

Technische Universität Ilmenau
Fakultät für Mathematik
und Naturwissenschaften
Institut für Mathematik
http://www.mathematik.tu-ilmenau.de/Math-Net/index_de.html

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 20/04

Abstiegsverfahren

Teil II

Werner Neundorf

Oktober 2004

‡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 in dieser mehrteiligen Arbeit sind grundlegende Abstiegsverfahren als Vertreter dieser Algorithmengruppe, die auf Minimierungsaufgaben nach Umformulierung eines regulären linearen Gleichungssystems führen.

Unter bestimmten Voraussetzungen an die Matrix des Gleichungssystems werden geeignete Funktionale konstruiert und damit der Weg des Abstiegs illustriert.

Das Verhalten der Abstiegsverfahren ist in den normalen gutartigen Fällen hinreichend bekannt und untersucht worden. Hier soll zunächst eine Gegenüberstellung zu anderen Situationen gemacht werden, wo man unter veränderten Voraussetzungen arbeitet, und damit der typische Charakter der Minimierungsaufgabe nicht mehr vorhanden ist. Dabei ist teilweise noch mit zufrieden stellenden Ergebnissen zu rechnen, es können aber auch starke Abweichungen vom Normalfall auftreten. Diese Darstellungen findet der Leser in den Teilen I und II.

Des Weiteren betrachten wir in einem Teil III die Abstiegsverfahren als polynomiale Iterationsverfahren und untersuchen den Einfluss von Rundungsfehlern bei der Implementation der Verfahren im Zusammenhang mit ihrer (eventuell) theoretischen Endlichkeit sowie mit der Notation des Formelapparates und seiner numerischen Auswertung (Fehlerverhalten und Fehlererinnerung).

Praktische Rechnungen mit kleindimensionierten Beispielen, wo man dies auch gut illustrieren kann, demonstrieren die unterschiedlichen Situationen.

Eileen Caddy

Spuren auf dem Weg zum Licht

Hört auf zu versuchen,
alles gedanklich aufarbeiten,
das bringt euch nirgendwohin.
Lebt auf der Intuition und Inspiration,
und laßt euer ganzes Leben
eine Offenbarung sein.

Inhaltsverzeichnis

Teil I

1	Funktionale und Abstiegszenarien	1
2	Grundlagen der Abstiegsverfahren	23
2.1	Ansätze für quadratische Formen	25
3	Abstiegsverfahren – 1	30
3.1	Die quadratische Form $Q(x)$	30
3.2	Das Gradientenverfahren	35
3.2.1	Zur Konvergenz des Gradientenverfahrens	43
3.2.2	Krylov-Unterraum und Suchrichtungen	53
3.2.3	Beispiele zum Gradientenverfahren	56
3.3	Abstiegsverfahren mit linear unabhängigen Richtungen	74
3.4	Optimalität im Abstiegsverfahren	80
3.5	Abstiegsverfahren mit konjugierten Richtungen	82
3.6	Verfahren der konjugierten Gradienten	90
3.6.1	Beschreibung als (endliches) Iterationsverfahren	91
3.6.2	Wichtige Eigenschaften der Verfahrensgrößen	93
3.6.3	Varianten der Realisierung des Verfahrens	108
3.6.4	Zur Konvergenz des Verfahrens	114
3.6.5	Modellproblem mit Vergleich von Abstiegsverfahren	129
3.6.6	Beispiele zum Verfahren der konjugierten Gradienten	133

Teil II

4	Abstiegsverfahren – 2	1
----------	------------------------------	----------

4.1	Die quadratische Form $R(x)$	1
4.2	Verfahren der A-orthogonalen Residua	3
4.2.1	Beispiele zum Verfahren der A-orthogonalen Residua	7
4.2.2	Verfahren der A-orthogonalen Residua als Projektionsmethode	19
4.3	Verfahren der konjugierten Residua	20
4.3.1	Beispiele zum Verfahren der konjugierten Residua	28
5	Abstiegsverfahren unter veränderten Voraussetzungen	44
5.1	Gradientenverfahren	44
5.1.1	GV mit symmetrischer indefiniter Matrix	44
5.1.2	GV mit nicht symmetrischer Matrix	51
5.2	Verfahren der konjugierten Gradienten	67
5.2.1	CG mit symmetrischer indefiniter Matrix	67
5.2.2	CG mit nicht symmetrischer Matrix	73
5.3	Verfahren der A-orthogonalen Residua	86
5.3.1	Orthomin(0) mit symmetrischer indefiniter Matrix	87
5.3.2	Orthomin(0) mit nicht symmetrischer Matrix	98
5.4	Verfahren der konjugierten Residua	106
5.4.1	CR mit symmetrischer indefiniter Matrix	107
5.4.2	CR mit nicht symmetrischer Matrix	116
	Literaturverzeichnis	131

Kapitel 4

Abstiegsverfahren – 2

4.1 Die quadratische Form $R(x)$

Im Abschnitt 2.1 haben wir zum LGS (1.1) die quadratische Form $R(x)$ (2.19) unter verschiedenen Gesichtspunkten schon betrachtet.

Das zugehörige Minimierungsproblem ist

$$\min_{x \in \mathbb{R}^n} R(x) \Leftrightarrow \min_{x \in \mathbb{R}^n} \|r(x)\|_2. \quad (4.1)$$

Wie in den Beziehungen (2.16) unschwer zu erkennen ist, kommt man auf das Funktional über das Normalgleichungssystem (Symmetrisierung) zu (1.1).

$Ax = b, A = A^T > 0$	$Ax = b, \exists A^{-1}$
\Downarrow	$A^T Ax = A^T b, A^T A \text{ spd}$
\Downarrow	\Downarrow
$Q(x) = \frac{1}{2}x^T Ax - x^T b \rightarrow \min_x$	$R(x) = \frac{1}{2}x^T A^T Ax - x^T A^T b \rightarrow \min_x$
\Updownarrow	\Updownarrow
$\ e(x)\ _A \rightarrow \min_x$	$\ e(x)\ _{A^T A} \rightarrow \min_x$
	$(x^* - x)^T A^T A(x^* - x) \rightarrow \min$
	$(A(x^* - x))^T A(x^* - x) \rightarrow \min$
	$\ A(x^* - x)\ _2 \rightarrow \min$
	$\ b - Ax\ _2 \rightarrow \min$
	$\ r(x)\ _2 \rightarrow \min.$

Egal, ob man von der Untersuchung der Norm $\|b - Ax\|_2$ kommt oder vom Normalgleichungssystem $A^T Ax = A^T b$, die gemeinsame quadratische Form $R(x)$ hängt mit der Minimierung der Residuumnorm zusammen.

Zum Funktional $R(x)$ werden natürlich auch sein Gradient und seine Hesse-Matrix gebraucht. Wir haben also

$$\begin{aligned} R(x) &= \frac{1}{2}x^T A^T A x - x^T A^T b = x^T B x - x^T c, \\ B &= A^T A = B^T > 0, \quad c = A^T b, \\ \text{grad } R(x) &= \nabla R(x) = Bx - c, \\ \nabla^2 R(x) &= B > 0. \end{aligned} \tag{4.2}$$

Dazu ergibt sich wie in (2.22) die Fehlerbetrachtung

$$\begin{aligned} 0 \leq \|e(x)\|_{A^T A}^2 &= 2R(x) + x^{*T} c, \\ &\vdots \\ \|r(x)\|_2^2 &= 2R(x) + x^{*T} A^T b \\ &= 2R(x) + (Ax^*)^T b \\ &= 2R(x) + b^T b, \end{aligned}$$

und die Beziehungen

$$\begin{aligned} R(x) &= \frac{1}{2} \left(\|r(x)\|_2^2 - b^T b \right), \\ R(x^*) &= \min_{x \in \mathbb{R}^n} R(x) = -\frac{1}{2} b^T b, \\ x^* &= \arg \min_{x \in \mathbb{R}^n} R(x), \\ R(x^* + \Delta x) - R(x^*) &= \frac{1}{2} \|\Delta x\|_{A^T A}^2 = \frac{1}{2} \|A \Delta x\|_2^2 > 0, \quad \Delta x \neq 0. \end{aligned} \tag{4.3}$$

Formal ist natürlich der Fehler $\|e(x)\|_{A^T A}$ nicht so praktikabel und plausibel wie $\|r(x)\|_2$.

Die notwendige Bedingung am Minimum (Extremum) $\nabla R(x) = 0$ liefert die einzige globale Minimumstelle x^* von $R(x)$ und zugleich die eindeutige Lösung des LGS $Bx = c$ bzw. von (1.1). Es trifft also die Aussage von Satz 3.1 zu.

Die Konstruktion einer Folge von Näherungen zur Minimumstelle von $R(x)$ geht von einem Startvektor $x^{(0)}$ als ersten Näherungswert zu x^* aus. Weitere Iterierte $x^{(1)}, x^{(2)}, \dots$ eines AV findet man mit der Bedingung

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

also $\|r^{(m+1)}\|_2 < \|r^{(m)}\|_2$, $r^{(m)} = r(x^{(m)})$.

Dazu betrachtet man AV als Folge ‘‘eindimensionaler Probleme‘‘ mit einer Strahlenminimierung.

Bezüglich geeigneter Suchrichtungen (Richtungsvektoren) $p(x)$ haben wir in den bisherigen AV schon mehrere Varianten betrachtet, die auch hier verwendet werden können.

- (a) $r(x) = b - Ax$ als Richtung des steilsten Abstiegs zum Funktional $Q(x)$ im Fall $A = A^T$.
- (b) $r(x) = b - \frac{1}{2}(A + A^T)x$ zum Funktional $Q(x)$ im Fall $A \neq A^T$ (siehe (1.8)).
- (c) $\hat{r}(x) = A^T(b - Ax)$ als Richtung des steilsten Abstiegs zum Funktional $R(x)$ im Fall A regulär. Damit ist man beim GV für das Normalgleichungssystem.
- (d) $p(x)$ als Abstiegsrichtung zum Funktional $R(x)$ bei ihrer rekursiven Berechnung aus $r(x)$ und vorheriger Abstiegsrichtung analog zur Vorgehensweise im CG.

4.2 Verfahren der A-orthogonalen Residua

Man bezeichnet das Verfahren auch mit **Orthomin(0)**.

Das zu minimierende Funktional $R(x)$ mit seinen Merkmalen (4.2) verwendet als Suchrichtung den negativen Gradienten von $Q(x)$, der im Fall $A = A^T > 0$ auch Abstiegsrichtung ist.

$$p(x) = r(x), \quad r(x) = b - Ax. \quad (4.4)$$

Ausgehend von der Iterierten $x^{(m)}$ haben wir in einem Schritt die Beziehungen

$$\begin{aligned} x^{(m+1)} &= x^{(m)} + \alpha_m p^{(m)}, \\ R(x^{(m+1)}) &\leq R(x^{(m)}), \end{aligned} \quad (4.5)$$

sowie die Minimierungsaufgabe in kompakter Form

$$\begin{aligned} R(x^{(m+1)}) &= \min_{\alpha \in \mathbb{R}} R(x^{(m)} + \alpha p^{(m)}) \\ x^{(m+1)} &= \arg \min_{\alpha \in \mathbb{R}} R(x^{(m)} + \alpha p^{(m)}) \\ p^{(m)} &= r^{(m)} \end{aligned} \quad (4.6)$$

Vom Ansatz her wird somit das Konvergenzverhalten von Orthomin(0) dem des GV ähneln.

Für die Bestimmung des optimalen Parameters α brauchen wir nur die skalare Funktion $f(\alpha)$ und bezüglich α die notwendige Bedingung an ihrem Minimum auszuwerten.

Es gilt

$$\begin{aligned}
f(\alpha) &= R(x + \alpha p) \\
&= \frac{1}{2}(x + \alpha p)^T A^T A(x + \alpha p) - (x + \alpha p)^T A^T b \\
&= R(x) + \alpha(A^T(Ax - b))^T p + \frac{1}{2}\alpha^2(Ap)^T Ap \\
&= R(x) + \alpha(Ax - b)^T Ap + \frac{1}{2}\alpha^2(Ap)^T Ap, \tag{4.7} \\
0 = f'(\alpha) &= (Ax - b)^T Ap + \alpha(Ap)^T Ap, \quad r = b - Ax, \\
\alpha &= \frac{r^T Ap}{(Ap)^T Ap} = \frac{(r, Ap)}{(Ap, Ap)} = \frac{(Ap, r)}{(Ap, Ap)}, \\
f''(\alpha) &= (Ap)^T Ap = \|Ap\|_2^2 > 0, \quad p \neq 0.
\end{aligned}$$

Somit gilt im Orthomin(0)

$$\alpha = \alpha_m = \frac{r^{(m)T} Ar^{(m)}}{(Ar^{(m)})^T Ar^{(m)}} = \frac{(r^{(m)}, Ar^{(m)})}{(Ar^{(m)}, Ar^{(m)})} = \frac{r^{(m)T} Ar^{(m)}}{\|Ar^{(m)}\|_2^2} \stackrel{A=A^T > 0}{=} \frac{\|r^{(m)}\|_A^2}{\|Ar^{(m)}\|_2^2}, \tag{4.8}$$

wobei $\alpha_m \geq 0$ für A spd.

Der Iterationsschritt als Aufdatierungsformel ist

$$x^{(m+1)} = x^{(m)} + \alpha_m r^{(m)}. \tag{4.9}$$

Die Iterierten $r^{(m+1)}$ kann man wieder rekursiv bestimmen. Unter Verwendung der Kürzel (3.14) erhält man

$$\begin{aligned}
r^{(m+1)} &= b - Ax^{(m+1)} \quad (\text{direkte Berechnung}), \\
&= b - A(x^{(m)} + \alpha_m r^{(m)}) \\
&= b - (v^{(m)} + \alpha_m w^{(m)}), \quad w^{(m)} = Ar^{(m)} \\
r^{(m+1)} &= b - v^{(m+1)}, \quad \text{wobei } v^{(m+1)} = v^{(m)} + \alpha_m w^{(m)} \\
&\quad (1. \text{ rekursive Berechnung}), \\
&= b - Ax^{(m)} - \alpha_m Ar^{(m)} \\
&= r^{(m)} - \alpha_m Ar^{(m)} \\
r^{(m+1)} &= r^{(m)} - \alpha_m w^{(m)} \quad (2. \text{ rekursive Berechnung}).
\end{aligned} \tag{4.10}$$

Die Funktionswerte $R(x^{(m)})$ bzw. $\|r(x^{(m)})\|_2$ sind abnehmend und $R(x^*) = -\frac{1}{2}b^T b$.

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

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

mit Nenner $\neq 0$ zu machen.

Wenn $r^{(m)} = 0$ ist, ist die Iteration beendet. Für $r^{(m)} \neq 0$ gilt theoretisch $Ar^{(m)} \neq 0$, aber in der numerischen Berechnung könnte dieses Produkt ein Nullvektor sein und α_m ist nicht auswertbar. Das Verfahren bricht dann ab.

Die Suchrichtungen $r^{(m)}$ erfüllen eine Orthogonalitätsbedingung. Es gilt für aufeinander folgende Richtungen $r^{(m+1)} \perp Ar^{(m)}$, denn mit (4.10) haben wir

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

Damit sind gemäß Satz 3.8 die Richtungen $r^{(m+1)}$ und $r^{(m)}$ linear unabhängig. Die A-Orthogonalität und lineare Unabhängigkeit von $r^{(m+1)}$ und $r^{(m)}$ sind noch relativ schwache Eigenschaften, da dies zu den weiteren vorherigen Vektoren nicht gilt.

Die A-Orthogonalität macht eigentlich nur Sinn für symmetrische Matrizen.

Ist $A \neq A^T$, so gilt im Allgemeinen $(r^{(m+1)}, A^T r^{(m)}) = (r^{(m)}, Ar^{(m+1)}) \neq 0$.

Zusammenfassung und Bewertung der Eigenschaften von Orthomin(0)

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

- | | |
|---|--------|
| <ul style="list-style-type: none"> + Der Fehler $\ r^{(m)}\ _2 = \ b - Ax^{(m)}\ _2$ wird ständig verkleinert. + Einfache Implementation des Algorithmus. ± Das Abstiegszenario und die Minimierungsaufgabe sind dem Problem nur teilweise angepasst. ± Hauptaufwand in einem Iterationsschritt ist eine Matrix-Vektor-Multiplikation. ± Aufeinander folgende Residua $r^{(m+1)}$ und $r^{(m)}$ sind A-orthogonal und damit linear unabhängig. – Die Suchrichtungen $r^{(m)}$ sind aus $Q(x)$ abgeleitet und keine Richtungen des steilsten Abstiegs zu $R(x)$ (“schräge“ oder “schiefe“ Abstiegsrichtung zu $R(x)$). – Konvergenzverhalten | (4.11) |
| <p>Orthomin(0) ist i. Allg. ein unendliches Iterationsverfahren. Die A-Orthogonalität aufeinander folgender Residua $r^{(m+1)}$, $r^{(m)}$ bedeutet nicht unbedingt dasselbe für $r^{(m+1)}$ und $r^{(m-1)}$.</p> | |

Die Implementierung von Orthomin(0) erfolgt in Maple.

Ergänzt um die zusätzliche Abfrage des Nenners bei der Berechnung der Schrittzahl α , die Bestimmung von Werten der Funktionale sowie einige Ergebnisfelder und Zwischenausgaben entsteht dann die erweiterte Prozedur, die wir in den Beispielrechnungen verwenden.

```
> orthomin0:=proc(n::posint,A::matrix,b::vector,x0::vector,
                 maxiter::posint,etol::numeric,aus::name)

    local  i,k,m,x,r,r2,r20,w,v,
           alpha,alphan,alphas,b2,Q,R,fh,fh1,fh2,xv1,rv1,xx;
    global xv,rv,lp;

    k:=0:
    x:=evalm(x0):
    v:=evalm(A&*x):
    r:=evalm(b-v):  lp:=evalm(r):
    b2:=evalm(transpose(b)&*b):
    r2:=evalm(transpose(r)&*r):
    r20:=evalf(r2):
    Q:=0.5*evalm(transpose(x)&*v)-evalm(transpose(x)&*b):
    R:=0.5*evalm(transpose(v)&*v)-evalm(transpose(v)&*b):
    # R:=0.5*(r2-b2):

    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('Schritt k = %g\n',k);
        printf('Startvektor          x = [||fh||']\n',seq(x[i],i=1..n));
        printf('Residuum/Suchr. r = b-Ax = [||fh||']\n',seq(r[i],i=1..n));
        printf('Funktionswert          Q(x) = '||fh2||'\n',Q);
        printf('Funktionswert          R(x) = '||fh2||'\n',R);
        printf('Anfangsfehlerquadrat r'r = '||fh2||'\n\n',r2);
    end if;
    if r20<etol then RETURN(x,0); end if;

    while (r2/r20>etol) and (k<maxiter) do
        w:=evalm(A&*r):
        alphan:=evalm(transpose(w)&*r):
        alphas:=evalm(transpose(w)&*w);
```

```

if alphas=0 then
  lprint('Abbruch wegen Nenner (Ar)'Ar=0'):
  RETURN(x,k);
end if;

if alphas=0 then
  lprint('Abbruch wegen Zaehler (Ar)'r=0 -> alpha=0'):
  RETURN(x,k);
end if;

alpha:=alphas/alphas;
x:=evalm(x+alpha*r);
r:=evalm(r-alpha*w); # r:=evalm(b-A*x);
lp:=evalm(r);

r2:=evalm(transpose(r)*r);
k:=k+1;
v:=evalm(A*x);
Q:=0.5*evalm(transpose(x)*v)-evalm(transpose(x)*b);
R:=0.5*evalm(transpose(v)*v)-evalm(transpose(v)*b);
# R:=0.5*(r2-b2);

xv1:=concat(xv,x);
xv:=evalm(xv1);
rv1:=concat(rv,r);
rv:=evalm(rv1);

if aus=ja then
  printf('\n');
  printf('Schritt k = %g\n',k);
  printf('Suchschritt      alpha = '||fh2||'\n',alpha);
  printf('Iterationsvektor    x = '||fh||']\n',seq(x[i],i=1..n));
  printf('Residuum/Suchr. r = b-Ax = '||fh||']\n',seq(r[i],i=1..n));
  printf('Funktionswert      Q(x) = '||fh2||'\n',Q);
  printf('Funktionswert      R(x) = '||fh2||'\n',R);
  printf('Fehlernormquadrat  r'r = '||fh2||'\n\n',r2);
end if;

end do:

[x,k];
end:

```

4.2.1 Beispiele zum Verfahren der A-orthogonalen Residua

Für die Beispiele, die wir zum GV im Abschnitt 3.2.3 und CG im Abschnitt 3.6.6 betrachten haben, illustrieren wir nun den Iterationsverlauf von Orthomin(0) für das LGS (1.1) unter der üblichen Voraussetzung $A = A^T > 0$. Dabei bieten sich insbesondere Vergleiche zum GV an.

Beispiel 4.1

Gegeben sei das LGS mit der spd Tridiagonalmatrix

$$A = A(7, 7) = \text{tridiag}(-1, 2, -1) \quad \text{und} \quad b = (2, -7, 11, -13, 8, 2, 5)^T.$$

Die exakte Lösung des LGS ist $x^* = (1, 0, 6, 1, 9, 9, 7)^T$.

Für die Funktionale $Q(x)$ und $R(x)$ gilt $Q(x^*) = -90$, $R(x^*) = -218$.

Der Startvektor sei $x^{(0)} = 0$.

Die folgenden Rechnungen sind mit Maple bei `Digits:=16` gemacht worden.

Orthomin(0) - Iterationsverlauf mit Funktionalen und verschiedenen Fehlern

m	[x(m) [1], x(m) [2], x(m) [3], x(m) [4], x(m) [5], x(m) [6], x(m) [7]]	R(x(m))	Q(x(m))	e(m) _A	e(m) _2
0	[0.000, 0.000, 0.000, 0.000, 0.000, 0.000, 0.000]	0.000	0.000	13.416	15.780
1	[0.543, -1.900, 2.985, -3.528, 2.171, 0.543, 1.357]	-202.976	-63.233	7.317	13.543
2	[-0.070, -1.696, 2.738, -4.018, 2.590, 3.306, 3.117]	-212.407	-76.060	5.280	11.333
3	[0.111, -2.079, 2.660, -3.870, 3.448, 3.752, 3.960]	-214.911	-79.594	4.562	10.373
4	[-0.049, -2.117, 2.517, -3.420, 3.971, 4.763, 4.401]	-216.009	-82.290	3.927	9.340
5	[-0.057, -2.238, 2.692, -3.286, 4.542, 5.108, 4.793]	-216.622	-83.663	3.560	8.680
6	[-0.124, -2.178, 2.742, -2.849, 4.942, 5.714, 5.076]	-216.991	-84.956	3.176	7.900
7	[-0.095, -2.189, 2.945, -2.690, 5.350, 5.959, 5.310]	-217.225	-85.699	2.933	7.373
8	[-0.094, -2.061, 3.074, -2.313, 5.668, 6.373, 5.499]	-217.383	-86.476	2.655	6.726
9	[-0.040, -2.018, 3.277, -2.157, 5.976, 6.552, 5.659]	-217.497	-86.952	2.469	6.282
10	[-0.004, -1.862, 3.432, -1.832, 6.230, 6.856, 5.793]	-217.583	-87.479	2.246	5.729
20	[0.510, -0.889, 4.814, -0.263, 7.808, 8.103, 6.505]	-217.921	-89.504	0.996	2.550
30	[0.781, -0.397, 5.471, 0.438, 8.471, 8.603, 6.781]	-217.984	-89.902	0.443	1.134
40	[0.903, -0.177, 5.765, 0.750, 8.765, 8.823, 6.903]	-217.997	-89.981	0.197	0.504
50	[0.957, -0.079, 5.895, 0.889, 8.895, 8.921, 6.957]	-217.999	-89.996	0.088	0.224
60	[0.981, -0.035, 5.954, 0.951, 8.954, 8.965, 6.981]	-218.000	-89.999	0.039	0.100
70	[0.991, -0.016, 5.979, 0.978, 8.979, 8.984, 6.991]	-218.000	-90.000	0.017	0.044
80	[0.996, -0.007, 5.991, 0.990, 8.991, 8.993, 6.996]	-218.000	-90.000	0.008	0.020
90	[0.998, -0.003, 5.996, 0.996, 8.996, 8.997, 6.998]	-218.000	-90.000	0.003	0.009
100	[0.999, -0.001, 5.998, 0.998, 8.998, 8.999, 6.999]	-218.000	-90.000	0.002	0.004

Orthomin(0) konvergiert linear und etwas langsamer als das GV (vergl. Beispiel 3.10). Verwendet man die Suchrichtung $\hat{r}(x) = A^T(b - Ax)$, was auf das GV für das Normalgleichungssystem hinausläuft, so bemerkt man einen erheblichen Verlust an Konvergenzgeschwindigkeit. Das korrespondiert auch mit der schlechteren spektralen Kondition $\text{cond}_2(A^T A) = 638.782\dots$ im Vergleich mit $\text{cond}_2(A) = 25.274\dots$

GV fuer A'Ax=A'b - Iterationsverlauf mit Funktionalen und verschiedenen Fehlern

m	[x(m) [1], x(m) [2], x(m) [3], x(m) [4], x(m) [5], x(m) [6], x(m) [7]]	R(x(m))	Q(x(m))	e(m) _A	e(m) _2
0	[0.000, 0.000, 0.000, 0.000, 0.000, 0.000, 0.000]	0.000	0.000	13.416	15.780
1	[0.772, -1.895, 2.947, -3.158, 1.895, -0.632, 0.561]	-193.427	-53.944	8.492	14.661
2	[0.212, -1.497, 3.185, -3.846, 1.393, 0.654, 0.689]	-202.257	-60.436	7.689	14.199
3	[0.309, -1.571, 3.174, -4.197, 1.646, 0.559, 1.421]	-205.205	-63.201	7.321	13.942
4	[0.309, -1.499, 2.905, -4.135, 1.383, 1.095, 1.614]	-207.030	-65.281	7.031	13.719
10	[0.274, -1.807, 2.314, -4.629, 1.833, 2.271, 2.791]	-211.989	-72.189	5.968	12.783
20	[-0.127, -2.482, 1.787, -4.658, 2.596, 3.419, 3.682]	-214.793	-77.223	5.055	11.847
100	[-0.684, -3.112, 1.899, -3.453, 4.848, 5.815, 5.265]	-217.076	-83.942	3.481	8.920
200	[-0.249, -2.302, 2.983, -2.255, 5.983, 6.698, 5.750]	-217.506	-86.763	2.544	6.521
300	[0.086, -1.683, 3.794, -1.380, 6.794, 7.317, 6.086]	-217.736	-88.270	1.860	4.767
400	[0.332, -1.230, 4.387, -0.740, 7.387, 7.770, 6.332]	-217.859	-89.075	1.360	3.485
500	[0.512, -0.899, 4.821, -0.272, 7.821, 8.101, 6.512]	-217.925	-89.506	0.994	2.548
600	[0.643, -0.658, 5.138, 0.070, 8.138, 8.342, 6.643]	-217.960	-89.736	0.727	1.863

Beispiel 4.2

Sei

$$A(n, n) = \begin{pmatrix} 2 & -1 & 0 & \dots & 0 & 1 \\ -1 & 4 & -1 & \dots & 0 & 0 \\ 0 & -1 & 4 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 4 & -1 \\ 1 & 0 & 0 & \dots & -1 & 2 \end{pmatrix}, \quad b = \begin{pmatrix} n \\ 4 \\ 6 \\ \dots \\ 2(n-1) \\ n+2 \end{pmatrix}.$$

Damit ist $x^* = A^{-1}b = (1, 2, \dots, n)^T$.

Wir berechnen den Fehler $\|e^{(m)}\|_\infty = \|x^* - x^{(m)}\|_\infty$ und vergleichen diesen zwischen GV für $Ax = b$, Orthomin(0) und GV für $A^T Ax = A^T b$ für ausgewählte Schritte m (Rechnung mit 16 Mantissenstellen).

Bei $n = 10$ gilt $\text{cond}_2(A) = 8.835\dots$ und $\text{cond}_2(A^T A) = 78.070\dots$ sowie bei $n = 100$ $\text{cond}_2(A) = 8.998\dots$ und $\text{cond}_2(A^T A) = 80.973\dots$

Der Startvektor für GV und Orthomin(0) ist der Nullvektor.

m	n=10			n=100		
	GV Ax=b	Orthomin(0)	GV A'Ax=A'b	GV Ax=b	Orthomin(0)	GV A'Ax=A'b
0	1.000e+01	1.000e+01	1.000e+01	1.000e+02	1.000e+02	1.000e+02
1	4.202e+00	4.436e+00	7.235e+00	4.936e+01	4.971e+01	7.544e+01
2	2.946e+00	3.333e+00	6.351e+00	3.518e+01	3.805e+01	6.462e+01
3	2.223e+00	2.439e+00	5.821e+00	2.670e+01	2.896e+01	6.096e+01
4	1.714e+00	1.933e+00	5.474e+00	2.072e+01	2.285e+01	5.790e+01
5	1.332e+00	1.478e+00	5.149e+00	1.606e+01	1.786e+01	5.589e+01
6	1.053e+00	1.191e+00	4.918e+00	1.269e+01	1.419e+01	5.393e+01
7	8.221e-01	9.207e-01	4.675e+00	9.837e+00	1.113e+01	5.233e+01
8	6.583e-01	7.465e-01	4.502e+00	7.850e+00	8.885e+00	5.078e+01
9	5.130e-01	5.781e-01	4.309e+00	6.079e+00	6.979e+00	4.937e+01
10	4.134e-01	4.698e-01	4.171e+00	4.880e+00	5.586e+00	4.802e+01
20	4.043e-02	4.657e-02	3.077e+00	4.683e-01	5.647e-01	3.699e+01
30	3.949e-03	4.616e-03	2.353e+00	4.599e-02	5.826e-02	2.875e+01
40	3.858e-04	4.575e-04	1.806e+00	4.558e-03	6.062e-03	2.238e+01
50	3.769e-05	4.535e-05	1.386e+00	4.541e-04	6.339e-04	1.744e+01
60	3.681e-06	4.495e-06	1.064e+00	4.539e-05	6.648e-05	1.359e+01
70	3.596e-07	4.455e-07	8.173e-01	4.547e-06	6.987e-06	1.059e+01
80	3.513e-08	4.416e-08	6.275e-01	4.562e-07	7.355e-07	8.259e+00
90	3.432e-09	4.377e-09	4.818e-01	4.582e-08	7.751e-08	6.440e+00
100	3.353e-10	4.338e-10	3.699e-01	4.607e-09	8.177e-09	5.022e+00
200			2.634e-02			4.193e-01
300			1.875e-03			3.511e-02
400			1.335e-04			2.944e-03

Beispiel 4.3

Sei A 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 $Q(x^*) = R(x^*) = 0$.
Im Orthomin(0) werden wie im GV sehr viele Schritte ausgeführt.

Die Schritte des Orthomin(0) sind wie folgt.

$$\begin{aligned} x^{(0)} &= \left(\frac{9}{2}, 3\right)^T && \text{Startvektor,} \\ r^{(0)} &= b - Ax^{(0)} = \left(-\frac{9}{2}, -6\right)^T && \text{Anfangsresiduum, Such- und Abstiegsrichtung.} \end{aligned}$$

S1

$$\begin{aligned} w^{(0)} &= Ar^{(0)} = \left(-\frac{9}{2}, -12\right)^T, \\ \alpha_0 &= \frac{w^{(0)T}r^{(0)}}{\|w^{(0)}\|_2^2} = \frac{41}{73} = 0.561\dots, \\ x^{(1)} &= x^{(0)} + \alpha_0 r^{(0)} = \frac{9}{73}(16, -3)^T, \\ r^{(1)} &= b - Ax^{(1)} = -Ax^{(1)} = r^{(0)} - \alpha_0 Ar^{(0)} = \frac{18}{73}(-8, 3)^T. \end{aligned}$$

S2

$$\begin{aligned} w^{(1)} &= \frac{18}{73}(-8, 6)^T, \quad \alpha_1 = \frac{41}{50} = 0.82, \\ x^{(2)} &= \frac{216}{1825}(3, 2)^T, \quad r^{(2)} = -\frac{216}{1825}(3, 4)^T. \end{aligned}$$

S3

$$\begin{aligned} w^{(2)} &= -\frac{216}{1825}(3, 8)^T, \quad \alpha_2 = \frac{41}{73}, \\ x^{(3)} &= \frac{1\,296}{133\,225}(16, -3)^T, \quad r^{(3)} = \frac{2\,592}{133\,225}(-8, 3)^T. \end{aligned}$$

S4

$$\begin{aligned} w^{(3)} &= \frac{5\,184}{133\,225}(-4, 3)^T, \quad \alpha_3 = \frac{41}{50}, \\ x^{(4)} &= \frac{31\,104}{3\,330\,625}(3, 2)^T, \quad r^{(4)} = -\frac{31\,104}{3\,330\,625}(3, 4)^T. \end{aligned}$$

Dazu erhalten wir die Beziehungen

$$\begin{aligned} r^{(0)} \parallel r^{(2)} \parallel r^{(4)} \parallel \dots, \quad r^{(1)} \parallel r^{(3)} \parallel r^{(5)} \parallel \dots, \\ r^{(1)} \perp Ar^{(0)}, \quad r^{(2)} \perp Ar^{(1)}, \quad r^{(2)} \not\perp Ar^{(0)}, \dots \end{aligned}$$

Der Iterationsverlauf ist wie eine ‘‘Zick-Zack‘‘-Kurve.

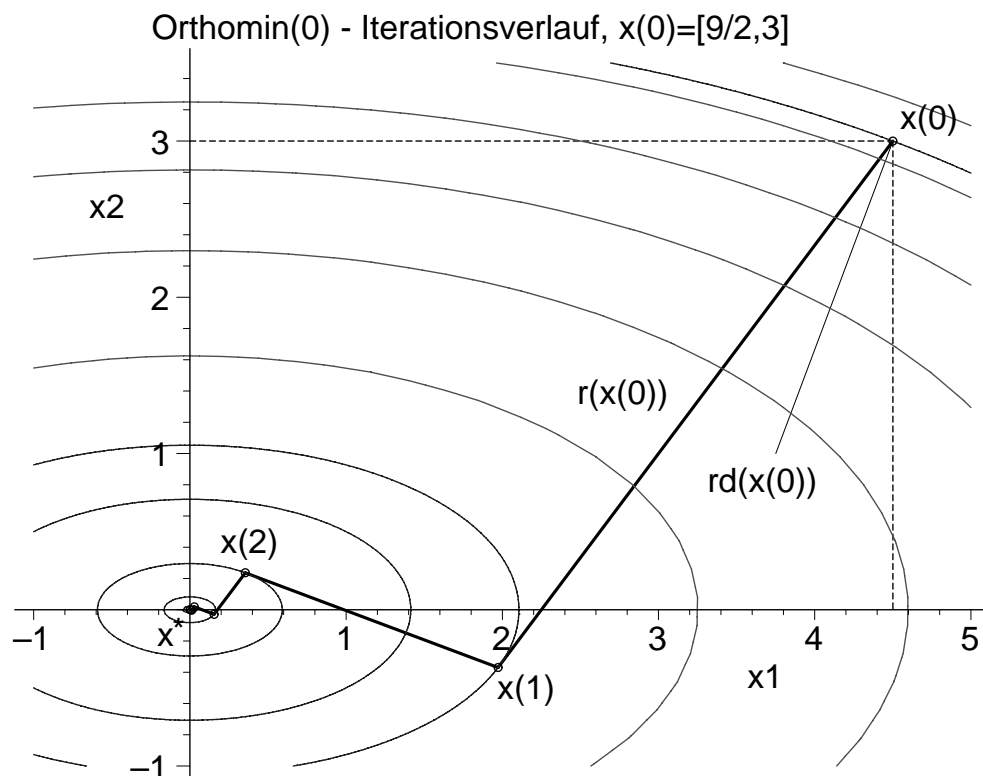


Abb. 4.1 Datei *abst301.ps*

Höhenlinienbild mit Iterationsverlauf $x^{(m)}$ von Orthomin(0) zu $R(x) = \frac{1}{2}x_1^2 + 2x_2^2$ mit Abstiegsrichtung $r(x)$, $r(x^{(0)})$, $\hat{r}(x^{(0)}) = A^T r(x^{(0)})$,
`contours=[28.125,2.219,1,0.1751,0.0138,0.00109,0]` und `contours=6`

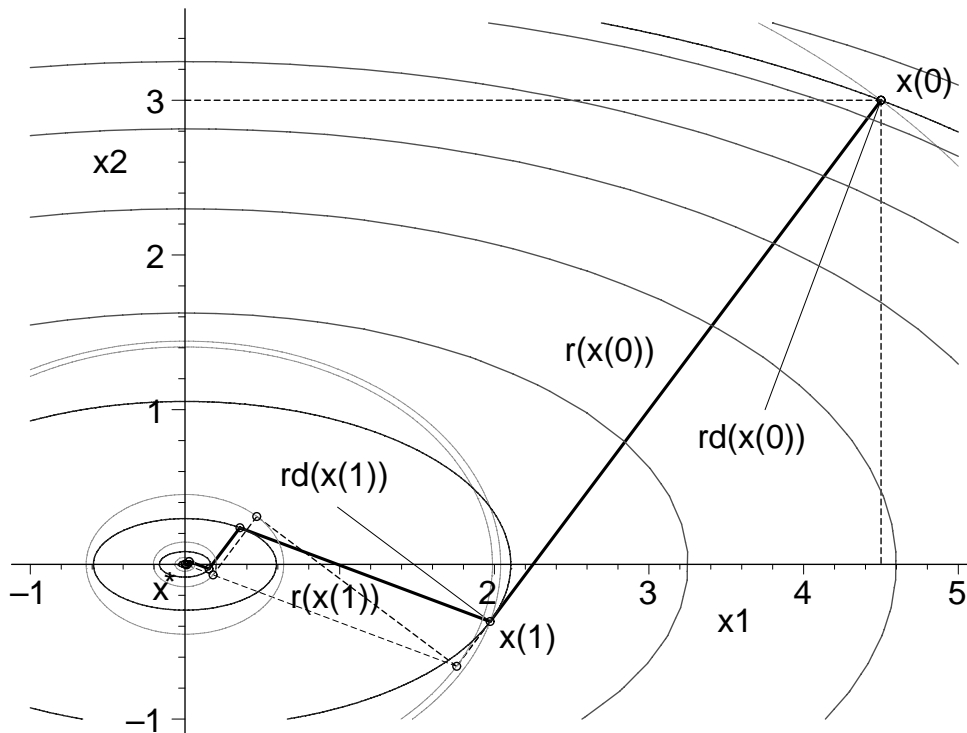
Für die Such- und Abstiegsrichtung gilt $r(x) = b - Ax \perp Q(x) = \text{const}$, $\nabla R(x) = \text{const}$, dagegen gilt Richtung $\hat{r}(x) = A^T r(x) = A^T(b - Ax) \perp R(x) = \text{const}$, was für den Vektor $\hat{r}(x^{(0)}) = \text{rd}(x^{(0)})$ in der Abbildung zu erkennen ist.

Berechnungen mit Maple und Ergebnisse $x^{(m)}$, $\|r^{(m)}\|_2^2$, $m = 0, 1, \dots, 7$

```

x(0): [4.5,          3          ] r(0)'r(0)=56.25
x(1): [1.9726027397260270, -0.36986301369863010 ] r(1)'r(1)= 4.438...
x(2): [0.35506849315068490,  0.23671232876712330 ] r(2)'r(2)= 0.350...
x(3): [0.15564646275098520, -0.029183711765809720 ] r(3)'r(3)= 0.027...
x(4): [0.028016363295177330,  0.018677575530118220 ] r(4)'r(4)= 0.002180...
x(5): [0.012281145554050340, -0.0023027147913844380] r(5)'r(5)= 0.000172...
x(6): [0.0022106061997290610, 0.0014737374664860400] r(6)'r(6)= 0.000013...
x(7): [0.0009690328546757526, -0.0001816936602517036] r(7)'r(7)= 0.000001...

```

Orthomin(0) - Iterationsverlauf, $x(0)=[9/2,3]$, Vergleich mit GV,CG**Abb. 4.2** Datei *abst302.ps*

Höhenlinienbild mit Iterationsverlauf $x^{(m)}$ von Orthomin(0) zu $R(x) = \frac{1}{2}x_1^2 + 2x_2^2$ im Vergleich mit GV (dick gestrichelt) und CG (dünn gestrichelt),

Orthomin(0): contours=[28.125,2.219,1,0.1751,0.0138,0.00109,0] und contours=6,

GV: contours=[19.125,1.976,0.204,0.021,0.0022,0],

zusätzliche Höhenlinie $Q(x) = Q(x^{(1)}) = 2.082379\dots$

Die ellipsenförmigen Höhenlinien zum Funktional $R(x)$ sind im Vergleich zu denen von $Q(x)$ lang gestreckter.

Die Such- und Abstiegsrichtung $r(x)$ vom Funktional $Q(x)$ (orthogonal zu den Höhenlinien von $Q(x)$) sind günstiger in der Strahlenminimierung als die zu den Höhenlinien von $R(x)$ orthogonalen Richtungen $\hat{r}(x)$. Deshalb wird in Orthomin(0) als Suchrichtung $r(x)$ genommen. Die ersten beiden Richtungen $r(x^{(0)})$, $r(x^{(1)})$ sowie $\hat{r}(x^{(0)})$, $\hat{r}(x^{(1)})$ in der Abbildung machen dies schon deutlich.

Die Nachbarschaft der Iterationsfolgen von Orthomin(0) und des GV weisen auf ein ähnliches Konvergenzverhalten hin. Das GV für das Normalgleichungssystem, also mit dem Funktional $R(x)$, weicht davon deutlich ab, denn bei gleichem Startvektor ist schon seine erste Iterierte $(\frac{864}{265}, -\frac{81}{265})^T = (3.260\dots, -0.305\dots)^T$ und ihr Residuumquadrat gleich 12.124...

CG ist im ersten Schritt wie GV und bei exakter Rechnung mit seiner zweiten Iterierten an der Lösung x^* .

Beispiel 4.4

Sei A eine Tridiagonalmatrix.

$$\begin{pmatrix} 4 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 6 \\ 2 \end{pmatrix}, \quad x^* = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \quad x^{(0)} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Die Funktionale sind

$$\begin{aligned} Q(x) &= 2x_1^2 - x_1x_2 + 2x_2^2 - x_2x_3 + 2x_3^2 - 2x_1 - 6x_2 - 2x_3, \\ R(x) &= \frac{17}{2}x_1^2 - 8x_1x_2 + x_1x_3 + 9x_2^2 - 8x_2x_3 + \frac{17}{2}x_3^2 - 2x_1 - 20x_2 - 2x_3. \end{aligned}$$

Am gemeinsamen eindeutigen Minimum an der Stelle x^* gilt $Q(x^*) = -8$ und $R(x^*) = -22$.

Die ersten Schritte von Orthomin(0) sind wie folgt.

$$\begin{aligned} x^{(0)} &= (0, 0, 0)^T && \text{Startvektor,} \\ r^{(0)} &= b - Ax^{(0)} = b = 2(1, 3, 1)^T && \text{Anfangsresiduum, Abstiegsrichtung.} \end{aligned}$$

S1

$$\begin{aligned} \alpha_0 &= \frac{\|r^{(0)}\|_A^2}{\|Ar^{(0)}\|_2^2} = \frac{16}{51} = 0.313\,725\dots \\ x^{(1)} &= x^{(0)} + \alpha_0 r^{(0)} = \frac{32}{51}(1, 3, 1)^T, \\ r^{(1)} &= b - Ax^{(1)} = r^{(0)} - \alpha_0 Ar^{(0)} = \frac{14}{51}(5, -1, 5)^T. \end{aligned}$$

S2

$$\begin{aligned} \alpha_1 &= \frac{16}{77} = 0.207\,792\dots, \\ x^{(2)} &= \frac{512}{561}(1, 2, 1)^T, \quad r^{(2)} = \frac{512}{561}(1, 3, 1)^T. \end{aligned}$$

S3

$$\begin{aligned} \alpha_2 &= \frac{16}{51} = 0.313\,725\dots, \\ x^{(3)} &= \frac{32}{28\,611}(865, 1\,779, 865)^T, \quad r^{(3)} = \frac{686}{28\,611}(5, -1, 5)^T. \end{aligned}$$

S4

$$\begin{aligned} \alpha_3 &= \frac{16}{77} = 0.207\,792\dots, \\ x^{(4)} &= \frac{312\,320}{314\,721}(1, 2, 1)^T, \quad r^{(4)} = \frac{4\,802}{314\,721}(1, 3, 1)^T. \end{aligned}$$

Startvektor $x^{(0)} = (x_1^{(0)}, x_2^{(0)}, x_3^{(0)})^T$ und Iterierte $x^{(k)}$, $k = 1, 2, \dots, 10$

k	x(k) [1],	x(k) [2],	x(k) [3]
0	[0,	0,	0]
1	[0.6274509803921568,	1.882352941176470,	0.6274509803921568]
2	[0.9126559714795008,	1.825311942959001,	0.9126559714795008]
3	[0.9674600678060884,	1.989724231938764,	0.9674600678060884]
4	[0.9923710206818101,	1.984742041363620,	0.9923710206818101]
5	[0.9971578312343998,	1.999102473021389,	0.9971578312343998]
6	[0.9993336542128496,	1.998667308425699,	0.9993336542128496]
7	[0.9997517535302772,	1.999921606377982,	0.9997517535302772]
8	[0.9999417986745625,	1.999883597349125,	0.9999417986745625]
9	[0.9999783171532683,	1.999993152785242,	0.9999783171532683]
10	[0.9999949164617709,	1.999989832923541,	0.9999949164617709]

Vergleich mit $x(10)$ von GV

10 [0.9999995108519213, 1.999999021703844, 0.9999995108519213]

Dazu erhalten wir die Beziehungen

$$Ar^{(0)} \perp r^{(1)}, Ar^{(1)} \perp r^{(2)}, Ar^{(0)} \not\perp r^{(2)}, Ar^{(2)} \perp r^{(3)}, \dots,$$

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

$$\begin{aligned} \mathcal{K}_2(A, r^{(0)}) &= \text{span}\{r^{(0)}, r^{(1)}\} = \text{span} \left\{ \begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix}, \begin{pmatrix} 5 \\ -1 \\ 5 \end{pmatrix} \right\} \\ &= \text{span}\{r^{(0)}, Ar^{(0)}\} = \text{span} \left\{ \begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 10 \\ 1 \end{pmatrix} \right\} \subset \mathbb{R}^3, \end{aligned}$$

$$x^* = \frac{77}{224}r^{(0)} + \frac{51}{224}r^{(1)} \in \text{span}\{r^{(0)}, r^{(1)}\}.$$

Orthomin(0) konvergiert wie das GV nur langsam.

Bei den Abstiegsrichtungen gibt es keine und braucht man auch keine dritte linear unabhängige Richtung.

Beispiel 4.5

Wir nehmen das LGS aus Beispiel 3.7 mit $A = I$, $b = (0, 0, 1)^T = x^*$ und dem Startvektor $x^{(0)} = (0, 0, 0)^T$.

Der Ablauf des Iterationsprozesses von Orthomin(0) ist wie beim CG (Beispiel 3.14) sehr kurz, denn mit

$$x^{(0)} = 0,$$

$$r^{(0)} = b - Ax^{(0)} = b$$

folgt

$$\begin{aligned} w^{(0)} &= Ar^{(0)} = r^{(0)} = b, \\ \alpha_0 &= \frac{w^{(0)T}r^{(0)}}{\|w^{(0)}\|_2^2} = 1, \\ x^{(1)} &= x^{(0)} + \alpha_0 r^{(0)} = r^{(0)} = b, \\ r^{(1)} &= b - Ax^{(1)} = r^{(0)} - \alpha_0 Ar^{(0)} = 0 \quad \text{und vorzeitiger Stopp mit } x^{(1)} = x^*. \end{aligned}$$

Beispiel 4.6

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

$$Q(x) = x_1^2 + x_1x_2 + \frac{3}{2}x_2^2 - x_1 - 2x_2, \quad R(x) = \frac{5}{2}x_1^2 + 5x_1x_2 + 5x_2^2 - 4x_1 - 7x_2.$$

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

Der erste Schritt von Orthomin(0) ist wie folgt.

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

S1

$$\begin{aligned} w^{(0)} &= Ar^{(0)} = -\frac{1}{2}(17, 21)^T, \\ \alpha_0 &= \frac{w^{(0)T}r^{(0)}}{\|w^{(0)}\|_2^2} = \frac{207}{730} = 0.283561\dots \\ x^{(1)} &= x^{(0)} + \alpha_0 r^{(0)} = \frac{1}{1460}(948, 425)^T, \\ r^{(1)} &= b - Ax^{(1)} = r^{(0)} - \alpha_0 Ar^{(0)} = \frac{41}{1460}(-21, 17)^T. \end{aligned}$$

Orthomin(0) - Iterationsverlauf mit Funktionalen und verschiedenen Fehlern

m	[x(m) [1], x(m) [2]]	R(x(m))	Q(x(m))	e(m) _A	e(m) _2
0	[1.5000000000, 1.0000000000]	5.125000000000	1.750000000000	2.213594362118	1.360147050874
1	[0.6493150685, 0.2910958904]	-2.212157534247	-0.493778617001	0.642217070777	0.545257535203
2	[0.2490747810, 0.6150999326]	-2.489133995398	-0.696508628029	0.083562814344	0.051345322087
3	[0.2169615682, 0.5883389219]	-2.499589810156	-0.699706124263	0.024243586256	0.020583380118
4	[0.2018525647, 0.6005700199]	-2.499984515402	-0.699995024621	0.003154482167	0.001938277261
5	[0.2006402963, 0.5995597963]	-2.499999415459	-0.699999581212	0.000915191298	0.000777019133
6	[0.2000699340, 0.6000215182]	-2.49999997934	-0.69999992910	0.000119081171	0.000073169640

7	[0.2000241711, 0.5999833824]	-2.499999999167	-0.699999999403	0.000034548317	0.000029332341
8	[0.2000026400, 0.600008123]	-2.499999999969	-0.699999999990	0.000004495294	0.000002762142
9	[0.2000009125, 0.5999993727]	-2.499999999999	-0.699999999999	0.000001304193	0.000001107291
10	[0.2000000997, 0.6000000307]	-2.500000000000	-0.700000000000	0.000000169697	0.000000104270

Vergleich mit $x(10)$ von GV

10	[0.2000050793, 0.6000015628]	-2.499999999884	-0.699999999963	0.000008648773	0.000015257781
----	------------------------------	-----------------	-----------------	----------------	----------------

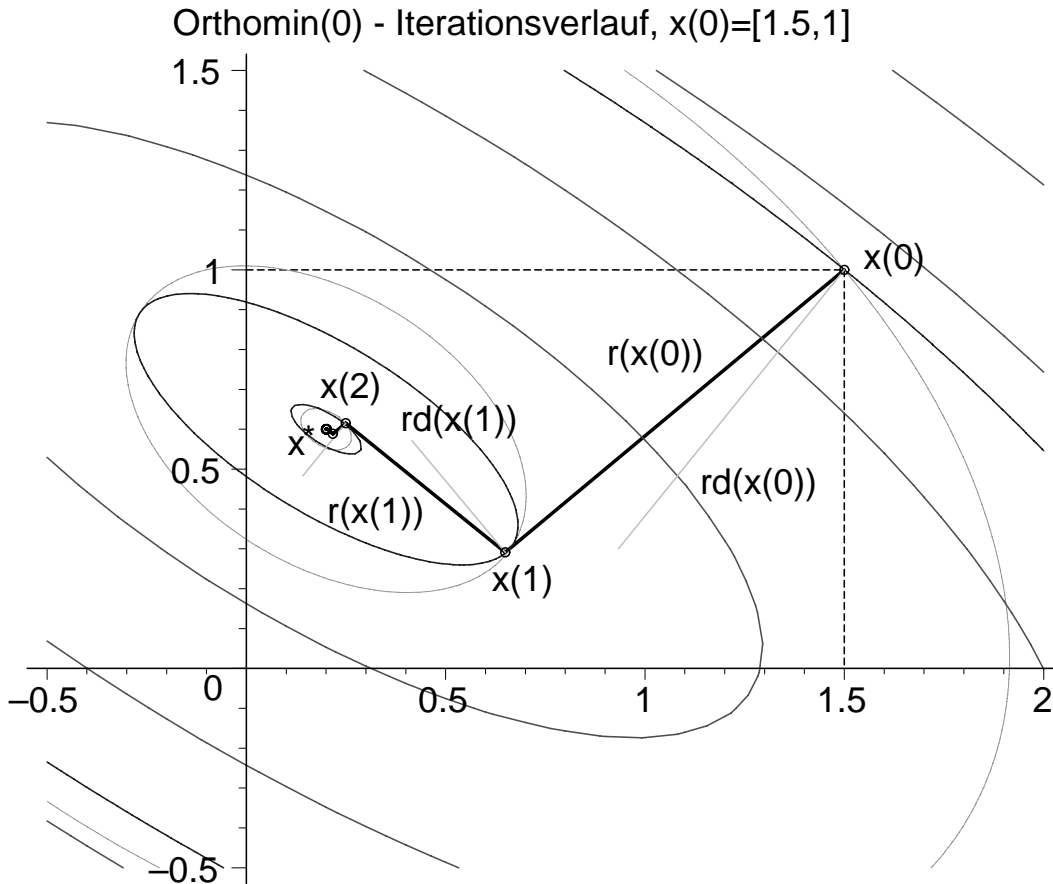


Abb. 4.3 Datei *abst303.ps*

Höhenlinienbild mit Iterationsverlauf $x^{(m)}$ von Orthomin(0)

zu $R(x) = \frac{5}{2}x_1^2 + 5x_1x_2 + 5x_2^2 - 4x_1 - 7x_2$ mit Abstiegsrichtung $r(x)$,

$r(x^{(0)}), r(x^{(1)}), \hat{r}(x^{(0)}) = A^T r(x^{(0)}), \hat{r}(x^{(1)}) = A^T r(x^{(1)})$,

mit `contours`=[-1,2,7,13,20, 5.13,-2.21,-2.49,-2.5],

dazu $Q(x) = x_1^2 + x_1x_2 + \frac{3}{2}x_2^2 - x_1 - 2x_2$ mit `contours`=[1.75,-0.49,-0.6965]

In der obigen Abbildung mit den Höhenlinien der Funktionale $R(x)$ und $Q(x)$ erkennen wir erneut, dass die Such- und Abstiegsrichtungen $r(x^{(m)}) = \mathbf{r}(x^{(m)})$ in der Strahlenminimierung von Orthomin(0) orthogonal zu den Höhenlinien von $Q(x)$, die anderen auch eingezeichneten Richtungsvektoren $\hat{r}(x^{(m)}) = \mathbf{rd}(x^{(m)})$ orthogonal zu den Höhenlinien von $R(x)$ sind.

Orthomin(0) für das LGS $Bx = A^T Ax = A^T b = c$ hat ähnliches Verhalten, der Iterationsverlauf ist wegen der Ausgeglichenheit der Elemente der Matrix $B = A^T A$ sogar etwas günstiger.

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

Orthomin(0) fuer A'Ax=A'b - Iterationsverlauf mit Funktionalen und verschiedenen Fehlern

m	[x(m) [1], x(m) [2]]	R(x(m))	Q(x(m))	e(m) _B	e(m) _2
0	[1.5000000000, 1.0000000000]	58.750000000000	5.125000000000	3.905124837953	1.360147050874
1	[0.8493401015, 0.1962436548]	-31.433375634518	-1.941673580870	1.056717955871	0.764631776555
2	[0.2151957444, 0.6046756137]	-32.487532191375	-2.498958169416	0.045647137561	0.015898805316
3	[0.2075901586, 0.5952804783]	-32.499854263359	-2.499923713897	0.012352012265	0.008937806942
4	[0.2001776236, 0.6000546534]	-32.499998296479	-2.499999857651	0.000533570949	0.000185841678
5	[0.2000887216, 0.5999448333]	-32.499999980087	-2.499999989577	0.000144383093	0.000104474331
6	[0.2000020762, 0.6000006388]	-32.49999999767	-2.499999999981	0.000006236929	0.000002172310
7	[0.2000010371, 0.5999993552]	-32.49999999997	-2.499999999999	0.000001687699	0.000001221204
8	[0.2000000243, 0.6000000075]	-32.500000000000	-2.500000000000	0.00000072904	0.00000025392
9	[0.2000000121, 0.5999999925]	-32.500000000000	-2.500000000000	0.000000019728	0.000000014275
10	[0.2000000003, 0.6000000001]	-32.500000000000	-2.500000000000	0.000000000852	0.000000000297

Beispiel 4.7

Gegeben sei das LGS aus den Beispielen 1.7, 3.6, 3.16 mit der regulären, aber nicht definiten Matrix A .

