Numerische Mathematik für Elektrotechniker

Lineare Ausgleichsrechnung

Benjamin Berkels

Karl-Heinz Brakhage, Thomas Jankuhn, Christian Löbbert

Institut für Geometrie und Praktische Mathematik
RWTH Aachen

Wintersemester 2019/2020

Zusammenfassung der letzten Vorlesung

Eine Matrix $Q \in \mathbb{R}^{m \times m}$ heißt orthogonal, falls

$$Q^T Q = I.$$

Eigenschaften:

- $Q^{-1} = Q^T$.
- $\|Qx\|_2 = \|x\|_2$, $\kappa_2(Q) = 1$.
- $\|QA\|_2 = \|A\|_2$, $\kappa_2(QA) = \kappa_2(A)$.

Satz QR-Zerlegung

Zu $A \in \mathbb{R}^{m \times n}$ existiert eine Orthogonale Matrix Q, so dass

$$A = Q\,R,$$

mit R obere Dreiecksmatrix.

Zusammenfassung der letzten Vorlesung

Grundaufgabe

Gegeben sei $(a, b)^T \in \mathbb{R}^2 \setminus \{0\}$. Finde $c, s, r \in \mathbb{R}$ mit

$$\begin{pmatrix} c & s \\ -s & c \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} r \\ 0 \end{pmatrix}$$

und

$$c^2 + s^2 = 1.$$

Die Lösung ist: $r = \pm\sqrt{a^2 + b^2}, \quad c := \frac{a}{r}, \quad s := \frac{b}{r}$

Beachte

Drehung verändert nicht die Euklidische Länge eines Vektors, d.h.

$$\|(r, 0)^T\|_2 = |r| = \sqrt{a^2 + b^2} = \|(a, b)^T\|_2.$$

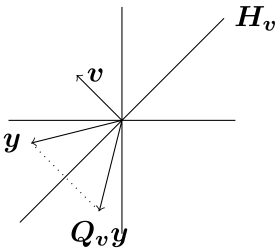
- Die obige (Rotations-)Matrix ist orthogonal.

Zusammenfassung der letzten Vorlesung

Definition

Für $v = (v_1, \dots, v_n)^T \in \mathbb{R}^n \setminus \{0\}$, ist die Householder-Transformation definiert als

$$Q_v = I - 2\,\frac{v\,v^T}{v^T\,v}.$$



$$H_v = \{x \in \mathbb{R}^n \mid x^T v = 0\}.$$

Lösung von $Q_v y = \pm\|y\|_2 e^1$

$$v = y + \alpha\, e^1$$
$$\alpha = \mathrm{sign}(y_1)\, \|y\|_2$$
$$Q_v y = -\alpha\, e^1$$

Heute in der Vorlesung

Themen: Dahmen & Reusken Kap. 4.1-4.4

- Lineare Ausgleichsrechnung
  1. Problemstellung
  2. Kondition
  3. Lösungsverfahren
     - über Normalgleichungen
     - über QR-Zerlegung

Was Sie mitnehmen sollten:

- Was ist ein lineares Ausgleichsproblem?
- Wie ist das lineare Ausgleichsproblem konditioniert?
- Welche Lösungsverfahren gibt es und wie stabil sind diese?

Problemstellung

Notation:

$\mathbb{R}^{m \times n}$ ist die Menge der $(m \times n)$-Matrizen

$$A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{pmatrix}$$

mit Einträgen

$$a_{i,j} \in \mathbb{R}.$$

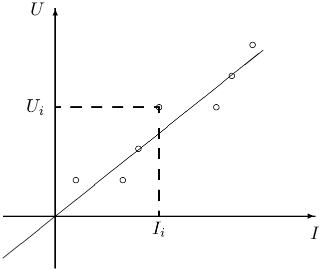
Bestimmung des elektrischen Widerstands (Beispiel 4.1.)

- Ohmsches Gesetz: $U = R\,I$
- Aufgabe: Bestimme Widerstand R im Stromkreis aus einer Reihe von Messungen:

$$(U_i, I_i) \quad \text{(Spannung, Stromstärke)}, \; i = 1, \ldots, m.$$

- Problem: Messungen (Daten) sind mit Fehlern behaftet, d.h.

$$U_i \neq R\,I_i, \; \text{für fast alle } i = 1, \ldots, m.$$



Beispiel 4.1.

