



Einführung in die Grundlagen der Numerik

Winter semester 2019/2020
Prof. Dr. Marc Alexander Schweitzer
Denis Düsseldorf



Übungsblatt 4.

Abgabe: 12.11.2019 vor der Vorlesung

Aufgabe 9. (Invarianzeigenschaften von Krylov-Räumen)

Aus der Vorlesung, bzw. vom vorherigen Aufgabenblatt kennen Sie bereits die Krylov-Unterraum Verfahren und die zugehörigen Krylov-Räume: Für einen beliebigen Startvektor $x_0 \in \mathbb{K}^n$, für $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ einer Krylov-Methode, die wir zum lösen des linearen Gleichungssystems $Ax = b$ mit $A \in \mathbb{K}^{n \times n}$, $b \in \mathbb{K}^n$ verwenden möchten, definieren wir den k -ten Krylov-Unterraum $K_k(A, r_0) \subset \mathbb{K}^n$ als

$$K_k = K_k(A, r_0) := \text{span}\{r_0, Ar_0, \dots, A^{k-1}r_0\},$$

wobei $r_0 = b - Ax_0$ das Startresiduum ist.
Zeigen Sie, dass für beliebiges $x \in \mathbb{K}^n$ gilt:

- Skalierung:** Für alle $\alpha, \beta \in \mathbb{K} \setminus \{0\}$ gilt $K_k(\alpha A, \beta x) = K_k(A, x)$
 - Translation:** Für alle $\alpha \in \mathbb{K}$ gilt $K_k(A - \alpha \mathbb{I}, x) = K_k(A, x)$
 - Basiswechsel:** Für $T \in \mathbb{K}^{n \times n}$ regulär gilt $K_k(TAT^{-1}, Tx) = TK_k(A, x)$
- (3 Punkte)

Aufgabe 10. (Herleitung des CG-Verfahrens)

Das CG-Verfahren ist eine effiziente numerische Methode um große lineare Gleichungssysteme $Ax = b$, mit symmetrisch positiv-definiter Systemmatrix A zu lösen. Es wurde 1952 von Stiefel und Hestenes vorgeschlagen.

In dieser Aufgabe leiten wir das CG-Verfahren her. Es sei $A \in \mathbb{R}^{n \times n}$ symmetrisch und positiv definit. Wir nennen zwei Vektoren *konjugiert*, bzw. A -orthogonal, falls sie orthogonal im durch A induzierten Skalarprodukt sind, d.h. wenn

$$\langle x, y \rangle_A := \langle x, Ay \rangle = x^T Ay = 0.$$

- a) Optimierungsverfahren werden i.d.R. formuliert zum lösen von Problemen der Form

$$\min_{x \in \mathbb{R}^n} f(x),$$

wobei $f : \mathbb{R}^n \rightarrow \mathbb{R}$ eine differenzierbare Funktion ist. Um die Verfahren auf lineare Gleichungssysteme anwenden zu können, müssen wir deren Lösung als Minimierungsproblem auffassen.

Zeigen Sie, dass das Lösen des Gleichungssystems $Ax = b$ mit spd Matrix A äquivalent zur Minimierung des Energiefunktional

$$J(x) = \frac{1}{2}x^T Ax - b^T x$$

ist.

- b) Das Verfahren des steilsten Abstiegs ist effektiv, solange keine parallelen Suchrichtungen benutzt werden. In diesen Fällen wird der Fehler kaum reduziert, wodurch das Verfahren deutlich verlangsamt wird. Skizzieren Sie einen für das Verfahren des steilsten Abstiegs ungünstigen Fall, der zu vielen Iterationen und parallelen Suchrichtungen führt.

Erinnerung: Verfahren des steilsten Abstiegs

Das Verfahren des Steilsten Abstiegs ist die Iteration $x_{k+1} = x_k + \lambda_k p_k$. Hierbei beschreibt p_k die Suchrichtung, in diesem Fall den negativen normierten Gradienten des Energiefunktional,

$$p_k := \begin{cases} \frac{Ax_k - b}{\|Ax_k - b\|}, & \text{für } Ax_k - b \neq 0 \\ 0, & \text{sonst} \end{cases}$$

und λ_k die genutzte Schrittweite. Für die optimale Schrittweite

$$\lambda_k = \lambda_k^* := \frac{\langle b - Ax_k, p_k \rangle}{\langle Ap_k, p_k \rangle}$$

erhält man im Falle einer SPD-Matrix A die Fehlerabschätzung

$$\|x_k - x^*\|_A \leq \left[\frac{\kappa(A) - 1}{\kappa(A) + 1} \right]^k \|x_0 - x^*\|_A.$$

- c) Um die Problematik aus dem vorherigen Aufgabenteil zu umgehen, führten Hestenes und Stiefel 1952 die Idee ein, solche fast parallelen Suchrichtungen zu vermeiden, indem statt Orthogonalität im euklidischen Skalarprodukt die Orthogonalität im A -induzierten Skalarprodukt gefordert wird. Bildlich gesprochen werden aus einer neuen Suchrichtung alle Anteile von Richtungen, in die bereits in vorherigen Schritten minimiert wurde, entfernt.

