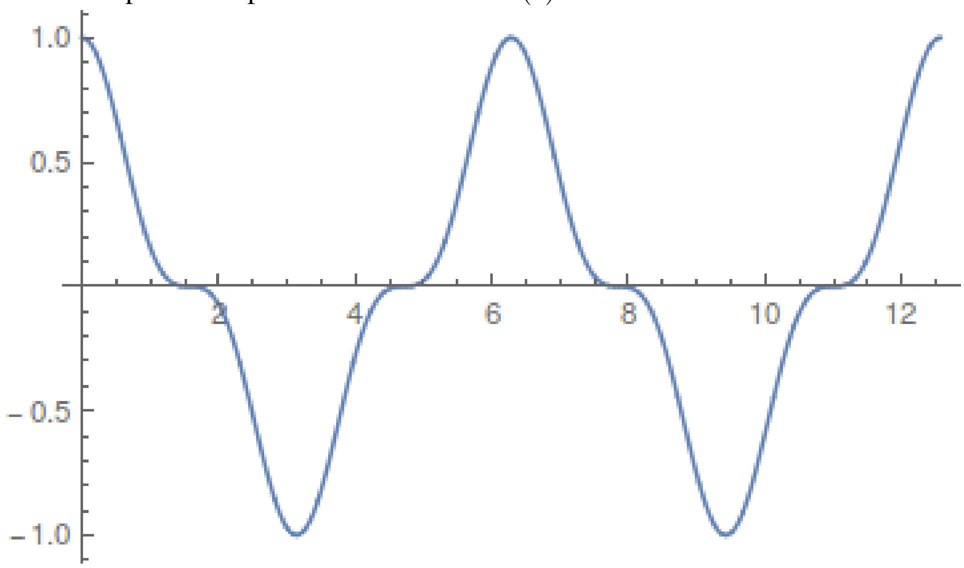
formen

$$A\cos(Nx) \qquad \text{og} \qquad A\sin(Nx), \tag{2.10}$$

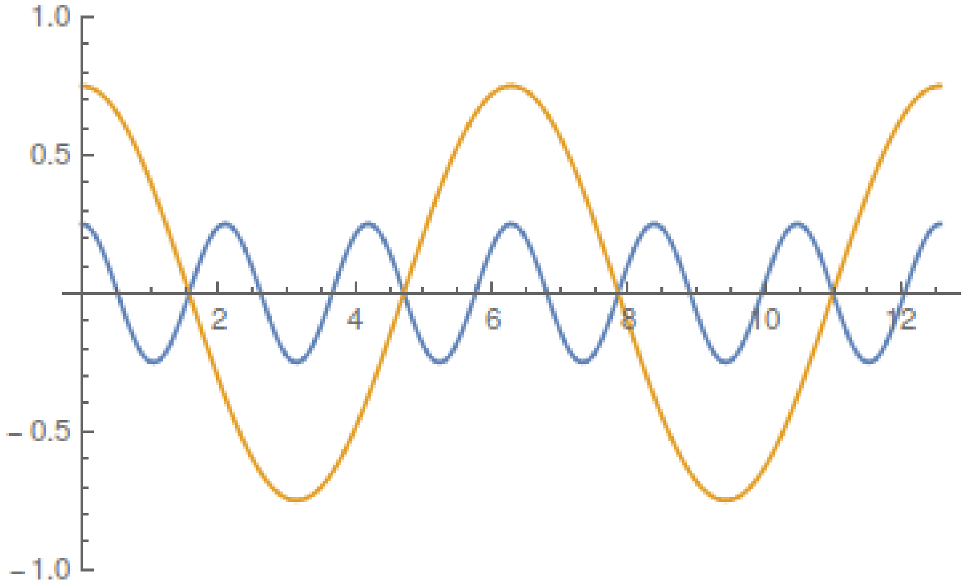
hvor A er et tal, som angiver højden (amplituden) af bølgerne og N er et tal, som beskriver antal bølger per tidsenhed (frekvensen). Cosinus og sinusfunktionerne i (2.10) samles under et i funktionerne Ae^{iNx}.

Disse funktioner er byggeklodser for naturligt forekommende periodiske fænomener.

For eksempel er den periodiske funktion $\cos(x)^3$:



sum af de to periodiske funktioner $\frac{1}{4}\cos(3x)$ og $\frac{3}{4}\cos(x)$:



Dette kan aflæses af formlen i (2.6), som vi netop fik ved hjælp af

$$(e^{ix})^3 = e^{i(3x)}.$$

For at få et indtryk af de komplekse tals nytte i signalbehandling opfordres du til at kigge nærmere på opgaverne 2.6.6 og 2.6.7.

2.6 Opgaver

2.6.1

2.15 Quizopgave

Hvad gælder om $z = (1-i)(2+i)$?

z svarer til at gange $2+i$ med $\sqrt{2}$ og derefter gange med

$$e^{i\pi/4}$$

.

$$z = 2 - i$$

$$z = 3 - i$$

z svarer til at gange $2+i$ med $\sqrt{2}$ og derefter med

$$e^{i7\pi/4}$$

. ♠

2.6.2

Gør rede for at $e^{ix}e^{-ix} = 1$, hvis $x \in \mathbb{R}$. Hvad er den polære form for $z^{-1} = 1/z$, hvis z har polær form

$$z = re^{i\theta}?$$

Hvad med den polære form for det konjugerede komplekse tal $\bar{z}$?

2.6.3

Quizopgave

Hvad gælder om $z = (2+i)/(1-i)$?

z svarer til at dividere $2+i$ med $\sqrt{2}$ og derefter gange med

$$e^{i\pi/4}$$

.

$$2z = 1+3i$$

$$z = -1+3i$$

z svarer til at gange $2+i$ med $\sqrt{2}$ og derefter med

$$e^{-i\pi/4}$$

. ♠

2.6.4

Løs andengradsligningen

$$z^2 - (3+2i)z + (1+3i) = 0.$$

2.6.5

Opskriv samtlige komplekse tal z, som løser ligningen

$$z^6 = 64.$$

Hvilke af disse løsninger er reelle tal?

2.6.6

Find reelle tal A og f så at

$$\cos(t) + \sin(t) = A\cos(t+f)$$

for alle t (A kaldes amplituden og f faseforskydningen af "signalet"på venstresiden).

Vink

Opfat venstresiden som realdelen af

$$e^{it} - ie^{it}$$

og højresiden som realdelen af

$$Ae^{i(t+f)}.$$

og regn med komplekse tal! ♠

2.6.7

Generaliser den foregående opgave til at finde C og f ud fra ω, A og B så

$$A\cos(\omega t) + B\sin(\omega t) = C\cos(\omega t + f).$$

2.6.8

Antag at $z = x + iy \in \mathbb{C}$. Gør rede for at $z = \bar{z}$ hvis og kun hvis $z \in \mathbb{R}$.

Kapitel 3

Lineære ligninger

VIDEO: https://youtu.be/JFTwV6KzgQI

Lineære ligninger optræder i et utal af anvendelser af matematikken og udgør det idémæssigt fundament for lineær algebra. Inden vi berører de mere abstrakte dele af den lineære algebra ser vi nærmere på lineære ligninger og deres anvendelser.

Lineære ligninger er ligninger, hvor de ubekendte optræder i første potens. For eksempel er $2x-3=1$ en lineær ligning i den ubekendte x, mens $x^2+x+1=0$ ikke er det, da den ubekendte x optræder i anden potens. Lineære ligninger dækker også over flere ligninger med flere ubekendte som f.eks. følgende tre ligninger med de tre ubekendte x, y og z.

$$\begin{aligned} x+y+z&=3\\ x-y+z&=1\\ x+y-z&=1 \end{aligned} \tag{3.1}$$

3.1 Quiz

Har følgende system af ligninger en og kun en løsning?

$$\begin{aligned} x+y&=1\\ x-y&=0 \end{aligned}$$

Ja. $x=1/2$ og $y=1/2$ er den eneste løsning.

Nej, systemet har uendeligt mange løsninger.

Nej, der er flere mulige løsninger, men der er kun endeligt mange.

Nej, der er ikke nogen løsning. ♠

3.2 Quiz

Har følgende system af ligninger en og kun en løsning?

$$\begin{aligned} x+y &= 1 \\ 2x+2y &= 2 \end{aligned}$$

Ja. $x = 1/2$ og $y = 1/2$ er den eneste løsning.

Nej, systemet har uendeligt mange løsninger.

Nej, der er flere mulige løsninger, men der er kun endeligt mange.

Nej, der er ikke nogen løsning. ♠

3.3 Quiz

Har følgende system af ligninger en og kun en løsning?

$$\begin{aligned} x+y &= 1 \\ 2x+2y &= 1 \end{aligned}$$

Ja. $x = 1/2$ og $y = 1/2$ er den eneste løsning.

Nej, systemet har uendeligt mange løsninger.

Nej, der er flere mulige løsninger, men der er kun endeligt mange.

Nej, der er ikke nogen løsning. ♠

3.4 Opgave

Gæt en løsning til ligningerne i (3.1) det vil sige gæt på tre tal x, y, z, som tilfredsstiller alle tre ligninger. Findes der mere end en løsning? Opskriv et system af tre ligninger med tre ubekendte, som ikke har en løsning. ♠

3.1 En ligning med en ubekendt

Der er nogle ganske enkle regler for løsning af lineære ligninger. Lad os, som eksempel, kigge nærmere på ligningen $2x - 3 = 1$. Processen for at isolere x er helt mekanisk:

$$\begin{gathered}
2x - 3 = 1 \\
\Updownarrow \\
2x - 3 + 3 = 1 + 3 \\
\Updownarrow \\
2x = 4 \\
\Updownarrow \\
\left(\frac{1}{2}\right) 2x = \left(\frac{1}{2}\right) 4 \\
\Updownarrow \\
x = 2
\end{gathered}$$

De overordnede regler vi har brugt er

$$\begin{array}{ccc}
a = b & \iff & a + c = b + c \\
a = b & \iff & ta = tb,
\end{array}$$

hvor a, b, c er tal og t et tal $\neq 0$. Disse regler gør at vi altid kan isolere den ubekendte på den ene side af lighedstegnet.

3.5 Opgave

Hvorfor bliver vi nødt til at kræve at $t \neq 0$ ovenfor? ♠

3.6 Quiz

Fysiologisk saltvand består af 0.9% salt. Du har 2 liter vand med en saltkoncentration på 9%. Hvor mange liter destilleret vand (0 procent salt) skal du tilsætte for at få fysiologisk saltvand?

4.5 liter

10 liter

18 liter ♠

3.2 Flere ligninger og flere ubekendte

Ligningen $2x - 3 = 1$ indeholder kun en ubekendt og har kun løsningen $x = 2$. *Hvis en lineær ligning indeholder mere end en ubekendt har den uendeligt mange løsninger*. Tag som eksempel ligningen $2x - 3y = 1$. Ved samme omskrivninger som ovenfor gælder

$$
\begin{gathered}
2x - 3y = 1 \\
\Updownarrow \\
x = \frac{1}{2} + \frac{3}{2}y
\end{gathered}
$$

Her kan vi altså vælge y frit på uendeligt mange måder, men når først y er valgt er x lagt fast.

Pedantens taletid

Påstanden ovenfor har faktisk en enkelt undtagelse – hvis alle ubekendte indgår med koefficient 0, som for eksempel

$$0 \cdot x + 0 \cdot y = 1$$

så er der måske slet ikke nogen løsning. ♠

3.2.1 Flere ligninger

Det giver også mening at betragte flere ligninger med flere ubekendte som f.eks.

$$
\begin{aligned}
x + y &= 3 \\
2x - 3y &= 1
\end{aligned}
$$

To tal x og y er en løsning hvis begge ligningerne er opfyldt. Fra eksemplet ovenfor ved vi at den anden ligning medfører at

$$x = \frac{1}{2} + \frac{3}{2}y. \tag{3.2}$$

Dette kan indsættes for x i den første ligning og vi får

$$3 = x + y = \frac{1}{2} + \frac{3}{2}y + y = \frac{1}{2} + \frac{5}{2}y.$$

Dette er en almindelig førstegradsligning kun i variablen y. Løsningen er $y = 1$, som så indsættes i ligningen (3.2). Her ses så at $x = 2$. Det vil sige de to ligninger har løsningen $x = 2$ og $y = 1$.

3.7 Quiz

Kona kaffe fra Hawaii er en udsøgt delikatesse til 200 kr for 400 gram. En standard pose Arabica bønner kan fås til 60 kroner for 500 gram. En forhandler vil gerne lave en blandingskaffe af de to bønner til en pris på 75 kroner for 400 gram. Hvilken af nedenstående procentsatser vil Konaindholdet i blandingskaffen ligge tættest på?

18%

5%

30%

12% ♠

3.3 Gauss elimination

Ved løsning af flere lineære ligninger, er det naturligt at fastholde en af ligningerne, isolere en variabel og så indsætte i de andre ligninger. Lad os studere denne operation via et eksempel med to ligninger med tre ubekendte:

$$\begin{aligned} x+2y+z&=8\\ 2x+y+z&=7 \end{aligned}$$

I den første ligning isoleres $x = 8-2y-z$, som så indsættes i den anden ligning:

$$7 = 2x+y+z = 2(8-2y-z)+y+z = -3y-z+16 \implies -3y-z = -9.$$

I ligningssystemet giver det også god mening at gange første ligning med 2 og trække fra anden ligning. Denne operation giver ligningen

$$-3y-z = -9.$$

At de to operationer giver samme ligning er ikke noget tilfælde. Det er indholdet af følgende resultat.

(3.8) SÆTNING.

Lad

$$\begin{aligned} a_1x_1+a_2x_2+\cdots+a_nx_n &= c_1\\ b_1x_1+b_2x_2+\cdots+b_nx_n &= c_2 \end{aligned}$$

være to lineære ligninger i de ubekendte $x_1,\dots,x_n$ *med* $a_1 \neq 0$. *Ligningen som fremkommer ved først at isolere* x_1 *i den første ligning og derefter indsætte udtrykket for*

x_1 i den anden ligning svarer præcis til ligningen, som fremkommer ved at gange første ligning med $-b_1/a_1$ og addere til anden ligning.

Bevis

Bevis. Multiplikation af første ligning med $\frac{-b_1}{a_1}$ med efterfølgende addition til anden ligning giver ligningen

$$\left(b_2 - \frac{b_1 a_2}{a_1}\right)x_2 + \cdots + \left(b_n - \frac{b_1 a_n}{a_1}\right)x_n = c_2 - \frac{b_1}{a_1}c_1 \tag{3.3}$$

Isolering af x_1 i første ligning med indsættelse i anden ligning giver ligningen

$$b_1\left(\frac{c_1}{a_1} - \frac{a_2}{a_1}x_2 - \cdots - \frac{a_n}{a_1}x_n\right) + b_2x_2 + \cdots + b_nx_n = c_2 \tag{3.4}$$

Vi kan skrive om på venstresiden i (3.4). Ved at udvikle parentesen får vi

$$\frac{b_1c_1}{a_1} - \frac{b_1a_2}{a_1}x_2 - \cdots - \frac{b_1a_n}{a_1}x_n + b_2x_2 + \cdots + b_nx_n = c_2$$

Nu flytter vi den første term om på højresiden, og samler termer med x_i. Resultatet er at vi får (3.3) tilbage. På tilsvarende måde kan vi skrive om (3.3) til (3.4). De to ligninger er altså helt ækvivalente. $\square$

♠

Operationen med at gange en ligning med et tal og addere til en anden ligning er umiddelbart nemmere at håndtere end substitutionsmetoden og vi har ovenfor vist at de er ens. Nedenfor er et gennemregnet eksempel.

3.9 Eksempel

Vi ønsker at løse ligningssystemet

$$\begin{array}{ccccccc} 2x & + & y & + & z & = 7 \\ x & + & 2y & + & z & = 8. \\ x & + & y & + & 2z & = 9 \end{array} \tag{3.5}$$

Første trin nedenfor består i at trække den tredje ligning fra den anden:

$$\begin{array}{ccccccc} 2x & + & y & + & z & = 7 \\ x & + & 2y & + & z & = 8 \\ x & + & y & + & 2z & = 9 \end{array} \iff \begin{array}{ccccccc} 2x & + & y & + & z & = 7 \\ & & y & - & z & = -1 \\ x & + & y & + & 2z & = 9 \end{array}$$

Derefter trækkes 2 gange tredje ligning fra den første:

$$\begin{array}{rcrcrl} 2x & + & y & + & z & = 7 \\ & & y & - & z & = -1 \\ x & + & y & + & 2z & = 9 \end{array} \iff \begin{array}{rcrcrl} & - & y & - & 3z & = -11 \\ & & y & - & z & = -1 \\ x & + & y & + & 2z & = 9 \end{array}$$

Til sidste lægges anden ligning til første ligning:

$$\begin{array}{rcrcrl} & - & y & - & 3z & = -11 \\ & & y & - & z & = -1 \\ x & + & y & + & 2z & = 9 \end{array} \iff \begin{array}{rcrcrl} & & & - & 4z & = -12 \\ & & y & - & z & = -1 \\ x & + & y & + & 2z & = 9 \end{array}$$

Vi har nu reduceret det oprindelige ligningssystem (3.5) til ligningssystemet

$$\begin{array}{rcrcrl} & & & - & 4z & = -12 \\ & & y & - & z & = -1\,, \\ x & + & y & + & 2z & = 9 \end{array}$$

hvor vi ud fra første ligning hurtigt ser at $z = 3$. Men så kan $z = 3$ sættes ind i anden ligning, som så bliver $y - 3 = -1$ med løsning $y = 2$. Til sidst sættes $y = 2$ og $z = 3$ ind i den tredje ligning og man får ligningen $x + 8 = 9$ eller $x = 1$. ♠

Eliminations- eller substitutionsmetoden til løsning af lineære ligningerne er en gammel kending. Sir Isaac Newton beskrev i 1720 metoden som følger.

> *And you are to know, that by each Æquation one unknown Quantity may be taken away, and consequently, when there are as many Æquations and unknown Quantities, all at length may be reduc'd into one, in which there shall be only one Quantity unknown.*

Den matematiske superstjerne Carl Friedrich Gauss benyttede metoden til at bestemme banen for asteroiden Pallas. Den matematiske behandling af observationerne ledte ham til mindste kvadraters metode og et ligningssystem med seks lineære ligninger og seks ubekendte.

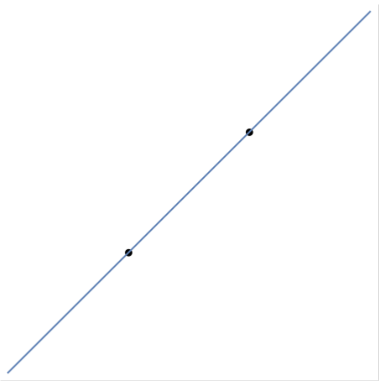
Selvom han langt fra var den første til at løse lineære ligninger ved proceduren ovenfor, er metoden blevet opkaldt efter ham. I nutiden kendes den ved navnet *Gauss elimination*.

3.4 Anvendelser

Vi har allerede ovenfor set et par quizeksempler på anvendelser af lineære ligninger. Her giver vi nogle flere.

3.4.1 Linjer, parabler og polynomier af højere grad

En linje i planen er karakteriseret ved dens ligning $y = ax + b$, hvor a er hældningskoefficienten og b skæringen med y-aksen. Gennem to punkter (x_1, y_1) og (x_2, y_2) med $x_1 \neq x_2$ går præcis en linje:



Linjen kan findes ved at løse to ligninger med to ubekendte:

$$x_1 a + b = y_1$$
$$x_2 a + b = y_2$$

Her er de ubekendte a og b. Lige i dette tilfælde kan vi benytte Gauss elimination og trække sidste ligning fra første og få $(x_1 - x_2)a = y_1 - y_2$ det vil sige

$$a = \frac{y_1 - y_2}{x_1 - x_2}.$$

Ved indsættelse af a i første ligning fås

$$b = \frac{x_1 y_2 - x_2 y_1}{x_1 - x_2}.$$

Vi kan helt eksplicit konstruere linjen gennem de to punkter som

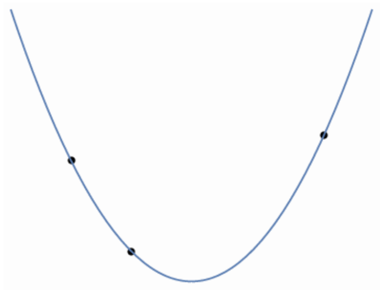
$$y = f(x) = y_1 \frac{x - x_2}{x_1 - x_2} + y_2 \frac{x - x_1}{x_2 - x_1}. \tag{3.6}$$

Funktionen $f(x)$ i (3.6) er et polynomium af grad en med $f(x_1) = y_1$ og $f(x_2) = y_2$.

Næsten analogt hermed går der en entydig parabel

$$y = ax^2 + bx + c$$

gennem tre punkter $(x_1, y_1), (x_2, y_2)$ og (x_3, y_3) med forskellige x-værdier:



Her giver punkterne følgende tre ligninger

$$
\begin{aligned}
x_1^2 a + x_1 b + c &= y_1 \\
x_2^2 a + x_2 b + c &= y_2 \\
x_3^2 a + x_3 b + c &= y_3
\end{aligned}
\tag{3.7}
$$

i de ubekendte a, b og c, men det er ikke helt oplagt at ligningerne har en løsning.

Vi kan helt eksplicit konstruere parablen gennem de tre punkter som

$$
y = f(x) = y_1 \frac{(x-x_2)(x-x_3)}{(x_1-x_2)(x_1-x_3)} + y_2 \frac{(x-x_1)(x-x_3)}{(x_2-x_1)(x_2-x_3)} + y_3 \frac{(x-x_1)(x-x_2)}{(x_3-x_1)(x_3-x_2)}
\tag{3.8}
$$

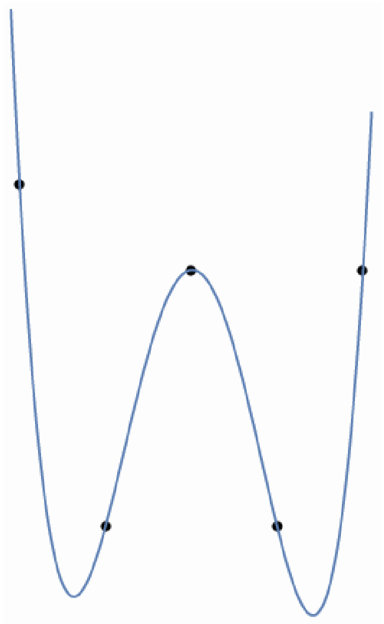
Læg igen mærke til dette fantastiske trick kopieret fra linjen ovenfor: funktionen $f(x)$ i (3.8) er et polynomium af grad to med $f(x_1) = y_1, f(x_2) = y_2$ og $f(x_3) = y_3$. Samtidig giver dette et bevis for at ligningerne i (3.7) faktisk kan løses!

VIDEO: https://youtu.be/$_kcg3rUegD4$

Den ultimative generalisering er at til $n+1$ punkter $(x_1, y_1), \cdots, (x_{n+1}, y_{n+1})$ med forskellige x-værdier går grafen for præcis en funktion af formen (et polynomium af grad højst n)

$$
y = f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0
$$

gennem punkterne. Tricket ovenfor i (3.8), som forøvrigt kaldes Lagrange interpolation, virker også i det generelle tilfælde. Nedenfor er et eksempel på 5 punkter, som definerer et fjerdegradspolynomium:



3.4.2 Kemisk ligevægt

I kemiske reaktioner er et grundliggende princip massebevarelse. I nedenstående proces reagerer methan med oxygen og der opstår kuldioxid og vand som følge, men der er ubalance mellem masserne på hver side af pilen.

$$CH_4 + O_2 \quad \longrightarrow \quad CO_2 + H_2O$$

På venstresiden er der f.eks. fire hydrogenatomer, mens der på højresiden kun er to. Vi kan afstemme reaktionen ved at indføre fire variable x, y, z, w, som hver for sig angiver mængden af de involverede molekyler:

$$xCH_4 + yO_2 \quad \longrightarrow \quad zCO_2 + wH_2O \tag{3.9}$$

Ved at benytte at antallet af de enkelte atomer skal være bevaret får vi følgende lineære ligninger

$$\begin{array}{rcrcrcrcl} x & & & - & z & & & = & 0 \\ 4x & & & & & - & 2w & = & 0 \\ & & 2y & - & 2z & - & w & = & 0 \end{array}$$

Igen er Gauss elimination nyttig. Vi ganger sidste ligning med 2 og trækker fra den næstsidste ligning for at eliminere w:

$$\begin{array}{rcrcrcrcl} x & & & - & z & & & = & 0 \\ 4x & - & 4y & + & 4z & & & = & 0 \\ & & 2y & - & 2z & - & w & = & 0 \end{array}$$

Dernæst ganger vi første ligning med 4 og lægger til anden ligning:

$$\begin{array}{rcrcrcrcl} x & & & - & z & & & = & 0 \\ 8x & - & 4y & + & & & & = & 0 \\ & & 2y & - & 2z & - & w & = & 0 \end{array}$$

Nu ses at løsningerne til ligningen kun afhænger af den frie variabel z:

$$\begin{array}{rcl} x & = & z \\ y & = & 2z \\ w & = & 2z \end{array}$$

Det vil sige der er uendeligt mange måder at balancere reaktionsskemaet (3.9) på afhængig af valget af z. For $z = 1$ balancerer reaktionsskemaet som

$$CH_4 + 2O_2 \quad \longrightarrow \quad CO_2 + 2H_2O.$$

3.5 En meget vigtig matematisk sætning

For at komme i gang med den lineære algebra, som egentlig blot er en fin ramme for studiet af lineære ligninger, er der specielt et vigtigt resultat som skal vises.

Lineære ligningssystemer med lutter nuller på højresiden kaldes *homogene*. Et eksempel kunne være

$$\begin{array}{rcrcrcl} x & + & y & + & 3z & = & 0 \\ x & - & y & + & 2z & = & 0 \end{array}$$

Sådanne ligningssystemer har altid løsningen, hvor alle de ubekendte er 0. Denne løsning kan være den eneste som i tilfældet

$$\begin{array}{rcrl} x & + & y & = 0 \\ x & - & y & = 0^{\cdot} \end{array}$$

VIDEO: https://youtu.be/pQ4kqc6ufZc

Det vigtige resultat er, at der altid er en løsning til et homogent ligningssystem forskellig fra nulløsningen, hvis *antallet af ubekendte er større end antallet af ligninger*, som f.eks. tilfældet $x + y = 0$ med kun en ligning og to ubekendte.

(3.10) SÆTNING.

Et homogent lineært ligningssystem har altid en løsning forskellig fra nulløsningen, hvis antallet af ubekendte er større end antallet af ligninger.

Bevis

Bevis. Lad os antage vi kun har en ligning med $n > 1$ ubekendte skrevet som

$$a_1x_1 + \cdots + a_nx_n = 0. \tag{3.10}$$

Hvis alle $a_i = 0$ for $i = 1,\ldots,n$ kan vi helt frit vælge $x_1,\ldots,x_n$ som så vil være en løsning til (3.10). Specielt har (3.10) en løsning forskellig for nulløsningen.

Vi kan derfor uden tab af generalitet antage at $a_1 \neq 0$ og isolere x_1 ud fra følgende formel

$$x_1 = \frac{1}{a_1}(-a_2x_2 - \cdots - a_nx_n). \tag{3.11}$$

Vi vælger nu de variable $x_2,\cdots,x_n$ så mindst en af dem er $\neq 0$ og fastlægger derefter x_1 via (3.11). Herefter har vi en løsning $x_1,\ldots,x_n$ til (3.10) forskellig fra nulløsningen.

Hvad gør vi med et ligningssystem med mere end en ligning? Lad os antage, at vi har m ligninger med n ubekendte, hvor $m > 1$ og $n > m$:

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0\\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0\\
&\vdots\\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= 0
\end{aligned}$$

Vi viser denne sætning ved brug af matematisk induktion

<u>Hvad er induktion?</u>

Induktion er en genial måde at organisere sine tanker på. Ideen er den følgende. Som det plejer at ske for os matematiker, er vi havnet i den situation at vi skal bevise noget. Det vi skal vise afhænger af et tal —i vores tilfælde af antallet af ligninger. I stedet for at prøve på at bevise sætningen for alle ligningssystemer i et slag, så vil vi vise vi den først for systemer med én ligning, derefter for systemer med to ligninger ... og så videre. Vi havde et problem, og nu har vi uendeligt mange problemer. Dette kalder vi matematiker at gøre fremskridt.

Vi begynder helt naivt med at løse problemet for systemer der kun består af en eneste ligning. Dette trin kaldes for ”induktionsstarten”. Det er måske ikke så svært, og da vi har gjort det har vi jo heldigvis kun uendeligt mange andre problemer tilbage at løse. Men nu kommer det smarte. I stedet for at vise det vi gerne vil have for systemer med 2 ligninger, og derefter gå i gang med systemer med 3 ligninger, så viser vi at hvis vi kan løse problemet for systemer med $m-1$ ligninger, så kan vi også løse det for systemer med m ligninger. Dette kaldes for ”induktionsskridtet”. Og så er beviset allerede helt færdigt! Fordi nu ruller logikken, og ingen magt i denne verden kan standse den: Det er OK for 1 ligning, altså også for 2 ligninger. Men hvis det er OK for 2 ligninger, så er det også OK for 3 ligninger. Og så videre. ♠

Lad os stiltiende antage at sætningen er sand for homogene ligningssystemer med færre end m ligninger (vi har ovenfor bevist sætningen for et homogent ligningssystem med kun en ligning det vil sige for $m = 1$).

Hvis alle $a_{i1} = 0$ for $i = 1, \ldots, m$, så er $x_1 = 1$ og $x_2 = \cdots = x_n = 0$ en løsning forskellig fra nulløsningen. Antag derfor at $a_{11} \neq 0$. Så kan vi ved Gauss elimination ud fra a_{11} i første ligning eliminere x_1 i ligningerne nedenunder. Dette giver ligningssystemet

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\ a'_{22}x_2 + \cdots + a'_{2n}x_n &= 0 \\ &\vdots \\ a'_{m2}x_2 + \cdots + a'_{mn}x_n &= 0, \end{aligned}$$

hvor første ligning er uændret, men hvor Gauss elimination har ændret de $m-1$ ligninger under den første. Vi kigger nu nærmere på det mindre ligningssystem

$$\begin{aligned} a'_{22}x_2 + \cdots + a'_{2n}x_n &= 0 \\ &\vdots \\ a'_{m2}x_2 + \cdots + a'_{mn}x_n &= 0. \end{aligned}$$

Dette er et homogent ligningssystem med $m-1$ ligninger og $n-1$ ubekendte. Da $m-1 < n-1$ ved vi per vores antagelse at der findes en løsning $x_2, \ldots, x_n$ forskellig fra nulløsningen til det mindre ligningssystem ovenfor. På samme måde som for $m = 1$ giver denne løsning en løsning $x_1, x_2, \ldots, x_n$ forskellig fra nulløsningen til det større ligningssystem. $\square$

♠

Bevismetoden i beviset ovenfor kendes under betegnelsen *matematisk induktion*. Vi vil ofte få brug for denne måde at argumentere på, så prøv på at forstå idéen.

3.11 Opgave

Er følgende argument korrekt?

(3.12) SÆTNING.

Lad F_m være summen af de første m naturlige tal, altså

$$F_m = 1 + 2 + 3 + \cdots + m$$

Da er F_m givet ved formlen

$$F_m = \frac{m(m+1)}{2}$$

Bevis. Induktionsstart: $F_1 = 1 = \frac{1 \cdot 2}{2}$.

Induktionsskridt: Antag at sætningen er rigtig for $m-1$. Vi regner:

$$\begin{aligned} F_m &= 1+2+\cdots+(m-1)+m \\ &= (1+2+\cdots+(m-1))+m \\ &= F_{m-1}+m \end{aligned}$$

Ifølge induktionsantagelsen kan vi nu fortsætte sådan:

$$\begin{aligned} &= \frac{(m-1)m}{2}+m \\ &= \frac{(m-1)m+2m}{2} \\ &= \frac{m(m+1)}{2} \end{aligned}$$

Dermed er sætningen bevist ved induktion. □

♠

3.13 Opgave

Er følgende argument korrekt?

(3.14) SÆTNING.

Hvis en ud af en endelig mængde af matematiker hedder Marcel, så hedder alle matematiker i denne mængde Marcel.

Bevis. Induktionsstart: Hvis der kun er en matematiker, så er sætningen åbenbart rigtig.

Induktionsskridt:Antag at den er rigtig for mængder der består af $m-1$ matematiker. Betragt nu en mængde X af matematiker der opfylder at en af dem, lad os kalde ham M, hedder Marcel. Vi skal altså vise at alle matematiker i mængden X hedder Marcel. Vi indfører nu to delmængder af X. Hvis X for eksempel består af A,B,C,D,M, så at $m=5$, lader vi X_1 være mængden bestående af A,B,C,M, og X_2 mængden bestående af A,B,D,M. Hver af disse delmængder indeholder en matematiker der hedder Marcel. Ifølge induktionsantagelsen brugt på X_1, som har $4=m-1$ elementer, hedder altså A,B,C alle Marcel. Det samme argument brugt på X_2 viser at det gør D også. Dermed er sætningen bevist ved induktion. □

♠

3.6 Opgaver

3.6.1

3.15 Quizopgave

Hvor mange løsninger har ligningssystemet

$$\begin{aligned} x+y+z &= 0 \\ x-y+z &= 0 \\ x+y-z &= 0? \end{aligned}$$

Ingen.

Præcis en.

Uendeligt mange. ♠

3.6.2

3.16 Quizopgave

Hvor mange løsninger har ligningssystemet

$$\begin{aligned} x+y+z &= 0 \\ x-y+z &= 0 \\ 5x+y+5z &= 0? \end{aligned}$$

Præcis en.

Præcis to.

Uendeligt mange. ♠

3.6.3

Find samtlige løsninger til ligningssystemet

$$\begin{array}{rcrcrcl} x & + & 3y & + & z & = & 2 \\ -2x & - & 5y & + & 3z & = & 4. \end{array}$$

3.6.4

Løs ligningssystemet

$$\begin{array}{rcrcrcl} x & + & y & + & z & = & 2 \\ x & + & y & - & iz & = & 3-i \\ ix & + & y & + & z & = & 1+i \end{array}$$

ved hjælp af Gauss elimination, som forklaret i dette kapitel.

3.6.5

En mand kaster en stålkugle lodret ned fra toppen af en skyskraber på en planet i vores solsystem. Fra nabobygningen måles kuglens højde efter givne tidsrum: Efter 4 sekunder har kuglen en højde på 426 meter, efter 6 sekunder har kuglen en højde på 369 meter og efter 9 sekunder har kuglen en højde på 256 meter.

Hvor stor en hastighed blev stålkuglen kastet med til at begynde med? Hvad er tyngdeaccelerationen på planeten i forhold til målingerne? Hvilken planet befinder manden sig højst sandsynligt på?

3.17 Vink

NB: En passende fysisk model er at højden til tiden t er givet ud fra formlen

$$h_0 - v_0 t - \frac{1}{2} a t^2,$$

hvor h_0 er højden af bygningen, v_0 begyndelseshastigheden og a tyngdeaccelerationen.

Og det er ikke på Pallas. ♠

3.6.6

Find $a_0, a_1, a_2, a_3 \in \mathbb{R}$ så at

$$\begin{aligned} f(-2) &= -1 \\ f(-1) &= 1 \\ f(1) &= 1 \\ f(2) &= 1, \end{aligned}$$

hvor

$$f(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3.$$

3.6.7

Gør rede for at der til $n+1$ punkter $(x_0,y_0),(x_1,y_1),\dots,(x_n,y_n)$ med forskellige x-værdier findes entydige tal $a_0,a_1,\dots,a_n$ så

$$f(x_0)=y_0, \quad f(x_1)=y_1, \quad \dots \quad ,f(x_n)=y_n,$$

hvor

$$y=f(x)=a_0+a_1x+\cdots+a_nx^n.$$

Hvorfor medfører det at ligningssystemet

$$\begin{array}{ccccccccc} a_0 & + & x_0a_1 & + & \cdots & + & x_0^na_n & = & y_0 \\ a_0 & + & x_1a_1 & + & \cdots & + & x_1^na_n & = & y_1 \\ & & & & & & & \vdots & \\ a_0 & + & x_na_1 & + & \cdots & + & x_n^na_n & = & y_n \end{array}$$

har en løsning i de ubekendte $a_0,a_1,\dots,a_n$? Findes der kun en løsning her?

3.18 Vink

Eksistensen af $a_0,\dots,a_n$ i $f(x)$ fremkommer ved at generalisere parabeltilfældet (3.8) ovenfor (prøv først med $n=3$). Et polynomium $f(x)=a_0+a_1x+\cdots+a_nx^n$, siges at have grad n, hvis $a_n\neq 0$. Et generelt resultat om polynomier siger at et polynomium $f(x)$ af grad $n\geq 0$ højst kan have n rødder (en rod er et nulpunkt for $f(x)$). Dette resultat kan bruges til at bevise entydigheden af $a_0,\dots,a_n$, for eksempel ved at antage eksistensen af et andet polynomium $g(x)$ af grad n, som opfylder $g(x_0)=y_0,\dots,g(x_n)=y_n$ og så betragte polynomiet $f(x)-g(x)$ (som har hvor mange rødder?). ♠

Kapitel 4

Matricer

VIDEO: https://youtu.be/KxIRozfLhHo

Når man har regnet med lineære ligninger et stykke tid opstår behovet for at forenkle notationen. For eksempel kan ligningerne

$$\begin{array}{ccccccc} & & 2y & + & 4z & = & -2 \\ 3x & + & 2y & + & 7z & = & 4 \end{array} \tag{4.1}$$

repræsenteres ved talskemaet

$$\begin{pmatrix} 0 & 2 & 4 & -2 \\ 3 & 2 & 7 & 4 \end{pmatrix} \tag{4.2}$$

og mange af de operationer vi foretager for at løse ligningerne kan lige så vel udføres på det tilsvarende talskema.

4.1 Matricer

4.1.1 Definitioner

Et rektangulært talskema kaldes en *matrix*. En matrix med m *rækker* og n *søjler* kaldes en $m \times n$ (læs: m gange n) matrix. Notation for en $m \times n$ matrix A er

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{pmatrix}, \tag{4.3}$$

hvor $A_{ij} = a_{ij}$ betegner tallet i i-te række og j-te søjle. Hvis vi kalder matricen i (4.2) for A, består den af 2 rækker og 4 søjler med $A_{14} = -2$.

1. En matrix kaldes *kvadratisk* hvis den har lige så mange rækker som søjler. For eksempel er de første to matricer nedenfor kvadratiske, mens den tredje ikke er det.

$$\begin{pmatrix}1\end{pmatrix}, \qquad \begin{pmatrix}1 & 2 & 3\\ 4 & 5 & 6\\ 7 & 8 & 9\end{pmatrix}, \qquad \begin{pmatrix}0 & 1 & 0\\ 1 & 0 & 1\end{pmatrix}.$$

2. *Diagonalen* i en matrix er defineret som indgangene i matricen med samme række- og søjlenummer. Nedenfor er angivet en 3×4 matrix, hvor diagonalelementerne er markerede

$$\begin{pmatrix}\color{red}{1} & 3 & 0 & 1\\ 3 & \color{red}{2} & 1 & 5\\ 1 & 0 & \color{red}{3} & 6\end{pmatrix}.$$

   En matrix kaldes en *diagonalmatrix*, hvis alle dens indgange udenfor diagonalen er $= 0$. Nedenfor er et eksempel på en kvadratisk diagonalmatrix

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

3. En matrix kaldes en *rækkevektor* hvis den kun har en række. For eksempel er

$$\begin{pmatrix}1 & 2 & 3\end{pmatrix}$$

   en rækkevektor med tre søjler.

4. En matrix kaldes en *søjlevektor* hvis den kun har en søjle. For eksempel er

$$\begin{pmatrix}1\\ 2\\ 3\end{pmatrix}$$

   en søjlevektor med tre rækker.

5. Rækkerne i en matrix kaldes matricens *rækkevektorer*. Den i-række i en matrix A betegnes A_i. For eksempel har matricen A i (4.2) rækkevektorerne

$$A_1 = \begin{pmatrix}0 & 2 & 4 & -2\end{pmatrix} \qquad \text{og} \qquad A_2 = \begin{pmatrix}3 & 2 & 7 & 4\end{pmatrix}.$$

6. Søjlerne i en matrix kaldes matricens *søjlevektorer*. Den j-te søjle i en matrix A betegnes A^j. For eksempel har matricen A i (4.2) søjlevektorerne

$$A^1 = \begin{pmatrix}0\\ 3\end{pmatrix}, \quad A^2 = \begin{pmatrix}2\\ 2\end{pmatrix}, \quad A^3 = \begin{pmatrix}4\\ 7\end{pmatrix} \quad \text{og} \quad A^4 = \begin{pmatrix}-2\\ 4\end{pmatrix}.$$

7. En række- eller søjlevektor refereres til som en *vektor*.

Vi vil senere give en mere abstrakt definition af vektorer som elementer i et såkaldt vektorrum.

4.2 Matrixmultiplikation

Antag vi har givet to ligningssystemer

$$\begin{array}{rcrl} u & + & 2v & = p \\ u & - & 2v & = q \end{array} \quad \text{og} \quad \begin{array}{rcrl} 2x & + & 3y & = u \\ -x & - & 2y & = v. \end{array}$$

i de variable u, v og x, y.

Vi får et nyt ligningssystem i x og y ved at sætte $u = 2x + 3y$ og $v = -x - 2y$ ind i det første ligningssystem:

$$\begin{array}{rcrcrcrcrcrcl} u & + & 2v & = & (2x+3y) & + & 2(-x-2y) & = & & - & y & = p \\ u & - & 2v & = & (2x+3y) & - & 2(-x-2y) & = & 4x & + & 7y & = q. \end{array}$$

Med matricer skriver vi

$$\begin{pmatrix} 1 & 2 \\ 1 & -2 \end{pmatrix} \begin{pmatrix} 2 & 3 \\ -1 & -2 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 4 & 7 \end{pmatrix} \tag{4.4}$$

Lad os prøve at skrive operationen i (4.4) ud generelt det vil sige antag vi har to ligningssystemer a la ovenfor:

$$\begin{array}{rcrl} a_{11}u & + & a_{12}v & = p \\ a_{21}u & + & a_{22}v & = q \end{array} \quad \text{og} \quad \begin{array}{rcrl} b_{11}x & + & b_{12}y & = u \\ b_{21}x & + & b_{22}y & = v \end{array}$$

men nu med generelle koefficienter. Ved substitution fås som før

$$\begin{array}{rcrcrcr} a_{11}u & + & a_{12}v & = & a_{11}(b_{11}x + b_{12}y) & + & a_{12}(b_{21}x + b_{22}y) \\ a_{21}u & + & a_{22}v & = & a_{21}(b_{11}x + b_{12}y) & + & a_{22}(b_{21}x + b_{22}y) \end{array}$$

som så er lig med

$$\begin{array}{rcrl} (a_{11}b_{11} + a_{12}b_{21})x & + & (a_{11}b_{12} + a_{12}b_{22})y & = p \\ (a_{21}b_{11} + a_{22}b_{21})x & + & (a_{21}b_{12} + a_{22}b_{22})y & = q \end{array}$$

Formuleret med matricer som i (4.4) skrives

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix} \tag{4.5}$$

Ligningen ovenfor er intet mindre end formlen for multiplikation af to 2×2 matricer, præcis som den blev indført historisk af Cayley omkring 1857. Ved nærmere eftersyn (og markeret med farver i (4.5) for $i = 2$ og $j = 1$) ses reglen at tallet i den i-te række og j-te søjle i produktmatricen er *række-søjle multiplikationen* mellem i-te række og j-te søjle i de to matricer.

Rækkesøjle multiplikationen mellem en rækkevektor

$$\mathbf{x} = (x_1 x_2 \ldots x_n)$$

og en søjlevektor

$$\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

med det samme antal indgange er defineret som

$$\mathbf{xy} = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n.$$

Hvis man er lidt pedantisk vil man måske i notationen skelne mellem tallet 5 og 1×1 matricen (5), men det er vi ikke.

(4.1) DEFINITION.

Lad A være en $m \times p$ matrix og B en $p \times r$ matrix. Så er produktet AB defineret som $m \times r$ matricen C givet ved

$$C_{ij} = A_i B^j$$

for $1 \leq i \leq m$ og $1 \leq j \leq r$.

Hvis A er en $m \times n$ matrix og B en $r \times s$ matrix giver matrixproduktet AB kun mening, hvis $n = r$: Antallet af søjler i A skal være lig med antallet af rækker i B.

4.2 Quiz

Lad matricerne

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}, \quad \text{og} \quad D = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

være givet. Hvilke af nedenstående matrixprodukter giver mening?

BA

AB

CD

DC

CA

AD ♠

VIDEO: https://youtu.be/nIeN$_N roisA$

Med formlen for matrix multiplikation kan ligningssystemet (4.1) nu skrives som

$$\begin{pmatrix} 0 & 2 & 4 \\ 3 & 2 & 7 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -2 \\ 4 \end{pmatrix}$$

Her ganger vi en 2×3 matrix sammen med en 3×1 matrix. Rækkesøjlemultiplikationen giver 2×1 matricen

$$\begin{pmatrix} 2y + 4z \\ 3x + 2y + 7z \end{pmatrix}.$$

Denne matrix skal netop være lig med 2×1 matricen på højresiden ovenfor for at ligningssystemet (4.1) er opfyldt.

4.3 Quiz

Lad

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 3 & x & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 0 & 1 & 1 \end{pmatrix} \quad \text{og} \quad C = AB$$

Hvilke af nedenstående udsagn er rigtige?

$C_{12} = 9$

$C_{23} = 4$

Hvis $C_{32} = 4$, så er $x = 0$.

Hvis $C_{31} = -1$, så er $x = -1$. ♠

Matrixmultiplikation optræder mange steder. Nedenfor et meget anvendeligt eksempel indenfor sandsynlighedsregning, som i generaliseret form leder til Googles berømte page rank algoritme.

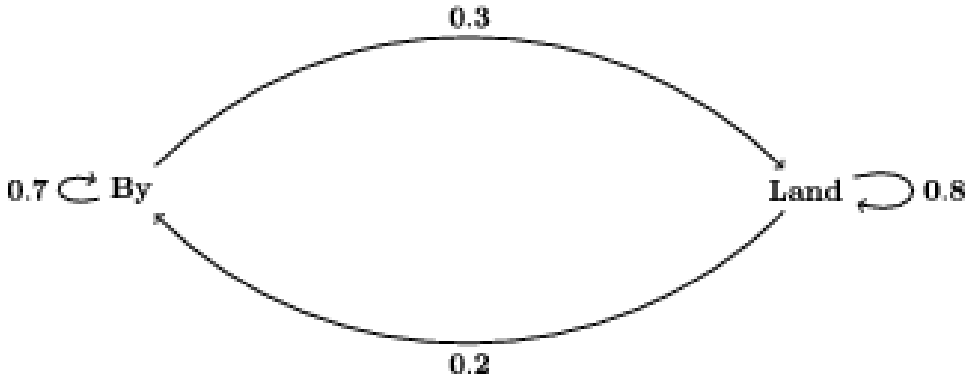
4.4 Eksempel

Matrixmultiplikation forekommer naturligt i sandsynlighedsregning. Lad os illustrere med et enkelt eksempel.

Lad os antage at rundt regnet 20% af de mennesker, der bor på landet, flytter til byen og at 30% af de mennesker, som bor i byen flytter til landet. Lad os også fastslå at disse procentsatser er opgjort per år og lige omformulere en smule:

1. Hvis man bor på landet er sandsynligheden for at man flytter til byen 0.2,
2. Hvis man bor på landet er sandsynligheden for at man bliver boende 0.8,
3. Hvis man bor i byen er sandsynligheden for at man flytter til landet 0.3,
4. Hvis man bor i byen er sandsynligheden for at man bliver boende 0.7,

når man ser på et år som tidsramme. Dette kan illustreres med nedenstående diagram



Dette giver anledning til lidt købmandsregning. Lad os sige at der i starten til tiden $t = 0$ år bor x_0 mennesker i byen og y_0 mennesker på landet. Hvor mange mennesker x_1 bor der så i byen og hvor mange mennesker y_1 bor der på landet til tiden $t = 1$ år?

Med ord bliver byen affolket med 30%, men der kommer tilflyttere, som udgør 20% af befolkningen på landet. Det vil sige

$$x_1 = 0.7x_0 + 0.2y_0.$$

Tilsvarende har vi for befolkningen på landet at

$$y_1 = 0.3x_0 + 0.8y_0.$$

Dette kan skrives via matrixmultiplikation som

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}.$$

Proceduren giver også mening for $t = 2$ år. Her bliver resultatet

$$\begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{pmatrix} \left(\begin{pmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} \right) \tag{4.6}$$

$$= \left(\begin{pmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{pmatrix} \begin{pmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{pmatrix} \right) \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = P^2 \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}, \tag{4.7}$$

hvor

$$P = \begin{pmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{pmatrix}. \tag{4.8}$$

Ovenstående kan generaliseres så vi har formlen

$$\begin{pmatrix} x_n \\ y_n \end{pmatrix} = P^n \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}, \tag{4.9}$$

som giver fordelingen af by- og landbefolkning til tiden $t = n$ år. For at kunne benytte formlen (4.9) skal vi altså udføre $n - 1$ matrixmultiplikationer, hvilket kan være lidt overvældende, for eksempel hvis vi ønsker at kende befolkningstallet på landet efter 50 år. Hver matrixmultiplikation indeholder 8 almindelige talmultiplikationer og 4 almindelige taladditioner. Vi vil senere i kapitlet se hvordan *egenvektorer* og *egenværdier* for matricer kan hjælpe med denne udregning.

