Technische Universität München
Institut für Informatik
Prof. Dr. Daniel Cremers
Dr. Rudolph Triebel
Nikola Tchipev, M.Sc.
Sebastian Rettenberger, M.Sc.

WiSe 2014/2015

Numerisches Programmieren, Übungen

Musterlösung 4. Übungsblatt: Lineares Ausgleichsproblem, Regularisierung

1) Ausgleichsgerade

a) Jeder Punkt (x_i, y_i) in die Gleichung $y = mx + t$ eingesetzt entspricht:

$$\underbrace{\begin{pmatrix} 1 & 2 \\ 1 & 0 \\ 1 & 1 \end{pmatrix}}_{=A} \underbrace{\begin{pmatrix} t \\ m \end{pmatrix}}_{=z} = \underbrace{\begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}}_{=b}$$

b)

$$A^\top A x = A^\top b$$

$$\Rightarrow \begin{pmatrix} 1 & 1 & 1 \\ 2 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} t \\ m \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} 3 & 3 \\ 3 & 5 \end{pmatrix} \begin{pmatrix} t \\ m \end{pmatrix} = \begin{pmatrix} 6 \\ 7 \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} 3 & 3 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} t \\ m \end{pmatrix} = \begin{pmatrix} 6 \\ 1 \end{pmatrix} \Rightarrow \begin{pmatrix} t \\ m \end{pmatrix} = \begin{pmatrix} 1.5 \\ 0.5 \end{pmatrix}$$

$y = 0.5x + 1.5$

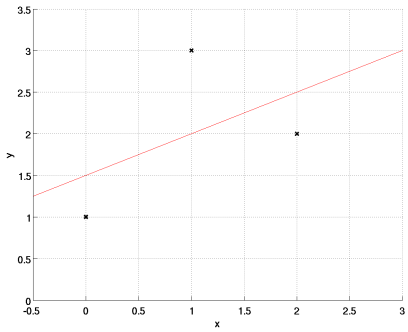

Abbildung 1: lineare Ausgleichsgerade

2) Lineares Ausgleichsproblem und QR-Zerlegung

a) Um Element (2,1) zu Null zu drehen, müssen wir Element (3,1) unverändert lassen und (1,1) als Diagonalelement wählen (also $a = A_{21} = 1$ und $d = A_{11} = 1$). Damit brauchen wir die Elementarmatrix $G_{2,1}$. Die konkreten Einträge ergeben sich zu

$$c = 1/\sqrt{1^2 + 1^2} = 1/\sqrt{2}, \quad s = 1/\sqrt{1^2 + 1^2} = 1/\sqrt{2}.$$

Damit erhalten wir:

$$\begin{aligned} G_{2,1} &= \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ \Rightarrow \quad A_1 := G_{2,1} \cdot A &= \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 1 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} \sqrt{2} & \sqrt{2} \\ 0 & -\sqrt{2} \\ 1 & 1 \end{pmatrix} \end{aligned}$$

Analoges Vorgehen für Eintrag (3,1) von A_1 mit Elementarmatrix $G_{3,1}$ und

$$c = \sqrt{2}/\sqrt{2+1} = \sqrt{2/3}, \quad s = 1/\sqrt{2+1} = \sqrt{1/3}$$

führt zu

$$\begin{aligned} G_{3,1} &= \begin{pmatrix} \sqrt{2/3} & 0 & \sqrt{1/3} \\ 0 & 1 & 0 \\ -\sqrt{1/3} & 0 & \sqrt{2/3} \end{pmatrix} \\ \Rightarrow \quad Q^T A := G_{3,1} \cdot A_1 &= \begin{pmatrix} \sqrt{2/3} & 0 & \sqrt{1/3} \\ 0 & 1 & 0 \\ -\sqrt{1/3} & 0 & \sqrt{2/3} \end{pmatrix} \cdot \begin{pmatrix} \sqrt{2} & \sqrt{2} \\ 0 & -\sqrt{2} \\ 1 & 1 \end{pmatrix} \\ &= \left(\begin{array}{cc} \sqrt{3} & \sqrt{3} \\ 0 & -\sqrt{2} \\ \hline 0 & 0 \end{array}\right) \end{aligned}$$

Normalerweise müsste man hier nun eine dritte Drehung $G_{3,2}$ anwenden, aber in unserem Fall entstand schon bei der vorehrigen Drehung zufällig eine 0 im Feld $(3, 2)$.

