Numerisches Rechnen

Lineare Gleichungssysteme

M. Grepl

J. Berger & R. O'Connor

Institut für Geometrie und Praktische Mathematik
RWTH Aachen

Wintersemester 2015/16

Vorlesungsinhalt

1. Fehleranalyse: Kondition, Rundungsfehler, Stabilität
   a) $\boldsymbol{y} = \boldsymbol{f}(\boldsymbol{x})$, Eingabefehler $\boldsymbol{\Delta x} \rightarrow$ Ausgabefehler $\boldsymbol{\Delta y}$
   b) Fehler aufgrund Gleitpunktdarstellung
   c) Fehler (durch Algorithmus) $\approx$ Fehler (durch Kondition)
2. **Lineare Gleichungssysteme, direkte Lösungsverfahren**
   geg.: $\boldsymbol{A} \in \mathbb{R}^{\boldsymbol{n} \times \boldsymbol{n}}$, $\boldsymbol{b} \in \mathbb{R}^{\boldsymbol{n}}$;
   ges.: $\boldsymbol{x} \in \mathbb{R}^{\boldsymbol{n}}$, so dass $\boldsymbol{Ax} = \boldsymbol{b}$
3. Lineare Ausgleichsrechnung
   geg.: $\boldsymbol{A} \in \mathbb{R}^{\boldsymbol{m} \times \boldsymbol{n}}$, $\boldsymbol{b} \in \mathbb{R}^{\boldsymbol{m}}$, $\boldsymbol{m} > \boldsymbol{n}$;
   ges.: $\boldsymbol{x}^* \in \mathbb{R}^{\boldsymbol{n}}$, so dass $\boldsymbol{x}^* = \arg\min_{\boldsymbol{x} \in \mathbb{R}^{\boldsymbol{n}}} \|\boldsymbol{Ax} - \boldsymbol{b}\|_{\boldsymbol{2}}$

Übersicht

Themen: Dahmen & Reusken Kap. 3.6-3.9

- Cholesky-Zerlegung
- Stabilität der LR- und Cholesky-Zerlegung
- QR-Zerlegung

Was Sie mitnehmen sollten:

- Was ist die Cholesky-Zerlegung und wann wird sie eingesetzt?
- Sind die LR- und Cholesky-Zerlegung stabil?
- Was ist die QR-Zerlegung und wie wird sie berechnet?

Matrix-Zerlegung

Aufgabe

Gegeben sei $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ $(\det \boldsymbol{A} \neq \boldsymbol{0})$ und $\boldsymbol{b} \in \mathbb{R}^{n}$, bestimme $\boldsymbol{x} \in \mathbb{R}^{n}$ so dass

$$\boldsymbol{A}\,\boldsymbol{x} = \boldsymbol{b}.$$

Vorgehensweise: Bestimme eine Faktorisierung (Zerlegung) von $\boldsymbol{A}$, so dass das Gleichungssystem "leichter" lösbar ist.

Verfahren:

- LR-Zerlegung: $\boldsymbol{A} = \boldsymbol{L}\,\boldsymbol{R}$, wobei $\boldsymbol{L}$ untere Dreiecksmatrix, $\boldsymbol{R}$ obere Dreiecksmatrix
- Cholesky-Zerlegung: $\boldsymbol{A} = \boldsymbol{L}\,\boldsymbol{D}\,\boldsymbol{L}^{\boldsymbol{T}}$, wobei $\boldsymbol{D}$ Diagonalmatrix
- QR-Zerlegung: $\boldsymbol{A} = \boldsymbol{Q}\,\boldsymbol{R}$, wobei $\boldsymbol{Q}$ orthogonale Matrix

Cholesky-Zerlegung

Definition 3.31.

$A \in \mathbb{R}^{n \times n}$ heißt symmetrisch positiv definit (s.p.d.), falls

$$A^T = A \quad \text{(Symmetrie)}$$

und

$$x^T A x > 0 \quad \text{(positiv definit)}$$

für alle $x \in \mathbb{R}^n, x \neq 0$, gilt.

Tritt bei vielen (physikalischen) Problemen auf:

- Netzwerke mit passiven Komponenten
- Diffusions-/Wärmeleitungsgleichung
- Normalgleichung (Lineare Ausgleichsrechnung)
- ...

Beispiel 3.32

1. $A = I$ (Identität) ist s.p.d. Die Symmetrie ist trivial und
$$x^T I x = x^T x = \|x\|_2^2 > 0,$$
falls $x \neq 0$.