Inden da, lad os blot eksperimentere med at udregne de første potenser af P:

$$\begin{aligned}
P^2 &= \begin{pmatrix} 0.55 & 0.3 \\ 0.45 & 0.7 \end{pmatrix} \\
P^3 = PP^2 &= \begin{pmatrix} 0.475 & 0.35 \\ 0.525 & 0.65 \end{pmatrix} \\
P^4 = PP^3 &= \begin{pmatrix} 0.4375 & 0.375 \\ 0.5625 & 0.625 \end{pmatrix} \\
&\vdots \\
P^{15} &= \begin{pmatrix} 0.400018 & 0.399951 \\ 0.599982 & 0.600012 \end{pmatrix} \\
P^{16} &= \begin{pmatrix} 0.400009 & 0.399994 \\ 0.599991 & 0.600006 \end{pmatrix}
\end{aligned}$$

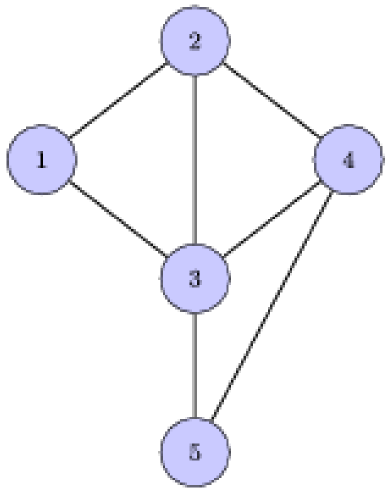
Umiddelbart ser det ud som om udregningerne stabiliseres på et stationært niveau, hvor 40% bor i byen og 60% på landet taget ud fra det samlede indbyggertal til at begynde med det vil sige til $t = 0$ år.

Matricen P er et eksempel på en stokastisk 2×2 matrix. Generelt kaldes en kvadratisk matrix en *stokastisk matrix* hvis alle dens indgange er ≥ 0 og dens søjlesummer er 1. ♠

Nedenfor et eksempel på anvendelser i netværksteori.

4.5 Eksempel

Matrixmultiplikation forekommer også i praktiske problemer, hvor netværk er involveret. Lad os antage vi har fem byer, som er forbundet med forskellige veje som nedenfor



Dette netværk har en 5×5 *incidensmatrix*, hvor by nummer i hører til i-te række og i-te søjle. Et 1-tal i matricen på plads (i, j) betyder at der er en vej fra by i til by j, mens et 0 betyder at by i og by j ikke er forbundet med en vej:

$$A = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{pmatrix}.$$

Her er

$$A^2 = \begin{pmatrix} 2 & 1 & 1 & 2 & 1 \\ 1 & 3 & 2 & 1 & 2 \\ 1 & 2 & 4 & 2 & 1 \\ 2 & 1 & 2 & 3 & 1 \\ 1 & 2 & 1 & 1 & 2 \end{pmatrix} \quad \text{og} \quad A^3 = \begin{pmatrix} 2 & 5 & 6 & 3 & 3 \\ 5 & 4 & 7 & 7 & 3 \\ 6 & 7 & 6 & 7 & 6 \\ 3 & 7 & 7 & 4 & 5 \\ 3 & 3 & 6 & 5 & 2 \end{pmatrix}.$$

Hvad er netværksfortolkningen af A^2,A^3 og generelt A^n? Det viser sig at fortolkningen af indgang (i, j) i matricen A^n netop er antallet af stier af længde n fra by i til by j. For eksempel ser vi ovenfor at der er 3 stier fra by 1 til by 5 af længde 3 svarende til $1245, 1345, 1235$. De 2 stier fra by 1 til by 1 af længde 3 er $1231, 1321$ og de 5 stier af længde 3 fra by 1 til 2 er $1342, 1242, 1323, 1212, 1232$.

Lad os antage at vi har et netværk med m byer og en tilhørende incidensmatrix A.

Det generelle bevis bygger på at en sti af længde n fra by i til by j må ende med en vej fra en naboby k til j. For hver af disse nabobyer kan vi så nøjes med at tælle antallet af stier af længde $n-1$ fra by i. Hvis vi nu antager at A_{gh}^{n-1} er antallet af stier af længde $n-1$ fra by g til by h, så siger matrixmultiplikation at

$$A_{ij}^{n} = A_{i1}^{n-1} A_{1j} + \cdots + A_{im}^{n-1} A_{mj}$$

Dette tal er antallet af stier af længde n fra by i til by j fordi $A_{kj} = 1$ netop når k er en naboby til j (og ellers 0). ♠

4.3 Matrixregning

Matrixmultiplikation er forskellig fra almindelig talmultiplikation på et helt centralt punkt: Faktorernes orden er ikke ligegyldig. Betragt matricerne

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

Så er

$$AB = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \qquad \text{og} \qquad BA = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$$

dvs $AB \neq BA$. Man siger også at matrixmultiplikation er *ikke-kommutativ*.

4.3.1 Addition af matricer

Man kan (næsten) regne med matricer som almindelige tal. Specielt giver det mening at lægge matricer med samme antal rækker og søjler sammen indgang for indgang:

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mn} \end{pmatrix} = \begin{pmatrix} a_{11}+b_{11} & \cdots & a_{1n}+b_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1}+b_{m1} & \cdots & a_{mn}+b_{mn} \end{pmatrix}.$$

Med hensyn til addition opfører matricer sig ligesom almindelige tal, det vil sige at $A + B = B + A$.

4.3.2 Skalarmultiplikation af matricer

En matrix kan på naturlig måde multipliceres med et tal λ ved at gange ind plads for plads:

$$\lambda \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} = \begin{pmatrix} \lambda a_{11} & \cdots & \lambda a_{1n} \\ \vdots & \ddots & \vdots \\ \lambda a_{m1} & \cdots & \lambda a_{mn} \end{pmatrix}.$$

4.6 Opgave

Findes et tal λ så

$$\lambda \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 4 & 6 \\ 8 & 10 & 15 \end{pmatrix}?$$

♠

4.3.3 Den distributive lov

For almindelige tal gælder at man kan gange ind i en parentes det vil sige $a(b+c) = ab+ac$. Denne regel gælder også for matricer og kaldes generelt for den distributive lov (gange bliver distribueret (fordelt) over plus).

(4.7) PROPOSITION.

Lad B og C være $m \times n$ matricer, A en $r \times m$ matrix og D en $n \times s$ matrix. Så gælder

$$A(B+C) = AB+AC \qquad \text{og} \qquad (B+C)D = BD+CD.$$

Bevis

Bevis. Man kan nøjes med at bevise den første påstand i tilfældet, hvor A er en rækkevektor og B, C søjlevektorer, fordi

$$(A(B+C))_{ij} = A_i(B+C)^j = A_i(B^j + C^j).$$

Tilsvarende kan den anden påstand bevises i tilfældet hvor B, C er rækkevektorer og D en søjlevektor, fordi

$$((B+C)D)_{ij} = (B+C)_i D^j = (B_i + C_i)D^j.$$

Begge disse tilfælde følger af den distributive lov for almindelige tal.

For eksempel, hvis $A = (a_1, \ldots, a_m)$ er en rækkevektor og

$$B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \qquad C = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{pmatrix}$$

søjlevektorer, så er $A(B+C)$ og $AB+AC$ begge 1×1 matricer, det vil sige at de kun har et eneste element. Helt præcis er

$$\begin{aligned}A(B+C) &= (\sum_i a_i(b_i+c_i))\\ &= (\sum_i (a_ib_i+a_ic_i))\\ &= (\sum_i a_ib_i)+(\sum_i a_ic_i)\\ &= AB+AC\end{aligned}$$

□

♠

4.3.4 Den mirakuløse associative lov

Giver det mening at gange tre matricer A, B og C sammen? Vi har faktisk kun defineret produktet af to matricer. Der er to naturlige måder at udregne produktet ABC på:

$$(AB)C \qquad \text{og} \qquad A(BC).$$

Vi kan begynde med at gange A sammen med B og så gange C på fra højre. Vi kan også først gange B sammen med C og så gange A på fra venstre. Det er slet ikke klart at de to måder leder frem til samme resultat! At det gælder er helt centralt når man regner med matricer. Resultatet kaldes den associative lov for matrixmultiplikation.

(4.8) SÆTNING.

Lad A være en $m\times n$ matrix, B en $n\times r$ matrix og C en $r\times s$ matrix. Så er

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

Det nedenstående bevis er ikke særlig informativt, men det er korrekt. Senere, når

vi har set sammenhængen mellem matricer og lineære afbildninger, vil vi være i stand til at give en meget bedre forklaring på hvorfor den associative lov er en selvfølge.

Bevis

Bevis. Vi skal bevise at

$$((AB)C)_{ij} = (A(BC))_{ij}$$

for $1 \leq i \leq m$ og $1 \leq j \leq s$. Venstresiden kan skrives

$$\begin{aligned}(AB)_i C^j &= (A_i B^1, \ldots, A_i B^r) C^j && (4.10)\\ &= (A_i B^1) C_{1j} + (A_i B^2) C_{2j} + \cdots + (A_i B^r) C_{rj}. && (4.11)\end{aligned}$$

Højresiden kan skrives som

$$A_i (BC)^j = A_i \begin{pmatrix} B_1 C^j \\ \vdots \\ B_n C^j \end{pmatrix} = A_{i1}(B_1 C^j) + \cdots + A_{in}(B_n C^j). \tag{4.12}$$

Ved at skrive rækkesøjle multiplikationerne i (4.10) ud fås

$$\begin{aligned}&A_{i1} B_{11} C_{1j} + \cdots + A_{in} B_{n1} C_{1j} + && (4.13)\\ &A_{i1} B_{12} C_{2j} + \cdots + A_{in} B_{n2} C_{2j} + && (4.14)\\ &\vdots && (4.15)\\ &A_{i1} B_{1r} C_{rj} + \cdots + A_{in} B_{nr} C_{rj}. && (4.16)\end{aligned}$$

Ved at skrive rækkesøjle multiplikationerne i (4.12) ud fås

$$\begin{aligned}&A_{i1} B_{11} C_{1j} + \cdots + A_{i1} B_{1r} C_{rj} + && (4.17)\\ &A_{i2} B_{21} C_{1j} + \cdots + A_{i2} B_{2r} C_{rj} + && (4.18)\\ &\vdots && (4.19)\\ &A_{in} B_{n1} C_{1j} + \cdots + A_{in} B_{nr} C_{rj}. && (4.20)\end{aligned}$$

Rækkerne i summen i (4.13) svarer til søjlerne i summen (4.17) og det ses at disse summer er ens. Derfor er $((AB)C)_{ij} = (A(BC))_{ij}$. □

♠

4.3.5 Opbygning af matricer fra søjler

Hvis vi har n søjlevektorer $\mathbf{v}_i$ som alle har højde n kan vi danne en $m \times n$ matrix $A = (\mathbf{v}_1, \mathbf{v}_2, \ldots \mathbf{v}_n)$ ved at sætte dem ved siden af hinanden. Så hvis vi har søjlevektorerne

$$\mathbf{v}_1 = \begin{pmatrix} 0 \\ 3 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 2 \\ 2 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} 4 \\ 7 \end{pmatrix} \quad \text{og} \quad \mathbf{v}_4 = \begin{pmatrix} -2 \\ 4 \end{pmatrix}.$$

så er

$$A = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n) = \begin{pmatrix} 0 & 2 & 4 & -2 \\ 3 & 2 & 7 & 4 \end{pmatrix}$$

Vi kan genfinde søjlevektorerne af A som $A^1 = \mathbf{v}_1$,$A^2 = \mathbf{v}_2$,$A^3 = \mathbf{v}_3$ og $A^4 = \mathbf{v}_4$. Vi vil senere få brug for følgende simple udregning.

(4.9) LEMMA.

Hvis B er en $p \times m$ matrix, så er

$$BA = B(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n) = (B\mathbf{v}_1, B\mathbf{v}_2, \ldots, B\mathbf{v}_n)$$

<u>**Bevis**</u>

Vi regner efter. På den ene side er

$$(BA)_{ij} = B_i A^j = B_i \mathbf{v}_j$$

På den anden side er

$$(B\mathbf{v}_1, B\mathbf{v}_2, \ldots, B\mathbf{v}_n)_{ij} = (B\mathbf{v}_j)_{ij} = B_i \mathbf{v}_j$$

♠

<u>**Illustration af beviset ved et eksempel**</u>

$$A = \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix}; \quad B = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 0 & -1 \end{pmatrix} = \left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ -1 \end{pmatrix} \right)$$

Nu er

$$\begin{aligned} AB &= \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 1 & 0 & -1 \end{pmatrix} \\ &= \begin{pmatrix} 1*1+2*1 & 1*2+2*0 & 1*3+2*(-1) \\ 1*1+0*1 & 1*2+0*0 & 1*3+0*(-1) \end{pmatrix} \\ &= \left(\begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 3 \\ -1 \end{pmatrix} \right) \end{aligned}$$

♠

4.3.6 Identitetsmatricen

Identitetsmatricen I_n af orden n er $n \times n$ diagonalmatricen med 1 i diagonalen. Nedenfor er identitetsmatricen af orden 5

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

Identitetsmatricen I_n har egenskaben at

$$I_n A = A I_n = A$$

for alle $n \times n$ matricer A.

4.10 Opgave

Gør rede for ovenstående egenskab det vil sige at identitetsmatricen ikke ændrer ved en kvadratisk matrix når den bliver ganget på enten fra venstre eller fra højre. ♠

4.3.7 Den inverse matrix

Man kan dividere med almindelige tal $\neq 0$. Giver det mening at dividere med matricer? Et almindeligt tal $a \neq 0$ har et "inverst"tal c så $ac = ca = 1$. Her kan vi bare sætte $c = 1/a = a^{-1}$. Vi kan umiddelbart overføre denne definition til kvadratiske matricer.

4.11 Opgave

Lad A, B og C være $n \times n$ matricer. Gør rede for at hvis

$$BA = I_n$$

og

$$AC = I_n$$

så må $B = C$. ♠

(4.12) DEFINITION.

En $n \times n$ matrix A siges at være invertibel, hvis der eksisterer en $n \times n$ matrix B så

$$AB = BA = I_n.$$

I givet fald kaldes B den inverse matrix og betegnes A^{-1}.

(4.13) BEMÆRKNING.

Man kunne jo spørge om det kan ske for en $m \times n$ matrix A hvor $m < n$ at der findes en $n \times m$ matrix B så at $BA = I_n$. Det kan desværre aldrig lade sig gøre, selv om det sagtens kan ske at der findes en $n \times m$ matrix B så at $AB = I_m$.

VIDEO: https://youtu.be/4aeyUedQi0k

Den inverse matrix kommer for eksempel ind i billedet ved løsning af lineære ligninger. Et lineært ligningssystem med n ligninger og n ubekendte:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n \end{aligned}$$

kan med matrixnotation skrives

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

eller mere kompakt som $A\mathbf{x} = \mathbf{b}$.

Hvis A er invertibel giver den associative lov følgende udregning:

$$\begin{aligned} A\mathbf{x} = \mathbf{b} &\implies \\ A^{-1}(A\mathbf{x}) = A^{-1}\mathbf{b} &\implies \\ (A^{-1}A)\mathbf{x} = A^{-1}\mathbf{b} &\implies \\ I\mathbf{x} = A^{-1}\mathbf{b} &\implies \\ \mathbf{x} = A^{-1}\mathbf{b}. & \end{aligned}$$

Den inverse matrix giver altså løsningen til ligningssystemet ud fra kun en matrixmultiplikation med højresiden. Bemærk at denne udregning gælder for alle højresider $\mathbf{b}$ i ligningssystemet.

4.14 Eksempel

Ligningssystemet

$$\begin{array}{rcrcr} 5x & + & 3y & = & 13 \\ 3x & + & 2y & = & 8 \end{array} \tag{4.21}$$

kan ved hjælp af matrixmultiplikation skrives som

$$A\mathbf{v} = \mathbf{b},$$

hvor

$$A = \begin{pmatrix} 5 & 3 \\ 3 & 2 \end{pmatrix}, \qquad \mathbf{v} = \begin{pmatrix} x \\ y \end{pmatrix} \qquad \text{og} \qquad \mathbf{b} = \begin{pmatrix} 13 \\ 8 \end{pmatrix}.$$

Jeg kan her afsløre at A rent faktisk er invertibel samt at

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

En enkel matrixmultiplikation:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 & -3 \\ -3 & 5 \end{pmatrix} \begin{pmatrix} 13 \\ 8 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

afslører som forventet løsningen til ligningssystemet i (4.21). ♠

Produktet af to invertible matricer (når produktet giver mening) er også en invertibel matrix. Dette er indholdet af følgende resultat, som bevises helt formelt ud fra definitionen og den associative lov.

(4.15) PROPOSITION.

Produktet AB af to invertible matricer A og B er invertibelt med $(AB)^{-1} = B^{-1}A^{-1}$.

Bevis

Bevis. Vi skal checke betingelserne i definitionen det vil sige at

$$(B^{-1}A^{-1})(AB) = I \qquad \text{og} \qquad AB(B^{-1}A^{-1}) = I.$$

Lad os checke den første betingelse ved brug af den associative lov:

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

hvor I betegner identitetsmatricen. Betingelsen $AB(B^{-1}A^{-1}) = I$ checkes analogt. □

♠

For de nysgerrige er her en udfordring. Vi har defineret en matrix A til at være invertibel, hvis der findes en matrix B så både $AB = I$ og $BA = I$. Kan vi umiddelbart konkludere at $BA = I$ hvis kun $AB = I$?

Vi vil vende tilbage til denne udfordring senere.

4.3.8 Den transponerede matrix

Den transponerede til en $m \times n$ matrix A er $n \times m$ matricen A^T givet ved

$$A^T_{ij} = A_{ji},$$

det vil sige matricen, som indeholder søjlerne i A som rækker (og rækkerne som søjler). For eksempel er

$$\begin{pmatrix} 0 & 2 & 4 & -2 \\ 3 & 2 & 7 & 4 \end{pmatrix}^T = \begin{pmatrix} 0 & 3 \\ 2 & 2 \\ 4 & 7 \\ -2 & 4 \end{pmatrix}.$$

Læg også mærke til at $(A^T)^T = A$ for en vilkårlig matrix A.

(4.16) PROPOSITION.

Lad A være en $m \times r$ matrix og B en $r \times n$ matrix. Så er

$$(AB)^T = B^T A^T.$$

Bevis

Bevis. Per definition er $(AB)^T_{ij} = (AB)_{ji}$. Denne indgang er givet som række-søjle multiplikation mellem j-te række i A og i-te søjle i B, hvilket er identisk med række-søjle multiplikation mellem i-række i B^T og j-te søjle i A^T. □

♠

4.17 Opgave

Lad A være en kvadratisk matrix. Gør rede for at A er invertibel hvis og kun hvis A^T er invertibel. ♠

4.18 Opgave

En kvadratisk matrix A kaldes symmetrisk hvis $A = A^T$. Gør rede for at

$$BB^T$$

er en symmetrisk matrix, hvor B er en vilkårlig matrix.

Vink

Brug proposition 16! ♠ ♠

4.4 Rækkeoperationer

Der er en række meget naturlige operationer man kan udføre på matricer, som præcis svarer til hvad man ville gøre på det tilsvarende system af ligninger:

1. Ombytning af to rækker.
2. Multiplikation af en række med et tal forskellig fra nul.
3. Addition af en række multipliceret med et tal til en anden række.

Disse operationer kaldes *rækkeoperationer*. Rækkeoperationer er invertible:

• Ved først at ombytte to rækker og dernæst ombytte de samme to rækker genfinder vi den oprindelige matrix.

• Ved først at gange en række med et tal $\lambda \neq 0$ og dernæst gange samme række med $1/\lambda$ genfinder vi den oprindelige matrix.

• Ved at addere λ gange rækken i til rækken j og dernæst addere $-\lambda$ gange rækken i til rækken j genfinder vi den oprindelige matrix.

(4.19) DEFINITION.

To matricer A og B kaldes rækkeækvivalente*, hvis man kan udføre en følge af rækkeoperationer på A og få B frem. Dette skrives $A \sim B$.*

4.20 Quiz

Betragt følgende fire 2×2 matricer

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad C = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix} \quad \text{og} \quad D = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

Hvilke af følgende udsagn er sande?

$A \sim A$

$A \sim B$

$A \sim C$

$B \sim D$ ♠

4.21 Opgave

Lad A, B og C være tre matricer med samme antal rækker og søjler. Gør rede for at

1. $A \sim A$.
2. $A \sim B$ medfører at $B \sim A$.
3. $A \sim B$ og $B \sim C$ medfører at $A \sim C$.

♠

4.22 Opgave

Vis at hvis $A \sim B$ så er $B \sim A$, det vil sige, hvis at man kan udføre en følge af rækkeoperationer på B så at man ender med at få A frem. ♠

Operationen (γ) svarer til Gauss elimination. At trække første ligning fra anden ligning i (4.1) svarer til at gange første række i (4.2) med -1 og addere til anden række. Efter denne operation på matricen (4.2) har vi matricen

$$\begin{pmatrix} 0 & 2 & 4 & -2 \\ 3 & 0 & 3 & 6 \end{pmatrix} \tag{4.22}$$

Vi benytter nu operationen (β) og ganger anden række med $\frac{1}{3}$ og får matricen

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

Tilsvarende ganger vi første række med $\frac{1}{2}$ og får matricen

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

Hvis vi omformulerer matricen ovenfor til ligninger, svarer den til

$$\begin{array}{ccccccc} & & y & + & 2z & = & -1 \\ x & & & + & z & = & 2 \end{array}$$

Intuitivt er rækkefølgen af ligningerne her forkert. Vi vil gerne have at ligningen indeholdende den første variabel x kommer først. Vi benytter operationen (α) og bytter rundt på første og anden række. Dermed har vi

$$\begin{pmatrix} 0 & 2 & 4 & -2 \\ 3 & 0 & 3 & 6 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 1 & 2 \\ 0 & 1 & 2 & -1 \end{pmatrix}. \tag{4.23}$$

Vi accepterer matricen til sidst i (4.23) som en specielt enkel form vi kan reducere den oprindelige matrix (4.2) til. Den enkle form af matricen afspejler sig i det tilsvarende ligningssystem ved at man umiddelbart kan aflæse løsningerne til at være

$$\begin{array}{ccccccc} x & & = & -z & + & 2 \\ & y & = & -2z & - & 1 \end{array}.$$

Det Vil Sige z er en fri variabel og bestemmer x og y som ovenfor. Den simple form vi har reduceret den oprindelige matrix til kaldes *reduceret række echelon form*.

4.5 Reduceret række echelon form (RREF)

En række i en matrix kaldes en *nulrække* hvis alle dens indgange er tallet 0.

(4.23) DEFINITION.

En matrix A siges at være på reduceret række echelon form *(RREF) hvis*

1. *Nulrækker er i bunden af matricen.*
2. *Hvis en række i A ikke er en nulrække, så er den første indgang* $\neq 0$ *i rækken tallet* 1. *Denne indgang kaldes et* pivotelement.
3. *Et pivotelement er længere til højre end pivotelementerne i de foregående rækker.*
4. *Et pivotelement er den eneste indgang* $\neq 0$ *i sin søjle.*

4.24 Quiz

Hvilke af nedenstående matricer er på RREF?

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

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

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

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

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

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

♠

(4.25) SÆTNING.

Enhver $m \times n$ matrix A er rækkeækvivalent med en entydig $m \times n$ matrix på RREF.

Bevis *

Bevis. **En konkret matrix**

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 2 & 4 & 10 \\ 0 & 2 & 1 & 1 \end{pmatrix}$$

$\spadesuit$

Lad j markere første søjle i A, som indeholder en indgang $\neq 0$. Efter ombytning af rækker kan vi antage at $a = A_{1j} \neq 0$.

<u>**Vi følger A**</u>

$$\begin{pmatrix} 0 & 2 & 4 & 10 \\ 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 1 \end{pmatrix}$$

$\spadesuit$

Ved at gange første række igennem med $1/a$ kan vi yderligere antage at $A_{1j} =$ 1.

<u>**Fortsætning**</u>

$$\begin{pmatrix} 0 & 1 & 2 & 5 \\ 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 1 \end{pmatrix}$$

$\spadesuit$

Vi kan så rækkereducere via Gauss elimination ud fra antagelsen $A_{1j} = 1$ og opnå at A_{1j} er eneste indgang $\neq 0$ i sin søjle. Lad os kalde første række R efter disse modifikationer af A.

<u>**Nu ser vi på R**</u>

$$\begin{pmatrix} 0 & 1 & 2 & 5 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -3 & -9 \end{pmatrix} \quad R = \begin{pmatrix} 0 & 1 & 2 & 5 \end{pmatrix}$$

$\spadesuit$

Denne procedure kan gentages på $(m-1) \times n$ matricen B bestående af de sidste $m-1$ rækker i A og vi kan antage at denne mindre matrix kan rækkereduceres til $(m-1) \times n$ matricen H på RREF.

<u>**Og på B og H**</u>

$$B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -3 & -9 \end{pmatrix} \quad H = \begin{pmatrix} 0 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix}$$

$\spadesuit$

Lad $m \times n$ matricen C bestå af R som første række og H som de sidste $m-1$ rækker.

Reduktion til C

$$C = \begin{pmatrix} 0 & 1 & 2 & 5 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$\spadesuit$

RREF for den oprindelige matrix A fremkommer nu ved at benytte pivotelementerne i H til at skabe 0 i første række i C i deres respektive søjler.

Endelig har vi en RREF

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

$\spadesuit$

Dermed har vi vist eksistensen af en RREF. Nu skitserer vi et bevis for entydigheden af RREF (fra en artikel af Thomas Yuster i Mathematics Magazine, March, 1984). Vi bruger et induktionsargument. Induktionen bruger antallet søjler.

Hvis $n = 1$ har A kun en søjle. Der er kun to muligheder. A kunne være nullvektoren $\mathbf{0}$. Men $\mathbf{0}$ er selv på RREF, og den ændrer sig ikke under rækkoperationer, så entydigheden er klar. Hvis $A \neq \mathbf{0}$ er RREF også entydig, fordi rækkereduktionen af A er nødvendigvis søjlevektoren med 1 på første indgang og 0 på de øvrige indgange. Nu har vi klaret induktionsstarten. For at give et fuldstændigt induktionsbevis for entydigheden er det nok at vise at hvis vi allerede har bevist entydighed af RREF for matricer på formen $m \times (n-1)$, så er entydigheden også gældende for matricer på formen $m \times n$.

Vi laver nu induktionskridtet, og antager at $n > 1$, og at sætningen gælder for alle $m \times (n-1)$ matricer. Lad A' betegne $m \times (n-1)$ matricen som fremkommer fra A ved at slette sidste søjle i A. Hvis nu B og C er to RREF, som begge hører til A, kan vi antage at de stemmer overens på de første $n-1$ søjler.

Stop! Hvorfor kan vi antage det?

Jo, fordi hvis vi tilsvarende sletter de sidste søjler i B og C får vi to $m \times (n-1)$ matricer B' og C'. Og B' og C' er begge RREF for A'. Ifølge vores induktionsantagelse er dermed $B' = C'$. $\spadesuit$

B og C kan kun adskille i den sidste søjle, så vores opgave er at vise at også de to sidste søjler er ens, det vil sige at $B^n = C^n$.

Vi skelner nu mellem to tilfælder. Det første tilfælde er at både B^n og C^n er pivotsøjler. Det andet tilfælde er at enten B^n ikke er en pivotsøjle i B, eller at C^n ikke er en pivotsøjle i C. Vi giver et argument der viser at $B = C$ i det første tilfælde, og et helt andet argument der viser at $B = C$ i det andet tilfælde. Tilsammen beviser de to argumenter sætningen.

Bevis for at B og C er ens i det første tilfælde

Antag at B^n og C^n begge er pivotsøjler. Vi kigger først på B. Vi bemærker at pivotsøjlen B^n er bestemt af B'. Hvis vi kender B' og ved at den sidste søjle er en pivotsøjle, så kender vi også B. Pivotelementet står nemlig i den første række i B der er en nullrække i B'.

Sådan her:

$$\begin{pmatrix} 1 & 0 & 6 & 0 & 0 \\ 0 & 1 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

♠ På den samme måde er C bestemt af C', fordi vi ved at C^n er en pivotsøjle.

Men da $B' = C'$ følger det at $B = C$. ♠

Bevis for at B og C er ens i det andet tilfælde

Enten er den sidste række i B eller den sidste række i C ikke en pivotsøjle. Der er ikke forskel på B og C i antagelserne, så vi kan også lige så godt antage at det er B^n der ikke er en pivotsøjle. Det efterfølgende argument virker lige så fint for C, vi skifter bare B ud mod C i notationen.

Vi antager altså at B^n ikke er en pivotsøjle.

Da den sidste række i B ikke er en pivotsøjle, kan vi finde ifølge bemærkning

4.26 Eksempel

Et eksempel til illustration af proceduren i beviset kunne være

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

hvor $j = 2$. Her er

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

og dermed

$$H = \begin{pmatrix} 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Derfor bliver

$$C = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & {\color{red}1} & 1 & 0 \\ 0 & 0 & 0 & 0 & {\color{red}1} \end{pmatrix}$$

og de markerede pivotelementer ovenfor bruges ved Gauss elimination til at give den endelige RREF

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

♠

4.5.1 Løsning af ligninger ved hjælp af RREF

Hvis en $m \times n$ matrix R er på RREF er ligningssystemet $R\mathbf{x} = \mathbf{b}$ med m ligninger og n variable specielt nemt at gå til. Pivotelementerne i R er de eneste indgange i deres søjle $\neq 0$. Deres søjlenumre B svarer til de såkaldte *bundne variable*. De øvrige søjlenumre F svarer til de såkaldte *frie variable*. Vi samler de frie variable i en vektor $\mathbf{x}_F$, og de bundne variable i en anden vektor $\mathbf{x}_B$.

Lad os se på et konkret eksempel. Lad

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

Da er R på RREF, og de tre pivoter står i søjlerne med nummer $2,4,5$. Hvis vi vil løse en ligning $R\mathbf{x} = \mathbf{b}$, så er $\mathbf{x} = (x_1, x_2, x_3, x_4, x_5.x_6)^T$, de bundne variable er x_2, x_4, x_5 og de frie variable er x_1, x_3, x_6. Vi skriver altså $\mathbf{x}_F = (x_1, x_3, x_6)^T$ og $\mathbf{x}_B = (x_2, x_4, x_5)^T$. Vi laver nu en lille omsortering af søjlerne i R. Vi flytter de tre pivot søjle foran. De resterende søjler der svarer til bundne variable samler vi til en matrix vi kalder

$$R' = \begin{pmatrix} 0 & 2 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 3 \end{pmatrix}.$$

Nu ser vi at følgende to ligninger er fuldstændigt ensbetydende:

$$\begin{pmatrix} 0 & 1 & 2 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 2 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} x_2 \\ x_4 \\ x_5 \\ x_1 \\ x_3 \\ x_6 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

Formuleret i matrixsprog siger det at ligningen

$$A = \mathbf{b}$$

er ensbetydende med at

$$I_3\mathbf{x}_B + R'\mathbf{x}_F = \mathbf{b}$$

som er ensbetydende med at

$$\mathbf{x}_B = \mathbf{b} - R'\mathbf{x}_F.$$

For eksempel er $\mathbf{x} = (x_1, x_2, x_3, x_4, x_5, x_6)^T$ en løsning til

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

hvis og kun hvis

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

Her er x_2, x_4, x_5 de bundne variable og x_1, x_3, x_6 de frie variable. Skrevet som ligninger svarer dette til ligningssystemet

$$\begin{array}{ccccccccc} x_2 & + & 2x_3 & & & + & x_6 & = & 1 \\ & & & x_4 & & + & x_6 & = & 2 \\ & & & & x_5 & + & 3x_6 & = & 3 \end{array}$$

med løsningsformlerne

$$\begin{aligned} x_2 &= 1 - 2x_3 - x_6 \\ x_4 &= 2 - x_6 \\ x_5 &= 3 - 3x_6. \end{aligned}$$

Læg mærke til at x_1 er en fri variabel, som her ikke indgår i formlerne for x_2, x_4, x_5.

(4.27) BEMÆRKNING.

Det betyder, for eksempel, at hvis R er en matrix på RREF, og hvis en søjle i matricen R med søjlenummer i ikke indeholder en pivot, så findes der en søjlevektor $\mathbf{u}$ så at hver indgang i $\mathbf{u}$ opfylder at $u_i = 1$, og desuden sådan at $R\mathbf{u} = 0$.

4.28 Opgave

Vi har allerede brugt denne bemærkning i beviset for den vigtige sætning (25). Men måske var det snyd at vi brugte et resultat der står senere i teksten? Lidt som at rejse tilbage i tiden og give sig selv de rigtige lottotal? Argumentér for at vi ikke har snydt (eller argumentér alternativt for at vi har snydt). ♠

VIDEO: https://youtu.be/vcp8J9TqmEs

4.6 Elementære matricer

Vi vil nu omfortolke rækkeoperationer ved hjælp af matrixmultiplikation. Hver af de tre typer af rækkeoperationer som vi beskrev i begyndelsen af

For eksempel er ombytning af rækkke 1 med række 2 i en $3 \times n$ matrix det samme som multiplikation fra venstre med

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

Multiplikation af den anden række med 5 er detsamme som produkt med

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

og operationen at gange den tredie række med -3 og lægge den til den første række er detsamme som multiplikation med matricen

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

4.29 Eksempel

For eksempel giver udtrykket for matrixmultiplikation

$$\begin{aligned} & \begin{pmatrix} 1 & 0 & -3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \\ = & \begin{pmatrix} 1*a+0*d-3*g & 1*b+0*e-3*h & 1*c+0*f-3*i \\ 0*a+1*d+0*g & 0*b+1*e+0*h & 0*c+1*f+0*i \\ 0*a+0*d+1*g & 0*b+0*e+1*h & 0*c+0*f+1*i \end{pmatrix} \\ = & \begin{pmatrix} a-3g & b-3h & c-3i \\ d & e & f \\ g & h & i \end{pmatrix} \end{aligned}$$

♠

4.30 Opgave

Vis de resterende to af de ovenstående påstande for 3×3 matricer ved direkte udregning! ♠

Vi siger at en elementær matrix fremkommer ved at udføre præcis en rækkeoperation på den kvadratiske $m \times m$ identitetsmatrix I_m. Hvis denne rækkeoperation er givet ved at multiplicere fra venstre med en matrix E, er den tilhørende elementære matrix altså $EI_m = E$.

Vi indfører betegnelser for de tre typer af elementære matricer. Lad P_{rs} være den matrix der fremkommer fra identitetsmatricen ved at bytte om på rækkerne med nummer r respektive s. Vi lader $D_i(\lambda)$ betegne matricen, som fremkommer ved at gange i-te række i identitetsmatricen af orden m med λ. Dette er stadig en diagonal matrix, lige som enhedsmatricen, det vil sige at hvis $j \neq k$ er $D_i(\lambda)_{jk} = 0$. Til sidst lader vi $E_{ij}(\lambda)$ betegne den elementære matrix, som fremkommer fra identitetsmatricen af orden m ved at gange j-te række med λ og addere til i-te række. Denne matrix er lig identitetsmatricen med undtagelse af at der i indgangen i i-te række og j-te søjle står λ i stedet for 0.

4.31 Quiz

Et

Lad $A = \begin{pmatrix} 1 & 1 \\ 0 & 3.14 \end{pmatrix}$ Hvad gælder om A?

$A = P_{12}$

$A = D_1(2)$

$A = E_{12}(1)$

A er ikke en elementær matrix. ♠

To

Lad $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ Hvad gælder om A?

$A = P_{12}$

$A = D_1(2)$

$A = E_{12}(1)$

A er ikke en elementær matrix. ♠

Tre

Lad $A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ Hvad gælder om A?

$A = P_{12}$

$A = D_1(2)$

$A = E_{12}(1)$

A er ikke en elementær matrix. ♠

Fire

Lad $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$ Hvad gælder om A?

$A = P_{23}$

$A = D_1(1)$

$A = E_{12}(1)$

A er ikke en elementær matrix. ♠

Fem

Lad $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ Hvad gælder om A?

$A = P_{12}$

$A = D_2(2)$

$A = E_{12}(1)$

A er ikke en elementær matrix. ♠

Seks

Lad $A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ Hvad gælder om A?

$A = P_{12}$

$A = D_2(2)$

$A = E_{12}(1)$

A er ikke en elementær matrix. ♠

Syv

Lad $A = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$ Hvad gælder om A?

$A = P_{13}$

$A = D_2(1)$

$A = E_{12}(0)$

A er ikke en elementær matrix. ♠

Otte

Lad $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3.14 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ Hvad gælder om A?

$A = P_{12}$

$A = D_2(3.14)$

$A = E_{12}(3.14)$

A er ikke en elementær matrix. ♠

♠

(4.32) PROPOSITION.

1. *At udføre en rækkeoperation på en $m \times n$ matrix A svarer til at gange den tilsvarende elementære $m \times m$ matrix på fra venstre.*
2. *En elementær matrix svarende til en rækkeoperation er invertibel. Dens inverse matrix er den elementære matrix svarende til den inverse rækkeoperation.*

Bevis

Bevis. Vi begynder med at bevis for (1). Vi betragter først tilfældet at A er en $m \times 1$ matrix, det vil sige at A er en søjlevektor. For at spare på det dyrbare papir plejer man at skrive en søjlevektor som $(v_1, v_2, \dots, v_m)^T$ hvor T står for transponering, og gamle vaner er svære at give slip på selv når man skriver for skærmen. Nu regner vi ved at bruge formlen for matrixmultiplikation. Følgende to produkter er nemme at beregne:

$$P_{ij}(v_1, v_2, \dots, v_m)^T = (\dots, v_{i-1}, v_j, v_{i+1}, \dots, v_{j-1}, v_i, v_{j+1}, \dots)^T$$
$$D_i(\lambda)(v_1, v_2, \dots, v_m)^T = (\dots, v_{i-1}, \lambda v_i, v_{i+1}, \dots)^T$$

Vi er lidt mere forsigtige i det tredie tilfælde.

$$E_{ij}(\lambda)(v_1, v_2, \dots, v_m)^T = (w_1, \dots, w_m)$$

hvor

$$w_r = \sum_s E_{ij}(\lambda)_{rs} v_s$$

Hvis $r \neq i$ er $E_{ij}(\lambda)_{rs} = 0$ for $r \neq s$, så at

$$w_r = E_{ij}(\lambda)_{rr} v_r = 1 \cdot v_r = v_r.$$

Hvis $r = i$ så er $E_{ij}(\lambda)_{is} = 0$ for $s \neq i$ eller $s \neq j$, så at

$$w_r = E_{ij}(\lambda)_{ri} v_i + E_{ij}(\lambda)_{rj} v_j = 1 \cdot v_i + \lambda \cdot v_j = v_i + \lambda v_j$$

Det vil sige,

$$E_{ij}(\lambda)(v_1, v_2, \dots, v_m)^T = (\dots, v_{i-1}, v_i + \lambda v_j, v_{i+1}, \dots)^T$$

Vi ser at i alle tre tilfælder er multiplikation med en elementær matrix EA detsamme som den tilsvarende rækkeoperation, hvis A er en søjlevektor.

I det generelle tilfælde kan vi skrive matricen A som opbygget af søjlevektorer af højde m, og bruge 9

$$\begin{aligned} A &= (A_1, \dots, A_n) \\ EA &= (EA_1, \dots, EA_n) \end{aligned}$$

Ifølge specialtilfældet brugt på hver søjle A_i, så fremkommer EA fra A ved at bruge den samme rækkeoperation på hver søjle i A. Men det er det samme som at bruge søjleoperationen på A.

Nu ser vi på del (2). Hvis E_1, E_2 er en elementære matricer der svarer til inverse søjleoperationer, så er $E_1 E_2 = E_1 E_2 I_m$ den matrix der fremkommer ved at udføre først søjleoperationen der svarer E_2 på identitetsmatricen, og derefter udføre søjleoperationen der svarer til E_1 på resultatet. Da disse søjleoperationer er inverse, ender vi med at få identitetsmatricen tilbage, det vil sige at $E_1 E_2 = I_m$. Tilsvarende er $E_2 E_1 = I_m$, så at E_1 og E_2 er inverse matricer.

□

♠

Nu er vi endelig i stand til at gengive en algoritme til at udregne den inverse matrix.

(4.33) SÆTNING.

En $n \times n$ matrix A er invertibel hvis og kun hvis dens RREF er I_n. Hvis A er invertibel er RREF for $n \times (2n)$ matricen

$$(A|I_n)$$

lig med $n \times (2n)$ matricen

$$\left(I_n|A^{-1}\right).$$

<u>**Bevis**</u>

Bevis. En $n \times n$ matrix på RREF som ikke er identitetsmatricen bliver nødt til at indeholde en nulrække. Med andre ord, hvis en matrix på RREF er invertibel, så er den nødt til at være identitetsmatricen. Lad os antage at A er invertibel. Som for enhver anden matrix kan vi finde et produkt E af elementære matricer så EA er RREF for A, men da E og EA er invertible bliver denne RREF altså nødt til at være lig identitetsmatricen.

Modsat hvis matricen har RREF lig identitetsmatricen så findes et produkt F af elementære matricer så $FA = I_n$ og A er invertibel med $A^{-1} = F$, da F som produkt af elementære matricer er invertibel.

Den sidste påstand følger ved at gange matricen F ovenfor på matricen $(A|I_n)$. Dette matrixprodukt giver $(FA|FI_n) = (I_n|F)$. Da multiplikation med F fra venstre giver rækkereduktion, og da $(I_n|F) = (I_n|A^{-1})$ er på RREF, er (I_n, A^{-1}) den entydigt bestemte RREF af $(A|I_n)$. □

♠

Ovenstående sætning giver en metode til at udregne den inverse matrix.

VIDEO: https://youtu.be/5hAasPSF3g8

(4.34) SÆTNING.

Lad A være en $n \times n$ matrix. Så er A invertibel hvis og kun hvis ligningssystemet

$$A\mathbf{v} = \mathbf{0}$$

kun har løsningen $\mathbf{v} = \mathbf{0}$.

<u>Bevis</u>

Bevis. Hvis A er invertibel fås

$$A\mathbf{v} = \mathbf{0} \implies A^{-1}(A\mathbf{v}) = A^{-1}\mathbf{0} = \mathbf{0} \implies (A^{-1}A)\mathbf{v} = \mathbf{0} \implies \mathbf{v} = \mathbf{0}.$$

Hvis A ikke er invertibel kan vi rækkereducere A til en matrix B med en nulrække til sidst (se Sætning 33 og Opgave 4.8.12). Men her gælder $A\mathbf{v} = \mathbf{0} \iff B\mathbf{v} = \mathbf{0}$ og ligningsystemet $B\mathbf{v} = \mathbf{0}$ har en løsning $\neq \mathbf{0}$, fordi det har mindst en fri variabel svarende til at den sidste søjle ikke indeholder et pivotelement (se afsnit 4.5.1). □

♠

4.7 Egenvektorer og egenværdier for en matrix

I eksemplet med stokastiske matricer havde vi brug for at udregne potenser $P^2, P^3, \ldots$ af en matrix P. Hvis P er en kvadratisk diagonalmatrix er disse operationer meget mere overkommelige.

(4.35) PROPOSITION.

Hvis

$$D = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}$$

er en kvadratisk diagonalmatrix, så er

$$D^m = \begin{pmatrix} \lambda_1^m & 0 & \ldots & 0 \\ 0 & \lambda_2^m & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n^m \end{pmatrix}.$$

Det vil sige en kvadratisk diagonalmatrix opløftes til en potens ved at opløfte diagonalelementerne til potensen.

Bevis

Bevis. Definition 1 (af matrixmultiplikation) for to diagonalmatricer giver

$$\begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix} \begin{pmatrix} \mu_1 & 0 & \ldots & 0 \\ 0 & \mu_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \mu_n \end{pmatrix} = \begin{pmatrix} \lambda_1\mu_1 & 0 & \ldots & 0 \\ 0 & \lambda_2\mu_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n\mu_n \end{pmatrix}.$$

Formlen for D^m er en konsekvens af dette. $\square$

♠

Det betaler sig at lave P om til en diagonalmatrix for at udregne P^m og jeg vil her kort forklare hvordan dette ofte kan lade sig gøre.

4.7.1 Konjugering

For en invertibel matrix T findes den inverse matrix T^{-1} og udregningen $T^{-1}AT$ giver mening for en kvadratisk matrix A med samme antal rækker som T. Denne operation kaldes *konjugering* med T og matricen $T^{-1}AT$ kaldes en konjugeret matrix til A.

4.36 Quiz

Lad

$$T = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \qquad \text{og} \qquad A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix},$$

hvor $\lambda_1 \neq 0$ og $\lambda_2 \neq 0$. Lad $B = T^{-1}AT$. Hvad er rigtigt af nedenstående?

$$B = \begin{pmatrix} a_{11} & \frac{\lambda_2}{\lambda_1}a_{12} \\ \frac{\lambda_1}{\lambda_2}a_{21} & a_{22} \end{pmatrix}.$$

$$B = \begin{pmatrix} \frac{\lambda_2}{\lambda_1}a_{11} & a_{12} \\ \frac{\lambda_1}{\lambda_2}a_{21} & a_{22} \end{pmatrix}.$$

$$AT = TA.$$

$$B = A$$

hvis $\lambda_1 = \lambda_2$. ♠

Konjugering viser sig at være central i lineær algebra. Lad os se på et eksempel.

Lad os antage et øjeblik at matricen P kan konjugeres til en diagonalmatrix det vil sige vi kan finde en invertibel matrix T så

$$T^{-1}PT = D,$$

hvor D er en diagonalmatrix. Så vil

$$P = TDT^{-1}$$

og dermed

$$P^2 = (TDT^{-1})(TDT^{-1}) = TD(T^{-1}T)DT^{-1} = TD^2T^{-1}.$$

Nøjagtig den samme udregning kan laves ikke bare for potensen 2, men for en generel potens m:

$$P^m = TD^mT^{-1}.$$

Det Vil Sige hvis vi er så heldige at finde en invertibel matrix T, således at $T^{-1}PT$ er en diagonalmatrix, så kan vi udregne potenser af P meget nemmere end ved almindelig matrixmultiplikation. Der er slet ikke sikkert at T findes, men vi kan prøve på at analysere hvad matricen P skal opfylde for at det lader sig gøre.

(4.37) SÆTNING.

Lad P være en $n \times n$ matrix, T en invertibel $n \times n$ matrix med søjlevektorer $\mathbf{v}_1, \ldots, \mathbf{v}_n$ og D diagonalmatricen

$$D = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}.$$

Så gælder $T^{-1}PT = D$ hvis og kun hvis

$$P\mathbf{v}_i = \lambda_i \mathbf{v}_i$$

for $i = 1, \ldots, n$.

<u>**4.38 Bevis**</u>

Bevis. $T^{-1}PT = D$ gælder hvis og kun hvis $PT = TD$. Per definition af matrix-multiplikation følger det at søjlevektorerne i PT er $P\mathbf{v}_i$ for $i = 1, \ldots, n$ samt at de tilsvarende søjlevektorer i TD er $\lambda_i \mathbf{v}_i$. □

♠

Disse overvejelser leder frem til følgende definitioner.

(4.39) DEFINITION.

Lad P være en kvadratisk matrix.

1. *P kaldes* diagonaliserbar *hvis der findes en invertibel matrix T så*

$$T^{-1}PT$$

er en diagonalmatrix.

2. *En vektor* $\mathbf{v}$ *kaldes en* egenvektor *for P, hvis* $\mathbf{v} \neq \mathbf{0}$ *og*

$$P\mathbf{v} = \lambda \mathbf{v},$$

for et tal λ *(som gerne må være 0). Dette tal kaldes for en* egenværdi *for P og* $\mathbf{v}$ *siges at være en egenvektor hørende til* λ*.*

Det er ikke oplagt med vores viden nu om en matrix overhovedet har endeligt mange egenværdier eller hvordan man bærer sig ad med at regne egenværdier ud. Lad os prøve at kigge på 2×2 matricer.

4.7.2 Hvad sker der for små matricer?

At finde egenværdier for en kvadratisk $n \times n$ matrix A kan omformuleres til at at finde et tal λ (en egenværdi), så der findes en vektor $\mathbf{v} \neq \mathbf{0}$ med $A\mathbf{v} = \lambda \mathbf{v}$. Dette er det samme som at der findes en vektor $\mathbf{v} \neq \mathbf{0}$ med

$$(A - \lambda I_n)\mathbf{v} = 0. \tag{4.24}$$

Lad os i dette lille afsnit foregribe begivenhedernes gang ved at kigge på en 2×2 matrix

$$B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$

og spørgsmålet: Hvornår findes en vektor $\mathbf{v} \neq 0$ så $B\mathbf{v} = 0$? Vi ved fra Sætning 34 at dette forekommer præcis når B ikke er invertibel.

Samtidig ved vi fra Sætning 33 at B er invertibel hvis og kun B er rækkeækvivalent med identitetsmatricen. Lad os eksperimentere: Hvis både b_{11} og b_{21} er 0 kan B ikke være invertibel. Hvis $b_{11} \neq 0$ så er

$$B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \sim \begin{pmatrix} b_{11} & b_{12} \\ 0 & b_{22} - \frac{b_{21}}{b_{11}} b_{12} \end{pmatrix}$$

og dermed er B invertibel hvis og kun hvis

$$b_{11}\left(b_{22} - \frac{b_{21}}{b_{11}} b_{12}\right) = b_{11}b_{22} - b_{21}b_{12} \neq 0.$$

Samme betingelse gør sig gældende ved rækkereduktioner ud fra antagelsen $b_{21} \neq 0$. Vi kalder $b_{11}b_{22} - b_{21}b_{12}$ for determinanten for B og betegner den $\det(B)$.