b) Es gilt für die QR-Zerlegung die Rücksubstitution: $Rz = \beta_1$. Dabei ist β_1 Teil des Vektors $Q^T b$, wobei Q^T die Hintereinanderausführung der Givensrotationen (zur Erzeugung von R) darstellt. In unserem Fall bedeutet das:

$$
\begin{aligned}
b_1 &:= G_{2,1} \cdot b = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ -1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 3/\sqrt{2} \\ -1/\sqrt{2} \\ 3 \end{pmatrix} \\
\begin{pmatrix} \beta_1 \\ \hline \beta_2 \end{pmatrix} &:= G_{3,1} \cdot b_1 = \begin{pmatrix} \sqrt{2/3} & 0 & \sqrt{1/3} \\ 0 & 1 & 0 \\ -\sqrt{1/3} & 0 & \sqrt{2/3} \end{pmatrix} \begin{pmatrix} 3/\sqrt{2} \\ -1/\sqrt{2} \\ 3 \end{pmatrix} \\
&= \begin{pmatrix} 2\sqrt{3} \\ -1/\sqrt{2} \\ \hline \sqrt{3/2} \end{pmatrix}
\end{aligned}
$$

Damit muss nur mehr $Rz = \beta_1$ per Rückwärtssubstitution gelöst werden:

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

$$
\Rightarrow m = 1/2, \qquad \Rightarrow \quad t = \left(2\sqrt{3} - \frac{\sqrt{3}}{2}\right) \cdot \frac{1}{\sqrt{3}} = 3/2
$$

c) Für den ersten Beweis nutzen wir die folgende besondere Eigenschaft orthogonaler Matrizen
$Q \in \mathbb{R}^{n \times n}$ aus:

$$
||Qx||_2 = ||x||_2 \; \forall x \in \mathbb{R}^n.
$$

Da Q^T ebenfalls orthogonal ist, ergibt sich:

$$
\begin{aligned}
||b - Az||_2^2 &= ||Q^T(b - Az)||_2^2 = \left\| \begin{pmatrix} \beta_1 - Rz \\ \beta_2 \end{pmatrix} \right\|_2^2 \\
&= ||\beta_1 - Rz||_2^2 + ||\beta_2||_2^2 \geq ||\beta_2||_2^2
\end{aligned}
$$

Daher ist $||b - Az||_2$ minimal genau dann wenn $||\beta_1 - Rz||_2 = 0$, d.h. wenn z Lösung des linearen Gleichungssystems $Rz = \beta_1$ ist.

Für den zweiten Beweis brauchen wir nur die Eigenschaft $Q^\top Q = I$. Es gilt:

$$
\begin{aligned}
A^\top A x &= A^\top b \\
[R^\top|0] Q^\top Q \begin{bmatrix} R \\ \hline \mathbf{0} \end{bmatrix} x &= [R^\top|0] Q^\top b \\
\stackrel{Q^\top Q = I}{\Rightarrow} [R^\top|0] \begin{bmatrix} R \\ \hline \mathbf{0} \end{bmatrix} x &= [R^\top|0] Q^\top b \\
R^\top R x &= R^\top \beta_1 \\
\Rightarrow R x &= \beta_1
\end{aligned}
$$

3) Computertomographie

a)

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

L ist singulär, denn wegen

$$L_{4,k} + L_{5,k} + L_{6,k} = L_{7,k} + L_{8,k} + L_{9,k}, \quad k = 1, \ldots, 9$$

sind die Zeilen der Matrix L linear abhängig. Folglich ist auch $L^T L$ sind singulär.

Auf Grund der Singularität lassen sich sowohl das Gleichungssystem $Lr = v$ als auch das Ausgleichsproblem $L^T L r = L^T v$ nicht eindeutig lösen. Es werden also mehr Strahlen benötigt, um eine eindeutige Lösung zu finden.

