Symmetrische Gleichungssysteme
Das Verfahren konjugierter Gradienten

Lineares Gleichungssystem

$$A\underline{x} = \underline{f}, \qquad \sum_{j=1}^{n} A[i,j]x_j = f_i \quad \text{für } i = 1, \ldots, n$$

Voraussetzungen

- Matrix A sei symmetrisch: $A[i,j] = A[j,i]$
- Matrix A sei positiv definit: $(A\underline{u}, \underline{u}) > 0$ für alle $\underline{u} \in \mathbb{R}^n$

Definition: Ein System $\{\underline{p}^k\}_{k=0}^{n-1}$ von n Vektoren $\underline{p}^k \in \mathbb{R}^n$ heißt A–orthogonal oder konjugiert, falls gilt:

$$(A\underline{p}^k, \underline{p}^j) = (\underline{p}^k, A\underline{p}^j) = 0 \quad \text{für } k \neq j, \quad (A\underline{p}^k, \underline{p}^k) > 0$$

Lineares Gleichungssystem

$$A\underline{x} = \underline{f}$$

Darstellung des Lösungsvektors

$$\underline{x} = \sum_{k=1}^{n} x_k \underline{e}^k$$

mit Einheitsvektoren $\underline{e}^k \in \mathbb{R}^n$ oder

$$\underline{x} = \underline{x}^0 - \sum_{\ell=0}^{n-1} \alpha_\ell \underline{p}^\ell$$

mit konjugierten Vektoren $\underline{p}^\ell \in \mathbb{R}^n$ und einem beliebigen Startvektor $\underline{x}^0 \in \mathbb{R}^n$.

Bestimmung der konjugierten Vektoren $\underline{p}^\ell$

Bestimmung der Zerlegungskoeffizienten α_ℓ

Lineares Gleichungssystem

$$A\underline{x} = \underline{f}, \qquad \underline{x} = \underline{x}^0 - \sum_{\ell=0}^{n-1} \alpha_\ell \underline{p}^\ell$$

Einsetzen

$$A\underline{x}^0 - \sum_{\ell=0}^{n-1} \alpha_\ell A\underline{p}^\ell = \underline{f}$$

Skalarprodukt mit konjugierten Vektoren $\underline{p}^j$

$$\sum_{\ell=0}^{n-1} \alpha_\ell (A\underline{p}^\ell, \underline{p}^j) = (A\underline{x}^0 - \underline{f}, \underline{p}^j)$$

Zerlegungskoeffizienten

$$\alpha_\ell = \frac{(A\underline{x}^0 - \underline{f}, \underline{p}^\ell)}{(A\underline{p}^\ell, \underline{p}^\ell)} \quad \text{für } \ell = 0, \ldots, n-1$$

Definition einer Näherungslösung

$$\underline{x}^k = \underline{x}^0 - \sum_{\ell=0}^{k-1} \alpha_\ell \underline{p}^\ell, \qquad \underline{x}^n = \underline{x}, \qquad \alpha_\ell = \frac{(A\underline{x}^0 - \underline{f}, \underline{p}^\ell)}{(A\underline{p}^\ell, \underline{p}^\ell)}$$

Rekursive Definition

$$\begin{aligned} \underline{x}^{k+1} &= \underline{x}^0 - \sum_{\ell=0}^{k} \alpha_\ell \underline{p}^\ell \\ &= \underline{x}^0 - \sum_{\ell=0}^{k-1} \alpha_\ell \underline{p}^\ell - \alpha_k \underline{p}^k = \underline{x}^k - \alpha_k \underline{p}^k, \quad \alpha_k = \frac{(A\underline{x}^0 - \underline{f}, \underline{p}^k)}{(A\underline{p}^k, \underline{p}^k)} \end{aligned}$$