Vorgehen:

- Fehler in Messung i (Residuum)

$$r_i = R\,I_i - U_i, \quad i = 1, \ldots, m$$

- Ein Maß für den Gesamtfehler: Summe der Fehlerquadrate

$$f(R) := \sum_{i=1}^{m} r_i^2 = \sum_{i=1}^{m} (R\,I_i - U_i)^2$$

- Bestimme Widerstand R^* so, dass Gesamtfehler minimal wird

$$R^* = \arg\min_R f(R)$$

- Extremum der quadratischen Funktion $f(R)$

$$f'(R^*) = 0 \quad \Rightarrow \quad R^* = \left(\sum_{i=1}^{m} U_i\,I_i\right) \Big/ \left(\sum_{i=1}^{m} I_i^2\right)$$

Fourierapproximation (Beispiel 4.2.)

In der Fourieranalyse wird eine T-periodische Funktion f durch eine Linearkombination der T-periodischen trigonometrischen Polynome

$$1, \cos(ct), \sin(ct), \cos(2ct), \sin(2ct), \ldots, \cos(Nct), \sin(Nct)$$

mit $c := \dfrac{2\pi}{T}$ in der Form

$$g_N(t) = \frac{1}{2}\alpha_0 + \sum_{k=1}^{N} \Big(\alpha_k \cos(kct) + \beta_k \sin(kct) \Big)$$

approximiert.

Beispiel 4.2.

Annahme: nicht f, sondern nur eine Reihe vom Messdaten

$$b_i \approx f(t_i), \quad 0 \leq t_1 < t_2 < \cdots < t_m \leq T,$$

ist bekannt, wobei $m > 2N + 1$.

Ansatz zur Bestimmung der Koeffizienten
α_0, α_1, β_1, α_2, $\beta_2 \ldots, \alpha_N$, β_N:

$$\sum_{i=1}^{m} \Big(g_N(t_i) - b_i \Big)^2 \to \min_{\alpha_k, \beta_k}.$$

g_N ist linear in den Koeffizienten!

Allgemeines lineares Ausgleichsproblem

Definition

Zu gegebenen $A \in \mathbb{R}^{m \times n}$ und $b \in \mathbb{R}^m$, bestimme $x^* \in \mathbb{R}^n$, für dass

$$\|A\,x^* - b\|_2 = \min_{x \in \mathbb{R}^n} \|A\,x - b\|_2$$

gilt. Diese Problemstellung heißt das lineare Ausgleichsproblem.

oder:

Lineares Ausgleichsproblem

Zu gegebenen $A \in \mathbb{R}^{m \times n}$ und $b \in \mathbb{R}^m$, bestimme $x^* \in \mathbb{R}^n$, so dass

$$x^* \in \arg\min_{x \in \mathbb{R}^n} \|A\,x - b\|_2^2.$$

Allgemeines lineares Ausgleichsproblem

Warum 2-Norm?

- $\|Ax - b\|_2^2$ ist differenzierbar und Ableitung ist linear
- Statistischer Hintergrund ("BLUE", erwartungstreu, Kap. 4.5).
- Euklidische Norm bleibt bei orthogonalen Transformationen erhalten, d.h. für jede orthogonale Matrix $Q \in \mathbb{R}^{m \times m}$ ist

$$\min_{x \in \mathbb{R}^n} \|Ax - b\|_2 = \min_{x \in \mathbb{R}^n} \|Q(Ax - b)\|_2$$

$\Rightarrow$ führt auf lineares Optimierungsproblem

Auch möglich:

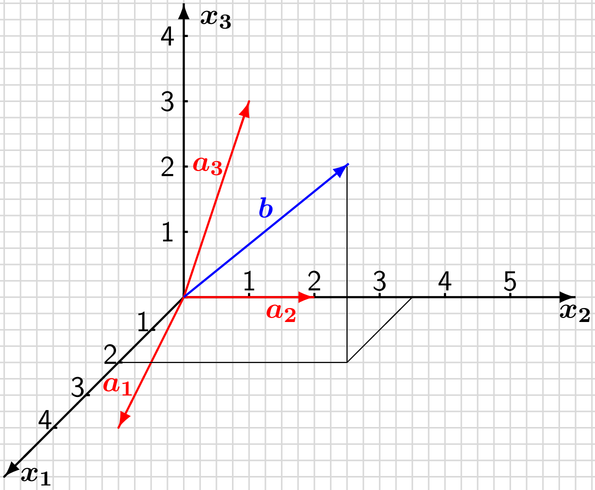
$$\min_{x \in \mathbb{R}^n} \|Ax - b\|_1 \text{ oder } \min_{x \in \mathbb{R}^n} \|Ax - b\|_\infty$$

