

Technische Universität Ilmenau
Fakultät für Mathematik
und Naturwissenschaften
Institut für Mathematik
<http://www.tu-ilmenau.de/math>

Postfach 10 05 65
D - 98684 Ilmenau
Germany
Tel.: 03677/69 3267
Fax: 03677/69 3272
Telex: 33 84 23 tuil d.
email: werner.neundorf@tu-ilmenau.de

Preprint No. M 13/05

Zur Konvergenz des Gradientenverfahrens

Werner Neundorf

September 2005

‡MSC (2000): 65F10, 65G50, 65-05, 65Y20

Zusammenfassung

Die Entwicklung moderner numerischer Algorithmen hat zu einem hohen Bedarf an effizienten, robusten iterativen Gleichungssystemlösern geführt. So entstand eine Vielzahl von Verfahren, die man zur Gruppe der Projektionsmethoden und Krylov-Unterraum-Methoden zählt.

Gegenstand der Betrachtungen ist das Gradientenverfahren als spezielles und grundlegendes Abstiegsverfahren und dazu die zahlreichen Aspekte seiner Konvergenz. Unter der üblichen Voraussetzung der Symmetrie und positiven Definitheit der Koeffizientenmatrix des linearen Gleichungssystems werden Konvergenzfragen und -untersuchungen zusammengestellt, erläutert und an Beispielen illustriert.

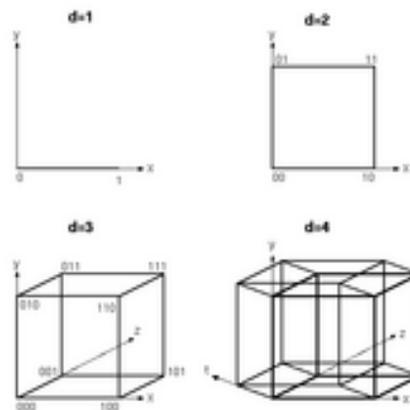
Dabei nehmen wir Bezug auf zahlreiche Hinweise, Vergleiche und Ergebnisse zu verschiedenen Abstiegsverfahren unter unterschiedlichen Voraussetzungen, wie sie Leser in den Preprints Abstiegsverfahren Teil I (Preprint No. M 19/04 IfMath TUI), II (Preprint No. M 20/04) und III (Preprint No. M 09/05) findet.

Richard Wesley Hamming (1915-1998)

Amerikanischer Mathematiker, der mit seinen Arbeiten wesentlichen Einfluss auf die Computerwissenschaften und Telekommunikation nahm. Begründer der algebraischen Codierungstheorie

Der Sinn von Berechnungen
liegt darin, Einsichten zu gewinnen
und nicht Zahlen.

Numerical Methods for Scientists and Engineers.
McGraw Hill 1973 bzw. Dover Books On Engineering archive 1986.
Coding and Information Theory. Prentice-Hall 1980, 1986.



Hyperwürfel mit $d = 1$ bis $d = 4$

Inhaltsverzeichnis

1	Funktionale und Abstiegsszenarien	1
2	Grundlagen der linearen Algebra	6
3	Grundlagen der Abstiegsverfahren	13
3.1	Ansätze für quadratische Formen	14
4	Abstiegsverfahren	17
4.1	Die quadratische Form $Q(x)$	17
4.2	Das Gradientenverfahren	22
4.2.1	Konvergenz des GV mittels Richardson-Extrapolation	30
4.2.2	Krylov-Unterraum und Suchrichtungen	40
4.2.3	Konvergenz des GV mittels Kantorovich-Ungleichung	43
4.3	Beispiele zum Gradientenverfahren	46
4.3.1	Beispiel 1	46
4.3.2	Beispiel 2	51
4.3.3	Beispiel 3	54
	Literaturverzeichnis	58

Kapitel 1

Funktionale und Abstiegszenarien

Wir betrachten im \mathbb{R}^n das lineare Gleichungssystem (LGS)

$$Ax = b \tag{1.1}$$

mit der regulären Matrix $A \in \mathbb{R}^{n,n} = (a_{ij})_{i,j=1}^n$ und der exakten Lösung $x^* \in \mathbb{R}^n$.

Die Überführung in ein Optimierungsproblem ist verknüpft mit der Definition geeigneter zu minimierender Funktionale, deren eindeutige Minimumstelle eben diese Lösung x^* sein sollte.

Unsere Vorstellungen im dreidimensionalen Raum sagen uns natürlich schon, dass Figuren wie ein Rotationsparaboloid mit kreis- oder ellipsenförmigen Querschnitt (Höhenlinien) eine eindeutige Minimumstelle besitzen. Aber auch andere Figuren, wie eine flache Schale oder die Bananenschale, besitzen ein Minimum. Die Oberflächenstruktur eines Geländeabschnitts hat möglicherweise ein globales Minimum, aber auch zahlreiche lokale Minima.

Als zu minimierende Funktionale nehmen wir die folgenden drei Funktionen. Ihre Entstehung, Bedeutung und Eigenschaften werden später noch eingehender untersucht. Zunächst wollen wir sie in unterschiedlichen Beispielen anwenden.

Die Funktionale leiten wir aus dem LGS ab.

Ihre einfachste direkt mit (1.1) verknüpfte Form ist

$$\begin{aligned} Q(x) &= \frac{1}{2}x^T Ax - x^T b \\ &= \frac{1}{2} \sum_{i,j=1}^n a_{ij} x_i x_j - \sum_{i=1}^n b_i x_i. \end{aligned} \tag{1.2}$$

Es gilt $Q(x^*) = \frac{1}{2}x^{*T}(Ax^*) - x^{*T}b = \frac{1}{2}x^{*T}b - x^{*T}b = -\frac{1}{2}x^{*T}b = -\frac{1}{2}x^{*T}Ax^*$.

Dieses Funktional ist die Basis für das Gradientenverfahren (GV).

Transformiert man das LGS (1.1) durch Multiplikation mit A^T , ohne dabei die Lösung

zu verändern, auf das so genannte Normalgleichungssystem (Gaußsche Normalgleichungen)

$$A^T Ax = A^T b, \text{ d. h. } Bx = c, \quad (1.3)$$

mit der symmetrischen und positiv definiten Koeffizientenmatrix (spd) $B = A^T A$, so folgt sofort das Funktional

$$\begin{aligned} R(x) &= \frac{1}{2}x^T Bx - x^T c \\ &= \frac{1}{2}x^T A^T Ax - x^T A^T b = \frac{1}{2}(Ax)^T Ax - (Ax)^T b \\ &= \frac{1}{2} \sum_{i=1}^n \sum_{j,k=1}^n a_{ij} a_{ik} x_j x_k - \sum_{i=1}^n \sum_{j=1}^n b_i a_{ij} x_j. \end{aligned} \quad (1.4)$$

An der Lösung ist $R(x^*) = \frac{1}{2}(Ax^*)^T Ax^* - (Ax^*)^T b = -\frac{1}{2}b^T b \leq 0$.

Die Matrix B hat im Allgemeinen eine schlechtere Kondition als A , so dass sich dies auf das Iterationsverfahren, dazu gehören die Abstiegsverfahren, ungünstig auswirken wird und somit dann mehr Schritte erforderlich sind.

Bei Implementierungen wird die Matrix B wegen des Aufwands von $2n^3$ Operationen nicht explizit ermittelt, sondern im Verfahren hat man anstelle einer Matrix-Vektor-Multiplikation in der Hauptschleife nun zwei.

Eine dritte mit $R(x)$ verwandte Form ist

$$f(x) = \sum_{i=1}^n [f_i(x)]^2 \geq 0, \quad (1.5)$$

wobei $f_i(x)$ die i -te Zeile des LGS (1.1) darstellt, also

$$f_i(x) = (b - Ax)_i = b_i - \sum_{j=1}^n a_{ij} x_j.$$

Man rechnet einfach nach, dass die Beziehungen

$$f(x) = 2R(x) + b^T b \quad (1.6)$$

und somit $f(x^*) = 2R(x^*) + b^T b = 0$ gelten. Die Funktionale $f(x)$ und $R(x)$ weisen deshalb keine qualitativen Unterschiede auf.

Nun berechnen wir die genannten Funktionale für einige LGS im \mathbb{R}^2 , wo wir sie auch grafisch darstellen können. Wir diskutieren die zugehörige Abstiegsituation und ihre eventuellen Besonderheiten. Dabei charakterisieren wir im Rahmen des Iterationsprozesses für einen Schritt die Wahl von möglichen Abstiegsrichtungen bzw. generell Suchrichtungen hin zur Lösung x^* . Der Grenzvektor x^* ist unter den für Abstiegsverfahren üblichen Voraussetzungen die Minimumstelle des Funktionals, anders jedoch bei abweichenden Bedingungen (siehe [36]).

Die Abbildungen zu den folgenden Beispielen sind in Maple erzeugt worden.

Beispiel 1.1

Sei $Ax = b$, $A = I$ Einheitsmatrix und $b = 0$.

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad x^* = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Die Funktionale sind

$$Q(x) = \frac{1}{2}(x_1^2 + x_2^2), \quad R(x) = \frac{1}{2}(x_1^2 + x_2^2), \quad f(x) = x_1^2 + x_2^2.$$

Am gemeinsamen eindeutigen Minimum bei $x^* = 0$ ist $Q(x^*) = R(x^*) = f(x^*) = 0$.

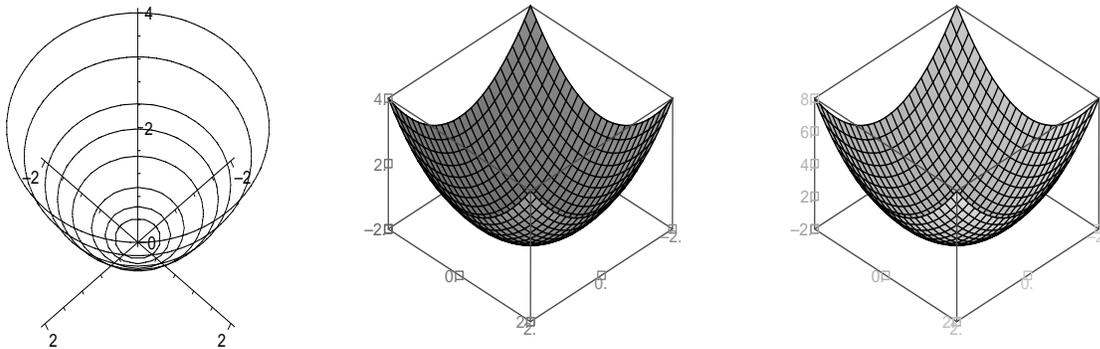


Abb. 1.1 Dateien *abt_10.ps*, *abt_11.ps*

3D-Portraits von $Q(x) = R(x) = \frac{1}{2}(x_1^2 + x_2^2)$ und $f(x) = x_1^2 + x_2^2$ (v.l.n.r.)

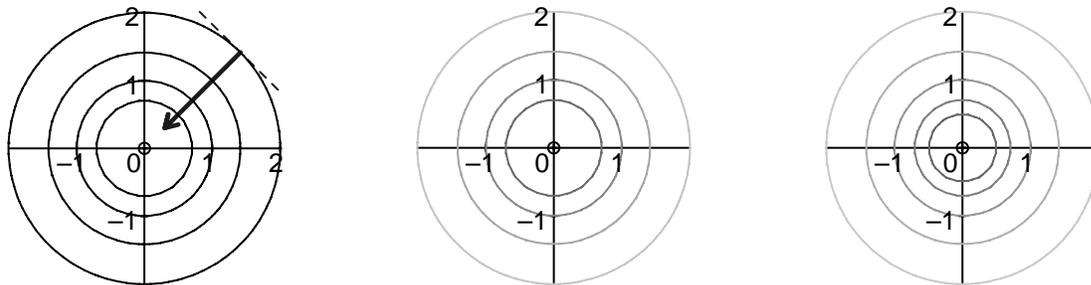


Abb. 1.2 Datei *abt_12.ps*

Höhenlinienbild von $Q(x)$, $R(x)$ mit
contours=[0,0.25,0.5,1,2]

und $f(x)$ mit
contours=[0,0.25,0.5,1,2,4]

Als Suchrichtung und Abstiegsrichtung in einem Schritt bietet sich wie beim GV die Richtung des steilsten Abstiegs an. Sie ist der negative Gradient des Funktionals

$$-\nabla Q(x) = -\left(\frac{\partial Q(x)}{\partial x_1}, \frac{\partial Q(x)}{\partial x_2}, \dots, \frac{\partial Q(x)}{\partial x_n}\right)^T = b - Ax = r(x) \quad (1.7)$$

und orthogonal zur Höhenlinie $Q(x)=\text{const}$, genauer gesagt: $\nabla Q(x) \perp$ zur Tangente an die Höhenlinie. Im Höhenlinienbild von $Q(x)$ in Abbildung 1.2 (linke Figur) ist eine solche Abstiegsrichtung eingetragen. Im Fall kreisförmiger Konturen kann ein Schritt in dieser Richtung genau zur Minimumstelle führen.

Beispiel 1.2

Sei $A = A^T > 0$ eine Diagonalmatrix und $b = 0$.

$$\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad x^* = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad x^{(0)} = \begin{pmatrix} 9/2 \\ 3 \end{pmatrix}.$$

Die Funktionale sind

$$Q(x) = \frac{1}{2}x_1^2 + x_2^2,$$

$$R(x) = \frac{1}{2}x_1^2 + 2x_2^2.$$

Am gemeinsamen eindeutigen Minimum an der Stelle $x^* = 0$ gilt $Q(x^*) = R(x^*) = 0$.

Als Suchrichtung und Abstiegsrichtung in einem Schritt nehmen wir die Richtung des steilsten Abstiegs $r(x) = -\nabla Q(x)$, die orthogonal zu den Höhenlinien $Q(x) = \text{const}$ sind.

Im GV werden unendlich viele Schritte ausgeführt. Der Iterationsverlauf ist in der Projektion wie eine rechtwinklige ‘‘Zick-Zack‘‘-Kurve.

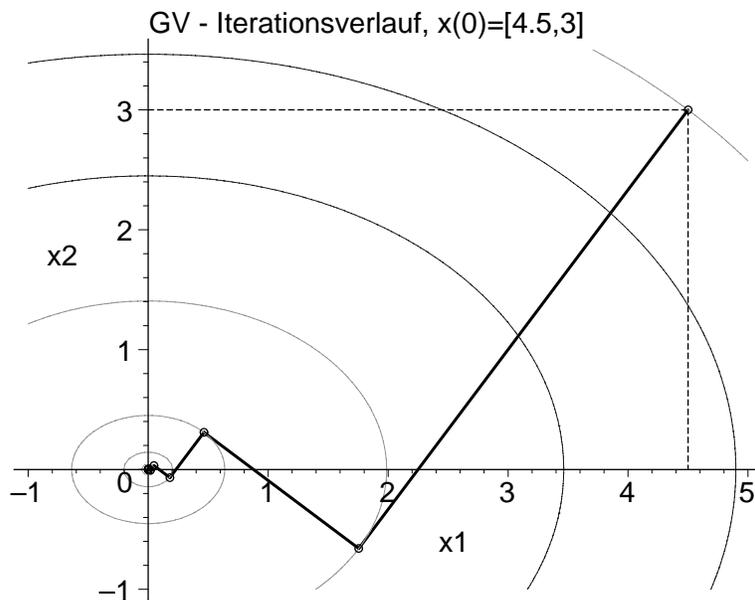


Abb. 1.3 Datei *abst101.ps*

Höhenlinienbild mit Iterationsverlauf des GV zu $Q(x) = \frac{1}{2}x_1^2 + x_2^2$ mit $\text{contours} = [19.125, 12, 6, 1.976, 0.204, 0.021, 0.0022, 0]$

Wir erhalten zum GV (Version 1, siehe (4.19)) die Beziehungen

$$r^{(0)} = \left(-\frac{9}{2}, -6 \right)^T \parallel (3, 4)^T, \quad r^{(1)} = \left(-\frac{72}{41}, \frac{54}{41} \right)^T \parallel (-4, 3)^T,$$

$$\alpha_0 = \frac{25}{41} = 0.609756\dots, \quad \alpha_1 = \frac{25}{34} = 0.735294\dots,$$

mit

$$r^{(k)} = b - Ax^{(k)}, \quad \|r^{(k)}\|^2 = r^{(k)T}r^{(k)},$$

$$r^{(0)} \perp r^{(1)} \perp r^{(2)} \perp r^{(3)} \perp \dots, \quad r^{(0)} \parallel r^{(2)} \parallel r^{(4)} \parallel \dots, \quad r^{(1)} \parallel r^{(3)} \parallel r^{(5)} \parallel \dots,$$

$$\mathcal{K}_2(A, r^{(0)}) = \text{span}\{r^{(0)}, Ar^{(0)}\} = \text{span}\{r^{(0)}, r^{(1)}\} = \text{span}\left\{\begin{pmatrix} 3 \\ 4 \end{pmatrix}, \begin{pmatrix} -4 \\ 3 \end{pmatrix}\right\} = \mathbb{R}^2,$$

$$x^* = 0 \cdot r^{(0)} + 0 \cdot r^{(1)} \in \text{span}\{r^{(0)}, r^{(1)}\}.$$

Dabei bedeutet $\mathcal{K}_k(A, x)$ den durch den Vektor $x \neq 0$ aufgespannten Raum von linear unabhängigen Vektoren (solange diese nicht verschwinden)

$$\mathcal{K}_k = \mathcal{K}_k(A, x) = \text{span}\{x, Ax, \dots, A^{k-1}x\}. \quad (1.8)$$

Er wird auch als Krylov-Unterraum oder Krylov-Teilraum bezeichnet.

Ergebnisse aus Berechnungen mit Maple

```
Startvektor          x = [+4.5000000000000000e+00 +3.0000000000000000e+00]
Residuum/SR p = r = b-Ax = [-4.5000000000000000e+00 -6.0000000000000000e+00]
Funktionswert       Q(x) = +1.9125000000000000e+01
Anfangsfehlerquadrat r'r = 5.6250000000000000e+01
```

k	Schrittzahl	alpha	Iterationsvektor x Residuum/neue Suchrichtung p=r=b-Ax	Funktionswert Q(x) Fehlernormquadrat r'r
1	6.0975609756097560e-01		[+1.7560975609756100e+00, -6.5853658536585400e-01] [-1.7560975609756100e+00, +1.3170731707317070e+00]	1.9756097560975620e+00 4.8185603807257580e+00
2	7.3529411764705890e-01		[+4.6484935437589700e-01, +3.0989956958393070e-01] [-4.6484935437589700e-01, -6.1979913916786200e-01]	2.0408020436014960e-01 6.0023589517691110e-01
3	6.0975609756097590e-01		[+1.8140462609791090e-01, -6.8026734786717000e-02] [-1.8140462609791090e-01, +1.3605346957343350e-01]	2.1081455830603760e-02 5.1418184952692060e-02
4	7.3529411764705830e-01		[+4.8018871614153000e-02, +3.2012581076101700e-02] [-4.8018871614153000e-02, -6.4025162152203900e-02]	2.1777113627022390e-03 6.4050334197125120e-03
5	6.0975609756097620e-01		[+1.8739071849425530e-02, -7.0271519435348600e-03] [-1.8739071849425530e-02, +1.4054303887069220e-02]	2.2495727132649200e-04 5.4867627152802140e-04
6	7.3529411764705750e-01		[+4.9603425483773700e-03, +3.3068950322513100e-03] [-4.9603425483773700e-03, -6.6137900645031300e-03]	2.3238053852949840e-05 6.8347217214563220e-05
7	6.0975609756097660e-01		[+1.9357434335131150e-03, -7.2590378756767800e-04] [-1.9357434335131150e-03, +1.4518075751348470e-03]	2.4004876289996720e-06 5.8548478756080670e-06
8	7.3529411764705690e-01		[+5.1240267357700500e-04, +3.4160178238441300e-04] [-5.1240267357700500e-04, -6.8320356476933400e-04]	2.4797002767263920e-07 7.2932361080238830e-07
9	6.0975609756097700e-01		[+1.9996201895687930e-04, -7.4985757109084300e-05] [-1.9996201895687930e-04, +1.4997151421766060e-04]	2.5615268281878260e-08 6.2476264102049340e-08
10	7.3529411764705700e-01		[+5.2931122665056600e-05 +3.5287415109783500e-05] [-5.2931122665056600e-05 -7.0574830220075000e-05]	2.6460535384218120e-09 7.7825104071756800e-09

Kapitel 2

Grundlagen der linearen Algebra

Dieses Kapitel dient der Einführung der später benötigten grundlegenden Begriffe und Aussagen der linearen Algebra für die Untersuchungen zum GV.

Ein Teil davon ist Gegenstand der mathematischen Grundlagenausbildung und wird deshalb nur übersichtsartig zusammengestellt (siehe auch [17], [18]).

Sei

$$Ax = b, \quad A = A^T > 0, \quad \text{d. h. } A \text{ ist spd.} \quad (2.1)$$

Damit ist A eine reguläre und invertierbare Matrix. Die Inverse A^{-1} ist ebenfalls spd wegen $y^T A^{-1} y = (Ax)^T A^{-1} Ax = x^T A^T x = (x^T A^T x)^T = x^T Ax > 0$.

Weiterhin gelten für $A(n, n) = (a_{ij})$:

$$\begin{aligned} (1) \quad & a_{ii} > 0 \quad \forall i, \\ (2) \quad & a_{ij}^2 < a_{ii} a_{jj} \quad \forall i \neq j, \\ (3) \quad & \exists k \text{ mit } a_{kk} = \max_{i,j} |a_{ij}|. \end{aligned} \quad (2.2)$$

Die Menge der n positiven Eigenwerte (EW) von A heißt Spektrum

$$\sigma(A) = \{\lambda\} = \{\lambda_1, \lambda_2, \dots, \lambda_n\}. \quad (2.3)$$

Es sei

$$0 < \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n, \quad \lambda_1 = \lambda_{\min}, \quad \lambda_n = \lambda_{\max}. \quad (2.4)$$

Damit gilt für das Spektrum der inversen Matrix

$$\sigma(A^{-1}): \quad 0 < \frac{1}{\lambda_n} \leq \dots \leq \frac{1}{\lambda_2} \leq \frac{1}{\lambda_1}. \quad (2.5)$$

Das Hauptachsentheorem, auch Spektralsatz oder Theorem von Schur genannt (Satz 2.31 in [18]), sagt nun, dass damit eine Orthogonalmatrix Q ($Q^T Q = Q Q^T = I$) existiert, so dass die Beziehung

$$Q^T A Q = Q^{-1} A Q = \Lambda = \text{diag}(\lambda_i) \quad (2.6)$$

erfüllt ist. Die Spalten von Q sind wegen $AQ = Q\Lambda$ die orthonormalen Eigenvektoren (EV) von A .

Das Skalarprodukt oder innere Produkt im Reellen, also von Vektoren aus \mathbb{R}^n , ist definiert durch

$$(x, y) = x^T y = \sum_{i=1}^n x_i y_i. \quad (2.7)$$

Der Zusammenhang zur euklidischen Vektornorm (l_2 -Norm) $\|\cdot\|_2$ ist

$$\|x\|_2^2 = (x, x) = x^T x = \sum_{i=1}^n x_i^2. \quad (2.8)$$

Von den zahlreichen Matrixnormen erwähnen wir

$$\begin{aligned} \|A\|_F = \|A\|_E &= \sqrt{\sum_{i=1}^n \sum_{j=1}^n a_{ij}^2} && \text{Frobenius-Norm, Schursche Norm,} \\ & && \text{euklidische Norm,} \\ &= \sqrt{\text{spur}(A^T A)} && (2.9) \\ \|A\|_2 &= \sqrt{\max_{i=1(1)n} \mu_i}, && 0 \leq \mu_i \in \sigma(A^T A) \text{ (Spektrum),} \\ &= \sqrt{\rho(A^T A)} && \text{Spektralnorm, Hilbert-Norm,} \\ & && \text{größter Singulärwert von } A. \end{aligned}$$

Es gelten $B = A^T A = B^T > 0$ und $\sigma(A^T A) = \sigma(AA^T)$. Falls $A = A^T$ ist, hat man

$$\|A\|_2 = \sqrt{\rho(A^2)} = \rho(A) = \max_{i=1(1)n} |\lambda_i(A)|. \quad (2.10)$$

Falls $A = A^T > 0$ ist, folgt sofort $\|A\|_2 = \max_{i=1(1)n} \lambda_i(A)$.

