OH till Föreläsning 5, Numme T1, 120921

GKN Kap 4.1D & 4.2C-1 Linjära Minstakvadratmetoden, (MKV)

Lösning saknas?

$$\begin{cases} x+y=5 \\ x+2y=6 \\ x+2y=7 \end{cases} \iff \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 5 \\ 6 \\ 7 \end{pmatrix} \iff Ac = b$$

Använder vi bara ekvation 1 och 2 får vi $x = 4$ och $y = 1$ vilket inte stämmer i ekvation 3. Använder vi istället ekvation 1 och 3 får vi $x = 3$ och $y = 2$ vilket inte stämmer i ekvation 2. Båda sätten ger $||r||_2 = 1$.

$$r = b - Ac = \begin{pmatrix} 5 \\ 6 \\ 7 \end{pmatrix} - \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 4 \\ 1 \end{pmatrix} = \begin{pmatrix} 5 \\ 6 \\ 7 \end{pmatrix} - \begin{pmatrix} 5 \\ 6 \\ 6 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

$$r = b - Ac = \begin{pmatrix} 5 \\ 6 \\ 7 \end{pmatrix} - \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 5 \\ 6 \\ 7 \end{pmatrix} - \begin{pmatrix} 5 \\ 7 \\ 7 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}$$

Hitta en lösning så att normen av felet minimeras. Hitta c så att $||r||_2$ minimeras.

Minimum fås då derivatan av $||r||_2^2$ med avseende på c är noll.

$$||r||_2^2 = r^T r = (b - Ac)^T (b - Ac) = b^T b + c^T A^T A\, c - 2c^T A^T b$$

$$\frac{\partial \left\{ ||r||_2^2 \right\}}{\partial c} = 2A^T A\, c - 2A^T b = 0 \implies A^T A\, c = A^T b$$

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 2 \end{pmatrix} \quad b = \begin{pmatrix} 5 \\ 6 \\ 7 \end{pmatrix} \quad A^T A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 3 & 5 \\ 5 & 9 \end{pmatrix} \quad A^T b = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \end{pmatrix} \begin{pmatrix} 5 \\ 6 \\ 7 \end{pmatrix} = \begin{pmatrix} 18 \\ 31 \end{pmatrix}$$

$$c = \left(A^T A\right)^{-1} A^T b = \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3.5 \\ 1.5 \end{pmatrix} \quad r = b - Ac = \begin{pmatrix} 5 \\ 6 \\ 7 \end{pmatrix} - \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 3.5 \\ 1.5 \end{pmatrix} = \begin{pmatrix} 5 \\ 6 \\ 7 \end{pmatrix} - \begin{pmatrix} 5 \\ 6.5 \\ 6.5 \end{pmatrix} = \begin{pmatrix} 0 \\ -0.5 \\ 0.5 \end{pmatrix}$$

$$||r||_2 = \sqrt{0^2 + (-0.5)^2 + 0.5^2} = 1/\sqrt{2} \approx 0.7071$$

När vi hittar en lösning så att kvadraten på normen av residualen minimeras har vi ”löst problemet” med **minsta-kvadrat-metoden**.

Anpassa en rät linje till givna mätdata, $p(x) = k\,x + m$

x	y
1	1
2	4
4	5
5	7

$$\begin{cases} k\,x_1 + m = y_1 \\ k\,x_2 + m = y_2 \\ k\,x_3 + m = y_3 \\ k\,x_4 + m = y_4 \end{cases} \quad \begin{cases} k \cdot 1 + m = 1 \\ k \cdot 2 + m = 4 \\ k \cdot 4 + m = 5 \\ k \cdot 5 + m = 7 \end{cases} \quad \begin{pmatrix} 1 & 1 \\ 2 & 1 \\ 4 & 1 \\ 5 & 1 \end{pmatrix} \begin{pmatrix} k \\ m \end{pmatrix} = \begin{pmatrix} 1 \\ 4 \\ 5 \\ 7 \end{pmatrix} \quad c = \left(A^T A\right)^{-1} A^T b$$

$$A^T A = \begin{pmatrix} 1 & 2 & 4 & 5 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 2 & 1 \\ 4 & 1 \\ 5 & 1 \end{pmatrix} = \begin{pmatrix} 46 & 12 \\ 12 & 4 \end{pmatrix} \quad A^T b = \begin{pmatrix} 1 & 2 & 4 & 5 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 4 \\ 5 \\ 7 \end{pmatrix} = \begin{pmatrix} 64 \\ 17 \end{pmatrix} \Rightarrow c = \begin{pmatrix} k \\ m \end{pmatrix} = \begin{pmatrix} 1.3 \\ 0.35 \end{pmatrix}$$