Geometrische Interpretation $\boldsymbol{Ax = b}$

$$\boldsymbol{A} = \begin{bmatrix} \mathbf{4} & \mathbf{0} & \mathbf{0} \\ \mathbf{1} & \mathbf{2} & \mathbf{1} \\ \mathbf{0} & \mathbf{0} & \mathbf{3} \end{bmatrix}$$

$$\boldsymbol{b} = \begin{bmatrix} \mathbf{2} \\ \mathbf{3.5} \\ \mathbf{3} \end{bmatrix}$$

$$\Rightarrow \boldsymbol{x} = \begin{bmatrix} \mathbf{0.5} \\ \mathbf{1} \\ \mathbf{1} \end{bmatrix}$$

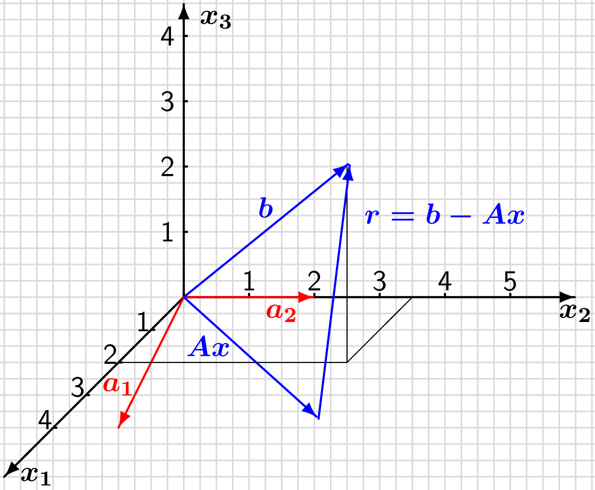


Geometrische Interpretation $\min_{x\in\mathbb{R}^2} \|Ax - b\|_2$

$$A = \begin{bmatrix} 4 & 0 \\ 1 & 2 \\ 0 & 0 \end{bmatrix}$$

$$b = \begin{bmatrix} 2 \\ 3.5 \\ 3 \end{bmatrix}$$

$$\Rightarrow x^* = ?$$



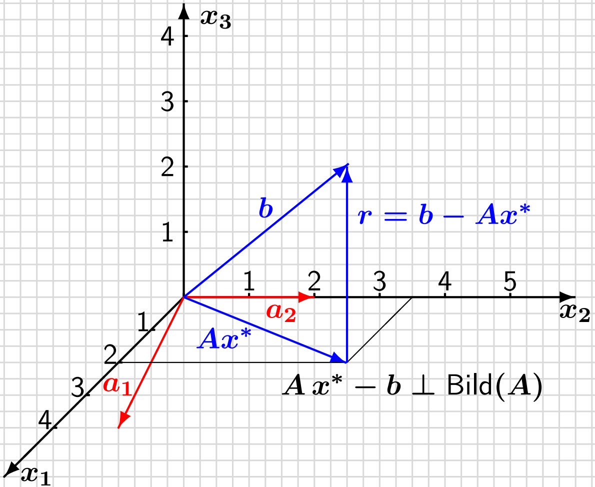
Geometrische Interpretation $\min_{x\in\mathbb{R}^2} \|Ax-b\|_2$

$$A = \begin{bmatrix} 4 & 0 \\ 1 & 2 \\ 0 & 0 \end{bmatrix}$$

$$b = \begin{bmatrix} 2 \\ 3.5 \\ 3 \end{bmatrix}$$

$$\Rightarrow x^* = \begin{bmatrix} 0.5 \\ 1.5 \end{bmatrix}$$

$$\Rightarrow \|r\|_2 = 3$$



Beispiel 4.3.

Man vermutet, dass die Messdaten

t	0	1	2	3
y	3	2.14	1.86	1.72

einer Gesetzmäßigkeit der Form

$$y = f(t) = \alpha \frac{1}{1+t} + \beta$$

mit noch zu bestimmenden Parametern $\alpha, \beta \in \mathbb{R}$ gehorchen.

Frage/Problem

- Wie lautet das zugehörige lineare Ausgleichsproblem?

Beispiel 4.3.