Für beliebige Matrizen stellt der Spektralradius $\rho(A)$ eine untere Schranke aller Matrixnormen dar.

Falls $A = A^T > 0$ ist, dann definiert man die energetische Norm (A -Norm)

$$x^T A x = (Ax, x) = (x, x)_A = \|x\|_A^2, \quad (2.11)$$

und es gilt $x^T A x > 0$ für $x \neq 0$.

Für $A = A^T > 0$ gelten auch die Beziehungen

$$\begin{aligned} \rho(A) &= \max_{x \neq 0} \frac{x^T A x}{x^T x} = \max_{\|x\|=1} x^T A x = \lambda_{max}, \\ \frac{1}{\rho(A^{-1})} &= \min_{x \neq 0} \frac{x^T A x}{x^T x} = \min_{\|x\|=1} x^T A x = \lambda_{min}. \end{aligned} \quad (2.12)$$

und die Normäquivalenz

$$\begin{aligned} \lambda_{min} x^T x &\leq x^T A x \leq \lambda_{max} x^T x \text{ bzw.} \\ \lambda_{min} \|x\|_2^2 &\leq \|x\|_A^2 \leq \lambda_{max} \|x\|_2^2. \end{aligned} \quad (2.13)$$

Der Spektralradius $\rho(A)$ ist

$$\rho(A) = \max_{\lambda \in \sigma(A)} |\lambda| = \max_{i=1,2,\dots,n} |\lambda_i| \leq \|A\|. \quad (2.14)$$

Ist die Matrix A eine spaltenreguläre Rechteckmatrix $A(m, n)$ mit $m \geq n$, so existiert eine eindeutige QR -Faktorisierung dieser gemäß

$$\begin{aligned} A &= QR, \\ Q &= Q(m, n), \quad Q^T Q = I(n, n), \quad Q Q^T = G(m, m), \\ R &= R(n, n) \text{ obere Dreiecksmatrix mit } r_{ii} > 0, \\ Ax &= QRx = b, \\ Rx &= Q^T b, \quad R = Q^T A, \\ x^* &= R^{-1} Q^T b. \end{aligned} \quad (2.15)$$

Der Vektor x^* ist die Lösung im Sinne der Methode der kleinsten Quadrate.

Er minimiert die euklidische Norm $\|b - Ax\|_2$.

Es gilt weiterhin

$$\begin{aligned} x^* &= R^{-1} Q^T b = (Q^T A)^{-1} (AR^{-1})^T b = (Q^T A)^{-1} (AR^{-1})^T b \\ &= [(AR^{-1})^T A]^{-1} (AR^{-1})^T b = (R^{-T} A^T A)^{-1} R^{-T} A^T b \\ &= (A^T A)^{-1} R^T R^{-T} A^T b = (A^T A)^{-1} A^T b. \end{aligned} \quad (2.16)$$

Die Matrix $A^+ = (A^T A)^{-1} A^T$ heißt Pseudoinverse.

Bei der QR -Faktorisierung ist es möglich die Matrix Q um weitere $m-n$ orthonormale Spalten zu erweitern und die Dreiecksmatrix R um $m-n$ Nullzeilen zu ergänzen, ohne die Gültigkeit der Beziehung $A = QR$ zu verletzen. In CAS wird oft diese erweiterte Orthonormalmatrix erzeugt.

Für $m = n$ erhält man mit $Q^T Q = Q Q^T = I$

$$\begin{aligned} \|Q^T b\|_2^2 &= (Q^T b, Q^T b) = (Q^T b)^T Q^T b = b^T Q Q^T b = \|b\|_2^2, \\ (QA)^T (QA) &= A^T (Q^T Q) A = A^T A, \\ R^T R &= (Q^T A)^T (Q^T A) = A^T (Q Q^T) A = A^T A, \\ \|R\|_2^2 &= \rho(R^T R) = \rho(A^T A) = \|A\|_2^2. \end{aligned} \quad (2.17)$$

Vektor- und Matrixnormen stehen miteinander in Beziehung.

Eine Matrixnorm $\|A\|_M$ heißt kompatibel (passend, verträglich, konsistent) zu einer Vektornorm $\|x\|_V$, wenn für alle A und x gilt

$$\|Ax\|_V \leq \|A\|_M \|x\|_V. \quad (2.18)$$

Vektornormen induzieren Matrixnormen. Für die induzierte Matrixnorm gilt

$$\|A\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \sup_{\|x\|=1} \|Ax\| = \max_{\|x\|=1} \|Ax\|. \quad (2.19)$$

Die induzierte Matrixnorm ist mit der zu Grunde liegenden Vektornorm zugleich kompatibel und unter allen mit dieser Vektornorm kompatiblen Matrixnormen die kleinste.

Die induzierte Matrixnorm ist wohl definiert im Sinne einer Norm, was leicht nachzuprüfen ist, wenn man die Eigenschaften der Vektornorm ausnutzt.

So induziert die euklidische Vektornorm die Spektralnorm. Zum Nachweis verwendet man für die Matrix

$$B = A^T A = B^T \geq 0$$

das Hauptachsentheorem.

Es gelten bei beliebigem $x \in \mathbb{R}^n$ die Beziehungen

$$\begin{aligned} Q^T B Q &= \text{diag}(\mu_i), \quad \mu_i \in \mathbb{R}, \quad \mu_i \geq 0, \quad Q^T Q = I, \\ B Q &= Q \text{diag}(\mu_i), \quad Q = (q_1, q_2, \dots, q_n), \quad q_i \text{ orthonormale Spalten,} \\ &\text{Eigenwertproblem (EWP),} \end{aligned}$$

$$\begin{aligned} x &= \sum_{i=1}^n \alpha_i q_i \in \mathbb{R}^n, \quad \alpha_i \in \mathbb{R}, \\ \|x\|_2^2 &= \sum_{i=1}^n \alpha_i^2, \\ Bx &= \sum_{i=1}^n \alpha_i B q_i = \sum_{i=1}^n \alpha_i \mu_i q_i, \\ \|Bx\|_2^2 &= (Bx)^T Bx = x^T B^T Bx = x^T B^2 x, \\ \|Ax\|_2^2 &= (Ax, Ax) = (x, A^T A x) = (x, Bx) \\ &= \left(\sum_{i=1}^n \alpha_i q_i, \sum_{i=1}^n \alpha_i \mu_i q_i \right) \\ &= \sum_{i=1}^n \alpha_i^2 \mu_i \\ &\leq \max_i \mu_i \sum_{i=1}^n \alpha_i^2 = \rho(B) \|x\|_2^2 = \rho(A^T A) \|x\|_2^2, \\ \frac{\|Ax\|_2^2}{\|x\|_2^2} &\leq \rho(A^T A) = \|A\|_2^2. \end{aligned}$$

Somit brauchen wir nur noch einen speziellen Vektor x für die Gleichheit.

Es ist $\mu_n = \max_i \mu_i \geq 0$. Sei $v_n \neq 0$ der zu μ_n gehörige EV.

Dann folgt

$$\begin{aligned} \frac{\|Av_n\|_2^2}{\|v_n\|_2^2} &= \frac{(Av_n, Av_n)}{\|v_n\|_2^2} = \frac{(v_n, A^T Av_n)}{\|v_n\|_2^2} = \frac{(v_n, Bv_n)}{\|v_n\|_2^2} = \frac{\mu_n(v_n, v_n)}{\|v_n\|_2^2} \\ &= \mu_n = \rho(A^T A) = \|A\|_2^2. \end{aligned}$$

Der Zusammenhang zum EWP ist auch der Grund für die Bezeichnung dieser induzierten Matrixnorm als Spektralnorm.

In der Formel (2.14) haben wir schon die Ungleichung $\rho(A) \leq \|A\|$ für beliebige Normen eingeführt.

Dies folgt aus dem EWP $Az = \lambda z$ mit λ als betragsgrößter EW und $z \neq 0$, $\|z\| = 1$, denn bei einer induzierten Norm ist

$$\|A\| = \max_{\|x\|=1} \|Ax\| \geq \|Az\| = \|\lambda z\| = |\lambda| \|z\| = |\lambda|,$$

bzw. bei einer kompatiblen Norm aus

$$|\lambda| \|z\| = \|\lambda z\| = \|Az\| \leq \|A\| \|z\|.$$

Ein EW λ_i von A besitzt die algebraische Vielfachheit $n_i \geq 1$, falls λ_i eine n_i -fache Nullstelle des charakteristischen Polynoms $p_n(\lambda) = \det(A - \lambda I)$ ist. Es gilt $\sum_i n_i = n$. λ_i besitzt die geometrische Vielfachheit $m_i \geq 1$, falls $m_i = n - \text{rang}(A - \lambda_i I)$ gilt. Man spricht vom Defekt von $A - \lambda_i I$. Es gilt stets $m_i \leq n_i$, wobei m_i die Anzahl der linear unabhängigen EV zu λ_i ist.

Eine wichtige Aussage für spätere Konvergenzbetrachtungen findet man im folgenden Satz (siehe [32]).

Satz 2.1 Für die Matrix $A = A^T > 0$ mit ihren EW $0 < \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ gilt bei beliebigem $x \neq 0$ die Ungleichung von Kantorovich

$$\frac{(x^T x)^2}{(x^T A x)(x^T A^{-1} x)} \geq \frac{4\lambda_1 \lambda_n}{(\lambda_1 + \lambda_n)^2}. \quad (2.20)$$

Beweis. Man braucht zunächst das Hauptachsentheorem, also

$$Q^T A Q = Q^{-1} A Q = \text{diag}(\lambda_i), \quad A Q = Q \text{diag}(\lambda_i)$$

mit $Q = (q_1, q_2, \dots, q_n)$, q_i orthonormale Spalten und EV von A , $Q^T Q = I$, sowie die EW $\frac{1}{\lambda_i}$ von A^{-1} .

Sei $x \in \mathbb{R}^n$ ein nicht verschwindender beliebiger, aber fester Vektor. Da die Vektoren q_i auch eine Basis des \mathbb{R}^n bilden, ist

$$x = \sum_{i=1}^n \alpha_i q_i, \quad \text{wobei} \quad \sum_{i=1}^n \alpha_i^2 > 0.$$

Wir formen mit x die linke Seite $F(x)$ der Ungleichung (2.20) um.

$$\begin{aligned}
 F(x) &= \frac{\left(\sum_{i=1}^n \alpha_i^2\right)^2}{\left(\sum_{i=1}^n \alpha_i^2 \lambda_i\right) \left(\sum_{i=1}^n \alpha_i^2 \lambda_i^{-1}\right)} \\
 &= \frac{1}{\left(\sum_{i=1}^n \underbrace{\frac{\alpha_i^2}{\sum_{j=1}^n \alpha_j^2}}_{\beta_i} \lambda_i\right) \left(\sum_{i=1}^n \underbrace{\frac{\alpha_i^2}{\sum_{j=1}^n \alpha_j^2}}_{\beta_i} \lambda_i^{-1}\right)} \\
 &= \frac{1}{\left(\sum_{i=1}^n \beta_i \lambda_i\right) \left(\sum_{i=1}^n \beta_i \lambda_i^{-1}\right)}
 \end{aligned}$$

mit

$$\beta_i = \frac{\alpha_i^2}{\sum_{j=1}^n \alpha_j^2} \geq 0 \quad \text{und} \quad \sum_{i=1}^n \beta_i = 1.$$

Die beiden Summen im Nenner von $F(x)$ sind damit konvexe Kombinationen der jeweiligen EW.

Es gilt

$$\lambda_1 \leq \tilde{\lambda} = \sum_{i=1}^n \beta_i \lambda_i \leq \lambda_n. \tag{2.21}$$

Für die Funktion $p(\lambda) = \frac{1}{\lambda}$, $\lambda > 0$, haben wir die Eigenschaften der abnehmenden Monotonie wegen $p'(\lambda) = -\lambda^{-2} < 0$ sowie der strikten Konvexität durch $p''(\lambda) = 2\lambda^{-3} > 0$.

Die Punkte $P_i = P(\lambda_i, \lambda_i^{-1})$, $i = 1, 2, \dots, n$, liegen auf der Kurve $p(\lambda)$. Folglich befinden sich die "inneren" Punkte P_2, P_3, \dots, P_{n-1} nicht oberhalb der Geraden $g(\lambda)$ durch die beiden Randpunkte P_1 und P_n .

Die Geradengleichung lautet

$$\begin{aligned}
 g = g(\lambda) &= \lambda_1^{-1} + \frac{\lambda_n^{-1} - \lambda_1^{-1}}{\lambda_n - \lambda_1} (\lambda - \lambda_1) \\
 &= \lambda_1^{-1} \lambda_n^{-1} \left[\lambda_n + \frac{\lambda_1^{-1} - \lambda_n^{-1}}{\lambda_n - \lambda_1} (\lambda - \lambda_1) \right] \\
 &= \lambda_1^{-1} \lambda_n^{-1} [\lambda_n - (\lambda - \lambda_1)] \\
 &= \frac{\lambda_1 + \lambda_n - \lambda}{\lambda_1 \lambda_n}.
 \end{aligned} \tag{2.22}$$

Der Punkt

$$Q = Q\left(\tilde{\lambda}, \sum_{i=1}^n \beta_i \lambda_i^{-1}\right) = \sum_{i=1}^n \beta_i P(\lambda_i, \lambda_i^{-1})$$

befindet sich somit in der konvexen Hülle von $\{P_1, P_2, \dots, P_n\}$.

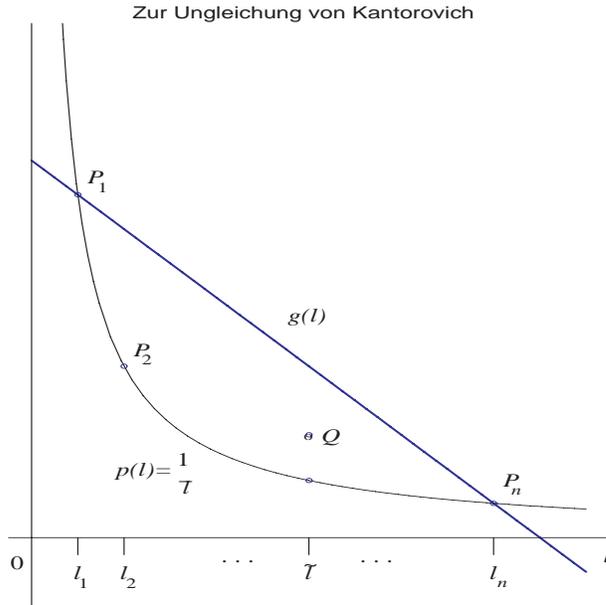


Abb. 2.1 Datei *kantor1.ps*
 Funktionen $p(\lambda)$ und $g(\lambda)$, $\lambda \equiv l$,
 sowie Lage der Punkte
 $P_i = P(\lambda_i, \lambda_i^{-1})$, $i = 1, 2, \dots, n$,
 und Q

Somit liegt Q jedenfalls nicht oberhalb der Geraden $g(\lambda)$ und damit ist

$$\sum_{i=1}^n \beta_i \lambda_i^{-1} \leq \frac{\lambda_1 + \lambda_n - \tilde{\lambda}}{\lambda_1 \lambda_n} = g(\tilde{\lambda}). \quad (2.23)$$

Daraus folgt

$$\begin{aligned} F(x) &= \frac{1}{\tilde{\lambda} \sum_{i=1}^n \beta_i \lambda_i^{-1}} \geq \frac{1}{\tilde{\lambda} \frac{\lambda_1 + \lambda_n - \tilde{\lambda}}{\lambda_1 \lambda_n}} = \\ &= \frac{\lambda_1 \lambda_n}{\tilde{\lambda} (\lambda_1 + \lambda_n - \tilde{\lambda})} \geq \min_{\lambda_1 \leq \lambda \leq \lambda_n} \frac{\lambda_1 \lambda_n}{\tilde{\lambda} (\lambda_1 + \lambda_n - \tilde{\lambda})} = \\ &= \frac{\lambda_1 \lambda_n}{\frac{\lambda_1 + \lambda_n}{2} (\lambda_1 + \lambda_n - \frac{\lambda_1 + \lambda_n}{2})} = \\ &= \frac{4 \lambda_1 \lambda_n}{(\lambda_1 + \lambda_n)^2}. \end{aligned}$$

Der vorletzte Schritt beruht auf der Erkenntnis, dass der Ausdruck $z(1-z)$, $z \in [0, 1]$, für $z = \frac{1}{2}$ am größten ist. \square

Kapitel 3

Grundlagen der Abstiegsverfahren

Gegeben sei das reguläre LGS (1.1) bzw. (2.1) mit der exakten Lösung $x^* = A^{-1}b \in \mathbb{R}^n$. Damit gilt $b - Ax^* = 0$. Ein Näherungsverfahren zu seiner Lösung liefert im Allgemeinen auch nur eine Näherungslösung \tilde{x} .

Somit entsteht ein Lösungsfehler

$$e = e(\tilde{x}) = x^* - \tilde{x} \quad (3.1)$$

und ein nicht verschwindender Rest bzw. Residuum (Residuenvektor)

$$r = r(\tilde{x}) = b - A\tilde{x}. \quad (3.2)$$

Zwischen Fehler und Residuum erkennt man den Zusammenhang

$$\begin{aligned} r &= b - A\tilde{x} = A(A^{-1}b - \tilde{x}) = A(x^* - \tilde{x}) = Ae, \\ e &= A^{-1}r. \end{aligned} \quad (3.3)$$

Somit kann nur bei Kenntnis spezieller zusätzlicher Eigenschaften der Koeffizientenmatrix A – eine davon ist die Kondition – aus der einen Größe auf die andere geschlossen werden.

Ein kleines Residuum im Sinne der Norm des Vektors bedeutet also nicht automatisch einen kleinen Lösungsfehler.

Der Grund ist, dass die Matrix A eine schlechte Kondition hat. Kennzeichen dafür sind unter anderem:

- Die Matrix A ist fast singulär.
- Die Determinante von A ist nahe Null.
- Wenn die Matrix A Elemente der Größenordnung $\mathcal{O}(1)$ besitzt, dann hat die inverse Matrix A^{-1} betragsmäßig große Elemente, oder umgekehrt.
- Das Spektrum der EW von A ist sehr “breit“. Es liegen Größenordnungen zwischen dem betragsmäßig kleinsten ($\neq 0$) und größten EW.

Zur Kondition gehört ihre wertemäßige Berechnung als Konditionszahl mit Hilfe der inversen Matrix und der Norm

$$\kappa(A) = \text{cond}_s(A) = \|A\|_s \|A^{-1}\|_s, \quad (3.4)$$

wobei durch den Index die Wahl einer konkreten Norm $\|A\|_s$ gemeint ist.

Ein schlechter Konditionswert charakterisiert im Groben die Reduzierung der gültigen Mantissenstellen der in Rechnungen mit A oder A^{-1} benutzten Gleitpunktarithmetik.

3.1 Ansätze für quadratische Formen

Die Idee zur Wahl von quadratischen Formen und ihre Einbindung in ein Optimierungsproblem für die Lösung von (1.1) kann man ganz einfach im eindimensionalen (skalaren) Fall nachvollziehen.

Für die Nullstellenaufgabe

$$f(x) = ax + b = 0, \quad a \neq 0,$$

mit der Nullstelle $x^* = -\frac{b}{a}$ ergibt sich sofort die daraus abgeleitete und zu minimierende Funktion

$$g(x) = [f(x)]^2.$$

Sie hat mindestens dieselben Glattheitseigenschaften wie $f(x)$ und an der Stelle x^* ihr eindeutiges absolutes Minimum.

Man schreibt

$$x^* = \arg \min_{x \in \mathbb{R}} g(x).$$

Bezüglich der Minimumstelle ändert sich nichts, wenn wir die Funktion $g(x)$ strecken und vertikal verschieben, also mit $c_1 g(x) + c_2$, $c_1 > 0$, arbeiten.

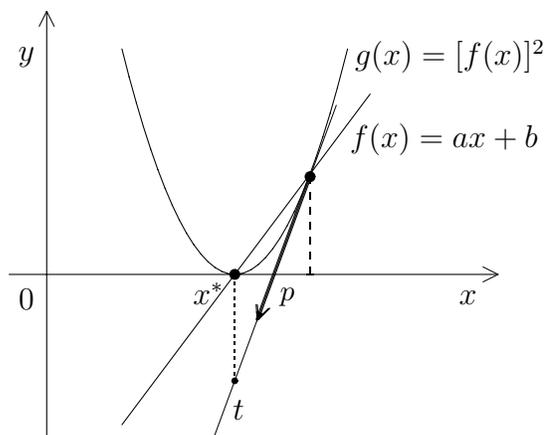


Abb. 3.1 Datei *abst1.pic*
Nullstellensituation $f(x) = 0$ und zugehörige Minimierungsaufgabe $\min_x [f(x)]^2$ im \mathbb{R}^1

Wenn man entlang der Tangente t an die Kurve $g(x)$ in Richtung des steilsten Abstiegs (Vektor p) geht, “fährt“ man sozusagen unter dem Minimum entlang und kann mittels einer Strahlenminimierung genau die Stelle darunter treffen.

Im mehrdimensionalen Fall werden im Allgemeinen mehrere aufeinander folgende Strahlenminimierungen in Richtung der Minimumstelle erforderlich sein.

Wir betrachten zwei wichtige quadratische Formen für das LGS (1.1) bzw. (2.1), entsprechend für die Koeffizientenmatrix A , die daraus abgeleiteten Funktionale sowie erste Eigenschaften dieser.

(1) Quadratische Form $\delta_Q(x)$ und Funktional $Q(x)$ auf der Basis von A

$$\begin{aligned}
0 \leq \delta_Q(x) &= \|e(x)\|_A^2 = \|x^* - x\|_A^2, \quad A = A^T > 0, \\
\delta_Q(x) &= (x^* - x)^T A (x^* - x) \quad (\text{allgemein}) \\
&= x^T A x - x^T A x^* - x^{*T} A x + x^{*T} A x^* \\
&= x^T A x - x^T A x^* - (x^{*T} A x)^T + x^{*T} b \\
&= x^T A x - x^T A x^* - x^T A^T x^* + x^{*T} b \\
&= x^T A x - x^T (A + A^T) x^* + x^{*T} b, \quad \tilde{A} = A + A^T = \tilde{A}^T \quad (3.5) \\
&= x^T A x - 2x^T A x^* + x^{*T} b, \quad \text{falls } A = A^T \\
&= x^T A x - 2x^T b + x^{*T} b \\
&= 2 \underbrace{\left(\frac{1}{2} x^T A x - x^T b \right)}_{= Q(x)} + x^{*T} b,
\end{aligned}$$

$$Q(x) = \frac{1}{2} x^T A x - x^T b. \quad (3.6)$$

Falls die Matrix nicht symmetrisch ist, kommt man nur bis zum Funktional

$$Q_{ns}(x) = \frac{1}{2} \left(x^T A x - x^T (A + A^T) x^* \right). \quad (3.7)$$

Außerdem gilt $\delta_Q(x) \geq 0$ natürlich nur im Fall $A = A^T > 0$.

