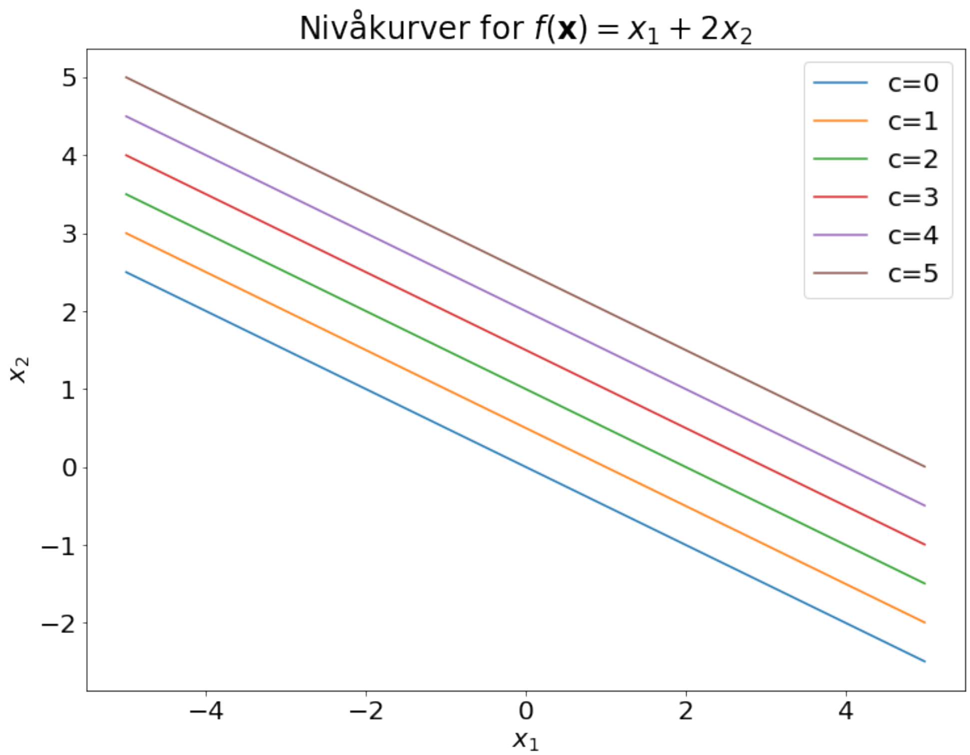
2-6 Regresjon LF

Astrid Mysterud, februar 2024

1) Nivåkurvene til $f(\mathbf{x}) = x_1 + 2x_2$ er på formen $x_1 + 2x_2 = c$ hvor c er en konstant. Dette kjenner vi igjen som rette linjer $x_2 = -\frac{1}{2}x_1 + \frac{c}{2}$ i x_1-x_2-planet med stigningstall $-\frac{1}{2}$ og konstantledd $\frac{c}{2}$. Her er nivåkurver for $c \in \{0, 1, 2, 3, 4, 5\}$ plottet i python.