Meßdaten

t	0	1	2	3
y	3	2.14	1.86	1.72

Gesetzmäßigkeit

$$y = f(t) = \alpha \frac{1}{1+t} + \beta$$

Das Ausgleichsproblem lautet $\|A\,x^* - b\|_2 = \min_{x \in \mathbb{R}^2} \|A\,x - b\|_2$,

wobei

$$x = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \quad A = \begin{pmatrix} \frac{1}{1+0} & 1 \\ \frac{1}{1+1} & 1 \\ \frac{1}{1+2} & 1 \\ \frac{1}{1+3} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ \frac{1}{2} & 1 \\ \frac{1}{3} & 1 \\ \frac{1}{4} & 1 \end{pmatrix}, \quad b = \begin{pmatrix} 3 \\ 2.14 \\ 1.86 \\ 1.72 \end{pmatrix}.$$

Matlab-Demo

Normalgleichungen

Die Lösung des linearen Ausgleichsproblems lässt sich auf die Lösung des linearen Gleichungssystems

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

reduzieren, das häufig als Normalgleichungen bezeichnet wird.

Bemerkung

- Für $A \in \mathbb{R}^{m \times n}$ ist $A^T A \in \mathbb{R}^{n \times n}$ stets symmetrisch.
- Falls $A \in \mathbb{R}^{m \times n}$ vollen (Spalten-)Rang n hat, so ist die Matrix $A^T A \in \mathbb{R}^{n \times n}$ symmetrisch positiv definit.

Annahme:

- Wir beschränken uns hier auf den Fall, dass A vollen Spaltenrang hat: $\mathrm{Rang}(A) = n$ (Fall $\mathrm{Rang}(A) < n$, siehe SVD).

Normalgleichungen

Die Lösung des linearen Ausgleichsproblems lässt sich auf die Lösung des linearen Gleichungssystems

$$\boldsymbol{A^T A\, x = A^T b}$$

reduzieren, das häufig als Normalgleichungen bezeichnet wird.

Satz 4.5.

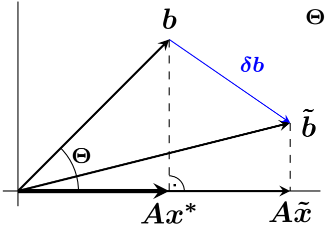
$\boldsymbol{x}^* \in \mathbb{R}^{\boldsymbol{n}}$ ist genau dann Lösung des linearen Ausgleichsproblems, wenn $\boldsymbol{x}^*$ Lösung der Normalgleichungen

$$\boldsymbol{A^T A\, x^* = A^T b}$$

ist. Das System der Normalgleichungen hat stets mindestens eine Lösung. Sie ist genau dann eindeutig, wenn $\mathbf{Rang}(\boldsymbol{A}) = \boldsymbol{n}$ gilt.

Kondition des linearen Ausgleichsproblems

Für $A \in \mathbb{R}^{m \times n (m \neq n)}$ sei $\kappa_2(A) := \max_{x \neq 0} \frac{\|A\,x\|_2}{\|x\|_2} \Big/ \min_{x \neq 0} \frac{\|A\,x\|_2}{\|x\|_2}$.



Θ: Winkel zwischen b und $A\,x^*$

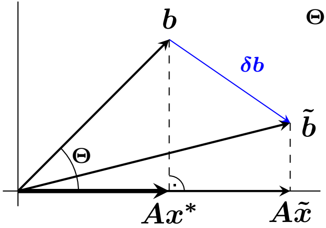
Satz 4.7.

Für die Kondition des linearen Ausgleichsproblems bezüglich Störungen in b gilt

$$\frac{\|\tilde{x} - x^*\|_2}{\|x^*\|_2} \leq \frac{\kappa_2(A)}{\cos\Theta} \frac{\|\tilde{b} - b\|_2}{\|b\|_2}.$$

Kondition des linearen Ausgleichsproblems

Für $A \in \mathbb{R}^{m\times n (m\neq n)}$ sei $\kappa_2(A) := \max_{x\neq 0} \frac{\|Ax\|_2}{\|x\|_2} \Big/ \min_{x\neq 0} \frac{\|Ax\|_2}{\|x\|_2}$.



Θ: Winkel zwischen b und $A\,x^*$

Satz 4.9.

Für die Kondition des linearen Ausgleichsproblems bezüglich Störungen in A gilt

$$\frac{\|\tilde{x} - x^*\|_2}{\|x^*\|_2} \leq \left(\kappa_2(A) + \kappa_2(A)^2 \tan\Theta\right) \frac{\|\tilde{A} - A\|_2}{\|A\|_2}$$

Beispiel 4.8.

