Numerische Mathematik für Elektrotechniker

Lineare Gleichungssysteme III

Benjamin Berkels

Karl-Heinz Brakhage, Thomas Jankuhn, Christian Löbbert

Institut für Geometrie und Praktische Mathematik
RWTH Aachen

Wintersemester 2019/2020

Zusammenfassung der letzten Vorlesung

Definition 3.31. (Symmetrisch postiv definite Matrix)

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

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

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

$$\boldsymbol{x}^T \boldsymbol{A} \boldsymbol{x} > \boldsymbol{0} \qquad \text{(positiv definit)}.$$

- $\boldsymbol{A}$ ist invertierbar, und $\boldsymbol{A}^{-1}$ ist s.p.d.
- $\boldsymbol{A}$ hat nur strikt positive (insbesondere reelle) Eigenwerte.
- $\boldsymbol{A}$ hat nur strikt positive Diagonaleinträge und der betragsgrößte Eintrag von $\boldsymbol{A}$ liegt auf der Diagonalen.
- Bei Gauß-Elimination ohne Pivotisierung sind alle Pivotelemente strikt positiv.

Heute in der Vorlesung

Themen: Dahmen & Reusken Kap. 3.6-3.9

- Cholesky-Verfahren
- Stabilität der LR- und Cholesky-Zerlegung
- QR-Zerlegung
- Methoden: Givens und Householder

Was Sie mitnehmen sollten:

- Was ist die Cholesky-Zerlegung und wie wird sie bestimmt?
- Sind die Gauß-Elimination und das Cholesky-Verfahren stabil?
- Was ist die QR-Zerlegung und wie wird sie berechnet?

Matrix-Zerlegung

Aufgabe

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

$$A\,x = b.$$

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

Verfahren:

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

Satz 3.34.

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 normierte 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.

Definition

Die obige Zerlegung $A = LDL^T$ heißt *Choleskey-Zerlegung*.

Beispiel 3.35. (Konstruktion der Cholesky-Zerlegung)

$$A = \begin{pmatrix} 2 & 6 & -2 \\ 6 & 21 & 0 \\ -2 & 0 & 16 \end{pmatrix}, L = \begin{pmatrix} 1 & 0 & 0 \\ \ell_{2,1} & 1 & 0 \\ \ell_{3,1} & \ell_{3,2} & 1 \end{pmatrix}, D = \begin{pmatrix} d_{1,1} & 0 & 0 \\ 0 & d_{2,2} & 0 \\ 0 & 0 & d_{3,3} \end{pmatrix}.$$

Es gilt

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

Die elementare Auswertung der Gleichung $LDL^T = A$ kann man aufgrund der Symmetrie auf den unteren Dreiecksteil beschränken.

Beispiel 3.35.

$$\begin{pmatrix} 2 & * & * \\ 6 & 21 & * \\ -2 & 0 & 16 \end{pmatrix} = \begin{pmatrix} d_{1,1} & 0 & 0 \\ \ell_{2,1}d_{1,1} & d_{2,2} & 0 \\ \ell_{3,1}d_{1,1} & \ell_{3,2}d_{2,2} & d_{3,3} \end{pmatrix} \begin{pmatrix} 1 & \ell_{2,1} & \ell_{3,1} \\ 0 & 1 & \ell_{3,2} \\ 0 & 0 & 1 \end{pmatrix}$$

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

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

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

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

$\Rightarrow d_{2,2} = 21 - 9 \cdot 2 \quad \Rightarrow \boxed{d_{2,2} = 3}$

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

$\Rightarrow \ell_{3,2} = -3 \cdot (-1) \cdot 2/3 \quad \Rightarrow \boxed{\ell_{3,2} = 2}$

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

$\Rightarrow d_{3,3} = 16 - (-1)^2 \cdot 2 - 2^2 \cdot 3 \Rightarrow \boxed{d_{3,3} = 2}$

Beispiel 3.35.