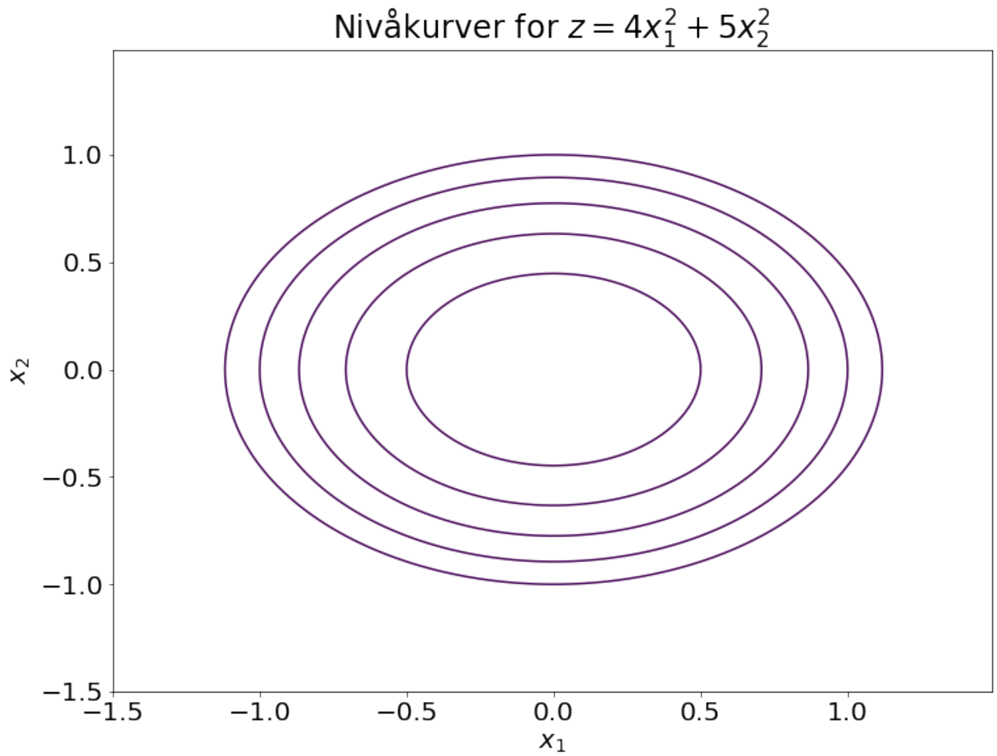
2) Nivåkurvene til $z = 4x_1^2 + 5x_2^2$ er gitt ved

$$4x_1^2 + 5x_2^2 = c$$

$$\frac{4}{c}x_1^2 + \frac{5}{c}x_2^2 = 1$$

$$\implies \left(\frac{x_1}{\frac{\sqrt{c}}{2}}\right)^2 + \left(\frac{x_2}{\sqrt{\frac{c}{5}}}\right)^2 = 1$$

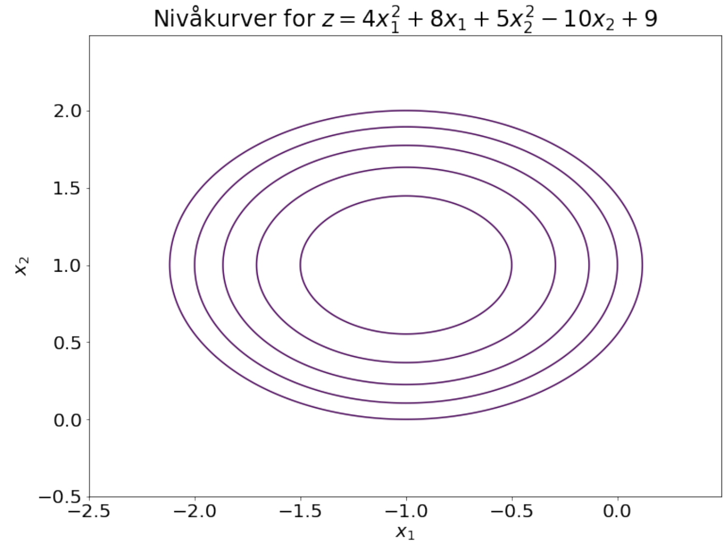
Dette kjenner vi igjen som likningen for en ellipse sentrert i origo med halvakser $\frac{\sqrt{c}}{2}$ og $\sqrt{\frac{c}{5}}$ for konstanter c. Plottet i python for $c \in \{1, 2, 3, 4, 5\}$ ser det slik ut.



