Numerische und stochastische Grundlagen

Institut für Parallele und Verteilte Systeme
Universität Stuttgart

Wintersemester 2011/2012

Lösung linearer Gleichungssysteme

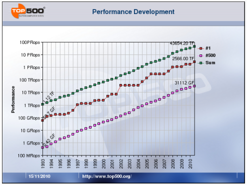
Häufigste Teilaufgabe der numerischen Simulation

Zu gegebenem $b \in \mathbb{R}^N$ (rechte Seite) und $A \in \mathbb{R}^{N \times N}$ (Systemmatrix), finde $x \in \mathbb{R}^N$, so dass gilt

$$A\,x = b.$$

(Koeffizientenbestimmung zu gewählter Basis)

- Interpolation
- Differentialgleichungen
- Integralgleichungen
- Lösung nicht-linearer Gleichungen



Linpack Benchmark für Super-Computer

http://www.top500.org/
http://www.netlib.org/benchmark/hpl/

Lösbarkeit

$$A\,x = b \qquad A \in \mathbb{R}^{M\times N},\ b \in \mathbb{R}^M,\ x \in \mathbb{R}^N$$

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,N} \\ a_{2,1} & a_{2,2} & \dots & a_{2,N} \\ \vdots & \ddots & \ddots & \vdots \\ a_{M,1} & a_{M,2} & \dots & a_{M,N} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_M \end{pmatrix}$$

- heißt ***unterbestimmt***, falls $M < N$
- heißt ***quadratisch***, falls $M = N$
- heißt ***überbestimmt***, falls $M > N$

Lösbarkeit

Für beliebige M und N ist $A\,x = b$ lösbar, wenn

$$\mathrm{Rang}(A) = \mathrm{Rang}((Ab))$$

(nicht notwendigerweise eindeutig!).

Lösbarkeit

Für quadratische Matrizen ($N = M$) sind folgende Aussagen äquivalent

- $A\,x = b$ ist für jedes $b \in \mathbb{R}$ eindeutig lösbar, d.h. ***regulär***.
- $\det(A) \neq 0$
- $\mathrm{Rang}(A) = N$
- Für alle Eigenwerte $\lambda \in \sigma(A)$ von A gilt $\lambda \neq 0$.

Normen

Normen auf dem $\mathbb{R}^N$

Eine Abbildung $\|\cdot\| : \mathbb{R}^N \to \mathbb{R}_+$ heißt Norm, wenn gilt

- $\|x\| > 0$ für alle $x \in \mathbb{R}^N \setminus 0$
- $\|\alpha\, x\| = |\alpha| \|x\|$ für alle $x \in \mathbb{R}^N$, $\alpha \in \mathbb{R}$
- $\|x + y\| \leq \|x\| + \|y\|$ für alle $x, y \in \mathbb{R}^N$

Betragssummennorm $\|x\|_1 := \sum_{i=1}^{N} |x_i|$

Euklidische Norm $\|x\|_2 := \sqrt{\sum_{i=1}^{N} |x_i|^2} = \sqrt{\boldsymbol{x}^\star \boldsymbol{x}}$

Maximumsnorm $\|x\|_\infty := \max_{i=1,\ldots,N} |x_i|$

Auf dem $\mathbb{R}^N$ sind alle Normen äquivalent.

Matrixnormen

Matrixnormen

Eine Norm $\|\cdot\|_O : \mathbb{R}^{M\times N} \to \mathbb{R}_+$ heißt Matrixnorm.
Im Fall $N = M$ heißt sie
verträglich mit der Norm $\|\cdot\|_V : \mathbb{R}^N \to \mathbb{R}_+$, wenn für alle $x \in \mathbb{R}^N$ gilt

$$\|Ax\|_V \leq \|A\|_O \, \|x\|_V.$$

.

Sie heißt submultiplikativ, wenn für alle A, B gilt

$$\|A\cdot B\|_O \leq \|A\|_O \|B\|_O.$$

Spaltensummenorm $\|A\|_1 := \max_{j=1,\ldots N} \sum_{i=1}^{M} |a_{i,j}|$

Zeilensummennorm $\|A\|_\infty := \max_{i=1,\ldots M} \sum_{j=1}^{N} |a_{i,j}|$

Frobeniusnorm $\|A\|_F := \sqrt{\sum_{i,j=1}^{M,N} |a_{i,j}|^2}$

Die Frobeniusnorm ist mit der Euklidischen Norm verträglich.

