1.3 Lösningar till kapitel 3

1: Ett linjärt minstakvadratproblem har alltid minst en lösning (om $\mathbf{A}$ har linjärt beroende kolonner så finns oändligt många). Följande linjära ekvationssystem saknar lösning (de två sista ekvationerna kan inte satisfieras):

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 1 \end{bmatrix} \mathbf{x} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

För att lösa minstakvadratproblemet ställer vi upp normalekvationerna:

$$\underbrace{\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}}_{\mathbf{A}^T} \underbrace{\begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 1 \end{bmatrix}}_{\mathbf{A}} \mathbf{x} = \underbrace{\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}}_{\mathbf{A}^T} \underbrace{\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}}_{\mathbf{b}}$$

eller

$$\begin{bmatrix} 3 & 2 \\ 2 & 2 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

som har den entydigt bestämda lösningen $\mathbf{x} = [-1, 1.5]^T$. Vi kan, på skoj, gå vidare och räkna ut residualvektorn:

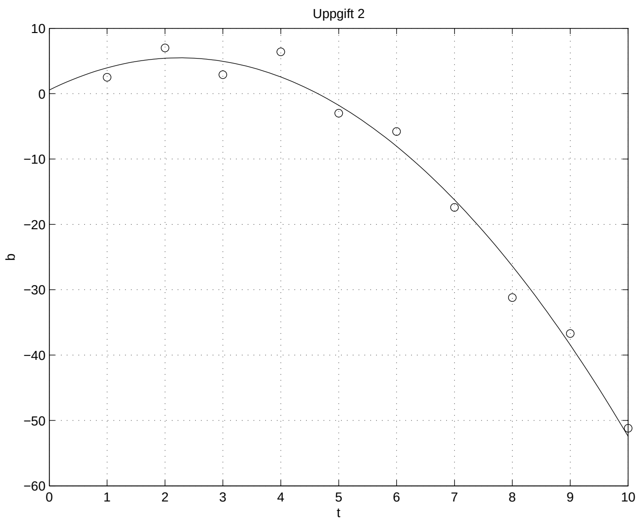
$$\mathbf{r} = \underbrace{\begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 1 \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} -1 \\ 1.5 \end{bmatrix}}_{\mathbf{x}} - \underbrace{\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}}_{\mathbf{b}} = \begin{bmatrix} 0 \\ 0.5 \\ -0.5 \end{bmatrix}$$

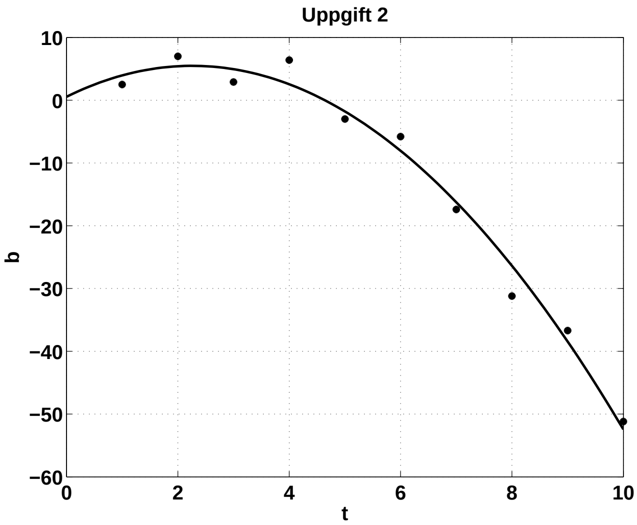
Vi noterar att $\mathbf{r}$ är ortogonal mot båda kolonnerna i $\mathbf{A}$ och därmed ortogonal mot alla linjärkombinationer av kolonnerna (bildrummet).

2:

```
b = [ 2.5 7.0 2.9 6.4 -3.0 -5.8 -17.4 -31.2 -36.7 -51.2 ]';
t = (1:10)';
A = [ones(10, 1) t t.^2];

x = A \ b;         % Minimera || A x - b ||_2 över x
r = b - A * x;     % beräkna residualvektorn
norm(r)            % blir  7.7023
norm(b)            % blir 73.4288

tfin = 0:0.1:10;   % t-värden för att plotta den heldragna kurvan

figure(1)          % aktivera figurfönster ett
hold off           % "lösgör" eventuella kurvor
plot(t, b, 'o')    % mätpunkterna
hold on

% plotta kurvan
plot(tfin, x(1) + x(2) * tfin + x(3) * tfin.^2)
xlabel('t')
ylabel('b')
title('Uppgift 2')
grid on
```

```
% Ett alternativ

x    = polyfit(t, b, 2);  % OBS annan ordning på x (blir RADvektor):
                          % x(1)*t^2 + x(2)*t + x(3)

tfin = linspace(0, 10);   % ger 100 punkter
figure(1)
hold off
plot(t, b, 'b*', tfin, polyval(x, tfin), 'r', 'Linewidth', 2)
xlabel('t', 'Fontsize', 16, 'Fontweight', 'Bold')
ylabel('b', 'Fontsize', 16, 'Fontweight', 'Bold')
title('Uppgift 2', 'Fontsize', 16, 'Fontweight', 'Bold')
set(gca, 'Fontsize', 16, 'Fontweight', 'Bold')
grid on
```





