Lineær algebra

Preben Alsholm
MAT, DTU

17. september 2002

1 Lineære ligningssystemer

Et system af m ligninger med n ubekendte $x_1, x_2, \ldots, x_n$ kaldes lineært, hvis det kan skrives på formen

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

hvor koefficienterne $a_{ij}, i = 1 \ldots m, j = 1 \ldots n$, og højresiderne $b_i, i = 1 \ldots m$, er konstanter, altså ikke afhængige af de ubekendte.

Ved løsningen af et lineært ligningssystem kan følgende operationer udføres på ligningssystemet uden at løsningsmængden derved ændres.

1. Et par af ligninger kan bytte plads.
2. En ligning kan ganges med en fra nul forskellig konstant.
3. Ligning nr. i kan erstattes med summen af ligning i og ligning j.
4. Ligning nr. i kan erstattes med summen af ligning i og et vilkårligt multiplum af ligning j.

Tilladeligheden af operation 4 følger ved kombination af operationerne 2 og 3.

Eksempel 1 *Betragt følgende system af 2 ligninger med 2 ubekendte*

$$\begin{aligned} 2x + 3y &= 6 \\ x - 4y &= 7 \end{aligned}$$

Vi ombytter ligning 1 og 2, hvorved vi får det dermed ækvivalente system

$$\begin{aligned} x - 4y &= 7 \\ 2x + 3y &= 6 \end{aligned}$$

Vi erstatter ligning 2 med ligning 2 minus 2 gange ligning 1. Herved får vi systemet

$$\begin{aligned} x - 4y &= 7 \\ 0x + 11y &= -8 \end{aligned}$$

Sidste ligning fortæller os åbenbart, at $y = -\frac{8}{11}$*. Herefter findes* x *af øverste ligning til* $x = 4y + 7 = -\frac{32}{11} + 7 = \frac{45}{11}$*. Der er altså netop én løsning til systemet og den er* $(x, y) = \left(\frac{45}{11}, -\frac{8}{11}\right)$*.*

Ved løsningen af lineære ligningssystemer opskrevet på den viste måde er det arbejdsbesparende at opskrive systemet i et rektangulært talarrangement (en matrix), hvor kun koefficienterne og højresiderne figurerer. Herved får det generelle ligningssystem vist ovenfor udseendet

$$T = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{pmatrix}$$

Man kalder denne matrix for systemets *totalmatrix*. Matricen bestående af kun koefficienterne til de ubekendte kaldes *koefficientmatricen* for systemet:

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

Denne sidste matrix A har m rækker og n søjler. Totalmatricen T har også m rækker, men har $n+1$ søjler. Man siger, at A er en $m \times n$ matrix, og at T er en $m \times (n+1)$ matrix. Rækkeantallet angives altså først. En matrix med lige mange rækker og søjler kaldes en kvadratisk matrix.

Eksempel 2 *For ligningssystemet i eksemplet ovenfor er totalmatricen*

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

og koefficientmatricen er

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

Den første er en 2×3 *matrix, den sidste er en* 2×2 *matrix og er dermed en kvadratisk matrix.*

De operationer nævnt ovenfor, som kan udføres på et ligningssystem uden at ændre på løsningsmængden, kan nu formuleres i matrixsprog på følgende måde.

Ved løsningen af et lineært ligningssystem kan følgende operationer udføres på totalmatricen uden at løsningsmængden derved ændres.

1. Et par rækker kan ombyttes.
2. En række kan ganges med en fra nul forskellig konstant.
3. Række nr. i kan erstattes med summen af række i og række j.
4. Række nr. i kan erstattes med summen af række i og et vilkårligt multiplum af række j, når $i \neq j$.

Idet række nr. i betegnes R_i kan disse operationer skrives således:

1. $R_i \longleftrightarrow R_j$
2. $R_i := kR_i$, hvor $k \neq 0$
3. $R_i := R_i + R_j$
4. $R_i := R_i + kR_j$ for alle k, når $i \neq j$.

Eksempel 3 *Vi vender tilbage til ligningssystemet*

$$\begin{aligned} 2x + 3y &= 6 \\ x - 4y &= 7 \end{aligned}$$

med toalmatricen

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

På denne udfører vi først rækkeoperationen $R_1 \longleftrightarrow R_2$*, hvorved vi får*

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

Herefter udfører vi rækkeoperationen $R_2 := R_2 - 2R_1$*, hvorved vi får*

$$\begin{pmatrix} 1 & -4 & 7 \\ 0 & 11 & -8 \end{pmatrix}$$

Vi gik ikke længere ovenfor, men vi kunne fortsætte med rækkeoperationen $R_2 := \frac{1}{11}R_2$*, hvorved vi får*

$$\begin{pmatrix} 1 & -4 & 7 \\ 0 & 1 & -\frac{8}{11} \end{pmatrix}$$

Herefter foretager vi rækkeoperationen $R_1 := R_1 + 4R_2$*, hvorved vi får*

$$\begin{pmatrix} 1 & 0 & \frac{45}{11} \\ 0 & 1 & -\frac{8}{11} \end{pmatrix}$$

Løsningen $(x, y) = \left(\frac{45}{11}, -\frac{8}{11}\right)$ *kan nu direkte aflæses i sidste søjle, idet den sidste totalmatrix gemmer på følgende ligningssystem*

$$\begin{aligned} x + 0y &= \frac{45}{11} \\ 0x + y &= -\frac{8}{11} \end{aligned}$$

Fremgangsmåden brugt i eksemplet ovenfor er et eksempel på Gauss' eliminationsmetode, der forløber som følger:

1. Ombyt om nødvendigt rækker, så række nr. 1 har første element forskellig fra nul. Hvis søjle 1 består af lutter nuller, er dette ikke muligt. I så fald går man blot til søjle 2 i stedet. I beskrivelsen nedenfor er $A = (a_{ij})$ den nye matrix.

2. Skaf nuller i resten af søjle 1 ved rækkeoperationerne $R_i := R_i - \frac{a_{i1}}{a_{11}} R_1$ for $i = 2 \dots m$. Ved disse operationer bruges hele tiden række 1 som *pivoteringsrække*. Elementet a_{11} er *pivoteringselement*.

3. Ombyt om nødvendigt række 2 med en af rækkerne $3 \dots m$, så den nye række 2 har andet element forskellig fra nul.Er alle elementerne $a_{i2} = 0$ for $i = 2 \dots m$ er dette ikke muligt. I så fald går man blot til søjle 3 i stedet. I beskrivelsen nedenfor er $A = (a_{ij})$ den nye matrix.

4. Skaf nuller i resten af søjle 2 ved rækkeoperationerne $R_i := R_i - \frac{a_{i2}}{a_{22}} R_2$ for $i = 3 \dots m$. Her bruges række 2 som pivoteringsrække. a_{22} er pivoteringselement.

5. Fortsæt denne procedure med søjlerne $3 \dots n$ i rækkefølge.

6. Løs det tilsvarende ligningssystem nedefra og op.

Ved denne fremgangsmåde opnås til sidst en matrix på såkaldt *echelonform*, dvs. en matrix hvor første søjle enten har lutter nuller eller også har nuller alle andre steder end i første række, anden søjle har enten nuller i de samme rækker som første søjle eller har nuller fra og med anden række, tredie søjle har enten nuller i samme rækker som anden søjle eller også starter den med nuller en række senere, osv. Vi giver nogle eksempler.

Eksempel 4 *Betragt ligningssystemet*

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 4 \\ 2x_1 + 5x_2 &= -2 \\ 3x_1 + 3x_2 + x_3 &= 2 \\ 2x_1 + 5x_2 + x_3 &= 3 \end{aligned}$$

Totalmatricen er

$$T = \begin{pmatrix} 1 & -2 & 1 & 4 \\ 2 & 5 & 0 & -2 \\ 3 & 3 & 1 & 2 \\ 2 & 5 & 1 & 3 \end{pmatrix}$$

Vi udfører Gausselimination som følger. Først operationerne $R_2 := R_2 - 2R_1, R_3 := R_3 - 3R_1, R_4 := R_4 - 2R_1$, hvor R_1 bliver brugt som pivoteringsrække. Herved *opnås*

$$T_1 = \begin{pmatrix} 1 & -2 & 1 & 4 \\ 0 & 9 & -2 & -10 \\ 0 & 9 & -2 & -10 \\ 0 & 9 & -1 & -5 \end{pmatrix}$$

Herefter følger operationerne $R_3 := R_3 - R_2, R_4 := R_4 - R_2$, hvor R_2 bliver brugt som pivoteringsrække. Herved opnås

$$T_2 = \begin{pmatrix} 1 & -2 & 1 & 4 \\ 0 & 9 & -2 & -10 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 5 \end{pmatrix}$$

For at få matricen på echelonform udføres nu yderligere operationen $R_3 \longleftrightarrow R_4$, hvorefter vi har

$$T_2 = \begin{pmatrix} 1 & -2 & 1 & 4 \\ 0 & 9 & -2 & -10 \\ 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Det tilsvarende ligningssystem er

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 4 \\ 9x_2 - 2x_3 &= -10 \\ x_3 &= 5 \\ 0 &= 0 \end{aligned}$$

Ved løsning nedefra fås af $x_3 = 5$, at $x_2 = \frac{1}{9}(2x_3 - 10) = 0$ og dernæst, at $x_1 = 2x_2 - x_3 + 4 = -1$. Lignigssystemet har altså netop én løsning, og den er $x = (-1, 0, 5)$.

Eksempel 5 *Betragt ligningssystemet*

$$\begin{aligned} x_1 + 2x_2 + 3x_3 + 4x_4 &= 5 \\ 6x_1 + 7x_2 + 8x_3 + 9x_4 &= 35 \\ -x_1 + 2x_2 - 3x_3 + 4x_4 &= 7 \\ x_1 + x_2 + x_3 + x_4 &= a \end{aligned}$$

hvor a er en reel konstant. Totalmatricen er

$$T = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 6 & 7 & 8 & 9 & 35 \\ -1 & 2 & -3 & 4 & 7 \\ 1 & 1 & 1 & 1 & a \end{pmatrix}$$

Vi udfører Gausselimination som følger. Først pivoteres med række 1 i operationerne $R_2 := R_2 - 6R_1, R_3 := R_3 + R_1, R_4 := R_4 - R_1$. Herved opnås matricen

$$T_1 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & -5 & -10 & -15 & 5 \\ 0 & 4 & 0 & 8 & 12 \\ 0 & -1 & -2 & -3 & a-5 \end{pmatrix}$$

For at gøre de følgende operationer simplere udføres nu operationen $R_2 := -\frac{1}{5}R_2$. Herved opnås

$$T_2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 1 & 2 & 3 & -1 \\ 0 & 4 & 0 & 8 & 12 \\ 0 & -1 & -2 & -3 & a-5 \end{pmatrix}$$

Dernæst operationerne $R_3 := R_3 - 4R_2, R_4 := R_4 + R_2$ med R_2 som pivoteringsrække. Herved opnås

$$T_3 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 1 & 2 & 3 & -1 \\ 0 & 0 & -8 & -4 & 16 \\ 0 & 0 & 0 & 0 & a-6 \end{pmatrix}$$

Matricen er nu på echelonform. Vi kan dog simplificere yderligere ved operationen $R_3 := -\frac{1}{8}R_3$. Herved opnås

$$T_4 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 1 & 2 & 3 & -1 \\ 0 & 0 & 1 & \frac{1}{2} & -2 \\ 0 & 0 & 0 & 0 & a-6 \end{pmatrix}$$

Det tilsvarende ligningssystem ser således ud

$$\begin{aligned} x_1 + 2x_2 + 3x_3 + 4x_4 &= 5 \\ x_2 + 2x_3 + 3x_4 &= -1 \\ x_3 + \frac{1}{2}x_4 &= -2 \\ 0 &= a - 6 \end{aligned}$$

Den nederste ligning indeholder slet ingen ubekendte. Denne ligning er opfyldt, hvis og kun hvis konstanten a (tilfældigvis) har værdien 6, ellers ikke. Vi må altså konkludere, at hvis $a \neq 6$, så har ligningssystemet ingen løsning. Hvis derimod $a = 6$, så er nederste ligning opfyldt, og vi må nu sørge for, at de andre 3 ligninger også er opfyldt. Vi kan åbenbart sætte x_4 til hvad som helst. Sæt altså $x_4 = t$, hvor $t \in R$. Herefter giver tredie ligning, at $x_3 = -2 - \frac{1}{2}x_4 = -2 - \frac{1}{2}t$. Anden ligning giver nu, at $x_2 = -1 - 2x_3 - 3x_4 = -1 - 2\left(-2 - \frac{1}{2}t\right) - 3t = 3 - 2t$.

Til sidst giver første ligning, at $x_1 = 5 - (2x_2 + 3x_3 + 4x_4) = 5 + \frac{3}{2}t$. *Altså er løsningerne givet ved*

$$\begin{aligned}(x_1, x_2, x_3, x_4) &= \left(5 + \frac{3}{2}t, 3 - 2t, -2 - \frac{1}{2}t, t\right) \\ &= (5, 3, -2, 0) + t\left(\frac{3}{2}, -2, -\frac{1}{2}, 1\right)\end{aligned}$$

hvor $t \in R$.

Bemærkning 6 *Hvis et ligningssystem skal løses med flere forskellige højre sider, men med fast koefficientmatrix, så betaler det sig at løse systemerne på én gang. Totalmatricen består så af koefficientmatricen fulgt af lige så mange søjler, som der er højresider.*

Bemærkning 7 *Når en totalmatrix for et ligningssystem er reduceret til echelonform, anses det normalt for mest arbejdsbesparende ikke at foretage yderligere rækkeoperationer, men blot fortsætte med løsning nedefra og op. Undertiden kan det dog være hensigtsmæssigt at fortsætte med rækkeoperationer, der (når det iøvrigt er muligt) reducerer koefficientmatricen til en matrix, hvis diagonalelementer består af ettaller og hvis øvrige elementer er nul. En sådan total reduktion kaldes Gauss-Jordan-reduktion. Den består af Gauss-elimination plus Gausselimination med "hovedet nedad". Vi illustrerer metoden med et eksempel nedenfor.*

Eksempel 8 *Vi vil løse ligningssystemet*

$$\begin{aligned} x + 5y - 2z &= a \\ 5x + 7y - 3z &= b \\ -2x - y &= c \end{aligned}$$

for to forskellige højre sider, nemlig for $(a, b, c) = (3, 0, 9)$ *og for* $(a, b, c) = (6, 0, -9)$. *Dette kan gøres enten ved at løse systemet med højresiden som bogstaver først og dernæst indsætte værdier, eller ved at medtage begge konkrete højresider i totalmatricen. Vi prøver det sidste. Desuden agter vi at bruge Gauss-Jordan-reduktion. Totalmatricen med de to højresider er*

$$T = \begin{pmatrix} 1 & 5 & -2 & 3 & 6 \\ 5 & 7 & -3 & 0 & 0 \\ -2 & -1 & 0 & 9 & -9 \end{pmatrix}$$

Vi udfører først sædvanlig Gausselimination. Ved rækkeoperationerne $R_2 := R_2 - 5R_1, R_3 := R_3 + 2R_1$ *opnås*

$$T_1 = \begin{pmatrix} 1 & 5 & -2 & 3 & 6 \\ 0 & -18 & 7 & -15 & -30 \\ 0 & 9 & -4 & 15 & 3 \end{pmatrix}$$

Dernæst rækkeoperationen $R_3 := R_3 + \frac{1}{2}R_2$, der giver

$$T_2 = \begin{pmatrix} 1 & 5 & -2 & 3 & 6 \\ 0 & -18 & 7 & -15 & -30 \\ 0 & 0 & -\frac{1}{2} & \frac{15}{2} & -12 \end{pmatrix}$$

Hermed er totalmatricen på echelonform, og de to ligningssystemer kunne løses nedefra og op. Vi går imidlertid videre. Vores formål er at skabe nuller uden for diagonalen i koefficientmatricen og ettaller i diagonalen. Vi begynder i sidste række med operationen $R_3 := -2R_3$, hvorved opnås

$$T_3 = \begin{pmatrix} 1 & 5 & -2 & 3 & 6 \\ 0 & -18 & 7 & -15 & -30 \\ 0 & 0 & 1 & -15 & 24 \end{pmatrix}$$

Herefter skaber vi nuller i tredie søjles anden og første række ved rækkeoperationerne $R_2 := R_2 - 7R_3, R_1 := R_1 + 2R_3$, hvor R_3 bruges som pivoteringsrække. Vi finder så

$$T_4 = \begin{pmatrix} 1 & 5 & 0 & -27 & 54 \\ 0 & -18 & 0 & 90 & -198 \\ 0 & 0 & 1 & -15 & 24 \end{pmatrix}$$

Vi skaber nu et ettal i position $(2, 2)$ ved operationen $R_2 := -\frac{1}{18}R_2$. Herved opnås

$$T_5 = \begin{pmatrix} 1 & 5 & 0 & -27 & 54 \\ 0 & 1 & 0 & -5 & 11 \\ 0 & 0 & 1 & -15 & 24 \end{pmatrix}$$

Vores sidste operation er nu $R_1 := R_1 - 5R_2$, der giver

$$T_6 = \begin{pmatrix} 1 & 0 & 0 & -2 & -1 \\ 0 & 1 & 0 & -5 & 11 \\ 0 & 0 & 1 & -15 & 24 \end{pmatrix}$$

De to tilsvarende ligningssystemer er nu

$$\begin{aligned} x + 0y + 0z &= -2 \\ 0x + y + 0z &= -5 \\ 0x + 0y + z &= -15 \end{aligned}$$

og

$$\begin{aligned} x + 0y + 0z &= -1 \\ 0x + y + 0z &= 11 \\ 0x + 0y + z &= 24 \end{aligned}$$

Som vi ser, er højresiderne simpelthen løsningerne til de to systemer.

Vi har i disse eksempler nu set de tilfælde (hvad løsningsmængden angår), der overhovedet kan forekomme ved løsning af lineære ligningssystemer. Der er enten ingen løsninger, præcis én løsning, eller uendeligt mange løsninger givet ved en eller flere parametre. For bedre at kunne beskrive situationen indfører vi begrebet rang af matrix.

Definition 9 *Lad A være en matrix. Antallet af fra nulrækken forskellige rækker i en echelonform for A kaldes matricens rang. Vi skal betegne den med $\rho(A)$.*

Bemærkning 10 *Da echelonformen for en matrix ikke er entydigt defineret, bør man selvfølgelig sikre sig, at rangen alligevel er veldefineret. Det undlader vi på dette sted. Se dog senere i dette afsnit.*

Sætning 11 *Lad A være en $m \times n$ matrix. Så gælder, at $\rho(A) \leq \min(m, n)$.*

Bevis. At $\rho(A) \leq m$ følger umiddelbart. At også $\rho(A) \leq n$ er klart hvis $n \geq m$. Hvis $n < m$ udnytter vi, at i en echelonform for A har række R_{i+1} mindst ét ledende nul mere end række R_i, altså har generelt række R_i mindst $i-1$ ledende nuller. Men dette betyder, at R_{n+1} og følgende rækker er nulrækker. ■

Sætning 12 *Betragt et lineært ligningssystem bestående af m ligninger med n ubekendte. Lad T være systemets totalmatrix og A dets koefficientmatrix. Så gælder*

1. *Hvis $\rho(T) > \rho(A)$, så har ligningssystemet ingen løsning.*
2. *Hvis $\rho(T) = \rho(A) = n$, så har systemet præcis én løsning.*
3. *Hvis $\rho(T) = \rho(A) < n$, så har systemet uendeligt mange løsninger, til hvis beskrivelse kræves $n - \rho(A)$ parametre.*

Bevis. Bringes totalmatricen T på echelonform T_e er samtidigt koefficientmatricen A bragt på echelonform A_e. Når $\rho(T) > \rho(A)$ må T_e indeholde en række bestående af $n-1$ nuller fulgt af et tal forskellig fra nul. Denne række svarer til en ligning, der siger $0x_1 + 0x_2 + \ldots + 0x_n = a \neq 0$, der åbenbart ikke kan opfyldes. Hvis $\rho(T) = \rho(A) = n$, må række R_i i A_e indeholde præcis $i-1$ ledende nuller. Der kommer altså netop et nyt nul med ved hver overgang fra række R_i til R_{i+1}. Heraf følger specielt, at den sidste fra nulrækken forskellige række R_n består af $n-1$ ledende nuller fulgt af et fra nul forskelligt tal. Denne række svarer til ligningen $0x_1 + 0x_2 + \ldots + 0x_{n-1} + ax_n = b$, hvor altså $a \neq 0$. Heraf fås $x_n = \frac{b}{a}$. Række R_{n-1} svarer til en ligning af formen $0x_1+0x_2+\ldots+0x_{n-2}+cx_{n-1}+dx_n = h$, hvor $c \neq 0$. Dermed er x_{n-1} bestemt. Vi ser, at vi kan fortsætte således. Hermed bliver $(x_1, x_2, \ldots, x_n)$ entydigt bestemt. Hvis $\rho(T) = \rho(A) < n$, så er der en eller flere "dobbeltindrykninger" i echelonformen, dvs. ved overgang fra R_i til R_{i+1} vil der for en eller flere værdier af i komme mere end et ekstra ledende nul. Ved løsningen af den ligning, der svarer til R_i, vil vi da savne kendskab til (mindst) én af de ubekendte. Denne sættes til en vilkårlig værdi (sættes lig en parameter). Antallet af ekstra nuller fremkommet ved dobbeltindrykninger er $n-\rho(A)$. Dette må også være det antal parametre vi får brug for ved beskrivelsen af løsningerne. ■

Korollar 13 *Hvis ligningssystemet bestående af m ligninger med n ubekendte er homogent, dvs. hvis højresiderne $b_1, b_2, \ldots, b_m$ alle er nul, så gælder, at hvis $\rho(A) = n$, så har systemet kun nulløsningen, og hvis $\rho(A) < n$, så har systemet uendeligt mange løsninger, til hvis beskrivelse kræves $n-\rho(A)$ parametre. Specielt gælder altså, at et homogent system med flere ubekendte end ligninger har uendeligt mange løsninger.*

2 Underrum, lineær uafhængighed, dimension

Et talsæt $(x_1, x_2, \ldots, x_n) \in R^n$ vil vi i disse sammenhænge kalde en vektor. Mængden af sådanne talsæt, altså R^n vil vi da kalde et vektorrum. På samme måde kan C^n betragtes som et vektorrum.