Induzierte Matrixnorm

$$\|A\| := \sup_{x\neq 0} \frac{\|A\,x\|}{\|x\|} = \sup_{\|x\|=1} \|A\,x\|$$

Diese ist verträglich mit $\|\cdot\|$ und submultiplikativ.

Spektralradius einer Matrix

$$\sigma(A) := \{\lambda \in \mathbb{C} \mid \det(A - \lambda\mathbb{I}) = 0\}$$

$$Av = \lambda v$$

$$\rho(A) := \max\{|\lambda| \mid \lambda \in \sigma(A)\}$$

Induzierte Euklidische Norm $\|A\|_2$

$$\|A\|_2 := \max_{\|x\|_2=1} \|Ax\| = \max_{\|x\|_2=1} \sqrt{(Ax)^\star(Ax)} = \max_{\|x\|_2=1} \sqrt{x^\star A^\star Ax}$$

$$\|A\|_2 = \sqrt{\rho(A^\star A)}$$

Kondition $A\,x = b$

Kondition einer Matrix

$$\kappa_M(A) := \mathrm{cond}_M := \|A\|_M \|A^{-1}\|_M$$

Einfacher Störungssatz

Es sei $A \in \mathbb{R}^{N\times N}$ regulär. Für eine gestörte Eingabe $b + \Delta b$ gilt für die Störung der Ausgabe $x + \Delta x$

$$\frac{\|\Delta x\|}{\|x\|} \leq \kappa(A) \frac{\|\Delta b\|}{\|b\|}.$$

$$(x+\Delta x) = A^{-1}(b+\Delta b), \quad \Delta x = A^{-1}\Delta b, \quad \frac{\|\Delta x\|}{\|x\|} \leq \|A\| \|A^{-1}\| \frac{\|\Delta b\|}{\|b\|}$$

Störung in A und b

Allgemeiner Störungssatz

Es sei $A \in \mathbb{R}^{N\times N}$ regulär und es gelte $\|\Delta A\| < \|A^{-1}\|^{-1}$, dann gilt für die Lösung des gestörten Problems

$$(A + \Delta A)(x + \Delta x) = (b + \Delta b)$$

die Abschätzung

$$\frac{\|\Delta x\|}{\|x\|} \leq \frac{\kappa(A)}{1 - \kappa(A)\frac{\|\Delta A\|}{\|A\|}} \left(\frac{\|\Delta b\|}{\|b\|} + \frac{\|\Delta A\|}{\|A\|} \right)$$

für ein beliebiges Paar verträglicher Matrix/Vektornormen.

Beispiel

Eingabefehler:

$$\frac{\|\Delta A\|}{\|A\|} \approx 10^{-k}, \quad \frac{\|\Delta b\|}{\|b\|} \approx 10^{-k}$$

Kondition von A:

$$\kappa(A) \approx 10^s \text{ mit } s \ll k$$

Ausgabefehler:

$$\frac{\|\Delta x\|}{\|x\|} \leq \frac{10^s}{1 - 10^s \cdot 10^{-k}} \cdot 2 \cdot 10^{-k} \approx 10^{s-k}$$

Verlust von s Stellen Genauigkeit!

Lösungsverfahren für $A\,x = b$

Direkte Verfahren

Realisierung von $A^{-1}b = x$ mit endlich vielen Operationen, z.B. mittels

$$A = SBT \qquad A^{-1} = T^{-1}B^{-1}S^{-1}$$

Leicht lösbare Gleichungssysteme

Diagonale Systemmatrix (lösbar in $O(N)$):

$$A\,x = b; \qquad A := \begin{pmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3 \end{pmatrix}, \quad b := \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

Dreiecks-Systemmatrix (lösbar in $O(N^2)$):

$$A\,x = b; \qquad A := \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ 0 & a_{2,2} & a_{2,3} \\ 0 & 0 & a_{3,3} \end{pmatrix}, \quad b := \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

Tridiagonal-Matrizen

Betrachte das Gleichungssystem

$$a_i x_{i-1} + b_i x_i + c_i x_{i+1} = d_i, \quad \text{mit } a_1 = 0 \text{ und } c_n = 0.$$

In Matrix-Vektor Notation

$$\begin{pmatrix} b_1 & c_1 & & & 0 \\ a_2 & b_2 & c_2 & & \\ & a_3 & b_3 & \cdot & \\ & & \cdot & \cdot & c_{n-1} \\ 0 & & & a_n & b_n \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_n \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ \cdot \\ \cdot \\ d_n \end{pmatrix}$$