Gegeben seien

$$\boldsymbol{A} := \begin{pmatrix} 1 & 1 \\ 0 & 0 \\ 0 & 1 \end{pmatrix} \text{ und } \boldsymbol{b} := \begin{pmatrix} 0.01 \\ 1 \\ 0 \end{pmatrix},$$

sowie eine gestörte rechte Seite $\tilde{\boldsymbol{b}} = (0.01, 1, 0.01)^T$. Bestimmen Sie $\boldsymbol{x}^*$ und $\tilde{\boldsymbol{x}}$, und diskutieren Sie die Kondition des linearen Ausgleichsproblems.

Die Lösung der Normalgleichungen liefert

$$\boldsymbol{x}^* = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{b} = \begin{pmatrix} 0.01 \\ 0 \end{pmatrix},$$

sowie für die gestörte rechte Seite

$$\tilde{\boldsymbol{x}} = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\tilde{\boldsymbol{b}} = \begin{pmatrix} 0 \\ 0.01 \end{pmatrix}.$$

Beispiel 4.8.

Daraus folgt

$$\frac{\|\tilde{x} - x^*\|_2}{\|x^*\|_2} \approx 141 \cdot \frac{\|\tilde{b} - b\|_2}{\|b\|_2},$$

also eine schlechte Kondition des linearen Ausgleichsproblems.

Mit Hilfe von Satz 4.7. erhält man aus

$$\kappa_2(A) \approx 2.62 \quad \text{und} \quad \cos\Theta = \frac{\|A\,x^*\|_2}{\|b\|_2} = 0.01$$

für die Kondition bezüglich Störungen in b

$$\frac{\kappa_2(A)}{\cos\Theta} = 262,$$

d.h. eine schlechte Kondition, obwohl $\kappa_2(A)$ klein ist.

Lösung über Normalgleichungen

Da die Matrix $A^T A$ symmetrisch positiv definit ist, ergibt sich folgende Methode:

Lösung über Normalgleichungen

- Berechne $A^T A$, $A^T b$.
- Berechne die Cholesky-Zerlegung von $A^T A$

$$LDL^T = A^T A$$

- Löse

$$L\,y = A^T\,b, \quad L^T\,x = D^{-1}\,y$$

durch Vorwärts- bzw. Rückwärtseinsetzen.

Lösung über Normalgleichungen — Nachteile

- Die Berechnung von $A^T A$ ist für große m aufwendig und birgt die Gefahr von Genauigkeitsverlust durch Auslöschungseffekte. Die Einträge von $A^T A$ sind also mit (möglicherweise erheblichen relativen) Fehlern behaftet.
- Bei der Lösung des Systems $A^T A\,x = A^T b$ über das Cholesky-Verfahren werden die Rundungsfehler in $A^T A$ und $A^T b$ mit

$$\kappa_2(A^T A)$$

verstärkt. Es gilt

$$\kappa_2(A^T A) = \kappa_2(A)^2.$$

Folglich wird die Rundungsfehlerverstärkung durch $\kappa_2(A)^2$ beschrieben.

Beispiel 4.12.

Gegeben seien

$$A = \begin{pmatrix} \sqrt{3} & \sqrt{3} \\ \delta & 0 \\ 0 & \delta \end{pmatrix}, \; b = \begin{pmatrix} 2\sqrt{3} \\ \delta \\ \delta \end{pmatrix}, \; 0 < \delta \ll 1.$$

Bestimmen Sie die Lösung des linearen Ausgleichsproblems über die Normalgleichungen und diskutieren Sie das Ergebnis.

- Das lineare Ausgleichsproblem hat die Lösung $x^* = (1,1)^T$ (für alle $\delta > 0$).
- Es gilt $\Theta = 0$ und damit $\cos\Theta = 1$, d.h. die Kondition des Problems wird ausschließlich durch $\kappa_2(A)$ beschrieben.
- Man rechnet einfach nach, dass

$$\kappa_2(A) = \sqrt{\frac{\lambda_{\max}(A^TA)}{\lambda_{\min}(A^TA)}} = \frac{\sqrt{6+\delta^2}}{\delta} \approx \frac{\sqrt{6}}{\delta}.$$

Beispiel 4.12.