Koll1 : $$r = b - A\,c = \begin{pmatrix}1\\4\\5\\7\end{pmatrix} - \begin{pmatrix}1&1\\2&1\\4&1\\5&1\end{pmatrix}\begin{pmatrix}1.3\\0.35\end{pmatrix} = \begin{pmatrix}1\\4\\5\\7\end{pmatrix} - \begin{pmatrix}1.65\\2.95\\5.55\\6.85\end{pmatrix} = \begin{pmatrix}-0.65\\1.05\\-0.55\\0.15\end{pmatrix}$$

Koll2 : $$A^T r = \begin{pmatrix}1&2&4&5\\1&1&1&1\end{pmatrix}\begin{pmatrix}-0.65\\1.05\\-0.55\\0.15\end{pmatrix} = 10^{-13}\cdot\begin{pmatrix}-0.1288\\-0.0400\end{pmatrix}$$

Anpassa ett andragradspolynom till givna mätdata, $p(x) = c_1 + c_2\,x + c_3\,x^2$

x	y
1	1
2	4
4	5
5	7

$$\begin{cases} c_1 + c_2\,x_1 + c_3\,x_1^2 = y_1\\ c_1 + c_2\,x_2 + c_3\,x_2^2 = y_2\\ c_1 + c_2\,x_3 + c_3\,x_3^2 = y_3\\ c_1 + c_2\,x_4 + c_3\,x_4^2 = y_4 \end{cases} \qquad \begin{cases} c_1 + c_2\cdot 1 + c_3\cdot 1^2 = 1\\ c_1 + c_2\cdot 2 + c_3\cdot 2^2 = 4\\ c_1 + c_2\cdot 4 + c_3\cdot 4^2 = 5\\ c_1 + c_2\cdot 5 + c_3\cdot 5^2 = 7 \end{cases} \qquad \begin{pmatrix}1&1&1^2\\1&2&2^2\\1&4&4^2\\1&5&5^2\end{pmatrix}\begin{pmatrix}c_1\\c_2\\c_3\end{pmatrix} = \begin{pmatrix}1\\4\\5\\7\end{pmatrix}$$

$$A^TA = \begin{pmatrix}4&12&46\\12&46&198\\46&198&898\end{pmatrix}\ A^Tb = \begin{pmatrix}17\\64\\272\end{pmatrix} \Rightarrow c = \begin{pmatrix}-0.7333\\2.3000\\-0.1667\end{pmatrix} \Rightarrow r = b - A\,c = \begin{pmatrix}-0.4\\0.8\\-0.8\\0.4\end{pmatrix} \Rightarrow A^Tr = 10^{-13}\begin{pmatrix}0.0400\\0.1643\\0.7860\end{pmatrix}$$

Tre parametrar innebär tre basfunktioner:

$$\begin{cases}\varphi_1(x) = 1\\ \varphi_2(x) = x\\ \varphi_3(x) = x^2\end{cases} \implies p(x) = c_1\,\varphi_1(x) + c_2\,\varphi_2(x) + c_3\,\varphi_3(x) \implies A = \begin{pmatrix}|&|&|\\ \varphi_1(x) & \varphi_2(x) & \varphi_3(x)\\ |&|&|\end{pmatrix}$$

Andragradspolynom med bättre basfunktioner (centrerade)

$$\begin{cases}\phi_1(x) = 1\\ \phi_2(x) = x - 3\\ \phi_3(x) = (x-3)^2\end{cases} \implies \begin{aligned} p(x) &= \alpha_1\,\phi_1(x) + \alpha_2\,\phi_2(x) + \alpha_3\,\phi_3(x) =\\ &= \alpha_1 + \alpha_2\,(x-3) + \alpha_3\,(x-3)^2 =\\ &= \alpha_1 + \alpha_2\,(x-3) + \alpha_3\,\left(x^2 - 6x + 9\right) =\\ &= (\alpha_1 - 3\alpha_2 + 9\alpha_3) + (\alpha_2 - 6\alpha_3)\,x + \alpha_3 x^2\\ &= c_1 + c_2\,x + c_3\,x^2 \end{aligned} \iff \begin{cases} c_1 = \alpha_1 - 3\alpha_2 + 9\alpha_3\\ c_2 = \alpha_2 - 6\alpha_3\\ c_3 = \alpha_3\end{cases}$$