Wir nehmen hier Bezug auf einige Eigenschaften des Systems.

$$\begin{pmatrix} 0.780 & 0.563 \\ 0.913 & 0.659 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0.217 \\ 0.254 \end{pmatrix}, \quad x^* = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Die Funktionale sind

$$\begin{aligned} R(x) &= 0.5(1.441\,969x_1 + 1.040\,807x_2)x_1 + 0.5(1.040\,807x_1 + 0.751\,250x_2)x_2 \\ &\quad - 0.401\,162x_1 - 0.289\,557x_2 \\ &= 0.720\,984\,5x_1^2 + 1.040\,807x_1x_2 + 0.375\,625x_2^2 - 0.401\,162x_1 - 0.289\,557x_2, \\ Q(x) &= 0.5(0.780x_1 + 0.913x_2)x_1 + 0.5(0.563x_1 + 0.659x_2)x_2 - 0.217x_1 - 0.254x_2 \\ &= 0.390x_1^2 + 0.738x_1x_2 + 0.3295x_2^2 - 0.217x_1 - 0.254x_2. \end{aligned}$$

Das Minimum von $R(x)$ ist bei der Lösung $x^* = (1, -1)^T$ und $R(x^*) = -0.055\,802\,5$ sowie $Q(x^*) = 0.018\,5$.

Aber $Q(x)$ hat die Oberfläche eines Sattels mit dem Sattelpunkt bei

$$z = \left(\frac{44\,449}{30\,624}, -\frac{6\,329}{5\,104} \right)^T = (1.451\,443\,312\,434\,691\,7, -1.240\,007\,836\,990\,595\,6)^T,$$

$$Q(z) = -\frac{37}{61\,248\,000} = -0.604\,101\,358\,411\,702\,9\,10^{-6},$$

und keine Minimumstelle.

Die Durchführung von Orthomin(0) mit der Suchrichtung $r(x)$ von einer Satteloberfläche $Q(x)$ birgt natürlich Probleme in sich. Man sieht das am Iterationsverlauf. Egal, ob man in Maple exakt arbeitet oder wie genau man in Maple numerisch rechnet, das Verfahren erzielt nach der 4. Iterierten keine Verbesserungen mehr, obwohl man noch nicht an der Lösung ist.

Orthomin(0) - Iterationsverlauf mit Funktionalen und Fehler

m	[x(m)[1], r(m)[1],	x(m)[2] r(m)[2]	R(x(m))	Q(x(m)) e(m) _2
0	[1.200000000000000000, -1.200000000000000000]		-0.053570400000000000	0.017760000000000000
	[-4.3400000000000000e-02, -5.0800000000000000e-02]			0.2828427124746190098
1	[1.16984017373116808048, -1.23530228512572952791]		-0.0558024999999808015	0.0299558971206173798
	[-1.4898452537856358e-07, 1.2728129930141152e-07]			0.2901945037357245874
2	[1.16984017373032291028, -1.23530228512500747734]		-0.0558024999999808015	0.0299558971206522670
	[-1.4898427266027551e-07, 1.2728159511048197e-07]			0.2901945037346444705
3	[1.16984017373032290338, -1.23530228512500747144]		-0.0558024999999808015	0.0299558971206522673
	[-1.4898427265821248e-07, 1.2728159511289676e-07]			0.2901945037346444617
4	[1.16984017373032290338, -1.23530228512500747144]		-0.0558024999999808015	0.0299558971206522673
	[-1.4898427265821247e-07, 1.2728159511289678e-07]			0.2901945037346444617
5	[1.16984017373032290338, -1.23530228512500747144]		-0.0558024999999808015	0.0299558971206522673
	[-1.4898427265821247e-07, 1.2728159511289678e-07]			0.2901945037346444617

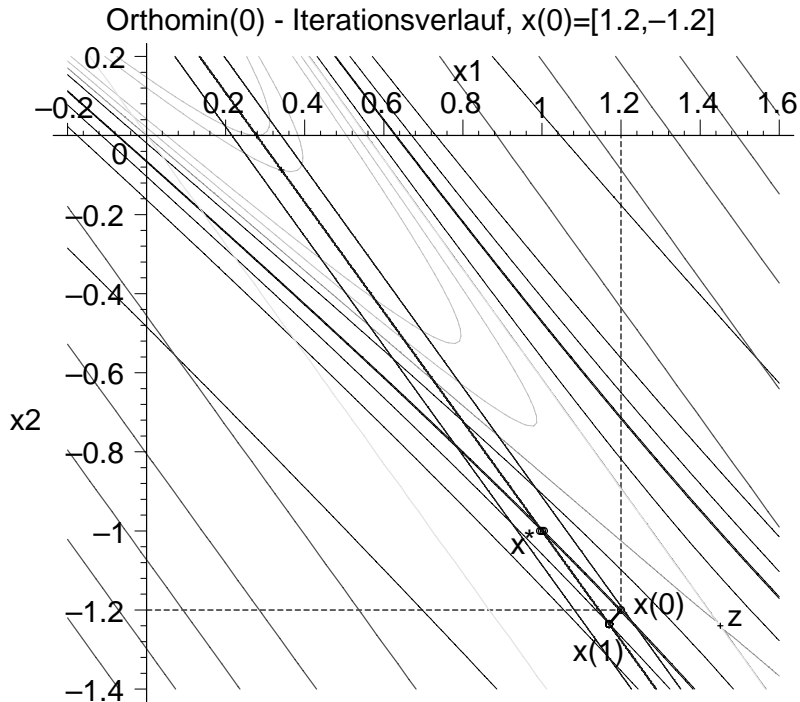


Abb. 4.4 Datei *abst304.ps*

Höhenlinienbild mit Iterationsverlauf $x^{(m)}$ von Orthomin(0)

zu $R(x) = 0.7209845x_1^2 + 1.040807x_1x_2 + 0.375625x_2^2 - 0.401162x_1 - 0.289557x_2$
mit $\text{contours}=[0, -0.05357, -0.0558]$ und $\text{contours}=5$,

dazu $Q(x) = 0.390x_1^2 + 0.738x_1x_2 + 0.3295x_2^2 - 0.217x_1 - 0.254x_2$ mit $\text{contours}=[-0.03018, -0.0259, -0.01, -0.005, 0, 0.005, 0.01776, 0.0185, 0.0299, 0.05, 0.2]$

Die Abstiegsrichtungen $r^{(m)}$ stehen senkrecht auf den Höhenlinien von $Q(x)$, damit fast senkrecht auf der Tallinie von $R(x)$ und sind fast parallel zu Tangenten an die Höhenlinien von $R(x)$. Damit sind aufeinander folgende $r^{(m)}$ "numerisch" parallel und es findet keine Fortbewegung in der Iterationsfolge mehr statt.

Um Orthomin(0) zu verwenden, muss man das LGS symmetrisieren.

$$Bx = A^T Ax = \begin{pmatrix} 1.441\,969 & 1.040\,807 \\ 1.040\,807 & 0.751\,250 \end{pmatrix} x = \begin{pmatrix} 0.401\,162 \\ 0.289\,557 \end{pmatrix} = A^T b = c.$$

Die Matrix B hat eine sehr schlechte Kondition, ihre spektrale Konditionszahl ist ungefähr $0.5 \cdot 10^{13}$. Darüber hinaus verwenden wir im Algorithmus auch $B^T B$.

Wir testen Orthomin(0) mit dem Startvektor $x^{(0)} = (1.2, -1.2)^T$.

In Maple wird erst ab `Digits:=27, 28, ...` die Konvergenz gegen $x^* = (1, -1)^T$ sichtbar.

Orthomin(0) - Iterationsverlauf mit Funktionalen und Fehler

m	[x(m) [1], r(m) [1],	x(m) [2] r(m) [2]	R(x(m))	Q(x(m)) e(m) _2
0	[1.20000000000000000000, -1.20000000000000000000] [-8.0232400000000000e-02, -5.7911400000000000e-02]		-0.1174916191166400000	-0.0535704000000000000 0.2828427124746190098
1	[1.16341797148391625296, -1.22640475027802150139] [-7.4510557989854300e-14, 1.0322943138741210e-13]		-0.1223871032465000000	-0.0558024999999822260 0.2792213178687675751
2	[1.00000001984505180306, -1.00000002749400180761] [3.5964205282060283e-17, 2.5977875119946671e-17]		-0.1223871032465000000	-0.0558025000000000000 0.0000000339079078751
3	[1.00000001984505181946, -1.00000002749400179577] [-9.0483676436093602e-21, 1.2535912654972169e-20]		-0.1223871032465000000	-0.0558025000000000000 0.0000000339079078751
4	[1.0000000000000000239445, -1.0000000000000000331735] [4.3673988884979544e-24, 3.1546851106156313e-24]		-0.1223871032465000000	-0.0558025000000000000 0.00000000000000040912
5	[1.0000000000000000239445, -1.0000000000000000331735] [-1.0988099877061574e-27, 1.5223282886857590e-27]		-0.1223871032465000000	-0.0558025000000000000 0.00000000000000040912

4.2.2 Verfahren der A-orthogonalen Residua als Projektionsmethode

Betrachten wir Orthomin(0) als AV im Sinne einer Projektionsmethode, so haben wir mit der Suchrichtung $r(x)$ die Beziehungen

$$\begin{aligned} x^{(m)} &= x^{(m)} + \alpha_{m-1} r^{(m-1)} \in x^{(m)} + K, \quad K = \text{span}\{r^{(m-1)}\}, \\ x^{(m)} &\in x^{(0)} + K_m = x^{(0)} + \mathcal{R}_m = x^{(0)} + \text{span}\{r^{(0)}, \dots, r^{(m-1)}\}, \\ r^{(m)} &\perp AL_m = \text{span}\{Ar^{(m-1)}\} = L, \\ r^{(m)} &\perp Ar^{(m-1)}, \quad r^{(m)} \notin K_m. \end{aligned}$$

Orthomin(0) stellt zumindest in jedem Schritt eine schiefe Projektionsmethode mit $K = \text{span}\{r^{(m-1)}\}$ und $L = \text{span}\{Ar^{(m-1)}\}$ dar.

Wie beim GV ist das natürlich recht wenig für ein gutes Konvergenzverhalten.

4.3 Verfahren der konjugierten Residua

Man bezeichnet das Verfahren auch mit **CR** (conjugate residual method).

Zunächst ist die Vorgehensweise wie bei Orthomin(0), mit dem Unterschied, dass die Suchrichtungen $p(x^{(m)})$ gemäß der Forderung der generellen A-Orthogonalität der Residua $r(x^{(m)})$ gewählt werden, d. h.

$$r^{(k)T} Ar^{(j)} = 0 \quad \forall k \neq j, \quad r^{(k)} = r(x^{(k)}). \quad (4.12)$$

Dafür ist mindestens die Symmetrie der Matrix A vorauszusetzen.

Falls $r^{(k)} \neq 0$ und $r^{(k)T} Ar^{(k)} \neq 0$ sind, bilden die Vektoren $r^{(0)}, r^{(1)}, \dots, r^{(m)}$ nach Satz 3.8 ein System linear unabhängiger Vektoren und spätestens gilt

$$\text{span}\{r^{(0)}, r^{(1)}, \dots, r^{(n-1)}\} = \mathbb{R}^n.$$

Aus $r^{(n)T} Ar^{(j)} = r^{(j)T} Ar^{(n)} = 0$, $j = 0, 1, \dots, n-1$, folgt für $r^{(n)}$

$$\begin{aligned} Ar^{(n)} &= 0, \\ r^{(n)} &= 0. \end{aligned}$$

Damit erhält man die Endlichkeit des CR.

Analogien zum CG sowie seine Anwendung auf das Normalgleichungssystem als CGNR (conjugate gradient normal residual) sind insofern möglich, wenn man die Beziehungen

$$\begin{aligned} Ax = b &\Rightarrow r^{(k)} \perp r^{(j)} \quad \forall k \neq j, \\ A^T Ax = A^T b &\Rightarrow r^{(k)} \perp Ar^{(j)} \quad \forall k \neq j \end{aligned}$$

sieht.

Mit der Strahlenminimierung (4.7) ergeben sich für CR die Formeln

$$\alpha_m = \frac{r^{(m)T} Ap^{(m)}}{(Ap^{(m)})^T Ap^{(m)}} = \frac{r^{(m)T} Ap^{(m)}}{\|Ap^{(m)}\|_2^2} \stackrel{A=A^T}{=} \frac{p^{(m)T} Ar^{(m)}}{\|Ap^{(m)}\|_2^2}. \quad (4.13)$$

Der Iterationsschritt als Aufdatierungsformel ist

$$x^{(m+1)} = x^{(m)} + \alpha_m p^{(m)}. \quad (4.14)$$

Die Iterierten $r^{(m+1)}$ kann man wieder rekursiv bestimmen. Unter Verwendung der Kürzel (3.14) bekommt man

$$\begin{aligned} r^{(m+1)} &= b - Ax^{(m+1)} \quad (\text{direkte Berechnung}), \\ &= b - A(x^{(m)} + \alpha_m p^{(m)}) \\ &= b - (v^{(m)} + \alpha_m w^{(m)}), \quad w^{(m)} = Ap^{(m)}, \end{aligned}$$

$$\begin{aligned}
r^{(m+1)} &= b - v^{(m+1)}, \quad \text{wobei } v^{(m+1)} = v^{(m)} + \alpha_m w^{(m)} \\
&\quad (1. \text{ rekursive Berechnung}), \\
&= b - Ax^{(m)} - \alpha_m Ap^{(m)} \\
&= r^{(m)} - \alpha_m Ap^{(m)} \\
r^{(m+1)} &= r^{(m)} - \alpha_m w^{(m)} \quad (2. \text{ rekursive Berechnung}).
\end{aligned} \tag{4.15}$$

Die Funktionswerte $R(x^{(m)})$ bzw. $\|r(x^{(m)})\|_2$ sind abnehmend und $R(x^*) = -\frac{1}{2}b^T b$.

Die Vektoren $r(x^{(m)})$ sind keine Richtungen des steilsten Abstiegs zu $R(x)$.

Die Berechnung der Suchrichtungen $p^{(m)}$ erfolgt analog zum CG mit dem Ansatz

$$p^{(0)} = r^{(0)} = b - Ax^{(0)}, \quad x^{(0)} \text{ gegeben,}$$

$$p^{(m)} = r^{(m)} + \sum_{j=0}^{m-1} \beta_j p^{(j)}, \quad m \geq 1.$$

Bei der Bestimmung der Koeffizienten β_j gehen wir wie im CG gleich von der verkürzten Rekursion aus und erhalten

$$\begin{aligned}
p^{(m)} &= r^{(m)} + \beta_{m-1} p^{(m-1)}, \\
p^{(m+1)} &= r^{(m+1)} + \beta_m p^{(m)}, \\
Ap^{(m+1)} &= Ar^{(m+1)} + \beta_m Ap^{(m)}, \quad \text{das ist } w^{(m+1)} = \underbrace{Ar^{(m+1)}}_{=t^{(m+1)}} + \beta_m w^{(m)}, \\
r^{(m)T} Ap^{(m+1)} &= \underbrace{r^{(m)T} Ar^{(m+1)}}_{=0} + \beta_m r^{(m)T} Ap^{(m)}, \\
r^{(m)T} Ap^{(m+1)} &= \beta_m r^{(m)T} Ap^{(m)}
\end{aligned}$$

und somit

$$\beta_m = \frac{r^{(m)T} Ap^{(m+1)}}{r^{(m)T} Ap^{(m)}}. \tag{4.16}$$

Neben der Wahl geeigneter Abbruchbedingungen ist im Fall beliebiger regulärer Matrizen noch ein zusätzlicher Test auf die Durchführbarkeit der Berechnung von β_m mit Nenner $r^{(m)T} Ap^{(m)} \neq 0$ zu machen, d. h. in α_m würde der Zähler verschwinden.

Für allgemeines A , insbesondere A indefinit, kann der Nenner Null sein, so dass man ein vorzeitiges Ende hat ($x^{(m+1)} = x^{(m)}$).

Die A-Orthogonalität der Suchrichtungen $p^{(m)}$ kann man nicht erfüllen.

Trotzdem ist zu erwarten, dass neben den schon genannten Eigenschaften der A-Orthogonalität von $r^{(m)}$ und der Endlichkeit des CR noch andere abgeleitet werden können, insbesondere wenn A spd ist.

Vergleich

CG	CR
$r^{(m)}$ orthogonal	$r^{(m)}$ A-orthogonal
$p^{(m)}$ A-orthogonal	$p^{(m)}$?
$\text{span}\{r^{(0)}, \dots, r^{(m)}\} = \text{span}\{p^{(0)}, \dots, p^{(m)}\}$?

Durch Rekursivität ergibt sich für die Unterräume

$$\text{span}\{p^{(0)}\} = \text{span}\{r^{(0)}\} \text{ wegen } p^{(0)} = r^{(0)} \neq 0,$$

$$p^{(1)} = r^{(1)} + \beta_0 p^{(0)}, \quad r^{(1)} \neq 0$$

$$= r^{(1)} + \beta_0 r^{(0)},$$

$$r^{(1)} = p^{(1)} - \beta_0 p^{(0)}, \text{ damit}$$

$$\text{span}\{p^{(0)}, p^{(1)}\} = \text{span}\{r^{(0)}, r^{(1)}\},$$

$$p^{(2)} = r^{(2)} + \beta_1 p^{(1)}, \quad r^{(2)} \neq 0$$

$$= r^{(2)} + \beta_0(\gamma_1 r^{(1)} + \gamma_0 r^{(0)}),$$

$$r^{(2)} = p^{(2)} - \beta_1 p^{(1)}, \text{ damit}$$

$$\text{span}\{p^{(0)}, p^{(1)}, p^{(2)}\} = \text{span}\{r^{(0)}, r^{(1)}, r^{(2)}\},$$

$$\text{allg. } \text{span}\{p^{(0)}, p^{(1)}, \dots, p^{(m)}\} = \text{span}\{r^{(0)}, r^{(1)}, \dots, r^{(m)}\},$$

weiter

$$p^{(m)} = r^{(m)} + \beta_{m-1} p^{(m-1)} = r^{(m)} + \beta_{m-1} \sum_{j=0}^{m-1} \gamma_j r^{(j)}, \quad r^{(m)} \neq 0,$$

$$r^{(m)} = p^{(m)} - \beta_{m-1} p^{(m-1)}.$$

Nun findet man damit systematisch einfachere Darstellungen der Koeffizienten α_m und β_m aus den Formeln (4.13) und (4.16).

- $r^{(k)T} A r^{(j)} = 0, \quad r^{(k)T} A^T r^{(j)} = 0, \quad k \neq j$
- $r^{(m)T} A p^{(m)} = r^{(m)T} A \left(r^{(m)} + \beta_{m-1} \sum_{j=0}^{m-1} \gamma_j r^{(j)} \right)$

$$= r^{(m)T} A r^{(m)} + \beta_{m-1} \sum_{j=0}^{m-1} \gamma_j \underbrace{r^{(m)T} A r^{(j)}}_{=0}$$

$$= r^{(m)T} A r^{(m)}$$

- $$\begin{aligned}
\bullet \quad r^{(m+1)T} Ap^{(m)} &= r^{(m+1)T} A \left(r^{(m)} + \beta_{m-1} \sum_{j=0}^{m-1} \gamma_j r^{(j)} \right) \\
&= \underbrace{r^{(m+1)T} Ar^{(m)}}_{=0} + \beta_{m-1} \sum_{j=0}^{m-1} \gamma_j \underbrace{r^{(m)T} Ar^{(j)}}_{=0} \\
&= 0
\end{aligned}$$
- $$\begin{aligned}
\bullet \quad r^{(m)T} Ap^{(m+1)} &= r^{(m)T} A \left(r^{(m+1)} + \beta_m \sum_{j=0}^m \gamma_j r^{(j)} \right) \\
&= \underbrace{r^{(m)T} Ar^{(m+1)}}_{=0} + \beta_m \sum_{j=0}^m \gamma_j \underbrace{r^{(m)T} Ar^{(j)}}_{=0 \text{ für } j \leq m-1} \\
&= \beta_m \gamma_m r^{(m)T} Ar^{(m)} \rightarrow \gamma_m = 1
\end{aligned}$$
- $$\begin{aligned}
\bullet \quad p^{(m)T} Ar^{(m)} &= \left(r^{(m)} + \beta_{m-1} \sum_{j=0}^{m-1} \gamma_j r^{(j)} \right)^T Ar^{(m)} \\
&= r^{(m)T} Ar^{(m)}
\end{aligned}$$
- $$\begin{aligned}
\bullet \quad (Ap^{(m+1)})^T Ap^{(m)} &= \frac{1}{\alpha_{m+1}} (r^{(m+1)} - r^{(m+2)}) Ap^{(m)} \\
&= \frac{1}{\alpha_{m+1}} (r^{(m+1)} - r^{(m+2)}) A \sum_{j=0}^m \gamma_j r^{(j)} \\
&= 0
\end{aligned}$$
- $$\begin{aligned}
\bullet \quad (Ap^{(m)})^T Ap^{(m+1)} &= \left(A \sum_{j=0}^m \gamma_j r^{(j)} \right)^T \frac{1}{\alpha_{m+1}} (r^{(m+1)} - r^{(m+2)}) \\
&= 0
\end{aligned}$$
- $$\begin{aligned}
\bullet \quad r^{(m)T} Ap^{(m+1)} &= (r^{(m+1)} + \alpha_m Ap^{(m)})^T Ap^{(m+1)} \\
&= r^{(m+1)T} Ap^{(m+1)} + \alpha_m \underbrace{(Ap^{(m)})^T Ap^{(m+1)}}_{=0} \\
&= r^{(m+1)T} Ap^{(m+1)} \\
&= r^{(m+1)T} Ar^{(m+1)}
\end{aligned}$$

Damit vereinfachen sich in den Fällen $A = A^T$ bzw. $A = A^T > 0$ die beiden Koeffizienten (Schrittzahlen) zu

$$\begin{aligned}
\alpha_m &= \frac{r^{(m)T} Ap^{(m)}}{(Ap^{(m)})^T Ap^{(m)}} = \frac{r^{(m)T} Ar^{(m)}}{\|Ap^{(m)}\|_2^2} = \frac{\|r^{(m)}\|_A^2}{\|Ap^{(m)}\|_2^2} = \frac{\|r^{(m)}\|_A^2}{\|w^{(m)}\|_2^2} > 0, \\
\beta_m &= \frac{r^{(m)T} Ap^{(m+1)}}{r^{(m)T} Ap^{(m)}} = \frac{r^{(m+1)T} Ar^{(m+1)}}{r^{(m)T} Ar^{(m)}} = \frac{\|r^{(m+1)}\|_A^2}{\|r^{(m)}\|_A^2} \geq 0.
\end{aligned} \tag{4.17}$$

Die Zusammenfassung der neuen Eigenschaften ergibt

$$\begin{aligned} r^{(m+1)T} A p^{(m)} &= 0, \\ (A p^{(m+1)})^T A p^{(m)} &= 0, \quad p^{(m+1)T} A^2 p^{(m)} = 0 \quad \text{für } A = A^T, \\ r^{(m)T} A p^{(m)} &= r^{(m)T} A r^{(m)} = r^{(m-1)T} A p^{(m)}. \end{aligned} \quad (4.18)$$

Eine A-Orthogonalität der Suchrichtungen $p^{(m)}$ sowie eine generelle Orthogonalität der Vektoren $A p^{(m)}$ liegen nicht vor.

Aus der rekursiven Bildung der Unterräume ergeben sich wie beim CG (siehe 3.87) die polynomiale Darstellung der Residua und Suchrichtungen.

Als lineare Unterräume der Vektoren des CR können wir also \mathcal{R}_m , \mathcal{P}_m und $\mathcal{K}_m(A, r^{(0)})$ gemäß der Formeln (3.88) verwenden.

Die Krylov-Unterräume \mathcal{K}_m bilden, wenn kein vorzeitiges Ende eintritt, eine aufsteigende Folge von Unterräumen des \mathbb{R}^n .

$$\mathcal{K}_1(A, r^{(0)}) \subset \mathcal{K}_2(A, r^{(0)}) \subset \dots \subset \mathcal{K}_n(A, r^{(0)}).$$

Das gemeinsame gleichmäßige Anwachsen der Dimension der genannten Unterräume entspricht der Aussage, dass $r^{(m)} = 0$ gdw. $p^{(m)} = 0$ ist.

Die m -te Iterierte $x^{(m)}$ liegt im affinen Unterraum $x^{(0)} + \mathcal{K}_m$, also

$$x^{(m)} = x^{(0)} + \sum_{k=0}^{m-1} \alpha_k p^{(k)} = x^{(0)} + \delta^{(m)} \in x^{(0)} + \mathcal{P}_m = x^{(0)} + \mathcal{K}_m \quad (4.19)$$

und ist so gewählt, dass das Residuum

$$r^{(m)} = r^{(0)} - \sum_{k=0}^{m-1} \alpha_k A p^{(k)} = r^{(0)} - A \delta^{(m)} \in \mathcal{K}_{m+1} \quad (4.20)$$

senkrecht auf $\text{span}\{A r^{(0)}, A r^{(1)}, \dots, A r^{(m-1)}\}$ steht.

Damit ist CR eine schiefe Projektionsmethode bzw. schiefe Krylov-Unterraum-Methode mit $K_m = \mathcal{K}_m = \mathcal{K}_m(A, r^{(0)})$ und

$$\begin{aligned} L_m &= \text{span}\{A r^{(0)}, A r^{(1)}, \dots, A r^{(m-1)}\} \\ &= A \text{span}\{r^{(0)}, r^{(1)}, \dots, r^{(m-1)}\} \\ &= A \text{span}\{p^{(0)}, p^{(1)}, \dots, p^{(m-1)}\} \\ &= A \text{span}\{r^{(0)}, A r^{(0)}, \dots, A^{m-1} r^{(0)}\} \\ &= \text{span}\{A r^{(0)}, A^2 r^{(0)}, \dots, A^m r^{(0)}\} \\ &= A K_m \\ &\subset \text{span}\{r^{(0)}, A r^{(0)}, A^2 r^{(0)}, \dots, A^m r^{(0)}\} = \mathcal{K}_{m+1}. \end{aligned}$$

Im Allgemeinen ist $r^{(m)} \not\perp r^{(0)}$.

Zudem gilt

$$x^{(m)} = \arg \min_{x \in x^{(0)} + K_m} R(x), \quad K_m = \mathcal{K}_m(A, r^{(0)}). \quad (4.21)$$

Damit haben wir den Iterierten wieder eine geometrische Interpretation gegeben.

Zusammenfassung und Bewertung der Eigenschaften des CR

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

- + Bei exakter Rechnung ist spätestens $x^{(n)} = A^{-1}b$, damit erhält man die Optimalität. Bei $r^{(m)} = 0$, $m \leq n - 1$, haben wir ein vorzeitiges Ende des CR.
Es gilt $r^{(m)} = 0$ gdw. $p^{(m)} = 0$.
- + Der Fehler $\|r^{(m)}\|_2 = \|b - Ax^{(m)}\|_2$ wird ständig verkleinert.
- + Einfache Implementation des Algorithmus.
- + Die Residua $r^{(m)}$ sind orthogonal zum Unterraum $AP_m = A \text{span}\{p^{(0)}, p^{(2)}, \dots, p^{(m-1)}\}$ und untereinander A-orthogonal, damit linear unabhängig.
- + Das Abstiegszenario und die Minimierungsaufgabe sind dem Problem angepasst.
- + Nur die letzten zwei Suchrichtungen sind zu merken. Die Suchrichtungen werden rekursiv im Iterationsverlauf erzeugt. (4.22)
- ± Hauptaufwand in einem Iterationsschritt ist eine Matrix-Vektor-Multiplikation.
- ± Aufeinander folgende Vektoren $Ap^{(m+1)}$ und $Ap^{(m)}$ sind orthogonal und damit linear unabhängig.
- ± Die Suchrichtungen $p^{(m)}$ sind keine Richtungen des steilsten Abstiegs zu $R(x)$ (jedoch "schräge" oder "schiefe" Abstiegsrichtung zu $R(x)$).
- Bei numerischen Rechnungen kann der Fall $r^{(n)} \neq 0$ eintreten, so dass eine Fortsetzung des CR sinnvoll ist, möglichst verbunden mit genaueren Rechnungen (verbesserte Gleitpunktarithmetik).

Die Implementierung des CR erfolgt in Maple.

Ergänzt um die zusätzliche Abfrage des Nenners bei der Berechnung des Parameters β , die Bestimmung von Werten der Funktionale sowie einige Ergebnisfelder und Zwischenausgaben entsteht dann die erweiterte Prozedur, die wir in den Beispielrechnungen verwenden.

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

    local k,i,x,p,t,v,w,r,r2,r20,alpha,alphan,alphas,beta,betar,
          b2,Q,R,fh,fh1,fh2,xv1,rv1,pv1,xx;
    global xv,rv,pv,lp,lr;

    k:=0:
    x:=evalm(x0):
    v:=evalm(A&*x):
    r:=evalm(b-v): lr:=evalm(r):

    p:=evalm(r): lp:=evalm(p):
    t:=evalm(A&*r):
    w:=evalm(t): # w:=evalm(A&*p);

    alphan:=evalm(transpose(t)&*r):
    alphas:=evalm(transpose(w)&*w):
    b2:=evalm(transpose(b)&*b):
    r2:=evalm(transpose(r)&*r):
    r20:=evalf(r2):
    Q:=0.5*evalm(transpose(x)&*v)-evalm(transpose(x)&*b):
    R:=0.5*evalm(transpose(v)&*v)-evalm(transpose(v)&*b):
    # R:=0.5*(r2-b2):

    xx:=matrix(n,0,[]):
    xv1:=concat(xx,x):
    xv:=evalm(xv1):
    rv1:=concat(xx,r):
    rv:=evalm(rv1):
    pv1:=concat(xx,p):
    pv:=evalm(pv1):

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

    if aus=ja then
        printf('\n'):
        printf('Schritt k = %g\n',k);
        printf('Startvektor      x = [||fh||]\n',seq(x[i],i=1..n));
        printf('Residuum/Suchr. r = b-Ax = [||fh||]\n',seq(r[i],i=1..n));
        printf('Funktionswert      Q(x) = '||fh2||'\n',Q);
        printf('Funktionswert      R(x) = '||fh2||'\n',R);
        printf('Anfangsfehlerquadrat r'r = '||fh2||'\n\n',r2);
    end if;
```



```

if r2<etol then RETURN(x,0); end if;

while (r2/r20>etol) and (k<maxiter) do
  if alphas=0 then
    lprint('Abbruch wegen Nenner r'Ar=0'):
    RETURN(x,k);
  end if;
  alpha:=alphas/alphas;
  x:=evalm(x+alpha*p);
  r:=evalm(r-alpha*w); lr:=evalm(r);
  r2:=evalm(transpose(r)&*r);
  k:=k+1;

  v:=evalm(A&*x):
  Q:=0.5*evalm(transpose(x)&*v)-evalm(transpose(x)&*b);
  R:=0.5*evalm(transpose(v)&*v)-evalm(transpose(v)&*b):
  # R:=0.5*(r2-b2):

  xv1:=concat(xv,x);
  xv:=evalm(xv1);
  rv1:=concat(rv,r);
  rv:=evalm(rv1);

  if aus=ja then
    printf('\n');
    printf('Schritt k = %g\n',k);
    printf('Suchrichtung          p = [||fh||']\n',seq(p[i],i=1..n));
    printf('Suchschritt          alpha = [||fh2||']\n',alpha);
    printf('Iterationsvektor      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) = [||fh2||']\n',Q);
    printf('Funktionswert      R(x) = [||fh2||']\n',R);
    printf('Fehlernormquadrat  r'r = [||fh2||']\n\n',r2);
  end if;

  betar:=1/alphas;
  t:=evalm(A&*r):
  alphas:=evalm(transpose(t)&*r);
  beta:=alphas*betar;
  p:=evalm(r+beta*p); lp:=evalm(p);
  w:=evalm(t+beta*w);
  alphas:=evalm(transpose(w)&*w);
  pv1:=concat(pv,p);
  pv:=evalm(pv1);

  if aus=ja then
    printf('Schritt          beta = [||fh2||']\n',beta);
    printf('neue Suchrichtung  p = [||fh||']\n\n',seq(p[i],i=1..n));
  end if;

end do:

[x,k];
end:

```

4.3.1 Beispiele zum Verfahren der konjugierten Residua

In den Beispielen, die wir zum CG im Abschnitt 3.6.6 betrachtet haben, illustrieren wir nun das CR für das LGS (1.1) unter der üblichen Voraussetzung $A = A^T > 0$. Dazu erfolgen zum Teil vergleichende Betrachtungen zu den anderen AV.

Beispiel 4.8

Gegeben sei das LGS mit der spd Tridiagonalmatrix

$$A = A(7, 7) = \text{tridiag}(-1, 2, -1) \quad \text{und} \quad b = (2, -7, 11, -13, 8, 2, 5)^T.$$

Die exakte Lösung des LGS ist $x^* = (1, 0, 6, 1, 9, 9, 7)^T$.

Für die Funktionale $Q(x)$ und $R(x)$ gilt $Q(x^*) = -90$, $R(x^*) = -218$.

Der Startvektor sei $x^{(0)} = 0$.

Bei exakter Rechnung mit dem CR werden genau $n = 7$ Schritte bis zu x^* ausgeführt, wobei im letzten Schritt die entscheidende Fehlerverbesserung von $x^{(6)} = (-0.174, -1.958, 3.768, -1.239, 6.958, 7.498, 6.244)^T$ zu $x^{(7)} = x^*$ auftritt (Angaben in $x^{(6)}$ auf 3 Stellen nach dem Komma abgeschnitten).

CG war mit $x^{(6)} = (0.126, -1.141, 5.043, 0.541, 8.232, 8.540, 6.978)^T$ jedoch etwas besser.

Im Iterationsverlauf bilden wie erwartet die Funktionswerte $R(x^{(m)})$ sowie die Normwerte $\|r^{(m)}\|_2 = \|b - Ax^{(m)}\|_2$ monoton abnehmende Folgen. Hier gilt dies auch für die Normen $\|e^{(m)}\|_A$, $\|e^{(m)}\|_2$ und Werte $Q(x^{(m)})$.

Wenn bei numerischer Rechnung das CR auf Grund von Rundungsfehlern nicht mit dem n -ten Schritt endet, werden jedoch die weiteren Iterierten in der Nähe von $x^{(n)}$ bleiben, vorausgesetzt die Kondition der Matrix A ist moderat. Bei schlechter Kondition kann der Iterationsverlauf sich auch etwas “wegbewegen“, insbesondere bei schwacher Gleitpunktarithmetik.

Die folgende Rechnung ist mit Maple bei `Digits:=16` gemacht worden.

CR - Iterationsverlauf mit Funktionalen und verschiedenen Fehlern

m	[x(m) [1], x(m) [2], x(m) [3], x(m) [4], x(m) [5], x(m) [6], x(m) [7]]	R(x(m))	Q(x(m))	e(m) _A	e(m) _2
0	[0.000, 0.000, 0.000, 0.000, 0.000, 0.000, 0.000]	0.000	0.000	13.416	15.780
1	[0.543, -1.900, 2.985, -3.528, 2.171, 0.543, 1.357]	-202.976	-63.233	7.317	13.543
2	[-0.073, -1.779, 2.871, -4.214, 2.716, 3.467, 3.269]	-212.866	-76.651	5.167	11.194
3	[-0.036, -2.164, 2.348, -3.772, 4.100, 5.134, 5.179]	-216.160	-82.800	3.795	9.173
4	[-0.092, -2.539, 2.464, -2.919, 5.352, 6.358, 5.572]	-217.119	-85.488	3.004	7.604
5	[-0.335, -2.396, 2.892, -2.194, 6.275, 6.941, 5.851]	-217.474	-86.883	2.497	6.354
6	[-0.174, -1.958, 3.768, -1.239, 6.958, 7.499, 6.245]	-217.657	-88.237	1.878	4.712
7	[1.000, -0.000, 6.000, 1.000, 9.000, 9.000, 7.000]	-218.000	-90.000	0.000	0.000
8	[1.000, -0.000, 6.000, 1.000, 9.000, 9.000, 7.000]	-218.000	-90.000	0.000	0.000
9	[1.000, -0.000, 6.000, 1.000, 9.000, 9.000, 7.000]	-218.000	-90.000	0.000	0.000
10	[1.000, -0.000, 6.000, 1.000, 9.000, 9.000, 7.000]	-218.000	-90.000	0.000	0.000

Die Iterierten $x^{(7)} \approx x^*, x^{(8)}, x^{(9)}, x^{(10)}$ liegen nahe beieinander.

$x^{(7)}$	$x^{(8)}$	$x^{(10)}$	x^*
9.9999999999999480e-01	9.9999999999999680e-01	9.9999999999999680e-01	1
-3.0000000000000000e-15	-6.6917217813141610e-15	-6.8418502630509300e-15	0
5.999999999999900e+00	5.999999999999950e+00	5.999999999999950e+00	6
9.999999999999800e-01	9.999999999999230e-01	9.999999999999190e-01	1
8.999999999999900e+00	8.999999999999950e+00	8.999999999999950e+00	9
9.000000000000010e+00	8.999999999999970e+00	8.999999999999970e+00	9
6.999999999999970e+00	6.999999999999990e+00	6.999999999999990e+00	7

Tab. 4.1 Iterierte $x^{(7)}, x^{(8)}, x^{(10)}$ des CR bei fortlaufender Rechnung

Im CR kann man bei mehr als n Iterationen einen Restart machen.

Dies bedeutet, dass die Iterierte $x^{(n)}$ mit $r^{(n)} \neq 0$ als neuer Startvektor $x_r^{(0)}$ genommen wird und dann als nächstes vor einer neuen Schleife das Residuum $r_r^{(0)} = b - Ax_r^{(0)}$ ermittelt und die Anfangssuchrichtung $p_r^{(0)} = r_r^{(0)}$ definiert wird. Damit werden die erste und auch weitere Iterierte nach dem Restart geringfügig von den ursprünglichen Iterationsvektoren bei fortlaufender Rechnung (vergl. Tabelle 4.1) abweichen.

$x_r^{(0)} = x^{(7)}$	$x_r^{(1)}$	$x_r^{(3)}$	x^*
9.9999999999999480e-01	9.9999999999999660e-01	9.999999999999930e-01	1
-3.0000000000000000e-15	-5.3462017434620170e-15	-1.0018475064768060e-15	0
5.999999999999900e+00	5.999999999999950e+00	6.0000000000000000e+00	6
9.999999999999800e-01	9.999999999999280e-01	9.999999999999620e-01	1
8.999999999999900e+00	8.999999999999950e+00	8.999999999999980e+00	9
9.000000000000010e+00	8.999999999999980e+00	9.000000000000020e+00	9
6.999999999999970e+00	7.0000000000000000e+00	7.0000000000000030e+00	7

Tab. 4.2 Iterierte $x_r^{(0)} = x^{(7)}, x_r^{(1)}, x_r^{(3)}$ des CR bei Restart nach n -tem Schritt

Beispiel 4.9

Sei

$$A(n, n) = \begin{pmatrix} 2 & -1 & 0 & \dots & 0 & 1 \\ -1 & 4 & -1 & \dots & 0 & 0 \\ 0 & -1 & 4 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 4 & -1 \\ 1 & 0 & 0 & \dots & -1 & 2 \end{pmatrix}, \quad b = \begin{pmatrix} n \\ 4 \\ 6 \\ \dots \\ 2(n-1) \\ n+2 \end{pmatrix}.$$

Damit ist $x^* = A^{-1}b = (1, 2, \dots, n)^T$.

Wir berechnen den Fehler $\|e^{(m)}\|_\infty = \|x^* - x^{(m)}\|_\infty$ und vergleichen diesen zwischen CR und CG für ausgewählte Schritte m (Rechnung mit 16 Mantissenstellen). Der Startvektor für CR und CG ist der Nullvektor.

Fehler $\ e(m)\ _\infty$				
m	n=10		n=100	
	CR	CG	CR	CG
0	1.000e+01	1.000e+01	1.000e+02	1.000e+02
1	4.436e+00	4.202e+00	4.971e+01	4.936e+01
2	3.111e+00	2.894e+00	3.770e+01	3.502e+01
3	2.163e+00	1.375e+00	2.642e+01	1.633e+01
4	7.783e-01	3.142e-01	9.400e+00	3.531e+00
5	1.404e-01	8.785e-02	1.212e+00	1.019e+00
6	2.000e-15	1.100e-14	2.548e-01	2.753e-01
7	2.000e-15	1.100e-14	6.861e-02	7.385e-02
8			1.841e-02	1.980e-02
9			4.936e-03	5.309e-03
10			1.324e-03	1.424e-03
20			2.541e-09	2.729e-09
30			9.000e-14	3.000e-13

Beispiel 4.10

Sei A 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 eine Richtung $p(x)$, wo $Ap(x)$ A -orthogonal zur vorherigen ist.

Im CR werden (theoretisch) endlich viele Schritte ausgeführt.

Die Schritte des CR sind wie folgt.

$$\begin{aligned} x^{(0)} &= \left(\frac{9}{2}, 3\right)^T && \text{Startvektor,} \\ r^{(0)} &= b - Ax^{(0)} = \left(-\frac{9}{2}, -6\right)^T = -\frac{3}{2}(3, 4)^T && \text{Anfangsresiduum,} \\ p^{(0)} &= r^{(0)} && \text{Abstiegsrichtung,} \\ R(x^{(0)}) &= \frac{225}{8} = 28.125. \end{aligned}$$

S1

$$\begin{aligned} t^{(0)} &= Ar^{(0)} = \left(-\frac{9}{2}, -12\right)^T = -\frac{3}{2}(3, 8)^T, \\ w^{(0)} &= Ap^{(0)} = \left(-\frac{9}{2}, -12\right)^T, \\ \alpha_0 &= \frac{\|r^{(0)}\|_A^2}{\|Ap^{(0)}\|_2^2} = \frac{r^{(0)T}t^{(0)}}{w^{(0)T}w^{(0)}} = \frac{41}{73} = 0.561\dots, \end{aligned}$$

$$\begin{aligned}
x^{(1)} &= x^{(0)} + \alpha_0 p^{(0)} = \frac{9}{73}(16, -3)^T, \\
r^{(1)} &= b - Ax^{(1)} = r^{(0)} - \alpha_0 w^{(0)} = \frac{18}{73}(-8, 3)^T, \\
R(x^{(1)}) &= \frac{162}{73} = 2.219\dots, \\
\beta_0 &= \frac{\|r^{(1)}\|_A^2}{\|r^{(0)}\|_A^2} = \frac{288}{5329} = 0.054\dots, \quad p^{(1)} = r^{(1)} + \beta_0 p^{(0)} = \frac{738}{5329}(-16, 3)^T.
\end{aligned}$$

S2

$$\begin{aligned}
t^{(1)} &= Ar^{(1)} = \frac{36}{73}(-4, 3)^T, \\
w^{(1)} &= Ap^{(1)} = t^{(1)} + \beta_0 w^{(0)} = \frac{1476}{5329}(-8, 3)^T, \\
\alpha_1 &= \frac{\|r^{(1)}\|_A^2}{\|Ap^{(1)}\|_2^2} = \frac{r^{(1)T}t^{(1)}}{w^{(1)T}w^{(1)}} = \frac{73}{82} = 0.890\dots, \\
x^{(2)} &= x^{(1)} + \alpha_1 p^{(1)} = (0, 0)^T, \\
r^{(2)} &= b - Ax^{(2)} = r^{(1)} - \alpha_1 w^{(1)} = (0, 0)^T \quad \text{und Stopp mit } x^{(2)} = x^*, \\
R(x^{(2)}) &= 0, \\
\beta_1 &= \frac{\|r^{(2)}\|_A^2}{\|r^{(1)}\|_A^2} = 0, \quad p^{(2)} = r^{(2)} + \beta_1 p^{(1)} = (0, 0)^T.
\end{aligned}$$

Bei $r^{(m)} = 0$ ist die Rechnung mit $x^{(m)} = x^*$, $\beta_{m-1} = 0$, $p^{(m)} = 0$ beendet. Dann wäre auch $\alpha_m = 0$ und β_m wegen verschwindenden Nenner nicht berechenbar. Dazu erhalten wir weiterhin die Beziehungen

$$\begin{aligned}
K_1 &= \mathcal{K}_1(A, r^{(0)}) = \text{span}\{r^{(0)}\} = \text{span}\{p^{(0)}\}, \\
L_1 &= \text{span}\{Ar^{(0)}\}, \\
r^{(1)} &\perp L_1, \quad r^{(1)} \perp Ar^{(0)}, \quad (Ap^{(1)}, Ap^{(0)}) = 0, \\
K_2 &= \mathcal{K}_2(A, r^{(0)}) = \text{span}\{r^{(0)}, Ar^{(0)}\} = \text{span}\{r^{(0)}, r^{(1)}\} = \text{span}\{p^{(0)}, p^{(1)}\} = \mathbb{R}^2, \\
L_2 &= \text{span}\{Ar^{(0)}, Ar^{(1)}\} = \mathbb{R}^2, \\
r^{(2)} &\perp L_2, \quad r^{(2)} \perp Ar^{(0)}, Ar^{(1)}, \quad (Ap^{(2)}, Ap^{(1)}) = 0, \\
&\rightarrow r^{(2)} = 0, \quad x^{(2)} = x^*.
\end{aligned}$$

Dies sind auch die Ergebnisse aus Rechnungen mit Maple in der Rationalarithmetik. Bei numerischen Rechnungen in Maple (`Digits:=16`) taucht in der Iterierten $x^{(2)}$ ein Fehler in der Größenordnung der Mantissengenauigkeit auf. Ist die Abbruchbedingung moderat, so endet das CR. Nimmt man jedoch eine extrem kleine Toleranz, beispielsweise $\varepsilon = 10^{-40}$, so wird der Iterationsprozess fortgesetzt.

Aber nach der n -ten Iteration ist auch ein Restart möglich, wobei sich die nächsten Iterierten vom ursprünglichen Iterationslauf ganz wenig unterscheiden werden.

Berechnungen mit Maple

```

Urspruengliche Iteration mit sehr kleiner Toleranz
x(0): [ 4.5,          3          ]
x(1): [ 1.972602739726027, -0.369863013698630 ]
x(2): [-0.1e-14,      0.9e-15     ] r(2)=[0.1e-14, -0.8e-15 ]
x(3): [-0.3595505617977530e-15, 0.3876404494382020e-15] r(3)=[0.359e-15, 0.224e-15]
x(4): [ 0,          0.5000000000000000e-15] r(4)=[0,      0      ]
-----
Restart mit x(2) --> x(0)
x(0): [-0.1e-14,      0.9e-15     ] r(0)=[0.1e-14, -0.18e-14 ]
x(1): [-0.4641833810888252e-15, -0.644699140401146e-16] r(1)=[0.464e-15, 0.128e-15]
x(2): [-0.1e-30,      -0.16e-30    ] r(2)=[0.1e-30, 0.1e-30 ]

```

CR - Iterationsverlauf, $x(0)=[9/2,3]$, Vergleich mit GV,CG,Orthomin(0)

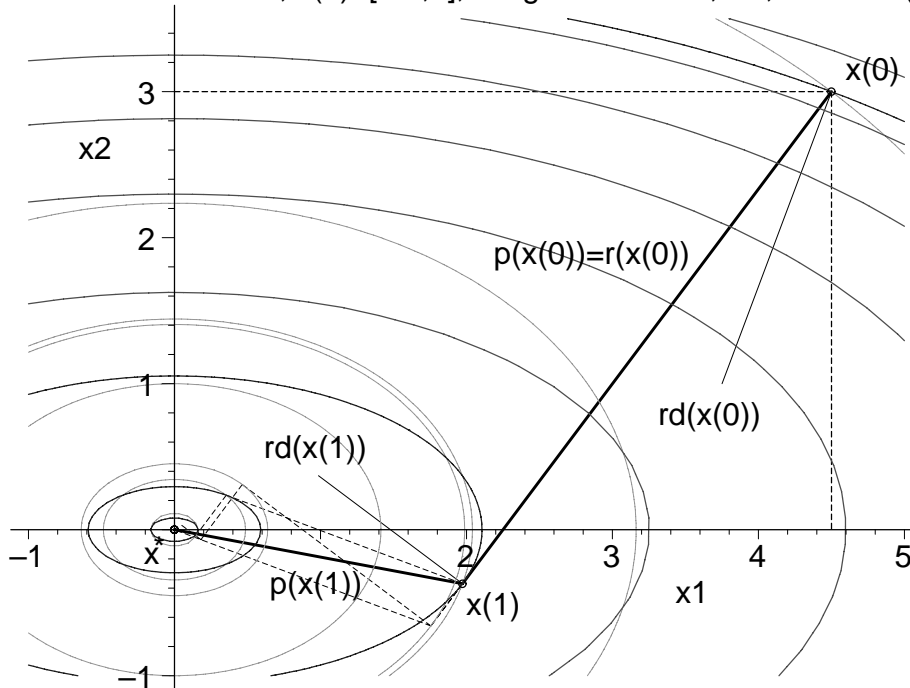


Abb. 4.5 Datei *abst401.ps*

Höhenlinienbild mit Iterationsverlauf $x^{(0)}, x^{(1)}, x^{(2)} = x^*$ von CR zu $R(x) = \frac{1}{2}x_1^2 + 2x_2^2$ mit Abstiegsrichtung $p(x)$, $\text{contours}=[28.125, 2.219, 1, 0.1751, 0.0138, 0]$ und $\text{contours}=6$, im Vergleich mit GV, CG und Orthomin(0); $Q(x)$ mit den Höhenlinien $\text{contours}=[19.125, 2.0824, 1.976, 0.11907, 0.2041, 0.01296, 1, 5]$

In der Abbildung 4.5 sind weitere Abstiegsrichtungen eingezeichnet.

Für die Suchrichtung und Abstiegsrichtung $p(x)$ gelten $p^{(0)} = r^{(0)} = b - Ax^{(0)} \perp Q(x)=\text{const}$ und $\nabla R(x)=\text{const}$, weitere $p(x)$ - gebildet aus $r(x) \perp Q(x)=\text{const}$ und p_{alt} - sind $\nabla Q(x)=\text{const}$, $\nabla R(x)=\text{const}$.

Der Vektor $r(x) = b - Ax \perp Q(x)=\text{const}$ ist eine "schiefe" Abstiegsrichtung bez. $R(x)$ und $\nabla R(x)=\text{const}$. Die Richtung $\hat{r}(x) = A^T r(x) = A^T(b - Ax)=\text{rd}(x)$ ist $\perp R(x)=\text{const}$.

Beispiel 4.11

Sei A eine Tridiagonalmatrix.

$$\begin{pmatrix} 4 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 6 \\ 2 \end{pmatrix}, \quad x^* = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \quad x^{(0)} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Die Funktionale sind

$$R(x) = \frac{17}{2}x_1^2 - 8x_1x_2 + x_1x_3 + 9x_2^2 - 8x_2x_3 + \frac{17}{2}x_3^2 - 2x_1 - 20x_2 - 2x_3,$$

$$Q(x) = 2x_1^2 - x_1x_2 + 2x_2^2 - x_2x_3 + 2x_3^2 - 2x_1 - 6x_2 - 2x_3.$$

Am gemeinsamen eindeutigen Minimum an der Stelle x^* gilt $R(x^*) = -22$ und $Q(x^*) = -8$.

Die Schritte des CR sind wie folgt.

$$\begin{aligned} x^{(0)} &= (0, 0, 0)^T && \text{Startvektor,} \\ r^{(0)} &= b - Ax^{(0)} = b = 2(1, 3, 1)^T && \text{Anfangsresiduum,} \\ p^{(0)} &= r^{(0)} && \text{Abstiegsrichtung,} \\ R(x^{(0)}) &= 0. \end{aligned}$$

S1

$$\begin{aligned} \alpha_0 &= \frac{16}{51} = 0.313\,725\dots, \\ x^{(1)} &= \frac{32}{51}(1, 3, 1)^T, \\ r^{(1)} &= \frac{14}{51}(5, -1, 5)^T, \quad R(x^{(1)}) = -\frac{1\,024}{51} = -20.078\dots, \\ \beta_0 &= \frac{343}{2\,601} = 0.131\,872\dots, \quad p^{(1)} = \frac{224}{2\,601}(19, 6, 19)^T. \end{aligned}$$

S2

$$\begin{aligned} \alpha_1 &= \frac{51}{224} = 0.227\,678\dots, \\ x^{(2)} &= (1, 2, 1)^T, \\ r^{(2)} &= 0 \quad \text{und vorzeitiger Stopp mit } x^{(2)} = x^*, \\ \beta_1 &= 0, \quad p^{(2)} = r^{(2)} = 0. \end{aligned}$$

Bei numerischer Rechnung in Maple mit `Digits:=16` hängt die Anzahl der ausgeführten Iterationen k von der Toleranz ε im Abbruchkriterium ab. Das trifft auch für andere Maple-Genauigkeiten zu.

Startvektor $x^{(0)} = (x_1^{(0)}, x_2^{(0)}, x_3^{(0)})^T$, Iterierte $x^{(i)}$ und Residua $r^{(i)}$, $i = 1, 2, \dots$, in Abhängigkeit von der Toleranz ε im Abbruchkriterium von CR

eps	k	[x(k) [1], [r(k) [1],	x(k) [2], r(k) [2],	x(k) [3] r(k) [3]]
	0	[0, [2,	0, 6,	0 2]
	1	[0.6274509803921568, [1.372549019607843,	1.882352941176470, -0.274509803921568,	0.6274509803921568 1.372549019607843]
1e-10, 1e-20, 1e-30	2	[1.0000000000000000, [0.0,	2.0000000000000000, 0.1e-15,	1.0000000000000000 0.0]
1e-35	3	[1.0000000000000000, [0.2222222222222222e-16,	2.0000000000000000, 0.1111111111111111e-16,	1.0000000000000000 0.2222222222222222e-16]]
1e-40, 1e-50, 1e-60	4	[1.0000000000000000, [0.2e-31,	2.0000000000000000, 0.1e-31,	1.0000000000000000 0.2e-31]
1e-65	5	[1.0000000000000000, [0.2e-46,	2.0000000000000000, 0.9999999999999999e-32,	1.0000000000000000 0.2e-46]
1e-70, 1e-80, 1e-90	6	[1.0000000000000000, [0.5499999999999999e-47,	2.0000000000000000, 0.6e-47,	1.0000000000000000 0.5499999999999999e-47]]
1e-96	7	[1.0000000000000000, [-0.467694566813e-48,	2.0000000000000000, 0.1151248164464e-47,	1.0000000000000000 -0.4676945668134e-48]]
1e-100	8	[1.0000000000000000, [0.28e-62,	2.0000000000000000, 0.4e-62,	1.0000000000000000 0.28e-62]

Schon $x^{(2)}$ ist sehr nahe bei x^* .

Aber es ist kein vorzeitiges Ende zu erwarten, die Iterationsanzahl geht über n hinaus. In der Prozedur für CR wird im Abbruchkriterium u. a. der relative Fehler der Residuumnorm einbezogen. Das Residuum $r^{(k)}$ selbst wird rekursiv berechnet. Aber man bemerke, dass genau diese rekursive Formel (4.15) dazu führt, dass sich im Iterationsprozess die Norm $\|r^{(k)}\|_2$ systematisch verkleinert, obwohl die Iterierten $x^{(k)}$ sich nicht mehr verändern und numerisch gleich x^* sind. Würde man die explizite Formel $r^{(k)} = b - Ax^{(k)}$ auswerten, hätte man sofort den Nullvektor. Auf diesen Sachverhalt werden wir in einem späteren Kapitel noch eingehen.

Beispiel 4.12

Wir nehmen das LGS aus Beispiel 3.7 mit $A = I$, $b = (0, 0, 1)^T = x^*$ und dem Startvektor $x^{(0)} = (0, 0, 0)^T$.

Ähnlich zum CG in Beispiel 3.14 gilt für das CR $p^{(0)} = r^{(0)} = b = (0, 0, 1)^T$.

Der Ablauf des Iterationsprozesses des CR ist ebenfalls sehr kurz, denn

$$\begin{aligned}\alpha_0 &= \frac{\|r^{(0)}\|_A^2}{\|Ap^{(0)}\|_2^2} = 1, \\ x^{(1)} &= x^{(0)} + \alpha_0 p^{(0)} = (0, 0, 1)^T, \\ r^{(1)} &= b - Ax^{(1)} = r^{(0)} - \alpha_0 Ap^{(0)} = 0 \quad \text{und vorzeitiger Stopp mit } x^{(1)} = x^*, \\ \beta_0 &= 0, \quad p^{(1)} = r^{(1)} = 0.\end{aligned}$$

Dazu erhalten wir u. a. die Beziehungen

$$\begin{aligned}Ar^{(0)} &= r^{(0)} = (0, 0, 1)^T, \\ r^{(1)} &\perp L_1 = \text{span}\{Ar^{(0)}\}, \quad r^{(1)} \perp Ar^{(0)}, \\ K_1 &= \mathcal{K}_1(A, r^{(0)}) = \text{span}\{r^{(0)}\} = \text{span}\{p^{(0)}\}.\end{aligned}$$

Beispiel 4.13

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

$$R(x) = \frac{5}{2}x_1^2 + 5x_1x_2 + 5x_2^2 - 4x_1 - 7x_2, \quad Q(x) = x_1^2 + x_1x_2 + \frac{3}{2}x_2^2 - x_1 - 2x_2.$$

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

Wir testen das CR nur für $Ax = b$.

Der Vorabschritt im CR ist wie folgt.

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

Bei exakter Rechnung haben wir zwei Schritte und die Iterationsvektoren $x^{(i)}, r^{(i)}, p^{(i)}$, $i = 0, 1, 2$, sind die Spalten der folgenden Felder.

$$\begin{aligned}
 m &= 2 \\
 xv &= \begin{bmatrix} \frac{3}{2} & \frac{237}{365} & \frac{1}{5} \\ 1 & \frac{85}{292} & \frac{3}{5} \end{bmatrix} \\
 rv &= \begin{bmatrix} -3 & \frac{-861}{1460} & 0 \\ \frac{-5}{2} & \frac{697}{1460} & 0 \end{bmatrix} \\
 pv &= \begin{bmatrix} -3 & \frac{-16974}{26645} & 0 \\ \frac{-5}{2} & \frac{93357}{213160} & 0 \end{bmatrix}
 \end{aligned}$$

Bei numerischen Rechnungen in Maple (`Digits:=16`) taucht in der Iterierten $x^{(2)}$ ein Fehler in der Größenordnung der Mantissengenauigkeit auf. Ist die Abbruchbedingung moderat, so endet das CR spätestens beim n -ten Schritt. Nimmt man jedoch eine extrem kleine Toleranz, z. B. $\varepsilon = 10^{-40}$, so wird der Iterationsprozess fortgesetzt. Aber nach der n -ten Iteration ist auch ein Restart möglich, wobei sich die nächsten Iterierten vom ursprünglichen Iterationslauf ganz wenig unterscheiden werden.

Berechnungen mit Maple

Urspruengliche Iteration mit sehr kleiner Toleranz

