

7 - SINGULÆRVERDIDEKOMPOSISJON - LF

1 Vi regner ut

$$A^T A = \begin{pmatrix} \|x\|^2 & n\bar{x} \\ n\bar{x} & n \end{pmatrix} \quad \text{og} \quad A^T y = \begin{pmatrix} x^T y \\ n\bar{y} \end{pmatrix}.$$

Den andre raden i $A^T A v = A^T y$ sier

$$n\bar{x}\beta_1 + n\beta_0 = n\bar{y},$$

altså $\beta_0 = \bar{y} - \beta_1 \bar{x}$ som på første side av økten. Den første raden sier

$$\|x\|^2 \beta_1 + n\bar{x}\beta_0 = x^T y,$$

og setter vi inn uttrykket for β_0 , kommer uttrykket for β_1 fra første side ut.

2 Kravet om at parabolen skal reise gjennom datapunktene gir likningssystemet $Av = y$ med

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 4 & 2 & 1 \\ 9 & 3 & 1 \\ 16 & 4 & 1 \\ 25 & 5 & 1 \end{pmatrix} \quad v = \begin{pmatrix} \beta_2 \\ \beta_1 \\ \beta_0 \end{pmatrix} \quad y = \begin{pmatrix} 2 \\ 3 \\ 4 \\ 5 \\ 1 \end{pmatrix}.$$

Normallikningene $A^T A v = A^T y$ blir

$$\begin{pmatrix} 979 & 225 & 55 \\ 225 & 55 & 15 \\ 55 & 15 & 5 \end{pmatrix} v = \begin{pmatrix} 155 \\ 45 \\ 15 \end{pmatrix}$$

med løsning $\beta_2 = -5/7$, $\beta_1 = 30/7$ og $\beta_0 = -2$:

$$y = -\frac{5}{7}x^2 + \frac{30}{7}x - 2.$$

3 Likningen $\dot{h} = ah$ har løsning

$$h(t) = Ce^{at}.$$

Tar vi logaritmen på begge sider, blir

$$\ln h = \ln C + at,$$

som er lineær i a og $\ln C$ - altså vanlig lineær regresjon med t som forklaringsvariabel og $\ln h$ som responsvariabel. Regresjonen gir

$$a \approx -0.035 \text{ s}^{-1} \quad \text{og} \quad C \approx 53 \text{ cm}.$$

Integrasjonskonstanten er starthøyden, $C = h(0)$. Regresjonen sier 53 cm og målingen sa 48 - allerede her aner vi at eksponentialmodellen ikke passer så godt.

- 4 Torricellis lov sier at $\sqrt{h}$ er lineær i t :

$$\sqrt{h} = \sqrt{h_0} \left(1 - \frac{t}{T}\right).$$

Lineær regresjon på t og $\sqrt{h}$ gir $\sqrt{h_0} \approx 6.95$, altså $h_0 \approx 48.3$ cm - nesten nøyaktig målingen - og tømmetid $T \approx 85$ s. Med $V_0 = \pi r^2 h_0$ og $r = 1$ m gir formelen for T

$$A = \frac{V_0}{T} \sqrt{\frac{2}{gh_0}} \approx 0.012 \text{ m}^2,$$

altså et tappehull med radius om lag 6 cm.

- 5 Andreordens regresjon som i oppgave 2 gir

$$h \approx \frac{9}{1400}t^2 - \frac{157}{140}t + \frac{337}{7} \approx 0.0064t^2 - 1.12t + 48.1,$$

mens Torricelli-parabelen fra forrige oppgave er

$$h_0 \left(1 - \frac{t}{T}\right)^2 \approx 0.0067t^2 - 1.13t + 48.3.$$

De to er så godt som identiske, og bunnpunktet til regresjonsparabelen ligger ved $t \approx 87$ s - omtrent tømmetiden T . Fysikkboken vinner.