Lösung Tridiagonaler Systeme

Die Unbekannten seien $x_1, \ldots, x_n$. Die zu lösenden Gleichungen seien:

$$b_1 x_1 + c_1 x_2 = d_1; \qquad i = 1 \tag{1}$$

$$a_i x_{i-1} + b_i x_i + c_i x_{i+1} = d_i; \qquad i = 2, \ldots, n-1 \tag{2}$$

$$a_n x_{n-1} + b_n x_n = d_n; \qquad i = n \tag{3}$$

Die zweite Gleichung ($i = 2$) wird nun wie folgt mit der ersten Gleichung modifiziert:

$$(\text{Gleichung } 2) \cdot b_1 - (\text{Gleichung } 1) \cdot a_2$$

Dies ergibt:

$$\begin{aligned} (a_2 x_1 + b_2 x_2 + c_2 x_3) b_1 - (b_1 x_1 + c_1 x_2) a_2 &= d_2 b_1 - d_1 a_2 \\ (b_2 b_1 - c_1 a_2) x_2 + c_2 b_1 x_3 &= d_2 b_1 - d_1 a_2 \end{aligned}$$

Dadurch wurde x_1 aus der zweiten Gleichung eliminiert. Analog ergibt sich aus der modifizierten zweiten und der dritten Gleichung mit $A := (b_2 b_1 - c_1 a_2)$

$$\begin{aligned} (a_3 x_2 + b_3 x_3 + c_3 x_4) A - (A x_2 + c_2 b_1 x_3) a_3 &= d_3 A - (d_2 b_1 - d_1 a_2) a_3 \\ (b_3 A - c_2 b_1 a_3) x_3 + c_3 A x_4 &= d_3 A - (d_2 b_1 - d_1 a_2) a_3 \end{aligned}$$

Jetzt wurde x_2 eliminiert. Etc. etc.

Thomas-Algorithmus

Definiere Hilfsvariablen

$$c'_i = \begin{cases} \dfrac{c_1}{b_1} & ; \quad i = 1 \\ \dfrac{c_i}{b_i - c'_{i-1} a_i} & ; \quad i = 2, 3, \dots, n-1 \end{cases}$$

und

$$d'_i = \begin{cases} \dfrac{d_1}{b_1} & ; \quad i = 1 \\ \dfrac{d_i - d'_{i-1} a_i}{b_i - c'_{i-1} a_i} & ; \quad i = 2, 3, \dots, n \end{cases}$$

Die Lösung ergibt sich dann durch Rückwärts-Einsetzen:

$$\begin{aligned} x_n &= d'_n \\ x_i &= d'_i - c'_i x_{i+1} \qquad i = n-1, n-2, \dots, 1 \end{aligned}$$

Aufwand

$$5(N-1) + 1 + 3(N-1) = O(N)$$

Direkte Gleichungslöser für volle Matrizen

Gleichungssystem

$$\begin{array}{rcrcrcr} 5x_1 & + & 2x_2 & + & x_3 & = & 1 \\ 2x_1 & + & 4x_2 & + & 2x_3 & = & 6 \\ x_1 & + & x_2 & + & 4x_3 & = & 3 \end{array}$$

in Matrix-Vektor Notation

$$A\,x = b; \qquad A := \begin{pmatrix} 5 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 1 & 4 \end{pmatrix}, \quad b := \begin{pmatrix} 1 \\ 6 \\ 3 \end{pmatrix}$$

oder kompakt

$$\left(\begin{array}{ccc|c} 5 & 2 & 1 & 1 \\ 2 & 4 & 2 & 6 \\ 1 & 1 & 4 & 3 \end{array} \right)$$

$$\left(\begin{array}{ccc|c} 5 & 2 & 1 & 1 \\ 2 & 4 & 2 & 6 \\ 1 & 1 & 4 & 3 \end{array}\right)$$

Elimination von x_1 in Zeile 2, 3 über $-\frac{2}{5}, -\frac{1}{5}$-fache von Zeile 1

$$\left(\begin{array}{ccc|c} 5 & 2 & 1 & 1 \\ 0 & \frac{16}{5} & \frac{8}{5} & \frac{28}{5} \\ 0 & \frac{3}{5} & \frac{19}{5} & \frac{14}{5} \end{array}\right)$$