```

x(0): [ 1.5, 1 ]
x(1): [ 0.6493150684931508, 0.2910958904109590 ]
x(2): [ 0.1999999999999997, 0.5999999999999996 ] r(2)=[ 0.6e-15, 0.13e-14]
x(3): [ 0.1999999999999999, 0.6000000000000000 ] r(3)=[-0.9e-16, 0.5e-16 ]
x(4): [ 0.1999999999999998, 0.6000000000000000 ] r(4)=[-0.8e-31,-0.5e-31 ]
-----

```

Restart mit x(2) --> x(0)

```

x(0): [ 0.1999999999999997, 0.5999999999999996 ] r(0)=[ 0.1e-14, 0.1e-14 ]
x(1): [ 0.2000000000000000, 0.5999999999999996 ] r(1)=[ 0.16e-15,-0.12e-15]
x(2): [ 0.20000000000000001, 0.5999999999999998 ] r(2)=[ 0, 0]

```

Numerische Berechnung in Maple mit `Digits:=16`

Man bemerke die kleine Ungenauigkeit bei $m = 2$ mit $R(x^{(2)}) < R(x^*)$.

Iterationsverlauf mit Funktionalen und verschiedenen Fehlern

m	[x(m)[1], x(m)[2]]	R(x(m))	Q(x(m))	e(m) _A	e(m) _2
0	[1.5000000000000000, 1.0000000000000000]	5.1250000000000000	1.7500000000000000 2.213594362117866	1.360147050873544	
1	[0.6493150684931508, 0.2910958904109590]	-2.212157534246575	-0.493778617001314 0.642217070776987	0.545257535203323	
2	[0.1999999999999997, 0.5999999999999996]	-2.5000000000000001	-0.7000000000000000 0.0000000000000001	0.0000000000000001	0.0000000000000001
3	[0.1999999999999999, 0.6000000000000000]	-2.5000000000000000	-0.7000000000000000 0.0000000000000000	0.0000000000000000	0.0000000000000000
4	[0.1999999999999998, 0.6000000000000000]	-2.5000000000000000	-0.7000000000000000 0.0000000000000000	0.0000000000000000	0.0000000000000000

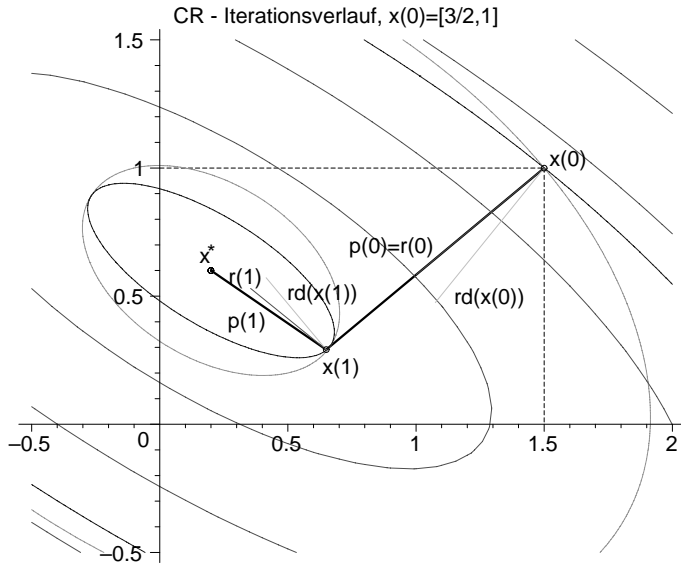


Abb. 4.6 Datei *abst402.ps*
 Höhenlinienbild mit
 Iterationsverlauf $x^{(0)}, x^{(1)}, x^*$
 des CR zu $R(x)$ mit
 $\text{contours}=[5.125, -2.212]$
 und $[-2.5, -1, 2, 7, 13, 20]$,
 dazu $Q(x)$ mit
 $\text{contours}=[-0.494, 1.75]$

In der Abbildung 4.6 sind weitere Abstiegsrichtungen eingezeichnet.

Für die Suchrichtung und Abstiegsrichtung $p(x)$ gelten $p^{(0)} = r^{(0)} = b - Ax^{(0)} \perp Q(x)=\text{const}$ und $\nabla R(x)=\text{const}$, weitere $p(x)$ – gebildet aus $r(x) \perp Q(x)=\text{const}$ und p_{alt} – sind $\nabla Q(x)=\text{const}$, $\nabla R(x)=\text{const}$. Der Vektor $r(x) = b - Ax \perp Q(x)=\text{const}$ ist eine “schiefe“ Abstiegsrichtung bez. $R(x)$ und $\nabla R(x)=\text{const}$. Die Richtung $\hat{r}(x) = A^T r(x) = A^T(b - Ax)=\text{rd}(x)$ ist $\perp R(x)=\text{const}$.

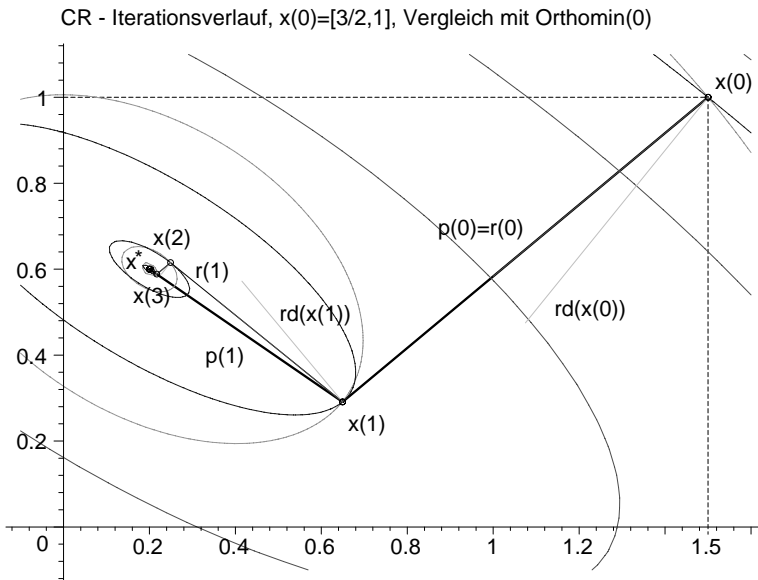


Abb. 4.7 Datei *abst403.ps*

Höhenlinienbild mit Iterationsverlauf $x^{(0)}, x^{(1)}, x^*$ des CR
 zu $R(x)$ mit $\text{contours}=[-2.4996, -2.489, -2.212, 5.125, -1, 2, 7]$,
 Vergleich mit Orthomin(0) (Iterierte $x^{(0)}, x^{(1)}, x^{(2)}, x^{(3)}$),
 dazu $Q(x)$ mit $\text{contours}=[-0.6997, -0.6965, -0.494, 1.75]$

Beispiel 4.14

Gegeben sei das LGS aus den Beispielen 1.7, 3.6, 3.16, 4.7 mit der regulären, aber nicht symmetrischen und nicht definiten Matrix A .

Wir nehmen hier Bezug auf die dort genannten Eigenschaften des Systems.

$$\begin{pmatrix} 0.780 & 0.563 \\ 0.913 & 0.659 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0.217 \\ 0.254 \end{pmatrix}, \quad x^* = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Die im Funktional $R(x)$ verwendete Matrix $B = A^T A$ hat eine sehr schlechte Kondition, ihre spektrale Konditionszahl ist ungefähr $0.5 \cdot 10^{13}$.

Die AV GV und CG sind anwendbar auf das LGS $A^T A x = A^T b$ und brauchen für akzeptable Ergebnisse eine entsprechend starke Gleitpunktarithmetik. Orthomin(0) konvergiert nicht für $Ax = b$ trotz hoher Genauigkeit – eine Grund mit ist die ungünstige Wahl von $r(x)$ als Suchrichtung – und tut sich auch schwer für $A^T A x = A^T b$.

Wie verhält sich nun das CR mit der veränderten Suchrichtung $p(x)$?

Wir testen den Iterationsverlauf des CR mit ausgewählten Startvektoren

$$x^{(0)} = \begin{pmatrix} 1.2 \\ -1.2 \end{pmatrix}, \begin{pmatrix} 0.999 \\ -1.001 \end{pmatrix}, \begin{pmatrix} 0.341 \\ -0.087 \end{pmatrix}, \begin{pmatrix} 0.991\ 891\ 566\ 446\ 068\ 04 \\ -1.005\ 852\ 632\ 339\ 509\ 328 \end{pmatrix}.$$

Die Durchführung des CR für $Ax = b$ mit der Suchrichtung $p(x)$, die von einer Richtung auf der Satteloberfläche $Q(x)$ mitbestimmt wird, birgt natürlich Probleme in sich. Man sieht das am Iterationsverlauf. Egal, ob man in Maple exakt arbeitet oder wie genau man in Maple numerisch rechnet, das Verfahren erzielt nach der 4. Iterierten keine Verbesserungen mehr, obwohl man noch nicht an der Lösung ist. Die Suchrichtungen unterliegen noch kleinen Veränderungen, nicht aber die anderen Vektoren.

Hier sind die ersten 10 Iterationen einer exakten Rechnung in Maple.

CR - Iterationsverlauf mit Funktionalen und Fehler

m	[x(m) [1], r(m) [1], p(m) [1],	x(m) [2] r(m) [2] p(m) [2]	R(x(m))	Q(x(m)) e(m) _2
0	[1.20000000000000000000, -1.20000000000000000000]		-0.0535704000000000000	0.0177600000000000000
	[-4.340000000000000000e-02, -5.080000000000000000e-02]			0.2828427124746190098
	[-4.340000000000000000e-02, -5.080000000000000000e-02]			
1	[1.16984017373116808048, -1.23530228512572952791]		-0.0558024999999808015	0.0299558971206173798
	[-1.4898452537856358e-07, 1.27281299301411152e-07]			0.2901945037357245874
	[-1.4898452537874384e-07, 1.2728129930120051e-07]			
2	[1.16984017373032291028, -1.23530228512500747734]		-0.0558024999999808015	0.0299558971206522670
	[-1.4898427266027551e-07, 1.2728159511048197e-07]			0.2901945037346444705
	[-1.4898548886135021e-07, 1.2728263414223783e-07]			
3	[1.16984017373032290338, -1.23530228512500747144]		-0.0558024999999808015	0.0299558971206522673
	[-1.4898427265821250e-07, 1.2728159511289674e-07]			0.2901945037346444617
	[-1.4898670508515293e-07, 1.2728367320594260e-07]			
4	[1.16984017373032290338, -1.23530228512500747144]		-0.0558024999999808015	0.0299558971206522673
	[-1.4898427265821247e-07, 1.2728159511289678e-07]			0.2901945037346444617
	[-1.4898792130315400e-07, 1.2728471225633664e-07]			

```

-----
5 [ 1.16984017373032290338,-1.23530228512500747144] -0.0558024999999808015 0.0299558971206522673
  [-1.4898427265821247e-07, 1.2728159511289678e-07] 0.2901945037346444617
  [-1.4898913752115485e-07, 1.2728575130673030e-07]
6 unveraendert
  unveraendert
  [-1.4899035373915570e-07, 1.2728679035712395e-07]
7 ...
  [-1.4899156995715655e-07, 1.2728782940751760e-07]
8 ...
  [-1.4899278617515740e-07, 1.2728886845791125e-07]
9 ...
  [-1.4899400239315825e-07, 1.2728990750830490e-07]
10 ...
  [-1.4899521861115910e-07, 1.2729094655869855e-07]

```

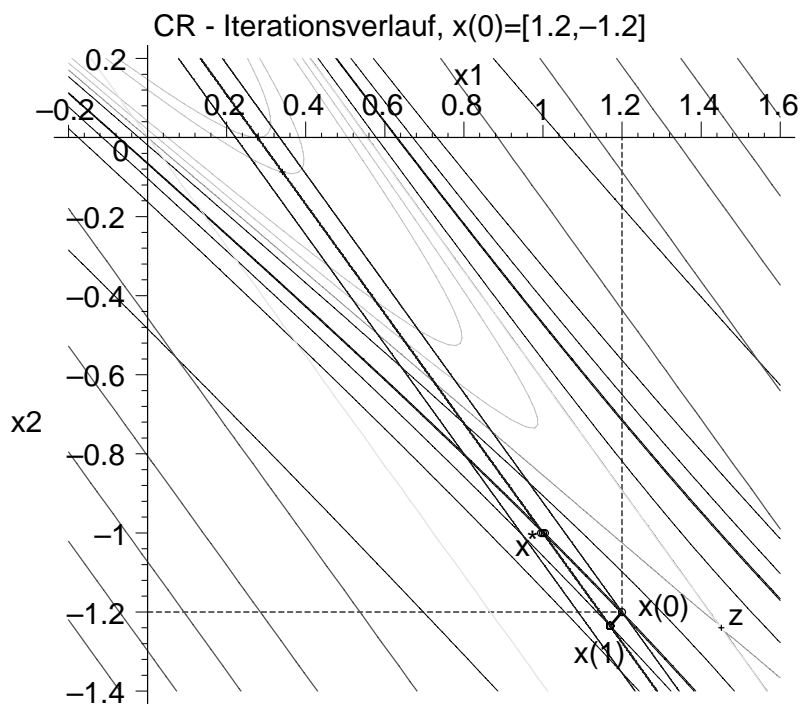


Abb. 4.8 Datei *abst404.ps*

Höhenlinienbild mit Iterationsverlauf $x^{(m)}$ des CR

zu $R(x) = 0.7209845x_1^2 + 1.040807x_1x_2 + 0.375625x_2^2 - 0.401162x_1 - 0.289557x_2$
mit `contours=[0,-0.05357,-0.0558]` und `contours=5`,

dazu $Q(x) = 0.390x_1^2 + 0.738x_1x_2 + 0.3295x_2^2 - 0.217x_1 - 0.254x_2$ mit `contours=`
`[-0.03018,-0.0259,-0.01,-0.005,0,0.005,0.01776,0.0185,0.0299,0.05,0.2]`

Die Suchrichtungen $p^{(m)}$ stehen wie die Vektoren $r^{(m)}$, die senkrecht zu den Höhenlinien von $Q(x)$ sind, fast senkrecht auf der Tallinie von $R(x)$. Damit findet keine Fortbewegung in der Iterationsfolge mehr statt ($x^* \neq x^{(2)} \approx x^{(1)}$).

Bei numerischen Rechnungen in Maple mit `Digits:=k` ändert sich nur wenig am Gesamtverhalten des CR.

Wir wenden nun CR auf das symmetrisierte LGS $Bx = c$ mit dem Startvektor $x^{(0)} = (1.2, -1.2)^T$ an.

Bei exakter Rechnung haben wir zwei Schritte und die Iterationsvektoren $x^{(i)}, \hat{r}^{(i)}, p^{(i)}$, $i = 0, 1, 2$, sind die Spalten der folgenden Felder.

$$\begin{aligned}
 & m = 2 \\
 xv &= \begin{bmatrix} \frac{6}{5} & \frac{6849130212700399614580066}{5887076167445197367298925} & 1 \\ -\frac{6}{5} & \frac{-7219938177003319170841344}{5887076167445197367298925} & -1 \end{bmatrix} \\
 &= \begin{bmatrix} 1.2 & 1.16341797148391625295701235051 & 1 \\ -1.2 & -1.22640475027802150139476414721 & -1 \end{bmatrix} \quad (30 \text{ Dezimalstellen}) \\
 rv &= \begin{bmatrix} \frac{-200581}{2500000} & \frac{-6853895783829914}{91985565116331208864045703125} & 0 \\ \frac{-289557}{5000000} & \frac{607719525299752963}{5887076167445197367298925000000} & 0 \end{bmatrix} \\
 &= \begin{bmatrix} -0.0802324 & -0.745105579898542854621338876337 \cdot 10^{-13} & 0 \\ -0.0579114 & 0.103229431387412095617712282320 \cdot 10^{-12} & 0 \end{bmatrix} \\
 pv &= \begin{bmatrix} \frac{-200581}{2500000} & \frac{-258236201748084525118476368884972160646021}{3465766580130123351047149376029084393495030615562500000} & 0 \\ \frac{-289557}{5000000} & \frac{357769113388328433218453459425056889438339}{3465766580130123351047149376029084393495030615562500000} & 0 \end{bmatrix} \\
 &= \begin{bmatrix} -0.0802324 & -0.745105579898542854621339152454 \cdot 10^{-13} & 0 \\ -0.0579114 & 0.103229431387412095617712262390 \cdot 10^{-12} & 0 \end{bmatrix}
 \end{aligned}$$

Bei numerischen Rechnungen muss das CR auf Grund der schlechten Kondition der Matrix $B = A^T A$ der Größenordnung 10^{12} mit einer entsprechend starken Gleitpunktarithmetik arbeiten, um relevante Ergebnisse zu erzielen. Es trifft das, was zum CG gesagt wurde, auch hier zu.

Ergebnisse aus Berechnungen (9 Iterationen) mit Maple mit `Digits:=13`.

m	[x(m)[1],	x(m)[2]] [rd(m)[1],	rd(m)[2]] [p(m)[1],	p(m)[2]]
0	[1.2,	-1.2] [-0.0802324,	-0.0579114] [-0.0802324,	-0.0579114]
1	[1.163417971484,	-1.226404750278] [-0.8e-13,	0.1e-12] [-0.8e-14,	0.1e-12]
2	[1.163417971484,	-1.226404750278] [-0.7485e-13,	0.1037e-12]	[-0.748...e-13,	0.103...e-12]	
3	[1.146224816551,	-1.202584775560] [-0.9764e-13,	0.7339e-13]	[-0.343...e-01,	0.476...e-01]	
4	[0.967862058594,	-0.955475015798] [0.304e-23,	0.436e-23]	[-0.656...e-21,	0.918...e-21]	
5	[0.967862058594,	-0.955475015798] [-0.1027e-23,	0.1423e-23]	[-0.102...e-23,	0.142...e-23]	
6	[0.967862058594,	-0.955475015798] [-0.1020e-23,	0.1428e-23]	[-0.143...e-15,	0.198...e-15]	
7	[0.967862058591,	-0.955475015794] [0.1499e-29,	0.1070e-29]	[-0.638...e-23,	0.885...e-23]	
8	[0.967862058591,	-0.955475015794] [0.5337e-32,	-0.7394e-32]	[0.533...e-32,	-0.739...e-32]	
9	[0.967862058591,	-0.955475015794] [0.5337e-32,	-0.7394e-32]	[0.640...e-32,	-0.887...e-32]	

Es gibt meistens einen Schritt, wo eine besonders gute Verbesserung der Iterierten $x^{(m)}$ auftritt. Je höher die Genauigkeit ist, desto früher wird das eintreten.

In Maple, das standardmäßig mit dem Gleitpunktformat `Digits:=10` arbeitet, ist also die Genauigkeit deutlich zu erhöhen.

Ergebnisse aus Berechnungen (erste 3 Iterationen) mit Maple mit `Digits:=24`.

```

Startvektor          x=[ 1.2000000000000000e+00 -1.2000000000000000e+00]
Residuum/SR  p = rd = c-Bx=[-8.0232400000000000e-02 -5.7911400000000000e-02]
Funktionswert      Q(x)= -5.3570400000000000e-02
Funktionswert      R(x)= -1.1749161911664000e-01
Anfangsfehlerquadrat rd'rd= 9.7909682597200000e-03

```

Schritt k = 1

```

Suchrichtung        p=[-8.0232400000000000e-02 -5.7911400000000000e-02]
Suchschritt         alpha= 4.5595081932091957e-01
Iterationsvektor    x=[ 1.1634179714839163e+00 -1.2264047502780215e+00]
Residuum            rd = c-Bx=[-7.4510557989900000e-14 1.0322943138740000e-13]
Funktionswert      Q(x)= -5.5802499999982226e-02
Funktionswert      R(x)= -1.2238710324650000e-01
Fehlernormquadrat  rd'rd= 1.6208138756532175e-26
Schritt            beta= 3.4414661657875583e-37
neue Suchrichtung  p=[-7.4510557989900000e-14 1.0322943138740000e-13]

```

Schritt k = 2

```

Suchrichtung        p=[-7.4510557989900000e-14 1.0322943138740000e-13]
Suchschritt         alpha= 5.7988106640981765e+11
Iterationsvektor    x=[ 1.1202107096579425e+00 -1.1665439575202169e+00]
Residuum            rd = c-Bx=[-9.2847817947022607e-15 1.0879588370821311e-13]
Funktionswert      Q(x)= -5.5802499999990382e-02
Funktionswert      R(x)= -1.2238710324650000e-01
Fehlernormquadrat  rd'rd= 1.1922751484826265e-26
Schritt            beta= 9.3554412140868651e+11
neue Suchrichtung  p=[-6.9707914510341267e-02 9.6575687690952216e-02]

```

Schritt k = 3

```

Suchrichtung        p=[-6.9707914510341267e-02 9.6575687690952216e-02]
Suchschritt         alpha= 1.7244915516749652e+00
Iterationsvektor    x=[ 9.999999999997825e-01 -9.999999999996986e-01]
Residuum            rd = c-Bx=[ 9.4000000000000000e-37 0.0000000000000000e-01]
Funktionswert      Q(x)= -5.5802500000000000e-02
Funktionswert      R(x)= -1.2238710324650000e-01
Fehlernormquadrat  rd'rd= 8.8360000000000000e-73
Schritt            beta= 1.8428764083274434e-46
neue Suchrichtung  p=[ 9.3999999998715369e-37 1.7797705646365491e-47]

```

Die Ergebnisse sind mit 16 Nachkommastellen angezeigt.

Die Iterierte $x^{(2)}$ ist noch nicht nahe bei x^* . Aber dann ist eine deutliche Verbesserung der Iterierten von $x^{(2)}$ zu $x^{(3)}$ zu bemerken. Weitere Rechnungen mit Maple können den Einfluss des Gleitpunktformats mit `Digits:=k` deutlich machen.

Ergebnisse aus Berechnungen mit Maple für Startvektor

$$x^{(0)} = (0.991\,891\,566\,446\,068\,04, -1.005\,852\,632\,339\,509\,328)^T$$

Der Vektor $x^* - x^{(0)}$ ist orthogonal zur Tallinie $t(x_1)$ mit der Genauigkeitsordnung $\mathcal{O}(10^{-17})$.

Somit muss der erste CR-Schritt, der ein Gradientenschritt ist, bis auf Ungenauigkeiten in den letzten Dezimalstellen in das Minimum führen, also $x^{(1)} \approx x^*$. Weitere Iterationen werden sich dann um die Minimumstelle bewegen, wobei die Situation zwischenzeitlich sich verschlechtern kann.

```

Startvektor          x = [ 9.9189156644606804e-01, -1.0058526323395093e+00]
Residuum            rd = c-Bx = [ 1.7783570530717400e-02,  1.2836104447023644e-02]
Funktionswert       Q(x) = -5.56928390500000023e-02
Funktionswert       R(x) = -1.2214659276740205e-01
Anfangsfehlerquadrat rd'rd =  4.8102095819590051e-04

```

k	Iterationsvektor x Residuum rd=c-Bx	Funktionswert Q(x) Funktionswert R(x) Fehlernormquadrat rd'rd
1	[0.9999999999999998743683, -0.9999999999999998259454] [1.4e-24, -1.9e-24]	-5.5802500000000000e-02 -1.2238710324650000e-01 5.5700000000000000e-48
2	[0.9999999999999998743684, -0.9999999999999998259455] [1.3812042025493872e-24, -1.9135667254679633e-24]	-5.5802500000000000e-02 -1.2238710324650000e-01 5.5694626619582725e-48
3	[1.00000000000302928204366, -1.00000000000419686915943] [-1.2274820475240600e-33, -8.85991244468583000e-34]	-5.5802500000000000e-02 -1.2238710324650000e-01 2.2916926626538050e-66

Zum Vergleich geben wir einige Iterationen des CR für $Ax = b$ an, das über eine kleine anfängliche Verbesserung nicht hinauskommt und nicht gegen den Lösungsvektor x^* konvergiert (Maple mit `Digits:=24`).

CR fuer Ax=b - Iterationsverlauf

m	[x(m) [1], x(m) [2]]	[r(m) [1], r(m) [2]]
0	[0.991891566446068040000000, -1.005852632339509328000000]	[9.619610e-03, 1.125988e-02]
1	[0.998576496651265889142867, -0.998027830615791487431756]	[1.248703e-09, -1.066799e-09]
2	[0.998576496651272972881763, -0.998027830615797539251489]	[1.248701e-09, -1.066802e-09]
3	[0.998576496651272972939590, -0.998027830615797539300892]	[1.248701e-09, -1.066802e-09]
4	[0.998576496651272972939591, -0.998027830615797539300893]	[1.248701e-09, -1.066802e-09]
5	[0.998576496651272972939591, -0.998027830615797539300893]	[1.248701e-09, -1.066802e-09]

Kapitel 5

Abstiegsverfahren unter veränderten Voraussetzungen

Wir betrachten die vier Basisvarianten von AV für (1.1) unter Voraussetzungen, die vom Normalfall $A = A^T > 0$ abweichen. Wir werden diese Annahme systematisch abschwächen. Dabei versuchen wir, insbesondere auf sich verändernde Eigenschaften und Verhalten im Iterationsablauf einzugehen. Die Bezeichnungen der AV behalten wir bei.

5.1 Gradientenverfahren

5.1.1 GV mit symmetrischer indefiniter Matrix

In Anlehnung an (2.13) haben wir die quadratische Form

$$\delta_Q(x) = (x^* - x)^T A(x^* - x) = e(x)^T A e(x) = 2Q(x) + x^{*T} b \quad (5.1)$$

mit

$$\begin{aligned} \delta_Q(x^*) &= 0, \\ Q(x) &= \frac{1}{2} x^T A x - x^T b = \frac{1}{2} (\delta_Q(x) - x^{*T} b) \quad (\text{indefinit}), \quad (5.2) \\ Q(x^* + \Delta x) - Q(x^*) &= \frac{1}{2} (\Delta x)^T A \Delta x \geq 0, \quad \Delta x \neq 0. \end{aligned}$$

Hier könnte der Gedanke entstehen, dass man eventuell die Größe $|\delta_Q(x)|$ möglichst klein machen sollte. Diese Idee wird sich gleich als irrelevant erweisen.

Welche neuen Aspekte treten nun auf?

1. Die Berechnung von $\sqrt{\delta_Q(x)}$ im Reellen ist für $\delta_Q(x) < 0$ nicht möglich. Man kann jedoch auf die Größe $\sqrt{|\delta_Q(x)|} = |\sqrt{\delta_Q(x)}|$ ausweichen.

2. Man gelangt zur Erkenntnis, dass $\delta_Q(x) = 0$ nicht zwingend $Ax = b$ nach sich zieht. Es kann viele Stellen z geben, wo $\delta_Q(z) = 0$ ist, aber die keine Lösung des LGS sind.

Beispiel 5.1

Sei $Ax = b$, A diagonal (damit symmetrisch) und indefinit.

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

Die Funktionale sind

$$\begin{aligned} Q(x) &= \frac{1}{2}x^T Ax - x^T b = \frac{1}{2}x^T Ax = \frac{1}{2}(x_1^2 - x_2^2), \quad Q(x^*) = 0, \\ \delta_Q(x) &= (x^* - x)^T A(x^* - x) = x^T Ax = x_1^2 - x_2^2. \end{aligned}$$

Es gilt $\delta_Q(x) = 0$, falls $|x_1| = |x_2|$, so dass unendlich viele Lösungen vorliegen. Die uns interessierende ist der Sattelpunkt $x^* = (0, 0)^T$, durch den zwei Null-Höhenlinien gehen und an welchem die Tangentialebene horizontal verläuft.

Man hat also keine typische Minimierungsaufgabe für $Q(x)$ mehr. Aber die notwendige Bedingung am Sattelpunkt ist genau wie die an der Minimumstelle, nämlich dass der Gradient des Funktionals $\nabla Q(x)$ an dieser Stelle verschwindet. Das bedeutet $\nabla Q(x) = (x_1, -x_2)^T = 0$ und führt auf x^* .

3. Bei der Berechnung der Schrittzahl

$$\alpha = \frac{r^T r}{(Ar)^T r} = \frac{r^T r}{r^T Ar} = \frac{\|r\|_2^2}{r^T Ar} \quad (5.3)$$

kann der Nenner Null sein, also $r^T Ar = 0$.

Dann ist α nicht berechenbar, d. h. die Strahlenminimierung der Funktion $f(\alpha) = Q(x + \alpha p)$, $p = r$, ist geometrisch nicht durchführbar bzw. die Gleichung $f'(\alpha) = 0$ nicht nach α auflösbar. In der Suchrichtung r gibt es keine Stelle α , wo $f'(\alpha) = 0$ wird.

Falls dieser Sonderfall nicht auftritt, stellt der Vektor r entweder eine Abstiegs- oder Anstiegsrichtung dar.

Beispiel 5.2

Sei $Ax = b$, $A = A^T$ und indefinit aus Beispiel 1.4.

$$\begin{pmatrix} 1 & 0 \\ 0 & -4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad x^* = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Wir rechnen mit verschiedenen Startvektoren den ersten Schritt oder mehrere Anfangsschritte des GV.

- (1) $x^{(0)} = (1, 2)^T$ Startvektor,
 $r^{(0)} = b - Ax^{(0)} = (-1, 8)^T$ Anfangsresiduum, Suchrichtung.

$$\begin{aligned} \underline{\text{S1}} \quad \alpha_0 &= \frac{\|r^{(0)}\|_2^2}{r^{(0)T}Ar^{(0)}} = \frac{65}{-255} = -\frac{13}{51}, \\ x^{(1)} &= x^{(0)} + \alpha_0 r^{(0)} = \frac{1}{51}(64, -2)^T, \\ r^{(1)} &= b - Ax^{(1)} = r^{(0)} - \alpha_0 Ar^{(0)} = -\frac{8}{51}(8, 1)^T. \end{aligned}$$

Das GV konvergiert mit Abstiegs- und Anstiegsrichtungen gegen x^* .

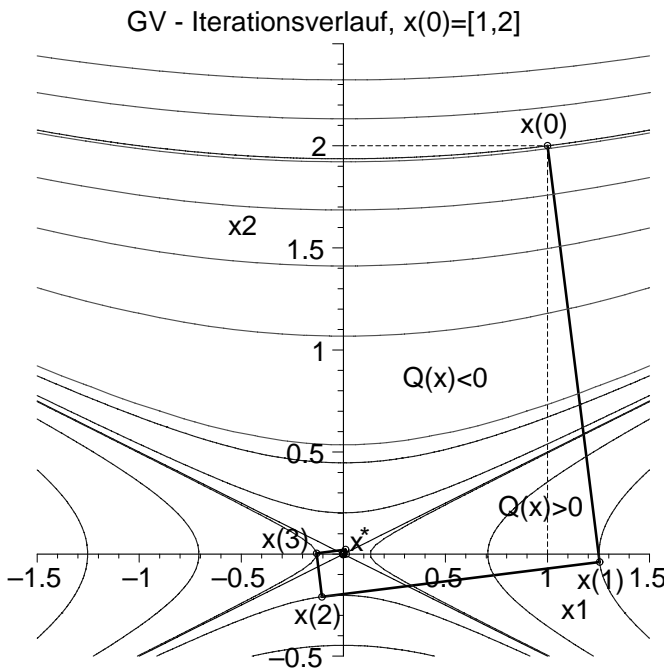


Abb. 5.1 Datei *gv_04.ps*

Höhenlinienbild

bei "einschachtelndem"

Iterationsverlauf des GV mit

$x^{(0)} = (1, 2)^T$
 zu $Q(x) = \frac{1}{2}x_1^2 - 2x_2^2$, $Q(x^*) = 0$,

mit `contours=[-7.5,0.7843,`
`-0.0820,0.00858,0,-0.4,0.25]`
 und `contours=7`

GV fuer $Ax=b$, A diagonal und indefinit - konvergenter Iterationsverlauf

m	[x(m) [1], x(m) [2]]	[r(m) [1], r(m) [2]]
0	[1.0000000000, 2.0000000000]	[-1.000e+00, 8.000e+00]
1	[1.2549019608, -0.0392156863]	[-1.255e+00, -1.569e-01]
2	[-0.1045751634, -0.2091503268]	[1.046e-01, -8.366e-01]
3	[-0.1312315776, 0.0041009868]	[1.312e-01, 1.640e-02]
4	[0.0109359648, 0.0218719296]	[-1.094e-02, 8.749e-02]
5	[0.0137235637, -0.0004288614]	[-1.372e-02, -1.715e-03]
6	[-0.0011436303, -0.0022872606]	[1.144e-03, -9.149e-03]
7	[-0.0014351439, 0.0000448482]	[1.435e-03, 1.794e-04]
8	[0.0001195953, 0.0002391907]	[-1.196e-04, 9.568e-04]
9	[0.0001500804, -0.0000046900]	[-1.501e-04, -1.876e-05]
10	[-0.0000125067, -0.0000250134]	[1.251e-05, -1.001e-04]

$$(2) \quad x^{(0)} = (1, 0)^T, \quad r^{(0)} = b - Ax^{(0)} = (-1, 0)^T.$$

$$\begin{aligned} \underline{\text{S1}} \quad \alpha_0 &= \frac{\|r^{(0)}\|_2^2}{r^{(0)T}Ar^{(0)}} = \frac{1}{1} = 1, \\ x^{(1)} &= x^{(0)} + \alpha_0 r^{(0)} = (0, 0)^T, \\ r^{(1)} &= b - Ax^{(1)} = r^{(0)} - \alpha_0 Ar^{(0)} = (0, 0)^T \Rightarrow \text{Stopp mit } x^{(1)} = x^*. \end{aligned}$$

Das GV braucht nur endlich viele Schritte.

$$(3) \quad x^{(0)} = (0, 1)^T, \quad r^{(0)} = (0, 4)^T.$$

$$\begin{aligned} \underline{\text{S1}} \quad \alpha_0 &= \frac{16}{-64} = -\frac{1}{4}, \\ x^{(1)} &= (0, 0)^T, \\ r^{(1)} &= (0, 0)^T \Rightarrow \text{Stopp wie in (2)}. \end{aligned}$$

$$(4) \quad x^{(0)} = (3, 1)^T, \quad r^{(0)} = (-3, 4)^T.$$

$$\begin{aligned} \underline{\text{S1}} \quad \alpha_0 &= -\frac{5}{11}, \\ x^{(1)} &= \frac{3}{11}(16, -3)^T, \\ r^{(1)} &= -\frac{12}{11}(4, 3)^T. \end{aligned}$$

Das GV divergiert. Die ersten Iterierten sind

$$m = 4$$

$$xv = \begin{bmatrix} 3 & \frac{48}{11} & \frac{108}{11} & \frac{1728}{121} & \frac{3888}{121} \\ 1 & \frac{-9}{11} & \frac{36}{11} & \frac{-324}{121} & \frac{1296}{121} \end{bmatrix}$$

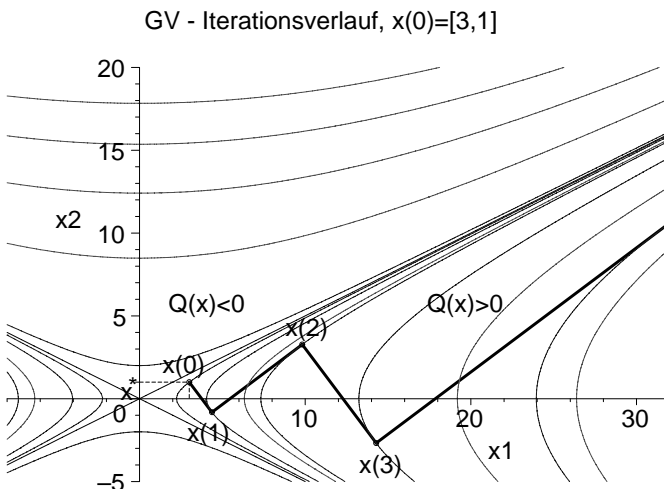


Abb. 5.2 Datei *gv_03.ps*

Höhenlinienbild

bei divergentem "Zick-Zack"-

Iterationsverlauf des GV mit

$x^{(0)} = (3, 1)^T$
zu $Q(x) = \frac{1}{2}x_1^2 - 2x_2^2$, $Q(x^*) = 0$,

mit `contours=[-8,0, 2.5,`

`8.18,26.78,87.63,286.8]`

und `contours=7`

$$(5) \quad x^{(0)} = (2, 1)^T, \quad r^{(0)} = (-2, 4)^T.$$

$$\begin{aligned} \underline{\text{S1}} \quad \alpha_0 &= -\frac{1}{3}, \\ x^{(1)} &= \frac{1}{3}(8, -1)^T, \\ r^{(1)} &= -\frac{4}{3}(2, 1)^T. \end{aligned}$$

In Richtung $r^{(1)}$ gelten für $Q(x)$ die Beziehungen

$$Q(x + \alpha r) = f(\alpha) = \frac{10}{3}(2\alpha + 1), \quad f'(\alpha) = \frac{20}{3} \neq 0$$

und man kann entlang der Geraden $g(\alpha) = x + \alpha r$ kein Extremum finden. Der Ausdruck $r^{(1)T} A r^{(1)}$ ist Null und die nächste Schrittzahl α nicht berechenbar.

```
Startvektor      x = [ 2.0000000000000000e+00  1.0000000000000000e+00]
Residuum        r = b-Ax = [-2.0000000000000000e+00  4.0000000000000000e+00]
Funktionswert   Q(x) =  0.0000000000000000e-01
Anfangsfehlerquadrat r'r =  2.0000000000000000e+01
```

Schritt k = 1

```
Suchrichtung    p=r=b-Ax = [-2.6666666666666667e+00 -1.3333333333333333e+00]
Suchschritt     alpha = -3.3333333333333333e-01
Iterationsvektor x = [ 2.6666666666666667e+00 -3.3333333333333333e-01]
```