Definition 14 *En ikke-tom delmængde V af vektorer fra R^n kaldes et underrum af R^n, hvis der gælder, at*

$$\begin{aligned} u, v &\in V \Longrightarrow u + v \in V \\ s &\in R \wedge u \in V \Longrightarrow su \in V \end{aligned}$$

Bemærkning 15 *Med ord siger man, at V er stabil overfor addition og multiplikation med skalar.*

Eksempel 16 *Lad $V = \{(x_1, x_2, 0) \,|x_1, x_2 \in R\}$. Så er V et underrum af R^3.*

Eksempel 17 *Lad $V = \{(x_1, x_2, 1) \,|x_1, x_2 \in R\}$. Så er V ikke et underrum af R^3.*

Eksempel 18 *Mængden bestående blot af nulvektoren er et underrum af R^n. R^n er et underrum af R^n.*

Sætning 19 *Lad $v_1, v_2, \ldots, v_k$ være vektorer fra R^n. Mængden af linearkombinationer*

$$span\,(v_1, v_2, \ldots, v_k) = \{c_1v_1 + c_2v_2 + \ldots + c_kv_k \,|c_1, c_2, \ldots, c_k \in R\}$$

er et underrum af R^n.

Bevis. Vi skal vise, at hvis $u, v \in span\,(v_1, v_2, \ldots, v_k)$ og s er et tal, så har vi, at $u + v \in span\,(v_1, v_2, \ldots, v_k)$ og $su \in span\,(v_1, v_2, \ldots, v_k)$. Med $u = c_1v_1 + c_2v_2 + \ldots + c_kv_k$ og $v = d_1v_1 + d_2v_2 + \ldots + d_kv_k$, så har vi

$$\begin{aligned} u + v &= c_1v_1 + c_2v_2 + \ldots + c_kv_k + d_1v_1 + d_2v_2 + \ldots + d_kv_k \\ &= (c_1 + d_1)\,v_1 + (c_2 + d_2)\,v_2 + \ldots + (c_k + d_k)\,v_k \in span\,(v_1, v_2, \ldots, v_k) \end{aligned}$$

og også

$$\begin{aligned} su &= s\left(c_1v_1 + c_2v_2 + \ldots + c_kv_k\right) \\ &= sc_1v_1 + sc_2v_2 + s\ldots + sc_kv_k \in span\left(v_1, v_2, \ldots, v_k\right) \end{aligned}$$

■

Bemærkning 20 *Vi skal kalde span* $(v_1, v_2, \ldots, v_k)$ *for underrummet udspændt af vektorerne* $v_1, v_2, \ldots, v_k$.

Definition 21 *Vektorerne* $v_1, v_2, \ldots, v_k$ *fra* R^n *kaldes lineært uafhængige, hvis*

$$c_1v_1 + c_2v_2 + \ldots + c_kv_k = \overline{0} \Longrightarrow c_1 = c_2 = \ldots = c_k = 0$$

altså hvis ingen linearkombination af vektorerne kan være lig med nulvektoren undtagen den trivielle linearkombination $0v_1 + 0v_2 + \ldots + 0v_k$. *Vektorerne kan lineært afhængige, hvis de ikke er lineært uafhængige.*

Bemærkning 22 *Et sæt vektorer* $v_1, v_2, \ldots, v_k$ *fra* R^n *er lineært afhængigt, hvis (mindst) én af dem kan skrives som en linearkombination af de øvrige. Hvis nemlig sættet ikke er lineært uafhængigt må der åbenbart findes konstanter* $c_1, c_2, \ldots, c_k$, *der ikke alle er nul, så* $c_1v_1 + c_2v_2 + \ldots + c_kv_k = \overline{0}$. *Hvis* $c_i \neq 0$, *har vi*

$$v_i = -\frac{c_1}{c_i}v_1 - \frac{c_2}{c_i}v_2 - \ldots - \frac{c_{i-1}}{c_i}v_{i-1} - \frac{c_{i+1}}{c_i}v_{i+1} - \ldots - \frac{c_k}{c_i}v_k$$

Bemærkning 23 *To vektorer er lineært afhængige, hvis og kun hvis de er proportionale, altså hvis og kun hvis den ene er en skalar gange den anden.*

Eksempel 24 *Vektorerne* $v_1 = (1, 2, 3, 4)\,, v_2 = (0, 4, 5, -3)\,, v_3 = (0, 0, -11, 2)$ *er lineært uafhængige, da*

$$\begin{aligned} c_1v_1 + c_2v_2 + c_3v_3 &= (0, 0, 0, 0) \Longrightarrow \\ c_1 &= 0 \wedge 2c_1 + 4c_2 = 0 \wedge 3c_1 + 5c_2 - 11c_3 = 0 \Longrightarrow \\ c_1 &= 0 \wedge c_2 = 0 \wedge c_3 = 0 \end{aligned}$$

Bemærkning 25 *På samme måde som i eksemplet indses det, at de fra nulrækken forskellige rækker i en matrix på echelonform er lineært uafhængige.*

Definition 26 *Lad* V *være et underrum af* R^n. *Mængden af vektorer* $\{v_1, v_2, \ldots, v_k\}$ *siges at udgøre en basis for* V, *hvis den er lineært uafhængig, og hvis enhver vektor i* V *kan skrives som en linearkombination af vektorerne* $v_1, v_2, \ldots, v_k$. *Hvis* $V = \left\{\overline{0}\right\}$, *vil vi sige, at basis er den tomme mængde af vektorer, i overensstemmelse med konventionen om, at en sum over den tomme mængde er nul* $\sum_{i \in \emptyset} v_i = 0$.

Sætning 27 *Ethvert underrum* V *af* R^n *har en basis. Enhver basis for* V *har samme antal vektorer.*

Bevis. Hvis $V = \{\overline{0}\}$ er basis den tomme mængde. Hvis $V \neq \{\overline{0}\}$, så vælger vi en vektor $v_1 \in V \setminus \{\overline{0}\}$. Hvis enhver anden vektor i V er proportional med v_1, så er $\{v_1\}$ en basis for V. Hvis ikke, så vælg en vektor v_2, der ikke er proportional med v_1. Sættet $\{v_1, v_2\}$ er åbenbart lineært uafhængigt. Hvis enhver vektor V kan skrives som en linearkombination af v_1, v_2, så er $\{v_1, v_2\}$ en basis for V. Hvis ikke, så vælg en vektor v_3, der ikke kan skrives som en linearkombination af v_1 og v_2. Sættet $\{v_1, v_2, v_3\}$ er lineært uafhængigt, da $c_1v_1 + c_2v_2 + c_3v_3 = \overline{0}$ åbenbart ikke kan være tilfældet med et $c_3 \neq 0$, men så har vi $c_1v_1 + c_2v_2 = \overline{0}$, hvoraf følger $c_1 = c_2 = 0$, da v_1, v_2 er lineært uafhængige. Således fortsættes. Processen må imidlertid stoppe, senest når n vektorer $v_1, v_2, \ldots, v_n$ er fundet. Thi antag, at $v_1, v_2, \ldots, v_n, v_{n+1}$ var et lineært uafhængigt sæt af vektorer. Så ville vektorligningen $c_1v_1 + c_2v_2 + \ldots + c_{n+1}v_{n+1} = \overline{0}$ kun have nulløsningen $(c_1, c_2, \ldots, c_{n+1}) = (0, 0, \ldots, 0)$. Men denne vektorligning kan skrives ud som et homogent system af n ligninger med $n + 1$ ubekendte. Et sådant system har uendeligt mange løsninger, ikke kun nulløsningen. Hermed er vist, at ethvert underrum har en basis.

Vi viser nu, at enhver basis for V har samme antal vektorer. Lad da $\{v_1, v_2, \ldots, v_k\}$ være en basis for V. Lad $\{u_1, u_2, \ldots, u_{k+1}\}$ være et sæt af $k + 1$ vektorer fra V. Vi vil vise, at dette sæt nødvendigvis er lineært afhængigt, og således ikke kan være en basis. Vi skal altså vise, at vektorligningen $c_1u_1 + c_2u_2 + \ldots + c_{k+1}u_{k+1} = \overline{0}$ har andre løsninger end nulløsningen $(c_1, c_2, \ldots, c_{k+1}) = (0, 0, \ldots, 0)$. Da $\{v_1, v_2, \ldots, v_k\}$ er en basis for V, findes der tal a_{ij}, så

$$u_i = \sum_{j=1}^{k} a_{ij} v_j$$

Ligningen $c_1u_1 + c_2u_2 + \ldots + c_{k+1}u_{k+1} = \overline{0}$ kan kortere skrives $\sum_{i=1}^{k+1} c_{i+1}u_i = \overline{0}$ og dermed

$$\overline{0} = \sum_{i=1}^{k+1} c_{i+1} \left(\sum_{j=1}^{k} a_{ij} v_j \right) = \sum_{j=1}^{k} \left(\sum_{i=1}^{k+1} c_{i+1} a_{ij} \right) v_j$$

Da nu $\{v_1, v_2, \ldots, v_k\}$ er lineært uafhængigt, fås deraf, at

$$\sum_{i=1}^{k+1} c_{i+1} a_{ij} = 0$$

for alle $i = 1 \ldots k$. Men dermed har vi et homogent ligningssystem bestående af k ligninger med $k + 1$ ubekendte. Et sådant system har uendeligt mange løsninger. ■

Definition 28 *Dimensionen af et underrum V af R^n er antallet af vektorer i en basis.*

Eksempel 29 *En simpel basis for R^n er $\{e_1, e_2, \ldots, e_n\}$, hvor $e_1 = (1, 0, 0, \ldots, 0)$, $e_2 = (0, 1, 0, \ldots, 0), \ldots, e_n = (0, 0, \ldots, 0, 1)$. Dimensionen af R^n er altså n. Denne basis for R^n kaldes undertiden den kanoniske basis.*

Eksempel 30 *Dimensionen af underrummet* $\{\overline{0}\}$ *af* R^n *er nul, da basis er den tomme mængde af vektorer.*

Sætning 31 *Det maksimale antal lineært uafhængige vektorer, der kan udtages fra et underrum* V *af* R^n *er lig med dimensionen. Ethvert maksimalt sæt af lineært uafhængige vektorer i* V *udgør en basis for* V.

Sætning 32 *Matricen* A *har rang* r, *hvis og kun hvis det maksimale antal lineært uafhængige rækker, der kan udtages fra* A *er* r.

Bevis. Er en samling rækker lineært uafhængige vil den nye samling, der opnås ved de tilladelige rækkeoperationer 1-3, stadig være lineært uafhængige. De fra nulrækken forskellige rækker i en echelonform for A udgør en maksimal mængde af lineært uafhængige vektorer blandt rækkerne i echelonformen. Heraf følger sætningen. ■

Bemærkning 33 *At rangen for en matrix er veldefineret følger af sætningen ovenfor, idet det maksimale antal lineært uafhængige rækkevektorer er et veldefineret begreb.*

Bemærkning 34 *Vil man afgøre om et sæt vektorer fra* R^n *er lineært uafhængige, kan man lade dem være rækkerne i en matrix og udføre Gausselimination på denne. På denne måde kan man også bestemme en basis for* $span\,(v_1, v_2, \ldots, v_k)$.

Eksempel 35 *Vektorerne* $v_1 = (1, 2, 3, 4, 5)\,, v_2 = (1, 1, 1, 1, 1)\,, v_3 = (1, 0, 1, 0, 1)\,, v_4 = (4, 5, 8, 9, 12)$ *udspænder et vektorrum* V. *Vi vil først finde en basis for V. Vi opskriver vektorerne som rækkerne i en matrix* A

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 4 & 5 & 8 & 9 & 12 \end{pmatrix}$$

Vi udfører sædvanlig Gausselimination og får

$$U = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & -1 & -2 & -3 & -4 \\ 0 & 0 & 2 & 2 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

En basis for $V = span\,(v_1, v_2, v_3, v_4)$ *består dermed af vektorerne* $u_1 = (1, 2, 3, 4, 5)\,, u_2 = (0, -1, -2, -3, -4)\,, u_3 = (0, 0, 2, 2, 4)$. *Dermed har vi, at* $\dim V = 3$. *Da ingen rækkeombytninger var nødvendige, må en anden basis være* $\{v_1, v_2, v_3\}$.

3 Matricer, decimaltal og computere

Her følger blot nogle advarende ord i form af to eksempler, der iøvrigt er taget fra Gilbert Strang, *Linear Algebra* (Han siger selv, at han har taget dem fra bøger af Noble og Forsythe & Moler).

Eksempel 36 *Rangen af matricen*

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1.0001 \end{pmatrix}$$

ses hurtigt ved rækkeoperationen $R_2 := R_2 - R_1$ *at være 2. Vi ser dog, at anden række i den reducerede matrix er* $(0, 0.0001)$. *Hvis den sidste decimal i* a_{22} *var nul i stedet for 1, ville rangen have været 1. Rangen af en matrix kan altså være meget følsom overfor tilfældige afrundingsfejl i ens beregninger. Betragt nu først ligningssystemet*

$$\begin{aligned} x + y &= 2 \\ x + 1.0001y &= 2 \end{aligned}$$

Løsningen ses at være $(x, y) = (2, 0)$. *Betragt så ligningssystemet*

$$\begin{aligned} x + y &= 2 \\ x + 1.0001y &= 2.0001 \end{aligned}$$

der kun afviger meget lidt fra det første. Løsningen til dette system er $(x, y) = (1, 1)$, *ganske langt fra løsningen til det første system.*

Eksempel 37 *Betragt nu matricen*

$$A = \begin{pmatrix} 0.0001 & 1 \\ 1 & 1 \end{pmatrix}$$

Denne matrix har klart rang lig med 2. Det har den også selv om 0.0001 *ændres til* 0. *Lad os løse ligningssystemet*

$$\begin{pmatrix} 0.0001 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

Laver vi standard Gausselimination på totalmatricen og benytter første række som pivoteringsrække, får vi

$$\begin{pmatrix} 0.0001 & 1 & 1 \\ 0 & -9999 & -9998 \end{pmatrix}$$

Heraf følger, at $y = \frac{9998}{9999} \cong 0.999\,90$ *(med 5 betydende cifre). Den ubekendte x findes af* $0.0001x + \frac{9998}{9999} = 1$ *til* $x = 10000\left(1 - \frac{9998}{9999}\right) \cong 1.000\,1$. *Dette er det resultat vi burde få. Men antag nu, at der regnes med 3 betydende cifre. Så ville totalmatricen efter Gausselimination se således ud:*

$$\begin{pmatrix} 0.0001 & 1 & 1 \\ 0 & -10000 & -10000 \end{pmatrix}$$

Heraf findes $y = 1$ *(helt fint!), men derefter fås af ligningen* $0.0001x + 1 = 1$, *at* $x = 0$, *hvilket ikke er så godt. Løsningen på dette numeriske problem er at bruge* *rækkepivotering eller delvis pivotering, som det kaldes. Hermed menes, at man ombytter rækkerne så den række, der skal bruges til pivoteringsrække indeholder søjlens numerisk største element. I det foreliggende tilfælde vil det sige, at række 1 og række 2 byttes om før vi iøvrigt går igang. Hermed er udgangspunktet*

$$\begin{pmatrix} 1 & 1 & 2 \\ 0.0001 & 1 & 1 \end{pmatrix}$$

der efter operationen $R_2 := R_2 - 0.0001R_1$ giver (når der regnes med 3 betydende cifre)

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

Det tilsvarende ligningssystem har løsningen $x = 1$ og $y = 1$. Fuldstændig pivotering betyder, at man også bytter om på søjlerne, så matricens numerisk største element bliver brugt som pivoteringselement. Ombytning af søjlerne betyder ved fortolkningen som ligningssystem, at de ubekendte har byttet plads, hvilket kan være uønskelig.

4 Matrixalgebra

Definition 38 *Lad A være en matrix, $A = (a_{ij})$ og s et tal. Så defineres sA som matricen $sA = (sa_{ij})$. Matricen sA fremkommer altså af matricen A ved multiplikation af hvert af A's elementer med tallet s.*

Definition 39 *Lad $A = (a_{ij})$ og $B = (b_{ij})$ være matricer med samme rækkeantal og samme søjleantal. Så defineres matricen $A + B$ som den matrix, hvis ij-element er lig med $a_{ij} + b_{ij}$. Matricen $A + B$ defineres altså ved elementvis addition.*

Eksempel 40 *Lad matricerne A og B være givet ved*

$$A = \begin{pmatrix} 7 & 9 & 13 \\ -2 & \pi & a \end{pmatrix}, B = \begin{pmatrix} 2 & -5 & 3 \\ b & 1 & 8 \end{pmatrix}$$

Så har vi, at

$$3A = \begin{pmatrix} 21 & 27 & 39 \\ -6 & 3\pi & 3a \end{pmatrix}$$

og

$$A + B = \begin{pmatrix} 9 & 4 & 16 \\ b-2 & \pi+1 & a+8 \end{pmatrix}$$

Definition 41 *Lad A være en $m \times n$ matrix, og lad B være en $n \times k$ matrix. Så defineres produktet AB som den matrix, hvis ij element er givet ved*

$$(AB)_{ij} = \sum_{p=1}^{n} a_{ip}b_{pj} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{in}b_{nj}$$

Element ij i produktet AB er altså skalarproduktet mellem den i'te række i A og den j'te søjle i B.

Eksempel 42 *Lad* A *og* B *være matricerne*

$$A = \begin{pmatrix} 7 & 9 & 13 \\ -2 & 3 & a \end{pmatrix}, B = \begin{pmatrix} 1 & -1 & 3 & -4 \\ 0 & 3 & -2 & 3 \\ 5 & -1 & 2 & 4 \end{pmatrix}$$

Så er produktet AB *givet ved*

$$AB = \begin{pmatrix} 72 & 7 & 29 & 51 \\ -2+5a & 11-a & -12+2a & 17+4a \end{pmatrix}$$

Bemærkning 43 *Da elementerne i produktet* AB *findes som skalarprodukter af rækkerne i* A *og søjlerne i* B*, følger det, at søjleantallet i* A *skal være det samme som rækkeantallet i* B*. Er* A *en* $m \times n$ *matrix og* B *en* $q \times k$ *matrix, så eksisterer produktet* AB*, hvis og kun hvis* $n = q$*. Produktet er i bekræftende fald en* $m \times k$ *matrix. Produktet* BA *eksisterer, hvis og kun hvis* $k = m$*. I bekræftende fald er* BA *en* $q \times n$ *matrix. Begge produkter (altså* AB *og* BA*) eksisterer, hvis og kun hvis* $n = q$ *og* $k = m$*, og i så fald er* AB *en* $m \times m$ *matrix og* BA *en* $n \times n$ *matrix. Forlanger vi foruden eksistensen af begge produkter også, at disse skal være af samme størrelse, ender vi altså med forlangendet om, at* $k = m = n = q$*, dvs. begge matricer skal være kvadratiske og af samme størrelse.*

Eksempel 44 *Lad matricerne* A *og* B *være givet ved*

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

Så fås

$$\begin{aligned} AB &= \begin{pmatrix} 2 & 1 \\ 4 & 3 \end{pmatrix} \\ BA &= \begin{pmatrix} 3 & 4 \\ 1 & 2 \end{pmatrix} \end{aligned}$$

Vi ser, at $AB \neq BA$*.*

Eksempel 45 *Lad matricen* N *være givet ved*

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

Så finder vi, at $N^2 = NN$ *og* $N^3 = NNN$ *er givet ved*

$$\begin{aligned} N^2 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \\ N^3 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

Idet vi (lidt upræcist) betegner enhver matrix bestående udelukkende af nuller som nulmatricen og blot skriver 0*, har vi altså, at* $N^3 = 0$*. Man siger, at matricen* N *er nilpotent. Bemærk specielt, at det altså ikke af* $AB = 0$ *følger, at enten* A *eller* B *er nulmatricen.*

Sætning 46 *Lad* A, B *og* C *være matricer. Så gælder*

$$\begin{aligned} A + B &= B + A \\ (A + B) + C &= A + (B + C) \\ A(B + C) &= AB + AC \\ (A + B)C &= AC + BC \\ (AB)C &= A(BC) \end{aligned}$$

når iøvrigt de forskellige konstruktioner giver mening, dvs. når matricerne har de fornødne dimensioner. Når s *og* t *er tal gælder også*

$$\begin{aligned} (s + t)A &= sA + tA \\ s(tA) &= (st)A \\ A(sB) &= s(AB) \end{aligned}$$

Bevis. De to øverste påstande er trivielle. Tænk blot på, at additionen af matricer er defineret som elementvis addition. En matrix kan i den sammenhæng opfattes som én lang vektor, der blot af pladshensyn er skrevet som et antal rækker. Matrixmultiplikationen er imidlertid lidt speciel - den kommutative lov gælder (som vi så ovenfor) jo ikke. Den distributive lov $A(B + C) = AB + AC$ vises således, når A er en $m \times n$ matrix, og B og C er $n \times k$ matricer:

$$(A(B + C))_{pq} = \sum_{j=1}^{n} a_{pj}(b_{jq} + c_{jq}) = \sum_{j=1}^{n} a_{pj}b_{jq} + \sum_{j=1}^{n} a_{pj}c_{jq} = (AB)_{pq} + (AC)_{pq}$$

Den anden distributive lov $(A + B)C = AC + BC$ vises ganske analogt. Den associative lov for multiplikation $(AB)C = A(BC)$ vises således, når A er en $m \times n$ matrix, B er en $n \times k$ matrix og C er en $k \times r$ matrix:

$$\begin{aligned} ((AB)C)_{pq} &= \sum_{j=1}^{k} (AB)_{pj} c_{jq} = \sum_{j=1}^{k} \left(\sum_{i=1}^{n} a_{pi}b_{ij} \right) c_{jq} \\ &= \sum_{i=1}^{n} \sum_{j=1}^{k} a_{pi}b_{ij}c_{jq} = \sum_{i=1}^{n} a_{pi} \left(\sum_{j=1}^{k} b_{ij}c_{jq} \right) \\ &= \sum_{i=1}^{n} a_{pi} (BC)_{iq} = (A(BC))_{pq} \end{aligned}$$

De tre påstande, der involverer skalarmultiplikation, er alle simple at eftervise.
■