Es ist

$$\begin{aligned} (A\underline{x}^0 - \underline{f}, \underline{p}^k) &= (A\underline{x}^0 - \underline{f}, \underline{p}^k) - \sum_{\ell=0}^{k-1} \alpha_\ell (A\underline{p}^\ell, \underline{p}^k) \\ &= (A\underline{x}^0 - \sum_{\ell=0}^{k-1} \alpha_\ell A\underline{p}^\ell - \underline{f}, \underline{p}^k) = (A\underline{x}^k - \underline{f}, \underline{p}^k) = (\underline{r}^k, \underline{p}^k) \end{aligned}$$

Rekursionsvorschrift

$$\underline{x}^{k+1} = \underline{x}^k - \alpha_k \underline{p}^k, \quad \alpha_k = \frac{(\underline{r}^k, \underline{p}^k)}{(A\underline{p}^k, \underline{p}^k)}$$

Residuum

$$\underline{r}^k = A\underline{x}^k - \underline{f}$$

Rekursionsvorschrift

$$\begin{aligned} \underline{r}^{k+1} &= A\underline{x}^{k+1} - \underline{f} \\ &= A(\underline{x}^k - \alpha_k \underline{p}^k) - \underline{f} \\ &= A\underline{x}^k - \underline{f} - \alpha_k A\underline{p}^k = \underline{r}^k - \alpha_k A\underline{p}^k \end{aligned}$$

Bestimmung der konjugierten Vektoren $\underline{p}^k$

System linear unabhängiger Vektoren

$$\{\underline{w}^\ell\}_{\ell=0}^{n-1}, \qquad \underline{w}^\ell = \underline{e}^{\ell+1}$$ Einheitsvektoren [Fox, Huskey, Wilkinson 1948]

System konjugierter Vektoren

$$\{\underline{p}^\ell\}_{\ell=0}^{n-1} \; : \; (A\underline{p}^\ell, \underline{p}^j) = (\underline{p}^\ell, A\underline{p}^j) = 0 \quad \text{für } \ell \neq j$$

Orhogonalisierungsverfahren nach Gram–Schmidt

$$\begin{aligned}
\underline{p}^0 &= \underline{w}^0 \\
\underline{p}^{k+1} &= \underline{w}^{k+1} - \sum_{\ell=0}^{k} \beta_{k\ell}\underline{p}^\ell \quad \text{für } k = 0, \ldots, n-2 \\
0 \stackrel{!}{=} (A\underline{p}^{k+1}, \underline{p}^j) &= (A\underline{w}^{k+1}, \underline{p}^j) - \sum_{\ell=0}^{k} \beta_{k\ell}(A\underline{p}^\ell, \underline{p}^j) \\
&= (A\underline{w}^{k+1}, \underline{p}^j) - \beta_{kj}(A\underline{p}^j, \underline{p}^j), \quad \beta_{kj} = \frac{(A\underline{w}^{k+1}, \underline{p}^j)}{(A\underline{p}^j, \underline{p}^j)}
\end{aligned}$$

Bestimmung der linear unabhängigen Ausgangsvektoren $\underline{w}^k$

Iterationsvorschrift für Suchrichtungen

$$\underline{p}^0 = \underline{w}^0, \quad \underline{p}^{k+1} = \underline{w}^{k+1} - \sum_{\ell=0}^{k} \beta_{k\ell}\underline{p}^\ell, \quad \beta_{k\ell} = \frac{(A\underline{w}^{k+1}, \underline{p}^\ell)}{(A\underline{p}^\ell, \underline{p}^\ell)}$$

Iterationsvorschrift für Näherungslösungen

$$\underline{x}^{k+1} = \underline{x}^k - \alpha_k \underline{p}^k, \quad \underline{r}^{k+1} = \underline{r}^k - \alpha_k A\underline{p}^k, \quad \alpha_k = \frac{(\underline{r}^k, \underline{p}^k)}{(A\underline{p}^k, \underline{p}^k)}$$

Es gilt

$$\begin{aligned}
(\underline{r}^{k+1}, \underline{p}^k) &= (\underline{r}^k - \alpha_k A\underline{p}^k, \underline{p}^k) = (\underline{r}^k, \underline{p}^k) - \alpha_k (A\underline{p}^k, \underline{p}^k) = 0 \\
(\underline{r}^{k+1}, \underline{p}^{k-1}) &= (\underline{r}^k, \underline{p}^{k-1}) - \alpha_k (A\underline{p}^k, \underline{p}^{k-1}) = 0
\end{aligned}$$