- 6 Newtons avkjølingslov gir

$$T(t) = 22 + Ce^{-\alpha t},$$

så $\ln(T - 22) = \ln C - \alpha t$, og vi kjører lineær regresjon som i oppgave 3. Den gir $\alpha \approx 0.014 \text{ min}^{-1}$. Minimerer man kvadratsummen direkte på temperaturene istedenfor på logaritmene - et ikkelineært problem - får man $\alpha \approx 0.018 \text{ min}^{-1}$. Forskjellen skyldes at logaritmen vekter de små differansene på slutten av måleserien hardt. Det er altså ikke helt opplagt hva som er "riktig" α , selv med minste kvadraters metode.

- 7 En projeksjon er en lineæroperator som tilfredsstiller $P^2 = P$. Matrisen J med enere overalt oppfyller $J^2 = nJ$, siden hvert element i J^2 er en sum av n enere. Da blir

$$\begin{aligned} C_n^2 &= \left(I - \frac{1}{n}J\right) \left(I - \frac{1}{n}J\right) \\ &= I - \frac{2}{n}J + \frac{1}{n^2}J^2 \\ &= I - \frac{2}{n}J + \frac{1}{n}J \\ &= I - \frac{1}{n}J. \end{aligned}$$

Gjennomsnittet til den sentrerte variabelen er null. Det er det som er poenget. Summen av komponentene er

$$1^T C_n x = 1^T x - \frac{1}{n} 1^T J x = n\bar{x} - \frac{1}{n} \cdot n \cdot n\bar{x} = 0,$$

siden $1^T J = n1^T$.

- 8 Med sentrerte variable der $\bar{x} = \bar{y} = 0$, får vi

$$\beta_1 = \frac{x^T y}{x^T x} \quad \text{og} \quad \beta_0 = 0.$$

Nå kjenner vi igjen β_1 som projeksjonen av y på x .

- 9 Vi har

$$y = \beta_{10}x_1 + \beta_{01}x_2 + \beta_0 1$$

altså

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \beta_{10} \begin{pmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{n1} \end{pmatrix} + \beta_{01} \begin{pmatrix} x_{12} \\ x_{22} \\ \vdots \\ x_{n2} \end{pmatrix} + \beta_0 \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}.$$

La nå

$$A = \begin{pmatrix} x_{11} & x_{12} & 1 \\ x_{21} & x_{22} & 1 \\ \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & 1 \end{pmatrix}.$$

Vi ganger med A^T fra venstre for å få normallikningene:

$$A^T y = A^T A \beta$$

Skrevet ut på komponentform blir dette

$$\begin{pmatrix} y^T x_1 \\ y^T x_2 \\ n\bar{y} \end{pmatrix} = \begin{pmatrix} \|x_1\|^2 & x_1^T x_2 & n\bar{x}_1 \\ x_1^T x_2 & \|x_2\|^2 & n\bar{x}_2 \\ n\bar{x}_1 & n\bar{x}_2 & n \end{pmatrix} \begin{pmatrix} \beta_{10} \\ \beta_{01} \\ \beta_0 \end{pmatrix}$$

Dersom $\overline{x_1} = \overline{x_2} = \overline{y} = 0$, blir dette mye enklere:

$$\begin{pmatrix} y^T x_1 \\ y^T x_2 \\ 0 \end{pmatrix} = \begin{pmatrix} \|x_1\|^2 & x_1^T x_2 & 0 \\ x_1^T x_2 & \|x_2\|^2 & 0 \\ 0 & 0 & n \end{pmatrix} \begin{pmatrix} \beta_{10} \\ \beta_{01} \\ \beta_0 \end{pmatrix}$$

og vi ser lett at $\beta_0 = 0$, slik at likningssystemet blir

$$\begin{pmatrix} y^T x_1 \\ y^T x_2 \end{pmatrix} = \begin{pmatrix} \|x_1\|^2 & x_1^T x_2 \\ x_1^T x_2 & \|x_2\|^2 \end{pmatrix} \begin{pmatrix} \beta_{10} \\ \beta_{01} \end{pmatrix}$$