Med indførslen af matrixmultiplikationen kan det lineære ligningssystem

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

skrives på formen

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

altså på formen $Ax = b$, når A er koefficientmatricen og de ubekendte repræsenteres ved en søjlevektor (en $n \times 1$ matrix)

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

og hvor højresiderne repræsenteres ved søjlevektoren ($m \times 1$ matricen)

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

Bemærkning 47 *Fra nu af vil alle vektorer fra R^n blive anset som søjlevektorer, dvs. matricer af typen $n \times 1$, med mindre naturligvis andet siges eksplicit.*

Sætning 48 *Hvis systemet $Ax = b$ overhovedet har en løsning, så er den fuldstændige løsning til ligningssystemet $Ax = b$ lig med summen af en partikulær løsning og den fuldstændige løsning til det homogene system $Ax = 0$.*

Bemærkning 49 *Med symbolet 0 skal vi undertiden mene nulvektoren (af variende længde), undertiden nulmatricen (af variende dimensioner) og undertiden blot skalaren 0. Det skulle gerne af sammenhængen fremgå, hvad meningen er.*

Bevis. Lad x_p være en løsning til $Ax = b$. Vi skal først vise, at hvis x_0 er en løsning til $Ax = 0$, så er $x_p + x_0$ løsning til $Ax = b$. Men dette følger således

$$A(x_p + x_0) = Ax_p + Ax_0 = b + 0 = b$$

Dernæst skal vi vise, at hvis y er en anden løsning til den inhomogene ligning $Ax = b$, så er forskellen $y - x_p$ løsning til den homogene ligning. Dette følger lige så let

$$A\left(y - x_p\right) = Ay - Ax_p = b - b = 0.$$

■

Bemærkning 50 *Denne sætning minder jo kraftigt om en sætning fra teorien om lineære differentialligninger.*

Sætning 51 *Ligningssystemet $Ax = b$ har en løsning, hvis og kun hvis søjlevektoren b kan skrives som en linearkombination af søjlerne i A.*

Bevis. Idet vi betegner søjlerne i $m \times n$ matricen A med $S_1, S_2, \ldots, S_n$ har vi

$$\begin{aligned} Ax &= \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \end{pmatrix} \\ &= x_1S_1 + x_2S_2 + \ldots + x_nS_n \end{aligned}$$

Heraf fremgår påstanden. ■

Definition 52 *Ved den transponerede af matricen A forstås den matrix, hvis rækker består af søjlerne for A i samme rækkefølge. Den transponerede til A vil i disse noter blive betegnet med A^t. Vi har altså, hvis $A = (a_{ij})$, at $(A^t)_{ij} = a_{ji}$.*

Bemærkning 53 *Andre betegnelser for den transponerede til A er A^T, A' eller $\widetilde{A}$ (som i den blå bog).*

Bemærkning 54 *Hvis A er en $m \times n$ matrix, så er A^t en $n \times m$ matrix.*

Eksempel 55 *Lad A være matricen*

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}$$

så er den transponerede givet ved

$$A^t = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{pmatrix}$$

Sætning 56 *Der gælder*

$$\begin{aligned} \left(A^t\right)^t &= A \\ (A + B)^t &= A^t + B^t \\ (AB)^t &= B^t A^t \end{aligned}$$

Bevis. De to første er trivielle. Det er den sidste ikke. Vi har, når A er en $m \times n$ matrix og B en $n \times k$ matrix

$$\begin{aligned} \left((AB)^t\right)_{pq} &= (AB)_{qp} = \sum_{j=1}^{n} a_{qj} b_{jp} \\ &= \sum_{j=1}^{n} \left(A^t\right)_{jq} \left(B^t\right)_{pj} = \left(B^t A^t\right)_{pq} \end{aligned}$$

■

Sætning 57 *En matrix A og dens transponerede A^t har samme rang.*

Bevis. Lad U være en til echelonform reduceret version af A. Hermed har vi, at $Ax = 0 \iff Ux = 0$. Antallet af fra nulrækken forskellige rækker i U er pr. definition lig med $\rho(A)$. Dette er også det maksimale antal lineært uafhængige rækker i U og i A. Men på grund af den form U har, er dette antal også lig med det maksimale antal lineært uafhængige søjler i U. Betragter vi nu et sæt af lineært uafhængige søjler i U, lad os sige $S_1^U, S_3^U, \ldots, S_k^U$ så er de tilsvarende søjler i A, altså $S_1, S_3, \ldots, S_k$ også lineært uafhængige. Antag nemlig, at $c_1 S_1 + c_3 S_3 + \ldots + c_k S_k = 0$. Vi har, at

$$c_1 S_1 + c_3 S_3 + \ldots + c_k S_k = Ac$$

hvor $c^t = (c_1, 0, c_3, \ldots, c_k, 0, \ldots, 0)$. Altså $Ac = 0$. Men heraf følger $Uc = 0$, dvs. $c_1 S_1^U + c_3 S_3^U + \ldots + c_k S_k^U = 0$, hvoraf følger $c_1 = c_3 = \ldots = c_k = 0$. Da det omvendt også gælder, at de til et sæt lineært uafhængige søjler fra A svarende søjler i U er lineært uafhængige, følger det, at det maksimale antal lineært uafhængige søjler i A er lig med det maksimale antal lineært uafhængige søjler i U, altså lig med $\rho(A)$. ■

Bemærkning 58 *Sætningen udtrykkes ofte ved: Rækkerang er lig med søjlerang.*

Definition 59 *Ved skalarproduktet $x \cdot y$ mellem to søjlevektor $x, y \in R^n$ forstås matrixproduktet $x^t y$. To vektorer x og y siges at være ortogonale, hvis deres skalarprodukt er nul. Hvis $M \subseteq R^n$, så defineres det ortogonale komplement $M^\perp$ som mængden af vektorer i R^n, der er vinkelrette på alle vektorerne i M. En basis for et underrum af R^n kaldes ortogonal, hvis vektorerne er indbyrdes ortogonale. Basen kaldes ortonormal, hvis den er ortogonal, og hvis basisvektorerne desuden alle har længde 1.*

Bemærkning 60 *Vi ser, at $x \cdot x = \|x\|^2$ for alle $x \in R^n$, når $\|x\|$ betegner den sædvanlige euklidiske norm af x.*

Bemærkning 61 *I vektorrummet C^n defineres skalarproduktet $x \cdot y$ ved $\overline{x}^t y$. Her er $\overline{x} = (\overline{x_1}, \overline{x_2}, \ldots, \overline{x_n})$.*

Sætning 62 *Hvis M og N er delmængder af R^n og $M \subseteq N$, så gælder $M^\perp \supseteq N^\perp$.*

Sætning 63 *Hvis et sæt vektorer* $\{v_1, v_2, \ldots, v_k\}$ *er indbyrdes ortogonal, og hvis* $v_i \neq 0$ *for alle* $i = 1 \ldots k$*, så er sættet lineært uafhængigt.*

Bevis. Antag, at

$$c_1 v_1 + c_2 v_2 + \ldots + c_k v_k = 0$$

Ved skalarmultiplikation med vektoren v_i fås

$$c_i v_i \cdot v_i = 0$$

hvoraf følger, at $c_i = 0$. Da dette gælder for alle i, må sættet være lineært afhængigt. ■

Sætning 64 *Ethvert underrum* V *af* R^n *har en ortogonal (og dermed en ortonormal) basis. En ortonormal basis kan findes udfra en given basis ved Gram-Schmidts ortogonaliseringsprocedure.*

Bevis. Lad V være et underrum af R^n, og lad $\{v_1, v_2, \ldots, v_k\}$ være en basis for V. Vi skal udfra denne basis konstruere en ortonormal basis $\{u_1, u_2, \ldots, u_k\}$. Dette gøres ved Gram-Schmidts ortogonaliseringsprocedure. Vi sætter

$$u_1 = \frac{v_1}{\|v_1\|}$$

Hermed har u_1 længde 1. Dernæst sætter vi

$$w_2 = v_2 - (v_2 \cdot u_1)\, u_1$$

Hermed har vi, at $w_2 \perp u_1$, da $w_2 \cdot u_1 = (v_2 - (v_2 \cdot u_1)\, u_1) \cdot u_1 = v_2 \cdot u_1 - (v_2 \cdot u_1)\,(u_1 \cdot u_1) = v_2 \cdot u_1 - v_2 \cdot u_1 = 0$. Hvis ikke $k = 2$ kan w_2 ikke være nulvektoren. Sæt nu

$$u_2 = \frac{w_2}{\|w_2\|}$$

Hermed får u_2 længden 1. Vi fortsætter med

$$w_3 = v_3 - (v_3 \cdot u_1)\, u_1 - (v_3 \cdot u_2)\, u_2$$

Hermed har vi, at $w_3 \perp u_1$, da $w_3 \cdot u_1 = (v_3 - (v_3 \cdot u_1)\, u_1 - (v_3 \cdot u_2)\, u_2) \cdot u_1 = v_3 \cdot u_1 - (v_3 \cdot u_1)\,(u_1 \cdot u_1) = 0$, hvor vi har udnyttet, at $u_1 \cdot u_2 = 0$. At også $w_3 \perp u_2$ følger på samme måde. Hvis ikke $k = 3$ kan w_3 ikke være nulvektoren. Sæt nu

$$u_3 = \frac{w_3}{\|w_3\|}$$

Hermed har u_3 længden 1. Således fortsættes indtil u_k er defineret. At de fundne vektorer udgør en basis for V følger af, at antal vektorer er k og at de er lineært uafhængige. ■

Sætning 65 *(Projektionssætningen) Lad* V *være et underrum af* R^n*. Så kan enhver vektor i* R^n *på én og kun én måde skrives som en sum af en vektor fra* V *og en vektor fra* $V^\perp$*.*

Bevis. Lad $x \in R^n$. Vi viser først entydigheden. Antag, at $x = v_1 + u_1 = v_2 + u_2$, hvor $v_1, v_2 \in V$ og $u_1, u_2 \in V^{\perp}$. Så har vi, at $v_1 - v_2 = u_2 - u_1$. Men venstre side af denne ligning tilhører V og højre side tilhører $V^{\perp}$. Vektoren $v_1 - v_2 = u_2 - u_1$ er dermed ortogonal på sig selv, dvs. $\|v_1 - v_2\| = 0$, hvormed $v_1 = v_2$ og også $u_2 = u_1$. Vi viser eksistensen af opsplitningen af x. Hvis $x \in V$, så er opsplitningen simpelthen $x = x + 0$. Antag derfor, at $x \notin V$. Vælg en ortonormal basis $\{v_1, v_2, \ldots, v_k\}$ i V. Brug Gram-Schmidts ortogonaliseringsmetode til at finde en vektor w der er ortogonal på alle vektorerne i $\{v_1, v_2, \ldots, v_k\}$. Vi sætter altså

$$w = x - \sum_{i=1}^{k} (x \cdot v_i) \, v_i$$

Men hermed har vi

$$\begin{aligned} x &= v + w \\ v &= \sum_{i=1}^{k} (x \cdot v_i) \, v_i \end{aligned}$$

hvor $v \in V$ og $w \in V^{\perp}$. ■

Sætning 66 *Lad $M \subseteq R^n$. Så gælder, at $M^{\perp}$ er et underrum af R^n. Desuden gælder, hvis M er et underrum af R^n, at $\left(M^{\perp}\right)^{\perp} = M$.*

Bevis. Åbenbart ligger nulvektoren i $M^{\perp}$, så $M^{\perp} \neq \emptyset$. Lad $x, y \in M^{\perp}$. Vi skal vise, at $x + y$ og sx ligger i $M^{\perp}$, når s er en skalar. Vi har for alle $m \in M$

$$(x + y) \cdot m = x \cdot m + y \cdot m = 0 + 0 = 0$$

så $x + y \in M^{\perp}$. Vi har også, at $(sx) \cdot m = s\,(x \cdot m) = 0$, derfor gælder $sx \in M^{\perp}$. Vi skal nu vise, at hvis M er et underrum, så gælder, at $\left(M^{\perp}\right)^{\perp} = M$. Vi betragter først den ene inklusion: $M \subseteq \left(M^{\perp}\right)^{\perp}$. Lad altså $m \in M$. Vi har åbenbart, at $m \cdot v = 0$ for alle $v \in M^{\perp}$. Men dette medfører, at $v \perp M^{\perp}$, altså $v \in M^{\perp\perp}$. Vi har altså $M \subseteq \left(M^{\perp}\right)^{\perp}$ også uanset om M er et underrum. Den omvendte inklusion $\left(M^{\perp}\right)^{\perp} \subseteq M$ er ikke helt så triviel. Lad da $x \in \left(M^{\perp}\right)^{\perp}$. Da M er et underrum af R^n, kan x skrives på formen $x = m + y$, hvor $m \in M$ og $y \in M^{\perp}$. Vi skal vise, at $y = 0$. Da $x \in \left(M^{\perp}\right)^{\perp}$ har vi, at $x \cdot y = 0$, dvs. at $(m + y) \cdot y = 0$. Men $m \cdot y = 0$, så $y \cdot y = 0$. Heraf følger, at $y = 0$. ■

4.1 Afbildninger defineret ved en matrix

Lad A være en $m \times n$ matrix. For enhver søjlevektor $x \in R^n$ er da Ax en søjlevektor i R^m. Sætter vi

$$T_A\,(x) = Ax$$

for alle $x \in R^n$ er T_A hermed en afbildning fra R^n til R^m. Denne afbildning er lineær, dvs. den opfylder følgende to krav

$$\begin{aligned} T_A(x+y) &= T_A(x) + T_A(y) \\ T_A(sx) &= sT_A(x) \end{aligned}$$

for alle $x, y \in R^n$ og alle skalarer (tal) s. Disse to krav er opfyldt fordi $T_A(x+y) = A(x+y) = Ax + Ay = T_A(x) + T_A(y)$ og $T_A(sx) = A(sx) = sAx = sT_A(x)$.

Sætning 67 *Lad A være en $m \times n$ matrix og lad afbildningen T_A være givet ved $T_A(x) = Ax$ for alle $x \in R^n$. Billedmængden $T_A(R^n)$ er det underrum af R^m, der udspændes af matricen A's søjler.*

Bevis. Som tidligere observeret har vi, idet vi betegner søjlerne i $m \times n$ matricen A med $S_1, S_2, \ldots, S_n$

$$\begin{aligned} Ax &= \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \end{pmatrix} \\ &= x_1S_1 + x_2S_2 + \ldots + x_nS_n \end{aligned}$$

heraf følger påstanden. ∎

Definition 68 *Billedmængden $T_A(R^n)$ betegnes normalt med $R(A)$ (R for engelsk range).*

Sætning 69 *Lad A være en $m \times n$ matrix og lad afbildningen T_A være givet ved $T_A(x) = Ax$ for alle $x \in R^n$. Mængden $N(A)$ af løsninger til $T_A(x) = 0$, altså til $Ax = 0$, er et underrum af R^n.*

Bevis. Vi skal blot vise, at hvis $x, y \in N(A)$ og hvis s er et tal, så gælder $x + y \in N(A)$ og $sx \in N(A)$. Men dette følger af

$$\begin{aligned} A(x+y) &= Ax + Ay = 0 + 0 = 0 \\ A(sx) &= sAx = s0 = 0 \end{aligned}$$

hvor alle nullerne er nulvektorer. ∎

Sætning 70 *Lad A være en $m \times n$ matrix. Så gælder*

$$\begin{aligned} \dim R(A) &= \dim R\left(A^t\right) = \rho(A) \\ \dim R(A) + \dim N(A) &= n \\ \dim R\left(A^t\right) + \dim N\left(A^t\right) &= m \end{aligned}$$

Bevis. Den første påstand er kun en gentagelse af en allerede vist påstand. Af de to sidste er det åbenbart nok at vise den første. Lad en basis for billedrummet $R(A)$ være $Av_1, Av_2, \ldots, Av_r$, hvor $r = \rho(A)$. Lad en basis for nulrummet $N(A)$ være $n_1, n_2, \ldots, n_k$. Vi viser, at $n_1, n_2, \ldots, n_k, v_1, v_2, \ldots, v_r$ udgør en basis for R^n, hvoraf følger, at $k + r = n$. Først viser vi, at vektorerne er lineært uafhængige. Antag, at

$$c_1 n_1 + c_2 n_2 + \ldots + c_k n_k + d_1 v_1 + d_2 v_2 + \ldots + d_r v_r = 0$$

Gang med A fra venstre, så fås, da $An_i = 0$

$$d_1 A v_1 + d_2 A v_2 + \ldots + d_r A v_r = 0$$

Men vektorerne $Av_1, Av_2, \ldots, Av_r$ er lineært uafhængige, så $d_1 = d_2 = \ldots = d_r = 0$. Hermed har vi

$$c_1 n_1 + c_2 n_2 + \ldots + c_k n_k = 0$$

Da også $n_1, n_2, \ldots, n_k$ er lineært uafhængige følger, at $c_1 = c_2 = \ldots = c_k = 0$. Altså er $n_1, n_2, \ldots, n_k, v_1, v_2, \ldots, v_r$ lineært uafhængige. Vi skal nu vise, at enhver given vektor kan skrives som en linearkombination af vektorerne $n_1, n_2, \ldots, n_k, v_1, v_2, \ldots, v_r$. Lad derfor $v \in R^n$ være en vilkårlig vektor.Da $Av_1, Av_2, \ldots, Av_r$ udgør en basis for $R(A)$ kan Av skrives som en linearkombination af disse vektorer, altså

$$Av = d_1 A v_1 + d_2 A v_2 + \ldots + d_r A v_r$$

Men så har vi

$$A(d_1 v_1 + d_2 v_2 + \ldots + d_r v_r - v) = 0$$

dvs. at $d_1 v_1 + d_2 v_2 + \ldots + d_r v_r - v \in N(A)$, men da $n_1, n_2, \ldots, n_k$ er en basis for $N(A)$ følger, at der findes tal $c_1, c_2, \ldots, c_k$ så

$$d_1 v_1 + d_2 v_2 + \ldots + d_r v_r - v = c_1 n_1 + c_2 n_2 + \ldots + c_k n_k$$

men dvs. at v kan skrives som en linearkombination af $n_1, n_2, \ldots, n_k, v_1, v_2, \ldots, v_r$.
■

Sætning 71 *Lad A være en $m \times n$ matrix. Så gælder*

$$\begin{aligned} R(A) &= N\left(A^t\right)^{\perp} \\ R\left(A^t\right) &= N(A)^{\perp} \\ N\left(A^t\right) &= R(A)^{\perp} \\ N(A) &= R\left(A^t\right)^{\perp} \end{aligned}$$

Bevis. Vi viser den første påstand. $R(A) = N(A^t)^{\perp}$. Lad først $y = Ax \in R(A)$ og lad $z \in N(A^t)$. Så gælder, at $Ax \cdot z = (Ax)^t z = x^t A^t z = 0$, altså gælder, at $y = Ax \in N(A^t)^{\perp}$. Hermed er $R(A) \subseteq N(A^t)^{\perp}$ vist. Vi skal nu vise, at den omvendte inklusion gælder, altså $R(A) \supseteq N(A^t)^{\perp}$. Vi foretrækker i stedet at vise det ækvivalente udsagn $R(A)^{\perp} \subseteq N(A^t)$. Lad altså $y \in R(A)^{\perp}$. Så har vi, at $y \cdot Ax = 0$ for alle $x \in R^n$, og dermed også for $x = A^t y$. Hermed har vi $0 = y \cdot AA^t y = y^t AA^t y = (A^t y)^t (A^t y) = A^t y \cdot A^t y$ hvoraf følger, at $A^t y = 0$, altså $y \in N(A^t)$. Hermed er det første udsagn vist. Det andet udsagn fås ved ombytning af A med A^t. De to sidste fås af de to første ved at betragte ortogonale komplementer af begge sider. ∎

4.2 Kvadratiske matricer

Definition 72 *Diagonalen i en kvadratisk ($n \times n$) matrix A består af elementerne $a_{ii}, i = 1 \ldots n$. En diagonalmatrix er en matrix, der har lutter nuller uden for diagonalen.*

Definition 73 *En enhedsmatrix E er en diagonalmatrix, hvis diagonalelementer alle er 1.*

Bemærkning 74 *Uanset størrelsen af enhedsmatricen vil den blive betegnet med E. I mange lærebøger betegnes den i stedet med I.*

Sætning 75 *Lad A være en kvadratisk matrix og E enhedsmatricen. Så gælder, at $AE = EA = A$.*

Bevis. Stil matricerne op ved siden af hinanden og gang ud. ∎

Sætning 76 *Lad A være en kvadratisk matrix. Der er højst én matrix B, der opfylder*

$$AB = BA = E$$

Bevis. Antag, at vi foruden $AB = BA = E$ også har $AC = CA = E$. Så fås

$$C = CE = C(AB) = (CA)B = EB = B$$

∎

Bemærkning 77 *Bemærk, at beviset kun udnyttede $AB = E$ og $CA = E$. Heraf alene konkluderedes, at $C = B$. En matrix B, der opfylder $AB = E$, kaldes en højreinvers til A. Tilsvarende kaldes en matrix C, der opfylder $CA = E$ for en venstreinvers til A.*

Definition 78 *Lad A være en kvadratisk matrix. Hvis der findes en matrix B, der opfylder*

$$AB = BA = E$$

vil vi kalde B den inverse matrix til A, og vi vil sige, at A er invertibel. Den inverse matrix til A vil blive betegnet med A^{-1}.