3: Om $\mathbf{A}$ har ortogonala kolonner så är $\mathbf{A}^T\mathbf{A} = \mathbf{D}$, där $\mathbf{D} = \text{diag}(\mathbf{a}_1^T\mathbf{a}_1, \mathbf{a}_2^T\mathbf{a}_2, \ldots, \mathbf{a}_n^T\mathbf{a}_n)$. Antag att ingen kolonn har längden noll, då är matrisen $\mathbf{D}$ ickesingulär med invers $\mathbf{D}^{-1} = \text{diag}(1/||\mathbf{a}_1||_2^2, 1/||\mathbf{a}_2||_2^2, \ldots, 1/||\mathbf{a}_n||_2^2)$. Lösningen till minstakvadratproblemet kan skrivas (normalekvationerna) $\mathbf{x} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b}$ (ty om $\mathbf{A}$ har ortogonala kolonner och inga nollkolonner så måste den ha full kolonnrang. Varför?) Så, $\mathbf{x} = \mathbf{D}^{-1}\mathbf{A}^T\mathbf{b}$ eller $x_k = \mathbf{a}_k^T\mathbf{b}/||\mathbf{a}_k||_2^2$.

4: Problemet är att N_0 och λ ej ingår **linjärt** i modellen varför vi inte kan använda $\min_{\mathbf{x}} ||\mathbf{A}\mathbf{x} - \mathbf{b}||_2$ direkt. Idé: logaritmera

$$\log N(t) = \underbrace{\log N_0}_{x_1} - \underbrace{\lambda}_{x_2} t$$

Vi får då ett linjärt problem i de nya variablerna x_1 och x_2.

$$\min_{x_1,x_2} \left\| \underbrace{\begin{bmatrix} 1 & -t_1 \\ 1 & -t_2 \\ \vdots & \vdots \\ 1 & -t_n \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}}_{\mathbf{x}} - \underbrace{\begin{bmatrix} \log N_1 \\ \log N_2 \\ \vdots \\ \log N_n \end{bmatrix}}_{\mathbf{b}} \right\|_2$$

Vi sätter sedan $N_0 = e^{x_1}$ och $\lambda = x_2$.

5: Vi börjar med tvånormen. $\mathbf{A}$ och $\mathbf{b}$ har $2n+1$ rader och ser ut som:

$$\mathbf{A} = \begin{bmatrix} 1 & -n \\ 1 & -(n-1) \\ \vdots & \vdots \\ 1 & -1 \\ 1 & 0 \\ \vdots & \vdots \\ 1 & n \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} \quad \leftarrow \text{rad } n+1$$

Vi noterar att vi kan utnyttja föregående uppgift eftersom $\mathbf{A}$ har ortogonala kolonner. Lösningen blir alltså

$$\mathbf{x} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b} = \begin{bmatrix} 2n+1 & 0 \\ 0 & 2n(n+1)(2n+1)/6 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/(2n+1) \\ 0 \end{bmatrix}$$

Den räta linjen ges alltså av $f(t) = 1/(2n+1)$ (konstant oberoende av t). Nu till maxnormen. Man gissar att $f(t) = 1/2$ ty då är $|1 - f(t)| = |0 - f(t)|$. Det måste gälla att $f(0) = 1/2$ ty annars är $|1 - f(0)| > 1/2$ eller $|f(0)| > 1/2$. Derivatan måste vidare vara noll annars är någon av $|f(1)|$ eller $|f(-1)|$ större än $1/2$. Slutligen ettnormen. Man gissar att $f(t) = 0$, som ger normen ett. **Om** $f(t) \equiv \delta$ så är $||\mathbf{Ax} - \mathbf{b}||_1 = 2n\delta + 1 - \delta > 1$. **Om** $f'(t) \neq 0$ så blir $|f(-1)| + |f(1)| > 1$.
Slutsatsen är att vi får tre parallella linjer och att maxnormslinjen påverkas mest av utliggaren. Ettnormslinjen påverkas minst och tvånormslinjen ligger emellan de andra två.