2. Sei $B \in \mathbb{R}^{m \times n}$, $m \geq n$, und B habe vollen Rang. Dann ist $A := B^T B \in \mathbb{R}^{n \times n}$ s.p.d., denn:
$$A^T = (B^T B)^T = B^T (B^T)^T = B^T B = A.$$
Sei $x \in \mathbb{R}^n$, $x \neq 0$. Dann gilt
$$x^T A x = x^T B^T B x = (Bx)^T (Bx) = \|Bx\|_2^2 \geq 0.$$
Es gilt $x^T A x = \|Bx\|_2^2 = 0$ nur falls $Bx = 0$ gilt. Da B vollen Rang hat, muß daher $x = 0$ sein.

Satz 3.33

$\boldsymbol{A} \in \mathbb{R}^{n \times n}$ sei s.p.d. Dann gelten folgende Aussagen:

1. $\boldsymbol{A}$ ist invertierbar, und $\boldsymbol{A}^{-1}$ ist s.p.d.
2. $\boldsymbol{A}$ hat nur strikt positive (insbesondere reelle) Eigenwerte.
3. Jede Hauptuntermatrix von $\boldsymbol{A}$ ist s.p.d.
4. Die Determinante von $\boldsymbol{A}$ ist positiv (und damit die Determinante aller Hauptuntermatrizen von $\boldsymbol{A}$)
5. $\boldsymbol{A}$ hat nur strikt positive Diagonaleinträge und der betragsgrößte Eintrag von $\boldsymbol{A}$ liegt auf der Diagonalen.
6. Bei Gauß-Elimination ohne Pivotisierung sind alle Pivotelemente strikt positiv.

Cholesky-Zerlegung

Satz 3.34

Jede s.p.d. Matrix $A \in \mathbb{R}^{n \times n}$ besitzt eine eindeutige Zerlegung

$$A = LDL^T,$$

wobei L eine nominierte untere Dreiecksmatrix und D eine Diagonalmatrix mit Diagonaleinträgen

$$d_{i,i} > 0, i = 1, \ldots, n,$$

ist. Umgekehrt ist jede Matrix der Form LDL^T, wobei D eine Diagonalmatrix ist, die $d_{i,i} > 0$ erfüllt, und L eine normierte untere Dreiecksmatrix ist, symmetrisch positiv definit.

Beachte: Aufgrund von Satz 3.33 (5) ist bei s.p.d. Matrizen Gauß-Elimination *ohne Pivotisierung* durchführbar.

$\Rightarrow$ "Symmetrische" LR-Zerlegung N3.9

Konstruktion der Cholesky-Zerlegung

Beispiel: Für $n = 3$ ergibt sich

$$L = \begin{pmatrix} 1 & 0 & 0 \\ \ell_{2,1} & 1 & 0 \\ \ell_{3,1} & \ell_{3,2} & 1 \end{pmatrix}, \quad D = \begin{pmatrix} d_{1,1} & 0 & 0 \\ 0 & d_{2,2} & 0 \\ 0 & 0 & d_{3,3} \end{pmatrix}.$$

und damit

$$\begin{aligned} LDL^T &= \begin{pmatrix} 1 & 0 & 0 \\ \ell_{2,1} & 1 & 0 \\ \ell_{3,1} & \ell_{3,2} & 1 \end{pmatrix} \begin{pmatrix} d_{1,1} & 0 & 0 \\ 0 & d_{2,2} & 0 \\ 0 & 0 & d_{3,3} \end{pmatrix} \begin{pmatrix} 1 & \ell_{2,1} & \ell_{3,1} \\ 0 & 1 & \ell_{3,2} \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} d_{1,1} & 0 & 0 \\ \ell_{2,1}d_{1,1} & d_{2,2} & 0 \\ \ell_{3,1}d_{1,1} & \ell_{3,2}d_{2,2} & d_{3,3} \end{pmatrix} \begin{pmatrix} 1 & \ell_{2,1} & \ell_{3,1} \\ 0 & 1 & \ell_{3,2} \\ 0 & 0 & 1 \end{pmatrix} \\ &= \dots \end{aligned}$$

Konstruktion der Cholesky-Zerlegung

$$\ldots = \begin{pmatrix} d_{1,1} & \ldots & \ldots \\ \ell_{2,1}d_{1,1} & \ell_{2,1}^2 d_{1,1} + d_{2,2} & \ldots \\ \ell_{3,1}d_{1,1} & \ell_{2,1}\ell_{3,1}d_{1,1} + \ell_{3,2}d_{2,2} & \ell_{3,1}^2 d_{1,1} + \ell_{3,2}^2 d_{2,2} + d_{3,3} \end{pmatrix}$$

$$= \begin{pmatrix} a_{1,1} & \ldots & \ldots \\ a_{2,1} & a_{2,2} & \ldots \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix} = A$$