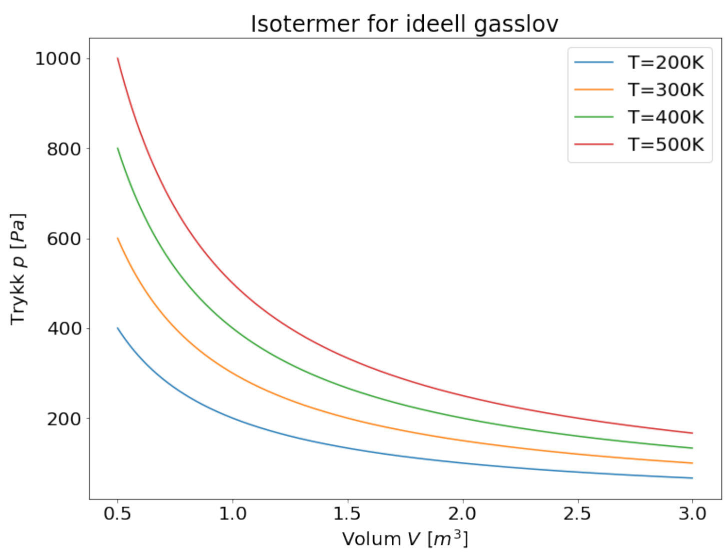
3) Nivåkurvene til $z = 4x_1^2 + 8x_1 + 5x_2^2 - 10x_2 + 9$ er gitt ved

$$\begin{aligned} 4x_1^2 + 8x_1 + 5x_2^2 - 10x_2 + 9 &= c \\ 4(x_1^2 + 2x_1) + 5(x_2^2 - 2x_2) + 9 &= c \\ 4(x_1^2 + 2x_1 + 1) + 5(x_2^2 - 2x_2 + 1) - 4 - 5 + 9 &= c \\ 4(x_1 + 1)^2 + 5(x_2 - 1)^2 &= c \\ \frac{4}{c}(x_1 + 1)^2 + \frac{5}{c}(x_2 - 1)^2 &= 1 \\ \implies \left(\frac{x_1 + 1}{\frac{\sqrt{c}}{2}}\right)^2 + \left(\frac{x_2 - 1}{\sqrt{\frac{c}{5}}}\right)^2 &= 1. \end{aligned}$$

Dette kjenner vi igjen som en ellipse med sentrum i $(-1, 1)$ og halvakser $\frac{\sqrt{c}}{2}$ og $\sqrt{\frac{c}{5}}$ for konstanter c. Følgende plott viser nivåkurvene for $c \in \{1, 2, 3, 4, 5\}$, som da ser veldig like ut som i sta, men med annet sentrum.



4) Den ideelle gasslov er gitt ved $pV = NkT$. Trykket som en funksjon av volum beskrives da ved $p = \frac{NkT}{V}$. Vi plotter isotermer ved å bestemme oss for noen temperaturer og plotte $p = \frac{NkT}{V}$ for disse temperaturene i et pV-diagram. Med $Nk = 1$ får vi følgende isotermer.



Følgende viser temperatur T som en funksjon av trykk p og volum V.

Her kommer det et plot

5) Trykket i Steinkjer finner vi ved å se på trykket langs nivåkurven (isobaren) nærmest Steinkjer. Nivåkurven merket med 980 (jeg tror enheten er hPa) går ish gjennom Steinkjer. Det er dermed rimelig å si at trykket på Steinkjer denne dagen var 980 hPa.

6) Partiellderiverte til $f(\mathbf{x}) = x_1^2 + x_1x_2 + x_2^2 + x_1 + x_2$ er

$$\begin{aligned} \frac{\partial f}{\partial x_1} &= 2x_1 + x_2 + 1 \\ \text{og} \quad \frac{\partial f}{\partial x_2} &= x_1 + 2x_2 + 1. \end{aligned}$$

7) De partiellderiverte til $T(p, V) = pV$ er

$$\frac{\partial T}{\partial p} = V \quad \text{og} \quad \frac{\partial T}{\partial V} = p.$$

8) Gradienten finner vi ved å putte de partiellderiverte inn i en vektor, for oppgave 6 blir gradienten

$$\begin{aligned} \nabla f(x_1, x_2) &= \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}\right) \\ &= (2x_1 + x_2 + 1,\ x_1 + 2x_2 + 1), \end{aligned}$$

og for oppgave 7 blir det

$$\begin{aligned} \nabla T(p, V) &= \left(\frac{\partial T}{\partial p}, \frac{\partial T}{\partial V}\right) \\ &= (V,\ p). \end{aligned}$$