6: Vi vet att $\mathbf{r}$ skall vara ortogonal mot $\mathbf{A}$:s bildrum. Endast c-vektorn är det.

7: Vi kollar: $\mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} = (\mathbf{A}\mathbf{x})^T(\mathbf{A}\mathbf{x}) = ||\mathbf{A}\mathbf{x}||_2^2$. Det sista uttrycket kan inte vara negativt (det är ju en norm). Så frågan är om det kan vara noll även om $\mathbf{x} \neq \mathbf{0}$. $||\mathbf{A}\mathbf{x}||_2 = 0 \Rightarrow \mathbf{A}\mathbf{x} = \mathbf{0}$ vilket inte kan inträffa eftersom $\mathbf{A}$ har full kolonnrang enligt förutsättningarna.

8: a) Antag att $\mathbf{A}$ är reell och undertriangulär. Eftersom $\mathbf{A}$ dessutom är ortogonal gäller att $\mathbf{A}\mathbf{A}^T = \mathbf{I}$. Innerprodukterna mellan första raden i $\mathbf{A}$ och kolonnerna i $\mathbf{A}^T$ blir $a_{1.1}a_{1,1}$, $a_{1,1}a_{2,1}$, $a_{1,1}a_{3,1}, \ldots, a_{1,1}a_{n,1}$ eftersom första raden i $\mathbf{A}$ endast innehåller ett element skilt från noll (nämligen $a_{1,1}$). Men eftersom första raden i enhetsmatrisen har nollor utom i första positionen måste $a_{2,1} = a_{3,1} \ldots a_{n,1} = 0$. Vi kan nu utnyttja induktion och upprepa resonemanget för andra kolonnen i $\mathbf{A}$ osv.
b) Det måste gälla att $a_{k,k}^2 = 1$ så att $a_{k,k} = \pm 1$.

9: Detta är en form av generalisering av föregående övning. Vi kollar:

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{0} & \mathbf{C} \end{bmatrix}^T \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{0} & \mathbf{C} \end{bmatrix} = \begin{bmatrix} \mathbf{A}^T & \mathbf{0} \\ \mathbf{B}^T & \mathbf{C}^T \end{bmatrix} \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{0} & \mathbf{C} \end{bmatrix} = \begin{bmatrix} \mathbf{A}^T\mathbf{A} & \mathbf{A}^T\mathbf{B} \\ \mathbf{B}^T\mathbf{A} & \mathbf{B}^T\mathbf{B} + \mathbf{C}^T\mathbf{C} \end{bmatrix}$$

---

vilket skall vara lika med enhetsmatrisen. Detta innebär att $\mathbf{A}^T\mathbf{A} = \mathbf{I}$, så att $\mathbf{A}$ är ortogonal. $\mathbf{A}^T\mathbf{B}$ skall vara nollmatrisen vilket, eftersom $\mathbf{A}$ är ortogonal och därmed ickesingulär, medför att $\mathbf{B} = \mathbf{0}$. Detta medför slutligen att $\mathbf{C}$ är ortogonal, ty $\mathbf{B}^T\mathbf{B} + \mathbf{C}^T\mathbf{C} = \mathbf{I}$.

10: a) Låt oss numrera likheterna 1), 2) och 3). 1)+2) $\Rightarrow$ 3) eftersom 2) övergår i 3) om vi använder symmetrin. 1) + 3) $\Rightarrow$ 2) eftersom vi kan byta ut det första $\mathbf{A}$:t i 3) med $\mathbf{A}^T$. 2) + 3) $\Rightarrow$ 1) ty 2) medför att $\mathbf{A}$ är ickesingulär. Därmed gäller (multiplicera med $\mathbf{A}^{-1}$ från höger) att $\mathbf{A}^T = \mathbf{A}^{-1}$. Gör vi samma operation med 3) får vi $\mathbf{A} = \mathbf{A}^{-1}$. dessa tillsammans medför 1).
b) Ett exempel ges av

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & \sin\phi \\ 0 & \sin\phi & -\cos\phi \end{bmatrix}$$