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

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

(3,1)-Element: $\ell_{3,1} d_{1,1} = a_{3,1} \Rightarrow \ell_{3,1} = \frac{a_{3,1}}{d_{1,1}} \Rightarrow \boxed{\ell_{3,1} = -1}$

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

$\Rightarrow d_{2,2} = a_{2,2} - \ell_{2,1}^2 d_{1,1} \Rightarrow \boxed{d_{2,2} = 3}$

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

$\Rightarrow \ell_{3.2} = -a_{3,2} - \ell_{2,1}\ell_{3,1} d_{1,1}/d_{2,2} \Rightarrow \boxed{\ell_{3,2} = 2}$

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

$\Rightarrow d_{3,3} = a_{3,3}\text{-}(\ell_{3,1}^2 d_{1,1} + \ell_{3,2}^2 d_{2,2}) \Rightarrow \boxed{d_{3,3} = 2}$

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

$$\begin{pmatrix} d_{1,1} & & \\ l_{2,1} d_{1,1} & d_{2,2} & \\ l_{3,1} d_{1,1} & l_{3,2} d_{2,2} & d_{3,3} \end{pmatrix} \begin{pmatrix} 1 & l_{2,1} & l_{3,1} \\ & 1 & l_{3,2} \\ & & 1 \end{pmatrix}$$

$$a_{kk} = (LDL^T)_{kk} = \sum_{j=1}^{k-1} l_{k,j} d_{j,j} l_{k,j} + d_{k,k}$$

$$i > k \qquad a_{ik} = \sum_{j=1}^{k-1} l_{i,j} d_{j,j} l_{k,j} + l_{i,k} d_{k,k}$$

Cholesky-Verfahren

Ergebnis für die Einträge von L und D

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

$$d_{k,k} = a_{k,k} - \sum_{j=1}^{k-1} \ell_{k,j}^2 d_{j,j},$$

$$\ell_{i,k} = \frac{a_{i,k} - \sum_{j=1}^{k-1} \ell_{i,j} d_{j,j} \ell_{k,j}}{d_{k,k}}$$

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.

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

- Die Lösung des Problems $\boldsymbol{Ax = b}$ reduziert sich auf

$$\boldsymbol{L \underbrace{DL^T x}_{=y} = b}, \text{ d.h. } \boldsymbol{Ly = b} \text{ und } \boldsymbol{L^T x = D^{-1} y}.$$