Nu kan vi svare på hvornår

$$\begin{pmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{pmatrix} v = 0$$

har en løsning $v \neq 0$ for en 2×2 matrix

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

i (4.24). Dette gælder hvis og kun hvis

$$\det \begin{pmatrix} a_{11}-\lambda & a_{12} \\ a_{21} & a_{22}-\lambda \end{pmatrix} = \lambda^2 - (a_{11}+a_{22})\lambda + \det(A) = 0. \tag{4.25}$$

Polynomiet i (4.25) kaldes for det karakteristiske polynomium hørende til A. Det vi har vist er altså at en 2×2 matrix A har mindst en egenvektor hørende til egenværdien λ hvis og kun hvis λ er en rod i det karakteristiske polynomium. I næste kapitel kommer vi ind på hvad der sker for større matricer ved at definere determinanten af en generel $n\times n$ matrix.

4.7.3 Differentialligninger som eksempel

Egenværdier og egenvektorer er ekstremt nyttige ved løsning af koblede differentialligninger som

$$x_1'(t) = ax_1(t) + bx_2(t) \tag{4.26}$$
$$x_2'(t) = cx_1(t) + dx_2(t), \tag{4.27}$$

hvor $a,b,c,d \in \mathbb{R}$.

Tilfældet med kun en ubekendt funktion kendes fra radioaktivt henfald. Her støder vi på differentialligningen

$$x'(t) = \lambda x(t),$$

som har løsningen $x(t) = Ce^{\lambda t}$, hvor C er en konstant. Hvis man arbejder ud fra hypotesen om at (4.26) har løsninger af formen

$$x_1(t) = Ae^{\lambda t} \qquad \text{og} \qquad x_2(t) = Be^{\lambda t}$$

så kan man indsætte i (4.26) og komme frem til at

$$\begin{pmatrix} A \\ B \end{pmatrix} \quad \text{erenegenvektorfor} \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

hørende til egenværdien λ. Dette er gennemgået i videoen nedenfor.

VIDEO: https://youtu.be/oKziKRz0KzU

4.8 Opgaver

4.8.1

Lad

$$P = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix}$$

være en stokastisk matrix det vil sige alle indgangene i matricen er ≥ 0 og $p_{11} + p_{21} = 1$ samt $p_{12} + p_{22} = 1$.

Antag at $p_{21} + p_{12} > 0$ og lad v være vektoren

$$\begin{pmatrix} \dfrac{p_{12}}{p_{21} + p_{12}} \\ \dfrac{p_{21}}{p_{21} + p_{12}} \end{pmatrix}$$

Hvorfor er $Pv = v$? Hvordan relaterer det til Eksempel 4.2 om stokastiske matricer?

4.8.2

Lad A og B være invertible $n \times n$ matricer. Gør detaljeret rede for at $AB(B^{-1}A^{-1}) = I$ ved brug af den associative lov.

4.8.3

Forklar hvorfor matricen

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

ikke er invertibel.

4.8.4

For hvilke tal x er matricen

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

invertibel.

4.8.5

Lad

$$A = \begin{pmatrix} 1 & -2 & -7 & -8 & -9 \\ 3 & -4 & -13 & -14 & -15 \end{pmatrix}.$$

1. Find den reducerede række echelon form for A.

2. Find samtlige løsninger til ligningssystemet

$$\begin{aligned} x - 2y - 7z - 8w &= -9 \\ 3x - 4y - 13z - 14w &= -15. \end{aligned}$$

4.8.6

Udregn den inverse matrix til matricen

$$\begin{pmatrix} i & 1 & 1 \\ 1 & i & 1 \\ 1 & 1 & -i \end{pmatrix}$$

og gør rede for alle trin i din beregning.

4.8.7

Skriv matricen

$$\begin{pmatrix} 5 & 3 \\ 3 & 2 \end{pmatrix}$$

som et produkt af elementære matricer.

4.8.8

Giv alle detaljer i udregningen (4.25).

4.8.9

Lad

$$A = \begin{pmatrix} 7 & -2 \\ 15 & -4 \end{pmatrix}.$$

Bestem egenværdierne for A og egenværdierne for A^{10}. Hvad er egenvektorerne for A^{10}? Begrund dine svar.

4.8.10

Find, ved at bruge teorien i dette kapitel, to funktioner $x_1(t), x_2(t) : \mathbb{R} \to \mathbb{R}$ som udgør en løsning til systemet

$$\begin{aligned} x_1'(t) &= x_1(t) + x_2(t) \\ x_2'(t) &= -x_1(t) + x_2(t), \end{aligned}$$

af differentialligninger og som opfylder $x_1(0) = 1$ og $x_2(0) = 0$. Skitser din metode grundigt og henvis kun til materialet i disse noter.

4.8.11

I udregningen (4.6) i eksemplet om stokastiske matricer er der urent trav i forklaringerne. Hvad er der galt?

4.8.12

Gør rede for at en kvadratisk matrix forskellig fra identitetsmatricen bliver nødt til at indeholde en nulrække hvis den er på RREF.

4.8.13

Lad A og B være to 2×2 matricer. Er det rigtigt at

$$(A+B)^2 = (A+B)(A+B) = A^2 + B^2 + 2AB?$$

Hvad med

$$(A+B)(A-B) = A^2 - B^2?$$

Kapitel 5

Determinanter

VIDEO: https://youtu.be/hP7MruBdSOA

I sidste kapitel fandt vi ud af at determinanten for en 2×2 matrix

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc \tag{5.1}$$

er forskellig fra nul hvis og kun hvis matricen er invertibel.

I dette kapitel vil vi definere *determinanten*, $\det(A)$ for en generel $n \times n$ matrix A, som viderefører denne magiske egenskab: A er invertibel hvis og kun hvis $\det(A) \neq 0$. Fremstillingen er inspireret af Artin's legendariske algebrabog.

Dette leder til sidst i kapitlet frem til en metode til udregning af egenværdier for en vilkårlig kvadratisk matrix.

5.1 Definition

Vi definerer determinanten for en 1×1 matrix som $\det(a) = a$ for derefter at definere determinanten $\det(A)$ for en $n \times n$ matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \tag{5.2}$$

induktivt ud fra formlen

$$a_{11} \det(A_{11}) - \cdots + (-1)^{i+1} a_{i1} \det(A_{i1}) + \cdots + (-1)^{1+n} a_{n1} \det(A_{n1}), \tag{5.3}$$

hvor A_{ij} er $(n-1) \times (n-1)$ matricen, som fremkommer ved at slette i-te række og j-te søjle.

Definitionen er "induktiv", fordi før vi kan beregne determinanten af en $n \times n$ matrix med formlen, så er vi nødt til at kunne beregne determinanten af A_{i1}, det vil sige determinanten af en $(n-1) \times (n-1)$ matrix. Vi kunne sådan set skrive ned en formel for determinanten som ikke er induktiv. Ulempen ved dette er at denne formel bliver meget lang og stor.

Hvis en matrix ikke er kvadratisk, altså for eksempel en 2×3 matrix og den slags, definerer vi ikke dens determinant

5.1 Quiz

Lad

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

Hvilke af nedenstående udsagn er rigtige?

$$A_{12} = \begin{pmatrix} 0 & 0 & 1 \\ 3 & 1 & 0 \\ 1 & 3 & 4 \end{pmatrix}$$

$$A_{22} = \begin{pmatrix} 1 & 2 & 1 \\ 3 & 2 & 1 \\ 1 & 2 & 3 \end{pmatrix}$$

$$A_{31} = \begin{pmatrix} 2 & 1 & 2 \\ 1 & 0 & 1 \\ 2 & 3 & 4 \end{pmatrix}$$

$$A_{21} = \begin{pmatrix} 2 & 1 & 2 \\ 2 & 1 & 0 \\ 2 & 3 & 4 \end{pmatrix}$$

♠

Læg mærke til fortegnsmønsteret $+-+-\cdots$ i (5.3). For en 2×2 matrix ses ud fra (5.3) at

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

i fin overensstemmelse med (5.1).

For 3×3 matricer, det vil sige $n = 3$ i (5.3), udleder vi en formel som ikke er induktiv på følgende måde:

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{11}(a_{22}a_{33} - a_{32}a_{23}) - a_{21}(a_{12}a_{33} - a_{32}a_{13}) + a_{31}(a_{12}a_{23} - a_{22}a_{13})$$

$$= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}.$$

Det virker nu klart hvorfor man ikke har lyst til at gentage denne spøg for 4×4 matricer (her er der hele 24 led i formlen for determinanten. For 3×3 matricer var der kun 6). Formlen for determinanten vokser eksplosivt. For eksempel, for at beregne determinanten af en 10×10 matrix består den tilsvarende formel af 3628800 termer, og alle dem gider vi lægge sammen. Heldigvis er der langt hurtigere metoder til at udregne determinanten, men det kræver vi kigger nærmere på dens egenskaber.

5.2 Egenskaber

Vi udleder her nogle helt fundamentale egenskaber for determinanten, specielt at den ikke ændrer sig ved addition af et multiplum af en række til en anden række. I det følgende betegner $A_1, \ldots, A_n$ rækkerne i $n \times n$ matricen A.

5.2.1 Determinanten af identitetsmatricen

Ud fra definitionen (5.3), ses for identitetsmatricen I_n at

$$\det(I_n) = \det(I_{n-1}),$$

hvor $n > 1$. Ud fra $\det(I_1) = 1$ sluttes altså generelt at

$$\det(I_n) = 1.$$

> Determinanten af identitetsmatricen er 1.

5.2.2 Multiplikation af en række med et tal

Hvis en række multipliceres med et tal ganges determinanten med tallet:

$$\det \begin{pmatrix} A_1 \\ \vdots \\ \lambda A_i \\ \vdots \\ A_n \end{pmatrix} = \lambda \det \begin{pmatrix} A_1 \\ \vdots \\ A_i \\ \vdots \\ A_n \end{pmatrix},$$

hvor λ er et tal. For eksempel er

$$\det\begin{pmatrix}1 & 2 & 3\\ 21 & 9 & 12\\ 0 & 0 & 1\end{pmatrix} = 3\cdot\det\begin{pmatrix}1 & 2 & 3\\ 7 & 3 & 4\\ 0 & 0 & 1\end{pmatrix}.$$

Denne er påstand er rigtig for $n = 1$. Ved at antage at den er rigtig for $(n-1)\times(n-1)$ matricer fås ved indsættelse i (5.3), at

$$a_{11}\lambda\det(A_{11}) - \cdots + (-1)^{i+1}\lambda a_{i1}\det(A_{1i}) + \cdots + a_{n1}\lambda\det(A_{n1}) = \lambda\det(A),$$

hvilket viser påstanden for $n\times n$ matricer.

Igen, prins Knud har lige et spørgsmål

Vi skriver ud eksemplet ovenfor, fordi det skulle være nok til at være helt overbevisende.

$$\det\begin{pmatrix}1 & 2 & 3\\ 3\cdot 7 & 3\cdot 3 & 3\cdot 4\\ 0 & 0 & 1\end{pmatrix} = 1\cdot\det\begin{pmatrix}3\cdot 3 & 3\cdot 4\\ 0 & 1\end{pmatrix} - 3\cdot 7\cdot\begin{pmatrix}2 & 3\\ 0 & 1\end{pmatrix} + 0\cdot\begin{pmatrix}2 & 3\\ 3\cdot 3 & 3\cdot 4\end{pmatrix}$$

Nu bruger vi induktionsantagelsen på termerne nummer 1 og 3, og får at determinaten er det samme som

$$\begin{aligned}
1\cdot 3\cdot\det\begin{pmatrix}3 & 4\\ 0 & 1\end{pmatrix} - 3\cdot 7\cdot\det\begin{pmatrix}2 & 3\\ 0 & 1\end{pmatrix} + 0\cdot 3\cdot\det\begin{pmatrix}2 & 3\\ 3 & 4\end{pmatrix}&\\
= 3\cdot\left(1\cdot\det\begin{pmatrix}3 & 4\\ 0 & 1\end{pmatrix} - 7\cdot\det\begin{pmatrix}2 & 3\\ 0 & 1\end{pmatrix} + 0\cdot\det\begin{pmatrix}2 & 3\\ 3 & 4\end{pmatrix}\right)&\\
= 3\cdot\det\begin{pmatrix}1 & 2 & 3\\ 7 & 3 & 4\\ 0 & 0 & 1\end{pmatrix}&
\end{aligned}$$

♠

Dette lille argument er et nyt simpelt eksempel på et induktionbevis.

> Når en række i en matrix ganges med et tal ændres determinanten ved at gange med tallet.

5.2 Quiz

Benyt den viden du allerede har nu til at afgøre hvilke af følgende udsagn, der er rigtige.

$$\det\begin{pmatrix}1&0&0\\0&2&0\\0&0&3\end{pmatrix}=6.$$

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

$$\det\begin{pmatrix}1&1&1\\0&2&1\\0&0&3\end{pmatrix}=6.$$

$$\det\begin{pmatrix}1&1&1\\1&1&1\\1&1&1\end{pmatrix}=1.$$

♠

5.2.3 Additivitet af rækker

Determinanten er additiv i rækkerne i den forstand at

$$\det(B)=\det\begin{pmatrix}A_1\\\vdots\\A_i+A_i'\\\vdots\\A_n\end{pmatrix}=\det\begin{pmatrix}A_1\\\vdots\\A_i\\\vdots\\A_n\end{pmatrix}+\det\begin{pmatrix}A_1\\\vdots\\A_i'\\\vdots\\A_n\end{pmatrix},$$

hvor A_i' er en rækkevektor.

For eksempel er

$$\det\begin{pmatrix}111&222&3333\\1+4&2+5&3+6\\777&888&999\end{pmatrix}=\det\begin{pmatrix}111&222&3333\\1&2&3\\777&888&999\end{pmatrix}+\det\begin{pmatrix}111&222&3333\\4&5&6\\777&888&999\end{pmatrix}.$$

Denne er påstand er rigtig for $n=1$. Ved at antage at den er rigtig for $(n-1)\times(n-1)$ matricer fås ved indsættelse i (5.3), at determinanten af den første matrix er

$$\begin{aligned}\det(B)=a_{11}(\det(A_{11})+\det(A'_{11}))-\cdots+(-1)^i a_{(i-1)1}(\det(A_{(i-1)1})+\det(A'_{(i-1)1}))\\+(-1)^{i+1}(a_{i1}+a'_{i1})\det(A_{1i})\\+(-1)^{i+2}a_{(i+1)1}(\det(A_{(i+1)1})+\det(A'_{(i+1)1}))\cdots+(-1)^{n+1}a_{n1}(\det(A_{n1})+\det(A'_{n1}))\\=\det(A)+\det(A'),\end{aligned}$$

hvor A' er matricen A med i-te række udskiftet med rækken A'_i. For at producere det sidste lighedstegn i denne udregning har vi brugt at $A_{1i} = A'_{1i}$ og $a_{j1} = a'_{j1}$ for $j \neq i$. Dette viser påstanden for $n \times n$ matricer.

Altså, et induktionbevis igen.

5.2.4 To ens naborækker

Hvis en matrix består af rækkerne $A_1, A_2, \ldots, A_n$ og $A_i = A_{i+1}$ for et $i = 1, \ldots, n-1$, så er $\det(A) = 0$.

Påstanden er korrekt for 2×2 matricer. Ved at antage at den er rigtig for $m \times m$ matricer med $m < n$ fås ved indsættelse i (5.3), at

$$\det(A) = (-1)^{i+1} a_{i1} \det(A_{i1}) + (-1)^{i+2} a_{i+1,1} \det(A_{i+1,1}) = 0,$$

da $A_{i1} = A_{i+1,1}$ og $a_{i1} = a_{i+1,1}$. Dermed er påstanden også korrekt for $n \times n$ matricer.

5.2.5 Naborækkeoperation

Resultatet i afsnit 5.2.4 medfører at

$$\det \begin{pmatrix} A_1 \\ \vdots \\ A_i \\ A_{i+1} + \lambda A_i \\ \vdots \\ A_m \end{pmatrix} = \det(A) + \lambda \det \begin{pmatrix} A_1 \\ \vdots \\ A_i \\ A_i \\ \vdots \\ A_m \end{pmatrix} = \det(A),$$

hvor λ er et tal.

5.2.6 Fortegnsskift ved ombytning af naborækker

Resultaterne i afsnit 5.2.2, 5.2.3 og 5.2.5 medfører at determinanten skifter fortegn, når man ombytter to naborækker:

$$\det \begin{pmatrix} A_1 \\ \vdots \\ A_i \\ A_{i+1} \\ \vdots \\ A_m \end{pmatrix} = \det \begin{pmatrix} A_1 \\ \vdots \\ A_i \\ A_{i+1} - A_i \\ \vdots \\ A_m \end{pmatrix} = \det \begin{pmatrix} A_1 \\ \vdots \\ A_{i+1} \\ A_{i+1} - A_i \\ \vdots \\ A_m \end{pmatrix} = \det \begin{pmatrix} A_1 \\ \vdots \\ A_{i+1} \\ -A_i \\ \vdots \\ A_m \end{pmatrix} = -\det \begin{pmatrix} A_1 \\ \vdots \\ A_{i+1} \\ A_i \\ \vdots \\ A_m \end{pmatrix}.$$

5.2.7 To ens rækker

Resultatet i afsnit 5.2.6 medfører følgende.

Determinanten til en matrix med to ens rækker er 0.

Hvis en matrix har to ens rækker kan vi nemlig efter ombytning mellem naborækker opnå at disse to ens rækker er naborækker. Så følger resultatet af afsnit 5.2.4.

5.2.8 Rækkeoperation

På samme måde som i afsnit 5.2.5 medfører afsnit 5.2.7 nu at determinanten ikke ændres ved addition af et multiplum af en række til en anden række helt generelt:

$$\det\begin{pmatrix} A_1 \\ \vdots \\ A_i \\ \vdots \\ A_j + \lambda A_i \\ \vdots \\ A_m \end{pmatrix} = \det(A) + \lambda \det\begin{pmatrix} A_1 \\ \vdots \\ A_i \\ \vdots \\ A_i \\ \vdots \\ A_m \end{pmatrix} = \det(A),$$

hvor λ er et tal.

Determinanten ændres ikke når et multiplum af en række adderes til en anden række.

5.2.9 Ombytning af to rækker

På præcis samme måde som i afsnit 5.2.6 kan vi nu vise at

$$\det\begin{pmatrix} A_1 \\ \vdots \\ A_i \\ \vdots \\ A_j \\ \vdots \\ A_m \end{pmatrix} = -\det\begin{pmatrix} A_1 \\ \vdots \\ A_j \\ \vdots \\ A_i \\ \vdots \\ A_m \end{pmatrix}.$$

Determinanten skifter fortegn når to rækker ombyttes.

5.3 Udregning af determinanten

Vi har nu udledt nok egenskaber ved determinanten til at angive en meget effektiv måde at regne den ud på. Vi følger ganske enkelt rækkeoperationerne på vej til dens RREF. De eneste rækkeoperationer, som ændrer determinanten er ombytning af rækker samt multiplikation af en række med et tal.

Lad os se denne procedure i et par eksempler:

1.

$$\det\begin{pmatrix}1 & 2 & 3\\ 4 & 5 & 6\\ 7 & 8 & 9\end{pmatrix} = \det\begin{pmatrix}1 & 2 & 3\\ 0 & -3 & -6\\ 0 & -6 & -12\end{pmatrix} = \det\begin{pmatrix}1 & 2 & 3\\ 0 & -3 & -6\\ 0 & 0 & 0\end{pmatrix} = 0,$$

idet en matrix med en nulrække har determinant 0 (Se Opgave 5.7.2).

2.

$$\det\begin{pmatrix}0 & 1 & 1\\ 1 & 0 & 1\\ 1 & 1 & 0\end{pmatrix} = -\det\begin{pmatrix}1 & 0 & 1\\ 0 & 1 & 1\\ 1 & 1 & 0\end{pmatrix} = -\det\begin{pmatrix}1 & 0 & 1\\ 0 & 1 & 1\\ 0 & 1 & -1\end{pmatrix}$$

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

I videoen nedenfor prøver vi kræfter med et lidt større eksempel og giver også en lille ekstra udfordring.

VIDEO: https://youtu.be/vPNLbMtiDfs

5.3 Quiz

Lad

$$A = \begin{pmatrix}1 & 1 & 1\\ 1 & 1 & x\\ x & 1 & 1\end{pmatrix}.$$

Hvilke af nedenstående udsagn er rigtige?

$\det(A) = (x-1)^2$.

$\det(A) = x^2 + 2x - 3$.

$\det(A) = x^2 + x - 2$. ♠

(5.4) SÆTNING.

En matrix A er invertibel hvis og kun hvis $\det(A) \neq 0$.

Bevis

Bevis. En sætning siger at matricen A er invertibel hvis og kun hvis den via rækkeoperationer kan omformes til identitetsmatricen. Under hver af rækkeoperationerne ændres determinanten ved multiplikation af et tal $\neq 0$. Da determinanten af identitetsmatricen er 1 følger resultatet. □

♠

5.4 Determinanten af elementære matricer

I sidste kapitel introducerede vi tre typer elementære matricer $E_{ij}(\lambda), P_{ij}$ og $D_i(\lambda)$:

1. $E_{ij}(\lambda)$ betegner den elementære matrix, som fremkommer via identitetsmatricen af orden n ved at gange j-te række med λ og addere til i-te række. Denne matrix er lig identitetsmatricen med undtagelse af indgangen λ der står i i-te række og j-te søjle.
2. P_{ij} betegner matricen, som fremkommer ved at bytte rundt på i-te række og j-te række i identitetsmatricen af orden n.
3. $D_i(\lambda)$ betegne matricen, som fremkommer ved at gange i-te række i identitetsmatricen af orden n med $\lambda \neq 0$.

Alle invertible matricer kan skrives som et produkt af disse. Ud fra determinantens egenskaber ovenfor ser vi at

1. $\det(E_{ij}(\lambda)) = 1$.
2. $\det(P_{ij}) = -1$.
3. $\det(D_i(\lambda)) = \lambda$

Specielt ses at

$$\det(EA) = \det(E)\det(A), \tag{5.4}$$

hvor A er en vilkårlig $n \times n$ matrix og E en elementær matrix. Dette leder frem til følgende.

(5.5) SÆTNING.

Lad A og B være $n \times n$ matricer. Så gælder

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

<u>Bevis</u>

Bevis. Generelt findes en matrix C, som er et produkt af elementære matricer så at CA er på RREF. Hvis A ikke er invertibel indeholder både CA og CAB en nulrække og dermed er $\det(A) = 0$ og $\det(AB) = 0$. Det vil sige formlen gælder i dette tilfælde. Hvis A er invertibel kan A skrives som et produkt af elementære matricer. Dermed er

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

ved gentagen anvendelse af (5.4). □

♠

En anvendelse af Sætning 5.5 med $B = A^{-1}$ giver følgende resultat.

> Hvis A er en invertibel matrix, så er
>
> $$\det(A^{-1}) = \frac{1}{\det(A)}.$$

(5.6) SÆTNING.

Lad A være en kvadratisk matrix. Så er

$$\det(A^T) = \det(A).$$

<u>Bevis</u>

Bevis. I en nem opgave har vi set at A er invertibel hvis og kun hvis A^T er invertibel. Formlen gælder derfor hvis A ikke er invertibel, da vi dermed har $\det(A^T) = 0$ og $\det(A) = 0$. Hvis A er invertibel kan A skrives som et produkt af elementære matricer. En elementær matrix E opfylder at

$$\det(E^T) = \det(E).$$

Derfor har vi $\det(A^T) = \det(A)$ ved gentagen anvendelse af Sætning 5.5 i dette tilfælde. □

♠

Søjleoperationer kan identificeres med rækkeoperationer på den transponerede matrix. Læg mærke til at Sætning 5.6 viser at determinanten ændres på præcis samme måde ved søjleoperationer som ved rækkeoperationer. For eksempel har vi

1. determinanten er uændret ved addition af multiplum af en søjle til en anden søjle:
$$\det\begin{pmatrix}1 & 1 & 1\\ 1 & 2 & 4\\ 1 & 3 & 9\end{pmatrix} = \begin{pmatrix}1 & 1 & 0\\ 1 & 2 & 3\\ 1 & 3 & 8\end{pmatrix}$$

2. determinanten skifter fortegn ved ombytning af to søjler:
$$\det\begin{pmatrix}1 & 1 & 1\\ 1 & 2 & 4\\ 1 & 3 & 9\end{pmatrix} = -\det\begin{pmatrix}1 & 1 & 1\\ 4 & 2 & 1\\ 9 & 3 & 1\end{pmatrix}$$

3. hvis en søjle ganges med et tal ganges determinanten med tallet:
$$\det\begin{pmatrix}2 & 1 & 1\\ 2 & 2 & 4\\ 2 & 3 & 9\end{pmatrix} = 2\det\begin{pmatrix}1 & 1 & 1\\ 1 & 2 & 4\\ 1 & 3 & 9\end{pmatrix}$$

5.5 Polynomier af grad n gennem $n+1$ punkter

Under brug af determinanter kan vi nu give et bevis for at der gennem $n+1$ punkter

$$(x_0, y_0), (x_1, y_1), \ldots, (x_n, y_n) \in \mathbb{R}^2 \tag{5.5}$$

altid findes et og kun et polynomium

$$f(x) = a_0 + a_1 x + \cdots + a_n x^n \tag{5.6}$$

af grad n, som går gennem punkterne forudsat at x-værdierne er forskellige det vil sige $x_i \neq x_j$ for $i \neq j$.

At polynomiet (5.6) går gennem punkterne i (5.5) betyder at $f(x_i) = y_i$ for $i = 0, \ldots, n$. Denne betingelse kan oversættes til ligningssystemet

$$\begin{array}{ccccccccccc} a_0 & + & a_1 x_0 & + & a_2 x_0^2 & + & \cdots & + & a_n x_0^n & = & y_0 \\ a_0 & + & a_1 x_1 & + & a_2 x_1^2 & + & \cdots & + & a_n x_1^n & = & y_1 \\ & & & & & & & & & \vdots & \\ a_0 & + & a_1 x_n & + & a_2 x_n^2 & + & \cdots & + & a_n x_n^n & = & y_n \end{array} \tag{5.7}$$

med $n+1$ ligninger i de $n+1$ ubekendte $a_0, a_1, \dots, a_n$. Ligningssystemet (5.7) skrives op på matrixform som

$$\begin{pmatrix} 1 & x_0 & x_0^2 & \dots & x_0^n \\ 1 & x_1 & x_1^2 & \dots & x_1^n \\ \vdots & \vdots & \ddots & \vdots & \\ 1 & x_n & x_n^2 & \dots & x_n^n \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

Matricen

$$\begin{pmatrix} 1 & x_0 & x_0^2 & \dots & x_0^n \\ 1 & x_1 & x_1^2 & \dots & x_1^n \\ \vdots & \vdots & \ddots & \vdots & \\ 1 & x_n & x_n^2 & \dots & x_n^n \end{pmatrix} \tag{5.8}$$

kaldes Vandermonde matricen (efter Alexandre-Théophile Vandermonde) med hensyn til $x_0, x_1, \dots, x_n$. Man kan vise generelt at determinanten af Vandermonde matricen i (5.8) er

$$(x_1 - x_0)(x_2 - x_0)(x_2 - x_1)(x_3 - x_0)(x_3 - x_1)(x_3 - x_2) \cdots (x_n - x_{n-1}). \tag{5.9}$$

Specielt er determinanten af Vandermonde matricen $\neq 0$ hvis og kun hvis x-erne er forskellige det vil sige $x_i \neq x_j$ for $i \neq j$.

Ideen til beviset for formlen (5.9) kommer fra tilfældet $n = 2$. Alt hvad man benytter er reglerne 5.2.8 og 5.2.2 samt Sætning 5.6:

$$\det \begin{pmatrix} 1 & x_0 & x_0^2 \\ 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \end{pmatrix} = \det \begin{pmatrix} 1 & x_0 & 0 \\ 1 & x_1 & x_1^2 - x_0 x_1 \\ 1 & x_2 & x_2^2 - x_0 x_2 \end{pmatrix} = \det \begin{pmatrix} 1 & 0 & 0 \\ 1 & x_1 - x_0 & x_1^2 - x_0 x_1 \\ 1 & x_2 - x_0 & x_2^2 - x_0 x_2 \end{pmatrix}$$

$$= \det \begin{pmatrix} 1 & 0 & 0 \\ 0 & x_1 - x_0 & x_1^2 - x_0 x_1 \\ 0 & x_2 - x_0 & x_2^2 - x_0 x_2 \end{pmatrix}$$

$$= (x_1 - x_0) \det \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & x_1 \\ 0 & x_2 - x_0 & x_2^2 - x_0 x_2 \end{pmatrix}$$

$$= (x_1 - x_0)(x_2 - x_0) \det \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & x_1 \\ 0 & 1 & x_2 \end{pmatrix} = (x_1 - x_0)(x_2 - x_0)(x_2 - x_1).$$

Det generelle bevis for $x_0, x_1, \dots, x_n$ følger successivt det centrale trin i ovenstående tilfælde gennem n søjleoperationer ud fra første række. Se opgave 5.7.9.

I tilfældet med forskellige x-værdier er determinanten af Vandermonde matricen $\neq 0$. Dermed er den invertibel ifølge Sætning 5.4 og ligningsystemet (5.7) har en og kun en løsning. Derfor er der præcis et polynomium, som går gennem punkterne.

Man kan selvfølgelig bare gå i krig og regne på løsninger til (5.7), men matematikkens styrke er beviset for den helt generelle sætning. Matematikken viser at der er noget unikt at lede efter og regne på.

5.6 Udregning af egenværdierne for en matrix

(5.7) SÆTNING.

Lad A være en $n \times n$ matrix. Tallet λ er en egenværdi for A hvis og kun hvis

$$\det(A - \lambda I_n) = 0.$$

Bevis

Bevis. At λ er en egenværdi for A betyder per definition at der findes en vektor $v \neq 0$ så $Av = \lambda v$. Dette er det samme som at

$$Av - \lambda v = Av - \lambda I_n v = (A - \lambda I_n)v = 0$$

for $v \neq 0$. Dette er ensbetydende med at $A - \lambda I_n$ ikke er invertibel eller

$$\det(A - \lambda I_n) = 0$$

via Sætning 5.4. □

♠

(5.8) DEFINITION.

Lad A være en $n \times n$ matrix. Så er

$$\det(A - \lambda I_n) = 0$$

et polynomium af grad n i λ. Dette polynomium kaldes for det karakteristiske polynomium *for A.*

Sætning 5.7 kan benyttes til konkret at udregne egenværdierne for en kvadratisk matrix A ved at finde rødderne i det karakteristiske polynomium for A. Lad os tage tråden op fra afsnittet, hvor vi kiggede konkret på egenværdier for 2×2 matricer. Tag som eksempel

$$A = \begin{pmatrix} -7 & 6 \\ -12 & 10 \end{pmatrix}.$$

Det karakteristiske polynomium for A er

$$\det(A - \lambda I_2) = \det \begin{pmatrix} -7-\lambda & 6 \\ -12 & 10-\lambda \end{pmatrix} = (-7-\lambda)(10-\lambda) + 72 = \lambda^2 - 3\lambda + 2.$$

Dette er et andengradspolynomium i λ med diskriminant $9 - 8 = 1$ og rødderne og dermed egenværdierne $\lambda = 1$ og $\lambda = 2$.

For 3×3 matricer bliver det karakteristiske polynomium et tredjegradspolynomium. Metoden til at finde rødder for tredjegradspolynomier er ikke lige til at huske på. Oftest kender man en rod på forhånd og kan regne sig frem til de to andre.

VIDEO: https://youtu.be/IFhxuCp-OeQ

For et polynomium af grad ≥ 5 kan man bevise at der faktisk ikke findes en løsningsformel! I praksis er man tilfreds med approksimationer til egenværdierne.

Dog har vi følgende vigtige resultat.

(5.9) SÆTNING.

En kvadratisk matrix har altid en kompleks egenværdi.

Bevis

Bevis. Det følger af algebraens fundamentalsætning. □

♠

5.10 Quiz

Hvilke af nedenstående udsagn er korrekte?

Matricen

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

har ingen reelle egenværdier.

Matricen $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$ har $-\lambda^3 + 15\lambda^2 + 16\lambda$ som karakteristisk polynomium.

0 er en egenværdi for
$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}.$$

$\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ er en egenvektor for
$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}.$$

$\begin{pmatrix} i \\ -i \end{pmatrix}$ er en egenvektor for
$$\begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix}.$$

$\begin{pmatrix} i \\ 1 \end{pmatrix}$ er en egenvektor for
$$\begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix}.$$

♠

5.7 Opgaver

5.7.1

Udregn determinanterne af
$$\begin{pmatrix} 2 \end{pmatrix}, \quad \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \quad \begin{pmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{pmatrix} \quad \text{og} \quad \begin{pmatrix} 2 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ 1 & 1 & 2 & 1 \\ 1 & 1 & 1 & 2 \end{pmatrix}$$
ved at reducere matricerne til RREF. Skitser dine trin i udregningerne. Summen af de sidste to svar skulle gerne give 25.

5.7.2

Hvorfor medfører afsnit 5.2.2 at determinanten til en matrix med en nulrække er 0?

5.7.3

Hvad er determinanten af
$$\begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}?$$
Kan du generalisere til $n \times n$ matricer?

5.7.4

Lad $A_n = (a_{ij})$ betegne $n \times n$ matricen med $a_{ij} = (i-1)n + j$. Feks er

$$A_4 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{pmatrix}.$$

Gør rede for at $\det(A_n) = 0$ for $n > 2$.

5.7.5

Vi ved at AB ikke nødvendigvis er lig med BA for to $n \times n$ matricer A og B. Men hvorfor er
$$\det(AB) = \det(BA)?$$

5.7.6

Udregn egenværdierne for matricen

$$A = \begin{pmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \\ 6 & 7 & 8 \end{pmatrix}.$$

Summen af dem skulle gerne give 12. Kan du forresten gennemskue i denne her sammenhæng hvordan tallet 12 er relateret til matricen A?

5.7.7

Lad A være en $n \times n$ matrix.

1. Gør rede for at A og A^T har det samme karakteristiske polynomium.

   5.11 Vink

   Vis og benyt at
   $$A^T - \lambda I_n = (A - \lambda I_n)^T.$$

   ♠

2. Lad P være en stokastisk $n \times n$ matrix, det vil sige at alle indgange er ≥ 0 og alle søjlesummer er 1. Vis at der findes en vektor $\mathbf{v} \neq 0$ så $P\mathbf{v} = \mathbf{v}$.

5.7.8

Lad P være en invertibel $n \times n$ matrix og A en $n \times n$ matrix.

1. Gør rede for at
   $$\det(P^{-1}AP) = \det(A).$$

2. Vis at A og $P^{-1}AP$ har det samme karakteristiske polynomium.

   5.12 Vink

   $$P^{-1}AP - \lambda I_n = P^{-1}(A - \lambda I_n)P.$$

   ♠

5.7.9

Gør rede for alle trin i udregningen af determinanten

$$\det \begin{pmatrix} 1 & x_0 & x_0^2 & x_0^3 \\ 1 & x_1 & x_1^2 & x_1^3 \\ 1 & x_2 & x_2^2 & x_2^3 \\ 1 & x_3 & x_3^2 & x_3^3 \end{pmatrix}$$

for at nå frem til formlen

$$(x_1 - x_0)(x_2 - x_0)(x_2 - x_1)(x_3 - x_0)(x_3 - x_1)(x_3 - x_2).$$

5.7.10

Lad

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

Gør rede for at A er rod i sit eget karakteristiske polynomium dvs

$$A^2 - (a_{11} + a_{22})A + \det(A)I_2 = 0$$

Kapitel 6

Konkrete vektorer

VIDEO: https://youtu.be/C0YpAzNmU44

6.1 Konkrete vektorrum

Vi vil indføre (konkrete) vektorrum over F, hvor F enten er de reelle tal $\mathbb{R}$ eller de komplekse $\mathbb{C}$. Sjovt nok afhænger rammen eller beviserne ikke af om F er netop $\mathbb{R}$ eller $\mathbb{C}$ og det er da også en af grundene til at indføre vektorrum som abstrakt begreb senere. Faktisk kan F helt generelt være det man i matematikken kalder et legeme (engelsk: field) - et algebraisk system, hvor det er muligt at dividere med alle elementer $\neq 0$.

6.1.1 De reelle tal $\mathbb{R}$

En helt naturlig generalisering af vektorer i planen $\mathbb{R}^2$ er søjlevektorer af længde n med indgange i $\mathbb{R}$:

$$\mathbb{R}^n = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \middle| \, x_1, \dots, x_n \in \mathbb{R} \right\}.$$

To vektorer

$$\mathbf{u} = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} \qquad \text{og} \qquad \mathbf{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

kan adderes plads for plads som

$$\mathbf{u} + \mathbf{v} = \begin{pmatrix} u_1 + v_1 \\ \vdots \\ u_n + v_n \end{pmatrix} \tag{6.1}$$

og vektoren **u** kan ganges med et tal $\lambda \in \mathbb{R}$ som

$$\lambda \mathbf{u} = \begin{pmatrix} \lambda u_1 \\ \vdots \\ \lambda u_n \end{pmatrix}. \tag{6.2}$$

Vi vil ofte bruge følgende notation: Hvis

$$\mathbf{u} = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$$

er en søjlevektor, så er $(\mathbf{u})_i = u_i$ den indgang i **u** som står i række nummer i.

6.1.2 Geometri, linear algebra og vektorer i rummet.

Det er ofte en god hjælp for forståelsen at forestille sig vektorer i rummet. Vi plejer jo at visualisere $\mathbb{R}^3$ som det rum vi allesammen lever i. Metoden har visse ulemper som vi vil diskutere løbende alt efter at vi opdager dem.

Den første ulempe er at vi har brug for et punkt. I $\mathbb{R}^3$ er der en vektor der er speciel, nemlig nulvektoren $\mathbf{0} = (0,0,0)^T$, fordi den er den eneste vektor der opfylder at $\mathbf{v} + \mathbf{0} = \mathbf{v}$ for alle vektorer **v**. Men i det rum vi lever i findes der ikke et punkt som er verdens centrum. Marcels kat ville sige at det punkt findes, og at det er den selv, men nabokatten Alfred ville ikke være enig deri. Så det første vi skal gøre er at vælge et punkt i rummet som vi kalder for **0**, nulvektoren eller ”origo”. Det er ligegyldigt hvad for et punkt vi vælger, men når vi først har valgt et punkt, så er vi nødt til at holde fast i det valg.

Det næste valg vi gør er at vælge et koordinatsystem med origo i det punkt vi har valgt. Når vi har gjort alt dette, kan vi sige at en søjlevektor $(x,y,z)^T$ svarer til nøjagtig et punkt i rummet. Men husk det nu og glem det ikke - det punkt vil også afhænge af valget af origo og af valget af koordinatsystem.

Noget man skal bide mærke i er at skalering med et tal og addition af to vektorer ikke behøver referere til valget af koordinatsystem, det kan beskrives helt geometrisk!

En anden måde at udtrykke det på, er at sige at hvis vi udstyrer vores fysiske rum med et origo (for eksempel en kat), så danner det et *abstrakt vektorrum*.

Aksiomer for et abstrakt vektorrum

At V er et abstrakt vektorrum over et legeme F betyder:

1. Givet $\mathbf{v}, \mathbf{u} \in V$ og $\lambda \in F$ kan vi definere $\mathbf{u} + \mathbf{v}$ og $\lambda \mathbf{v}$.

2. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$.

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

4. Der findes en vektor $\mathbf{0} \in V$ så at $\mathbf{u}+\mathbf{0} = \mathbf{u}$ for alle $\mathbf{u} \in V$.

5. For $\lambda \in F$ og $\mathbf{uv} \in V$ er $\lambda(\mathbf{u}+\mathbf{v}) = \lambda\mathbf{u}+\lambda\mathbf{v}$.

6. For $\lambda, \mu \in F$ og $\mathbf{u} \in V$ er $(\lambda+\mu)\mathbf{u} = \lambda\mathbf{u}+\mu\mathbf{u}$.

7. For $\lambda, \mu \in F$ og $\mathbf{u} \in V$ er $\lambda(\mu\mathbf{u}) = (\lambda\mu)\mathbf{u}$.

8. For $\mathbf{v} \in V$ er $1 \cdot \mathbf{v} = \mathbf{v}$.

♠

Abstrakte vektorrum forekommer meget ofte både i matematikken og i anvendelser. Vi vil ikke diskutere dem videre i disse noter, men de underrum af F^n som vi kommer at beskæftige os en del med er fortræffelige eksempler på abstrakte vektorrum. Hvis man forstår dem, så er man også meget tæt på at forstå abstrakte vektorrum generelt.

Vi anbefaler i øvrigt meget stærkt at tage et kig på Essence of linear algebra af 3Blue1Brown.

6.1.3 De komplekse tal $\mathbb{C}$

På analog vis definerer vi for de komplekse tal

$$\mathbb{C}^n = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \middle| \; x_1, \ldots, x_n \in \mathbb{C} \right\},$$

hvor addition af vektorer defineres som i (6.1) og talmultiplikation som i (6.2), men nu med komplekse tal i stedet for reelle tal.

<u>**Hvordan ser det ud i rummet?**</u>

Marcel beklager det dybt, men der er ikke en god måde at visualisere $\mathbb{C}^n$ for $n \geq 2$. For $n = 1$ har vi allerede diskuteret en geometrisk beskrivelse af $\mathbb{C}^1 = \mathbb{C}$ til bevidstløshed i kapitlet "De komplekse tal". ♠

6.2 Underrum

Lad nu F betegne enten $\mathbb{R}$ eller $\mathbb{C}$. Bemærk at F^n ligesom $\mathbb{R}^3$ indeholder *nulvektoren* $\mathbf{0}$, som består af 0 på alle indgangene og at

$$\mathbf{0}+\mathbf{v} = \mathbf{v}$$

for alle $\mathbf{v} \in F^n$.

(6.1) DEFINITION.

Lad $n > 0$ være et naturligt tal. Et underrum *af F^n er en ikke tom delmængde*

$$V \subseteq F^n,$$

som opfylder at

1. $\mathbf{u}+\mathbf{v} \in V$*, hvis* $\mathbf{u}, \mathbf{v} \in V$.
2. $\lambda \mathbf{u} \in V$*, hvis* $\lambda \in F$ *og* $\mathbf{u} \in V$

Hvordan ser det ud i rummet?

Et underrum af $\mathbb{R}^3$ kan være af fire forskellige slags.

1. Det kan bestå af kun origo.
2. Det kan være en linje der indeholder origo.
3. Det kan være en plan der indeholder origo.
4. Det kan være hele $\mathbb{R}^3$ (som selvfølgelig automatisk indeholder origo).

♠

6.2 Quiz

Hvilke af nedenstående udsagn er rigtige?

Punkterne på linjen i $\mathbb{R}^2$ givet ved $y = 2x + 1$ er et underrum af $\mathbb{R}^2$.

Et underrum af F^n indeholder altid $\mathbf{0}$.

Hvis $\lambda, \mu \in F$ og $\mathbf{u}, \mathbf{v} \in V$, hvor V er et underrum af F^n så vil

$$\lambda \mathbf{u} + \mu \mathbf{v} \in V.$$

Punkterne på parablen $y = x^2$ i $\mathbb{R}^2$ er et underrum af $\mathbb{R}^2$. ♠

6.3 Quiz

Betragt delmængden $W = \{(x,y,z) \mid x+y-z=a\}$ af $\mathbb{R}^3$. For hvilke a er W et underrum af $\mathbb{R}^3$?

$a = 1$.

$a = -1$.

$a = 0$. ♠

6.2.1 Linearkombinationer og span af vektorer

Lad $\mathbf{v}_1, \ldots, \mathbf{v}_m \in F^n$. En vektor på formen

$$\lambda_1 \mathbf{v}_1 + \cdots + \lambda_m \mathbf{v}_m$$

med $\lambda_1, \ldots, \lambda_m \in F$ kaldes en *linearkombination* af $\mathbf{v}_1, \ldots, \mathbf{v}_m$.

Mængden

$$\operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_m) := \{\lambda_1 \mathbf{v}_1 + \cdots + \lambda_m \mathbf{v}_m \mid \lambda_1, \ldots, \lambda_m \in F\}$$

af alle mulige linearkombinationer af $\mathbf{v}_1, \ldots, \mathbf{v}_m$ er et underrum af F^n og kaldes for *span* af vektorerne $\mathbf{v}_1, \ldots, \mathbf{v}_m$.

Hvordan ser det ud i rummet?

Vi ved i hvert fald at span af nogle vektorer i rummet er et underrum, så det må være en af de fire typer af underrum: punkt, linje, plan eller hele rummet. Hvis vi har en mængde af vektorer i $\mathbb{R}^3$, hvad for type underrum vil vi de udspænde? Ja, det er et godt spørgmål, tak for at I stillede det. Vi skal lige tænke os om, og når vi har tænkt os om så vender vi tilbage til det. ♠

6.4 Quiz

Betragt ligningen

$$\lambda_1 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} + \lambda_3 \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} + \lambda_4 \begin{pmatrix} 4 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

hvor $\lambda_1, \lambda_2, \lambda_3, \lambda_4 \in \mathbb{R}$. Denne ligning er opfyldt for?

$$\lambda_1 = -6, \quad \lambda_2 = -2, \quad \lambda_3 = 3$$
$$\lambda_4 = 5.$$

$$2\lambda_1 = -5\lambda_4, \quad \lambda_2 = 3\lambda_4$$
$$2\lambda_3 = -3\lambda_4.$$

$$\lambda_1 = -2, \quad \lambda_2 = -1, \quad \lambda_3 = 3$$
$$\lambda_4 = 7.$$

$$\lambda_1 = 5, \quad \lambda_2 = -6, \quad \lambda_3 = 3$$
$$\lambda_4 = -2.$$

♠

Der er her tale om en relativt abstrakt definition og det er en rigtig god ide at forbinde den til de konkrete forhold i den følgende opgave.

6.5 Opgave

Lad

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \text{og} \quad \mathbf{v}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

være vektorer i $\mathbb{R}^3$. Forklar hvorfor

$$\text{span}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) = \mathbb{R}^3$$

det vil sige hvorfor alle vektorer i $\mathbb{R}^3$ er linearkombinationer af $\mathbf{v}_1, \mathbf{v}_2$ og $\mathbf{v}_3$.

Lad nu

$$\mathbf{v}_1 = \begin{pmatrix} 3 \\ 2 \end{pmatrix} \quad \text{og} \quad \mathbf{v}_2 = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

være vektorer i $\mathbb{R}^2$. Hvordan afgør man om $\text{span}(\mathbf{v}_1, \mathbf{v}_2) = \mathbb{R}^2$ i dette tilfælde? ♠

6.6 Opgave

Lad $\mathbf{v}_1, \dots, \mathbf{v}_m \in V$ være vektorer i et underrum V af F^n. Forklar hvorfor

$$\operatorname{span}(\mathbf{v}_1, \dots, \mathbf{v}_m) \subseteq V.$$

♠

En helt fundamental observation er at span ikke ændrer sig ved operationer svarende til rækkeoperationerne for en matrix. Læg også mærke til (4) i Proposition 6.7, som forklarer at span blot er en ramme for løsning af ligninger.

(6.7) PROPOSITION.

Lad $\mathbf{v}_1, \dots, \mathbf{v}_m \in F^n$ *og*

$$V = \operatorname{span}(\mathbf{v}_1, \dots, \mathbf{v}_i, \dots, \mathbf{v}_j, \dots, \mathbf{v}_m),$$

hvor $1 \le i < j \le m$. *Så gælder*

1. *(Ombytning af vektorer)*

$$V = \operatorname{span}(\mathbf{v}_1, \dots, \mathbf{v}_j, \dots, \mathbf{v}_i, \dots, \mathbf{v}_m).$$

2. *(Multiplikation af en vektor med et tal* $\neq 0$*)*

$$V = \operatorname{span}(\mathbf{v}_1, \dots, \lambda \mathbf{v}_i, \dots, \mathbf{v}_m)$$

   for $\lambda \neq 0$.

3. *(Addition af et multiplum af en vektor til en anden vektor)*

$$V = \operatorname{span}(\mathbf{v}_1, \dots, \mathbf{v}_i, \dots, \mathbf{v}_j + \lambda \mathbf{v}_i, \dots, \mathbf{v}_m)$$

   for alle λ.

4. *Lad* A *være matricen med søjler* $A^1 = \mathbf{v}_1, \dots, A^m = \mathbf{v}_m$. *Vektoren* $\mathbf{b} \in F^n$ *ligger i* V *hvis og kun hvis ligningssystemet*

$$A\mathbf{x} = \mathbf{b}$$

   har en løsning $\mathbf{x} \in F^m$.

Bevis

Bevis. Vi beviser kun (3) og (4) her. Lad

$$V' = \operatorname{span}(\mathbf{v}_1, \dots, \mathbf{v}_i, \dots, \mathbf{v}_j + \lambda \mathbf{v}_i, \dots, \mathbf{v}_m).$$

Vi skal bevise at $V = V'$. Hvis en vektor $\mathbf{v} \in V'$ så er

$$\mathbf{v} = \lambda_1 \mathbf{v}_1 + \cdots + \lambda_i \mathbf{v}_i + \cdots + \lambda_j(\mathbf{v}_j + \lambda \mathbf{v}_i) + \cdots + \lambda_m \mathbf{v}_m$$

for passende $\lambda_1, \dots, \lambda_m \in F$. En omskrivning giver

$$\mathbf{v} = \lambda_1 \mathbf{v}_1 + \cdots + (\lambda_i + \lambda_j \lambda)\mathbf{v}_i + \cdots + \lambda_j \mathbf{v}_j + + \cdots + \lambda_m \mathbf{v}_m,$$

hvilket viser at $\mathbf{v} \in V$. Modsat hvis nu $\mathbf{v} \in V$, så er

$$\mathbf{v} = \lambda_1 \mathbf{v}_1 + \cdots + \lambda_i \mathbf{v}_i + \cdots + \lambda_j \mathbf{v}_j + \cdots + \lambda_m \mathbf{v}_m$$

for passende $\lambda_1, \dots, \lambda_m \in F$. Her giver en lidt mere indviklet omskrivning at

$$\mathbf{v} = \lambda_1 \mathbf{v}_1 + \cdots + (\lambda_i - \lambda_j \lambda)\mathbf{v}_i + \cdots + \lambda_j(\mathbf{v}_j + \lambda \mathbf{v}_i) + \cdots + \lambda_m \mathbf{v}_m,$$

hvilket viser at $\mathbf{v} \in V'$. Derfor er $V = V'$.