(2) Quadratische Form $\delta_R(x)$ und Funktional $R(x)$ auf der Basis von $A^T A$

$$\begin{aligned}
0 \leq \delta_R(x) &= \|r(x)\|_2^2 = r(x)^T r(x), \quad r(x) = b - Ax \\
&= \|b - Ax\|_2^2 = (b - Ax)^T (b - Ax) \\
&= (Ax)^T Ax - (Ax)^T b - b^T Ax + b^T b \\
&= x^T A^T Ax - x^T A^T b - (b^T Ax)^T + b^T b \\
&= x^T A^T Ax - x^T A^T b - x^T A^T b + b^T b \\
&= x^T A^T Ax - 2x^T A^T b + b^T b
\end{aligned}$$

$$\begin{aligned}
0 \leq \delta_R(x) &= x^T A^2 x - 2x^T A b + b^T b, \quad \text{falls } A = A^T \\
&= (Ax, x)_A - 2(x, b)_A + (b, b), \quad \text{falls } A = A^T > 0 \\
&= 2 \left(\underbrace{\frac{1}{2} (Ax, x)_A - (x, b)_A}_{= R_s(x)} \right) + (b, b),
\end{aligned} \tag{3.8}$$

$$R_s(x) = \frac{1}{2} (Ax, x)_A - (x, b)_A = \frac{1}{2} x^T A^2 x - (Ax)^T b. \tag{3.9}$$

Wegen $b = Ax^*$, $e(x) = x^* - x$ gilt auch

$$\|r(x)\|_2^2 = \|A(x^* - x)\|_2^2 = [A^T A(x^* - x)]^T (x^* - x) = \|e(x)\|_{A^T A}^2. \tag{3.10}$$

Falls die Matrix nicht symmetrisch ist, und dafür ist dieser Zugang ja vorgesehen, denn er beinhaltet die Symmetrisierung, so kommt man zum Funktional

$$R(x) = \frac{1}{2} (Ax)^T A x - (Ax)^T b = \frac{1}{2} x^T A^T A x - x^T A^T b. \tag{3.11}$$

Im Fall $A = A^T$ ist eine zusätzliche Symmetrisierung von A und damit die Erzeugung von A^2 eigentlich überflüssig und man sollte mit $Q(x)$ operieren.

Wir haben zum LGS $Ax = b$, $A = A^T > 0$, die Beziehung $e(x^*) = 0$ und das Ziel $\|e(x)\|_A \rightarrow \min_x$ und zum LGS $A^T A x = A^T b$ die Beziehung $r(x^*) = 0$ und das Ziel $\|r(x)\|_2 = \|e(x)\|_{A^T A} \rightarrow \min_x$, sowie nach Transformation die Gleichungen

$$\begin{aligned}
\|e(x)\|_A^2 &= 2 \left(\underbrace{\frac{1}{2} x^T A x - x^T b}_{= Q(x) \rightarrow \min_x} \right) + \underbrace{x^{*T} b}_{= \text{const}}
\end{aligned} \tag{3.12}$$

und

$$\begin{aligned}
\|r(x)\|_2^2 &= 2 \left(\underbrace{\frac{1}{2} (Ax)^T A x - (Ax)^T b}_{= R(x) \rightarrow \min_x} \right) + \underbrace{b^T b}_{= \text{const}}
\end{aligned} \tag{3.13}$$

mit den neuen zu minimierenden Funktionalen $Q(x)$ und $R(x)$.

Zum Funktional $R(x)$ kommt man ebenfalls über das spd Normalgleichungssystem $A^T A x = A^T b$, indem man genau dafür das Funktional $Q(x)$ nimmt.

Als Maß für den Fehler der Iterierten in einem Abstiegsverfahren eignet sich natürlich vorrangig die Residuumnorm $\|r(x)\|_2$. In der Formel erscheint erstens nicht die exakte Lösung x^* und zweitens ist die bequemere euklidische Norm anstelle der A -Norm auszuwerten. Nur für akademische Kontrollrechnungen, wo die exakte Lösung bekannt ist, bietet sich die Berechnung beider Fehlerterme $\|r(x)\|_2$ und $\|e(x)\|_A$ an.

Kapitel 4

Abstiegsverfahren

4.1 Die quadratische Form $Q(x)$

Die Minimierung des Funktionals $\|e(x)\|_A$ geht konform mit der Minimierung des Funktionals und der quadratischen Form

$$Q(x) = \frac{1}{2}x^T Ax - x^T b, \quad A = A^T > 0. \quad (4.1)$$

Sein Gradient als Richtung des steilsten Anstiegs ist

$$\text{grad } Q(x) = \nabla Q(x) = Ax - b, \quad (4.2)$$

für die Hesse-Matrix gilt

$$\nabla^2 Q(x) = A = A^T > 0, \quad (4.3)$$

so dass $Q(x)$ konvex ist.

Wir bemerken, dass im Fall $A \neq A^T$ der Gradient $\nabla Q(x) = \frac{1}{2}(A + A^T)x - b$ ist.

Aus der Beziehung (3.12) erhält man

$$\begin{aligned} 0 \leq \|\varepsilon(x)\|_A^2 &= 2Q(x) + x^{*T}b = 2[Q(x) - (-\frac{1}{2}x^{*T}b)] = 2[Q(x) - Q(x^*)], \\ Q(x) &= \frac{1}{2}(\|e(x)\|_A^2 - x^{*T}b), \\ Q(x^*) &= \min_{x \in \mathbb{R}^n} Q(x) = -\frac{1}{2}x^{*T}b = -\frac{1}{2}x^{*T}Ax^* \leq 0, \\ Q(x^* + \Delta x) - Q(x^*) &= \frac{1}{2}\|\Delta x\|_A^2 > 0, \quad \Delta x \neq 0. \end{aligned} \quad (4.4)$$

Durch die Normäquivalenz (siehe [17]) kann man aus Abschätzungen für $\|e(x)\|_A$ auch zu solchen in der euklidischen Norm gelangen.

Die notwendige Bedingung am Minimum (Extremum) $\nabla Q(x) = 0$ entspricht der Lösung des LGS.

Der Zusammenhang zwischen Minimumstelle und Lösung soll als Satz formuliert werden.

Satz 4.1 *Unter der Voraussetzung $A = A^T > 0$ ist die einzige Minimumstelle x^* von $Q(x)$ zugleich die eindeutige Lösung des LGS $Ax = b$.*

Beweis.

(1) Für die eindeutige Lösbarkeit des LGS betrachten wir die Inverse A^{-1} .

Diese existiert. Wäre das nicht so, hätte das homogene LGS $Ax = 0$ eine nichttriviale Lösung x . Somit wäre auch $x^T Ax = 0$, $x \neq 0$, was der positiven Definitheit von A widerspricht.

(2) Für eine spd Matrix A existiert eine orthogonale Matrix U ($U^T U = I$), so dass die Matrix $T = U^T A U$ eine Diagonalmatrix ist mit den reellen positiven Diagonalelementen t_{ii} als Eigenwerte von A (siehe Hauptachsentheorem Satz 2.31 bzw. spezieller Fall des Satzes von SCHUR 2.15 in [18]).

(3) Zu zeigen ist nun der Zusammenhang mit der Extremwerteigenschaft von $Q(x)$, das heißt

$$Q(x^*) < Q(x) \text{ für alle } x \in \mathbb{R}^n \setminus \{x^*\}.$$

Betrachten wir zunächst den einfachen Fall, dass A eine Diagonalmatrix der Form $A = \text{diag}(d_1, d_2, \dots, d_n)$ ist. Aus der spd-Eigenschaft folgt sofort für ihre Diagonalelemente $d_i > 0$, $i = 1, 2, \dots, n$.

Das Funktional ist gegeben durch

$$Q(x) = \sum_{i=1}^n \left(\frac{1}{2} d_i x_i^2 - x_i b_i \right).$$

Das Extremum von $Q(x)$ erhält man durch die Bedingung

$$\nabla Q(x) = \left(\frac{\partial Q(x)}{\partial x_1}, \frac{\partial Q(x)}{\partial x_2}, \dots, \frac{\partial Q(x)}{\partial x_n} \right)^T = 0.$$

Wegen

$$\frac{\partial Q(x)}{\partial x_j} = \frac{\partial}{\partial x_j} \left(\frac{1}{2} d_j x_j^2 - x_j b_j \right) = d_j x_j - b_j = (Ax - b)_j = 0, \quad j = 1, 2, \dots, n,$$

ist es eindeutig und führt auf die Forderung $Ax - b = 0$, also auf die Lösung des LGS. Das Extremum ist wegen der Positivität von A ein Minimum.

Betrachten wir nun den allgemeinen Fall. Nach Teil (2) ist A ähnlich zu einer Diagonalmatrix $D = \text{diag}(d_1, d_2, \dots, d_n)$ mit einer orthogonalen Transformationsmatrix U ($U^T U = U U^T = I$), d. h. $D = U^T A U = U^{-1} A U$ oder $A = U D U^T$.

Es gilt $d_i > 0 \forall i$. Außerdem ist

$$Q(x) = \frac{1}{2} x^T U D U^T x - x^T U U^T b = \frac{1}{2} (U^T x)^T D (U^T x) - (U^T x)^T U^T b.$$

Wir definieren $z = U^T x$, $c = U^T b$ und

$$Q_D(z) = \frac{1}{2} z^T D z - z^T c.$$

Offenbar liegt das Minimum von $Q(x)$ genau dann in x , wenn das Minimum von $Q_D(z)$ in z angenommen wird. Das Minimum von $Q_D(z)$ wird gemäß Diagonalfall in dem Wert z erreicht, welcher Lösung des LGS $Dz = c$ ist.

Dann ist aber x genau die Lösung von $Ax = b$. \square

Die Minimaleigenschaft des Funktionals $Q(x)$ war schon in der letzten Beziehung von (4.4) zu erkennen.

Wie konstruiert man nun eine Folge von Näherungen zum Minimum von $Q(x)$?

Welche Form haben die Iterationsverfahren zu seiner Bestimmung?

Ausgehend von einem Startvektor $x^{(0)}$ als erste Näherung zu x^* sucht man weitere Iterierte $x^{(1)}$, $x^{(2)}$, ... eines Abstiegsverfahrens mit der Bedingung

$$Q(x^{(m+1)}) < Q(x^{(m)}),$$

also $\|e^{(m+1)}\|_A < \|e^{(m)}\|_A$, wobei $e^{(m)} = e(x^{(m)})$.

Dazu betrachtet man eine Folge "eindimensionaler Probleme" mit einer Strahlenminimierung.

Definition 4.1 Strahlenminimierung und Abstiegsverfahren (AV)

Gegeben sei die Iterierte $x^{(m)}$ des AV mit $Q(x^{(m)})$.

Mit einem Richtungsvektor $p^{(m)}$ ermittelt man diejenige Stelle $x^{(m+1)}$ auf der Geraden

$$g(\alpha) = x^{(m)} + \alpha p^{(m)}, \quad \alpha \in \mathbb{R}, \quad (4.5)$$

in welcher die Funktion

$$f(\alpha) = Q(x^{(m)} + \alpha p^{(m)}) \quad (4.6)$$

ihr Minimum annimmt.

$p^{(m)}$ ist eine lokal günstige Suchrichtung, der optimale Wert $\alpha = \alpha_{\min} = \alpha_m$ heißt Schrittzahl und die neue Iterierte im AV ist

$$x^{(m+1)} = x^{(m)} + \alpha_m p^{(m)}. \quad (4.7)$$

Wir notieren die Minimierungsaufgabe im m -ten Schritt in der kompakten Form

$$\boxed{\begin{aligned} Q(x^{(m+1)}) &= \min_{\alpha \in \mathbb{R}} Q(x^{(m)} + \alpha p^{(m)}) \\ x^{(m+1)} &= \arg \min_{\alpha \in \mathbb{R}} Q(x^{(m)} + \alpha p^{(m)}) \end{aligned}} \quad (4.8)$$

je nachdem, worauf wir den Schwerpunkt der Betrachtung legen.

Die Bestimmung des Parameters α und somit von $x^{(m+1)}$ in der Strahlenminimierung folgt aus dem folgenden Lemma.

Lemma 4.2 Gegeben seien zwei Vektoren x, p ($p \neq 0$).

Das Minimum der Funktion $f(\alpha) = Q(x + \alpha p)$ auf der Geraden $g(\alpha) = x + \alpha p$, $\alpha \in \mathbb{R}$, wird angenommen für

$$\alpha = \alpha_{min} = \frac{p^T r}{p^T A p}. \quad (4.9)$$

Beweis.

Dazu brauchen wir nur die skalare Funktion $f(\alpha)$ und bezüglich α die notwendige Bedingung an ihrem Minimum auszuwerten. Es gilt

$$\begin{aligned} f(\alpha) &= Q(x + \alpha p) = \frac{1}{2}(x + \alpha p)^T A(x + \alpha p) - (x + \alpha p)^T b \\ &= Q(x) + \alpha p^T (Ax - b) + \frac{1}{2} \alpha^2 p^T A p, \\ 0 = f'(\alpha) &= p^T (Ax - b) + \alpha p^T A p, \quad r = b - Ax, \\ \alpha &= \frac{p^T r}{p^T A p} = \frac{(p, r)}{(p, A p)} = \frac{(r, p)}{(A p, p)}, \\ f''(\alpha) &= p^T A p > 0, \quad p \neq 0. \end{aligned} \quad (4.10) \quad \square$$

Für die Schrittzahl haben wir verschiedene Darstellungen angegeben.

Wenn man also von $x^{(m)}$ in Richtung $p^{(m)}$ geht und zur Stelle $x^{(m+1)}$ kommt, so muss diese nicht unbedingt den geringsten Abstand von der Lösung x^* (im Sinne des üblichen euklidischen Abstands) haben. Es wäre durchaus denkbar, dass eine andere Stelle $x'^{(m+1)}$ näher an x^* liegt. In der nachfolgenden Abbildung ist z. B. das Lot von x^* auf die Gerade g gezeichnet worden, um $x'^{(m+1)}$ zu erhalten.

Der Abstand zwischen $x^{(m+1)}$ und x^* ist eventuell minimal im Sinne einer anderen Norm. Leider kennt man x^* nicht, um sein Lot (Projektion) auf die Gerade g zu fällen.

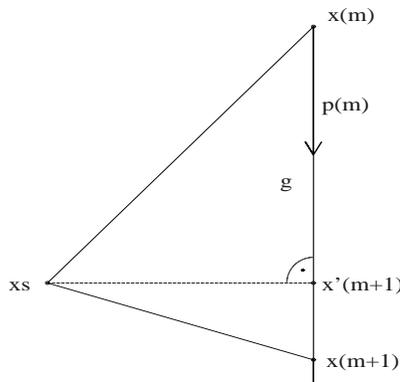


Abb. 4.1

Datei *abst_90.ps*

Geometrische Deutung

der Abstiegsituation

in der Draufsicht,

$x_s = x^*$, $z^{(m)} = z^{(m)}$

Wünschenswert in Bezug auf die Wahl von $x^{(m+1)}$ ist ein möglichst kleiner Abstand (euklidische Norm) $\|x^* - x^{(m+1)}\|_2$.

Anders ausgedrückt heißt dies, dass

$$x^* - x^{(m+1)} \perp p^{(m)} \quad \text{bzw.} \quad (x^* - x^{(m+1)}, p^{(m)}) = 0$$

ist. Damit ist der Vektor $x^{(m+1)} - x^{(m)}$ die Projektion von $x^* - x^{(m)}$ auf die Gerade g bzw. den Vektor $p^{(m)}$. Auf diesem Weg haben wir den Begriff der Projektionsmethode eingeführt.

In dem von uns betrachteten AV liegt jedoch das Augenmerk auf der Berechnung von Funktionswerten $Q(x)$ entlang eines Strahls, auf der Minimierung des Funktionals $Q(x)$ in Strahlenrichtung sowie damit wegen (4.4) auf der Minimierung von $\|e^{(m+1)}\|_A = \|x^* - x^{(m+1)}\|_A$. In [11] wurde gezeigt, dass bei $A = A^T > 0$ auch die Normwerte $\|e^{(m+1)}\|_2$ kontinuierlich kleiner werden.

In der folgenden Abbildung soll zumindest in einer etwas anderen Situation angedeutet werden, dass trotz einer Verkleinerung von $Q(x)$ bzw. $\|e(x)\|_A$ die Ungleichung $\|x^* - x^{(m+1)}\|_2 > \|x^* - x^{(m)}\|_2$ auftreten kann. Dabei stammt das Höhenlinienbild natürlich nicht von einer spd Matrix.

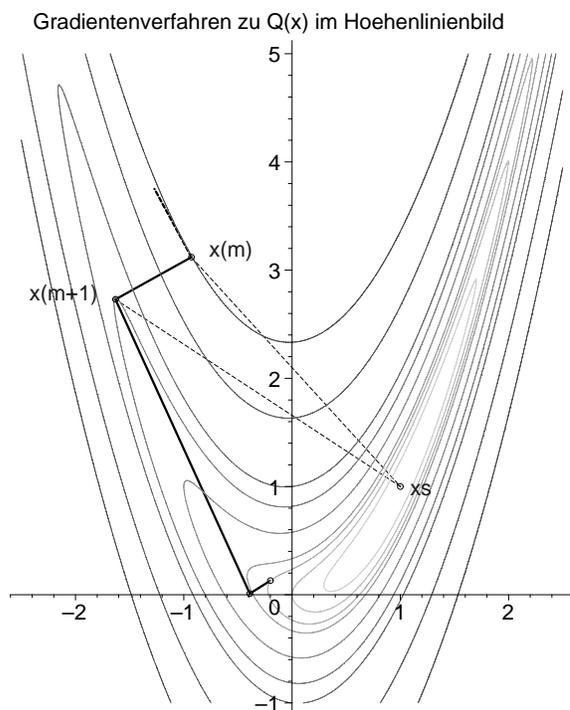


Abb. 4.2

Datei *abst_93.ps*

Abstiegssituation für $Q(x)$,
aber $\|x^* - x^{(m+1)}\|_2 > \|x^* - x^{(m)}\|_2$

Die Wahl der Suchrichtung p zum Funktional $Q(x)$ charakterisiert sowohl Algorithmengruppen als auch einzelne Methoden.

Unsere Betrachtungen beziehen sich auf das Gradientenverfahren, auch als Methode des steilsten Abstiegs bezeichnet.

4.2 Das Gradientenverfahren

Das klassische GV ist ein AV der Gestalt (4.7) und verwendet für das Funktional $Q(x)$ gemäß Formel (4.1) als Suchrichtung $p(x)$ die Richtung des steilsten Abstiegs des Funktionals (deepest descent method), die gleich dem negativen Gradient $-\nabla Q(x)$ ist und damit dem Residuum $r(x)$ entspricht.

$$p(x) = r(x) = -\nabla Q(x) = b - Ax. \quad (4.11)$$

Eine Normierung von $p(x)$ wäre denkbar, wird aber praktisch nicht vorgenommen. Die Durchführung der Strahlenminimierung mit der notwendigen Bedingung beim Minimum erlaubt die Berechnung der Schrittzahl (Suchschritt, Vielfache) α im GV.

$$\alpha = \alpha_m = \frac{r^{(m)T} r^{(m)}}{r^{(m)T} A r^{(m)}} = \frac{(r^{(m)}, r^{(m)})}{(A r^{(m)}, r^{(m)})} = \frac{\|r^{(m)}\|_2^2}{\|r^{(m)}\|_A^2} \geq 0. \quad (4.12)$$

Der Iterationsschritt als Aufdatierungsformel im GV ist

$$x^{(m+1)} = x^{(m)} + \alpha_m r^{(m)}. \quad (4.13)$$

Damit ergibt sich aus allen Schritten die monoton fallende Folge von Funktionswerten

$$Q(x^{(0)}) \geq Q(x^{(1)}) \geq \dots \geq Q(x^{(m)}) \geq Q(x^{(m+1)}) \geq \dots \geq Q(x^*), \quad (4.14)$$

bzw.

$$Q(x^{(m)}) > Q(x^{(m+1)}) \text{ bei } x^{(m)} \neq x^{(m+1)}.$$

Der wesentliche Aufwand in einem Schritt des Zyklus entsteht bei der Berechnung der Schrittzahl durch eine Matrix-Vektor-Multiplikation $A r^{(m)}$ sowohl für α_m als auch $r^{(m+1)}$, ansonsten sind höchstens Skalarprodukte auszuwerten.

Das Residuum selber kann man rekursiv bestimmen. Ausgehend von seiner expliziten Formel

$$r^{(m)} = b - A x^{(m)} = A x^* - A x^{(m)} = A e^{(m)} \quad (4.15)$$

bieten sich zwei Versionen der rekursiven Ermittlung an. Auf Unterschiede bei der Implementierung dieser Versionen der Residuumsberechnung wird in [37] eingegangen.

Für die rekursive Darstellung führen wir die Kürzel

$$v = Ax, \quad w = Ap, \quad t = Ar \quad (4.16)$$

ein.

Wegen $p = r$ ist im GV $w = Ar$ und die Größe t wird nicht benötigt.

$$\begin{aligned}
r^{(m+1)} &= b - Ax^{(m+1)} = b - A(x^{(m)} + \alpha_m r^{(m)}) \\
&= b - (Ax^{(m)} + \alpha_m Ar^{(m)}) \\
&= b - (v^{(m)} + \alpha_m w^{(m)}), \quad w^{(m)} = Ar^{(m)} \\
r^{(m+1)} &= b - v^{(m+1)}, \quad \text{wobei } v^{(m+1)} = v^{(m)} + \alpha_m w^{(m)} \\
&\quad (1. \text{ rekursive Berechnung}), \\
&= b - Ax^{(m)} - \alpha_m Ar^{(m)} \\
&= r^{(m)} - \alpha_m Ar^{(m)} \\
r^{(m+1)} &= r^{(m)} - \alpha_m w^{(m)} \quad (2. \text{ rekursive Berechnung}).
\end{aligned} \tag{4.17}$$

Mit den Abkürzungen notieren wir auch das Funktional

$$\begin{aligned}
Q(x) &= \frac{1}{2}x^T Ax - x^T b = -\frac{1}{2}x^T(2b - Ax) = -\frac{1}{2}x^T(b - Ax + b) \\
&= -\frac{1}{2}x^T s, \quad s = r + b = 2b - Ax = 2b - v \\
&= \frac{1}{2}(\|e(x)\|_A^2 - x^T b), \quad e(x) = x^* - x, \\
Q(x^{(m)}) &= -\frac{1}{2}x^{(m)T} s^{(m)} = -\frac{1}{2}x^{(m)T}(2b - v^{(m)}).
\end{aligned} \tag{4.18}$$

Die Funktionswerte $Q(x^{(m)})$ bzw. $\|e(x^{(m)})\|_A$ sind abnehmend und $Q(x^*) = -\frac{1}{2}x^{*T}b$. Den Verlauf von $Q(x^{(m)})$ kann man kontrollieren, genauso ist unter Beachtung der Beziehung (3.13) eine Kontrolle der Größen $\|r(x^{(m)})\|_2^2 = 2R(x^{(m)}) + b^T b$ möglich.

Bezüglich der Abbruchbedingungen für den Iterationsprozess eignen sich bei vorgegebener Toleranz $\varepsilon > 0$ wegen $r^{(m)} \rightarrow 0$ bei $e^{(m)} \rightarrow 0$ (siehe (4.15)) und der einfacheren Handhabung des Residuums beispielsweise folgende Kriterien:

- Iterationsanzahl begrenzen, $m \leq m_{max}$,
- $\|r^{(m)}\|_2 < \varepsilon$,
- $\|r^{(0)}\|_2^2 < \varepsilon$ bzw. $\frac{r^{(m)T} r^{(m)}}{r^{(0)T} r^{(0)}} = \frac{\|r^{(m)}\|_2^2}{\|r^{(0)}\|_2^2} < \varepsilon$ bei $m > 0$,
- $\|x^{(m+1)} - x^{(m)}\|_2 < \varepsilon$,
- $|Q(x^{(m+1)}) - Q(x^{(m)})| < \varepsilon$.

Meistens wird der relative Fehler mit dem Residuumnormquadrat verwendet.

Neben der Wahl geeigneter Abbruchbedingungen ist im Fall beliebiger Matrizen noch ein zusätzlicher Test auf die Durchführbarkeit der Berechnung von

$$\alpha_m = \frac{r^{(m)T} r^{(m)}}{r^{(m)T} Ar^{(m)}} = \frac{\|r^{(m)}\|_2^2}{r^{(m)T} Ar^{(m)}}$$

mit Nenner $\neq 0$ zu empfehlen.

Weiterhin kann sich eine Kontrolle zur Divergenz gemäß $\|x^{(m+1)} - x^{(m)}\|_2^2 > 10^k \gg 1$ als sinnvoll erweisen.