9) Hvis vi tenker oss at retning nord er gitt ved $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ og retning vest er gitt ved $\begin{pmatrix} -1 \\ 0 \end{pmatrix}$, så vil nordvest være retning $\mathbf{v} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$. Vi er altså ute etter den retningsderiverte i punktet $(1, 2)$ i retning $\mathbf{v}$. Vi starter med å finne gradienten til f, den er gitt ved

$$\nabla f(x, y) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) = (2x + y + 1,\ x + 2y + 1).$$

Gradienten i punktet $(1, 2)$ er dermed

$$\nabla f(1, 2) = (2 \cdot 1 + 2 + 1,\ 1 + 2 \cdot 2 + 1) = (5, 6).$$

For å finne den retningsderiverte i retning $\mathbf{v}$, må vi ta skalarproduktet mellom $\nabla f(1, 2)$ og enhetsvektoren

$\mathbf{n_v}$ i retning $\mathbf{v}$. Enhetsvektoren finner vi ved å dele på lengden til $\mathbf{v}$, altså

$$\mathbf{n_v} = \frac{\mathbf{v}}{||\mathbf{v}||} = \frac{\begin{pmatrix} -1 \\ 1 \end{pmatrix}}{\sqrt{1^2+1^2}} = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}.$$

Dermed blir den retningsderiverte i punktet $(1,2)$ i retning nordvest gitt ved

$$\nabla f(1,2) \cdot \mathbf{n_v} = (5,6) \cdot \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} = -\frac{5}{\sqrt{2}} + \frac{6}{\sqrt{2}} = \frac{1}{\sqrt{2}}.$$

10) Langs en ekvidistanselinje er funksjonsverdien konstant, den retningsderiverte vil altså være 0 i retning langs en ekvidistanselinje. Vi ønsker å finne retningen $\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$, hvor $\sqrt{u_1^2+u_2^2} = 1$, slik at vi går langs en ekvidistanselinje. Den retningsderiverte blir

$$\nabla f(1,2) \cdot \mathbf{u} = (5,6) \cdot \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = 5u_1 + 6u_2 = 0,$$

dermed blir $u_1 = -\frac{6}{5}u_2$. Ved å putte dette inn i $\sqrt{u_1^2+u_2^2} = 1$, får vi at

$$1 = \sqrt{\left(-\frac{6}{5}u_2\right)^2 + u_2^2} = \sqrt{\frac{36}{25}u_2^2 + u_2^2} = \sqrt{\left(\frac{36}{25}+1\right)u_2^2} = \sqrt{\frac{61}{25}}u_2.$$

Da blir $u_2 = \sqrt{\frac{25}{61}} = \frac{5}{\sqrt{61}}$, og dermed $u_1 = -\frac{6}{5} \cdot \frac{5}{\sqrt{61}} = -\frac{6}{\sqrt{61}}$. Altså må vi peke skiene i retning $\begin{pmatrix} -\frac{6}{\sqrt{61}} \\ \frac{5}{\sqrt{61}} \end{pmatrix}$.

11) Jeg tolker "kjøre rett utfor så bratt som mulig" som at vi ønsker å finne den retningen med mest negativ stigning (retningsderivert). I en retning $\mathbf{v}$ med enhetsvektor $\mathbf{n_v}$ er den retningsderiverte gitt ved

$$\nabla f(x,y) \cdot \mathbf{n_v} = ||\nabla f(x,y)||\,||\mathbf{n_v}|| \cos\theta = ||\nabla f(x,y)|| \cos\theta$$

hvor θ er vinkelen mellom $\nabla f(x,y)$ og $\mathbf{n_v}$. $||\nabla f(x,y)|| \cos\theta$ er mest negativ når $\cos\theta = -1$, altså når $\theta = \pi$. Dermed må $\mathbf{v}$ være vinkel π unna $\nabla f(x,y)$. Dette betyr at $\mathbf{v}$ er *antiparallell* med $\nabla f(x,y)$, altså at $\mathbf{v} = -\nabla f(x,y)$. Siden vi står i $(1,2)$ får vi $\mathbf{v} = -\nabla f(1,2) = -(5,6) = (-5,-6)$.