Den sidste påstand (4) følger naturligt af definitionen af matrixmultiplikation, idet

$$A\mathbf{x} = \lambda_1 A^1 + \cdots + \lambda_m A^m = \lambda_1 \mathbf{v}_1 + \cdots + \lambda_m \mathbf{v}_m,$$

hvor

$$\mathbf{x} = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_m \end{pmatrix}.$$

□

♠

6.8 Opgave

Giv et alternativt bevis for 6.7(3) på følgende måde. Vis først at enhver af vektorerne $\mathbf{v}_1, \dots, \mathbf{v}_j + \lambda \mathbf{v}_i, \dots$ er indeholdt i $\operatorname{span}(\mathbf{v}_1, \dots, \mathbf{v}_m)$. Brug opgave 6.6 for at konkludere at $\operatorname{span}(\mathbf{v}_1, \dots, \mathbf{v}_j + \lambda \mathbf{v}_i, \dots, \mathbf{v}_m) \subseteq \operatorname{span}(\mathbf{v}_1, \dots, \mathbf{v}_m)$ Vis omvendt at enhver af vektorerne $\mathbf{v}_1, \dots, \mathbf{v}_m$ er indholdt i $\operatorname{span}(\mathbf{v}_1, \dots, \mathbf{v}_j + \lambda \mathbf{v}_i, \dots, \mathbf{v}_m)$. Overvej at det nu følger at

$$\operatorname{span}(\mathbf{v}_1, \dots, \mathbf{v}_i, \dots, \mathbf{v}_j + \lambda \mathbf{v}_i, \dots, \mathbf{v}_m) = \operatorname{span}(\mathbf{v}_1, \dots, \mathbf{v}_i, \dots, \mathbf{v}_j, \dots, \mathbf{v}_m)$$

For en ekstra stjerne, overvej at man med densamme metode også kan bevise 6.7(1) og 6.7(2).

♠

6.9 Eksempel

Betragt nu vektorerne

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \qquad \text{og} \qquad \mathbf{v}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

i $\mathbb{R}^3$. Lad os undersøge om

$$\mathbf{b} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \in \operatorname{span}(\mathbf{v}_1, \mathbf{v}_2).$$

Her bruger vi Proposition 6.7(4) og opstiller ligningssystemet $A\mathbf{x} = \mathbf{b}$, hvor A er matricen med søjler $\mathbf{v}_1$ og $\mathbf{v}_2$. Da

$$\begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 2 \\ 0 & 1 & 3 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix},$$

hvor sidste række indikerer $0 = 2$ (hvad sker der her? hvorfor gør den det?), ses at ligningssystemet ikke har en løsning og dermed at $\mathbf{b} \notin \operatorname{span}(\mathbf{v}_1, \mathbf{v}_2)$. ♠

6.2.2 Nulrum, søjlerum og rækkerum for matricer

Tre helt fundamentale eksempler på underrum er knyttet til en matrix.

(6.10) PROPOSITION.

Lad A være en $m \times n$ matrix A med indgange i F.

1.
$$N(A) = \{\mathbf{v} \in F^n \mid A\mathbf{v} = 0\}$$
er et underrum af F^n.

2. *$C(A)$*
$$C(A) = \{A\mathbf{v} \mid \mathbf{v} \in F^n\}$$
er et underrum af F^m.

3. *$R(A)$*
$$R(A) = \{A^T\mathbf{v} \mid \mathbf{v} \in F^m\}$$
er et underrum af F^n.

Bevis

Bevis. Det følger at $N(A)$ er et underrum af F^n, fordi

$$A(\mathbf{u}+\mathbf{v}) = A\mathbf{u}+A\mathbf{v} = 0 \qquad \text{og} \qquad A(\lambda\mathbf{u}) = \lambda A\mathbf{u} = 0,$$

hvis $\mathbf{u},\mathbf{v} \in N(A)$ og $\lambda \in F$.

På næsten samme måde vises at $C(A)$ er et underrum af F^m: Hvis $\mathbf{u}',\mathbf{v}' \in C(A)$ og $\lambda \in F$ skal vi vise at $\mathbf{u}'+\mathbf{v}' \in C(A)$ og $\lambda\mathbf{u}' \in C(A)$. Men $\mathbf{u}',\mathbf{v}' \in C(A)$ betyder at $\mathbf{u}' = A\mathbf{u}, \mathbf{v}' = A\mathbf{v}$ for passende $\mathbf{u},\mathbf{v} \in F^n$. Derfor gælder

$$\mathbf{u}'+\mathbf{v}' = A\mathbf{u}+A\mathbf{v} = A(\mathbf{u}+\mathbf{v}) \in C(A)$$

og

$$\lambda\mathbf{u}' = \lambda A\mathbf{u} = A(\lambda\mathbf{u}) \in C(A).$$

Ved at sammenligne definitionerne af rækkerum og søjlerum ser vi at $R(A) = C(A^T)$. Ved at bruge (2) på $n \times m$ matricen A^T indser vi at $C(A^T)$ er et underrum af F^n. Men da $R(A) = C(A^T)$ er $R(A)$ er et underrum af F^n. $\square$

♠

Underrummet $N(A)$ kaldes for *nulrummet* for A, $C(A)$ kaldes for *søjlerummet* for A og $R(A)$ *rækkerummet* for A.

Hvordan ser det ud i rummet?

Der er virkeligt gode geometriske beskrivelser af disse tre underrum, men det er nemmere at forstå dette efter at vi har forklaret hvad en lineær transformation er. Vi vender tilbage. ♠

Disse definitioner kan være svære at forstå uden et konkret eksempel.

6.11 Eksempel

Lad os se på matricen

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}.$$

Her er søjlerummet

$$C(A) = \left\{ \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \middle| x_1,x_2,x_3 \in F \right\} = \left\{ \begin{pmatrix} x_1+2x_2+3x_3 \\ 4x_1+5x_2+6x_3 \end{pmatrix} \middle| x_1,x_2,x_3 \in F \right\}$$
$$= \left\{ x_1 \begin{pmatrix} 1 \\ 4 \end{pmatrix} + x_2 \begin{pmatrix} 2 \\ 5 \end{pmatrix} + x_3 \begin{pmatrix} 3 \\ 6 \end{pmatrix} \middle| x_1,x_2,x_3 \in F \right\}.$$

På samme måde ses rækkerummet at være

$$R(A) = \left\{ \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \middle| y_1, y_2 \in F \right\} = \left\{ y_1 \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + y_2 \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} \middle| y_1, y_2 \in F \right\}.$$

Nulrummet $N(A)$ er løsningerne til det homogene ligningssystem $Ax = 0$ det vil sige

$$N(A) = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in F^3 \middle| \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\}$$

$$= \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in F^3 \middle| \begin{array}{rrrcr} x_1+ & 2x_2+ & 3x_3 & = & 0 \\ 4x_1+ & 5x_2+ & 6x_3 & = & 0 \end{array} \right\}.$$

I Eksempel 6.6.2.2 giver vi et eksempel på hvordan man udregner $N(A)$ via RREF. Ikke overraskende er der tale om at løse det homogene ligningssystem ved hjælp af af bundne og frie variable. ♠

> Eksempel 6.6.2.2 viser at $R(A)$ er span af rækkevektorerne (transponeret) og $C(A)$ span af søjlevektorerne for en matrix A.

En særdeles vigtig observation er at nulrummet og rækkerummet ikke ændres ved rækkeoperationer på matricen. Desuden er søjlerummet relateret til pivotrækkerne i matricens RREF.

(6.12) PROPOSITION.

Lad A være en $m \times n$ matrix og B dens RREF. Så er

$$N(A) = N(B) \qquad \text{og} \qquad R(A) = R(B).$$

Søjlerummet $C(A)$ er span af søjlerne i A svarende til pivotsøjlerne i B det vil sige de søjler eller søjlenumre i B, som indeholder pivotelementerne.

Bevis

Bevis. Der findes en invertibel matrix E så $B = EA$. Da E er invertibel ses at $A\mathbf{x} = \mathbf{0}$ holder hvis og kun hvis $B\mathbf{x} = (EA)\mathbf{x} = E(A\mathbf{x}) = \mathbf{0}$. Dette oversættes umiddelbart til at $N(A) = N(B)$.

På samme måde får vi at $A^T\mathbf{x} = (EA)^T\mathbf{y} = B^T\mathbf{y}$, hvor $\mathbf{y} = (E^T)^{-1}\mathbf{x}$ og dermed er $R(A) = C(A^T) = C(B^T) = R(B)$.

Det var nemt nok. Det er lidt mere indviklet at vise påstanden om søjlerummet $C(A)$. Da B er på RREF så er pivotsøjlerne i B er simpelthen vores standard basisvektorer $\mathbf{e}_1, \mathbf{e}_2, \dots \mathbf{e}_m$ hvor

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

Span af disse pivotsøjler er altså alle vektorer $\mathbf{v}$ på formen

$$\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_{m-1} \\ v_m \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Søjlerummet af B er åbenbart span af pivotsøjlerne (hvorfor er dette egentlig så åbenbart?).

<u>Vink</u>

Hvis B er på RREF kunne den for eksempel se sådan ud:

$$\begin{pmatrix} 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0.01 \\ 0 & 0 & 0 & 0 & 1 & 3.2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

♠

Nu overvejer vi følgende. Antag at $\mathbf{v}_1, \dots \mathbf{v}_r$ er en mængde af vektorer, og at $\mathbf{w}_1, \dots \mathbf{w}_s$ er en anden mængde af vektorer. Hvis det nu er så heldigt at

$$\mathrm{span}(\mathbf{v}_1, \dots \mathbf{v}_r) = \mathrm{span}(\mathbf{w}_1, \dots \mathbf{w}_s)$$

og E er en invertibel matrix, så er også

$$\operatorname{span}(E^{-1}\mathbf{v}_1,\dots E^{-1}\mathbf{v}_r) = \operatorname{span}(E^{-1}\mathbf{w}_1,\dots E^{-1}\mathbf{w}_s).$$

Hvis nu $(\mathbf{v}_1,\dots \mathbf{v}_r)$ er pivotsøjlerne i B og $(\mathbf{w}_1,\dots \mathbf{w}_n)$ er *alle* søjler i B så er ifølge det vi lige har overvejet $\operatorname{span}(\mathbf{v}_1,\dots \mathbf{v}_r) = \operatorname{span}(\mathbf{w}_1,\dots \mathbf{w}_n)$ og dermed $\operatorname{span}(E^{-1}\mathbf{v}_1,\dots E^{-1}\mathbf{v}_r) = \operatorname{span}(E^{-1}\mathbf{w}_1,\dots E^{-1}\mathbf{w}_n)$. Men da $E^{-1}B = A$ er $E^{-1}\mathbf{w}_i$ jo netop den søjle i A der har søjlenummer i (overvej også det!), og vektorerne $E^{-1}\mathbf{v}_i$ er de søjler i A der har de samme søjlenummer som pivotsøjlerne i B. □

♠

Vi giver et eksempel på hvordan nulrummet $N(A)$, rækkerummet $R(A)$ og søjlerummet $C(A)$ udregnes for en matrix A.

6.13 Eksempel

Lad

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

Først rækkereducerer vi A til RREF:

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

I første $\sim$ adderes første række til anden række, trækkes fra tredie række og adderes til fjerde række. I anden $\sim$ trækkes anden række fra fjerde række. I tredje $\sim$ adderes anden række til første række.

Sidste matrix er på RREF. Lad os kalde den B. Nu ved vi fra Proposition 6.12 at $N(A) = N(B)$ og $R(A) = R(B)$. De bundne variable er x_1, x_2. De frie er x_3, x_4. Det vil sige et typisk element i $N(A)$ har formen

$$\begin{pmatrix} -3x_3 + 3x_4 \\ -2x_3 + x_4 \\ x_3 \\ x_4 \end{pmatrix} = x_3 \begin{pmatrix} -3 \\ -2 \\ 1 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} 3 \\ 1 \\ 0 \\ 1 \end{pmatrix}.$$

Heraf fremgår det at

$$N(A) = \operatorname{span}\left(\begin{pmatrix} -3 \\ -2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \\ 0 \\ 1 \end{pmatrix} \right)$$

samt

$$R(A) = \operatorname{span}\left(\begin{pmatrix} 1 \\ 0 \\ 3 \\ -3 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 2 \\ -1 \end{pmatrix}\right).$$

Læg mærke til at vi gik fra at have rækkerummet som span af 4 vektorer (de 4 rækker (transponeret) i A) til et span af kun 2 vektorer.

Svarende til pivotsøjlerne i B bliver

$$C(A) = \operatorname{span}\left(\begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \\ -1 \\ 2 \end{pmatrix}\right).$$

svarende til de to første søjler i A igen i følgende Proposition 6.12. ♠

6.3 Lineær uafhængighed

De to vektorer

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{og} \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

er specielle i og med at ingen af dem kan udelades fra

$$V = \operatorname{span}\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right)$$

uden at V bliver mindre eller ændres (fra $\mathbb{R}^2$ til x-aksen eller y-aksen). Det er helt anderledes med for eksempel vektorerne

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{og} \quad \begin{pmatrix} 2 \\ 2 \end{pmatrix}.$$

Her ændres

$$V = \operatorname{span}\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \end{pmatrix}\right)$$

ikke hvis en af dem udelades.

(6.14) DEFINITION.

Vektorerne $\mathbf{v}_1, \dots, \mathbf{v}_m \in F^n$ *kaldes* lineært uafhængige *hvis de er minimale i den forstand at ingen af dem kan udelades fra*

$$V = \operatorname{span}(\mathbf{v}_1, \dots, \mathbf{v}_m)$$

uden at V ændres. Vektorerne kaldes for lineært afhængige*, hvis de ikke er lineært uafhængige.*

Definition 6.14 er ækvivalent med at

$$\mathbf{v}_i \notin \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \ldots, \mathbf{v}_m)$$

for $i = 1, \ldots, m$. Faktisk gælder følgende.

(6.15) PROPOSITION.

1. *Vektorerne $\mathbf{v}_1, \ldots, \mathbf{v}_m \in F^n$ er lineært uafhængige hvis og kun hvis man af $\lambda_1 \mathbf{v}_1 + \cdots + \lambda_m \mathbf{v}_m = \mathbf{0}$ for $\lambda_1, \ldots, \lambda_m \in F$ kan slutte at*

$$\lambda_1 = \cdots = \lambda_m = 0.$$

2. *Lad $\mathbf{v}_1, \ldots, \mathbf{v}_n \in F^m$ og lad A betegne $m \times n$ matricen med søjler*

$$A^1 = \mathbf{v}_1, \ldots, A^n = \mathbf{v}_n.$$

Så er $\mathbf{v}_1, \ldots, \mathbf{v}_n \in F^m$ lineært uafhængige hvis og kun hvis ligningssystemet

$$A\mathbf{x} = \mathbf{0}$$

kun har løsningen $\mathbf{x} = \mathbf{0}$.

Bevis

Bevis. Del 1:

Antag først at $\mathbf{v}_1, \ldots, \mathbf{v}_n \in F^m$ lineært uafhængige. Hvis $\lambda_1 \mathbf{v}_1 + \cdots + \lambda_m \mathbf{v}_m = \mathbf{0}$ skal vi vise at $\lambda_i = 0$ for alle i. Vi argumenterer ved modstrid. Antag er der skulle findes et i så at $\lambda_i \neq 0$. Da kan vi dividere med λ_i (Men I må aldrig dividere med 0!!!), så at

$$\mathbf{v}_i = -(\lambda_1/\lambda_i)\mathbf{v}_1 - \cdots - (\lambda_{i-1}/\lambda_i)\mathbf{v}_{i-1} - (\lambda_{i+1}/\lambda_i)\mathbf{v}_{i+1} \cdots$$

Dermed er jo $\mathbf{v}_i \in \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \ldots, \mathbf{v}_m)$, hvad der ikke måtte ske fordi vi har antaget at $\mathbf{v}_1, \ldots, \mathbf{v}_m$ et lineært uafhængige. Vi har opnået en modstrid, det vil sige, vi ved nu at $\lambda_i = 0$ for alle i.

Del 2:

Antag nu omvendt at $\mathbf{v}_1, \ldots, \mathbf{v}_m$ ikke er lineært uafhængige. Vi skal finde $\lambda_i \in F$, ikke alle lige med 0, så at $\lambda_1 \mathbf{v}_1 + \cdots + \lambda_m \mathbf{v}_m = \mathbf{0}$. Men da vektorerne ikke er linært uafhængige så findes der et j så at $\mathbf{v}_j \in \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_{j-1}, \mathbf{v}_{j+1}, \ldots, \mathbf{v}_m)$. Det vil sige, der findes $\lambda_1, \ldots, \lambda_{j-1}, \lambda_{j+1}, \ldots, \lambda_m$ så at

$$\mathbf{v}_j = \lambda_1 \mathbf{v}_1 + \cdots \lambda_{j-1} \mathbf{v}_{j-1} + \lambda_{j+1} \mathbf{v}_{j+1}, \cdots, \lambda_m \mathbf{v}_m.$$

For så vidt kunne $\lambda_i = 0$ for alle $i \neq j$. Men hvis vi definerer $\lambda_j = -1$ så er i hvert fald $\lambda_j \neq 0$, og på den anden side er nu

$$\lambda_1 \mathbf{v}_1 + \cdots + \lambda_m \mathbf{v}_m = \mathbf{0}.$$

Del 3:
Sidste påstand følger helt naturligt af definitionen af matrixmultiplikation, idet

$$A\mathbf{x} = \lambda_1 A^1 + \cdots + \lambda_m A^m = \lambda_1 \mathbf{v}_1 + \cdots + \lambda_m \mathbf{v}_m,$$

hvor

$$\mathbf{x} = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_m \end{pmatrix}.$$

□

♠

Det er på høje tid med en quiz.

6.16 Quiz

Hvilke af nedenstående påstande er rigtige?

Vektorerne $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ og $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ i $\mathbb{R}^2$ er lineært uafhængige.

Vektorerne $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ og $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ i $\mathbb{R}^2$ er lineært uafhængige.

Vektorerne $\begin{pmatrix} 2 \\ 3 \end{pmatrix}$ og $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ i $\mathbb{R}^2$ er lineært uafhængige.

Vektorerne $\begin{pmatrix} 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ og $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ i $\mathbb{R}^2$ er lineært uafhængige.

Vektoren $\begin{pmatrix} 2 \\ 3 \end{pmatrix}$ i $\mathbb{R}^2$ ligger i $\operatorname{span}\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right)$.

Vektoren $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ i $\mathbb{C}^2$ ligger i span $\left(\begin{pmatrix} -i \\ i \end{pmatrix}\right)$. ♠

6.17 Eksempel

Betragt vektorerne

$$\begin{pmatrix} 2 \\ 3 \end{pmatrix}, \qquad \begin{pmatrix} 1 \\ 2 \end{pmatrix} \qquad \text{og} \qquad \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

i $\mathbb{R}^2$. For at afgøre om de er lineært uafhængige skal vi ifølge Proposition 6.15(2) undersøge ligningssystemet

$$\begin{pmatrix} 2 & 1 & 1 \\ 3 & 2 & 1 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Vi kan ret hurtigt se at RREF for koefficientmatricen er

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

Derfor er λ_3 en fri variabel og med $\lambda_3 = 1$ bliver $\lambda_1 = -1$ og $\lambda_2 = 1$ i fin overensstemmelse med at

$$(-1) \cdot \begin{pmatrix} 2 \\ 3 \end{pmatrix} + 1 \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} + 1 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Derfor er vektorerne ikke lineært uafhængige. ♠

Ved et nærmere kig på Proposition 6.15(2) antydes et helt centralt resultat i lineær algebra, som går tilbage til et helt centralt resultat om løsning af ligninger.

Der er nemlig en meget naturlig øvre grænse på hvor mange lineært uafhængige vektorer man kan have i F^n.

(6.18) SÆTNING.

Lad $\mathbf{v}_1, \ldots, \mathbf{v}_m \in F^n$. Hvis $m > n$, så er $\mathbf{v}_1, \ldots, \mathbf{v}_m$ lineært afhængige.

Bevis

Bevis. Vi skriver søjlevektoren $\mathbf{v}_i$ som

$$\mathbf{v}_i = \begin{pmatrix} v_{i1} \\ v_{i2} \\ \vdots \\ v_{in} \end{pmatrix}.$$

Vi skal vise at vi kan finde tal λ_i, ikke alle nul, så at

$$\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \cdots + \lambda_m \mathbf{v}_m = \mathbf{0}.$$

Vi skriver dette helt ud i en formel: vektorerne er lineært afhængige hvis og kun hvis vi kan finde tal λ_m som ikke alle er 0, så at

$$\lambda_1 \begin{pmatrix} v_{11} \\ v_{12} \\ \vdots \\ v_{1n} \end{pmatrix} + \lambda_2 \begin{pmatrix} v_{21} \\ v_{22} \\ \vdots \\ v_{2n} \end{pmatrix} + \cdots + \lambda_m \begin{pmatrix} v_{m1} \\ v_{m2} \\ \vdots \\ v_{mn} \end{pmatrix} = \mathbf{0}.$$

Det kan vi også skrive som

$$\begin{aligned} \lambda_1 v_{11} + \lambda_2 v_{12} + \cdots + \lambda_m v_{1m} &= 0 \\ \lambda_1 v_{21} + \lambda_2 v_{22} + \cdots + \lambda_m v_{2m} &= 0 \\ &\cdots \\ \lambda_1 v_{n1} + \lambda_2 v_{n2} + \cdots + \lambda_m v_{nm} &= 0. \end{aligned}$$

Fordi vi har antaget at $m > n$, så har dette system altid en løsning forskellig fra nulløsningen. Derfor er vektorerne lineært afhængige. □

♠

Hvordan ser det ud i rummet?

En enkelt vektor $\mathbf{v}_1$ er lineært uafhængig hvis og kun hvis den er forskellig fra nulvektoren. Da er span$(\mathbf{v}_1)$ en linje gennem origo. To vektorer $\mathbf{v}_1, \mathbf{v}_2$ er lineært uafhængige hvis de ikke ligger i forlængelse af hindanden, det vil sige, hvis de ikke ligger på den samme linje gennem origo. I dette tilfælde er span$(\mathbf{v}_1, \mathbf{v}_2)$ en plan der indeholder origo. De tre vektorer $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ er lineært uafhængige hvis de to vektorer $\mathbf{v}_1, \mathbf{v}_2$ er lineært uafhængige, og samtidigt $\mathbf{v}_3$ ikke ligger i den plan der er udspændt af $\mathbf{v}_1$ og $\mathbf{v}_2$. Hvis $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ er lineær uafhængige, så er span$(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ hele det tredimensionelle rum. ♠

6.4 Basis for og dimension af underrum

(6.19) DEFINITION.

Lad V være et underrum af F^n. Et ordnet sæt $(\mathbf{v}_1, \ldots, \mathbf{v}_m)$ med $\mathbf{v}_1, \ldots, \mathbf{v}_m \in V$ kaldes en basis *for V, hvis vektorerne $\mathbf{v}_1, \ldots, \mathbf{v}_m$ er lineært uafhængige og $V =$* $\operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_m)$.

Det vil sige: en basis for et underrum er en minimal udspændende mængde, hvor ingen vektorer kan udelades. Det kan ikke understreges nok at man for at forstå hvad en basis er, bliver nødt til at se på adskillige konkrete eksempler. Som et absolut minimum bør du løse følgende opgave.

6.20 Opgave

Forklar helt præcist hvorfor

$$B = \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right)$$

er en basis for $\mathbb{R}^2$. ♠

Sætning 6.18 medfører følgende centrale resultat, som siger at ethvert underrum har en basis og at antallet af vektorer i en basis er entydigt bestemt. Bemærk dog at et underrum har tonsvis af baser. Vi siger ikke at basen er entydigt bestemt, kun antallet af vektorer i den.

Begrebet basis er helt basalt for lineær algebra, så brug tid på at forstå det. I de fleste anvendelser vil man få brug for at lave konkrete udregninger, og de vil ofte afhænge af at man har valgt en basis for et underrum. I mere teoretiske overvejelser er det mange gange bedre at ikke lægge sig fast på en bestemt basis, med mindre denne basis er specielt egnet for netop det problem man betragter.

En matematikers tre regler for valg af basis

Regel 1 : Vælg aldrig en basis.

Regel 2 :

Seriøst - vælg aldrig en basis.

Regel 3 :

Hvis du alligevel kommer til at vælge en basis, gør det på en ekstremt snedig måde. ♠

♠

♠

(6.21) SÆTNING.

Til ethvert underrum $V \subseteq F^r$ findes en basis med n elementer, hvor $n \leq r$. Hvis $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ og $(\mathbf{u}_1, \ldots, \mathbf{u}_m)$ er to baser for V, så er $m = n$.

Bevis*

Bevis. Vi kan lige så godt antage at V ikke kun består af nulvektoren. Vælg $\mathbf{v} \in V$ så $\mathbf{v} \neq \mathbf{0}$. Sæt $\mathbf{v}_1 = \mathbf{v}$. Hvis $V = \text{span}(\mathbf{v}_1)$ er vi færdige. Hvis ikke findes der et $\mathbf{v}_2 \in V$ så at $\mathbf{v}_2 \notin \text{span}(\mathbf{v}_1)$. Det betyder at $\mathbf{v}_1$ og $\mathbf{v}_2$ er lineært uafhængige. Denne proces fortsættes. Hvis vi har fundet $\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}$ som er lineært uafhængige, og hvis $\mathbf{v}_n \notin \text{span}(\mathbf{v}_1, \ldots \mathbf{v}_{n-1})$ så er $\mathbf{v}_1, \ldots, \mathbf{v}_n$ lineært uafhængige, fordi hvis

$$\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \cdots \lambda_n \mathbf{v}_n = \mathbf{0},$$

så er

$$\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \cdots \lambda_{n-1} \mathbf{v}_{n-1} = -\lambda_n \mathbf{v}_n.$$

Hvis nu $\lambda_n \neq 0$ kunne man bruge denne ligning til at skrive $\mathbf{v}_n$ som linearkombination af $\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}$, men det kan vi jo ikke, fordi $\mathbf{v}_n$ ikke ligger i span af de andre vektorer. Derfor må $\lambda_n = 0$, og altså

$$\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \cdots \lambda_{n-1} \mathbf{v}_{n-1} = \mathbf{0}.$$

Fordi $\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}$ er lineært uafhængige er nu $\lambda_i = 0$ for alle i. Vi kan dermed slå fast at $\mathbf{v}_1, \ldots, \mathbf{v}_n$ er lineært uafhængige. Et stykke tid kan vi blive ved med at finde nye $\mathbf{v}_i$, men denne leg må nødvendigvis stoppe med n lineært uafhængige vektorer $\mathbf{v}_1, \ldots, \mathbf{v}_n$ med $V = \text{span}(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ og $n \leq r$ på grund af Sætning 18.

Hvis $(\mathbf{u}_1, \ldots, \mathbf{u}_m)$ er en anden basis for V, skal vi vise at $m = n$. Det er nok at vise at $m \leq n$ (fordi på den samme måde kan vi også vise at $n \leq m$). Vi argumenterer ved modstrid. Det vil sige, vi antager at $m > n$, og viser at dette fører til en modstrid mod forudsætningen at $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ og $(\mathbf{u}_1, \ldots, \mathbf{u}_m)$ begge er baser for V.

Til at begynde med bruger vi at vektorerne $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ danner en basis, så at vi kan skrive

$$\begin{aligned}
\mathbf{u}_1 &= a_{11}\mathbf{v}_1 + a_{21}\mathbf{v}_2 + \cdots + a_{n1}\mathbf{v}_n \\
\mathbf{u}_2 &= a_{12}\mathbf{v}_1 + a_{22}\mathbf{v}_2 + \cdots + a_{n2}\mathbf{v}_n \\
&\vdots \\
\mathbf{u}_m &= a_{1m}\mathbf{v}_1 + a_{2m}\mathbf{v}_2 + \cdots + a_{nm}\mathbf{v}_n.
\end{aligned}$$

Nu opstiller vi et ligningssystem med m ubekendte og n ligninger.

$$\begin{aligned} a_{11}\lambda_1 + a_{12}\lambda_2 + \cdots + a_{1m}\lambda_m &= 0 \\ a_{21}\lambda_1 + a_{22}\lambda_2 + \cdots + a_{2m}\lambda_m &= 0 \\ &\vdots \\ a_{n1}\lambda_1 + a_{n2}\lambda_2 + \cdots + a_{nm}\lambda_m &= 0. \end{aligned}$$

Da $m > n$ findes der en løsning til dette ligningssystem som ikke er nulløsningen. Lad $\lambda_1, \ldots, \lambda_m$ være en sådan løsning. Vi definerer en vektor $\Lambda = \lambda_1 \mathbf{u}_1 + \cdots + \lambda_m \mathbf{u}_m$. Nu er tiden kommet til at bruge at $(\mathbf{u}_1, \ldots, \mathbf{u}_m)$ også er en basis, fordi det betyder jo at $\Lambda \neq \mathbf{0}$. På den anden side er

$$\begin{aligned} \Lambda = & (a_{11}\lambda_1 + a_{12}\lambda_2 + \cdots + a_{1m}\lambda_m)\,\mathbf{v}_1 \\ & + \cdots + \\ & (a_{n1}\lambda_1 + a_{n2}\lambda_2 + \cdots + a_{nm}\lambda_m)\,\mathbf{v}_n = \mathbf{0}. \end{aligned}$$

Dette giver en modstrid mod antagelsen at såvel $(\mathbf{u}_1, \ldots, \mathbf{u}_m)$ som $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ er en basis. Sætningen er bevist.

□

♠

Med entydigheden af antallet af elementer i en basis for et underrum har vi nu et veldefineret mål for størrelsen eller omfanget af et underrum kaldet dimensionen.

(6.22) DEFINITION.

Lad V være et underrum af F^n. Dimensionen af V er antal vektorer i en basis for V og betegnes $\dim(V)$.

Hvordan ser det ud i rummet?

Det skulle helst ikke chokere nogen der har fulgt med hertil, at dimensionen af origo er 0, dimensionen af en linje der indeholder origo er 1, dimensionen af et plan der indeholder origo er 2 og dimensionen af hele rummet er 3. ♠

Hvordan finder vi så dimensionen af et underrum? Igen kommer den reducerede række echelon form til hjælp.

(6.23) SÆTNING.

Lad

$$V = \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_m)$$

være et underrum af F^n. En basis for V fås ud fra rækkerne $\neq \mathbf{0}$ i den reducerede række echelon form B for A, hvor A er matricen med rækker

$$A_1 = \mathbf{v}_1^T, \ldots, A_m = \mathbf{v}_m^T.$$

Dimensionen, $\dim(V)$*, er lig med antal pivotelementer i B.*

Bevis

Ifølge definitionen af A og definitionen af rækkerummet $R(A)$ er $V = R(A)$. Da rækkerummet ikke ændres ved rækkeoperationer er $R(A) = R(B)$. Så vi skal vise, at hvis B er på RREF, så har $R(B)$ basis der består af alle rækker i B, som ikke er nulrækker. Vi kan nu glemme alt om matricen A som måske ikke var på RREF, og koncentrere os om B. Hvis vi lader alle nulrækker i B helt væk, så får vi en ny matrix C på RREF med det samme rækkerum. Forskellen er kun at C ikke har nulrækker, og åbenbart er $R(B) = R(C)$. For at give et fuldstændigt bevis for sætningen er det altså nok at vise at hvis C er en matrix på RREF, som ikke indeholder nulrækker, så danner rækkerne i C en basis for $R(C)$. Ifølge definitionen af hvad det vil sige at være en basis, skal vi vise, at rækkerne i C er lineært uafhængige.

Hver række $\mathbf{u}_i$ i C indeholder et pivotelement. Lad k_i være nummeret på den søjle, der indeholder pivotelementet i række nummer i. Da er $(\mathbf{u}_i)_{k_i} = 1$, fordi det er et pivotelement, og hvis $j \neq i$ er $(\mathbf{u}_j)_{k_i} = 0$, da pivotelementet er den eneste indgang i sin søjle der er forskelligt fra 0. Lad

$$\mathbf{w} = \lambda_1 \mathbf{u}_1 + \lambda_2 \mathbf{u}_2 + \cdots + \lambda_m \mathbf{u}_m.$$

Hvis vi skriver $\mathbf{w} = (w_1, \ldots, w_n)$, så er altså $w_{k_i} = \lambda_i$.

For at vise at $\mathbf{u}_1, \ldots, \mathbf{u}_m$ er lineært uafhængige, skal vi vise: hvis $\mathbf{w} = \mathbf{0}$, så er $\lambda_i = 0$ for alle i. Men hvis $\mathbf{w} = 0$, så er specielt $w_{k_i} = 0$, så at $\lambda_i = w_{k_i} = 0$. ♠

6.24 Eksempel

Vi gennefører udregningerne i beviset ovenfor på et eksempel. Lad

$$V = \operatorname{span}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) = \operatorname{span}\left(\begin{pmatrix}1\\2\\3\end{pmatrix}, \begin{pmatrix}4\\5\\6\end{pmatrix}, \begin{pmatrix}7\\8\\9\end{pmatrix}\right)$$

være et underrum i $\mathbb{R}^3$. Vi transponerer søjlevektorerne $\mathbf{v}_i$, og bruger rækkevektorerne $\mathbf{v}_i^T$ som rækker i en matrix A. Det vi sige, vi opstiller matricen

$$A = \begin{pmatrix}1 & 2 & 3\\4 & 5 & 6\\7 & 8 & 9\end{pmatrix},$$

som så rækkereduceres til B.

$$A \sim \begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & -6 & -12 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 7 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{pmatrix} = B.$$

Vi ved at $R(A) = R(B)$, så for at bestemme $R(A)$ er det nok at bestemme $R(B)$. $R(B)$ er udspændt af de tre vektorer $(1,0,7),(0,1,-2),(0,0,0)$. Den sidste nul vektor gør hverken fra eller til, så $R(A) = R(B)$ er span af de to første vektorer $(1,0,7),(0,1,-2)$. Det betyder at $R(B) = R(C)$, hvor

$$C = \begin{pmatrix} 1 & 0 & 7 \\ 0 & 1 & -2 \end{pmatrix},$$

Vi påstår at $(u_1,u_2) = ((1,0,7),(0,1,-2))$ er en basis for $R(C) = \mathrm{span}(u_1,u_2)$. Men det er det samme som at sige at de to vektorer er lineært uafhængige. Så lad os antage at $\lambda_1 u_1 + \lambda_2 u_2 = 0$. Vi regner:

$$(0,0,0) = \lambda_1(1,0,7) + \lambda_2(0,1,-2) = (\lambda_1, \lambda_2, 7\lambda_1 - 2\lambda_2)$$

Det følger at $\lambda_1 = \lambda_2 = 0$, så at u_1, u_2 faktisk er lineært uafhængige.

Altså er $\dim(V) = 2$ og

$$(u_1^T, u_2^T) = \left(\begin{pmatrix} 1 \\ 0 \\ 7 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -2 \end{pmatrix} \right)$$

er en basis for V. ♠

Selvom jeg har set det mange gange før, synes jeg stadig at Sætning 25(1) nedenfor er ekstremt overraskende. Hvem skulle på forhånd tro at dimensionerne af række- og søjlerummene for en matrix havde noget med hinanden at gøre? Hvorfor skulle for eksempel dimensionerne af

$$\mathrm{span}\left(\begin{pmatrix} 1 \\ 5 \\ 9 \end{pmatrix}, \begin{pmatrix} 2 \\ 6 \\ 10 \end{pmatrix}, \begin{pmatrix} 3 \\ 7 \\ 11 \end{pmatrix}, \begin{pmatrix} 4 \\ 8 \\ 12 \end{pmatrix} \right)$$

og

$$\mathrm{span}\left(\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}, \begin{pmatrix} 5 \\ 6 \\ 7 \\ 8 \end{pmatrix}, \begin{pmatrix} 9 \\ 10 \\ 11 \\ 12 \end{pmatrix} \right)$$

være identiske?

Sætning 25(2) kaldes *dimensionssætningen* og giver en meget stærk og nyttig sammenhæng mellem dimensionerne af nulrummet og søjlerummet for en matrix.

(6.25) SÆTNING.

Lad A være en $m \times n$ matrix. Så gælder

1.

$$\dim C(A) = \dim R(A)$$

2.

$$n = \dim N(A) + \dim C(A)$$

Før vi går i gang med beviset for denne sætning, skal vi i anslutning til eksempel 6.2.2 kigge lidt nærmere på på nulrummet for en matrix A der er på RREF. Det er en god idé at genopfriske dette eksempel før man læser beviset for lemmaet.

(6.26) LEMMA.

Antag at A er en $m \times n$ matrix på RREF. Da er dimensionen af nulrummet $N(A)$ det samme som antallet af søjler i A som ikke *er pivotsøjler.*

Et lemma? Hvad er nu det for noget? Skal vi også kunne lemmaer til eksamen?

Et lemma eller en hjælpesætning er en sætning som ikke er så vigtig i sig selv, men som skal bruges lige om lidt i et bevis for en mere vigtig sætning. Til eksamen bliver der desværre ikke stillet opgaver til beviset for dette lemma. ♠

Bevis *

Bevis. At $\mathbf{v} \in N(A)$ betyder det samme som at $\mathbf{v}$ er en løsning til ligningen $A\mathbf{v} = \mathbf{0}$. I afsnit "Løsning af ligninger via RREF" beskrev vi hvordan man kan finde alle løsninger til dette system når B er på RREF. Vi betragter de n variable x_i, og inddeler disse variable x_i i to grupper, de frie variable x_i for en mængde af indekser $i \in F$, og de bundne variable x_j for en anden mængde af indekser $j \in B$. Enhver variabel x_i er enten fri eller bunden, så at tilsammen udgører F og B mængden af alle indekser $\{1, 2. \dots, n\}$. De bundne variable svarer til pivotsøjlerne, det vil sige at der er lige så mange bundne variable som der er pivotsøjler. De frie variable svarer til de søjler som ikke er pivotsøjler. Resultatet af vores overvejelser var at hvis man angiver værdiene λ_i for hver af de frie variable, så findes der netop en løsning til ligningen $B\mathbf{x} = \mathbf{0}$ så at $x_i = \lambda_i$ for alle frie variable, det vil sige, for $i \in B$.

Nu definerer vi følgende vektorer. For hvert $k \in F$ lader vi $\mathbf{v}_k$ være den eneste ene løsning til $A\mathbf{v} = \mathbf{0}$ hvor den frie variabel $v_{ki} = (\mathbf{v}_k)_i$ med $i \in F$ antager værdien 1, og alle andre frie variable $v_{ki'}$ med $i' \in F, i' \neq i$ antager værdien 0. Vi påstår at $N(A)$ er udspændt af vektorerne $\mathbf{v}_i$ for $i \in F$. Vi skal altså vise at enhver vektor $\mathbf{u} = (u_1, u_2, \dots, u_n)$ der opfylder at $A\mathbf{u} = \mathbf{0}$ er en linearkombination af vektorerne $\mathbf{v}_i$.

Til dette formål definerer vi nu vektoren

$$\mathbf{u}' = \cdots + u_k\mathbf{v}_k + \cdots ,$$

hvor vi bruger alle frie variable $k \in F$. I denne sum forekommer altså ikke nogen j med $j \in B$. I mange kloge bøger ville denne sum i øvrigt skrives som

$$\mathbf{u}' = \sum_{k \in F} u_k \mathbf{v}_k.$$

Vi påstår at $\mathbf{u} = \mathbf{u}'$. For at vise dette bruger vi at der er netop en løsning til ligningen $A\mathbf{x} = \mathbf{0}$ sådan at $x_i = u_i$ for $i \in F$, og den løsning kender vi, fordi det er jo $\mathbf{u}$. For at vise at $\mathbf{u}' = \mathbf{u}$ er det altså fuldstændigt nok at vise at for alle $i \in F$ så er $(\mathbf{u})_i = (\mathbf{u}')_i$. Det er en lille udregning. Antag at $i \in F$.

$$\begin{aligned}(\mathbf{u}')_i &= (\cdots)_i + u_i(\mathbf{v}_i)_i + (\cdots)_i \\ &= 0 + u_i + 0 \\ &= u_i.\end{aligned}$$

Dermed har vi vist at vektorerne $\mathbf{v}_i$ udspænder $N(A)$. Der er lige så mange af dem som antallet søjler i A der ikke er pivotsøjler. Så for at afslutte beviset skal v vise at vektorerne $\mathbf{v}_i$ er lineært uafhængige. Antag altså at $\cdots + \lambda_i \mathbf{v}_i + \cdots = \mathbf{0}$ (hvor selvfølgelig $i \in F$). Da er

$$\begin{aligned}0 &= (\mathbf{0})_i \\ &= (\cdots + \lambda_i \mathbf{v}_i + \cdots)_i \\ &= 0 + \lambda_i + 0 = \lambda_i\end{aligned}$$

Altså er $\lambda_i = 0$ for alle i. Derfor er vektorerne $\mathbf{v}_i$ lineært uafhængige. □

♠

Nu kan vi bevise den vigtige sætning 25.

Bevis

Bevis. Vi viser først sætningen for en matrix B på RREF. Lad os sige at B har p pivotsøjler. Pivotsøjlerne er de første p enhedsvektorer $\mathbf{e}_i$ for $i \leq p$, så at $\text{span}(\mathbf{e}_1, \ldots, \mathbf{e}_p) \subset C(B)$. En søjle i B,

$$\mathbf{v} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix},$$

opfylder at $\lambda_i = 0$ for $i > p$, da B er på RREF. Men det vil sige at

$$\mathbf{v} = \lambda_1 \mathbf{e}_1 + \cdots \lambda_p \mathbf{e}_p \in \text{span}(\mathbf{e}_1, \ldots, \mathbf{e}_p).$$

Det betyder at $C(B) \subset \text{span}(\mathbf{e}_1, \ldots, \mathbf{e}_p)$, og dermed at faktisk $C(B) = \text{span}(\mathbf{e}_1, \ldots, \mathbf{e}_p)$. Men $\text{span}(\mathbf{e}_1, \ldots, \mathbf{e}_p)$ har basis $\mathbf{e}_1, \ldots \mathbf{e}_p$, så at $\dim C(B) = p$.

Rækkerummet $R(B)$ er rummet udspændt af de rækker i B der indeholder pivoter. Vi kan lade de nulrækker væk der ikke indeholder pivotelementer, og vedtage at de resterende rækker hedder $\mathbf{v}_1^T, \ldots, \mathbf{v}_p^T$, så at $R(B) = \text{span}(\mathbf{v}_1^T, \ldots, \mathbf{v}_p^T)$. Sætning 23 siger at $R(B)$ har en basis der består af de rækker som ikke er nulrækker i den reducerede række echelon form for B. Men RREF af B er jo bare B selv, så det betyder at $(\mathbf{v}_1, \ldots, \mathbf{v}_p)$ er lineært uafhængige, og $\dim R(B) = p$. Vi har altså vist at $\dim C(B) = \dim R(B)$.

Lemma 26 siger at $\dim(N(B)) = n - p$. Ved at sætte ind at $\dim C(B) = p$ får vi at $\dim N(B) + \dim C(B) = n$. Dermed er sætningen bevist for de matricer B som er på RREF. Vi skal selvfølgelig også kigge på det generelle tilfælde

Hvis A er en eller anden $m \times n$ matrix, så kan den i hvert fald rækkereduceres til en $m \times n$ matrix B på RREF. Og nu gælder det at $N(A) = N(B)$ og $C(A) = C(B)$. Der er ingen grund til at tro at $R(A) = R(B)$, men i det mindste følger det fra sætning 23 at $\dim R(A)$ er lig med antallet p af pivotelementer i B. På samme måde er $\dim R(B)$ er lig med antallet p af pivotelementer i rækkereduktionen af B, men da B rækkereducerer til sig selv, er $\dim R(B) = p = \dim R(A)$. Nu bruger vi at vi allerede ved at sætningen er rigtig for B til at konkludere

$$\begin{aligned} \dim C(A) &= \dim C(B) = \dim R(B) = \dim R(A) \\ \dim N(A) + \dim C(A) &= \dim N(B) + \dim C(B) = n. \end{aligned}$$

□

♠

Dimensionen af søjlerummet $C(A)$ for en matrix kaldes *rangen* af matricen A og betegnes $\text{rk}(A)$. Med denne betegnelse kan dimensionssætningen for en $m \times n$ matrix udtrykkes som

$$n = \dim N(A) + \text{rk}(A).$$

6.27 Opgave

I denne opgave gælder det om at regne mindst muligt. Hvad er dimensionen af nulrummet for matricen

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

ud fra udregningerne i Eksempel 6.4? ♠

En løs snak om abstrakte vektorrum

Det kan ske at et abstrakt vektorrum V har en basis der består af endeligt mange vektorer. Antallet af elementer i denne basis kaldes for dimensionen af V, og vi siger at V har endelig dimension. I stort set alt der gælder om underrum af F^n vil også gælde for vektorrum af endelig dimension. For eksempel vil alle baser for V have det samme antal elementer, og de efterfølgende afsnit om koordinater og lineære transformationer vil også kunne kopieres næsten uden ændringer. Det betyder jo også at det ikke giver så meget mere indsigt at formulere disse afsnit for abstrakte vektorrum, så det vil vi ikke gøre.

Der findes også abstrakte vektorrum som ikke har en endelig basis. De spiller en stor rolle indenfor matematik, men er ikke så vigtige i anvendelser. I kvantfysik bruges dog ofte “Hilbertrum” som er vektorrum over de kompleks tal, og ikke plejer at have en endelig dimension. ♠

6.5 Koordinater

(6.28) DEFINITION.

Lad V være et underrum af F^n og $B = (\mathbf{v}_1, \dots, \mathbf{v}_r)$ en basis for V. Hvis

$$\mathbf{v} = x_1\mathbf{v}_1 + \cdots + x_r\mathbf{v}_r \in V$$

for $x_1, \dots, x_r \in F$, så kaldes $\begin{pmatrix} x_1 \\ \vdots \\ x_r \end{pmatrix}$ for koordinatvektoren eller koordinaterne for $\mathbf{v}$ med hensyn til basen B.

Læg mærke til at en vektor $\mathbf{v} \in V$ ikke kan have to forskellige koordinatvektorer med hensyn til en basis B som ovenfor. Hvis

$$\mathbf{v} = x_1\mathbf{v}_1 + \cdots + x_r\mathbf{v}_r = y_1\mathbf{v}_1 + \cdots + y_r\mathbf{v}_r,$$

så er

$$(x_1 - y_1)\mathbf{v}_1 + \cdots + (x_r - y_r)\mathbf{v}_r = \mathbf{0},$$

og dermed er $x_1 = y_1, \dots, x_r = y_r$, fordi basisvektorerne er lineært uafhængige.

Der er myriader af baser for et underrum. Det er rigtigt nyttigt at kunne regne om fra koordinater i en basis til koordinater i en anden basis, men det gælder om at holde tungen lige i munden.

(6.29) DEFINITION.

Lad $B_1 = (\mathbf{u}_1, \dots, \mathbf{u}_n)$ og $B_2 = (\mathbf{v}_1, \dots, \mathbf{v}_n)$ være to baser for et underrum V af dimension n. Basisskiftmatricen fra B_1 til B_2 er $n \times n$ matricen $T = (a_{ij})$, hvis *søjler er koordinaterne for basisvektorerne i B_1 med hensyn til basen B_2, det vil sige*

$$\mathbf{u}_i = a_{1i}\mathbf{v}_1 + \cdots + a_{ni}\mathbf{v}_n$$

for $i = 1, \ldots, n$.

6.30 Eksempel

Hvis $B_1 = (\mathbf{u}_1, \ldots, \mathbf{u}_n)$ er en basis for F^n og $B_2 = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ er standardbasen for F^n , så er basisskiftmatricen fra B_1 til B_2 matricen hvis søjler er vektorerne $\mathbf{u}_i$, fordi

$$\mathbf{u}_i = (\mathbf{u}_i)_1\mathbf{e}_1 + (\mathbf{u}_i)_2\mathbf{e}_2 + \cdots + (\mathbf{u}_i)_n\mathbf{e}_n.$$

♠

(6.31) PROPOSITION.

Hvis

$$\mathbf{w} = x_1\mathbf{u}_1 + \cdots + x_n\mathbf{u}_n \tag{6.3}$$

har koordinaterne

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

med hensyn til basen B_1 og det samme $\mathbf{w} = y_1\mathbf{v}_1 + \cdots + y_n\mathbf{v}_n$ har koordinaterne

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

med hensyn til basen B_2, så gælder

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = T \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Matricen T er invertibel.

Hvis $V = F^n$ er $T = B_2^{-1}B_1$, hvor baserne opfattes som søjler i en $n \times n$ matrix.