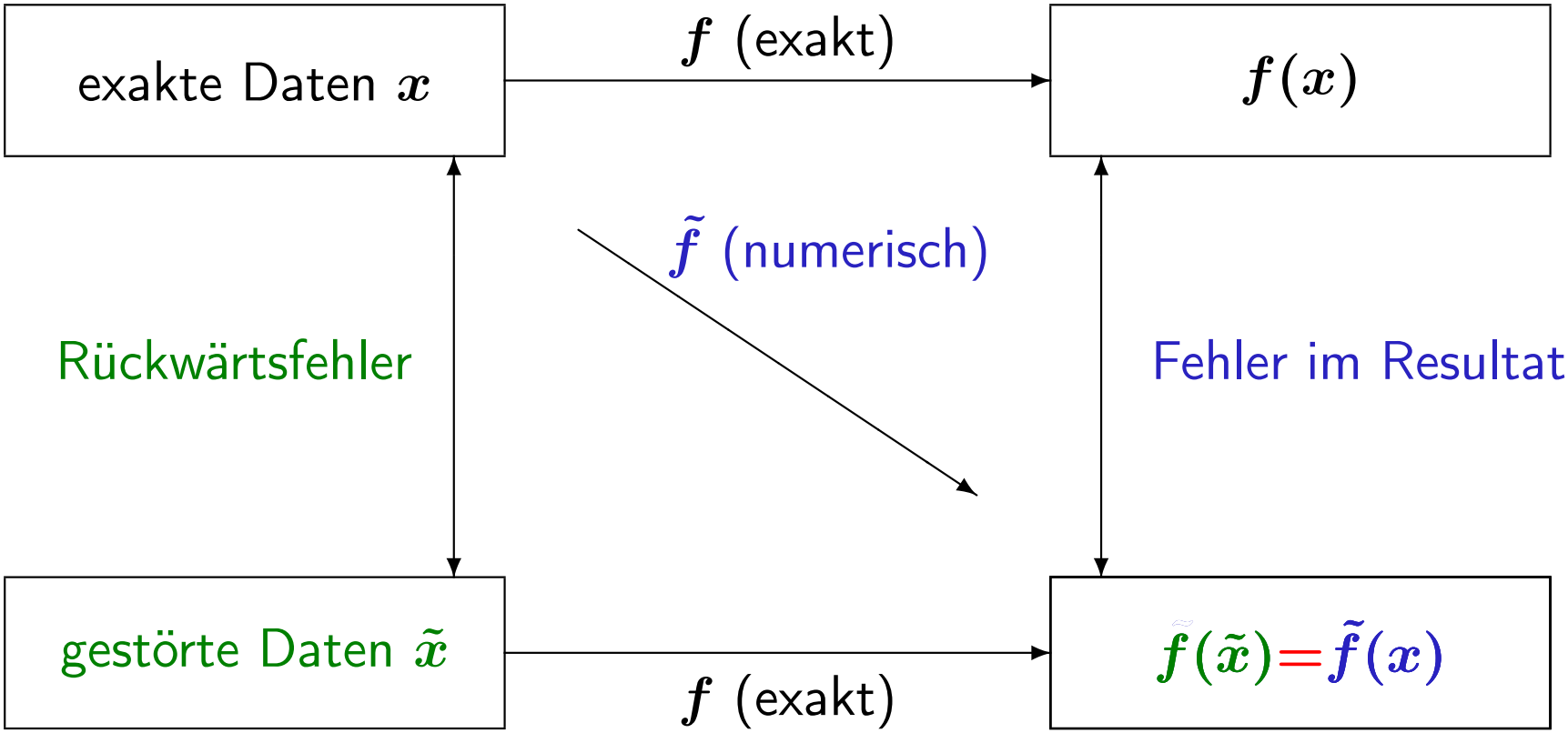
- In obiger Version enthält das Verfahren die Abfrage

$$\mathbf{diag} < \mathbf{10^{-5}} \boldsymbol{a_{k,k}}$$

Falls dies gilt, kann nicht mehr gewährleistet werden, dass das entsprechende Pivotelement strikt positiv ist.
In diesem Sinne testet das Verfahren Positiv-Definitheit.

Stabilitätsanalyse Cholesky-Verfahren und Gauß-Elimination

Wiederholung: Rückwärtsanalyse



Stabilitätsanalyse Cholesky-Verfahren und Gauß-Elimination

Nach dem Prinzip der Rückwärtsanalyse wird das Ergebnis der Rechnung als Ergebnis exakter Rechnung zu gestörten Eingabedaten $\boldsymbol{A} + \boldsymbol{\Delta A}$ interpretiert.
Sei

$$\boldsymbol{Ax} = \boldsymbol{b}$$

mit $\boldsymbol{A}$ symmetrisch positiv definit.
Das Cholesky-Verfahren wird eingesetzt.
Man kann zeigen, dass die berechnete Lösung $\tilde{\boldsymbol{x}}$ die exakte Lösung eines Systems

$$(\boldsymbol{A} + \boldsymbol{\Delta A})\tilde{\boldsymbol{x}} = \boldsymbol{b}$$

ist, mit

$$\frac{\|\boldsymbol{\Delta A}\|_\infty}{\|\boldsymbol{A}\|_\infty} \lesssim \boldsymbol{c_n} \cdot \mathbf{eps}.$$

Hierbei ist $\boldsymbol{c_n}$ eine "kleine" Zahl und **eps** die Maschinengenauigkeit.