```
Residuum       r = b-Ax = [-2.6666666666666667e+00 -1.3333333333333333e+00]
Funktionswert  Q(x) =  3.3333333333333333e+00
Fehlernormquadrat r'r =  8.8888888888888889e+00
```

Abbruch wegen Nenner $r'Ar=0$

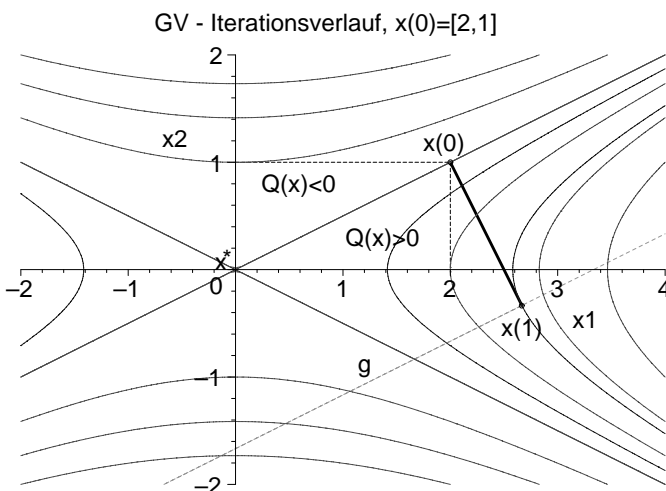


Abb. 5.3 Datei *gv_02.ps*

Höhenlinienbild

mit Abbruch bei $x^{(1)}$ des

Iterationsverlaufs des GV mit

$g(\alpha)$ und $x^{(0)} = (2, 1)^T$

zu $Q(x) = \frac{1}{2}x_1^2 - 2x_2^2$, $Q(x^*) = 0$,

mit `contours=[3.333,1,0]`

und `contours=7`

$$(6) \quad x^{(0)} = (-8, 1)^T, \quad r^{(0)} = (8, 4)^T = 4(2, 1)^T.$$

In Richtung des Vektors $(2, 1)^T$ war beim vorhergehenden Test gerade das Problem der Nulldivision im ersten Schritt aufgetreten, was hier nun schon bei der Anfangsrechnung passiert. Das GV bricht ab.

$$x^{(0)} = (-8, 1.1)^T \text{ oder } x^{(0)} = (-8, 0.9)^T.$$

Für beide Startvektoren ist die Rechnung durchführbar, aber das GV divergiert.

GV - Iterationsverlaufe, $x(0)=[-8,1], [-8,1.1], [-8,0.9]$

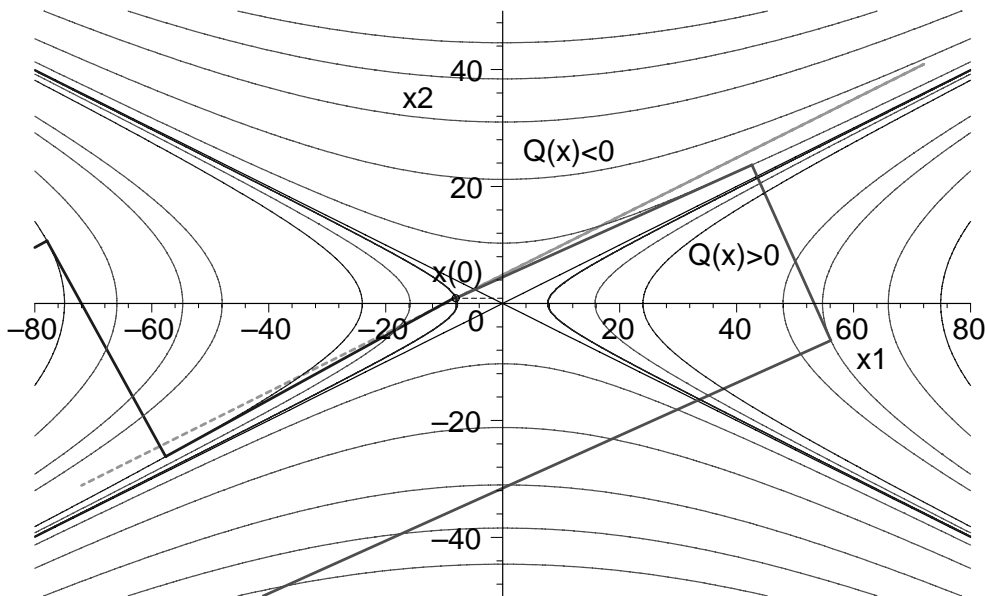


Abb. 5.4 Datei *gv_05.ps*

Höhenlinienbild und Iterationsverlauf des GV mit

$x^{(0)} = (-8, 1)^T$ (Abbruch), $x^{(0)} = (-8, 1.1)^T$, $(-8, 0.9)^T$ (Divergenz)

zu $Q(x) = \frac{1}{2}x_1^2 - 2x_2^2$, $Q(x^*) = 0$,

mit `contours=[30, 29.58,288,2806, 30.38,-213,1496, 0]` und `contours=7`

GV fuer $Ax=b$, A diagonal und indefinit - 2 divergente Iterationsverlaufe

m	[x(m) [1], x(m) [2]]	[x(m) [1], x(m) [2]]
0	[-8.000, 1.100]	[-8.000, 0.900]
1	[-57.619, -26.190]	[42.632, 23.684]
2	[-77.916, 10.713]	[56.131, -6.315]
3	[-561.182, -255.083]	[-299.120, -166.178]
4	[-758.867, 104.344]	[-393.838, 44.307]
5	[-5465.651, -2484.387]	[2098.740, 1165.967]

Beispiel 5.3 Sei $Ax = b$, $A = A^T$ und indefinit aus Beispiel 1.5.

$$\begin{pmatrix} 2 & 1 \\ 1 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \quad x^* = \begin{pmatrix} 10/7 \\ 1/7 \end{pmatrix}, \quad x^{(0)} = \begin{pmatrix} 3/2 \\ 1 \end{pmatrix}.$$

Die Funktionale sind zu $Ax = b$

$$Q(x_1, x_2) = \frac{1}{2}x^T Ax - x^T b = x_1^2 + x_1 x_2 - \frac{3}{2}x_2^2 - 3x_1 - x_2, \quad Q(x^*) = -\frac{31}{14},$$

und zu $Bx = A^T Ax = A^T b = c$

$$R(x_1, x_2) = \frac{1}{2}(Ax)^T Ax - (Ax)^T b = \frac{5}{2}x_1^2 - x_1 x_2 + 5x_2^2 - 7x_1, \quad R(x^*) = -5.$$

x^* ist Sattelpunkt von $Q(x)$ und Minimumstelle von $R(x)$.

Die Suchrichtung $r(x) = b - Ax \perp Q(x)=\text{const}$, $\nabla R(x)=\text{const}$, ist Anstiegs- bzw. Abstiegsrichtung bezüglich $Q(x)$.

Der Vektor $\hat{r}(x) = A^T r(x) = A^T(b - Ax) \perp R(x)=\text{const}$.

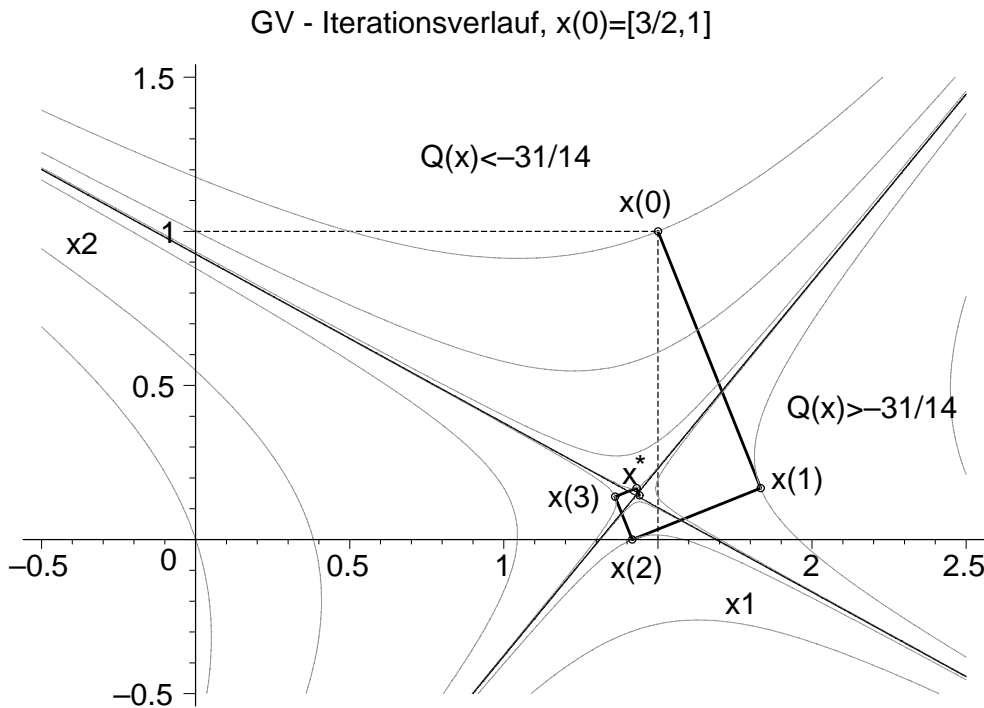


Abb. 5.5 Datei *gv_002.ps*

Höhenlinienbild bei "einschachtelndem" Iterationsverlauf des GV mit $x^{(0)} = (\frac{3}{2}, 1)^T$ zu $Q(x) = x_1^2 + x_1 x_2 - \frac{3}{2}x_2^2 - 3x_1 - x_2$, $Q(x^*) = -\frac{31}{14} = -2.214285\dots$,
mit `contours=[-2.5,0,-1.0,-3.25,-2.0417,-2.2431,-2.2095,-2.2151]`
und `contours=[-31/14]`

$x^{(0)} = \left(\frac{3}{2}, 1\right)^T$ ist ein günstiger Startvektor.

Die ersten Iterierten sowie die zugehörigen und gegen den Nullvektor tendierenden Residua sind

$$m = 5$$

$$xv = \begin{bmatrix} \frac{3}{2} & \frac{11}{6} & \frac{17}{12} & \frac{49}{36} & \frac{103}{72} & \frac{311}{216} \\ 1 & \frac{1}{6} & 0 & \frac{5}{36} & \frac{1}{6} & \frac{31}{216} \end{bmatrix}$$

$$rv = \begin{bmatrix} -1 & \frac{-5}{6} & \frac{1}{6} & \frac{5}{36} & \frac{-1}{36} & \frac{-5}{216} \\ \frac{5}{2} & \frac{-1}{3} & \frac{-5}{12} & \frac{1}{18} & \frac{5}{72} & \frac{-1}{108} \end{bmatrix}$$

Andere Startvektoren werden wir nicht testen. Aber Situationen, wie sie in Beispiel 5.2 aufgetreten sind, können auch hier vorkommen.

5.1.2 GV mit nicht symmetrischer Matrix

In den einführenden Beispielen dazu mit definiter Matrix (Beispiel 1.6) bzw. indefiniter Matrix (Beispiel 1.7) und mit den Beziehungen (1.8) und (1.9) haben wir die sich verändernde Situation schon etwas beleuchtet.

Mit dem “nicht symmetrischen“ Funktional $Q_{ns}(x)$ (2.15), in dem die exakte Lösung noch erscheint, kann man aber nicht weiter rechnen. Also muss man auch hier mit $Q(x)$, $R(x)$ oder anderen Funktionalen arbeiten. Das werden wir im Weiteren auch probieren.

(a) Der übliche Zugang für $A \neq A^T$ geht über die Minimierung der Residuumnorm $\|r(x)\|_2$ (2.18), was konform ist mit der Symmetrisierung des LGS (Betrachtung des Normalgleichungssystems $A^T A x = A^T b$, $B = A^T A = B^T > 0$) und der Minimierung des Funktionals $R(x)$ (2.19).

Aus der notwendigen Bedingung am Minimum $\nabla R(x) = 0$ (4.2) leitet man die Lösung des LGS (1.1) ab. Der Gradient $\nabla R(x)$ ist senkrecht zur Höhenlinie $R(x) = \text{const}$ und liefert die Abstiegsrichtung $\hat{r}(x) = -\nabla R(x) = A^T(b - Ax) \perp R(x) = \text{const}$.

(b) Betrachten wir das Funktional $Q(x)$ bei $A \neq A^T$.

Für seinen Gradienten gilt

$$\nabla Q(x) = \frac{1}{2} \tilde{A}x - b, \quad \tilde{A} = A + A^T = \tilde{A}^T. \quad (5.4)$$

Als Suchrichtung müsste man dann

$$\tilde{r}(x) = b - \frac{1}{2} \tilde{A}x \quad (5.5)$$

wählen. Damit ist $\tilde{r}(x)$ orthogonal zu den Höhenlinien von $Q(x)$, aber der Vektor $r(x) = b - Ax \not\perp Q(x) = \text{const}$.

Führt man das GV mit $\tilde{r}(x)$ aus und gelangt an die Stelle, wo $\tilde{r}(x) = 0$ ist, dann bedeutet das

$$0 = \tilde{r}(z) = b - \frac{1}{2}\tilde{A}z, \quad (5.6)$$

z löst das LGS $\frac{1}{2}\tilde{A}x = b$ und damit ist

$$z = \left(\frac{1}{2}\tilde{A}\right)^{-1}b \quad (5.7)$$

und i. Allg. $z \neq x^* = A^{-1}b$.

Zum dem LGS $\frac{1}{2}\tilde{A}x = b$ definiert man formal das Funktional

$$\begin{aligned} T(x) &= \frac{1}{2}x^T\left(\frac{1}{2}\tilde{A}\right)x - x^Tb \\ &= \frac{1}{4}x^T(A + A^T)x - x^Tb \\ &= \frac{1}{2}\left(\frac{1}{2}x^T Ax - x^Tb\right) + \frac{1}{2}\left(\frac{1}{2}x^T A^T x - x^Tb\right) \\ &= \frac{1}{2}\left(\frac{1}{2}x^T Ax - x^Tb\right) + \frac{1}{2}\left(\frac{1}{2}x^T Ax - \underbrace{b^T x}_{=x^Tb}\right)^T \\ &= \frac{1}{2}Q(x) + \frac{1}{2}Q(x)^T, \quad Q = Q^T \\ &= Q(x). \end{aligned} \quad (5.8)$$

Also spielen neben der Symmetrie von \tilde{A} noch weitere Eigenschaften eine Rolle.

Wie steht es mit ihrer positiven Definitheit?

Wenn $A > 0$ ist, dann gilt auch $\tilde{A} > 0$. Das führt letztendlich auf die zwei Fälle mit A positiv definit oder nicht.

(c) Sei $\tilde{A} = A + A^T = \tilde{A}^T > 0$.

Dann ist $Q(x) = T(x) = \frac{1}{4}x^T\tilde{A}x - x^Tb$ eine positive quadratische Form.

Diese hat an der Stelle z (5.7) ein Minimum und

$$\begin{aligned} \nabla Q(x) &= \frac{1}{2}\tilde{A}x - b, \quad \nabla Q(z) = 0, \\ Q(z) &= \min_{x \in \mathbb{R}^n} Q(x), \\ z &= \arg \min_{x \in \mathbb{R}^n} Q(x). \end{aligned} \quad (5.9)$$

Mit dem GV und $\tilde{r}(x)$ als Abstiegsrichtung gelangt man zu diesem Minimum bei z .

Das GV zu $Q(x)$, jedoch mit $r(x)$ als Suchrichtung liegt in der "Nachbarschaft" vom GV zu $R(x)$ mit $\hat{r}(x)$ als Abstiegsrichtung und kann damit zu x^* konvergieren.

(d) Sei $\tilde{A} = \tilde{A}^T$ indefinit.

Dann hat die quadratische Form $Q(x) = T(x)$ an der Stelle z einen Sattelpunkt und

$$\nabla Q(x) = \frac{1}{2}\tilde{A}x - b,$$

$$\nabla Q(z) = 0.$$

Das GV zu $Q(x)$ mit $r(x)$ als Suchrichtung (Abstieg-/Anstiegsrichtung) hat mit dem GV zu $R(x)$ mit $\hat{r}(x)$ als Abstiegsrichtung i. Allg. nichts zu tun und divergiert.

(e) Einige Beziehungen zwischen verschiedenen Funktionalen und ihren Funktionswerten an ausgewählten Stellen.

LGS, Lösung	Funktional	Eigenschaften
$Ax = b$ $Ax^* = b$	$Q(x) = \frac{1}{2}x^T Ax - x^T b$ $= Q(x)^T$	$Q(\frac{x}{2}) = \frac{1}{2}(\frac{1}{4}x^T Ax - x^T b)$ $\nabla Q(x) = \frac{1}{2}\tilde{A}x - b, \tilde{A} = A + A^T$ $\nabla Q(z) = 0, \frac{1}{2}\tilde{A}z = b$ z Minimumstelle oder Sattelpunkt $z = (\frac{1}{2}\tilde{A})^{-1}b \neq x^* = A^{-1}b$ $Q(z) \geq Q(x^*) = -\frac{1}{2}x^{*T}b$ $\nabla Q(x^*) \neq 0,$ weil mit $Ax^* = b, x^{*T}A^T = b^T$ folgen $x^{*T}Ax^* = x^{*T}b, x^{*T}A^T x^* = b^T x^* = x^{*T}b$ $x^{*T}(A + A^T)x^* = 2x^{*T}b$ $\frac{1}{2}x^{*T}(A + A^T)x^* - x^{*T}b = 0$ $\frac{1}{2}x^{*T}\tilde{A}x^* - x^{*T}b = 0$ $x^{*T}(\frac{1}{2}\tilde{A}x^* - b) = 0$ $\nrightarrow \frac{1}{2}\tilde{A}x^* - b = 0$
$\frac{1}{2}\tilde{A}x = b$ $\frac{1}{2}\tilde{A}z = b$	$T(x) = \frac{1}{2}x^T(\frac{1}{2}\tilde{A})x - x^T b$ $= Q(x)$	$\nabla T(x) = \nabla Q(x) = \frac{1}{2}\tilde{A}x - b$
$A^T Ax = A^T b$ $A^T Ax^* = A^T b$	$R(x) = \frac{1}{2}x^T A^T Ax - x^T A^T b$	$\nabla R(x) = A^T Ax - A^T b = A^T(Ax - b)$ $\nabla R(x^*) = 0$ $x^* = A^{-1}b$ absolute Minimumstelle $R(x) > R(x^*) = -\frac{1}{2}\ b\ _2^2$ für $x \neq x^*$

Wir definieren unter Verwendung der bisherigen Funktionale ein zusätzliches $S(x)$.

$\tilde{A}x = b$ $\tilde{A}t = b$	$S(x) = \frac{1}{2}x^T \tilde{A}x - x^T b$ $= 2Q(x) + x^T b$	$\nabla S(x) = \tilde{A}x - b = 2\nabla Q(x) + b$ $\nabla S(t) = 0, \quad \tilde{A}t = b$ t Minimumstelle oder Sattelpunkt $t = \tilde{A}^{-1}b \neq x^*, z$ $0 = \nabla S(t) = 2\nabla Q(t) + b$ $= 2(\frac{1}{2}\tilde{A}t - b) + b$ $= \tilde{A}t - b$ $= \frac{1}{2}\tilde{A}(2t) - b$ $\rightarrow z = 2t, \quad t = \frac{1}{2}z$ $S(t) = 2Q(t) + t^T b$ $= 2Q(\frac{1}{2}z) + (\frac{1}{2}z)^T b$ $= 2[\frac{1}{2}(\frac{1}{4}z^T Az - z^T b)] + \frac{1}{2}z^T b$ $= \frac{1}{4}z^T Az - \frac{1}{2}z^T b$ $= \frac{1}{2}(\frac{1}{2}z^T Az - z^T b)$ $= \frac{1}{2}Q(z)$ d. h. mehrere Varianten zur Berechnung von $S(t)$ $S(t) \geq \frac{1}{2}Q(x^*)$ $S(0) = 0$ $S(x^*) = \frac{1}{2}x^{*T} \tilde{A}x^* - x^{*T} b = 0$ $S(z) = \frac{1}{2}z^T \tilde{A}z - z^T b = z^T (\frac{1}{2}\tilde{A}z - b) = z^T \cdot 0 = 0$
--------------------------------------	---	---

(f) Die Durchführung der Strahlenminimierung für $Q(x) = T(x) = \frac{1}{2}x^T (\frac{1}{2}\tilde{A})x - x^T b$ gemäß $f(\alpha) = Q(x + \alpha\tilde{r}) \rightarrow \min_{\alpha}$ mit der Suchrichtung $\tilde{r}(x) = -\nabla Q(x) = b - \frac{1}{2}\tilde{A}x \perp Q(x) = \text{const}$ bringt uns nicht zu x^* , sondern führt zu z .

Nehmen wir jedoch alle Informationen vom "symmetrischen Fall", also die Ausgangsmatrix A ,

$$r(x) = b - Ax \perp Q(x) = \text{const},$$

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

dann ist $r(x)$ eine andere Suchrichtung bezüglich $Q(x)$ und das Verfahren kann im Fall $A > 0$ langsam gegen x^* konvergieren oder sonst divergieren.

Ein anderes Funktional anzugeben, das sein eindeutiges Minimum in x^* besitzt und nicht ganz so aufwändig wie $R(x)$ zu berechnen ist, aber natürlich eine gewisse "Verwandtschaft" zu diesem haben wird, und mit der Suchrichtung $r(x)$ arbeitet, ist nicht ohne Weiteres möglich.

Und wie steht es um die Schrittzahl im nicht symmetrischen Fall?

In (3.10) wurde ihre Herleitung bei $A = A^T$ gemacht. Davon abweichend erhält man bei $A \neq A^T$

$$\begin{aligned}
 f(\alpha) &= Q(x + \alpha p) = \frac{1}{2}(x + \alpha p)^T A(x + \alpha p) - (x + \alpha p)^T b \\
 &= Q(x) + \frac{1}{2}\alpha(x^T A p + x^T A^T p) + \frac{1}{2}\alpha^2 p^T A p - \alpha p^T b, \\
 &= Q(x) + \alpha[x^T(\frac{1}{2}(A + A^T))p - b^T p] + \frac{1}{2}\alpha^2 p^T A p, \\
 &= Q(x) + \alpha[x^T(\frac{1}{2}\tilde{A}) - b^T]p + \frac{1}{2}\alpha^2 p^T A p, \\
 0 = f'(\alpha) &= [x^T(\frac{1}{2}\tilde{A}) - b^T]p + \alpha p^T A p, \\
 \bar{\alpha} &= \frac{[b^T - x^T(\frac{1}{2}\tilde{A})]p}{p^T A p} = \frac{(b - \frac{1}{2}\tilde{A}x)^T p}{p^T A p}, \\
 f''(\alpha) &= p^T A p.
 \end{aligned} \tag{5.10}$$

Ein Iterationslauf mit der Suchrichtung $r(x)$ und der Schrittzahl $\bar{\alpha}$ anstelle von α bringt gar nichts, weder eine Konvergenz gegen x^* noch gegen z . Was z betrifft, zeigt sich dies auch dadurch, dass die Richtungen $r(x)$ fast senkrecht zu den Suchrichtungen $\tilde{r}(x)$ stehen bzw. fast parallel zu den Höhenlinien von $Q(x)$ verlaufen und damit sich kein Fortschritt in Richtung der Stelle z einstellt. Außerdem werden die Schrittzahlen $\bar{\alpha}$ betragsmäßig sehr klein.

Beispiel 5.4 Sei $Ax = b$, $A \neq A^T$, $A > 0$, aus Beispiel 1.6.

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

Die Funktionale sind zu $Bx = A^T A x = A^T b = c$

$$R(x_1, x_2) = \frac{1}{2}(Ax)^T Ax - (Ax)^T b = 2x_1^2 + 2x_1x_2 + 5x_2^2 - 6x_1 - 12x_2$$

mit seinem Minimum bei $x^* = (1, 1)^T$ und $R(x^*) = -9$,

zu $Ax = b$

$$Q(x_1, x_2) = \frac{1}{2}x^T Ax - x^T b = x_1^2 + \frac{1}{2}x_1x_2 + \frac{3}{2}x_2^2 - 3x_1 - 3x_2, \quad Q(x)^T = Q(x),$$

wobei sein Minimum bei $z = (\frac{30}{23}, \frac{18}{23})^T = (1.304\,347\dots, 0.782\,608\dots)^T$ liegt und $Q(z) = -\frac{72}{23} = -3.130\,434\dots$, $Q(x^*) = -3$, $R(z) = -\frac{4608}{529} = -8.710\,775\dots$ ist.

Zum LGS $\frac{1}{2}\tilde{A}x = \frac{1}{2}(A + A^T)x = b$ haben wir das Funktional

$$T(x_1, x_2) = \frac{1}{4}x^T(A + A^T)x - x^Tb = \frac{1}{4}x^T\tilde{A}x - x^Tb = Q(x_1, x_2), \quad \tilde{A} = \tilde{A}^T > 0,$$

eingeführt, sowie zu $\tilde{A}x = b$ das zusätzliche Funktional

$$S(x_1, x_2) = \frac{1}{2}x^T\tilde{A}x - x^Tb = 2x_1^2 + x_1x_2 + 3x_2^2 - 3x_1 - 3x_2 = 2Q(x_1, x_2) + x^Tb.$$

Das Minimum von $S(x)$ ist bei $t = \frac{z}{2} = (\frac{15}{23}, \frac{9}{23})^T$ und

$$S(t) = -\frac{36}{23} = \frac{1}{2}Q(z) = 2Q(t) + t^Tb = -1.565\ 217\dots, \quad Q(t) = -\frac{54}{23}, \quad R(t) = -\frac{3429}{529}.$$

Auf der Höhenlinie $S(x) = 0$ liegen u. a. die Punkte x^* , z und 0 .

GV - Höhenlinien von $R(x), Q(x), S(x)$ in $[-1, 2] \times [-1, 2]$

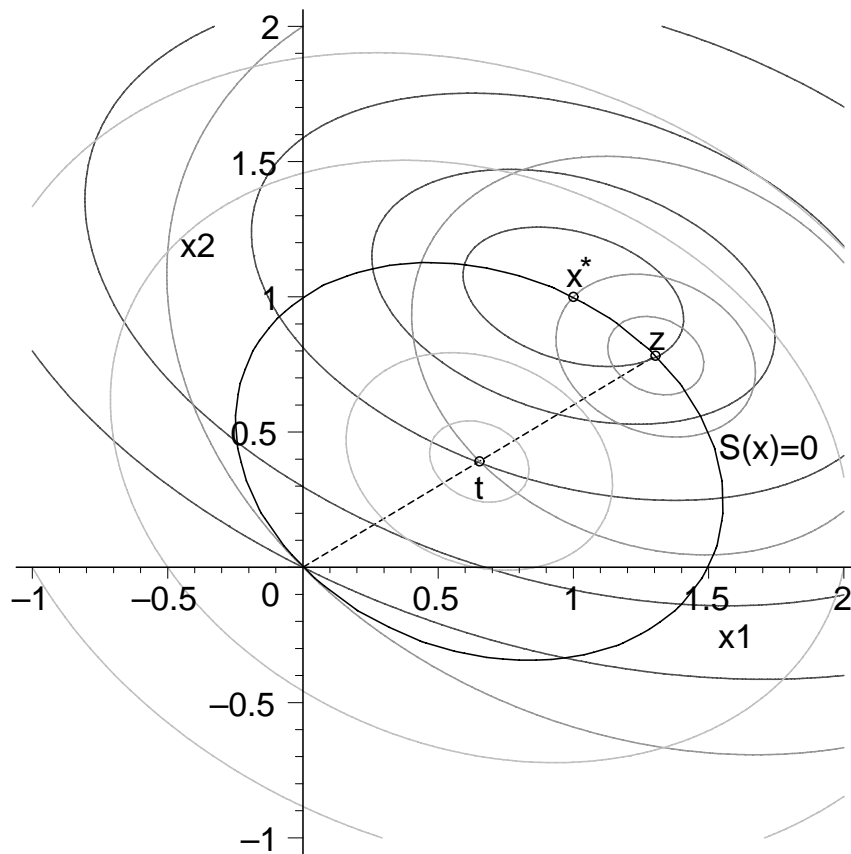


Abb. 5.6 Datei *gv_0030.ps*

Höhenlinienbilder (Ellipsenscharen) zu den Funktionalen

$Q(x)$, Minimum bei $z = (\frac{1}{2}\tilde{A})^{-1}b \neq x^*$, contours=[-3.130, -3.1, -3, -2.35, 0],

$R(x)$, Minimum bei $x^* = A^{-1}b$, contours=[-9, -8.7, -8, -6.45, -3.13, 0] und

$S(x)$, Minimum bei $t = \tilde{A}^{-1}b = \frac{z}{2}$, contours=[-1.565, -1.5, -1.1, 0, 2, 5]

Das GV für $A^T A x = A^T b$, also zum Funktional $R(x)$, mit der Richtung des steilsten Abstiegs $\hat{r}(x) = A^T(b - Ax) = \text{rd}(x)$ konvergiert. Dabei muss man aber die allgemein schlechtere Kondition der spd Matrix $A^T A$ in Kauf nehmen.

GV fuer $A^T A x = A^T b$ - Iterationsverlauf, $x(0)=[1,-1]$

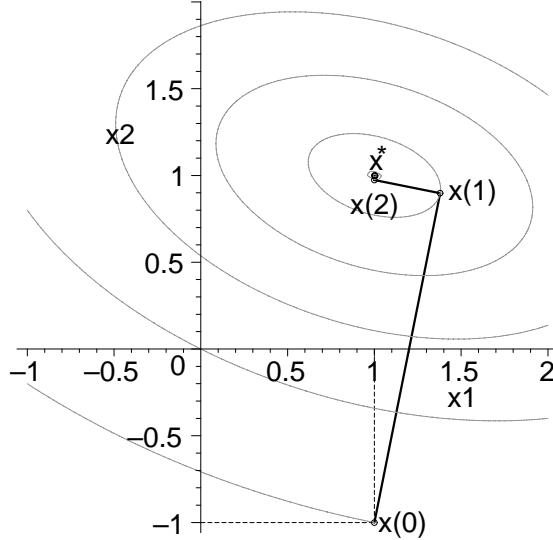


Abb. 5.7 Datei *gv_0031.ps*
 Höhenlinienbild und GV zu $R(x)$,
 Minimum bei $x^* = A^{-1}b = (1, 1)^T$,
 $R(x^*) = -9$,
 Abstiegsrichtung $\hat{r}(x) = A^T(b - Ax)$,
 $\text{contours} = [-9, -8.999, -8.9965,$
 $-8.7372, -7.5, -5, 0, 11]$

GV fuer $A^T A x = A^T b$, $A \gg A'$, $A > 0$ - konvergenter Iterationsverlauf

m	[x(m) [1], x(m) [2]]	[rd(m) [1], rd(m) [2]]
0	[1.0000000000, -1.0000000000]	[4.000e+00, 2.000e+01]
1	[1.3795620440, 0.8978102190]	[-1.314e+00, 2.628e-01]
2	[1.0000000000, 0.9737226277]	[5.255e-02, 2.628e-01]
3	[1.0049869470, 0.9986573605]	[-1.726e-02, 3.453e-03]
4	[1.0000000000, 0.9996547499]	[6.905e-04, 3.453e-03]
5	[1.0000655220, 0.9999823595]	[-2.268e-04, 4.536e-05]
6	[1.0000000000, 0.9999954639]	[9.072e-06, 4.536e-05]

Das GV für $Ax = b$, also zum Funktional $Q(x)$, mit der Suchrichtung $r(x) = b - Ax$ konvergiert ebenfalls gegen x^* .

Dabei ist sowohl $r(x) \perp Q(x) = \text{const}$ als auch $r(x) \perp R(x) = \text{const}$.

Ersteres ist sinnvoll, denn sonst würden wir mit $r(x)$ zur Minimumstelle z kommen. Die andere Bedingung bedeutet, dass $r(x)$ eine "schräge" Abstiegsrichtung bez. $R(x)$ darstellt und sich nahe seiner Richtung des steilsten Abstiegs $\hat{r}(x) = A^T(b - Ax)$ bewegt. Der Vektor $\hat{r}(x)$ ist in der folgenden Abbildung nicht eingezeichnet.

Die Abstiegsrichtung $\tilde{r}(x) = b - \frac{1}{2}\tilde{A}x = \text{rs}(x)$, senkrecht zu den Höhenlinien von $Q(x) = T(x)$, ist die Richtung des steilsten Abstiegs zum LGS $\frac{1}{2}\tilde{A}x = b$ und liefert eine gegen die Stelle z konvergierende Folge.

GV fuer $Ax=b$ und $1/2(A+A')x=b$ - Iterationsverlaeuft, $x(0)=[1,-1]$

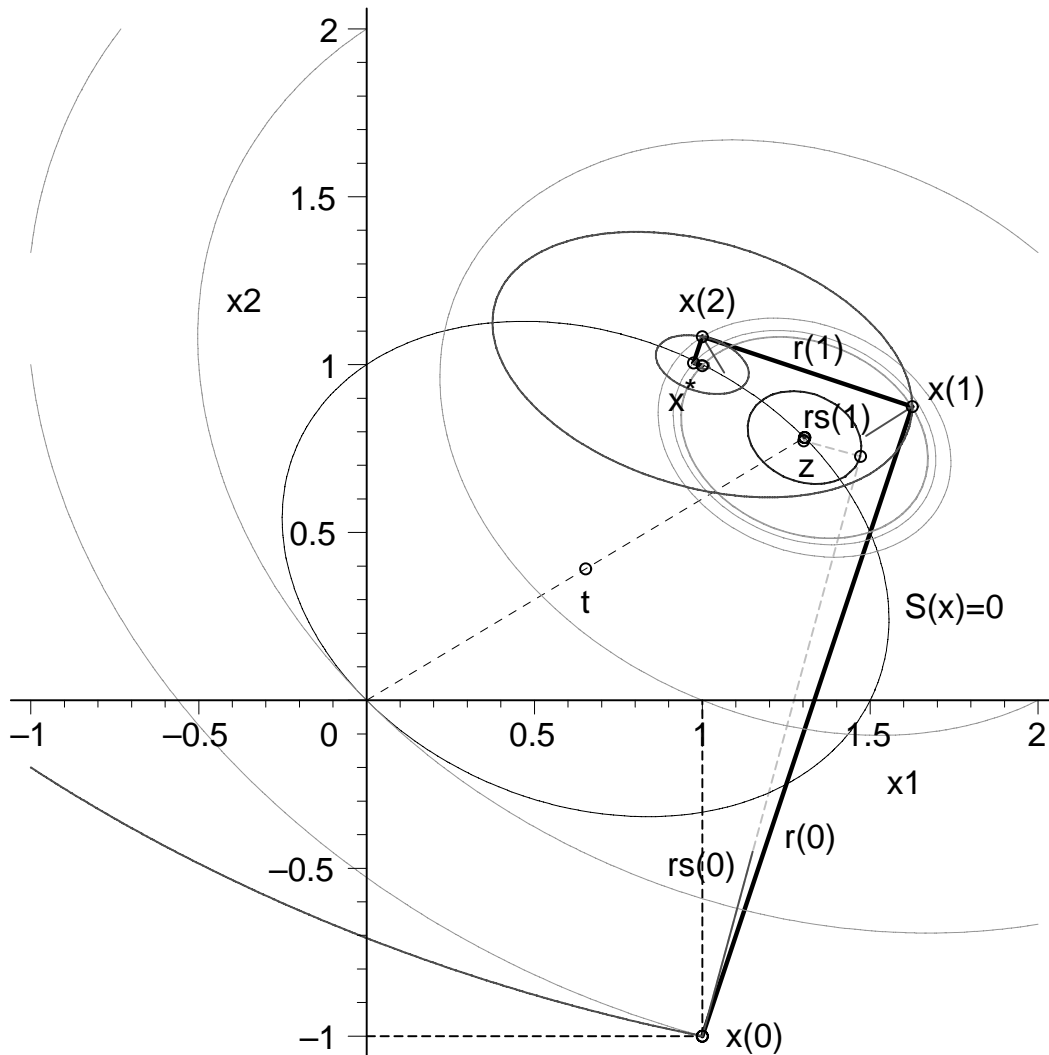


Abb. 5.8 Datei gv_003.ps

Höhenlinienbild und GV zu $Q(x)$ mit den Iterierten $x^{(m)}$ und Grenzvektor x^*

(Iterationsverlauf fett gezeichnet), Such-/Abstiegsrichtung $r(x) = b - Ax$,

$Q(x^*) = -3$, $Q(z) = -\frac{72}{23} = -3.130434$, Höhenlinien zu $Q(x)$ mit

contours=[2,0,-2,-2.9479,-2.9837,-3,-3.0017,-3.10265,-3.13,-3.1302].

Zusätzlich eingezeichnet:

- Vektor $\tilde{r}(x) = \mathbf{rs}(x) \perp Q(x) = \text{const}$ in den ersten Iterierten,
- Iterationsverlauf für $\frac{1}{2}\tilde{A}x = b$ (Funktional $Q(x) = T(x)$) mit Abstiegsrichtung $\tilde{r}(x)$, Konvergenz gegen z (dünn gestrichelt),
- Höhenlinien von $R(x)$ an den Iterierten, contours=[11,-8.2969,-8.9653,-8.999],
- Funktional $S(x)$ mit Minimumstelle $t = \frac{z}{2}$ und Höhenlinie $S(x) = 0$.

Berechnungen (2 Iterationsverläufe) in Maple mit `Digits:=10`

GV fuer $Ax=b$, $A \ll A'$, $A > 0$ - konvergenter Iterationsverlauf

m	[x(m) [1], x(m) [2]]	[r(m) [1], r(m) [2]]
0	[1.0000000000, -1.0000000000]	[2.000e+00, 6.000e+00]
1	[1.6250000000, 0.8750000000]	[-1.125e+00, 3.750e-01]
2	[1.0000000000, 1.0833333330]	[-8.333e-02, -2.500e-01]
3	[0.9739583334, 1.0052083330]	[4.688e-02, -1.563e-02]
4	[1.0000000000, 0.9965277774]	[3.472e-03, 1.042e-02]
5	[1.0010850690, 0.9997829858]	[-1.953e-03, 6.510e-04]
6	[0.9999999995, 1.0001446760]	[-1.447e-04, -4.340e-04]

GV fuer $1/2(A+A')x=b$, $A \ll A'$, $A > 0$ - Iterationsverlauf

m	[x(m) [1], x(m) [2]]	[rs(m) [1], rs(m) [2]]
0	[1.0000000000, -1.0000000000]	[1.500e+00, 5.500e+00]
1	[1.4710144930, 0.7270531400]	[-3.056e-01, 8.333e-02]
2	[1.3026999920, 0.7729570952]	[8.121e-03, 2.978e-02]
3	[1.3052502120, 0.7823079007]	[-1.654e-03, 4.512e-04]
4	[1.3043389050, 0.7825564392]	[4.397e-05, 1.612e-04]
5	[1.3043527130, 0.7826070673]	[-8.957e-06, 2.443e-06]
6	[1.3043477790, 0.7826084130]	[2.380e-07, 8.730e-07]
z	[1.3043478260, 0.7826086956]	

Wir machen noch einen Test mit dem modifizierten GV für $Ax = b$, also zum Funktional $Q(x)$, mit der Suchrichtung $r(x) = b - Ax$ und der Schrittzahl $\bar{\alpha}$ (5.10). Dieses AV konvergiert weder gegen $x^* = (1, 1)^T$ noch gegen $z = (1.304\ 347\dots, 0.782\ 608\dots)^T$.

Modifiziertes GV fuer $Ax=b$, $A \ll A'$, $A > 0$ - Iterationsverlauf

m	[x(m) [1], x(m) [2]]	[r(m) [1], r(m) [2]]
0	[1.0000000000, -1.0000000000]	[2.000e+00, 6.000e+00]
1	[1.5625000000, 0.6875000000]	[-8.125e-01, 9.375e-01]
2	[1.4284077020, 0.8422218826]	[-6.990e-01, 4.733e-01]
3	[1.3858530230, 0.8710366422]	[-6.427e-01, 3.869e-01]
4	[1.3765936480, 0.8766101947]	[-6.298e-01, 3.702e-01]
5	[1.3747517940, 0.8776927618]	[-6.272e-01, 3.669e-01]
6	[1.3743921230, 0.8779031764]	[-6.267e-01, 3.663e-01]
7	[1.3743221420, 0.8779440793]	[-6.266e-01, 3.662e-01]
8	[1.3743085360, 0.8779520307]	[-6.266e-01, 3.661e-01]
9	[1.3743058910, 0.8779535765]	[-6.266e-01, 3.661e-01]
10	[1.3743053770, 0.8779538768]	[-6.266e-01, 3.661e-01]

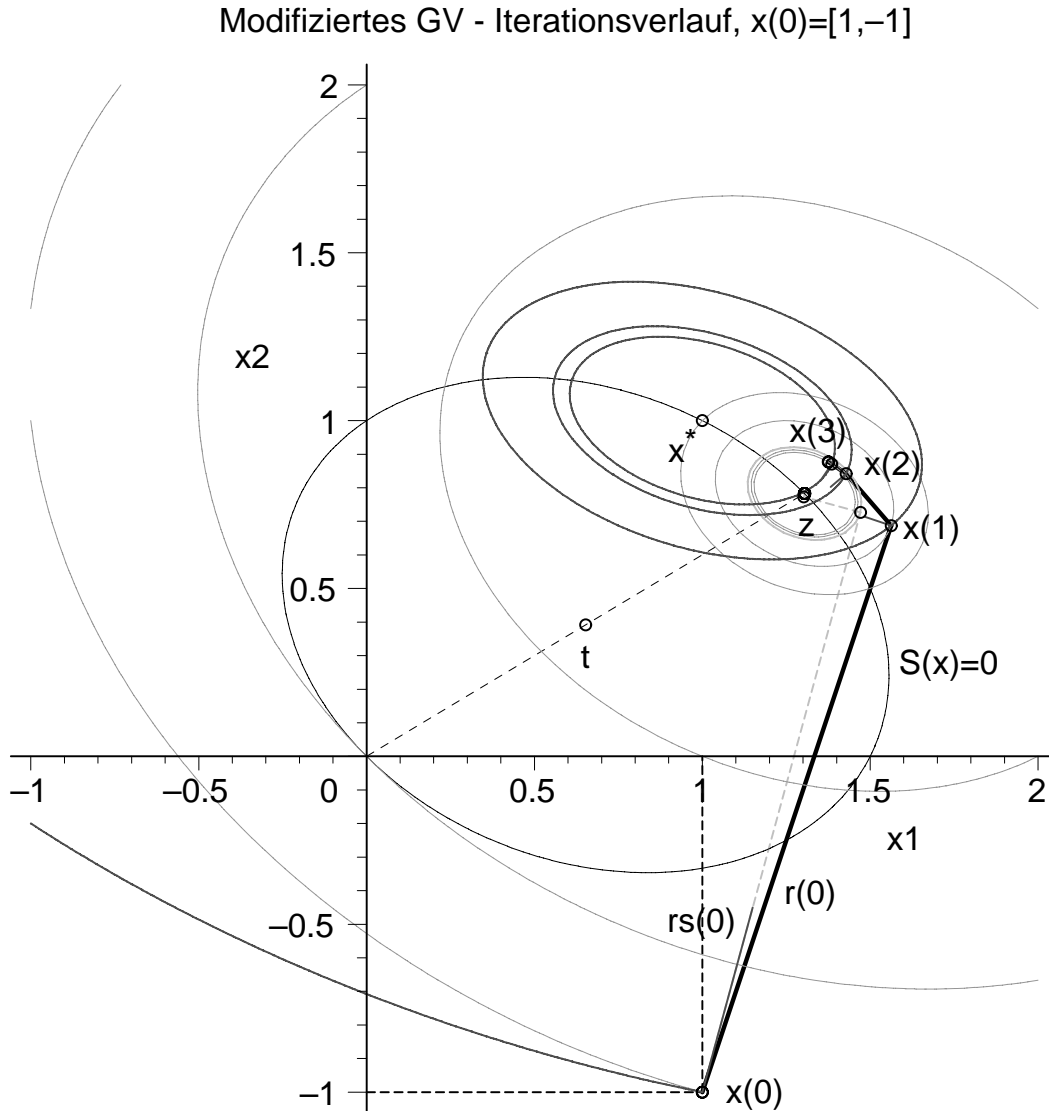


Abb. 5.9 Datei *lgs10_02.ps*

Höhenlinienbild und modifiziertes GV zu $Q(x)$ mit den Iterierten $x^{(m)}$

(Iterationsverlauf fett gezeichnet), Suchrichtung $r(x) = b - Ax$, Schrittzahl $\bar{\alpha}$,
 $Q(x^*) = -3$, $Q(z) = -3.130434$, Höhenlinien zu $Q(x)$ mit
 $\text{contours}=[2, 0, -2, -3, -3.0625, -3.10265, -3.1060, -3.1085]$.

Zusätzlich eingezeichnet:

- Vektor $\tilde{r}(x) = \text{rs}(x) \perp Q(x) = \text{const}$ in den ersten Iterierten,
- Iterationsverlauf für $\frac{1}{2}\tilde{A}x = b$ (Funktional $Q(x) = T(x)$) mit Abstiegsrichtung $\tilde{r}(x)$,
Konvergenz gegen z (dünn gestrichelt),
- Höhenlinien von $R(x)$ an den Iterierten, $\text{contours}=[11, -8.230, -8.644, -8.719]$,
- Funktional $S(x)$ mit Minimumstelle $t = \frac{z}{2}$ und Höhenlinie $S(x) = 0$.

Im folgenden Beispiel erhöhen wir unter Beibehaltung der Matrixeigenschaften die Dimension des LGS um Eins.

Beispiel 5.5 Sei $Ax = b$, $A \neq A^T$, $A > 0$.

$$\begin{pmatrix} 2 & 1 & 0 \\ 0 & 3 & 1 \\ 1 & 2 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \\ 7 \end{pmatrix}, \quad x^* = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad x^{(0)} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}.$$

Zu $Bx = A^T Ax = A^T b = c$ ist das Funktional $R(x)$ mit seinem Minimum bei $x^* = (1, 1, 1)^T$ und $R(x^*) = -37$, zu $Ax = b$ ist es $Q(x)$, wobei sein Minimum bei $z = \left(\frac{38}{37}, \frac{16}{37}, \frac{54}{37}\right)^T = (1.027\,027\dots, 0.432\,432\dots, 1.459\,459\dots)^T$ liegt und $Q(z) = -\frac{278}{37} = -7.513\,513\dots$, $Q(x^*) = -7$, $R(z) = -\frac{49050}{1369} = -35.829\,072\dots$ sind.

Das GV konvergiert langsam zur Lösung x^* .

Die Iterationsvektoren $x^{(i)}, r^{(i)}$, $i = 0, 1, 2, 3$, aus den ersten drei Schritten sind die Spalten der folgenden Felder.

$$xv = \begin{bmatrix} 1 & \frac{281}{203} & \frac{3622049}{2732264} & \frac{1754920830948748060433}{1696699048723738646552} \\ -1 & \frac{10}{29} & \frac{10360655}{9562924} & \frac{247651939900413696913}{242385578389105520936} \\ 0 & \frac{312}{203} & \frac{2164383}{2390731} & \frac{2919010180273619953755}{3393398097447477293104} \end{bmatrix}$$

$$rv = \begin{bmatrix} 2 & \frac{-23}{203} & \frac{-3513113}{4781462} & \frac{-153308095029176059601}{1696699048723738646552} \\ 7 & \frac{10}{7} & \frac{-212543}{1366132} & \frac{253200733698913948315}{3393398097447477293104} \\ 8 & \frac{-248}{203} & \frac{-2176283}{19125848} & \frac{28166378998772096591}{58506863749094436088} \end{bmatrix}$$

Rechnung mit Maple bei `Digits:=16`

GV fuer Ax=b, A<>A', A>0 - konvergenter Iterationsverlauf

m	[x(m)[1],	x(m)[2],	x(m)[3]] [r(m)[1],	r(m)[2],	r(m)[3]
0	[1.000000000000,	-1.000000000000,	0.000000000000]	[2.000e+00,	7.000e+00,	8.000e+00]
1	[1.384236453202,	0.344827586207,	1.536945812808]	[-1.133e-01,	1.429e+00,	-1.222e+00]
2	[1.325658501521,	1.083419150879,	0.905322681640]	[-7.347e-01,	-1.556e-01,	-1.138e-01]
3	[1.034314737354,	1.021727206488,	0.860202692537]	[-9.036e-02,	7.462e-02,	4.814e-01]
4	[1.012516415105,	1.039728053689,	0.976344076787]	[-6.476e-02,	-9.553e-02,	2.651e-03]
5	[0.991491148401,	1.008713858735,	0.977204805683]	[8.304e-03,	-3.346e-03,	8.226e-02]
6	[0.993591623354,	1.007867384386,	0.998013129492]	[4.949e-03,	-2.162e-02,	-1.379e-03]
7	[0.995295473273,	1.000426193743,	0.997538431425]	[8.983e-03,	1.183e-03,	1.370e-02]
8	[0.997501302204,	1.000716687839,	1.000902210652]	[4.281e-03,	-3.052e-03,	-2.544e-03]
9	[0.999127090799,	0.999557451371,	0.999936195955]	[2.188e-03,	1.391e-03,	2.013e-03]
10	[0.999624140102,	0.999873494891,	1.000393464272]	[8.782e-04,	-1.395e-05,	-9.450e-04]
11	[0.999963102068,	0.999868111120,	1.000028734645]	[2.057e-04,	3.669e-04,	1.857e-04]
12	[1.000009140699,	0.999950241892,	1.000070308377]	[3.148e-05,	7.897e-05,	-1.909e-04]
13	[1.000020818943,	0.999979539222,	0.999999497734]	[-2.118e-05,	6.188e-05,	2.211e-05]
14	[1.000014762251,	0.999997238331,	1.000005821714]	[-2.676e-05,	2.463e-06,	-3.253e-05]
15	[1.000007135008,	0.999997940354,	0.999996552069]	[-1.221e-05,	9.627e-06,	1.078e-05]
16	[1.000003172359,	1.000001064576,	1.00000049222]	[-7.409e-06,	-3.243e-06,	-5.498e-06]
17	[1.000001310049,	1.000000249467,	0.999998667211]	[-2.870e-06,	5.844e-07,	3.522e-06]
18	[1.000000331167,	1.000000448817,	0.99999868714]	[-1.111e-06,	-1.215e-06,	-7.037e-07]
19	[1.000000068752,	1.000000161837,	0.999999702535]	[-2.993e-07,	-1.880e-07,	7.974e-07]
20	[0.999999965066,	1.000000096702,	0.999999978751]	[-2.683e-08,	-2.689e-07,	-7.347e-08]
25	[0.999999997792,	0.999999998586,	1.000000000282]	[5.830e-09,	3.959e-09,	3.906e-09]
30	[1.000000000049,	0.999999999989,	1.000000000022]	[-8.640e-11,	1.192e-11,	-1.128e-10]
35	[1.000000000000,	1.000000000001,	0.999999999999]	[-1.117e-12,	-5.920e-13,	2.697e-12]

Die Residua genügen der Beziehung $r^{(k-1)T} r^{(k)} = 0$.

Ein bisschen komplizierter wird es für den allgemeinen Fall, wenn die Matrix nicht symmetrisch und indefinit ist. Die genannten Stellen z und t werden nun Sattelpunkte. Wir erläutern die Situation anhand des Beispiels 1.7.

Beispiel 5.6

Sei $A \neq A^T$ und A indefinit aus Beispiel 1.7.

$$\begin{pmatrix} 0.780 & 0.563 \\ 0.913 & 0.659 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0.217 \\ 0.254 \end{pmatrix}, \quad x^* = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad x^{(0)} = \begin{pmatrix} 1.2 \\ -1.2 \end{pmatrix}.$$

Die Funktionale $Q(x)$ und $R(x)$ haben wir im Beispiel 1.7 schon beschrieben. Dazu kommt noch das Funktional $S(x)$.

Wir fassen diese in einer Übersicht zusammen.

Funktional	$x^* = A^{-1}b$ Lösung $x^* = (1, -1)^T$	$z = (\frac{1}{2}\tilde{A})^{-1}b$ Sattelpunkt von Q $z = (\frac{44449}{30624}, -\frac{6329}{5104})^T$ $= \begin{pmatrix} 1.4514433124346917 \\ -1.2400078369905956 \end{pmatrix}$
		$t = \tilde{A}^{-1}b$ Sattelpunkt von S $t = \frac{z}{2} = (\frac{44449}{61248}, -\frac{6329}{10208})^T$ $= \begin{pmatrix} 0.7257216562173458 \\ -0.6200039184952978 \end{pmatrix}$
$Q(x) = \frac{1}{2}x^T Ax - x^T b$ $= \frac{1}{2}x^T (\frac{1}{2}\tilde{A})x - x^T b$ $= T(x)$	$Q(x^*) = -\frac{1}{2}x^{*T}b$ $= \frac{37}{2000}$ $= 0.0185$	$Q(z) = -\frac{37}{61248000}$ $= -0.6041013584117029 \cdot 10^{-6}$ $Q(t) = -\frac{37}{81664000}$ $= -0.4530760188087774 \cdot 10^{-6}$
$R(x) = \frac{1}{2}(Ax)^T Ax - (Ax)^T b$	$R(x^*) = -\frac{1}{2}\ b\ _2^2$ $= -\frac{22321}{400000}$ $= -0.0558025$	$R(z) = \frac{357444929}{375131750400000}$ $= 0.952851707750301 \cdot 10^{-6}$ $R(t) = -\frac{83733158005183}{1500527001600000}$ $= -0.0558024999989330$
$S(x) = \frac{1}{2}x^T \tilde{A}x - x^T b$ $= 2Q(x) + x^T b$	$S(x^*) = 0$ $S(0) = 0$	$S(z) = 0$ $S(t) = \frac{1}{2}Q(z) = -\frac{37}{122496000}$ $= -0.3020506792058516 \cdot 10^{-6}$

Tab. 5.1

Funktionale $Q(x)$, $R(x)$ und $S(x)$ und Funktionswerte an ausgewählten Stellen

Im Höhenlinienbild mit den drei Funktionalen treten einige Besonderheiten auf. Die Höhenlinie $S(x) = 0$ ist im Gegensatz zum Beispiel 5.4 keine Ellipse mehr. Die Niveaulinien $S(x) = S(t) = -0.000\,000\,302\dots$ sind sich kreuzenden Kurven und $S(x) = 0$ zwei sich fast berührende Hyperbeläste. Sie sind lokal nur mit feiner Auflösung zu unterscheiden. Die Punkte 0 und x^* liegen auf dem linken unteren Ast, während der Punkt z auf dem rechten oberen sich befindet. Die Tallinie der sehr lang gestreckten Ellipsenschar zum Funktional $R(x)$ liegt nahe und fast parallel zu einem Zweig der Höhenlinien von $S(x) = S(t)$ und $R(t) \approx R(x^*)$. Mit dem anderen Zweig fällt ein Ast der Höhenlinie $Q(x) = Q(z) = -0.000\,000\,604\dots$ fast zusammen.

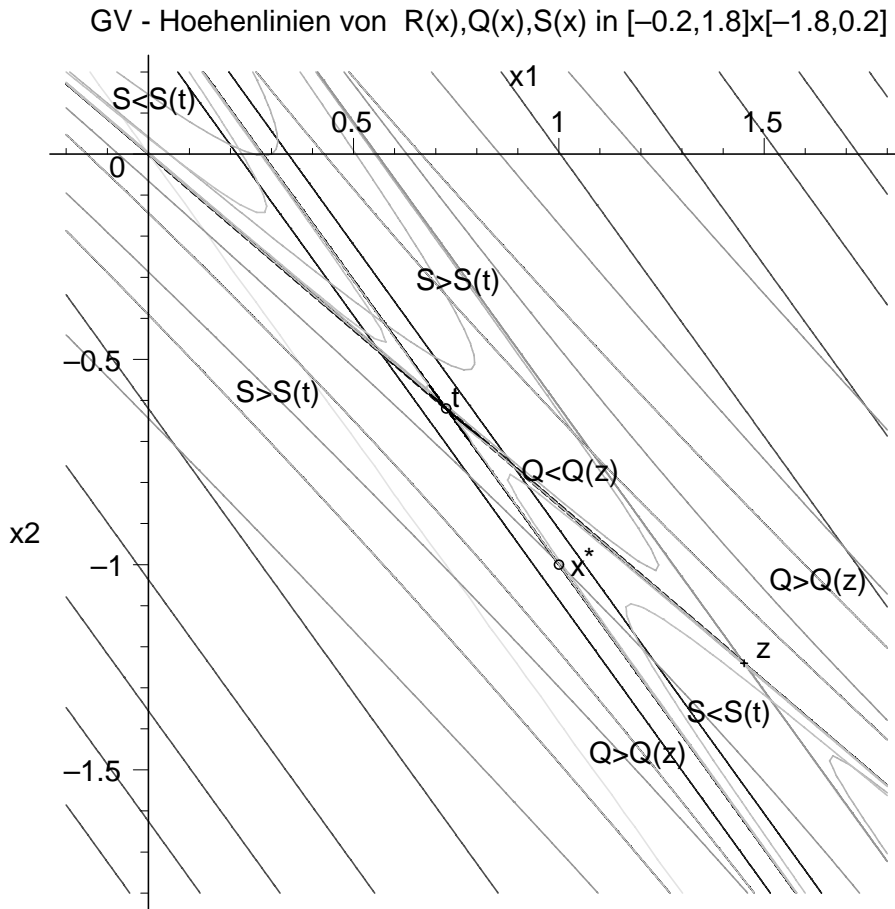


Abb. 5.10 Datei *gv_0040.ps*

Höhenlinienbilder zu den Funktionalen

$Q(x)$ (Hyperbelschar), Sattel bei $z = (\frac{1}{2}\tilde{A})^{-1}b \neq x^*$,

`contours=[0.2,0.1,0.01776,0,-0.00000060410,-0.001,-0.01,-0.03]`,

$R(x)$ (Ellipsenschar), Minimum bei $x^* = A^{-1}b$,

`contours=[-0.0558025,-0.0558,-0.053,0]` und `contours=5`,

$S(x)$ (Hyperbelschar), Sattel bei $t = \tilde{A}^{-1}b = \frac{z}{2}$,

`contours=[0.2,0.05,0,-0.00000030205,-0.0000261,-0.001,-0.00888]`

Das GV für $A^T A x = A^T b$, also zum Funktional $R(x)$, mit der Richtung des steilsten Abstiegs $\hat{r}(x) = A^T(b - Ax)$ konvergiert. Dabei muss man aber die allgemein schlechtere Kondition der spd Matrix $A^T A$ in Kauf nehmen.

GV fuer A'Ax=A'b - Iterationsverlauf, x(0)=[1.2,-1.2]

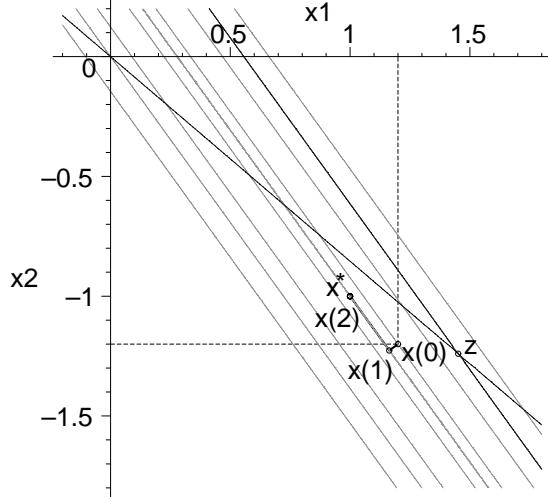


Abb. 5.11 Datei *gv_0041.ps*
 Höhenlinienbild und GV zu $R(x)$,
 Minimum bei $x^* = A^{-1}b = (1, -1)^T$,
 $R(x^*) = -0.0558025$,
 Abstiegsrichtung $\hat{r}(x) = A^T(b - Ax)$,
 $\text{contours}=[0.05, 0, -0.03, -0.05357, -0.0558, -0.0558025]$,
 dazu Höhenlinie
 $Q(x) = Q(z) = -0.000000604\dots$

Berechnungen in Maple mit `Digits:=24`

GV fuer $A'Ax=A'b$, $A \gg A'$, $A > 0$ - konvergenter Iterationsverlauf

m	$x(m) [1]$,	$x(m) [2]$	$rd(m) [1]$, $rd(m) [2]$
0	[1.200000000000000000000000]	[-1.200000000000000000000000]	[-8.0e-02, -5.7e-02]
1	[1.16341797148391625295701,	-1.22640475027802150139476]	[-7.4e-14, 1.0e-13]
2	[1.00000000000215594235871,	-1.00000000000214027421180]	[-8.8e-13, -6.3e-13]
3	[1.0000000000175416320875,	-1.00000000000243027666798]	[-1.0e-24, 2.0e-24]
4	[1.0000000000175416320873,	-1.00000000000243027666794]	[-1.2e-23, -6.0e-24]
5	[1.0000000000175416320872,	-1.00000000000243027666794]	[-1.0e-24, 2.0e-24]
6	[1.0000000000175416320870,	-1.00000000000243027666790]	[-1.2e-23, -6.0e-24]

Die schlechte Kondition der Matrix $\text{cond}_2(A^T A) \approx 0.5 \cdot 10^{13}$ gestattet keine weitere signifikante Verbesserung der Iterierten (nur Veränderungen in den letzten Dezimalstellen). Im Beispiel 3.6, wo wir das GV für das Normalgleichungssystem testeten, haben wir den Einfluss des Gleitpunktformats `Digits:=k` in Maple auf die numerische Rechnung verdeutlicht. Wir mussten schon über `Digits:=16` hinausgehen, um den Verlust von ungefähr 12 Dezimalstellen durch die Matrixkondition zu kompensieren und zumindest nach wenigen Schritten in den Iterierten einige Nullen nach dem Komma zu garantieren.

Das GV divergiert für $Ax = b$, also zum Funktional $Q(x)$, mit der Suchrichtung $r(x) = b - Ax$, die sowohl $\nabla Q(x) = \text{const}$ als auch $\nabla R(x) = \text{const}$ ist. Der Vektor $\tilde{r}(x) = b - \frac{1}{2} \tilde{A}x \perp$ Höhenlinien von $Q(x) = T(x)$ ist die Suchrichtung (Abstieg/Anstieg) zum LGS $\frac{1}{2} \tilde{A}x = b$ und liefert eine gegen z konvergierende Folge.

GV fuer $Ax=b$ und $1/2(A+A')x=b$ - Iterationsverlauf, $x(0)=[1.2,-1.2]$

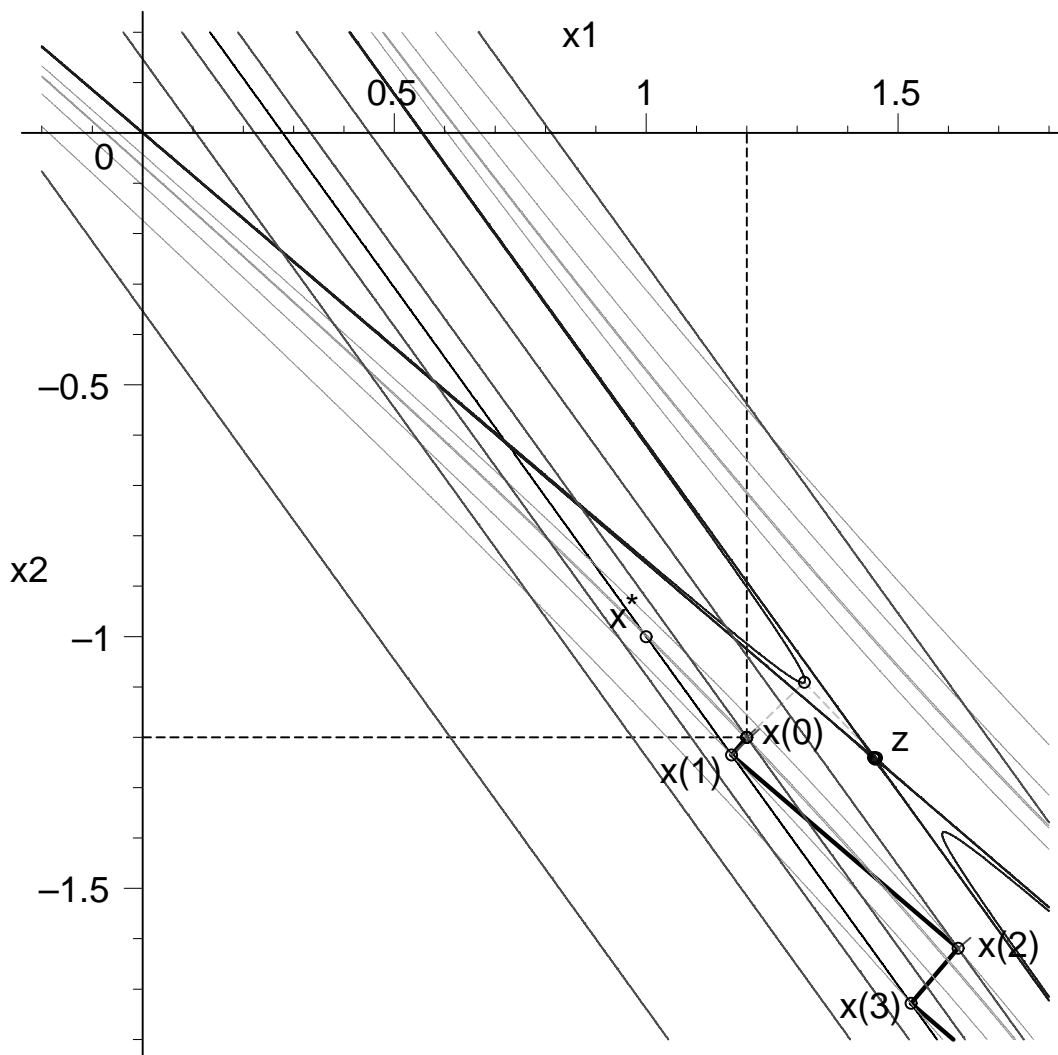


Abb. 5.12 Datei *gv_004.ps*

Höhenlinienbild und GV zu $Q(x)$ mit den Iterierten $x^{(m)}$

(divergenter "Zick-Zack"- Iterationsverlauf fett gezeichnet),

$Q(x^*) = 0.0185$, $Q(z) = -0.000\,000\,604\dots$, Suchrichtung $r(x) = b - Ax$,

Vektor $\tilde{r}(x) \perp Q(x) = \text{const}$ in den Iterierten eingetragen,

Höhenlinien zu $Q(x)$ mit `contours`=[-0.0495,-0.000431,-0.000000604,
0.00000984,0.0114,0.01776,0.0185,0.02995,0.05397].

Zusätzlich eingezeichnet:

- Iterationsverlauf für $\frac{1}{2}\tilde{A}x = b$ (Funktional $Q(x) = T(x)$) mit Abstiegsrichtung $\tilde{r}(x)$, Konvergenz gegen z (dünn gestrichelt),
- Höhenlinien von $R(x)$ an den Iterierten,
`contours`=[-0.0558025,-0.05357,-0.034405,0.14931],
- Funktional $S(x)$ mit Höhenlinie $S(x) = 0$.

Berechnungen (2 Iterationsverläufe) in Maple mit `Digits:=24`
(Ausgabe der Iterierten mit 10 Nachkommastellen)

Divergenz des GV für $Ax = b$

GV fuer $Ax=b$, $A \succ A'$, A indefinit - divergenter Iterationsverlauf

m	[x(m) [1], x(m) [2]]	[r(m) [1], r(m) [2]]
0	[1.2000000000, -1.2000000000]	[-4.340e-02, -5.080e-02]
1	[1.1698401737, -1.2353022851]	[-1.490e-07, 1.273e-07]
2	[1.6192287222, -1.6192287222]	[-1.344e-01, -1.573e-01]
3	[1.5258495688, -1.7285296668]	[-4.613e-07, 3.941e-07]
4	[2.9172210521, -2.9172210521]	[-4.160e-01, -4.870e-01]
5	[2.6281057829, -3.2556324733]	[-1.428e-06, 1.220e-06]
6	[6.9359917117, -6.9359917117]	[-1.288e+00, -1.508e+00]
7	[6.0408493179, -7.9837620713]	[-4.422e-06, 3.778e-06]
8	[19.3786828140, -19.3786828140]	[-3.988e+00, -4.668e+00]
xs	[1, -1]	

Konvergenz des GV für $\frac{1}{2}\tilde{A}x = b$ gegen z , $\tilde{r}(x) = b - \frac{1}{2}\tilde{A}x = rs(x)$

GV fuer $1/2(A+A')x=b$, $A \succ A'$, A indefinit - Iterationsverlauf

m	[x(m) [1], x(m) [2]]	[rs(m) [1], rs(m) [2]]
0	[1.2000000000, -1.2000000000]	[1.666e-01, 1.592e-01]
1	[1.3141497896, -1.0909204892]	[-2.938e-03, 3.074e-03]
2	[1.4575423387, -1.2409782698]	[-4.041e-03, -3.862e-03]
3	[1.4547735136, -1.2436241099]	[7.125e-05, -7.456e-05]
4	[1.4512953740, -1.2399842981]	[9.802e-05, 9.367e-05]
5	[1.4513625348, -1.2399201204]	[-1.728e-06, 1.809e-06]
6	[1.4514469008, -1.2400084080]	[-2.378e-06, -2.272e-06]
7	[1.4514452718, -1.2400099646]	[4.192e-08, -4.387e-08]
8	[1.4514432254, -1.2400078231]	[5.767e-08, 5.511e-08]
9	[1.4514432649, -1.2400077854]	[-1.017e-09, 1.064e-09]
10	[1.4514433145, -1.2400078373]	[-1.399e-09, -1.337e-09]
11	[1.4514433136, -1.2400078382]	[2.467e-11, -2.581e-11]
12	[1.4514433124, -1.2400078370]	[3.393e-11, 3.242e-11]
z	[1.4514433124, -1.2400078369]	

5.2 Verfahren der konjugierten Gradienten

5.2.1 CG mit symmetrischer indefiniter Matrix

Man bemerkt eine gewisse Ähnlichkeit zu den Betrachtungen des GV im Kapitel 5.1.1. Das betrifft zunächst die Geometrie und Sattelpunktsituation der Funktionale $Q(x)$ und $\delta_Q(x)$ an der Lösung x^* sowie die Lösungsmannigfaltigkeit.

Bei der Berechnung der Schrittzahl

$$\alpha = \frac{p^T r}{(Ap)^T p} = \frac{r^T p}{p^T Ap} = \frac{\|r\|_2^2}{p^T Ap} \quad (5.11)$$

kann der Nenner Null sein, also $p^T Ap = 0$.

Dann ist α nicht berechenbar, d. h. die Strahlenminimierung der Funktion $f(\alpha) = Q(x + \alpha p)$, erstes $p = r$, weitere Suchrichtungen $p = r + \beta p_{alt}$, ist geometrisch nicht durchführbar.

Dann ist auch in der Formel für die Größe

$$\beta = -\frac{p^T Ar_{neu}}{(Ap)^T p} = -\frac{p^T Ar_{neu}}{p^T Ap} = \frac{\|r_{neu}\|_2^2}{\|r\|_2^2} \quad (5.12)$$

der Nenner gleich Null. Das CG bricht ab.

Tritt dieser Sonderfall nicht auf, endet das CG normal und bei exakter Rechnung spätestens nach n Schritten an der Lösung x^* . Die Suchvektoren p sind A-orthogonal. Der Vektor p stellt entweder eine Abstiegs- oder Anstiegsrichtung dar.

Beispiel 5.7

Sei $Ax = b$, $A = A^T$ und indefinit aus den Beispielen 1.4, 5.2.

$$\begin{pmatrix} 1 & 0 \\ 0 & -4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad x^* = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Wir rechnen mit verschiedenen Startvektoren den ersten Schritt oder mehrere Anfangsschritte des CG.

$$\begin{aligned} (1) \quad x^{(0)} &= (1, 2)^T && \text{Startvektor,} \\ r^{(0)} &= b - Ax^{(0)} = (-1, 8)^T && \text{Anfangsresiduum,} \\ p^{(0)} &= r^{(0)} = (-1, 8)^T && \text{Anfangssuchrichtung.} \end{aligned}$$

$$\begin{aligned} \underline{\mathbf{S1}} \quad \alpha_0 &= \frac{\|r^{(0)}\|_2^2}{p^{(0)T} Ap^{(0)}} = \frac{65}{-255} = -\frac{13}{51}, \\ x^{(1)} &= x^{(0)} + \alpha_0 p^{(0)} = \frac{2}{51}(32, -1)^T = (1.254\dots, -0.039\dots)^T, \\ r^{(1)} &= b - Ax^{(1)} = r^{(0)} - \alpha_0 Ar^{(0)} = -\frac{8}{51}(8, 1)^T, \end{aligned}$$

$$\beta_0 = \frac{\|r^{(1)}\|_2^2}{\|r^{(0)}\|_2^2} = \frac{64}{2601},$$

$$p^{(1)} = r^{(1)} + \beta_0 p^{(0)} = \frac{104}{2601}(-32, 1)^T.$$

S2

$$\alpha_1 = \frac{51}{52},$$

$$x^{(2)} = (0, 0)^T,$$

$$r^{(2)} = (0, 0)^T \rightarrow \text{Stopp mit } x^{(2)} = x^*.$$

Das CG konvergiert mit Abstiegs- und Anstiegsrichtungen gegen x^* .

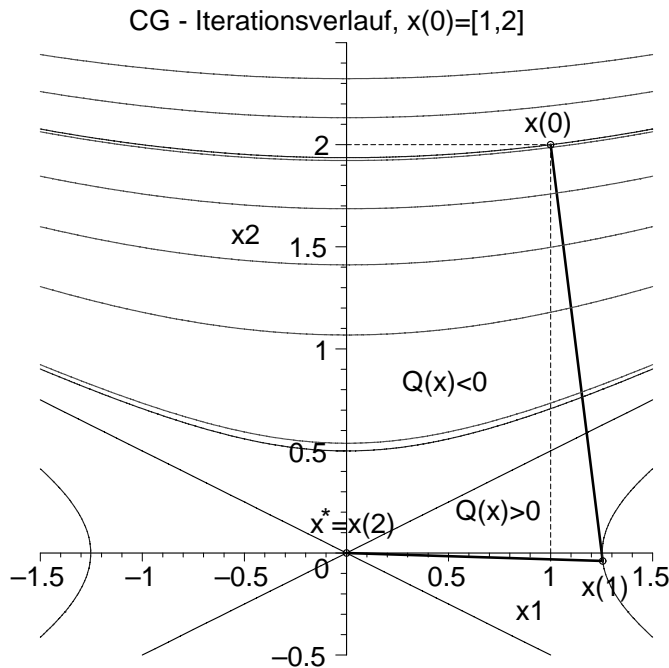


Abb. 5.13 Datei *cg_04.ps*
 Höhenlinienbild
 bei endlichem
 Iterationsverlauf des CG mit
 $x^{(0)} = (1, 2)^T$
 zu $Q(x) = \frac{1}{2}x_1^2 - 2x_2^2$, $Q(x^*) = 0$,
 mit contours=[-7.5, -0.5, 0,
 0.7843] und contours=7