Et klassisk triks for å kjapt løse 2×2 -matriser, er å kjenne til formelen

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

slik at

$$\begin{aligned} \beta &= (A^T A)^{-1} A^T y \\ &= \frac{1}{\|x_1\|^2 \|x_2\|^2 - (x_1^T x_2)^2} \begin{pmatrix} \|x_2\|^2 & -x_1^T x_2 \\ -x_1^T x_2 & \|x_1\|^2 \end{pmatrix} \begin{pmatrix} y^T x_1 \\ y^T x_2 \end{pmatrix} \end{aligned}$$

I dataanalyse er det også vanlig å dele ut standardavvikene. Dette er det samme som å normalisere vektorer i lineæralgebra. Gjør vi det, blir det enda penere:

$$\beta = \frac{1}{1 - (x_1^T x_2)^2} \begin{pmatrix} 1 & -x_1^T x_2 \\ -x_1^T x_2 & 1 \end{pmatrix} \begin{pmatrix} y^T x_1 \\ y^T x_2 \end{pmatrix}$$

- 10 Empirisk kovarians mellom to datasett x og y ser slik ut

$$\frac{1}{n-1} \sum_k (x_k - \bar{x})(y_k - \bar{y})$$

og dersom vi sentrerer, blir dette bare skalarproduktet mellom x og y delt på $n-1$. Hvis du ser nøye på $\frac{1}{n-1} X^T X$ vil du se at hvert element er den empiriske kovariansen mellom de forskjellige kolonnene i X siden $X^T X$ bare er skalarprodukter mellom kolonner i X .

- 11 Alle matriser på formen $A^T A$ er positivt definitte om A har full kolonnerang, og positivt semidefinitte om den ikke har det.¹ Dersom A har full kolonnerang, er

$$0 < \|Ax\|^2 = x^T A^T A x$$

så lenge $x \neq 0$. Dersom A ikke har full kolonnerang, er

$$0 \leq \|Ax\|^2 = x^T A^T A x$$

siden det finnes $x \neq 0$ slik at $Ax = 0$. Med $A = X$ og faktoren $\frac{1}{n-1} > 0$ er den empiriske kovariansmatrisen altså alltid positivt semidefinit.

- 12 PCA er en teknikk for å projisere dataskyen inn på en retning i rommet slik at variansen til det resulterende envariable datasettet blir maksimert:

https://en.wikipedia.org/wiki/Principal_component_analysis

Jeg har plaget dere med dette fordi det er sentralt i kunstig intelligens, og skal du skjønne noe av det er du nødt til å forstå lineæralgebra og statistikk. La oss anta for enkelhets skyld at X kun har to kolonner. Vi ønsker å finne den retningen p i koordinatsystemet der datamengden er plotta slik at datamengden projisert på p har så stor varians som mulig. Produktet Xp er en lang kolonnevektor der element k er projeksjonen av datapunkt k på p , og dersom vi ønsker å maksimere summen av kvadratene av disse, må vi maksimere

$$\|Xp\|^2 = p^T X^T X p$$

under bibetingelsen $\|p\| = 1$. Nå er det et standardresultat i numerisk analyse at om denne skal maksimeres, må man velge p som den egenvektoren til $X^T X$ med størst egenverdi.