Die elementweise Auswertung der Gleichung $LDL^T = A$, die man aufgrund der Symmetrie auf den unteren Dreiecksteil beschränken kann, ergibt dann zum Beispiel

$$\begin{aligned} d_{1,1} &= a_{1,1} \\ \ell_{2,1} &= a_{2,1}/d_{1,1} \\ \ell_{3,1} &= a_{3,1}/d_{1,1} \\ d_{2,2} &= \ldots \end{aligned}$$

Konstruktion der Cholesky-Zerlegung

$$\ldots = \begin{pmatrix} d_{1,1} & \ldots & \ldots \\ \ell_{2,1}d_{1,1} & \ell_{2,1}^2 d_{1,1} + d_{2,2} & \ldots \\ \ell_{3,1}d_{1,1} & \ell_{2,1}\ell_{3,1}d_{1,1} + \ell_{3,2}d_{2,2} & \ell_{3,1}^2 d_{1,1} + \ell_{3,2}^2 d_{2,2} + d_{3,3} \end{pmatrix}$$

$$= \begin{pmatrix} a_{1,1} & \ldots & \ldots \\ a_{2,1} & a_{2,2} & \ldots \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix} = A$$

Für die aufeinanderfolgenden Spalten, $k = 1, 2, \ldots, n$ hat man explizite Formeln für $d_{k,k}$ und $\ell_{i,k} (i > k)$:

$$\begin{aligned} d_{k,k} &= a_{k,k} - \sum_{j=1}^{k-1} \ell_{k,j}^2 d_{j,j}, \\ \ell_{i,k} &= \left(a_{i,k} - \sum_{j=1}^{k-1} \ell_{i,j} d_{j,j} \ell_{k,j} \right) \Big/ d_{k,k}. \end{aligned}$$

Beispiel 3.35

Bestimme die Cholesky-Zerlegung von

$$A = \begin{pmatrix} 2 & 6 & -2 \\ 6 & 21 & 0 \\ -2 & 0 & 16 \end{pmatrix}$$

1. Spalte:

(1,1)-Element: $d_{1,1} = a_{1,1} = 2 \quad \Longrightarrow \quad \boxed{d_{1,1} = 2}$

(2,1)-Element: $\ell_{2,1} d_{1,1} = a_{2,1} = 6 \quad \Longrightarrow \quad \boxed{\ell_{2,1} = 3}$

(3,1)-Element: $\ell_{3,1} d_{1,1} = a_{3,1} = -2 \quad \Longrightarrow \quad \boxed{\ell_{3,1} = -1}$

2. Spalte:

(2,2)-Element: $\ell_{2,1}^2 d_{1,1} + d_{2,2} = a_{2,2} = 21 \quad \Rightarrow \quad \boxed{d_{2,2} = 3}$

(3,2)-Element: $\ell_{2,1} \ell_{3,1} d_{1,1} + \ell_{3,2} d_{2,2} = a_{3,2} = 0 \quad \Rightarrow \quad \boxed{\ell_{3,2} = 2}$

Beispiel 3.35

Bestimme die Cholesky-Zerlegung von

$$A = \begin{pmatrix} 2 & 6 & -2 \\ 6 & 21 & 0 \\ -2 & 0 & 16 \end{pmatrix}$$

3. Spalte:

(3,3)-Element: $\ell_{3,1}^2 d_{1,1} + \ell_{3,2}^2 d_{2,2} + d_{3,3} = a_{3,3} = 16$

$$\Longrightarrow \boxed{d_{3,3} = 2}$$

Damit ergibt sich schließlich

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ -1 & 2 & 1 \end{pmatrix} \text{ und } D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Programmentwurf Cholesky-Verfahren

Für $k = 1, 2, \ldots, n$:

diag $\leftarrow a_{k,k} - \sum_{j<k} a_{k,j}^2 a_{j,j}$;

falls diag $< 10^{-5} a_{k,k}$ Abbruch

$a_{k,k} \leftarrow$ diag,

für $i = k+1, \ldots, n$

$a_{i,k} \leftarrow \left(a_{i,k} - \sum_{j<k} a_{i,j} a_{j,j} a_{k,j}\right) / a_{k,k}$;

Rechenaufwand 3.36.