b) Zur Bestimmung des Identitätsverlusts v ist lediglich das Produkt $v = Lr$ mit

$$r = (1, 0, 1, 0, 1, 0, 1, 0, 1)^T$$

zu berechnen. Hierfür müssen die ungeraden Spalten der Matrix L aufsummiert werden. Es ergibt sich

$$v = Lr = (0, 3\sqrt{2}, 0, 2, 1, 2, 2, 1, 2)^T.$$

c) Wir führen noch einen Regularisierungsterm ein, der beim Minimierungsproblem verhindert, dass r zu groß wird:

$$\min_r (||Lr - v||_2^2 + \gamma^2 ||r||_2^2)$$

Dies ist äquivalent zum Minimierungsproblem

$$\min_r \left\| \begin{pmatrix} \gamma I_n \\ L \end{pmatrix} r - \begin{pmatrix} 0 \\ v \end{pmatrix} \right\|_2^2,$$

denn

$$\begin{aligned} \left\| \begin{pmatrix} \gamma I_n \\ L \end{pmatrix} r - \begin{pmatrix} 0 \\ v \end{pmatrix} \right\|_2^2 &= ||\gamma I_n r||_2^2 + ||Lr - v||_2^2 \\ &= \gamma^2 ||I_n r||_2^2 + ||Lr - v||_2^2 = \gamma^2 ||r||_2^2 + ||Lr - v||_2^2 \end{aligned}$$

Auf Grund des Terms $I_n r$ ist das regularisierte Ausgleichsproblem nun stets regulär.

d) Wir erweitern die Matrix aus (i) um drei weitere Zeilen:

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

Um das Minimierungsproblem

$$\min_r ||Lr - v||_2$$

zu lösen, lösen wir die Normalengleichung

$$L^T Lr = L^T v.$$

In Zahlen sieht das wie folgt aus:

$$\begin{pmatrix} 4 & 1 & 1 & 1 & 2 & 0 & 1 & 0 & 2 \\ 1 & 6 & 1 & 2 & 1 & 2 & 0 & 1 & 0 \\ 1 & 1 & 4 & 0 & 2 & 1 & 2 & 0 & 1 \\ 1 & 2 & 0 & 6 & 1 & 1 & 1 & 2 & 0 \\ 2 & 1 & 2 & 1 & 6 & 1 & 2 & 1 & 2 \\ 0 & 2 & 1 & 1 & 1 & 6 & 0 & 2 & 1 \\ 1 & 0 & 2 & 1 & 2 & 0 & 4 & 1 & 1 \\ 0 & 1 & 0 & 2 & 1 & 2 & 1 & 6 & 1 \\ 2 & 0 & 1 & 0 & 2 & 1 & 1 & 1 & 4 \end{pmatrix} \cdot \begin{pmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \rho_4 \\ \rho_5 \\ \rho_6 \\ \rho_7 \\ \rho_8 \\ \rho_9 \end{pmatrix} = \begin{pmatrix} 4 \\ 1 \\ 1 \\ 1 \\ 2 \\ 0 \\ 1 \\ 0 \\ 2 \end{pmatrix}$$

Als Lösung ergibt sich

$$r = (1, 0, 0, 0, 0, 0, 0, 0, 0)^T.$$

Anmerkung:
Dazu muss man in unserem Beispiel hier gar nicht viel rechnen, wenn man genau hinschaut. Denn die linke Seite des Gleichungssystems entspricht genau der ersten Spalte der Matrix $L^T L$.

Wiederholung: Lineare Gleichungssysteme

a) (1) falsch (z.B. $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \Rightarrow A^T = A$, aber $A^T A = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix} \neq \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$)

(2) falsch

(3) wahr

(4) falsch ($\frac{2}{3}n^3$)

b) i) Ansatz:

$$f(x_i) = a + bx_i + cx_i^2 = y_i, \; i = 1, \ldots, 4$$

als LGS:

$$\begin{pmatrix} 1 & -2 & 4 \\ 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 3 \\ 7 \end{pmatrix}$$

ii) Um ein überbestimmtes LGS $Ax = b$ in ein Minimierungsproblem zu überführen, nutzen wir die Normalengleichung

$$A^T A x = A^T b,$$

mit

$$x = (a, b, c)^T.$$

Dabei gilt:

$$A^T A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ -2 & -1 & 0 & 1 \\ 4 & 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -2 & 4 \\ 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 4 & -2 & 6 \\ -2 & 6 & -8 \\ 6 & -8 & 18 \end{pmatrix}$$

$$A^T b = \begin{pmatrix} 1 & 1 & 1 & 1 \\ -2 & -1 & 0 & 1 \\ 4 & 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 2 \\ 3 \\ 7 \end{pmatrix} = \begin{pmatrix} 14 \\ 1 \\ 17 \end{pmatrix}$$

(Die Lösung ist $(a, b, c)^T = (3.3, 2.6, 1)^T$)