Bevis *

Bevis. Man udtrykker som skrevet basisvektorerne i B_1 via basisvektorerne i B_2:

$$\begin{aligned}\mathbf{u}_1 &= a_{11}\mathbf{v}_1 + \cdots + a_{n1}\mathbf{v}_n \\ &\vdots \\ \mathbf{u}_n &= a_{1n}\mathbf{v}_1 + \cdots + a_{nn}\mathbf{v}_n\end{aligned} \tag{6.4}$$

og sætter derefter (6.4) ind i (6.3):

$$\begin{aligned}\mathbf{w} = x_1(a_{11}\mathbf{v}_1 + \cdots + a_{n1}\mathbf{v}_n) &+ \\ \vdots& \\ x_n(a_{1n}\mathbf{v}_1 + \cdots + a_{nn}\mathbf{v}_n)&.\end{aligned} \tag{6.5}$$

Ved nu at summere (6.5) lodret får vi koordinaterne for $\mathbf{w}$ i basen B_2 som

$$\begin{aligned}\mathbf{w} = (a_{11}x_1 + \cdots + a_{1n}x_n)\mathbf{v}_1 &+ \\ \vdots& \\ +(a_{n1}x_1 + \cdots + a_{nn}x_n)\mathbf{v}_n&.\end{aligned}$$

Hvis vi skriver $\mathbf{y} = T\mathbf{x}$, så at $y_j = a_{j1}x_1 + a_{j2}x_2 + \cdots + a_{jn}x_j$, får vi altså

$$\mathbf{w} = y_1\mathbf{v}_1 + y_2\mathbf{v}_2 + \cdots + y_n\mathbf{v}_n.$$

Herved ses at matrixmultiplikationen giver koordinaterne for $\mathbf{w}$ i basen B_2.

Hvad er nulvektorens koordinater i en basis B? Ja,

$$\mathbf{0} = 0\mathbf{v}_1 + 0\mathbf{v}_2 + \cdots + 0\mathbf{v}_n,$$

så de er allesammen 0. Omvendt, hvis en vektor $\mathbf{v}$ har alle koordinater lige med 0, så er

$$\mathbf{v} = 0\mathbf{v}_1 + 0\mathbf{v}_2 + \cdots + 0\mathbf{v}_n = \mathbf{0}.$$

Og det gælder for enhver basis. Vi vil bruge dette lille faktum til at vise at T er invertibel. Hvis den ikke var, så ville der findes en vektor $\mathbf{x} \in F^n$, $\mathbf{x} \neq \mathbf{0}$, så at $T\mathbf{x} = \mathbf{0}$. Det vil sige, $T\mathbf{x}$ repræsenterer vektoren $\mathbf{0}$ i basen B_2. Men ifølge det vi lige har sagt om repræsentationen af $\mathbf{0}$ i en basis, betyder det at $\mathbf{x}$ også repræsenterer vektoren $\mathbf{0}$ i basen B_1, Det følger at $\mathbf{x} = \mathbf{0}$, så at T er invertibel.

Hvad med den sidste påstand? Ved vi egentlig at B_2 er invertibel, så at det i det hele taget giver det mening at skrive B_2^{-1}? Jo, det ved vi, fordi ifølge eksempel 30 er B_2 også en basisskiftmatrix. Og vi har jo lige vist, at enhver basisskiftmatrix er invertibel!

Hvis vi nu skriver $B_1 = (u_{ij})$, $B_2 = (v_{ij})$ ser vi at $u_{ij} = (\mathbf{u}_j)_i$ og $v_{ij} = (\mathbf{v}_j)_i$. Læg mærke til at de to indekser i og j ”bytter plads”, og sådan skal det også være. Nu kan vi skrive om på (6.4):

$$\begin{aligned} u_{ij} &= (\mathbf{u}_j)_i \\ &= a_{1j}(\mathbf{v}_1)_i + \cdots + a_{nj}(\mathbf{v}_n)_i \\ &= a_{1j}v_{i1} + \cdots + a_{nj}v_{in} \\ &= v_{i1}a_{1j} + \cdots + v_{in}a_{nj} \end{aligned}$$

Dette genkender vi som et matrixprodukt, så at vi har vist at

$$B_1 = B_2 T$$

Nu multiplicerer vi med B_2^{-1} på venstresiden, og får at

$$B_2^{-1}B_1 = B_2^{-1}(B_2 T) = \left(B_2^{-1}B_2\right)T = T.$$

□

♠

6.32 Eksempel

Det betaler sig at se et helt konkret eksempel på anvendelsen af Proposition 31. Lad

$$B_1 = \left(\begin{pmatrix}5\\3\end{pmatrix}, \begin{pmatrix}2\\1\end{pmatrix}\right) \qquad \text{og} \qquad B_2 = \left(\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\right)$$

være to baser for $\mathbb{R}^2$. Her er T heldigvis nem at regne ud: Da

$$\begin{pmatrix}5\\3\end{pmatrix} = 5\begin{pmatrix}1\\0\end{pmatrix} + 3\begin{pmatrix}0\\1\end{pmatrix} \qquad \text{og}$$

$$\begin{pmatrix}2\\1\end{pmatrix} = 2\begin{pmatrix}1\\0\end{pmatrix} + \begin{pmatrix}0\\1\end{pmatrix}$$

bliver

$$T = \begin{pmatrix}5 & 2\\3 & 1\end{pmatrix}.$$

Så hvis x_1, x_2 er koordinaterne til en vektor i basen B_1, så er koordinaterne y_1, y_2 til vektoren i basen B_2 givet ved

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

Men vent! Det gælder også den anden vej ved at invertere T: Hvis y_1, y_2 er koordinaterne til en vektor i basen B_2, så er koordinaterne x_1, x_2 til vektoren i basen B_1 givet ved

$$\begin{pmatrix}x_1\\x_2\end{pmatrix} = \begin{pmatrix}5 & 2\\3 & 1\end{pmatrix}^{-1}\begin{pmatrix}y_1\\y_2\end{pmatrix} = \begin{pmatrix}-1 & 2\\3 & -5\end{pmatrix}\cdot\begin{pmatrix}y_1\\y_2\end{pmatrix}.$$

For eksempel har vektoren $\begin{pmatrix}1\\1\end{pmatrix}$ koordinaterne $\begin{pmatrix}1\\-2\end{pmatrix}$ med hensyn til basen B_1. Lad os checke påstanden:

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

♠

6.33 Quiz

Hvad gælder om λ_1 og λ_2 hvis

$$\lambda_1 \begin{pmatrix}1\\2\end{pmatrix} + \lambda_2 \begin{pmatrix}3\\4\end{pmatrix} = \begin{pmatrix}2\\3\end{pmatrix}?$$

$\lambda_1 + \lambda_2 = 0$

$\lambda_1 = 2$

$2\lambda_1^2 = \lambda_2$ ♠

6.34 Quiz

Hvis koordinaterne i basen

$$\begin{pmatrix}1\\2\end{pmatrix} \quad \text{og} \quad \begin{pmatrix}3\\4\end{pmatrix}$$

til en vektor er $(2,-1)$, hvad gælder så om koordinaterne (λ_1, λ_2) til vektoren i basen

$$\begin{pmatrix}5\\6\end{pmatrix} \quad \text{og} \quad \begin{pmatrix}7\\8\end{pmatrix}?$$

$\lambda_1 < 0.$

$\lambda_2 < 0.$

$\lambda_1 + \lambda_2 = 1$ ♠

VIDEO: https://youtu.be/pr6uEx9kmSA

6.6 Lineære transformationer

Ligesom definitionen af underrum var forbavsende enkel, har vi her også en ret enkel definition af lineære afbildninger (transformationer) mellem underrum

(6.35) DEFINITION.

Lad $U \subseteq F^m$ og $V \subseteq F^n$ være underrum En lineær transformation fra U til V er en afbildning

$$L : U \to V,$$

som opfylder

1.
$$L(\mathbf{u}+\mathbf{v}) = L(\mathbf{u}) + L(\mathbf{v})$$

2.
$$L(\lambda \mathbf{u}) = \lambda L(\mathbf{u}),$$

for $\mathbf{u}, \mathbf{v} \in U$ og $\lambda \in F$.

Det er ekstremt vigtigt at bemærke at en lineær transformation $L : U \to V$ er givet entydigt ud fra dens værdier

$$L(\mathbf{u}_1), \dots, L(\mathbf{u}_r)$$

på en basis $\mathbf{u}_1, \dots, \mathbf{u}_r$ for U: Enhver vektor $\mathbf{u} \in U$ kan skrives som

$$\mathbf{u} = \lambda_1 \mathbf{u}_1 + \cdots + \lambda_r \mathbf{u}_r$$

for entydigt bestemte $\lambda_1, \dots, \lambda_r \in F$. Dette følger ved at bruge egenskaberne (1) og (2) i Definition 35:

$$\begin{aligned} L(\mathbf{u}) &= L(\lambda_1 \mathbf{u}_1 + \cdots + \lambda_r \mathbf{u}_r) \\ &= L(\lambda_1 \mathbf{u}_1) + \cdots + L(\lambda_r \mathbf{u}_r) \\ &= \lambda_1 L(\mathbf{u}_1) + \cdots + \lambda_r L(\mathbf{u}_r). \end{aligned}$$

6.36 Quiz

Hvilke af nedenstående påstande er rigtige?

Afbildningen $f : \mathbb{R} \to \mathbb{R}$ givet ved $f(x) = x + 1$ er en lineær transformation.

Afbildningen $f: \mathbb{R} \to \mathbb{R}$ givet ved $f(x) = x^2$ er en lineær transformation.

Afbildningen $f: \mathbb{R} \to \mathbb{R}$ givet ved $f(x) = 2x$ er en lineær transformation.

Afbildningen $f: \mathbb{R}^2 \to \mathbb{R}^2$ givet ved $f(x,y) = (x+y, x-y)$ er en lineær transformation. ♠

6.6.1 Repræsentation ved en matrix

Nu kommer vi til et meget centralt punkt i disse noter: sammenhænget mellem matricer og lineære afbildninger. Vi vil gerne kunne sige at "en lineær afbildning er det samme som en matrix", og det er ikke helt forkert, men det er heller ikke helt rigtigt, og for at komme videre er man nødt til at forstå hvorfor. Den lille men vigtige forskel er at før vi kan sige det på denne måde har vi har brug for at vælge baser!

Hvis vi for eksempel er kommet i besiddelse af en lineær transformation $L : U \to V$, og gerne vil "oversætte" denne lineære transformation til en matrix gør vi følgende. Betragt baserne $B^U = (\mathbf{u}_1, \ldots, \mathbf{u}_n)$ for U og $B^V = (\mathbf{v}_1, \ldots, \mathbf{v}_m)$ for V. Per ovenstående noterer vi at

$$\begin{aligned} L(\mathbf{u}_1) &= a_{11}\mathbf{v}_1 + \cdots + a_{m1}\mathbf{v}_m \\ &\vdots \\ L(\mathbf{u}_n) &= a_{1n}\mathbf{v}_1 + \cdots + a_{mn}\mathbf{v}_m \end{aligned}$$

Og nu har vi lavet en matrix.

(6.37) DEFINITION.

Matricen

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

i (6.6) *siges at* repræsentere *L med hensyn til basen B^U for U og basen B^V for V.*

Hvis vi allerede har valgt de to baser B^U og B^V så er matricen (a_{ij}) entydigt bestemt

af L. Hvis vi vil holde styr på at den opstår ud fra L så skriver vi $(a_{ij}) = M(L)$. Nogle mennesker ville endog være omhyggelige nok at skrive $(a_{ij}) = {}_{B^V}M(L)_{B^U}$ for at være sikre på at vi ikke glemmer at vores matrix også afhænger af vores valg af baser.

En forvirrende detalje er at vi ofte har favoritbaser for U og V, og så vælger vi bare dem. For eksempel har vi standardbasen $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ for F^n, så hvis $U = \mathbb{R}^n$

og $V=\mathbb{R}^m$ kan vi altid bare vælge $B^U=(\mathbf{e}_1,\ldots,\mathbf{e}_n)$ og $B^V=(\mathbf{e}_1,\ldots,\mathbf{e}_m)$, og sige at en linear transformation $L:\mathbb{R}^n\to\mathbb{R}^m$ bestemmer en matrix. Simpelthen.

Der er to problemer med denne lemfældige tilgang til situationen. Det ene er at vi sagtens kan komme ud for at betragte lineære afbildninger mellem underrum af $\mathbb{R}^n$, og så findes der ikke mere en standardbasis. Det andet og mere afgørende problem er at vi ofte vil bruge en basis eller måske endog flere forskellige baser der er *tilpasset til den situation vi betragter*.

Så vi må leve med at hvis vi vil kunne gå uhindret fra en lineær afbildning $L:U\to V$ til en matrix, så *skal* U og V være forsynede med baser B^U og B^V.

Vi ved altså at vi kan gå fra lineære transformationer til matricer, men kan vi gå den anden vej? Selvfølgelig kun under forudsætning af at vi har givet baser. Så lad $B^U=(\mathbf{u}_1,\ldots,\mathbf{u}_n)$ være en basis for U og $B^V=(\mathbf{v}_1,\ldots,\mathbf{v}_m)$ en basis for for V. Hvis (a_{ij}) er en $m\times n$ matrix, så laver vi en afbildning $L:U\to V$ sådan her: Vektoren $x_1\mathbf{u}_1+\cdots+x_n\mathbf{u}_n\in U$ afbildes over i vektoren med koordinater

$$\begin{pmatrix}\lambda_{11} & \cdots & \lambda_{1n}\\ \vdots & \ddots & \vdots\\ \lambda_{m1} & \cdots & \lambda_{mn}\end{pmatrix}\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix} \tag{6.6}$$

med hensyn til basen B^V.

Vi formulerer nu en sætning der siger at de to begreber "lineær transformation"og "matrix"kun er to forskellige avatarer af det samme begreb.

(6.38) SÆTNING.

Lad $U\subset F^p$ og $V\subset F^q$ være underrum med baser B^U henholdsvis B^V. Hvis M er en $m\times n$ matrix og L er afbildningen beskrevet ved (6.6)*, så er L en lineær transformation, entydigt bestemt af M. Desuden er $M=M(L)$.*

Ikke flere beviser nu

Beviset for dettes minder i høj grad om beviset for Proposition 31 og vi vil ikke gå i detaljer her. ♠

6.39 Quiz

Antag at den lineære transformation $L:U\to V$, hvor $U=\mathbb{R}^2$ og $V=\mathbb{R}^2$ er givet ved

$$L\begin{pmatrix}x\\ y\end{pmatrix}=\begin{pmatrix}x+y\\ x-y\end{pmatrix}.$$

Hvad er matricen som repræsenterer L med hensyn til

$$B_1^U = B_1^V = \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right)?$$

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

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

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

♠

Nedenfor en forelæsning fra 2012 forklarende matrixrepræsentationer af lineære transformationer med ekstern bistand.

VIDEO: https://youtu.be/31iL3IU7mxQ

Matricen i Definition 37 kan oftest udregnes ved hjælp af nedenstående omformulering af definitionen.

(6.40) PROPOSITION.

Lad $L : U \to V$ være en lineær transformation, $B_1 = (\mathbf{u}_1, \dots, \mathbf{u}_n)$ en basis for U og $B_2 = (\mathbf{v}_1, \dots, \mathbf{v}_m)$ en basis for V. Søjlen A^j i $m \times n$ matricen A med som repræsenterer L med hensyn til baserne B_1 og B_2 er den vektor

$$A^j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

som er entydigt fastlagt af at

$$a_{1j}\mathbf{v}_1 + a_{2j}\mathbf{v}_2 + \cdots a_{mj}\mathbf{v}_m = L(\mathbf{u}_j).$$

En helt konkret anvendelse af resultatet ovenfor er gennemgået i nedenstående video.

VIDEO: https://youtu.be/-HGxo0fj-dY

6.7 Sammensætning af lineære transformationer

De sker at vi har tre underrum U, V, W og lineære transformationer $G : U \to V$ og $F : V \to W$. Da kan vi sammensætte og får en ny lineær transformation FG defineret ved at

$$FG(\mathbf{v}) = F(G(\mathbf{v})).$$

Lad os lægge mærke til at denne sammensætning lige som alle andre sammensætninger af afbildninger overholder den associative lov:

$$((FG)H)(\mathbf{v}) = F(G(H(\mathbf{v}))) = F(GH)(\mathbf{v})$$

Hvis de tre rum har baser B^U, B^V, B^W så kan vi ifølge sætning 38 om matrixrepræsentationer lige så godt betragte de to tre matricer $M(F), M(G)$ og $M(FG)$.

(6.41) SÆTNING.

$$M(FG) = M(F)M(G)$$

<u>Bevis</u>

Vi kan genkende matricen $M(FG)$ på at dens indgange er givet ved at

$$FG(\mathbf{u}_j) = M(FG)_{1j}\mathbf{w}_1 + \cdots + M(FG)_{mj}\mathbf{w}_m \tag{6.7}$$

På den anden side er

$$\begin{aligned}
FG(\mathbf{u}_j) &= F(G(\mathbf{u}_j)) \\
&= F(M(G)_{1j}\mathbf{v}_1 + M(G)_{2j}\mathbf{v}_2 + \cdots M(G)_{kj}\mathbf{v}_k \\
&= M(G)_{1j}F(\mathbf{v}_1) + M(G)_{2j}F(\mathbf{v}_2) + \cdots M(G)_{kj}F(\mathbf{v}_k) \\
&= M(G)_{1j}(M(F)_{11}\mathbf{w}_1 + M(F)_{21}\mathbf{w}_2 + \cdots M(F)_{m1}\mathbf{w}_m) \\
&\quad + M(G)_{2j}(M(F)_{12}\mathbf{w}_1 + M(F)_{22}\mathbf{w}_2 + \cdots M(F)_{m2}\mathbf{w}_m) \\
&\quad \ldots
\end{aligned}$$

Nu samler vi koefficienterne, hvilket i dette tilfælde betyder at vi summerer vertikalt:

$$\begin{aligned}
FG(\mathbf{u}_j) = {} & (M(F)_{11}M(G)_{1j} + M(F)_{12}M(G)_{2j} + \cdots + M(F)_{1k}M(G)_{kj})\mathbf{w}_1 \\
& + (M(F)_{21}M(G)_{1j} + M(F)_{22}M(G)_{2j} + \cdots + M(F)_{2k}M(G)_{kj})\mathbf{w}_2 \\
& \ldots
\end{aligned}$$

Ved at sammenligne koefficienterne i denne ligning med koefficienterne i (6.7) ser vi at

$$M(FG)_{ij} = M(F)_{i1}M(G)_{1j} + M(F)_{12}M(G)_{2j} + \cdots + M(F)_{ik}M(G)_{kj}$$

Her genkender vi formlen for matrixmultiplikation, og slutter at $M(FG) = M(F)M(G)$.
♠

Hvad er spørgsmålet som har svaret "sætning 6.42"?

Hvor kommer formlen for matrixmultiplikation fra? ♠

(6.42) BEMÆRKNING.

Nu kan vi endelig forklare hvorfor det er blændende indlysende at matrixmultiplikation er associativ - det er den fordi den bare er en anden måde at skrive sammensætning af lineære transformationer, og sammensætning af afbildninger er altid associativ! I formler: Lad M_1, M_2, M_3 være tre matricer som kan ganges sammen i denne rækkefølge. Da er

$$\begin{aligned} L((M_1M_2)M_3) &= L(M_1M_2)L(M_3) \\ &= (L(M_1)L(M_2))L(M_3) \\ &= L(M_1)(L(M_2)L(M_3)) \\ &= L(M_1(M_2M_3)) \end{aligned}$$

Det vil sige, $L((M_1M_2)M_3)) = L(M_1(M_2M_3))$ og derfor er $(M_1M_2)M_3 = M_1(M_2M_3)$.

(6.43) PROPOSITION.

Lad A være matricen, som repræsenterer en lineær transformation $L : U \to V$ med hensyn til basen B_1^U for U og basen B_1^V for V.

Matricen, som repræsenterer L med hensyn til basen B_2^U for U og basen B_2^V for V er givet ved

$$SAT,$$

hvor T er basiskiftmatricen fra B_2^U til B_1^U og S er basisskiftmatricen fra B_1^V til B_2^V.

Bevis

Bevis. Vi skal holde styr på fire baser i den her sætning. Det er simpelthen *for meget* for vores hjernekapacitet, så vi vil dele det lidt op. Lad os skrive matricen der repræsenterer L med hensyn til baserne B^U og B^V som $A(B^U, B^V)$. Vi skriver altså $A = A(B_1^U, B_1^V)$ og skal vise at $A(B_2^U, B_2^V) = SAT$. Beviset går nu i to trin.

1. $A(B_2^U, B_1^V) = A(B_1^U, B_1^V)T$
2. $A(B_2^U, B_2^V) = SA(B_2^U, B_1^V)$

Hvis vi kan bevise begge disse to trin så følger sætningen, fordi

$$A(B_2^U, B_2^V) = SA(B_2^U, B_1^V) = S\left(A(B_1^U, B_1^V)T\right) = SAT$$

Beviset for de to trin minder meget om hinanden, og også stærkt om beviset for Sætning 41. Begge dele består af hjernedøde udregninger, som vi drevet af en absurd pligtfølelse nu reproducerer.

HD udregning 1

Vi har tre baser at holde styr på.

$$\begin{aligned} B_1^U &= (\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n) \\ B_2^U &= (\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n) \\ B_1^V &= (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m) \end{aligned}$$

Lad os skrive $T = (t_{ij})$ og $A(B_1^U, B_1^V) = (a_{ij})$ Ifølge proposition 40 kan vi kende matricen $A' = A(B_2^U, B_1^V) = (a'_{ij})$ på at dens søjle med søjlenummer j

$$A'^j = \begin{pmatrix} a'_{1j} \\ a'_{2j} \\ \vdots \\ a'_{mj} \end{pmatrix}$$

opfylder at

$$a'_{1j}\mathbf{v}_1 + a'_{2j}\mathbf{v}_2 + \cdots a'_{mj}\mathbf{v}_m = L(\mathbf{w}_j). \tag{6.8}$$

Men det kan vi faktisk regne på.

$$\begin{aligned} L(\mathbf{w}_j) &= L(t_{1j}\mathbf{u}_1 + t_{2j}\mathbf{u}_2 + \cdots t_{nj}\mathbf{u}_n) \\ &= t_{1j}L(\mathbf{u}_1) + t_{2j}L(\mathbf{u}_2) + \cdots t_{nj}L(\mathbf{u}_n) \\ &= t_{1j}\left(a_{11}\mathbf{v}_1 + \cdots + a_{n1}\mathbf{v}_n\right) + \\ &\quad + t_{2j}\left(a_{12}\mathbf{v}_1 + \cdots + a_{n2}\mathbf{v}_n\right) + \\ &\quad \vdots \\ &\quad + t_{nj}\left(a_{1n}\mathbf{v}_1 + \cdots + a_{nn}\mathbf{v}_n\right) \\ &= (a_{11}t_{1j} + a_{12}t_{2j} + \cdots + a_{1n}t_{nj})\mathbf{v}_1 + \cdots \end{aligned}$$

Ved at sammenligne med (6.8) ser vi at

$$\begin{aligned}
a'_{1j} &= a_{11}t_{1j} + a_{12}t_{2j} + \cdots + a_{1n}t_{nj} \\
a'_{2j} &= a_{21}t_{1j} + a_{22}t_{2j} + \cdots + a_{2n}t_{nj} \\
&\cdots \\
a'_{mj} &= a_{m1}t_{1j} + a_{m2}t_{2j} + \cdots + a_{mn}t_{nj}
\end{aligned}$$

Dette genkender vi som formlen for matrixmultiplikation, så at $A(B_2^U, B_1^V) = A' = AT$. ♠

HD udregning 2

Vi har tre baser at holde styr på.

$$\begin{aligned}
B_2^U &= (\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n) \\
B_1^V &= (\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m) \\
B_2^V &= (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m)
\end{aligned}$$

Lad os skrive $S = (s_{ij})$ og $A(B_2^U, B_1^V) = (a'_{ij})$ Ifølge proposition 40 kan vi kende matricen $A'' = A(B_2^U, B_2^V) = (a''_{ij})$ på at dens søjle med søjlenummer j

$$A''^j = \begin{pmatrix} a''_{1j} \\ a''_{2j} \\ \vdots \\ a''_{mj} \end{pmatrix}$$

opfylder at

$$a''_{1j}\mathbf{x}_1 + a''_{2j}\mathbf{x}_2 + \cdots a''_{mj}\mathbf{x}_m = L(\mathbf{w}_j). \tag{6.9}$$

Men det kan vi faktisk regne på.

$$\begin{aligned}
L(\mathbf{w}_j) &= a'_{1j}\mathbf{v}_1 + a'_{2j}\mathbf{v}_2 + \cdots a'_{mj}\mathbf{v}_m \\
&= a'_{1j}(s_{11}\mathbf{x}_1 + \cdots + s_{m1}\mathbf{x}_m) + \\
&+ a'_{2j}(s_{12}\mathbf{x}_1 + \cdots + s_{m2}\mathbf{x}_m) + \\
&\vdots \\
&+ a'_{mj}(s_{1n}\mathbf{x}_1 + \cdots + s_{mn}\mathbf{x}_m) \\
&= (s_{11}a'_{1j} + s_{12}a'_{2j} + \cdots + s_{1m}a'_{mj})\mathbf{x}_1 + \cdots
\end{aligned}$$

Ved at sammenligne med (6.9) ser vi at

$$\begin{aligned} a''_{1j} &= s_{11}a'_{1j} + s_{12}a'_{2j} + \cdots + s_{1m}a'_{mj} \\ a''_{2j} &= s_{21}a'_{1j} + s_{22}a'_{2j} + \cdots + s_{2m}a'_{mj} \\ &\cdots \\ a''_{mj} &= s_{m1}a'_{1j} + s_{m2}a'_{2j} + \cdots + s_{mm}a'_{mj} \end{aligned}$$

Dette genkender vi som formlen for matrixmultiplikation, så at $A(B_2^U, B_2^V) = A'' = SA'$.

♠

□

♠

Som man hurtigt ser er proposition 43 er en anelse langhåret og notationstungt, og vi skal slet ikke snakke om beviset. Lad os prøve at kigge på den i et helt konkret tilfælde.

6.44 Eksempel

I vektorrummet $V = \mathbb{R}^2$ har vi den naturlige basis

$$B = \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right).$$

Den lineære transformation $f : \mathbb{R}^2 \to \mathbb{R}^2$ givet ved

$$f \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 7x + 2y \\ -15x - 4y \end{pmatrix}$$

repræsenteres af matricen

$$A = \begin{pmatrix} 7 & 2 \\ -15 & -4 \end{pmatrix}$$

med hensyn til basen B for $\mathbb{R}^2$ (her er $U = V = \mathbb{R}^2$ og $B_1^U = B_1^V = B$ med hensyn til notationen i Proposition 43). Hvad er matricen, som repræsenterer f med hensyn til basen

$$B' = \left(\begin{pmatrix} -2 \\ 5 \end{pmatrix}, \begin{pmatrix} -1 \\ 3 \end{pmatrix} \right)?$$

I henhold til Proposition 31 bliver basisskiftmatricen fra B' til B netop

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

Tilsvarende bliver basisskiftmatricen fra B til B'

$$T^{-1} = \begin{pmatrix} -3 & -1 \\ 5 & 2 \end{pmatrix}.$$

Samlet bliver matricen, som repræsenterer f i basen B' så

$$T^{-1}AT = \begin{pmatrix} -3 & -1 \\ 5 & 2 \end{pmatrix} \begin{pmatrix} 7 & 2 \\ -15 & -4 \end{pmatrix} \begin{pmatrix} -2 & -1 \\ 5 & 3 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}.$$

♠

Dette lille eksempel leder os naturligt frem til næste afsnit.

6.8 Egenvektorer og diagonalisering af kvadratiske matricer

Husk på definitionen af diagonaliserbare matricer og egenvektorer fra foregående kapitel. Der definerede vi en kvadratisk matrix A til at være diagonaliserbar, hvis der fandtes en invertibel matrix T så

$$T^{-1}AT$$

er en diagonalmatrix. Samtidig så vi at søjlerne i T blev nødt til at være egenvektorer for A.

Tallet $\lambda \in F$ er en egenværdi for A hvis og kun hvis

$$A\mathbf{v} = \lambda \mathbf{v}$$

for en vektor $\mathbf{v} \neq 0$ i F^n. Hvis λ er en egenværdi lader vi

$$F^n_\lambda = N(A - \lambda I_n).$$

Læg mærke til at F^n_λ er nulrummet for matricen $A - \lambda I_n$ og at dette nulrum netop er mængden af egenvektorer hørende til λ kaldet *egenrummet* hørende til λ.

Vi siger at en vektor $\mathbf{v} \in F^n_\lambda$ *som ikke er nulvektoren* er en *egenvektor hørende til egenværdien* λ. Underrummet $F^n_\lambda \subset F^n$ består altså af alle egenvektorer til egenværdien λ sammen med nulvektoren $\mathbf{0}$.

Med nogle få tricks kan man vise følgende

(6.45) PROPOSITION.

Egenvektorer hørende til forskellige egenværdier er lineært uafhængige.

Bevis

Bevis. Helt præcis vil vi vise, at hvis $\mathbf{v}_1, \dots, \mathbf{v}_n$ er egenvektorer til A med egenværdier $\lambda_1, \dots, \lambda_n$ som alle er forskellige, så er $\mathbf{v}_1, \dots, \mathbf{v}_n$ lineært uafhængige. Vi begynder med at vise det for to vektorer $\mathbf{v}_1, \mathbf{v}_2$. Hvis vi har en linearkombination

$$b_1\mathbf{v}_1 + b_2\mathbf{v}_2 = \mathbf{0} \tag{6.10}$$

så kan vi gange med λ_1 og få

$$\mathbf{0} = \lambda_1 b_1\mathbf{v}_1 + \lambda_1 b_2\mathbf{v}_2 \tag{6.11}$$

Man vi kan også anvende A på (6.10) og få

$$\mathbf{0} = A\mathbf{0} = b_1A\mathbf{v}_1 + b_2A\mathbf{v}_2 = \lambda_1 b_1\mathbf{v}_1 + \lambda_2 b_2\mathbf{v}_2 \tag{6.12}$$

Ved at trække (6.12) fra (6.11) fås

$$(\lambda_1 - \lambda_2)b_2\mathbf{v}_2 = 0,$$

Nu var jo $\mathbf{v}_2$ en egenvektor, så at specielt er $\mathbf{v}_2 \neq \mathbf{0}$. Dette giver $b_2 = 0$, da $\lambda_1 \neq \lambda_2$. På den samme måde er $b_1 = 0$.

Nu har vi starten på en induktion. Induktionsstarten er at sætningen er rigtig for $n = 2$. Induktionsskridtet er at vise at hvis sætningen er rigtig for $n-1$ egenvektorer, så er den også rigtig for n egenvektorer. Hvis vi har en linearkombination

$$b_1\mathbf{v}_1 + b_2\mathbf{v}_2 + \cdots + b_n\mathbf{v}_n = \mathbf{0},$$

får vi på samme måde som før to ligninger:

$$\mathbf{0} = \lambda_1 b_1\mathbf{v}_1 + \lambda_1 b_2\mathbf{v}_2 + \cdots + \lambda_1 b_n\mathbf{v}_n$$

og

$$\mathbf{0} = \lambda_1 b_1\mathbf{v}_1 + \lambda_2 b_2\mathbf{v}_2 + \cdots + \lambda_n b_n\mathbf{v}_n.$$

Dette medfører at

$$b_2(\lambda_1 - \lambda_2)\mathbf{v}_2 + \cdots + b_n(\lambda_1 - \lambda_n)\mathbf{v}_n = 0.$$

Ved at bruge sætningen på de $n-1$ egenvektorer $\mathbf{v}_2, \dots, \mathbf{v}_n$ finder vi at $0 = b_2 = \cdots = b_n$. Det følger dermed også at $b_1\mathbf{v}_1 = \mathbf{0}$, så at $b_1 = 0$, og induktionen er fuldstændig. □

♠

Med hensyn til et konkret eksempel på udregning af egenvektorer for en matrix det vil sige udregning af egenrummene F^n_λ henvises til sidste del (cirka fra 7:25) af videoen (som skamløst er blevet genbrugt fra Kapitel 4) nedenfor.

VIDEO: https://youtu.be/oKziKRz0KzU

Med vores nye viden om underrum, baser og dimension kan vi nu sammenfatte dette i følgende

(6.46) SÆTNING.

En $n \times n$ matrix A er diagonaliserbar hvis og kun hvis F^n har en basis bestående af egenvektorer for A.

At F^n har en basis bestående af egenvektorer er ækvivalent med at

$$\dim F^n_{\mu_1} + \cdots + \dim F^n_{\mu_r} = n,$$

hvor $\mu_1, \ldots, \mu_r$ er de forskellige egenværdier for A.

<u>Bevis</u>

Bevis. Hvis F^n har en basis $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ bestående af egenvektorer for A hørende til respektive egenværdier $\lambda_1, \lambda_2, \ldots, \lambda_n$, så vil matricen T med søjler $\mathbf{v}_1, \ldots, \mathbf{v}_n$ være invertibel og

$$T^{-1}AT = D,$$

hvor

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

Det har vi overejet. Modsat hvis T er invertibel med ovenstående egenskab, så vil søjlerne i T udgøre en basis af egenvektorer med tilhørende egenværdier i diagonalen. For at vise det skriver vi ud de to matrixmultiplikationer AT og TD, og overvejer at er identiske. Det er jo nok at tjekke dette på vektorerne i standardbasen, så det er det vi gør.

$$AT\mathbf{e}_j = A\mathbf{v}_j = \lambda_j \mathbf{v}_j = A(\lambda_j \mathbf{e}_j) = AD\mathbf{e}_j.$$

Vi mangler den sidste påstand i sætningen. Som vi plejer, deler vi beviset op i lidt mindre stykker. Til at begynde med kan vi for hvert μ_i vælge en basis for $F^n_{\mu_i}$, lad os sige at den hedder $\mathbf{u}^i_1, \ldots, \mathbf{u}^i_{m_i}$. Da antallet elementer i en basis er det samme som dimensionen er $m_i = \dim F^n_{\mu_i}$.

Vi påstår nu at hvis vi tager alle disse vektorer sammen, så er de stadig lineært uafhængige i F^n. Hvorfor nu det? Jo, hvis vi har en linearkombination

$$\mathbf{0} = \left(b^1_1 \mathbf{u}^1_1 + \cdots\right) + \left(b^2_1 \mathbf{u}^2_1 + \cdots\right) + \cdots \tag{6.13}$$

samler vi alle termer der hører til det samme egenrum, det vil sige

$$\begin{aligned}
\mathbf{w}_1 &= b_1^1\mathbf{u}_1^1 + \cdots \\
\mathbf{w}_2 &= b_1^2\mathbf{u}_1^2 + \cdots \\
&\cdots
\end{aligned}$$

Da er $\mathbf{w}_i$ enten en egenvektor med egenværdi μ_i, eller også er $\mathbf{w}_i = 0$. Disse μ_i er forskellige, og

$$\mathbf{w}_1 + \mathbf{w}_2 + \cdots = \mathbf{0}$$

Ifølge (6.13) er altså alle $\mathbf{w}_i = 0$, så at

$$\mathbf{0} = \mathbf{w}_i = b_1^i\mathbf{u}_1^i + \cdots \in F_{\mu_i}^n.$$

Da $\mathbf{u}_1^i, \mathbf{u}_2^i, \ldots$ er en basis for $F_{\mu_i}^n$ så følger det at alle $b_j^i = 0$.

Vi har nu lavet en mængde af lineært uafhængige vektorer i F^n. Da de er lineært uafhængige kan der ikke være flere end allerhøjst n af dem, og hvis der er nøjagtig n af dem, så danner de en basis for F^n. Nu tæller vi op hvor mange $\mathbf{u}_j^i$ vi har. Alt i alt er der

$$m_1 + m_2 + \cdots + m_r = \dim F_{\mu_1}^n + \cdots + F_{\mu_r}^n.$$

Der er altså nu to alternativer. Enten er

$$\dim F_{\mu_1}^n + \cdots + \dim F_{\mu_r}^n = n$$

og $\{\mathbf{u}_j^i\}$ danner en basis af egenvektorer for F^n eller også er

$$\dim F_{\mu_1}^n + \cdots + \dim F_{\mu_r}^n < n$$

Nu følger den første halvdel af vores påstand:

Hvis

$$\dim F_{\mu_1}^n + \cdots + \dim F_{\mu_r}^n = n,$$

så er vi i tilfældet hvor der findes en basis for F^n der udelukkende består af egenvektorer for A. For at fuldstændiggøre beviset skal vi vise den omvendte implikation, nemlig at hvis der findes en basis af egenvektorer, så er

$$\dim F_{\mu_1}^n + \cdots + \dim F_{\mu_r}^n = n.$$

Så antag at der findes en sådan basis λ_j.

Vi kan omsortere egenværdierne λ_j så at der er tal $n_0 = 0 < n_1 < n_2 < \cdots < n_{r-1} < n_r = n$ og

$$\begin{aligned}
\lambda_j &= \mu_1 \quad \text{hvis} \quad 0 < j \le n_1, \\
\lambda_j &= \mu_2 \quad \text{hvis} \quad n_1 < j \le n_2, \\
&\cdots \\
\lambda_j &= \mu_r \quad \text{hvis} \quad n_{r-1} < j \le n.
\end{aligned}$$

Der er altså $n_i - n_{i-1}$ af de lineært uafhængige vektorer $\mathbf{v}_j$ der er egenvektorer med egenværdi μ_i. Det betyder at $\dim F^n_{\mu_j} \geq n_i - n_{i-1}$, så at

$$\dim F^n_{\mu_1} + \cdots + \dim F^n_{\mu_r} \geq (n_1 - n_0) + (n_2 - n_1) + \cdots (n - n_{r-1}) = n.$$

Dette udelukker vores andet alternativ. så at vi faktisk har en basis for F^n der består af egenvektorer for A.

□

♠

Det er hændt at jeg er stødt på kandidatstuderende i matematik, som ikke kan give et eksempel på en ikke-diagonaliserbar matrix efter at have været gennem større kurser i lineær algebra og avancerede kurser i topologi og abstrakt algebra. Det får mig til at tænke på at matematikundervisningen tit fokuserer alt for meget på at skrive tingene fint og fejlfrit ned og ofte er alt for langt fra de konkrete tiltag, hvor de fleste mennesker har en reel mulighed for at få en dyb forståelse.

Her er en helt konkret opgave.

6.47 Opgave

Gør detaljeret rede for at matricen

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

ikke er diagonaliserbar. ♠

6.8.1 Egenværdier via potensmetoden

Hånden på hjertet. Vi har reelt kun nu det karakteristiske polynomium til at bestemme egenværdierne for en matrix. For store matricer bliver det helt uoverkommeligt for ikke at sige umuligt at udregne det karakteristiske polynomium.

Der er brug for andre metoder til udregning af egenværdier. Her giver jeg et eksempel på en sådan klassisk metode.

Antag at A er diagonaliserbar med en basis $\mathbf{v}_1, \ldots, \mathbf{v}_n$ af egenvektorer hørende til egenværdierne $\lambda_1, \ldots, \lambda_n$. Antag yderligere at $\dim F^n_{\lambda_1} = 1$ og at

$$|\lambda_1| > |\lambda_i|$$

for $i > 1$.

Begynd med en vektor

$$\mathbf{v}^0 = x_1 \mathbf{v}_1 + \cdots + x_n \mathbf{v}_n$$

og antag $x_1 \neq 0$. Herefter itereres

$$\mathbf{v}^k := A^k \mathbf{v}_0 = x_1 \lambda_1^k \mathbf{v}_1 + \cdots + x_n \lambda_n^k \mathbf{v}_n.$$

Antag at den i-te koordinat i $x_1 \mathbf{v}_1$ er $\neq 0$. Så vil

$$(\mathbf{v}^k)_i = \lambda_1^k x_1 v_{1i} + \cdots + \lambda_n^k x_n v_{ni},$$

fordi $A^k \mathbf{v}_i = \lambda_i^k \mathbf{v}_i$ og dermed

$$\frac{(v^{k+1})_i}{(v^k)_i} = \frac{\lambda_1 x_1 v_{1i} + \varepsilon_1}{x_1 v_{1i} + \varepsilon_2},$$

hvor

$$\begin{aligned}
\varepsilon_1 &= \lambda_2 \left(\frac{\lambda_2}{\lambda_1}\right)^k x_2 v_{2i} + \cdots + \lambda_n \left(\frac{\lambda_n}{\lambda_1}\right)^k x_n v_{ni} \\
\varepsilon_2 &= \left(\frac{\lambda_2}{\lambda_1}\right)^k x_2 v_{2i} + \cdots + \left(\frac{\lambda_n}{\lambda_1}\right)^k x_n v_{ni}
\end{aligned}$$

Derfor vil

$$\lambda^{(k)} = \frac{(v^k)_i}{(v^{k-1})_i} \to \lambda_1$$

for $k \to \infty$.

Lad os illustrere metoden med matricen

$$A = \begin{pmatrix} 7 & 2 \\ -15 & -4 \end{pmatrix}$$

og startvektoren $v^0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

Nedenfor er angivet de første 7 iterationer af metoden. Første søjle angiver k, anden søjle er x-koordinaten for v^k, tredje søjle er y-koordinaten for v^k, mens sidste søjle angiver $\lambda^{(k)}$ for $i = 1$ det vil sige med hensyn til x-koordinaten.

Iterationerne ser ud til at indikere at $\lambda = 2$ er en egenværdi for A, hvilket viser sig at være korrekt.

Der findes modeller i anvendelser hvor en vektor $\mathbf{v}$ angiver tilstanden til et tidspunkt, og den lineære afbildning A fortæller hvordan $\mathbf{v}$ udvikles i løbet af en bestemt tidsenhed (et sekund, en dag eller et år etc..). Det betyder at tilstanden efter at n tidsenheder er forløbet er angivet af vektoren $A^n \mathbf{v}$. Ifølge ovenstående betragtning vil det betyde at hvis den største numerisk største egenvektor λ_1 opfylder at $|\lambda_1| > 1$ så forudsiger modellen eksponentiel vækst. Det vil normalt betyde at på et tidspunkt vil modellen bryde sammen, og ikke mere være en god beskrivelse af virkeligheden

6.9 Gershgorins cirkelsætning

Gershgorins cirkelsætning (efter Semyon Aranovich Gershgorin) udtrykker hvor langt vi kan forvente egenværdierne for en matrix ligger fra diagonalelementer.

Lad $A = (a_{ij})$ være en $n \times n$ matrix og lad

$$D(a_{ii}, R_i) = \{z \in \mathbb{C} \mid |z - a_{ii}| \leq R_i\},$$

hvor

$$R_i = \sum_{j \neq i} |a_{ij}|$$

være cirkelskiven med centrum i a_{ii} og radius R_i i de komplekse tal. Bemærk at R_i er summen af de absolutte værdier af indgangene udenfor diagonalen i i-te række for A.

(6.48) SÆTNING.

Enhver egenværdi for A ligger i mindst en af cirkelskiverne $D(a_{ii}, R_i)$ for $i = 1, \ldots, n$.

Bevis

Bevis. Lad λ være en egenværdi for A og vælg en egenvektor v hørende til λ med en koordinat $x_i = 1$, hvor $|x_j| \leq 1$ for $j \neq i$. Med definitionen af matrixmultiplikation og $Av = \lambda v$ følger

$$\sum_{j \neq i} a_{ij} x_j + a_{ii} = \lambda$$

og dermed

$$|\lambda - a_{ii}| = \left| \sum_{j \neq i} a_{ij} x_j \right| \leq \sum_{j \neq i} |a_{ij}| \, |x_j| \leq \sum_{j \neq i} |a_{ij}| = R_i.$$

□

♠

(6.49) DEFINITION.

En $n \times n$ matrix $A = (a_{ij})$ kaldes strengt diagonaldominant *hvis*

$$|a_{ii}| > \sum_{j \neq i} |a_{ij}|$$

for $i = 1, \ldots, n$.

(6.50) SÆTNING.

En strengt diagonaldominant $n \times n$ matrix A er invertibel.

Bevis

Bevis. Lad a_{ij} betegne indgangene i A. Det er nok at vise at 0 ikke er en egenværdi for A. Hvis 0 var en egenværdi måtte vi have $0 \in D(a_{ii}, R_i)$ for et eller andet $i = 1, \ldots, n$ på grund af Sætning 48. Dette er umuligt, da

$$|a_{ii}| > \sum_{j \neq i} |a_{ij}| = R_i.$$

□

♠

(6.51) EKSEMPEL.

Matricen

$$\begin{pmatrix} 8 & 1 & 2 & 3 \\ 0 & 4 & 1 & 1 \\ 4 & 2 & 7 & 0 \\ 1 & 2 & 1 & 5 \end{pmatrix}$$

er invertibel ifølge Sætning 50. For denne matrix er

$$\begin{aligned} R_1 &= 1+2+3 = 6 \\ R_2 &= 0+1+1 = 2 \\ R_3 &= 4+2+0 = 6 \\ R_4 &= 1+2+1 = 3. \end{aligned}$$

6.10 Opgaver

6.10.1

Lad $v_1, \ldots, v_m \in F^n$. Vis ud fra Definition 1 at

$$\mathrm{span}(v_1, \ldots, v_m)$$

er et underrum i F^n.

Vink

Begynd med $m = 1$ og $m = 2$ for at få ideer. ♠

6.10.2

Lad $e_1, \dots, e_n$ være den naturlige basis for $\mathbb{R}^n$ det vil sige

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1 e_1 + \cdots + x_n e_n$$

for $x_1, \dots, x_n \in \mathbb{R}$. Hvis V er et underrum af $\mathbb{R}^n$ og $e_1, \dots, e_n$ alle ligger i V, hvorfor gælder så at $V = \mathbb{R}^n$?

6.10.3

Er

$$\left(\begin{pmatrix} 2 \\ 3 \\ 5 \end{pmatrix}, \quad \begin{pmatrix} 7 \\ 11 \\ 13 \end{pmatrix}, \quad \begin{pmatrix} 17 \\ 19 \\ 23 \end{pmatrix} \right)$$

en basis for $\mathbb{R}^3$? Begrund dit svar.

6.10.4

Lad

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

Find baser for $R(A), C(A)$ og $N(A)$ som underrum af $\mathbb{R}^3$.

6.10.5

Lad A være en 4×5 matrix med rang 3. Hvad kan du sige om $\dim N(A)$? Opskriv et eksempel på en matrix A med disse egenskaber.

6.10.6

Opgave om taxa fordeling.

6.10.7

Lad $L : U \to V$, hvor $U = V = \mathbb{R}^2$ være givet ved

$$L \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ -y \end{pmatrix}.$$

Find matrixrepræsentationen af L med hensyn til basen

$$\left(\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix} \right)$$

for U og basen

$$\left(\begin{pmatrix}2\\3\end{pmatrix},\begin{pmatrix}3\\2\end{pmatrix}\right)$$

for V.

6.10.8

Lad

$$A=\begin{pmatrix}11 & -2 & 2\\ -36 & 10 & -8\\ -81 & 18 & -16\end{pmatrix}.$$

Det opgives at $\lambda = 1$ og $\lambda = 2$ er egenværdierne for A. Undersøg om $\mathbb{R}^3$ har en basis af egenvektorer for A det vil sige om A er diagonaliserbar. Find i givet fald en invertibel matrix T så

$$T^{-1}AT$$

er en diagonalmatrix.

6.10.9

Gør rede for at matricen

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

ikke er diagonaliserbar ud fra oplysningen om at dens karakteristiske polynomium er

$$-\lambda^3+7\lambda^2-11\lambda+5=(\lambda-1)^2(\lambda-5).$$

6.10.10

Kan en invertibel matrix have 0 som egenværdi?

6.10.11

Sandsynliggør at 2 er en egenværdi for

$$A=\begin{pmatrix}1 & 1\\ 1 & 1\end{pmatrix}$$

ved at benytte potensmetoden beskrevet i afsnit 6.8.1. Er A diagonaliserbar?

6.10.12

Hvorfor vil n lineært uafhængige vektorer i et underrum af dimension n altid udgøre en basis?

Kapitel 7

Prikprodukter

VIDEO: https://youtu.be/sxLS7uWh3XU

I planen $\mathbb{R}^2$ har vi stiftet bekendtskab med prikproduktet

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = x_1 y_1 + x_2 y_2$$

for to vektorer. Ved hjælp af prikproduktet indførte vi begreber som længden af en vektor og ortogonalitet. Længden af en vektor $\mathbf{u} \in \mathbb{R}^2$ blev defineret som $\sqrt{\mathbf{u} \cdot \mathbf{u}}$ og to vektorer $\mathbf{u}, \mathbf{v} \in \mathbb{R}^2$ blev kaldt vinkelrette eller ortogonale hvis $\mathbf{u} \cdot \mathbf{v} = 0$.

I dette kapitel indfører vi den oplagte generalisering af prikproduktet fra planen $\mathbb{R}^2$ til vilkårlige søjlevektorer $\mathbb{R}^n$ ved

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = x_1 y_1 + \cdots + x_n y_n. \tag{7.1}$$

Næsten alle definitioner fra $\mathbb{R}^2$ generaliserer. Læg mærke til at prikproduktet kan formuleres via matrixmultiplikationen

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

svarende til $\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T \mathbf{v}$ for to vektorer $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$.

For at udvide prikproduktet i (7.1) fra $\mathbb{R}^n$ til $\mathbb{C}^n$ kan man ikke overtage definitionen uden ændringer.

For naturligt at kunne definere længden af en vektor $\mathbf{u}$ har vi brug for at $\mathbf{u} \cdot \mathbf{u} \geq 0$. Dette er ikke opfyldt for $\mathbf{u} \in \mathbb{C}^n$ med definitionen i (7.1). Allerede for $n = 1$ gælder $\mathbf{u} \cdot \mathbf{u} = -1$ med $\mathbf{u} = (i)$. Heldigvis kan vi reparere denne defekt ret enkelt ved at benytte den konjugerede

$$\bar{z} = x - iy$$

til et komplekst tal $z = x + iy$. Her er $z\bar{z} = x^2 + y^2 = |z|^2$. Med disse forberedelser udvider prikproduktet til $\mathbb{C}^n$ som

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = x_1\bar{y}_1 + \cdots + x_n\bar{y}_n,$$

hvor indgangene i den anden vektor er konjugerede.