```
Startvektor      x=[ 1.0000000000000000e+00  2.0000000000000000e+00]
Residuum        r = b-Ax=[-1.0000000000000000e+00  8.0000000000000000e+00]
Funktionswert   Q(x)= -7.5000000000000000e+00
Anfangsfehlerquadrat r'r=  6.5000000000000000e+01
```

Schritt k = 1

```
Suchrichtung    p=[-1.0000000000000000e+00  8.0000000000000000e+00]
Suchschritt     alpha= -2.5490196078431373e-01
Iterationsvektor x=[ 1.2549019607843137e+00 -3.9215686274509804e-02]
```

```
Residuum        r = b-Ax=[-1.2549019607843137e+00 -1.5686274509803922e-01]
Funktionswert   Q(x)=  7.8431372549019608e-01
Fehlernormquadrat r'r=  1.5993848519800077e+00
Schritt         beta=  2.4605920799692426e-02
neue Suchrichtung p=[-1.2795078815840062e+00  3.9984621299500192e-02]
```

Schritt k = 2

Suchrichtung $p = [-1.2795078815840062e+00 \quad 3.9984621299500192e-02]$
 Suchschritt $\alpha = 9.8076923076923077e-01$
 Iterationsvektor $x = [0.0000000000000000e-01 \quad 0.0000000000000000e-01]$

Residuum $r = b - Ax = [0.0000000000000000e-01 \quad 0.0000000000000000e-01]$
 Funktionswert $Q(x) = 0.0000000000000000e-01$
 Fehlernormquadrat $r^T r = 0.0000000000000000e-01$

CG fuer $Ax=b$, $A=A'$ indefinit - konv. endlicher Iterationsverlauf

m	[$x^{(m)}[1]$, $x^{(m)}[2]$]	[$r^{(m)}[1]$, $r^{(m)}[2]$]
0	[1.0000000000, 2.0000000000]	[-1.000e+00, 8.000e+00]
1	[1.2549019608, -0.0392156863]	[-1.255e+00, -1.569e-01]
2	[0.0000000000, 0.0000000000]	[0.000e-01, 0.000e-01]

$$(2) \quad x^{(0)} = (1, 0)^T, \quad r^{(0)} = b - Ax^{(0)} = (-1, 0)^T = p^{(0)}.$$

$$\begin{aligned} \underline{\mathbf{S1}} \quad \alpha_0 &= \frac{\|r^{(0)}\|_2^2}{p^{(0)T} A p^{(0)}} = \frac{1}{1} = 1, \\ x^{(1)} &= x^{(0)} + \alpha_0 p^{(0)} = (0, 0)^T, \\ r^{(1)} &= b - Ax^{(1)} = r^{(0)} - \alpha_0 A p^{(0)} = (0, 0)^T \Rightarrow \text{Stopp mit } x^{(1)} = x^*. \end{aligned}$$

Das CG braucht weniger als $n = 2$ Schritte.

$$(3) \quad x^{(0)} = (0, 1)^T, \quad r^{(0)} = (0, 4)^T = p^{(0)}.$$

$$\begin{aligned} \underline{\mathbf{S1}} \quad \alpha_0 &= \frac{16}{-64} = -\frac{1}{4}, \\ x^{(1)} &= (0, 0)^T, \\ r^{(1)} &= (0, 0)^T \Rightarrow \text{Stopp wie in (2)}. \end{aligned}$$

$$(4) \quad x^{(0)} = (3, 1)^T, \quad r^{(0)} = (-3, 4)^T = p^{(0)}.$$

$$\begin{aligned} \underline{\mathbf{S1}} \quad \alpha_0 &= -\frac{5}{11}, \\ x^{(1)} &= \frac{3}{11}(16, -3)^T = (4.363636\dots, -0.818181\dots)^T, \\ r^{(1)} &= -\frac{12}{11}(4, 3)^T, \\ \beta_0 &= \frac{144}{121}, \quad p^{(1)} = \frac{60}{121}(-16, 3)^T. \end{aligned}$$

$$\begin{aligned} \underline{\mathbf{S2}} \quad \alpha_1 &= \frac{11}{20}, \\ x^{(2)} &= (0,0)^T, \quad r^{(2)} = (0,0)^T \rightarrow \text{Stopp mit } x^{(2)} = x^*. \end{aligned}$$

Das CG konvergiert in $n = 2$ Schritten mit Anstiegs- und Abstiegsrichtung gegen x^* .

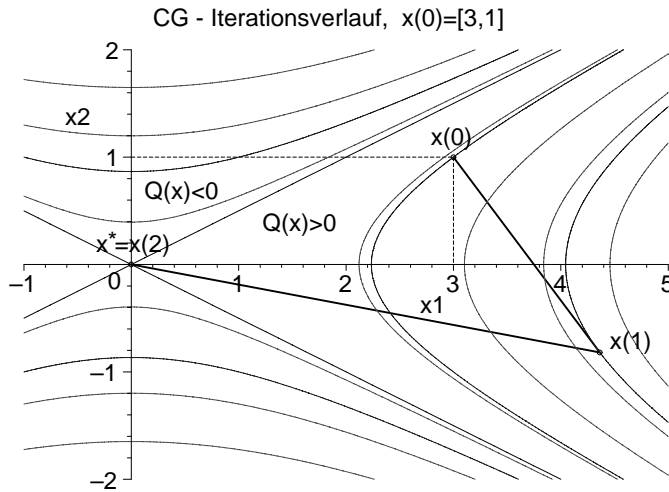


Abb. 5.14 Datei *cg_03.ps*
Höhenlinienbild
bei endlichem
Iterationsverlauf des CG mit
 $x^{(0)} = (3, 1)^T$
zu $Q(x) = \frac{1}{2}x_1^2 - 2x_2^2$, $Q(x^*) = 0$,
mit contours=[-1.5, 2.5,
8.1818] und contours=7

$$(5) \quad x^{(0)} = (2, 1)^T, \quad r^{(0)} = 2(-1, 2)^T = p^{(0)}.$$

$$\begin{aligned} \underline{\mathbf{S1}} \quad \alpha_0 &= -\frac{1}{3}, \\ x^{(1)} &= \frac{1}{3}(8, -1)^T, \quad r^{(1)} = -\frac{4}{3}(2, 1)^T, \\ \beta_0 &= \frac{4}{9}, \quad p^{(1)} = \frac{4}{9}(-8, 1)^T. \end{aligned}$$

$$\begin{aligned} \underline{\mathbf{S2}} \quad \alpha_1 &= \frac{3}{4}, \\ x^{(2)} &= (0,0)^T, \quad r^{(2)} = (0,0)^T \rightarrow \text{Stopp mit } x^{(2)} = x^*. \end{aligned}$$

Das CG konvergiert in $n = 2$ Schritten mit Anstiegs- und Abstiegsrichtung gegen x^* .

Notwendig zur Durchführung des CG ist die Bedingung $p^{(m)T} A p^{(m)} \neq 0$.

Für $p^{(0)} = 2(-1, 2)^T$, $p^{(1)} = \frac{4}{9}(-8, 1)^T$ erhält man $p^{(0)T} A p^{(0)} = -60$, $p^{(1)T} A p^{(1)} = \frac{320}{27}$ sowie die A-Orthogonalität gemäß

$$p^{(0)T} A p^{(1)} = (p^{(0)T} A p^{(1)})^T = p^{(1)T} A p^{(0)} = 2(-1, 2)^T \begin{pmatrix} 1 & 0 \\ 0 & -4 \end{pmatrix} \begin{pmatrix} -8 \\ 1 \end{pmatrix} \frac{4}{9} = 0.$$

Das GV bricht nach dem ersten Schritt wegen $r^{(1)T} A r^{(1)} = 0$ ab.

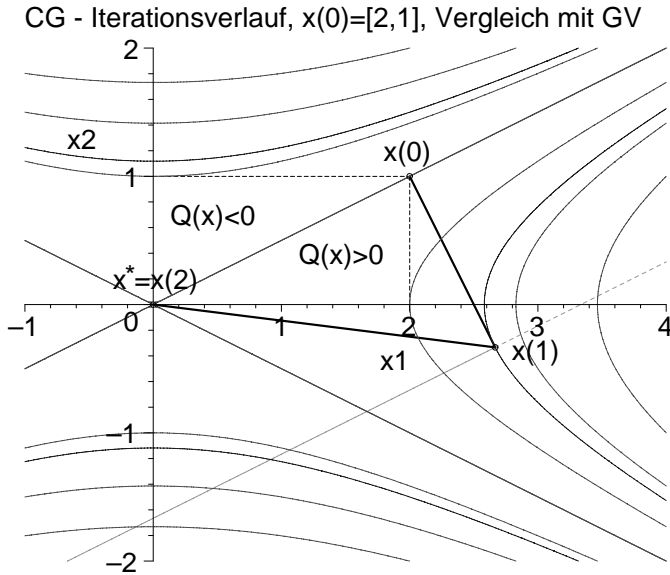


Abb. 5.15 Datei *cg_02.ps*
 Höhenlinienbild
 bei endlichem
 Iterationsverlauf des CG mit
 $x^{(0)} = (2, 1)^T$
 zu $Q(x) = \frac{1}{2}x_1^2 - 2x_2^2$, $Q(x^*) = 0$,
 mit `contours=[-2.5,3.333]`
 und `contours=7`,
 Abbruch von GV bei $x^{(1)}$

(6) $x^{(0)} = (-8, 1)^T$, $r^{(0)} = 4(2, 1)^T = p^{(0)}$.

Beim GV mit dem Startvektor $x^{(0)} = (2, 1)^T$ war von $x^{(1)} = \frac{1}{3}(8, -1)^T$ aus in Richtung des Vektors $r^{(1)} = -\frac{4}{3}(2, 1)^T$ das Problem der Nulldivision aufgetreten, was hier nun schon bei der Anfangsrechnung passiert. Das CG bricht wegen $p^{(0)T} A p^{(0)} = 0$ im Nenner von α_0 ab. In Richtung $p^{(0)}$ ist die Funktion $f(\alpha) = Q(x^{(0)} + \alpha p^{(0)})$ streng monoton fallend und $f'(\alpha) \neq 0$.

$$x^{(0)} = (-8, 1.1)^T \text{ oder } x^{(0)} = (-8, 0.9)^T.$$

Für beide Startvektoren ist die Rechnung durchführbar und CG konvergiert.

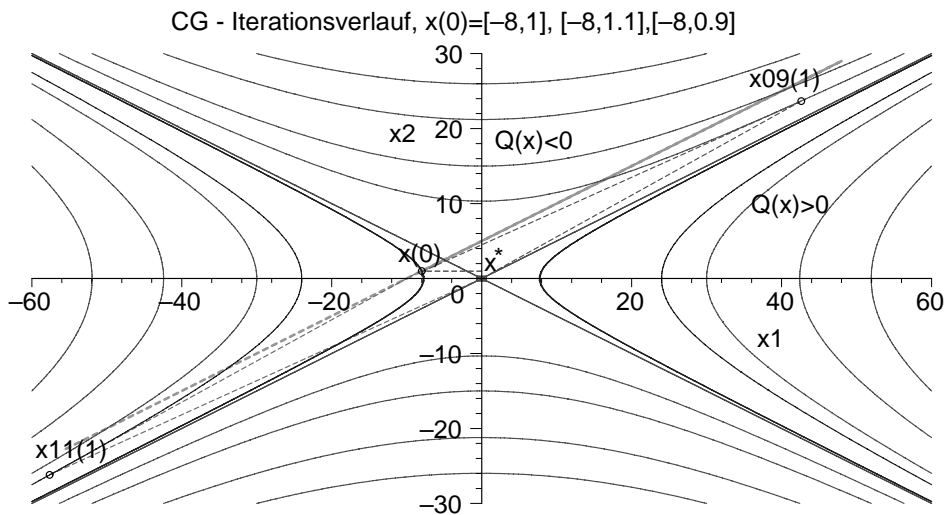


Abb. 5.16 Datei *cg_05.ps*
 Höhenlinienbild und Iterationsverlauf des CG mit
 $x^{(0)} = (-8, 1)^T$ (Abbruch), $x^{(0)} = (-8, 1.1)^T$, $(-8, 0.9)^T$ (Konvergenz)
 mit `contours=[30, 29.58,288.09, 30.38,-213.16, 0]` und `contours=7`

CG fuer $Ax=b$, $A=A'$ indefinit - 2 konvergente Iterationsverlaeufe

m	$[x(m) [1], x(m) [2]]$	$Q(x(m))$	$[x(m) [1], x(m) [2]]$	$Q(x(m))$
0	[-8.000, 1.100]	29.58	[-8.000, 0.900]	30.38
1	$x_{11}(1)=[-57.619, -26.190]$	288.095	$x_{09}(1)=[42.632, 23.684]$	-213.157
2	[0, 0]	0	[0, 0]	0

Beispiel 5.8 Sei $Ax = b$, $A = A^T$ und indefinit aus den Beispielen 1.5, 5.3.

$$\begin{pmatrix} 2 & 1 \\ 1 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \quad x^* = \begin{pmatrix} 10/7 \\ 1/7 \end{pmatrix}.$$

Wir testen das CG mit dem Startvektor $x^{(0)} = (\frac{3}{2}, 1)^T$.

Bei Durchführbarkeit ($p^{(m)T} A p^{(m)} \neq 0$) und exakter Arithmetik werden maximal zwei Schritte gemacht, wobei $p^{(0)} = r^{(0)} \perp Q(x)=\text{const}$ eine Anstiegs- und $p^{(1)} = r^{(1)} + \beta_0 p^{(0)} \not\perp Q(x)=\text{const}$ eine Abstiegsrichtung bezüglich $Q(x)$ sind.

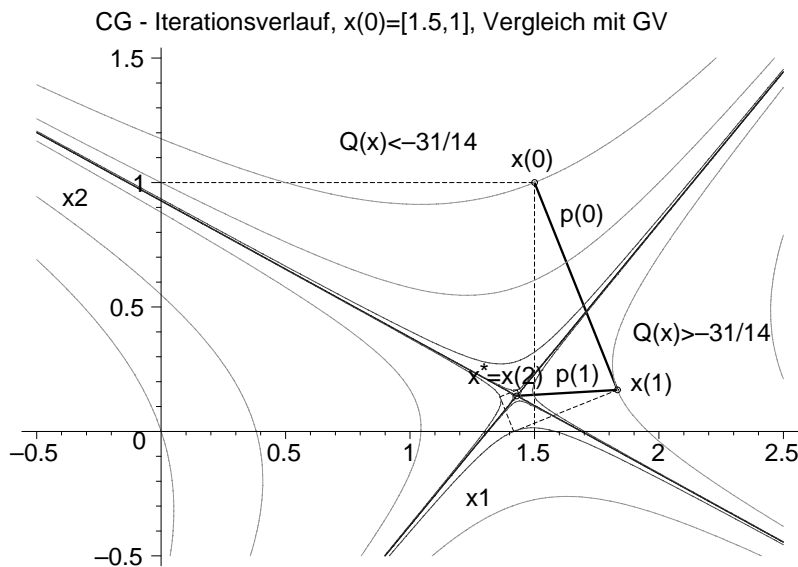


Abb. 5.17 Datei *cg_002.ps*

Höhenlinienbild bei endlichem Iterationsverlauf des CG mit $x^{(0)} = (1.5, 1)^T$

zu $Q(x) = x_1^2 + x_1 x_2 - \frac{3}{2} x_2^2 - 3x_1 - x_2$, $Q(x^*) = -\frac{31}{14} = -2.214285\dots$,

mit contours=[-2.5, 0, -1.0, -3.25, -2.0417, -2.2431, -2.2095, -2.2151]

und contours=[-31/14], Vergleich mit GV (dünn gestrichelter Linienzug)

$x^{(0)} = (\frac{3}{2}, 1)^T$ ist ein günstiger Startvektor. Andere Startvektoren werden wir nicht testen. Aber Situationen, wie sie in Beispiel 5.7 aufgetreten sind, können auch hier vorkommen.

Das CG für $Bx = A^T Ax = A^T b = c$, verbunden mit dem Funktional $R(x)$ und seiner Richtung des steilsten Abstiegs $\hat{r}(x) = A^T (b - Ax)$ liefert keine neuen Erkenntnisse.

5.2.2 CG mit nicht symmetrischer Matrix

Auch hier kann man die Funktionale $Q(x) = T(x), R(x), S(x)$ betrachten, die wir im Kapitel 5.1.2 für das GV mit nicht symmetrischer Matrix verwendet haben.

Als erste Besonderheit ist wie zum CG für $A = A^T$ und indefinit (Kap. 5.2.1) bei der Berechnung der Schrittzahlen α und β gemäß (5.11) bzw. (5.12) die Durchführbarkeit zu beachten. Das bedeutet für die Größe im Nenner der Test

$$(Ap)^T p = 0 \quad \text{bzw.} \quad p^T Ap = 0.$$

Weiterhin gibt es für nicht symmetrische Matrizen Probleme bei der Konstruktion der A-orthogonalen Suchvektoren p , die im Normalfall einen Vektorraum mit stetig wachsender Dimension gebildet haben und die Endlichkeit des CG bewirkten sowie letztendlich eine Vereinfachung der hergeleiteten Schrittzahlen in der Form

$$\alpha = \frac{\|r\|_2^2}{p^T w}, \quad w = Ap, \quad \beta = \frac{\|r_{neu}\|_2^2}{\|r\|_2^2} \quad (5.13)$$

erlaubten.

Was passiert, wenn man bei $A \neq A^T$ im CG einfach mit diesen Werten rechnet? Oder ist es zumindest sinnvoll, dann mit den originalen Formeln

$$\alpha = \frac{p^T r}{p^T w}, \quad \beta = -\frac{p^T A r_{neu}}{p^T w}, \quad w = Ap, \quad (5.14)$$

zu arbeiten, wie sie bei der Herleitung des Verfahrens in Kapitel 3.6.2 zunächst dargestellt worden sind.

Was bleibt dann von der A-Orthogonalität der Suchrichtungen übrig? Nach den Betrachtungen zur Optimalität von AV im Kapitel 3.4 sollte im Fall von $A \neq A^T$ und (5.14) zumindest die Bedingung (3.59) in den abgeschwächten Formen, also

$$p^{(j)T} A p^{(m)} = 0, \quad j = 0, 1, \dots, m-1, \quad \text{bzw.} \quad p^{(m-1)T} A p^{(m)} = 0 \quad (5.15)$$

erfüllt sein. Erhält man dann die Endlichkeit des Verfahrens?

Diese Fragestellungen und Vergleiche zum GV in den Varianten $A > 0$ bzw. A indefinit sollen an den schon mehrfach genannten Beispielen erläutert und illustriert werden.

Dabei wird der Iterationsverlauf verfolgt. Neben der Suchrichtung $p(x)$ – erste Suchrichtung $p(x) = r(x)$ ist eine Richtung des steilsten Abstiegs/Anstiegs zu $Q(x)$, $p(x) = r(x) + \beta p_{alt}(x)$ sind weitere Suchrichtungen – werden auch andere Richtungsvektoren angezeigt. So sind die Vektoren

- $r(x) = b - Ax \perp$ zu $Q(x) = \text{const}$ und \perp zu $R(x) = \text{const}$,
- $\tilde{r}(x) = b - \frac{1}{2} \tilde{A}x \perp$ zu Höhenlinien von $Q(x)$, Richtung des steilsten Ab-/Anstiegs,
- $\hat{r}(x) = A^T(b - Ax)$ als Richtung des steilsten Abstiegs zu $R(x)$

von Interesse.

Beispiel 5.9 Sei $Ax = b$, $A \neq A^T$, $A > 0$ aus den Beispielen 1.6, 5.4.

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

Wir erinnern uns an das Höhenlinienbild mit den verschiedenen Minimumstellen.

CG - Höhenlinien von $R(x), Q(x), S(x)$ in $[-1,2] \times [-1,2]$

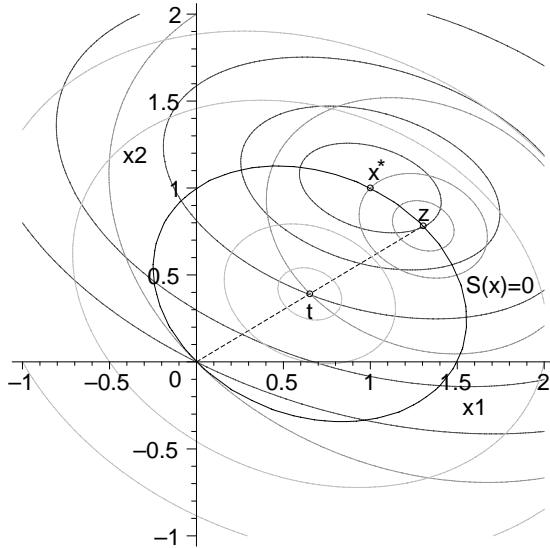


Abb. 5.18 Datei *cg_0030.ps*

Höhenlinienbilder (Ellipsenscharen) zu den Funktionalen

$Q(x)$, Minimum bei $z = (\frac{1}{2}\tilde{A})^{-1}b \neq x^*$,
 $R(x)$, Minimum bei $x^* = A^{-1}b$, und
 $S(x)$, Minimum bei $t = \tilde{A}^{-1}b = \frac{z}{2}$

Das CG für $A^T A x = A^T b$, also zu $R(x)$, mit der Abstiegsrichtung $p(x)$ (1. Schritt wie im GV) konvergiert (theoretisch) maximal nach zwei Schritten. Dabei muss man aber die allgemein schlechtere Kondition der spd Matrix $A^T A$ in Kauf nehmen.

CG fuer $A^T A x = A^T b$, $A \ll A^T$, $A > 0$ - konv. endlicher Iterationsverlauf

m [x(m) [1], x(m) [2]] [rd(m) [1], rd(m) [2]]

0 [1.0000000000, -1.0000000000] [4.000e+00, 2.000e+01]

1 [1.3795620440, 0.8978102190] [-1.314e+00, 2.628e-01]

2 [1, 1] [0, 0]

Das übliche CG für $Ax = b$, also zum Funktional $Q(x)$, mit der Suchrichtung $p(x)$ und den Schrittzahlen α und β gemäß (5.13) zeigt zwar einen anfänglichen starken Trend zur Lösung x^* hin, tut aber letztendlich nicht konvergieren.

Dabei ist sowohl $p(x) \not\perp Q(x) = \text{const}$ als auch $p(x) \not\perp R(x) = \text{const}$.

Ersteres ist sinnvoll, denn sonst würden wir mit $p(x)$ zur Minimumstelle z kommen. Die andere Bedingung bedeutet, dass $p(x)$ eine "schräge" Abstiegsrichtung bez. $R(x)$ darstellt, aber sich nicht nahe genug seiner Richtung des steilsten Abstiegs $\hat{r}(x) = A^T(b - Ax)$ bewegt. Der Vektor $\hat{r}(x)$ ist in der folgenden Abbildung nicht eingezeichnet.

Die Abstiegsrichtung $\tilde{r}(x) = b - \frac{1}{2}\tilde{A}x = \text{rs}(x)$, senkrecht zu den Höhenlinien von $Q(x) = T(x)$, ist die Richtung des steilsten Abstiegs zum LGS $\frac{1}{2}\tilde{A}x = b$ und liefert eine gegen die Stelle z konvergierende Folge.

CG fuer $Ax=b$ und $1/2(A+A')x=b$ - Iterationsverlauf, $x(0)=[1,-1]$

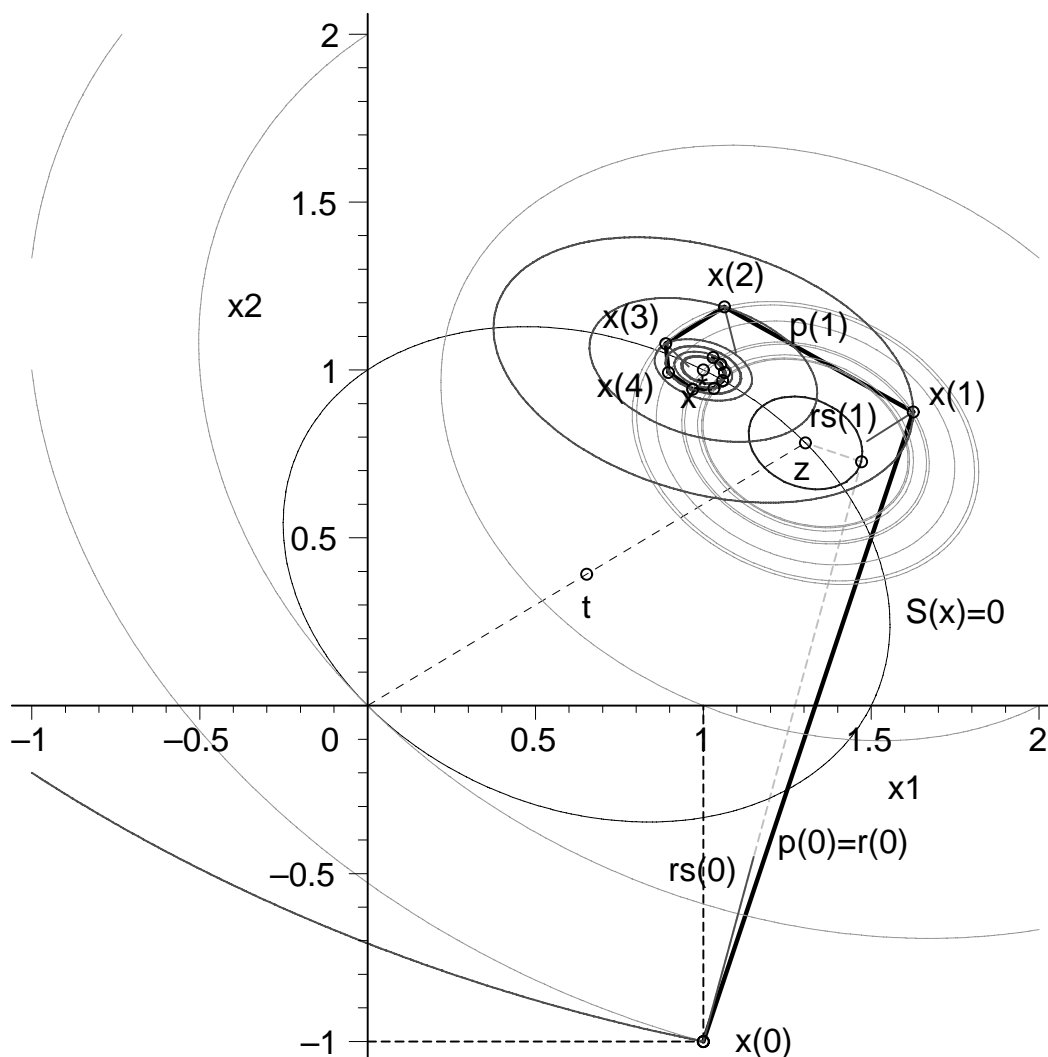


Abb. 5.19 Datei cg_003.ps

Höhenlinienbild und CG zu $Q(x)$ mit den Iterierten $x^{(m)}$ (Iterationsverlauf fett gezeichnet), Suchrichtung $p(x)$, $Q(x^*) = -3$, $Q(z) = -3.130434$,
Höhenlinien zu $Q(x)$ mit contours=[2,0,-2,-3,-2.875,-2.8874,-2.9408,-3.0059,-3.0383,-3.0409,-3.0311,-3.10265,-3.1304].

Zusätzlich eingezeichnet:

- Vektor $\tilde{r}(x) = rs(x) \perp Q(x) = \text{const}$ in den ersten Iterierten,
- Iterationsverlauf für $\frac{1}{2}Ax = b$ (Funktional $Q(x) = T(x)$) mit Abstiegsrichtung $\tilde{r}(x)$, Konvergenz gegen z in 2 Schritten (dünn gestrichelt),
- Höhenlinien von $R(x)$ an den Iterierten, contours=[11,-8.2969,-8.793,-8.962,-8.9768,-8.9777,-8.9864,-8.9919,-8.9929],
- Funktional $S(x)$ mit Minimumstelle $t = \frac{z}{2}$ und Höhenlinie $S(x) = 0$.

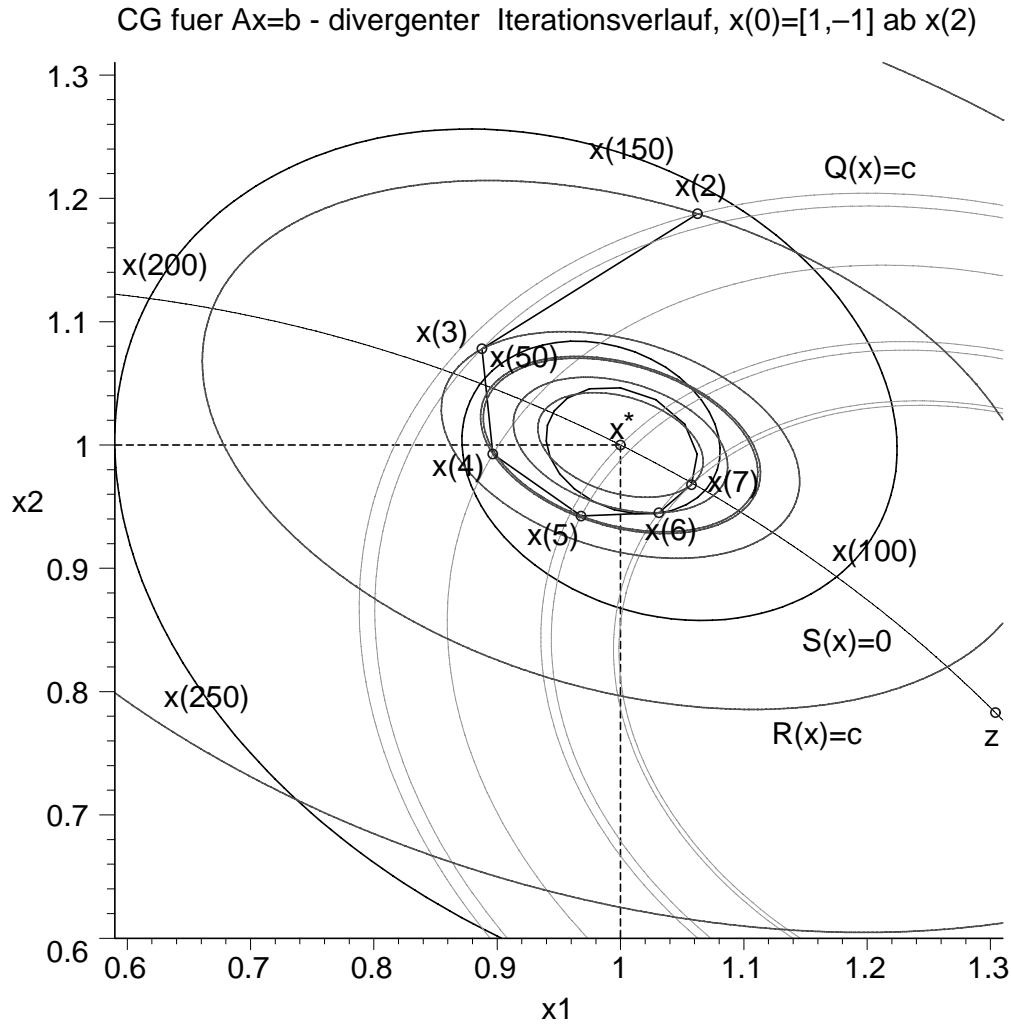


Abb. 5.20 Datei *cg_0031.ps*

Höhenlinienbild und CG zu $Q(x)$ mit den Iterierten $x^{(m)}$, $m = 2, 3, \dots$,
bei lokaler Betrachtung, Suchrichtung $p(x)$,

$Q(x^*) = -3$, $Q(z) = -3.130434$, Höhenlinien zu $Q(x)$ mit

`contours=[-2.875,-2.8874,-2.9408,-3,-3.0059,-3.0383,-3.0409]`

Zusätzlich eingezeichnet:

– Höhenlinien von $R(x)$ an den Iterierten,

`contours=[-8.2969,-8.793,-8.962,-8.9768,-8.9777,-8.9864,-8.9919]`

CG fuer $Ax=b$, $A \succ A'$, $A \succ 0$ - divergenter Iterationsverlauf

m	$[x^{(m)}[1], x^{(m)}[2]]$	$[r^{(m)}[1], r^{(m)}[2]]$
0	$[1.00000000, -1.00000000]$	$[2.000e+00, 6.000e+00]$
1	$[1.62500000, 0.87500000]$	$[-1.125e+00, 3.750e-01]$

2	[1.062500000, 1.187500000]	[-3.125e-01, -5.625e-01]
3	[0.887667521, 1.078069849]	[1.466e-01, -2.342e-01]
4	[0.896524084, 0.992596330]	[2.144e-01, 2.221e-02]
5	[0.968121823, 0.942340808]	[1.214e-01, 1.730e-01]
6	[1.031071008, 0.945115413]	[-7.257e-03, 1.647e-01]
7	[1.057521370, 0.967750713]	[-8.279e-02, 9.675e-02]

8	[1.062276737, 0.992504011]	[-1.171e-01, 2.249e-02]
9	[1.052009082, 1.016941051]	[-1.210e-01, -5.082e-02]
10	[1.028721803, 1.036815328]	[-9.426e-02, -1.104e-01]
11	[0.999899098, 1.046381094]	[-4.618e-02, -1.391e-01]
12	[0.974778094, 1.045463875]	[4.980e-03, -1.364e-01]
13	[0.957104379, 1.038084017]	[4.771e-02, -1.143e-01]
14	[0.946277135, 1.027687474]	[7.976e-02, -8.306e-02]
15	[0.940791943, 1.015891492]	[1.025e-01, -4.767e-02]
16	[0.939749721, 1.003214075]	[1.173e-01, -9.642e-03]
17	[0.942985227, 0.989926815]	[1.241e-01, 3.022e-02]
18	[0.950642324, 0.976604813]	[1.221e-01, 7.019e-02]
19	[0.962567322, 0.964295108]	[1.106e-01, 1.071e-01]
20	[0.977871179, 0.954258070]	[9.000e-02, 1.372e-01]
30	[1.080384141, 0.988754254]	[-1.495e-01, 3.374e-02]
40	[1.011543145, 1.075459799]	[-9.855e-02, -2.264e-01]
50	[0.902263592, 1.065342300]	[1.301e-01, -1.960e-01]
60	[0.871469759, 0.999551097]	[2.575e-01, 1.347e-03]
70	[0.913108920, 0.917707117]	[2.561e-01, 2.469e-01]
80	[1.018170478, 0.862819929]	[1.008e-01, 4.115e-01]
90	[1.123113653, 0.865978369]	[-1.122e-01, 4.021e-01]
100	[1.190017270, 0.906869275]	[-2.869e-01, 2.794e-01]
150	[1.007722566, 1.234321962]	
200	[0.623995852, 1.129342167]	
250	[0.664786953, 0.793829157]	
300	[1.106340551, 0.534751570]	

Wir kontrollieren noch einige Eigenschaften der auftretenden Vektoren.

Bei exakter Rechnung haben wir aus den ersten drei Schritten die Iterationsvektoren $x^{(i)}, r^{(i)}, p^{(i)}$, $i = 0, 1, 2, 3$, als Spalten der folgenden Felder.

$$\begin{aligned}
 & m = 3 \\
 xv &= \begin{bmatrix} 1 & \frac{13}{8} & \frac{17}{16} & \frac{85841}{96704} \\ -1 & \frac{7}{8} & \frac{19}{16} & \frac{312761}{290112} \end{bmatrix} \\
 rv &= \begin{bmatrix} 2 & \frac{-9}{8} & \frac{-5}{16} & \frac{42529}{290112} \\ 6 & \frac{3}{8} & \frac{-9}{16} & \frac{-22649}{96704} \end{bmatrix} \\
 pv &= \begin{bmatrix} 2 & \frac{-135}{128} & \frac{-319}{512} & \frac{5339196521}{168329945088} \\ 6 & \frac{75}{128} & \frac{-599}{1536} & \frac{-154583637359}{504989835264} \end{bmatrix}
 \end{aligned}$$

Wegen $A \neq A^T$ ist die A-Orthogonalität der Suchrichtungen $p^{(i)}$ nicht erfüllt.

$$\begin{aligned} p^{(0)T} Ap^{(1)} &= 7.5 & p^{(1)T} Ap^{(0)} &= p^{(0)T} A^T p^{(1)} = 0 \\ p^{(0)T} Ap^{(2)} &= -10.291\dots & p^{(2)T} Ap^{(0)} &= -13.25 \\ p^{(0)T} Ap^{(3)} &= -5.995\dots & p^{(3)T} Ap^{(0)} &= -5.192\dots \\ p^{(1)T} Ap^{(2)} &= 1.040\dots & p^{(2)T} Ap^{(1)} &= 0.263\dots \\ p^{(1)T} Ap^{(3)} &= -0.282\dots & p^{(3)T} Ap^{(1)} &= -0.586\dots \\ p^{(2)T} Ap^{(3)} &= 0.509\dots & p^{(3)T} Ap^{(2)} &= 0.306\dots \end{aligned}$$

Genauso geht die Orthogonalität der Vektoren $r^{(i)}$ verloren.

$$\begin{aligned} r^{(0)T} r^{(1)} &= r^{(1)T} r^{(0)} = 0 \\ r^{(0)T} r^{(2)} &= r^{(2)T} r^{(0)} = -4 \\ r^{(0)T} r^{(3)} &= r^{(3)T} r^{(0)} = -1.112\dots \\ r^{(1)T} r^{(2)} &= r^{(2)T} r^{(1)} = 0.140\dots \\ r^{(1)T} r^{(3)} &= r^{(3)T} r^{(1)} = -0.252\dots \\ r^{(2)T} r^{(3)} &= r^{(3)T} r^{(2)} = 0.085\dots \end{aligned}$$

Wir testen noch zwei Versionen des CG mit kleinen Veränderungen im Algorithmus. Wir machen zunächst eine Rechnung mit dem modifizierten CG für $Ax = b$, also zum Funktional $Q(x)$, mit der Suchrichtung $p(x)$ und der Schrittzahl $\bar{\alpha}$ (5.10), d. h.

$$\bar{\alpha}_k = \frac{(b - \frac{1}{2} \tilde{A} x^{(k)})^T p^{(k)}}{p^{(k)T} A p^{(k)}}.$$

Dieses AV konvergiert weder gegen $x^* = (1, 1)^T$ noch gegen $z = (1.304\ 347\dots, 0.782\ 608\dots)^T$.

Modifiziertes CG fuer $Ax=b$, $A \ll A'$, $A > 0$ - Iterationsverlauf

m	[x(m) [1], x(m) [2]]	[r(m) [1], r(m) [2]]
0	[1.000000000, -1.000000000]	[2.000e+00, 6.000e+00]
1	[1.562500000, 0.687500000]	[-8.125e-01, 9.375e-01]
2	[1.472667192, 0.830192474]	[-7.755e-01, 5.094e-01]
3	[1.437257179, 0.864886870]	[-7.394e-01, 4.053e-01]
4	[1.418492076, 0.879790509]	[-7.168e-01, 3.606e-01]
5	[1.406976864, 0.887879530]	[-7.018e-01, 3.364e-01]
6	[1.399220614, 0.892900365]	[-6.913e-01, 3.213e-01]
7	[1.393647759, 0.896301993]	[-6.836e-01, 3.111e-01]
8	[1.389450293, 0.898752698]	[-6.777e-01, 3.037e-01]
9	[1.386173514, 0.900600225]	[-6.729e-01, 2.982e-01]
10	[1.383542772, 0.902042258]	[-6.691e-01, 2.939e-01]
100	[1.361040094, 0.912911116]	[-6.350e-01, 2.613e-01]
200	[1.359484500, 0.913567388]	[-6.325e-01, 2.593e-01]

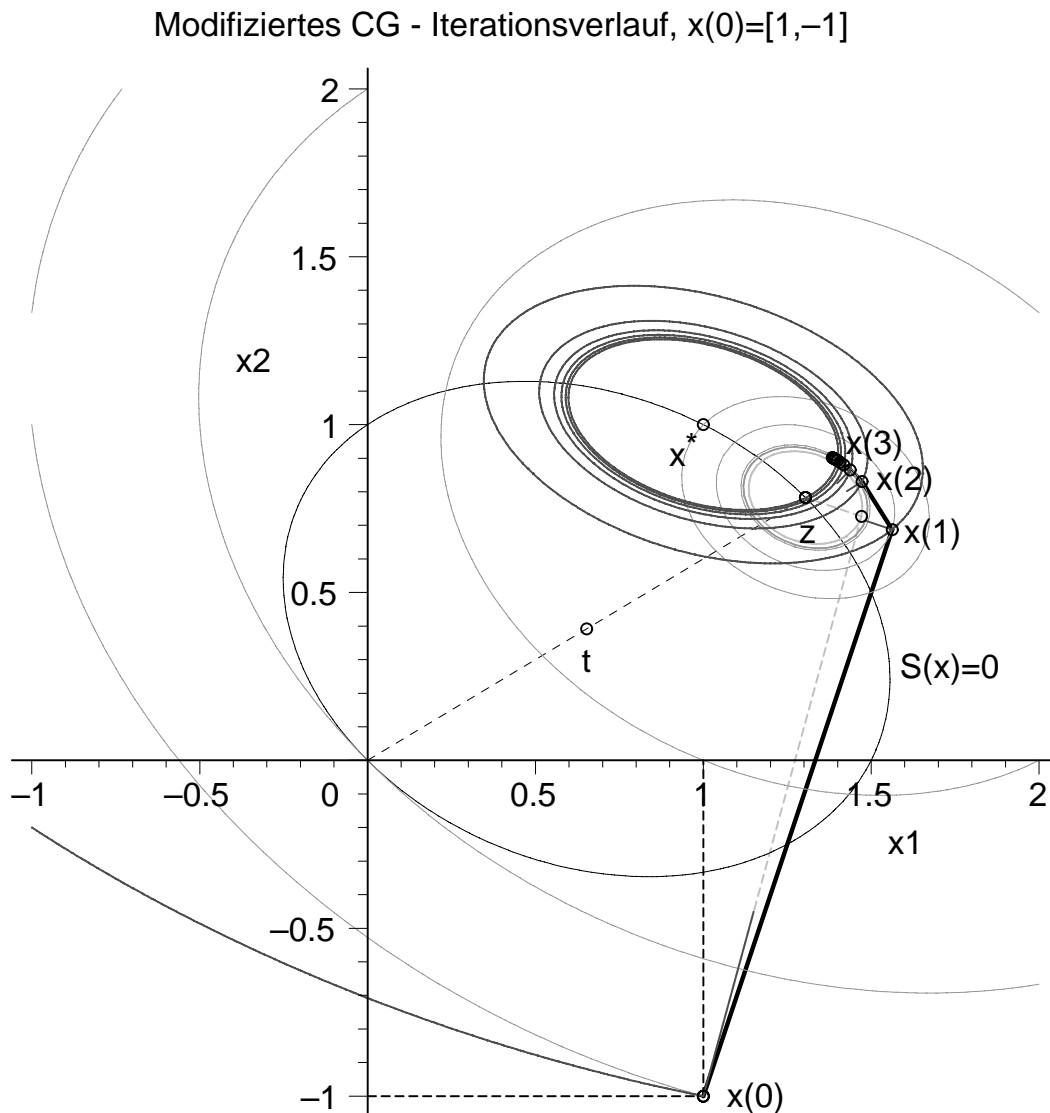


Abb. 5.21 Datei *lgs11.02.ps*

Höhenlinienbild und modifiziertes CG zu $Q(x)$ mit den Iterierten $x^{(m)}$

(Iterationsverlauf fett gezeichnet), Suchrichtung $p(x)$, Schrittzahl $\bar{\alpha}$,

$Q(x^*) = -3$, $Q(z) = -3.130434$, Höhenlinien zu $Q(x)$ mit

`contours=[2,0,-2,-3,-3.0625,-3.0947,-3.0971,-3.0977,-3.0979,-3.10265]`.

Zusätzlich eingezeichnet:

- Vektor $\tilde{r}(x) \perp Q(x) = \text{const}$ in den ersten Iterierten,
- Iterationsverlauf für $\frac{1}{2}\tilde{A}x = b$ (Funktional $Q(x) = T(x)$) mit Abstiegsrichtung $\tilde{r}(x)$, Konvergenz gegen z (dünn gestrichelt),
- Höhenlinien von $R(x)$ an den Iterierten,
`contours=[11,-8.2305,-8.5695,-8.6445,-8.6781,-8.6971,-8.7094]`
- Funktional $S(x)$ mit Minimumstelle $t = \frac{z}{2}$ und Höhenlinie $S(x) = 0$.

Als zweiter Versuch bleibt die Verwendung der Schrittzahlen α und β gemäß (5.14) anstelle von (5.13).

Kann man dann im so “veränderten“ CG die Konvergenz in einem unendlichen Prozess oder vielleicht sogar nach endlich vielen Schritten erreichen? Wie steht es um die A-Orthogonalität der Suchrichtungen bzw. Orthogonalität der Residua?

Unser Beispiel zeigt sowohl mit dem Startvektor $(1, -1)^T$ als auch mit anderen ein gutartiges Verhalten und das “veränderte“ CG konvergiert nach zwei Schritten.

```

Startvektor          x=[ 1.0000000000000000e+00 -1.0000000000000000e+00]
Residuum             r = b-Ax=[ 2.0000000000000000e+00  6.0000000000000000e+00]
Funktionswert       Q(x)=  2.0000000000000000e+00
Anfangsfehlerquadrat r'r=  4.0000000000000000e+01

```

Schritt k = 1

```

Suchrichtung        p=[ 2.0000000000000000e+00  6.0000000000000000e+00]
Suchschritt         alpha=  3.1250000000000000e-01
Iterationsvektor    x=[ 1.6250000000000000e+00  8.7500000000000000e-01]

```

```

Residuum            r = b-Ax=[-1.1250000000000000e+00  3.7500000000000000e-01]
Funktionswert       Q(x)= -3.0000000000000000e+00
Fehlernormquadrat   r'r=  1.4062500000000000e+00
Schritt             beta= -2.3437500000000000e-02
neue Suchrichtung   p=[-1.1718750000000000e+00  2.3437500000000000e-01]

```

Schritt k = 2

```

Suchrichtung        p=[-1.1718750000000000e+00  2.3437500000000000e-01]
Suchschritt         alpha=  5.3333333330000000e-01
Iterationsvektor    x=[ 1.0000000000000000e+00  1.0000000000000000e+00]

```

```

Residuum            r = b-Ax=[ 0.0000000000000000e-01  0.0000000000000000e-01]
Funktionswert       Q(x)= -3.0000000000000000e+00
Fehlernormquadrat   r'r=  0.0000000000000000e-01

```

Es gilt $x^{(2)} = x^*$ und $r^{(2)} = 0$, $p^{(2)} = 0$. Die Residua $r^{(k)}$, $k = 0, 1, 2$, sind orthogonal. Die Suchrichtungen $p^{(k)}$ erfüllen jedoch nicht die allgemeine A-Orthogonalität $p^{(j)T} A p^{(k)} = 0$ für $j \neq k$. Es ist nämlich $p^{(1)T} A p^{(0)} = -7.5$, während alle anderen Funktionalwerte verschwinden.

Unser LGS ist zu klein, um weiter gehende Aussagen zu treffen.

Aber schon mit einer Erhöhung seiner Dimension um Eins können wir ausreichend Informationen darüber sammeln, wie der Iterationsverlauf ist, ob Konvergenz des CG erreichbar ist und was von der A-Orthogonalität der Suchrichtungen übrig bleibt.

Beispiel 5.10 Sei $Ax = b$, $A \neq A^T$, $A > 0$, aus Beispiel 5.5.

$$\begin{pmatrix} 2 & 1 & 0 \\ 0 & 3 & 1 \\ 1 & 2 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \\ 7 \end{pmatrix}, \quad x^* = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad x^{(0)} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}.$$

Das ‘‘veränderte‘‘ CG konvergiert und sogar schneller als das GV (Beispiel 5.5), aber die exakte Rechnung ist nicht nach $n = 3$ Schritten an der Lösung.

Die Iterationsvektoren $x^{(i)}, r^{(i)}, p^{(i)}$, $i = 0, 1, 2, 3$, aus den ersten drei Schritten sind die Spalten der folgenden Felder.

$$xv = \begin{bmatrix} 1 & \frac{281}{203} & \frac{51127}{38842} & \frac{307752202}{307530877} \\ -1 & \frac{10}{29} & \frac{19876}{19421} & \frac{303984152}{307530877} \\ 0 & \frac{312}{203} & \frac{17076}{19421} & \frac{1225021983}{1230123508} \end{bmatrix}$$

$$pv = \begin{bmatrix} 2 & \frac{-5577}{41209} & \frac{-7041884395}{10938081989} & \frac{896385910462806051577900}{54810329687894111410654869} \\ 7 & \frac{7956}{5887} & \frac{-780172120}{10938081989} & \frac{1235253189047975624176975}{31320188393082349377517068} \\ 8 & \frac{-53976}{41209} & \frac{5203880535}{21876163978} & \frac{1338156255041595013661325}{36540219791929407607103246} \end{bmatrix}$$

$$rv = \begin{bmatrix} 2 & \frac{-23}{203} & \frac{-12740}{19421} & \frac{3104075}{307530877} \\ 7 & \frac{10}{7} & \frac{980}{19421} & \frac{1643525}{42418052} \\ 8 & \frac{-248}{203} & \frac{4655}{38842} & \frac{11973650}{307530877} \end{bmatrix}$$

Rechnung mit Maple bei `Digits:=16`

Veraendertes CG fuer Ax=b, A<>A', A>0 - konvergenter Iterationsverlauf

m	[x(m)[1],	x(m)[2],	x(m)[3]] [r(m)[1],	r(m)[2],	r(m)[3]
0	[1.000000000000,	-1.000000000000,	0.000000000000]	[2.000e+00,	7.000e+00,	8.000e+00]
1	[1.384236453202,	0.344827586207,	1.536945812808]	[-1.133e-01,	1.429e+00,	-1.222e+00]
2	[1.316281344936,	1.023428247773,	0.879254415324]	[-6.560e-01,	5.046e-02,	1.198e-01]
3	[1.000719683832,	0.988467093013,	0.995852835128]	[1.009e-02,	3.875e-02,	3.893e-02]
4	[1.003879593648,	0.996087421607,	1.002928671053]	[-3.847e-03,	8.809e-03,	-7.769e-03]
5	[1.001861606144,	1.000417213000,	0.998839262541]	[-4.140e-03,	-9.090e-05,	1.947e-03]
6	[1.000011982757,	1.000109017964,	1.000017216368]	[-1.330e-04,	-3.443e-04,	-2.989e-04]
7	[0.999954153391,	1.000034723333,	0.999977048900]	[5.697e-05,	-8.122e-05,	6.820e-05]
8	[0.999985509859,	0.999997684626,	1.000013167543]	[3.130e-05,	-6.221e-06,	-3.355e-05]
9	[0.999997773168,	0.999994520624,	1.000001664840]	[9.933e-06,	1.477e-05,	6.526e-06]
10	[1.000002627765,	0.99998062746,	1.000001493089]	[-3.318e-06,	4.319e-06,	-4.726e-06]
11	[1.000000730469,	0.999999647954,	0.999999426392]	[-1.109e-06,	1.630e-06,	2.268e-06]
12	[1.000000192615,	1.000000286669,	0.999999776326]	[-6.719e-07,	-6.363e-07,	1.287e-07]
13	[0.999999856474,	1.000000082100,	0.999999865200]	[2.050e-07,	-1.115e-07,	5.185e-07]
14	[0.99999925841,	1.000000057680,	1.00000008411]	[9.064e-08,	-1.814e-07,	-7.484e-08]
15	[0.99999980510,	0.99999980623,	1.00000020002]	[5.836e-08,	3.813e-08,	-2.176e-08]
16	[1.000000003749,	1.000000000982,	1.000000009867]	[-8.480e-09,	-1.281e-08,	-4.518e-08]
17	[1.000000002415,	0.999999998680,	1.000000000161]	[-3.511e-09,	3.799e-09,	-4.185e-10]
18	[1.000000000439,	1.000000000630,	0.999999999444]	[-1.508e-09,	-1.335e-09,	5.238e-10]
19	[1.000000000024,	0.999999999938,	0.999999999712]	[1.453e-11,	4.755e-10,	1.254e-09]
20	[1.000000000002,	0.999999999999,	1.000000000000]	[-3.863e-12,	2.032e-12,	-7.260e-13]

Die Suchrichtungen $p^{(k)}$ erfüllen die spezielle A-Orthogonalität $p^{(k-1)T} A p^{(k)} = 0$, und es ist $p^{(k)T} A p^{(k-1)} = p^{(k-1)T} A^T p^{(k)} \neq 0$.

Die Residua genügen der Beziehung $r^{(k-1)T} r^{(k)} = 0$.

Damit sind die Bedingungen i. Allg. zu schwach für die Endlichkeit des CG.

Beispiel 5.11

Sei $A \neq A^T$ und A indefinit aus den Beispielen 1.7, 5.6.

$$\begin{pmatrix} 0.780 & 0.563 \\ 0.913 & 0.659 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0.217 \\ 0.254 \end{pmatrix}, \quad x^* = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad x^{(0)} = \begin{pmatrix} 1.2 \\ -1.2 \end{pmatrix}.$$

Die Funktionale $Q(x)$, $R(x)$ und $S(x)$ haben wir im Beispiel 5.6 schon beschrieben. Wir erinnern uns an das Höhenlinienbild mit den verschiedenen Minimum- und Sattelpunktstellen zu diesen Funktionalen.

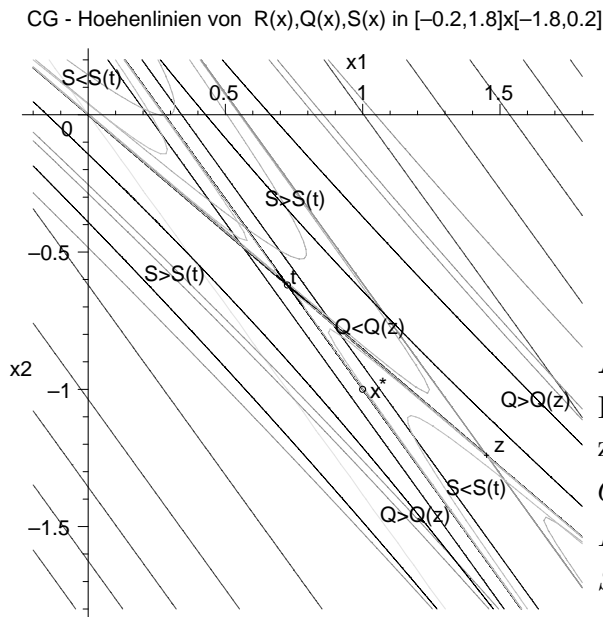


Abb. 5.22 Datei *cg_0040.ps*

Höhenlinienbilder
zu den Funktionalen

$Q(x)$, Sattelpunkt bei $z = (\frac{1}{2}\tilde{A})^{-1}b \neq x^*$,
 $R(x)$, Minimum bei $x^* = A^{-1}b$, und
 $S(x)$, Sattelpunkt bei $t = \tilde{A}^{-1}b = \frac{z}{2}$

Das CG für $A^T A x = A^T b$, also zu $R(x)$, mit der Abstiegsrichtung $p(x)$ – der 1. Schritt ist wie im GV – konvergiert (theoretisch) maximal nach zwei Schritten. Dabei muss man aber die allgemein schlechtere Kondition der spd Matrix $A^T A$ in Kauf nehmen. Bei numerischen Rechnungen kann die Schrittzahl bei entsprechender Abbruchbedingung und sehr kleiner Toleranz damit größer als die Dimension n werden, obwohl $x^{(n)} \approx x^*$ ist.

Rechnung mit Maple bei `Digits:=22`

CG fuer $A^T A x = A^T b$, $A \ll A^T$, A indefinit - Iterationsverlauf