Sætning 79 *Lad A og B være invertible matricer. Så gælder, at A^{-1}, A^t og AB er invertible, og vi har*

$$\begin{aligned} \left(A^{-1}\right)^{-1} &= A \\ \left(A^t\right)^{-1} &= \left(A^{-1}\right)^t \\ (AB)^{-1} &= B^{-1}A^{-1} \end{aligned}$$

Bevis. Vi har

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

men det kan jo læses som, at der eksisterer en matrix (nemlig A), som ganget på A^{-1} giver E (og fra begge sider!). Derfor har A^{-1} en invers, og A er denne inverse. Vi har også

$$\begin{aligned} A^t\left(A^{-1}\right)^t &= \left(A^{-1}A\right)^t = E^t = E \\ \left(A^{-1}\right)^t A^t &= \left(AA^{-1}\right)^t = E^t = E \end{aligned}$$

derfor er A^t invertibel med $\left(A^t\right)^{-1} = \left(A^{-1}\right)^t$. Vi har videre dels

$$(AB)\left(B^{-1}A^{-1}\right) = A\left(BB^{-1}\right)A^{-1} = AEA^{-1} = AA^{-1} = E$$

men også

$$\left(B^{-1}A^{-1}\right)(AB) = B^{-1}\left(A^{-1}A\right)B = B^{-1}EB = B^{-1}B = E$$

Hermed har vi altså, at $(AB)\left(B^{-1}A^{-1}\right) = \left(B^{-1}A^{-1}\right)(AB) = E$, hvoraf påstanden følger. ■

Sætning 80 *Lad A være en invertibel matrix. Så har ligningssystemet $Ax = b$ én løsning for ethvert givet b, nemlig $x = A^{-1}b$.*

Bevis. Ligningssystemet $Ax = b$ er ækvivalent med det ligningssystem, som vi opnår ved multiplikation fra venstre med A^{-1} på begge sider, altså $A^{-1}(Ax) = A^{-1}b$. Vi kan jo komme tilbage til det oprindelige ved multiplikation med A fra venstre på begge sider af lighedstegnet. Men systemet $A^{-1}(Ax) = A^{-1}b$ er jo blot $x = A^{-1}b$. ■

Sætning 81 *Lad A være en kvadratisk ($n \times n$) matrix. Så er A invertibel, hvis og kun hvis $\rho(A) = n$. I bekræftende fald kan den inverse findes ved Gauss-Jordan-reduktion på systemet $AX = E$.*

Bevis. Hvis A er invertibel, så har systemet $Ax = 0$ netop én løsning. Men så må $\rho(A) = n$. Antag omvendt, at $\rho(A) = n$. Vi vil nu løse matrixligningen $AX = E$. Lad $X = [X_1, X_2, \ldots, X_n]$ og $E = [e_1, e_2, \ldots, e_n]$ være matricerne X og E opskrevet som følger af søjler. Så er systemet $AX = E$ ækvivalent med systemerne $AX_i = e_i, i = 1 \ldots n$. Men disse kan løses på én gang ved Gauss-Jordan-elimination, hvor totalmatricen T opnås ved sammenstilling af A og alle højresiderne $e_1, e_2, \ldots, e_n$, altså E. Da rangen af A er n kan denne Gauss-Jordan-elimination gennemføres. Totalmatricen $T = [A\,|E]$ er altså blevet reduceret til en matrix af formen $[E\,|B]$. Hermed opfylder B ligningen $AX = E$, med andre ord, vi har $AB = E$. Matricen B må imidlertid have rang n, da den ved rækkeoperationer jo kan føres tilbage til E. Altså gælder det om B, at ligningen $BX = E$ har en løsning, lad os kalde denne løsning C. Vi har altså, at $BC = E$. Vi har nu både $AB = E$ og $BC = E$, dvs. at B har både en venstre- og en højreinvers. Som tidligere set er disse nødvendigvis ens. Altså har vi, at $C = A$, hvoraf følger, at $AB = BA = E$, således at A er invertibel med invers B. ■

Korollar 82 *En kvadratisk matrix A er invertibel, hvis og kun hvis den har enten en venstre- eller en højreinvers.*

Bevis. Hvis $n \times n$ matricen A har en invers, så er denne jo både højre- og venstreinvers. Vi viser det omvendte. Hvis A har en højreinvers, så har $AX = E$ en løsning, dvs. $AX = b$ har en løsning for ethvert b. Men så må $\rho(A) = n$. Af sætningen ovenfor følger, at A er invertibel. Hvis A har en venstreinvers B, så har B en højreinvers (nemlig A) og derfor en invers (også A). Men så gælder $AB = BA = E$, hvorfor A er invertibel. ■

Eksempel 83 *Lad os sige, at én påstår, at matricen*

$$A = \begin{pmatrix} 1 & 2 \\ 7 & 9 \end{pmatrix}$$

har som invers matricen B givet ved

$$B = \begin{pmatrix} -\frac{9}{5} & \frac{2}{5} \\ \frac{7}{5} & -\frac{1}{5} \end{pmatrix}$$

Ønsker vi at kontrollere denne påstand, så behøver vi blot kontrollere, om $AB = E$ eller $BA = E$. Udfører man den ene af disse kontroludregninger, finder man, at svaret er ja, A er invertibel med $A^{-1} = B$.

Eksempel 84 *Antag, at både A og $E + A$ er invertible. Lad den inverse til $E + A$ være $(E + A)^{-1} = B$. Vi vil vise, at også $E + A^{-1}$ har en invers, og at den er lig med BA. Vi behøver blot kontrollere, om enten $BA\left(E + A^{-1}\right) = E$ eller om $\left(E + A^{-1}\right) BA = E$. Den første ser lettest ud, så den kigger vi på. Vi har $BA\left(E + A^{-1}\right) = BAE + BAA^{-1} = BA + B = B\,(A. + E) = E$, hvor den sidste lighed, skyldes, at B er den inverse til $E + A$.*

Eksempel 85 *Lad A være matricen*

$$A = \begin{pmatrix} 2 & 3 & -4 \\ 6 & 1 & 1 \\ -2 & 2 & -4 \end{pmatrix}$$

Vi vil undersøge, om A er invertibel. I bekræftende fald vil vi bestemme den inverse matrix. I håbet om, at svaret er ja, stiller vi totalmatricen $[A\,|E]$ op:

$$T = \begin{pmatrix} 2 & 3 & -4 & 1 & 0 & 0 \\ 6 & 1 & 1 & 0 & 1 & 0 \\ -2 & 2 & -4 & 0 & 0 & 1 \end{pmatrix}$$

Undervejs i Gauss-Jordan-reduktionen finder vi jo ud af, om rangen af A er 3. Hvis ikke, så har A ingen invers og den del af arbejdet, der vedrørte E er (desværre) spildt. De første operationer er $R_2 := R_2 - 3R_1, R_3 := R_3 + R_1$, der giver

$$T_1 = \begin{pmatrix} 2 & 3 & -4 & 1 & 0 & 0 \\ 0 & -8 & 13 & -3 & 1 & 0 \\ 0 & 5 & -8 & 1 & 0 & 1 \end{pmatrix}$$

Derefter operationen $R_3 := R_3 + \frac{5}{8}R_2$, der giver

$$T_2 = \begin{pmatrix} 2 & 3 & -4 & 1 & 0 & 0 \\ 0 & -8 & 13 & -3 & 1 & 0 \\ 0 & 0 & \frac{1}{8} & -\frac{7}{8} & \frac{5}{8} & 1 \end{pmatrix}$$

Vi ser nu, ar $\rho(A) = 3$, hvorfor A er invertibel. Vi går videre med operationen $R_3 := 8R_3$, der giver

$$T_3 = \begin{pmatrix} 2 & 3 & -4 & 1 & 0 & 0 \\ 0 & -8 & 13 & -3 & 1 & 0 \\ 0 & 0 & 1 & -7 & 5 & 8 \end{pmatrix}$$

Herefter følger operationerne $R_2 := R_2 - 13R_3, R_1 := R_1 + 4R_3$, hvorved fås

$$T_4 = \begin{pmatrix} 2 & 3 & 0 & -27 & 20 & 32 \\ 0 & -8 & 0 & 88 & -64 & -104 \\ 0 & 0 & 1 & -7 & 5 & 8 \end{pmatrix}$$

Operationen $R_2 := -\frac{1}{8}R_2$ giver os det ønskede ettal i diagonalens position $(2,2)$

$$T_5 = \begin{pmatrix} 2 & 3 & 0 & -27 & 20 & 32 \\ 0 & 1 & 0 & -11 & 8 & 13 \\ 0 & 0 & 1 & -7 & 5 & 8 \end{pmatrix}$$

Operationen $R_1 := R_1 - 3R_2$ giver

$$T_6 = \begin{pmatrix} 2 & 0 & 0 & 6 & -4 & -7 \\ 0 & 1 & 0 & -11 & 8 & 13 \\ 0 & 0 & 1 & -7 & 5 & 8 \end{pmatrix}$$

Vores sidste operation er $R_1 := \frac{1}{2}R_1$, der giver

$$T_7 = \begin{pmatrix} 1 & 0 & 0 & 3 & -2 & -\frac{7}{2} \\ 0 & 1 & 0 & -11 & 8 & 13 \\ 0 & 0 & 1 & -7 & 5 & 8 \end{pmatrix}$$

Vi har hermed, at den inverse matrix til A er givet ved

$$A^{-1} = \begin{pmatrix} 3 & -2 & -\frac{7}{2} \\ -11 & 8 & 13 \\ -7 & 5 & 8 \end{pmatrix}$$

Ved håndregning er det næsten umuligt ikke at lave en regnefejl på et eller andet punkt i udregningerne. Det er derfor en særdeles god idé, at kontrollere, om $AA^{-1} = E$, som det jo skal være.

Bemærkning 86 *Vil man løse et ligningssystem $Ax = b$, hvor koefficientmatricen er en kvadratisk og invertibel matrix, kan man gøre det ved at bestemme A^{-1} først og derefter udregne $A^{-1}b$. Arbejdet forbundet med bestemmelsen af den inverse er imidlertid stort og multiplikationen i $A^{-1}b$ koster også arbejde, så normalt betaler det sig ikke at bruge denne fremgangsmåde. Man bruger i stedet almindelig Gausselimination på totalmatricen med påfølgende løsning af det reducerede system nedefra.*

4.3 LU-faktorisering. Gauss-eliminationsprocessen som matrixprodukt

Rækkeoperationerne $R_2 := R_2 - 5R_1, R_3 := R_3 + 2R_1$ på matricen A givet ved

$$A = \begin{pmatrix} 1 & 5 & -2 \\ 5 & 7 & -3 \\ -2 & -1 & 0 \end{pmatrix}$$

vil skaffe nuller under ettallet i søjle 1. Herved opnås matricen

$$A_1 = \begin{pmatrix} 1 & 5 & -2 \\ 0 & -18 & 7 \\ 0 & 9 & -4 \end{pmatrix}$$

Denne delvist reducerede matrix A_1kan også opnås ved multiplikation af A fra venstre med matricen

$$L_1 = \begin{pmatrix} 1 & 0 & 0 \\ -5 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}$$

Vi har nemlig, at $A_1 = L_1A$. Den næste operation i Gausseliminationen er $R_3 := R_3 + \frac{1}{2}R_2$, der giver os

$$A_2 = \begin{pmatrix} 1 & 5 & -2 \\ 0 & -18 & 7 \\ 0 & 0 & -\frac{1}{2} \end{pmatrix}$$

Denne rækkeoperation kan opnås ved multiplikation af A_1 med matricen

$$L_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \frac{1}{2} & 1 \end{pmatrix}$$

Vi har nemlig, at $A_2 = L_2 A_1$. Hermed har vi, at den færdige Gausseliminerede matrix A_2 er givet ved $A_2 = L_2 L_1 A$. De to matricer L_1 og L_2 er invertible med inverse givet ved

$$L_1^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 5 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix}, L_2^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\frac{1}{2} & 1 \end{pmatrix}$$

Bemærk, at disse fremkommer af L_1 og L_2 ved blot at ændre fortegnene på elementerne uden for diagonalen. De inverse svarer til de inverse rækkeoperationer. Hermed har vi, at $A = L_1^{-1} L_2^{-1} A_2$. Vi sætter $L = L_1^{-1} L_2^{-1}$. Bemærk nu videre, at

$$L = L_1^{-1} L_2^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 5 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\frac{1}{2} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 5 & 1 & 0 \\ -2 & -\frac{1}{2} & 1 \end{pmatrix}$$

Matricen L kan i virkeligheden skrives direkte ned udfra kendskabet til de udførte rækkeoperationer, der jo var $R_2 := R_2 - 5R_1, R_3 := R_3 + 2R_1$ og $R_3 := R_3 + \frac{1}{2} R_2$. Det i dette eksempel viste resultat er (med modifikationer) generelt.

Definition 87 *Ved en øvre (nedre) triangulær matrix forstås en kvadratisk matrix, hvor elementerne under (over) diagonalen er lutter nuller.*

Sætning 88 *Lad A være en kvadratisk og invertibel matrix. Hvis rækkeombytninger ved Gausseliminationen ikke er nødvendige, så kan A skrives som et produkt af en nedre triangulær matrix L med ettaller i diagonalen og en øvre triangulær matrix U, altså $A = LU$.*

Bemærkning 89 *Der findes en mere generel version af denne sætning. Se f.eks. Jens Eising, Lineær Algebra.*

En sådan faktorisering kaldes en LU-faktorisering. Bogstaverne L og U stammer fra de engelske betegnelser for nedre og øvre, nemlig *lower* og *upper*.

En løsning af systemet $Ax = b$, når LU-faktorisering allerede er udført, kan udføres i to tempi således Løs først systemet $Ly = b$. Løs derefter systemet $Ux = y$. Det første systems løsning foregår let oppefra. Det sidste foregår let nedefra.

Eksempel 90 *LU-faktoriseringen af matricen*

$$A = \begin{pmatrix} 1 & 5 & -2 \\ 5 & 7 & -3 \\ -2 & -1 & 0 \end{pmatrix}$$

blev tidligere fundet til

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 5 & 1 & 0 \\ -2 & -\frac{1}{2} & 1 \end{pmatrix}, U = \begin{pmatrix} 1 & 5 & -2 \\ 0 & -18 & 7 \\ 0 & 0 & -\frac{1}{2} \end{pmatrix}$$

Ønsker vi at løse systemet $Ax = b$, hvor $b = (3, 0, 9)^t$, kan dette gøres således. Vi løser først $Ly = b$ for y:

$$\begin{aligned} y_1 &= 3 \\ 5y_1 + y_2 &= 0 \\ -2y_1 - \frac{1}{2}y_2 + y_3 &= 9 \end{aligned}$$

Vi finder $y = \left(3, -15, \frac{15}{2}\right)^t$. Herefter løses $Ux = y$ for x:

$$\begin{aligned} x_1 + 5x_2 - 2x_3 &= 3 \\ -18x_2 + 7x_3 &= -15 \\ -\frac{1}{2}x_3 &= \frac{15}{2} \end{aligned}$$

Her finder vi $x = (-2, -5, -15)^t$.

5 Overbestemt ligningssystem

Der er i praksis mange tilfælde, hvor et ligningssystem $Ax = b$ ikke har nogen løsning, men hvor man gerne stiller sig tilfreds med en bedste erstatning for en løsning.

Eksempel 91 *(Lineær regression). Givet m målepunkter i planen*

$$(x_1, y_1), (x_2, y_2), \ldots, (x_m, y_m)$$

Find den rette linie $y = ax + b$, der på bedst mulig måde repræsenterer datapunkterne. Bestemmes skal altså a og b. Til bestemmelse af disse to ubekendte har vi følgende m ligninger

$$\begin{aligned} ax_1 + b &= y_1 \\ ax_2 + b &= y_2 \\ &\vdots \\ ax_m + b &= y_m \end{aligned}$$

Hvis $m > 2$ må man formode, at ligningssystemet ikke har nogen løsning, også i det tilfælde hvor punkterne teoretisk set skulle ligge på en linie. Målinger er jo altid behæftet med en vis usikkerhed. Man siger, at man har et overbestemt ligningssystem. Der er altså for mange krav til de ubekendte. Vi skal så i stedet finde den bedste rette linie, der repræsenterer datapunkterne. Hvad der menes med 'bedst', må naturligvis præciseres. Vi vil ikke kræve, at linien nødvendigvis går igennem et eneste af punkterne, kun at linien tilgodeser alle punkterne bedst muligt. Dette er stadig meget vagt. Det viser sig beregningsmæssigt praktisk, at definere "den bedst mulige linie " som den, for hvilken summen af kvadraterne *på afvigelserne i y-retningen mellem punkterne og linien er mindst mulig, altså for hvilken*

$$f(a,b)=\sum_{i=1}^{m}(ax_i+b-y_i)^2$$

er mindst mulig. Minimering af denne funktion af to variable kan ske ved bestemmelse af stationære punkter:

$$\begin{aligned} f_a(a,b) &= 2\sum_{i=1}^{m}(ax_i+b-y_i)\,x_i \\ &= 2\left(a\sum_{i=1}^{m}x_i^2+b\sum_{i=1}^{m}x_i-\sum_{i=1}^{m}x_iy_i\right) \\ f_b(a,b) &= 2\sum_{i=1}^{m}(ax_i+b-y_i) \\ &= 2\left(a\sum_{i=1}^{m}x_i+bm-\sum_{i=1}^{m}y_i\right) \end{aligned}$$

Vi finder af kravet $\nabla f(a,b)=(0,0)$ de to lineære ligninger

$$\begin{aligned} a\sum_{i=1}^{m}x_i^2+b\sum_{i=1}^{m}x_i &= \sum_{i=1}^{m}x_iy_i \\ a\sum_{i=1}^{m}x_i+bm &= \sum_{i=1}^{m}y_i \end{aligned}$$

Med indførelsen af gennemsnittene $\overline{x}=\frac{1}{m}\sum_{i=1}^{m}x_i,\overline{x^2}=\frac{1}{m}\sum_{i=1}^{m}x_i^2$ osv. kan det stationære punkt udtrykkes således

$$\begin{aligned} a &= \frac{\overline{xy}-\overline{x}\cdot\overline{y}}{\overline{x^2}-(\overline{x})^2} \\ b &= \frac{\overline{x^2}\cdot\overline{y}-\overline{x}\cdot\overline{xy}}{\overline{x^2}-(\overline{x})^2} \end{aligned}$$

Der mangler så et argument for, at disse værdier for a og b svarer til mindsteværdi for f. Et sådant argument forsøger vi ikke her. Der kommer alligevel et generelt argument senere.

Eksempel 92 *Givet m målepunkter i planen $(x_1,y_1),(x_2,y_2),\ldots,(x_m,y_m)$. Find den kurve af formen $y=a_1f_1(x)+a_2f_2(x)+\ldots+a_kf_k(x)$, der på bedst mulig måde repræsenterer datapunkterne. Her tænkes det, at $f_1,f_2,\ldots,f_k$ er kendte konkrete funktioner. Bestemmes skal $a_1,a_2,\ldots,a_k$. Vi tænker os specielt på tilfældet, hvor $m>k$ Til bestemmelse af disse ubekendte har vi følgende m*

ligninger

$$\begin{aligned}
a_1 f_1(x_1) + a_2 f_2(x_1) + \ldots + a_k f_k(x_1) &= y_1 \\
a_1 f_1(x_2) + a_2 f_2(x_2) + \ldots + a_k f_k(x_2) &= y_2 \\
&\vdots \\
a_1 f_1(x_m) + a_2 f_2(x_m) + \ldots + a_k f_k(x_m) &= y_m
\end{aligned}$$

Dette ligningssystem er lineært og har formodentlig ingen løsninger. Systemet er overbestemt. Totalmatricen

$$\begin{pmatrix}
f_1(x_1) & f_2(x_1) & \ldots & f_k(x_1) & y_1 \\
f_1(x_2) & f_2(x_2) & \ldots & f_k(x_2) & y_2 \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
f_1(x_m) & f_2(x_m) & \ldots & f_k(x_m) & y_m
\end{pmatrix}$$

har formodentlig rang større end k. Vi skal da finde det bedste substitut for en løsning. Vi går ikke videre med dette eksempel, men vender os mod den generelle situation.

Vi tænker os et system af m ligninger med n ubekendte $Ax = b$, hvor koefficientmatricen altså er en $m \times n$ matrix, og hvor højresiden er en $m \times 1$ matrix, altså en søjlevektor. Vi tænker os den situation, hvor systemet ikke har nogen løsning. Vi opgiver imidlertid ikke så let, men stiller os tilfreds med det bedste substitut vi kan få for en løsning. En erstatning for en rigtig løsning er den vektor x, der minimerer afstanden mellem Ax og b. Når en rigtig løsning findes, vil denne afstand være nul, så beskrivelsen passer også på det tilfælde. Vi må nu afgøre, hvilket afstandsbegreb (hvilken norm), vi skal bruge. Der er mange muligheder, men den simpleste at have med at gøre *i denne sammenhæng*, er 2-normen, nemlig den sædvanlige euklidiske norm givet ved

$$\|y\| = \sqrt{y_1^2 + y_2^2 + \ldots + y_m^2}$$

når $y = (y_1, y_2, \ldots, y_m)^t$. Vi vil altså forsøge at finde den vektor $x \in R^n$, der minimerer størrelsen

$$\|Ax - b\| = \sqrt{(Ax-b)_1^2 + (Ax-b)_2^2 + \ldots + (Ax-b)_m^2}$$

eller ækvivalent hermed, minimerer størrelsen

$$\|Ax - b\|^2 = \sum_{i=1}^{m} (Ax-b)_i^2$$

Derfor tales der også i denne sammenhæng om mindste kvadraters metode. På forhånd og helt generelt kan det ikke udelukkes, at dette problem har flere løsninger for x.

Vi skal give en geometrisk udledning af formlen for løsningen (løsningerne). Vi udnytter, at Ax kan skrives som en linearkombination af A's søjler, $S_1, S_2, \ldots, S_n$, nemlig simpelthen således

$$Ax = x_1 S_1 + x_2 S_2 + \ldots + x_n S_n$$

Vi skal altså finde de tal $x_1, x_2, \ldots, x_n$, der sikrer, at $\|x_1 S_1 + x_2 S_2 + \ldots + x_n S_n - b\|$ er mindst mulig. Tænk på vektorerne $S_1, S_2, \ldots, S_n$ og b som pile udgående fra begyndelsespunktet. Den søgte minimale afstand er så afstanden fra spidsen af b til planen (rummet, nemlig $R(A)$) udspændt af pilene $S_1, S_2, \ldots, S_n$. Hvis tallene $x_1, x_2, \ldots, x_n$ vælges, så $x_1 S_1 + x_2 S_2 + \ldots + x_n S_n$ er projektionen af b ned i planen udspændt af $S_1, S_2, \ldots, S_n$, så er afstanden $\|x_1 S_1 + x_2 S_2 + \ldots + x_n S_n - b\|$ mindst mulig. For at finde denne projektion, bemærker vi, at vi må forlange, at

$$x_1 S_1 + x_2 S_2 + \ldots + x_n S_n - b \perp S_i$$

for alle $i = 1 \ldots n$. Men dette betyder, at skalarprodukterne er nul, dvs.

$$S_i \cdot (x_1 S_1 + x_2 S_2 + \ldots + x_n S_n - b) = 0$$

for alle $i = 1 \ldots n$. Skrevet med matrixprodukt i stedet har vi altså (idet vi igen erindrer, at $Ax = x_1 S_1 + x_2 S_2 + \ldots + x_n S_n$)

$$S_i^t (Ax - b) = 0$$

for alle $i = 1 \ldots n$. Men disse n ligninger kan under et skrives som

$$A^t (Ax - b) = 0$$

Hermed har vi altså fundet, at det bedste substitut x for en løsning skal opfylde ligningssystemet

$$A^t A x = A^t b$$

Dette ligningssystem udtrykker faktisk kun, at $Ax - b$ er ortogonal på alle søjlevektorerne i A. Systemet kaldes *normalligningssystemet* svarende til det overbestemte system $Ax = b$. Bemærk, at matricen $A^t A$ er kvadratisk. Når A er en $m \times n$ matrix, så er $A^t A$ en $n \times n$ matrix. Bemærk desuden, at $A^t A$ er en symmetrisk matrix, dvs. at den er lig med sin transponerede, $(A^t A)^t = A^t A$.