Die Such- und Abstiegsrichtungen $r^{(m)}$ erfüllen eine Orthogonalitätsbedingung. Es gilt für aufeinander folgende Richtungen $r^{(m+1)} \perp r^{(m)}$, denn mit (4.17) haben wir

$$\begin{aligned} (r^{(m+1)}, r^{(m)}) &= (r^{(m)}, r^{(m)}) - \alpha_m (Ar^{(m)}, r^{(m)}) \\ &= (r^{(m)}, r^{(m)}) - \frac{(r^{(m)}, r^{(m)})}{(Ar^{(m)}, r^{(m)})} (Ar^{(m)}, r^{(m)}) \\ &= 0. \end{aligned}$$

Die Orthogonalität ist aber nicht transitiv, so dass aus $r^{(m+1)} \perp r^{(m)}$ und $r^{(m)} \perp r^{(m-1)}$ nicht notwendigerweise $r^{(m+1)} \perp r^{(m-1)}$ folgt.

Varianten der Realisierung des GV als (unendliches) Iterationsverfahren

Version 1 GV mit Indizierung

- $x^{(0)}$ Startvektor,
- $r^{(0)} = b - Ax^{(0)}$ Anfangsresiduum, Abstiegsrichtung, negativer Gradient,
- ε Toleranz für den Test auf Abbruch der Iteration.

m = 0,1,2,...

falls $\|r^{(m)}\| < \varepsilon$, dann break

$$\alpha_m = \frac{r^{(m)T} r^{(m)}}{r^{(m)T} w^{(m)}}, \quad \text{wobei } w^{(m)} = Ar^{(m)}$$

$$x^{(m+1)} = x^{(m)} + \alpha_m r^{(m)}$$

$$r^{(m+1)} = r^{(m)} - \alpha_m w^{(m)} \quad \text{oder} \quad r^{(m+1)} = b - Ax^{(m+1)}$$

(4.19)

end m

Näherungslösung $x^* = x^{(m)}$

Version 2 GV mit Test auf relativen Fehler

- K maximale Iterationsanzahl,
- $x^{(0)}$ Startvektor,
- $v^{(0)} = Ax^{(0)}$ Hilfsvektor,
- $r^{(0)} = b - v^{(0)}$ Anfangsresiduum, Abstiegsrichtung, negativer Gradient,
- $\gamma[0..K]$ Vektor der Normquadrate der euklidischen Norm, $\gamma_m = \|r^{(m)}\|_2^2$,
- ε Toleranz für den Test auf Abbruch der Iteration,
- falls $\gamma_0 < \varepsilon$, dann $x^* = x^{(0)}$ und Stopp.

m = 0,1,...,K

$$\alpha_m = \frac{\gamma_m}{r^{(m)T}w^{(m)}}, \text{ wobei } w^{(m)} = Ar^{(m)}$$

$$x^{(m+1)} = x^{(m)} + \alpha_m r^{(m)}$$

$$\left\{ \begin{array}{l} v^{(m+1)} = v^{(m)} + \alpha_m w^{(m)} \\ r^{(m+1)} = b - v^{(m+1)} \end{array} \right\} \text{ oder } r^{(m+1)} = b - Ax^{(m+1)} \quad (4.20)$$

$$\gamma_{m+1} = \|r^{(m+1)}\|_2^2, \text{ falls } \frac{\gamma_{m+1}}{\gamma_0} < \varepsilon, \text{ dann break}$$

end m

Näherungslösung $x^* = x^{(m+1)}$

Bemerkung 4.1 Das Produkt $w^{(m)} = Ar^{(m)}$ tritt in den Formeln für α_m und $r^{(m+1)}$ auf und ist nur einmal zu berechnen.

Version 3 GV ohne Indizierung

K maximale Iterationsanzahl,

x Startvektor,

$v = Ax$ Hilfsvektor,

$r = b - v$ Anfangsresiduum, Abstiegsrichtung, negativer Gradient,

$\gamma[0..K]$ Vektor der Normquadrate der euklidischen Norm von r ,

$\gamma_0 = r^T r = \|r\|_2^2$,

ε Toleranz für den Test auf Abbruch der Iteration,

falls $\gamma_0 < \varepsilon$, dann $x^* = x$ und Stopp.

m = 1,2,...,K

$w = Ar$

$\alpha = \frac{\gamma_{m-1}}{d}$, wobei $d = r^T w = w^T r$

$x = x + \alpha r$

$$\left\{ \begin{array}{l} v = v + \alpha w \\ r = b - v \end{array} \right\} \text{ oder } r = b - Ax \text{ oder } r = r - \alpha w \quad (4.21)$$

$\gamma_m = \|r\|_2^2$, falls $\frac{\gamma_m}{\gamma_0} < \varepsilon$, dann break

end m

Näherungslösung $x^* = x$

Wenn A nur regulär ist, dann bildet man das Normalgleichungssystem $A^T Ax = Bx = c = A^T b$ mit $B = B^T > 0$, auf welches das GV angewendet werden kann. Von der Kondition der Koeffizientenmatrix her ist aber die Symmetrisierung des LGS nicht unbedingt der günstigste Weg, auch wenn sie die Matrix in gewisser Weise skaliert. Daraus ergibt sich das zu minimierende Funktional

$$\begin{aligned} R(x) &= \frac{1}{2}x^T Bx - x^T c = -\frac{1}{2}(Ax)^T(2b - Ax) \\ &= -\frac{1}{2}(Ax)^T s, \quad s = r + b = 2b - Ax = 2b - v, \\ &= \frac{1}{2}(\|r(x)\|_2^2 - b^T b), \quad r(x) = b - Ax. \end{aligned}$$

Die Matrix B hat im Allgemeinen eine schlechtere Kondition, so dass meistens mehr Gradientenschritte erforderlich sind.

Bei Implementierungen wird die Matrix $B = A^T A$ wegen des Aufwands von $2n^3$ Operationen nicht explizit ermittelt, sondern im GV hat man anstelle einer Matrix-Vektor-Multiplikation in der Hauptschleife nun zwei.

Version 4 GV mit Symmetrisierung für beliebiges A

K maximale Iterationsanzahl,

x Startvektor,

$v = A^T(Ax)$, $c = A^T b$ Hilfsvektoren,

$\tilde{r} = c - v$ Anfangsresiduum, Abstiegsrichtung, negativer Gradient,

$\gamma[0..K]$ Vektor der Normquadrate der euklidischen Norm von \tilde{r} ,

$\gamma_0 = \tilde{r}^T \tilde{r} = \|\tilde{r}\|_2^2$,

ε Toleranz für den Test auf Abbruch der Iteration,

falls $\gamma_0 < \varepsilon$, dann $x^* = x$ und Stopp.

m = 1,2,...,K

$$w = A^T(A\tilde{r})$$

$$\alpha = \frac{\gamma_{m-1}}{d}, \quad \text{wobei } d = \tilde{r}^T w = w^T \tilde{r}$$

$$x = x + \alpha \tilde{r}$$

$$\left\{ \begin{array}{l} v = v + \alpha w \\ \tilde{r} = c - v \end{array} \right\} \quad \text{oder } \tilde{r} = c - A^T(Ax) \quad \text{oder } \tilde{r} = \tilde{r} - \alpha w$$

$$\gamma_m = \|\tilde{r}\|_2^2, \quad \text{falls } \frac{\gamma_m}{\gamma_0} < \varepsilon, \quad \text{dann break}$$

end m

Näherungslösung $x^* = x$

(4.22)

Da LGS mit Koeffizientenmatrizen gleich oder nahe der Einheitsmatrix der ideale Fall sind, betrachten wir noch kurz das GV mit Vorkonditionierung (preconditioned gradient method). Mit einer einfachen regulären Vorkonditionierungsmatrix C (preconditioner) transformieren wir das LGS durch beidseitige Multiplikation gemäß

$$C^{-1}Ax = Bx = c = C^{-1}b. \quad (4.23)$$

Im GV selbst soll der Mehraufwand, der nun durch die zusätzliche Lösung eines Gleichungssystems entstehen wird, von der Größenordnung nicht den einer Matrix-Vektor-Multiplikation übersteigen. Das wäre dann der Fall, wenn z. B. die Matrix C in Produktform $C = LL^T > 0$ mit der Dreiecksmatrix L vorliegt und die Lösung des Gleichungssystems auf gestaffelte Systeme führt.

Die Lösung des Gleichungssystems $Cs = r$ erfolgt dann in zwei Schritten mit den Dreieckssystemen $Lz = r$ und $L^T s = z$.

Version 5 Vorkonditioniertes GV mit Indizierung

- $x^{(0)}$ Startvektor,
- $r^{(0)} = b - Ax^{(0)}$ Anfangsresiduum, negativer Gradient,
- ε Toleranz für Test auf Abbruch der Iteration,
- C Vorkonditionierungsmatrix.

m = 0,1,2,...

falls $\|r^{(m)}\| < \varepsilon$, dann break

$$Cs^{(m)} = r^{(m)}, \quad w^{(m)} = As^{(m)}, \quad \alpha_m = \frac{r^{(m)T} s^{(m)}}{w^{(m)T} s^{(m)}}$$

$$x^{(m+1)} = x^{(m)} + \alpha_m s^{(m)}$$

$$r^{(m+1)} = r^{(m)} - \alpha_m w^{(m)}$$

(4.24)

end m

Näherungslösung $x^* = x^{(m)}$

Wie man zu den leicht modifizierten Darstellungen in Version 5 im Vergleich mit Version 1 gelangt, soll kurz erläutert werden.

Ausgangspunkt ist einfach das Funktional $Q(x + \alpha s)$ zum LGS (2.1) und die neue Suchrichtung $s = C^{-1}r$, $r = b - Ax$, in der Strahlenminimierung.

Damit ergibt sich wie in Lemma 4.2 die Schrittzahl

$$\alpha = \frac{s^T r}{s^T A s} = \frac{r^T s}{(A s)^T s},$$

so dass wir mit $Cs = r$ und $w = As$ auf die Beziehungen in Version 5 kommen.

Eine Implementierung des GV soll sich auf die Versionen 2 (4.20) und 3 (4.21) stützen. Dieser Algorithmus liefert in Maple die folgenden Kommandos.

```
# GV, Versionen 2,3
x:=evalm(x0):
v:=evalm(A&*x):
r:=evalm(b-v):
gamma0:=evalm(transpose(r)&*r):

if gamma0<etol then RETURN(x,0); end if;
gamman:=gamma0:

for k from 1 by 1 to maxiter do
  w:=evalm(A&*r);
  d:=evalm(transpose(r)&*w);

  alpha:=gamman/d;
  x:=evalm(x+alpha*r);
  v:=evalm(v+alpha*w);
  r:=evalm(b-v);
  gamman:=evalm(transpose(r)&*r);

  if gamman/gamma0<etol then break; end if;  # relativer Fehler
end do:

if k>maxiter then k:=k-1; end if;
[x,k];
```

Ergänzt um die zusätzliche Abfrage des Nenners bei der Berechnung der Schrittzahl α , die Bestimmung von Werten des Funktionals sowie einige Ergebnisfelder und Zwischenausgaben entsteht dann die erweiterte Maple-Prozedur in Anlehnung an diese Versionen des GV, die wir in den Beispielrechnungen verwenden.

```
> gv:=proc(n::posint,A::matrix,b::vector,x0::vector,
           maxiter::posint,etol::numeric,aus::name)

  local k,i,m,v,r,w,d,gamma0,gamman,x,alpha,Q,fh,fh1,fh2,xv1,rv1,xx;
  global xv,rv,lr;

  x:=evalm(x0):
  v:=evalm(A&*x):
  r:=evalm(b-v):  lr:=evalm(r):
  gamma0:=evalm(transpose(r)&*r):
  Q:=0.5*evalm(transpose(x)&*A&*x)-evalm(transpose(x)&*b);

  fh2:='%+.16e'; # Ausgabeformate einstellen
  fh1:='%+.16e';
  fh :=fh1;
  for m from 2 to n do
    fh:=cat(fh,' ',fh1);
  end do;
```

```

xx:=matrix(n,0,[]):
xv1:=concat(xx,x);
xv:=evalm(xv1);
rv1:=concat(xx,r);
rv:=evalm(rv1);

if aus=ja then
  printf('\n'):
  printf('Startvektor          x = [||fh||']\n',seq(x[i],i=1..n));
  printf('Residuum           r = b-Ax = [||fh||']\n',seq(r[i],i=1..n));
  printf('Funktionswert      Q(x) = [||fh1||']\n',Q);
  printf('Anfangsfehlerquadrat r'r = [||fh2||']\n\n',gamma0);
end if;

if gamma0<etol then RETURN(x,0); end if;
gamman:=gamma0;

for k from 1 by 1 to maxiter do
  w:=evalm(A&*r);
  d:=evalm(transpose(r)&*w);
  if d=0 then
    lprint('Abbruch wegen Nenner r'Ar=0'):
    RETURN(x,k-1);
  end if;

  alpha:=gamman/d;
  x:=evalm(x+alpha*r);
  v:=evalm(v+alpha*w);
  r:=evalm(b-v);  lr:=evalm(r);
  gamman:=evalm(transpose(r)&*r);

  Q:=0.5*evalm(transpose(x)&*A&*x)-evalm(transpose(x)&*b);
  xv1:=concat(xv,x);
  xv:=evalm(xv1);
  rv1:=concat(rv,r);
  rv:=evalm(rv1);

  if aus=ja then
    printf('Schritt k = %g\n',k);
    printf('Suchrichtung    p=r=b-Ax = [||fh||']\n',seq(r[i],i=1..n));
    printf('Suchschritt      alpha = [||fh1||']\n',alpha);
    printf('Iterationsvektor    x = [||fh||']\n\n',seq(x[i],i=1..n));
    printf('Residuum           r = b-Ax = [||fh||']\n',seq(r[i],i=1..n));
    printf('Funktionswert      Q(x) = [||fh1||']\n',Q);
    printf('Fehlernormquadrat    r'r = [||fh2||']\n\n',gamman);
  end if;

  if gamman/gamma0<etol then break; end if;  # relativer Fehler
end do;

if k>maxiter then k:=k-1; end if;
[x,k];

end:

```

4.2.1 Konvergenz des GV mittels Richardson-Extrapolation

Um Konvergenzaussagen zum GV zu machen, sind einige vorbereitende Überlegungen zu Iterationsverfahren anzustellen.

Das LGS (2.1) kann man zunächst als Nullstellengleichung $0 = b - Ax$ notieren und daraus die Gestalt einer Fixpunktgleichung

$$x = x + b - Ax = (I - A)x + b, \quad I \text{ Einheitsmatrix,} \quad (4.25)$$

ableiten. Damit ergeben sich mehrere äquivalente Formen der Iterationsformel

$$\begin{aligned} x^{(m+1)} &= (I - A)x^{(m)} + b \\ &= x^{(m)} - I^{-1}(Ax^{(m)} - b) \\ &= x^{(m)} + r^{(m)}, \quad r^{(m)} = r(x^{(m)}) = b - Ax^{(m)} \quad (\text{Residualform}), \end{aligned} \quad (4.26)$$

mit

$$\begin{aligned} H &= I - A \quad \text{als Iterationsmatrix,} \\ W &= I \quad \text{als Wichtung, Skalierungsmatrix.} \end{aligned}$$

Diese Umformungen sind jedoch trivial, so dass wir es als naives Iterationsverfahren bezeichnen wollen.

Eleganter und gebräuchlicher ist das Gesamtschrittverfahren (GSV, Jacobi-Iterationsverfahren), das unter der Voraussetzung $a_{ii} \neq 0$, $i = 1, 2, \dots, n$, in der Matrix-Vektor-Form

$$\begin{aligned} x^{(m+1)} &= (I - D^{-1}A)x^{(m)} + D^{-1}b, \quad D = \text{diag}(A) = \text{diag}(a_{11}, a_{22}, \dots, a_{nn}) \\ &= x^{(m)} - D^{-1}(Ax^{(m)} - b) \\ &= x^{(m)} + D^{-1}r^{(m)} \end{aligned} \quad (4.27)$$

notiert werden kann, wobei

$$\begin{aligned} H &= J = I - D^{-1}A \quad \text{Iterationsmatrix,} \\ W &= D \quad \text{Wichtung, Skalierungsmatrix,} \\ &\quad D^{-1}r^{(m)} \quad \text{gewichtetes Residuum (weighted residual),} \\ &\quad \text{Verbesserung (update) für } x^{(m)}. \end{aligned}$$

Auf diese Art und Weise können weitere Basisverfahren definiert werden.

Das allgemeine Iterationsverfahren

$$x^{(m+1)} = Hx^{(m)} + c \quad (4.28)$$

konvergiert für jede beliebige Startnäherung $x^{(0)}$ gegen die Lösung x^* , wenn $\|H\| < 1$ gilt. Hinreichendes und notwendiges Konvergenzkriterium ist $\rho(H) < 1$. Je kleiner die Norm oder der Spektralradius der Iterationsmatrix H sind, desto schneller konvergiert das Verfahren. Dies kann man mit einer günstigen Wahl der Skalierungsmatrix eventuell bewirken.

Hier wollen wir jedoch die Wichtung möglichst einfach wählen.

Somit liegt nahe, mit der parameterabhängigen Matrix $W = \frac{1}{\omega}I$, $\omega > 0$, zu arbeiten. Den Parameter ω nennt man Relaxationsparameter. Die Vorgehensweise beschreibt die so genannte Richardson-Iteration (auch semiiterative Richardson-Methode oder Richardson-Relaxation genannt) mit festem Parameter (RF)

$$\begin{aligned} x^{(m+1)} &= x^{(m)} - \left(\frac{1}{\omega}I\right)^{-1} (Ax^{(m)} - b) \\ &= (I - \omega A)x^{(m)} + \omega b \\ &= x^{(m)} + \omega r^{(m)}, \quad r^{(m)} = b - Ax^{(m)}, \end{aligned} \quad (4.29)$$

wobei

$$\begin{aligned} H &= I - \omega A \quad \text{Iterationsmatrix,} \\ W &= \frac{1}{\omega}I \quad \text{Wichtung, Skalierungsmatrix.} \end{aligned}$$

Bei geeigneter Wahl dieses Relaxationsparameters wird das Iterationsverfahren konvergieren und möglicherweise besonders schnell.

Unter der Voraussetzung $A = A^T > 0$ untersuchen wir die symmetrische Iterationsmatrix $H = I - \omega A$, speziell das Verhalten ihrer Eigenwerte.

Wir gehen von den reellen Eigenwerten der spd Matrix A aus, die der Bedingung (Ungleichungskette)

$$0 < \lambda = \lambda_1(A) \leq \lambda_2(A) \leq \dots \leq \lambda_n(A) = \Lambda = \rho(A) \quad (4.30)$$

genügen. Weiter ist die Spektralnorm $\|A\|_2 = \rho(A)$.

Die Eigenwerte von $H = H(\omega)$ sind dann

$$\lambda(H(\omega)) = 1 - \omega \lambda(A). \quad (4.31)$$

Für ihr Verhalten in Abhängigkeit vom Parameter $\omega \geq 0$ braucht man nur die beiden "äußeren" Eigenwerte zu betrachten. Es gilt

$$\lambda_n(H(\omega)) = 1 - \omega \lambda_n(A) \leq \lambda_i(H(\omega)) \leq \lambda_1(H(\omega)) = 1 - \omega \lambda_1(A) \leq 1, \quad i = 2, 3, \dots, n-1.$$

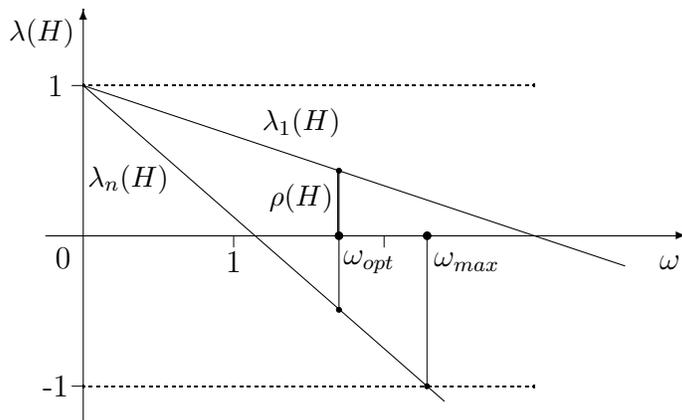


Abb. 4.3 Datei *abst2.pic*
Verlauf der Funktionen
 $\lambda_i(H) = 1 - \omega \lambda_i(A)$, $i = 1, n$,
 $H = H(\omega)$,
in Abhängigkeit
vom Parameter $\omega \geq 0$

Die Abschätzung der Eigenwerte $\lambda(H(\omega))$ hat sowohl das Ziel, den Spektralradius $\rho(H(\omega)) < 1$ zu machen, als auch noch möglichst klein durch eine günstige Wahl des Parameters ω .

Dazu genügt es das Verhalten des Eigenwerts $\lambda_n(H(\omega)) = 1 - \omega\lambda_n(A)$ als Geradengleichung in Abhängigkeit von $\omega \geq 0$ bei abfallendem Verlauf der Geraden für den ersten Grenzfall $\lambda_n(H(\omega)) = -1$ zu untersuchen. Daraus erhält man

$$\omega_{max} = \frac{2}{\lambda_n(A)} = \frac{2}{\Lambda} = \frac{2}{\rho(A)} \quad (4.32)$$

und die Folgerung $\rho(H(\omega)) < 1$ für $\omega \in (0, \omega_{max})$.

Der Spektralradius $\rho(H(\omega))$ wird jedoch am kleinsten im zweiten Grenzfall $-\lambda_n(H(\omega)) = \lambda_1(H(\omega))$, was auf den optimalen Parameter

$$\omega_{opt} = \frac{2}{\lambda_1(A) + \lambda_n(A)} = \frac{2}{\lambda + \Lambda} < \omega_{max} \quad (4.33)$$

führt.

Ist $\lambda_1(A) \approx 0$, so liegt ω_{opt} nahe bei ω_{max} .

Spezielle Werte von ω sind z. B. $\omega = \frac{1}{\lambda_n(A)} = \frac{1}{\rho(A)} < \omega_{opt}$ oder $\omega = \frac{1}{\|A\|} \leq \frac{1}{\|A\|_2} = \frac{1}{\rho(A)}$.

Bei optimaler Situation der Konvergenz beträgt der Spektralradius der Iterationsmatrix

$$\eta = \rho(H(\omega_{opt})) = 1 - \omega_{opt} \lambda_1(A) = \frac{\Lambda - \lambda}{\Lambda + \lambda} = \frac{\frac{\Lambda}{\lambda} - 1}{\frac{\Lambda}{\lambda} + 1} = \frac{\kappa(A) - 1}{\kappa(A) + 1} < 1. \quad (4.34)$$

Das Ergebnis fassen wir als Satz zusammen.

Satz 4.3 Falls $A = A^T > 0$ gilt, ist die Richardson-Iteration (4.29) für einen bestimmten Parameterbereich $\omega \in (0, \omega_{max})$, $\omega_{max} > 0$, konvergent. Unter diesen gibt es einen optimalen Wert $\omega = \omega_{opt}$.

Unter den gleichen Voraussetzungen arbeiten das GV (4.13) mit variablem Parameter

$$x^{(m+1)} = x^{(m)} + \alpha_m r^{(m)}$$

und die Richardson-Iteration (4.29) mit festem Parameter

$$x^{(m+1)} = x^{(m)} + \omega_{opt} r^{(m)} = (I - \omega_{opt} A)x^{(m)} + \omega_{opt} b = Hx^{(m)} + c \quad (4.35)$$

mit derselben Struktur der Iterationsformel sowie jeweils optimaler Wahl der Schrittzahl in der Abstiegsrichtung $r^{(m)} = b - Ax^{(m)}$. Somit kann das GV als spezielle Richardson-Iteration mit variablem Parameter interpretiert werden und es entstehen vergleichbare Iterationsfolgen. Damit übertragen sich Konvergenzaussagen, Fehleranalyse und Abschätzungen zur Richardson-Iteration (4.35) auf das GV.

Satz 4.4 Für den Fehlervektor $e^{(m)}$ des GV (4.13) mit (4.11), (4.12) gilt bei dem Faktor η aus Formel (4.34) die Abschätzung

$$\|e^{(m)}\|_A \leq \eta^m \|e^{(0)}\|_A. \quad (4.36)$$

Beweis.

Da der Startvektor $x^{(0)} \in \mathbb{R}^n$ beliebig ist, reicht der Nachweis der Abschätzung (4.36) für $m = 1$. Sei also $x^{(1)} = x^{(0)} + \alpha_0 r^{(0)}$.

