Das Verfahren der konjugierten Gradienten

Michael Bauer

7. November 2013

1.1 Motivation

Löse ein Gleichungssystem $Ax = b$, wobei $A \in \mathbb{R}^n$ s.p.d., $x, b \in \mathbb{R}^n$ und n sehr groß.

1.2 Definition (A-orthogonal)

Sei A eine symmetrische, nicht singuläre Matrix. Zwei Vektoren x und y heißen **konjugiert** oder **A-orthogonal**, wenn $x^T Ay = 0$ ist.

1.3 Lemma

Sei U_k ein k-dimensionaler Teilraum des $\mathbb{R}^n$ $(k \leq n)$, und $p^0, p^1, ..., p^{k-1}$ eine *A-orthogonale Basis* dieses Teilraums, also $\langle p^i, p^j \rangle_A = 0 \;\; für \;\; i \neq j$. Sei $v \in \mathbb{R}^n$, dann gilt für $u^k \in U_k$:

$$\|u^k - v\|_A = \min_{u \in U_k} \|u - v\|_A \tag{1}$$

genau dann, wenn u^k die *A-orthogonale Projektion* von v auf $U_k = span\{p^0, ..., p^{k-1}\}$ ist. Außrdem hat $\mathbf{u}^k$ die Darstellung

$$P_{U_{k,\langle\cdot,\cdot\rangle}}(v) = \mathbf{u}^k = \sum_{j=0}^{k-1} \frac{\langle v, p^j \rangle_A}{\langle p^j, p^j \rangle_A} p^j \tag{2}$$

1.4 Algorithmus

Die folgenden Teilschritte definieren die Vorgehensweise zur Erzeugung der Lösung x^* durch Näherungen $x^1, x^2, ...$

$U_1 := span\{r^0\}$, wobei $r^0 = b - Ax^0$
dann gilt für $k = 1, 2, 3, ...$, falls $r^{k-1} = b - Ax^{k-1} \neq 0$:

CG_a**:** Bestimme A-orthogonale Basis

$$p^0, ..., p^{k-1} \quad \text{von} \quad U_k \tag{3}$$

CG_b**:** Bestimme $x^k \in U_k$, so dass

$$\|x^k - x^*\|_A = \min_{u \in U_k} \|x - x^*\|_A \tag{4}$$

CG_c**:** Erweitung des Teilraumes:

$$U_{k+1} := span\{p^0, ..., p^{k-1}, r^k\} \;\; wobei \;\; r^k := b - Ax^k \tag{5}$$

d.h.

$$x^k = \sum_{j=0}^{k-1} \frac{\langle x^*, p^j \rangle_A}{\langle p^j, p^j \rangle_A} p^j \tag{6}$$

1.5 Lemma

Sei x^* die Lösung von Gleichung (4) und x^k die optimale Approximation von x^* in U_k. Dann kann x^k wie folgt berechnet werden:

$$x^k = x^{k-1} + \alpha_{k-1} p^{k-1}, \;\; mit \;\; \alpha_{k-1} := \frac{\langle r^0, p^{k-1} \rangle}{\langle p^{k-1}, Ap^{k-1} \rangle} \tag{7}$$

1.6 Lemma

Um U_{k+1} zu erhalten, also den Teilraum zu erweitern, muss lediglich das neue Residuum $r^k = b - Ax^k$ berechnet werden. Dieses erhält man durch:

$$r^k = r^{k-1} - \alpha_{k-1} A p^{k-1} \tag{8}$$

Wobei α_{k-1} wie in (7).

1.7 Lemma (Zusammenhang zu Krylovräumen)

Man kann U_k auch in folgender Form schreiben:

$$U_k := span\{r^0, r^1, ..., r^{k-1}\} = span\{p^0, p^1, ..., p^{k-1}\} = span\{r^0, Ar^0, ..., A^{k-1} r^0\} \tag{9}$$

1.8 Satz (Bestimmung einer A-orthogonalen Basis)

Durch

$$p^{k-1} = r^{k-1} - \sum_{j=0}^{k-2} \frac{\langle r^{k-1}, p^j \rangle_A}{\langle p^j, p^j \rangle_A} p^j \tag{10}$$

wird die A-orthogonale Basis zum Vektor r^{k-1} bestimmt.

1.9 Lemma

Für jedes $\mathbf{r}^{k-1}$ und $\mathbf{p}^j$ gilt:

$$\langle r^{k-1}, p^j \rangle_A = 0 \;\; für \;\; 0 \leq j \leq k-3$$

1.10 Algorithmus

Gegeben: $A \in \mathbb{R}^n$ s.p.d., $b \in \mathbb{R}^n$, Startvektor $x^0 \in \mathbb{R}^n$, $\beta_{-1} := 0$. Berechne $r^0 = b - Ax^0$. Für $k = 1, 2, ...$, falls $r^{k-1} \neq 0$:

$$p^{k-1} = r^{k-1} + \beta_{k-2} p^{k-2}, \;\; wobei \;\; \beta_{k-2} = \frac{\langle r^{k-1}, r^{k-1} \rangle}{\langle r^{k-2}, r^{k-2} \rangle} \;\; mit \;\; (k \geq 2), \tag{11a}$$

$$x^k = x^{k-1} + \alpha_{k-1} p^{k-1}, \;\; mit \;\; \alpha_{k-1} = \frac{\langle r^{k-1}, r^{k-1} \rangle}{\langle p^{k-1}, Ap^{k-1} \rangle} \tag{11b}$$

$$r^k = r^{k-1} - \alpha_{k-1} A p^{k-1} \tag{11c}$$