Eksempel 93 *Betragt systemet*

$$\begin{aligned}
x_1 + x_3 &= 1 \\
-2x_1 + 3x_2 - x_3 &= -2 \\
x_1 + x_2 &= 6 \\
-x_1 + x_3 &= 4 \\
-3x_1 + x_2 + 2x_3 &= 3
\end{aligned}$$

Koefficientmatricen A og højresiden b er givet ved

$$A = \begin{pmatrix} 1 & 0 & 1 \\ -2 & 3 & -1 \\ 1 & 1 & 0 \\ -1 & 0 & 1 \\ -3 & 1 & 2 \end{pmatrix}, b = \begin{pmatrix} 1 \\ -2 \\ 6 \\ 4 \\ 3 \end{pmatrix}$$

Ved Gausselimination vil man finde, at rangen af A er 3, mens rangen af totalmatricen er 4. Systemet har altså ingen løsninger. Vi finder i stedet den vektor x, der minimerer $\|Ax - b\|$, dvs. vi løser normalligningssystemet $A^t Ax = A^t b$. Vi finder

$$A^t A = \begin{pmatrix} 16 & -8 & -4 \\ -8 & 11 & -1 \\ -4 & -1 & 7 \end{pmatrix}, A^t b = \begin{pmatrix} -2 \\ 3 \\ 13 \end{pmatrix}$$

Dette system løses ved Gausselimination på sædvanlig måde. Vi finder

$$x = \left(\frac{4}{3}, \frac{3}{2}, \frac{17}{6} \right)^t$$

Som mål for, hvor god løsningen er, tager vi $\|Ax - b\|$. Først udregnes $Ax - b$ til

$$Ax - b = \left(\frac{19}{6}, 1, -\frac{19}{6}, -\frac{5}{2}, \frac{1}{6} \right)^t$$

hvorefter vi finder

$$\|Ax - b\| = \frac{1}{3}\sqrt{246} \cong 5.23$$

I dette eksempel, som er et rent regneeksempel løsrevet fra al sammenhæng, er det svært at sige om dette tal er stort eller småt.

Eksempel 94 *Vi vender tilbage til det første eksempel. Givet m målepunkter i planen $(x_1, y_1), (x_2, y_2), \ldots, (x_m, y_m)$. Find den rette linie $y = ax + b$, der på bedst mulig måde repræsenterer datapunkterne. Bestemmes skal a og b. Til bestemmelse af disse to ubekendte har vi følgende m ligninger*

$$\begin{aligned} ax_1 + b &= y_1 \\ ax_2 + b &= y_2 \\ &\vdots \\ ax_m + b &= y_m \end{aligned}$$

Koefficientmatricen A er

$$A = \begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_m & 1 \end{pmatrix}$$

Højresiden er $y = (y_1, y_2, \ldots, y_m)^t$. *Normalligningssystemet* $A^t A \begin{pmatrix} a \\ b \end{pmatrix} = A^t y$ *ser således ud*

$$\begin{pmatrix} \sum_{i=1}^m x_i^2 & \sum_{i=1}^m x_i \\ \sum_{i=1}^m x_i & m \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^m x_i y_i \\ \sum_{i=1}^m y_i \end{pmatrix}$$

Dette er netop ligningssystemet fundet (og også løst) i første eksempel.

Eksempel 95 *Fra bogen Scott & Crossman, Freshwater Fishes of Canada, Fisheries Research Board of Canada, 1973, henter vi en tabel (p.195) over længder i mm af laks (Salmo salar) fra Gambo Pond, Newfoundland.*

Alder	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*
Længde	*80*	*157*	*205*	*322*	*367*	*349*	*421*	*437*	*532*

Vi vil først finde den bedste rette linie i et Alder-Længde-koordinatsystem, der repræsenterer data bedst forstået på den måde, at summen af kvadraterne på afvigelserne fra linien er mindst mulig. Derefter vil vi finde det bedste polynomium af grad højst 2.

Vores første model til repræsentation af data er altså $L = at + b$, *hvor* L *er længden i mm, og* t *er alderen i år. For denne model har vi færdige formler for* a *og* b. *Vi angiver nogle mellemregninger. Antal ligninger er åbenbart* $m = 9$.

$$\sum_{i=1}^{9} t_i = 45, \sum_{i=1}^{9} t_i^2 = 285, \sum_{i=1}^{9} L_i = 2870, \sum_{i=1}^{9} t_i L_i = 17457$$

Hermed findes $a = \frac{3107}{60} \cong 51.8$ *og* $b = \frac{2159}{36} \cong 60.0$. *Hermed er den søgte sammenhæng*

$$L = 51.8t + 60.0$$

Kvadratroden af summen af kvadraterne på afvigelserne er

$$\sqrt{\sum_{i=1}^{9} (51.8t_i + 60.0 - L_i)^2} = 91.5$$

Vi forsøger os nu med modellen $L = at^2 + bt + c$. *De ligninger, der skal løses, er dermed*

$$at_i^2 + bt_i + c = L_i$$

hvor $i = 1 \ldots 9$. *Koefficientmatricen* A *har formen*

$$\begin{pmatrix} t_1^2 & t_1 & 1 \\ t_2^2 & t_2 & 1 \\ \vdots & \vdots & \vdots \\ t_9^2 & t_9 & 1 \end{pmatrix}$$

Højresiden er $L = (L_1, L_2, \ldots, L_9)$. Ved udregning findes A^tA til

$$A^tA = \begin{pmatrix} 15333 & 2025 & 285 \\ 2025 & 285 & 45 \\ 285 & 45 & 9 \end{pmatrix}$$

og højresiden af normalligningssystemet findes til $A^tL = (121133, 17457, 2870)^t$. Løsning af normalligningssystemet fører til

$$(a, b, c) = \left(-\frac{2461}{924}, \frac{120763}{1540}, \frac{78}{7}\right) \cong (-2.66, 78.4, 11.1)$$

Den søgte sammenhæng er altså

$$L = -2.66t^2 + 78.4t + 11.1$$

Kvadratroden af summen af kvadraterne på afvigelserne er

$$\sqrt{\sum_{i=1}^{9} \left(-2.66t_i^2 + 78.4t_i + 11.1 - L_i\right)^2} = 78.7$$

Det kan ikke undre, at introduktion af en parameter mere i modellen giver bedre repræsentation af data. Der er imidlertid også et ønske om at data repræsenteres af en simpel model. Man må her finde en balance. Til sidst bør bemærkes til begge modeller, at summen af afvigelsernes kvadrat -alt andet lige - vokser med med antallet af datapunkter. Skal denne sum derfor bruges ved sammenligning mellem to sæt af data med forskellige antal datapunkter, bør den justeres for antallet. Man kunne dividere med antal punkter m. I statistik bruger man at dividere med $m - 2$ i den lineære model og med $m - 3$ i den anden model.

Glemmer man udledningen og spørger om ethvert system af formen $A^tAx = A^tb$ er sikker på at have en løsning, må man altså spørge, om A^tb ligger i $R(A^tA)$. Skal dette gælde for alle b, må vi spørge, om $R(A^t) \subseteq R(A^tA)$. Svaret er ja, der gælder faktisk følgende sætning.

Sætning 96 *Lad A være en $m \times n$ matrix. Så gælder*

$$\begin{aligned} N\left(A^tA\right) &= N(A) \\ R\left(A^tA\right) &= R\left(A^t\right) \\ \rho\left(A^tA\right) &= \rho(A) = \rho\left(A^t\right) \end{aligned}$$

Bevis. Det er klart, at $N(A^tA) \sqsupseteq N(A)$. Vi skal vise den modsatte inklusion. Så lad $x \in N(A^tA)$, dvs. $A^tAx = 0$. Men så gælder også, at $0 = x^t A^tAx = (Ax)^t (Ax) = \|Ax\|^2$, dvs. $Ax = 0$, altså $x \in N(A)$. Den anden påstand følger af den første således.

$$R\left(A^tA\right) = N\left(A^tA\right)^\perp = N(A)^\perp = R\left(A^t\right)$$

Sidste påstand følger således

$$\rho\left(A^tA\right) = \dim R\left(A^tA\right) = \dim R\left(A^t\right) = \rho(A) = \rho\left(A^t\right)$$

■

Hvis rangen af A^tA er n (når A er en $m \times n$ matrix), så er A^tA invertibel, og normalligningssystemet $A^tAx = A^tb$ har den entydigt bestemte løsning

$$x = \left(A^tA\right)^{-1} A^tb$$

Fra udledningen af normalligningssystemet husker vi, at Ax skulle være projektionen af b ned i underrummet udspændt af søjlerne i A, altså $R(A)$. Når disse n søjler er lineært uafhængige er A^tA netop invertibel. Da $Ax = A(A^tA)^{-1}A^tb$ er projektionen Pb af vektoren b ned i underrummet $R(A)$ altså givet ved

$$Pb = A\left(A^tA\right)^{-1} A^tb$$

Bemærk, at $P = A(A^tA)^{-1}A^t$ har egenskaberne $P^2 = P$ og $P^t = P$, idet

$$\begin{aligned}
P^2 &= PP = \left(A\left(A^tA\right)^{-1}A^t\right)\left(A\left(A^tA\right)^{-1}A^t\right) \\
&= A\left(A^tA\right)^{-1}A^tA\left(A^tA\right)^{-1}A^t = AE\left(A^tA\right)^{-1}A^t \\
&= A\left(A^tA\right)^{-1}A^t = P
\end{aligned}$$

og videre

$$\begin{aligned}
P^t &= \left(A\left(A^tA\right)^{-1}A^t\right)^t = A\left(\left(A^tA\right)^{-1}\right)^tA^t \\
&= A\left(\left(A^tA\right)^t\right)^{-1}A^t = A\left(A^tA\right)^{-1}A^t = P
\end{aligned}$$

6 Determinanter

Vi kender determinantbegrebet for 2×2 matricer. Med

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

er determinanten $\det A$ givet ved

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

Vi kan ved direkte udregning hurtigt kontrollere, at den herved definerede determinantfunktion det har følgende egenskaber

$$\begin{aligned}
\begin{vmatrix} a & b \\ c+k & d+h \end{vmatrix} &= \begin{vmatrix} a & b \\ c & d \end{vmatrix} + \begin{vmatrix} a & b \\ k & h \end{vmatrix} \\
\begin{vmatrix} a & b \\ sc & sd \end{vmatrix} &= s\begin{vmatrix} a & b \\ c & d \end{vmatrix} \\
\begin{vmatrix} a & b \\ c & d \end{vmatrix} &= -\begin{vmatrix} c & d \\ a & b \end{vmatrix} \\
\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} &= 1
\end{aligned}$$

De to første betyder, at determinantfunktionen er en lineær funktion af anden række. Det samme gælder også for første række, men er ikke vist. Derfor siger man, at determinantfunktionen er multilineær. Den tredie ligning betyder, at determinantfunktionen er alternerende. Hermed menes blot, at fortegnet skifter, når to rækker ombyttes.

Vi kan også opskrive disse observationer således, idet vi opfatter determinantfunktionen som en funktion af matricens rækker

$$\begin{aligned}
\det(R_1, R_2' + R_2'') &= \det(R_1, R_2') + \det(R_1, R_2'') \\
\det(R_1, sR_2) &= s\det(R_1, R_2) \\
\det(R_2, R_1) &= -\det(R_1, R_2) \\
\det(E) &= \det(e_1, e_2) = 1
\end{aligned}$$

hvor $e_1 = (1,0)$ og $e_2 = (0,1)$. Vi skal nu generalisere determinantfunktionen til $n \times n$ matricer. Den vil blive defineret som en multilineær og alternerende funktion på rækkerne.

Definition 97 *Ved en determinant af en kvadratisk $n \times n$ matrix forstås en reel funktion* det *af matricens rækker, der er lineær i hver række, alternerende og opfylder, at* $\det(E) = 1$. *Dvs.*

$$\begin{aligned}
\det(R_1, R_2, \ldots, R_i' + R_i'', \ldots, R_n) &= \det(R_1, R_2, \ldots, R_i', \ldots, R_n) \\
&\quad + \det(R_1, R_2, \ldots, R_i'', \ldots, R_n) \\
\det(R_1, R_2, \ldots, sR_i, \ldots, R_n) &= s\det(R_1, R_2, \ldots, R_i, \ldots, R_n) \\
\det(R_1, R_2, \ldots, R_i, \ldots, R_j, \ldots, R_n) &= -\det(R_1, R_2, \ldots, R_j, \ldots, R_i, \ldots, R_n), \quad (i \neq j) \\
\det(E) &= \det(e_1, e_2, \ldots, e_n) = 1
\end{aligned}$$

Bemærkning 98 *Lineariteten kan udtrykkes kortere således*

$$\begin{aligned}
\det(R_1, R_2, \ldots, sR_i' + tR_i'', \ldots, R_n) &= s\det(R_1, R_2, \ldots, R_i', \ldots, R_n) \\
&\quad + t\det(R_1, R_2, \ldots, R_i'', \ldots, R_n)
\end{aligned}$$

Bemærkning 99 *Determinantfunktionen kunne ligeså godt være defineret som en multilineær og alternerende funktion af matricens søjler. Her har vi valgt rækkerne. Det skal senere vise sig, at* $\det(A^t) = \det(A)$. *Det er således uden betydning for det endelige resultat, hvilket synspunkt man anlægger.*

Sætning 100 *Hvis matricen* A, *har to ens rækker, så gælder, at* $\det(A) = 0$. *Hvis matricen* A_1 *fremkommer af matricen* A *ved rækkeoperationen* $R_i := R_i + kR_j$, *hvor* $j \neq i$, *så gælder* $\det(A_1) = \det(A)$.

Bevis. Første påstanden følger af, at opbytning af de to ens rækker ændrer fortegnet på determinanten. Anden påstand følger således

$$\begin{aligned}
\det(A_1) &= \det(R_1, R_2, \ldots, R_i + kR_j, \ldots, R_j, \ldots, R_n) \\
&= \det(R_1, R_2, \ldots, R_i, \ldots, R_j, \ldots, R_n) + \\
&\quad k\det(R_1, R_2, \ldots, R_j, \ldots, R_j, \ldots, R_n) \\
&= \det(A) + 0 = \det(A)
\end{aligned}$$

■

Det første vi må tage stilling til er, om der for en given værdi af n overhovedet eksisterer en multilineær og alternerende reel funktion, der også opfylder $\det(E) = 1$. Og hvis der gør, om der så er mere end én.

Eksempel 101 *Eksistensen for* $n = 2$ *er afklaret, det var der vi hentede inspirationen. Vi vil nu vise, at hvis* det *er multilineær, alternerende og opfylder* $\det(E) = 1$, *så må* $\det(A)$ *udregnes, som vi kender det. Med* $e_1 = (1, 0)$ *og* $e_2 = (0, 1)$ *har vi*

$$\begin{aligned}
\det\begin{pmatrix} a & b \\ c & d \end{pmatrix} &= \det([a, b], [c, d]) = a\det(e_1, [c, d]) + b\det(e_2, [c, d]) \\
&= a\{c\det(e_1, e_1) + d\det(e_1, e_2)\} + b\{c\det(e_2, e_1) + d\det(e_2, e_2)\} \\
&= ad\det(E) + bc\det(e_2, e_1) = ad - bc\det(E) = ad - bc
\end{aligned}$$

Eksempel 102 *Tilsvarende regninger kan udføres på*

$$\det\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

Arbejdets omfang er dog betydeligt større. Resultatet er, at determinanten, hvis den skal være multilineær og alternerende og opfylde $\det(E) = 1$, *nødvendigvis må være givet ved*

$$a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{21}a_{12}a_{33} + a_{21}a_{13}a_{32} + a_{31}a_{12}a_{23} - a_{31}a_{13}a_{22}$$

Hermed er vist, at der højst er én determinantfunktion. Man bør derefter undersøge om den, hvis den er således givet, er multilineær, alternerende og opfylder $\det(E) = 1$. *Det er ikke svært, men vi udelader regningerne. Det konkrete* *udtryk, som her er fundet for en 3×3 determinant, kan beregnes v.hj. Sarrus' regel (P.F.Sarrus, 1798-1861):*

$$\left| \begin{array}{ccc} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{array} \right| \begin{array}{cc} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{array}$$

Hvert af de tre led i udtrykket for determinanten, der har '+' som fortegn, fås som produktet af elementerne i en af de tre diagonaler, der vender mod SØ. Hvert af de tre led med '-' fås tilsvarende som produktet af elementerne i en de tre diagonaler, der vender mod SV. ***Sarrus' regel gælder kun for determinanter af tredie orden.***

For at kunne formulere det generelle resultat behøver vi et par begreber.

Definition 103 *Ved en permutation af tallene $1, 2, 3, \ldots, n$ forstås en bijektiv afbildning af mængden $\{1, 2, 3, \ldots, n\}$ på sig selv. Mængden af sådanne permutationer betegner vi med S_n. En permutation $s \in S_n$ kan beskrives ved følgen $s(1), s(2), s(3), \ldots, s(n)$ af funktionsværdier skrevet i rækkefølge. Ved en inversion i $s \in S_n$ forstås et par $s(i), s(j)$ der opfylder $i < j$ og $s(i) > s(j)$. Ved inversionstallet $I(s)$ for s forstås antallet af inversioner i s. En permutation kaldes lige, hvis dens inversionstal er et lige tal, ellers kaldes den ulige.*

Eksempel 104 *Lad $n = 5$. Ved følgen $4, 2, 5, 1, 3$ defineres en permutation s, for hvilken $s(1) = 4, s(2) = 2, s(3) = 5, s(4) = 1$ og $s(5) = 3$. I permutationen findes følgende inversioner: $(4, 2), (4, 1), (4, 3), (2, 1), (5, 1), (5, 3)$. Inversionstallet er dermed 6. Da dette tal er lige, er permutationen lige.*

Sætning 105 *For ethvert naturligt tal n findes præcis én funktion, der opfylder kravene til determinanten, nemlig, at den som funktion af rækkerne er multilineær, alternerende og har værdien 1 på enhedsmatricen. Denne determinantfunktion er givet ved regneudtrykket*

$$\det(A) = \sum_{s \in S_n} (-1)^{I(s)} a_{1s(1)} a_{2s(2)} \ldots a_{ns(n)}$$

Bevis. Som for $n = 2$ fås i første omgang

$$\det(A) = \sum_{s \in S_n} \det\left(e_{s(1)}, e_{s(2)}, \ldots, e_{s(n)}\right) a_{1s(1)} a_{2s(2)} \ldots a_{ns(n)}$$

hvor $e_1 = (1, 0, 0, \ldots, 0), e_2 = (0, 1, 0, \ldots, 0), \ldots, e_n = (0, 0, \ldots, 0, 1)$. Ved succesive ombytninger af to naboer i talfølgen $s(1), s(2), s(3), \ldots, s(n)$ kan denne omformes til talfølgen $1, 2, 3, \ldots, n$. Ved ombytning af to naboer enten forsvinder eller opstår en inversion. Naboombytningen har ingen indflydelse på eksistensen af evt. andre inversioner. Ved naboombytning ændres inversionstallet altså med præcis én, op eller ned. Dvs. at inversionstallets *paritet* ændres: Lige bliver til ulige, og omvendt. Ved hver ombytning skifter determinanten fortegn. Ved omformningen fra talfølgen $s(1), s(2), s(3), \ldots, s(n)$ til talfølgen $1, 2, 3, \ldots, n$ sker der altså samme antal paritetsændringer som ændringer i fortegnet ved omformningen fra det $\left(e_{s(1)}, e_{s(2)}, \ldots, e_{s(n)}\right)$ til det $(e_1, e_2, \ldots, e_n)$. Heraf følger, at det $\left(e_{s(1)}, e_{s(2)}, \ldots, e_{s(n)}\right) = (-1)^{I(s)} \det(e_1, e_2, \ldots, e_n) = (-1)^{I(s)}$. ■

Bemærkning 106 *Udregning af determinanten efter formlen i sætningen ovenfor er en håbløs sag, når n blot er nogenlunde stor. Antallet af permutationer i S_n er nemlig $n!$. Så summen indeholder $n!$ led. Dermed skal der udregnes $n!(n-1)$ multiplikationer. Er eksempelvis $n = 10$, så er antallet af permutationer $10! = 3628\,800$. Det antal multiplikationer, der skal udføres, er $10! \cdot 9 = 32\,659\,200$. Sætningen er derfor hovedsagelig af interesse ved at garantere os, at determinantfunktionen eksisterer og er entydigt bestemt. Vi skal dog også udnytte sætningen i beviset for den næste sætning.*

Sætning 107 *For enhver kvadratisk matrix A gælder* $\det(A^t) = \det(A)$.