Man kann das Cholesky-Verfahren mit ca. $\frac{1}{6} n^3$ Multiplikationen und etwa ebenso vielen Additionen realisieren. Der Rechenaufwand beträgt also etwa die Hälfte des Aufwands der LR-Zerlegung.

Bemerkung 3.37

- LDL^T entspricht der LR-Zerlegung für $R = DL^T$. Bei s.p.d. Matrizen ist Pivotisierung weder nötig noch sinnvoll. Beachte, dass Pivotisierung die Symmetrie der Matrix zerstören würde.

- Die Lösung des Problems $Ax = b$ reduziert sich auf
$$L \underbrace{DL^T x}_{=y} = b, \text{ d.h. } Ly = b \text{ und } L^T x = D^{-1} y.$$

- In obiger Version enthält das Verfahren die Abfrage
$$\text{diag} < 10^{-5} a_{k,k}$$
Falls dies gilt, kann nicht mehr gewährleistet werden, dass das entsprechende Pivotelement strikt positiv ist. In diesem Sinne *testet* das Verfahren Positiv-Definitheit.

LR-Zerlegung ohne Pivotisierung

Satz

Sei $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ eine reguläre Matrix. Es wird eine LR-Zerlegung ohne Pivotisierung auf einem Rechner mit Maschinengenauigkeit $\mathbf{eps}$ berechnet. Wenn die LR-Zerlegung

$$\boldsymbol{A} = \boldsymbol{L}\,\boldsymbol{R}$$

existiert, dann erfüllen die berechneten Matrizen $\tilde{\boldsymbol{L}}$, $\tilde{\boldsymbol{R}}$

$$\tilde{\boldsymbol{L}}\,\tilde{\boldsymbol{R}} = \boldsymbol{A} + \boldsymbol{\Delta A}, \qquad \frac{\|\boldsymbol{\Delta A}\|}{\|\boldsymbol{L}\|\,\|\boldsymbol{R}\|} = \mathcal{O}(\mathbf{eps})$$

für ein $\boldsymbol{\Delta A} \in \mathbb{R}^{n\times n}$.

- LR-Zerlegung ist stabil, wenn $\|\boldsymbol{L}\|\,\|\boldsymbol{R}\| = \mathcal{O}(\|\boldsymbol{A}\|)$
- LR-Zerlegung kann instabil sein, wenn $\|\boldsymbol{L}\|\,\|\boldsymbol{R}\| \neq \mathcal{O}(\|\boldsymbol{A}\|)$

N3.10

LR-Zerlegung mit Pivotisierung

Satz

Sei $\boldsymbol{P\,A = L\,R}$ die LR-Zerlegung mit Pivotisierung einer Matrix $\boldsymbol{A} \in \mathbb{R}^{\boldsymbol{n\times n}}$, dann erfüllen die berechneten Matrizen $\boldsymbol{\tilde{L}}$, $\boldsymbol{\tilde{R}}$

$$\boldsymbol{\tilde{L}\,\tilde{R} = P\,A + \Delta A}, \qquad \frac{\|\boldsymbol{\Delta A}\|}{\|\boldsymbol{A}\|} = \mathcal{O}(\boldsymbol{\rho}\,\mathbf{eps})$$

für ein $\boldsymbol{\Delta A} \in \mathbb{R}^{\boldsymbol{n\times n}}$ und

$$\boldsymbol{\rho} = \frac{\max\limits_{i,j} |r_{i,j}|}{\max\limits_{i,j} |a_{i,j}|} \quad \text{(Wachstumsfaktor)}.$$

Annahme: $\boldsymbol{P}$ eindeutig, sonst $\boldsymbol{P \neq \tilde{P}}$.

- LR-Zerlegung mit Pivotisierung ist stabil, wenn $\boldsymbol{\rho \approx 1}$ N3.10
- Stabilität "in der Praxis" vs. theoretische Instabilität N3.11

Cholesky-Zerlegung

Satz

Sei $\boldsymbol{A = L\,D\,L^T}$ die Cholesky-Zerlegung einer s.p.d. $\boldsymbol{A} \in \mathbb{R}^{\boldsymbol{n \times n}}$, die auf einem Rechner mit Maschinengenauigkeit $\mathbf{eps}$ berechnet wird. Dann erfüllen die berechneten Matrizen $\boldsymbol{\tilde{L}}$, $\boldsymbol{\tilde{D}}$

$$\boldsymbol{\tilde{L}\,\tilde{D}\,\tilde{L}^T = A + \Delta A}, \qquad \frac{\|\boldsymbol{\Delta A}\|}{\|\boldsymbol{A}\|} = \mathcal{O}(\boldsymbol{c_n}\,\mathbf{eps})$$

für eine "kleine" Zahl $\boldsymbol{c_n}$, die nur von der Dimension von $\boldsymbol{A}$ abhängt.

