Einführung in die Numerische Mathematik

Wintersemester 2011/12
Prof. Dr. Carsten Burstedde
Jutta Adelsberger, Daniel Wissel

universitätbonn

Übungsblatt 7.

Abgabe am **Dienstag, 6.12.2011**

Aufgabe 19. (Residual Norm Steepest Descent) (4 Punkte)

Zeige anhand der Algorithmen, dass für die Verfahren Steepest Descent (SD) und Residual Norm Steepest Descent (RNSD) folgender Zusammenhang gilt:

$$\mathrm{RNSD}(A, b, x, \eta) = \mathrm{SD}(A^T A, A^T b, x, \eta).$$

Aufgabe 20. (Full Orthogonalization Method/Arnoldi-Verfahren) (5 Punkte)

Löse das System $Ax = b$,

$$A = \begin{pmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 3 & 2 & 1 & 1 \end{pmatrix}, \qquad b = \begin{pmatrix} 0 \\ 2 \\ 1 \\ 3 \end{pmatrix},$$

mit Hilfe der "Full Orthogonalization Method" $\mathrm{FOM}(A, b, x_0, m)$ zum Startwert $x_0 = \left(0, \frac{1}{2}, \frac{1}{2}, \frac{3}{2}\right)^T$ und $m = 4$.

Aufgabe 21. (Arnoldi-Verfahren für schiefsymmetrische Matrizen) (5 Punkte)

Das Arnoldi-Verfahren entspricht einer Orthogonaltransformation $H_m = V_m^T A V_m$ von $A \in \mathbb{R}^{n\times n}$ in eine obere Hessenbergmatrix $H_m \in \mathbb{R}^{m\times m}$. Sei nun A schiefsymmetrisch, d.h. $A^T = -A$. Welche Vereinfachungen ergeben sich daraus für die Arnoldi-Iteration? Geben Sie den Algorithmus an.

Aufgabe 22. (Alternative Herleitung des CG-Verfahrens) (28 Punkte)

Sei $A \in \mathbb{R}^{n\times n}$ symmetrisch und positiv definit. Zwei Vektoren heißen *konjugiert* oder *A-orthogonal*, wenn $\langle x, y\rangle_A := \langle x, Ay\rangle = x^T Ay = 0$ gilt.

a) Beim Verfahren des steilsten Abstiegs stellte man fest, dass fast parallele Suchrichtungen den Fehler kaum reduzieren und daher das Verfahren deutlich verlangsamen. Skizzieren Sie analog zu Aufgabe 17 a) einen solchen extrem ungünstigen Fall.

Die Idee des CG-Verfahrens (conjugate gradients; Hestenes und Stiefel 1952) ist nun, solche fast parallelen Suchrichtungen zu vermeiden, indem sie statt dessen orthogonal bezüglich A gewählt werden. D.h. aus den Suchrichtungen werden die Komponenten (Richtungen) herausgenommen, in die bereits minimiert wurde.

b) Zeigen Sie, dass k paarweise konjugierte Richtungen $\{x_1, \ldots, x_k\}$ mit $x_i \neq 0$ für $1 \leq i \leq k$ linear unabhängig sind. Insbesondere bilden also n paarweise konjugierte Vektoren aus $\mathbb{R}^n \backslash \{0\}$ eine Basis des $\mathbb{R}^n$.
Bemerkung: Der Begriff „konjugierte Gradienten“ ist irreführend; die Suchrichtungen sind konjugiert.

c) Sei $\{p_0, \dots, p_{n-1}\}$ eine solche A-orthogonale Basis des $\mathbb{R}^n$. Außerdem sei x^* die exakte Lösung von $Ax = b$ und $x_0 \in \mathbb{R}^n$ beliebig. Leiten Sie für das Iterationsverfahren $x_{k+1} = x_k + \alpha_k p_k$ die Berechnungsvorschrift

$$\alpha_k = -\frac{r_k^T p_k}{p_k^T A p_k}$$

mit dem Residuum $r_k := Ax_k - b$ her.
Hinweis: Stellen Sie dazu $x^* - x_0$ zu obiger Basis dar und betrachten Sie $p_k^T A(x^* - x_0)$.

d) Zeigen Sie, dass das Verfahren der konjugierten Richtungen aus Teilaufgabe c) bei exakter Rechnung kein Iterationsverfahren sondern ein direktes Verfahren ist.
Bemerkung: Dies ist allerdings nicht praxisrelevant, da in der Regel nicht exakt gerechnet werden kann. Außerdem müssen bisher die A-orthogonalen Richtungen im Voraus bekannt sein.

e) Die Idee bei CG ist es, die Suchrichtungen so zu konstruieren, dass sie den steilsten Abstieg ausnutzen, dabei jedoch stets die konjugierten Richtungen herauszunehmen, in die bereits minimiert wurde. Dies führt zum Ansatz

$$p_{k+1} = -r_{k+1} + \beta_k p_k$$

mit $p_0 := -r_0$. Bestimmen Sie nun β_k so, dass stets $\langle p_{k+1}, p_k \rangle_A = 0$.
Bemerkung: Es werden also keine Richtungen vorgegeben, sondern während der Iteration berechnet (kein komplizierter Orthogonalisierungsprozess).

f) Zeigen sie, dass sich auch das Residuum rekursiv über

$$r_{k+1} = r_k + \alpha_k A p_k$$

konstruieren lässt. Dabei stammt α_k aus $x_{k+1} = x_k + \alpha_k p_k$.

g) Schreiben Sie nun den CG-Algorithmus in Pseudocode auf.
Bemerkung: In der Praxis formuliert man α_k und β_k noch in

$$\alpha_k = \frac{r_k^T r_k}{p_k^T A p_k} \quad \text{und} \quad \beta_k = \frac{r_{k+1}^T r_{k+1}}{r_k^T r_k}$$

um. Diese Darstellung ist numerisch stabiler, also weniger anfällig für Rundungsfehler.

h) Für alle Iterationsverfahren der Form $x_{k+1} = x_k + \alpha_k(b - Ax_k)$ minimiert das CG-Verfahren den Fehler $\|x_k - x^*\|_A$. Insbesondere gilt die Abschätzung

$$\|x_k - x^*\|_A \leq 2 \left(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^k \|x_0 - x^*\|_A$$

mit der Kondition $\kappa = \kappa(A)$. Vergleichen Sie diese Fehlerabschätzung mit derjenigen von Steepest Descent. Schreiben Sie diese dafür zunächst mittels $\kappa = \lambda_{\max}/\lambda_{\min}$ um.

i) Das CG-Verfahren lässt sich nur auf symmetrische Matrizen anwenden. Welche Möglichkeit besteht, auch lineare Gleichungssysteme mit unsymmetrischen Matrizen mittels CG-Verfahren zu lösen? Was bedeutet das für die Kondition und somit die Konvergenzgeschwindigkeit?

Gesamtpunktzahl: 42 Punkte