Hvis for eksempel $\mathbf{u} = (a+bi, c+di)^T \in \mathbb{C}^2$, så vil $\mathbf{u} \cdot \mathbf{u} = a^2 + b^2 + c^2 + d^2$ og generelt kan man sige at $\mathbb{C}^n$ i denne forstand svarer til $\mathbb{R}^{2n}$.

Moralen i dette kapitel er at lineær algebra bliver væsentligt mere kraftfuldt (og spændende!), når vi har et prikprodukt at arbejde med.

VIDEO: https://youtu.be/l2vSlep$_D Bc$

7.1 Definitioner og uligheder

(7.1) DEFINITION.

For en $m \times n$ matrix A med indgange i de komplekse tal defineres den komplekst konjugerede matrix $\bar{A}$ som matricen A med indgangene konjugerede det vil sige

$$\bar{A}_{ij} = \overline{A_{ij}}.$$

Læg mærke til at $\bar{A} = A$, hvis A har indgange i de reelle tal. Vi er nu i stand til at generalisere næsten alle begreberne vi har for prikproduktet i planen.

(7.2) DEFINITION.

Lad $\mathbf{u}, \mathbf{v} \in \mathbb{C}^n$. Prikproduktet mellem $\mathbf{u}$ og $\mathbf{v}$ er defineret som

$$\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T \bar{\mathbf{v}} = x_1\bar{y}_1 + \cdots + x_n\bar{y}_n,$$

for $\mathbf{u} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ *og* $\mathbf{v} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$.

1. *Vektorerne $\mathbf{u}$ og $\mathbf{v}$ kaldes vinkelrette eller ortogonale og betegnes $\mathbf{u} \perp \mathbf{v}$, hvis*

$$\mathbf{u} \cdot \mathbf{v} = 0.$$

2. *Længden af* $\mathbf{u}$ *er defineret som*

$$|\mathbf{u}| = \sqrt{\mathbf{u} \cdot \mathbf{u}}.$$

3. *En vektor* $\mathbf{u}$ *kaldes en* enhedsvektor *hvis* $|\mathbf{u}| = 1$.

7.3 Quiz

Hvilke af nedenstående udsagn er rigtige, når prikproduktet er defineret som i Definition 2?

Hvis $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ er

$$\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T \mathbf{v}.$$

$$\begin{pmatrix} 1 \\ i \end{pmatrix} \cdot \begin{pmatrix} 1 \\ i \end{pmatrix} = 0.$$

$$\begin{pmatrix} 1 \\ i \end{pmatrix} \cdot \begin{pmatrix} 1 \\ i \end{pmatrix} = 2.$$

$$\begin{pmatrix} 1 \\ 2 \end{pmatrix} \cdot \begin{pmatrix} 3 \\ 4 \end{pmatrix} = 11.$$

$$\overline{\begin{pmatrix} 1-i & -i \\ 1-i & -i \end{pmatrix}} = \begin{pmatrix} 1+i & i \\ -1+i & -i \end{pmatrix}.$$

$$\overline{\begin{pmatrix} 1-i & -i \\ 1-i & i \end{pmatrix}} = \begin{pmatrix} 1+i & i \\ 1+i & -i \end{pmatrix}.$$

For to matricer A og B, hvor matrixproduktet AB giver mening gælder

$$\overline{AB} = A\bar{B}.$$

For to matricer A og B, hvor matrixproduktet AB giver mening gælder

$$\overline{AB} = \bar{A}\bar{B}.$$

♠

(7.4) PROPOSITION.

Lad $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{C}^n$ *og* $\lambda \in \mathbb{C}$. *Så gælder*

1. $(\lambda \mathbf{u}) \cdot \mathbf{v} = \lambda(\mathbf{u} \cdot \mathbf{v})$
2. $\mathbf{u} \cdot (\lambda \mathbf{v}) = \bar{\lambda}(\mathbf{u} \cdot \mathbf{v})$
3. $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$
4. $\mathbf{u} \cdot \mathbf{v} = \overline{\mathbf{v} \cdot \mathbf{u}}$
5. $\mathbf{u} \cdot \mathbf{u} \geq 0$ *og* $\mathbf{u} \cdot \mathbf{u} = 0$ *hvis og kun hvis* $\mathbf{u} = 0$.

__Bevis__

Med prikproduktet defineret som $\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T \bar{\mathbf{v}}$ følger resultaterne af regneregler for matrixmultiplikation og konjugering af komplekse tal og overlades til læseren.

♠

Ud fra kun definitionen af prikproduktet kan vi lave overraskende meget nyttig matematik, bland andet nedenstående resultat, som kaldes Cauchy-Schwarz ulighed

(7.5) SÆTNING.

Lad $\mathbf{u}, \mathbf{v} \in \mathbb{C}^n$. *Så gælder*

$$|\mathbf{u} \cdot \mathbf{v}| \leq |\mathbf{u}||\mathbf{v}|.$$

Lighedstegnet $|\mathbf{u} \cdot \mathbf{v}| = |\mathbf{u}||\mathbf{v}|$ *gælder hvis og kun hvis en af* $\mathbf{u}$ *og* $\mathbf{v}$ *er et multiplum af den anden, det vil sige, hvis* $\mathbf{u}$ *og* $\mathbf{v}$ *ikke er lineært uafhængige.*

__Bevis__

Bevis. Med det komplekse prikprodukt $\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T \bar{\mathbf{v}}$ har vi

$$\begin{aligned} 0 \leq (\mathbf{u} + \lambda \mathbf{v}) \cdot (\mathbf{u} + \lambda \mathbf{v}) &= |\mathbf{u}|^2 + \mathbf{u} \cdot (\lambda \mathbf{v}) + (\lambda \mathbf{v}) \cdot \mathbf{u} + |\lambda|^2 |\mathbf{v}|^2 && (7.2) \\ &= |\mathbf{u}|^2 + \mathbf{u} \cdot (\lambda \mathbf{v}) + \overline{\mathbf{u} \cdot (\lambda \mathbf{v})} + |\lambda|^2 |\mathbf{v}|^2 && (7.3) \\ &= |\mathbf{u}|^2 + 2\mathrm{Re}\, \mathbf{u} \cdot (\lambda \mathbf{v}) + |\lambda|^2 |\mathbf{v}|^2 && (7.4) \\ &= |\mathbf{u}|^2 + 2\mathrm{Re}\, \bar{\lambda}(\mathbf{u} \cdot \mathbf{v}) + |\lambda|^2 |\mathbf{v}|^2, && (7.5) \end{aligned}$$

for et vilkårligt komplekst tal λ, fordi $z+\bar{z}=2\mathrm{Re}(z)$ for et komplekst tal z. Hvis $|\mathbf{v}|\neq 0$ kan vi indsætte

$$\lambda=-\frac{\mathbf{u}\cdot\mathbf{v}}{|\mathbf{v}|^2}$$

i (7.2) og få

$$0\leq|\mathbf{u}+\lambda\mathbf{v}|^2=|\mathbf{u}|^2-2\frac{|\mathbf{u}\cdot\mathbf{v}|^2}{|\mathbf{v}|^2}+\frac{|\mathbf{u}\cdot\mathbf{v}|^2|\mathbf{v}|^2}{|\mathbf{v}|^4}=|\mathbf{u}|^2-\frac{|\mathbf{u}\cdot\mathbf{v}|^2}{|\mathbf{v}|^2},\qquad(7.6)$$

hvoraf uligheden følger. Vi skal nu overveje hvornår denne ulighed faktisk er en lighed.

Hvis $\mathbf{v}=\mathbf{0}$ gælder lighed, og også at $\mathbf{u},\mathbf{v}$ er lineært afhængige.

Vi kan altså antage at $\mathbf{v}\neq\mathbf{0}$. Hvis $\mathbf{u},\mathbf{v}$ er lineært afhængige kan vi skrive $\mathbf{u}=\lambda\mathbf{v}$ for $\lambda\in\mathbb{C}$. Da er

$$|\mathbf{u}\cdot\mathbf{v}|=|\lambda|\cdot|\mathbf{v}|^2=|\lambda\mathbf{v}|\cdot|\mathbf{v}|=|\mathbf{u}|\cdot|\mathbf{v}|$$

så at vi har lighed. Antag omvendt at $|\mathbf{u}\cdot\mathbf{v}|=|\mathbf{u}|\cdot|\mathbf{v}|$. Da $\mathbf{v}\neq\mathbf{0}$ kan vi dividere med $|\mathbf{v}|^2$, så at vi kan definere $\lambda=-\frac{\mathbf{u}\cdot\mathbf{v}}{|\mathbf{v}|^2}$. Ifølge (7.6) er $|\mathbf{u}+\lambda\mathbf{v}|^2=0$ så at $\mathbf{u}=-\lambda\mathbf{v}$. Altså er $\mathbf{u}$ et multiplum af $\mathbf{v}$. □

♠

7.6 Opgave

Lad $x_1,\dots,x_n\in\mathbb{R}$. Prøv at bevise

$$(x_1+x_2+\cdots+x_n)^2\leq n(x_1^2+x_2^2+\cdots+x_n^2)$$

uden at bruge Sætning 5. Lykkedes det? Hvis ikke, prøv med. ♠

7.7 Eksempel

Cauchy-Schwarz ulighed giver at

$$-1\leq\frac{u\cdot\mathbf{v}}{|\mathbf{u}|\,|\mathbf{v}|}\leq 1$$

for to vektorer $\mathbf{u},\mathbf{v}\in\mathbb{R}^n$ og det giver mening at definere vinklen θ mellem dem ud fra

$$\cos(\theta)=\frac{\mathbf{u}\cdot\mathbf{v}}{|\mathbf{u}|\,|\mathbf{v}|}.$$

Dette tal benyttes under betegnelsen *cosine similarity* ofte som et mål for korrelationen mellem to datasæt givet i form af de to vektorer $\mathbf{u}$ og $\mathbf{v}$.

For eksempel kan man benytte det som et mål for ligheden mellem to tekster. Lad os som eksempel se på to tekster

Matematik	1	1
er	2	2
sjovt	1	0
morsomt	0	1
og	1	1
lineær	1	1
algebra	1	1
anvendeligt	1	0
brugbart	0	1

1. Matematik er sjovt og lineær algebra er anvendeligt
2. Matematik er morsomt og lineær algebra er brugbart

Ud fra ordene i de to tekster laver vi følgende to vektorer i $\mathbb{R}^9$:

hvor hvert ord i de to tekster har en koordinat, som tæller antal forekomster af ordet i teksten. Et mål for ligheden mellem de to tekster er cosinus til vinklen mellem de to vektorer. Jo tættere cosinus til vinklen kommer på 1 (svarende til en vinkel på 0 grader), jo tættere anser vi de to tekster at være på hinanden.

I ovenstående tilfælde er $\cos(\theta) = 0.8$. ♠

Ud fra Cauchy-Schwarz ulighed kan vi udlede trekantsuligheden.

(7.8) PROPOSITION.

For to vektorer $\mathbf{u}, \mathbf{v} \in \mathbb{C}^n$ *gælder*

$$|\mathbf{u}+\mathbf{v}| \leq |\mathbf{u}| + |\mathbf{v}|.$$

Lighedstegnet gælder hvis og kun hvis $\mathbf{u} = \mu \mathbf{v}$ *for* $\mu \geq 0 \in \mathbb{R}$.

<u>Bevis</u>

Bevis. Beviset følger af Sætning 5 på følgende måde:

$$\begin{aligned} |\mathbf{u}+\mathbf{v}|^2 = (\mathbf{u}+\mathbf{v}) \cdot (\mathbf{u}+\mathbf{v}) &= |\mathbf{u}|^2 + |\mathbf{v}|^2 + 2\mathrm{Re}\, \mathbf{u} \cdot \mathbf{v} \\ &\leq |\mathbf{u}|^2 + |\mathbf{v}|^2 + 2|\mathbf{u} \cdot \mathbf{v}| \\ &\leq |\mathbf{u}|^2 + |\mathbf{v}|^2 + 2|\mathbf{u}||\mathbf{v}| = (|\mathbf{u}| + |\mathbf{v}|)^2. \end{aligned}$$

Ved lighedstegn må begge de to uligheder være ligheder, så at

1. $\mathrm{Re}\, \mathbf{u} \cdot \mathbf{v} = |\mathbf{u} \cdot \mathbf{v}|$
2. $|\mathbf{u} \cdot \mathbf{v}| = |\mathbf{u}||\mathbf{v}|$

Ifølge (2) og Sætning 5 må $\mathbf{u} = \mu\mathbf{v}$ for et $\mu \in \mathbb{C}$. Ved at kigge nærmere på beviset, ser vi at

$$\mu = \frac{\mathbf{u}\cdot\mathbf{v}}{|\mathbf{v}|^2}.$$

Vi sætter dette ind i (1):

$$\operatorname{Re}\mathbf{u}\cdot\mathbf{v} = |\mathbf{u}\cdot\mathbf{v}| = |\mu|\cdot|\mathbf{v}|^2 = |\mathbf{u}\cdot\mathbf{v}|.$$

Men hvis $z = a+bi$ er et komplekst tal og $a = \operatorname{Re} z = |z| = \sqrt{a^2+b^2}$ så er $b = 0$ og $a \geq 0$, så $\mathbf{u}\cdot\mathbf{v} \in \mathbb{R}$ med $\mathbf{u}\cdot\mathbf{v} \geq 0$. Dermed er

$$\mu = \frac{\mathbf{u}\cdot\mathbf{v}}{|\mathbf{v}|^2} > 0$$

□

♠

7.2 Matricer og prikproduktet

Prikproduktet giver anledning til at indføre nogle vigtige klasser af matricer. Først en vigtig definition for komplekse matricer.

(7.9) DEFINITION.

Den konjugerede transponerede A^* af en $m\times n$ matrix A med komplekse indgange er defineret som

$$A^* = \bar{A}^T$$

det vil sige $A^*_{ij} = \bar{A}_{ji}$ for $i = 1,\dots,n$ og $j = 1,\dots,m$. Matricen A kaldes hermitesk *hvis*

$$A = A^*.$$

En hermitesk matrix med reelle indgange kaldes symmetrisk.

(7.10) EKSEMPEL.

For eksempel er

$$A^* = \begin{pmatrix} -i & 1 \\ 1-i & i \end{pmatrix} \qquad \text{for} \qquad A = \begin{pmatrix} i & 1+i \\ 1 & -i \end{pmatrix}.$$

Læg også mærke til at $B^* = B^T$ hvis og kun hvis B har reelle tal som indgange, netop fordi $z = \bar{z}$ gælder hvis og kun hvis $z \in \mathbb{R}$. En symmetrisk matrix B opfylder altså også at $B = B^T$.

7.11 Quiz

Hvilke af nedenstående udsagn er rigtige?

En hermitesk matrix er kvadratisk.

En hermitesk matrix kan godt have det komplekse tal $2+3i$ som indgang i diagonalen.

$$H = \begin{pmatrix} i & 1+i \\ 1 & -i \end{pmatrix}$$

er en hermitesk matrix.

$$H^*H$$

er en hermitesk matrix.

$$\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$

er en hermitesk matrix.

$$\begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix}$$

er en hermitesk matrix. ♠

7.2.1 Egenværdier og egenvektorer for hermiteske matricer

Nedenstående resultat viser hvordan prikproduktet opfører sig under matrixmultiplikation på en af vektorerne. Beviserne for formlerne gør brug af formlen $(AB)^T = B^T A^T$ for matricer A og B hvor produktet AB giver mening.

(7.12) PROPOSITION.

1. *Hvis en $n \times n$ matrix A opfylder at*

$$\mathbf{u} \cdot (A\mathbf{v}) = \mathbf{u} \cdot \mathbf{v}$$

for alle $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ (eller $\mathbb{C}^n$) så må $A = I_n$.

2. *Hvis A er en $n \times n$ matrix med reelle indgange, så gælder*

$$(A\mathbf{u}) \cdot \mathbf{v} = \mathbf{u} \cdot (A^T \mathbf{v})$$

for prikproduktet på $\mathbb{R}^n$.

3. *Hvis A er en $n \times n$ matrix med komplekse indgange, så gælder*

$$(A\mathbf{u}) \cdot \mathbf{v} = \mathbf{u} \cdot (A^* \mathbf{v})$$

for prikproduktet på $\mathbb{C}^n$.

Bevis

Bevis. Påstanden i (1) følger af at

$$\mathbf{e}_i \cdot (A\mathbf{e}_j) = A_{ij},$$

hvor $\mathbf{e}_1, \dots, \mathbf{e}_n$ er standardbasen for $\mathbb{R}^n$ (eller $\mathbb{C}^n$).

Påstanden i (2) følger af udregningen

$$(A\mathbf{u}) \cdot \mathbf{v} = (A\mathbf{u})^T \mathbf{v} = (\mathbf{u}^T A^T)\mathbf{v} = \mathbf{u}^T (A^T \mathbf{v}) = \mathbf{u} \cdot (A^T \mathbf{v}).$$

Påstanden i (3) følger på den samme måde af udregningen

$$(A\mathbf{u}) \cdot \mathbf{v} = (A\mathbf{u})^T \bar{\mathbf{v}} = (\mathbf{u}^T A^T)\bar{\mathbf{v}} = \mathbf{u}^T (A^T \bar{\mathbf{v}}) = \mathbf{u}^T (\overline{\bar{A}^T \mathbf{v}}) = \mathbf{u} \cdot (A^* \mathbf{v}).$$

□

♠

En ekstremt vigtig observation er at selvom hermiteske matricer generelt har komplekse tal som indgange så er deres egenværdier reelle tal og egenvektorer hørende til forskellige egenværdier er ortogonale. Dette er en meget stærk anvendelse af prikproduktet og faktisk er beviset slet ikke så svært.

(7.13) SÆTNING.

Lad A være en hermitesk matrix og λ en egenværdi for A. Så er λ reel.

Lad μ være en egenværdi for A med $\mu \neq \lambda$. Hvis $\mathbf{u}$ er en egenvektor hørende til μ og v en egenvektor hørende til λ, så gælder

$$\mathbf{u} \cdot \mathbf{v} = 0.$$

Det vil sige at egenvektorer hørende til forskellige egenrum for A er ortogonale.

Bevis

Bevis. For en egenvektor $\mathbf{v}$ hørende til λ har vi

$$\lambda(\mathbf{v}\cdot\mathbf{v}) = (A\mathbf{v})\cdot\mathbf{v} = \mathbf{v}\cdot(A^*\mathbf{v}) = \mathbf{v}\cdot(A\mathbf{v}) = \mathbf{v}\cdot(\lambda\mathbf{v}) = \bar{\lambda}(\mathbf{v}\cdot\mathbf{v}).$$

Dermed er $\lambda = \bar{\lambda}$ og vi må have $\lambda \in \mathbb{R}$, fordi $\mathbf{v} \neq 0$ og dermed $\mathbf{v}\cdot\mathbf{v} \neq 0$

Samme type argument giver ortogonaliteten af egenvektorerne $\mathbf{u}$ og $\mathbf{v}$:

$$\mu(\mathbf{u}\cdot\mathbf{v}) = (A\mathbf{u})\cdot\mathbf{v} = \mathbf{u}\cdot(A\mathbf{v}) = \lambda(\mathbf{u}\cdot\mathbf{v}).$$

Denne identitet medfører ligningen

$$(\mu-\lambda)(\mathbf{u}\cdot\mathbf{v}) = 0,$$

hvoraf $\mathbf{u}\cdot\mathbf{v} = 0$, da $\mu \neq \lambda$. □

♠

Faktisk er det ret utroligt at hvis man stanger en vilkårlig symmetrisk matrix som for eksempel

$$\begin{pmatrix} 1 & 3 & 7 & 5 \\ 3 & 1 & 2 & 4 \\ 7 & 2 & 2 & 6 \\ 5 & 4 & 6 & 1 \end{pmatrix}$$

ud, at dens karakteristiske polynomium kun kan have reelle rødder. Det er en af konsekvenserne af Sætning 13.

At vise at en hermitesk matrix faktisk har en egenværdi overhovedet er mere kompliceret. Her bliver vi nødt til at henvise til algebraens fundamentalsætning, som siger at ethvert polynomium af grad ≥ 1 har en kompleks rod.

VIDEO: https://youtu.be/0EJLfADstog

7.14 Opgave

Gør helt eksplicit rede for at en symmetrisk 2×2 matrix

$$\begin{pmatrix} a & b \\ b & c \end{pmatrix},$$

med $a, b, c \in \mathbb{R}$ faktisk har reelle egenværdier. ♠

7.2.2 Ortogonale og unitære matricer

Ud fra Proposition 12 fremkommer følgende resultat.

(7.15) PROPOSITION.

1. *Hvis A er en $n \times n$ matrix med reelle indgange og*

$$(A\mathbf{u}) \cdot (A\mathbf{v}) = \mathbf{u} \cdot \mathbf{v}$$

for alle $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, så må $A^T A = I_n$.

2. *Hvis A er en $n \times n$ matrix med komplekse indgange og*

$$(A\mathbf{u}) \cdot (A\mathbf{v}) = \mathbf{u} \cdot \mathbf{v}$$

for alle $\mathbf{u}, \mathbf{v} \in \mathbb{C}^n$, så må $A^* A = I_n$.

Matricer, som bevarer prikproduktet i ovenstående forstand, er så vigtige at de har specielle navne.

(7.16) DEFINITION.

1. *En $n \times n$ matrix Q med reelle indgange kaldes* ortogonal *hvis*

$$Q^T Q = I_n.$$

2. *En $n \times n$ matrix U med komplekse indgange kaldes* unitær *hvis*

$$U^* U = I_n.$$

(7.17) PROPOSITION.

En $n \times n$ matrix er ortogonal/unitær netop når dens søjler er ortogonale enhedsvektorer med hensyn til prikproduktet.

<u>Bevis</u>

Bevis. Dette følger af definitionen på matrixmultiplikation opfattet på følgende måde. Lad A og B være reelle $n \times n$ matricer og lad $\mathbf{b}_1, \dots, \mathbf{b}_n$ være søjlerne i B. Så er søjlerne i AB netop

$$A\mathbf{b}_1, \dots, A\mathbf{b}_n.$$

Altså er søjlerne i $B^T B$ netop $B^T \mathbf{b}_i$ og dermed er

$$(B^T B)_{ij} = (B^T)^i B_j = \mathbf{b}_i \cdot \mathbf{b}_j.$$

Hvis B er en kompleks matrix, så er tilsvarende indgangene i $B^* B$ netop $\mathbf{b}_i \cdot \mathbf{b}_j$ (med det komplekse prikprodukt). □

♠

7.18 Eksempel

1. Lad os analysere hvordan en ortogonal 2×2 matrix Q tager sig ud. Antag at

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

Så er

$$Q^T Q = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a^2 + c^2 & ab + cd \\ ab + dc & b^2 + d^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Derfor er søjlevektorerne i Q ortogonale enhedsvektorer. Sætter vi $a = \cos(\theta)$ og $c = \sin(\theta)$ har vi altså to muligheder for Q:

$$Q_1 = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \qquad \text{og} \qquad Q_2 = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix}.$$

Geometrisk svarer Q_1 til en drejning af planen, mens Q_2 svarer til en spejling. Ved eksplicit udregning ser man at 1 er en egenværdi for Q_2. En tilhørende egenvektor er retningsvektor for linjen, som der spejles i (som viser sig at være linjen gennem $(0,0)$ med vinklen $\theta/2$ med x-aksen).

2. En unitær 1×1 matrix er netop et komplekst tal z med $z\bar{z} = 1$. Det vil sige $z = e^{i\theta}$ for en passende vinkel θ. For to komplekse tal $a, b \in \mathbb{C}$ med $|a|^2 + |b|^2 = 1$ er

$$U = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}$$

et eksempel på en unitær 2×2 matrix. Her kan man for eksempel benytte

$$a = e^{i\varphi_1} \cos\theta$$
$$b = e^{i\varphi_2} \sin\theta$$

for vilkårlige vinkler $\theta, \varphi_1, \varphi_2$. For eksempel er

$$\begin{pmatrix} \frac{1}{2}+\frac{1}{2}i & \frac{1}{2}+\frac{1}{2}i \\ & \\ -\frac{1}{2}+\frac{1}{2}i & \frac{1}{2}-\frac{1}{2}i \end{pmatrix}$$

en unitær 2×2 matrix svarende til $\theta = \varphi_1 = \varphi_2 = \pi/4$.

Tre berømte unitære og hermiteske matricer er

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \text{og} \qquad \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

I kapitlet om anvendelser vil prøve på at forklare hvorfor disse matricer er så berømte.

♠

7.19 Quiz

Hvilke af nedenstående udsagn er rigtige?

Determinanten af en ortogonal matrix kan være alle tal $\neq 0$.

Determinanten af en ortogonal matrix er enten 1 eller -1.

Matricen, som repræsenterer en rotation omkring en akse gennem 0 i $\mathbb{R}^3$ med hensyn til den naturlige basis er en ortogonal matrix.

Matricen

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{pmatrix}$$

er ortogonal for en vinkel θ. ♠

7.3 Anvendelser af ortogonalitet

At have et prikprodukt giver sig specielt udslag i udregningen af baser for underrum. Pæne baser består af ortogonale vektorer.

Den første indikation på at ortogonale vektorer opfører sig specielt pænt kommer her.

(7.20) PROPOSITION.

Hvis $\mathbf{v}_1, \dots, \mathbf{v}_m \in \mathbb{C}^n$ *er vektorer med* $\mathbf{v}_i \neq 0$ *og* $\mathbf{v}_i \perp \mathbf{v}_j$ *for* $i \neq j$ *og* $i, j = 1, \dots, m$*, så er* $\mathbf{v}_1, \dots, \mathbf{v}_m$ *lineært uafhængige det vil sige hvis*

$$\lambda_1 \mathbf{v}_1 + \cdots + \lambda_m \mathbf{v}_m = 0$$

for $\lambda_1, \dots, \lambda_m \in \mathbb{C}$ *så medfører det* $\lambda_1 = \cdots = \lambda_m = 0$.

Bevis

Bevis. Lad

$$\mathbf{v} = \lambda_1 \mathbf{v}_1 + \cdots + \lambda_m \mathbf{v}_m.$$

Så vil

$$\begin{aligned} \mathbf{v} \cdot \mathbf{v}_i &= (\lambda_1 \mathbf{v}_1 + \cdots + \lambda_m \mathbf{v}_m) \cdot \mathbf{v}_i \\ &= \lambda_1 (\mathbf{v}_1 \cdot \mathbf{v}_i) + \cdots + \lambda_i (\mathbf{v}_i \cdot \mathbf{v}_i) + \cdots + \lambda_n (\mathbf{v}_n \cdot \mathbf{v}_i) \\ &= \lambda_i (\mathbf{v}_i \cdot \mathbf{v}_i) = \lambda_i |\mathbf{v}_i|^2 = 0, \end{aligned}$$

hvilket medfører at $\lambda_i = 0$ for alle $i = 1, \dots, m$, fordi $\mathbf{v}_i \neq \mathbf{0}$. □

♠

7.3.1 Ortogonal- og ortonormalbaser

(7.21) DEFINITION.

Lad V *være et underrum i* $\mathbb{C}^n$*. En basis* $(\mathbf{u}_1, \dots, \mathbf{u}_m)$ *for* V *kaldes en* ortogonalbasis*, hvis* $\mathbf{u}_i \cdot \mathbf{u}_j = 0$ *for* $i \neq j$*. Hvis alle vektorerne i en ortogonalbasis er enhedsvektorer kaldes basen en* ortonormalbasis.

Et naturligt eksempel på en ortonormalbasis for $\mathbb{R}^n$ består af standardbasisvektorerne

$$\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix},$$

men der er masser af andre spændende ortonormalbaser.

Nedenstående resultat viser at en ortonormalbasis opfører sig ligesom den kanoniske basis i to vigtige henseender.

(7.22) PROPOSITION.

Lad $\mathbf{e}_1, \dots, \mathbf{e}_m$ *være en ortonormal basis for et underrum* V *og lad* $\mathbf{v} \in V$*. Så er*

1.
$$\mathbf{v} = (\mathbf{v} \cdot \mathbf{e}_1)\mathbf{e}_1 + \cdots + (\mathbf{v} \cdot \mathbf{e}_m)\mathbf{e}_m.$$

2.
$$|\mathbf{v}|^2 = |\mathbf{v} \cdot \mathbf{e}_1|^2 + \cdots + |\mathbf{v} \cdot \mathbf{e}_m|^2$$

Bevis

Bevis. Vi ved at
$$\mathbf{v} = x_1\mathbf{e}_1 + \cdots + x_m\mathbf{e}_m$$
for entydigt bestemte x_i. Nu følger

$$\mathbf{v} \cdot \mathbf{e}_i = (x_1\mathbf{e}_1 + \cdots + x_m\mathbf{e}_m) \cdot \mathbf{e}_i = x_1(\mathbf{e}_1 \cdot \mathbf{e}_i) + \cdots + x_i(\mathbf{e}_i \cdot \mathbf{e}_i) + \cdots + x_m(\mathbf{e}_m \cdot \mathbf{e}_i) = x_i,$$

hvilket viser første påstand, fordi $\mathbf{e}_i \cdot \mathbf{e}_j = 0$ for $i \neq j$ og $\mathbf{e}_i \cdot \mathbf{e}_i = 1$.

Den anden påstand følger af regneregler for prikproduktet:

$$|\mathbf{v}|^2 = \mathbf{v} \cdot \mathbf{v} = (x_1\mathbf{e}_1 + \cdots + x_m\mathbf{e}_m) \cdot (x_1\mathbf{e}_1 + \cdots + x_m\mathbf{e}_m) = x_1^2 + \cdots + x_m^2,$$

hvor vi på samme måde har benyttet at $\mathbf{e}_i \cdot \mathbf{e}_j = 0$ for $i \neq j$ og $\mathbf{e}_i \cdot \mathbf{e}_i = 1$. $\square$

♠

7.23 Eksempel

De tre vektorer

$$\mathbf{v}_1 = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}, \quad \text{og} \quad \mathbf{v}_3 = \begin{pmatrix} -\frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$$

udgør en ortonormal basis for deres span $V = \text{span}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ i $\mathbb{R}^4$. Det er nu under brug af Proposition 22 specielt nemt at afgøre om en forelagt vektor $\mathbf{u} \in \mathbb{R}^4$ ligger i V. Dette sker hvis og kun hvis

$$\mathbf{u} = (\mathbf{u} \cdot \mathbf{v}_1)\mathbf{v}_1 + (\mathbf{u} \cdot \mathbf{v}_2)\mathbf{v}_2 + (\mathbf{u} \cdot \mathbf{v}_3)\mathbf{v}_3.$$

Hvis for eksempel

$$\mathbf{u} = \begin{pmatrix} 1 \\ 1 \\ 2 \\ 1 \end{pmatrix}$$

er

$$(\mathbf{u} \cdot \mathbf{v}_1)\mathbf{v}_1 + (\mathbf{u} \cdot \mathbf{v}_2)\mathbf{v}_2 + (\mathbf{u} \cdot \mathbf{v}_3)\mathbf{v}_3 = \tfrac{5}{2}\,\mathbf{v}_2 + \tfrac{1}{2}\,\mathbf{v}_3 = \begin{pmatrix} 1 \\ 1 \\ \frac{3}{2} \\ \frac{3}{2} \end{pmatrix}$$

og det fremgår at $\mathbf{u} \notin V$. Derimod har vi med

$$\mathbf{u}' = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}$$

at

$$\mathbf{u}' = (\mathbf{u}' \cdot \mathbf{v}_1)\mathbf{v}_1 + (\mathbf{u}' \cdot \mathbf{v}_2)\mathbf{v}_2 + (\mathbf{u}' \cdot \mathbf{v}_3)\mathbf{v}_3 = \mathbf{v}_2 + \mathbf{v}_3$$

det vil sige $\mathbf{u}' \in V$. ♠

7.24 Quiz

Lad V være underrummet i $\mathbb{R}^3$ med en basis B bestående af vektorerne

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \quad \text{og} \quad \mathbf{v}_2 = \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix}.$$

Hvilke af nedenstående udsagn er rigtigt?

B er en ortogonalbasis for V.

B er en ortonormalbasis for V.

$V = \mathbb{R}^3$.

$$\left(\frac{1}{\sqrt{6}}\mathbf{v}_1, \frac{1}{\sqrt{3}}\mathbf{v}_2\right)$$

er en ortonormalbasis for V.

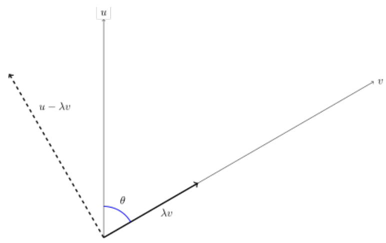
$$\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \in V.$$

♠

7.3.2 Gram-Schmidt algoritmen

Vi har set ovenfor at ortonormalbaser for underrum er specielt nemme at regne med. Spørgsmålet er om alle underrum har en ortonormalbasis. Her er svaret ja og grunden ligger i en klassisk og smuk algoritme tilmed opfundet af en dansker. Ideen kommer fra formlen for projektion af en vektor på en anden i planen.

Lad os kaste et blik igen på tegningen



som forekom i Kapitel 1 i forbindelse med formlen for projektion af en vektor på en anden vektor. Givet to vektorer $\mathbf{u}, \mathbf{v}$ med $\mathbf{v} \neq 0$ fandt vi ud af at hvis et tal λ opfylder

$$\mathbf{u} - \lambda\mathbf{v} \perp \mathbf{v}, \tag{7.7}$$

så må

$$\lambda = \frac{\mathbf{u} \cdot \mathbf{v}}{|\mathbf{v}|^2}.$$

Argumentet for dette var løsning af en førstegradsligning! Her er førstegradsligningen, som kommer fra betingelsen (7.7) og hvis løsning er ovenstående λ:

$$(\mathbf{u} - \lambda\mathbf{v}) \cdot \mathbf{v} = (\mathbf{u} \cdot \mathbf{v}) - \lambda(\mathbf{v} \cdot \mathbf{v}) = (\mathbf{u} \cdot \mathbf{v}) - \lambda|\mathbf{v}|^2 = 0.$$

Læg mærke til at disse udregninger er fuldstændigt uafhængige af om vi regner i de reelle eller komplekse tal.

Teknikken ovenfor kan benyttes til at omforme to vektorer $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{C}^n$ til to vektorer

$$\begin{aligned}
\mathbf{u}_1 &= \mathbf{v}_1 \\
\mathbf{u}_2 &= \mathbf{v}_2 - \frac{\mathbf{v}_1 \cdot \mathbf{v}_2}{|\mathbf{v}_1|^2}\mathbf{v}_1
\end{aligned}$$

så $\mathbf{u}_1$ er ortogonal på $\mathbf{u}_2$ samt

$$\operatorname{span}(\mathbf{v}_1, \mathbf{v}_2) = \operatorname{span}(\mathbf{u}_1, \mathbf{u}_2).$$

Man kan endda gå et skridt videre med tre eller flere vektorer: Lad os antage for tre vektorer $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ at $\mathbf{v}_1$ og $\mathbf{v}_2$ allerede er ortogonale det vil sige $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$.
Kan vi erstatte $\mathbf{v}_3$ med en vektor $\mathbf{u}_3$ i $\operatorname{span}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ så $\mathbf{u}_3$ bliver ortogonal på $\mathbf{v}_1$ og $\mathbf{v}_2$? Med input fra tilfældet med to vektorer viser formlen sig at blive

$$\mathbf{u}_3 = \mathbf{v}_3 - \frac{\mathbf{v}_3 \cdot \mathbf{v}_1}{|\mathbf{v}_1|^2}\mathbf{v}_1 - \frac{\mathbf{v}_3 \cdot \mathbf{v}_2}{|\mathbf{v}_2|^2}\mathbf{v}_2.$$

Det overlades som en øvelse at indse at $\mathbf{u}_3 \cdot \mathbf{v}_1 = 0$ og $\mathbf{u}_3 \cdot \mathbf{v}_2 = 0$.

Ovenstående procedure kan generaliseres fra to og tre vektorer til et vilkårligt antal vektorer. Denne generalisering blev først fundet af danskeren Jørgen Pedersen Grami 1883 og kendes i dag under navnet Gram-Schmidt algoritmen. Algoritmen er angivet i sætningen nedenfor. Det er en af de helt fundamentale algoritmer i lineær algebra.

(7.25) SÆTNING.

Lad $\mathbf{v}_1, \dots, \mathbf{v}_m$ være lineært uafhængige vektorer i $\mathbb{C}^n$ og lad

$$V = \operatorname{span}(\mathbf{v}_1, \dots, \mathbf{v}_m).$$

Ved algoritmen

$$\begin{aligned}
\mathbf{u}_1 &= \mathbf{v}_1 \\
\mathbf{u}_2 &= \mathbf{v}_2 - \frac{\mathbf{v}_2 \cdot \mathbf{u}_1}{|\mathbf{u}_1|^2}\mathbf{u}_1 \\
\mathbf{u}_3 &= \mathbf{v}_3 - \frac{\mathbf{v}_3 \cdot \mathbf{u}_1}{|\mathbf{u}_1|^2}\mathbf{u}_1 - \frac{\mathbf{v}_3 \cdot \mathbf{u}_2}{|\mathbf{u}_2|^2}\mathbf{u}_2 \\
&\vdots \\
\mathbf{u}_m &= \mathbf{v}_m - \frac{\mathbf{v}_m \cdot \mathbf{u}_1}{|\mathbf{u}_1|^2}\mathbf{u}_1 - \cdots - \frac{\mathbf{v}_m \cdot \mathbf{u}_{m-1}}{|\mathbf{u}_{m-1}|^2}\mathbf{u}_{m-1}
\end{aligned}$$

opnås ortogonale vektorer $\mathbf{u}_1, \dots, \mathbf{u}_m \in V$ så at $\mathbf{u}_i \neq \mathbf{0}$ og

$$V = \operatorname{span}(\mathbf{u}_1, \dots, \mathbf{u}_m).$$

Således har ethvert underrum en ortogonalbasis og (dermed) en ortonormalbasis.

Bevis

Bevis. Undervejs i algoritmen viser man ved ren og skær udregning med prikproduktet at den nyligt konstruerede vektor $\mathbf{u}_i$ opfylder $\mathbf{u}_i \cdot \mathbf{u}_1 = 0, \dots, \mathbf{u}_i \cdot \mathbf{u}_{i-1} = 0$ samt

$$\operatorname{span}(\mathbf{v}_1, \dots, \mathbf{v}_i) = \operatorname{span}(\mathbf{u}_1, \dots, \mathbf{u}_i).$$

At $\mathbf{u}_i = 0$ ikke kan forekomme er på grund af antagelsen om lineær uafhængighed blandt vektorerne $\mathbf{v}_1, \dots, \mathbf{v}_m$. Hvis $\mathbf{u}_i = 0$ ville

$$\mathbf{v}_i \in \operatorname{span}(\mathbf{u}_1, \dots, \mathbf{u}_{i-1}) = \operatorname{span}(\mathbf{v}_1, \dots, \mathbf{v}_{i-1}),$$

i strid med at vektorerne $\mathbf{v}_1, \dots, \mathbf{v}_i$ er lineært uafhængige. □

♠

Hvis Sætning 25 benyttes på et sæt af vektorer, som ikke er lineært uafhængige, vil algoritmen undervejs afsløre dette og give $\mathbf{u}_i = 0$, hvor

$$\mathbf{v}_i \in \operatorname{span}(\mathbf{v}_1, \dots, \mathbf{v}_{i-1}).$$

Algoritmen kan modificeres ret enkelt ved at springe trin med $\mathbf{u}_i = 0$ over og arbejde videre med $\mathbf{v}_{i+1}$ ud fra de allerede fundne ortogonale vektorer $\mathbf{u}_1, \dots, \mathbf{u}_{i-1}$.

En anden enkel modifikation er at normalisere vektorerne undervejs til enhedsvektorer ved at udskifte $\mathbf{u}_i$ med $(1/|\mathbf{u}_i|)\mathbf{u}_i$. Dette leder frem til en ortonormalbasis. Man kan også normalisere de ortogonale vektorer til sidst som i eksemplet nedenfor.

VIDEO: https://youtu.be/iyzRFkOqbws

7.26 Eksempel

Vi betragter underrummet V i $\mathbb{R}^5$ med basis B bestående af vektorerne

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 6 \\ 2 \\ 3 \\ 0 \\ 4 \end{pmatrix} \quad \text{og} \quad \mathbf{v}_3 = \begin{pmatrix} 5 \\ 3 \\ 3 \\ 1 \\ 5 \end{pmatrix}$$

og vil benytte Gram-Schmidt algoritmen til at finde en ortonormalbasis for V. Vi indleder med at finde en ortogonalbasis for V. Første trin er $\mathbf{u}_1 = \mathbf{v}_1$. Dernæst udregner vi

$$\mathbf{u}_2 = \mathbf{v}_2 - \frac{\mathbf{v}_2 \cdot \mathbf{u}_1}{|\mathbf{u}_1|^2}\mathbf{u}_1 = \begin{pmatrix} 6 \\ 2 \\ 3 \\ 0 \\ 4 \end{pmatrix} - 2 \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 4 \\ 3 \\ -2 \\ 2 \end{pmatrix}$$

Så vidt så godt. Man checker at $\mathbf{u}_1 \cdot \mathbf{u}_2 = 0$. Sidste skridt er nu udregningen af $\mathbf{u}_3$ via

$$\mathbf{u}_3 = \mathbf{v}_3 - \frac{\mathbf{v}_3 \cdot \mathbf{u}_1}{|\mathbf{u}_1|^2}\mathbf{u}_1 - \frac{\mathbf{v}_3 \cdot \mathbf{u}_2}{|\mathbf{u}_2|^2}\mathbf{u}_2$$

$$= \begin{pmatrix} 5 \\ 3 \\ 3 \\ 1 \\ 5 \end{pmatrix} - 2 \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 4 \\ 4 \\ 3 \\ -2 \\ 2 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}.$$

Hermed er $(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$ en ortogonalbasis for V. Da $|\mathbf{u}_1| = 2, |\mathbf{u}_2| = 7$ og $|\mathbf{u}_3| = 2$ vil

$$\left(\tfrac{1}{2}\mathbf{u}_1, \tfrac{1}{7}\mathbf{u}_2, \tfrac{1}{2}\mathbf{u}_3\right)$$

være en ortonormalbasis for V. ♠

7.3.3 Den modificerede Gram-Schmidt algoritme

Gram-Schmidt algoritmen som angivet ovenfor er numerisk ustabil i praksis. Ved en lille modifikation med hensyn til udregningen af $\mathbf{u}_i$ fås en numerisk stabil algoritme, som er mindre følsom overfor afrundingsfejl. Dette modificerede trin består i at udregne $\mathbf{u}_i$ gennem følgende kæde af operationer:

$$\begin{aligned} \mathbf{u}_i^{(1)} &= \mathbf{v}_i - (\mathbf{v}_i \cdot \mathbf{u}_1)\mathbf{u}_1 \\ \mathbf{u}_i^{(2)} &= \mathbf{u}_i^{(1)} - (\mathbf{u}_i^{(1)} \cdot \mathbf{u}_2)\mathbf{u}_2 \\ &\vdots \\ \mathbf{u}_i^{(i-1)} &= \mathbf{u}_i^{(i-2)} - (\mathbf{u}_i^{(i-2)} \cdot \mathbf{u}_{i-1})\mathbf{u}_{i-1} \end{aligned}$$

med $\mathbf{u}_i = \frac{\mathbf{u}_i^{(i-1)}}{|\mathbf{u}_i^{(i-1)}|}$ som resultat.

7.27 Eksempel

Lad os benytte vektorerne

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 6 \\ 2 \\ 3 \\ 0 \\ 4 \end{pmatrix} \quad \text{og} \quad \mathbf{v}_3 = \begin{pmatrix} 5 \\ 3 \\ 3 \\ 1 \\ 5 \end{pmatrix}$$

fra sidste eksempel som input til den modificerede Gram-Schmidt algoritme. Første skridt giver

$$\mathbf{u}_1 = \mathbf{v}_1/|\mathbf{v}_1| = \frac{1}{2}\begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \\ 1 \end{pmatrix}.$$

Udregningen af $\mathbf{u}_2$ foregår som

$$\mathbf{u}_2^{(1)} = \mathbf{v}_2 - (\mathbf{v}_2 \cdot \mathbf{u}_1)\mathbf{u}_1 = \begin{pmatrix} 6 \\ 2 \\ 3 \\ 0 \\ 4 \end{pmatrix} - 4 \cdot \frac{1}{2}\begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 4 \\ 3 \\ -2 \\ 2 \end{pmatrix}$$

med

$$\mathbf{u}_2 = \frac{1}{7}\begin{pmatrix} 4 \\ 4 \\ 3 \\ -2 \\ 2 \end{pmatrix}.$$

Udregningen af $\mathbf{u}_3$ foregår som

$$\mathbf{u}_3^{(1)} = \mathbf{v}_3 - (\mathbf{v}_3 \cdot \mathbf{u}_1)\mathbf{u}_1 = \begin{pmatrix} 5 \\ 3 \\ 3 \\ 1 \\ 5 \end{pmatrix} - 4 \cdot \frac{1}{2}\begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 5 \\ 3 \\ -1 \\ 3 \end{pmatrix}$$

og

$$\mathbf{u}_3^{(2)} = \mathbf{u}_3^{(1)} - (\mathbf{u}_3^{(1)} \cdot \mathbf{u}_2)\mathbf{u}_2 = \begin{pmatrix} 3 \\ 5 \\ 3 \\ -1 \\ 3 \end{pmatrix} - 7 \cdot \frac{1}{7}\begin{pmatrix} 4 \\ 4 \\ 3 \\ -2 \\ 2 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}$$

med endeligt resultat

$$\mathbf{u}_3 = \frac{1}{2}\begin{pmatrix} -1 \\ 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}$$

i fin overensstemmelse med Eksempel 7.3.2. ♠

Denne algoritme kaldes den modificerede Gram-Schmidt algoritme. Den numeriske stabilitet illustreres i eksemplet (hentet fra MIT OpenCourseWare) nedenfor.

(7.28) EKSEMPEL.

For at uddybe hvad der egentlig menes med numerisk stabil eller mindre følsom overfor afrundingsfejl kan man som eksempel afprøve Gram-Schmidt algoritmen og den modificerede Gram-Schmidt algoritme på vektorerne

$$\begin{pmatrix} 1 \\ \varepsilon \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 0 \\ \varepsilon \\ 0 \end{pmatrix} \quad \text{og} \quad \begin{pmatrix} 1 \\ 0 \\ 0 \\ \varepsilon \end{pmatrix}$$

med afrundingsfejlen $1+\varepsilon^2 \approx 1$. Hvis en lommeregner for eksempel kan vise 10 cifre i displayet og $\varepsilon = 0.000001$ så er $1+\varepsilon^2 = 1$ på lommeregneren.

Med afrunding og normering i hvert trin giver den klassiske Grams-Schmidt algoritme resultatet

$$\mathbf{u}_1 = \begin{pmatrix} 1 \\ \varepsilon \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{u}_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \end{pmatrix} \quad \text{og} \quad \mathbf{u}_3 = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ -1 \\ 0 \\ 1 \end{pmatrix}.$$

Læg mærke til at $\mathbf{u}_2 \cdot \mathbf{u}_3 = \frac{1}{2}$, hvilket er en grim fejl som følge af afrundingen.

Derimod giver den modificerede Grams-Schmidt algoritme resultatet

$$\mathbf{u}_1 = \begin{pmatrix} 1 \\ \varepsilon \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{u}_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \end{pmatrix} \quad \text{og} \quad \mathbf{u}_3 = \frac{1}{\sqrt{6}}\begin{pmatrix} 0 \\ -1 \\ -1 \\ 2 \end{pmatrix}.$$

Her er $\mathbf{u}_2 \cdot \mathbf{u}_3 = 0$.