Stabilitätsanalyse Cholesky-Verfahren und Gauß-Elimination

Deshalb ist das Resultat $\tilde{x}$ mit einem Fehler von der Größenordnung des durch die Kondition bedingten unvermeidbaren Fehlers behaftet:

$$\frac{\|\tilde{x}-x\|_\infty}{\|x\|_\infty} \leq \frac{\kappa_\infty(A)\frac{\|\Delta A\|_\infty}{\|A\|_\infty}}{1-\kappa_\infty(A)\frac{\|\Delta A\|_\infty}{\|A\|_\infty}} \lesssim \frac{\kappa_\infty(A)c_n\mathrm{eps}}{1-\kappa_\infty(A)c_n\mathrm{eps}}$$

$$\approx \kappa_\infty(A)c_n\mathrm{eps}$$

wenn $\kappa_\infty(A)c_n\mathrm{eps} \ll 1$.

Damit ist das Lösen eines s.p.d. Systems über das Cholesky-Verfahren **stabil**.

Stabilitätsanalyse Cholesky-Verfahren und Gauß-Elimination

Falls $\boldsymbol{A}$ nicht symmetrisch positiv definit ist, kann man zur Lösung des Problems $\boldsymbol{Ax} = \boldsymbol{b}$ auf eine $\boldsymbol{LR}$-Zerlegung der Matrix $\boldsymbol{A}$ zurückgreifen. Dazu wird die Gauß-Elimination eingesetzt.

Falls Gauß-Elimination **mit Spaltenpivotisierung** angewendet wird, kann man Resultate wie beim Cholesky-Verfahren zeigen.

Damit kann man die Lösung eines linearen Gleichungssystems über die Gauß-Elimination mit Spaltenpivotisierung als ein **stabiles** Verfahren einstufen (für praktische Anwendungen).

Die Lösung über die Gauß-Elimination ohne Pivotisierung ist im allgemeinen nicht stabil.

Hilbert-Matrix

$$H_n = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{n} \\ \frac{1}{2} & \frac{1}{3} & & & \frac{1}{n+1} \\ \frac{1}{3} & & \ddots & \iddots & \vdots \\ \vdots & & \iddots & \frac{1}{2n-3} & \frac{1}{2n-2} \\ \frac{1}{n} & \frac{1}{n+1} & \cdots & \frac{1}{2n-2} & \frac{1}{2n-1} \end{pmatrix}$$

$n \times n$-Matrix
regulär
symmetrisch
positiv-definit

$$x := \begin{pmatrix} 3 \\ 5 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^n, \qquad \tilde{b} := H_n * x, \qquad \text{(in Matlab, } \mathbf{eps} \approx 10^{-16}\text{)}$$

$H_n \tilde{x} = \tilde{b}$ gelöst mit stabiler Matlab-Methode. Matlab-Demo

Matrix-Zerlegung

Aufgabe

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

$$A\,x = b.$$

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

Verfahren:

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

Warum QR?

Orthogonale Matrizen

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

$$Q^T Q = I.$$

Das bedeutet,

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

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

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-Zerlegung auch für rechteckige Matrizen durchführbar.

Satz 3.41.

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

(i) Q^T ist orthogonal.

$$(Q^T)^T Q^T = Q\, Q^{-1} = I$$

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 $x \in \mathbb{R}^n$.

$$\|Qx\|_2^2 = \langle Qx, Qx \rangle = \langle Q^T Qx, x \rangle = \langle x, x \rangle = \|x\|_2^2$$

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) $\kappa_2(Q) = 1$.

$$\|Q\|_2 = \max_{x \neq 0} \frac{\|Qx\|_2}{\|x\|_2} = 1$$

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.