```
m [  x(m) [1],                x(m) [2]                ] [  rd(m) [1],  rd(m) [2]]
-----
0 [1.20000000000000000000, -1.20000000000000000000] [-8.023e-02, -5.791e-02]
1 [1.163417971483916252957, -1.226404750278021501395] [-7.451e-14, 1.032e-13]
2 [1.000000000120819184363, -1.000000000171221117411] [ 3.990e-12, 2.880e-12]
3 [1.000000000071868460864, -1.000000000099568981227] [ 0.000e-01, 0.000e-01]
```


Das übliche CG für $Ax = b$, also zum Funktional $Q(x)$, mit der Suchrichtung $p(x)$ und den Schrittzahlen α und β gemäß (5.13) zeigt zwar einen anfänglichen Trend zur Lösung x^* hin, tut aber letztendlich nicht konvergieren.

Dabei ist sowohl $p(x) \not\perp Q(x) = \text{const}$ als auch $p(x) \not\perp R(x) = \text{const}$.

Ersteres ist sinnvoll, denn sonst würden wir mit $p(x)$ eventuell zum Sattelpunkt z kommen. Die andere Bedingung bedeutet, dass $p(x)$ eine "schräge" Abstiegsrichtung bez. $R(x)$ darstellt, aber sich nicht nahe genug seiner Richtung des steilsten Abstiegs $\hat{r}(x) = A^T(b - Ax)$ bewegt. Der Vektor $\hat{r}(x)$ ist in der Abbildung 5.23 nicht eingezeichnet.

Die Abstiegsrichtung $\tilde{r}(x) = b - \frac{1}{2}\tilde{A}x = \text{rs}(x)$, senkrecht zu den Höhenlinien von $Q(x) = T(x)$, ist die Richtung des steilsten Abstiegs/Anstiegs zum LGS $\frac{1}{2}\tilde{A}x = b$ und liefert eine gegen die Stelle z konvergierende Folge.

Rechnungen mit Maple bei `Digits:=22`

CG fuer $Ax=b$, $A <> A'$, A indefinit - divergenter Iterationsverlauf

```
m [ x(m) [1], x(m) [2] ] [ r(m) [1], r(m) [2]]
-----
0 [ 1.2000000000000000, -1.2000000000000000] [-4.340e-02, -5.080e-02]
1 [ 1.1698401737309086, -1.2353022851260331] [-1.490e-07, 1.273e-07]
2 [ 0.9554513923142095, -1.0521444532302245] [ 6.411e-02, 7.504e-02]
-----
3 [-26261.606990830694, 22435.818403333609] [ 7.853e+03, 9.192e+03]
4 [-1608604450.4136749, 1374277543.9722173] [ 4.810e+08, 5.630e+08]
5 [-68626430490127.186, 58629554541817.570] [ 2.052e+13, 2.402e+13]
-----
xs [ 1, -1 ]
```

Der Iterationsverlauf für $\frac{1}{2}\tilde{A}x = b$ (Funktional $Q(x) = T(x)$) mit der Abstiegsrichtung $\tilde{r}(x)$ führt uns theoretisch in zwei Schritten zum Sattelpunkt z . Bei numerischer Rechnung können es ein paar Schritte mehr sein.

CG fuer $1/2(A+A')x=b$, $A <> A'$, A indefinit - Iterationsverlauf

```
m [ x(m) [1], x(m) [2] ] [ rs(m) [1], rs(m) [2]]
-----
0 [ 1.2000000000000000, -1.2000000000000000] [ 1.666e-01, 1.592e-01]
1 [ 1.314149789593230163, -1.090920489176217035] [-2.938e-03, 3.074e-03]
2 [ 1.451443312434691745, -1.240007836990595611] [ 7.000e-22, 7.000e-22]
-----
3 [ 1.451443312434691745, -1.240007836990595611] [ 0.000e-01, 0.000e-01]
-----
z [ 1.451443312434691745037, -1.240007836990595611285]
```

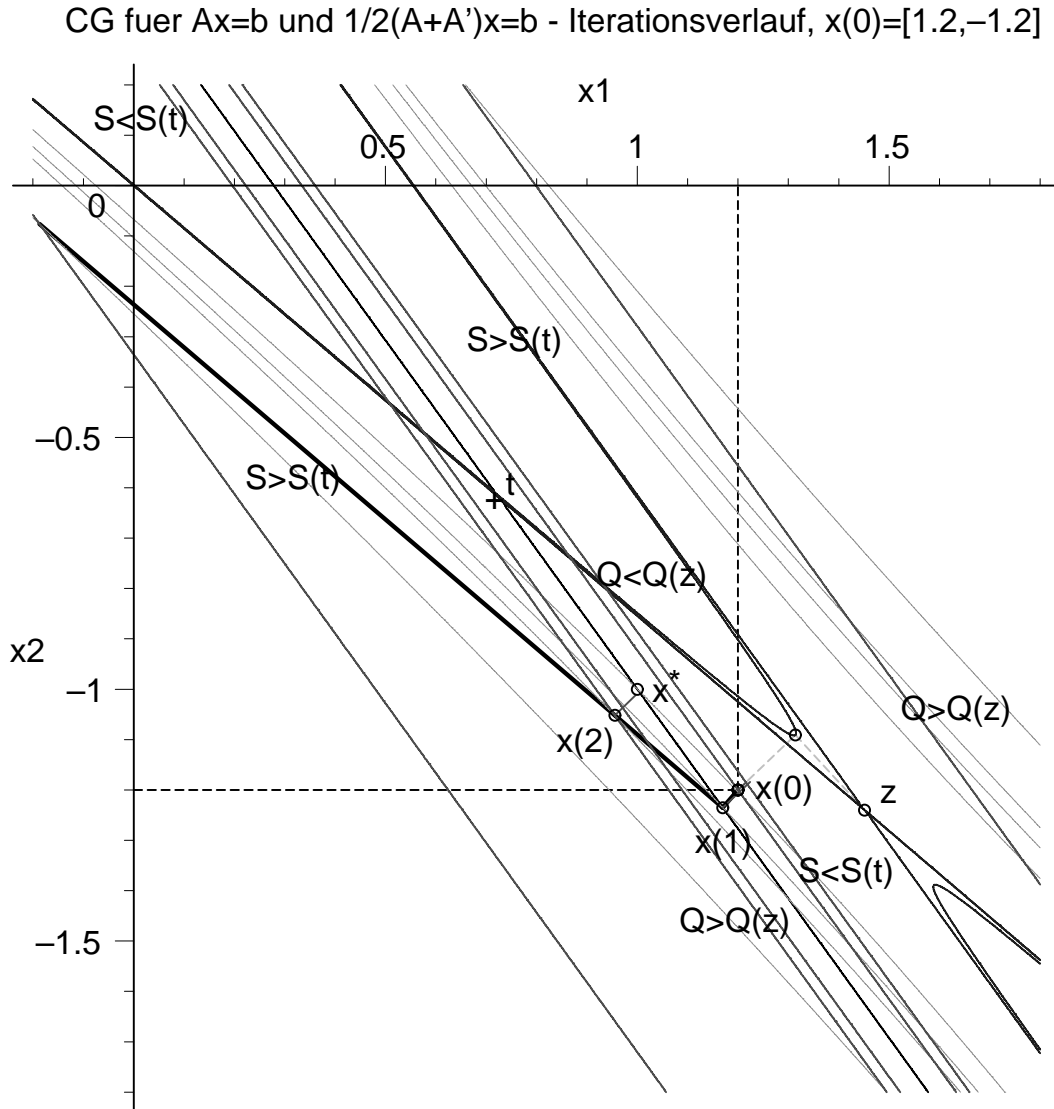


Abb. 5.23 Datei *cg-004.ps*

Höhenlinienbild und CG zu $Q(x)$ mit den Iterierten $x^{(m)}$ (Iterationsverlauf fett gezeichnet), Suchrichtung $p(x)$, $Q(x^*) = 0.0185$, $Q(z) = -0.604101 \cdot 10^{-6}$, Höhenlinien zu $Q(x)$ mit $\text{contours} = [-0.000431, -0.000000604, 0.0185, 0.0860, 0.02995, 0.038805, -0.0495]$.

Zusätzlich eingezeichnet:

- Vektor $\tilde{r}(x) \perp Q(x) = \text{const}$ in den ersten Iterierten,
- Iterationsverlauf für $\frac{1}{2}\tilde{A}x = b$ (Funktional $Q(x) = T(x)$) mit Abstiegsrichtung $\tilde{r}(x)$, Konvergenz gegen z in 2 Schritten (dünn gestrichelt),
- Höhenlinien von $R(x)$ an den Iterierten,
 $\text{contours} = [-0.05357, -0.0558025, -0.0509325, 0.1393]$
- Funktional $S(x)$ mit Minimumstelle $t = \frac{z}{2}$ und Höhenlinie $S(x) = 0$.

Wir testen noch zwei Versionen des CG mit kleinen Veränderungen im Algorithmus. Eine Rechnung mit dem modifizierten CG für $Ax = b$, also zum Funktional $Q(x)$, mit der Suchrichtung $p(x)$ und der Schrittzahl $\bar{\alpha}$ (5.10) hat einen konvergenten Iterationsverlauf mit dem Grenzvektor nahe $(1.303\,486, -1.078\,869)^T$.

Dieses AV konvergiert also weder gegen die Lösung $x^* = (1, -1)^T$ noch gegen den Sattelpunkt $z = (1.451\,443\dots, -1.240\,007\dots)^T$.

Als zweiter Versuch bleibt die Verwendung der Schrittzahlen α und β gemäß (5.14) anstelle von (5.13).

Kann man dann im so “veränderten“ CG die Konvergenz in einem unendlichen Prozess oder vielleicht sogar nach endlich vielen Schritten erreichen? Wie steht es um die A-Orthogonalität der Suchrichtungen bzw. Orthogonalität der Residua?

Unser Beispiel zeigt sowohl mit dem Startvektor $(1.2, -1.2)^T$ als auch mit anderen ein gutartiges Verhalten und das “veränderte“ CG konvergiert nach zwei Schritten. Bei exakter Rechnung haben wir aus den ersten zwei Schritten die Iterationsvektoren $x^{(i)}, r^{(i)}, p^{(i)}$, $i = 0, 1, 2$, als Spalten der folgenden Felder.

$$\begin{aligned}
 xv &= \begin{bmatrix} \frac{6}{5} & \frac{939379507}{802998160} & 1 \\ -\frac{6}{5} & \frac{-495972731}{401499080} & -1 \end{bmatrix} \\
 rv &= \begin{bmatrix} \frac{-217}{5000} & \frac{-59817}{401499080000} & 0 \\ \frac{-127}{2500} & \frac{102207}{802998160000} & 0 \end{bmatrix} \\
 pv &= \begin{bmatrix} \frac{-217}{5000} & \frac{-3044168046387}{25792241798535424000} & 0 \\ \frac{-127}{2500} & \frac{2108746363971}{12896120899267712000} & 0 \end{bmatrix}
 \end{aligned}$$

Numerische Rechnung mit Maple bei `Digits:=22`

Veraendertes CG fuer $Ax=b$, $A \succ A'$, A indefinit - konv. Iterationsverlauf

```

m   [   x(m) [1],       x(m) [2]   ]   [   r(m) [1],   r(m) [2]]
-----
0   [ 1.200000000000, -1.200000000000]   [-4.340e-02, -5.080e-02]
1   [ 1.169840173731, -1.235302285126]   [-1.490e-07,  1.273e-07]
2   [ 1.000000000000, -1.000000000000]   [ 0.000e-01,  0.000e-01]

```

In diesem zweidimensionalen Beispiel brauchen wir nur zwei Schritte auszuführen. Es gilt $x^{(2)} = x^*$ und $r^{(2)} = 0$, $p^{(2)} = 0$. Die Residua $r^{(k)}$, $k = 0, 1, 2$, sind orthogonal. Die Suchrichtungen $p^{(k)}$ erfüllen jedoch nicht die allgemeine A-Orthogonalität $p^{(j)T} A p^{(k)} = 0$ für $j \neq k$. Es ist nämlich $p^{(1)T} A p^{(0)} = -\frac{73592337}{16059963200000000}$, während alle anderen Funktionalwerte verschwinden.

Wie schon am Beispiel 5.10 erkennbar ist, kann man jedoch auch hier für beliebige reguläre LGS folgende weiter gehende Aussagen machen. Das “veränderte“ CG konvergiert i. Allg. in einem unendlichen Iterationsprozess. Die Suchrichtungen $p^{(k)}$ erfüllen die spezielle A-Orthogonalität $p^{(k-1)T} A p^{(k)} = 0$, und es ist $p^{(k)T} A p^{(k-1)} = p^{(k-1)T} A^T p^{(k)} \neq 0$. Die Residua genügen der Beziehung $r^{(k-1)T} r^{(k)} = 0$.

5.3 Verfahren der A-orthogonalen Residua

Das auch mit **Orthomin(0)** bezeichnete Verfahren (vergl. Kapitel 4.2)

$$x^{(m+1)} = x^{(m)} + \alpha_m r^{(m)}, \quad m = 0, 1, \dots, \quad x^{(0)} \text{ gegeben,}$$

$$r^{(m)} = b - Ax^{(m)}, \quad \alpha_m = \frac{r^{(m)T} Ar^{(m)}}{(Ar^{(m)})^T Ar^{(m)}},$$

basiert auf dem zu minimierenden Funktional $R(x)$ sowie der spd Matrix $B = A^T A$ mit seinen Merkmalen (4.2), das sein eindeutiges absolutes Minimum an der Lösung x^* besitzt. Das ist ein wichtiger Vorzug gegenüber dem GV mit dem "schwächeren" Funktional $Q(x)$ im Fall, dass die Bedingung $A = A^T > 0$ nicht gegeben ist.

Als Suchrichtung verwendet Orthomin(0) den negativen Gradienten $r(x) = b - Ax$ vom Funktional $Q(x)$. Unter der Voraussetzung A spd korrespondiert die Suchrichtung $r(x)$ mit der Richtung $\hat{r}(x)$ des steilsten Abstiegs für $R(x)$ und man erhält ein (unendliches) konvergentes Iterationsverfahren.

Im Fall einer nicht spd Matrix A treten jedoch gewisse Probleme auf. Man kann sich leicht vorstellen, dass eine solche ungünstige Richtung verknüpft ist mit einem Vektor auf oder fast parallel zu einer Tangente (Tangentialebene) an die Höhenlinien von $R(x)$. Im Beispiel 4.7 haben wir einen ersten Versuch mit Orthomin(0) für eine nicht spd Matrix aufgezeigt, das Problem gesehen und kurz beschrieben.

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

$$\alpha_m = \frac{r^{(m)T} Ar^{(m)}}{(Ar^{(m)})^T Ar^{(m)}} = \frac{r^{(m)T} Ar^{(m)}}{\|Ar^{(m)}\|_2^2} \quad (5.16)$$

zu machen.

Wenn $r^{(m)} = 0$ ist, gilt $x^* = x^{(m)}$ und die Iteration ist vorzeitig beendet.

Für $r^{(m)} \neq 0$ gilt theoretisch $Ar^{(m)} \neq 0$ und für den Nenner $\|Ar^{(m)}\|_2^2 > 0$. Aber in der numerischen Berechnung könnte das Produkt $Ar^{(m)}$ ein Nullvektor sein. Dann wäre auch der Zähler gleich Null und α_m ist nicht auswertbar.

Unabhängig davon kann für allgemeines A , insbesondere A indefinit, der Zähler $r^{(m)T} Ar^{(m)}$ bei $r^{(m)} \neq 0$ trotzdem Null sein, so dass ein vorzeitiger Abbruch des Verfahrens mit $x^{(m+1)} = x^{(m)}$ eintritt.

In Ausnahmefällen kann Orthomin(0) auch endlich sein.

Es ist also sinnvoll, diese Kontrollen in den Algorithmus mit entsprechenden Fehlermeldungen einzubauen.

Die Suchrichtungen $r^{(m)}$ erfüllen eine Orthogonalitätsbedingung. Es gilt für aufeinanderfolgende Richtungen die Beziehung $r^{(m+1)} \perp Ar^{(m)}$, wobei diese spezielle A-Orthogonalität eigentlich nur Sinn für symmetrische Matrizen macht.

5.3.1 Orthomin(0) mit symmetrischer indefiniter Matrix

Einige dieser beschriebenen Situationen des AV lassen sich an einem einfachen Beispiel nachvollziehen.

Beispiel 5.12

Sei $Ax = b$, A diagonal (damit symmetrisch) und indefinit aus Beispiel 5.1.

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

Das zu minimierende Funktional ist $R(x) = \frac{1}{2}x^T A^T Ax - x^T A^T b = \frac{1}{2}(x_1^2 + x_2^2)$, $R(x^*) = 0$, und die Suchrichtungen werden aus der Satteloberfläche $Q(x) = \frac{1}{2}x^T Ax - x^T b = \frac{1}{2}(x_1^2 - x_2^2)$, $Q(x^*) = 0$, abgeleitet.

(1) Für den Startvektor $x^{(0)} = (1, 0)^T$ erhält man

$$r^{(0)} = b - Ax^{(0)} = (-1, 0)^T, \quad w^{(0)} = Ar^{(0)} = (-1, 0)^T, \quad \alpha_0 = \frac{r^{(0)T} w^{(0)}}{w^{(0)T} w^{(0)}} = 1$$

und mit $x^{(1)} = x^{(0)} + \alpha_0 r^{(0)} = (0, 0)^T = x^*$, $r^{(1)} = 0$, einen endlichen Iterationsverlauf.

(2) Für den Startvektor $x^{(0)} = (1, 1)^T$ haben wir eine Abbruchsituation.

Der Suchvektor $r^{(0)} = (-1, 1)^T$ ist orthogonal zur Höhenlinie $Q(x) = 0$ im Punkt $(1, 1)$ und liegt gleichzeitig auf der Tangente zur Höhenlinie $R(x) = 1$.

Das AV kommt in der Suchrichtung nicht vorwärts. Es ist

$$w^{(0)} = (-1, -1)^T, \quad \alpha_0 = 0 \quad \text{und} \quad x^{(1)} = x^{(0)} = (1, 1)^T.$$

(3) Mit dem Startvektor $x^{(0)} = (1, \frac{1}{2})^T$ lässt sich ein unendlicher konvergenter Iterationsprozess beschreiben.

Die ersten Iterierten sowie die zugehörigen und gegen den Nullvektor tendierenden Residua sind

$$m = 5$$

$$xv = \begin{bmatrix} 1 & \frac{2}{5} & \frac{16}{25} & \frac{32}{125} & \frac{256}{625} & \frac{512}{3125} \\ \frac{1}{2} & \frac{4}{5} & \frac{8}{25} & \frac{64}{125} & \frac{128}{625} & \frac{1024}{3125} \end{bmatrix}$$

$$rv = \begin{bmatrix} -1 & \frac{-2}{5} & \frac{-16}{25} & \frac{-32}{125} & \frac{-256}{625} & \frac{-512}{3125} \\ \frac{1}{2} & \frac{4}{5} & \frac{8}{25} & \frac{64}{125} & \frac{128}{625} & \frac{1024}{3125} \end{bmatrix}$$

Dazu erhalten wir die Beziehungen

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

$$r^{(1)} \perp Ar^{(0)}, \quad r^{(2)} \perp Ar^{(1)}, \quad r^{(2)} \not\perp Ar^{(0)}, \dots$$

Der Iterationsverlauf für $x^{(0)} = (1, \frac{1}{2})^T$ ist wie eine "Zick-Zack"-Kurve.

Orthomin(0) - Höhenlinien und Iterationsverlauf, $x(0)=[1,1/2], [1,1], [1,0]$

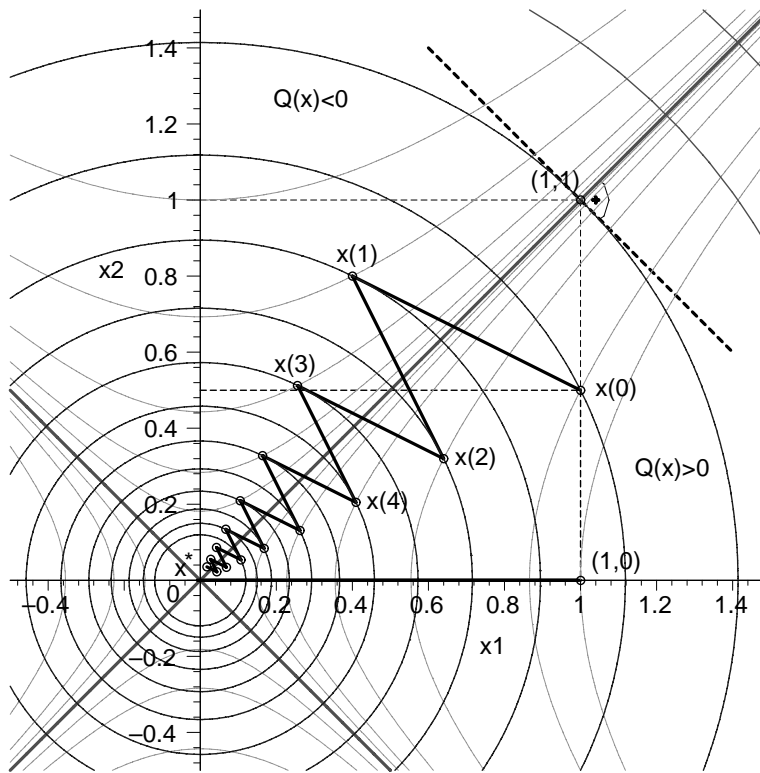


Abb. 5.24 Datei *or_005.ps*

Höhenlinienbild mit Iterationsverlauf $x^{(m)}$ von Orthomin(0) für 3 verschiedene Startvektoren zu $R(x) = \frac{1}{2}(x_1^2 + x_2^2)$ (Höhenlinien sind konzentrische Kreise),

`contours=[2,1.5,1,0.625,0.4,0.256,0.16384,0.10486,0.06711,0.04295,0.02749,0.01759,0.01126,0.00721],`

mit Abstiegsrichtung $r(x)$ von $Q(x) = \frac{1}{2}(x_1^2 - x_2^2)$, `contours=[0.5,-0.5,0.375,-0.24,0.1536,-0.09830,0.06291,-0.04026,0.02577,-0.01649,0.010555]`

Rechnung mit Maple bei `Digits:=16`

Orthomin(0) fuer $Ax=b$, A diagonal und indefinit - konv. Iterationsverlauf

m	[x(m) [1], x(m) [2]]	[r(m) [1], r(m) [2]]
0	[1.0000000000, 0.5000000000]	[-1.000e+00, 5.000e-01]
1	[0.4000000000, 0.8000000000]	[-4.000e-01, 8.000e-01]
2	[0.6400000000, 0.3200000000]	[-6.400e-01, 3.200e-01]
3	[0.2560000000, 0.5120000000]	[-2.560e-01, 5.120e-01]
4	[0.4096000000, 0.2048000000]	[-4.096e-01, 2.048e-01]

5	[0.1638400000, 0.3276800000]	[-1.638e-01, 3.277e-01]
6	[0.2621440000, 0.1310720000]	[-2.621e-01, 1.311e-01]
7	[0.1048576000, 0.2097152000]	[-1.049e-01, 2.097e-01]
8	[0.1677721600, 0.0838860800]	[-1.678e-01, 8.389e-02]
9	[0.0671088640, 0.1342177280]	[-6.711e-02, 1.342e-01]
10	[0.1073741824, 0.0536870912]	[-1.074e-01, 5.369e-02]

20	[0.0115292150, 0.0057646075]	[-1.153e-02, 5.765e-03]
30	[0.0012379400, 0.0006189700]	[-1.238e-03, 6.190e-04]
40	[0.0001329228, 0.0000664614]	[-1.329e-04, 6.646e-05]
50	[0.0000142725, 0.0000071362]	[-1.427e-05, 7.136e-06]
60	[0.0000015325, 0.0000007662]	[-1.532e-06, 7.662e-07]
70	[0.0000001646, 0.0000000823]	[-1.646e-07, 8.228e-08]
80	[0.0000000177, 0.0000000088]	[-1.767e-08, 8.834e-09]
90	[0.0000000019, 0.0000000009]	[-1.897e-09, 9.486e-10]
100	[0.0000000002, 0.0000000001]	[-2.037e-10, 1.019e-10]

Die weiteren Beispiele sind wie im Abschnitt 5.1.1.

Beispiel 5.13

Sei $Ax = b$, $A = A^T$ und indefinit aus den Beispielen 1.4, 5.2, 5.7.

$$\begin{pmatrix} 1 & 0 \\ 0 & -4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad x^* = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Wir rechnen mit verschiedenen Startvektoren die ersten Schritte von Orthomin(0).

$$\begin{aligned} (1) \quad x^{(0)} &= (1, 2)^T && \text{Startvektor,} \\ r^{(0)} &= b - Ax^{(0)} = (-1, 8)^T && \text{Anfangsresiduum, Suchrichtung,} \\ w^{(0)} &= Ar^{(0)} = -(1, 32)^T. \end{aligned}$$

S1

$$\begin{aligned} \alpha_0 &= \frac{r^{(0)T} w^{(0)}}{w^{(0)T} w^{(0)}} = -\frac{51}{205}, \\ x^{(1)} &= x^{(0)} + \alpha_0 r^{(0)} = \frac{2}{205} (128, 1)^T, \\ r^{(1)} &= b - Ax^{(1)} = r^{(0)} - \alpha_0 Ar^{(0)} = \frac{8}{205} (-32, 1)^T. \end{aligned}$$

Orthomin(0) konvergiert mit Abstiegs- und Anstiegsrichtungen gegen x^* .

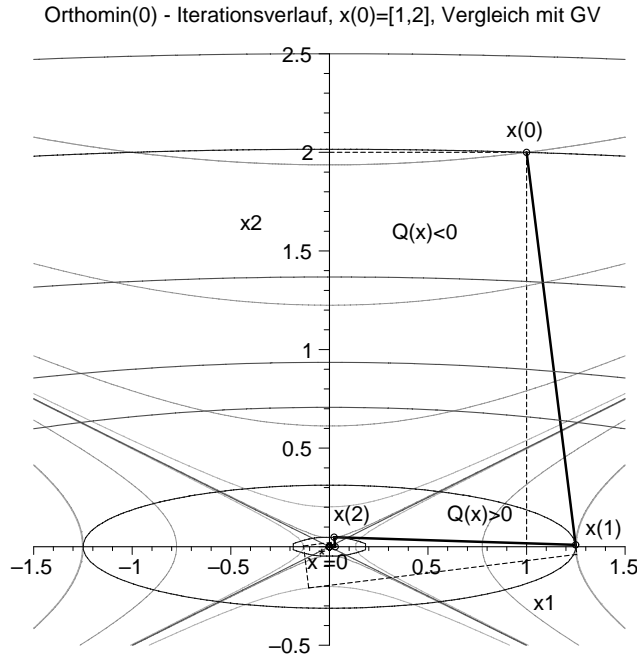


Abb. 5.25 Datei *or_006.ps*

Höhenlinienbild

bei "Zick-Zack"-Iterationsverlauf von Orthomin(0) mit $x^{(0)} = (1, 2)^T$ zu $R(x) = \frac{1}{2}x_1^2 + 8x_2^2$, $R(x^*) = 0$,

mit $\text{contours} = [50, 15, 7, 4,$

$32.5, 0.7805, 0.0187]$,

dazu die Höhenlinien von $Q(x)$ an den Iterierten $x^{(m)}$,

Vergleich mit GV zu $Q(x)$

(dick gestrichelter Verlauf) mit

$\text{contours} = [-7.5, 0.7843,$

$-0.0820, 0.00858]$

Orthomin(0) fuer $Ax=b$, A diagonal und indefinit - konv. Iterationsverlauf

m	[$x(m)$ [1], $x(m)$ [2]]	[$r(m)$ [1], $r(m)$ [2]]
0	[1.0000000000, 2.0000000000]	[-1.000e+00, 8.000e+00]
1	[1.2487804878, 0.0097560976]	[-1.249e+00, 3.902e-02]
2	[0.0240150094, 0.0480300188]	[-2.402e-02, 1.921e-01]
3	[0.0299894751, 0.0002342928]	[-2.999e-02, 9.372e-04]
4	[0.0005767207, 0.0011534414]	[-5.767e-04, 4.614e-03]
5	[0.0007201975, 0.0000056265]	[-7.202e-04, 2.251e-05]
6	[0.0000138500, 0.0000276999]	[-1.385e-05, 1.108e-04]
7	[0.0000172956, 0.0000001351]	[-1.730e-05, 5.405e-07]
8	[0.0000003326, 0.0000006652]	[-3.326e-07, 2.661e-06]
9	[0.0000004154, 0.0000000032]	[-4.154e-07, 1.298e-08]
10	[0.0000000080, 0.0000000160]	[-7.988e-09, 6.390e-08]

Orthomin(0) konvergiert etwas schneller als das GV.

$$(2) \quad x^{(0)} = (1, 0)^T, \quad r^{(0)} = (-1, 0)^T.$$

$$\mathbf{S1} \quad \alpha_0 = 1,$$

$$x^{(1)} = (0, 0)^T,$$

$$r^{(1)} = (0, 0)^T \rightarrow \text{Stopp mit } x^{(1)} = x^*.$$

Orthomin(0) braucht nur endlich viele Schritte.

$$(3) \quad x^{(0)} = (0, 1)^T, \quad r^{(0)} = (0, 4)^T.$$

S1

$$\alpha_0 = -\frac{1}{4},$$

$$x^{(1)} = (0, 0)^T, \quad r^{(1)} = (0, 0)^T \rightarrow \text{Stopp wie in (2)}.$$

$$(4) \quad x^{(0)} = (3, 1)^T, \quad r^{(0)} = (-3, 4)^T.$$

S1

$$\alpha_0 = -\frac{11}{53},$$

$$x^{(1)} = \frac{3}{53}(64, 3)^T, \quad r^{(1)} = \frac{12}{53}(16, 3)^T.$$

Orthomin(0) konvergiert mit "Zick-Zack"-Verlauf, während das GV divergent war.
Die ersten Iterierten sind

$$xv = \begin{bmatrix} 3 & \frac{192}{53} & \frac{432}{265} & \frac{27648}{14045} & \frac{62208}{70225} & \frac{3981312}{3721925} \\ 1 & \frac{9}{53} & \frac{144}{265} & \frac{1296}{14045} & \frac{20736}{70225} & \frac{186624}{3721925} \end{bmatrix}.$$

Berechnungen in Maple mit `Digits:=20`

Orthomin(0) fuer $Ax=b$, A diagonal und indefinit - konv. Iterationsverlauf

m	[x(m) [1], x(m) [2]]	[r(m) [1], r(m) [2]]
0	[3.0000000000, 1.0000000000]	[-3.000e+00, 4.000e+00]
1	[3.6226415094, 0.1698113208]	[-3.623e+00, 6.792e-01]
2	[1.6301886792, 0.5433962264]	[-1.630e+00, 2.174e+00]
3	[1.9685297259, 0.0922748309]	[-1.969e+00, 3.691e-01]
4	[0.8858383766, 0.2952794589]	[-8.858e-01, 1.181e+00]
5	[1.0696916246, 0.0501417949]	[-1.070e+00, 2.006e-01]
6	[0.4813612311, 0.1604537437]	[-4.814e-01, 6.418e-01]
7	[0.5812663923, 0.0272468621]	[-5.813e-01, 1.090e-01]
8	[0.2615698765, 0.0871899588]	[-2.616e-01, 3.488e-01]
9	[0.3158579641, 0.0148058421]	[-3.159e-01, 5.922e-02]
10	[0.1421360838, 0.0473786946]	[-1.421e-01, 1.895e-01]
20	[0.0067342221, 0.0022447407]	[-6.734e-03, 8.979e-03]
30	[0.0003190587, 0.0001063529]	[-3.191e-04, 4.254e-04]
40	[0.0000151166, 0.0000050389]	[-1.512e-05, 2.016e-05]
50	[0.0000007162, 0.0000002387]	[-7.162e-07, 9.549e-07]
60	[0.0000000339, 0.0000000113]	[-3.393e-08, 4.524e-08]
70	[0.0000000016, 0.0000000005]	[-1.608e-09, 2.144e-09]
80	[0.0000000001, 0.0000000000]	[-7.617e-11, 1.016e-10]

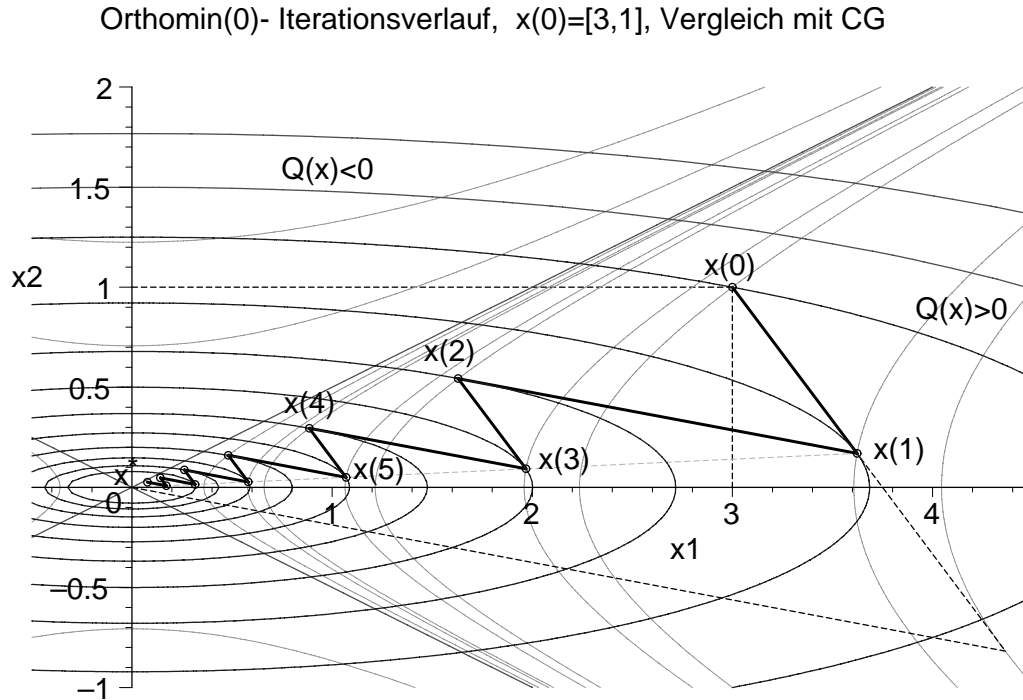


Abb. 5.26 Datei or_007.ps

Höhenlinienbild bei konvergentem “Zick-Zack“-Iterationsverlauf von Orthomin(0) mit $x^{(0)} = (3, 1)^T$ zu $R(x) = \frac{1}{2}x_1^2 + 8x_2^2$, $R(x^*) = 0$,

contours=[12.5,6.792,3.691,2.005,1.0899,0.5922,0.3218,0.1749,0.0950,0.0516, 25,18],

dazu auch die Höhenlinien von $Q(x)$ an den Iterierten $x^{(m)}$ mit

contours=[2.5,8.18,6.504,0.738,1.9205,0.2179,0.567,0.0644,0.1674,-1,-3],

Vergleich mit CG zu $Q(x)$ (dick gestrichelter Verlauf) mit 2 Schritten bis x^*

Das später einbezogene Verfahren CR braucht wie CG auch nur 2 Schritte von $x^{(0)}$ über $x^{(1)}$ und dann zu $x^{(2)} = x^*$.

$$(5) \quad x^{(0)} = (2, 1)^T, \quad r^{(0)} = 2(-1, 2)^T.$$

$$\mathbf{S1} \quad \alpha_0 = -\frac{3}{13},$$

$$x^{(1)} = \frac{1}{13}(32, 1)^T, \quad r^{(1)} = \frac{4}{13}(-8, 1)^T.$$

$$\mathbf{S2} \quad \alpha_1 = \frac{3}{4},$$

$$x^{(2)} = \frac{4}{13}(2, 1)^T, \quad r^{(2)} = \frac{8}{13}(-1, 2)^T.$$

Orthomin(0) konvergiert mit “Zick-Zack“-Verlauf, während das GV abgebrochen ist.

Die ersten Iterierten sind

$$xv = \begin{bmatrix} 2 & \frac{32}{13} & \frac{8}{13} & \frac{128}{169} & \frac{32}{169} & \frac{512}{2197} \\ 1 & \frac{1}{13} & \frac{4}{13} & \frac{4}{169} & \frac{16}{169} & \frac{16}{2197} \end{bmatrix}.$$

Dazu erhalten wir die Beziehungen

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

Berechnungen in Maple mit `Digits:=20`

Orthomin(0) fuer Ax=b, A diagonal und indefinit - konv. Iterationsverlauf

m	[x(m) [1], x(m) [2]]	[r(m) [1], r(m) [2]]
0	[2.0000000000, 1.0000000000]	[-2.000e+00, 4.000e+00]
1	[2.4615384615, 0.0769230769]	[-2.462e+00, 3.077e-01]
2	[0.6153846154, 0.3076923077]	[-6.154e-01, 1.231e+00]
3	[0.7573964497, 0.0236686391]	[-7.574e-01, 9.467e-02]
4	[0.1893491124, 0.0946745562]	[-1.893e-01, 3.787e-01]
5	[0.2330450614, 0.0072826582]	[-2.330e-01, 2.913e-02]
6	[0.0582612654, 0.0291306327]	[-5.826e-02, 1.165e-01]
7	[0.0717061728, 0.0022408179]	[-7.171e-02, 8.963e-03]
8	[0.0179265432, 0.0089632716]	[-1.793e-02, 3.585e-02]
9	[0.0220634378, 0.0006894824]	[-2.206e-02, 2.758e-03]
10	[0.0055158594, 0.0027579297]	[-5.516e-03, 1.103e-02]
20	[0.0000152124, 0.0000076062]	[-1.521e-05, 3.042e-05]
30	[0.0000000420, 0.0000000210]	[-4.195e-08, 8.391e-08]
40	[0.0000000001, 0.0000000001]	[-1.157e-10, 2.314e-10]

Orthomin(0) - Iterationsverlauf, x(0)=[2,1], Vergleich mit GV,CG

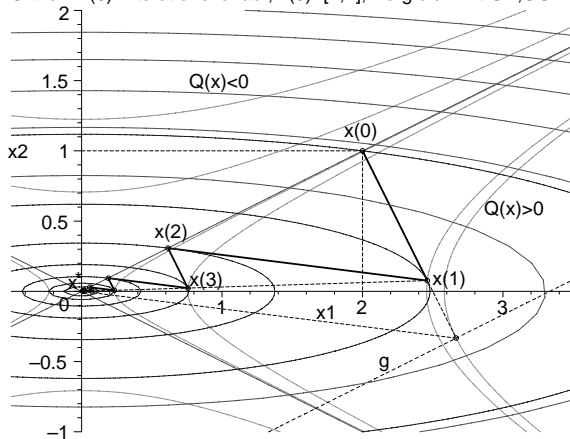


Abb. 5.27 Datei or_008.ps

Höhenlinienbild bei konv. "Zick-Zack"-Iterationsverlauf von Orthomin(0) mit $x^{(0)} = (2, 1)^T$ zu $R(x)$, $R(x^*) = 0$, `contours=[10,3.0769,0.9467,0.2913,0.0896,0.02758,0.0085]`, dazu auch die Höhenlinien von $Q(x)$ an den Iterierten $x^{(m)}$ mit `contours=[3.333,3.0177,0.2857,0.027,0,-1,-3]`, Vergleich mit CG zu $Q(x)$ (dick gestrichelter Verlauf) mit 2 Schritten bis x^* , sowie mit GV (Abbruch nach 1. Schritt)

CR braucht nur 2 Schritte von $x^{(0)}$ über $x^{(1)}$ und dann zu $x^{(2)} = x^*$.

$$(6) \quad x^{(0)} = (-8, 1)^T, \quad r^{(0)} = (8, 4)^T = 4(2, 1)^T.$$

In Richtung des Vektors $(2, 1)^T$ tritt bei der Berechnung der Schrittzahl α das Problem der Null im Zähler auf. Das führt zum Abbruch von Orthomin(0).

Die Suchrichtung im Punkt $x^{(0)}$ ist senkrecht zur Höhenlinie $Q(x) = 30$ und gleichzeitig tangential zur Höhenlinie $R(x) = 40$. Entlang der Geraden $g(\alpha) = x^{(0)} + \alpha r^{(0)}$ lässt sich kein kleinerer Wert des Funktionals $R(x)$ finden.

Schritt $k = 0$

Startvektor $x = [-8.0000000000e+00 \quad 1.0000000000e+00]$

Residuum/Suchr. $r = b - Ax = [8.0000000000e+00 \quad 4.0000000000e+00]$

Funktionswert $Q(x) = 3.0000000000000000e+01$

Funktionswert $R(x) = 4.0000000000000000e+01$

Anfangsfehlerquadrat $r'r = 8.0000000000000000e+01$

Abbruch wegen Zaehler $(Ar)'r=0 \rightarrow \alpha=0$

$$x^{(0)} = (-8, 1.1)^T \text{ oder } x^{(0)} = (-8, 0.9)^T.$$

Für beide Startvektoren ist die Rechnung durchführbar, aber Orthomin(0) konvergiert sehr langsam. In der nachfolgenden Abbildung ist nur der anfängliche "Zick-Zack"-Iterationsverlauf von Orthomin(0) eingezeichnet.

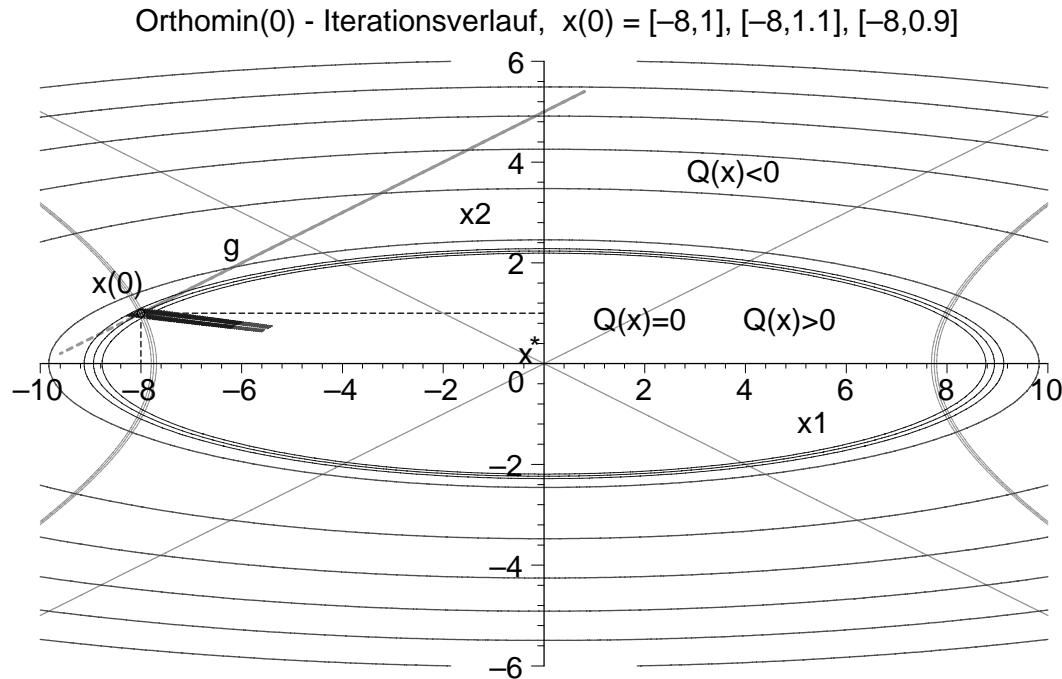


Abb. 5.28 Datei *or_05.ps*

Höhenlinienbild und Iterationsverlauf von Orthomin(0) zu $R(x)$, $R(x^*) = 0$, mit $x^{(0)} = (-8, 1)^T$ (Abbruch), $x^{(0)} = (-8, 1.1)^T$, $(-8, 0.9)^T$ (Konvergenz),

mit $\text{contours}=[40, 41.68, 38.48]$ und $\text{contours}=6$,

dazu die Höhenlinien von $Q(x)$ bei $x^{(0)}$ mit $\text{contours}=[30, 29.58, 30.38]$

Orthomin(0) fuer $Ax=b$, A diagonal, indefinit - 2 konv. Iterationsverlaeufe

m	[x(m) [1], x(m) [2]]	[x(m) [1], x(m) [2]]
0	[-8.000000, 1.100000] R(x(0)) = 41.68 Q(x(0)) = 29.58	[-8.000000, 0.900000] R(x(0)) = 38.48 Q(x(0)) = 30.38
1	[-8.287671, 0.941781]	[-7.641509, 1.061321]
2	[-7.953619, 1.093623]	[-7.943357, 0.893628]
3	[-8.239623, 0.936321]	[-7.587405, 1.053806]
4	[-7.907507, 1.087282]	[-7.887115, 0.887300]
5	[-8.191853, 0.930892]	[-7.533683, 1.046345]
6	[-7.861663, 1.080979]	[-7.831271, 0.881018]
7	[-8.144360, 0.925495]	[-7.480342, 1.038936]
8	[-7.816084, 1.074712]	[-7.775823, 0.874780]
9	[-8.097142, 0.920130]	[-7.427378, 1.031580]
10	[-7.770770, 1.068481]	[-7.720767, 0.868586]
20	[-7.548107, 1.037865]	[-7.451281, 0.838269]
30	[-7.331825, 1.008126]	[-7.191201, 0.809010]
40	[-7.121741, 0.979239]	[-6.940198, 0.780772]
50	[-6.917676, 0.951180]	[-6.697957, 0.753520]
60	[-6.719458, 0.923925]	[-6.464171, 0.727219]
70	[-6.526920, 0.897451]	[-6.238545, 0.701836]
80	[-6.339899, 0.871736]	[-6.020794, 0.677339]
90	[-6.158237, 0.846758]	[-5.810644, 0.653697]
100	[-5.981780, 0.822495]	[-5.607828, 0.630881]
200	[-4.472711, 0.614998]	[-3.930967, 0.442234]
300	[-3.344347, 0.459848]	[-2.755524, 0.309996]
400	[-2.500643, 0.343838]	[-1.931563, 0.217301]
500	[-1.869787, 0.257096]	[-1.353984, 0.152323]
600	[-1.398082, 0.192236]	[-0.949114, 0.106775]
700	[-1.045377, 0.143739]	[-0.665309, 0.074847]
800	[-0.781652, 0.107477]	[-0.466367, 0.052466]
900	[-0.584459, 0.080363]	[-0.326913, 0.036778]
1000	[-0.437013, 0.060089]	[-0.229159, 0.025780]
1100	[-0.326764, 0.044930]	[-0.160635, 0.018071]
1200	[-0.244329, 0.033595]	[-0.112602, 0.012667]
1300	[-0.182690, 0.025119]	[-0.078931, 0.008879]
1400	[-0.136601, 0.018782]	[-0.055329, 0.006224]
1500	[-0.102140, 0.014044]	[-0.038784, 0.004363]
2000	[-0.023872, 0.003282]	[-0.006564, 0.000738]
2500	[-0.005579, 0.000767]	[-0.001110, 0.000124]
3000	[-0.001304, 0.000179]	[-0.000188, 0.000021]

Wir wollen den anfänglichen konvergenten, aber sehr spitzen “Zick-Zack“-Iterationsverlauf von Orthomin(0) für die Startvektoren $x^{(0)} = (-8, 1.1)^T$ bzw. $x^{(0)} = (-8, 0.9)^T$ noch etwas besser auflösen.

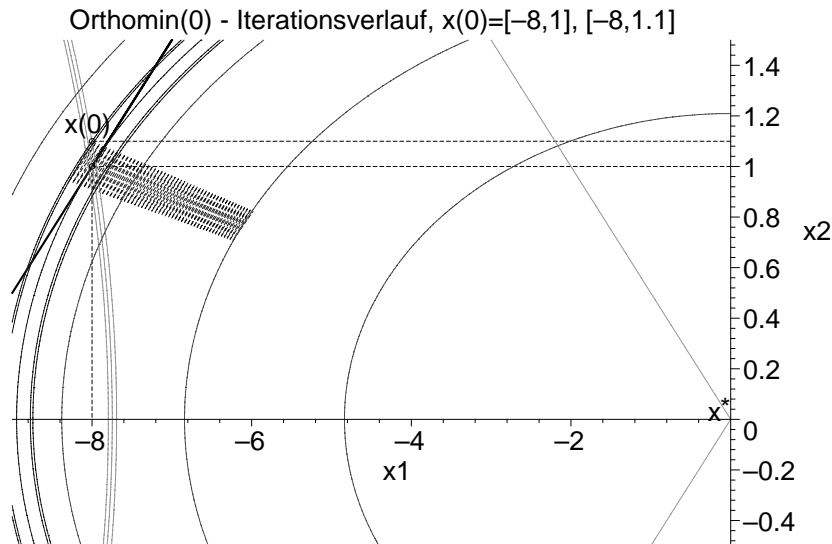


Abb. 5.29 Datei *or_06.ps*

Höhenlinienbild und Iterationsverlauf von Orthomin(0) zu $R(x)$, $R(x^*) = 0$, mit $x^{(0)} = (-8, 1)^T$ (Abbruch), $x^{(0)} = (-8, 1.1)^T$ (Konvergenz), mit $\text{contours}=[40, 41.68, 41.43, 38.48, 38.20]$ und $\text{contours}=4$, dazu Höhenlinien von $Q(x)$ bei $x^{(0)}$ mit $\text{contours}=[30, 29.58, 30.38]$ und $Q(x) = 0$

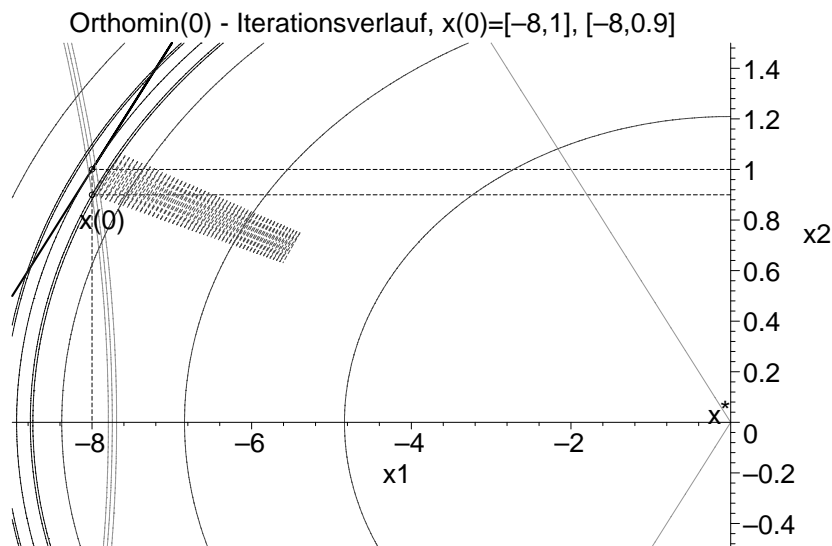


Abb. 5.30 Datei *or_07.ps*

Höhenlinienbild und Iterationsverlauf von Orthomin(0) zu $R(x)$, $R(x^*) = 0$, mit $x^{(0)} = (-8, 1)^T$ (Abbruch), $x^{(0)} = (-8, 0.9)^T$ (Konvergenz), mit $\text{contours}=[40, 41.68, 41.43, 38.48, 38.20]$ und $\text{contours}=4$, dazu Höhenlinien von $Q(x)$ bei $x^{(0)}$ mit $\text{contours}=[30, 29.58, 30.38]$ und $Q(x) = 0$

Beispiel 5.14 Sei $Ax = b$, $A = A^T$ und indefinit aus den Beispielen 1.5, 5.3, 5.8.

$$\begin{pmatrix} 2 & 1 \\ 1 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \quad x^* = \begin{pmatrix} 10/7 \\ 1/7 \end{pmatrix}, \quad x^{(0)} = \begin{pmatrix} 3/2 \\ 1 \end{pmatrix}.$$

Die Funktionale sind $R(x_1, x_2) = \frac{5}{2}x_1^2 - x_1x_2 + 5x_2^2 - 7x_1$, $R(x^*) = -5$, und $Q(x_1, x_2) = x_1^2 + x_1x_2 - \frac{3}{2}x_2^2 - 3x_1 - x_2$, $Q(x^*) = -\frac{31}{14} = -2.214285\dots$

x^* ist Minimumstelle von $R(x)$ und Sattelpunkt von $Q(x)$.

Die Suchrichtung $r(x) = b - Ax \perp Q(x) = \text{const}$, $\nabla R(x) = \text{const}$ ist eine ‘‘schräge‘‘ Abstiegsrichtung bez. $R(x)$ und die Richtung $\hat{r}(x) = A^T r(x) = A^T(b - Ax) \perp R(x) = \text{const}$.

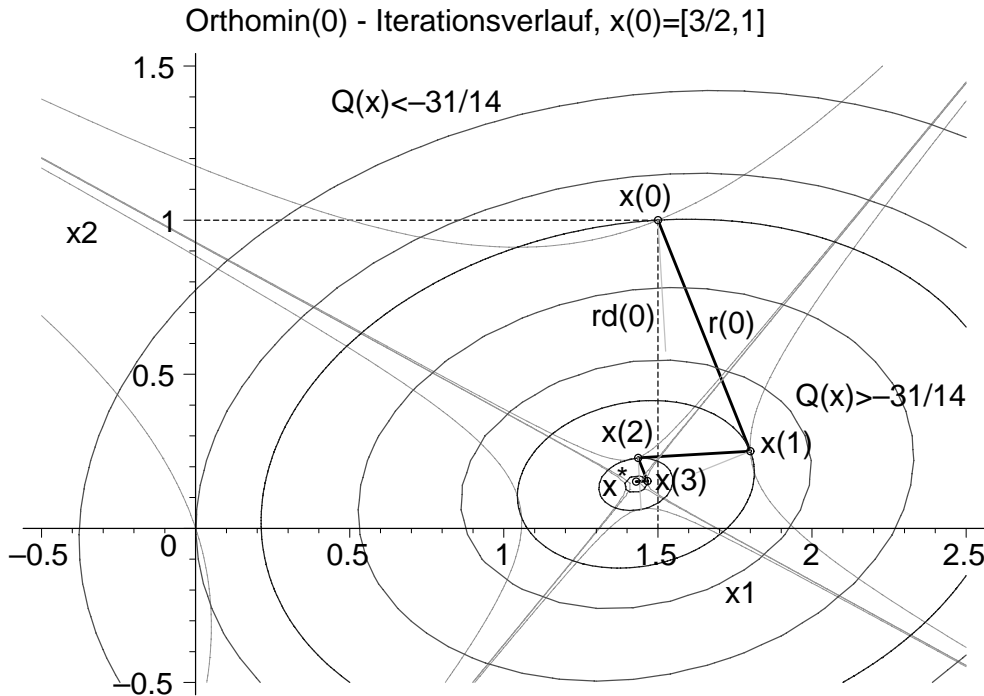


Abb. 5.31 Datei *or_002.ps*

Höhenlinienbild bei ‘‘Zick-Zack‘‘-Iterationsverlauf von Orthomin(0) zu $R(x)$ mit $x^{(0)} = (\frac{3}{2}, 1)^T$, $\text{contours} = [-4.996, -4.964, -4.6375, -4.2, -3, -1.375, 0, 3]$ und in den Iterierten $x^{(m)}$ eingezeichneten Richtungen $\hat{r}(x) \perp R(x)$ (dünne Striche), Höhenlinien zu $Q(x)$ mit $\text{contours} = [-3.25, -2.054, -2.225, -2.214, -2.213, 0]$

$x^{(0)} = (\frac{3}{2}, 1)^T$ ist ein günstiger Startvektor. Die ersten Iterierten sowie die zugehörigen und gegen den Nullvektor tendierenden Residua sind

$$xv = \begin{bmatrix} \frac{3}{2} & \frac{9}{5} & \frac{201}{140} & \frac{513}{350} & \frac{2001}{1400} & \frac{5013}{3500} \\ 1 & \frac{1}{4} & \frac{8}{35} & \frac{43}{280} & \frac{53}{350} & \frac{403}{2800} \end{bmatrix}, \quad rv = \begin{bmatrix} -1 & \frac{-17}{20} & \frac{-1}{10} & \frac{-17}{200} & \frac{-1}{100} & \frac{-17}{2000} \\ \frac{5}{2} & \frac{-1}{20} & \frac{1}{4} & \frac{-1}{200} & \frac{1}{40} & \frac{-1}{2000} \end{bmatrix}.$$

Andere Startvektoren werden nicht getestet. Aber Situationen, wie sie in Beispiel 5.13 aufgetreten sind, können auch hier vorkommen.

5.3.2 Orthomin(0) mit nicht symmetrischer Matrix

Wir betrachten hier das zu minimierende Funktional $R(x)$ sowie die Funktionale $Q(x) = T(x), S(x)$, die wir in den Kapiteln 5.1.2 und 5.2.2 für das GV bzw. CG mit nicht symmetrischer Matrix verwendet haben.

Das entscheidende Problem wird jedoch bei einer nicht spd Matrix A das Verhalten der Suchrichtung $r(x) = b - Ax$ sein, möglichst ähnlich zur Abstiegsrichtung $\hat{r}(x)$ für $R(x)$ und die nicht in Richtung einer Tangente an eine Höhenlinie von $R(x)$ zeigen sollte.

Beispiel 5.15 Sei $Ax = b, A \neq A^T, A > 0$ aus den Beispielen 1.6, 5.4, 5.9.

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

Wir erinnern uns an das Höhenlinienbild mit den verschiedenen Minimumstellen.

Orthomin(0) - Höhenlinien von $R(x), Q(x), S(x)$ in $[-1,2] \times [-1,2]$

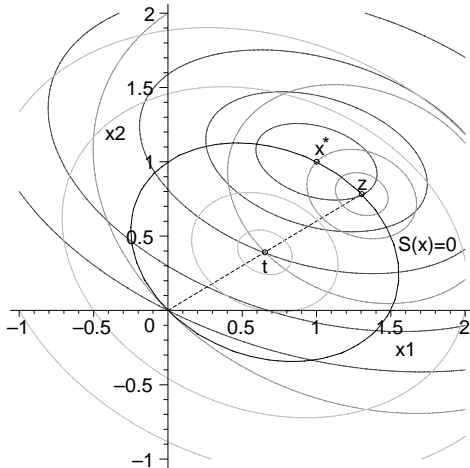


Abb. 5.32 Datei *or_0030.ps*

Höhenlinienbilder (Ellipsenscharen) zu den Funktionalen

$R(x)$, Minimum bei $x^* = A^{-1}b$, und
 $Q(x)$, Minimum bei $z = (\frac{1}{2}\tilde{A})^{-1}b \neq x^*$,
 $S(x)$, Minimum bei $t = \tilde{A}^{-1}b = \frac{z}{2}$

x^* ist Minimumstelle von $R(x)$, $R(x^*) = -9$, und Stelle mit $S(x^*) = 0$.

Die Suchrichtung $r(x) = b - Ax$ ist eine "schräge" Abstiegsrichtung bez. $R(x)$, es gilt $r(x) \not\perp Q(x) = \text{const}$ und $\not\perp R(x) = \text{const}$. Weitere Richtungen sind $\hat{r}(x) = A^T r(x) = A^T(b - Ax) \perp R(x) = \text{const}$ sowie $\tilde{r}(x) = b - \frac{1}{2}\tilde{A}x = b - \frac{1}{2}(A + A^T)x \perp Q(x) = \text{const}$.

Für den Startvektor $x^{(0)} = (1, -1)^T$ sind die Richtungsvektoren

$$\begin{aligned} r^{(0)} &= b - Ax^{(0)} = 2(1, 3)^T, \\ \hat{r}^{(0)} &= A^T(b - Ax^{(0)}) = A^T r^{(0)} = 4(1, 5)^T, \\ \tilde{r}^{(0)} &= b - \frac{1}{2}\tilde{A}x^{(0)} = \frac{1}{2}(3, 11)^T. \end{aligned}$$

In der Abbildung 5.34 sind neben der Suchrichtung $r^{(0)} = \mathbf{r}(0)$ auch die beiden anderen Richtungen eingetragen. $\hat{r}^{(0)} = \mathbf{rd}(0)$ ist als dicker Strich gezeichnet und $\tilde{r}^{(0)} = \mathbf{rs}(0)$ liegt zwischen diesem und $r^{(0)}$.

Um die Konvergenz von Orthomin(0) zu erzwingen, kann man das LGS gemäß $Bx = A^T Ax = A^T b = c$ symmetrisieren. Das zugehörige Funktional ist

$$R_B(x) = \frac{1}{2}(Bx)^T Bx - (Bx)^T c.$$

Die Abstiegsrichtung ist die Richtung des steilsten Abstiegs $\hat{r}(x) = A^T(b - Ax) = c - Bx = \text{rd}(x)$ zum Funktional $R(x)$. Dabei muss man aber die allgemein sehr schlechte Kondition der spd Matrix $(A^T A)^2$ in Kauf nehmen.

Orthomin(0) fuer A'Ax=A'b - Iterationsverlauf, x(0)=[1,-1]

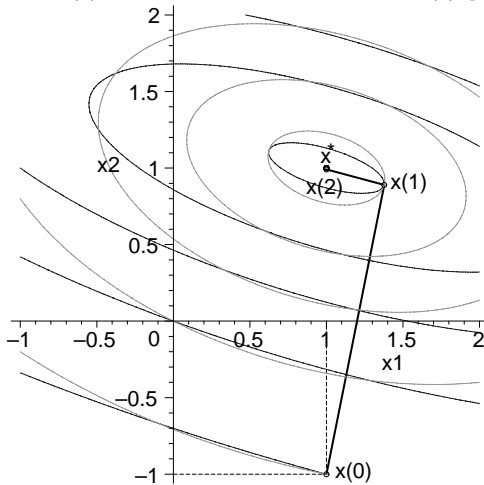


Abb. 5.33 Datei or_0031.ps

Höhenlinienbild und Orthomin(0) zu $R_B(x)$, Minimum bei $x^* = B^{-1}c = A^{-1}b = (1, -1)^T$, $R_B(x^*) = -90$, Abstiegsrichtung $\hat{r}(x) = A^T(b - Ax) = c - Bx$, contours=[118,0,-50,-75,-89.1062,-89.9961,-89.99998], dazu Höhenlinien von $R(x)$, $R(x^*) = -9$, contours=[11,0,-5,-7.5,-8.7368,-8.9996,-8.99999]

Berechnungen in Maple mit Digits:=16

Orthomin(0) fuer A'Ax=A'b, A<>A', A>0 - konvergenter Iterationsverlauf

m	[x(m) [1], x(m) [2]]	[rd(m) [1], rd(m) [2]]
0	[1.0000000000, -1.0000000000]	[4.000e+00, 2.000e+01]
1	[1.3779310345, 0.8896551724]	[-1.291e+00, 3.476e-01]
2	[1.0000000000, 0.9914058355]	[1.719e-02, 8.594e-02]
3	[1.0016240007, 0.9995258392]	[-5.548e-03, 1.494e-03]
4	[1.0000000000, 0.9999630702]	[7.386e-05, 3.693e-04]
5	[1.0000069785, 0.9999979625]	[-2.384e-05, 6.418e-06]
6	[1.0000000000, 0.9999998413]	[3.174e-07, 1.587e-06]
7	[1.000000300, 0.999999912]	[-1.024e-07, 2.758e-08]
8	[1.000000000, 0.999999993]	[1.364e-09, 6.819e-09]
9	[1.0000000001, 1.000000000]	[-4.402e-10, 1.185e-10]
10	[1.000000000, 1.000000000]	[5.860e-12, 2.930e-11]

Orthomin(0) für $Ax = b$, also zum Funktional $R(x)$, mit der Suchrichtung $\hat{r}(x) = A^T(b - Ax)$ konvergiert ebenfalls gegen x^* .

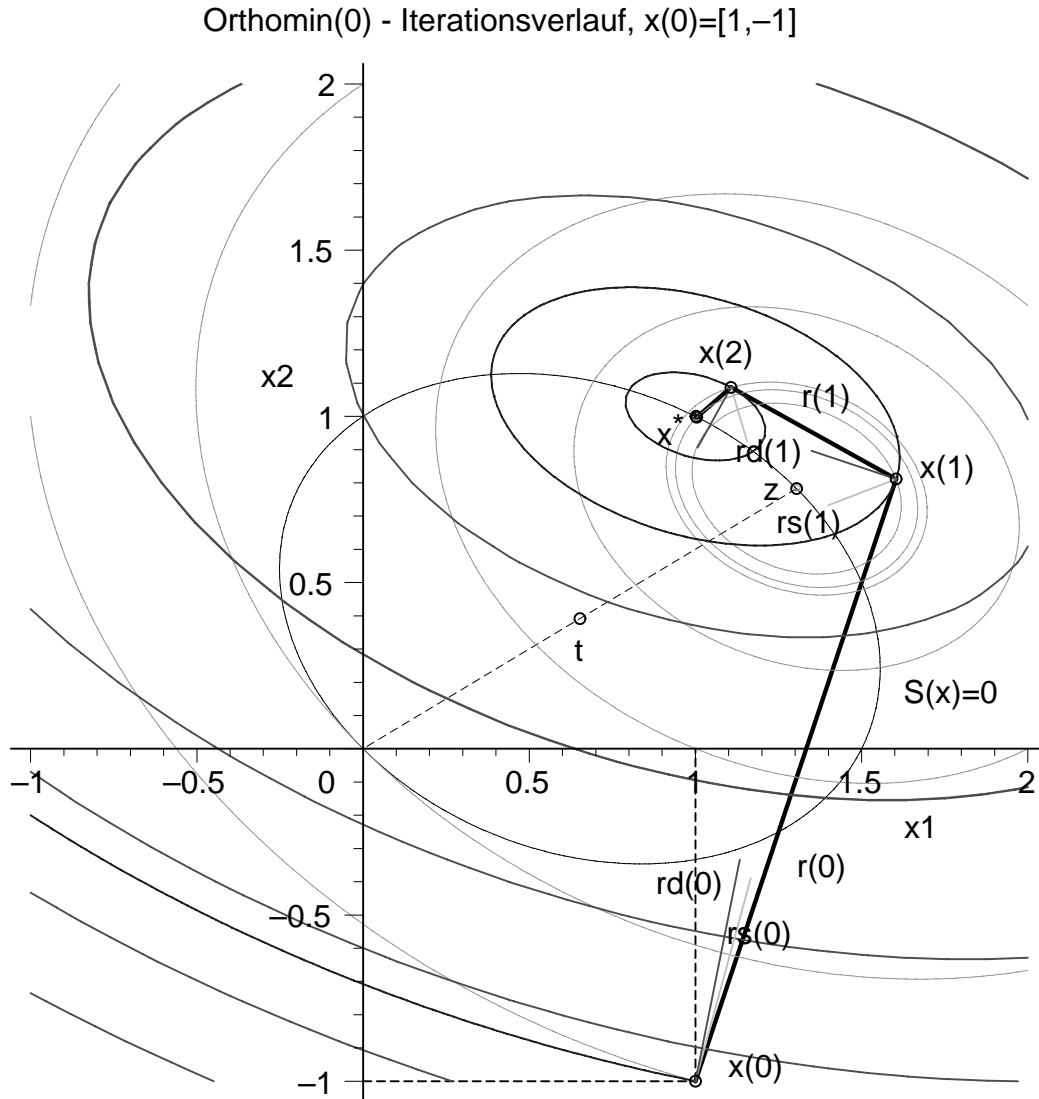


Abb. 5.34 Datei *or_003.ps*

Höhenlinienbild und Orthomin(0) zu $R(x)$ mit den Iterierten $x^{(m)}$ (Iterationsverlauf fett gezeichnet), Suchrichtung $r(x)$, $R(x^*) = -9$, $R(z) = -\frac{4608}{529} = -8.710775$, Höhenlinien zu $R(x)$ mit $\text{contours}=[11,-3,-7,-8.32,-8.92,-8.999]$ und $\text{contours}=5$.

Zusätzlich eingezeichnet:

- Vektor $\tilde{r}(x) \perp Q(x) = \text{const}$ in den ersten Iterierten (dünner Strich),
- Vektor $\hat{r}(x) \perp R(x) = \text{const}$ in den ersten Iterierten (dicker Strich),
- Höhenlinien von $Q(x)$, $Q(x^*) = -3$, $Q(z) = -\frac{72}{23} = -3.130434$, an den Iterierten, $\text{contours}=[2,0,-2,-2.7,-3.13,-3.0352,-2.9825,-3.003]$,
- Funktional $S(x)$ mit Minimumstelle $t = \frac{z}{2}$ und Höhenlinie $S(x) = 0$.

Orthomin(0) fuer $Ax=b$, $A \succ A'$, $A > 0$ - konvergenter Iterationsverlauf

m	[x(m) [1], x(m) [2]]	[r(m) [1], r(m) [2]]
0	[1.0000000000, -1.0000000000]	[2.000e+00, 6.000e+00]
1	[1.6037735849, 0.8113207547]	[-1.019e+00, 5.660e-01]
2	[1.1072694186, 1.0871564026]	[-3.017e-01, -2.615e-01]
3	[1.0041379894, 0.9977758307]	[-6.052e-03, 6.673e-03]
4	[1.0017966647, 1.0003572913]	[-3.951e-03, -1.072e-03]
5	[1.0001053942, 0.9998984194]	[-1.092e-04, 3.047e-04]
6	[1.0000705342, 0.9999956953]	[-1.368e-04, 1.291e-05]
7	[0.9999993299, 1.0000024188]	[-1.079e-06, -7.257e-06]
8	[0.9999990075, 1.0000002501]	[1.735e-06, -7.502e-07]
9	[0.999998994, 0.9999998644]	[3.367e-07, 4.069e-07]
10	[1.0000000085, 0.9999999961]	[-1.311e-08, 1.160e-08]

Wir kontrollieren noch einige Eigenschaften der auftretenden Vektoren.

Bei exakter Rechnung haben wir aus den ersten drei Schritten die Iterationsvektoren $x^{(i)}, r^{(i)}$, $i = 0, 1, 2, 3$, als Spalten der folgenden Felder.

$$\begin{array}{c}
 m = 3 \\
 xv = \begin{bmatrix} 1 & \frac{85}{53} & \frac{11561}{10441} & \frac{3533177}{3518617} \\ -1 & \frac{43}{53} & \frac{11351}{10441} & \frac{3510791}{3518617} \end{bmatrix} \\
 rv = \begin{bmatrix} 2 & \frac{-54}{53} & \frac{-3150}{10441} & \frac{-21294}{3518617} \\ 6 & \frac{30}{53} & \frac{-2730}{10441} & \frac{23478}{3518617} \end{bmatrix}
 \end{array}$$

Trotz $A \neq A^T$ ist die A-Orthogonalität aufeinander folgender Suchrichtungen $r^{(i)}$ erfüllt im Sinne von $r^{(k+1)T} A r^{(k)} = 0$, mehr aber nicht.

Weiter ist z. B.

$$\begin{aligned}
 r^{(2)T} A r^{(0)} &= -\frac{80640}{10441}, \quad r^{(0)T} A r^{(1)} = -\frac{384}{53}, \quad r^{(0)T} A r^{(2)} = -\frac{67200}{10441}, \\
 r^{(1)T} A r^{(2)} &= \frac{241920}{553373}, \quad r^{(2)T} A r^{(3)} = -\frac{132088320}{36737880097}.
 \end{aligned}$$

Unser LGS ist zu klein, um weiter gehende Aussagen zu treffen.

Auch eine Erhöhung seiner Dimension um Eins liefert außer Informationen zur Konvergenzgeschwindigkeit keine neuen Erkenntnisse.

Beispiel 5.16 Sei $Ax = b$, $A \neq A^T$, $A > 0$ aus den Beispielen 5.5, 5.10.

$$\begin{pmatrix} 2 & 1 & 0 \\ 0 & 3 & 1 \\ 1 & 2 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \\ 7 \end{pmatrix}, \quad x^* = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad x^{(0)} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}.$$

Orthomin(0) basiert auf dem Funktional $R(x)$ mit seinem Minimum bei $x^* = (1, 1, 1)^T$ und $R(x^*) = -37$, und der Suchrichtung $r(x) = b - Ax$ von $Q(x)$, wobei sein Minimum bei $z = (\frac{38}{37}, \frac{16}{37}, \frac{54}{37})^T = (1.027\ 027\dots, 0.432\ 432\dots, 1.459\ 459\dots)^T$ liegt.

Weiter sind $Q(z) = -\frac{278}{37} = -7.513\,513\dots$, $Q(x^*) = -7$, $R(z) = -\frac{49050}{1369} = -35.829\,072\dots$

Orthomin(0) konvergiert fast so langsam zur Lösung x^* wie das GV.

Die Iterationsvektoren $x^{(i)}, r^{(i)}$, $i = 0, 1, 2, 3$, aus den ersten drei Schritten sind die Spalten der folgenden Felder.

$$\begin{aligned}
 xv &= \begin{bmatrix} 1 & \frac{2242}{1633} & \frac{408103387253}{301565213483} & \frac{59478566174796329302643921999848306}{49718239186648330433881354938383401} \\ -1 & \frac{997}{3266} & \frac{24052214947}{26223062042} & \frac{94253369937420357915773713444315117}{99436478373296660867762709876766802} \\ 0 & \frac{2436}{1633} & \frac{339718136268}{301565213483} & \frac{47328318999310746449086899186439446}{49718239186648330433881354938383401} \end{bmatrix} \\
 rv &= \begin{bmatrix} 2 & \frac{-167}{3266} & \frac{-376223211895}{603130426966} & \frac{-33858199516715692523061271813407935}{99436478373296660867762709876766802} \\ 7 & \frac{5201}{3266} & \frac{73482603985}{603130426966} & \frac{20329165682304076825555900801242965}{99436478373296660867762709876766802} \\ 8 & \frac{-1552}{1633} & \frac{-209220381725}{301565213483} & \frac{4982462197078640022404252378762600}{49718239186648330433881354938383401} \end{bmatrix}
 \end{aligned}$$

Rechnung mit Maple bei `Digits:=16` und Vergleich

Orthomin(0) fuer $Ax=b$, $A \ll A'$, $A > 0$ - konvergenter Iterationsverlauf

m	[x(m) [1],	x(m) [2],	x(m) [3]] [r(m) [1],	r(m) [2],	r(m) [3]]
0	[1.000000000000,	-1.000000000000,	0.000000000000]	[2.000e+00,	7.000e+00,	8.000e+00]
1	[1.372933251684,	0.305266380894,	1.491733006736]	[-5.113e-02,	1.592e+00,	-9.504e-01]
2	[1.353284029479,	0.917216109563,	1.126516325754]	[-6.238e-01,	1.218e-01,	-6.938e-01]
3	[1.196312804874,	0.947875181013,	0.951930715439]	[-3.405e-01,	2.044e-01,	1.002e-01]
4	[1.069868608831,	1.023794900539,	0.989144946309]	[-1.635e-01,	-6.053e-02,	-7.404e-02]
5	[1.033823523832,	1.010453201499,	0.972825748181]	[-7.810e-02,	-4.185e-03,	5.397e-02]
6	[0.999643435074,	1.008621507906,	0.996444105605]	[-7.908e-03,	-2.231e-02,	-2.663e-03]
7	[0.998088915538,	1.004236386219,	0.995920674815]	[-4.142e-04,	-8.630e-03,	9.756e-03]
8	[0.997908208247,	1.000471517287,	1.000176678541]	[3.712e-03,	-1.591e-03,	4.420e-04]
9	[0.999797972299,	0.999661442834,	1.000401716821]	[7.426e-04,	6.140e-04,	-7.277e-04]
10	[1.000130757283,	0.999936572701,	1.000075603251]	[-1.981e-04,	1.147e-04,	-3.063e-04]
11	[1.000074758090,	0.999968992310,	0.999989007925]	[-1.185e-04,	1.040e-04,	3.123e-05]
12	[1.000036781803,	1.000002324175,	0.999999014229]	[-7.589e-05,	-5.987e-06,	-3.749e-05]
13	[1.000017729245,	1.000000821127,	0.999989602641]	[-3.628e-05,	7.934e-06,	2.222e-05]
14	[1.000003751538,	1.000003877907,	0.999998162704]	[-1.138e-05,	-9.796e-06,	-4.158e-06]
15	[1.000001458020,	1.000001903712,	0.999997324742]	[-4.820e-06,	-3.036e-06,	5.436e-06]
16	[0.999999291273,	1.000000538916,	0.99999768341]	[8.785e-07,	-1.385e-06,	5.575e-07]
17	[0.999999657337,	0.99999961785,	1.000000006650]	[7.235e-07,	1.140e-07,	4.165e-07]
18	[0.999999826504,	0.99999988437,	1.000000098027]	[3.586e-07,	-6.334e-08,	-1.955e-07]
19	[0.99999976014,	0.99999962026,	1.00000016513]	[8.595e-08,	9.741e-08,	3.388e-08]
20	[0.99999992857,	0.99999981116,	1.00000023153]	[3.317e-08,	3.350e-08,	-4.770e-08]
25	[0.99999999972,	0.99999999958,	1.00000000022]	[9.780e-11,	1.028e-10,	2.400e-11]
30	[1.000000000000,	1.000000000000,	1.000000000000]	[-7.808e-13,	-1.040e-12,	-3.048e-13]
35	[1.000000000000,	1.000000000000,	1.000000000000]	[1.962e-15,	6.133e-15,	-7.160e-15]
GV						
20	[0.999999965066,	1.000000096702,	0.999999978751]	[-2.683e-08,	-2.689e-07,	-7.347e-08]
CG						
20	[1.000000000002,	0.999999999999,	1.000000000000]	[-3.863e-12,	2.032e-12,	-7.260e-13]

Die Suchrichtungen und Residua genügen der Beziehung $r^{(k+1)T} A r^{(k)} = 0$.