7.3.4 QR dekomposition

Hvis vi benytter Gram-Schmidt algoritmen på søjlerne $\mathbf{a}_1, \ldots, \mathbf{a}_n$ i en invertibel $n \times n$ matrix A fås

$$\begin{aligned}
\mathbf{u}_1 &= \mathbf{a}_1 \\
\mathbf{u}_2 &= \mathbf{a}_2 - \frac{\mathbf{a}_2 \cdot \mathbf{u}_1}{|\mathbf{u}_1|^2}\mathbf{u}_1 \\
\mathbf{u}_3 &= \mathbf{a}_3 - \frac{\mathbf{a}_3 \cdot \mathbf{u}_1}{|\mathbf{u}_1|^2}\mathbf{u}_1 - \frac{\mathbf{a}_3 \cdot \mathbf{u}_2}{|\mathbf{u}_2|^2}\mathbf{u}_2 \\
&\vdots \\
\mathbf{u}_n &= \mathbf{a}_n - \frac{\mathbf{a}_n \cdot \mathbf{u}_1}{|\mathbf{u}_1|^2}\mathbf{u}_1 - \cdots - \frac{\mathbf{a}_n \cdot \mathbf{u}_{n-1}}{|\mathbf{u}_{n-1}|^2}\mathbf{u}_{n-1}.
\end{aligned}$$

Ved at ortonormalisere basen og sætte $\mathbf{e}_i = (1/|\mathbf{u}_i|)\mathbf{u}_i$ fås nu

$$\mathbf{a}_1 = (\mathbf{a}_1 \cdot \mathbf{e}_1)\mathbf{e}_1 \tag{7.8}$$

$$\mathbf{a}_2 = (\mathbf{a}_2 \cdot \mathbf{e}_1)\mathbf{e}_1 + (\mathbf{a}_2 \cdot \mathbf{e}_2)\mathbf{e}_2 \tag{7.9}$$

$$\mathbf{a}_3 = (\mathbf{a}_3 \cdot \mathbf{e}_1)\mathbf{e}_1 + (\mathbf{a}_3 \cdot \mathbf{e}_2)\mathbf{e}_2 + (\mathbf{a}_3 \cdot \mathbf{e}_3)\mathbf{e}_3 \tag{7.10}$$

$$\vdots \tag{7.11}$$

$$\mathbf{a}_n = (\mathbf{a}_n \cdot \mathbf{e}_1)\mathbf{e}_1 + (\mathbf{a}_n \cdot \mathbf{e}_2)\mathbf{e}_2 + \cdots + (\mathbf{a}_n \cdot \mathbf{e}_{n-1})\mathbf{e}_{n-1}. \tag{7.12}$$

Oversat giver det matrix identiteten

$$A = QR,$$

hvor Q er matricen med søjler $\mathbf{e}_1, \ldots, \mathbf{e}_n$ og

$$R = \begin{pmatrix} (\mathbf{a}_1 \cdot \mathbf{e}_1) & (\mathbf{a}_2 \cdot \mathbf{e}_1) & \cdots & (\mathbf{a}_n \cdot \mathbf{e}_1) \\ 0 & (\mathbf{a}_2 \cdot \mathbf{e}_2) & \cdots & (\mathbf{a}_n \cdot \mathbf{e}_2) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & (\mathbf{a}_n \cdot \mathbf{e}_n) \end{pmatrix}.$$

Læg mærke til at matricen Q er ortogonal fordi dens søjler udgør en ortonormal basis samt at R er en øvre trekantsmatrix (den har nuller under diagonalen). Vi har vist følgende.

(7.29) SÆTNING.

En invertibel matrix kan skrives som produkt af en ortogonal matrix og en øvre trekantsmatrix.

En faktorisering af en matrix $A = QR$ som produkt af en ortogonal matrix Q og en øvre trekantsmatrix R kaldes en QR dekomposition af A.

7.30 Eksempel

Matricen

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

er en invertibel 2×2 matrix. Ved hjælp af Gram-Schmidt algoritmen finder man med input fra søjlerne i A matricen Q indeholdende den ortonormale basis

$$Q = \begin{pmatrix} \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix}$$

med QR dekompositionen

$$A = \begin{pmatrix} \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} \sqrt{5} & \frac{4}{\sqrt{5}} \\ 0 & \frac{3}{\sqrt{5}} \end{pmatrix}$$

som følge. ♠

7.3.5 Den mirakuløse QR-algoritme

Det meste lineære algebra er langtidsholdbart matematik og flere hundrede år gammelt. Det hænder dog at ekstremt vigtige nye opdagelser bliver gjort for eksempel ved at eksperimentere med computere. Følgende næsten halvnaive algoritme til at udregne egenværdier for en kvadratisk matrix A blev opdaget sidst i 1950'erne. Den hedder QR algoritmen og bygger netop på QR dekompositionen.

Indledningsvis sættes $A_0 = A$ og algoritmen udregner nye QR dekompositioner i hvert trin med hensyn til det modsatte produkt af Q og R fra foregående trin:

$$\begin{aligned} A_0 &= Q_0 R_0 \\ A_1 &= R_0 Q_0 = Q_1 R_1 \\ A_2 &= R_1 Q_1 = Q_2 R_2 \\ &\vdots \end{aligned}$$

Læg her mærke til at A_i og A_{i+1} har de samme egenværdier, fordi de er "konjugerede":

$$A_{i+1} = R_i Q_i = (Q_i^{-1} Q_i)(R_i Q_i) = Q_i^{-1}(Q_i R_i) Q_i = Q_i^{-1} A_i Q_i$$

Her har vi brugt at vi ved at den ortogonale matrix Q_i er inverterbar. Hvorfra ved vi forresten det? Jo, fordi $Q_i Q_i^T = Q_i^T Q = I_n$, så $Q_i^{-1} = Q_i^T$. Under alle omstændigheder, hvis $\mathbf{u}$ er en egenvektor til A_i med egenværdi λ, så er $Q_i^{-1}\mathbf{u}$ en egenvektor til A_{i+1} med den samme egenværdi λ.

Oftest vil diagonalelementerne i R_n konvergere mod egenværdierne i den oprindelige matrix A.

(7.31) EKSEMPEL.

Betragt matricen

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

fra Eksempel 7.3.4. Man kan ret nemt regne ud at A har egenværdierne $\lambda = 1$ og $\lambda = 3$.

Lad os afprøve QR algoritmen rent numerisk på A. De første trin giver

$$\begin{aligned}
A_0 = Q_0R_0 &= \begin{pmatrix} 0.894427 & -0.447214 \\ 0.447214 & 0.894427 \end{pmatrix} \begin{pmatrix} 2.23607 & 1.78885 \\ 0 & 1.34164 \end{pmatrix} \\
A_1 = R_0Q_0 = Q_1R_1 &= \begin{pmatrix} 0.977802 & 0.209529 \\ 0.209529 & 0.977802 \end{pmatrix} \begin{pmatrix} 2.86356 & 0.83812 \\ 0 & 1.04765 \end{pmatrix} \\
A_2 = R_1Q_1 = Q_2R_2 &= \begin{pmatrix} 0.99729 & -0.0735706 \\ 0.0735706 & 0.99729 \end{pmatrix} \begin{pmatrix} 2.9837 & 0.29428 \\ 0 & 1.00546 \end{pmatrix} \\
A_3 = R_2Q_2 = Q_3R_3 &= \begin{pmatrix} 0.999696 & -0.0246726 \\ 0.0246726 & 0.999696 \end{pmatrix} \begin{pmatrix} 2.99817 & 0.098690 \\ 0 & 1.00061 \end{pmatrix}.
\end{aligned}$$

Eksperimentet synes at bekræfte at diagonalelementerne (markeret med rødt) i følgen af de øvrige trekantsmatricer i QR dekompositionerne konvergerer mod egenværdierne af den oprindelige matrix.

7.3.6 Ortogonalkomplement og ortogonalprojektion

Man har ofte brug for en anelse mere terminologi omkring underrum og ortogonalitet, herunder ortogonalkomplement og ortogonalprojektion.

(7.32) DEFINITION.

For et underrum V i $\mathbb{R}^n$ knytter der sig et komplementært underrum med hensyn til prikproduktet. Dette underrum er defineret ved

$$V^\perp = \{\mathbf{v} \in \mathbb{R}^n \mid \mathbf{u} \cdot \mathbf{v} = 0, \textit{for alle } \mathbf{u} \in V\}.$$

og kaldes ortogonalkomplementet *til V.*

At $V^\perp$ er et underrum følger af egenskaberne ved prikproduktet givet i Proposition 4.

(7.33) BEMÆRKNING.

Et underrum V i $\mathbb{R}^n$ er givet ved en basis $(\mathbf{u}_1, \dots, \mathbf{u}_m)$. Ud fra denne basis kan en basis for ortogonalkomplementet $V^\perp$ udregnes som en basis for $N(A^T)$, hvor A er matricen med søjler $\mathbf{u}_1, \dots, \mathbf{u}_m$.

Dette følger af observationen at $\mathbf{v} \in \mathbb{R}^n$ opfylder $\mathbf{u} \cdot \mathbf{v} = 0$ for alle $\mathbf{u} \in V$ hvis og kun hvis

$$\mathbf{u}_1 \cdot \mathbf{v} = 0 \tag{7.13}$$

$$\vdots \tag{7.14}$$

$$\mathbf{u}_m \cdot \mathbf{v} = 0, \tag{7.15}$$

fordi alle $\mathbf{u} \in V$ kan skrives $\mathbf{u} = \lambda_1 \mathbf{u}_1 + \cdots + \lambda_m \mathbf{u}_m$ og

$$\mathbf{u} \cdot \mathbf{v} = (\lambda_1 \mathbf{u}_1 + \cdots + \lambda_m \mathbf{u}_m) \cdot \mathbf{v} = \lambda_1 (\mathbf{u}_1 \cdot \mathbf{v}) + \cdots + \lambda_m (\mathbf{u}_m \cdot \mathbf{v}) = \mathbf{0},$$

hvis $\mathbf{v}$ er en løsning i (7.13).

Skrevet i matrixnotation svarer ligningssystemet (7.13) præcis til $A^T \mathbf{v} =< nul$. Derfor er $V^\perp = N(A^T)$.

(7.34) PROPOSITION.

Lad V være et underrum af $\mathbb{R}^n$. Enhver vektor $\mathbf{u} \in \mathbb{R}^n$ kan på entydig måde skrives

$$\mathbf{u} = \mathbf{v} + \mathbf{v}', \tag{7.16}$$

hvor $\mathbf{v} \in V$ og $\mathbf{v}' \in V^\perp$: Hvis $\mathbf{v}_1, \mathbf{v}_2 \in V$ og $\mathbf{v}'_1, \mathbf{v}'_2 \in V^\perp$ med

$$\mathbf{v}_1 + \mathbf{v}'_1 = \mathbf{v}_2 + \mathbf{v}'_2,$$

så er $\mathbf{v}_1 = \mathbf{v}_2$ og $\mathbf{v}'_1 = \mathbf{v}'_2$.

Bevis

Bevis. Hvis $\mathbf{v}_1, \dots, \mathbf{v}_m$ er en ortonormal basis for V, så gælder

$$\mathbf{v}^\perp = \mathbf{u} - ((\mathbf{u} \cdot \mathbf{v}_1)\mathbf{v}_1 + \cdots + (\mathbf{u} \cdot \mathbf{v}_m)\mathbf{v}_m) \in V^\perp.$$

Derfor gælder

$$\mathbf{u} = \mathbf{v} + \mathbf{v}^\perp \tag{7.17}$$

med

$$\mathbf{v} = (\mathbf{u} \cdot \mathbf{v}_1)\mathbf{v}_1 + \cdots + (\mathbf{u} \cdot \mathbf{v}_m)\mathbf{v}_m) \in V.$$

Fremstillingen i (7.17) er entydig. Hvis $\mathbf{v}_1 + \mathbf{v}'_1 = \mathbf{v}_2 + \mathbf{v}'_2$ så er $\mathbf{v}_1 - \mathbf{v}_2 = \mathbf{v}'_1 - \mathbf{v}'_2 \in V^\perp$. Men da $\mathbf{v}_1 - \mathbf{v}_2 \in V$, så er $|\mathbf{v}_1 - \mathbf{v}_2|^2 = (\mathbf{v}_1 - \mathbf{v}_2) \cdot (\mathbf{v}_1 - \mathbf{v}_2) = 0$. Det følger at $\mathbf{v}_1 = \mathbf{v}_2$. Formuleret på en anden måde er $V \cap V^\perp = \{< nul\}$. $\square$

♠

(7.35) DEFINITION.

I opskrivningen (7.16) *kaldes* $\mathbf{v}$ *ortogonalprojektionen af* $\mathbf{u}$ *på* V.

Begreberne kan endnu en gang forstås gennem den centrale figur nedenfor i tilfældet $\mathbb{R}^2$.

Med $V = \operatorname{span}(\mathbf{v})$ er $\lambda\mathbf{v}$ ortogonalprojektion af $\mathbf{u}$ på V og ortogonalkomplementet $V^\perp$ til V er underrummet $\operatorname{span}(\mathbf{u} - \lambda\mathbf{v})$.

(7.36) BEMÆRKNING.

Ortogonalprojektion af en vektor $\mathbf{u} \in \mathbb{R}^n$ på et underrum V fås ved hjælp af en ortonormalbasis for $\mathbf{v}_1, \ldots, \mathbf{v}_m$ for V som

$$\mathbf{v} = (\mathbf{u} \cdot \mathbf{v}_1)\mathbf{v}_1 + \cdots + (\mathbf{u} \cdot \mathbf{v}_m)\mathbf{v}_m.$$

Da $\mathbf{u} - \mathbf{v} \in V^\perp$, fordi $(\mathbf{u} - \mathbf{v}) \cdot \mathbf{v}_1 = 0, \ldots, (\mathbf{u} - \mathbf{v}) \cdot \mathbf{v}_m = 0$, bliver den entydige opskrivning fra (7.16) hermed

$$\mathbf{u} = \mathbf{v} + (\mathbf{u} - \mathbf{v}).$$

(7.37) EKSEMPEL.

I Eksempel 7.3.1 betragtede vi underrummet $V = \operatorname{span}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ af $\mathbb{R}^4$ med

$$\mathbf{v}_1 = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}, \quad \text{og} \quad \mathbf{v}_3 = \begin{pmatrix} -\frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}.$$

Bemærk igen at $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ er en ortonormalbasis for V. Her er

$$V^\perp = \operatorname{span}\left(\begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix}\right),$$

da vi ved at $\dim V^{\perp} = 4 - \dim V = 1$ (hvorfor?). I Eksempel 7.3.1 fandt vi ud af at $\mathbf{u} \notin V$, hvor

$$\mathbf{u} = \begin{pmatrix} 1 \\ 1 \\ 2 \\ 1 \end{pmatrix}.$$

Vi udregnede i eksemplet også

$$(\mathbf{u} \cdot \mathbf{v}_1)\mathbf{v}_1 + (\mathbf{u} \cdot \mathbf{v}_2)\mathbf{v}_2 + (\mathbf{u} \cdot \mathbf{v}_3)\mathbf{v}_3 = \tfrac{5}{2}\mathbf{v}_2 + \tfrac{1}{2}\mathbf{v}_3 = \begin{pmatrix} 1 \\ 1 \\ \frac{3}{2} \\ \frac{3}{2} \end{pmatrix}$$

Denne vektor er ortogonalprojektionen af $\mathbf{u}$ på V i henhold til Bemærkning 36.

7.4 Mindste kvadraters metode

VIDEO: https://youtu.be/P8vqgYK3bic

Mindste kvadraters metode kendes nok bedst fra problemet om at bestemme den bedste rette linje gennem nogle punkter i planen. Faktisk er den langt mere generel og drejer sig om at finde approksimative løsninger til overbestemte ligningssystemer. Metoden blev opfundet af Gauss og brugt første gang i forbindelse med bestemmelse af baner for himmellegemer ud fra tre observationer.

Det er ikke altid et ligningssystem

$$A\mathbf{x} = \mathbf{b}$$

med A en $m \times n$ matrix og $\mathbf{b}$ en $m \times 1$ matrix har en løsning, og det er heller ikke altid at denne løsning er entydig. Hvis for eksempel $m \gg n$, det vil sige at der er mange flere ligninger end ubekendte, og hvis A og $\mathbf{b}$ er "tilfældige", så vil med overvældende sandsynlighed følgende to ting gælde:

1. Ligningen $A\mathbf{x} = \mathbf{b}$ har ikke nogen løsning.
2. Den homogene ligning $A\mathbf{x} = \mathbf{0}$ har kun den ene løsning $\mathbf{x} = \mathbf{0}$.

Det lyder jo ikke særlig lovende.

I denne situation kan vores forventning til at vi kan finde en løsning ligge på et meget lille sted, og det bedste vi kan gøre er at søge en vektor $\mathbf{x}$ så at $A\mathbf{x}$ kommer tættest muligt på $\mathbf{b}$.

7.38 Eksempel

Ligningssystemet

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix},$$

er et ligningssystem med 3 ligninger og 2 ubekendte. Det har ingen løsninger. ♠

For at præcisere hvad der menes med tættest på indføres følgende definition.

(7.39) DEFINITION.

En mindste kvadraters løsning *til ligningssystemet* $A\mathbf{x} = \mathbf{b}$*, med* A *en* $m \times n$ *matrix og* $\mathbf{b}$ *en* $m \times 1$*, er en vektor* $\mathbf{x}_0 \in \mathbb{R}^n$*, som opfylder*

$$|\mathbf{b} - A\mathbf{x}_0|^2 \leq |\mathbf{b} - A\mathbf{x}|^2$$

for alle $\mathbf{x} \in \mathbb{R}^n$*.*

Læg mærke til at en mindste kvadraters løsning opfylder at $|\mathbf{b} - A\mathbf{x}_0| \leq |\mathbf{b} - A\mathbf{x}|$ for alle $\mathbf{x} \in \mathbb{R}^n$. $\mathbf{x}_0$ er det valg af $\mathbf{x} = \mathbf{x}_0$ så at $A\mathbf{x}_0$ er tættest på $\mathbf{b}$ for alle mulige valg af $\mathbf{x}$.

Jeg bliver lige overrasket hver gang over at dette minimeringsproblem har så smuk en løsning ved hjælp af lineær algebra. En vigtig ingrediens er generaliseringen af Pythagoras til n dimensioner.

(7.40) PROPOSITION.

Lad $\mathbf{u}, \mathbf{v} \in \mathbb{C}^n$*. Hvis* $\mathbf{u}$ *og* $\mathbf{v}$ *er ortogonale, så er*

$$|\mathbf{u} + \mathbf{v}|^2 = |\mathbf{u}|^2 + |\mathbf{v}|^2.$$

Bevis

Bevis. Dette følger af egenskaber ved prikproduktet:

$$|\mathbf{u} + \mathbf{v}|^2 = (\mathbf{u} + \mathbf{v}) \cdot (\mathbf{u} + \mathbf{v}) = \mathbf{u} \cdot \mathbf{u} + \mathbf{u} \cdot \mathbf{v} + \mathbf{v} \cdot \mathbf{u} + \mathbf{v} \cdot \mathbf{v} = |\mathbf{u}|^2 + |\mathbf{v}|^2.$$

□

♠

(7.41) SÆTNING.

1. *Et ligningssystem $A\mathbf{x} = \mathbf{b}$, hvor A er en $m \times n$ matrix og $\mathbf{b}$ en $m \times 1$ matrix, har altid en mindste kvadraters løsning $\mathbf{x}_0 \in \mathbb{R}^n$. En løsning $\mathbf{x}_0$ kan findes som løsning til ligningssystemet*

$$A^T A\mathbf{x} = A^T\mathbf{b}.$$

2. *Antag nu at ligningen $A\mathbf{x} = \mathbf{0}$ ikke har andre løsninger end $\mathbf{x} = \mathbf{0}$. Da er der kun en eneste mindste kvadraters løsning. Desuden er matricen $A^T A$ invertibel, og den entydigt bestemte løsning kan udregnes ved formlen*

$$\mathbf{x}_0 = (A^T A)^{-1} A^T\mathbf{b}. \tag{7.18}$$

(7.42) BEMÆRKNING.

Betingelse at $A\mathbf{x} = \mathbf{0}$ kun har nulløsningen kan på en mere fancy måde udtrykkes ved at dimensionen af nulrummet $N(A)$ er 0. Ifølge dimensionssætningen i Kapitel 6 er dette det samme som at rangen af matricen A er n.

Bevis*

Bevis. Vi viser at der altid findes en løsning. En helt central observation er at vi kan finde $\mathbf{x}_0 \in \mathbb{R}^n$ så $\mathbf{b} - A\mathbf{x}_0$ er ortogonal på søjlerne i A.

Dette er en følge af Gram-Schmidt algoritmen: Hvis $\mathbf{e}_1, \dots, \mathbf{e}_r$ er en ortonormal basis for søjlerummet for A, så vil

$$\mathbf{y} = (\mathbf{b} \cdot \mathbf{e}_1)\mathbf{e}_1 + \cdots + (\mathbf{b} \cdot \mathbf{e}_r)\mathbf{e}_r$$

opfylde at $\mathbf{b} - \mathbf{y}$ er ortogonal på $\mathbf{e}_1, \dots, \mathbf{e}_r$ og dermed ortogonal på søjlerne i A. Da $\mathbf{y}$ ligger i søjlerummet for A kan $\mathbf{y}$ skrives som $A\mathbf{x}_0$ for passende $\mathbf{x}_0 \in \mathbb{R}^n$. Et sådant $\mathbf{x}_0$ vil være en mindste kvadraters løsning. Vi forklarer hvorfor. Da vektoren $\mathbf{b} - A\mathbf{x}_0$ er ortogonal på alle søjler i A, er den ortogonal på alle vektorer af formen $A\mathbf{v}$. Specielt gælder det at for enhver vektor $\mathbf{x} \in \mathbb{R}^n$ er de to vektorer $\mathbf{b} - A\mathbf{x}_0$ og $A(\mathbf{x}_0 - \mathbf{x})$ ortogonale. Ved at bruge Proposition 40 finder vi at

$$\begin{aligned} |\mathbf{b} - A\mathbf{x}|^2 &= |\mathbf{b} - A\mathbf{x}_0 + A\mathbf{x}_0 - A\mathbf{x}|^2 \\ &= |\mathbf{b} - A\mathbf{x}_0 + A(\mathbf{x}_0 - \mathbf{x})|^2 \\ &= |\mathbf{b} - A\mathbf{x}_0|^2 + |A(\mathbf{x}_0 - \mathbf{x})|^2. \end{aligned}$$

Derfor gælder

$$|\mathbf{b} - A\mathbf{x}_0|^2 \leq |\mathbf{b} - A\mathbf{x}|^2$$

for alle $\mathbf{x} \in \mathbb{R}^n$.

Nu vil vi vise at det $\mathbf{x}_0$ *som vi lige har fundet løser ligningen* $(A^TA)\mathbf{x} = A^T\mathbf{b}$. Vi arbejder videre ud fra vores nyvundne viden om at $\mathbf{b} - A\mathbf{x}_0$ skal være ortogonal på alle vektorer af formen $A\mathbf{y}$. Dette betyder at, for alle $\mathbf{y} \in \mathbb{R}^n$,

$$A\mathbf{y} \cdot (\mathbf{b} - A\mathbf{x}_0) = 0,$$

og dermed (husk at $\mathbf{y} \cdot A\mathbf{v} = A^T\mathbf{y} \cdot \mathbf{v}$)

$$\mathbf{y} \cdot (A^T\mathbf{b} - A^TA\mathbf{x}_0) = \mathbf{y} \cdot A^T(\mathbf{b} - A\mathbf{x}_0) = A\mathbf{y} \cdot (\mathbf{b} - A\mathbf{x}_0) = 0.$$

Dette vil gælde for alle $\mathbf{y} \in \mathbb{R}^n$. Nu bruger vi et beskidt men nyttigt kneb: Vi sætter ind $\mathbf{y} = (A^T\mathbf{b} - A^TA\mathbf{x}_0)$ og får at

$$|A^T\mathbf{b} - A^TA\mathbf{x}_0|^2 = (A^T\mathbf{b} - A^TA\mathbf{x}_0) \cdot (A^T\mathbf{b} - A^TA\mathbf{x}_0) = 0.$$

Det følger at

$$A^T\mathbf{b} - A^TA\mathbf{x}_0 = 0$$

eller

$$A^TA\mathbf{x}_0 = A^T\mathbf{b}.$$

Fra nu af antager vi at $A\mathbf{x} = \mathbf{0}$ kun har den trivielle nulløsning.

Vi vil vise påstanden om entydighed. Hvis $\mathbf{x}_0$ er den løsning som vi fandt i de foregående trin, og $\mathbf{x}_0'$ er en anden mindste kvadraters løsninger, så har vi jo lige set at $\mathbf{b} - A\mathbf{x}_0$ er ortogonal på $A(\mathbf{x}_0 - \mathbf{x}_0')$. Igen følger det af Pythagoras:

$$|\mathbf{b} - A\mathbf{x}_0'|^2 = |\mathbf{b} - A\mathbf{x}_0 + A(\mathbf{x}_0 - \mathbf{x}_0')|^2 = |\mathbf{b} - A\mathbf{x}_0|^2 + |A(\mathbf{x}_0 - \mathbf{x}_0')|^2.$$

Hermed er $|A(\mathbf{x}_0 - \mathbf{x}_0')|^2 = 0$ og dermed $A\mathbf{x}_0 = A\mathbf{x}_0'$ og $A(\mathbf{x}_0 - \mathbf{x}_0') = \mathbf{0}$. Så hvis ligningen $A\mathbf{x} = \mathbf{0}$ kun har den trivielle løsning, så er $\mathbf{x}_0' = \mathbf{x}_0$.

Til sidst skal vi vise at A^TA *er invertibel.* Dette gør vi ved at vise at antage at

$$(A^TA)\mathbf{x} = \mathbf{0}$$

og vise at dette medfører at $\mathbf{x} = \mathbf{0}$ (se den relevante sætning i Kapitel 4). Det kan vi gøre med denne fikse udregning:

$$A\mathbf{x} \cdot A\mathbf{x} = \mathbf{x} \cdot A^T(A\mathbf{x}) = \mathbf{x} \cdot (A^TA)\mathbf{x} = \mathbf{x} \cdot \mathbf{0} = 0.$$

Da $|A\mathbf{x}|^2 = A\mathbf{x} \cdot A\mathbf{x} = 0$ er $A\mathbf{x} = \mathbf{0}$, og det følger fra vores antagelse om A at $\mathbf{x} = \mathbf{0}$. □

♠

7.43 Eksempel

Lad os finde en mindste kvadraters løsning for ligningssystemet

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix},$$

fra Eksempel 7.4. Her er

$$A = \begin{pmatrix} 1 & 1 \\ 1 & -1 \\ -1 & 1 \end{pmatrix} \qquad \text{og} \qquad b = \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix},$$

som ved indsættelse i (7.18) giver mindste kvadraters løsningen

$$x_0 = \left(A^T A\right)^{-1} A^T b = \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}^{-1} \begin{pmatrix} 3 \\ 3 \end{pmatrix} = \begin{pmatrix} \frac{3}{2} \\ \frac{3}{2} \end{pmatrix}.$$

♠

(7.44) EKSEMPEL.

Den helt klassiske anvendelse af mindste kvadraters løsninger er at finde den bedste linje $y = \alpha x + \beta$ gennem nogle givne punkter

$$(x_1, y_1), \quad (x_1, y_1), \quad \dots \quad , (x_n, y_n)$$

i planen $\mathbb{R}^2$.

At det oftest ikke er muligt at finde en perfekt linje, som går gennem alle punkterne kan oversættes til at ligningssystemet

$$\begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

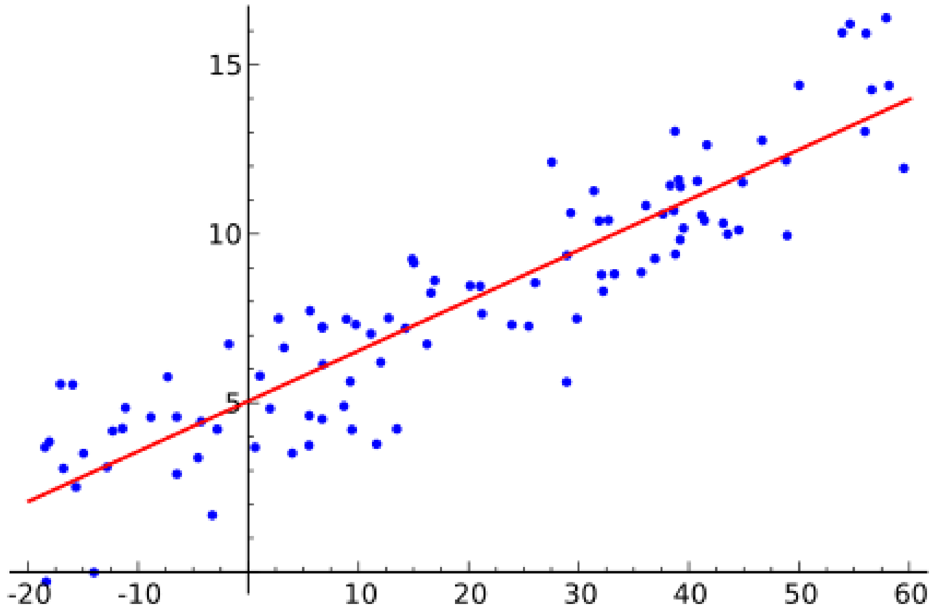
ikke har løsninger. Her kan vi så arbejde med en mindste kvadraters løsning, som giver den bedste linje $y = \alpha x + \beta$ gennem punkterne i den forstand at summen af den vertikale kvadratafstand fra linjen til y-erne

$$(y_1 - \alpha x_1 - \beta)^2 + (y_2 - \alpha x_2 - \beta)^2 + \cdots + (y_n - \alpha x_n - \beta)^2$$

bliver minimal.

Faktisk kunne vi ligeså godt have spurgt om den bedste parabel

$$y = \alpha x^2 + \beta x + \gamma$$



Figur 7.1: Bedste fit af linje til tilfældige punkter fra Wikipedia.

gennem punkterne

$$(x_1, y_1), \quad (x_1, y_1), \quad \ldots \quad , (x_n, y_n)$$

i $\mathbb{R}^2$. Dette ville med samme metode give os ligningssystemet

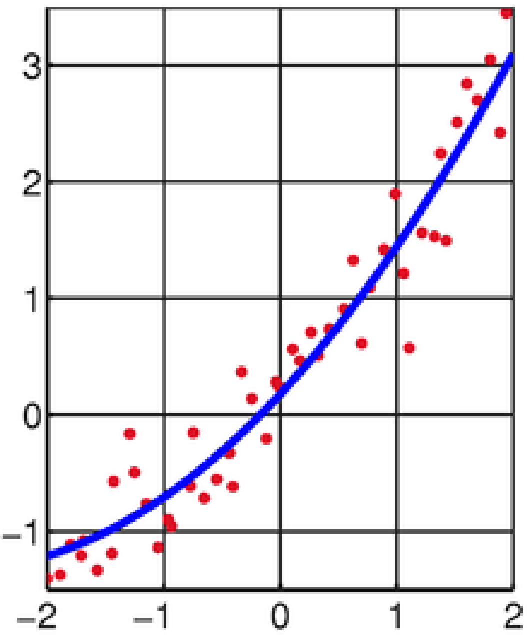
$$\begin{pmatrix} x_1^2 & x_1 & 1 \\ x_2^2 & x_2 & 1 \\ \vdots & \vdots & \\ x_n^2 & x_n & 1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

og mindste kvadraters løsningen til dette ligningssystem ville give den bedste parabel gennem punkterne med summen af den vertikale kvadratafstand fra parablen til punkterne minimeret.

Helt generelt kan den samme metode bruges til at finde det bedste eller fitte et m-te grads polynomium

$$y = a_m x^m + a_{m-1} x^{m-1} + \cdots + a_1 x + a_0$$

til en række punkter i planen.



Figur 7.2: Bedste fit af parabel til tilfældige punkter fra Wikipedia.

7.5 Opgaver

7.5.1

Gør rede for at 1 og -1 er egenværdier for matricen

$$\begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix},$$

hvor θ er en vilkårlig vinkel. Forsøg at bestemme egenvektorerne.

7.5.2

Gør rede for at

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

hvor A og B er to matricer, hvor produktet AB giver mening. Vis yderligere at

$$(A^*)^* = A.$$

7.5.3

Find den bedste mindste kvadraters linje $y = \alpha x + \beta$ gennem punkterne $(1,2), (2,1)$ og $(4,3)$ ved at benytte Sætning 41.

7.5.4

En cirkel med centrum i (a,b) og radius r i planen har ligningen

$$(x-a)^2 + (y-b)^2 = r^2. \tag{7.19}$$

1. Gør rede for hvordan (7.19) kan omskrives til ligningen

   $$2ax + 2by + c = x^2 + y^2, \tag{7.20}$$

   hvor $c = r^2 - a^2 - b^2$.

2. Forklar hvordan (7.20) i mindste kvadraters kontekst leder frem til ligningssystemet

   $$\begin{pmatrix} 2x_1 & 2y_1 & 1 \\ 2x_2 & 2y_2 & 1 \\ \vdots & \vdots & \vdots \\ 2x_n & 2y_n & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} x_1^2 + y_1^2 \\ x_2^2 + y_2^2 \\ \vdots \\ x_n^2 + y_n^2 \end{pmatrix},$$

   i forsøget på at tilpasse en cirkel til punkterne

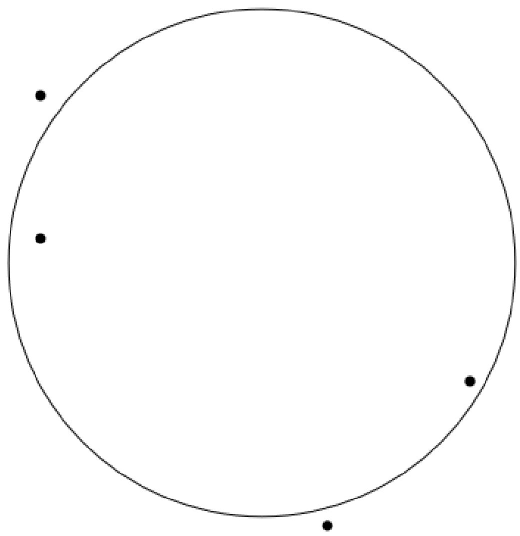
   $$(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$$

   i planen.

3. Find den bedste cirkel i mindste kvadraters forstand gennem punkterne

   $$(0,2), \quad (0,3), \quad (2,0) \quad \text{og} \quad (3,1)$$

   ved at angive centrumkoordinaterne og radius med to decimaler.



4. Diskuter hvornår der går en entydig cirkel gennem tre givne punkter $(x_1, y_1), (x_2, y_2)$ og (x_3, y_3) i henhold til egenskaber ved matricen

$$\begin{pmatrix} 2x_1 & 2y_1 & 1 \\ 2x_2 & 2y_2 & 1 \\ 2x_3 & 2y_3 & 1 \end{pmatrix}.$$

5. Kan man finde en cirkel gennem tre ikke sammenfaldende punkter, som ligger på en linje?

7.5.5

Bestem vinklen mellem vektorerne

$$\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \qquad \text{og} \qquad \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}.$$

7.5.6

Hvor tæt er de to sætninger

1. Jeg vil gerne have et glas vand
2. Jeg vil gerne have et glas vin

på hinanden?

7.5.7

Find den inverse til matricen

$$P = \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ -1 & 0 & 0 \end{pmatrix}$$

uden at regne.

7.5.8

Gør detaljeret rede for at

$$\left(\begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right)$$

er en ortogonalbasis for underrummet

$$V = \operatorname{span}\left(\begin{pmatrix}1\\2\\1\end{pmatrix}, \begin{pmatrix}1\\1\\1\end{pmatrix}\right)$$

af $\mathbb{R}^3$. Benyt derefter Proposition 22 (og Eksempel 7.3.1) til at afgøre om

$$\begin{pmatrix}1\\0\\1\end{pmatrix} \in V.$$

7.5.9

Find en ortonormalbasis for underrummet

$$\operatorname{span}\left(\begin{pmatrix}1\\1\\0\\0\end{pmatrix}, \begin{pmatrix}1\\0\\1\\0\end{pmatrix}, \begin{pmatrix}1\\0\\0\\1\end{pmatrix}\right)$$

af $\mathbb{R}^4$.

7.5.10

Skriv matricen

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

som et produkt QR, hvor Q er en ortogonal matrix og R en øvre trekantsmatrix (en matrix, som har nuller under diagonalen). Gør rede for dine udregninger og metoder.

Find en ortogonal matrix P så

$$P^T A P = \begin{pmatrix}2 & 0 & 0\\0 & -1 & 0\\0 & 0 & 1\end{pmatrix}.$$

Kapitel 8

Symmetriske matricer

VIDEO: https://youtu.be/RzXWbeYF-iE

Blandt de komplekse matricer fortjener de hermiteske en særstatus. En hermitesk matrix A er en kvadratisk matrix, som opfylder

$$A = A^*, \tag{8.1}$$

hvor

$$A^* = \bar{A}^T$$

Hvis A er en reel matrix er $A^* = A^T$ og A kaldes her reel symmetrisk det vil sige $A = A^T$.

Vi ved fra sidste kapitel at egenværdierne for en hermitesk matrix er reelle. Symmetrien (8.1) af matricen er her central. For eksempel har

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

ikke reelle egenværdier.

Mirakuløst er hermiteske matricer også diagonaliserbare endda med en ortonormal basis af egenvektorer! Dette resultat kaldes spektralsætningen og er et af højdepunkterne i noterne. Ortogonaliteten af egenvektorer hørende til forskellige egenværdier så vi allerede i det foregående kapitel. Igen er symmetrien central. For eksempel er matricen

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

slet ikke diagonaliserbar.

Hvis A er en vilkårlig reel $m \times n$ matrix af rang r kan man vise at

$$A^T A$$

en reel symmetrisk matrix med ikke-negative egenværdier

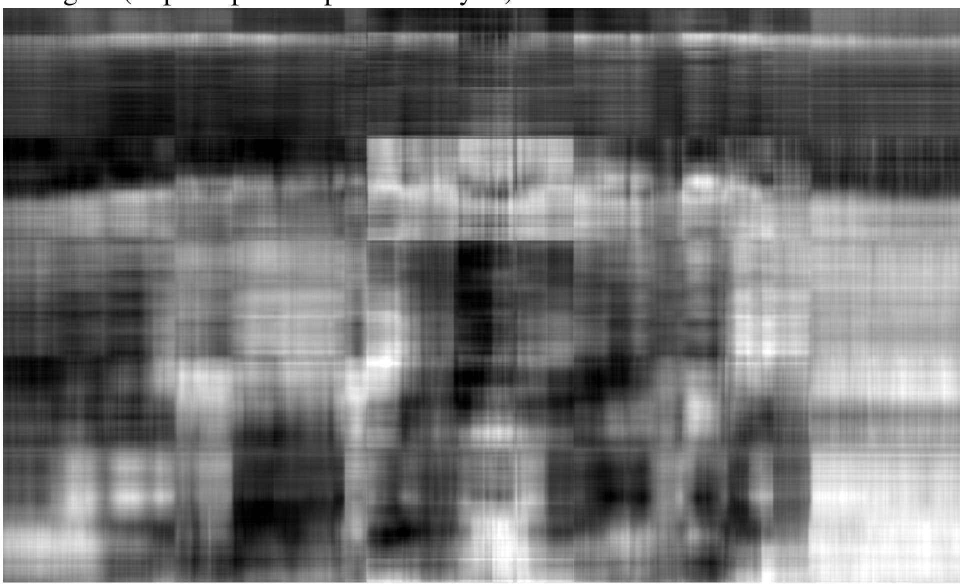
$$\lambda_1 \geq \cdots \geq \lambda_r > 0 = \lambda_{r+1} = \cdots = \lambda_n.$$

Den ortogonale diagonalisering af $A^T A$ kan benyttes til at fremskaffe ortonormale søjlevektorer $u_1, \dots, u_r$ i $m \times r$ matricen P og ortononormale søjlevektorer $\mathbf{v}_1, \dots, \mathbf{v}_r$ i $n \times r$ matricen Q så

$$A = P\Sigma Q^T = \sigma_1 u_1 \mathbf{v}_1^T + \cdots + \sigma_r u_r \mathbf{v}_r^T, \tag{8.2}$$

hvor $\sigma_i = \sqrt{\lambda_i}$ for $i = 1, \dots, r$ er de *singulære værdier* for A (disse optræder som diagonalelementer i $r \times r$ diagonalmatricen Σ ovenfor).

Dekompositionen (8.2) kaldes en *singulær værdi dekomposition* for A og er ekstremt anvendelig: Ved at afkorte summen i (8.2) opnås approksimationer til matricen A, som kan benyttes ved komprimering af billeder og data mining af store datamængder (se principal component analysis).



Grov approksimation til et grayscale billede i form af en 1536×2560 matrix ud fra få led i (8.2)

<u>8.1 Quiz</u>

Hvorfor har en kvadratisk $n \times n$ matrix altid en egenværdi, når $n > 1$?

Fordi nulrum kan udregnes over de komplekse tal.

Fordi det karakteristiske polynomium altid har ulige grad.

Fordi algebraens fundamentalsætning medfører at det karakteristiske polynomium har en rod.

Fordi RREF for en kvadratisk matrix er identitetsmatricen. ♠

8.1 Schurs lemma

Den matematiske indgang til diagonalisering af hermiteske matricer er et klassisk resultat af Issai Schur. Hvis du kigger fremad til beviset for Sætning 5 vil du helt klart opdage meningen med resultatet nedenfor, som kaldes *Schurs lemma*.

(8.2) LEMMA.

Lad A være en kvadratisk kompleks $n \times n$ matrix. Så findes en unitær matrix U så

$$U^*AU$$

er en øvre trekantsmatrix.

Bevis

Bevis. Lad $\mathbf{v}_1$ være en egenvektor for A med $|\mathbf{v}_1| = 1$ hørende til egenværdien λ. Ved hjælp af Gram-Schmidt algoritmen kan vi finde $\mathbf{v}_2, \dots, \mathbf{v}_n$ så $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ udgør en ortonormalbasis. Lad U_1 være matricen med søjler $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$. Specielt er $U_1\mathbf{e}_1 = \mathbf{v}_1$, hvor $\mathbf{e}_1$ er den første standard basisvektor. Dermed er

$$(U_1^*AU_1)\mathbf{e}_1 = U_1^{-1}AU_1(\mathbf{e}_1) = U_1^{-1}A(\mathbf{v}_1) = U_1^{-1}(\lambda\mathbf{v}_1) = \lambda\mathbf{e}_1,$$

idet $U_1^* = U_1^{-1}$. Dette medfører at

$$U_1^*AU_1 = \begin{pmatrix} \lambda & ? & \cdots & ? \\ 0 & & & \\ \vdots & & B & \\ 0 & & & \end{pmatrix},$$

hvor B er en $(n-1) \times (n-1)$ matrix. Per induktion findes en $(n-1) \times (n-1)$ unitær matrix V så V^*BV er en øvre trekantsmatrix. Vi udvider nu V til den unitære $n \times n$ matrix

$$U_2 = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & V & \\ 0 & & & \end{pmatrix}.$$

Hermed er

$$U_2^* \left(U_1^* A U_1\right) U_2 = \begin{pmatrix} \lambda & ? & \cdots & ? \\ 0 & & & \\ \vdots & & V^*BV & \\ 0 & & & \end{pmatrix}$$

en øvre trekantsmatrix. Nu er $U = U_1 U_2$ en unitær matrix sådan at $U^*AU = U_2^* U_1^* A U_1 U_2$ er en øvre trekantsmatrix. □

♠

På nøjagtig samme måde kan man bevise følgende.

(8.3) LEMMA.

Lad A være en kvadratisk reell $n \times n$ matrix. Så findes en ortogonal matrix Q så

$$Q^T A Q$$

er en øvre trekantsmatrix.

Bevis

Nu har Marcel jo allerede fortalt at beviset er det samme som beviset for Schurs lemma. Man skal bare erstatte ordet "unitær" i beviset for Schurs lemma med ordet "ortogonal", matricen U med matricen Q, og så virker det samme bevis ord for ord. ♠

8.4 Quiz

Hvilke af nedenstående matricer er øvre trekantsmatricer?

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

$$\begin{pmatrix} 1 & i \\ 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 \end{pmatrix}$$

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

♠

8.2 Spektralsætningen

Ud fra Schurs lemma bliver beviset for følgende sætning ekstremt kort.

(8.5) SÆTNING.

Lad A være en hermitesk matrix. Så findes en unitær matrix U så

$$U^*AU$$

er en diagonalmatrix med reelle indgange.

(8.6) BEMÆRKNING.

Indgangerne på diagonalen i diagonalmatricen er jo netop den hermiteske matrices egenværdier. At de er reelle har vi allerede konstateret kapitel 7.

Bevis

Bevis. Dette er et smukt og kort bevis, som benytter Schurs lemma: Fra Schurs lemma ved vi at der findes en unitær matrix U så

$$U^*AU$$

er en øvre trekantsmatrix. Men U^*AU er en hermitesk matrix, da

$$(U^*AU)^* = U^*A^*(U^*)^* = U^*AU.$$

Men en øvre trekantsmatrix som er hermitesk bliver nødt til at være en diagonalmatrix med reelle indgange. □

♠

På den samme måde får vi følgende resultat for symmetriske reelle matricer.

(8.7) PROPOSITION.

Lad A være en (reel) symmetrisk $n \times n$ matrix. Så findes en ortogonal matrix Q så

$$Q^TAQ$$

er en diagonalmatrix. Søjlerne i Q er egenvektorer for A.

Bevis

Bevis. Ifølge lemma 3 kan vi finde en ortogonal matrix Q så at $Q^T AQ$ er en øvre trekantmatrix. Men en øvre trekantmatrix som er symmetrisk er en diagonal matrix.

Se eventuelt yderligere detaljer om denne opskrivning i et tidligere kapitel. □

♠

VIDEO: https://youtu.be/CH1VLIy3VhQ

(8.8) EKSEMPEL.

Betragt den symmetriske matrix

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

Det oplyses at A har egenværdierne $\lambda = -1$ og $\lambda = 5$. Find ud fra dette en ortogonal matrix Q så $Q^T AQ$ er en diagonalmatrix.

Her er det nok at finde en ortonormal basis af egenvektorer for A. Dette gøres ved at finde en ortonormal basis for hvert egenrum og kombinere dem til en basis for $\mathbb{R}^3$. Den samlede basis bliver ortonormal, fordi egenvektorer hørende til forskellige egenværdier er ortogonale. Vi kigger først på egenværdien $\lambda = -1$. Her regner man sig frem til at

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \qquad \text{og} \qquad \mathbf{v}_2 = \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} \tag{8.3}$$

er en basis for egenrummet $N(A + I_3)$ ved at løse ligningssystemet

$$\begin{pmatrix} 1 & 2 & -1 \\ 2 & 4 & -2 \\ -1 & -2 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \tag{8.4}$$

Se resultatet om hvordan dette relaterer til RREF og det efterfølgende eksempel for at forstå hvordan man kommer fra ligningssystemet (8.4) til en basis for $N(A + I_3)$.

Vi benytter nu Gram-Schmidt algoritmen på vektorerne i (8.3) og kommer frem til ortonormalbasen

$$u_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{pmatrix} \qquad \text{og} \qquad u_2 = \begin{pmatrix} -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix} \tag{8.5}$$

Nu ved vi at $\dim N(A-5I_3)=1$ for egenværdien $\lambda=5$ (hvorfor?). Som ovenfor finder vi at

$$\mathbf{v}_3=\begin{pmatrix}-1\\-2\\1\end{pmatrix}$$

er en basis for $N(A-5I_3)$ og dermed at

$$u_3=\begin{pmatrix}-\frac{1}{\sqrt{6}}\\ \\ -\frac{2}{\sqrt{6}}\\ \\ \frac{1}{\sqrt{6}}\end{pmatrix}$$

er en ortonormal basis for $N(A-5I_3)$. Samlet bliver (u_1,u_2,u_3) en ortonormalbasis for $\mathbb{R}^3$ bestående af egenvektorer for A. Matricen Q med søjler u_1,u_2 og u_3 er dermed en ortogonal matrix, som opfylder at

$$Q^TAQ=\begin{pmatrix}-1&0&0\\0&-1&0\\0&0&5\end{pmatrix}.$$

8.9 Quiz

Lad A være en vilkårlig $m\times n$ matrix. Hvad er rangen $\mathrm{rk}(A^TA)$ af $n\times n$ matricen

A^TA?

$\mathrm{rk}(A)-1$

$\mathrm{rk}(A)+1$

$\mathrm{rk}(A)$

♠

8.3 Singulær værdi dekomposition

Inden vi behandler det centrale emne i dette afsnit først en nyttig omformulering af matrixmultiplikation.

(8.10) PROPOSITION.

Lad A være en $m \times n$ matrix med søjlevektorerne $\mathbf{u}_1, \dots, \mathbf{u}_n$ og B en $n \times r$ matrix med rækkevektorerne $\mathbf{v}_1, \dots, \mathbf{v}_n$. Så er

$$AB = \mathbf{u}_1\mathbf{v}_1 + \cdots + \mathbf{u}_n\mathbf{v}_n.$$

Bevis

Bevis. Ud fra definitionen på matrixmultiplikation følger at

$$\begin{aligned}(\mathbf{u}_1\mathbf{v}_1 + \cdots + \mathbf{u}_n\mathbf{v}_n)_{ij} &= (\mathbf{u}_1)_i(\mathbf{v}_1)^j + \cdots + (\mathbf{u}_n)_i(\mathbf{v}_n)^j \\ &= A_{i1}B_{1j} + \cdots + A_{in}B_{nj} = A_iB^j = (AB)_{ij}.\end{aligned}$$

□

♠

8.11 Quiz

Lad $\mathbf{u}, \mathbf{v} \in \mathbb{C}^n$ være vilkårlige vektorer. Hvilke udsagn nedenfor omkring rangen $\mathrm{rk}(\mathbf{u}\mathbf{v}^T)$ af $n \times n$ matricen

$$B = \mathbf{u}\mathbf{v}^T$$

er korrekte.

$$\mathrm{rk}(B) = 1$$

$$\mathrm{rk}(B) \le 1$$

$$\mathrm{rk}(B) \ge 2$$

♠

(8.12) LEMMA.

For en vilkårlig reel $m \times n$ matrix A er

$$A^T A$$

en symmetrisk $n \times n$ matrix med ikke negative egenværdier.

Bevis

Bevis. $A^T A$ er symmetrisk på grund af udregningen

$$(A^T A)^T = A^T (A^T)^T = A^T A.$$

Derved er alle egenværdier for $A^T A$ reelle.

Lad nu $\mathbf{v}$ være en egenvektor for $A^T A$ hørende til egenværdien λ. Så gælder

$$(A^T A)\mathbf{v} = \lambda \mathbf{v} \implies$$
$$|A\mathbf{v}|^2 = (A\mathbf{v}) \cdot (A\mathbf{v}) = \mathbf{v} \cdot (A^T A\mathbf{v}) = \mathbf{v} \cdot (\lambda \mathbf{v}) = \lambda (\mathbf{v} \cdot \mathbf{v}) = \lambda |\mathbf{v}|^2$$

og derfor

$$\lambda = \frac{|A\mathbf{v}|^2}{|\mathbf{v}|^2} \geq 0.$$

□

♠

(8.13) DEFINITION.