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 $\boldsymbol{A}$

Givens-Rotationen

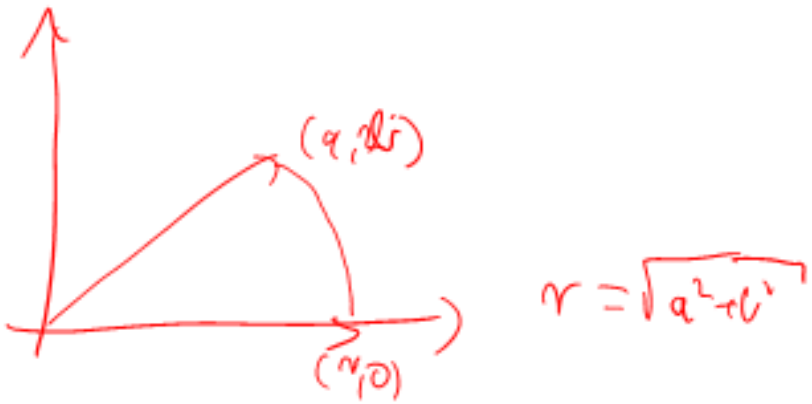
Grundaufgabe

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

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



Givens-Rotationen

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.

Givens-Rotations-Matrix

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

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

Givens-Rotations-Matrix

. . . und damit ergibt sich

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

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

Beispiel 3.43.

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

$r = \pm\sqrt{16+9} = \pm 5$

$c = \frac{4}{5} \quad s = \frac{-3}{5}$

mit

$c = \frac{a}{r}$, $s = \frac{b}{r}$, $r = \pm\sqrt{a^2+b^2}$

Beispiel 3.43.

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

mit

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

$r = \pm\sqrt{26}$

$c = \frac{5}{\sqrt{26}}$ $s = \frac{1}{\sqrt{26}}$

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

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

Beispiel 3.45.

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

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

- 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.

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

wobei

$$G_{1,4} = \begin{pmatrix} \frac{3}{5} & 0 & 0 & \frac{4}{5} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\frac{4}{5} & 0 & 0 & \frac{3}{5} \end{pmatrix},$$

$(1,2)$: $\frac{3}{5} 5 + \frac{4}{5} \cdot 5 = 7$

Beispiel 3.46.

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

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

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

Merke:

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

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-Rotation wird der Aufwand etwa halbiert $\left(\sim \frac{2}{3}n^3, \text{ falls } n \approx m; \quad \sim mn^2, \text{ falls } m \gg n \right)$.

Berechnung der QR-Zerlegung

Wir haben mehrere Möglichkeiten:

- **Gram-Schmidt Orthogonalisierung**
  Idee: Schrittweise Orthogonalisierung der Spalten von A
- **Givens-Rotation**
  Idee: Schrittweise ebene Drehungen der Spalten von A
- **Householder-Transformation**
  Idee: Schrittweise Spiegelungen der Spalten von A

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

$$= I - 4\frac{vv^T}{v^Tv} + 4\frac{\overbrace{vv^Tvv^T}^{=v^Tvv}}{(v^Tv)^2} = I$$

Householder-Transformationen

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 (Tensorprodukt), $v\,v^T$, gegeben ist durch