- 13 Det karakteristiske polynomet er

$$\begin{aligned} \det \begin{pmatrix} 1-\lambda & 2 & 2 \\ 2 & 6-\lambda & 2 \\ 2 & 2 & 6-\lambda \end{pmatrix} &= (1-\lambda)((6-\lambda)^2 - 4) - 2(2(6-\lambda) - 4) + 2(4 - 2(6-\lambda)) \\ &= (1-\lambda)(4-\lambda)(8-\lambda) - 4(4-\lambda) - 4(4-\lambda) \\ &= (4-\lambda)((1-\lambda)(8-\lambda) - 8) \\ &= (4-\lambda)(\lambda^2 - 9\lambda) \\ &= \lambda(4-\lambda)(\lambda-9), \end{aligned}$$

¹Full rang betyr maksimalt antall lineært uavhengige rader eller kolonner:
[https://en.wikipedia.org/wiki/Rank_\(linear_algebra\)](https://en.wikipedia.org/wiki/Rank_(linear_algebra))

så egenverdiene er 0, 4 og 9. Matrisen er altså singulær. Matrisen endrer seg ikke om vi bytter om andre og tredje koordinat, så vi prøver oss frem:

$$\begin{pmatrix} 1 & 2 & 2 \\ 2 & 6 & 2 \\ 2 & 2 & 6 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 4 \\ -4 \end{pmatrix} \quad \text{og} \quad \begin{pmatrix} 1 & 2 & 2 \\ 2 & 6 & 2 \\ 2 & 2 & 6 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 9 \\ 18 \\ 18 \end{pmatrix}.$$

Den siste egenvektoren må stå ortogonalt på begge disse, og $(-4, 1, 1)^T$ gjør det. Den sjekker vi:

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

Egenverdiene er forskjellige, så egenvektorene kommer automatisk ut ortogonale (se slutten av løsningen på neste oppgave), og vi trenger ingen Gram-Schmidt. Normaliserer vi, blir

$$V = \begin{pmatrix} 1/3 & 0 & -4/(3\sqrt{2}) \\ 2/3 & 1/\sqrt{2} & 1/(3\sqrt{2}) \\ 2/3 & -1/\sqrt{2} & 1/(3\sqrt{2}) \end{pmatrix} \quad D = \begin{pmatrix} 9 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

og matrisen er VDV^T .

14 Vi beregner først det karakteristiske polynomet

$$\begin{aligned} \det \begin{pmatrix} 2-\lambda & 1 & 1 \\ 1 & 2-\lambda & 1 \\ 1 & 1 & 2-\lambda \end{pmatrix} &= (2-\lambda)((2-\lambda)^2 - 1) - (2-\lambda-1) + (1-(2-\lambda)) \\ &= (2-\lambda)(3-4\lambda+\lambda^2) - (1-\lambda) - (1-\lambda) \\ &= (2-\lambda)(3-\lambda)(1-\lambda) - (1-\lambda) - (1-\lambda) \\ &= (1-\lambda)((2-\lambda)(3-\lambda) - 2) \\ &= (1-\lambda)(4-5\lambda+\lambda^2) \\ &= (1-\lambda)(1-\lambda)(4-\lambda) \end{aligned}$$

og ser at vi har en enkel egenverdi $\lambda = 4$ og en dobbel egenverdi $\lambda = 1$. La oss begynne med den enkle egenverdien. Vi kunne gausset i vei, men siden vi elsker tall og observerer at radsummen er 4 i alle radene, må egenvektoren være

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Så den doble. Siden

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

ser vi at alle egenvektorer ligger i planet til likningen

$$x_1 + x_2 + x_3 = 0$$

så for eksempel

$$v_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \quad \text{og} \quad v_3 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$$

duger. Merk at begge er ortogonale på v_1 , men ikke på hverandre. Heldigvis er vi flinke i projeksjon, og ser at hvis vi bytter ut v_2 med

$$v_2 - \frac{v_2^T v_3}{v_3^T v_3} v_3 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$$

så er denne fortsatt egenvektor, men også ortogonal på både v_1 og v_3 . Dersom vi normaliserer alle egenvektorer og setter dem opp som kolonner i en matrise

$$V = \begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & 0 & -2/\sqrt{6} \\ 1/\sqrt{3} & -1/\sqrt{2} & 1/\sqrt{6} \end{pmatrix}$$

vil $A = VDV^T$, der

$$D = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Merk at egenvektorene som korresponderte til forskjellige egenverdier automatisk kom ut ortogonale. Dette var ikke tilfeldig, for