Wir haben im RF mit optimalem Gewichtsparameter ω_{opt} die Größen $H = H(\omega_{opt})$, $\|H\|_2 = \rho(H) = \eta$, $x_{RF}^{(1)} = x^{(0)} + \omega_{opt} r^{(0)}$ und den Fehlervektor

$$x^* - x_{RF}^{(1)} = e_{RF}^{(1)} = H e^{(0)}.$$

Zudem gilt für beliebiges $q \in \mathbb{R}$ die Gleichheit $HA^q = A^q H$ und es folgt mit

$$\tilde{e}^{(0)} = A^{1/2} e^{(0)}, \quad \tilde{e}_{RF}^{(1)} = A^{1/2} e_{RF}^{(1)}$$

die Beziehung

$$\tilde{e}_{RF}^{(1)} = A^{1/2} e_{RF}^{(1)} = A^{1/2} H e^{(0)} = H A^{1/2} e^{(0)} = H \tilde{e}^{(0)}.$$

Hiermit erhält man

$$\begin{aligned} \|e_{RF}^{(1)}\|_A &= (A e_{RF}^{(1)}, e_{RF}^{(1)})^{1/2} = (A^{1/2} e_{RF}^{(1)}, A^{1/2} e_{RF}^{(1)})^{1/2} \\ &= \|\tilde{e}_{RF}^{(1)}\|_2 = \|H \tilde{e}^{(0)}\|_2 \leq \\ &\leq \|H\|_2 \|\tilde{e}^{(0)}\|_2 = \|H\|_2 \|A^{1/2} e^{(0)}\|_2 \\ &= \|H\|_2 \|e^{(0)}\|_A = \eta \|e^{(0)}\|_A. \end{aligned}$$

Wegen

$$x^{(1)} = \arg \min_{x \in x^{(0)} + \text{span}\{r^{(0)}\}} Q(x) = \arg \min_{x \in x^{(0)} + \text{span}\{r^{(0)}\}} \left(\frac{1}{2} \|x^* - x\|_A^2 - \frac{1}{2} x^{*T} b \right)$$

im GV folgt

$$\|x^* - x^{(1)}\|_A \leq \|x^* - x_{RF}^{(1)}\|_A = \|e_{RF}^{(1)}\|_A$$

und

$$\|e^{(1)}\|_A \leq \eta \|e^{(0)}\|_A.$$

Das GV ist also lokal besser als das optimale RF, so dass für den m -ten Fehlervektor $e^{(m)}$ die Schranke η^m auftritt. \square

Aus (4.4) und der Abschätzung (4.36) ergeben sich auch

$$\begin{aligned} \|e^{(m)}\|_A^2 &\leq \eta^{2m} \|e^{(0)}\|_A^2, \\ Q(x^{(m)}) - Q(x^*) &\leq \eta^{2m} [Q(x^{(0)}) - Q(x^*)], \end{aligned} \quad (4.37)$$

$$Q(x^{(m)}) \leq \eta^2 Q(x^{(m-1)}) + (1 - \eta^2)Q(x^*) \leq \eta^2 Q(x^{(m-1)}) \leq \eta^{2m} Q(x^{(0)}),$$

wobei die letzte Ungleichung nur für $(Q(x^*) =) 0 \leq Q(x^{(m)}) \leq Q(x^{(0)})$ brauchbar ist. Damit erkennt man jedoch allgemein die lineare Konvergenz des GV.

Wir wollen noch einige Bemerkungen zum Iterationsfehler von (4.35) $e_{RF}^{(m)} = x^* - x_{RF}^{(m)}$ machen, woraus man auch auf das Verhalten des GV dann schließen kann. Er genügt mit $H = H(\omega_{opt})$ der Beziehung

$$e_{RF}^{(m)} = x^* - x_{RF}^{(m)} = Hx^* + c - (Hx_{RF}^{(m-1)} + c) = He_{RF}^{(m-1)} = H^m e^{(0)}. \quad (4.38)$$

Die symmetrische Matrix H hat n linear unabhängige Eigenvektoren u_1, u_2, \dots, u_n (Basis in \mathbb{R}^n) mit den Eigenwerten $\lambda_1(H), \lambda_2(H), \dots, \lambda_n(H)$, wobei $1 > \eta = \rho(H) = \lambda_1(H) = |\lambda_1(H)| \geq |\lambda_i(H)|$, $i = 2, 3, \dots, n$.

Der Ausgangsfehler lässt sich in der Basis $\{u_i\}$ darstellen, und die weiteren Fehlervektoren genügen den folgenden Formeln.

$$\begin{aligned} e^{(0)} &= \sum_{i=1}^n c_i u_i, \\ e_{RF}^{(1)} &= He^{(0)} = H \sum_{i=1}^n c_i u_i = \sum_{i=1}^n c_i \lambda_i(H) u_i, \\ e_{RF}^{(m)} &= He_{RF}^{(m-1)} = H^m e^{(0)} = \sum_{i=1}^n c_i [\lambda_i(H)]^m u_i. \end{aligned} \quad (4.39)$$

Die Konvergenz $e_{RF}^{(m)} \rightarrow 0$ bei $m \rightarrow \infty$ und beliebigem Startvektor folgt aus den Ungleichungen $|\lambda_i(H)| < 1$ für alle i bzw. aus $\eta < 1$. Für großes m dominiert dann in (4.39) der Fehlerterm $c_1 [\lambda_1(H)]^m u_1$.

Für beliebige kompatible Normen folgt aus (4.39) die Abschätzung

$$\|e_{RF}^{(m)}\| = \|H^m e^{(0)}\| \leq \|H^m\| \|e^{(0)}\| \leq \|H\|^m \|e^{(0)}\|. \quad (4.40)$$

Bezüglich der Norm folgt ebenfalls die Konvergenz unter Verwendung des folgenden Lemmas.

Lemma 4.5 *Zu einer gegebenen Matrix C mit ihrem Spektralradius $\rho(C)$ gibt es für ein beliebiges $\varepsilon > 0$ eine Matrixnorm $\|\cdot\|$ mit $\|C\| \leq \rho(C) + \varepsilon$*

Beweis. Satz 2.52 in [18] zu Matrixnorm und Spektralradius.

Das heißt insbesondere, wenn wir von $\eta = \rho(H) < 1$ ausgehen, dann gilt auch für die Norm die Ungleichung $\|H\| < 1$, weil ja ε beliebig klein sein kann.

Damit ist wegen (4.40) und $\lim_{m \rightarrow \infty} \|H\|^m = 0$ die Beziehung

$$\lim_{m \rightarrow \infty} \|e_{RF}^{(m)}\| = 0$$

für alle $e^{(0)}$ erfüllt. Mit den Normeigenschaften bedeutet das

$$\lim_{m \rightarrow \infty} x_{RF}^{(m)} = x^*.$$

Für die symmetrische Iterationsmatrix H kann die in Lemma 4.5 benötigte Norm leicht angegeben werden. Es ist nämlich die Spektralnorm $\|\cdot\|_2$, denn

$$\|H\|_2 = \sqrt{\rho(H^T H)} = \sqrt{\rho(H^2)} = \sqrt{\rho(H)^2} = \eta < \eta + \varepsilon.$$

Für $A = A^T > 0$ möchten wir noch eine Fehlerabschätzung in der A -Norm $\|\cdot\|_A$ erhalten, weil die Minimierung des Funktionals $Q(x)$ auch die von $\|e(x)\|_A$ bedeutet. Dazu brauchen wir noch eine weitere Hilfsaussage zur Matrix $A^{1/2}$.

Lemma 4.6 *Sei A symmetrisch und positiv definit. Dann gibt es eine eindeutige Matrix F , ebenfalls symmetrisch und positiv definit, mit $F^2 = A$.*

Beweis. Satz 2.18 in [18] allgemein für A hermitesch.

Damit gelingt uns der Nachweis der Verträglichkeit mit der A -Norm.

Wegen

$$\begin{aligned} e_{RF}^{(m+1)} &= H e_{RF}^{(m)}, \\ \|e_{RF}^{(m+1)}\|_2 &\leq \|H\|_2 \|e_{RF}^{(m)}\|_2 = \eta \|e_{RF}^{(m)}\|_2 \end{aligned} \tag{4.41}$$

und $\|A^{1/2}x\|_2^2 = (A^{1/2}x, A^{1/2}x) = (Ax, x) = (x, x)_A = \|x\|_A^2$ erhält man für die Vektoren $\tilde{e}_{RF}^{(m)} = A^{1/2}e_{RF}^{(m)}$ die Beziehungen

$$\begin{aligned} \tilde{e}_{RF}^{(m+1)} &= A^{1/2}e_{RF}^{(m+1)} = A^{1/2}H e_{RF}^{(m)} = A^{1/2}(I - \omega_{opt}A)e_{RF}^{(m)} \\ &= (A^{1/2} - \omega_{opt}A^{1/2}A)e_{RF}^{(m)} = (A^{1/2} - \omega_{opt}AA^{1/2})e_{RF}^{(m)} \\ &= (I - \omega_{opt}A)A^{1/2}e_{RF}^{(m)} = H\tilde{e}_{RF}^{(m)}, \end{aligned}$$

$$\|\tilde{e}_{RF}^{(m+1)}\|_2 \leq \eta \|\tilde{e}_{RF}^{(m)}\|_2,$$

und somit in verschiedenen Varianten die Fehlerschätzungen bzw. Abschätzungen

$$\|A^{1/2}e_{RF}^{(m+1)}\|_2 \leq \eta \|A^{1/2}e_{RF}^{(m)}\|_2,$$

$$\|e_{RF}^{(m+1)}\|_A \leq \eta \|e_{RF}^{(m)}\|_A,$$

$$\|e_{RF}^{(m)}\|_A \leq \eta^m \|e^{(0)}\|_A,$$

$$\|e_{RF}^{(m)}\|_A^2 \leq \eta^{2m} \|e^{(0)}\|_A^2.$$

Die Abschätzung $\|e_{RF}^{(m+1)}\|_2 \leq \eta \|e_{RF}^{(m)}\|_2$ aus (4.41) ist scharf.

Man kann also eine Situation konstruieren, wo die Gleichheit zutrifft. Sei

$$\begin{aligned} v_1 & \text{ Eigenvektor von } A \text{ zu } \Lambda \text{ mit } \|v_1\|_2 = \lambda, \\ v_2 & \text{ Eigenvektor von } A \text{ zu } \lambda \text{ mit } \|v_2\|_2 = \Lambda \end{aligned}$$

und

$$e^{(0)} = x^* - x^{(0)} = v_1 \pm v_2.$$

Dann erhält man mit (4.34) für die weiteren Fehler

$$\begin{aligned} e_{RF}^{(1)} &= He^{(0)} = (I - \omega_{opt}A)e^{(0)} = \left(I - \frac{2}{\Lambda + \lambda}A\right)e^{(0)} \\ &= \left(I - \frac{2}{\Lambda + \lambda}A\right)(v_1 \pm v_2) \\ &= \left(I - \frac{2}{\Lambda + \lambda}A\right)v_1 \pm \left(I - \frac{2}{\Lambda + \lambda}A\right)v_2 \\ &= \left(1 - \frac{2}{\Lambda + \lambda}\Lambda\right)v_1 \pm \left(1 - \frac{2}{\Lambda + \lambda}\lambda\right)v_2 \\ &= \frac{-\Lambda + \lambda}{\Lambda + \lambda}v_1 \pm \frac{\Lambda - \lambda}{\Lambda + \lambda}v_2 \\ &= -\eta v_1 \pm \eta v_2, \\ e_{RF}^{(2)} &= He_{RF}^{(1)} = -\eta H v_1 \pm \eta H v_2 \\ &= -\eta(-\eta v_1) \pm \eta \eta v_2 = \eta^2(v_1 \pm v_2) = \eta^2 e^{(0)}, \\ e_{RF}^{(2k)} &= \eta^{2k} e^{(0)}, \\ \|e_{RF}^{(2k)}\|_2 &= \eta^{2k} \|e^{(0)}\|_2. \end{aligned}$$

Das trifft auch auf die A -Norm zu. Wir haben sowohl in der euklidischen Norm als auch in der A -Norm die Fehlerabschätzung

$$\|e_{RF}^{(m)}\| \leq \eta^m \|e^{(0)}\| \text{ mit dem Konvergenzfaktor } \eta = \rho(H) = \frac{\Lambda - \lambda}{\Lambda + \lambda} < 1. \quad (4.42)$$

Bemerkung 4.2 Einige Bemerkungen zur Konvergenzgeschwindigkeit und Konvergenzrate.

(1) Mit der Ungleichung $\|e_{RF}^{(m)}\| \leq \|H^m\| \|e^{(0)}\|$ und der Matrixnormdefinition erhält man

$$\frac{\|e_{RF}^{(m)}\|}{\|e^{(0)}\|} \leq \|H^m\| = \sup_{x \neq 0} \frac{\|H^m x\|}{\|x\|}.$$

$\|H^m\|$ ist somit ein Maß für die Verringerung des Fehlers nach m Iterationen. Wegen $\lim_{m \rightarrow \infty} \|H^m\| = 0$ gibt es einen Wert m , so dass eine gegebene Toleranz $\varepsilon > 0$ erreicht wird, d. h.

$$\begin{aligned} \|H^m\| &\leq \varepsilon, \\ m \ln(\|H^m\|) &\leq m \ln(\varepsilon), \\ m &\geq \frac{-\ln(\varepsilon)}{-m^{-1} \ln(\|H^m\|)}. \end{aligned}$$

Die Größe im Nenner

$$R_m(H) = -\frac{1}{m} \ln(\|H^m\|) = -\ln(\sqrt[m]{\|H^m\|})$$

heißt mittlere Konvergenzrate (average rate of convergence), die näherungsweise durch

$$\tilde{R}_m(H) = -\ln(\|H\|), \quad \|H\| < 1,$$

abgeschätzt werden kann.

Auf die Schätzung für m mit $\tilde{R}_m(H)$ kommt man auch über $\|H^m\| \leq \|H\|^m \leq \varepsilon$ und der Umstellung nach m .

(2) Wenn man die allgemein gültige Beziehung

$$\eta = \rho(H) = \lim_{m \rightarrow \infty} \sqrt[m]{\|H^m\|} \quad (4.43)$$

(siehe Satz 2.52 in [18]) verwendet, kommt man im Grenzübergang auf die asymptotische Konvergenzrate, auch einfach Konvergenzrate genannt (asymptotic average rate of convergence).

$$R(H) = \lim_{m \rightarrow \infty} R_m(H) = -\ln(\rho(H)) > 0, \quad (4.44)$$

die mit fallendem Spektralradius wächst.

Die Anzahl der erforderlichen Iterationen zur Erreichung der Genauigkeit ε beträgt dann bei Rundung auf die nächst größere ganze Zahl

$$m = \left\lceil \frac{-\ln(\varepsilon)}{R(H)} + 1 \right\rceil. \quad (4.45)$$

Je kleiner ε ist, desto mehr Iterationen sind auszuführen. Je größer die Konvergenzrate ist, desto weniger Iterationen werden gebraucht.

Im Sonderfall $\eta^m = 10^{-k}$ ist $m \approx k/(-\log_{10}(\eta))$ und m wächst linear mit der geforderten Genauigkeit.

(3) In der Abschätzung (4.42) mit dem Spektralradius kann man auch die Matrixkondition einbeziehen.

Es gilt in der Spektralnorm mit (3.4) und

$$\kappa = \kappa(A) = \text{cond}_2(A) = \frac{\max \lambda(A)}{\min \lambda(A)} = \frac{\Lambda}{\lambda} \quad (4.46)$$

die Formel

$$\eta = \frac{\Lambda - \lambda}{\Lambda + \lambda} = \frac{\frac{\Lambda}{\lambda} - 1}{\frac{\Lambda}{\lambda} + 1} = \frac{\kappa - 1}{\kappa + 1} < 1. \quad (4.47)$$

(4) Mit der Beziehung (4.47) drückt man die Iterationsanzahl m in Abhängigkeit von der Kondition aus. Unter Verwendung von

$$-\ln\left(\frac{z-1}{z+1}\right) = \ln\left(\frac{z+1}{z-1}\right) = 2\left(\frac{1}{z} + \frac{1}{3z^3} + \frac{1}{5z^5} + \dots\right), \quad |z| > 1,$$

erhält man ähnlich zu (4.44) die Schätzung

$$m \approx \frac{-\ln(\varepsilon)}{-\ln(\eta)} = \frac{-\ln(\varepsilon)}{-\ln\left(\frac{\kappa-1}{\kappa+1}\right)} \approx \frac{\kappa}{2} \ln\left(\frac{1}{\varepsilon}\right), \quad (4.48)$$

die für κ nahe Eins natürlich zu grob ausfallen kann.

Aber man sieht in den Beziehungen, dass eine Verschlechterung der Kondition von A mehr Iterationen erforderlich macht.

Als eine wichtige Erkenntnis aus den Bemerkungen nehmen wir mit, dass in einem AV nicht nur die Bestimmung einer optimalen Schrittzahl α in der Suchrichtung wichtig ist, sondern auch eine Verbesserung der Kondition der Matrix des LGS anzustreben ist.

Die Verbesserung der Kondition haben wir im vorkonditionierten GV (Version 5) durch den Zugang (4.23) und somit durch eine sich ergebende Veränderung der Suchrichtung beschrieben.

Für die Vorkonditionierungsmatrix C fordert man im Allgemeinen die Bedingung der Spektraläquivalenz

$$0 < \gamma C \leq A \leq \Gamma C, \quad \text{wobei } 0 < \gamma \leq \Gamma. \quad (4.49)$$

Das führt für die Matrix $B = C^{-1}A$ auf die Konditionszahl

$$\kappa' = \kappa(B) = \frac{\Gamma}{\gamma}$$

und wegen

$$\kappa' = \frac{\Gamma}{\gamma} \leq \frac{\Lambda}{\lambda} = \kappa$$

ist dann

$$\eta' = \frac{\kappa' - 1}{\kappa' + 1} \leq \frac{\kappa - 1}{\kappa + 1} = \eta < 1.$$

Damit erhält man in der Fehlerschätzung

$$\|e^{(m)}\|_A \leq \eta'^m \|e^{(0)}\|_A \quad (4.50)$$

einen kleineren Konvergenzfaktor.

Bezüglich detaillierter Konvergenzaussagen verweisen wir auf [8].

Zusammenfassung und Bewertung der Eigenschaften des GV

in der Skala $\{+, \pm, -\}$ unter der Voraussetzung $A = A^T > 0$

- + Das Abstiegszenario und die Minimierungsaufgabe sind dem Problem angepasst.
- + Einfache Implementation des Algorithmus.
- ± Hauptaufwand in einem Iterationsschritt ist eine Matrix-Vektor-Multiplikation.
- ± Der Lösungsfehler $\|e^{(m)}\|_A = \|x^* - x^{(m)}\|_A$ wird ständig verkleinert, aber nicht unbedingt $\|r^{(m)}\|_2 = \|b - Ax^{(m)}\|_2$. Für ein Abbruchkriterium kann man natürlich die Größe $\|e^{(m)}\|_A$ wegen der darin enthaltenen unbekanntenen Lösung x^* nicht verwenden.
- ± Aufeinander folgende Residua $r^{(m+1)} \perp r^{(m)}$ sind orthogonal und damit linear unabhängig.
- Konvergenzverhalten
Das GV ist i. Allg. ein unendliches Iterationsverfahren.
Die Orthogonalität aufeinander folgender Residua $r^{(m+1)} \perp r^{(m)}$ bedeutet nicht unbedingt die lineare Unabhängigkeit von $r^{(m+1)}$ und $r^{(m-1)}$.
Wünschenswert wäre z. B., dass das System der Suchrichtungen nur aus linear unabhängigen Vektoren besteht.

(4.51)

4.2.2 Krylov-Unterraum und Suchrichtungen

Aus linear unabhängigen Vektoren $x_i \in \mathbb{R}^n$, $i = 1, 2, \dots$, kann man $k \leq n$ Vektoren auswählen, die dann eine Basis $X = \{x_1, x_2, \dots, x_k\}$ des linearen Unterraums

$$\mathcal{X} = [X] = \text{span } X = \text{span}\{x_1, x_2, \dots, x_k\} \subset \mathbb{R}^n, \quad (4.52)$$

bilden. Wir notieren hierbei verschiedene gebräuchliche Bezeichnungen des Unterraums. Seine Dimension $k \leq n$ kann durch einen zusätzlichen Index gekennzeichnet werden. Nimmt man die Vektoren x_i als Spalten einer Matrix, so erhält man

$$X = X(n, k) = (x_1, x_2, \dots, x_k) = [x_1, x_2, \dots, x_k]. \quad (4.53)$$

Definition 4.2 Krylov-Unterraum und Krylov-Matrix

Ein spezielles mittels der Matrix A und einem Vektor $x \neq 0$ generiertes Vektorsystem ist $\{x, Ax, \dots, A^m x\}$, $m \geq 0$.

Sind seine ersten k Vektoren linear unabhängig, so bezeichnet man den durch sie aufgespannten Raum als Krylov-Unterraum oder Krylov-Teilraum

$$\mathcal{K}_k = \mathcal{K}_k(A, x) = \text{span}\{x, Ax, \dots, A^{k-1}x\}. \quad (4.54)$$

Damit ergibt sich die so genannte Krylov-Matrix

$$K^{(k)} = (x, Ax, \dots, A^{k-1}x). \quad (4.55)$$

Ist x ein Eigenvektor der Matrix A , d. h. $Ax = \lambda x$, $x \neq 0$, so erhält man wegen $A^m x = \lambda^m x$, $m = 1, 2, 3, \dots$, die einfache Situation mit den Krylov-Unterräumen $\mathcal{K}_1 = \mathcal{K}_2 = \mathcal{K}_3 = \dots = \text{span}\{x\}$.

Für eine singuläre Matrix macht die Konstruktion von \mathcal{K}_k wenig Sinn.

Beispiel 4.1 Krylov-Unterräume für eine singuläre Matrix A

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

(1) $x = (1, 0, 0)^T$ Eigenvektor (EV) zu $\lambda = 0$

$$\mathcal{K}_1 = \text{span}\{x\}, \quad Ax = 0 = 0 \cdot x$$

(2) $x = (0, 1, 0)^T$ kein EV

$$\mathcal{K}_2 = \text{span}\{x, Ax\}, \quad Ax = (1, 0, 0)^T, \quad A^2x = 0$$

(3) $x = (0, 0, 1)^T$ kein EV

$$\mathcal{K}_3 = \text{span}\{x, Ax, A^2x\}, \quad Ax = (1, 1, 0)^T, \quad A^2x = (1, 0, 0)^T, \quad A^3x = 0$$

Für reguläre Matrizen ist zu prüfen, wie groß maximal die Dimension von \mathcal{K}_k werden kann.

Beispiel 4.2 Krylov-Unterräume für eine reguläre Matrix A

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

(1) $x = (1, 0, 0)^T$ EV

$$\mathcal{K}_1 = \text{span}\{x\}, \quad Ax = x = 1 \cdot x$$

(2) $x = (1, 1, 0)^T$ kein EV

$$\mathcal{K}_2 = \text{span}\{x, Ax\}, \quad Ax = (2, 1, 0)^T, \quad A^m x = (c_1, c_2, 0)^T, \quad m \geq 2, \quad \mathcal{K}_3 = \mathcal{K}_2$$

(3) $x = (0, 0, 1)^T$ kein EV

$$\mathcal{K}_3 = \text{span}\{x, Ax, A^2 x\}, \quad Ax = (1, 1, 1)^T, \quad A^2 x = (3, 2, 1)^T$$

Die Kontrolle, ob bei Hinzunahme des nächsten Vektors $A^m x$ die Dimension des Vektorraums um Eins wächst, erfolgt in Anwendungen durch die Transformation des Vektorsystems auf eine orthogonale Basis $\{q_1, q_2, \dots, q_m, q_{m+1}\}$, $q_i^T q_j = \delta_{ij}$, unter Verwendung z. B. des Orthogonalisierungsverfahrens von Gram-Schmidt. Kann der Vektor q_{m+1} nicht erzeugt werden, ist der maximale Krylov-Unterraum $\mathcal{K}_m(A, x)$.