Elimination von x_2 in Zeile 3 über $-\frac{3}{16}$-fache von Zeile 2

$$\left(\begin{array}{ccc|c} 5 & 2 & 1 & 1 \\ 0 & \frac{16}{5} & \frac{8}{5} & \frac{28}{5} \\ 0 & 0 & \frac{280}{80} & \frac{140}{80} \end{array}\right)$$

Rückwärts-Substitution

$$x_3 = \frac{140}{80}\frac{80}{280} = \frac{1}{2}, \quad x_2 = \frac{5}{16}\left(\frac{28}{5} - \frac{8}{5}\cdot\frac{1}{2}\right) = \frac{3}{2}, \quad x_1 = \frac{1}{5}\left(1 - 1\cdot\frac{1}{2} - 2\cdot\frac{3}{2}\right) = -\frac{1}{2}.$$

Gauss-Elimination

Idee

Transformiere das Gleichungssystem $A\,x = b$ äquivalent zu $T\,x = \tilde{b}$ mit oberer Dreiecksmatrix T.

Verfahren

$A \in \mathbb{R}^{N\times N}$ mit $\det(A) \neq 0$ und $b \in \mathbb{R}^N$. Setze $S^{(1)} := A$, $\tilde{b}^{(1)} := b$ und es gelte $S^{(r)}_{r,r} \neq 0$ für $r = 1, \ldots, N$, dann liefert

$$l_{i,r} := (S^{(r)}_{r,r})^{-1} S^{(r)}_{i,r}, \quad S^{(r+1)}_{i,j} := S^{(r)}_{i,j} - l_{i,r} S^{(r)}_{r,j}, \quad \tilde{b}^{(r+1)}_i := \tilde{b}^{(r)}_i - l_{i,r} b^{(r)}_r$$

für $i, j = r+1, \ldots, N$ und $r = 1, \ldots, N-1$ die Lösung $x \in \mathbb{R}^N$ von $A\,x = b$ mittels

$$x_N = (S^{(N)}_{N,N})^{-1} \tilde{b}^{(N)}_N, \quad x_i = (S^{(N)}_{i,i})^{-1} \Big(\tilde{b}^{(N)}_i - \sum_{j=i+1}^{N} S^{(N)}_{i,j} x_j \Big)$$

für $i = N-1, \ldots, 1$.

Beobachtung

Verfahrenschritte:

$$
\begin{array}{ccccccc}
S^{(1)} & \to & \begin{pmatrix} S^{(1)}_{1,j} & j=1,\ldots,N \\ 0 & \\ \vdots & S^{(2)} \\ 0 & \end{pmatrix} & \to & \begin{pmatrix} S^{(1)}_{1,j} & & j=1,\ldots,N \\ 0 & S^{(2)}_{2,j} & j=2,\ldots,N \\ \vdots & 0 & \\ 0 & \vdots & S^{(3)} \\ 0 & 0 & \end{pmatrix} & \to \\
T^{(1)} & \to & T^{(2)} & \to & T^{(3)} & \to
\end{array}
$$

Für $r = 1, \ldots, N-1$ gilt

$$
T^{(r+1)}_{i,j} = \begin{cases} 0 & j = 1, \ldots, r; i = j+1, \ldots, N \\ S^{(i)}_{i,j} & i = 1, \ldots, r; j = i, \ldots, N \\ S^{(r+1)}_{i,j} & i = r+1, \ldots, N; j = r+1, \ldots, N \end{cases}
$$

LU-Zerlegung

Es gilt

$$T^{(r+1)} = L^{(r)} T^{(r)}$$

mit der Matrix

$$L^{(r)} := \begin{pmatrix} 1 & 0 & & & & & \\ 0 & \ddots & 0 & & & & \\ & 0 & 1 & 0 & & & \\ & & -l_{r+1,r} & \ddots & & & \\ & & -l_{r+2,r} & 0 & \ddots & & \\ & & \vdots & 0 & \dots & \ddots & \\ & & -l_{N,r} & 0 & \dots & 0 & 1 \end{pmatrix}$$

Diese Multiplikation eliminiert die Einträge $T_{i,r}^{(r)}$ für $i = r+1, \dots, N$ und $T^{(N)}$ ist schließlich eine obere Dreiecksmatrix.