$$\lambda_1 v_1^T v_2 = (\lambda_1 v_1)^T v_2 = (Av_1)^T v_2 = v_1^T A^T v_2 = v_1^T (Av_2) = v_1^T (\lambda_2 v_2) = \lambda_2 v_1^T v_2$$

slik at

$$\lambda_1 v_1^T v_2 = \lambda_2 v_1^T v_2$$

eller

$$(\lambda_1 - \lambda_2) v_1^T v_2 = 0.$$

Dersom A er symmetrisk og $\lambda_1 \neq \lambda_2$ må altså $v_1^T v_2 = 0$.

- 15** Dersom X har dimensjon $n \times p$, har $X^T X$ og V dimensjon $p \times p$ mens XX^T og U er $n \times n$. Siden rangen til X er høyst p , kan høyst p av egenverdiene til XX^T være forskjellige fra null.

- 16** Dersom v er en egenvektor til $X^T X$ med egenverdi $\lambda \neq 0$, kan vi gange likningen

$$X^T X v = \lambda v$$

med X og få

$$XX^T X v = \lambda X v$$

som sier at Xv er en egenvektor til XX^T med den samme egenverdien. Dersom $\lambda = 0$ bryter dette sammen siden $Xv = 0$ og nullvektoren ikke klassifiserer som egenvektor. Resonnementet den andre veien er likt, og egenvektorene til XX^T kaller vi u . Egenverdiene kan ikke være negative, siden

$$0 \leq \|Xv\|^2 = v^T X^T X v = \lambda \|v\|^2.$$

- 17 Beregningen i oppgaven over viser at dersom $\|v\| = 1$ er

$$\|Xv\| = \sqrt{\lambda}.$$

- 18 Matrisen Σ har samme dimensjon som X , altså $n \times p$, med singularverdiene $\sigma_1 \geq \sigma_2 \geq \dots$ på diagonalen og nuller ellers. I de foregående oppgavene har vi vist at

$$Xv = \sqrt{\lambda}u$$

der v er en egenvektor til $X^T X$ og u er en egenvektor til XX^T , begge normaliserte og med egenverdi $\lambda \neq 0$. Setter vi opp alt dette i en matriselikning, får vi

$$XV = U\Sigma$$

der Σ nå er en kvadratisk diagonalmatrise med de positive singularverdiene på diagonalen. Dette kalles den **reduserte svd-faktoriseringen** til X . Dersom $X^T X$ eller XX^T har egenverdier som er null kan vi utvide faktoriseringen til å inkludere de korresponderende egenvektorene, men disse er sjelden nyttige i praktiske anvendelser. Svd-faktoriseringen til

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

finner vi ved å beregne egenverdier og egenvektorer til

$$X^T X = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

som er

$$\lambda_1 = 3 \quad v_1 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} \quad \text{og} \quad \lambda_2 = 1 \quad v_2 = \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}$$

og for

$$XX^T = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

som er

$$\lambda_1 = 3 \quad u_1 = \begin{pmatrix} 1/\sqrt{6} \\ 1/\sqrt{6} \\ 2/\sqrt{6} \end{pmatrix} \quad \text{og} \quad \lambda_2 = 1 \quad u_2 = \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \end{pmatrix} \quad \text{og} \quad \lambda_3 = 0 \quad u_3 = \begin{pmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ -1/\sqrt{3} \end{pmatrix}$$

Du kan nå selv sjekke at

$$X = U\Sigma V^T = \begin{pmatrix} 1/\sqrt{6} & 1/\sqrt{2} \\ 1/\sqrt{6} & -1/\sqrt{2} \\ 2/\sqrt{6} & 0 \end{pmatrix} \begin{pmatrix} \sqrt{3} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix}$$

Dette kalles egentlig **redusert svd**. Vi kan inkludere den siste egenvektoren til XX^T , og skrive

$$X = \begin{pmatrix} 1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \\ 1/\sqrt{6} & -1/\sqrt{2} & 1/\sqrt{3} \\ 2/\sqrt{6} & 0 & -1/\sqrt{3} \end{pmatrix} \begin{pmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix}$$

som kalles **full svd**.