Die Such- und Abstiegsrichtungen $r^{(m)}$ im GV (4.13) sind orthogonal, was zwei aufeinander folgende Richtungen betrifft. Das genügt jedoch nicht für ihre lineare Unabhängigkeit insgesamt. Es gilt jedoch der folgende Satz.

Satz 4.7 Solange im GV die nicht verschwindenden Suchrichtungen $r^{(m)}$ linear unabhängig sind, gilt

$$\text{span}\{r^{(0)}, r^{(1)}, \dots, r^{(k)}\} = \text{span}\{r^{(0)}, Ar^{(0)}, \dots, A^{k-1}r^{(0)}\} = \mathcal{K}_k(A, r^{(0)}). \quad (4.56)$$

Beweis.

Wir zeigen zunächst die Behauptung für $k = 1, 2$ unter Verwendung der rekursiven Berechnungsformeln der Suchrichtungen (4.17).

(1) $k = 1$

Wir haben $r^{(1)} \perp r^{(0)}$ und $\dim(\text{span}\{r^{(0)}, r^{(1)}\}) = 2$.

Aus $r^{(1)} = r^{(0)} - \alpha_0 Ar^{(0)}$ folgen unmittelbar die beiden Beziehungen mit reellen Koeffizienten

$$\begin{aligned} r^{(1)} &= \delta_0 r^{(0)} + \delta_1 Ar^{(0)} \in \text{span}\{r^{(0)}, Ar^{(0)}\}, \\ Ar^{(0)} &= \gamma_0 r^{(0)} + \gamma_1 r^{(1)} \in \text{span}\{r^{(0)}, r^{(1)}\}. \end{aligned}$$

(2) $k = 2$

Wir haben $r^{(2)} \perp r^{(1)} \perp r^{(0)}$, $r^{(2)} \not\parallel r^{(0)}$ und $\dim(\text{span}\{r^{(0)}, r^{(1)}, r^{(2)}\}) = 3$.

Aus $r^{(2)} = r^{(1)} - \alpha_1 A r^{(1)}$ und Teil (1) folgen die Beziehungen

$$\begin{aligned}
 r^{(2)} &= r^{(1)} - \alpha_1 A(r^{(0)} - \alpha_0 A r^{(0)}) \\
 &= r^{(1)} - \alpha_1 A r^{(0)} + \alpha_0 \alpha_1 A^2 r^{(0)} \\
 &= r^{(0)} - \alpha_0 A r^{(0)} - \alpha_1 A r^{(0)} + \alpha_0 \alpha_1 A^2 r^{(0)} \\
 &= r^{(0)} - (\alpha_0 + \alpha_1) A r^{(0)} + \alpha_0 \alpha_1 A^2 r^{(0)} \\
 &= \delta_0 r^{(0)} + \delta_1 A r^{(0)} + \delta_2 A^2 r^{(0)} \in \text{span}\{r^{(0)}, A r^{(0)}, A^2 r^{(0)}\}, \\
 A^2 r^{(0)} &= \frac{1}{\alpha_0 \alpha_1} (\alpha_1 A r^{(0)} - r^{(1)} + r^{(2)}) \\
 &= \frac{1}{\alpha_0 \alpha_1} \left(\alpha_1 \frac{1}{\alpha_0} (r^{(0)} - r^{(1)}) - r^{(1)} + r^{(2)} \right) \\
 &= \frac{1}{\alpha_0^2} r^{(0)} - \left(\frac{1}{\alpha_0^2} + \frac{1}{\alpha_0 \alpha_1} \right) r^{(1)} + \frac{1}{\alpha_0 \alpha_1} r^{(2)} \\
 &= \gamma_0 r^{(0)} + \gamma_1 r^{(1)} + \gamma_2 r^{(2)} \in \text{span}\{r^{(0)}, r^{(1)}, r^{(2)}\}.
 \end{aligned}$$

Diese Vorgehensweise ist auf beliebige k übertragbar. □

In der praktischen Durchführung des GV kann es sein, dass

$$r^{(m+1)} \in \text{span}\{r^{(0)}, r^{(1)}, \dots, r^{(m)}\},$$

das bedeutet auch $r^{(m+1)} \notin \text{span}\{r^{(0)}, r^{(1)}, \dots, r^{(m)}\}$, obwohl $r^{(m+1)} \perp r^{(m)}$.

Die Dimension des Unterraums der Suchrichtungen muss also nicht ständig wachsen, was natürlich konform geht mit der Unendlichkeit des GV.

Wenn $\dim(\text{span}\{r^{(0)}, r^{(1)}, \dots, r^{(n-1)}\}) = n$ ist, müssen die n -te Suchrichtung $r^{(n)}$ und weitere im Vektorraum $\text{span}\{r^{(0)}, r^{(1)}, \dots, r^{(n-1)}\} = \mathbb{R}^n$ liegen.

Aus $x^{(m+1)} = x^{(m)} + \alpha_m r^{(m)}$, $m = 0, 1, \dots$, ergibt sich

$$x^{(m+1)} = x^{(0)} + \alpha_0 r^{(0)} + \alpha_1 r^{(1)} + \dots + \alpha_m r^{(m)}$$

und somit

$$x^{(m+1)} \in x^{(0)} + \mathcal{R}_{m+1}, \quad \mathcal{R}_{m+1} = \text{span}\{r^{(0)}, r^{(1)}, \dots, r^{(m)}\}. \quad (4.57)$$

Die m -te Iterierte $x^{(m)}$ im GV wird als dasjenige Element aus dem Unterraum $x^{(0)} + \text{span}\{r^{(0)}, r^{(1)}, \dots, r^{(m-1)}\}$ bestimmt, für welches das Funktional $Q(x)$ seinen minimalen Wert annimmt.

4.2.3 Konvergenz des GV mittels Kantorovich-Ungleichung

Da das Funktional $Q(x)$ in einem AV eingebettet ist, also seine Funktionswerte an den Iterierten kontinuierlich kleiner werden, wird auch die A -Norm $\|e(x)\|_A$ systematisch kleiner (siehe (4.14), (4.18)).

Praktische Abschätzungen mit dem Fehler $e(x)$ sind wegen der Kenntnis der exakten Lösung x^* nur bedingt möglich. Wünschenswert wären natürlich auch Aussagen zum Verhalten von $r(x)$ bzw. in anderen Normen, wie der euklidischen Norm.

Deshalb versuchen wir hier, eventuell Bedingungen mit den Normen

$$\|e(x)\|_2, \|r(x)\|_2, \|r(x)\|_A, \|r(x)\|_{A^{-1}}$$

aufzustellen.

Aus dem Vergleich des GV mit der Richardson-Iteration und wegen (4.4), (4.18)

$$\|e^{(m)}\|_A^2 = 2Q(x^{(m)}) + x^{*T}b = 2[Q(x^{(m)}) - Q(x^*)]$$

haben wir bisher folgende Abschätzungen erhalten.

$$\text{GV:} \quad \|e^{(m+1)}\|_A \leq \eta \|e^{(m)}\|_A, \quad \eta = \frac{\kappa-1}{\kappa+1} \leq 1, \quad \kappa = \frac{\lambda}{\lambda'}$$

$$\|e^{(m+1)}\|_A^2 \leq \eta^2 \|e^{(m)}\|_A^2,$$

$$\|e^{(m)}\|_A \leq \eta^m \|e^{(0)}\|_A,$$

$$Q(x^{(m+1)}) - Q(x^*) \leq \eta^2 [Q(x^{(m)}) - Q(x^*)],$$

$$Q(x^{(m+1)}) \leq \eta^2 Q(x^{(m)}), \quad \text{brauchbar nur bei } Q(x^*) = 0 \leq Q(x^{(m)}),$$

$$Q(x^{(m)}) \leq \eta^{2m} Q(x^{(0)}),$$

$$\text{RF:} \quad \|e_{RF}^{(m+1)}\|_A \leq \eta \|e_{RF}^{(m)}\|_A, \quad \text{Abschätzungen sind scharf,}$$

$$\|e_{RF}^{(m+1)}\|_2 \leq \eta \|e_{RF}^{(m)}\|_2.$$

Damit sieht man die lineare Konvergenz und die Ungleichungsketten

$$Q(x^{(0)}) > Q(x^{(1)}) > Q(x^{(2)}) > \dots \rightarrow Q(x^*) \leq 0,$$

$$\|e^{(0)}\|_A > \|e^{(1)}\|_A > \|e^{(2)}\|_A > \dots \rightarrow 0.$$

Zunächst eine Abschätzung für die energetische Norm $\|r(x)\|_{A^{-1}}$.

Es gilt

$$\begin{aligned}
\|r(x)\|_{A^{-1}}^2 &= r(x)^T A^{-1} r(x) \\
&= (b - Ax)^T A^{-1} (b - Ax) \\
&= (b^T - x^T A^T)(A^{-1}b - x) \\
&= (b^T - x^T A)(x^* - x) \\
&= b^T x^* - b^T x - x^T A x^* + x^T A x, \quad Ax^* = b, \\
&= b^T x^* - b^T x - x^T b + x^T A x \\
&= x^{*T} b - 2x^T b + x^T A x \\
&= 2[\frac{1}{2}x^T A x - x^T b - (-\frac{1}{2}x^{*T} b)] \\
&= 2[Q(x) - Q(x^*)] \quad \text{wegen (4.1), (4.4)} \\
&= \|e(x)\|_A^2.
\end{aligned}$$

Somit folgt

$$\|r^{(m)}\|_{A^{-1}} = \|e^{(m)}\|_A \quad (4.58)$$

und insbesondere

$$\frac{1}{2}r^{(m)T} A^{-1} r^{(m)} = Q(x^{(m)}) - Q(x^*). \quad (4.59)$$

Nun soll die Ungleichung von Kantorovich (2.20) verwendet werden, um die Differenz $Q(x^{(m+1)}) - Q(x^*)$ abzuschätzen.

Gemäß GV (4.13), (4.12) erhält man

$$\begin{aligned}
Q(x^{(m+1)}) &= Q(x^{(m)} + \alpha_m r^{(m)}) \\
&= Q(x^{(m)} + \frac{r^{(m)T} r^{(m)}}{r^{(m)T} A r^{(m)}} r^{(m)}), \\
Q(x + \alpha r) &= Q(x) - \alpha r^T r + \frac{1}{2} \alpha^2 r^T A r, \quad \text{siehe (4.10)} \\
&= Q(x^{(m)}) - \frac{r^{(m)T} r^{(m)}}{r^{(m)T} A r^{(m)}} r^{(m)T} r^{(m)} + \frac{1}{2} \left(\frac{r^{(m)T} r^{(m)}}{r^{(m)T} A r^{(m)}} \right)^2 r^{(m)T} A r^{(m)} \\
&= Q(x^{(m)}) - \frac{(r^{(m)T} r^{(m)})^2}{r^{(m)T} A r^{(m)}} + \frac{1}{2} \frac{(r^{(m)T} r^{(m)})^2}{r^{(m)T} A r^{(m)}} \\
&= Q(x^{(m)}) - \frac{1}{2} \frac{(r^{(m)T} r^{(m)})^2}{r^{(m)T} A r^{(m)}}
\end{aligned}$$

und daraus

$$Q(x^{(m+1)}) - Q(x^*) = Q(x^{(m)}) - Q(x^*) - \frac{1}{2} \frac{(r^{(m)T} r^{(m)})^2}{r^{(m)T} A r^{(m)}}. \quad (4.60)$$

Aus (4.59) folgt

$$1 = \frac{Q(x^{(m)}) - Q(x^*)}{\frac{1}{2}r^{(m)T}A^{-1}r^{(m)}},$$

so dass damit die Beziehung (4.60) weiter umgeformt werden kann.

$$\begin{aligned} Q(x^{(m+1)}) - Q(x^*) &= Q(x^{(m)}) - Q(x^*) - \frac{1}{2} \frac{(r^{(m)T}r^{(m)})^2}{r^{(m)T}Ar^{(m)}} \cdot 1 \\ &= Q(x^{(m)}) - Q(x^*) - \frac{1}{2} \frac{(r^{(m)T}r^{(m)})^2}{r^{(m)T}Ar^{(m)}} \frac{Q(x^{(m)}) - Q(x^*)}{\frac{1}{2}r^{(m)T}A^{-1}r^{(m)}} \\ &= \left(1 - \frac{(r^{(m)T}r^{(m)})^2}{(r^{(m)T}Ar^{(m)})(r^{(m)T}A^{-1}r^{(m)})}\right) [Q(x^{(m)}) - Q(x^*)] \end{aligned}$$

Jetzt nutzen wir die Ungleichung von Kantorovich (2.20) und (4.30), (4.34), um darin den Bruch zu ersetzen und erhalten

$$\begin{aligned} Q(x^{(m+1)}) - Q(x^*) &\leq \left(1 - \frac{4\lambda\Lambda}{(\lambda + \Lambda)^2}\right) [Q(x^{(m)}) - Q(x^*)] = \\ &= \left(\frac{\Lambda - \lambda}{\Lambda + \lambda}\right)^2 [Q(x^{(m)}) - Q(x^*)] = \\ &= \left(\frac{\kappa - 1}{\kappa + 1}\right)^2 [Q(x^{(m)}) - Q(x^*)] = \\ &= \eta^2 [Q(x^{(m)}) - Q(x^*)]. \end{aligned}$$

Damit bestätigen wir die Fehlerabschätzungen (4.36), (4.37). Die Verwendung der Normäquivalenz (2.13) in der Form

$$\lambda\|x\|_2^2 \leq \|x\|_A^2 \leq \Lambda\|x\|_2^2$$

gestattet nur eine grobe Einschachtelung des Fehlers gemäß

$$\begin{aligned} \lambda\|e^{(m+1)}\|_2^2 &\leq \|e^{(m+1)}\|_A^2 \leq \eta^2\|e^{(m)}\|_A^2 \leq \eta^2\Lambda\|e^{(m)}\|_2^2, \\ \|e^{(m+1)}\|_2^2 &\leq \kappa\eta^2\|e^{(m)}\|_2^2 \leq \kappa\eta^2\kappa\eta^2\|e^{(m-1)}\|_2^2 \leq \dots, \\ \|e^{(m)}\|_2^2 &\leq \kappa^m\eta^{2m}\|e^{(0)}\|_2^2, \\ \|e^{(m)}\|_2 &\leq \sqrt{\kappa^m}\eta^m\|e^{(0)}\|_2, \quad \eta = \frac{\kappa-1}{\kappa+1}. \end{aligned}$$

Wegen $\kappa = \frac{\Lambda}{\lambda} > 1$ und $\kappa^m \gg 1$ hat diese Abschätzung wenig Nutzen. Geschickter ist die Vorgehensweise in [31]. Mit der Ungleichungskette

$$\|e^{(m+1)}\|_A^2 \leq \eta^2\|e^{(m)}\|_A^2 \leq \eta^4\|e^{(m-1)}\|_A^2 \leq \dots$$

und der Normäquivalenz folgt nun

$$\begin{aligned}\lambda \|e^{(m)}\|_2^2 &\leq \|e^{(m)}\|_A^2 \leq \eta^{2m} \|e^{(0)}\|_A^2 \leq \eta^{2m} \Lambda \|e^{(0)}\|_2^2, \\ \|e^{(m)}\|_2^2 &\leq \kappa \eta^{2m} \|e^{(0)}\|_2^2, \\ \|e^{(m)}\|_2 &\leq \sqrt{\kappa} \eta^m \|e^{(0)}\|_2.\end{aligned}\tag{4.61}$$

Der Faktor $\sqrt{\kappa}$ ist zwar größer als Eins, aber für hinreichend große Iterationszahlen m strebt der zweite Faktor η^m gegen Null und die Konvergenz in der euklidischen Norm ist gesichert. Mehr noch: In [11] wurde gezeigt, dass bei $A = A^T > 0$ auch die Normwerte $\|e^{(m)}\|_2$ kontinuierlich kleiner werden.

Es gilt natürlich $\lim_{m \rightarrow \infty} \|r^{(m)}\|_{2,A} = 0$ wegen (4.15), aber die monotone Konvergenz der Größen $\|r^{(m)}\|_2$ und $\|r^{(m)}\|_A$ wird sich bei einfachen LGS zwar einstellen, muss im Allgemeinen jedoch nicht erfüllt sein.

4.3 Beispiele zum Gradientenverfahren

4.3.1 Beispiel 1

Zunächst ein sehr gutartiges kleines Beispiel. Sei $A = A^T > 0$.

$$\begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad x^* = \begin{pmatrix} 1/5 \\ 3/5 \end{pmatrix}, \quad x^{(0)} = \begin{pmatrix} 3/2 \\ 1 \end{pmatrix}.$$

Die Funktionale sind

$$\begin{aligned}Q(x) &= x_1^2 + x_1 x_2 + \frac{3}{2} x_2^2 - x_1 - 2x_2, \\ R(x) &= \frac{5}{2} x_1^2 + 5x_1 x_2 + 5x_2^2 - 4x_1 - 7x_2.\end{aligned}$$

Am gemeinsamen eindeutigen Minimum an der Stelle x^* gilt $Q(x^*) = -\frac{7}{10}$ und $R(x^*) = -\frac{5}{2}$.

Als Suchrichtung und Abstiegsrichtung in einem Schritt nehmen wir die Richtung des steilsten Abstiegs $r(x) = -\nabla Q(x)$, die orthogonal zu den Höhenlinien $Q(x) = \text{const}$ sind.

Die Eigenwerte von A sind

$$\lambda_{1,2} = \frac{5 \mp \sqrt{5}}{2} = 1.381\,966\,011\dots, \quad 3.618\,033\,988\dots,$$

woraus die Kondition

$$\kappa = \frac{\lambda_2}{\lambda_1} = \frac{5 + \sqrt{5}}{5 - \sqrt{5}} = 2.618\,033\dots$$

und der Konvergenzfaktor

$$\eta = \frac{\kappa - 1}{\kappa + 1} = 0.447\,213\dots < 1$$

folgen.

So wie die Folge der Funktionalwerte $Q(x^{(m)})$ streng monoton fallend gegen $-\frac{7}{10}$ strebt, so tun dies auch die Fehler $e^{(m)} = e(x^{(m)}) = x^* - x^{(m)}$ in der A -Norm, aber natürlich gegen Null.

Man prüft leicht nach, dass die Fehlerschätzung

$$\|e^{(m)}\|_A \leq \eta^m \|e^{(0)}\|_A$$

erfüllt ist. So haben wir für $m = 10$ die Abschätzung

$$\|e^{(10)}\|_A = 0.000\,008\,648\dots \leq 0.000\,708\,350\dots = \eta^{10} \|e^{(0)}\|_A.$$

Die ersten Schritte des GV sind wie folgt.

$$\begin{aligned} x^{(0)} &= \left(\frac{3}{2}, 1\right)^T && \text{Startvektor,} \\ r^{(0)} &= b - Ax^{(0)} = \frac{1}{2}(-6, -5)^T && \text{Anfangsresiduum, Abstiegsrichtung,} \\ Ar^{(0)} &= -\frac{1}{2}(17, 21)^T. \end{aligned}$$

$$\begin{aligned} \mathbf{S1} \quad \alpha_0 &= \frac{\|r^{(0)}\|_2^2}{\|r^{(0)}\|_A^2} = \frac{61}{207} = 0.294\,685\dots \\ x^{(1)} &= x^{(0)} + \alpha_0 r^{(0)} = \frac{1}{414}(255, 109)^T, \\ r^{(1)} &= b - Ax^{(1)} = r^{(0)} - \alpha_0 Ar^{(0)} = \frac{41}{414}(-5, 6)^T. \end{aligned}$$

$$\begin{aligned} \mathbf{S2} \quad \alpha_1 &= \frac{61}{98} = 0.622\,448\dots, \\ x^{(2)} &= \frac{1}{40\,572}(12\,485, 25\,688)^T, \quad r^{(2)} = \frac{1\,681}{40\,572}(-6, -5)^T. \end{aligned}$$

$$\begin{aligned} \mathbf{S3} \quad \alpha_2 &= \frac{61}{207}, \\ x^{(3)} &= \left(\frac{93\,769}{399\,924}, \frac{4\,804\,711}{8\,398\,404}\right)^T, \quad r^{(3)} = \frac{68\,921}{8\,398\,404}(-5, 6)^T. \end{aligned}$$

$$\begin{aligned} \mathbf{S4} \quad \alpha_3 &= \frac{61}{98}, \\ x^{(4)} &= \left(\frac{171\,955\,697}{823\,043\,592}, \frac{124\,021\,691}{205\,760\,898}\right)^T, \quad r^{(4)} = \frac{2\,825\,761}{823\,043\,592}(-6, -5)^T. \end{aligned}$$

Dazu erhalten wir die Beziehungen

$$r^{(0)} \perp r^{(1)} \perp r^{(2)} \perp r^{(3)} \perp \dots, \quad r^{(0)} \parallel r^{(2)} \parallel r^{(4)} \parallel \dots, \quad r^{(1)} \parallel r^{(3)} \parallel r^{(5)} \parallel \dots,$$

$$\mathcal{K}_2(A, r^{(0)}) = \text{span}\{r^{(0)}, r^{(1)}\} = \text{span}\{r^{(0)}, Ar^{(0)}\} = \text{span}\left\{\begin{pmatrix} -6 \\ -5 \end{pmatrix}, \begin{pmatrix} -5 \\ 6 \end{pmatrix}\right\} \subset \mathbb{R}^2,$$

$$x^* = \frac{-42}{305}r^{(0)} + \frac{5382}{12505}r^{(1)} \in \text{span}\{r^{(0)}, r^{(1)}\}.$$

Das GV konvergiert nur langsam.