Bevis. Vi udnytter sætningen ovenfor. Leddet $a_{1s(1)}a_{2s(2)} \ldots a_{ns(n)}$ kan også skrives $a_{t(1)1}a_{t(2)2} \ldots a_{t(n)n}$, hvor t er den inverse permutation til s, $t = s^{-1}$, dvs. hvis $s(i) = k$, så $t(k) = i$. Vi har, at $I\left(s^{-1}\right) = I(s)$, hvilket ses som følger: Hvis $s(i), s(j)$ er en inversion i s, så har vi $i < j$ og $s(i) > s(j)$. Men med $k = s(i)$ og $m = s(j)$ har vi, $i = t(k)$ og $j = t(m)$, og dermed $m < k$ og $t(m) > t(k)$, hvormed $t(m), t(k)$ er en inversion i t. Til enhver inversion i s svarer altså en inversion i t. Da det omvendte også gælder, følger påstanden. Hermed kan det fulde led $(-1)^{I(s)} a_{1s(1)}a_{2s(2)} \ldots a_{ns(n)}$ også skrives $(-1)^{I(t)} a_{t(1)1}a_{t(2)2} \ldots a_{t(n)n}$, hvoraf sætningen følger. ■

Lemma 108 *Der gælder, at*

$$\begin{vmatrix} 1 & 0 & \ldots & 0 \\ a_{21} & a_{22} & \ldots & a_{2n} \\ a_{31} & a_{32} & \ldots & a_{3n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{22} & \ldots & a_{2n} \\ a_{32} & \ldots & a_{3n} \\ \vdots & \vdots & \vdots \\ a_{n2} & \ldots & a_{nn} \end{vmatrix}$$

Bevis. På determinanten til venstre udføres rækkeoperationerne $R_i := R_i - a_{i1}R_1$ med $i = 2 \ldots n$ Determinantens værdi ændres ikke herved. Dermed opnås følgende determinant

$$\begin{vmatrix} 1 & 0 & \ldots & 0 \\ 0 & a_{22} & \ldots & a_{2n} \\ 0 & a_{32} & \ldots & a_{3n} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & a_{n2} & \ldots & a_{nn} \end{vmatrix}$$

Men denne funktion af $(n-1)$-rækkevektorerne

$$(a_{22}, a_{23}, \ldots, a_{2n}), (a_{32}, a_{33}, \ldots, a_{3n}), \ldots, (a_{n2}, a_{n3}, \ldots, a_{nn})$$

er multilineær og alternerende. Endvidere er værdien på enhedsmatricen af $(n-1)$'te orden lig med 1. Da determinantfunktionen er entydigt bestemt ved disse egenskaber, følger lemmaets påstand. ■

Sætning 109 *Determinanten af en øvre eller nedre triangulær matrix er lig med produktet af diagonalelementerne.*

Bevis. Da $\det(A^t) = \det(A)$ behøver vi kun betragte en nedre triangulær matrix. Lad altså A være nedre triangulær. Så har $\det(A)$ formen

$$\det(A) = \begin{vmatrix} a_{11} & 0 & 0 & \dots & 0 \\ a_{21} & a_{22} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n-1,1} & a_{n-1,2} & \dots & a_{n-1,n-1} & 0 \\ a_{n1} & a_{n2} & \dots & a_{n,n-1} & a_{nn} \end{vmatrix}$$

Ved brug af lineariteten i første række fås

$$\det(A) = a_{11} \begin{vmatrix} 1 & 0 & 0 & \dots & 0 \\ a_{21} & a_{22} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n-1,1} & a_{n-1,2} & \dots & a_{n-1,n-1} & 0 \\ a_{n1} & a_{n2} & \dots & a_{n,n-1} & a_{nn} \end{vmatrix}$$

Herefter giver lemmaet ovenfor, at

$$\det(A) = a_{11} \begin{vmatrix} a_{22} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ a_{n-1,2} & \dots & a_{n-1,n-1} & 0 \\ a_{n2} & \dots & a_{n,n-1} & a_{nn} \end{vmatrix}$$

Men herefter kan igen lineariteten i første række udnyttes, hvorved a_{22} kommer ud foran. Lemmaet bruges igen, osv. Man slutter med 1×1 determinanten (tallet) a_{nn}. Hermed har vi, at $\det(A) = a_{11}a_{22}\dots a_{nn}$. ■

Eksempel 110 *Vi vil udregne determinanten af følgende matrix*

$$A = \begin{pmatrix} 7 & 9 & 13 & 14 \\ -14 & -18 & -24 & -23 \\ 21 & 24 & 45 & 56 \\ 35 & 54 & 75 & 93 \end{pmatrix}$$

Vi forsøger standard Gausselimination på rækkerne. Vi udfører operationerne $R_2 := R_2 + 2R_1, R_3 := R_3 - 3R_1, R_4 := R_4 - 5R_1$. *Herved ændres værdien af determinanten ikke. Vi har derfor*

$$\det(A) = \begin{vmatrix} 7 & 9 & 13 & 14 \\ 0 & 0 & 2 & 5 \\ 0 & -3 & 6 & 14 \\ 0 & 9 & 10 & 23 \end{vmatrix}$$

Vi kunne nu udnytte lemmaet i forbindelse med $\det(A^t) = \det(A)$ *til at få resultatet*

$$\det(A) = 7 \begin{vmatrix} 0 & 2 & 5 \\ -3 & 6 & 14 \\ 9 & 10 & 23 \end{vmatrix}$$

men her foretrækker vi at fortsætte med 4×4 *determinanten. Vi udfører rækkeoperationen* $R_2 \longleftrightarrow R_3$*, der ændrer fortegnet på determinanten. Herved opnås*

$$\det(A) = - \begin{vmatrix} 7 & 9 & 13 & 14 \\ 0 & -3 & 6 & 14 \\ 0 & 0 & 2 & 5 \\ 0 & 9 & 10 & 23 \end{vmatrix}$$

Rækkeoperationen $R_4 := R_4 + 3R_2$ *ændrer ikke determinantens værdi. Hermed har vi*

$$\det(A) = - \begin{vmatrix} 7 & 9 & 13 & 14 \\ 0 & -3 & 6 & 14 \\ 0 & 0 & 2 & 5 \\ 0 & 0 & 28 & 65 \end{vmatrix}$$

Til sidst udføres rækkeoperationen $R_4 := R_4 - 14R_3$*, der heller ikke ændrer determinantens værdi. Hermed har vi*

$$\det(A) = - \begin{vmatrix} 7 & 9 & 13 & 14 \\ 0 & -3 & 6 & 14 \\ 0 & 0 & 2 & 5 \\ 0 & 0 & 0 & -5 \end{vmatrix} = -(7 \cdot (-3) \cdot 2 \cdot (-5)) = -210$$

hvor vi til slut har brugt, at vi har determinanten af en øvre triangulær matrix.

Sætning 111 *For enhver kvadratisk matrix* A *gælder, at følgende udsagn er ækvivalente:*

1. *det*$(A) \neq 0$
2. A *er invertibel*
3. A *har fuld rang*
4. *Ligningen* $Ax = 0$ *har kun løsningen* $x = 0$

Bevis. De tre sidste påstande er tidligere vist at være ækvivalente. Derfor behøver vi blot vise, at påstand 1 og påstand 3 er ækvivalente. Ved Gausselimination kan A under alle omstændigheder reduceres til en matrix Q på echelonform. Af de sædvanlige Gaussoperationer er der kun rækkeombytninger, der ændrer determinanten, og det kun dens fortegn. Altså har vi, at $\det(A) = k \det(Q)$, hvor $k \neq 0$. Men Q er en øvre triangulær matrix, så $\det(Q)$ er lig med produktet af diagonalelementerne i Q. Dette produkt er forskellig fra nul, hvis og kun hvis A har fuld rang. Hermed er sætningen vist. ∎

Sætning 112 *For ethvert par af kvadratiske matricer A og B gælder, at*

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

Bevis. Hvis $\det(B) = 0$, så findes der iflg. den foregående sætning en vektor $u \neq 0$, med $Bu = 0$. Men så fås, at $ABu = 0$, hvoraf omvendt sluttes, at $\det(AB) = 0$. Hermed gælder sætningens påstand, når $\det(B) = 0$. Antag nu, at $\det(B) \neq 0$. Vi vil vise, at udtrykket $\frac{\det(AB)}{\det(B)}$ betragtet som funktion af rækkerne i A er multilineær og alternerende og desuden har værdien 1 på enhedsmatricen. Kan vi nemlig vise disse tre ting, så må dette udtryk være determinanten af A. Vi begynder med det simpleste. Erstatter vi A med E i $\frac{\det(AB)}{\det(B)}$ får vi åbenbart 1. For at vise de to andre bemærker vi, at produktet AB kan skrives

$$AB = \begin{pmatrix} R_1 \\ R_2 \\ \vdots \\ R_n \end{pmatrix} B = \begin{pmatrix} R_1 B \\ R_2 B \\ \vdots \\ R_n B \end{pmatrix}$$

Derfor har vi, at

$$\frac{\det(AB)}{\det(B)} = \frac{1}{\det(B)} \begin{vmatrix} R_1 B \\ R_2 B \\ \vdots \\ R_n B \end{vmatrix}$$

der (betragtet som funktion af rækkerne $R_1, R_2, \ldots, R_n$) tydeligvis er multilineær og alternerende. Hermed er sætningen vist. ■

Korollar 113 *Hvis A er invertibel gælder*

$$\det\left(A^{-1}\right) = \det(A)^{-1}$$

Bemærkning 114 *Der gælder ikke nogen simpel formel for* $\det(A+B)$. *Men der gælder, at* $\det(sA) = s^n \det(A)$, *når A er en $n \times n$ matrix og s er et tal.*

Den følgende løsningsformel for ligningssystemet $Ax = b$ kaldes normalt Cramers regel (Gabriel Cramer, 1704-1752).

Sætning 115 ***Cramers regel.*** *Lad A være en invertibel $n \times n$ matrix. Så er løsningen $x = (x_1, x_2, \ldots, x_n)^t$ til ligningssystemet $Ax = b$ givet ved*

$$x_i = \frac{\det(B_i)}{\det(A)}$$

for alle $i = 1 \ldots n$, hvor matricen B_i er den matrix, der fås ved i matricen A at erstatte søjle nr. i med ligningssystemets højre side b.

Bevis. Skriv A udtrykt ved sine søjler som $A = (S_1, S_2, \ldots, S_n)$. Så kan B_i skrives

$$B_i = (S_1, S_2, \ldots, S_{i-1}, b, S_{i+1}, \ldots, S_n)$$

Lad i være et vilkårligt af tallene $1 \ldots n$. Lad X_i være den matrix, der udtrykt ved søjlerne er $X_i = (e_1, e_2, \ldots, e_{i-1}, x, e_{i+1}, \ldots, e_n)$, hvor $e_1, e_2, \ldots, e_n$ er de sædvanlige (her søjle-)enhedsvektorer,

$$e_1 = (1, 0, 0, \ldots, 0)^t, e_2 = (0, 1, 0, \ldots, 0)^t, \ldots, (0, 0, \ldots, 0, 1)^t$$

Da $Ae_j = S_j$ for alle j, ses det, at

$$\begin{aligned} AX_i &= (Ae_1, Ae_2, \ldots, Ae_{i-1}, Ax, Ae_{i+1}, \ldots, Ae_n) \\ &= (S_1, S_2, \ldots, S_{i-1}, Ax, S_{i+1}, \ldots, S_n) \end{aligned}$$

Men heraf ses, at $AX_i = B_i$, hvis og kun hvis $Ax = b$. Er x altså løsning til $Ax = b$, så følger, at $\det(AX_i) = \det(B_i)$. Men heraf fås så, at $\det(A)\det(X_i) = \det(B_i)$. Men $\det(X_i) = x_i$. Hermed har vi Cramers regel. ■

Eksempel 116 *Vi vil løse systemet $Ax = b$ givet ved*

$$\begin{pmatrix} 8 & 4 & -5 \\ -5 & 3 & -5 \\ 8 & 5 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \\ -4 \end{pmatrix}$$

Vi ønsker at bruge Cramers regel. Vi skal altså udregne 4 determinanter. Først finder vi $\det(A) = 241$. *Derefter*

$$\begin{aligned} \det(B_1) &= \begin{vmatrix} 0 & 4 & -5 \\ 3 & 3 & -5 \\ -4 & 5 & -1 \end{vmatrix} = -43 \\ \det(B_2) &= \begin{vmatrix} 8 & 0 & -5 \\ -5 & 3 & -5 \\ 8 & -4 & -1 \end{vmatrix} = -164 \\ \det(B_3) &= \begin{vmatrix} 8 & 4 & 0 \\ -5 & 3 & 3 \\ 8 & 5 & -4 \end{vmatrix} = -200 \end{aligned}$$

Hermed er løsningen til ligningssystemet

$$x = \left(-\frac{43}{241}, -\frac{164}{241}, -\frac{200}{241}\right)$$

Bemærkning 117 *Som det vist allerede fremgår af dette eksempel, er Cramers regel altfor arbejdskrævende, med mindre systemet er meget småt. For et $n \times n$ system skal $n + 1$ determinanter regnes ud. Det vil normalt være billigst at udregne disse determinanter ved brug af Gausselimination. Men hvorfor så ikke bruge Gausselimination på totalmatricen for systemet $Ax = b$?*

Da den inverse matrix til matricen A jo er løsningen til matrixligningen $AX = E$, er det klart, at vi v.hj.a. Cramers regel kan finde en formel for den inverse udtrykt ved determinanter. Selv om denne formel kun er nyttig i specielle tilfælde, skal vi alligevel udlede den, idet vi i beviset må indføre et nyttigt begreb, de såkaldte komplementer til en matrix.

Definition 118 *Lad A være en $n \times n$ matrix. Ved matricens (i, j)-komplement forstås determinanten*

$$K_{ij} = \det(R_1, R_2, \ldots, R_{i-1}, e_j, R_{i+1}, \ldots, R_n)$$

hvor $R_1, R_2, \ldots, R_n$ er rækkerne i A, og hvor e_j er en rækkevektor med et 1-tal i den j'te søjle og nuller i de andre.

Sætning 119 *Lad A være en $n \times n$ matrix. Så gælder:*

1. *Komplementet K_{ij} kan udregnes som følger*

$$K_{ij} = (-1)^{i+j} D_{ij}$$

   hvor underdeterminanten D_{ij} er den $(n-1) \times (n-1)$ determinant, der fås ved at stryge den i'te række og den j'te søjle i $\det(A)$.

2. *Determinanten af $A = (a_{ij})$ kan udregnes v.hj.a. komplementerne som følger*

$$\det(A) = \sum_{j=1}^{n} a_{ij} K_{ij}$$

   *hvor $i = 1 \ldots n$ kan vælges vilkårligt. (*Udvikling langs i'te række*).*

3. *Determinanten af $A = (a_{ij})$ kan også udregnes v.hj.a. komplementerne som følger*

$$\det(A) = \sum_{i=1}^{n} a_{ij} K_{ij}$$

   *hvor $j = 1 \ldots n$ kan vælges vilkårligt. (*Udvikling langs j'te søjle*).*

Bevis. Ved i $K_{ij} = \det(R_1, R_2, \ldots, R_{i-1}, e_j, R_{i+1}, \ldots, R_n)$ at foretage $i-1$ naboombytninger af vektoren e_j opnås

$$K_{ij} = (-1)^{i-1} \det(e_j, R_1, R_2, \ldots, R_{i-1}, R_{i+1}, \ldots, R_n)$$

Da determinanter også er alternerende som funktion af søjlerne, får vi følgende ved $j-1$ naboombytninger af den j'te søjle

$$K_{ij} = (-1)^{i+j-2} \begin{vmatrix} 1 & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ a_{1j} & a_{11} & a_{12} & \dots & a_{1,j-1} & a_{1,j+1} & \dots & a_{1n} \\ a_{2j} & a_{21} & a_{22} & \dots & a_{2,j-1} & a_{2,j+1} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i-1,j} & a_{i-1,1} & a_{i-1,2} & \dots & a_{i-1,,j-1} & a_{i-1,j+1} & \dots & a_{i-1,n} \\ a_{i+1,j} & a_{i+1,1} & a_{i+1,2} & \dots & a_{i+1,,j-1} & a_{i+1,j+1} & \dots & a_{i+1,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{nj} & a_{n1} & a_{n2} & \dots & a_{n,j-1} & a_{n,j+1} & \dots & a_{nn} \end{vmatrix}$$

Da nu $(-1)^{i+j-2} = (-1)^{i+j}$ fås ved udnyttelse af lemmaet, at

$$K_{ij} = (-1)^{i+j} \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1,j-1} & a_{1,j+1} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2,j-1} & a_{2,j+1} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i-1,1} & a_{i-1,2} & \dots & a_{i-1,,j-1} & a_{i-1,j+1} & \dots & a_{i-1,n} \\ a_{i+1,1} & a_{i+1,2} & \dots & a_{i+1,,j-1} & a_{i+1,j+1} & \dots & a_{i+1,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{n,j-1} & a_{n,j+1} & \dots & a_{nn} \end{vmatrix}$$

hvilket netop er den søgte formel $K_{ij} = (-1)^{i+j} D_{ij}$.

For at vise udviklingen af determinanten langs en række benytter vi, at den i'te række R_i i A kan skrives

$$R_i = \sum_{j=1}^{n} a_{ij} e_j$$

Hermed følger af lineariteten af determinanten

$$\begin{aligned} \det(A) &= \sum_{j=1}^{n} a_{ij} \det(R_1, R_2, \dots, R_{i-1}, e_j, R_{i+1}, \dots, R_n) \\ &= \sum_{j=1}^{n} a_{ij} K_{ij} \end{aligned}$$

Formlen for udvikling af determinanten langs den j'te søjle kan fås af ovenstående formel ved at udvikle $\det(A^t)$ langs den j'te række og udnytte, at (j,i)-komplementet for A^t er lig med (i,j)-komplementet for A. Dette ses let af det konkrete udtryk for K_{ij} givet i den første del af beviset. ■

Vi er nu istand til at give den tidligere omtalte formel for A^{-1} udtrykt ved determinanter.

Sætning 120 *For enhver kvadratisk matrix A gælder*

$$AK^t = \det(A)\, E$$

hvor $K = (K_{ij})$ betegner matricen bestående af komplementerne for A. Er A en invertibel matrix, så kan A^{-1} findes ved formlen

$$A^{-1} = \det(A)^{-1} K^t$$

Bevis. Vi benytter som ovenfor, at den i'te række R_i i A kan skrives

$$R_i = \sum_{j=1}^{n} a_{ij} e_j$$

Hermed følger af lineariteten af determinanten, når $i < k$

$$\begin{aligned} & \det(R_1, R_2, \ldots, R_{i-1}, R_i, R_{i+1}, \ldots, R_{k-1}, R_i, R_{k+1}, \ldots, R_n) \\ = & \sum_{j=1}^{n} a_{ij} \det(R_1, R_2, \ldots, R_{i-1}, R_i, R_{i+1}, \ldots, R_{k-1}, e_j, R_{k+1}, \ldots, R_n) \\ = & \sum_{j=1}^{n} a_{ij} K_{kj} \end{aligned}$$

Men da denne determinant har to ens rækker, er den nul. For $i > k$ gælder det samme resultat. Hermed har vi

$$\left(AK^t\right)_{ik} = \sum_{j=1}^{n} a_{ij} \left(K^t\right)_{jk} = \sum_{j=1}^{n} a_{ij} K_{kj}$$

der altså er nul for $i \neq k$. For $i = k$ er værdien lig med $\det(A)$ iflg. en tidligere sætning. Hermed er sætningen vist. ∎

Bemærkning 121 *Vi valgte at give et bevis, der ikke byggede på Cramers regel, idet et sådant bevis kun ville gælde for invertible matricer. Forudsætter vi imidlertid, at A er invertibel, så kan ovenstående sætning vises således v.hj.a. Cramers regel:*

$$\left(A^{-1}\right)_{ij} = \left(A^{-1} e_j\right)_i = \frac{\det(B_{ij})}{\det(A)} = \frac{K_{ji}}{\det(A)}$$

hvor matricen B_{ji} fremkommer ved at erstatte den i'te søjle i A med søjlen e_j. Men så gælder netop $\det(B_{ij}) = K_{ji}$.

Bemærkning 122 *Omvendt kan Cramers regel bevises ud fra ovenstående sætning således, hvor vi vil løse $Ax = b$, dvs. finde $x = A^{-1}b$:*

$$\begin{aligned} x_i &= \det(A)^{-1} \left(K^t b\right)_i = \det(A)^{-1} \sum_{j=1}^{n} \left(K^t\right)_{ij} b_j \\ &= \det(A)^{-1} \sum_{j=1}^{n} K_{ji} b_j \end{aligned}$$

Men udvikles $\det(B_i)$ langs den i'te søjle, fås $\det(B_i) = \sum_{j=1}^{n} K_{ji} b_j$.

Bemærkning 123 *Fastholder man en eller flere rækker i en kvadratisk matrix og betragter* $\det(A)$ *som en funktion af de øvrige rækker, så er det klart, at denne funktion bliver multilineær og alternerende. Omvendt kan det let vises, at hvis en reel funktion på* $R^n \times R^n \times \ldots \times R^n$ *(k faktorer) er multilineær og alternerende, så kan den, når* $k < n$*, skrives som* $n \times n$ *determinant, der er opnået ved tilføjelse af* $n - k$ *faste rækker. Når* $k > n$*, kan den skrives som en* $k \times k$ *determinant ved tilføjelse af* $k - n$ *faste søjler. Konklusionen er da den, at en udvidelse af determinantbegrebet til også at omfatte rektangulære matricer egentlig ikke fører til noget interessant nyt.*

7 Egenværdier og egenvektorer

7.1 Diagonalmatricer, Similaritet, Diagonaliserbarhed

Diagonalmatricer er specielt simple at regne med. Er Λ en diagonalmatrix, har den formen

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 & 0 \\ 0 & \lambda_2 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & \lambda_{n-1} & 0 \\ 0 & 0 & \ldots & 0 & \lambda_n \end{pmatrix}$$

Dette vil vi, når det er belejligt, skrive således $\Lambda = diag\,(\lambda_1, \lambda_2, \ldots, \lambda_n)$. Produktet af to diagonalmatricer $\Lambda = diag\,(\lambda_1, \lambda_2, \ldots, \lambda_n)$ og $M = diag\,(\mu_1, \mu_2, \ldots, \mu_n)$ er igen en diagonalmatrix, og man finder

$$\Lambda M = M\Lambda = diag\,(\lambda_1\mu_1, \lambda_2\mu_2, \ldots, \lambda_n\mu_n)$$

Derfor har vi også, at

$$\Lambda^k = diag\left(\lambda_1^k, \lambda_2^k, \ldots, \lambda_n^k\right)$$

for alle ikke negative hele tal k. Da en diagonalmatrix Λ allerede er på echelonform, ser vi med samme, at Λ er invertibel, hvis og kun hvis alle diagonalelementerne er forskellige fra nul, dvs. $\lambda_i \neq 0$ for $i = 1 \ldots n$. I så fald er den inverse Λ^{-1} diagonalmatricen $\Lambda^{-1} = diag\left(\lambda_1^{-1}, \lambda_2^{-1}, \ldots, \lambda_n^{-1}\right)$.