Samma typ av funktion ty samma beroende av x! Bestäm α_i med MKV och beräkna sedan därur de ursprungliga parametrarna c_i.

$$A = \begin{pmatrix}|&|&|\\ \phi_1(x) & \phi_2(x) & \phi_3(x)\\ |&|&|\end{pmatrix} = \begin{pmatrix}|&|&|\\ 1 & x-3 & (x-3)^2\\ |&|&|\end{pmatrix} = \begin{pmatrix}1 & 1-3 & (1-3)^2\\ 1 & 2-3 & (2-3)^2\\ 1 & 4-3 & (4-3)^2\\ 1 & 5-3 & (5-3)^2\end{pmatrix} = \begin{pmatrix}1&-2&4\\1&-1&1\\1&1&1\\1&2&4\end{pmatrix}$$

$$A^TA = \begin{pmatrix}4&0&10\\0&10&0\\10&0&34\end{pmatrix}\ A^Tb = \begin{pmatrix}17\\13\\41\end{pmatrix} \Rightarrow \alpha = \begin{pmatrix}4.6667\\1.3000\\-0.1667\end{pmatrix} \Rightarrow r = b - A\alpha = \begin{pmatrix}-0.4\\0.8\\-0.8\\0.4\end{pmatrix} \Rightarrow A^Tr = 10^{-14}\begin{pmatrix}-0.0444\\-0.0888\\-0.7105\end{pmatrix}$$

$$\begin{cases} c_1 = \alpha_1 - 3\alpha_2 + 9\alpha_3 = -0.7333 \\ c_2 = \alpha_2 - 6\alpha_3 = 2.3000 \\ c_3 = \alpha_3 = -0.1667 \end{cases} \qquad \begin{aligned} \kappa(A^TA_{\varphi}) &= 9.24 \cdot 10^3 = 10^{3.97} \\ \kappa(A^TA_{\phi}) &= 3.81 \cdot 10^1 = 10^{1.58} \end{aligned}$$

Vilka omskrivningar är tillåtna?

Omskrivningen skall resultera i en funktion som har samma beroende av den oberoende variabeln som den ursprungliga. De ”nya” parametrarna skall vara konstanter och inte bero av någon variabel.

$$\text{Ex1}: \quad \begin{aligned} y = c_1 + c_2\,x + c_3\,x^2 &= \alpha_1 + \alpha_2\,(x-m) + \alpha_3\,(x-m)^2 \\ &= \alpha_1 + \alpha_2\,(x-m) + \alpha_3\left(x^2 - 2mx + m^2\right) \\ &= \left\{\alpha_1 - m\alpha_2 + m^2\alpha_3\right\} + \{\alpha_2 - 2m\alpha_3\}\,x + \{\alpha_3\}\,x^2 \end{aligned} \quad \Rightarrow \begin{cases} c_1 = \alpha_1 - m\alpha_2 + m^2\alpha_3 \\ c_2 = \alpha_2 - 2m\alpha_3 \\ c_3 = \alpha_3 \end{cases}$$

$$\text{Ex2}: \quad \begin{aligned} y &= c_1 + c_2\cos(x + c_3) = c_1 + c_2\,(\cos x \cos c_3 - \sin x \sin c_3) \\ &= c_1 + \{c_2 \cos c_3\}\cos x + \{-c_2 \sin c_3\}\sin x \\ &= \alpha_1 + \alpha_2 \cos x + \alpha_3 \sin x \end{aligned} \quad \Longrightarrow \begin{cases} \alpha_1 = c_1 \\ \alpha_2 = c_2 \cos c_3 \\ \alpha_3 = -c_2 \sin c_3 \end{cases}$$

$$\text{Ex3}: \quad \begin{aligned} y = c_1 + c_2\,x^2 &\neq \alpha_1 + \alpha_2\,(x-m)^2 \\ &= \alpha_1 + \alpha_2\left(x^2 - 2mx + m^2\right) \\ &= \left\{\alpha_1 + m^2\alpha_2\right\} + \{\alpha_2\}\,x^2 + \{-2m\alpha_2\}\,x \quad \Longrightarrow \quad \{x\text{−termen är fel!}\} \end{aligned}$$

$$\text{Ex4}: \qquad y = c_1 + \frac{c_2}{x} = \alpha_1 + \frac{\alpha_2}{x-m} \qquad \text{OK?}$$

