

Numerische Mathematik für Maschinenbauer

Lineare Gleichungssysteme

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. 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: Übersicht

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

$$Ax = b.$$

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

Verfahren:

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

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 normierte untere Dreiecksmatrix und D eine Diagonalmatrix mit Diagonaleinträgen

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

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.

Konstruktion der Cholesky-Zerlegung

Beispiel 3.35.

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

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

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

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

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

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

$$(2,2)\text{-Element: } \ell_{2,1}^2 d_{1,1} + d_{2,2} = a_{2,2} = 21 \\ \Rightarrow d_{2,2} = 21 - 9 * 2 \Rightarrow \boxed{d_{2,2} = 3}$$

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

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

Beispiel 3.35

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

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

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

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

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

$$(3,3)\text{-Element: } \ell_{3,1}^2 d_{1,1} + \ell_{3,2}^2 d_{2,2} + d_{3,3} = 16 \quad \Rightarrow \quad \boxed{d_{3,3} = 2}$$

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

Beispiel 3.35

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

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

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

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

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

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

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

Cholesky-Verfahren

Ergebnis für die Einträge von L und D

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

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

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

Programmwurf Cholesky-Verfahren

Für $k = 1, 2, \dots, n$:

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

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

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

für $i = k + 1, \dots, n$

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

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

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

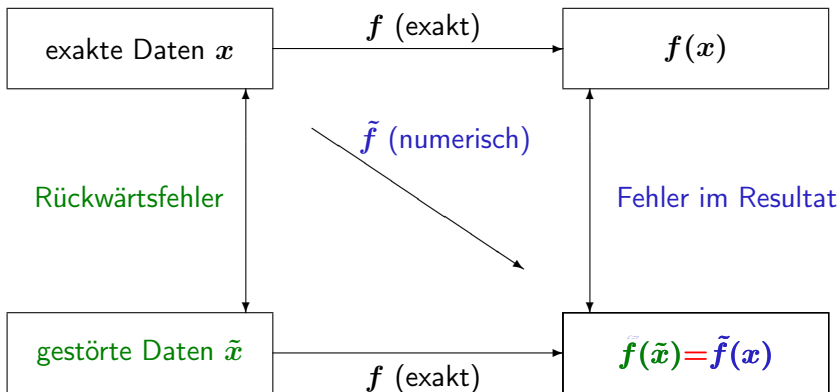
- ▶ In obiger Version enthält das Verfahren die Abfrage

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

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

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 $A + \Delta A$ interpretiert.

Sei

$$Ax = b$$

mit A symmetrisch positiv definit. Das Cholesky-Verfahren wird eingesetzt.

Man kann zeigen, dass die berechnete Lösung \tilde{x} die exakte Lösung eines Systems

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

ist, mit

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

Hierbei ist c_n eine "kleine" Zahl und eps die Maschinengenauigkeit.

Stabilitätsanalyse Cholesky-Verfahren und Gauß-Elimination

Deshalb ist das Resultat \tilde{x} mit einem Fehler behaftet, der in der Größenordnung des durch die Kondition des Problems bedingten **unvermeidbaren** Fehlers bleibt:

$$\begin{aligned} \frac{\|\tilde{x} - x\|_\infty}{\|x\|_\infty} &\leq \frac{\kappa_\infty(A) \frac{\|\Delta A\|_\infty}{\|A\|_\infty}}{1 - \kappa_\infty(A) \frac{\|\Delta A\|_\infty}{\|A\|_\infty}} \lesssim \frac{\kappa_\infty(A) c_n \text{eps}}{1 - \kappa_\infty(A) c_n \text{eps}} \\ &\approx \kappa_\infty(A) c_n \text{eps}, \quad \text{wenn } \kappa_\infty(A) c_n \text{eps} \ll 1. \end{aligned}$$

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

Stabilitätsanalyse Cholesky-Verfahren und Gauß-Elimination

Falls A nicht symmetrisch positiv definit ist, kann man zur Lösung des Problems $Ax = b$ auf eine LR -Zerlegung der Matrix 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.