```

1 [ 1.1698401737311680804846441523551893907368,
   -1.2353022851257295279119833424045248606123] [-1.490e-07, 1.273e-07]
2 [ 1.1698401737303229102782684402708033731191,
   -1.2353022851250074773384422266204371505730] [-1.490e-07, 1.273e-07]
3 [ 1.1698401737303229033788188763531772365449,
   -1.2353022851250074714440418969814363952947] [-1.490e-07, 1.273e-07]
4 [ 1.1698401737303229033787625533963593140732,
   -1.2353022851250074714439937786421153971384] [-1.490e-07, 1.273e-07]
5 [ 1.1698401737303229033787625529365725679902,
   -1.2353022851250074714439937782493062139338] [-1.490e-07, 1.273e-07]
6 [ 1.1698401737303229033787625529365688145677,
   -1.2353022851250074714439937782493030072758] [-1.490e-07, 1.273e-07]
7 [ 1.1698401737303229033787625529365688145371,
   -1.2353022851250074714439937782493030072497] [-1.490e-07, 1.273e-07]
8 [ 1.1698401737303229033787625529365688145371,
   -1.2353022851250074714439937782493030072497] [-1.490e-07, 1.273e-07]

```

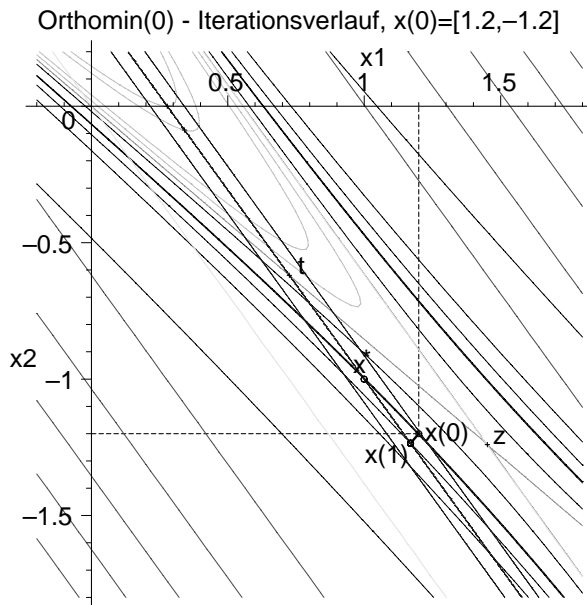


Abb. 5.36 Datei *or_004.ps*

Höhenlinienbild und Orthomin(0)

zu $R(x)$ mit den Iterierten

$x^{(0)}, x^{(1)}, x^{(2,3,\dots)} \approx x^{(1)}$, (nicht gegen x^* konvergierender Iterationsverlauf fett gezeichnet),

Suchrichtung $r(x)$ von $Q(x)$,

Höhenlinien zu $R(x)$ mit `contours=5`

und `contours=[-0.053, -0.0558, 0]`,

$R(x^*) = -0.0558025$, $R(z) \approx 0.9 \cdot 10^{-6}$,

$Q(x^*) = 0.0185$, $Q(z) \approx -0.6 \cdot 10^{-6}$

Zusätzlich sind in der Abbildung 5.36 eingezeichnet:

– Höhenlinien zu $Q(x)$ mit `contours=[-0.03018, -0.0259, -0.01, -0.005, 0, 0.005, 0.01776, 0.0185, 0.0299, 0.05, 0.2]`,

– Sattelpunkt z von Funktional $Q(x)$, Minimumstelle $t = \frac{z}{2}$ von Funktional $S(x)$.

Die Suchrichtungen und Residua genügen der Beziehung $r^{(k+1)T} A r^{(k)} = 0$.

Orthomin(0) für $A^T A x = A^T b$ mit der Abstiegsrichtung $\hat{r}(x)$ konvergiert.

Dabei muss man aber die sehr schlechte Kondition der spd Matrix $(A^T A)^2$ in Kauf nehmen und bei numerischen Rechnungen in Maple wird die Konvergenz so ab `Digits:=24,25,...`, mit grober Genauigkeit erst bei späteren Iterierten, erkennbar.

Rechnung mit Maple bei Digits:=27

```

Orthomin(0) fuer A'Ax=A'b, A<>A', A indefinit - Iterationsverlauf
m [ x(m) [1], x(m) [2] ] [ rd(m) [1], rd(m) [2]]
-----
0 [ 1.200000000000000000000000,
   -1.200000000000000000000000] [-8.023e-02, -5.791e-02]
1 [ 1.1634179714839162529570124,
   -1.2264047502780215013947642] [-7.451e-14, 1.032e-13]
2 [ 1.0000000198450518030645532,
   -1.0000000274940018076140880] [ 3.596e-17, 2.598e-17]
3 [ 1.0000000198450518194624621,
   -1.0000000274940017957694545] [-9.048e-21, 1.254e-20]
4 [ 1.000000000000000023944482834,
   -1.000000000000000033173491349] [ 4.367e-24, 3.155e-24]
5 [ 1.000000000000000023944482854,
   -1.000000000000000033173491335] [-1.099e-27, 1.522e-27]
6 [ 0.9999999999999999845176356,
   -0.9999999999999999785502120] [ 5.304e-31, 3.831e-31]
7 [ 0.9999999999999999845176356,
   -0.9999999999999999785502120] [-1.334e-34, 1.849e-34]
8 [ 0.9999999999999999845173429,
   -0.9999999999999999785498065] [ 6.441e-38, 4.652e-38]

```

Berechnung von $x^{(10)}$, $x^{(100)}$ und ev. $x^{(200)}$ zu Orthomin(0) für $A^T Ax = A^T b$ bei $x^{(0)} = (1.2, -1.2)^T$ und verschiedenen Genauigkeiten von Maple

```

Digits [ x(10) [1], x(10) [2] ]
        [ x(100) [1], x(100) [2] ]
ev. [ x(200) [1], x(200) [2] ]
-----
22 [ 1.163124778171131158080, -1.225998550408069406742 ]
    [ 1.160509596690955248782, -1.222375390087484944836 ]
23 [ 1.1628865231094641596490, -1.2256684638377278687512 ]
    [ 1.1581805911646708903331, -1.2191487075519809587076 ]
    [ 1.1531110635798911919781, -1.2121252136459111459074 ]
-----
24 [ 1.03519788304360539061793, -1.04876433019234296359013 ]
    [ 1.00000003511287863639610, -1.00000004864656222954178 ]
    [ 0.999999999999953667237807, -0.99999999999935809033912 ]
25 [ 1.000000753684912837040224, -1.000001044180410082575306 ]
    [ 0.999999999999932959319929, -0.99999999999907119588591 ]
26 [ 1.00000000000000003141670834, -1.00000000000000004352576367 ]
    [ 1.00000000000000003141670565, -1.00000000000000004352575995 ]
27 [ 0.999999999999999984517635589, -0.999999999999999978550211973]
    [ 0.999999999999999984517342933, -0.999999999999999978549806518]

```

5.4 Verfahren der konjugierten Residua

Das mit **CR** bezeichnete (theoretisch) endliche AV (vergl. Kapitel 4.3)

$$x^{(0)} \text{ gegeben, } p^{(0)} = r^{(0)} = b - Ax^{(0)},$$

$$m = 0, 1, \dots$$

$$\begin{aligned} x^{(m+1)} &= x^{(m)} + \alpha_m p^{(m)}, \\ \alpha_m &= \frac{r^{(m)T} Ap^{(m)}}{(Ap^{(m)})^T Ap^{(m)}} = \frac{r^{(m)T} Ar^{(m)}}{(Ap^{(m)})^T Ap^{(m)}}, \end{aligned} \quad (5.17)$$

$$\begin{aligned} p^{(m+1)} &= r^{(m+1)} + \beta_m p^{(m)}, \\ \beta_m &= \frac{r^{(m)T} Ap^{(m+1)}}{r^{(m)T} Ap^{(m)}} = \frac{r^{(m+1)T} Ar^{(m+1)}}{r^{(m)T} Ar^{(m)}}, \end{aligned}$$

basiert auf der spd Matrix $B = A^T A$ sowie dem zu minimierenden Funktional $R(x)$ sowie mit seinen Merkmalen (4.2), das sein eindeutiges absolutes Minimum an der Lösung x^* besitzt. Das ist ein wichtiger Vorzug gegenüber dem CG mit dem "schwächeren" Funktional $Q(x)$ im Fall, dass die Bedingung $A = A^T > 0$ nicht gegeben ist.

Als Such- und Abstiegsrichtung verwendet CR im Gegensatz zu Orthomin(0), das den negativen Gradienten $r(x) = b - Ax$ vom Funktional $Q(x)$ benutzt, einen Vektor, der in Analogie zum CG konstruiert wird. So ist der erste Suchvektor $p = r$, weitere Suchrichtungen entstehen rekursiv gemäß $p = r + \beta p_{alt}$.

Bei der Berechnung der Schrittzahlen α und β kann der Ausdruck $r^T Ar$ Null sein, so dass in β der Nenner Null wird, d. h. in α würde der Zähler verschwinden. Damit wird ein vorzeitiges Ende der Iteration mit $x^{(m+1)} = x^{(m)}$ signalisiert, ohne Konvergenz zu haben. Geometrisch bedeutet dies, wenn p – gebildet als Richtung des "optimalsten" Abstiegs aus $r \perp Q(x) = \text{const}$ und p_{alt} – eine Tangente zu $R(x) = \text{const}$ ist, dann bleibt die Iteration am Ort stehen.

Neben der Wahl anderer geeigneter Abbruchbedingungen ist also ein zusätzlicher Test auf die Durchführbarkeit der Berechnung zu machen.

Tritt der Sonderfall des Abbruchs nicht auf, endet das CR normal sowie bei exakter Rechnung und $A = A^T$ spätestens nach n Schritten an der Lösung x^* . Die Residua r sind dabei A-orthogonal.

Für nicht symmetrische Matrizen besteht natürlich auch die Gefahr eines vorzeitigen Abbruchs des CR. Mehr jedoch interessieren uns wie beim CG die Fragen, ob Endlichkeit, langsame Konvergenz in unendlichen vielen Schritten oder Divergenz des CR vorkommen.

Die Residua $r^{(m)}$ erfüllen eine A-Orthogonalitätsbedingung, wobei die A-Orthogonalität eigentlich nur Sinn für symmetrische Matrizen macht. Für die Suchrichtungen $p^{(m)}$ gilt dann bei aufeinander folgenden Richtungen die Beziehung $(Ap^{(m+1)})^T Ap^{(m)} = 0$. Andere Eigenschaften findet man noch in den Formeln (4.18).

5.4.1 CR mit symmetrischer indefiniter Matrix

Im symmetrischen Fall ist im Allgemeinen ein endlicher konvergenter Iterationsverlauf des CR zu erwarten. Die Suchrichtungen p sind dabei Abstiegsrichtungen.

Die Besonderheit besteht in der Gefahr eines vorzeitigen Abbruchs, falls $r^T A r = 0$ ist.

Beispiel 5.18

Sei $Ax = b$, A diagonal (damit symmetrisch) und indefinit aus den Beispielen 5.1, 5.12.

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

Das zu minimierende Funktional ist $R(x) = \frac{1}{2}x^T A^T A x - x^T A^T b = \frac{1}{2}(x_1^2 + x_2^2)$, $R(x^*) = 0$, und die Suchrichtungen werden als Linearkombination aus der letzten und dem Residuum abgeleitet.

(1) Für den Startvektor $x^{(0)} = (1, 0)^T$ erhält man

$$p^{(0)} = r^{(0)} = b - Ax^{(0)} = (-1, 0)^T, \quad w^{(0)} = t^{(0)} = Ar^{(0)} = (-1, 0)^T, \quad \alpha_0 = \frac{r^{(0)T} t^{(0)}}{w^{(0)T} w^{(0)}} = 1$$

und mit $x^{(1)} = x^{(0)} + \alpha_0 p^{(0)} = (0, 0)^T = x^*$, $r^{(1)} = 0$, einen endlichen Iterationsverlauf.

(2) Für den Startvektor $x^{(0)} = (1, 1)^T$ haben wir eine Abbruchsituation.

Der Suchvektor $p^{(0)} = r^{(0)} = (-1, 1)^T$ ist orthogonal zur Höhenlinie $Q(x) = 0$ im Punkt $(1, 1)$ und liegt gleichzeitig auf der Tangente zur Höhenlinie $R(x) = 1$.

Das AV kommt in der Suchrichtung nicht vorwärts. Es ist

$$w^{(0)} = t^{(0)} = Ar^{(0)} = (-1, -1)^T, \quad \alpha_0 = 0 \quad \text{und} \quad x^{(1)} = x^{(0)} = (1, 1)^T.$$

(3) Mit dem Startvektor $x^{(0)} = (1, \frac{1}{2})^T$ lässt sich ein endlicher konvergenter Iterationsprozess beschreiben.

Die beiden Iterierten, die zugehörigen gegen den Nullvektor tendierenden Residua sowie die Suchrichtungen sind

$$m = 2$$

$$xv = \begin{bmatrix} 1 & \frac{2}{5} & 0 \\ \frac{1}{2} & \frac{4}{5} & 0 \end{bmatrix}$$

$$rv = \begin{bmatrix} -1 & \frac{-2}{5} & 0 \\ \frac{1}{2} & \frac{4}{5} & 0 \end{bmatrix}$$

$$pv = \begin{bmatrix} -1 & \frac{6}{25} & 0 \\ \frac{1}{2} & \frac{12}{25} & 0 \end{bmatrix}$$

Dazu erhalten wir unter Berücksichtigung von (4.18) sowie $p^{(0)} = r^{(0)}$, $A^T A = I$ und $r^{(2)} = p^{(2)} = 0$ die Beziehungen

$$\begin{aligned} r^{(k)T} A r^{(j)} &= 0 \quad \forall k \neq j, \quad (A p^{(1)})^T A p^{(0)} = p^{(1)T} A^T A p^{(0)} = p^{(1)T} p^{(0)} = 0, \\ r^{(1)T} A p^{(0)} &= 0, \quad r^{(1)T} A r^{(1)} = r^{(1)T} A p^{(1)} = r^{(0)T} A p^{(1)} = -\frac{12}{25}. \end{aligned}$$

Damit sind die ersten beiden Suchrichtungen selber orthogonal.

Der Iterationsverlauf für $x^{(0)} = (1, \frac{1}{2})^T$ führt mit zwei Schritten $x^{(1)}$ und $x^{(2)} = x^*$ zum Minimum von $R(x)$ und damit zur Lösung des LGS.

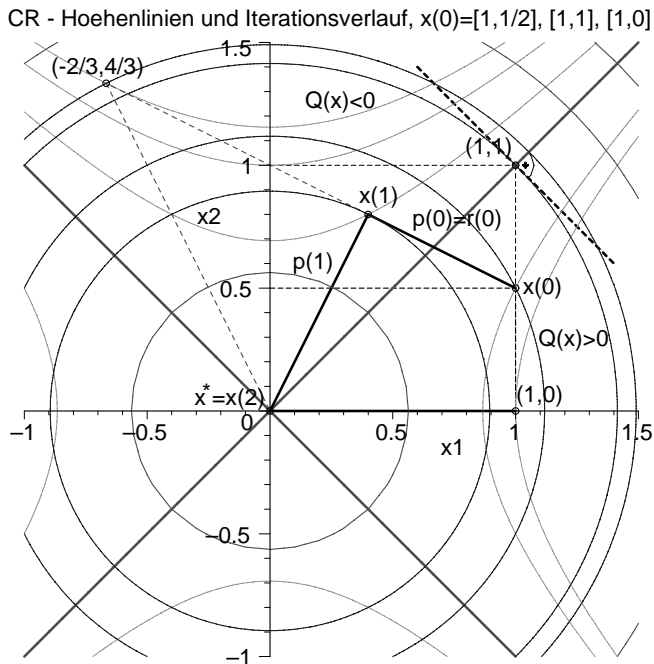


Abb. 5.37 Datei *cr_005.ps*

Höhenlinienbild mit Iterationsverlauf $x^{(m)}$, $m = 0, 1, 2$, von CR für 3 verschiedene Startvektoren zu $R(x) = \frac{1}{2}(x_1^2 + x_2^2)$ (Höhenlinien sind konzentrische Kreise), $\text{contours}=[2, 1.5, 1.111, 1, 0.625, 0.4, 0.16]$, Abstiegsrichtung $p(x)$, dazu Höhenlinien von $Q(x) = \frac{1}{2}(x_1^2 - x_2^2)$, $\text{contours}=[-0.667, 0.5, -0.5, 0.375, -0.24]$ und Vergleich mit Verlauf des CG gemäß $x^{(0)}, (-\frac{2}{3}, \frac{4}{3})^T, x^{(2)} = x^*$

Rechnung mit Maple bei `Digits:=16`

CR fuer $Ax=b$, A diagonal und indefinit - konv. endlicher Iterationsverlauf

m	[x(m) [1], x(m) [2]]	[r(m) [1], r(m) [2]]
0	[1.0000000000, 0.5000000000]	[-1.000e+00, 5.000e-01]
1	[0.4000000000, 0.8000000000]	[-4.000e-01, 8.000e-01]
2	[0.0000000000, -0.0000000000]	[-3.000e-16, -3.000e-16]

Beispiel 5.19

Sei $Ax = b$, $A = A^T$ und indefinit aus den Beispielen 1.4, 5.2, 5.7, 5.13.

$$\begin{pmatrix} 1 & 0 \\ 0 & -4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad x^* = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Wir rechnen mit verschiedenen Startvektoren die ersten Schritte des CR.

$$\begin{aligned} (1) \quad x^{(0)} &= (1, 2)^T && \text{Startvektor,} \\ r^{(0)} &= b - Ax^{(0)} = (-1, 8)^T && \text{Anfangsresiduum,} \\ p^{(0)} &= r^{(0)} = (-1, 8)^T && \text{1. Suchrichtung,} \\ t^{(0)} &= Ar^{(0)} = -(1, 32)^T, \quad w^{(0)} = Ap^{(0)} = -(1, 32)^T. \end{aligned}$$

S1

$$\begin{aligned} \alpha_0 &= \frac{r^{(0)T}t^{(0)}}{w^{(0)T}w^{(0)}} = -\frac{51}{205} = -0.248\,780\dots, \\ x^{(1)} &= x^{(0)} + \alpha_0 p^{(0)} = \frac{2}{205}(128, 1)^T, \\ r^{(1)} &= b - Ax^{(1)} = r^{(0)} - \alpha_0 Aw^{(0)} = \frac{8}{205}(-32, 1)^T, \\ \beta_0 &= \frac{r^{(1)T}Ar^{(1)}}{r^{(0)T}Ar^{(0)}} = -\frac{256}{42025} = -0.006\,091\dots, \\ p^{(1)} &= r^{(1)} + \beta_0 p^{(0)} = -\frac{408}{42025}(128, 1)^T. \end{aligned}$$

S2

$$\begin{aligned} \alpha_1 &= \frac{r^{(1)T}t^{(1)}}{w^{(1)T}w^{(1)}} = \frac{205}{204} = 1.004\,901\dots, \\ x^{(2)} &= (0, 0)^T, \\ r^{(2)} &= (0, 0)^T, \quad \beta_1 = 0, \quad p^{(2)} = (0, 0)^T \rightarrow \text{Stopp mit } x^{(2)} = x^*. \end{aligned}$$

CR konvergiert in zwei Schritten mit Abstiegsrichtungen gegen x^* .

Rechnung mit Maple bei `Digits:=16`

CR fuer $Ax=b$, A diagonal und indefinit - konv. endlicher Iterationsverlauf

m	[x(m) [1], x(m) [2]]	[r(m) [1], r(m) [2]]
0	[1.0000000000, 2.0000000000]	[-1.000e+00, 8.000e+00]
1	[1.2487804878, 0.0097560976]	[-1.249e+00, 3.902e-02]
2	[0.0000000000, 0.0000000000]	[0.000e-01, 0.000e-01]

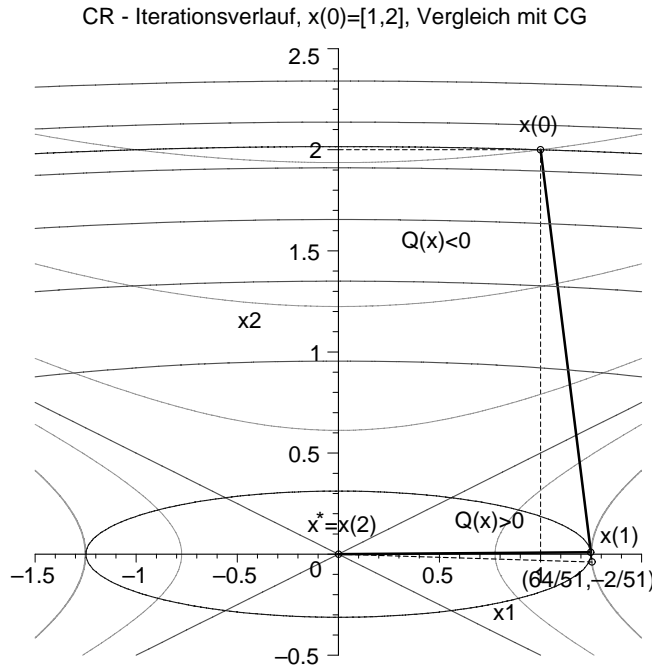


Abb. 5.38 Datei *cr_006.ps*
 Höhenlinienbild bei endlichem
 Iterationsverlauf des CR mit
 $x^{(0)} = (1, 2)^T$
 zu $R(x) = \frac{1}{2}x_1^2 + 8x_2^2$, $R(x^*) = 0$,
 mit `contours=[32.5,0.7805,0]`
 und `contours=6`,
 dazu die Höhenlinien von $Q(x)$
 an den Iterierten $x^{(m)}$,
 Vergleich mit CG zu $Q(x)$ gemäß
 $x^{(0)}, (\frac{64}{51}, -\frac{2}{51})^T, x^{(2)} = x^*$
 (dick gestrichelter Verlauf), mit
`contours=[-7.5,-3,-0.75,0,`
`0.3,0.7795,0.7843]`

$$(2) \quad x^{(0)} = (1, 0)^T, \quad r^{(0)} = (-1, 0)^T = p^{(0)}.$$

$$\underline{\mathbf{S1}} \quad \alpha_0 = 1,$$

$$x^{(1)} = (0, 0)^T,$$

$$r^{(1)} = (0, 0)^T, \beta_0 = 0, p^{(1)} = (0, 0)^T \rightarrow \text{Stopp mit } x^{(1)} = x^*.$$

CR braucht nur einen Schritt.

$$(3) \quad x^{(0)} = (0, 1)^T, \quad r^{(0)} = (0, 4)^T = p^{(0)}.$$

$$\underline{\mathbf{S1}} \quad \alpha_0 = -\frac{1}{4},$$

$$x^{(1)} = (0, 0)^T, \quad r^{(1)} = (0, 0)^T \rightarrow \text{Stopp wie in (2)}.$$

$$(4) \quad x^{(0)} = (3, 1)^T, \quad r^{(0)} = (-3, 4)^T = p^{(0)}.$$

$$\underline{\mathbf{S1}} \quad \alpha_0 = -\frac{11}{53} = -0.207547\dots,$$

$$x^{(1)} = \frac{3}{53}(64, 3)^T, \quad r^{(1)} = \frac{12}{53}(-16, 3)^T,$$

$$\beta_0 = -\frac{576}{2809} = -0.205055\dots,$$

$$p^{(1)} = -\frac{132}{2809}(64, 3)^T.$$

$$\begin{aligned} \underline{\text{S2}} \quad \alpha_1 &= \frac{53}{44} = 1.204545\dots, \\ x^{(2)} &= (0,0)^T, \\ r^{(2)} &= (0,0)^T, \quad \beta_1 = 0, \quad p^{(2)} = (0,0)^T \rightarrow \text{Stopp mit } x^{(2)} = x^*. \end{aligned}$$

CR konvergiert in zwei Schritten mit Abstiegsrichtungen gegen x^* .

Rechnung mit Maple bei `Digits:=16`

CR fuer $Ax=b$, A diagonal und indefinit - konv. endlicher Iterationsverlauf

m	[x(m)[1], x(m)[2]]	[r(m)[1], r(m)[2]]
0	[3.0000000000, 1.0000000000]	[-3.000e+00, 4.000e+00]
1	[3.6226415094, 0.1698113208]	[-3.623e+00, 6.792e-01]
2	[0.0000000000, 0.0000000000]	[0.000e-01, 0.000e-01]

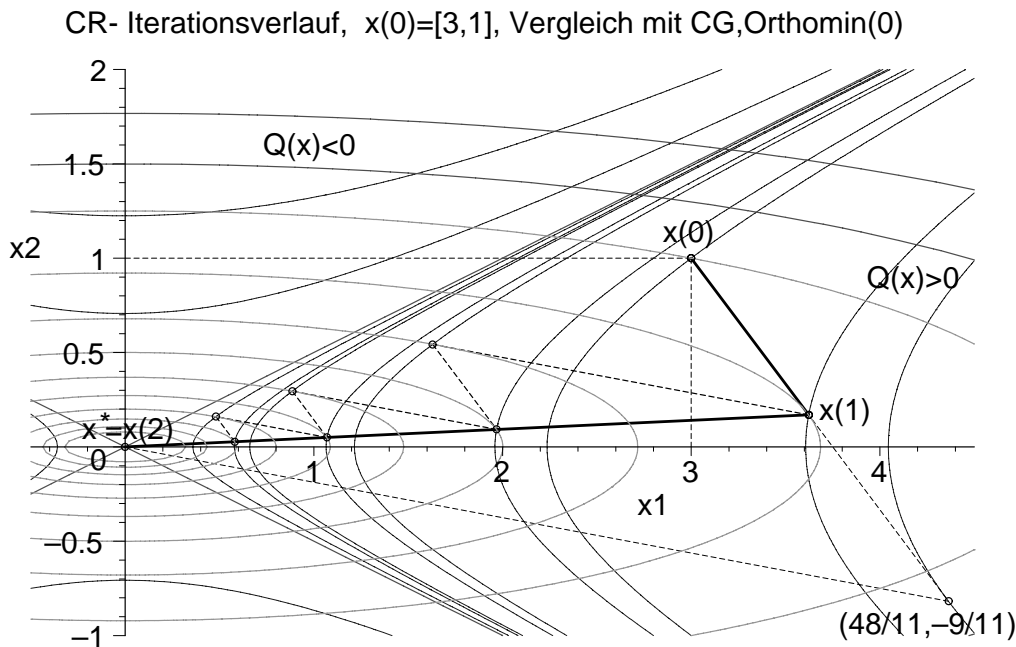


Abb. 5.39 Datei `cr_007.ps`

Höhenlinienbild bei konvergentem endlichem Iterationsverlauf des CR mit $x^{(0)} = (3,1)^T$ zu $R(x) = \frac{1}{2}x_1^2 + 8x_2^2$, $R(x^*) = 0$, `contours`=[12.5,6.792,3.691,2.005,1.0899,0.5922,0.3218,0.1749,0.0950,0.0516, 25,18], dazu auch die Höhenlinien von $Q(x)$ an den Iterierten $x^{(m)}$ mit `contours`=[2.5,8.18,6.504,0.738,1.9205,0.2179,0.567,0.0644,0.1674,-1,-3], Vergleich mit CG zu $Q(x)$ gemäß $x^{(0)}$, $(\frac{48}{11}, -\frac{9}{11})^T$, $x^{(2)} = x^*$ (dick gestrichelter Verlauf), sowie mit Orthomin(0) zu $R(x)$ bei konvergentem "Zick-Zack"-Iterationsverlauf

(5) $x^{(0)} = (2, 1)^T, r^{(0)} = 2(-1, 2)^T = p^{(0)}.$

S1

$$\alpha_0 = -\frac{3}{13} = -0.230769\dots,$$

$$x^{(1)} = \frac{1}{13}(32, 1)^T, r^{(1)} = \frac{4}{13}(-8, 1)^T,$$

$$\beta_0 = -\frac{16}{169} = -0.094674\dots,$$

$$p^{(1)} = -\frac{12}{169}(32, 1)^T.$$

S2

$$\alpha_1 = \frac{13}{12} = 1.083333\dots,$$

$$x^{(2)} = (0, 0)^T,$$

$$r^{(2)} = (0, 0)^T, \beta_1 = 0, p^{(2)} = (0, 0)^T \rightarrow \text{Stopp mit } x^{(2)} = x^*.$$

CR konvergiert in zwei Schritten mit Abstiegsrichtungen gegen x^* .

Berechnungen in Maple mit `Digits:=16`

CR fuer $Ax=b$, A diagonal und indefinit - konv. endlicher Iterationsverlauf

m	[x(m) [1], x(m) [2]]	[r(m) [1], r(m) [2]]
0	[2.0000000000, 1.0000000000]	[-2.000e+00, 4.000e+00]
1	[2.4615384615, 0.0769230769]	[-2.462e+00, 3.077e-01]
2	[0.0000000000, 0.0000000000]	[0.000e-01, 0.000e-01]

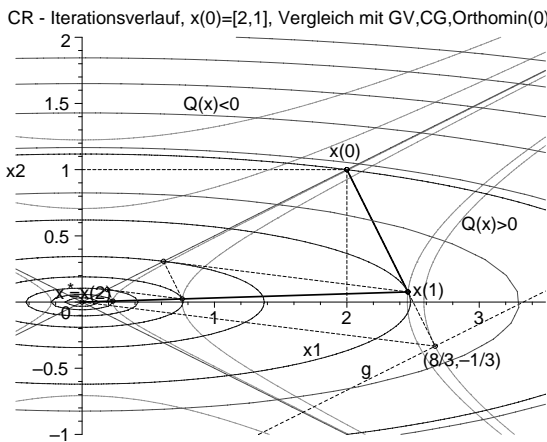


Abb. 5.40 Datei *cr_008.ps*
 Höhenlinienbild bei endlichem Iterationsverlauf des CR mit $x^{(0)} = (2, 1)^T$ zu $R(x), R(x^*) = 0$,
`contours=[10, 3.0769, 0.9467, 0.2913, 0.0896, 0.02758, 0.0085]`,
 dazu auch die Höhenlinien von $Q(x)$ an den Iterierten $x^{(m)}$ mit `contours=[3.333, 3.0177, 0.2857, 0.027, 0, -1, -3]`,
 Vergleich mit GV (Abbruch nach 1. Schritt), mit CG zu $Q(x)$ gemäß $x^{(0)}, (\frac{8}{3}, -\frac{1}{3})^T, x^*$ (dick gestrichelter Verlauf), sowie mit Orthomin(0) (konv. "Zick-Zack"-Iteration)

$$\begin{aligned}
 (6) \quad x^{(0)} &= (-8, 1)^T && \text{Startvektor,} \\
 r^{(0)} &= b - Ax^{(0)} = 4(2, 1)^T && \text{Anfangsresiduum,} \\
 p^{(0)} &= r^{(0)} = 4(2, 1)^T && \text{1. Suchrichtung,} \\
 t^{(0)} &= Ar^{(0)} = 8(1, -2)^T, \quad w^{(0)} = Ap^{(0)} = 8(1, -2)^T.
 \end{aligned}$$

$$\begin{aligned}
 \underline{\mathbf{S1}} \quad \alpha_0 &= \frac{r^{(0)T} t^{(0)}}{w^{(0)T} w^{(0)}} = 0, \\
 x^{(1)} &= x^{(0)} + \alpha_0 p^{(0)} = x^{(0)}, \quad r^{(1)} = b - Ax^{(1)} = r^{(0)}, \\
 \beta_0 &= \frac{r^{(1)T} Ar^{(1)}}{r^{(0)T} Ar^{(0)}} = \frac{\dots}{0}, \\
 &\rightarrow \text{Abbruch mit } x^{(0)} \text{ wegen } r^{(0)T} Ar^{(0)} = 0.
 \end{aligned}$$

In Richtung des Vektors $(2, 1)^T$ tritt bei der Berechnung der Schrittzahlen α und β das Problem der Null im Zähler bzw. Nenner auf. Das führt zum Abbruch von CR. Die erste Suchrichtung im Punkt $x^{(0)}$ ist senkrecht zur Höhenlinie $Q(x) = 30$ und gleichzeitig tangential zur Höhenlinie $R(x) = 40$. Entlang der Geraden $g(\alpha) = x^{(0)} + \alpha p^{(0)}$ lässt sich kein kleinerer Wert des Funktionals $R(x)$ finden.

Schritt k = 0

