Numerische Mathematik für Maschinenbauer

Lineare Ausgleichsrechnung

A. Reusken

K.-H. Brakhage, I. Voulis, H. Saß

Institut für Geometrie und Praktische Mathematik
RWTH Aachen

Sommersemester 2017

Heute in der Vorlesung

Themen: Dahmen & Reusken Kap. 4.1-4.4

- Householder-Transformation
- 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

Berechnung der QR-Zerlegung

Wir haben mehrere Möglichkeiten:

- **Gram-Schmidt Orthogonalisierung**
  Idee: Schrittweise Orthogonalisierung der Spalten von $\boldsymbol{A}$
- **Givens-Rotation**
  Idee: Schrittweise ebene Drehungen der Spalten von $\boldsymbol{A}$
- **Householder-Transformation**
  Idee: Schrittweise Spiegelungen der Spalten von $\boldsymbol{A}$

Householder-Transformationen

Definition

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

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

wobei die Dyade, $v\,v^T$, gegeben ist durch

$$v\,v^T = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} (v_1, \dots, v_n) = \begin{pmatrix} v_1 v_1 & \dots & v_1 v_n \\ \vdots & & \vdots \\ v_n v_1 & \dots & v_n v_n \end{pmatrix}.$$

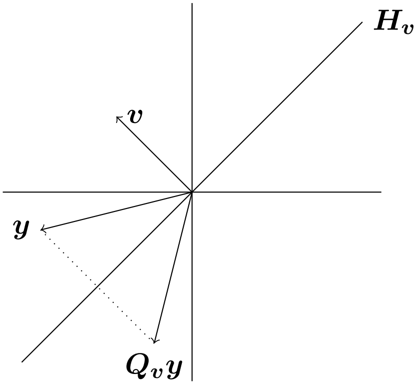
- Die Householder Transformation Q_v ist orthogonal, d.h.

$$Q_v^T Q_v = I, \quad Q_v^{-1} = Q_v^T$$

Householder-Transformationen

Geometrische Interpretation: Spiegelung

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



Householder-Transformationen

Eigenschaften 3.50.

- $Q_v = Q_v^T$
- $Q_v^2 = I$

  Zweimalige Spiegelung ergibt den ursprünglichen Punkt.
- $Q_{\alpha v} = Q_v, \quad \alpha \in \mathbb{R},\ \alpha \neq 0$

  Skalierung des Normalenvektors ändert nicht die Spiegelebene.
- $Q_v y = y \iff y^T v = 0$

  Urspünglicher und gespiegelter Punkt sind nur identisch, wenn der Punkt in der Spiegelebene liegt.
- $Q_v v = -v.$

  Spiegelung des Normalenvektors vertauscht das Vorzeichen.

Householder-Transformationen

Grundaufgabe

Zu $y \in \mathbb{R}^n, y \notin \mathrm{span}(e^1)$, finde $v \in \mathbb{R}^n$, so dass

$$Q_v y = \pm \|y\|_2 \, e^1$$

gilt.

- Die Lösung der Grundaufgabe ist:

$$v = y \pm \|y\|_2 \, e^1$$

- Um Auslöschung zu vermeiden, wählt man

$$v = y + \mathrm{sign}(y_1) \|y\|_2 e^1, \text{ mit } \mathrm{sign}(0) := 1.$$

Zusammenfassend

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

Beispiel 3.51

Zu $y = (2, 2, 1)^T$ wird $v \in \mathbb{R}^3$ gesucht, so dass

$$Q_v y = \pm \|y\|_2 e^1 = \pm 3 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

gilt.