- "Worst-case Instabilität" kann bei Cholesky-Zerlegung nicht auftreten
- Die Lösung eines s.p.d. Systems über das Cholesky-Verfahren ist stabil.

Beispiel 3.40.

Wir betrachten die Lösung von $A\,x = b$, wobei $A \in \mathbb{R}^{n\times n}$ eine sogenannte *Hilbertmatrix* ist

$$A = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{n} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \cdots & \frac{1}{n+1} \\ \vdots & \vdots & \vdots & & \vdots \\ \frac{1}{n} & \frac{1}{n+1} & \frac{1}{n+2} & \cdots & \frac{1}{2n-1} \end{pmatrix},$$

Sei $b = \left(\frac{1}{n}, \frac{1}{n+1}, \ldots, \frac{1}{2n-1}\right)^T$, d.h. $x = (0, 0, \ldots, 0, 1)^T$.

Für $n = 12$ erhält man mit dem Cholesky-Verfahren auf einer Maschine mit $\mathrm{eps} \approx 10^{-16}$ einen Fehler im berechneten Resultat von

$$\frac{\|x - \tilde{x}\|_\infty}{\|x\|_\infty} \approx 1.6 \times 10^{-2}.$$

Erklärung: Konditionszahl $\kappa_\infty(A) = \|A\|_\infty \|A^{-1}\|_\infty \approx 10^{16}.$

Matrix-Zerlegung

Aufgabe

Gegeben sei $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ ($\det \boldsymbol{A} \neq \boldsymbol{0}$) und $\boldsymbol{b} \in \mathbb{R}^{n}$, bestimme $\boldsymbol{x} \in \mathbb{R}^{n}$ so dass

$$\boldsymbol{A}\,\boldsymbol{x} = \boldsymbol{b}.$$

Vorgehensweise: Bestimme eine Faktorisierung (Zerlegung) von $\boldsymbol{A}$, so dass das Gleichungssystem "leichter" lösbar ist.

Verfahren:

- LR-Zerlegung: $\boldsymbol{A} = \boldsymbol{L}\,\boldsymbol{R}$, wobei $\boldsymbol{L}$ untere Dreiecksmatrix, $\boldsymbol{R}$ obere Dreiecksmatrix
- Cholesky-Zerlegung: $\boldsymbol{A} = \boldsymbol{L}\,\boldsymbol{D}\,\boldsymbol{L}^{\boldsymbol{T}}$, wobei $\boldsymbol{D}$ Diagonalmatrix
- QR-Zerlegung: $\boldsymbol{A} = \boldsymbol{Q}\,\boldsymbol{R}$, wobei $\boldsymbol{Q}$ orthogonale Matrix

Warum QR?

Orthogonale Matrizen

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

$$\boldsymbol{Q}^T \boldsymbol{Q} = \boldsymbol{I}.$$

Das bedeutet,

- die Spalten von $\boldsymbol{Q}$ bilden eine Orthonormalbasis des $\mathbb{R}^m$;
- die Inverse von $\boldsymbol{Q}$ ist einfach zu bestimmen

$$\boldsymbol{Q}^{-1} = \boldsymbol{Q}^T.$$

Warum QR?

QR-Zerlegung

Gegeben sei eine rechteckige Matrix $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, bestimme eine orthogonale Matrix $\boldsymbol{Q} \in \mathbb{R}^{m \times m}$ und eine obere Dreiecksmatrix $\boldsymbol{R} \in \mathbb{R}^{m \times n}$, so dass

$$\boldsymbol{A} = \boldsymbol{Q}\,\boldsymbol{R}.$$

- Lösung eines linear Gleichungssystems ($\boldsymbol{A} \in \mathbb{R}^{n \times n}$ regulär)

$$\boldsymbol{A}\,\boldsymbol{x} = \boldsymbol{b} \;\Leftrightarrow\; \boldsymbol{Q}\,\boldsymbol{R}\,\boldsymbol{x} = \boldsymbol{b} \;\Leftrightarrow\; \underbrace{\boldsymbol{R}\,\boldsymbol{x} = \underbrace{\boldsymbol{Q}^T\,\boldsymbol{b}}_{\text{Matrix-Vektor-Produkt}}}_{\text{Rückwärtseinsetzen}}$$

- QR-Zerl. auch für rechteckige Matrizen durchführbar

N3.12

Satz 3.41

Sei $Q \in \mathbb{R}^{n\times n}$ orthogonal, dann gilt:

(i) Q^T ist orthogonal.