12) Mulige toppunkt er punkter hvor $\nabla f(x_1,x_2) = \mathbf{0}$, altså kritiske punkter. Dermed starter vi med å finne kritiske punkter. Vi har at

$$\nabla f(x_1,x_2) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}\right) = (-2x_1 - x_2 + 1,\ -2x_2 - x_1 + 1).$$

Vi er ute etter punkter hvor $(-2x_1 - x_2 + 1,\, -2x_2 - x_1 + 1) = (0, 0)$, altså

$$\begin{cases} -2x_1 - x_2 + 1 = 0 \\ -2x_2 - x_1 + 1 = 0 \end{cases}$$

Videre, løser vi for x_1 og x_2. Fra første likning har vi at $x_2 = 1 - 2x_1$. Ved å putte dette inn i andre likning får vi

$$\begin{aligned} -2(1 - 2x_1) - x_1 + 1 &= 0 \\ 3x_1 - 1 = 0 \implies x_1 &= \frac{1}{3} \end{aligned}$$

Da blir $x_2 = 1 - 2 \cdot \frac{1}{3} = \frac{1}{3}$. Altså har vi kun et kritisk punkt, nemlig $\left(\frac{1}{3},\, \frac{1}{3}\right)$. Dette må da være toppunktet det blir spurt om i oppgaven (det gir også mening at vi kun har et toppunkt ettersom vi har et andregradspolynom som går mot $-\infty$ siden vi har negative koeffisienter foran x_1^2 og x_2^2). Toppunktet er altså gitt ved $\left(\frac{1}{3},\, \frac{1}{3}, f(\frac{1}{3},\, \frac{1}{3})\right) = \left(\frac{1}{3},\, \frac{1}{3}, \frac{4}{3}\right)$.

13) Gradienten er gitt ved

$$\begin{aligned} \nabla f(\beta_0,\, \beta_1) &= \left(\frac{\partial f}{\partial \beta_0},\, \frac{\partial f}{\partial \beta_1}\right) \\ &= (2n\beta_0 + 2\beta_1 n\overline{\mathbf{x}} - 2n\overline{\mathbf{y}},\, 2\beta_1 ||\mathbf{x}||^2 + 2\beta_0 n\overline{\mathbf{x}} - 2\mathbf{x}^T\mathbf{y}). \end{aligned}$$

Ved å sette lik $\mathbf{0}$, får vi likningene

$$\begin{cases} 2n\beta_0 + 2\beta_1 n\overline{\mathbf{x}} - 2n\overline{\mathbf{y}} = 0 \\ 2\beta_1 ||\mathbf{x}||^2 + 2\beta_0 n\overline{\mathbf{x}} - 2\mathbf{x}^T\mathbf{y} = 0 \end{cases}$$

Ved å dele første likning på $2n$ får vi at $\beta_0 = \overline{\mathbf{y}} - \beta_1\overline{\mathbf{x}}$. Vi putter dette inn i andre likning og får

$$\begin{aligned} 2\beta_1 ||\mathbf{x}||^2 + 2(\overline{\mathbf{y}} - \beta_1\overline{\mathbf{x}})n\overline{\mathbf{x}} - 2\mathbf{x}^T\mathbf{y} &= 0 \\ \beta_1 ||\mathbf{x}||^2 + n\overline{\mathbf{x}}\,\overline{\mathbf{y}} - n\beta_1(\overline{\mathbf{x}})^2 - \mathbf{x}^T\mathbf{y} &= 0 \\ \beta_1 \left(||\mathbf{x}||^2 - n(\overline{\mathbf{x}})^2\right) + n\overline{\mathbf{x}}\,\overline{\mathbf{y}} - \mathbf{x}^T\mathbf{y} &= 0 \\ \implies \beta_1 = \frac{\mathbf{x}^T\mathbf{y} - n\overline{\mathbf{x}}\,\overline{\mathbf{y}}}{||\mathbf{x}||^2 - n(\overline{\mathbf{x}})^2} \end{aligned}$$