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

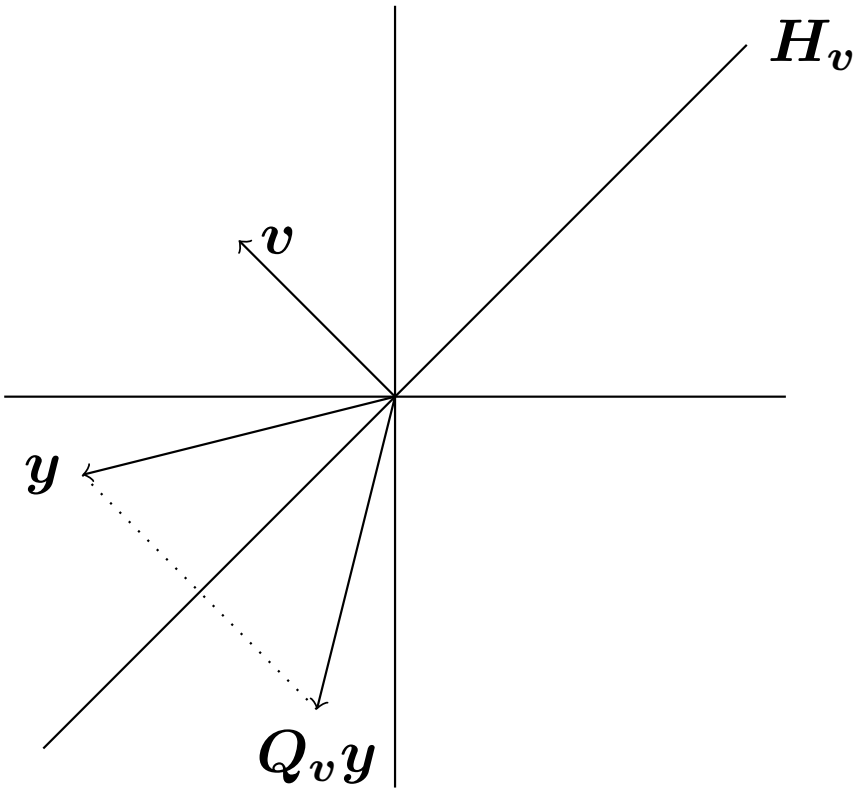
- Die Householder Transformation Q_v ist orthogonal, d.h.

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

Householder-Transformationen

Geometrische Interpretation: Spiegelung an der Ebene normal zu v

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



Householder-Transformationen

Eigenschaften 3.50.

- $Q_v = Q_v^T$ symmetrisch, da $(vv^T)^T = vv^T$
- $Q_v^2 = I$

  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.
- $Q_v y = y \iff y^T v = 0$

  Ursprünglicher 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 gilt:

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

- Die Lösung der Grundaufgabe ist:

$$v = y \mp \|y\|_2 \cdot e^1$$

$$a a^T y = \left(\sum_{k=1}^{m} a_i a_k y_k \right)_i = a^T y \; a$$

Householder-Transformationen

Grundaufgabe

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

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

- Die Lösung der Grundaufgabe ist:

$$v = y \mp \|y\|_2 \cdot e^1$$

- Um Auslöschung zu vermeiden, wählt man

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

Zusammenfassend

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

Beispiel 3.51.

Aufgabe

Zu $y = \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}$ wird $v \in \mathbb{R}^3$ gesucht, so dass gilt:

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

$v = y \mp \|y\|_2 e^1 = \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix} \mp 3 e^1 = \begin{pmatrix} 2 \mp 3 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 5 \\ 2 \\ 1 \end{pmatrix}$

Beispiel 3.51.

Aufgabe

Zu $y = \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}$ wird $v \in \mathbb{R}^3$ gesucht, so dass gilt:

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

- Wir erhalten $\alpha = 3$, und $v = y + \alpha \cdot e^1 = \begin{pmatrix} 5 \\ 2 \\ 1 \end{pmatrix}$, und damit

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

Beispiel 3.51.

Zur Berechnung von $Q_v y$ wird die explizite Form von Q_v

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

nicht benötigt.

Beachte

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

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:

$$v^1 = a^1 + \mathbf{sign}(a_{1,1}) \cdot \|a^1\|_2 \cdot e^1 \;\rightarrow\; Q_1 := Q_{v^1} \;\rightarrow\; Q_1 A$$

- Sei $\tilde{a}^{(2),1}$ die erste Spalte der Matrix $\tilde{A}^{(2)}$ ...

$A = A^{(1)}$

$$\begin{pmatrix} * & * & \cdots & \cdots & * \\ * & * & \cdots & \cdots & * \\ \vdots & \vdots & & & \vdots \\ \vdots & \vdots & & & \vdots \\ * & * & \cdots & \cdots & * \end{pmatrix}$$

$Q_1 A = A^{(2)}$

$$\begin{pmatrix} * & * & \cdots & \cdots & * \\ 0 & * & \cdots & \cdots & * \\ \vdots & \vdots & \tilde{A}^{(2)} & & \vdots \\ \vdots & \vdots & & & \vdots \\ 0 & * & \cdots & \cdots & * \end{pmatrix}$$

$Q_2 Q_1 A = A^{(3)}$

$$\begin{pmatrix} * & * & * & \cdots & * \\ 0 & * & * & \cdots & * \\ 0 & 0 & * & \cdots & * \\ \vdots & \vdots & \vdots & \tilde{A}^{(3)} & \vdots \\ 0 & 0 & * & \cdots & * \end{pmatrix}$$

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

Beispiel 3.52.