(ii) $\|Qx\|_2 = \|x\|_2$ für alle $x \in \mathbb{R}^n$.

(iii) $\kappa_2(Q) = 1$.

(iv) Für beliebiges $A \in \mathbb{R}^{n\times m}$ bzw. $A \in \mathbb{R}^{m\times n}$, $m \in \mathbb{N}$ beliebig, gilt $\|A\|_2 = \|QA\|_2 = \|AQ\|_2$.

(v) Es gilt (für A wie vorhin) $\kappa_2(A) = \kappa_2(QA) = \kappa_2(AQ)$.

(vi) Sei $\tilde{Q} \in \mathbb{R}^{n\times n}$ orthogonal, dann ist $Q\tilde{Q}$ orthogonal.

Vergleich LR- und QR-Zerlegung

LR-Zerlegung

Wir haben

$$L_{n-1} \dots L_2 L_1 A = R$$

und damit

$$A = L\, R$$

mit

$$L = L_1^{-1} L_2^{-1} \dots L_{n-1}^{-1}$$

einer unteren Dreiecksmatrix.

QR-Zerlegung

Wir haben

$$Q_{n-1} \dots Q_2 Q_1 A = R$$

und damit

$$A = Q\, R$$

mit

$$\begin{aligned} Q &= Q_1^{-1} Q_2^{-1} \dots Q_{n-1}^{-1} \\ &= Q_1^T Q_2^T \dots Q_{n-1}^T \end{aligned}$$

einer orthogonalen Matrix.

. . . im Detail N3.13

Berechnung der QR-Zerlegung

Wir haben mehrere Möglichkeiten:

- **Gram-Schmidt Orthogonalisierung** N3.14

  Idee: Schrittweise Orthogonalisierung der Spalten von $\boldsymbol{A}$

- **Givens-Rotation** N3.15

  Idee: Schrittweise ebene Drehungen der Spalten von $\boldsymbol{A}$

- **Householder-Transformation** N3.15

  Idee: Schrittweise Spiegelungen der Spalten von $\boldsymbol{A}$

Givens-Rotationen

Grundaufgabe

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

Givens-Rotations-Matrix

N3.17

Die orthogonale Matrix $G_{i,k} \in \mathbb{R}^{m \times m}$ ist gegeben durch

$$G_{i,k} = \begin{array}{rc}
 & \begin{array}{ccccccccccc} & & & i\downarrow & & & & k\downarrow & & & \end{array} \\
\begin{array}{r} \\ \\ \\ i \rightarrow \\ \\ \\ \\ k \rightarrow \\ \\ \\ \\ \end{array} &
\begin{pmatrix}
1 & & & & & & & & & & \\
 & \ddots & & & & & & & & & \\
 & & 1 & & & & & & & & \\
 & & & c & 0 & \cdots & 0 & s & & & \\
 & & & 0 & 1 & & & 0 & & & \\
 & & & \vdots & 0 & \ddots & & \vdots & & & \\
 & & & 0 & & & 1 & 0 & & & \\
 & & & -s & 0 & \cdots & 0 & c & & & \\
 & & & & & & & & 1 & & \\
 & & & & & & & & & \ddots & \\
 & & & & & & & & & & 1
\end{pmatrix}
\end{array}$$

Givens-Rotations-Matrix

…und damit ergibt sich

$$\begin{pmatrix} 1 &&&&&&&&&& \\ & \ddots &&&&&&&&& \\ && 1 &&&&&&&& \\ &&& c & 0 & \cdots & 0 & s &&& \\ &&& 0 & 1 && & 0 &&& \\ &&& \vdots & 0 & \ddots && \vdots &&& \\ &&& 0 &&& 1 & 0 &&& \\ &&& -s & 0 & \cdots & 0 & c &&& \\ &&&&&&&& 1 && \\ &&&&&&&&& \ddots & \\ &&&&&&&&&& 1 \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_{i-1} \\ x_i \\ x_{i+1} \\ \vdots \\ x_{k-1} \\ x_k \\ x_{k+1} \\ \vdots \\ x_m \end{pmatrix} = \begin{pmatrix} x_1 \\ \vdots \\ x_{i-1} \\ r \\ x_{i+1} \\ \vdots \\ x_{k-1} \\ 0 \\ x_{k+1} \\ \vdots \\ x_m \end{pmatrix}$$

für $\quad r = \pm\sqrt{{x_i}^2 + {x_k}^2}, \quad c = \frac{x_i}{r}, \quad s = \frac{x_k}{r}.$