Wir erhalten

$$\alpha = 3$$

$$v = y + \alpha e^1 = \begin{pmatrix} 5 \\ 2 \\ 1 \end{pmatrix}$$

und damit

$$Q_v y = \begin{pmatrix} -3 \\ 0 \\ 0 \end{pmatrix}$$

Beispiel 3.51

$$Q_v = I - 2\frac{vv^T}{v^Tv} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - 2\frac{\begin{pmatrix} 5 \\ 2 \\ 1 \end{pmatrix}(5,2,1)}{(5,2,1)\begin{pmatrix} 5 \\ 2 \\ 1 \end{pmatrix}}$$

$$= \frac{1}{15}\begin{pmatrix} -10 & -10 & -5 \\ -10 & 11 & -2 \\ -5 & -2 & 14 \end{pmatrix}$$

Beachte

$$Q_v w = \left(I - 2\frac{vv^T}{v^Tv}\right) w = w - \frac{2v^Tw}{v^Tv} v.$$

Reduktion auf obere Dreiecksform

- Sei a^1 die erste Spalte der Matrix $A \in \mathbb{R}^{m\times n}, m \geq n$.
- Man wendet die Grundaufgabe mit $y = a^1$ an:

$$v^1 = a^1 + \mathrm{sign}(a_{1,1})\|a^1\|_2\, e^1, \qquad Q_1 := Q_{v^1},$$

  und berechnet $Q_1 A$.
- Sei $a^{(2)^1}$ die erste Spalte ...

$A = A^{(1)}$

$$\begin{array}{|ccccc|}\hline * & * & \cdots & \cdots & * \\ * & * & \cdots & \cdots & * \\ \vdots & \vdots & & & \vdots \\ \vdots & \vdots & & & \vdots \\ * & * & \cdots & \cdots & * \\ \hline\end{array} \qquad \begin{array}{|c|cccc|}\hline * & * & \cdots & \cdots & * \\ \hline 0 & * & \cdots & \cdots & * \\ \vdots & \vdots & \tilde{A}^{(2)} & & \vdots \\ \vdots & \vdots & & & \vdots \\ 0 & * & \cdots & \cdots & * \\ \hline\end{array} \qquad \begin{array}{|cc|ccc|}\hline * & * & * & \cdots & * \\ 0 & * & * & \cdots & * \\ \hline 0 & 0 & * & \cdots & * \\ \vdots & \vdots & \vdots & \tilde{A}^{(3)} & \vdots \\ 0 & 0 & * & \cdots & * \\ \hline\end{array}$$

$Q_{n-1}\dots Q_2 Q_1 A = R$, bzw. $A = Q_1^T Q_2^T \dots Q_{n-1}^T R = QR$

Beispiel 3.52.

Bestimmen Sie die QR-Zerlegung der Matrix

$$\boldsymbol{A} = \begin{pmatrix} \mathbf{1} & \mathbf{1} \\ \mathbf{2} & \mathbf{0} \\ \mathbf{2} & \mathbf{0} \end{pmatrix}$$

mittels Householder-Transformation.

1. Grundaufgabe mit $\boldsymbol{y} = \boldsymbol{a}^{\mathbf{1}}$ (erste Spalte von $\boldsymbol{A}$) ergibt

$$\boldsymbol{v}^{\mathbf{1}} = \begin{pmatrix} \mathbf{1} \\ \mathbf{2} \\ \mathbf{2} \end{pmatrix} + \mathbf{3}\boldsymbol{e}^{\mathbf{1}} = \begin{pmatrix} \mathbf{4} \\ \mathbf{2} \\ \mathbf{2} \end{pmatrix}, \quad \boldsymbol{Q}_{\mathbf{1}} = \boldsymbol{Q}_{\boldsymbol{v}^{\mathbf{1}}}.$$

2. Für die zwei Spalten der Matrix $\boldsymbol{Q}_{\mathbf{1}}\boldsymbol{A}$ ergibt sich

$$\boldsymbol{Q}_{\mathbf{1}} \begin{pmatrix} \mathbf{1} \\ \mathbf{2} \\ \mathbf{2} \end{pmatrix} = \begin{pmatrix} \mathbf{-3} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix} \quad \text{(siehe Beispiel 3.51)}$$

Beispiel 3.52.

$$Q_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - \frac{2}{(v^1)^T v^1} v^1 (v^1)^T \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - \frac{1}{3} v^1 = \begin{pmatrix} -\frac{1}{3} \\ -\frac{2}{3} \\ -\frac{2}{3} \end{pmatrix}$$