- Ein stabiles Verfahren sollte ein Resultat $\tilde{x}$ liefern, mit

$$\frac{\|\tilde{x}-x^*\|_2}{\|x^*\|_2} \lessapprox \kappa_2(A)\,\mathrm{eps}.$$

(Satz 4.7+4.9 mit $\theta = 0$)

- Die Lösung über die Normalgleichungen und das Cholesky-Verfahren auf einer Maschine mit **eps** $\approx 10^{-16}$ ergibt jedoch:

$$\delta = 10^{-4}: \quad \frac{\|\tilde{x}-x^*\|_2}{\|x^*\|_2} \approx 2\cdot 10^{-8} \approx \frac{1}{3}\kappa_2(A)^2\,\mathrm{eps}$$

$$\delta = 10^{-6}: \quad \frac{\|\tilde{x}-x^*\|_2}{\|x^*\|_2} \approx 2\cdot 10^{-4} \approx \frac{1}{3}\kappa_2(A)^2\,\mathrm{eps}$$

Lösung über QR-Zerlegung

Zur Erinnerung:

- Für $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, $m > n$, mit $\mathbf{Rang}(\boldsymbol{A}) = n$, folgt aus der QR-Zerlegung von $\boldsymbol{A}$, dass $\boldsymbol{Q} \in \mathbb{R}^{m \times m}$ orthogonal ex. mit

$$\boldsymbol{Q}\,\boldsymbol{A} = \boldsymbol{R} = \begin{pmatrix} \widetilde{\boldsymbol{R}} \\ \mathbf{0} \end{pmatrix} \begin{array}{l} \} \;\; n \\ \} \;\; m-n \end{array},$$

  wobei die obere Dreiecksmatrix $\widetilde{\boldsymbol{R}} \in \mathbb{R}^{n \times n}$ regulär ist.
- Multiplikation mit (einer orthogonalen Matrix) $\boldsymbol{Q}$ verändert nicht die euklidische Länge eines Vektors, d.h.

$$\|\boldsymbol{x}\|_2 = \|\boldsymbol{Q}\boldsymbol{x}\|_2 \text{ für alle } \boldsymbol{x} \in \mathbb{R}^n.$$

- Das lineare Ausgleichsproblem: bestimme $\boldsymbol{x}^* \in \mathbb{R}^n$, für dass

$$\|\boldsymbol{A}\,\boldsymbol{x}^* - \boldsymbol{b}\|_2 = \min_{\boldsymbol{x} \in \mathbb{R}^n} \|\boldsymbol{A}\,\boldsymbol{x} - \boldsymbol{b}\|_2$$

  gilt.

Lösung über QR-Zerlegung

Daraus folgt:

$$\begin{aligned} \|A\,x^* - b\|_2 &= \min_{x\in\mathbb{R}^n} \|A\,x - b\|_2 = \min_{x\in\mathbb{R}^n} \|Q(Ax - b)\|_2 \\ &= \min_{x\in\mathbb{R}^n} \|Q\,A\,x - Q\,b\|_2 \\ &= \min_{x\in\mathbb{R}^n} \|R\,x - Q\,b\|_2, \end{aligned}$$

mit $R = \begin{pmatrix} \widetilde{R} \\ 0 \end{pmatrix} \begin{matrix} \}n \\ \}m-n \end{matrix}, \quad Q\,b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \begin{matrix} \}n \\ \}m-n \end{matrix},$ erhält man

$$\begin{aligned} \|A\,x^* - b\|_2^2 &= \min_{x\in\mathbb{R}^n} \left\| \begin{pmatrix} \widetilde{R} \\ 0 \end{pmatrix} x - \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \right\|_2^2 \\ &= \min_{x\in\mathbb{R}^n} \left(\|\widetilde{R}\,x - b_1\|_2^2 + \|b_2\|_2^2 \right) \\ &= \|b_2\|_2^2 \quad \text{für} \quad \widetilde{R}\,x = b_1 \end{aligned}$$

Lösung über QR-Zerlegung

Satz 4.13.

Seien $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ mit $\mathbf{Rang}(\boldsymbol{A}) = n$ und $\boldsymbol{b} \in \mathbb{R}^{m}$.
Sei $\boldsymbol{Q} \in \mathbb{R}^{m \times m}$ eine orthogonale Matrix und
$\widetilde{\boldsymbol{R}} \in \mathbb{R}^{n \times n}$ eine obere Dreiecksmatrix, so dass

$$\boldsymbol{Q}\,\boldsymbol{A} = \boldsymbol{R} := \begin{pmatrix} \widetilde{\boldsymbol{R}} \\ \boldsymbol{0} \end{pmatrix} \begin{matrix} \}n \\ \}m-n \end{matrix}.$$

Dann ist die Matrix $\widetilde{\boldsymbol{R}}$ regulär. Schreibt man

$$\boldsymbol{Q}\,\boldsymbol{b} = \begin{pmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \end{pmatrix} \begin{matrix} \}n \\ \}m-n \end{matrix},$$

dann ist $\boldsymbol{x}^* = \widetilde{\boldsymbol{R}}^{-1}\,\boldsymbol{b}_1$ die Lösung des linearen Ausgleichsproblems.
Die Norm $\|\boldsymbol{A}\,\boldsymbol{x}^* - \boldsymbol{b}\|_2$ ist gerade durch $\|\boldsymbol{b}_2\|_2$ gegeben.