- 19 <https://folk.ntnu.no/mortano/python/katt/>

GAMLE EKSAMENSOPPGAVER

1 Vi beregner først

$$A^T A = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

og

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

Egenvektormatrisen til $A^T A$ kan vi sette til identitetsmatrisen siden $A^T A$ er diagonal, mens en egenvektormatrise til $A A^T$ er

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

Eigenverdiene er en dobbel 2 (disse står på diagonalen i $A^T A$) og en dobbel 0 (siden $A^T A$ er 2×2 og $A A^T$ er 4×4), så en redusert svd-faktorisering er

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1/\sqrt{2} & 0 \\ 0 & 1/\sqrt{2} \\ 1/\sqrt{2} & 0 \\ 0 & 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} \sqrt{2} & 0 \\ 0 & \sqrt{2} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

mens en full er

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

2 Matrisen er blokkdiagonal med blokkene 1,

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

og 4, så vi kan invertere og diagonalisere hver blokk for seg. Siden $B^2 = 2I$, er $B^{-1} = B/2$, og

$$M^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1/2 & 1/2 & 0 \\ 0 & 1/2 & -1/2 & 0 \\ 0 & 0 & 0 & 1/4 \end{pmatrix}.$$

Det karakteristiske polynomet til B er $\lambda^2 - 2$, så eigenverdiene er $\pm\sqrt{2}$. For $\lambda = \sqrt{2}$ gir den første raden $(1 - \sqrt{2})v_1 + v_2 = 0$, altså $v = (1, \sqrt{2} - 1)^T$, og for $\lambda = -\sqrt{2}$ får vi $v = (1, -\sqrt{2} - 1)^T$. Disse står ortogonalt på hverandre, som seg hør og bør for en symmetrisk matrise. Det er

penest å legge merke til at $\sqrt{2} - 1 = \tan(\pi/8)$, så normaliserte egenvektorer er $(\cos \frac{\pi}{8}, \sin \frac{\pi}{8})^T$ og $(-\sin \frac{\pi}{8}, \cos \frac{\pi}{8})^T$. Da er $M = VDV^T$ med

$$V = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \frac{\pi}{8} & -\sin \frac{\pi}{8} & 0 \\ 0 & \sin \frac{\pi}{8} & \cos \frac{\pi}{8} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{og} \quad D = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \sqrt{2} & 0 & 0 \\ 0 & 0 & -\sqrt{2} & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}.$$

Blokken B er $\sqrt{2}$ ganger en speiling om linja med vinkel $\pi/8$, og V er en rotasjon på $\pi/8$ i x_2x_3 -planet.

- 3 Med $p(x) = \beta_1 x + \beta_0$ blir normallikningene $A^T A \beta = A^T y$ med

$$A = \begin{pmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{pmatrix}, \quad A^T A = \begin{pmatrix} 14 & 6 \\ 6 & 3 \end{pmatrix}, \quad A^T y = \begin{pmatrix} 23 \\ 10 \end{pmatrix},$$

som har løsning $\beta_1 = 3/2$ og $\beta_0 = 1/3$:

$$p(x) = \frac{3}{2}x + \frac{1}{3}.$$

Formlene fra første side av økten gir det samme, med $\bar{x} = 2$ og $\bar{y} = 10/3$:

$$\beta_1 = \frac{x^T y - n \bar{x} \bar{y}}{x^T x - n \bar{x}^2} = \frac{23 - 20}{14 - 12} = \frac{3}{2} \quad \text{og} \quad \beta_0 = \bar{y} - \beta_1 \bar{x} = \frac{1}{3}.$$