Bestimmen Sie die QR-Zerlegung der Matrix

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

mittels Householder-Transformation.

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

$$v^1 = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} + 3 \cdot e^1 = \begin{pmatrix} 4 \\ 2 \\ 2 \end{pmatrix} \rightarrow Q_1 = Q_{v^1}.$$

2. Für die zwei Spalten der Matrix $Q_1 A$ ergibt sich

$$Q_1 \cdot \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} = \begin{pmatrix} -3 \\ 0 \\ 0 \end{pmatrix} \quad \text{(siehe Beispiel 3.51.)}$$

Beispiel 3.52.

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

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 $\tilde{A}^{(2)}$, d.h. $y = (-\frac{2}{3}, -\frac{2}{3})^T$, ergibt

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

Beispiel 3.52.

$$\widetilde{Q}_2 \begin{pmatrix} -\frac{2}{3} \\ -\frac{2}{3} \end{pmatrix} = \begin{pmatrix} -\frac{2}{3} \\ -\frac{2}{3} \end{pmatrix} - \frac{2 \cdot (v^2)^T \begin{pmatrix} -\frac{2}{3} \\ -\frac{2}{3} \end{pmatrix}}{(v^2)^T v^2} v^2 = \begin{pmatrix} \frac{2}{3}\sqrt{2} \\ 0 \end{pmatrix}$$

5. Damit ergibt sich
$$\widetilde{Q}_2 \begin{pmatrix} -\frac{2}{3} \\ -\frac{2}{3} \end{pmatrix} = \begin{pmatrix} \frac{2}{3}\sqrt{2} \\ 0 \end{pmatrix}$$

6. Insgesamt erhält man
$$\underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & & \\ 0 & & \widetilde{Q}_2 \end{pmatrix}}_{Q_2} Q_1 A = \begin{pmatrix} -3 & -\frac{1}{3} \\ 0 & \frac{2}{3}\sqrt{2} \\ 0 & 0 \end{pmatrix}.$$

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

- Cholesky-Zerlegung: $\boldsymbol{A = LDL^T}$
  - Nur für symmetrisch positiv definite Matrizen ("symmetrische" LR-Zerlegung)
  - Berechnung durch elementweise Auswertung von $\boldsymbol{A = LDL^T}$: Cholesky-Verfahren
- QR-Zerlegung: $\boldsymbol{A = QR}$
  - Existiert auch für rechteckige Matrizen
  - Berechnung über Givens-Rotation, Housholder-Transformation (oder Gram-Schmidt)

Verständnisfragen

f Für die Matrix $A = \begin{pmatrix} -1 & 0 \\ 0 & 3 \end{pmatrix}$ ex. eine Cholesky-Zerlegung

Es sei $A = LDL^T$ mit $L = \begin{pmatrix} 1 & 0 \\ 10 & 1 \end{pmatrix}$ und $D = \begin{pmatrix} 2 & 0 \\ 0 & 2.5 \end{pmatrix}$.

Geben Sie $\det A^{-1}$ an. **0.2**

w Für jede symmetrische orthogonale Matrix Q gilt $Q^2 = I$.

Es seien $v, x \in \mathbb{R}^2$ mit $v \neq 0$, $x = \begin{pmatrix} -1 \\ 3 \end{pmatrix}$, und $Q_v = I - 2\frac{vv^T}{v^T v}$ eine Householder-Transformation.

Geben Sie $\|Q_v x\|_2^2$ an. **10**