11: Matrisen är en så kallad speglingsmatris eller Householdermatris (efter den kände numerikern Alston S. Householder). $\mathbf{H}$ används för att beräkna QR-faktoriseringen på ett stabilt sätt. Sätt $\gamma = 2/(\mathbf{v}^T\mathbf{v})$, då är $\mathbf{H} = \mathbf{I} - \gamma\mathbf{v}\mathbf{v}^T$. Symmetri? $\mathbf{H}^T = (\mathbf{I} - \gamma\mathbf{v}\mathbf{v}^T)^T = \mathbf{I}^T - (\gamma\mathbf{v}\mathbf{v}^T)^T = \mathbf{I} - \gamma(\mathbf{v}^T)^T\mathbf{v}^T = \mathbf{I} - \gamma\mathbf{v}\mathbf{v}^T = \mathbf{H}$. Ortogonal? $\mathbf{H}^T\mathbf{H} = \mathbf{H}^2 = (\mathbf{I}-\gamma\mathbf{v}\mathbf{v}^T)(\mathbf{I}-\gamma\mathbf{v}\mathbf{v}^T) = \mathbf{I}-2\gamma\mathbf{v}\mathbf{v}^T + \gamma^2\mathbf{v}\mathbf{v}^T\mathbf{v}\mathbf{v}^T$. Notera att $\mathbf{v}^T\mathbf{v}$ är en skalär och att $\gamma^2\mathbf{v}^T\mathbf{v} = 2\gamma$. Alltså $\mathbf{H}^T\mathbf{H} = \mathbf{I} - 2\gamma\mathbf{v}\mathbf{v}^T + 2\gamma\mathbf{v}\mathbf{v}^T = \mathbf{I}$.

12: Vi sätter upp normalekvationerna och får

$$\begin{bmatrix} 1 & 0 \\ 0 & \sigma^2 \end{bmatrix}\mathbf{x} = \begin{bmatrix} b_1 \\ \sigma b_2 \end{bmatrix} \quad \Rightarrow \quad \mathbf{x} = \begin{bmatrix} b_1 \\ b_2/\sigma \end{bmatrix}$$

Vi sätter nu upp normalekvationerna för det störda problemet, dvs. $(\mathbf{A}+\mathbf{E})^T(\mathbf{A}+\mathbf{E})\mathbf{y} = (\mathbf{A}+\mathbf{E})^T\mathbf{b}$. Så

$$\begin{bmatrix} 1 & 0 \\ 0 & \sigma^2+\epsilon^2 \end{bmatrix}\mathbf{y} = \begin{bmatrix} b_1 \\ \sigma b_2 + \epsilon b_3 \end{bmatrix} \quad \Rightarrow \quad \mathbf{y} = \begin{bmatrix} b_1 \\ \frac{\sigma b_2+\epsilon b_3}{\sigma^2+\epsilon^2} \end{bmatrix} \approx \begin{bmatrix} b_1 \\ \frac{\sigma b_2+\epsilon b_3}{\sigma^2} \end{bmatrix}$$

Så

$$\mathbf{y} - \mathbf{x} \approx \frac{\epsilon}{\sigma^2}\begin{bmatrix} 0 \\ b_3 \end{bmatrix}$$

vilket man kan visa är lika med

$$||\mathbf{E}||_2\kappa_2^2(\mathbf{A})\begin{bmatrix} 0 \\ b_3 \end{bmatrix}$$

Notera kvadraten på konditionstalet. b_3 är den del av $\mathbf{b}$ som inte tillhör matrisen $\mathbf{A}$:s bildrum.
Nu till den svårare delen. Eftersom $\text{rang}(\mathbf{A}) = 2$ i vårt exempel vill vi hitta $\mathbf{F}$ så att $\text{rang}(\mathbf{A}+\mathbf{F}) = 1$. Det medför att det existerar $\mathbf{z} \neq \mathbf{0}$ så att $(\mathbf{A}+\mathbf{F})\mathbf{z} = \mathbf{0}$, dvs. så att $\mathbf{A}\mathbf{z} = -\mathbf{F}\mathbf{z}$. Tag normer, $||\mathbf{A}\mathbf{z}||_2 = ||-\mathbf{F}\mathbf{z}||_2 \leq ||\mathbf{F}||_2||\mathbf{z}||_2$. Så $||\mathbf{F}||_2 \geq ||\mathbf{A}\mathbf{z}||_2/||\mathbf{z}||_2$. Men för varje tänkbar $\mathbf{z}$ gäller att

$$||\mathbf{F}||_2^2 \geq \frac{||\mathbf{A}\mathbf{z}||_2^2}{||\mathbf{z}||_2^2} = \frac{z_1^2 + \sigma^2 z_2^2}{z_1^2 + z_2^2} \geq \sigma^2$$

Så varje tänkbar $\mathbf{F}$ måste uppfylla $||\mathbf{F}||_2 \geq \sigma$. Vi testar nu med den speciella $\mathbf{F}$ som är noll förutom $(2,2)$-elementet som är $-\sigma$. Med detta val är $\text{rang}(\mathbf{A}+\mathbf{F}) = 1$ och $||\mathbf{F}||_2 = \sigma$. Eftersom $||\mathbf{A}||_2 = 1$ är då $\kappa_2(\mathbf{A}) = 1/\sigma$.