Iterationsverlauf mit verschiedenen Fehlern

m	[x(m)[1], x(m)[2]]	Q(x(m))	[r(m)[1], r(m)[2]]	x(m) _2^2
0	[1.5000000000, 1.0000000000]	1.750000000000	[-3.0000000000, -2.5000000000]	15.250000000000
1	[0.6159420290, 0.2632850242]	-0.496980676329	[-0.4951690821, 0.5942028986]	0.598269504539
2	[0.3077245391, 0.6331460120]	-0.683176797639	[-0.2485950902, -0.2071625752]	0.104715851430
3	[0.2344670487, 0.5720981034]	-0.698605944830	[-0.0410322009, 0.0492386411]	0.004108085282
4	[0.2089265972, 0.6027466453]	-0.699884481576	[-0.0205998396, -0.0171665330]	0.000719043249
5	[0.2028561130, 0.5976879085]	-0.699990427562	[-0.0034001346, 0.0040801615]	0.000028208633
6	[0.2007397027, 0.6002276008]	-0.699999206780	[-0.0017070063, -0.0014225053]	0.000004937392
7	[0.2002366719, 0.5998084085]	-0.699999934270	[-0.0002817523, 0.0003381027]	0.000000193698
8	[0.2000612955, 0.6000188602]	-0.699999994553	[-0.0001414511, -0.0001178759]	0.000000033903
9	[0.2000196118, 0.5999841238]	-0.699999999549	[-0.0000233474, 0.0000280169]	0.000000001330
10	[0.2000050793, 0.6000015628]	-0.699999999963	[-0.0000117214, -0.0000097678]	0.000000000233

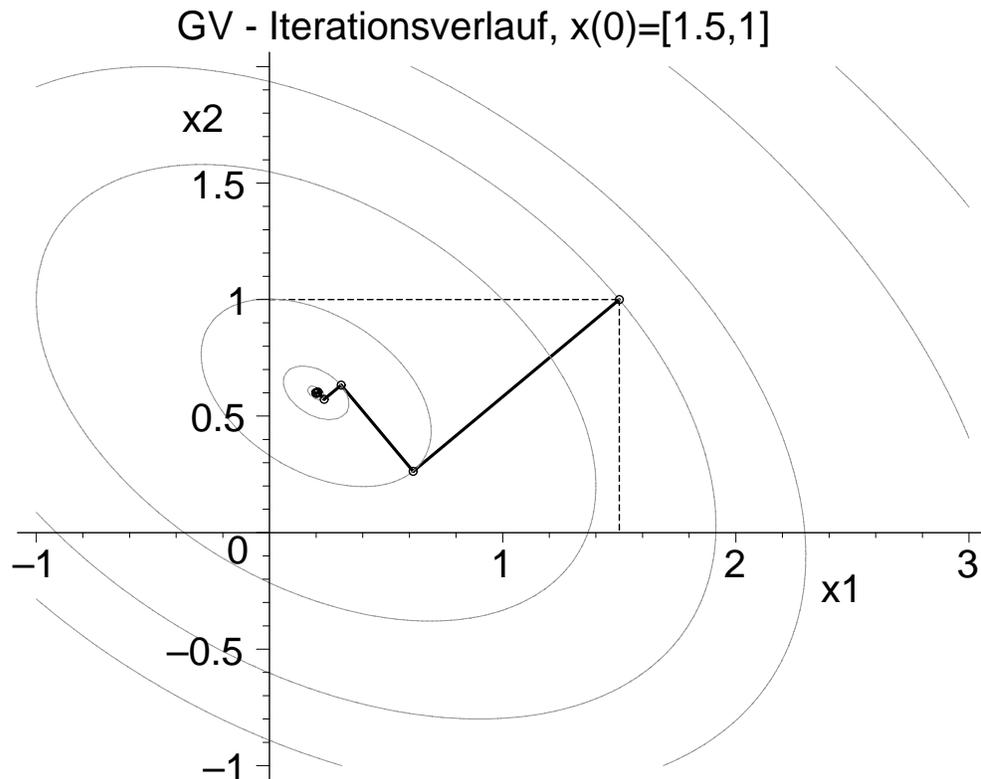


Abb. 4.4 Datei gv_001.ps

Höhenlinienbild mit Iterationsverlauf des GV zu $Q(x) = x_1^2 + x_1x_2 + \frac{3}{2}x_2^2 - x_1 - 2x_2$, contours=[1.75,0.5,-0.497,-0.6832,-0.6986,-0.6999,-0.7] und contours=3

Fehlerverhalten im Iterationsverlauf

m	Q(x(m))	xs-x(m) _A^2	xs-x(m) _2^2	r(m) _2^2	r(m) _A^2
0	1.750000000000	4.900000000000	1.850000000000	15.250000000000	51.750000000000
1	-0.496980676328	0.406038647343	0.286384746435	0.598269504539	0.961154285981
2	-0.683176797639	0.033646404722	0.012703234436	0.104715851430	0.355347233540
3	-0.698605944830	0.002788110339	0.001966493283	0.004108085282	0.006599874716
4	-0.699884481576	0.000231036847	0.000087228197	0.000719043249	0.002440032007
5	-0.699990427562	0.000019144875	0.000013503149	0.000028208633	0.000045318787
6	-0.699999206780	0.000001586441	0.000000598962	0.000004937392	0.000016754756
7	-0.699999934270	0.000000131460	0.000000092721	0.000000193698	0.000000311187
8	-0.699999994553	0.000000010893	0.000000004113	0.000000033903	0.000000115048
9	-0.699999999549	0.000000000903	0.000000000637	0.000000001330	0.000000002137
10	-0.699999999963	0.000000000075	0.000000000028	0.000000000233	0.000000000790

Wir erhalten hier für alle Fehler einen monoton abnehmenden Verlauf.

Das GV für $A^T A x = A^T b$ ist mit dem Funktional $R(x)$ verknüpft.

Es bringt eine Verschlechterung der Kondition des Systems durch Multiplikation mit A^T , aber für beliebiges reguläres A ist die Matrix $B = A^T A$ spd.

Im Allgemeinen ergeben sich dadurch mehr Iterationsschritte, aber nicht generell.

Wir betrachten also das LGS

$$Bx = \begin{pmatrix} 5 & 5 \\ 5 & 10 \end{pmatrix} x = \begin{pmatrix} 4 \\ 7 \end{pmatrix} = c, \quad B = B^T > 0.$$

Die Richtung $\hat{r}(x) = A^T r(x) = A^T (b - Ax)$ ist orthogonal zu $R(x) = \text{const.}$

Die Eigenwerte von B sind

$$\mu_{1,2} = \frac{15 \mp 5\sqrt{5}}{2} = \lambda_{1,2}^2 = 1.909\,830\,056\dots, \quad 13.090\,169\,943\dots,$$

woraus die Kondition

$$\kappa' = \frac{\mu_2}{\mu_1} = \frac{15 + 5\sqrt{5}}{15 - 5\sqrt{5}} = 6.854\,101\dots$$

und der Konvergenzfaktor

$$\eta' = \frac{\kappa' - 1}{\kappa' + 1} = 0.745\,355\dots < 1$$

folgen.

So wie die Folge der Funktionalwerte $R(x^{(m)})$ streng monoton fallend gegen $-\frac{5}{2}$ strebt, so tun dies auch die Residua $r^{(m)} = r(x^{(m)}) = b - Ax^{(m)}$ in der euklidischen Norm sowie wegen (3.10) auch $\|e^{(m)}\|_{A^T A}$, aber gegen Null.

Man prüft auch leicht nach, dass die Fehlerschätzung

$$\|e^{(m)}\|_{A^T A} \leq \eta'^m \|e^{(0)}\|_{A^T A}$$

erfüllt ist. Auch die Norm des Residuums $\hat{r}(x) = A^T r(x) = A^T (b - Ax)$ verkleinert sich stetig.

Iterationsverlauf mit verschiedenen Fehlern

m	[x(m)[1], x(m)[2]]	R(x(m))	$\ r(m)\ _2 = \ xs-x(m)\ _B$	$\ xs-x(m)\ _2$	$\ rd(m)\ _2$
0	[1.5000000000, 1.0000000000]	5.12500000000	15.25000000000	1.85000000000	182.50000000000
1	[0.8416445623, 0.1867374005]	-1.942639257294	1.114721485411	0.582493720493	2.158479268833
2	[0.2950254381, 0.6292385963]	-2.459258885573	0.081482228855	0.009884729402	0.975115197772
3	[0.2469019659, 0.5697919542]	-2.497021967502	0.005956064996	0.003112320435	0.011532964050
4	[0.2069460261, 0.6021372388]	-2.499782316275	0.000435367450	0.000052815068	0.005210135063
5	[0.2034283691, 0.5977918979]	-2.499984088084	0.000031823833	0.000016629430	0.000061621745
6	[0.2005077301, 0.6001562247]	-2.499998836895	0.000002326211	0.000000282196	0.000027838257
7	[0.2002506018, 0.5998385955]	-2.499999914981	0.000000170038	0.000000088853	0.000000329251
8	[0.2000371133, 0.6000114195]	-2.499999993785	0.000000012429	0.000000001508	0.000000148743
9	[0.2000183181, 0.5999882019]	-2.499999999546	0.000000000909	0.000000000475	0.000000001759
10	[0.2000027129, 0.6000008347]	-2.499999999967	0.000000000066	0.000000000008	0.000000000795

Im Iterationsverlauf und Konvergenzverhalten des GV erkennt man hier bezüglich des LGS und des Normalgleichungssystems keine wesentlichen Unterschiede. Das liegt mit an der kleinen Dimension des Problems und der Ausgeglichenheit der Elemente der Matrix $B = A^T A$.

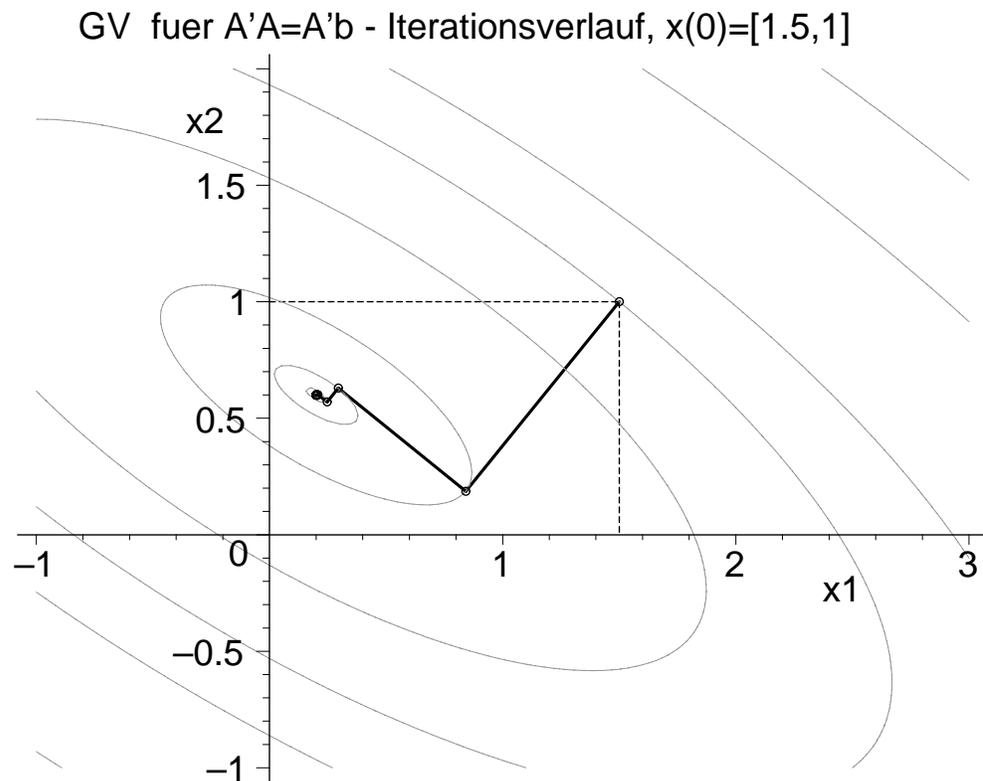


Abb. 4.5 Datei *gv_0011.ps*

Höhenlinienbild mit Iterationsverlauf des GV zu $R(x) = \frac{5}{2}x_1^2 + 5x_1x_2 + 5x_2^2 - 4x_1 - 7x_2$, $\text{contours}=[5.125, 1, -1.9426, -2.4593, -2.4970, -2.4998, -2.5]$ und $\text{contours}=3$

4.3.2 Beispiel 2

Wir betrachten nun ein LGS mit einer spd Tridiagonalmatrix, die eine etwas schlechtere Kondition hat.

Beispiel 2, $A=A'>0$, tridiagonal (1D-Laplace-Matrix)

```
n:=10:
A:=matrix(n,n,
  [[ 2, -1, 0, 0, 0, 0, 0, 0, 0, 0],
   [-1, 2, -1, 0, 0, 0, 0, 0, 0, 0],
   [ 0, -1, 2, -1, 0, 0, 0, 0, 0, 0],
   [ 0, 0, -1, 2, -1, 0, 0, 0, 0, 0],
   [ 0, 0, 0, -1, 2, -1, 0, 0, 0, 0],
   [ 0, 0, 0, 0, -1, 2, -1, 0, 0, 0],
   [ 0, 0, 0, 0, 0, -1, 2, -1, 0, 0],
   [ 0, 0, 0, 0, 0, 0, -1, 2, -1, 0],
   [ 0, 0, 0, 0, 0, 0, 0, -1, 2, -1],
   [ 0, 0, 0, 0, 0, 0, 0, 0, -1, 2]]):
b:=vector(n,[1,0,0,0,0,0,0,0,0,1]):
xs:=vector(n,[1,1,1,1,1,1,1,1,1,1]):
```

Bzw. einfache, aber etwas allgemeinere Definition

Tridiagonalmatrix mit Shift c

```
c := 0:      # Shift = Verstaerkung der Diagonalen
n := 10:
A := band([-1,2,-1],n):
A := evalm(A+c*diag(1$n)):
```

Loesung

```
xs := evalm(vector(n,[1$n])):
```

rechte Seite

```
b := evalm(vector(n,[1+c,c$(n-2),1+c])): # b := evalm(A&*xs):
```

EW $\lambda_i(A) = 4 \sin^2\left(\frac{\pi i}{2(n+1)}\right)$, $i = 1, 2, \dots, n$,

$\sigma(A) = \{0.081014, 0.317492, 0.690278, 1.169169, 1.715370,$
 $2.284629, 2.830830, 3.309721, 3.682507, 3.918985\}$,

Kondition mittels Spektralnorm $\kappa(A) = 48.374150$, Konvergenzfaktor $\eta = 0.959492$.

Für das spezielle LGS berechnen wir mit dem GV auch die absoluten Fehler

$$\|e^{(k)}\|_A, \|e^{(k)}\|_2, \|r^{(k)}\|_A, \|r^{(k)}\|_2, \quad k = 0, 1, 2, \dots,$$

die man als zusätzliche Ergebnisvektoren `eva[1..maxiter+1]`, `ev2[1..maxiter+1]`
bzw. `rva[1..maxiter+1]`, `rv2[1..maxiter+1]` in die Prozedur `gv` einbauen kann.

Die Werte $\|e^{(k)}\|_A$ und $\|e^{(k)}\|_2$ bilden monoton abnehmende Folgen, was für $\|r^{(k)}\|_A$ und $\|r^{(k)}\|_2$ nicht zutrifft, obwohl letztere schneller gegen Null tendieren.

Dann folgen grafische Vergleiche der relativen Fehler des GV, die aus Gründen der besseren Übersicht als Funktionen $\log_{10}()$ in Abhängigkeit von der Iterationszahl k dargestellt werden, also

$$f(k) = \log_{10} \left(\frac{\|e^{(k)}\|_A}{\|e^{(0)}\|_A} \right), \log_{10} \left(\frac{\|e^{(k)}\|_2}{\|e^{(0)}\|_2} \right), \log_{10} \left(\frac{\|r^{(k)}\|_A}{\|r^{(0)}\|_A} \right), \log_{10} \left(\frac{\|r^{(k)}\|_2}{\|r^{(0)}\|_2} \right).$$

Die numerische Rechnung mit `Digits:=16` erfolgt in Maple, die bei der Dimension $n = 10$ des Problems ohne Weiteres machbar ist, auch wenn das GV für eine hohe Genauigkeit viele Iterationen ausführt.

Der Startvektor ist $x^{(0)} = (1, 0, \dots, 0)^T$.

GV: Iterationsverlauf mit verschiedenen Fehlern

m	Q(x(m))= (e(m) ^2_A-x's'b)/2	e(m) _A= xs-x(m) _A	e(m) _2= xs-x(m) _2	r(m) _A= b-Ax(m) _A	r(m) _2= b-Ax(m) _2
0	0.000000000000	1.414213562373	3.000000000000	2.828427	1.732051
1	-0.562500000000	0.935414346693	2.814582562299	0.829156	0.612372
2	-0.664772727273	0.818812887939	2.660620189615	0.717980	0.474015
3	-0.713740756306	0.756649514232	2.556779804844	0.448842	0.336895
4	-0.745712276701	0.713144758516	2.434759429869	0.495434	0.330547
5	-0.770030522546	0.678187993781	2.344553350773	0.353482	0.266997
6	-0.790366392447	0.647508467208	2.235020537292	0.422037	0.282116
7	-0.808148254256	0.619438044914	2.153257234038	0.312524	0.236197
8	-0.824081292912	0.593158843966	2.053383274852	0.380497	0.254310
9	-0.838526320173	0.568284576295	1.978623769234	0.284422	0.214862
10	-0.851699224602	0.544611375933	1.887164023937	0.347833	0.232390
20	-0.936154591208	0.357338519592	1.239301373708	0.227598	0.151957
30	-0.972442019142	0.234767889022	0.814251087745	0.149509	0.099816
40	-0.988103025821	0.154252871474	0.535000829982	0.098233	0.065582
50	-0.994863934651	0.101351520455	0.351521240958	0.064544	0.043091
60	-0.997782697745	0.066592826264	0.230966375385	0.042408	0.028313
70	-0.999042763435	0.043754692653	0.151756027590	0.027864	0.018603
80	-0.999586749239	0.028748939489	0.099711015895	0.018308	0.012223
90	-0.999821594580	0.018889437263	0.065514937688	0.012029	0.008031
100	-0.999922980193	0.012411269649	0.043046468053	0.007904	0.005277
110	-0.999966749605	0.008154801658	0.028283601836	0.005193	0.003467
120	-0.999985645397	0.005358097274	0.018583688023	0.003412	0.002278
130	-0.999993802942	0.003520527856	0.012210377679	0.002242	0.001497
140	-0.999997324654	0.002313156285	0.008022805962	0.001473	0.000983
150	-0.999998845020	0.001519855038	0.005271369747	0.000968	0.000646
160	-0.999999501381	0.000998617928	0.003463543696	0.000636	0.000425
170	-0.999999784740	0.000656140054	0.002275714949	0.000418	0.000279
180	-0.999999907070	0.000431115603	0.001495254278	0.000275	0.000183
190	-0.999999959881	0.000283263706	0.000982454044	0.000180	0.000120
191	-0.999999963113	0.000271613062	0.000946874108	0.000135	0.000102
192	-0.999999966085	0.000260441610	0.000903299320	0.000166	0.000111
193	-0.999999968818	0.000249729639	0.000870586001	0.000124	0.000094
194	-0.999999971330	0.000239458253	0.000830521962	0.000152	0.000102
195	-0.999999973640	0.000229609328	0.000800444303	0.000114	0.000086
196	-0.999999975764	0.000220165491	0.000763608158	0.000140	0.000094
197	-0.999999977716	0.000211110078	0.000735953807	0.000105	0.000079
198	-0.999999979512	0.000202427115	0.000702085490	0.000129	0.000086
199	-0.999999981162	0.000194101282	0.000676659205	0.000097	0.000073
200	-0.999999982680	0.000186117891	0.000645519604	0.000119	0.000079
...					
	-1 = Q(xs)				

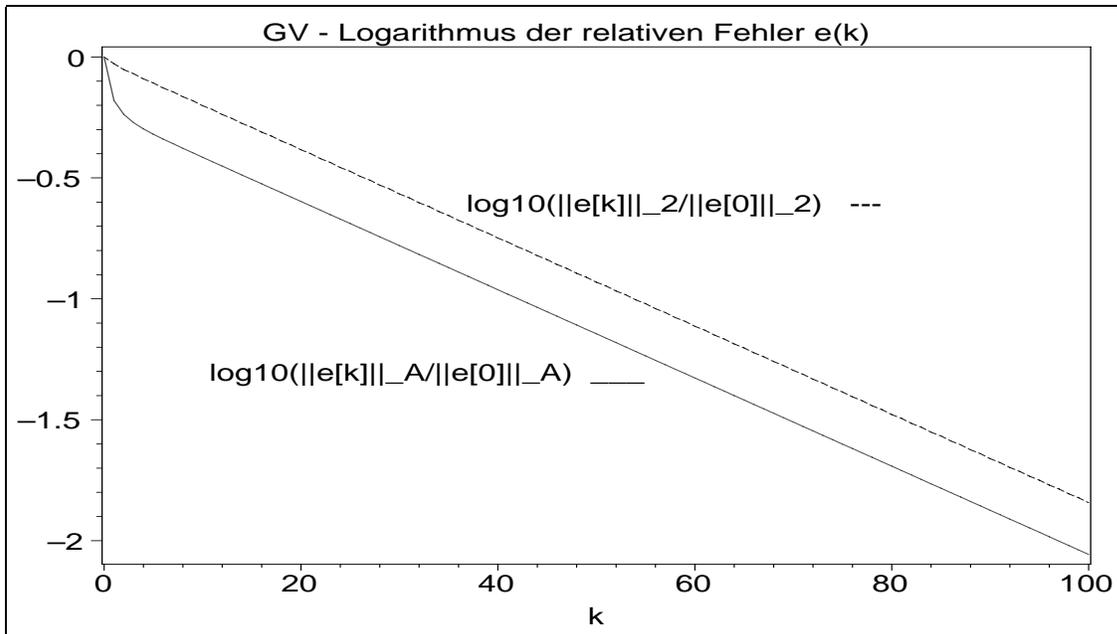


Abb. 4.6 Dateien *gv1_e02.ps*, Matrix $A(10, 10)$ zu Beispiel 2,
 GV: numerische Rechnung, `Digits:=16`,
 Verlauf der relativen Fehler $\log_{10}(\frac{\|e^{(k)}\|_{A,2}}{\|e^{(0)}\|_{A,2}})$, $k = 0, 1, \dots, 100$

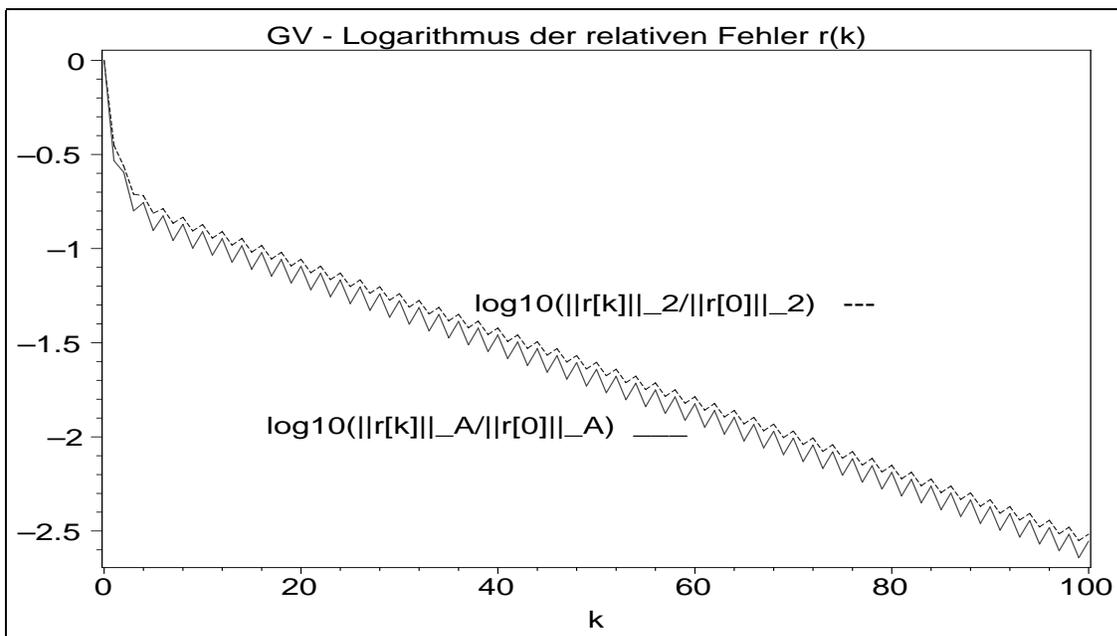


Abb. 4.7 Dateien *gv1_r02.ps*, Matrix $A(10, 10)$ zu Beispiel 2,
 GV: numerische Rechnung, `Digits:=16`,
 Verlauf der relativen Fehler $\log_{10}(\frac{\|r^{(k)}\|_{A,2}}{\|r^{(0)}\|_{A,2}})$, $k = 0, 1, \dots, 100$

4.3.3 Beispiel 3

Wir betrachten nun ein LGS mit einer spd Matrix, die eine sehr schlechte Kondition hat.

```
# Beispiel 3, A=A'>0
n:=8:
A:=matrix(n,n,
  [[ 168, 24, 338, 27, 27, 53, -7, 80],
   [ 24, 178, 169, 72, 53, -103, 17, 80]
   [ 338, 169, 1177, 192, -62, -108, -48, 180],
   [ 27, 72, 192, 125, 2, -24, 36, 180],
   [ 27, 53, -62, 2, 222, 70, 46, 100],
   [ 53, -103, -108, -24, 70, 178, 34, 100],
   [ -7, 17, -48, 36, 46, 34, 34, 100],
   [ 80, 80, 180, 180, 100, 100, 100, 400]]):
b:=vector(n,[ 229, 129, 790, -214, 106, -276, -216, -600]):
# Loesung
xs:=vector(n,[1,-1,1,-1,2,-2,2,-2]):
```

EW $\lambda(A) = \{0.006\,305, 1.359\,344, 5.859\,472, 44.417\,725, 199.730\,179, 282.983\,844, 548.279\,552, 1\,399.363\,575\}$,

Kondition mittels Spektralnorm $\kappa(A) = 221\,911.791$, Konvergenzfaktor $\eta = 0.999\,990$.

Für das spezielle LGS berechnen wir mit dem GV auch die absoluten Fehler

$$\|e^{(k)}\|_A, \|e^{(k)}\|_2, \|r^{(k)}\|_A, \|r^{(k)}\|_2, \quad k = 0, 1, 2, \dots,$$

die man als zusätzliche Ergebnisvektoren `eva[1..maxiter+1]`, `ev2[1..maxiter+1]` bzw. `rva[1..maxiter+1]`, `rv2[1..maxiter+1]` in die Prozedur `gv` einbauen kann. Die Werte $\|e^{(k)}\|_A$ und $\|e^{(k)}\|_2$ bilden monoton abnehmende Folgen, was für $\|r^{(k)}\|_A$ und $\|r^{(k)}\|_2$ nicht zutrifft. Man bemerke auch die Größenordnung dieser Werte.