Definition 124 *To matricer kvadratiske matricer af samme størrelse* A *og* B *kaldes similære, hvis der findes en invertibel matrix* S*, så*

$$B = S^{-1}AS$$

Bemærkning 125 *Manglen på symmetri i definitionen er kun tilsyneladende, thi hvis* $B = S^{-1}AS$*, så gælder også* $A = SBS^{-1} = \left(S^{-1}\right)^{-1} BS^{-1}$ *og omvendt.*

Er to matricer similære, kan regning med den ene let reduceres til regning med den anden. Antag, at A og B er similære med $B = S^{-1}AS$. Så gælder, at $B^k = S^{-1}A^kS$ for alle ikke-negative hele tal k. For at indse dette, tag f.eks. $k = 3$. Så har vi

$$\begin{aligned} B^3 &= \left(S^{-1}AS\right)\left(S^{-1}AS\right)\left(S^{-1}AS\right) \\ &= S^{-1}A\left(SS^{-1}\right)A\left(SS^{-1}\right)AS \\ &= S^{-1}AEAEAS = S^{-1}A^3S \end{aligned}$$

Er A invertibel, så er også B invertibel, og $B^{-1} = S^{-1}A^{-1}S$. Dette ses således:

$$\begin{aligned} B\left(S^{-1}A^{-1}S\right) &= \left(S^{-1}AS\right)\left(S^{-1}A^{-1}S\right) \\ &= S^{-1}A\left(SS^{-1}\right)A^{-1}S \\ &= S^{-1}\left(AA^{-1}\right)S = S^{-1}ES = S^{-1}S = E \end{aligned}$$

Da regning med diagonalmatricer er særligt simpel, og da regning med en af to similære matricer kan reduceres til regning med den anden, kan det ikke undre, at spørgsmålet, om en given matrix er similær med en diagonalmatrix, er godt at få besvaret.

Definition 126 *En kvadratisk matrix A siges at kunne diagonaliseres (eller at være diagonaliserbar), hvis den er similær med en diagonalmatrix Λ. Hvis S er en invertibel matrix med*

$$\Lambda = S^{-1}AS$$

så vil vi kalde S en diagonaliserende matrix for A.

7.2 Egenværdier og egenvektorer

Definition 127 *Lad A være en kvadratisk matrix. Et tal λ kaldes en egenværdi for A, hvis der findes en vektor $v \neq 0$, så $Av = \lambda v$. En sådan vektor kaldes en egenvektor for A hørende til egenværdien λ.*

Bemærkning 128 *At vi uanset værdien af λ altid har $A0 = \lambda 0$, gør ikke λ til en egenværdi for A. Men hvis λ er en egenværdi for A, så er det mest bekvemt at acceptere, at nulvektoren er en af egenvektorerne. Er $v \neq 0$ en egenvektor for A hørende til egenværdien λ, så er også tv en egenvektor for alle værdier af skalaren t. Vi har jo $A(tv) = tAv = t(\lambda v) = \lambda(tv)$.*

Bemærkning 129 *Vi skal normalt forudsætte, at såvel matricer som vektorer har reelle elementer. I så fald kan ikke-reelle værdier af λ ikke accepteres.*

Eksempel 130 *Lad matricen A være givet ved*

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

Vi vil finde egenværdier og egenvektorer for A. Vi skal altså afgøre for hvilke tal λ ligningssystemet $Av - \lambda v = 0$ har mindst én ikke-triviel løsning v. Systemet kan skrives $(A - \lambda E)\, v = 0$. Skrevet ud fås

$$\begin{pmatrix} 2-\lambda & -2 \\ -1 & 3-\lambda \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Dette homogene system har en ikke-triviel løsning, hvis og kun hvis ligningssystemets determinant er lig med nul, dvs. hvis og kun hvis

$$\begin{vmatrix} 2-\lambda & -2 \\ -1 & 3-\lambda \end{vmatrix} = 0$$

Egenværdierne for A er altså løsningerne til følgende ligning

$$\lambda^2 - 5\lambda + 4 = 0$$

som vi vil kalde karakterligningen for A. Løsningerne er 4 og 1. Disse to tal er altså matricens egenværdier. Vi bestemmer nu de tilhørende egenvektorer. Først betragter vi egenværdien 4. Egenvektorerne $v = (v_1, v_2)^t$ skal opfylde

$$\begin{pmatrix} -2 & -2 \\ -1 & -1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Dette system er ækvivalent med den ene ligning $v_1 + v_2 = 0$. Samtlige egenvektorer hørende til egenværdien 4 er derfor givet ved

$$v = t \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad t \in R$$

Til slut bestemmer vi egenvektorerne hørende til egenværdien 1. Sådanne vektorer er de ikke-trivielle løsninger v til ligningssystemet

$$\begin{pmatrix} 1 & -2 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Dette system er ækvivalent med ligningen $v_1 - 2v_2 = 0$. Samtlige egenvektorer hørende til egenværdien 1 er altså givet ved

$$v = t \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad t \in R$$

Den i eksemplet anvendte metode til bestemmelse af egenværdier og egenvektorer kan umiddelbart generaliseres:

Sætning 131 *Et tal λ er egenværdi for den kvadratiske matrix A, hvis og kun hvis karakterligningen* $\det(A - \lambda E) = 0$ *er opfyldt. De til egenværdien λ hørende egenvektorer er løsningerne v til ligningssystemet $(A - \lambda E)\, v = 0$. Dvs. $N(A - \lambda E)$ er mængden af egenvektorer.*

Definition 132 *Polynomiet* $\det(A - \lambda E)$ *kaldes karakterpolynomiet for* A. *Mængden ef egenvektorer hørende til egenværdien* λ, *altså* $N(A - \lambda E)$ *kaldes egenrummet hørende til egenværdien* λ. *Hvis* $\dim N(A - \lambda E) = k$, *så siges egenværdien* λ *at have geometrisk multiplicitet* k. *Hvis karakterligningen har* λ *som en rod af multiplicitet* q, *så siges egenværdien* λ *at have algebraisk multiplicitet* q.

Eksempel 133 *Lad matricen* A *være givet ved*

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

Karakterpolynomiet er

$$\begin{aligned} \begin{vmatrix} 3-\lambda & -2 & 0 \\ -2 & 3-\lambda & 0 \\ 0 & 0 & 5-\lambda \end{vmatrix} &= (5-\lambda)\begin{vmatrix} 3-\lambda & -2 \\ -2 & 3-\lambda \end{vmatrix} \\ &= -(\lambda-1)(\lambda-5)^2 \end{aligned}$$

Egenværdierne er altså 1 *med algebraisk multiplicitet 1 (en* simpel *rod) og* 5 *med algebraisk multiplicitet 2 (en dobbeltrod). Vi finder egenvektorer.*

1. *Egenvektorerne hørende til egenværdien 1: Ligningssystemet* $(A - E)v = 0$ *har udseendet*

$$\begin{pmatrix} 2 & -2 & 0 \\ -2 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

   og giver ved Gausselimination

$$\begin{pmatrix} 2 & -2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 4 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

   altså $v_1 - v_2 = 0$ *og* $v_3 = 0$. *Samtlige egenvektorer er dermed givet ved*

$$v = t\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad t \in R$$

   Egenrummet $N(A - E)$ *har åbenbart dimension 1, så den geometriske multiplicitet af egenværdien 1 er 1.*

2. *Egenvektorerne hørende til egenværdien 5: Ligningssystemet* $(A - 5E)v = 0$ *har udseendet*

$$\begin{pmatrix} -2 & -2 & 0 \\ -2 & -2 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

og giver ved Gausselimination

$$\begin{pmatrix} -2 & -2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

altså har vi kun en ligning, nemlig $v_1 + v_2 = 0$*. Samtlige egenvektorer er dermed givet ved*

$$v = \begin{pmatrix} t \\ -t \\ s \end{pmatrix} = s \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + t \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad s, t \in R$$

Egenrummet $N(A - 5E)$ *har åbenbart dimension 2, så den geometriske multiplicitet af egenværdien 5 er 2.*

Definition 134 *En egenværdi med samme algebraiske og geometriske multiplicitet kaldes semisimpel. Hvis* A *er en reel kvadratisk matrix og hvis samtlige rødder i karakterpolynomiet* $\det(A - \lambda E)$ *er reelle, så siges* A *at være (reelt) semisimpel, hvis alle dens egenværdier er enten simple eller semisimple.*

Eksempel 135 *Begge egenværdierne i eksemplet ovenfor er semisimple. Egenværdien 1 er også simpel. Matricen i eksemplet er derfor semisimpel.*

Eksempel 136 *Lad matricen* A *være givet ved*

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

Karakterpolynomiet er

$$\begin{vmatrix} -\lambda & 0 \\ 1 & \lambda \end{vmatrix} = \lambda^2$$

0 er altså dobbeltrod. Dvs. at 0 er egenværdi med algebraisk multiplicitet 2. Egenvektorerne findes ved løsning af $(A - 0E)v = Av = 0$*, der reducerer til* $v_1 = 0$*. Altså er samtlige egenvektorer givet ved*

$$v = t \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad t \in R$$

Egenrummet $N(A)$ *er altså éndimensionalt, så den geometriske multiplicitet af egenværdien 0 er 1 og dermed mindre end den algebraiske multiplicitet. Matricen er ikke semisimpel.*

Sætning 137 *Den geometriske multiplicitet for en egenværdi er altid mindre end eller lig med egenværdiens algebraiske multiplicitet.*

Eksempel 138 *Lad* A *være en kvadratisk matrix. Så har* A *og* A^t *de samme egenværdier, men ikke nødvendigvis de samme egenvektorer. Vi har jo, at* $\det(A^t - \lambda E) = \det\left((A - \lambda E)^t\right) = \det(A - \lambda E)$*, så* A *og* A^t *har samme karakterpolynomium. At egenvektorerne ikke nødvendigvis er de samme, kan ses ved at betragte en vilkårlig konkret ikke-symmetrisk matrix* A*.*

7.3 Diagonalisation

Sætning 139 *To similære matricer har samme karakterpolynomium og dermed de samme egenværdier med samme algebraiske multipliciteter. Men matricernes egenværdier har også de samme geometriske multipliciteter.*

Bevis. Lad A og B være similære med $B = S^{-1}AS$. Så gælder

$$B - \lambda E = S^{-1}AS - \lambda E = S^{-1}(A - \lambda E)S$$

og dermed

$$\begin{aligned} \det(B - \lambda E) &= \det\left(S^{-1}(A - \lambda E)S\right) \\ &= \det\left(S^{-1}\right)\det(A - \lambda E)\det(S) \\ &= \det(A - \lambda E) \end{aligned}$$

Hermed har vi vist den første påstand. Lad nu $v \neq 0$ være en egenvektor for A hørende til egenværdien λ. Så er $u = S^{-1}v$ en egenvektor ($\neq 0$) for B hørende til egenværdien λ, idet vi har

$$Bu = BS^{-1}v = \left(S^{-1}AS\right)S^{-1}v = S^{-1}Av = S^{-1}\lambda v = \lambda S^{-1}v = \lambda u$$

Er en basis for egenrummet $N(A - \lambda E)$ givet ved $\{v_1, v_2, \ldots, v_k\}$, så er $\left\{S^{-1}v_1, S^{-1}v_2, \ldots, S^{-1}v_k\right\}$ en basis for egenrummet $N(B - \lambda E)$. For at vise dette skal vi først vise, at $\left\{S^{-1}v_1, S^{-1}v_2, \ldots, S^{-1}v_k\right\}$ er et lineært uafhængigt sæt af vektorer. Antag derfor, at

$$c_1S^{-1}v_1 + c_2S^{-1}v_2 + \ldots + c_kS^{-1}v_k = 0$$

Vi skal vise, at $c_1 = c_2 = \ldots = c_k = 0$. Ved multiplikation fra venstre med S fås

$$c_1v_1 + c_2v_2 + \ldots + c_kv_k = 0$$

men vektorerne $\{v_1, v_2, \ldots, v_k\}$ er lineært uafhængige, derfor følger, at $c_1 = c_2 = \ldots = c_k = 0$. Vi har nu vist, at

$$\dim N(A - \lambda E) \leq \dim N(B - \lambda E)$$

Men da det lige viste også gælder med A og B ombyttede, gælder ligeledes, at $\dim N(B - \lambda E) \leq \dim N(A - \lambda E)$. Altså gælder lighed. Men dermed er $\left\{S^{-1}v_1, S^{-1}v_2, \ldots, S^{-1}v_k\right\}$ en basis for $N(B - \lambda E)$. ■

Bemærkning 140 *Den omvendte sætning gælder ikke, som det ses i det følgende eksempel.*

Eksempel 141 *Lad matricerne A og B være givet ved*

$$\begin{aligned} A &= \begin{pmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_1 & 1 & 0 \\ 0 & 0 & \lambda_1 & 1 \\ 0 & 0 & 0 & \lambda_1 \end{pmatrix} \\ B &= \begin{pmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_1 & 1 \\ 0 & 0 & 0 & \lambda_1 \end{pmatrix} \end{aligned}$$

Karakterpolynomiet er for begge matricer $(\lambda - \lambda_1)^4$. *For begge matricer gælder altså, at tallet* λ_1 *er eneste egenværdi. Algebraisk multiplicitet for* λ_1 *er åbenbart 4. Egenvektorerne for* A *findes til*

$$v_A = s \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad s, t \in R$$

og for B *til*

$$v_B = s \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad s, t \in R$$

Heraf følger, at den geometriske multiplicitet for egenværdien λ_1 *er 2 for begge matricer. Imidlertid skal vi nu vise, at* A *og* B *ikke er similære. Antag nemlig, at* $B = S^{-1}AS$, *så har vi også (ved regninger som vi har udført ofte ovenfor), at*

$$(B - \lambda_1 E)^2 = \left(S^{-1}AS - \lambda_1 E\right)^2 = S^{-1} (A - \lambda_1 E)^2 S$$

Ved konkret udregning findes $(B - \lambda_1 E)^2 = 0$ *og dermed skulle* $(A - \lambda_1 E)^2 = S (B - \lambda_1 E)^2 S^{-1} = 0$. *Men vi finder ligeledes ved konkret regning*

$$(A - \lambda_1 E)^2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Bemærkning 142 *Vi skal senere se, at hvis to* semisimple *matricer har samme egenværdier med de samme multipliciteter, så er de similære. Da den algebraiske og den geometriske multiplicitet for hver matricerne ovenfor er forskellige (de er nemlig henholdsvis 4 og 2), er ingen af de to matricer semisimple.*

Sætning 143 *Den kvadratiske* $n \times n$ *matrix* A *kan diagonaliseres, hvis og kun hvis* A *har* n *lineært uafhængige egenvektorer* $v_1, v_2, \ldots, v_n$. *I bekræftende fald er en diagonaliserende matrix* S *givet ved*

$$S = (v_1, v_2, \ldots, v_n)$$

hvor det understreges, at vektorerne $v_1, v_2, \ldots, v_n$ *skal være* søjlerne *i* S.

Bevis. Antag først, at der findes en invertibel matrix S, så $\Lambda = S^{-1}AS$ er en diagonalmatrix, $\Lambda = diag(\lambda_1, \lambda_2, \ldots, \lambda_n)$. Vi har, at $AS = S\Lambda$. Udtryk S ved sine søjler $S = (v_1, v_2, \ldots, v_n)$. Vi har da, at AS udtrykt ved sine søjler er givet ved

$$AS = (Av_1, Av_2, \ldots, Av_n)$$

og $S\Lambda$ udtrykt ved sine søjler er givet ved

$$S\Lambda = (\lambda_1 v_1, \lambda_2 v_2, \ldots, \lambda_n v_n)$$

Af $AS = S\Lambda$ følger derfor, at

$$Av_1 = \lambda_1 v_1, Av_2 = \lambda_2 v_2, \ldots, Av_n = \lambda_n v_n$$

med andre ord, den i'te søjle i S er egenvektor for A hørende til egenværdien λ_i, der er det i'te element i diagonalen for Λ. At egenvektorerne $v_1, v_2, \ldots, v_n$ er lineært uafhængige følger af, at S er invertibel. Antag nu omvendt, at A har n lineært uafhængige egenvektorer $v_1, v_2, \ldots, v_n$ hørende til egenværdierne $\lambda_1, \lambda_2, \ldots, \lambda_n$. Lad $\Lambda = diag\,(\lambda_1, \lambda_2, \ldots, \lambda_n)$ og $S = (v_1, v_2, \ldots, v_n)$. Så er S invertibel, og ved samme regninger som ovenfor fås, at $AS = S\Lambda$, altså også $\Lambda = S^{-1}AS$. ■

Eksempel 144 *For matricen*

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

fandt vi tidligere egenvektoren $(1, -1)^t$ *hørende til egenværdien 4 og* $(2, 1)^t$ *hørende til egenværdien 1. Da disse to egenvektorer åbenbart er lineært uafhængige, konkluderer vi, at* A *er diagonaliserbar med en diagonaliserende matrix givet ved*

$$S = \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix}$$

Der gælder nu, at $\Lambda = S^{-1}AS$ *er diagonalmatricen* $\Lambda = diag\,(4, 1)$.

Eksempel 145 *For matricen* A *givet ved*

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

fandt vi tidligere egenvektoren $v_1 = (1, 1, 0)^t$ *hørende til egenværdien 1 og egenvektorerne* $v_2 = (1, -1, 0)^t$ *og* $v_3 = (0, 0, 1)^t$ *hørende til egenværdien 5. Disse 3 egenvektorer er lineært uafhængige, så* A *er diagonaliserbar med en diagonaliserende matrix givet ved*

$$S = \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Vi har da uden yderligere regning, at $\Lambda = S^{-1}AS = diag\,(1, 5, 5)$.

Eksempel 146 *Vi så tidligere, at matricen*

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

kun havde egenvektorerne $s\,(0, 1)^t, \quad s \in R$. A *kan altså ikke diagonaliseres.*

Eksempel 147 *Ingen af de to 4×4 matricer A og B nævnt i eksemplet om similaritet og algebraisk og geometrisk multiplicitet er diagonaliserbare.*

Sætning 148 *Lad A være en $n \times n$ matrix. Hvis A har n indbyrdes forskellige egenværdier, så er A diagonaliserbar.*

Påstanden følger umiddelbart af følgende lemma.

Lemma 149 *Lad $\lambda_1, \lambda_2, \ldots, \lambda_k$ være indbyrdes forskellige egenværdier for matricen A. Lad $v_1, v_2, \ldots, v_k$ være tilhørende egenvektorer, alle $\neq 0$. Så er $v_1, v_2, \ldots, v_k$ lineært uafhængige.*

Bevis. Vi viser påstanden ved induktion i antallet k. Påstanden er åbenbart sand for $k = 1$. Antag, at påstanden er sand for et antal $k \geq 1$. Vi vil vise, at påstanden også er sand, når antallet er $k + 1$. Antag derfor, at $\lambda_1, \lambda_2, \ldots, \lambda_{k+1}$ er indbyrdes forskellige egenværdier for matricen A. Lad $v_1, v_2, \ldots, v_{k+1}$ være tilhørende egenvektorer, alle $\neq 0$. Antag, at

$$c_1 v_1 + c_2 v_2 + \ldots + c_k v_k + c_{k+1} v_{k+1} = 0$$

Vi skal vise, at $c_1, c_2, \ldots, c_k, c_{k+1} = 0$. Gang ligningen med A fra venstre. Hermed fås

$$c_1 \lambda_1 v_1 + c_2 \lambda_2 v_2 + \ldots + c_k \lambda_k v_k + c_{k+1} \lambda_{k+1} v_{k+1} = 0$$

Fratræk $\lambda_{k+1} (c_1 v_1 + c_2 v_2 + \ldots + c_k v_k + c_{k+1} v_{k+1}) = 0$ fra denne ligning. Hermed fås

$$c_1 (\lambda_1 - \lambda_{k+1}) v_1 + c_2 (\lambda_2 - \lambda_{k+1}) v_2 + \ldots + c_k (\lambda_k - \lambda_{k+1}) v_k = 0$$

Ifølge induktionsantagelsen er $v_1, v_2, \ldots, v_k$ lineært uafhængige. Derfor gælder, at

$$c_1 (\lambda_1 - \lambda_{k+1}) = c_2 (\lambda_2 - \lambda_{k+1}) = \ldots = c_k (\lambda_k - \lambda_{k+1}) = 0$$

Men da λ_i'erne er forskellige, fås deraf, at $c_1 = c_2 = \ldots = c_k = 0$. Indsættes disse værdier i ligningen $c_1 v_1 + c_2 v_2 + \ldots + c_k v_k + c_{k+1} v_{k+1} = 0$ følger det, at også $c_{k+1} = 0$. ∎

Dette lemma og dets bevis kan let (men omstændeligt) generaliseres til følgende.