13: Ickelinjärt problem:

$$\min_\alpha \phi(\alpha), \quad \phi(\alpha) = (e^{\alpha t_1} - b_1)^2 + (e^{\alpha t_2} - b_2)^2$$

Vi deriverar; $\phi'(\alpha) = t_1 e^{\alpha t_1}(e^{\alpha t_1} - b_1) + t_2 e^{\alpha t_2}(e^{\alpha t_2} - b_2)$. Vi vill lösa $\phi'(\alpha) = 0$ då $t_1 = 1$, $t_2 = 2$ och $b_2 = 1/2$ och får ekvationen: $0 = e^\alpha(e^\alpha - b_1) + 2e^{2\alpha}(e^{2\alpha} - 1/2) = e^{2\alpha} - b_1 e^\alpha + 2e^{4\alpha} - e^{2\alpha} = 2e^{4\alpha} - b_1 e^\alpha$. Så att,

$\alpha = (\log b_1 - \log 2)/3$. För att få ett linjärt problem logaritmerar vi (antar att $b_1 > 0$) och får $\alpha t_k \approx \log b_k$. Minstakvadratproblemet blir därför

$$\min_{\alpha} \, (t_1\alpha - \log b_1)^2 + (t_2\alpha - \log b_2)^2$$

varför (via normalekvationerna eller derivering): $\alpha = (t_1 \log b_1 + t_2 \log b_2)/(t_1^2 + t_2^2) = (\log b_1 - \log 4)/5$. Sammafattningsvis får vi de två α-värdena:

$$\alpha_{\text{ickelin}} = \frac{\log b_1 - \log 2}{3}, \quad \text{och} \quad \alpha_{\text{lin}} = \frac{\log b_1 - \log 4}{5}$$

14: Eftersom $fl(1+\epsilon^2) = 1$ så blir $\mathbf{A}^T\mathbf{A}$ en matris av idel ettor. Den exakta QR-faktoriseringen ges av följande uttryck (med $\omega = \sqrt{1+\epsilon^2}$).

$$\mathbf{q}_1 = \begin{bmatrix} 1 \\ \epsilon \\ 0 \end{bmatrix} / r_{1,1}, \quad \mathbf{q}_2 = \epsilon \begin{bmatrix} \epsilon \\ -1/\omega^2 \\ 1 \end{bmatrix} / r_{2,2}, \quad \mathbf{R} = \begin{bmatrix} \omega & 1/\omega \\ 0 & (\epsilon/\omega)\sqrt{1+\omega^2} \end{bmatrix}$$

Med flyttalsräkning får vi: $r_{1,1} = fl(||\mathbf{a}_1||_2) = 1$ så att $fl(\mathbf{q}_1) = \mathbf{a}_1$. $r_{1,2} = fl(\mathbf{q}_1^T\mathbf{a}_2) = 1$ (pga blandningen av ettor och nollor). $\mathbf{t} = fl(\mathbf{a}_2 - r_{1,2}\mathbf{q}_1) = [0, -\epsilon, \epsilon]^T$. $r_{2,2} = fl(||\mathbf{t}||_2) \approx \epsilon\sqrt{2}$. $\mathbf{q}_2$ slutligen blir nära $[0, -1/\sqrt{2}, 1/\sqrt{2}]^T$. Så här ser det ut i Matlab:

```
% QR
>> e = 2^-27
e = 7.450580596923828e-09

>> 1 + e*e - 1 % Kontroll
ans = 0

>> A = [1 e 0; 1 0 e]'
A =  1.000000000000000e+00     1.000000000000000e+00
     1.490116119384766e-08                         0
                         0     1.490116119384766e-08


>> R(1, 1) = norm(A(:, 1));
>> Q(:, 1) = A(:, 1) / R(1, 1);
>> R(1, 2) = Q(:, 1)' * A(:, 2);
>> t = A(:, 2) - Q(:, 1) * R(1, 2)
t =  0
    -1.490116119384766e-08
     1.490116119384766e-08

>> R(2, 2) = norm(t);
>> Q(:, 2) = t / R(2, 2)

Q = 1.000000000000000e+00                         0
    1.490116119384766e-08    -7.071067811865475e-01
                        0     7.071067811865475e-01
>> R
R = 1.000000000000000e+00     1.000000000000000e+00
                        0     2.107342425544702e-08
```

```
% Normalekvationerna går inte att använda här.
>> A' * A
ans =
     1     1   % Notera att matrisen är singulär
     1     1
```