- 4 Punktene ligger på den rette linja $y = x + 1$, så den ligger parabolen $p(x) = 0 \cdot x^2 + x + 1$ nøyaktig gjennom alle fire, med residual null. Bedre enn null blir det ikke, og siden $A^T A$ er inverterbar (fire forskjellige x -verdier), er dette den eneste løsningen av normallikningene. Andreordensleddet er altså null, og skissen er linja gjennom punktene. Kjører du normallikningene for hånd, får du

$$A^T A = \begin{pmatrix} 18 & 8 & 6 \\ 8 & 6 & 2 \\ 6 & 2 & 4 \end{pmatrix} \quad \text{og} \quad A^T y = \begin{pmatrix} 14 \\ 8 \\ 6 \end{pmatrix},$$

og $(\beta_2, \beta_1, \beta_0) = (0, 1, 1)$ passer.

- 5 Vi skriver punktene som (x_1, x_2, y) og søker $p(x) = \beta_1 x_1 + \beta_2 x_2 + \beta_0$. Da er

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \quad A^T A = \begin{pmatrix} 2 & 1 & 2 \\ 1 & 2 & 2 \\ 2 & 2 & 4 \end{pmatrix}, \quad A^T y = \begin{pmatrix} 2 \\ 2 \\ 3 \end{pmatrix}.$$

Symmetrien i de to første likningene gir $\beta_1 = \beta_2$, så $3\beta_1 + 2\beta_0 = 2$ og $4\beta_1 + 4\beta_0 = 3$, som gir $\beta_1 = \beta_2 = 1/2$ og $\beta_0 = 1/4$:

$$p(x) = \frac{1}{2}x_1 + \frac{1}{2}x_2 + \frac{1}{4}.$$

Planet treffer ingen av de fire punktene, men bommer med $1/4$ på alle.

- 6 Kall matrisen P . Rad én og tre er like, og rad to og fire er like, så rangen er 2 og egenverdien 0 har multiplisitet 2. Matrisen ganger $(1, 0, 1, 0)^T$ og $(0, 1, 0, 1)^T$ med 2, og sender $(1, 0, -1, 0)^T$

og $(0, 1, 0, -1)^T$ til null. De fire vektorene er innbyrdes ortogonale, så normaliserer vi dem, blir $P = VDV^T$ med

$$V = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix} \quad \text{og} \quad D = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Legg merke til at $P = AA^T$ for matrisen A i oppgave 1, og at $P/2$ er en projeksjon: $P^2 = 2P$.

- 7 Denne matrisen er den transponerte av matrisen i oppgave 1, så vi kan bytte rollene til U og V der. Vi regner likevel ut:

$$AA^T = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \quad \text{og} \quad A^T A = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix},$$

altså matrisen fra forrige oppgave. Singulærverdiene er $\sqrt{2}$ og $\sqrt{2}$, $U = I$ siden AA^T er diagonal, og kolonnene i V er egenvektorene fra forrige oppgave. En redusert svd-faktoriseringen er

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{2} & 0 \\ 0 & \sqrt{2} \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix},$$

og en full er

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{2} & 0 & 0 & 0 \\ 0 & \sqrt{2} & 0 & 0 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix},$$

der den siste matrisen er V^T .

- 8 Matrisen er symmetrisk med egenverdier 3 og 1 og egenvektorer $(1, 1)^T$ og $(1, -1)^T$, så $A = VDV^T$ med $V = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ og $D = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}$. Da er $A^{10} = VD^{10}V^T$, siden alle $V^T V$ i midten blir I :

$$A^{10} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 3^{10} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 3^{10} + 1 & 3^{10} - 1 \\ 3^{10} - 1 & 3^{10} + 1 \end{pmatrix} = \begin{pmatrix} 29525 & 29524 \\ 29524 & 29525 \end{pmatrix}.$$

Den store egenverdien tar over fullstendig: A^{10} er nesten 3^{10} ganger projeksjonen på $(1, 1)^T$.