14) Vi ønsker å løse $A^TA\mathbf{v} = A^T\mathbf{y}$. Vi starter med å regne ut A^TA

$$
\begin{aligned}
A^TA &= \begin{pmatrix} x_1 & x_2 & \dots & x_n \\ 1 & 1 & \dots & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{pmatrix} \\
&= \begin{pmatrix} x_1^2 + x_2^2 + \cdots + x_n^2 & x_1 + x_2 + \cdots + x_n \\ x_1 + x_2 + \cdots + x_n & 1 + 1 + \cdots + 1 \end{pmatrix} \\
&= \begin{pmatrix} ||\mathbf{x}||^2 & n\overline{\mathbf{x}} \\ n\overline{\mathbf{x}} & n \end{pmatrix}
\end{aligned}
$$

Videre, regner vi ut $A^T\mathbf{y}$

$$
\begin{aligned}
A^T\mathbf{y} &= \begin{pmatrix} x_1 & x_2 & \dots & x_n \\ 1 & 1 & \dots & 1 \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \\
&= \begin{pmatrix} x_1y_1 + x_2y_2 + \cdots + x_ny_n \\ y_1 + y_2 + \cdots + y_n \end{pmatrix} \\
&= \begin{pmatrix} \mathbf{x}^T\mathbf{y} \\ n\overline{\mathbf{y}} \end{pmatrix}
\end{aligned}
$$

Likningssystemet $A^TA\mathbf{v} = A^T\mathbf{y}$ blir altså

$$
A^TA\mathbf{v} = A^T\mathbf{y}
$$

$$
\begin{pmatrix} ||\mathbf{x}||^2 & n\overline{\mathbf{x}} \\ n\overline{\mathbf{x}} & n \end{pmatrix} \cdot \begin{pmatrix} \beta_1 \\ \beta_0 \end{pmatrix} = \begin{pmatrix} \mathbf{x}^T\mathbf{y} \\ n\overline{\mathbf{y}} \end{pmatrix}
$$

Ved å utføre matrisemultiplikasjonen på venstre side, ser vi at vi kan skrive likningssystemet slik

$$
\begin{cases} \beta_1||\mathbf{x}||^2 + \beta_0 \cdot n\overline{\mathbf{x}} = \mathbf{x}^T\mathbf{y} \\ \beta_1 \cdot n\overline{\mathbf{x}} + \beta_0 \cdot n = n\overline{\mathbf{y}} \end{cases}
$$

Ved å dele på n i nederste likning og løse for β_0 får vi $\beta_0 = \overline{\mathbf{y}} - \beta_1\overline{\mathbf{x}}$. Vi putter dette inn i første likning og løser for β_1

$$\begin{aligned}
\beta_1||\mathbf{x}||^2 + (\overline{\mathbf{y}} - \beta_1\overline{\mathbf{x}}) \cdot n\overline{\mathbf{x}} &= \mathbf{x}^T\mathbf{y} \\
\beta_1\left(||\mathbf{x}||^2 - n(\overline{\mathbf{x}})^2\right) + n\overline{\mathbf{x}}\,\overline{\mathbf{y}} &= \mathbf{x}^T\mathbf{y} \\
\implies \beta_1 &= \frac{\mathbf{x}^T\mathbf{y} - n\overline{\mathbf{x}}\,\overline{\mathbf{y}}}{||\mathbf{x}||^2 - n(\overline{\mathbf{x}})^2}
\end{aligned}$$

Som er de samme uttrykkene for β_0 og β_1 som tidligere.