```

Startvektor           x = [-8.0000000000e+00  1.0000000000e+00]
Residuum/Suchr. r = b-Ax = [ 8.0000000000e+00  4.0000000000e+00]
Funktionswert        Q(x) =  3.0000000000000000e+01
Funktionswert        R(x) =  4.0000000000000000e+01
Anfangsfehlerquadrat r'r =  8.0000000000000000e+01
Abbruch wegen Nenner r'Ar=0

```

$$x^{(0)} = (-8, 1.1)^T \text{ oder } x^{(0)} = (-8, 0.9)^T.$$

Für beide Startvektoren ist die Rechnung durchführbar und endet nach zwei Schritten.

CR fuer Ax=b, A diagonal, indefinit - 2 konv. endliche Iterationsverlaeufe

m	[x(m) [1], x(m) [2]]	[x(m) [1], x(m) [2]]
0	[-8.000000, 1.100000] R(x(0)) = 41.68 Q(x(0)) = 29.58	[-8.000000, 0.900000] R(x(0)) = 38.48 Q(x(0)) = 30.38
1	[-8.287671, 0.941781] R(x(1)) = 41.44 Q(x(1)) = 32.57	[-7.641509, 1.061321] R(x(1)) = 38.21 Q(x(1)) = 26.94
2	[0, 0]	[0, 0]

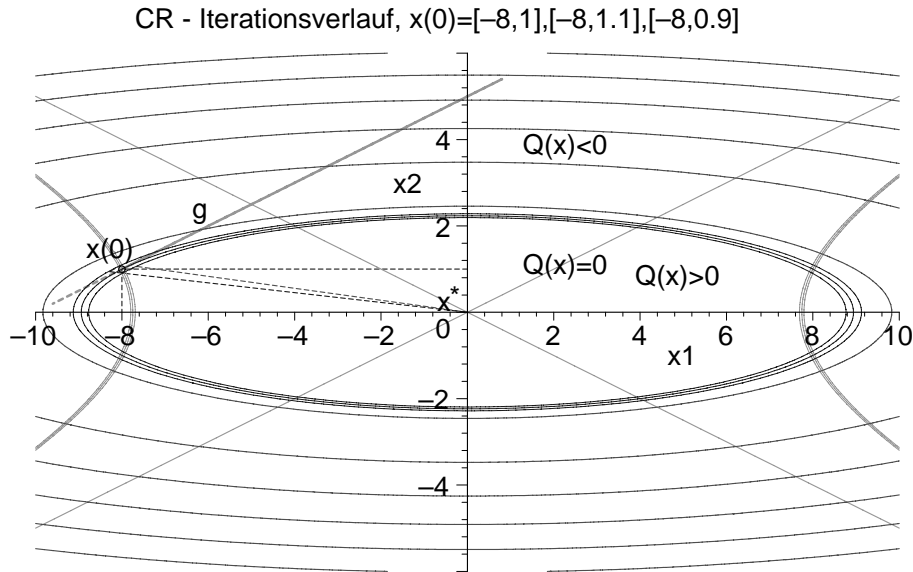


Abb. 5.41 Datei *cr_05.ps*

Höhenlinienbild und Iterationsverlauf von CR zu $R(x)$, $R(x^*) = 0$, mit $x^{(0)} = (-8, 1)^T$ (Abbruch), $x^{(0)} = (-8, 1.1)^T$, $(-8, 0.9)^T$ (Konvergenz), mit $\text{contours}=[40, 41.68, 38.48]$ und $\text{contours}=6$, dazu Höhenlinien von $Q(x)$ bei $x^{(0)}$ mit $\text{contours}=[30, 29.58, 30.38]$ und $Q(x) = 0$

Wir wollen den konvergenten endlichen Iterationsverlauf des CR für die Startvektoren $x^{(0)} = (-8, 1.1)^T$ bzw. $x^{(0)} = (-8, 0.9)^T$ noch etwas besser auflösen.

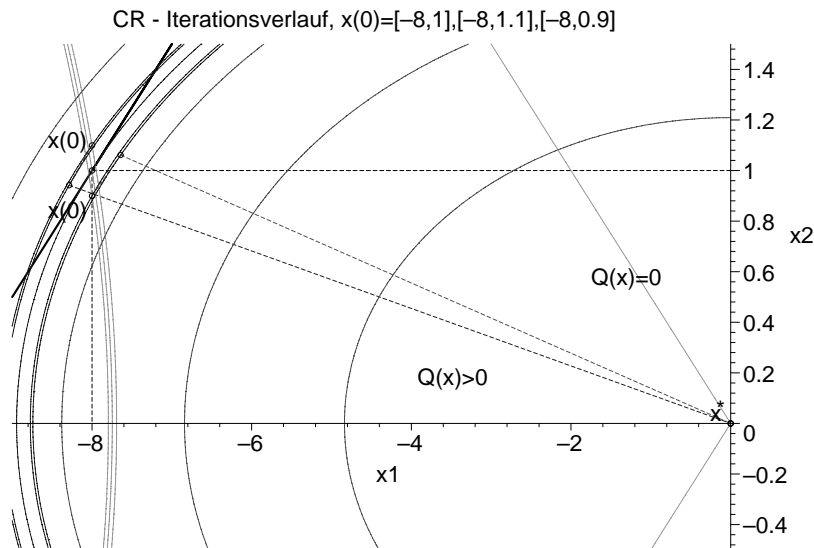


Abb. 5.42 Datei *cr_06.ps*

Höhenlinienbild und Iterationsverlauf von CR zu $R(x)$, $R(x^*) = 0$, mit $x^{(0)} = (-8, 1)^T$ (Abbruch), $x^{(0)} = (-8, 1.1)^T$, $(-8, 0.9)^T$ (Konvergenz), mit $\text{contours}=[40, 41.68, 41.43, 38.48, 38.20]$ und $\text{contours}=4$, dazu Höhenlinien von $Q(x)$ bei $x^{(0)}$ mit $\text{contours}=[30, 29.58, 30.38]$ und $Q(x) = 0$

Beispiel 5.20

Sei $Ax = b$, $A = A^T$ und indefinit aus den Beispielen 1.5, 5.3, 5.8, 5.14.

$$\begin{pmatrix} 2 & 1 \\ 1 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \quad x^* = \begin{pmatrix} 10/7 \\ 1/7 \end{pmatrix}, \quad x^{(0)} = \begin{pmatrix} 3/2 \\ 1 \end{pmatrix}.$$

Die Funktionale sind $R(x_1, x_2) = \frac{5}{2}x_1^2 - x_1x_2 + 5x_2^2 - 7x_1$, $R(x^*) = -5$, und $Q(x_1, x_2) = x_1^2 + x_1x_2 - \frac{3}{2}x_2^2 - 3x_1 - x_2$, $Q(x^*) = -\frac{31}{14} = -2.214285\dots$

x^* ist Minimumstelle von $R(x)$ und Sattelpunkt von $Q(x)$.

Die Richtungen (Anstieg/Abstieg bez. $Q(x)$) $r(x) = b - Ax \perp Q(x) = \text{const}$, $\nabla R(x) = \text{const}$ sind A-orthogonal. Die Richtung (Abstieg) $\hat{r}(x) = A^T r(x) = A^T(b - Ax) \perp R(x) = \text{const}$. Die Such- und Abstiegsrichtungen sind $p^{(0)} = r^{(0)} \perp Q(x) = \text{const}$, $\nabla R(x) = \text{const}$, und weiter $p(x)$, gebildet aus $r(x) \perp Q(x) = \text{const}$ und p_{alt} , die $\nabla Q(x) = \text{const}$, $\nabla R(x) = \text{const}$ sind.

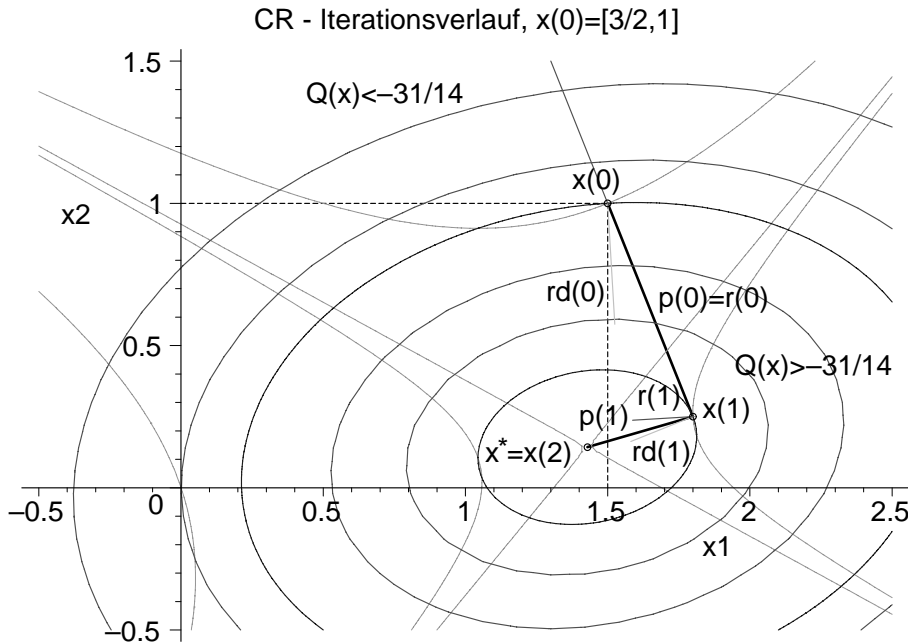


Abb. 5.43 Datei *cr_002.ps*

Höhenlinienbild bei endlichem Iterationsverlauf des CR zu $R(x)$ mit $x^{(0)} = (\frac{3}{2}, 1)^T$, $\text{contours} = [-4.6375, -4, -3, -1.375, 0, 3]$ und in den Iterierten $x^{(m)}$ eingezeichneten Richtungen $r(x) \perp Q(x)$ (dünne Striche), $\hat{r}(x) \perp R(x)$ (ganz dünne Striche), Höhenlinien zu $Q(x)$ mit $\text{contours} = [-3.25, -2.054, -2.214, 0]$

$x^{(0)} = (\frac{3}{2}, 1)^T$ ist ein günstiger Startvektor mit endlichem Iterationsverlauf. Die Iterierten, die zugehörigen Residua und Suchrichtungen sind

$$xv = \begin{bmatrix} \frac{3}{2} & \frac{9}{5} & \frac{10}{7} \\ 1 & \frac{1}{4} & \frac{1}{7} \end{bmatrix}, \quad rv = \begin{bmatrix} -1 & \frac{-17}{20} & 0 \\ \frac{5}{2} & \frac{-1}{20} & 0 \end{bmatrix}, \quad pv = \begin{bmatrix} -1 & \frac{-39}{50} & 0 \\ \frac{5}{2} & \frac{-9}{40} & 0 \end{bmatrix}.$$

Andere Startvektoren werden nicht getestet. Aber Situationen mit Abbruch, wie sie in Beispiel 5.19 aufgetreten sind, können auch hier vorkommen.

5.4.2 CR mit nicht symmetrischer Matrix

Auch hier betrachtet man die Funktionale $R(x)$, $Q(x)$ und $S(x)$.

Als erste Besonderheit ist wie zum CR für $A = A^T$ und indefinit (Kap. 5.4.1) bei der Berechnung der Schrittzahlen α und β gemäß (4.17) die Durchführbarkeit zu beachten. Das bedeutet für die Größen der Test

$$r^T A p = r^T A r = 0 \quad \text{bzw.} \quad (A p)^T r = 0.$$

Weiterhin gibt es für nicht symmetrische Matrizen Probleme bei der Konstruktion der A-orthogonalen Residua r , die im Normalfall einen Vektorraum mit stetig wachsender Dimension gebildet haben und die Endlichkeit des CR bewirkten sowie letztendlich eine Vereinfachung der hergeleiteten Schrittzahlen in der Form

$$\alpha = \frac{\|r\|_A^2}{\|w\|_2^2}, \quad w = A p, \quad \beta = \frac{\|r_{neu}\|_A^2}{\|r\|_A^2} \quad (5.18)$$

erlaubt. Man rechnet im CR, auch bei $A \neq A^T$, einfach mit den Skalarprodukten und nicht mit den Normen. Mit den originalen Formeln

$$\alpha = \frac{w^T r}{w^T w}, \quad \beta = \frac{w_{neu}^T r}{w^T r}, \quad w = A p, \quad (5.19)$$

zu arbeiten, ohne die weitergehenden Beziehungen zwischen den Größen zu verwenden, ist nicht möglich. Denn das bedeutet, dass z. B. β_m die Größe $w^{(m+1)} = A p^{(m+1)}$ und damit $p^{(m+1)}$ braucht, aber $p^{(m+1)}$ eine rekursive Darstellung hat, in deren rechte Seite wiederum das β_m auftaucht. Also gehen wir davon aus, dass die dritte Bedingung in (4.18) $r^T A w = r^T A r$ zutrifft.

Was bleibt aber dann von der A-Orthogonalität der Residua übrig? Nach den Betrachtungen im Kapitel 4.3 sollten im Fall von $A \neq A^T$ und (5.19) zumindest die A-Orthogonalität der Residua und die Bedingungen (4.18) in abgeschwächten Formen erfüllt sein, also beispielsweise

$$r^{(j)T} A r^{(m)} = 0, \quad j = 0, 1, \dots, m-1, \quad \text{bzw.} \quad r^{(m-1)T} A r^{(m)} = 0. \quad (5.20)$$

Erhält man dann die Endlichkeit des Verfahrens?

Diese Fragestellungen und Vergleiche zum CG und Orthomin(0) in den Varianten $A > 0$ bzw. A indefinit werden an den schon mehrfach genannten Beispielen erläutert und illustriert.

Dabei wird der Iterationsverlauf verfolgt. Neben der Suchrichtung $p(x)$ – erste Suchrichtung $p(x) = r(x)$ ist eine Abstiegsrichtung zu $R(x)$, $p(x) = r(x) + \beta p_{alt}(x)$ sind weitere Such-/Abstiegsrichtungen – werden auch andere Richtungsvektoren angezeigt. So sind die Vektoren

- $r(x) = b - Ax \not\perp$ zu $Q(x) = \text{const}$ und $\not\perp$ zu $R(x) = \text{const}$,
- $\tilde{r}(x) = b - \frac{1}{2} \tilde{A} x \perp$ zu Höhenlinien von $Q(x)$, Richtung des steilsten Ab-/Anstiegs,
- $\hat{r}(x) = A^T(b - Ax)$ als Richtung des steilsten Abstiegs zu $R(x)$

von Interesse.

Beispiel 5.21 Sei $Ax = b$, $A \neq A^T$, $A > 0$ aus den Beispielen 1.6, 5.4, 5.9, 5.15.

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

Wir erinnern uns wiederum an die Funktionale $R(x)$, $Q(x)$ und $S(x)$ sowie an das Höhenlinienbild mit den verschiedenen Minimumstellen.

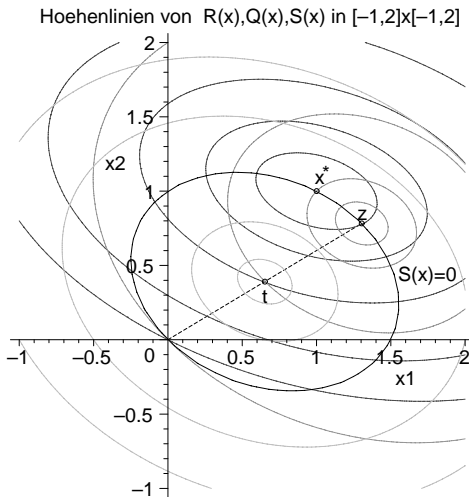


Abb. 5.44 Datei *cr_0030.ps*

Höhenlinienbilder (Ellipsenscharen) zu den Funktionalen

$R(x)$, Minimum bei $x^* = A^{-1}b$, und
 $Q(x)$, Minimum bei $z = (\frac{1}{2}\tilde{A})^{-1}b \neq x^*$,
 $S(x)$, Minimum bei $t = \tilde{A}^{-1}b = \frac{z}{2}$

x^* ist Minimumstelle von $R(x)$, $R(x^*) = -9$, und Stelle mit $S(x^*) = 0$.

Die Suchrichtung $p(x)$ ist eine "schräge" Abstiegsrichtung bez. $R(x)$, es gilt $p(x) \not\perp Q(x) = \text{const}$ und $\not\perp R(x) = \text{const}$. Weitere Richtungen sind $\hat{r}(x) = A^T r(x) = A^T(b - Ax) \perp R(x) = \text{const}$ sowie $\tilde{r}(x) = b - \frac{1}{2}\tilde{A}x = b - \frac{1}{2}(A + A^T)x \perp Q(x) = \text{const}$.

Für den Startvektor $x^{(0)} = (1, -1)^T$ sind die Richtungsvektoren

$$\begin{aligned} p^{(0)} &= r^{(0)} = b - Ax^{(0)} = 2(1, 3)^T, \\ \hat{r}^{(0)} &= A^T(b - Ax^{(0)}) = A^T r^{(0)} = 4(1, 5)^T, \\ \tilde{r}^{(0)} &= b - \frac{1}{2}\tilde{A}x^{(0)} = \frac{1}{2}(3, 11)^T. \end{aligned}$$

In der Abbildung 5.45 sind neben der Suchrichtung $p^{(0)} = p(0)$ auch die beiden anderen Richtungen eingetragen. $\hat{r}^{(0)} = \mathbf{rd}(0)$ ist als dicker Strich gezeichnet und $\tilde{r}^{(0)} = \mathbf{rs}(0)$ liegt zwischen diesem und $r^{(0)}$.

Um die Konvergenz mit endlich vielen Schritten von CR zu erzwingen, kann man das LGS gemäß $Bx = A^T Ax = A^T b = c$ symmetrisieren. Das zugehörige Funktional ist $R_B(x) = \frac{1}{2}(Bx)^T Bx - (Bx)^T c$.

Die Iterierten sind $x^{(0)} = (1, -1)^T$, $x^{(1)} = (\frac{999}{725}, \frac{129}{145})^T$, $x^{(2)} = x^*$.

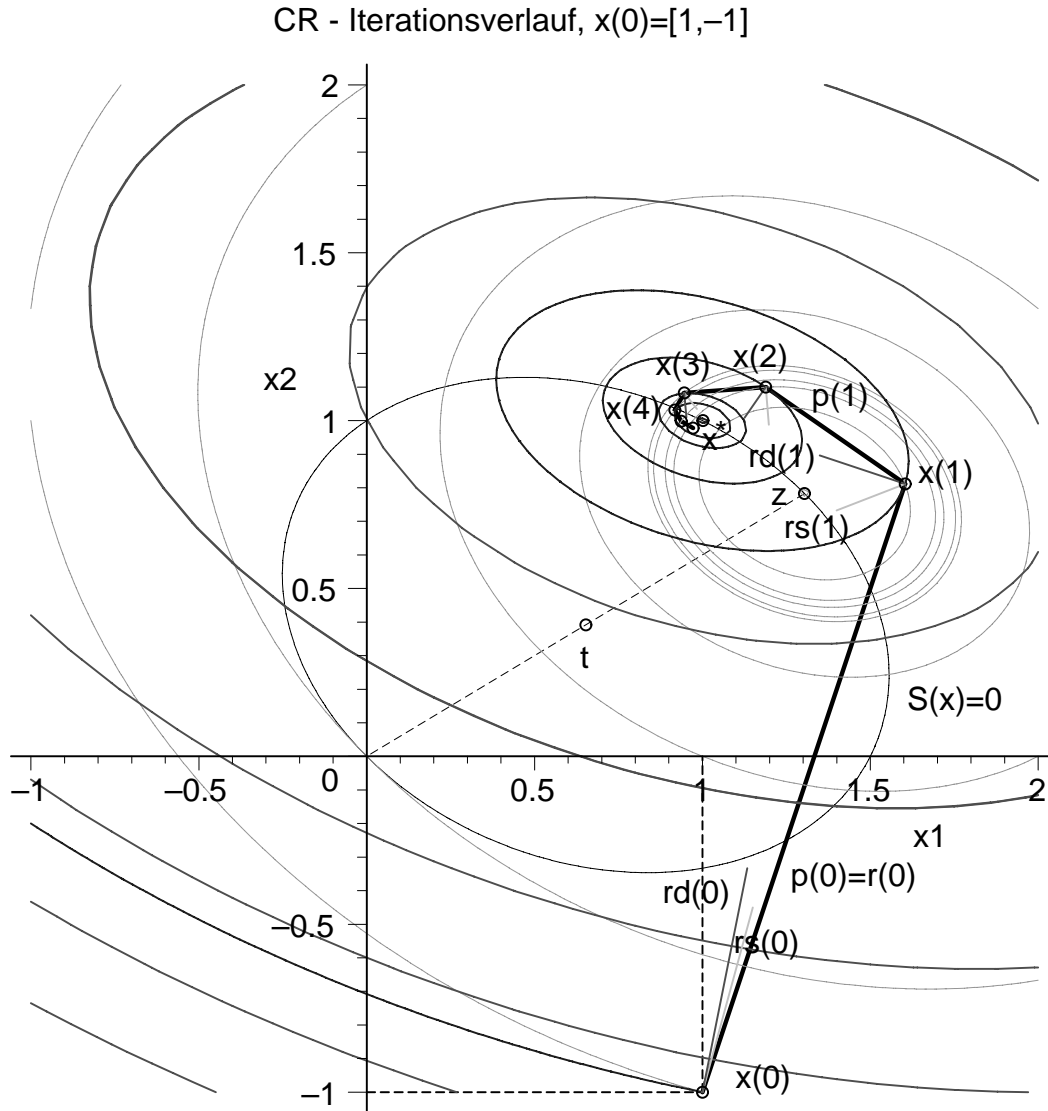


Abb. 5.45 Datei *cr_003.ps*

Höhenlinienbild und CR zu $R(x)$ mit den Iterierten $x^{(m)}$ (Iterationsverlauf fett gezeichnet), Suchrichtung $p(x)$, $R(x^*) = -9$, $R(z) = -\frac{4608}{529} = -8.710775$,

Höhenlinien zu $R(x)$ mit $\text{contours}=[11, -3, -7, -8.3207, -8.841, -8.969, -8.9867, -8.999]$ und $\text{contours}=5$.

Zusätzlich eingezeichnet:

- Vektor $\tilde{r}(x) \perp Q(x) = \text{const}$ in den ersten Iterierten (dünner Strich),
- Vektor $\hat{r}(x) \perp R(x) = \text{const}$ in den ersten Iterierten (dicker Strich),
- Höhenlinien von $Q(x)$, $Q(x^*) = -3$, $Q(z) = -\frac{72}{23} = -3.130434$, an den Iterierten, $\text{contours}=[2, 0, -2, -2.7, -3.13, -3.0352, -2.9846, -1.9211, -2.9364, -2.9651]$,
- Funktional $S(x)$ mit Minimumstelle $t = \frac{z}{2}$ und Höhenlinie $S(x) = 0$.

Der anfängliche Iterationsverlauf ist

$$x^{(0)} = (1, -1)^T, \quad r^{(0)} = 2(1, 3)^T = p^{(0)}.$$

$$\begin{aligned} \mathbf{S1} \quad \alpha_0 &= \frac{16}{53} = 0.301\,886\dots, \\ x^{(1)} &= \frac{1}{53}(85, 43)^T, \quad r^{(1)} = \frac{6}{53}(-9, 5)^T, \\ \beta_0 &= \frac{54}{2809} = 0.019\,223\dots, \\ p^{(1)} &= \frac{6}{2809}(-459, 319)^T. \end{aligned}$$

$$\begin{aligned} \mathbf{S2} \quad \alpha_1 &= \frac{5088}{12025} = 0.423\,118\dots, \\ x^{(2)} &= \frac{1}{12025}(14297, 13223)^T, \quad r^{(2)} = -\frac{6}{12025}(957, 599)^T, \\ \beta_0 &= \frac{50932788}{144600625} = 0.352\,230\dots, \\ p^{(2)} &= -\frac{6}{144600625}(19830513, 1418867)^T. \end{aligned}$$

Es ist zu vermuten, dass das CR in unendlich vielen Schritten mit Abstiegsrichtungen sehr langsam gegen x^* konvergiert.

Rechnung mit Maple bei `Digits:=16`

CR fuer $Ax=b$, $A <> A'$, $A > 0$ - unendlicher Iterationsverlauf

m	[x(m) [1], x(m) [2]]	[r(m) [1], r(m) [2]]
0	[1.0000000000, -1.0000000000]	[2.000e+00, 6.000e+00]
1	[1.6037735849, 0.8113207547]	[-1.019e+00, 5.660e-01]
2	[1.1889397089, 1.0996257796]	[-4.775e-01, -2.989e-01]
3	[0.9460911484, 1.0822500411]	[2.557e-02, -2.468e-01]
4	[0.9178424233, 1.0311555678]	[1.332e-01, -9.347e-02]
5	[0.9365512816, 0.9981077835]	[1.288e-01, 5.677e-03]
6	[0.9724291360, 0.9779929077]	[7.715e-02, 6.602e-02]
7	[1.0055052182, 0.9740116788]	[1.498e-02, 7.796e-02]
8	[1.0229506387, 0.9796942203]	[-2.560e-02, 6.092e-02]
9	[1.0288296762, 0.9872299135]	[-4.489e-02, 3.831e-02]
10	[1.0288143829, 0.9942597001]	[-5.189e-02, 1.722e-02]
20	[0.9826216317, 1.0040612486]	[3.070e-02, -1.218e-02]
30	[1.0040085494, 0.9908411451]	[1.142e-03, 2.748e-02]
40	[1.0132735398, 0.9974517616]	[-2.400e-02, 7.645e-03]

m	x(m) [1],	x(m) [2]	r(m) [1],	r(m) [2]	p(m) [1..2]
50	1.0067196129,	1.0050729770]	[-1.851e-02,	-1.522e-02]	
60	0.9951366715,	1.0075811293]	[2.146e-03,	-2.274e-02]	
70	0.9893002930,	1.0050022394]	[1.640e-02,	-1.501e-02]	
80	0.9885651480,	1.0013680650]	[2.150e-02,	-4.104e-03]	
90	0.9909058208,	0.9977078629]	[2.048e-02,	6.876e-03]	
100	0.9956567545,	0.9946558460]	[1.403e-02,	1.603e-02]	
200	1.0021868339,	1.0057745764]	[-1.015e-02,	-1.732e-02]	
300	0.9897520187,	1.0010786756]	[1.942e-02,	-3.236e-03]	
400	0.9999308268,	0.9938503695]	[6.288e-03,	1.845e-02]	
500	1.0091878583,	0.9954956515]	[-1.387e-02,	1.351e-02]	
600	1.0096815644,	0.9999449109]	[-1.931e-02,	1.653e-04]	
700	1.0051148974,	1.0042369158]	[-1.447e-02,	-1.271e-02]	
800	0.9979998170,	1.0063539394]	[-2.354e-03,	-1.906e-02]	
900	0.9926896981,	1.0056469053]	[8.974e-03,	-1.694e-02]	
1000	0.9903269700,	1.0036622729]	[1.568e-02,	-1.099e-02]	[-1.25, -2.31]
1001	0.9903149903,	1.0036401444]	[1.573e-02,	-1.092e-02]	
1500	1.0005161843,	0.9938713133]	[5.096e-03,	1.839e-02]	
2000	1.0099488711,	0.9972198752]	[-1.712e-02,	8.340e-03]	[1.12, 4.97]
2001	1.0099515496,	0.9972316732]	[-1.713e-02,	8.305e-03]	
3000	0.9974316980,	1.0063225757]	[-1.186e-03,	-1.897e-02]	[-16.34, 0.64]
4000	0.9899958672,	1.0018528944]	[1.816e-02,	-5.559e-03]	[0.45, -10.62]
5000	0.9949672511,	0.9958204193]	[1.425e-02,	1.254e-02]	[24.27, -13.31]
6000	1.0044624232,	0.9937400117]	[-2.665e-03,	1.878e-02]	[28.99, 2.89]
7000	1.0093042717,	0.9959495710]	[-1.456e-02,	1.215e-02]	[11.37, 15.15]
8000	1.0098273773,	0.9991209240]	[-1.878e-02,	2.637e-03]	[-6.53, 22.52]
9000	1.0076032590,	1.0023686682]	[-1.758e-02,	-7.106e-03]	[-30.33, 27.38]
10000	1.0031877016,	1.0050430976]	[-1.142e-02,	-1.513e-02]	[-54.54, 21.91]

Um den Konvergenzverlauf noch besser zu erfassen, betrachten wir das Verhalten von Fehlernormen und kontrollieren u. a., ob wie im Normalfall des CR (Kap. 4.3) der Fehler $\|r^{(m)}\|_2 = \|b - Ax^{(m)}\|_2$ eine stetig abnehmende Folge bildet. Das ist der Fall, während dasselbe auf den Fehler $\|x^{(m)} - x^*\|_2$ nicht zutrifft.

Normen von Fehlern zum Iterationsverlauf

m	$\ x(m) - x_s\ _2$	$\ r(m)\ _2$
0	2.0000000000	6.3245553203
500	0.0102325899	0.0193653622
1000	0.0103431017	0.0191491870
1500	0.0061503859	0.0190792993
2000	0.0103300112	0.0190413937
2500	0.0073503427	0.0190230653
3000	0.0068243049	0.0190047677

3500	0.0098369125	0.0189933714
4000	0.0101742759	0.0189872710
4500	0.0087362196	0.0189824485
5000	0.0065419765	0.0189774160
5500	0.0059796733	0.0189723428
6000	0.0076876963	0.0189680929
6500	0.0093298790	0.0189650416
7000	0.0101476819	0.0189629236
7500	0.0102663455	0.0189613360
8000	0.0098666164	0.0189599862
8500	0.0090610520	0.0189586998
9000	0.0079636760	0.0189573856
9500	0.0067968279	0.0189560167
10000	0.0059660938	0.0189546211

Die langsamen Veränderungen der Iterierten korrespondieren natürlich auch mit den entsprechenden Schrittzahlen α und β .

So ist im Schritt

$$x^{(10000)} = x^{(9999)} + \alpha_{9999}p^{(9999)}$$

der Parameter $\alpha_{9999} = 0.940\dots \cdot 10^{-7}$ und für

$$p^{(10000)} = r^{(10000)} + \beta_{9999}p^{(9999)}$$

der Wert $\beta_{9999} = 1.000\ 119\dots$, so dass die Iterierten $x^{(m)}$ nur geringfügig sich verändern und die Suchrichtung $p^{(m)}$ kaum noch von $r^{(m)}$ beeinflusst sondern durch die vorherige Richtung $p^{(m-1)}$ dominiert wird.

Hier noch ein Hinweis auf den Iterationsverlauf unter Verwendung der ‘‘Restart-Variante‘‘.

Wir machen nach jeweils 1000 Iterationen einen Neustart des CR mit dem jeweils letzten Iterationsvektor. Damit wird im Algorithmus zunächst die übliche fortlaufende rekursive Residuumberechnung $r^{(m+1)} = r^{(m)} - \alpha_m w^{(m)}$ beim Neustart durch die explizite Formel $r^{(m+1)} = b - Ax^{(m+1)}$ ersetzt. Das ist jedoch nicht ausschlaggebend, zumindest solange man nicht sehr große Genauigkeiten (z. B. abgeleitet von der Mantissenlänge) fordert. Viel wichtiger ist, dass der erste Suchvektor $p^{(0)}$ als $r^{(0)}$ definiert ist und damit seine rekursive Ermittlung $p^{(m+1)} = r^{(m+1)} + \beta_m p^{(m)}$ an der Stelle unterbrochen wird.

Das hat eine positive Auswirkung, denn im Ergebnis erhält man eine schneller konvergierende Iterationsfolge $x^{(m)}$. Wenn schon der Iterationsprozess nicht endlich ist, dann möchte man in einem solchen Fall eigentlich nicht viel schlechter als Orthomin(0) sein, was bedeutet, dass man gelegentlich als Suchrichtung eben den Vektor $r(x)$ nehmen sollte. In welchen Abständen das zu erfolgen hat, ist sicher schwierig festzustellen und wurde nicht weiter getestet.

“Restart-Variante“ von CR nach jeweils 1000 Iterationen

CR fuer $Ax=b$, $A \ll A'$, $A > 0$ - konvergenter unendlicher Iterationsverlauf

m	[x(m) [1], x(m) [2]]	[r(m) [1], r(m) [2]]
0	[1.0000000000, -1.0000000000]	[2.000e+00, 6.000e+00]
1000	[0.9903269700, 1.0036622729]	[1.568e-02, -1.099e-02]
1001	[0.9974471570, 0.9986744340]	[6.431e-03, 3.977e-03]
1500	[0.9999242146, 1.0000077520]	[1.438e-04, -2.326e-05]
2000	[1.0000720856, 0.9999998366]	[-1.440e-04, 4.902e-07]
2001	[0.9999999596, 1.0000000821]	[-1.260e-09, -2.463e-07]
3000	[1.0000000008, 0.9999999996]	[-1.202e-09, 1.097e-09]

Wir kontrollieren noch einige Eigenschaften der auftretenden Vektoren.

Bei exakter Rechnung haben wir die folgenden ersten Iterierten, die zugehörigen Suchrichtungen sowie die gegen den Nullvektor tendierenden Residua.

$$\begin{aligned}
 xv &= \begin{bmatrix} 1 & \frac{85}{53} & \frac{14297}{12025} & \frac{67098761437}{70922089853} & \frac{811824582678250508529191705}{884492329020045639594713401} \\ -1 & \frac{43}{53} & \frac{13223}{12025} & \frac{76755434659}{70922089853} & \frac{912049189734288504939824231}{884492329020045639594713401} \end{bmatrix} \\
 pv &= \begin{bmatrix} 2 & \frac{-2754}{2809} & \frac{-118983078}{144600625} & \frac{-719750607702127705674}{5029942829117005561609} & \frac{0.731562 \dots 10^{53}}{0.782326 \dots 10^{54}} \\ 6 & \frac{1914}{2809} & \frac{-8513202}{144600625} & \frac{-1301838511449496488222}{5029942829117005561609} & \frac{-0.129224 \dots 10^{54}}{0.782326 \dots 10^{54}} \end{bmatrix} \\
 rv &= \begin{bmatrix} 2 & \frac{-54}{53} & \frac{-5742}{12025} & \frac{1813312026}{70922089853} & \frac{117778631969347396785932562}{884492329020045639594713401} \\ 6 & \frac{30}{53} & \frac{-3594}{12025} & \frac{-17500034418}{70922089853} & \frac{-82670582142728596035332490}{884492329020045639594713401} \end{bmatrix}
 \end{aligned}$$

Bezüglich von Orthogonalitätseigenschaften der verschiedenen Vektoren ist, von Ausnahmen abgesehen, weder die A-Orthogonalität der Residua $r^{(k)}$, also $r^{(k)T} Ar^{(j)} = 0$, $k \neq j$, noch die Orthogonalität von $Ap^{(k)}$, also $(Ap^{(k+1)})^T Ap^{(k)} = (Ap^{(k)})^T Ap^{(k+1)} = 0$ in irgendeiner Weise erfüllt.

$$\begin{aligned}
 r^{(0)T} Ar^{(1)} &= \frac{384}{53}, \quad r^{(0)T} Ar^{(2)} = \frac{-7296}{925}, \quad r^{(0)T} Ar^{(3)} = \frac{-342747440256}{70922089853}, \\
 r^{(1)T} Ar^{(2)} &= \frac{490752}{637325}, \quad r^{(1)T} Ar^{(3)} = \frac{-15581866752}{70922089853}, \\
 r^{(2)T} Ar^{(3)} &= \frac{268346493416448}{852838130482325}, \\
 r^{(1)T} Ar^{(0)} &= 0, \quad r^{(2)T} Ar^{(0)} = \frac{-122112}{12025}, \quad r^{(3)T} Ar^{(0)} = \frac{-296867499264}{70922089853}, \\
 r^{(2)T} Ar^{(1)} &= \frac{124416}{637325}, \quad r^{(3)T} Ar^{(1)} = \frac{-1716441435648}{3758870762209}, \\
 r^{(3)T} Ar^{(2)} &= \frac{161344252366848}{852838130482325}, \\
 (Ap^{(1)})^T Ap^{(0)} &= 24, \quad (Ap^{(2)})^T Ap^{(1)} = \frac{5112}{2809}, \quad (Ap^{(3)})^T Ap^{(2)} = \frac{154162584}{144600625}.
 \end{aligned}$$

Die fehlende Symmetrie der Matrix A ist also entscheidend für diese Bedingungen. Im Gegensatz zum CG, das hier divergierte, und dem “veränderten“ CG, das in zwei Schritten an der Lösung war, haben wir im CR bei nicht eintretendem Abbruch wahrscheinlich eine sehr langsame Konvergenz. Die Konvergenzgeschwindigkeit des CR wäre noch genauer zu untersuchen.

Beispiel 5.22 Sei $Ax = b$, $A \neq A^T$, $A > 0$ aus den Beispielen 5.5, 5.10, 5.16.

$$\begin{pmatrix} 2 & 1 & 0 \\ 0 & 3 & 1 \\ 1 & 2 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \\ 7 \end{pmatrix}, \quad x^* = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad x^{(0)} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}.$$

CR basiert auf dem Funktional $R(x)$ mit seinem Minimum bei $x^* = (1, 1, 1)^T$ und $R(x^*) = -37$, und der Suchrichtung $p(x)$ als Linearkombination der vorhergehenden und des Residuum $r(x)$ von $Q(x)$, wobei sein Minimum bei $z = (\frac{38}{37}, \frac{16}{37}, \frac{54}{37})^T = (1.027027\dots, 0.432432\dots, 1.459459\dots)^T$ liegt.

Weiter sind $Q(z) = -\frac{278}{37} = -7.513513\dots$, $Q(x^*) = -7$, $R(z) = -\frac{49050}{1369} = -35.829072\dots$

CR bleibt zwar nahe der Lösung, konvergiert aber nicht wie die anderen Abstiegsverfahren GV, "verändertes" CG oder Orthomin(0).

Die Iterationsvektoren $x^{(i)}, p^{(i)}, r^{(i)}$, $i = 0, 1, 2, 3$, aus den ersten drei Schritten sind die Spalten der folgenden Felder.

$m = 3$

$xv =$	$\begin{bmatrix} 1 & \frac{2242}{1633} & \frac{469762071419}{344631242141} & \frac{2009610009788004601400421437498658092}{1619204409443714629596642426099338045} \\ -1 & \frac{997}{3266} & \frac{597217475021}{689262484282} & \frac{1788631059511371628010685568505735266}{1619204409443714629596642426099338045} \\ 0 & \frac{2436}{1633} & \frac{414088758300}{344631242141} & \frac{297472673126683073305900420432606239}{323840881888742925919328485219867609} \end{bmatrix}$
$pv =$	$\begin{bmatrix} 2 & \frac{-156185}{5333378} & \frac{-72454786261754341277323}{118770693059648574263881} & \frac{-2.027859\dots \cdot 10^{72}}{2.621822\dots \cdot 10^{72}} \\ 7 & \frac{4450537}{2666689} & \frac{282956947354186452764627}{237541386119297148527762} & \frac{0.345761\dots \cdot 10^{72}}{2.621822\dots \cdot 10^{72}} \\ 8 & \frac{-2301364}{2666689} & \frac{-168082709968414005298405}{118770693059648574263881} & \frac{-0.058473\dots \cdot 10^{72}}{0.104872\dots \cdot 10^{72}} \end{bmatrix}$
$rv =$	$\begin{bmatrix} 2 & \frac{-167}{3266} & \frac{-408478307851}{689262484282} & \frac{-190047570151247388404320233041007463}{323840881888742925919328485219867609} \\ 7 & \frac{5201}{3266} & \frac{137219995465}{689262484282} & \frac{-376438906392671732174989103282884813}{1619204409443714629596642426099338045} \\ 8 & \frac{-1552}{1633} & \frac{-310915884653}{344631242141} & \frac{-201894725238406916363304000466887089}{1619204409443714629596642426099338045} \end{bmatrix}$

Rechnung mit Maple bei `Digits:=16` und Vergleich

CR fuer Ax=b, A<>A', A>0 - unendlicher Iterationsverlauf

m	[x(m) [1],	x(m) [2],	x(m) [3]]	[r(m) [1],	r(m) [2],	r(m) [3]]
0	[1.000000000000,	-1.000000000000,	0.000000000000]		[2.000e+00,	7.000e+00,	8.000e+00]
1	[1.372933251684,	0.305266380894,	1.491733006736]		[-5.113e-02,	1.592e+00,	-9.504e-01]
2	[1.363086145355,	0.866458698449,	1.201541554177]		[-5.926e-01,	1.991e-01,	-9.022e-01]
3	[1.241109521483,	1.104635739058,	0.918576652187]		[-5.869e-01,	-2.325e-01,	-1.247e-01]
4	[1.142991093940,	1.121365474300,	0.847845244778]		[-4.073e-01,	-2.119e-01,	2.229e-01]
5	[1.050436147318,	1.101917369679,	0.845792856391]		[-2.028e-01,	-1.515e-01,	3.626e-01]
6	[0.957961724704,	1.070921830182,	0.882708569369]		[1.315e-02,	-9.547e-02,	3.694e-01]
7	[0.892327848715,	1.041312616162,	0.937451121666]		[1.740e-01,	-6.139e-02,	2.752e-01]
8	[0.864176387870,	1.020521458280,	0.985245251674]		[2.511e-01,	-4.681e-02,	1.538e-01]
9	[0.861100168980,	1.006742671729,	1.019208005961]		[2.711e-01,	-3.944e-02,	4.858e-02]

10	[0.871658829044, 0.996221591888, 1.042898477660]	[2.605e-01, -3.156e-02, -3.570e-02]
20	[1.091490822787, 0.962120446514, 1.008099119971]	[-1.451e-01, 1.055e-01, -4.813e-02]
30	[1.013459681907, 1.044836155917, 0.944409425260]	[-7.176e-02, -7.892e-02, 1.192e-01]
40	[0.933926526240, 1.040154058951, 0.984271858761]	[9.199e-02, -1.047e-01, 4.868e-02]
50	[0.932098070163, 0.996294038181, 1.024323641332]	[1.395e-01, -1.321e-02, -2.198e-02]
60	[0.967351822392, 0.964629356921, 1.043652703355]	[1.007e-01, 6.246e-02, -7.122e-02]
70	[1.007722560048, 0.953917827433, 1.044443454334]	[3.064e-02, 9.380e-02, -9.333e-02]
80	[1.038615477752, 0.957629812204, 1.034086737282]	[-3.486e-02, 9.302e-02, -9.022e-02]
90	[1.057614845980, 0.968277028013, 1.018999111315]	[-8.351e-02, 7.617e-02, -7.017e-02]
100	[1.066235779571, 0.981203623095, 1.003202641041]	[-1.137e-01, 5.319e-02, -4.145e-02]
200	[0.964960335156, 1.038322533614, 0.965947179582]	[3.176e-02, -8.091e-02, 9.461e-02]
300	[0.937912487833, 1.001009721145, 1.020509887612]	[1.232e-01, -2.354e-02, -2.197e-02]
400	[0.983607303330, 0.966461470752, 1.043810552524]	[6.632e-02, 5.681e-02, -9.177e-02]
500	[1.033667456227, 0.959251459001, 1.033376578380]	[-2.659e-02, 8.887e-02, -8.568e-02]
600	[1.059736267584, 0.971671055711, 1.010476035054]	[-9.114e-02, 7.451e-02, -4.498e-02]
700	[1.064418213786, 0.990696244437, 0.988508286192]	[-1.195e-01, 3.940e-02, 1.562e-04]
800	[1.054942162153, 1.009595300419, 0.971677076844]	[-1.195e-01, -4.630e-04, 3.916e-02]
900	[1.037015652270, 1.025078368632, 0.961188955022]	[-9.911e-02, -3.642e-02, 6.807e-02]
1000	[1.015390469375, 1.035565838223, 0.957007149249]	[-6.635e-02, -6.370e-02, 8.545e-02]
1500	[0.941127323226, 1.027496594492, 0.988394181655]	[9.025e-02, -7.088e-02, 5.030e-02]
2000	[0.943643350484, 0.991867887106, 1.026314643853]	[1.208e-01, -1.918e-03, -3.264e-02]
2500	[0.979574411548, 0.967262672213, 1.042612477251]	[7.359e-02, 5.560e-02, -8.455e-02]
3000	[1.019141775884, 0.959601106610, 1.038968491746]	[2.115e-03, 8.223e-02, -9.422e-02]
3500	[1.047142692842, 0.964703806514, 1.024437404003]	[-5.899e-02, 8.145e-02, -7.430e-02]
4000	[1.061223297729, 0.976112345416, 1.007076013984]	[-9.856e-02, 6.459e-02, -4.175e-02]
4500	[1.064256519602, 0.989575990654, 0.990894996463]	[-1.181e-01, 4.038e-02, -6.988e-03]
5000	[1.059357126646, 1.002943794635, 0.977410588520]	[-1.217e-01, 1.376e-02, 2.511e-02]
5500	[1.048913950719, 1.015054019937, 0.967200201213]	[-1.129e-01, -1.236e-02, 5.218e-02]
6000	[1.034876847801, 1.025159564757, 0.960477374690]	[-9.491e-02, -3.596e-02, 7.289e-02]
6500	[1.018990315083, 1.032795550425, 0.957181575943]	[-7.078e-02, -5.557e-02, 8.669e-02]
7000	[1.002785462103, 1.037789830168, 0.956981786060]	[-4.336e-02, -7.035e-02, 9.371e-02]
7500	[0.987467614351, 1.040249605989, 0.959338325369]	[-1.518e-02, -8.009e-02, 9.468e-02]
8000	[0.973842405703, 1.040486832952, 0.963617378549]	[1.183e-02, -8.508e-02, 9.071e-02]
8500	[0.962333352333, 1.038914806803, 0.969204007734]	[3.642e-02, -8.595e-02, 8.302e-02]
9000	[0.953067302847, 1.035958119954, 0.975574898971]	[5.791e-02, -8.345e-02, 7.272e-02]
9500	[0.945979815602, 1.031996880985, 0.982324630275]	[7.604e-02, -7.832e-02, 6.073e-02]
10000	[0.940905565260, 1.027344159804, 0.989159324130]	[9.084e-02, -7.119e-02, 4.777e-02]
GV		
20	[0.999999965066, 1.000000096702, 0.999999978751]	[-2.683e-08, -2.689e-07, -7.347e-08]
Veraendertes CG		
20	[1.000000000002, 0.999999999999, 1.000000000000]	[-3.863e-12, 2.032e-12, -7.260e-13]
Orthomin(0)		
20	[0.999999992857, 0.999999981116, 1.000000023153]	[3.317e-08, 3.350e-08, -4.770e-08]

Bezüglich von Orthogonalitätseigenschaften der verschiedenen Vektoren des CR ist weder die A-Orthogonalität der Residua $r^{(k)}$, ausgenommen $r^{(1)T}Ar^{(0)} = 0$, noch die Orthogonalität von $Ap^{(k)}$ erfüllt.

Von den Bedingungen in (4.18) gelten noch $r^{(k+1)T}Ap^{(k)} = 0$ und $r^{(k)T}Ar^{(k)} = r^{(k)T}Ap^{(k)}$.

Wir betrachten auch das Verhalten von Fehlernormen. Dabei stellen wir fest, dass der Fehler $\|r^{(m)}\|_2 = \|b - Ax^{(m)}\|_2$ eine stetig abnehmende aber nicht gegen Null tendierende Folge bildet. Die Fehlerwerte $\|x^{(m)} - x^*\|_2$ zeigen kein monotones Verhalten.

Normen von Fehlern zum Iterationsverlauf

m	$\ x(m) - x^*\ _2$	$\ r(m)\ _2$
0	2.2360679775	10.8166538264
1	0.9292660338	1.8552155838
2	0.4362153439	1.0976135684
3	0.2751585012	0.6434238761
4	0.2415100429	0.5104257550
5	0.1916006220	0.4421963381
6	0.1433680658	0.3817268679
7	0.1311959854	0.3313820002
8	0.1381552983	0.2981771885
9	0.1403836678	0.2781858148
10	0.1353736013	0.2647833836
500	0.0625134960	0.1262736594
1000	0.0578807447	0.1255458266
1500	0.0660056800	0.1252986145
2000	0.0627268976	0.1251899075
2500	0.0574870465	0.1251208174
3000	0.0593044814	0.1250718709
3500	0.0637608146	0.1250379836
4000	0.0660982768	0.1250142877
4500	0.0657302166	0.1249966504
5000	0.0635784234	0.1249824916
6000	0.0584075006	0.1249599036
7000	0.0573271035	0.1249423724
8000	0.0603912123	0.1249290205
9000	0.0639706965	0.1249190664
9500	0.0652257573	0.1249150594
9990	0.0659993349	0.1249116116
9991	0.0660004517	0.1249116050
9992	0.0660015667	0.1249115984
9993	0.0660026798	0.1249115918
9994	0.0660037911	0.1249115851
9995	0.0660049005	0.1249115785
9996	0.0660060081	0.1249115719
9997	0.0660071138	0.1249115653
9998	0.0660082177	0.1249115587
9999	0.0660093198	0.1249115521
10000	0.0660104200	0.1249115455


```

1 [ 1.1698401737311680804846441523551893907368,
   -1.2353022851257295279119833424045248606123] [-1.490e-07, 1.273e-07]
2 [ 1.1698401737303229102782772604952379553070,
   -1.2353022851250074773384518326433555373394] [-1.490e-07, 1.273e-07]
3 [ 1.1698401737303229033788790976345162922435,
   -1.2353022851250074714440954166149440792810] [-1.490e-07, 1.273e-07]
4 [ 1.1698401737303229033787664548187828934225,
   -1.2353022851250074714439991825829301990139] [-1.490e-07, 1.273e-07]
5 [ 1.1698401737303229033787664520602126936962,
   -1.2353022851250074714439991802262034856006] [-1.490e-07, 1.273e-07]
6 [ 1.1698401737303229033787664520601226195730,
   -1.2353022851250074714439991802261265326586] [-1.490e-07, 1.273e-07]
7 [ 1.1698401737303229033787664520601226158966,
   -1.2353022851250074714439991802261265295178] [-1.490e-07, 1.273e-07]
8 [ 1.1698401737303229033787664520601226158964,
   -1.2353022851250074714439991802261265295176] [-1.490e-07, 1.273e-07]
9 [ 1.1698401737303229033787664520601226158964,
   -1.2353022851250074714439991802261265295176] [-1.490e-07, 1.273e-07]

```

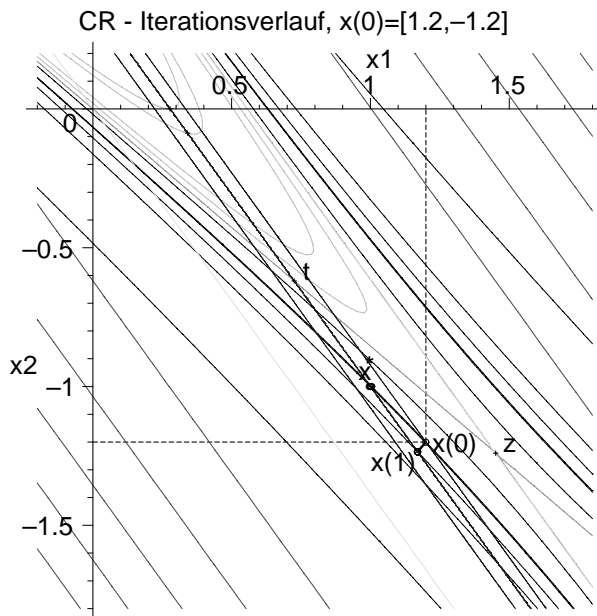


Abb. 5.47 Datei *cr_004.ps*

Höhenlinienbild und CR

zu $R(x)$ mit den Iterierten

$x^{(0)}, x^{(1)}, x^{(2,3,\dots)} \approx x^{(1)}$, (nicht gegen x^*

konvergierender Iterationsverlauf fett

gezeichnet),

Suchrichtung $p(x)$,

Höhenlinien zu $R(x)$ mit `contours=5`

und `contours=[-0.05357,-0.0558,0]`,

$R(x^*) = -0.0558025$, $R(z) \approx 0.9 \cdot 10^{-6}$,

$Q(x^*) = 0.0185$, $Q(z) \approx -0.6 \cdot 10^{-6}$

Zusätzlich sind in der Abbildung 5.47 eingezeichnet:

– Höhenlinien zu $Q(x)$ mit `contours=[-0.03018,-0.0259,-0.01,-0.005,0,0.005,0.01776,0.0185,0.0299,0.05,0.2]`,

– Sattelpunkt z von Funktional $Q(x)$, Minimumstelle $t = \frac{z}{2}$ von Funktional $S(x)$.

Bezüglich von Orthogonalitätseigenschaften der verschiedenen Vektoren ist, von Ausnahmen abgesehen, weder die A-Orthogonalität der Residua $r^{(k)}$ noch die Orthogonalität aufeinander folgender Vektoren $Ap^{(k)}$ erfüllt.

CR für $A^T Ax = A^T b$ konvergiert bei exakter Rechnung nach höchstens zwei Schritten. Die Iterationsvektoren $x^{(i)}, p^{(i)}, r^{(i)}, i = 0, 1, 2$, sind die Spalten der folgenden Felder.

$$\begin{aligned}
 xv &= \begin{bmatrix} \frac{6}{5} & \frac{6849130212700399614580066}{5887076167445197367298925} & 1 \\ \frac{-6}{5} & \frac{-7219938177003319170841344}{5887076167445197367298925} & -1 \end{bmatrix} \\
 pv &= \begin{bmatrix} \frac{-200581}{2500000} & \frac{-258236201748084525118476368884972160646021}{3465766580130123351047149376029084393495030615562500000} & 0 \\ \frac{-289557}{5000000} & \frac{357769113388328433218453459425056889438339}{3465766580130123351047149376029084393495030615562500000} & 0 \end{bmatrix} \\
 rv &= \begin{bmatrix} \frac{-200581}{2500000} & \frac{-6853895783829914}{91985565116331208864045703125} & 0 \\ \frac{-289557}{5000000} & \frac{607719525299752963}{5887076167445197367298925000000} & 0 \end{bmatrix}
 \end{aligned}$$

Bei numerischen Rechnungen in Maple muss man aber die sehr schlechte Kondition der spd Matrix $(A^T A)^2$ in Kauf nehmen und je nach Rechengenauigkeit gemäß `Digits:=...` und geforderter Toleranz werden einige zusätzliche Schritte ausgeführt. Wir vergleichen zu CR für $A^T Ax = A^T b$ bei $x^{(0)} = (1.2, -1.2)^T$ einige Ergebnisse bei `Digits:=40,30,27,24,22,20,16` mit dem exakten Iterationsverlauf, insbesondere betrachten wir die zweiten und folgenden Iterierten. Die schlechte Kondition der Koeffizientenmatrix gestattet bei geringer Genauigkeit schon nach wenigen Iterationen keine Verbesserung mehr.

Rechnung mit Maple bei `Digits:=k`

Lösungsvektoren $x^{(m)}, m = 1, 2, \dots$, im Vergleich mit exakten Vektoren (Ausgabe mit 40 Dezimalstellen)

Digits	[x(1) [1], [x(2) [1], ev. [x(3..5) [1],	x(1) [2]] x(2) [2]] x(3..5) [2]]
exakt	[1.1634179714839162529570123505095126912516314526096, -1.2264047502780215013947641472111442102174466960398] [1, -1]	
40	[1.163417971483916252957012350509512691252, -1.226404750278021501394764147211144210217] [0.9999999999999999999999999999999862343805436, -1.00000000000000000000000022087647505] [1.00000000000000000000000005738521252, -1.00000000000000000000000007950340217]	
30	[1.16341797148391625295701235051, -1.22640475027802150139476414721] [1.00000000000039349612022076008, -1.0000000000054516290451946210] [0.99999999999999999999998988150507, -0.999999999999999999999986129247212] [0.99999999999999999999998988150507, -0.999999999999999999999986129247212]	
27	[1.16341797148391625295701235, -1.22640475027802150139476415] [1.00000001984505194960573694, -1.00000002749400201063716934] [0.9999999999999999984517312345, -0.99999999999999999978549764145] [0.9999999999999999984517312345, -0.99999999999999999978549764145]	
24	[1.16341797148391625295701, -1.22640475027802150139476] [1.12021070965794248159583, -1.166543957520216886956439] [0.999999999999999978246157013, -0.9999999999999999969861494763] [0.999999999999999978246157013, -0.9999999999999999969861494763]	

Rechnung mit Maple bei `Digits:=k`

Residua $r^{(m)}$, $m = 1, 2, \dots$, im Vergleich mit exakten Vektoren

Digits	[r(1) [1], [r(2) [1], ev. [r(3..5) [1],	r(1) [2]] r(2) [2]] r(3..5) [2]]
exakt	[-0.74510557989854285462133887633710722931475493741532e-13, 0.10322943138741209561771228232005310364876532146333e-12] [0,	0]
40	[-0.7451055798985428546213388764e-13, [0.4283857465959084379242420e-28, [-0.1406358574064110e-52,	0.10322943138741209561771228231e-12] 0.309206982783533520223733e-28] 0.1948416437133888e-52]
30	[-0.745105579898542855e-13, [0.1601855785928863501505458e-18, [-0.28125e-43, [-0.281253235350605373785439903517e-43,	0.1032294313874120956e-12] 0.115621637262274468907464e-18] 0.38966e-43] 0.389657664737774714936891552557e-43]
27	[-0.745105579898543e-13, [0.359642054148115266612624e-16, [0.2717e-39, [0.271689408900298343553406132e-39,	0.1032294313874121e-12] 0.25977875215906741431281e-16] -0.3764e-39] -0.376407644609929559487228410e-39]
24	[-0.745105579899e-13, [-0.92847817947022606914699e-14, [0.94e-36, [0.321981069833688020316187e-36,	0.1032294313874e-12] 0.108795883708213109725770e-12] 0.0] -0.446083396140699746415338e-36]
22	[-0.7451055799e-13, [-0.7643959871645731616090e-13, [0.211259e-29, [-0.29110100678945781381e-31,	0.10322943139e-12] 0.1017806780324014303368e-12] 0.15862e-29] 0.40330111890009529661e-31]
20	[-0.74510557e-13, [-0.74512910896632868559e-13, [0.28269827577606e-19, [-0.465075582880789e-24,	0.103229432e-12] 0.10322773288596192863e-12] 0.2040603131610e-19] 0.644330972149721e-24]
16	[-0.7452e-13, [-0.7450931625397471e-13, [-0.7101298775523313e-13, [-0.674631e-23, [-0.1005458481235698e-21,	0.10322e-12] 0.1032277114817325e-12] -0.23516409892245e-14] 0.207003699e-21] 0.1392995971345603e-21]

Die Resultate sind auch ein Hinweis darauf, dass Berechnungen mit Tools in höheren Programmiersprachen wie Pascal oder C bei einer Gleitpunktarithmetik *double* oder *extended* zu "schwach" sind.

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] FADDEJEW, 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.

Anschrift:

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.mathematik.tu-ilmenau.de/~neundorf/index_de.html