LU-Zerlegung

Es gilt daher offensichtlich

$$U := T^{(N)} = L^{(N-1)} L^{(N-2)} \dots L^{(1)} A$$

Wir können also schreiben

$$A = (L^{(N-1)} L^{(N-2)} \dots L^{(1)})^{-1} U = (L^{(1)})^{-1} \cdots (L^{(N-1)})^{-1} U,$$

falls $(L^{(r)})^{-1}$ existiert. Es gilt

$$\begin{pmatrix} 1 & 0 & & & & & \\ 0 & \ddots & 0 & & & & \\ & 0 & 1 & 0 & & & \\ & & -l_{r+1,r} & \ddots & & & \\ & & -l_{r+2,r} & 0 & \ddots & & \\ & & \vdots & 0 & \dots & \ddots & \\ & & -l_{N,r} & 0 & \dots & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & & & & & \\ 0 & \ddots & 0 & & & & \\ & 0 & 1 & 0 & & & \\ & & l_{r+1,r} & \ddots & & & \\ & & l_{r+2,r} & 0 & \ddots & & \\ & & \vdots & 0 & \dots & \ddots & \\ & & l_{N,r} & 0 & \dots & 0 & 1 \end{pmatrix} = \mathbb{I}$$

LU-Zerlegung

Des weiteren gilt

$$(L^{(r)})^{-1}(L^{(r+1)})^{-1} = \begin{pmatrix} 1 & 0 & & & & & \\ 0 & \ddots & 0 & & & & \\ & 0 & 1 & 0 & & & \\ & & l_{r+1,r} & 1 & & & \\ & & l_{r+2,r} & l_{r+2,r+1} & \ddots & & \\ & & \vdots & \vdots & \dots & \ddots & \\ & & l_{N,r} & l_{N,r+1} & \dots & 0 & 1 \end{pmatrix}$$

Und daher ist $L := (L^{(1)})^{-1} \cdots (L^{(N-1)})^{-1}$ eine untere Dreiecksmatrix mit 1 auf der Diagonale und wir erhalten die Zerlegung

$$A = LU$$

mit zwei leicht zu lösenden (invertierbaren) Dreiecksmatrixen L & U.

Aufwand & Pivotisierung

Aufwand der *LU*-Zerlegung

$$\sum_{k=1}^{N-1} .... = \frac{2}{3}N^3 + O(N^2)$$

Pivotisierung

Eine wesentliche Annahme in obiger Konstruktion war $\tilde{T}_{r,r}^{(r)} \neq 0$ für $r = 1, \ldots, N$. Dies ist nicht automatisch garantiert für beliebige Anordnungen einer regulären Matrix, aber es gibt immer einfache Permutationsmatrizen P und Q, so dass

$$A = PLUQ$$

P vertauscht Zeilen und Q vertauscht Spalten der Matrix.

Cholesky-Zerlegung & Inversenbildung

Cholesky Zerlegung

Ist A symmetrisch ($A^T = A$) und positiv definit ($x^T Ax > 0$ mit $x \neq 0$), dann gilt

$$A^T = A = LU = L(DL^T) = (LD^{\frac{1}{2}})(D^{\frac{1}{2}}L^T) = (LD^{\frac{1}{2}})(LD^{\frac{1}{2}})^T = \tilde{L}\tilde{L}^T$$

ohne Pivotisierung!

Inversenbildung

Berechne $A = LU$ und löse für $i = 1, \dots, N$ die Gleichungssysteme $LU\alpha_i = e_i$. Es gilt α_i ist die i-te Spalte von A^{-1}. Der Aufwand ist $\frac{2}{3}N^3 + N \cdot 2N^2 = \frac{8}{3}N^3$.

Definition

Zu gegebenem $b \in \mathbb{R}^M$ und $A \in \mathbb{R}^{M \times N}$ mit $M \geq N$, heißt das Problem: Finde $x \in \mathbb{R}^N$, so dass gilt

$$\|b - A\,x\|_2 \to \min$$

lineares Ausgleichsproblem. Die Lösung heißt Ausgleichslösung oder kleinste-Quadrate-Lösung.

Direkte Verallgemeinerung des klassischen Lösungsbegriffs ($M = N$).

$$\|b - A\,x\|_2 \to \min \Leftrightarrow A\,x = b$$

Normalengleichungen

Satz

Die Ausgleichslösungen $\|b - A\,x\|_2 \to \min$ sind genau die Lösungen der Gaußschen Normalengleichungen

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

insbesondere existiert eine Lösung x. Falls z auch eine Lösung ist, dann gilt $A\,z = A\,x$. Das Residuum

$$r = b - A\,x$$

ist also eindeutig bestimmt und erfüllt

$$A^\star\, r = 0.$$

Eigenschaften der Systemmatrix

Satz

Die Matrix $A^\star A$ ist symmetrisch und positiv semidefinit. Sie ist genau dann positiv definit, wenn der Kern von A trivial ist, d.h. $ker(A) = \{0\}$. Dies gilt genau dann, wenn die Spalten von A linear unabhängig sind.