Lemma 150 *Lad $\lambda_1, \lambda_2, \ldots, \lambda_k$ være indbyrdes forskellige egenværdier for matricen A. Vælg for ethvert λ_i en basis for egenrummet $N(A - \lambda_i)$. Så gælder, at foreningsmængden af disse baser er en lineært uafhængig mængde af vektorer.*

Sætning 151 *Hvis matricen A er (reelt) semisimpel, så er A diagonaliserbar (og omvendt).*

Bevis. Når $n \times n$ matricen A er semisimpel gælder, at for samtlige dens egenværdier gælder, at algebraisk og geometrisk multiplicitet stemmer overens. Summerer vi samtlige algebraiske multipliciteter, får vi graden af karakterpolynomiet, altså n. Ifølge lemmaet ovenfor findes der et lineært uafhængigt sæt af egenvektorer for A med antal lig med summen af samtlige geometriske multipliciteter, altså også n. Resultatet følger nu af en tidligere sætning. ■

Bemærkning 152 *I en pose har vi samtlige reelle tal. Med tilbagelægning udtager vi n^2 tal og danner heraf en $n \times n$ matrix. Vi nedskriver dens karakterpolynomium og finder dets rødder. Antag, at disse alle er reelle. Hvad er da sandsynligheden for, at to af dem er ens? Svar 0. Hvorfor så overhovedet interessere sig for det tilfælde? Vi kan nævne 3 grunde:*

1. *En matrix med to tætliggende egenværdier opfører sig i nogen grad næsten som en matrix med to ens egenværdier.*
2. *Det kan i visse anvendelser af teorien tænkes, at matricer med flere ens egenværdier spiller en afgørende rolle, altså at matricerne ikke tages tilfældigt op af en pose.*
3. *Hvis der i en matrix A optræder en parameter p, som kan varieres i et interval, så vil egenværdierne også afhænge af p. De værdier af p, for hvilke to af egenværdierne er sammenfaldende, er da af interesse, idet der her kan ske en afgørende ændring i matricens egenskaber. Man taler om, at der sker en bifurkation.*

7.3.1 Symmetriske matricer

Definition 153 *En matrix A kaldes symmetrisk, hvis $A^t = A$.*

Definition 154 *En matrix S kaldes ortogonal, hvis den opfylder $S^tS = SS^t = E$.*

Sætning 155 *Lad A være en reel symmetrisk matrix. Så gælder:*

1. *Samtlige løsninger til karakterligningen er reelle (”samtlige egenværdier er reelle”).*
2. *Egenvektorer hørende til forskellige egenværdier er ortogonale.*
3. *A er (reelt) semisimpel.*
4. *A er diagonaliserbar. Der kan altid som diagonaliserende matrix S bestemmes en ortonal matrix.*

Bevis. Selv om A er en reel matrix, kan vi betragte den som en kompleks matrix. Lad $\lambda \in C$ være en egenværdi for A, og lad $v \in C^n$ vær en tilhørende egenvektor med $v \neq 0$. Længden (2-normen) af en vektor i C^n er kvadratroden af størrelsen

$$\overline{v}^t v = \sum_{i=1}^{n} \overline{v_i} v_i = \sum_{i=1}^{n} |v_i|^2 = \|v\|^2$$

Så har vi af $Av = \lambda v$, at også $A\overline{v} = \overline{Av} = \overline{\lambda v} = \overline{\lambda} \cdot \overline{v}$. Dermed følger dels

$$(A\overline{v})^t v = \overline{\lambda} \cdot (\overline{v})^t v = \overline{\lambda} \|v\|^2$$

og dels

$$(A\overline{v})^t v = (\overline{v})^t A^t v = (\overline{v})^t Av = \lambda (\overline{v})^t v = \lambda \|v\|^2$$

Da nu $\|v\|^2 \neq 0$ fås, at $\overline{\lambda} = \lambda$, altså at $\lambda \in R$.

Antag, at $Av_1 = \lambda_1 v_1$ og $Av_2 = \lambda_2 v_2$, hvor $\lambda_1 \neq \lambda_2$. Så har vi

$$\begin{aligned} \lambda_2 v_1 \cdot v_2 &= \lambda_2 v_1^t v_2 = v_1^t (\lambda_2 v_2) = v_1^t A v_2 \\ &= (Av_1)^t v_2 = \lambda_1 v_1^t v_2 = \lambda_1 v_1 \cdot v_2 \end{aligned}$$

Altså har vi, at

$$(\lambda_2 - \lambda_1) v_1 \cdot v_2 = 0$$

hvoraf følger, at skalarproduktet $v_1 \cdot v_2 = 0$. Altså $v_1 \perp v_2$.

Det sværeste er at vise, at A er semisimpel. Opskriv egenværdierne for A med gentagelser efter *geometrisk* multiplicitet, eksempelvis skal en egenværdi med geometrisk multiplicitet 2 altså forekomme to gange: $\lambda_1, \lambda_2, \ldots, \lambda_k$. Lad $v_1, v_2, \ldots, v_k$ være tilhørende ikke-trivielle egenvektorer. Hvis en egenværdi har geometrisk multiplicitet p, kan vi tænke os de p egenvektorer valgt indbyrdes ortogonale. Hermed er $v_1, v_2, \ldots, v_k$ et ortogonalt system. Vi skal vise, at $k = n$. Antag, at dette ikke var tilfældet. Så har vi altså $k < n$. Vi supplerer ortogonalsystemet op til en ortogonal basis $\{v_1, v_2, \ldots, v_k, v_{k+1}, \ldots, v_n\}$. Vi kan antage, at det faktisk er en ortonormal basis. Lad S være den matrix, hvis søjler består af basisvektorerne, altså $S = (v_1, v_2, \ldots, v_k, v_{k+1}, \ldots, v_n)$. Så er S en ortogonal matrix, og vi har

$$\begin{aligned} Q &= S^t AS = S^t (Av_1, Av_2, \ldots, Av_k, Av_{k+1}, \ldots, Av_n) \\ &= S^t (\lambda_1 v_1, \lambda_2 v_2, \ldots, \lambda_k v_k, Av_{k+1}, \ldots, Av_n) \\ &= \left(\lambda_1 S^t v_1, \lambda_2 S^t v_2, \ldots, \lambda_k S^t v_k, S^t Av_{k+1}, \ldots, S^t Av_n\right) \end{aligned}$$

men $S^t v_1 = (1, 0, 0, \ldots, 0)^t, S^t v_2 = (0, 1, 0, \ldots, 0), \ldots, S^t v_k = (0, 0, \ldots, 0, 1, 0, \ldots, 0)$. Desuden gælder, at

$$\begin{aligned} S^t Av_{k+i} &= (AS)^t v_{k+i} = (Av_1, Av_2, \ldots, Av_k, Av_{k+1}, \ldots, Av_n)^t v_{k+i} \\ &= (\lambda_1 v_1, \lambda_2 v_2, \ldots, \lambda_k v_k, Av_{k+1}, \ldots, Av_n)^t v_{k+i} \\ &= \left(0, 0, \ldots, 0, (Av_{k+1})^t v_{k+i}, \ldots, (Av_n)^t v_{k+i}\right) \end{aligned}$$

for $i > 0$. Hermed er vist, at $Q = S^t AS$ har formen

$$Q = S^t AS = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \lambda_k & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & b_{11} & \dots & b_{1,n-k} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & b_{n-k,1} & \dots & b_{n-k,n-k} \end{pmatrix}$$

med andre ord Q har formen $Q = diag\,(\lambda_1, \dots, \lambda_k, B)$, hvor $B = (b_{ij})$ er en $(n-k) \times (n-k)$ matrix. Denne matrix B er symmetrisk, da Q er symmetrisk. Derfor har B en egenværdi $\lambda_0 \in R$ og tilhørende ikke-triviel egenvektor $u \in R^{n-k}$. Lad $v = (0, 0, \dots, 0, u^t)^t$. Så gælder, at $Qv = \lambda_0 v$, og dermed $ASv = \lambda_0 Sv$. Altså er Sv en egenvektor for A hørende til egenværdien λ_0. Men nu er $\{v_1, v_2, \dots, v_k, Sv\}$ lineært uafhængigt, da mængden $\{S^t v_1, S^t v_2, \dots, S^t v_k, v\} = \{e_1, e_2, \dots, e_k, v\}$ klart er lineært uafhængig (her er $e_1, e_2, \dots, e_k$ de sædvanlige enhedsvektorer). Hermed har vi en modstrid. Vi må altså konkludere, at $k = n$.

Sætningens sidste påstand følger af den tredie.' ■

7.4 Komplekse egenværdier

Hidtil har vi omhyggeligt undgået det noget kildne problem: Antag, at A nok er en reel matrix, men at karakterpolynomiet $\det(A - \lambda E)$ har imaginære rødder. Hvad gør vi så? Dette er selvfølgelig kun et problem, hvis vi ikke kan acceptere, at egenvektorer og dermed en eventuel diagonaliserende matrix har imaginære elementer. Der gælder følgende substitut for en diagonaliseringssætning.

Sætning 156 *Lad A være en reel kvadratisk matrix. Antag, at A som kompleks matrix betragtet er semisimpel. Lad $\lambda_1, \lambda_2, \dots, \lambda_r$ være de reelle egenværdier for A, og lad $\mu_1, \overline{\mu_1}, \mu_2, \overline{\mu_2}, \dots, \mu_k, \overline{\mu_k}$ være de imaginære egenværdier, i begge tilfælde med gentagelser efter multipliciteter. Lad $\mu_j = a_j + ib_j$, hvor $a_j, b_j \in R$. Så findes der en invertibel, reel matrix S, så $S^{-1}AS = \Lambda$, hvor Λ er "næsten" diagonal,*

$$\Lambda = diag\left(\lambda_1, \lambda_2, \dots, \lambda_r, \begin{pmatrix} a_1 & b_1 \\ -b_1 & a_1 \end{pmatrix}, \begin{pmatrix} a_2 & b_2 \\ -b_2 & a_2 \end{pmatrix}, \dots, \begin{pmatrix} a_k & b_k \\ -b_k & a_k \end{pmatrix}\right)$$

Søjlerne i S er egenvektorerne hørende til de reelle egenværdier fulgt af skiftevis real- og imaginærdel af egenvektorerne hørende til de imaginære egenværdier $\mu_1, \mu_2, \dots, \mu_k$.

Bevis. Beviset er ikke svært, men udelades. ■

Eksempel 157 *Lad A være matricen*

$$A = \begin{pmatrix} 3 & -8 & -4 \\ 0 & 3 & 4 \\ 0 & -2 & -1 \end{pmatrix}$$

Karakterpolynomiet er $\det(A - \lambda E) = (\lambda - 3)\left(\lambda^2 - 2\lambda + 5\right)$. *Så indenfor de komplekse tal er egenværdierne* 3 *og* $1 \pm 2i$. *Tilhørende egenvektorer er til egenværdien 1,* $v_1 = (1,0,0)^t$. *Til egenværdien* $1{+}2i$ *findes* $v_2 = \left(1 + 2i, 1, -\frac{1}{2} + \frac{1}{2}i\right)$. *Vi skal ikke bruge en egenvektor hørende til* $1 - 2i$, *men som en sådan kan tages den komplekst konjugerede af* v_2. *Hermed er den i sætningen omtalte matrix* S *givet ved*

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

og vi har

$$\Lambda = S^{-1}AS = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & -2 & 1 \end{pmatrix}$$

8 Lineært differentialligningssystem

Betragt følgende system af n førsteordens differentialligninger med n ubekendte funktioner $x_1, x_2, \ldots, x_n$

$$\begin{aligned} \dot{x}_1(t) &= a_{11}x_1(t) + a_{12}x_2(t) + \ldots + a_{1n}x_n(t) \\ \dot{x}_2(t) &= a_{21}x_1(t) + a_{22}x_2(t) + \ldots + a_{2n}x_n(t) \\ &\vdots \\ \dot{x}_n(t) &= a_{n1}x_1(t) + a_{n2}x_2(t) + \ldots + a_{nn}x_n(t) \end{aligned}$$

Et system, der kan skrives på denne form, kaldes lineært. Koefficienterne behøver ikke være konstante. Er de imidlertid konstante, så er løsningen af systemet i høj grad "kun" en øvelse i lineær algebra. Systemet kan skrives på kompakt matrixform således

$$\dot{x}(t) = Ax(t)$$

når A er koefficientmatricen $A = (a_{ij})$ og $x(t) = (x_1(t), x_2(t), \ldots, x_n(t))^t$.

Er A reelt semisimpel (dvs. diagonaliserbar indenfor de reelle tal), så er løsningen simpel. Der gælder nemlig følgende sætning.

Sætning 158 *Lad* A *være en reelt semisimpel* $n \times n$ *matrix. Så er den fuldstændige løsning til differentialligningssystemet* $\dot{x}(t) = Ax(t)$ *givet ved*

$$x(t) = c_1 e^{\lambda_1 t} v_1 + c_2 e^{\lambda_2 t} v_2 + \ldots + c_n e^{\lambda_n t} v_n$$

$c_1, c_2, \ldots, c_n \in R$, *når* $\lambda_1, \lambda_2, \ldots, \lambda_n$ *er egenværdierne for* A, *og når* $v_1, v_2, \ldots, v_n$ *er tilhørende (lineært uafhængige) egenvektorer. Løsningen kan også angives på formen*

$$x(t) = \Phi(t)\, c$$

hvor $c = (c_1, c_2, \ldots, c_n)^t$ *og* $\Phi(t) = \left(e^{\lambda_1 t} v_1, e^{\lambda_2 t} v_2, \ldots, e^{\lambda_n t} v_n\right)$. *En sådan matrix* Φ *kaldes en fundamental matrix for differentialligningssystemet. Udtrykt ved begyndelserværdierne* $x(0)$, *kan løsningen skrives*

$$x(t) = \Phi(t)\,\Phi(0)^{-1}\,x(0)$$

Bevis. A er diagonaliserbar. En diagonaliserende matrix er $S = (v_1, v_2, \ldots, v_n)$. Vi har $A = S\Lambda S^{-1}$, hvor $\Lambda = diag(\lambda_1, \lambda_2, \ldots, \lambda_n)$. Hermed har vi

$$\dot{x}(t) = Ax(t) = S\Lambda S^{-1}x(t)$$

Med $y(t) = S^{-1}x(t)$ finder vi dermed, at

$$\dot{y}(t) = \frac{d}{dt}\left(S^{-1}x(t)\right) = S^{-1}\dot{x}(t) = \Lambda S^{-1}x(t) = \Lambda y(t)$$

Men systemet $\dot{y}(t) = \Lambda y(t)$ er ekstremt simpelt, nemlig blot

$$\begin{aligned}
\dot{y}_1(t) &= \lambda_1 y_1(t) \\
\dot{y}_2(t) &= \lambda_2 y_2(t) \\
&\vdots \\
\dot{y}_n(t) &= \lambda_n y_n(t)
\end{aligned}$$

Dette system er jo blot n differentialligninger, der intet har med hinanden at gøre. Der er ingen "kobling" mellem ligningerne. Løsningerne er $y_i(t) = c_i e^{\lambda_i t}$, $i = 1 \ldots n$. Men hermed følger

$$\begin{aligned}
x(t) &= Sy(t) = (v_1, v_2, \ldots, v_n)\,y(t) \\
&= (v_1, v_2, \ldots, v_n)\begin{pmatrix} c_1 e^{\lambda_1 t} \\ c_2 e^{\lambda_2 t} \\ \vdots \\ c_n e^{\lambda_{ni} t} \end{pmatrix} \\
&= c_1 e^{\lambda_1 t} v_1 + c_2 e^{\lambda_2 t} v_2 + \ldots + c_n e^{\lambda_n t} v_n \\
&= \left(e^{\lambda_1 t} v_1, e^{\lambda_2 t} v_2, \ldots, e^{\lambda_n t} v_n\right) c = \Phi(t)\,c
\end{aligned}$$

Er $x(0)$ givet, fås af $x(0) = \Phi(0)\,c$, at $c = \Phi(0)^{-1}\,x(0)$, hvoraf $x(t) = \Phi(t)\,\Phi(0)^{-1}\,x(0)$. ■

Bemærkning 159 *Bemærk udfra beviset, at* $\Phi(0) = S$, *den brugte diagonaliserende matrix.*

Eksempel 160 *Betragt differentialligningssystemet*

$$\begin{aligned}
\dot{x}_1(t) &= 2x_1(t) + 10x_2(t) - 20x_n(t) \\
\dot{x}_2(t) &= 16x_1(t) - x_2(t) + 2x_n(t) \\
\dot{x}_n(t) &= -17x_1(t) - 4x_2(t) + 8x_n(t)
\end{aligned}$$

Koefficientmatricen er

$$A = \begin{pmatrix} 2 & 10 & -20 \\ 16 & -1 & 2 \\ -17 & -4 & 8 \end{pmatrix}$$

Karakterpolynomiet er $\det(A - \lambda E) = \lambda^3 - 9\lambda^2 - 486\lambda$. *Egenværdierne findes til* $0, -18$ *og* 27. *Tilhørende egenvektorer er i samme rækkefølge,* $(0,2,1)^t, (2,-2,1)^t$ *og* $(2,1,-2)^t$. *Den fuldstændige løsning er dermed*

$$x(t) = c_1 \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix} + c_2 e^{-18t} \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix} + c_3 e^{27t} \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix}$$

hvor $c_1, c_2, c_3 \in R$. *Skrevet ud har den fuldstændige løsning formen*

$$\begin{aligned} x_1(t) &= 2c_2 e^{-18t} + 2c_3 e^{27t} \\ x_2(t) &= 2c_1 - 2c_2 e^{-18t} + c_3 e^{27t} \\ x_3(t) &= c_1 + c_2 e^{-18t} - 2c_3 e^{27t} \end{aligned}$$

hvor $c_1, c_2, c_3 \in R$. *En fundamental matrix er*

$$\Phi(t) = \begin{pmatrix} 0 & 2e^{-18t} & 2e^{27t} \\ 2 & -2e^{-18t} & e^{27t} \\ 1 & e^{-18t} & -2e^{27t} \end{pmatrix}$$

Den løsning, der opfylder begyndelsesbetingelsen $x(0) = (1,2,3)^t$ *er givet ved*

$$x(t) = \Phi(t)\,\Phi(0)^{-1} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

Men den bestemmes sikkert lettere direkte ved indsættelse i løsningen på udskrevet form, dvs. af ligningerne

$$\begin{aligned} 1 &= x_1(0) = 2c_2 + 2c_3 \\ 2 &= x_2(0) = 2c_1 - 2c_2 + c_3 \\ 3 &= x_3(0) = c_1 + c_2 - 2c_3 \end{aligned}$$

Dette system løses ved Gausselimination. Løsningen er $c = \left(\frac{25}{16}, \frac{13}{16}, -\frac{5}{16}\right)^t$. *Hermed finder vi*

$$\begin{aligned} x_1(t) &= \frac{13}{8} e^{-18t} - \frac{5}{8} e^{27t} \\ x_2(t) &= \frac{25}{8} - \frac{13}{8} e^{-18t} - \frac{5}{16} e^{27t} \\ x_3(t) &= \frac{25}{16} + \frac{13}{16} e^{-18t} + \frac{5}{8} e^{27t} \end{aligned}$$

Lemma 161 *Systemet*

$$\dot{x}(t) = \begin{pmatrix} a & b \\ -b & a \end{pmatrix} x(t)$$

hvor $a, b \in R$, *har den fuldstændige løsning*

$$x(t) = e^{at} \begin{pmatrix} \cos bt & \sin bt \\ -\sin bt & \cos bt \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$$

hvor $c = (c_1, c_2)^t \in R^2$.

Bevis. Matricen

$$A = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$$

har egenværdierne $a \pm ib$. Svarende til egenværdien $a + ib$ har vi egenvektoren $(1, i)^t$. Derfor er $e^{(a+ib)t}(1, i)^t$ løsning. Men da A er reel, er dermed denne løsning realdel og dens imaginærdel også løsninger. Disse er givet ved

$$\begin{aligned} \operatorname{Re}\left(e^{(a+ib)t} \begin{pmatrix} 1 \\ i \end{pmatrix}\right) &= e^{at} \begin{pmatrix} \cos bt \\ -\sin bt \end{pmatrix} \\ \operatorname{Im}\left(e^{(a+ib)t} \begin{pmatrix} 1 \\ i \end{pmatrix}\right) &= e^{at} \begin{pmatrix} \sin bt \\ \cos bt \end{pmatrix} \end{aligned}$$

Den fuldstændige løsning er nu (som det kan vises) givet som samtlige reelle linearkombinationer af disse, dvs.

$$\begin{aligned} x(t) &= c_1 e^{at} \begin{pmatrix} \cos bt \\ -\sin bt \end{pmatrix} + c_2 e^{at} \begin{pmatrix} \sin bt \\ \cos bt \end{pmatrix} \\ &= e^{at} \begin{pmatrix} \cos bt & \sin bt \\ -\sin bt & \cos bt \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \end{aligned}$$

hvor $c = (c_1, c_2)^t \in R^2$. ■

Er den reelle koefficientmatrix i systemet $\dot{x}(t) = Ax(t)$ ikke reelt semisimpel, men dog komplekst semisimpel, så kan løsningen af systemet i princippet udføres som i det reelle tilfælde. Er man interesseret i at løse begyndelsesværdiproblemet med $x(0) \in R^n$ givet, så vil den fundne løsning automatisk være reel, selv om den måske ikke ser sådan ud. Vil man derimod have den fuldstændige reelle løsning udtrykt ved reelle funktioner, kan man ud af bidraget $e^{\lambda t}v$ bestående af en kompleks eksponentialfunktion og en kompleks egenvektor at tage realdel og imaginærdel ganske som gjort i beviset for lemmaet ovenfor. Disse to indgår så i stedet for bidragene $e^{\lambda t}v$ og $e^{\overline{\lambda} t}\overline{v}$ i udtrykket for den fuldstændige løsning. Alternativt kan sætningen om næsten-diagonalisering af en komplekst semisimpel reel matrix bruges i forbindelse med lemmaet. Vi giver et eksempel.

Eksempel 162 *Betragt differentialligningssystemet* $\dot{x}(t) = Ax(t)$, *hvor koefficientmatricen* A *er givet ved*

$$A = \begin{pmatrix} 3 & -8 & -4 \\ 0 & 3 & 4 \\ 0 & -2 & -1 \end{pmatrix}$$

For denne matrix fandtes tidligere, at egenværdierne er $3, 1 \pm 3i$. *Desuden fandtes, at med*

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

gjaldt

$$\Lambda = S^{-1}AS = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & -2 & 1 \end{pmatrix}$$

Med $y(t) = S^{-1}x(t)$ *opfylder* y *ligningssystemet* $\dot{y}(t) = \Lambda y(t)$. *I dette system er* y_1 *afkoblet fra* y_2 *og* y_3, *mens disse to er koblet som i lemmaet. Hermed kan den fuldstændige løsning for* $\dot{y}(t) = \Lambda y(t)$ *skrives*

$$y(t) = \begin{pmatrix} e^{3t} & 0 & 0 \\ 0 & e^t \cos 2t & e^t \sin 2t \\ 0 & -e^t \sin 2t & e^t \cos 2t \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}$$

Den fuldstændige løsning til det oprindelige system er dermed $x(t) = Sy(t)$, *dvs.*

$$x(t) = \begin{pmatrix} e^{3t} & e^t \cos 2t - 2e^t \sin 2t & e^t \sin 2t + 2e^t \cos 2t \\ 0 & e^t \cos 2t & e^t \sin 2t \\ 0 & -\frac{1}{2}e^t \cos 2t - \frac{1}{2}e^t \sin 2t & -\frac{1}{2}e^t \sin 2t + \frac{1}{2}e^t \cos 2t \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}$$

hvor $c_1, c_2, c_3 \in R$.