Zeigen Sie, dass k paarweise konjugierte Richtungen $\{d_0, \dots, d_{k-1}\}$ mit $d_i \in \mathbb{R}^n \setminus \{0\}$ für $0 \leq i < k$ linear unabhängig sind. Insbesondere bilden also n paarweise konjugierte Vektoren aus $\mathbb{R}^n \setminus \{0\}$ eine Basis des \mathbb{R}^n .

Beim CG Verfahren sind die Suchrichtungen konjugiert zueinander, nicht die Gradienten!

- d) Es sei $\{d_0, \dots, d_{n-1}\}$ eine solche A -orthogonale Basis des \mathbb{R}^n . Mit $x_* \in \mathbb{R}^n$ bezeichnen wir die exakte Lösung des linearen Gleichungssystems $Ax = b$ und $x_0 \in \mathbb{R}^n$ sei beliebig. Stellen Sie $x_* - x_k$ in der obigen Basis $\{d_i\}_{i=0}^{n-1}$ dar und betrachten Sie $d_k^T A(x_* - x_k)$, um für das Iterationsverfahren $x_{k+1} = x_k + \alpha_k d_k$ die Berechnungsvorschrift

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

herzuleiten. Hierbei bezeichnet $r_k = b - Ax_k$ das Residuum.

- e) In der Numerik gibt es keine exakte Arithmetik, d.h. wir machen Rundungsfehler. Zeigen Sie trotzdem: Die Iteration $x_{k+1} = x_k + \alpha_k d_k$ für das α_k aus d) ist in exakter Arithmetik ein direktes (kein iteratives!) Lösungsverfahren.
- f) Die Idee des CG Verfahrens ist es, die Suchrichtungen so zu konstruieren dass der steilste Abstieg ausgenutzt, aber der Anteil an den vorherigen, konjugierten Richtungen entfernt wird. Es wird also in jede Suchrichtung nur in einem Schritt minimiert. Das führt zum Ansatz

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

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

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

Die Richtungen werden also nicht vorgegeben (was einen Orthogonalisierungsprozess involvieren würde), sondern in jedem CG-Schritt berechnet.

g) Zeigen Sie die folgende rekursive Berechnungsvorschrift des Residuums mit dem α_k aus obiger Teilaufgabe:

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

h) Zeigen Sie, dass es sich beim CG Verfahren um ein Krylov-Unterraum Verfahren handelt. Das bedeutet, dass das Verfahren in exakter Arithmetik nach maximal n Iterationen mit der exakten Lösung $x_n = x^*$ konvergiert.

i) Zeigen Sie, dass $\tilde{\alpha}_k = \alpha_k$ und $\tilde{\beta}_k = \beta_k$ mit

$$\tilde{\alpha}_k = \frac{r_k^T r_k}{d_k^T A d_k}, \quad \tilde{\beta}_k = \frac{r_{k+1}^T r_{k+1}}{r_k^T r_k}.$$

Diese Berechnungsvorschriften sind weniger anfällig für Rundungsfehler (numerisch stabiler!).

j) Für das CG-Verfahren gilt die Fehlerabschätzung

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

Vergleichen Sie die diese Schranke mit der vom Verfahren des steilsten Abstiegs. Die Konditionszahl der SPD Matrix A ist explizit gegeben durch $\kappa = \frac{\lambda_{max}}{\lambda_{min}}$.

(8 Punkte)

Programmieraufgabe 3. (Verfahren des steilsten Abstiegs)

Implementieren Sie das Verfahren des steilsten Abstiegs. Benutzen Sie die optimale Abstiegsrichtung

$$q_k = \begin{cases} \frac{Ax_k - b}{\|Ax_k - b\|}, & \text{wenn } Ax_k - b \neq 0 \\ 0, & \text{sonst} \end{cases}$$

und die optimale Schrittweite

$$\lambda_k^* := \frac{\langle b - Ax_k, q_k \rangle}{\langle Aq_k, q_k \rangle}.$$

Benutzen Sie als Abbruchkriterium eine Schranke für das Residuum, $\|r_k\| < 10^{-7}$, sowie eine Begrenzung der Iterationszahl $iter_{max} = 1e4$. Testen Sie das Verfahren mit der Matrix `matrix_500.txt` und der rechten Seite `rhs_500.txt`, die Sie als Textdatei auf der Webseite finden.

(5 Punkte)

Abgabe per Mail an die Tutoren