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

Hilbertmatrix

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

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

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

$H_n \tilde{x} = \tilde{b}$ gelöst mit stabiler MATLAB-Methode

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

$$Ax = b.$$

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

Verfahren:

- ▶ LR-Zerlegung: $PA = LR$, wobei L untere Dreiecksmatrix, R obere Dreiecksmatrix
- ▶ Cholesky-Zerlegung: $A = LDL^T$, wobei D Diagonalmatrix
- ▶ QR-Zerlegung: $A = QR$, 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 $A \in \mathbb{R}^{m \times n}$, bestimme eine orthogonale Matrix $Q \in \mathbb{R}^{m \times m}$ und eine obere Dreiecksmatrix $R \in \mathbb{R}^{m \times n}$, so dass

$$A = QR.$$

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

$$Ax = b \Leftrightarrow QRx = b \Leftrightarrow Rx = \underbrace{Q^T b}_{\substack{\text{Matrix-Vektor-} \\ \text{Produkt}}}$$

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.
- (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 A

- ▶ **Givens-Rotation**

Idee: Schrittweise ebene Drehungen der Spalten von A

- ▶ **Householder-Transformation**

Idee: Schrittweise Spiegelungen der Spalten von 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

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

$$G_{i,k} = \begin{matrix} & & i \downarrow & & k \downarrow & & \\ & & & & & & \\ & & & & & & \\ & & & & & & \\ i \rightarrow & \begin{pmatrix} 1 & & & & & & \\ & \dots & & & & & \\ & & 1 & & & & \\ & & & c & 0 & \dots & 0 & s \\ & & & 0 & 1 & & & 0 \\ & & & \vdots & 0 & \ddots & & \vdots \\ & & & 0 & & & 1 & 0 \\ k \rightarrow & & & -s & 0 & \dots & 0 & c \\ & & & & & & & & 1 \\ & & & & & & & & & \dots \\ & & & & & & & & & & 1 \end{pmatrix} & & & & & & \end{matrix}$$

Beispiel 3.43

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

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

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

- ▶ 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.

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

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 $A \in \mathbb{R}^{m \times n}$ eine **QR-Zerlegung existiert**.

Satz 3.47.

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

$$A = Q R.$$

Merke: Bei der Implementierung der QR-Zerlegung über Givens-Rotation werden die Matrizen $G_{i,k}$ **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 **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** ($\sim \frac{2}{3}n^3$, 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 A

- ▶ **Givens-Rotation**

Idee: Schrittweise ebene Drehungen der Spalten von A

- ▶ **Householder-Transformation**

Idee: Schrittweise Spiegelungen der Spalten von 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

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

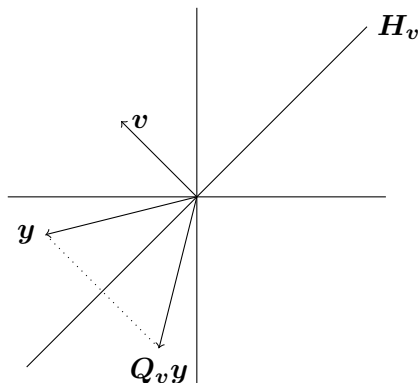
- Die Householder Transformation Q_v ist orthogonal, d.h.

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

Householder-Transformationen

Geometrische Interpretation: Spiegelung

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



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.

▶ $Q_v y = y \iff y^T v = 0$

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

$$Q_v \mathbf{y} = \pm \|\mathbf{y}\|_2 \mathbf{e}^1$$

gilt.

- Die Lösung der Grundaufgabe ist:

$$\mathbf{v} = \mathbf{y} \pm \|\mathbf{y}\|_2 \mathbf{e}^1$$

- Um Auslöschung zu vermeiden, wählt man

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

Zusammenfassend

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

Beispiel 3.51

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

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

gilt.

- Wir erhalten $\alpha = 3$, $\mathbf{v} = \mathbf{y} + \alpha \mathbf{e}^1 = (5, 2, 1)^T$, und damit $Q_v \mathbf{y} = (-3, 0, 0)^T$.

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

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

Zusammenfassung

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

Verständnisfragen

Für die Matrix A existiert eine Cholesky-Zerlegung mit

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

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.

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.