Durch vollständige Induktion folgt

$$(\underline{r}^{k+1}, \underline{p}^\ell) = 0 \quad \text{für } \ell = 0, \ldots, k$$

Orthogonalität

$$(\underline{r}^{k+1}, \underline{p}^{\ell}) \;=\; 0 \quad \text{für } \ell = 0, \ldots, k$$

Iterationsvorschrift für Suchrichtungen

$$\underline{p}^0 = \underline{w}^0, \quad \underline{p}^{k+1} = \underline{w}^{k+1} - \sum_{\ell=0}^{k} \beta_{k\ell} \underline{p}^{\ell}, \quad \beta_{k\ell} = \frac{(A\underline{w}^{k+1}, \underline{p}^{\ell})}{(A\underline{p}^{\ell}, \underline{p}^{\ell})}$$

Dann gilt

$$(\underline{r}^{k+1}, \underline{w}^{\ell}) = (\underline{r}^{k+1}, \underline{p}^{\ell} + \sum_{j=0}^{\ell-1} \beta_{\ell-1,j} \underline{p}^{j}) = (\underline{r}^{k+1}, \underline{p}^{\ell}) + \sum_{j=0}^{\ell-1} \beta_{\ell-1,j} (\underline{r}^{k+1}, \underline{p}^{j}) = 0$$

Orthogonalität

$$(\underline{r}^{k+1}, \underline{w}^{\ell}) \;=\; 0 \quad \text{für } \ell = 0, \ldots, k$$

Die Vektoren $\underline{w}^0, \ldots, \underline{w}^k, \underline{r}^{k+1}$ sind linear unabhängig: $\underline{w}^{k+1} = \underline{r}^{k+1}$

Initialisierung

$$\underline{x}^0, \qquad \underline{r}^0 = A\underline{x}^0 - \underline{f}, \qquad \underline{p}^0 = \underline{r}^0$$

Iterationsvorschrift

$$\underline{x}^{k+1} = \underline{x}^k - \alpha_k \underline{p}^k, \quad \underline{r}^{k+1} = \underline{r}^k - \alpha_k A\underline{p}^k, \quad \alpha_k = \frac{(\underline{r}^k, \underline{p}^k)}{(A\underline{p}^k, \underline{p}^k)}$$

$$\underline{p}^{k+1} = \underline{r}^{k+1} - \sum_{\ell=0}^{k} \beta_{k\ell} \underline{p}^\ell, \quad \beta_{k\ell} = \frac{(A\underline{r}^{k+1}, \underline{p}^\ell)}{(A\underline{p}^\ell, \underline{p}^\ell)}$$

Orthogonalitäten

$$(\underline{r}^{k+1}, \underline{p}^\ell) = 0, \quad (\underline{r}^{k+1}, \underline{w}^\ell) = (\underline{r}^{k+1}, \underline{r}^\ell) = 0 \quad \text{für } \ell = 0, \ldots, k$$

Dann gilt

$$(\underline{r}^k, \underline{p}^k) = (\underline{r}^k, \underline{r}^k - \sum_{\ell=0}^{k-1} \beta_{k-1,\ell} \underline{p}^\ell) = (\underline{r}^k, \underline{r}^k) - \sum_{\ell=0}^{k-1} \beta_{k-1,\ell} (\underline{r}^k, \underline{p}^\ell) = (\underline{r}^k, \underline{r}^k)$$

Iterationsvorschrift

$$\underline{r}^{k+1} = \underline{r}^k - \alpha_k A\underline{p}^k, \quad \alpha_k = \frac{(\underline{r}^k, \underline{r}^k)}{(A\underline{p}^k, \underline{p}^k)} > 0, \quad A\underline{p}^k = \frac{1}{\alpha_k}[\underline{r}^k - \underline{r}^{k+1}]$$

