Übungsblatt 4
Gradienten- und CG-Verfahren

Schreiben Sie ein Programm, welches das lineare Gleichungssystem $Ax = y$ mit dem Gradientenverfahren und der Methode der konjugierten Gradienten löst.

Gradientenverfahren

Das Gradientenverfahren löst ausgehend von einem Startwert x_0 das LGS $Ax = y$ über Minimierung des zugehörigen quadratischen Potentials $f(x) = \frac{1}{2}x^T Ax - y^T x$. Dazu wird in jedem Iterationsschritt die zuvor gefundene Lösung mit der Richtung des steilsten Abstiegs $-\nabla f(x) = y - Ax = r$ modifiziert.

(Gradientenverfahren 1) Setze $x_0 = 0,\ k = 0, r_0 = y$.

(Gradientenverfahren 2) Berechne:

$$\alpha_{k+1} = \frac{r_k{}^\mathsf{T} r_k}{r_k{}^\mathsf{T} A r_k},$$
$$x_{k+1} = x_k + \alpha_{k+1} r_k,$$
$$r_{k+1} = y - Ax_{k+1}$$

(Gradientenverfahren 3) Inkrementiere k. Relativer Fehler $err = \frac{\|r_k\|_2}{\|y\|_2} < \varepsilon_{\mathrm{TOL}}$?
Ja. Setze $x = x_k$ ENDE
Nein. Gehe zu (CG 4).

(Gradientenverfahren 4) Falls $k < N_{\max}$ gehe zu (CG 2). Sonst: **FEHLER.** Keine Konvergenz in $N_{\max}$ Schritten.

CG-Verfahren

Das CG-Verfahren funktioniert nur für symmetrische, positiv definite Matrizen. Gerade solche Matrizen treten in der rechnergestützten Mechanik jedoch meist auf. Beachten Sie auch die unter ILIAS bereitgestellten Zusatzmaterialien.

Herleitung des CG-Algorithmus (Beweisskizze)

(1) Darstellung der k-ten iterierten Lösung als Linearkombination:

$$x_k = x_0 + \sum_{i=1}^{k} \alpha_i p_i$$

(p_i: Abstiegsrichtungen; α_i: Koeffizienten)

(2) Definition des Residuums:

$$r_k = y - Ax_k = y - A\left(x_0 + \sum_{i=1}^{k} \alpha_i p_i\right)$$

(3) *Forderung:* Orthogonalität der Abstiegsrichtungen bezüglich der Systemmatrix A:

$$p_i{}^{\mathsf{T}} A p_j = \delta_{ij}\, p_i{}^{\mathsf{T}} A p_i$$

(4) Definition einer Zielfunktion, die minimiert werden soll

$$f(x_k) = \frac{1}{2} x_k{}^{\mathsf{T}} A x_k - x_k{}^{\mathsf{T}} y \underbrace{=}_{(3)} \frac{1}{2} x_0{}^{\mathsf{T}} A x_0 - x_0{}^{\mathsf{T}} y + \sum_{i=1}^{k} \alpha_i p_i{}^{\mathsf{T}} \left(A x_0 - y\right) + \sum_{i=1}^{k} \frac{\alpha_i^2}{2} p_i{}^{\mathsf{T}} A p_i.$$

Notwendige Bedingung für ein Minimum von f (mit $x = A^{-1} b$):

$$\frac{\partial f}{\partial \alpha_i} = 0 \underbrace{=}_{(3)} \alpha_i p_i{}^{\mathsf{T}} A p_i + p_i{}^{\mathsf{T}} \underbrace{(A x_0 - b)}_{=-r_0} \;\Rightarrow\; \alpha_i = \frac{p_i{}^{\mathsf{T}} A (x - x_0)}{p_i{}^{\mathsf{T}} A p_i} = \frac{p_i{}^{\mathsf{T}} r_0}{p_i{}^{\mathsf{T}} A p_i}$$

(5) Für das Residuum im Folgeschritt gilt:

$$r_{k+1} = y - A x_{k+1} = y - A x_k - \alpha_{k+1} p_{k+1} = r_k - \alpha_{k+1}\, A p_{k+1}$$

(6) *Folgerung:* Orthogonalität zwischen Residuum r_k und Abstiegsrichtungen p_i ($i < k$):

$$p_i{}^{\mathsf{T}} r_k = p_i{}^{\mathsf{T}} \left(b - A\left(x_0 + \sum_{j=1}^{k} \alpha_j p_j\right)\right) = p_i{}^{\mathsf{T}} \left(r_0 - \sum_{j=1}^{k} \alpha_j A p_j\right) = p_i{}^{\mathsf{T}} r_0 - \sum_{j=1}^{k} \alpha_j p_j{}^{\mathsf{T}} A p_i$$

$$\underbrace{=}_{(3)} p_i{}^{\mathsf{T}} r_0 - \alpha_i p_i{}^{\mathsf{T}} A p_i \underbrace{=}_{(4)} p_i{}^{\mathsf{T}} r_0 - \frac{r_0{}^{\mathsf{T}} p_i}{p_i{}^{\mathsf{T}} A p_i} p_i{}^{\mathsf{T}} A p_i = 0$$

D.h. es muss keine Abstiegsrichtung ein zweites Mal durchlaufen werden.