15) Det er ikke noe spesielt man skal gjøre her annet enn å teste forståelsen. Her er det viktig å forstå at når du ganger to matriser sammen, så tar du masse skalarprodukter. Altså når vi ser på produktet $A^T(A\mathbf{v}-\mathbf{y})$ så vil resultatet, som er en vektor, inneholde skalarproduktene mellom *radene* i A^T og kolonnevektoren $A\mathbf{v}-\mathbf{y}$. *Radene* i A^T er *kolonnene* i A. For at vektoren $A\mathbf{v}-\mathbf{y}$ skal stå ortogonalt på *alle kolonnene i* A, må alle disse skalarproduktene være 0. Dermed må matriseproduktet $A^T(A\mathbf{v}-\mathbf{y})$ være lik nullvektoren $\mathbf{0}$. (Hver 0 i $\mathbf{0}$ er et skalarprodukt mellom en kolonne i A, altså en rad i A^T, og vektoren $A\mathbf{v}-\mathbf{y}$).

16) Vi ønsker å finne det kvadratiske regresjonspolynomet til datasettet oppgitt. Et kvadratisk polynom er på formen $\beta_2 x^2 + \beta_1 x + \beta_0$. Oppgaven er å finne koeffisientene β_2, β_1 og β_0. Vi finner disse ved å løse $A^T A\mathbf{v} = A^T\mathbf{y}$ hvor A er matrisen

$$A = \begin{pmatrix} x_1^2 & x_1 & 1 \\ x_2^2 & x_2 & 1 \\ x_3^2 & x_3 & 1 \\ x_4^2 & x_4 & 1 \\ x_5^2 & x_5 & 1 \end{pmatrix} = \begin{pmatrix} 1^2 & 1 & 1 \\ 2^2 & 2 & 1 \\ 3^2 & 3 & 1 \\ 4^2 & 4 & 1 \\ 5^2 & 5 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 4 & 2 & 1 \\ 9 & 3 & 1 \\ 16 & 4 & 1 \\ 25 & 5 & 1 \end{pmatrix}$$

$\mathbf{v}$ er vektoren $\mathbf{v} = \begin{pmatrix} \beta_2 & \beta_1 & \beta_0 \end{pmatrix}^T$ og $\mathbf{y}$ er vektoren

$$\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ 4 \\ 5 \\ 1 \end{pmatrix}$$

Vi starter med å beregne $A^T A$

$$A^T A = \begin{pmatrix} 1 & 4 & 9 & 16 & 25 \\ 1 & 2 & 3 & 4 & 5 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 1 \\ 4 & 2 & 1 \\ 9 & 3 & 1 \\ 16 & 4 & 1 \\ 25 & 5 & 1 \end{pmatrix} = \begin{pmatrix} 979 & 225 & 55 \\ 225 & 55 & 15 \\ 55 & 15 & 5 \end{pmatrix}$$

Videre, beregner vi $A^T \mathbf{y}$

$$A^T \mathbf{y} = \begin{pmatrix} 1 & 4 & 9 & 16 & 25 \\ 1 & 2 & 3 & 4 & 5 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 3 \\ 4 \\ 5 \\ 1 \end{pmatrix} = \begin{pmatrix} 155 \\ 45 \\ 15 \end{pmatrix}$$

Likningssystemet $A^T A \mathbf{v} = A^T \mathbf{y}$ blir altså

$$\begin{pmatrix} 979 & 225 & 55 \\ 225 & 55 & 15 \\ 55 & 15 & 5 \end{pmatrix} \begin{pmatrix} \beta_2 \\ \beta_1 \\ \beta_0 \end{pmatrix} = \begin{pmatrix} 155 \\ 45 \\ 15 \end{pmatrix}$$

som har totalmatrise

$$\left(\begin{array}{ccc|c} 979 & 225 & 55 & 155 \\ 225 & 55 & 15 & 45 \\ 55 & 15 & 5 & 15 \end{array}\right)$$

Ved å gausseliminere (eller putte inn i WolframAlpha) får man at

$$\mathbf{v} = \begin{pmatrix} -\frac{5}{7} \\ \frac{30}{7} \\ -2 \end{pmatrix}$$

Altså er det kvadratiske regresjonspolynomet gitt ved $-\frac{5}{7}x^2 + \frac{30}{7}x - 2$.

To be continued...