Suchrichtungen

$$\underline{p}^{k+1} = \underline{r}^{k+1} - \sum_{\ell=0}^{k} \beta_{k\ell}\underline{p}^\ell, \quad \beta_{k\ell} = \frac{(A\underline{r}^{k+1}, \underline{p}^\ell)}{(A\underline{p}^\ell, \underline{p}^\ell)}$$

Berechnung von $\beta_{k\ell}$

$$\begin{aligned} (A\underline{r}^{k+1}, \underline{p}^\ell) &= (\underline{r}^{k+1}, A\underline{p}^\ell) \\ &= \frac{1}{\alpha_\ell}(\underline{r}^{k+1}, \underline{r}^\ell - \underline{r}^{\ell+1}) = \begin{cases} 0 & \text{für } \ell < k, \\ -\dfrac{1}{\alpha_k}(\underline{r}^{k+1}, \underline{r}^{k+1}) & \text{für } \ell = k \end{cases} \end{aligned}$$

Damit ist

$$\beta_{k\ell} = 0 \quad \text{für } \ell < k, \quad \beta_{kk} = -\frac{(\underline{r}^{k+1}, \underline{r}^{k+1})}{\alpha_k (A\underline{p}^k, \underline{p}^k)}$$

Suchrichtung

$$\underline{p}^{k+1} = \underline{r}^{k+1} + \beta_k \underline{p}^k, \quad \beta_k = \frac{(\underline{r}^{k+1}, \underline{r}^{k+1})}{\alpha_k (A\underline{p}^k, \underline{p}^k)}$$

Erinnerung

$$\underline{r}^{k+1} = \underline{r}^k - \alpha_k A\underline{p}^k, \quad (\underline{r}^{k+1}, \underline{p}^\ell) = 0 \quad \text{für } \ell = 0, \ldots, k$$

Wegen

$$\alpha_k (A\underline{p}^k, \underline{p}^k) = (\underline{r}^k - \underline{r}^{k+1}, \underline{p}^k) = (\underline{r}^k, \underline{p}^k) = (\underline{r}^k, \underline{r}^k + \beta_{k-1}\underline{p}^{k-1}) = (\underline{r}^k, \underline{r}^k)$$

ist schließlich

$$\underline{p}^{k+1} = \underline{r}^{k+1} + \beta_k \underline{p}^k, \quad \beta_k = \frac{(\underline{r}^{k+1}, \underline{r}^{k+1})}{(\underline{r}^k, \underline{r}^k)}$$

Iterationsvorschrift des Verfahrens konjugierter Gradienten

[Hestenes, Stiefel 1952]

Für eine beliebig gegebene Startnäherung $\underline{x}^0 \in \mathbb{R}^n$ sei $\underline{r}^0 = A\underline{x}^0 - \underline{f}$. Setze $\underline{p}^0 := \underline{r}^0$ und berechne $\varrho_0 = (\underline{r}^0, \underline{r}^0)$. Stoppe, falls $\varrho_0 < \varepsilon^2$ mit einer vorgegebenen Fehlergenauigkeit ε erreicht ist.

Berechne für $k = 0, 1, \ldots, n-2$:

$$\underline{s}^k = A\underline{p}^k, \; \sigma_k = (\underline{s}^k, \underline{p}^k), \; \alpha_k = \frac{\varrho_k}{\sigma_k}$$

$$\underline{x}^{k+1} := \underline{x}^k - \alpha_k \underline{p}^k$$

$$\underline{r}^{k+1} := \underline{r}^k - \alpha_k \underline{s}^k$$

$$\varrho_{k+1} := (\underline{r}^{k+1}, \underline{r}^{k+1})$$

Stoppe, falls $\varrho_{k+1} < \varepsilon^2 \varrho_0$ mit einer vorgegebenen Fehlergenauigkeit ε erreicht ist. Berechne andernfalls die neue Suchrichtung

$$\underline{p}^{k+1} := \underline{r}^{k+1} + \beta_k \underline{p}^k, \; \beta_k := \frac{\varrho_{k+1}}{\varrho_k}$$