(7) Für die Residuen gelten die folgende Beziehungen *(ohne Beweis)*

$$r_i{}^{\mathsf{T}} r_j = \delta_{ij} r_i{}^{\mathsf{T}} r_i \qquad (i, j = 1, \ldots, k),$$
$$r_i{}^{\mathsf{T}} A p_j = 0 \qquad (i = 1, \ldots, k;\; j < i).$$

(8) *Ansatz für Abstiegsrichtung* (p_{k+1})

Da x Lösung des Minimierungsproblems $\min_x \; f(x)$ ist, folgt unmittelbar, dass die Abstiegsrichtung p_{k+1} möglichst parallel zu $-\mathrm{D} f(x_k) = r_k$ zu wählen ist. Mit (3) muss jedoch $p_{k+1}{}^{\mathsf{T}} A p_j = 0$ für $1 \leq j \leq k$ gelten. Daraus folgt nach Projektion

$$p_{k+1} = r_k - \sum_{j=1}^{k} \frac{r_k{}^{\mathsf{T}} A p_j}{p_j{}^{\mathsf{T}} A p_j}\, p_j \underbrace{=}_{(7)} r_k - \underbrace{\frac{r_k{}^{\mathsf{T}} A p_k}{p_k{}^{\mathsf{T}} A p_k}}_{=\beta_{k+1}}\, p_k.$$

(9) Es gilt:

$$(x-x_0)^\mathsf{T} A p_{k+1} = (x - x_k + x_k - x_0)^\mathsf{T} A p_{k+1} = (x-x_k)^\mathsf{T} A p_{k+1} + \underbrace{(x_k - x_0)^\mathsf{T} A p_{k+1}}_{(6)=0}$$

$$= (Ax - Ax_k)^\mathsf{T} p_{k+1} = {r_k}^\mathsf{T} p_{k+1} = {r_k}^\mathsf{T} r_k - \frac{{r_k}^\mathsf{T} A p_k}{{p_k}^\mathsf{T} A p_k} \underbrace{{r_k}^\mathsf{T} p_k}_{=0} = {r_k}^\mathsf{T} r_k$$

(10) Aus (4) und (8) folgt

$$\alpha_{k+1} = \frac{(x-x_0)^\mathsf{T} A p_{k+1}}{{p_{k+1}}^\mathsf{T} A p_{k+1}} = \frac{{r_k}^\mathsf{T} r_k}{{p_{k+1}}^\mathsf{T} A p_{k+1}}$$

(11) Nebenrechnung:

$${r_k}^\mathsf{T} A \left(-\alpha_k p_k\right) \underbrace{=}_{(5)} {r_k}^\mathsf{T} \left(r_k - r_{k-1}\right) = {r_k}^\mathsf{T} r_k - \underbrace{{r_k}^\mathsf{T} r_{k-1}}_{=0} = {r_k}^\mathsf{T} r_k$$

(12) Mit (7), (9) und (10):

$$\beta_{k+1} = -\frac{{r_k}^\mathsf{T} A p_k}{{p_k}^\mathsf{T} A p_k} = \frac{{r_k}^\mathsf{T} r_k}{\alpha_k \, {p_k}^\mathsf{T} A p_k} = \frac{{r_k}^\mathsf{T} r_k}{{r_{k-1}}^\mathsf{T} r_{k-1}}$$

Der CG-Algorithmus lautet:

(CG 1) Setze $x_0 = 0,\ k = 0, r_0 = y$.

(CG 2) Berechne:

$$\begin{aligned}
\beta_{k+1} &= \frac{{r_k}^\mathsf{T} r_k}{{r_{k-1}}^\mathsf{T} r_{k-1}} \qquad (\beta_1 = 0) && \text{aus } (11),\\
p_{k+1} &= r_k + \beta_{k+1} p_k && \text{aus } (8),\\
d_{k+1} &= A p_{k+1}, &&\\
\alpha_{k+1} &= \frac{{r_k}^\mathsf{T} r_k}{{d_{k+1}}^\mathsf{T} p_{k+1}} && \text{aus } (10),\\
x_{k+1} &= x_k + \alpha_{k+1} p_{k+1} && \text{aus } (1),\\
r_{k+1} &= r_k - \alpha_{k+1} d_{k+1} && \text{aus } (5)
\end{aligned}$$

(CG 3) Inkrementiere k. Relativer Fehler $err = \frac{\|r_k\|_2}{\|y\|_2} < \varepsilon_{\mathrm{TOL}}$?
Ja. Setze $x = x_k$ ENDE
Nein. Gehe zu (CG 4).

(CG 4) Falls $k < N_{\max}$ gehe zu (CG 2). Sonst: **FEHLER.** Keine Konvergenz in $N_{\max}$ Schritten.

1. Implementieren Sie das in der Übung vorgestellte Gradientenverfahren. Lösen Sie hiermit das lineare Gleichungssystem $Ax = y$ über eine Minimierung des Potentials/ der quadratischen Form $f(x) = \frac{1}{2}x^T Ax - x^T y$.

2. Implementieren Sie das CG-Verfahren nach dem auf dem Übungsblatt diskutierten Algorithmus.

3. Testen Sie beide Verfahren, an der in der Vorlage gegebenen rechten Seite und der 2×2 Testmatrix B. Diskutieren Sie die Visualisierung.

4. Eine weitere Testmatrix können Sie aus der Datei `sparse_matrix2.mat` importieren. Die rechte Seite des Gleichungssystems kann z.B. über einen Zufallsvektor gestaltet werden. Vergleichen Sie die Anzahl der benötigten Iterationen für das CG-Verfahren mit dem `bicg`-Löser, der in Matlab implementiert ist. Stimmen die Ergebnisse überein?
   In diesem Aufgabenteil ist die Visualiserung nicht möglich, kommentieren Sie diese deshalb aus.

Kontakt

Dipl.-Ing. Johannes Ruck johannes.ruck@kit.edu
M.Sc. Hannes Erdle hannes.erdle@kit.edu
Sprechstunde Do. 13:00-14:00 Uhr (Geb. 10.23, Raum 302.3)