Dann folgen grafische Vergleiche der relativen Fehler des GV, die aus Gründen der besseren Übersicht als Funktionen $\log_{10}()$ in Abhängigkeit von der Iterationszahl k dargestellt werden, also

$$f(k) = \log_{10} \left(\frac{\|e^{(k)}\|_A}{\|e^{(0)}\|_A} \right), \log_{10} \left(\frac{\|e^{(k)}\|_2}{\|e^{(0)}\|_2} \right), \log_{10} \left(\frac{\|r^{(k)}\|_A}{\|r^{(0)}\|_A} \right), \log_{10} \left(\frac{\|r^{(k)}\|_2}{\|r^{(0)}\|_2} \right).$$

Die numerische Rechnung mit `Digits:=16` erfolgt in Maple, die bei der Dimension $n = 8$ des Problems ohne Weiteres machbar ist, auch wenn das GV für eine hohe Genauigkeit sehr viele Iterationen ausführt. Natürlich sollte bei dieser schlechten Matrixkondition von $\approx 2 \cdot 10^5$ die Computerarithmetik hinreichend stark sein, um abnehmendes Fehlerverhalten überhaupt zu erhalten.

Der Startvektor ist $x^{(0)} = (1, 0, \dots, 0)^T$.

GV: Iterationsverlauf mit verschiedenen Fehlern

m	$Q(x(m)) =$ $(\ e(m)\ ^2_{A^{-1}} - x^T b) / 2$	$\ e(m)\ _{A^{-1}} =$ $\ x^s - x(m)\ _{A^{-1}}$	$\ e(m)\ _{-2} =$ $\ x^s - x(m)\ _{-2}$	$\ r(m)\ _{A^{-1}} =$ $\ b - Ax(m)\ _{A^{-1}}$	$\ r(m)\ _{-2} =$ $\ b - Ax(m)\ _{-2}$
0	-145.000000000000	48.435524153249	4.358898943541	21918.640537	947.467150
1	-983.686052280891	25.857840115489	3.656099578092	16915.791279	528.125735
2	-1119.622116113098	19.918729070245	3.511209064653	6811.889128	287.675025
3	-1193.419930268122	15.784807235559	3.343143508574	9188.730109	288.114026
4	-1234.225234329088	12.944092526779	3.281298868820	3773.849690	161.096233
5	-1257.870399247182	10.966275644248	3.200385994632	5328.163489	168.426165
6	-1272.043117801318	9.587166651173	3.167090402675	2270.920830	98.750030
7	-1281.262778913838	8.571723407362	3.117327008117	3471.050202	110.934984
8	-1287.548025306617	7.804098243024	3.095233020972	1558.696560	69.092634
9	-1292.238036209978	7.178016967105	3.059339466667	2594.184321	83.602545
10	-1295.867532356602	6.653189858015	3.043036998582	1213.208761	54.380569
11	-1298.838335877535	6.190583837162	3.015649252255	2125.259693	68.748493
12	-1301.311191291818	5.777336533072	3.003183601509	1012.562512	45.583278
13	-1303.416658630783	5.400618736629	2.982005977484	1810.105828	58.629487
14	-1305.219786122933	5.055732167959	2.972363447648	867.957187	39.128389
15	-1306.775541375680	4.738028835775	2.955905137513	1561.937466	50.611897
16	-1308.120327372845	4.445148507565	2.948399169895	750.468916	33.847002
17	-1309.285486978356	4.174808503787	2.935542935962	1353.347456	43.858809
18	-1310.295620371509	3.925399248100	2.929662823158	650.672497	29.350843
19	-1311.172070918560	3.695383358040	2.919552075088	1174.253808	38.057007
20	-1311.932720429405	3.483469411548	2.914909029624	564.716448	25.475740
...					
40	-1316.694048822868	1.616138098760	2.856058849230	144.160913	6.520724
60	-1317.067494839672	1.365653807029	2.832676247286	56.586500	2.591264
80	-1317.173754818441	1.285492264899	2.812921158314	43.848513	2.020246
100	-1317.253309560465	1.222039638911	2.794990924012	40.239285	1.855144
120	-1317.321903883604	1.164556667917	2.778636255632	37.615834	1.734412
140	-1317.382223266327	1.111554527383	2.763673412923	35.368842	1.630932
160	-1317.435780660118	1.062279944159	2.749951961151	33.393751	1.539948
180	-1317.483684279328	1.016184747644	2.737346395370	31.631264	1.458740
190	-1317.505787673603	0.994195480172	2.731428120776	30.815641	1.421153
191	-1317.507935460519	0.992032801354	2.730660035357	60.847771	1.994426
192	-1317.510072202197	0.989876555739	2.730273652790	30.657170	1.413850
193	-1317.512197986136	0.987726696880	2.729511896991	60.535730	1.984201
194	-1317.514312883703	0.985583194152	2.729128697885	30.500172	1.406614
195	-1317.516416980256	0.983446002324	2.728373211802	60.226568	1.974071
196	-1317.518510345599	0.981315091499	2.727993166042	30.344618	1.399445
197	-1317.520593063015	0.979190417626	2.727243890568	59.920225	1.964033
198	-1317.522665200786	0.977071951509	2.726866968444	30.190477	1.392341
199	-1317.524726840197	0.974959650247	2.726123845253	59.616643	1.954086
200	-1317.526778048051	0.972853485320	2.725750017443	30.037719	1.385301
...					
250	-1317.617455484033	0.874693678915	2.700596193848	26.597610	1.226719
300	-1317.689049245282	0.788607322711	2.680030924914	23.703763	1.093283
350	-1317.746110380661	0.712586302617	2.663159882062	21.192825	0.977485
400	-1317.791801526251	0.645288266979	2.649276129269	18.977939	0.875333
450	-1317.828472364341	0.585709203713	2.637811110602	17.007805	0.784466
500	-1317.857937075356	0.533034566693	2.628304305875	15.248120	0.703304
600	-1317.900675163402	0.445701327347	2.613740409133	12.262377	0.565591
700	-1317.928323843768	0.378618954177	2.603362569855	9.864361	0.454985
800	-1317.946219792130	0.327964046409	2.595712131127	7.937120	0.366093
900	-1317.957809007504	0.290485774163	2.589835435893	6.388292	0.294655
1000	-1317.965319387949	0.263365191515	2.585111013665	5.143955	0.237261
1500	-1317.977656416013	0.211393396239	2.568664857770	1.777166	0.081974
2000	-1317.979237606687	0.203776315173	2.555338364269	0.720891	0.033259
3000	-1317.979822287211	0.200886598804	2.529597732401	0.428953	0.019799
4000	-1317.980228251351	0.198855468362	2.504157499861	0.419725	0.019373
5000	-1317.980623961770	0.196855470991	2.478973580545	0.415440	0.019176
10000	-1317.982486717084	0.187153856043	2.356802514100	0.394965	0.018231
15000	-1317.984170392287	0.177930366793	2.240652395088	0.375500	0.017332
20000	-1317.985692203936	0.169161438060	2.130226493533	0.356994	0.016477
...					
-1318	= Q(x _s)				

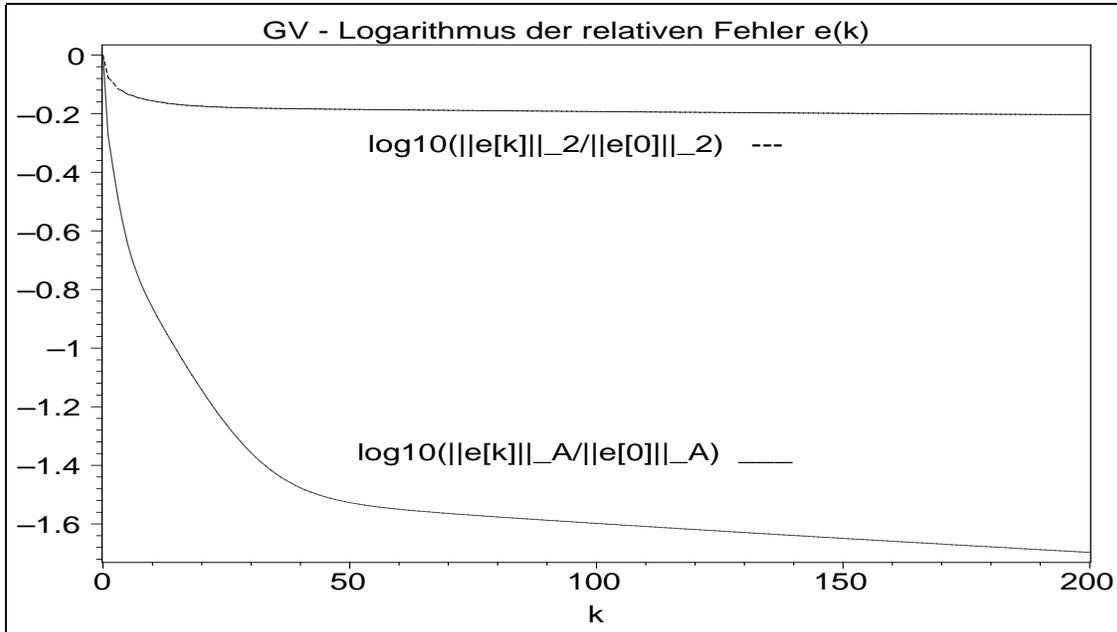


Abb. 4.8 Dateien *gv1_e03.ps*, Matrix $A(8,8)$ zu Beispiel 3,
 GV: numerische Rechnung, `Digits:=16`,
 Verlauf der relativen Fehler $\log_{10}\left(\frac{\|e^{(k)}\|_{A,2}}{\|e^{(0)}\|_{A,2}}\right)$, $k = 0, 1, \dots, 200$

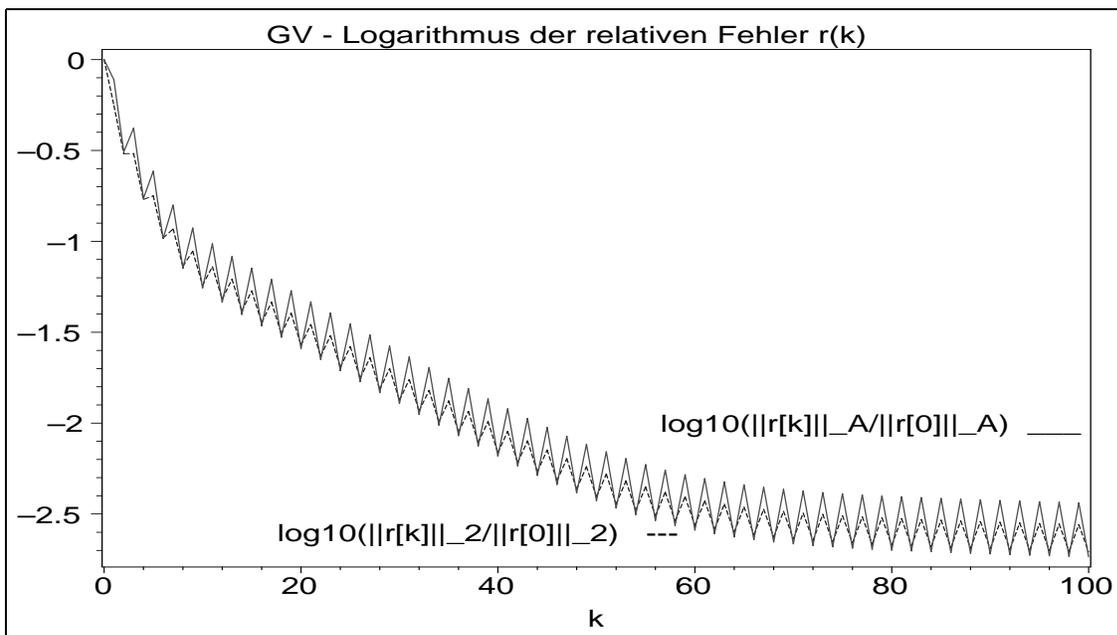


Abb. 4.9 Dateien *gv1_r03.ps*, Matrix $A(8,8)$ zu Beispiel 3,
 GV: numerische Rechnung, `Digits:=16`,
 Verlauf der relativen Fehler $\log_{10}\left(\frac{\|r^{(k)}\|_{A,2}}{\|r^{(0)}\|_{A,2}}\right)$, $k = 0, 1, \dots, 100$,
 stärkere Oszillation in der A -Norm

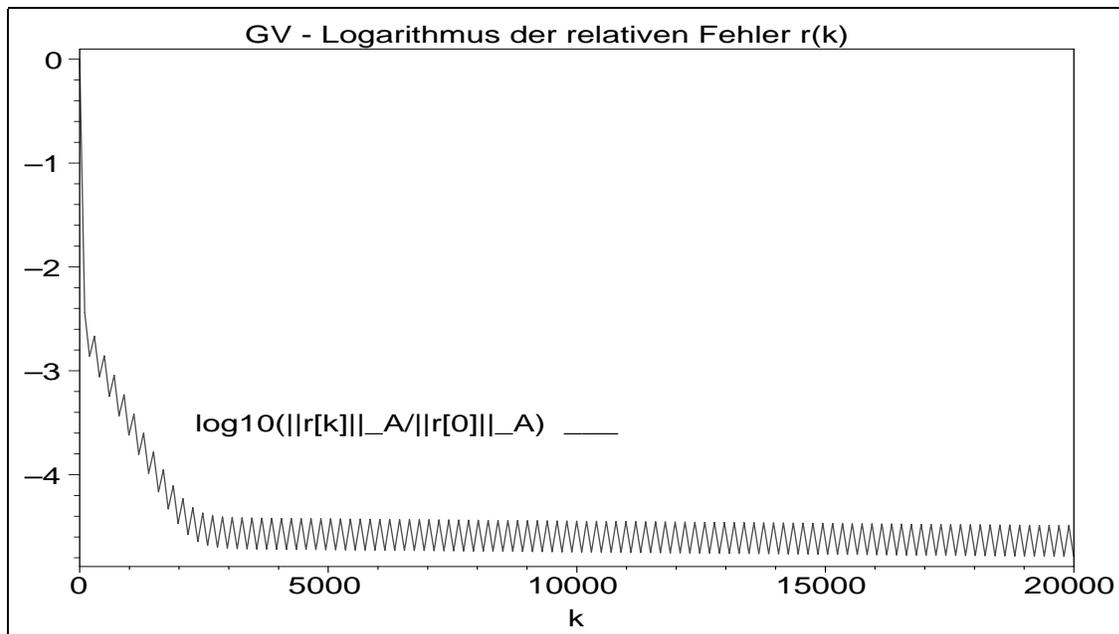


Abb. 4.10 Dateien *gv1_r03a.ps*, Matrix $A(8,8)$ zu Beispiel 3,
 GV: numerische Rechnung, `Digits:=16`,
 Verlauf des relativen Fehlers $\log_{10}\left(\frac{\|r^{(k)}\|_A}{\|r^{(0)}\|_A}\right)$, $k = 0, 99, 198, \dots \leq 20\,000$

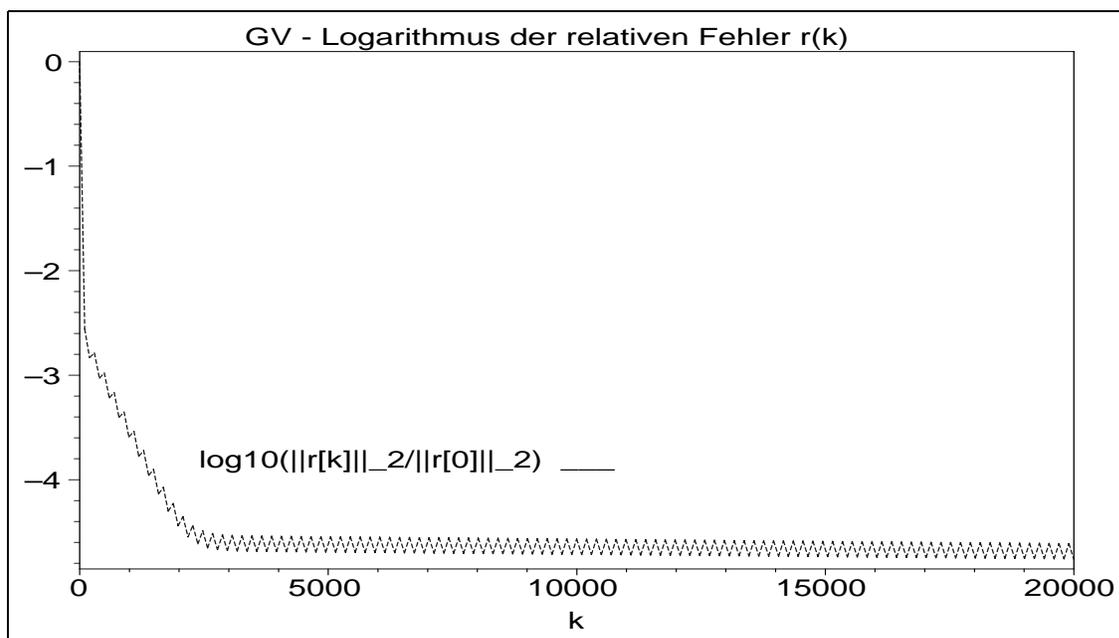


Abb. 4.11 Dateien *gv1_r032.ps*, Matrix $A(8,8)$ zu Beispiel 3,
 GV: numerische Rechnung, `Digits:=16`,
 Verlauf des relativen Fehlers $\log_{10}\left(\frac{\|r^{(k)}\|_2}{\|r^{(0)}\|_2}\right)$, $k = 0, 99, 198, \dots \leq 20\,000$

Literaturverzeichnis

- [1] AXELSSON, O.: *Iterative Solution Methods*. Cambridge University Press Cambridge 1994, 1996.
- [2] BERESIN, I. S. und N. P. SHIDKOW: *Numerische Methoden*. Bd. 1,2. DVW Berlin 1970, 1971.
- [3] BREZINSKI, C.: *Projection Methods for Systems of Equations*. Studies in Computational Mathematics. Elsevier Amsterdam 1997.
- [4] DEUFLHARD, P. und H. HOHMANN: *Numerische Mathematik*. 1: Eine algorithmisch orientierte Einführung. 3. überarbeitete und erweiterte Auflage, Lehrbuch. Walter de Gruyter Berlin 2002.
- [5] FADDEJEV, D. K. und W. N. FADDEJEWA: *Numerische Methoden der linearen Algebra*. Math. für Naturwiss. und Technik, Bd. 10. DVW Berlin 1973.
- [6] FISCHER, B.: *Polynomial Based Iteration Methods for Symmetric Linear Systems*. Advances in Numerical Mathematics. Wiley-Teubner Stuttgart 1996.
- [7] GREENBAUM, A.: *Iterative Methods for Solving Linear Systems*. SIAM Philadelphia 1997.
- [8] HACKBUSCH, W.: *Iterative Lösung großer schwachbesetzter Gleichungssysteme*. Leitfäden der angewandten Mathematik und Mechanik Band 69. B. G. Teubner Stuttgart 1991, 1993.
- [9] HÄMMERLIN G. und K.-H. HOFFMANN: *Numerische Mathematik*. Grundwissen Mathematik 7. Springer-Verlag Berlin 1991.
- [10] HERMANN, M.: *Numerische Mathematik*. R. Oldenbourg Verlag München 2001.
- [11] HESTENES, J. R. und E. STIEFEL: *Methods of conjugate gradients for solving linear systems*. Journ. Res. Nat. Bur. Stand. 49 (1952) 409-436.
- [12] KELLEY, C. T.: *Iterative Methods for Linear and Nonlinear Equations*. Frontiers in Applied Mathematics. SIAM Philadelphia 1995.
- [13] KIELBASIŃSKI, A. und H. SCHWETLICK: *Numerische lineare Algebra*. DVW Berlin 1988.
- [14] MAESS, G.: *Vorlesungen über Numerische Mathematik I, II*. Akademie-Verlag Berlin 1984, 1988.
- [15] MEISTER, A.: *Numerik linearer Gleichungssysteme*. Ein Einführung in moderne Verfahren. Vieweg Braunschweig 1999.
- [16] MEURANT, G.: *Computer Solution of Large Linear Systems*. Studies in Mathematics and Its Applications, Vol 28. Elsevier Science B. V. 1999.

- [17] NEUNDORF, W.: *Numerische Mathematik*. Vorlesungen, Übungen, Algorithmen und Programme. Shaker Verlag Aachen 2002.
- [18] NEUNDORF, W.: *Grundlagen der numerischen linearen Algebra*. Preprint No. M 04/04 IfMath der TU Ilmenau, Februar 2004.
- [19] OPFER, G.: *Numerische Mathematik für Anfänger*. Vieweg Studium Grundkurs Mathematik Wiesbaden 1993, 3. überarbeitete und erw. Auflage 2001.
- [20] PLATO, R.: *Numerische Mathematik kompakt*. Grundlagenwissen für Studium und Praxis. Vieweg Wiesbaden 2000.
- [21] QUARTERONI, A., R. SACCO und F. SALERI: *Numerische Mathematik*. Band 1, 2. Springer-Verlag Berlin 2002.
- [22] RALSTON, A.: *A First Course in Numerical Analysis*. McGraw-Hill New York 1965.
- [23] ROOS, H.-G. und H. SCHWETLICK: *Numerische Mathematik*. Das Grundwissen für jedermann. B. G. Teubner Stuttgart 1999.
- [24] SAAD, Y.: *Iterative Methods for Sparse Linear Systems*. PWS Publishing Company Boston 1995.
- [25] SCHWARZ, H. R.: *Numerische Mathematik*. B. G. Teubner Stuttgart 1988.
- [26] SCHWARZ, H. R., H. RUTISHAUSER und E. STIEFEL: *Numerik symmetrischer Matrizen*. B. G. Teubner Stuttgart 1972.
- [27] STOER, J. und R. BULIRSCH: *Einführung in die Numerische Mathematik II*. Heidelberger Taschenbücher 114. Springer-Verlag Berlin 1990.
- [28] TREFETHEN, L. N. und D. BAU: *Numerical Linear Algebra*. SIAM Philadelphia 1997.
- [29] ÜBERHUBER, C.: *Computer-Numerik 1,2*. Springer-Verlag Berlin 1995.
- [30] VAN DER VORST, H. A.: *Iterative Krylov Methods for Large Linear Systems*. Cambridge University Press Cambridge 2003.
- [31] KANZOW, CH.: *Numerik linearer Gleichungssysteme*. Direkte und iterative Verfahren. Springer-Verlag Berlin Heidelberg 2005.
- [32] GEIGER, C. und CH. KANZOW: *Numerische Verfahren zur Lösung unrestringierter Optimierungsaufgaben*. Springer-Verlag Berlin 1999.
- [33] WATKINS, DAVID S.: *Fundamentals of Matrix Computations*. 2nd Ed. Pure and Applied Mathematics. A Wiley-Interscience Series of Texts, Monographs, and Tracts. A John Wiley & Sons, Inc., Publication, New York 2002.
- [34] NEUNDORF, W.: *Zu Orthogonalsystemen von Polynomen und ihrer Rekursion*. Preprint No. M 08/05 IfMath der TU Ilmenau, Mai 2005.
- [35] NEUNDORF, W.: *Abstiegsverfahren*. Teil I. Preprint No. M 19/04 IfMath der TU Ilmenau, Oktober 2004.
- [36] NEUNDORF, W.: *Abstiegsverfahren*. Teil II. Preprint No. M 20/04 IfMath der TU Ilmenau, Oktober 2004.
- [37] NEUNDORF, W.: *Abstiegsverfahren*. Teil III. Preprint No. M 09/05 IfMath der TU Ilmenau, Mai 2005.

Anschrift:

Privatdozent Dr. rer. nat. habil. Werner Neundorf
Technische Universität Ilmenau, Institut für Mathematik
PF 10 05 65
D - 98684 Ilmenau

E-mail : werner.neundorf@tu-ilmenau.de

Homepage : <http://www.tu-ilmenau.de/site/math/neundorf.html>