Lösung der Normalengleichung

Prinzipiell mittels Cholesky-Zerlegung möglich, falls $ker(A) = \{0\}$. Dies kann aber zu Stabilitätsproblemen führen. Betrachte $M = N$ und A^{-1} existiere, dann gilt für A

$$\mathrm{cond}_2(A) = \|A\|_2 \|A^{-1}\|_2 = \sqrt{\rho(A^\star A)}\sqrt{\rho((A^{-1})^\star A^{-1})} = \sqrt{\frac{\lambda_{\max}(A^\star A)}{\lambda_{\min}(A^\star A)}}$$

aber für die Normalengleichung

$$\mathrm{cond}_2(A^\star A) = \frac{\lambda_{\max}(A^\star A)}{\lambda_{\min}(A^\star A)}.$$

Ausgleichsproblem: Betrachte A oder $A^\star A$?

Gestalt von A

$A \in \mathbb{R}^{M \times N}$ ist rechteckig mit $M \geq N$ Zeilen und N Spalten. Wir nehmen an, dass $\mathrm{Rang}(A) = N$. So ein A könnte also folgende Gestalt haben:

$$A_0 = \begin{pmatrix} q & r & s \\ v & u & t \\ w & h & j \\ e & v & n \\ z & p & k \end{pmatrix}, \quad A_1 = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & j \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} \alpha & \beta & \gamma \\ 0 & \delta & \eta \\ 0 & 0 & \tau \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Unitäre Matrizen

Definition

Eine Matrix $Q \in \mathbb{R}^{N \times N}$ heißt unitär, falls gilt

$$Q^\star Q = \mathbb{I}.$$

Die Spalten von Q bilden also eine Orthonormalbasis des $\mathbb{R}^N$.

Eigenschaften:

- $\|Qx\|_2^2 = \|x\|_2^2$ für alle x
- $\|Q\|_2 = \|Q^\star\|_2 = \|Q^{-1}\|_2 = 1$
- $\mathrm{cond}_2(Q) = 1$
- P und Q unitär, dann auch PQ

Householder-Transformation

Definition

Sei $v \neq 0$, dann heißt die Matrix

$$P = \mathbb{I} - \frac{2}{\|v\|_2^2} v v^\star$$

Householder-Transformation zu v.

Eigenschaften:

- $P = P^\star$
- $P^2 = \mathbb{I}$
- $P\, v = -v$
- $P\, w = w$ für alle $w \perp v$

Transformation $x \to e_1$

Lemma

Gegeben sei $x \neq 0$ und $x \notin \operatorname{span}(e_1)$. Für

$$v := x + \sigma e_1, \quad \sigma := \frac{x_1}{|x_1|}\|x\|_2$$

(falls $x_1 = 0$, setze $\sigma := \|x\|_2$) gilt

$$(\mathbb{I} - \frac{2}{\|v\|_2^2} v v^\star)x = -\sigma e_1.$$

Satz

Sei $A \in \mathbb{R}^{M \times N}$ mit $\operatorname{Rang}(A) = N$, dann existiert ein unitäres Q und eine rehcte obere Dreiecksmatrix R, so dass

$$A = QR.$$

Eine rechte obere Dreiecksmatrix hat die Gestalt

$$R = \begin{pmatrix} \alpha & \beta & \gamma \\ 0 & \delta & \eta \\ 0 & 0 & \tau \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Methode der Orthogonalisierung

Satz

Sei $A \in \mathbb{R}^{M \times N}$ mit $\mathrm{Rang}(A) = N$ mit $A = QR$, wobei $R = \begin{pmatrix} \tilde{R} \\ 0 \end{pmatrix}$.
Dann gilt für beliebiges $b \in \mathbb{R}^M$ mit

$$Q^\star b =: \begin{pmatrix} c \\ d \end{pmatrix},$$

dass die Lösung des Ausgleichsproblems $\|b - A\,x\|_2 \to \min$ durch

$$Rx = c$$

eindeutig bestimmt ist. Für das Residuum $r = b - A\,x$ gilt $\|r\|_2 = \|d\|_2$.

Wir können also das Ausgleichsproblems mit Hilfe der QR-Zerlegung von A ohne Verschlechterung der Kondition lösen.