3. Daraus folgt

$$Q_1 A = \begin{pmatrix} -3 & -\frac{1}{3} \\ 0 & -\frac{2}{3} \\ 0 & -\frac{2}{3} \end{pmatrix}$$

4. Grundaufgabe mit y gleich erster Spalte von $A^{(2)}$, d.h. $y = (-\frac{2}{3}, -\frac{2}{3})^T$, ergibt

$$v^2 = \begin{pmatrix} -\frac{2}{3} \\ -\frac{2}{3} \end{pmatrix} - \frac{2}{3}\sqrt{2} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -\frac{2}{3}(1+\sqrt{2}) \\ -\frac{2}{3} \end{pmatrix}, \quad \tilde{Q}_2 = \tilde{Q}_{v^2}.$$

Beispiel 3.52.

5. Damit ergibt sich

$$\tilde{Q}_2 \begin{pmatrix} -\frac{2}{3} \\ -\frac{2}{3} \end{pmatrix} = \begin{pmatrix} \frac{2}{3}\sqrt{2} \\ 0 \end{pmatrix}$$

6. Insgesamt erhält man

$$\underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & & \\ 0 & \multicolumn{2}{c}{\tilde{Q}_2} \end{pmatrix}}_{Q_2} Q_1\, A = \begin{pmatrix} -3 & -\frac{1}{3} \\ 0 & \frac{2}{3}\sqrt{2} \\ 0 & 0 \end{pmatrix}.$$

Householder-Transformation: Zusammenfassung

- Die QR-Zerlegung über Householder-Transformationen ist ebenfalls sehr stabil.
- Gesonderte Pivotisierung ist nicht erforderlich.
- Der Aufwand für die QR-Zerlegung einer vollbesetzten $m \times n$-Matrix über Householder-Transformationen ist etwa $\frac{2}{3}n^3$ Operationen, falls $m \approx n$, und etwa mn^2 Operationen, falls $m \gg n$.

Wichtige Anwendungen der QR-Zerlegung:

Ausgleichsrechnung (siehe nächste Folie), Berechnung von Eigenwerten

Problemstellung

Bisher: Lineare Gleichungssysteme

- geg.: $\boldsymbol{A} \in \mathbb{R}^{n \times n}$, $\boldsymbol{b} \in \mathbb{R}^{n}$;
  ges.: $\boldsymbol{x} \in \mathbb{R}^{n}$, so dass $\boldsymbol{Ax} = \boldsymbol{b}$
- Annahme: $\det \boldsymbol{A} \neq 0$
  $\Rightarrow$ Spalten von $\boldsymbol{A}$ bilden eine Basis in $\mathbb{R}^{n}$
  $\Rightarrow$ $\boldsymbol{Ax} = \boldsymbol{b}$ eindeutig lösbar.

Jetzt: Lineare Ausgleichsrechnung

- geg.: $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, $\boldsymbol{b} \in \mathbb{R}^{m}$, $m > n$;
  ges.: $\boldsymbol{x} \in \mathbb{R}^{n}$, so dass $\boldsymbol{Ax} = \boldsymbol{b}$?
  $\Rightarrow$ im Allgemeinen nicht lösbar!
- Lösung: Bestimme $\boldsymbol{x}^* \in \mathbb{R}^{n}$, so dass

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

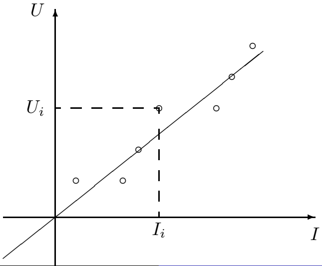
Bestimmung des elektrischen Widerstands (Beispiel 4.1)

- 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.

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



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 $c := \frac{2\pi}{T}$ 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:

$$\sum_{i=1}^{m} \Big(g_N(t_i) - b_i \Big)^2 = \text{min.}$$

Allgemeines lineares Ausgleichsproblem

Definition

Zu gegebenen $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ und $\boldsymbol{b} \in \mathbb{R}^{m}$, 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. Diese Problemstellung heißt das lineare Ausgleichsproblem.

oder:

Lineares Ausgleichsproblem

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

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

Definition

Warum 2-Norm?