Beispiel 3.43

$$\begin{pmatrix} 4 \\ -3 \\ 1 \end{pmatrix} \overset{G_{1,2}}{\rightsquigarrow} \begin{pmatrix} 5 \\ 0 \\ 1 \end{pmatrix} \overset{G_{1,3}}{\rightsquigarrow} \begin{pmatrix} \sqrt{26} \\ 0 \\ 0 \end{pmatrix}$$

mit

$$G_{1,2} = \begin{pmatrix} \frac{4}{5} & -\frac{3}{5} & 0 \\ \frac{3}{5} & \frac{4}{5} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \text{ und } G_{1,3} = \begin{pmatrix} \frac{5}{\sqrt{26}} & 0 & \frac{1}{\sqrt{26}} \\ 0 & 1 & 0 \\ -\frac{1}{\sqrt{26}} & 0 & \frac{5}{\sqrt{26}} \end{pmatrix}.$$

N3.18

Beispiel 3.45

$$\begin{pmatrix} * & * & * \\ * & * & * \\ * & * & * \\ * & * & * \end{pmatrix} \overset{G_{1,2}}{\rightsquigarrow} \begin{pmatrix} \circledast & \circledast & \circledast \\ 0 & \circledast & \circledast \\ * & * & * \\ * & * & * \end{pmatrix} \overset{G_{1,3}}{\rightsquigarrow} \begin{pmatrix} \circledast & \circledast & \circledast \\ 0 & * & * \\ 0 & \circledast & \circledast \\ * & * & * \end{pmatrix} \overset{G_{2,3}}{\rightsquigarrow} \begin{pmatrix} * & * & * \\ 0 & \circledast & \circledast \\ 0 & 0 & \circledast \\ * & * & * \end{pmatrix}$$

$$\overset{G_{1,4}}{\rightsquigarrow} \begin{pmatrix} \circledast & \circledast & \circledast \\ 0 & * & * \\ 0 & 0 & * \\ 0 & \circledast & \circledast \end{pmatrix} \overset{G_{2,4}}{\rightsquigarrow} \begin{pmatrix} * & * & * \\ 0 & \circledast & \circledast \\ 0 & 0 & * \\ 0 & 0 & \circledast \end{pmatrix} \overset{G_{3,4}}{\rightsquigarrow} \begin{pmatrix} * & * & * \\ 0 & * & * \\ 0 & 0 & \circledast \\ 0 & 0 & 0 \end{pmatrix}$$

- Mit $\circledast$ werden die Einträge angedeutet, die bei der Anwendung von $G_{i,k}$ neu berechnet werden müssen.
- Die Reihenfolge $G_{1,2}, G_{1,3}, G_{1,4}, G_{2,3}, G_{2,4}, G_{3,4}$ wäre auch möglich.

Beispiel 3.46.

$$\begin{pmatrix} 3 & 5 \\ 0 & 2 \\ 0 & 0 \\ 4 & 5 \end{pmatrix} \underset{\rightsquigarrow}{G_{1,4}} \begin{pmatrix} 5 & 7 \\ 0 & 2 \\ 0 & 0 \\ 0 & -1 \end{pmatrix} \underset{\rightsquigarrow}{G_{2,4}} \begin{pmatrix} 5 & 7 \\ 0 & \sqrt{5} \\ 0 & 0 \\ 0 & 0 \end{pmatrix},$$

wobei

$$G_{1,4} = \begin{pmatrix} \frac{3}{5} & 0 & 0 & \frac{4}{5} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\frac{4}{5} & 0 & 0 & \frac{3}{5} \end{pmatrix}, \text{ und } G_{2,4} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{2}{\sqrt{5}} & 0 & -\frac{1}{\sqrt{5}} \\ 0 & 0 & 1 & 0 \\ 0 & \frac{1}{\sqrt{5}} & 0 & -\frac{2}{\sqrt{5}} \end{pmatrix}.$$

QR-Zerlegung über Givens-Rotation

Die obige Konstruktion mit Givens-Rotationen zeigt, dass für *jede* Matrix $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ eine QR-Zerlegung *existiert*.