Ickelinjär modell, (GKN 4.2C-1, sid 153-154)

I modellen $f(x) = c_1\,\mathrm{e}^{c_2\,x}$ förekommer inte de sökta parametrarna c_i linjärt. Ett sätt att hitta en MKV-lösning är då att linearisera modellen:

$$y = c_1\,\mathrm{e}^{c_2\,x} \quad \Longrightarrow \quad \ln y = \ln\left(c_1\,\mathrm{e}^{c_2\,x}\right) = \ln c_1 + \ln\left(\mathrm{e}^{c_2\,x}\right) = \ln c_1 + c_2\,x = \alpha_1 + \alpha_2\,x$$

Sökta parametrar blir då α_i, basfunktionerna är $\phi_1(x) = 1$ och $\phi_2(x) = x$ och högerledet är $b = \ln y$.

x	y
1	1
2	4
4	5
5	7

$$A = \begin{pmatrix} | & | \\ \phi_1(x) & \phi_2(x) \\ | & | \end{pmatrix} = \begin{pmatrix} | & | \\ 1 & x \\ | & | \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 4 \\ 1 & 5 \end{pmatrix} \qquad b = \begin{pmatrix} \ln 1 \\ \ln 4 \\ \ln 5 \\ \ln 7 \end{pmatrix} \quad \Rightarrow \quad A^TA = \begin{pmatrix} 4 & 12 \\ 12 & 46 \end{pmatrix}$$

$$A^Tb = \begin{pmatrix} 4.9416 \\ 18.9399 \end{pmatrix} \Longrightarrow \alpha = \left(A^TA\right)^{-1} A^Tb = \begin{pmatrix} 9.2145 \cdot 10^{-4} \\ 4.1150 \cdot 10^{-1} \end{pmatrix} \quad \begin{cases} \alpha_1 = \ln c_1 \\ \alpha_2 = c_2 \end{cases} \Rightarrow \begin{cases} c_1 = \mathrm{e}^{\alpha_1} = 1.0009 \\ c_2 = \alpha_2 = 0.4115 \end{cases}$$

Vi har nu egentligen två olika sätt att beräkna residualen. Dels i det lineariserade problemet, dels i det ursprungliga. Den residual vi har minimerat är den i det lineariserade problemet.

$$r_\ell = b - A\,\alpha = \{\ln y\} - \{\alpha_1 + \alpha_2\,x\} = \begin{pmatrix} \ln 1 \\ \ln 4 \\ \ln 5 \\ \ln 7 \end{pmatrix} - \begin{pmatrix} 0.4124 \\ 0.8239 \\ 1.6469 \\ 2.0584 \end{pmatrix} = \begin{pmatrix} -0.4124 \\ 0.5624 \\ -0.0375 \\ -0.1125 \end{pmatrix} \qquad ||r_\ell||_2 = 0.7074$$

$$r_u = y - c_1\,\mathrm{e}^{c_2\,x} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} - \begin{pmatrix} c_1\,\mathrm{e}^{c_2\,x_1} \\ c_1\,\mathrm{e}^{c_2\,x_2} \\ c_1\,\mathrm{e}^{c_2\,x_3} \\ c_1\,\mathrm{e}^{c_2\,x_4} \end{pmatrix} = \begin{pmatrix} 1 \\ 4 \\ 5 \\ 7 \end{pmatrix} - \begin{pmatrix} 1.5105 \\ 2.2794 \\ 5.1909 \\ 7.8335 \end{pmatrix} = \begin{pmatrix} -0.5105 \\ 1.7206 \\ -0.1909 \\ -0.8335 \end{pmatrix} \qquad ||r_u||_2 = 1.9880$$

Att veta **hur** de skiljer behöver är lite överkurs men det kan ju ändå vara bra att veta **att** de skiljer. Hur man gör för att minimera residualen i det ursprungliga problemet lär vi oss på föreläsning 10 som handlar om ickelinjära ekvationssystem. (Gauss-Newtons metod som vi lär oss då ger oss $c_1 = 1.5381$ och $c_2 = 0.3086$ vilket ger $||r_u||_2 = 1.6103$, något bättre!)

Utnyttja residualen!

Residualvektorn är skillnaden mellan det önskade högerledet och det som vår modell ger. Det betyder att om vi kan "se en funktion" i residualen så antyder det att vi har glömt att ta med den i modellen och att vår modell skulle stämma bättre överens med mätdata om vi tog med den nya funktionen i modellen.