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

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

Auch möglich:

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

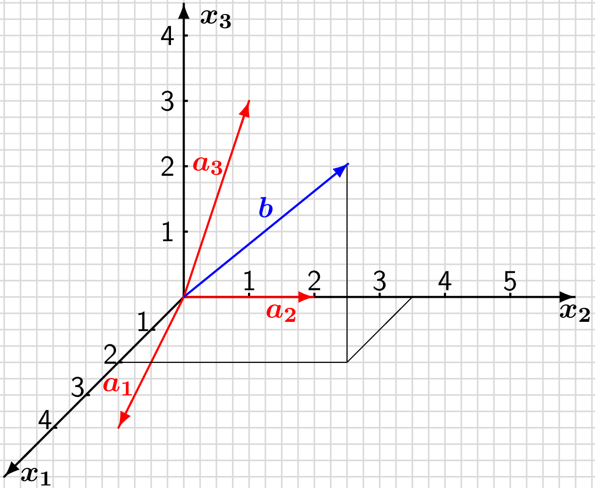
$\Rightarrow$ führt auf lineares Optimierungsproblem

Geometrische Interpretation $Ax = b$

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

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

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

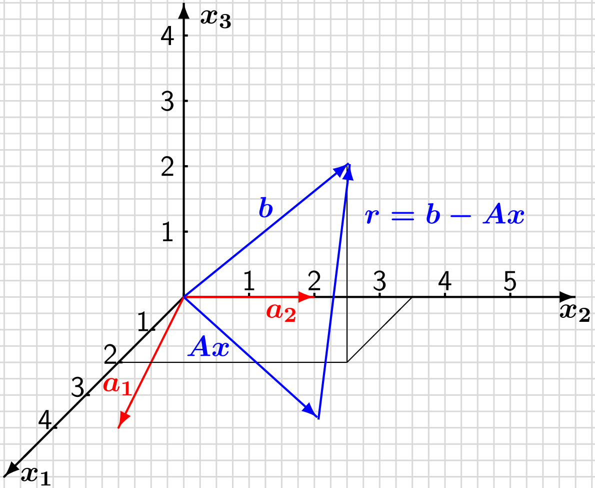


Geometrische Interpretation $\min_{x\in\mathbb{R}^2} \|Ax - b\|_2$

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

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

$$\Rightarrow x^* = ?$$



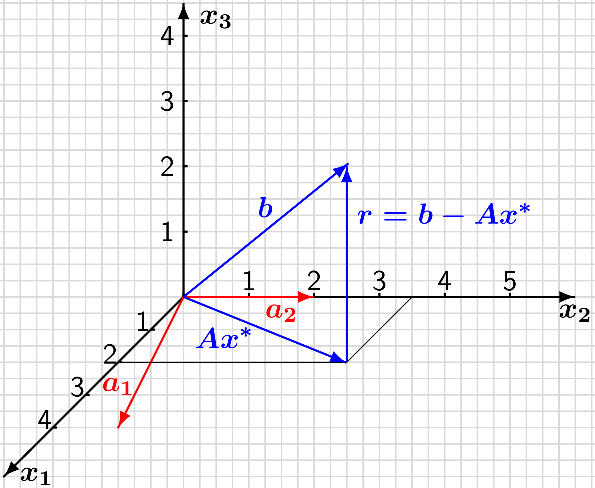
Geometrische Interpretation $\min_{x \in \mathbb{R}^2} \|Ax - b\|_2$

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

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

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

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



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 die Matrix $A^T A \in \mathbb{R}^{n \times n}$ stets quadratisch.
- 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: Rang $(A) = n$ (Fall Rang $(A) < n$, siehe SVD).

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.

Satz 4.5

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

$$A^T A x^* = A^T b$$

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

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.

Wie lautet das zugehörige lineare Ausgleichsproblem?