Iterationsvorschrift des Verfahrens konjugierter Gradienten

- Aufwand pro Iterationsschritt
  - eine Matrix–Vektor Multiplikation $\underline{s}^k = A\underline{p}^k$
  - zwei Skalarprodukte zur Berechnung von σ_k und ϱ_{k+1}
  - drei Vektoradditionen zur Berechnung von $\underline{x}^{k+1}$, $\underline{r}^{k+1}$, $\underline{p}^{k+1}$
- Speicherbedarf
  - Matrix A, rechte Seite $\underline{f}$, Lösungsvektor $\underline{x}$
  - Residuum $\underline{r}$, Suchrichtung $\underline{p}$, Hilfsvektor $\underline{s}$
- Konvergenz
  - Berechnung der exakten Lösung $\underline{x} = \underline{x}^n$
  - Interpretation als direktes Verfahren
  - Konvergenzabschätzung
  - Wie viele Iterationsschritte sind ausreichend, um eine vorgegebene Genauigkeit zu erreichen?

Lösung

$$\underline{x} = \underline{x}^0 - \sum_{\ell=0}^{n-1} \alpha_\ell \underline{p}^\ell$$

Näherungslösung

$$\underline{x}^{k+1} = \underline{x}^0 - \sum_{\ell=0}^{k} \alpha_\ell \underline{p}^\ell$$

Fehler

$$\underline{e}^{k+1} = \underline{x}^{k+1} - \underline{x} = \sum_{\ell=k+1}^{n-1} \alpha_\ell \underline{p}^\ell$$

Es gilt

$$\begin{aligned}\|\underline{e}^{k+1}\|_A^2 = (A\underline{e}^{k+1}, \underline{e}^{k+1}) &= (A \sum_{\ell=k+1}^{n-1} \alpha_\ell \underline{p}^\ell, \sum_{j=k+1}^{n-1} \alpha_j \underline{p}^j) \\ &= \sum_{\ell=k+1}^{n-1} \sum_{j=k+1}^{n-1} \alpha_\ell \alpha_j (A\underline{p}^\ell, \underline{p}^j) = \sum_{\ell=k+1}^{n-1} \alpha_\ell^2 (A\underline{p}^\ell, \underline{p}^\ell)\end{aligned}$$

Fehler

$$\|\underline{x}^{k+1} - \underline{x}\|_A^2 = \sum_{\ell=k+1}^{n-1} \alpha_\ell^2 (A\underline{p}^\ell, \underline{p}^\ell)$$

Für

$$\underline{x} = \underline{x}^0 - \sum_{\ell=0}^{n-1} \alpha_\ell \underline{p}^\ell, \quad \underline{u} = \underline{x}^0 - \sum_{\ell=0}^{k} \gamma_\ell \underline{p}^\ell$$

ergibt sich analog

$$\|\underline{u} - \underline{x}\|_A^2 = \sum_{\ell=0}^{k} [\alpha_\ell - \gamma_\ell]^2 (A\underline{p}^\ell, \underline{p}^\ell) + \sum_{\ell=k+1}^{n-1} \alpha_\ell^2 (A\underline{p}^\ell, \underline{p}^\ell)$$

Minimierungsproblem

$$\|\underline{x}^{k+1} - \underline{x}\|_A^2 = \min_{\underline{u}=\underline{x}^0 - \sum_{\ell=0}^{k} \gamma_\ell \underline{p}^\ell} \|\underline{u} - \underline{x}\|_A^2 = \min_{\gamma_0,\ldots,\gamma_k} \|\underline{e}^0 - \sum_{\ell=0}^{k} \gamma_\ell \underline{p}^\ell\|_A^2$$

Darstellung von Residuum und Suchrichtung durch Matrix–Polynome

$$\underline{r}^k = \varphi_k(A)\underline{r}^0, \quad \underline{p}^k = \psi_k(A)\underline{r}^0$$

Initialisierung

$$\underline{r}^0 = A\underline{x}^0 - \underline{f} = \varphi_0(A)\underline{r}^0, \quad \underline{p}^0 = \underline{r}^0 = \psi_0(A)\underline{r}^0, \quad \varphi_0(A) = \psi_0(A) = I$$