Lösung über QR-Zerlegung

Aus Satz 4.13. ergibt sich nun folgende Methode:

- Bestimme die $\boldsymbol{QR}$-Zerlegung von $\boldsymbol{A}$

$$\boldsymbol{QA} = \begin{pmatrix} \widetilde{\boldsymbol{R}} \\ \boldsymbol{0} \end{pmatrix} \qquad (\widetilde{\boldsymbol{R}} \in \mathbb{R}^{n \times n}),$$

  z.B. mittels Givens-Rotationen oder Householder-Spiegelungen und berechne

$$\boldsymbol{Qb} = \begin{pmatrix} \boldsymbol{b_1} \\ \boldsymbol{b_2} \end{pmatrix}.$$

- Löse $\widetilde{\boldsymbol{R}}\boldsymbol{x} = \boldsymbol{b_1}$ mittels Rückwärtseinsetzen.
- Die Norm des Residuums $\min\limits_{\boldsymbol{x} \in \mathbb{R}^n} \|\boldsymbol{Ax} - \boldsymbol{b}\|_2 = \|\boldsymbol{Ax}^* - \boldsymbol{b}\|_2$ ist gerade durch $\|\boldsymbol{b_2}\|_2$ gegeben.

Beispiel 4.15.

Gegeben seien

$$\boldsymbol{A} = \begin{pmatrix} 3 & 7 \\ 0 & 12 \\ 4 & 1 \end{pmatrix}, \quad \boldsymbol{b} = \begin{pmatrix} 10 \\ 1 \\ 5 \end{pmatrix},$$

d.h. $\boldsymbol{m = 3, n = 2.}$
Man bestimme die Lösung $x^* \in \mathbb{R}^2$ des zugehörigen linearen Ausgleichsproblem über QR-Zerlegung mittels Givens-Rotation.

Annullierung von $\boldsymbol{a_{3,1}}$:

$$\boldsymbol{A^{(2)} = G_{1,3}\, A} = \begin{pmatrix} 5 & 5 \\ 0 & 12 \\ 0 & -5 \end{pmatrix}, \quad \boldsymbol{b^{(2)} = G_{1,3}\, b} = \begin{pmatrix} 10 \\ 1 \\ -5 \end{pmatrix}.$$

Zur Erinnerung: die Transformationen $\boldsymbol{G_{1,3}A}$ und $\boldsymbol{G_{1,3}b}$ werden in der Praxis ausgeführt, ohne dass $\boldsymbol{G_{1,3}}$ explizit berechnet wird.

Beispiel 4.15.

Annullierung von $a^{(2)}_{3,2}$:

$$A^{(3)} = G_{2,3}\, A^{(2)} = \begin{pmatrix} 5 & 5 \\ 0 & 13 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} \widetilde{R} \\ 0 \end{pmatrix}, \quad b^{(3)} = G_{2,3}\, b^{(2)} = \begin{pmatrix} 10 \\ \frac{37}{13} \\ -\frac{55}{13} \end{pmatrix}$$

Lösung von

$$\begin{pmatrix} 5 & 5 \\ 0 & 13 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 10 \\ \frac{37}{13} \end{pmatrix}$$

durch Rückwärtseinsetzen liefert

$$x^* = \left(\frac{301}{169}, \frac{37}{169} \right)^T.$$

Als Norm des Residiums ergibt sich:

$$\|b_2\|_2 = \frac{55}{13}.$$

Lösung über QR-Zerlegung — Stabilität

Beachte

- Wegen Satz 3.41. gilt

$$\kappa_2(A) = \kappa_2(\tilde{R}),$$

  d.h. das Quadrieren der Kondition, das bei den Normalgleichungen auftritt, wird vermieden.

- Die Berechnung der QR-Zerlegung über Givens- oder Householder-Transformationen ist ein sehr stabiles Verfahren, wobei die Fehlerverstärkung durch $\kappa_2(A)$ (und nicht $\kappa_2(A)^2$) beschrieben wird.

Beispiel 4.16.

