Einführung in die Grundlagen der Numerik

Wintersemester 2014/2015
Prof. Dr. Marc Alexander Schweitzer
Sa Wu

Übungsblatt 4.

Abgabe am **04.11, vor der Vorlesung**.

Aufgabe 9. (Invarianzeigenschaften von Krylov-Räumen)

Zeigen Sie die folgenden Eigenschaften des Raums $\mathcal{K}_k(A, v)$.

a) $\mathcal{K}_k(\alpha A, \beta v) = \mathcal{K}_k(A, v)$ für alle $\alpha, \beta \neq 0$ (Skalierung)

b) $\mathcal{K}_k(A - \alpha I, v) = \mathcal{K}_k(A, v)$ für alle $\alpha \in \mathbb{C}$ (Translation)

c) $\mathcal{K}_k(TAT^{-1}, Tv) = T\mathcal{K}_k(A, v)$ für jedes reguläre $T \in \mathbb{K}^{n\times n}$ (Basiswechsel)

(3 Punkte)

Aufgabe 10. (Herleitung des CG-Verfahrens)

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 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.

a) Zeigen Sie, dass k paarweise konjugierte Richtungen $\{d_1, \ldots, d_k\}$ mit $d_i \neq 0$ 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.

b) Sei $\{d_0, \ldots, d_{n-1}\}$ 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 $x_{k+1} = x_k + \alpha_k d_k$ die Berechnungsvorschrift

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

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

c) 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.

d) 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

$$d_{k+1} = r_{k+1} + \beta_k d_k$$

mit $d_0 := r_0$. Zeigen Sie, dass aus $\langle d_{k+1}, d_k \rangle_A = 0$ folgt, dass

$$\beta_k = -\frac{r_{k+1}^T A d_k}{d_k^T A d_k} \; .$$

Bemerkung: Es werden also keine Richtungen vorgegeben, sondern während der Iteration berechnet (kein komplizierter Orthogonalisierungsprozess).

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

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

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

f) Zeigen Sie, dass

$$\alpha_k = \frac{r_k^T r_k}{d_k^T A d_k} \, , \qquad\qquad \beta_k = \frac{r_{k+1}^T r_{k+1}}{r_k^T r_k} \; .$$

Diese Form ist numerisch stabiler, also weniger anfällig für Rundungsfehler.

g) 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.

h) 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?

(6 Punkte)

Aufgabe 11. (Vorkonditionierung: Spaltenäquilibrierung)

Unter einer Spaltenskalierung versteht man die Multiplikation eines Gleichungssystems mit einer Diagonalmatrix $D = \operatorname{diag}\{d_j\}$ von rechts: $Ax = b \to AD\tilde{x} = b$, $x = D\tilde{x}$. Eine solche Skalierung wird mit dem Ziel durchgeführt, die Kondition des Problems zu verringern, sodass also $\operatorname{cond}(AD) \leq \operatorname{cond}(A)$.

Eine Spaltenäquilibrierung ist eine spezielle Form der Spaltenskalierung, bei der

$$d_j := \frac{\|A\|_1}{\sum_i |a_{ij}|}, \qquad j = 1, \ldots, n,$$

gewählt wird. Die Matrix $A \in \mathbb{K}^{n \times n}$ wird im folgenden als regulär vorausgesetzt. Zeige:

a) Es gilt $\|D\|_\infty^{-1} \operatorname{cond}_1(A) \leq \operatorname{cond}_1(AD) \leq \operatorname{cond}_1(A)$.

b) Mit einer beliebigen Diagonalmatrix $C \in \mathbb{K}^{n \times n}$ gilt $\operatorname{cond}_1(AD) \leq \operatorname{cond}_1(AC)$.

(3 Punkte)

Programmieraufgabe 6. (B-Splines Programm, Teil 2)

a) Implementieren Sie das Gradientenverfahren für lineare Gleichungssysteme und das CG-Verfahren.

b) Wenden Sie sie für $n = 10, 20$ auf die Gram-Matrix aus Teil 1 der Programmieraufgabe mit rechter Seite und Startvektor

$$b = (93, 119, 120, 120, \dots, 120, 120, 119, 93) \qquad\qquad x_0 = 0$$

an.

c) Verwenden Sie als Abbruchkriterium eine Schranke fürs Residuum $\|r_k\| < 10^{-7}$. Plotten Sie semilogarithmisch die Norm des Residuums gegen die Iterationszahl.

(4 Punkte)

Die Abgabe der Programmieraufgabe am 03-05.11. im CIP Pool, Wegelerstraße 6.