Iterationsvorschriften

$$\begin{aligned}
\underline{r}^{k+1} &= \underline{r}^k - \alpha_k A\underline{p}^k \\
&= \varphi_k(A)\underline{r}^0 - \alpha_k A\psi_k(A)\underline{r}^0 = [\varphi_k(A) - \alpha_k A\psi_k(A)]\underline{r}^0 = \varphi_{k+1}(A)\underline{r}^0 \\
\underline{p}^{k+1} &= \underline{r}^{k+1} + \beta_k\underline{p}^k \\
&= \varphi_{k+1}(A)\underline{r}^0 + \beta_k\underline{\psi}_k(A)\underline{r}^0 = [\varphi_{k+1}(A) + \beta_k\psi_k(A)]\underline{r}^0 = \psi_{k+1}(A)\underline{r}^0
\end{aligned}$$

Dann ist

$$\sum_{\ell=0}^{k} \gamma_\ell \underline{p}^\ell = \sum_{\ell=0}^{k} \gamma_\ell \psi_\ell(A)\underline{r}^0 = q_k(A)\underline{r}^0$$

mit einem Matrix–Polynom $q_k(A)$ mit Polynomgrad k

Darstellung des Residuums

$$\underline{r}^k = A\underline{x}^k - \underline{f} = A\underline{x}^k - A\underline{x} = A(\underline{x}^k - \underline{x}) = A\underline{e}^k$$

Minimierungsproblem

$$\begin{aligned}
\|\underline{x}^{k+1} - \underline{x}\|_A^2 &= \min_{\gamma_0,\ldots,\gamma_k} \|\underline{e}^0 - \sum_{\ell=0}^{k} \gamma_\ell \underline{p}^\ell\|_A^2 = \min_{q_k(A)} \|\underline{e}^0 - q_k(A)\underline{r}^0\|_A^2 \\
&= \min_{q_k(A)} \|(I - q_k(A)A)\underline{e}^0\|_A^2 = \min_{p_{k+1}(A),p_{k+1}(0)=1} \|p_{k+1}(A)\underline{e}^0\|_A^2
\end{aligned}$$

Fehlerabschätzung

$$\begin{aligned}
\|\underline{x}^{k+1} - \underline{x}\|_A &= \min_{p_{k+1}(A),p_{k+1}(0)=1} \|p_{k+1}(A)\underline{e}^0\|_A \\
&\le \min_{p_{k+1}(A),p_{k+1}(0)=1} \max_{j=1,\ldots,n} |p_{k+1}(\lambda_j(A))|\, \|\underline{e}^0\|_A \\
&\le \min_{p_{k+1}(A),p_{k+1}(0)=1} \max_{\lambda\in[\lambda_{\min}(A),\lambda_{\max}(A)]} |p_{k+1}(\lambda)|\, \|\underline{e}^0\|_A \\
&= \max_{\lambda\in[\lambda_{\min}(A),\lambda_{\max}(A)]} |\widetilde{T}_{k+1}(\lambda)|
\end{aligned}$$

Satz

$$\|\underline{x}^k - \underline{x}\|_A \le \frac{2q^k}{1+q^{2k}} \|\underline{x}^0 - \underline{x}\|_A$$

mit

$$q = \frac{\sqrt{\kappa_2(A)}+1}{\sqrt{\kappa_2(A)}-1}, \quad \kappa_2(A) = \|A\|_2 \|A^{-1}\|_2 = \frac{\lambda_{\max}(A)}{\lambda_{\min}(A)}$$

Krylov Raum

$$\begin{aligned} S_k(A, \underline{r}^0) &= \text{span}\{\underline{p}^0, \underline{p}^1, \ldots, \underline{p}^k\} \\ &= \text{span}\{\underline{r}^0, A\underline{r}^0, A^2\underline{r}^0, \ldots, A^k\underline{r}^0\} \end{aligned}$$

Krylov Raum Verfahren