De singulære værdier for en reel matrix A er givet som

$$\sigma_i = \sqrt{\lambda_i},$$

hvor $i = 1, \ldots, r$ og

$$\lambda_1 \geq \cdots \geq \lambda_r > 0$$

er egenværdierne > 0 for $A^T A$ i aftagende rækkefølge.

Nu kan vi introducere og bevise singulær værdi dekompositionen af en reel matrix.

(8.14) SÆTNING.

Lad A være en reel $m \times n$ matrix af rang r. Så kan A faktoriseres som

$$A = P\Sigma Q^T, \tag{8.6}$$

hvor P er en $m \times r$ matrix med ortonormale søjler $\mathbf{u}_1, \ldots, \mathbf{u}_r$, Q er en $n \times r$ matrix med ortonormale søjler $\mathbf{v}_1, \ldots, \mathbf{v}_r$ og Σ en diagonalmatrix med de singulære værdier

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$$

i diagonalen. Faktoriseringen (8.6) *kan også skrives*

$$A = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \cdots + \sigma_r \mathbf{u}_r \mathbf{v}_r^T. \tag{8.7}$$

Bevis

Bevis. Vi ved fra Proposition 7, at der findes en ortogonal matrix V så

$$V^T(A^TA)V = D,$$

hvor D er en diagonal matrix med egenværdierne $\lambda_1, \ldots, \lambda_n$ svarende til de singulære værdier i diagonalen. Vi skriver $V = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$, så at vektorerne $\mathbf{v}_i \in \mathbb{R}^n$ er søjlerne i V. Den første påstand vi skal bruge er at $A\mathbf{v}_i = \mathbf{0}$ for $i \geq r+1$. Det følger af af

$$(A\mathbf{v}_i) \cdot (A\mathbf{v}_i) = (AV\mathbf{e}_i) \cdot (AV\mathbf{e}_i) = (V^TA^TAV\mathbf{e}_i) \cdot (\mathbf{e}_i) = (D\mathbf{e}_i) \cdot \mathbf{e}_i = (\lambda_i\mathbf{e}_i) \cdot \mathbf{e}_i = \lambda_i,$$

hvor $\mathbf{e}_i$ er standard basisvektorerne i $\mathbb{R}^n$. Hvis $i \geq r+1$ er altså $|A\mathbf{v}_i|^2 = \lambda_i = 0$, så at $A\mathbf{v}_i = \mathbf{0}$.

Vi vil definere de to matricer P og Q ved at skrive deres søjler ned.

$$Q = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r)$$

består af de første r søjler i V. Men nu er det jo Q^T vi skal bruge, så vi må vide noget om denne $r \times n$ matrix. Om lidt vil det vise sig at det er godt hvis vi kan finde $Q^T\mathbf{v}_i$. Vi skriver denne vektor udtrykt i standardbasen $(\mathbf{e}_i)$ for $\mathbb{R}^r$ som

$$Q^T\mathbf{v}_i = a_{i1}\mathbf{e}_1 + a_{i2}\mathbf{e}_2 + \cdots + a_{ir}\mathbf{e}_r.$$

Vi kan beregne koefficienterne a_{ij} på følgende måde:

$$a_{ij} = (a_1\mathbf{e}_1 + \cdots a_r\mathbf{e}_r) \cdot \mathbf{e}_j = (Q^T\mathbf{v}_i) \cdot \mathbf{e}_j = \mathbf{v}_i \cdot (Q\mathbf{e}_j) = \mathbf{v}_i \cdot \mathbf{v}_j.$$

Det følger at $a_{ii} = 1$ og $a_{ij} = 0$ hvis $i \neq j$. Sagt på en anden måde er

$$\begin{aligned} Q^T(\mathbf{v}_i) &= \mathbf{e}_i \qquad \text{hvis } i \leq r \\ Q^T(\mathbf{v}_i) &= 0 \qquad \text{hvis } i \geq r+1. \end{aligned}$$

Vi lader nu $P = (\frac{1}{\sigma_1}A(\mathbf{v}_1), \frac{1}{\sigma_2}A(\mathbf{v}_2), \dots, \frac{1}{\sigma_r}A(\mathbf{v}_r))$. Vores påstand er at $A = P\Sigma Q^T$. Ved at bruge linearitet er det nok at tjekke dette på en basis for $\mathbb{R}^n$. Nu er vi kommet til et punkt hvor det er vigtigt at vi vælger denne basis på en meget snedig måde. Vi er meget snedige, og vi vælger $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ som basis, og skal altså vise at $A\mathbf{v}_i = P\Sigma Q^T \mathbf{v}_i$ for alle $i \leq n$. Det gør vi bare sådan her:

$$\begin{aligned} P\Sigma Q^T(\mathbf{v}_i) = P\Sigma\mathbf{e}_i = P(\sigma_i(\mathbf{e}_i)) = \sigma_i \frac{1}{\sigma_i} A(\mathbf{v}_i) = A(\mathbf{v}_i) & \quad \text{hvis } i \leq r, \\ P\Sigma Q^T(\mathbf{v}_i) = P\Sigma(\mathbf{0}) = \mathbf{0} = A(\mathbf{v}_i) & \quad \text{hvis } i \geq r+1. \end{aligned}$$

□

♠

(8.15) BEMÆRKNING.

Singulær værdi dekompositionen af en $m \times n$ matrix A regnes ifølge beviset ovenfor ud på følgende måde:

Den symmetriske $n \times n$ matrix $A^T A$ diagonaliseres via den ortogonale matrix Q_1 så

$$Q_1^T (A^T A) Q_1 = D,$$

hvor egenværdierne for $A^T A$ er arrangeret i aftagende rækkefølge

$$\lambda_1 \geq \cdots \geq \lambda_r > \lambda_{r+1} = \cdots = \lambda_n = 0$$

i diagonalen i diagonalmatricen D. Matricen Q dannes ud fra de første r søjler $\mathbf{v}_1, \dots, \mathbf{v}_r$ i $n \times n$ matricen Q_1. Matricen P bliver så $m \times r$ matricen med søjler

$$\frac{1}{\sigma_i} A\mathbf{v}_i.$$

(8.16) EKSEMPEL.

Betraget eksemplet (fra Olver, Shakiban, *Applied Linear Algebra*)

$$A = \begin{pmatrix} 3 & 5 \\ 4 & 0 \end{pmatrix}.$$

Her er

$$A^T A = \begin{pmatrix} 25 & 15 \\ 15 & 25 \end{pmatrix}$$

med egenværdierne $\lambda_1 = 40$ og $\lambda_2 = 10$ og tilhørende egenvektorer

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{og} \quad \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

og deraf følgende ortonormalbasis

$$\mathbf{v}_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} \qquad \text{og} \qquad \mathbf{v}_2 = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$$

af egenvektorer. Dermed er

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

i den singulære værdi dekomposition.

Søjlevektorerne i P bliver hermed

$$\mathbf{u}_1 = \frac{1}{\sigma_1} A\mathbf{v}_1 = \frac{1}{\sqrt{40}} \begin{pmatrix} 4\sqrt{2} \\ 2\sqrt{2} \end{pmatrix} = \begin{pmatrix} \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{pmatrix}$$
$$\mathbf{u}_2 = \frac{1}{\sigma_2} A\mathbf{v}_2 = \frac{1}{\sqrt{10}} \begin{pmatrix} \sqrt{2} \\ -2\sqrt{2} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{5}} \\ -\frac{2}{\sqrt{5}} \end{pmatrix}$$

og

$$P = \begin{pmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \end{pmatrix}.$$

Den singulære værdi dekomposition af A bliver hermed

$$A = \begin{pmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} \sqrt{40} & 0 \\ 0 & \sqrt{10} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} = P\Sigma Q^T.$$

8.17 Opgave

Lad A være en invertibel 2×2 matrix. Giv en geometrisk fortolkning af singulær værdi dekompositionen for A ud fra eksempelmaterialet i sidste kapitel. ♠

8.4 Approksimation af matricer via svd

Singulær værdi dekompositionen er et redskab til at approksimere en $m \times n$ matrix A af rang r med $m \times n$ matricer af lavere rang k, hvor

$$0 < k < r.$$

Her kan man vise at en optimal approksimation af rang k til A fremkommer som

$$\sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \cdots + \sigma_k \mathbf{u}_k \mathbf{v}_k^T$$

ud fra singulær værdi dekompositionen af A i (8.7).

Optimal betyder her optimal eller tættest på med hensyn til afstanden udmålt ud fra det naturlige prikprodukt

$$A \cdot B = \operatorname{tr}(A^T B)$$

på mængden af $m \times n$ matricer det vil sige $\mathbb{R}^{mn}$. Her betegner tr sporet af en matrix det vil sige summen af matricens diagonalelementer.

8.4.1 Eksperimenter med grayscale billeder

Adskillige computeralgebra systemer har funktioner til at konvertere billedfiler til grayscale og håndtere billedfiler som matricer. Nedenfor kode i Mathematica (tilsvarende eksperimenter kan relativt let også implementeres i for eksempel MATLAB eller Octave).

```
SVDimage[mat_, tol_]:=
  Module[{p, s, q},
  {p, s, q} = SingularValueDecomposition[mat, Tolerance->tol];
  Return[Image[p.s.Transpose[q]]]
]

image = ColorConvert[Import["Google.jpg"], "Grayscale"]
matr = ImageData[image]; (* Extract matrix from image *)

Export["Google0.1.jpg", SVDimage[matr, 0.1]]
Export["Google0.05.jpg", SVDimage[matr, 0.05]]
Export["Google0.01.jpg", SVDimage[matr, 0.01]]

(* Ranks of approximations: *)
Length[SingularValueList[matr, Tolerance->0.1]]
Length[SingularValueList[matr, Tolerance->0.05]]
Length[SingularValueList[matr, Tolerance->0.01]]
```

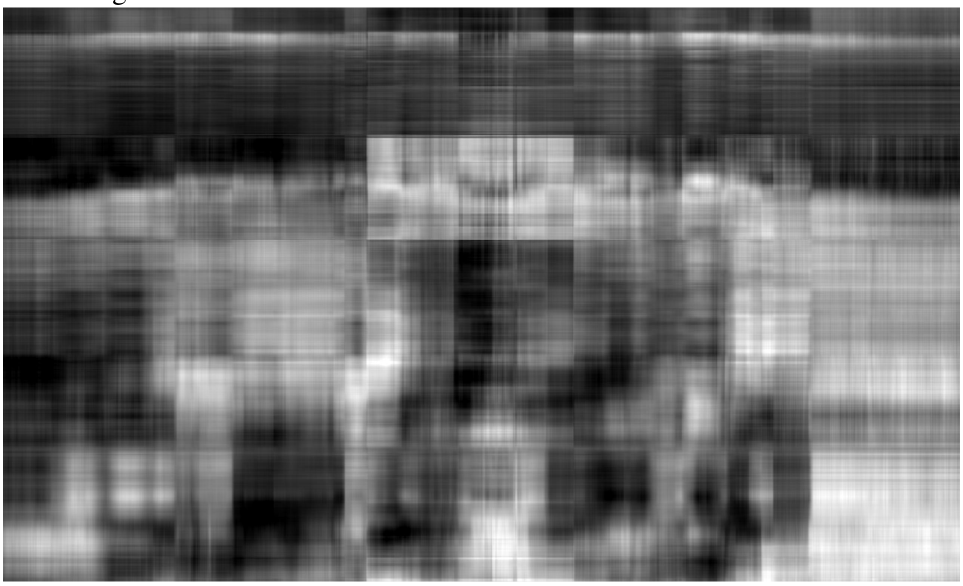
Input til ovenstående efter grayscaling er billedet

Originalt billede med matrix af rang $r = 1536$

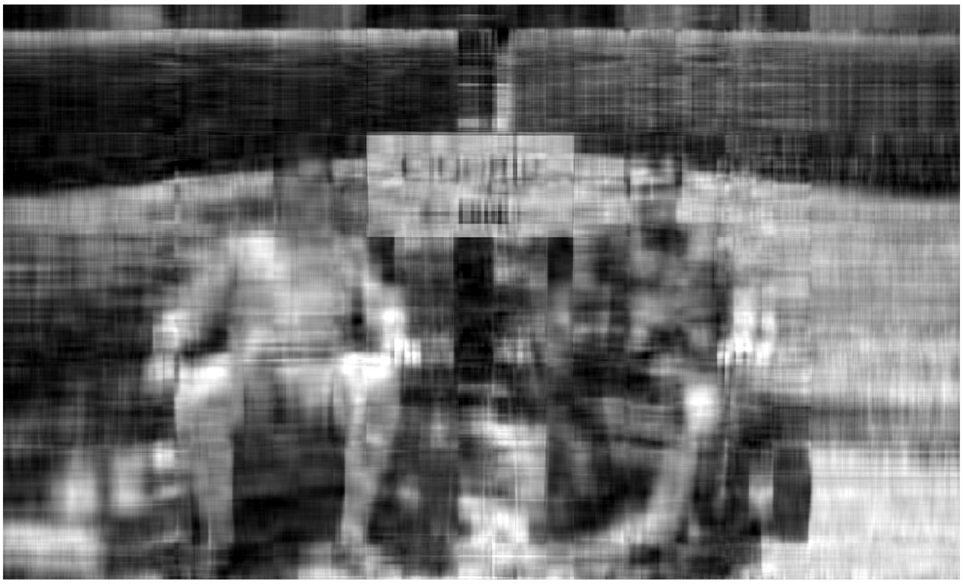
Herefter approksimeres matricen for billedet ud fra

$$\sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \cdots + \sigma_k \mathbf{u}_k \mathbf{v}_k^T$$

med forskellige værdier af $k < 1536$.



Approksimation med $k = 6$



Approksimation med $k = 15$

Approksimation med $k = 177$

Se også anvendelser fra kemi med hensyn til principal component analysis.

8.5 Opgaver

8.5.1

Vis at produktet af to unitære matricer er en unitær matrix.

8.5.2

Lad T være en øvre trekantsmatrix. Hvorfor bliver T nødt til at være en diagonalmatrix, hvis den er hermitesk?

8.5.3

Lad

$$A = \begin{pmatrix} 2 & 1-i \\ 1+i & 1 \end{pmatrix}.$$

Find en unitær matrix, som diagonaliserer A.

8.5.4

Udregn en ortonormal basis af egenvektorer for matricen

$$A = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

Find dernæst en ortogonal matrix P så $P^T AP$ er en diagonalmatrix.

8.5.5

Gør rede for at det karakteristiske polynomium til matricen

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

er $\lambda^2(3-\lambda)$. Benyt dette til at vise at

$$A^n = 3^{n-1}A$$

for $n \geq 1$. Find dernæst en ortogonal matrix P så $P^T AP$ er en diagonalmatrix.

8.5.6

(Fra Leon)

Find matricerne af rang 1 og 2 tættest på matricen

$$\begin{pmatrix} -2 & 8 & 20 \\ 14 & 19 & 10 \\ 2 & -2 & 1 \end{pmatrix} = \begin{pmatrix} \frac{3}{5} & -\frac{4}{5} & 0 \\ \frac{4}{5} & \frac{3}{5} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 30 & 0 & 0 \\ 0 & 15 & 0 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \end{pmatrix}$$

med udgangspunkt i afsnit 8.4.

Kapitel 9

Anvendelser af lineær algebra

9.1 Differentialligninger

Rutherfords og Soddys lov for radioaktivt henfald siger at en bestemt brøkdel λ af et grundstof vil henfalde per tidsenhed. Hvis vi med $N(t)$ betegner antallet af atomer til tiden t af grundstoffet, så kan loven skrives

$$N(t+\Delta) - N(t) = -\lambda \Delta N(t). \tag{9.1}$$

Ved at tidsintervallet Δ går mod 0 fås derfor differentialligningen

$$N'(t) = -\lambda N(t),$$

som har løsningerne $N(t) = Ce^{-\lambda t}$, hvor C er en konstant, som er givet ved antallet af atomer til tiden $t = 0$ i.e., $C = N(0)$.

I mange anvendelser støder man på systemer af differentialligninger som

$$x_1'(t) = a_{11}x_1(t) + \cdots + a_{1n}x_n(t) \tag{9.2}$$

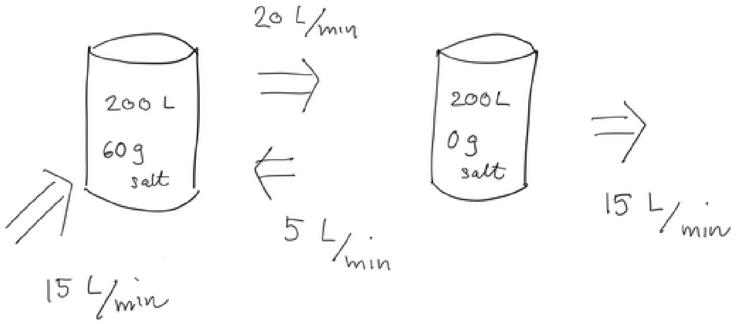
$$\vdots \tag{9.3}$$

$$x_n'(t) = a_{n1}x_1(t) + \cdots + a_{nn}x_n(t), \tag{9.4}$$

hvor opgaven så går ud på at at finde funktioner $x_1(t), \dots, x_n(t)$, som opfylder (9.2). I (9.2) er koefficienterne a_{ij} konstanter (reelle tal).

(9.1) EKSEMPEL.

Følgende eksempel er lånt fra Leons bog om anvendt lineær algebra. Vi har to forbundne kar hver indeholdende 200 L vand. Til at begynde med indeholder det første kar 60 gram salt og det andet 0 gram salt. Når hanerne tændes løber der vand gennem systemet som indikeret nedenfor (per minut). Hanen, som tilfører 15 L per minut, indeholder ferskvand.



Hvad er indholdet af salt i de to beholdere til tiden t?

Her opstiller vi to funktioner $x_1(t)$ og $x_2(t)$, som angiver saltindhold (i gram) i hver beholder til tiden t. Lad os sige at $x_1(t)$ angiver saltindhold i beholderen, som indeholder 60 gram til at starte med. Ved at betrage et lille tidsrum Δ får vi ved almindelig købmandsregning:

$$\begin{aligned} x_1(t+\Delta) - x_1(t) &= -0.1\,\Delta x_1(t) + 0.025\,\Delta x_2(t) \\ x_2(t+\Delta) - x_2(t) &= 0.1\,\Delta x_1(t) - 0.1\,\Delta x_2(t), \end{aligned}$$

som giver systemet

$$\begin{aligned} x_1'(t) &= -0.1\,x_1(t) + 0.025\,x_2(t) \\ x_2'(t) &= 0.1\,x_1(t) - 0.1\,x_2(t) \end{aligned}$$

af differentialligninger med begyndelsesbetingelserne $x_1(0) = 60$ og $x_2(0) = 0$.

9.1.1 Løsning via egenværdier og egenvektorer

Lad os kalde $\mathbf{x}(t) = (x_1(t), \dots, x_n(t))^T$ for en løsningsvektor til (9.2), hvis $x_1(t), \dots, x_n(t)$ tilfredsstiller (9.2). Nedenstående siger at løsningsvektorerne minder om (er) et vektorrum.

(9.2) PROPOSITION.

Hvis $\mathbf{x}(t)$ og $\mathbf{y}(t)$ er løsningsvektorer til (9.2) *og $\lambda, \mu \in \mathbb{R}$, så er*

$$\lambda \mathbf{x}(t) + \mu \mathbf{y}(t)$$

også en løsningsvektor til (9.2).

9.3 Bevis

Bevis. Beviset følger ved direkte indsætning under brug af

$$(\lambda \mathbf{x}(t) + \mu \mathbf{y}(t))' = \lambda \mathbf{x}'(t) + \lambda \mathbf{y}'(t).$$

□

♠

Differentialligningssystemet (9.2) kan skrives via matrixmultiplikation som

$$\begin{pmatrix} x_1'(t) \\ \vdots \\ x_n'(t) \end{pmatrix} = A \begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix}, \tag{9.5}$$

hvor

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}.$$

Egenværdier og egenvektorer for A indgår overraskende i løsningen af (9.2). Essensen er følgende resultat, som følger ved en ret enkel udregning.

(9.4) PROPOSITION.

Hvis λ er en egenværdi for A og $\mathbf{v}$ en egenvektor hørende til λ, så er

$$\mathbf{x}(t) = e^{\lambda t}\mathbf{v}$$

en løsningsvektor til (9.2).

9.5 Bevis

Bevis. For $\mathbf{x}(t)$ har vi via (9.5) at

$$\mathbf{x}'(t) = \lambda e^{\lambda t}\mathbf{v} = e^{\lambda t}A\mathbf{v},$$

som ved division med $e^{\lambda t}$ netop giver det ønskede. □

♠

Vi har ikke defineret vektorrum generelt, men man kan vise at løsningsvektorerne til (9.2) udgør et n-dimensionalt underrum i vektorrummet af alle vektorfunktioner med passende egenskaber.

(9.6) BEMÆRKNING.

Hvis man i tilknytning til problemet om at løse (9.2) forlanger at begyndelsesbetingelserne

$$\begin{gathered} x_1(0) = y_1 \\ \vdots \\ x_n(0) = y_n \end{gathered}$$

skal være opfyldt, så findes kun en løsning med disse egenskaber for en given vektor

$$\mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

Hvis A har n forskellige egenværdier $\lambda_1, \ldots, \lambda_n$ med tilhørende egenvektorer $\mathbf{v}_1, \ldots, \mathbf{v}_n$ kan denne løsning findes ved at bestemme konstanter $C_1, \ldots, C_n$ så

$$\mathbf{x}(t) = C_1 e^{\lambda_1 t}\mathbf{v}_1 + \cdots + C_n e^{\lambda_n t}\mathbf{v}_n.$$

Disse konstanter bestemmes ved at løse ligningssystemet

$$C_1\mathbf{v}_1 + \cdots + C_n\mathbf{v}_n = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix},$$

som skrevet ud bliver et ligningssystem med n ligninger og n ubekendte: $C_1, \ldots, C_n$. Dette ligningssystem har en entydig løsning, da $\mathbf{v}_1, \ldots, \mathbf{v}_n$ udgør en basis.

Lad os nu afprøve vores observationer på Eksempel 1. Matricen

$$\begin{pmatrix} -0.1 & 0.025 \\ 0.1 & -0.1 \end{pmatrix}$$

har egenværdierne -0.15 og -0.05 med tilhørende egenvektorer

$$\begin{pmatrix} -1 \\ 2 \end{pmatrix} \qquad \text{og} \qquad \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

Via Proposition 4 og Proposition 2 er

$$\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = C_1 e^{-0.15t}\begin{pmatrix} -1 \\ 2 \end{pmatrix} + C_2 e^{0.05t}\begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

derfor en løsningsvektor for vilkårlige $C_1, C_2 \in \mathbb{R}$. Med begyndelsesbetingelserne $x_1(0) = 60$ og $x_2(0) = 0$ kan vi derfor bestemme konstanterne C_1 og C_2 ud fra ligningssystemet

$$\begin{pmatrix} 60 \\ 0 \end{pmatrix} = C_1\begin{pmatrix} -1 \\ 2 \end{pmatrix} + C_2\begin{pmatrix} 1 \\ 2 \end{pmatrix},$$

som har løsningerne $C_1 = -30$ og $C_2 = 30$. De endelige løsninger bliver derfor

$$\begin{aligned} x_1(t) &= 30e^{-0.15t} + 30e^{-0.05t} \\ x_2(t) &= -60e^{-0.15t} + 60e^{-0.05t}. \end{aligned}$$

Efter en time er der kun 1.49731 gram salt i den første beholder $((x_1(t))$ og 2.97982 gram salt i den anden beholder $((x_2(t))$.

9.1.2 Oscillerende løsninger

Hvordan fortolkes en kompleks egenværdi $\lambda \in \mathbb{C}$ og en tilhørende egenvektor $\mathbf{v} \in \mathbb{C}^n$ for A med hensyn til løsningen af differentialligningssystemet

$$\mathbf{x}'(t) = A\mathbf{x}(t)? \tag{9.6}$$

Det giver stadig god mening matematisk at sætte $\mathbf{x}(t) = e^{\lambda t}\mathbf{v}$ ind i (9.6) og se at ligningen er opfyldt, men her er $\mathbf{x}(t)$ en løsningsvektor med komplekse koordinater for t.

Lad os antage at $\lambda = a + ib$ og skrive $\mathbf{v} = \mathbf{v}_1 + i\mathbf{v}_2$ med $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^n$. Da A har reelle indgange er A invariant med hensyn til kompleks konjugering det vil sige $\bar{A} = A$. Derfor er $\bar{\lambda} = a - ib$ også en egenværdi for A med egenvektor $\bar{\mathbf{v}} = \mathbf{v}_1 - i\mathbf{v}_2$, da

$$\bar{\lambda}\bar{\mathbf{v}} = \overline{\lambda\mathbf{v}} = \overline{A\mathbf{v}} = \bar{A}\bar{\mathbf{v}} = A\bar{\mathbf{v}}.$$

Opdagelsen er at realdelen og imaginærdelen

$$\mathrm{Re}\, e^{\lambda t}\mathbf{v} \qquad \text{og} \qquad \mathrm{Im}\, e^{\lambda t}\mathbf{v}$$

er løsninger til $\mathbf{x}'(t) = A\mathbf{x}(t)$. Dette følger nemlig ud fra identiteterne

$$\begin{aligned} \mathrm{Re}\, e^{\lambda t}\mathbf{v} &= \frac{1}{2}\left(e^{\lambda t}\mathbf{v} + e^{\bar{\lambda} t}\bar{\mathbf{v}}\right) \\ \mathrm{Im}\, e^{\lambda t}\mathbf{v} &= \frac{1}{2}\left(e^{\lambda t}\mathbf{v} - e^{\bar{\lambda} t}\bar{\mathbf{v}}\right) \end{aligned}$$

ved at benytte Proposition 2. Real- og imaginærdelene kan udregnes via almindelig multiplikation af komplekse tal ud fra opspaltningen

$$e^{\lambda t}\mathbf{v} = e^{at}(\cos(bt) + i\sin(bt))(\mathbf{v}_1 + i\mathbf{v}_2)$$

Her får man

$$\mathrm{Re}\, e^{\lambda t}\mathbf{v} = e^{at}\left(\cos(bt)\mathbf{v}_1 - \sin(bt)\mathbf{v}_2\right) \tag{9.7}$$

$$\mathrm{Im}\, e^{\lambda t}\mathbf{v} = e^{at}\left(\cos(bt)\mathbf{v}_2 + \sin(bt)\mathbf{v}_1\right) \tag{9.8}$$

(9.7) EKSEMPEL.

Differentialligningen $y'' = -ky$, hvor $k > 0$ er en konstant, kan ved et klassisk trick omskrives til

$$\begin{aligned} y' &= x \\ x' &= -ky. \end{aligned}$$

Her er differentialligningsystemet

$$\begin{pmatrix} y' \\ x' \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -k & 0 \end{pmatrix} \begin{pmatrix} y \\ x \end{pmatrix}$$

og dermed er

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

med egenværdien $\lambda = \sqrt{k}\, i$ og tilhørende egenvektor

$$\mathbf{v} = \begin{pmatrix} 1 \\ \sqrt{k}\, i \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + i \begin{pmatrix} 0 \\ \sqrt{k} \end{pmatrix}.$$

Nu kan en løsning til den oprindelige differentialligning udregnes via (9.7) til

$$y(t) = C_1 \cos(\sqrt{k}t) + C_2 \sin(\sqrt{k}t)$$

for passende konstanter C_1, C_2.

9.2 Principal component analysis

Nedenstående anvendelse kommer fra kemi og er skrevet af Frank Jensen.

Principal Component Analysis (PCA) er en hyppigt anvendt metode til at uddrage information fra store datamængder.

Et eksempel er en computer simulering på et microsekund af et protein indeholdende 10^4 atomer, hvor den rå datamængde er 10^5 tidsmæssige adskilte positioner af 10^4 atom koordinater. Langt størstedelen af denne information er tilfældige termiske bevægelser, som ikke er interessante, mens koordinerede atombevægelser, der ændrer proteinets struktur er interessante.

Et andet eksempel er korrelationen mellem molekylers struktur og deres biologiske virkning, hvor et molekyles vekselvirkninger med det biologiske target kan kvantificeres ved dets egenskaber i f.eks. 10^4 punkter i det 3-dimensionale rum omkring molekylet. I en Quantitative Structure Activity Relationship (QSAR) forsøger man at finde hvilke områder omkring et molekyle, der er vigtig for den biologiske virkning, ud fra informationen om hvordan de 10^4 beskrivende variable for f.eks. 50 molekyler korrelerer med deres biologiske aktivitet.

Den rå information kan arrangeres i en rektangulær matrix M, og vi er interesseret i korrelationen mellem elementerne i matricen. Dette kan vi finde ud fra en analyse af $A = M^t M$, som er en symmetrisk kvadratisk matrix. En sådan matrix kan ifølge noterne altid diagonaliseres ved en unitær transformation, hvor U-matricen indeholder egenvektorerne.

$$A = M^t M, \qquad \Lambda = U^t A U. \tag{9.9}$$

Den omvendte transformation betyder at den originale matrix kan konstrueres ud fra egenværdierne og egenvektorerne.

$$A = U \Lambda U^t \tag{9.10}$$

Ideen i PCA er at repræsentere informationen i A matricen som en approksimation ved kun nogle få egenvektorer. Egenværdierne fra diagonalisering af A matricen, relativt til summen af alle egenværdier, giver et mål for, hvor stor en brøkdel af den originale information et given antal egenvektorer kan repræsentere. Egenvektorerne kaldes Principal Components, og egenvektoren svarende til den største egenværdi beskriver den største variation af de originale variable, egenvektoren med den næst-største egenværdi beskriver den næst-største variation af de original variable, etc. Ofte kan man repræsentere 80-90% af informationen i en matrix med dimension 10^4 med kun nogle få (1-5) Principal Components.

9.3 Spin

Kvantefysik er en ekstremt nøjagtig matematisk model. Den beskriver verden på partikel (eller felt) niveau, så vi er altså nede i det mest submikroskopiske. Teorien strider mod al sund fornuft, og den sunde fornuft taber altid kampen.

I denne teori beskrives et fysisk systems tilstand som en vektor $\mathbf{v}$ af længde 1 i et vektorrum med prikprodukt over de komplekse tal. Teorien kan forudsige hvad resultat af forskellige målinger kan give.

De egenskaber som vi kan måle kaldes observable. Systemets position, hastighed, momentum, energi og så videre er alle eksempler på observable. Opskriften er at alle målinger vi kan udføre, altså alle observable, er repræsenterede af hermiteske matricer. Resultatet af en måling er en egenværdi af denne matrix. Hvis V er et endeligt dimensionelt vektorrum er der altså kun endelig mange mulige resultater (hvorfor?)!

Hvis vi i praksis udfører en måling på systemet forudsiger teorien ikke hvad for en egenværdi vi får ud af målingen, men giver kun sandsynligheder. Hvis tilstandsvektoren $\mathbf{v} = c_1\psi_1 + c_2\psi_2$ er en linearkombination af to egenvektorer ψ_1 og ψ_2 hørende til to forskellige egenværdier E_1 og E_2 for den hermiteske energioperator kunne teorien for eksempel forudsige “med 33 procents sandsynlighed giver målingen af energien E_1, og med 67 procents sandsynlighed giver målingen af energien

E_2". Disse sandsynligheder p_1 og p_2 kan beregnes med en simpel formel:

$$p_i = |c_i|^2 |\psi_i|^2$$

For at gøre formlen endnu mere simpel plejer man at vælge eigenvektorer af længde 1, det vil sige så at de også er tilstandsfunktioner, og så at $|\psi_i| = 1$. Hvis man gør det er altså $p_i = |c_i|^2$.

9.8 Opgave

Vis at $p_1 + p_2 = 1$! Der skal helt sikkert bruges at $|\mathbf{v}| = 1$. ♠

Dette lyder selvfølgelig som det rene nonsens, og I skal ikke have nogle illusioner om at det senere vil blive mindre underligt bare fordi man lærer mere om det, men det som teorien forudsiger er faktisk det man får ud af at gentage målingen mange gange og tælle op.

En beroligende bemærkning er at selv om de Hermiteske matricer har komplekse indgange, så giver disse målinger altid reelle tal, fordi vi har jo set at hvis A er en hermitesk matrix, så er dens egenværdier relle. Det ville jo føles lidt uhyggeligt at sidde i et fly hvis piloten fortalte i speakeren at vi nu flyver i en højde af $10+3i$ kilometer.

Nu kan man spørge: Hvis en partikels position er en observabel, hvordan kan det være at der kun er endeligt mange muligheder for en partikels position? Kan vi ikke flytte den x meter i en bestemt retning, hvor x er et vilkårligt reelt tal? Det fulde svar på dette spørgsmål er ret indviklet, men den korte version er at i dette tilfælde skal man betragte Hermiteske operatorer på vektorrum som ikke er af endelig dimension, det vil sige, der findes ikke en basis for vektorrummet der består af endeligt mange vektorer.

Nu til et konkret eksempel. Spin er en observabel som vi ikke kender fra den makroskopiske verden. Vi kan udføre følgende type måling. Den minder dog en del om impulsmoment, der fortæller noget om hvor hurtig noget roterer omkring en akse. Givet en enhedsvektor $\mathbf{u}$ i rummet kan vi måle en partikels spin i denne retning ved at sende den gennem et passende magnetfelt, og se hvor meget den bliver afledet af magnetfeltet. Den sunde fornuft siger nu at denne afledning kan ske med et vinkel der variere kontinuert. Man det er ikke det der sker. Partiklen bliver afledet, enten opad eller nedad, med en bestemt konstant vinkel. Denne vinkel er given af partiklens "spin", som kan opfattes som en form for impulsmoment (angulært moment). Dette var et af de tidlige eksperimenter (Stern-Gerlach, 1922) som viste nødvændigheden af at betragte kvantisering, i dette tilfælde en kvantisering af "magnetisk impulsmoment", som altså kun kan antage to værdier. I beskrivelsen af eksperimentet har jeg tilladt mig nogle hvide løgne og udeladelserfor at koncentrere på det væsentlige.

Her er en model for dette fenomen. Lad os sige at vi betragter en elektron. Vi vil lave en måling af elektronens spin. For det første skal vi vælge enhedsvektoren

$\mathbf{u} = (u_x, u_y, u_z)^T \in \mathbb{R}^3$, $|\mathbf{u}| = 1$. For at beregne partiklens spin i retning $\mathbf{u}$, skal vi bruge en hermitesk operator.

Elektronen har en indre struktur som er givet som en tilstandsvektor $\mathbf{y} = (z_1, z_2) \in \mathbb{C}^2$. Denne vektor betegnes som elektronens spin. Observablen "spin" svarer til en bestemt hermitesk operator, men hvad for en? Vi indfører de tre berømte Paulimatricer. De er komplekse 2×2 matricer

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

9.9 Opgave

Vis at de tre Paulimatricer er hermiteske matricer. Vis at enhver hermitesk 2×2 matrix M kan skrives på en og kun en måde som

$$H = u_x \sigma_x + u_y \sigma_y + u_z \sigma_z + \lambda I_2$$

hvor u_x, u_u, u_z, λ er relle tal og $I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Overvej om Paulimatricerne kommuterer, det vil sige hvorvidt $P_i P_j = P_j P_i$ for forskellige i, j. ♠

Det viser sig at den hermiteske operator som beregner det spin vi observerer i retningen $\mathbf{u} = (u_x, u_y, u_z)^T \in \mathbb{R}^3$ er

$$H = u_x \sigma_x + u_y \sigma_y + u_z \sigma_z.$$

(9.10) EKSEMPEL.

Antag at $\mathbf{u} = \begin{pmatrix} 4/9 \\ 4/9 \\ 7/9 \end{pmatrix}$. Vi antager at en elektronen er i en tilstand som beskrives ved vektoren $\mathbf{v} = (\frac{1+i}{2}, \frac{1-i}{2})^T$. Vi ønsker at beregne det forventede udfald af at måle elektronens spin i retning $\mathbf{u}$.

Ved et rent held er begge de to vektorer enhedsvektorer. Målingen vil altså resultere i en egenvektor til matricen

$$H = \frac{4}{9}\sigma_x + \frac{4}{9}\sigma_y + \frac{7}{9}\sigma_z = \begin{pmatrix} \frac{7}{9} & \frac{4}{9} - \frac{4}{9}i \\ \frac{4}{9} + \frac{4}{9}i & -\frac{7}{9} \end{pmatrix}$$

Det karakteristiske polynomium for H er

$$\det(H - \lambda I_2) = \det \begin{pmatrix} \frac{7}{9} - \lambda & \frac{4}{9} - \frac{4}{9}i \\ \frac{4}{9} + \frac{4}{9}i & -\frac{7}{9} - \lambda \end{pmatrix} = \lambda^2 - 1.$$

Rødderne i det karakteristiske polynomium er $\lambda = 1$ og $\lambda = -1$. Der er altså kun to mulige resultater af målingen, nemlig ± 1. Lad os vedtage at $\lambda = 1$ svarer til at spinnet er rettet langs **u**, og $\lambda = -1$ til at spinnet er rettet modsat **u**. Denne vedtægt er i al væsentlighed et spørgsmål om notation.

Vi finder tilhørende egenvektorer ved at løse ligninger:

$$\psi_1 = (4, 1+i)^T$$
$$\psi_2 = (1, -2-2i)^T$$

ψ_1 er en egenvektor tilhørende egenværdi 1, og ψ_2 er en egenvektor hørende til egenværdi -1. For en sikkerheds skyld laver vi prøve, som man altid skal.

$$\begin{aligned}
\begin{pmatrix} \frac{7}{9} & \frac{4}{9} - \frac{4}{9}i \\ \frac{4}{9} + \frac{4}{9}i & -\frac{7}{9} \end{pmatrix} \begin{pmatrix} 4 \\ 1+i \end{pmatrix} &= \frac{1}{9} \begin{pmatrix} 7 \cdot 4 + (4-4i)(1+i) \\ (4+4i) \cdot 4 + (-7)(1+i) \end{pmatrix} \\
&= \frac{1}{9} \begin{pmatrix} 28 + (8) \\ 16 + 16i + (-7-7i) \end{pmatrix} \\
&= \begin{pmatrix} 4 \\ 1+i \end{pmatrix},
\end{aligned}$$

og

$$\begin{aligned}
\begin{pmatrix} \frac{7}{9} & \frac{4}{9} - \frac{4}{9}i \\ \frac{4}{9} + \frac{4}{9}i & -\frac{7}{9} \end{pmatrix} \begin{pmatrix} 1 \\ -2-2i \end{pmatrix} &= \frac{1}{9} \begin{pmatrix} 7 + (4-4i)(-2-2i) \\ 4 + 4i + (-7)(-2-2i) \end{pmatrix} \\
&= \frac{1}{9} \begin{pmatrix} 7 + (-16) \\ 4 + 4i + (14+14i) \end{pmatrix} \\
&= \begin{pmatrix} -1 \\ 2+2i \end{pmatrix}.
\end{aligned}$$

Vi udtrykker nu **v** i den ortogonale basis for $\mathbb{C}^2$ der består af ψ_1 og ψ_2.

$$\begin{pmatrix} \frac{1+i}{2} \\ \frac{1-i}{2} \end{pmatrix} = c_1 \begin{pmatrix} 4 \\ 1+i \end{pmatrix} + c_2 \begin{pmatrix} 1 \\ -2-2i \end{pmatrix}$$

Vektorerne ψ_i er ikke enhedsvektorer, og kvadraterne på deres normer er $|\psi_1|^2 = 16 + 2 = 18$, $|\psi_2|^2 = 1 + 8 = 9$. Ved at løse ligninger finder vi at $c_1 = \frac{1}{18}(2+i)$, $|c_1|^2 = 5/324$ og $c_2 = \frac{1}{18}(1+5i)$, $|c_2|^2 = 26/324$ Vi laver selvfølgelig prøve, fordi det gør man:

$$\begin{aligned}
&\frac{1}{18}(2+i) \begin{pmatrix} 4 \\ 1+i \end{pmatrix} + \frac{1}{18}(1+5i) \begin{pmatrix} 1 \\ -2-2i \end{pmatrix} \\
&\qquad = \frac{1}{18} \left[\begin{pmatrix} (2+i) \cdot 4 \\ (2+i)(1+i) \end{pmatrix} + \begin{pmatrix} 1+5i \\ (1+5i)(-2-2i) \end{pmatrix} \right] \\
&\qquad = \frac{1}{18} \left[\begin{pmatrix} 8+4i \\ 2-1+2i+i \end{pmatrix} + \begin{pmatrix} 1+5i \\ -2+10-2i-10i \end{pmatrix} \right] \\
&\qquad = \frac{1}{18} \begin{pmatrix} 9+9i \\ 9-9i \end{pmatrix} \\
&\qquad = \begin{pmatrix} \frac{1+i}{2} \\ \frac{1-i}{2} \end{pmatrix}
\end{aligned}$$

Modellen siger nu at sandsynligheden for at elektronens spin bliver målt til at være rettet langs **u** er $|c_1|^2|\psi_1|^2 = \frac{5}{364} \cdot 18 = 5/18$ og at sandsynligheden for at spinnet bliver modsat mod **u** er $|c_2|^2|\psi_2|^2 = \frac{26}{364} \cdot 9 = 13/18$.

Dem der synes at dette er alt for intuitivt og indlysende og har brug for noget mere hjernefrysende anbefales at fortsætte studiet med Bells ulighed.

9.4 En støtte vektor maskine

Artificiell intelligens er et stort område som bruger mange forskellige metoder. Nogle af de mest effektive af disse er inspireret fra biologi og neurologi, men vi vil ikke gå ind på den side af sagen. En almindelig forudsætning er at man har adgang til en meget stor database, og at hver enkelt element i denne database kan beskrives som en vektor i et vektorrum F^N af meget høj dimension. Det er altså tale om en stor mængde af vektorer $v_i \in F^N$. Et eksempel ville for eksempel være en samling af billeder, repræsenterede ved farven i hver enkelt pixel. Opgaven er at foretage sig noget intelligent med denne meget store datamængde. Fordi data er givet ved vektorer er det ikke så forbavsende at lineær algebra næsten altid spiller en stor rolle. Som et eksempel som er blevet brugt i seriøse anvendelser, vil vi diskutere en "support vector machine". For en mere udførlig beskrivelse af matematikken bag dette anbefaler vi en forelæsning på MIT af Patrick Winston. Selv om man ikke er interesseret i matematikken er historien spændende, Winston fortæller den fra ca 46:19.

For at forklare metoden vil vi arbejde med et konkret men lidt kunstigt eksempel. Antag at vi har en liste af byer, og at vi for hver by på listen kender dens longitud og lattitud (u_1,u_2) og desuden ved hvad land byen ligger i. Opgaven er at skrive et program som helt automatisk kan placere en by som ikke er på vores liste i det rigtige land. For at gøre det simpelt antager vi at vi kun betragter to lande.

I nogle situationer kunne det være nemt. For eksempel følger en stor del af grænsen mellem USA og Canada den 49ende parallel, så hvis vi kun interessere os for byer i dette område, kunne vi bruge funktionen $f(u_1,u_2) = u_2 - 49$. Hvis for en by med kordinater $\mathbf{u} = (u_1,u_2)^T$ gælder at $f(u_1,u_2) < 0$ så ligger byen i USA, og hvis $f(u_1,u_2) > 0$ så er den i Canada.

Mere generelt kunne man forestille sig lande med snorlige grænser som ikke nødvendigvis er longitud storcirkle eller latitud paraleller. Så kunne man finde en lineær funktion (med konstantterm b) $f(u_1,u_2) = a_1u_1 + a_2u_2 + b$ så at betingelsen $f(u_1,u_2) > 0$ afgjorde en bys tilhørighedsforhold. Vi kan bruge vores liste af kendte byer til at finde gode vurderinger af a_1,a_2,b.

Som et lidt vanskeligere eksempel kan vi prøve at skelne byer i Danmark fra byer i Norge. Der er ikke en oplagt grænse mellem Danmark og Norge, men vi kan stadig lægge en linje gennem Skagerak der har Danmark på den ene side og Norge på den anden. Det kan vi også fortolke som en lineær funktion $f(u_1,u_2) = a_1u_1 + a_2u_2 + b$. Vi vil gerne vælge denne funktion så at den skelner mellem de to muligheder så godt som det er muligt.

Vi vil ikke gå i detaljer med hvordan man finder denne funktion, men der eksisterer gode algoritmer som kan gøre det. De hører under noget der kaldes konveks optimering, som sådan set også bygger ovenpå lineær algebra. Man kan altså give listen med byer til en computer, og den producerer helt af sig selv en funktion f sådan at hvis $\mathbf{u}_i$ er koordinater for en af byerne på vores liste så er $f(\mathbf{u}_i) > 0$ hvis og kun hvis den tilsvarende by er en by i Danmark. Og når vi har funktionen f kan vi bruge den til at lave et rimeligt gæt på om en by med koordinater (u_1, u_2) er i Danmark eller Norge. Dette gæt er nemlig afgjort af fortegnet af $f(\mathbf{u})$, og det virker lige så fint hvis vi har to mængder D og N i et Euklidisk run $\mathbb{R}^n$. Men forudsætningen er at de to mængder D og N rent faktisk kan skelnes ad ved en lineær afbildning Hvis det ikke kan lade sig gøre bliver algoritmen ved i det uendelige med at prøve på at løse et uløseligt problem.

Men antag nu at vi er interesserede i byer der ligger enten i Sverige eller i Danmark. Selv om de to lande er skilt fra hinanden af Øresunds og Kattegats vande er det umuligt at skelne dem ad med en lineær funktion f. Specielt er Bornholm problematisk. Vi kan altså ikke bruge den gode algoritme.

Men vi kan gøre noget andet. Vi finder en ikke-lineær afbinding $\phi : \mathbb{R}^2 \to \mathbb{R}^k$. For eksempel kunne vi vælge $\phi(u_1, u_2) = (u_1u_1, u_1u_2, u_2u_1, u_2u_2)$, eller skrevet på matrix form $\phi(\mathbf{u}) = \mathbf{u}\mathbf{u}^T$.

9.11 Opgave

Vis at med det ϕ vi har valgt ovenfor er $\phi(\mathbf{u}) \cdot \phi(\mathbf{v}) = (\mathbf{u} \cdot \mathbf{v})^2$. ♠

Selv om vi ikke kan skelne mængden D fra S med en lineær afbildning, så kan vi måske skelne deres billeder $\phi(D) \subset \mathbb{R}^k$ og $\phi(S) \subset \mathbb{R}^k$ fra hinanden med en lineær afbildning. En pointe er nu at det væsentlige input i den fine algoritme som kan finde en lineær afbildning der skelner D fra S er netop de indre produkter $\mathbf{u}_i \cdot \mathbf{u}_j$. Så hvis vi kan skelne $\phi(D)$ fra $\phi(S)$ med en lineær afbildning, og hvis kan angive en formel for de indre produkter $\phi(\mathbf{u}_i) \cdot \phi(\mathbf{u}_j)$ så kan algoritmen fortælle os nøjagtig hvad for et $f : \mathbb{R}^k \to \mathbb{R}$ vi skal bruge. Det vil sige, hvis $\mathbf{u}_i$ er en af vores byer så er $\mathbf{u}_i$ i Danmark hvis $f(\phi(\mathbf{u}_i)) > 0$ og i Sverige hvis $f(\phi(\mathbf{u}_i)) < 0$.

Antag nu at $\mathbf{u}$ for eksempel er koordinaterne for Hesselø. Selv om Hesselø (Hesselö) ikke er på vores liste af kendte byer $\mathbf{u}_i$, kan vi lave det rimelige gæt at hvis $f(\phi(\mathbf{u})) > 0$ så hører Hesselø til Danmark. Vi kan selvfølgelig ikke vide det helt sikkert.

9.12 Opgave

Lad $D = \{(0,1), (1,0) \in \mathbb{R}^2\}$ og $S = \{(0,0), (1,1)\} \in \mathbb{R}^2\}$. Overvej at man ikke kan finde reelle tal a_1, a_2, b så at hvis $(u_1, u_2) \in D$ så er $a_1u_1 + a_2u_2 + b > 0$ og hvis $(u_1, u_2) \in S$ så er $a_1u_2 + a_2u_2 + b < 0$. Find reelle tal $a_{11}, a_{12}, a_{21}, a_{22}, b$ sådanne at hvis $(u_1, u_2) \in D$ så er

$$a_{11}u_1u_1 + a_{12}u_1u_2 + a_{21}u_2u_1 + a_{22}u_2u_2 + b > 0$$

men hvis $(u_1, u_2) \in S$ så er

$$a_{11}u_1u_1 + a_{12}u_1u_2 + a_{21}u_2u_1 + a_{22}u_2u_2 + b < 0$$

♠

9.5 Opgaver

Vi vil illustrere principperne ved følgende A matrix.

$$A = \begin{pmatrix} 38.5 & 28.5 & 33.0 \\ 28.5 & 38.5 & 33.0 \\ 33.0 & 33.0 & 34.0 \end{pmatrix}$$

9.5.1

Diagonaliser A matricen og vis at den originale A matrix kan rekonstrueres ud fra dens egenværdier og egenvektorer fra (9.10).

9.5.2

Udregn matricen A_1 ved den omvendte transformation (9.10) af kun egenvektoren med den største egenværdi (dvs. sæt de to andre egenværdier = 0). Beregn %-fejlen ud fra en sammenligning af alle elementerne af A og A_1 matricen.

9.5.3

Beregn %-fejlen ud fra en sammenligning af alle elementerne af A og matricen A_2, svarende til den omvendte transformation (9.10) med egenvektorer hørende til de to største egenværdier.

Indeks