Beispiel 4.3

Meßdaten

t	0	1	2	3
y	3	2.14	1.86	1.72

Gesetzmäßigkeit

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

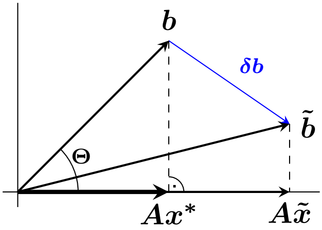
Das Ausgleichsproblem lautet $\|A\,x^* - b\|_2 = \min_{x \in \mathbb{R}^2} \|A\,x - b\|_2$,

wobei

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

Kondition des linearen Ausgleichsproblems

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

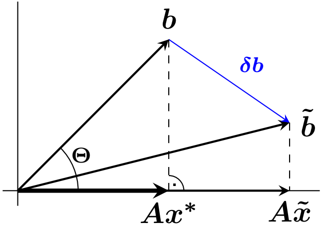


Satz 4.7

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



Satz 4.9.

Für die Kondition des linearen Ausgleichsproblems bezüglich Störungen in $\boldsymbol{A}$ 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

$$\frac{\|\tilde{x} - x^*\|_2}{\|x^*\|_2} \approx 100 \frac{\|\tilde{b} - b\|_2}{\|b\|_2},$$

also eine schlechte Kondition des linearen Ausgleichsproblems.

Mit Hilfe von Satz 4.7 erhält man aus

$$\kappa_2(A) \approx 2.62$$

und

$$\cos \Theta = \frac{\|A\, x^*\|_2}{\|b\|_2} = 0.01$$

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
$$LDL^T = A^T A$$
von $A^T A$
- 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

$$\kappa_2(A) \approx \frac{\sqrt{6}}{\delta}.$$

Beispiel 4.12.

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

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

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

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

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

Lösung über QR-Zerlegung

Zur Erinnerung:

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

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

  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 $R = \begin{pmatrix} \widetilde{R} \\ 0 \end{pmatrix} \begin{array}{l} \}n \\ \}m-n \end{array}$, $\quad Q\,b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \begin{array}{l} \}n \\ \}m-n \end{array}$, 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 $\mathrm{Rang}(\boldsymbol{A}) = n$ und $\boldsymbol{b} \in \mathbb{R}^{m}$. Seien $\boldsymbol{Q} \in \mathbb{R}^{m \times m}$ eine orthogonale Matrix und $\widetilde{\boldsymbol{R}} \in \mathbb{R}^{n \times n}$ eine obere Dreiecksmatrix, so dass

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

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 $\min_{\boldsymbol{x} \in \mathbb{R}^n} \|\boldsymbol{Ax} - \boldsymbol{b}\|_2 = \|\boldsymbol{Ax}^* - \boldsymbol{b}\|_2$ 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. $m = 3, n = 2.$ Man bestimme die Lösung $x^* \in \mathbb{R}^2$ des zugehörigen linearen Ausgleichsproblem über QR-Zerlegung mittels Givens-Rotation.

- Annullierung von $a_{3,1}$:

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

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

Beispiel 4.15.

- Annullierung von $a_{3,2}^{(2)}$:

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

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 3.14 gilt

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

  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 $\mathbf{eps} \approx 10^{-16}$ erhält man

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

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

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 $\kappa_2(A) \gg 1$ und $\theta \ll \frac{\pi}{2}$	stabil

Zusammenfassung

- Aufgabe:

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

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

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

- 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 $Q\,A = (\widetilde{R}, 0)^T$ gilt. Weiter seien $x^* \in \mathbb{R}^n$ die eindeutige Minimalstelle des Minimierungsproblems $\min_{x \in \mathbb{R}^n} \|A\,x - b\|_2$ und $\Theta \in \left[0, \frac{\pi}{2}\right)$ der Winkel zwischen $A\,x^*$ und b.

- X Es gilt $\det \widetilde{R} \neq 0$.
- ☐ Es gilt $\widetilde{R}x^* = Qb$.
- X 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 $\min_{x\in\mathbb{R}^n} \|A\,x - b\|_2$ sowie $\Theta \in \left[0, \frac{\pi}{2}\right)$ der Winkel zwischen $A\,x^*$ und b.

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