Satz 3.47.

Gegeben sei $\boldsymbol{A} \in \mathbb{R}^{m\times n}$. Dann existiert eine orthogonale Matrix $\boldsymbol{Q} \in \mathbb{R}^{m\times m}$ und eine obere Dreiecksmatrix $\boldsymbol{R} \in \mathbb{R}^{m\times n}$ mit

$$\boldsymbol{A} = \boldsymbol{Q}\,\boldsymbol{R},$$

wobei $\boldsymbol{R}$ im Sinne von N3.12 zu verstehen ist.

Merke: Bei der Implementierung der QR-Zerlegung über Givens-Rotation werden die Matrizen $\boldsymbol{G}_{i,k}$ *nie explizit berechnet*.

Givens-Rotation: Zusammenfassung

QR-Zerlegung über Givens-Rotationen:

- Das Verfahren ist *sehr stabil*. Pivotisierung ist *nicht* erforderlich.
- Durch Berücksichtigung von schon vorhandenen $\mathbf{0}$-Einträgen bei dünnbesetzten Matrizen läßt sich das Verfahren flexibel an die Struktur einer Matrix anpassen.
- Der Aufwand für die QR-Zerlegung einer vollbesetzten $m \times n$-Matrix über Givens-Rotationen beträgt etwa $\frac{4}{3}n^3$ Operationen, falls $m \approx n$, und etwa $2mn^2$ Operationen, falls $m \gg n$. Zu beachten ist aber, dass für dünnbesetzte Matrizen der Aufwand wesentlich niedriger ist.
- Bei der sogenannten *schnellen* Givens-Rotationen wird der Aufwand etwa halbiert ($\sim \frac{2}{3}n^3$, falls $n \approx m; \sim mn^2$, falls $m \gg n$).

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

N3.19

Definition

Für $v = (v_1, \ldots, v_n)^T \in \mathbb{R}^n, v \neq 0$ 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

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

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

$$Q_v^{-1} = Q_v^T$$

- Geometrische Interpretation: Spiegelung

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

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

gilt.

N3.20 D:SC

- **Die Lösung der Grundaufgabe ist:**

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

- Um Auslöschung zu vermeiden, wählt man

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

Zusammenfassend

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

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 $\alpha = 3$, $v = y + \alpha e^1 = (5,2,1)^T$, und damit $Q_v y = (-3,0,0)^T$.
- Beachte: explizite Form von Q_v wird zur Berechnung von $Q_v y$ **nicht** benötigt

$$Q_v = I - 2\frac{vv^T}{v^T v} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - 2\frac{\begin{pmatrix} 5 \\ 2 \\ 1 \end{pmatrix}(5,2,1)}{(5,2,1)\begin{pmatrix} 5 \\ 2 \\ 1 \end{pmatrix}} = \frac{1}{15}\begin{pmatrix} -10 & -10 & -5 \\ -10 & 11 & -2 \\ -5 & -2 & 14 \end{pmatrix}$$

Reduktion auf obere Dreiecksform

N3.21

- 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, d.h., man setzt

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

  und berechnet $Q_1 A$.
- Sei $\tilde{a}^{(2),1}$ 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

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

mittels Householder-Transformation.

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

$$\boldsymbol{v}^1 = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} + 3\boldsymbol{e}^1 = \begin{pmatrix} 4 \\ 2 \\ 2 \end{pmatrix}, \quad \boldsymbol{Q}_1 = \boldsymbol{Q}_{\boldsymbol{v}^1}.$$

2. Für die zwei Spalten der Matrix $\boldsymbol{Q}_1\boldsymbol{A}$ ergibt sich

$$\boldsymbol{Q}_1 \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} = \begin{pmatrix} -3 \\ 0 \\ 0 \end{pmatrix} \qquad \text{(siehe Beispiel 3.51)}$$

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 nächstes Kapitel), Berechnung von Eigenwerten

Zusammenfassung lineare Gleichungssysteme

- Cholesky-Zerlegung: $A = LDL^T$
  - Nur für symmetrisch positiv definite Matrizen ("symmetrische" LR-Zerlegung)
  - Berechnung durch elementweise Auswertung von $A = LDL^T$
- QR-Zerlegung: $A = QR$
  - Existiert auch für rechteckige Matrizen
  - Berechnung über Givens-Rotation, Housholder-Transformation (oder Gram-Schmidt)
- Übersicht Lösungsverfahren N3.22