Gegeben seien

$$A = \begin{pmatrix} \sqrt{3} & \sqrt{3} \\ \delta & 0 \\ 0 & \delta \end{pmatrix}, \; b = \begin{pmatrix} 2\sqrt{3} \\ \delta \\ \delta \end{pmatrix}, \; 0 < \delta \ll 1.$$

Bestimmen Sie die Lösung des linearen Ausgleichsproblems über die QR-Zerlegung und diskutieren Sie das Ergebnis.

Auf einer Maschine mit **eps** $\approx 10^{-16}$ erhält man

$$\delta = 10^{-4}: \; \frac{\|\tilde{x} - x^*\|_2}{\|x^*\|_2} \approx 2.2 \cdot 10^{-16},$$

$$\delta = 10^{-6}: \; \frac{\|\tilde{x} - x^*\|_2}{\|x^*\|_2} \approx 1.6 \cdot 10^{-16}.$$

Wegen der sehr guten Stabilität dieser Methode sind die Resultate viel besser als in Beispiel 4.12.

Zusammenfassung

	Normalgleichungen	QR-Zerlegung
Rechenaufwand $(m \gg n)$	ca. $\frac{1}{2}mn^2$	ca. mn^2 (Householder)
Stabilität	instabil, wenn $\kappa_2(A) \gg 1$ stabil, wenn $\kappa_2(A)$ moderat	stabil

Zusammenfassung

- Aufgabe:

$$\min_{x \in \mathbb{R}^n} \|A\,x - b\|_2 \quad \Leftrightarrow \quad A^T A\,x = A^T b$$

- Eindeutige Lösung $\Leftrightarrow$ $\mathbf{Rang}(A) = n$
- Kondition (nur Störung in b):

$$\frac{\|\tilde{x} - x^*\|_2}{\|x^*\|_2} \leq \frac{\kappa_2(A)}{\cos\Theta} \cdot \frac{\|\tilde{b} - b\|_2}{\|b\|_2}$$

- Lösungsverfahren:
  - über Normalgleichungen $A^T A\,x = A^T b$ (Cholesky-Verfahren)
  - über QR-Zerlegung (Householder, Givens)

Verständnisfragen

Es seien $A \in \mathbb{R}^{m\times n}$, mit $\mathbf{Rang}(A) = n < m$, und $b \in \mathbb{R}^m$. Weiter seien $Q \in \mathbb{R}^{m\times m}$ eine orthogonale Matrix und $\widetilde{R} \in \mathbb{R}^{n\times n}$ eine obere Dreiecksmatrix so, dass $Q\,A = (\widetilde{R}^T, 0)^T$ gilt. Weiter seien $x^* \in \mathbb{R}^n$ die eindeutige Minimalstelle des Minimierungsproblems $\arg\min_{x\in\mathbb{R}^n} \|A\,x - b\|_2$ und $\Theta \in \left[0, \frac{\pi}{2}\right)$ der Winkel zwischen $A\,x^*$ und b.

- w | Es gilt $\det \widetilde{R} \neq 0$.
- f | Es gilt $\widetilde{R}x^* = Q\,b$.
- w | Es gilt $\kappa_2(A) = \kappa_2(\widetilde{R})$.

Es seien $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 1 \end{pmatrix}$, $b = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$.

Bestimmen Sie Θ. 0

Verständnisfragen

Es seien $A \in \mathbb{R}^{m\times n}$ mit $\mathbf{Rang}(A) = n < m$, $b \in \mathbb{R}^m$, $Q \in \mathbb{R}^{m\times m}$ eine orthogonale Matrix und $R \in \mathbb{R}^{m\times n}$ eine obere Dreiecksmatrix so, dass $Q\,A = R$ gilt. Weiterhin seien $x^* \in \mathbb{R}^n$ die eindeutige Minimalstelle des Minimierungsproblems $\arg\min_{x\in\mathbb{R}^n} \|A\,x - b\|_2$ sowie $\Theta \in \left[0, \frac{\pi}{2}\right)$ der Winkel zwischen $A\,x^*$ und b.

- f — Je kleiner der Winkel Θ, desto schlechter ist das Problem konditioniert.
- w — Es gilt $\|A\,x - b\|_2 = \|R\,x - Q\,b\|_2$ für beliebiges $x \in \mathbb{R}^n$.
- f — Die Matrix R kann man über Gauß-Elimination mit Spaltenpivotisierung bestimmen.
- w — Es gilt $A\,x^* - b \perp \mathrm{Bild}(A)$.