

Iteratives Lösen linearer Gleichungssysteme - Übung 4

Tobias Danczul

11. Mai 2017

1. Beispiel

Geg.: $A \in \mathbb{R}^{n \times n}$ invertierbar.

a)


Z.z.: x_n aus dem n -ten Schritt des MINRES Algorithmus ist die exakte Lösung x_* von $Ax = b$.

Beweis: Nach (8.1b) gilt: $A^{-1}r_0 \in \mathcal{K}_n = \text{span}\{r_0, Ar_0, \dots, A^{n-1}r_0\}$. Nach (10.3b) gilt außerdem für jeden Schritt des MINRES Algorithmus:

$$\|r_m\|_2 = \|b - Ax_m\|_2 = \min_{x \in x_0 + \mathcal{K}_m} \|b - Ax\|_2 = \min_{r \in r_0 + A\mathcal{K}_m} \|r\|_2$$

Also gilt

$$\begin{aligned} A^{-1}r_0 \in \mathcal{K}_n &\iff r_0 \in A\mathcal{K}_n \\ &\iff 0 \in r_0 + A\mathcal{K}_n \\ &\iff \min_{r \in r_0 + A\mathcal{K}_n} \|r\|_2 = 0 \\ &\iff x_n = x_* \end{aligned}$$

womit die Behauptung folgt. 


b)

Z.z.: Es gelte für $m \leq n$ $\mathcal{K}_m = \mathcal{K}_n$. Daraus folgt

$$x_m = x_{m+1} = \dots = x_*$$

Beweis: Per Definition gilt: $\mathcal{K}_m \subset \mathcal{K}_{m+1} \subset \dots \subset \mathcal{K}_n$ damit folgt aus $\mathcal{K}_m = \mathcal{K}_n$ die Gleichheit aller Krylovräume von m bis n . Also gilt

$$\min_{r \in r_0 + A\mathcal{K}_m} \|r\|_2 = \min_{r \in r_0 + A\mathcal{K}_{m+1}} \|r\|_2 = \dots = \min_{r \in r_0 + A\mathcal{K}_n} \|r\|_2 \stackrel{a)}{=} 0$$

Und damit die Gleichheit der $x_m = \dots = x_n$. 

5. Beispiel

Für GMRES benötigt man die QR -Zerlegung der Matrix \bar{H}_m . Diese kann für das upgedatete \bar{H}_{m+1} auch mittels Update der Zerlegung aus dem vorherigen Schritt bestimmt werden:

Wir gehen von einer Zerlegung für den Schritt m aus

$$Q_m^T \bar{H}_m = \bar{R}_m = \begin{pmatrix} R_m \\ 0 \end{pmatrix}$$

mit einer oberen $m \times m$ Dreiecksmatrix R_m und einer orthogonalen $m + 1 \times m + 1$ Matrix Q_m . Die nächste Matrix ist

$$\bar{H}_{m+1} = \begin{pmatrix} \bar{H}_m & h_{m+1} \\ 0 & h_{m+2,m+1} \end{pmatrix}$$

Wir wählen den Ansatz

$$\tilde{Q}_{m+1}^T \bar{H}_{m+1} = \begin{pmatrix} Q_m^T & 0 \\ 0 & 1 \end{pmatrix} \bar{H}_{m+1} = \begin{pmatrix} R_m & r_{m+1} \\ 0 & \rho \\ 0 & \sigma \end{pmatrix}$$

Damit dies eine QR -Zerlegung liefert muss $\sigma = 0$ erfüllt sein. Dies realisieren wir mit einer Givens-Rotation G_m , die auf der rechten Seite den Eintrag σ eliminieren und dabei nur ρ verändern kann und außerdem die Orthogonalität auf der linken Seite erhält.

$$G_m = \begin{pmatrix} I_m & 0 & 0 \\ 0 & c_m & s_m \\ 0 & -s_m & c_m \end{pmatrix}$$

mit

$$c_m = \frac{\rho}{\sqrt{\rho^2 + \sigma^2}} \quad \text{und} \quad s_m = \frac{\sigma}{\sqrt{\rho^2 + \sigma^2}}$$

Damit ist

$$Q_{m+1}^T = G_m \tilde{Q}_{m+1}^T$$

also insbesondere orthogonal und

$$\bar{R}_{m+1} = \begin{pmatrix} R_m & r_{m+1} \\ 0 & \sqrt{\rho^2 + \sigma^2} \\ 0 & 0 \end{pmatrix}$$



Iteratives Lösen linearer Gleichungssysteme - Übung 4

Michael Neunteufel

16. Mai 2017

2. Beispiel


Mit den Bezeichnungen aus dem Skriptum für die Arnoldi-Iteration und den Voraussetzungen

$$h_{j+1,j} \neq 0, \quad j = 1, \dots, m-1$$

soll gezeigt werden:

a)

Z.z.: $\mathcal{K}_m = \text{span}\{v_1, \dots, v_m\}$ und $\dim \mathcal{K}_m = m$.

Beweis: Aus $h_{m,m-1} \neq 0$ folgt, dass $m-1$ Schritte durchgeführt wurden, also dass v_1, \dots, v_m ausgerechnet wurden, die eine ONB für $\mathcal{K}_m = \text{span}\{v_1, \dots, v_m\}$ bilden und woraus $\dim \mathcal{K}_m = m$ folgt. 

b)

Z.z.: $\text{rank} \bar{H}_m = m$.

Beweis: Nach Theorem 9.1 gilt

$$H_m = V_m^T A V_m$$

mit Matrizen mit vollem Rang $V_m \in \mathbb{R}^{m \times m}$, $A \in \mathbb{R}^{n \times n}$, woraus $\text{rank} H_m = m$ folgt. Für die Matrix $\bar{H}_m \in \mathbb{R}^{m+1 \times m}$, gilt (als Erweiterung der Matrix H_m), dass

$$\text{rank} \bar{H}_m = m$$



c)

Z.z.: $h_{m+1,m} = 0 \implies A^m r_0 \in \mathcal{K}_m$

Gemäß der Arnoldi-Iteration gilt:

$$0 = h_{m+1,m} = \|w_m\|_2 \quad \text{und damit} \quad w_m = 0$$

Man folgert unmittelbar:

$$\begin{aligned} 0 = w_m &= A v_m - \sum_{i=1}^m h_{im} v_i \\ &= A v_m - \sum_{i=1}^m \langle A v_m, v_i \rangle v_i \end{aligned}$$

Also ist $A v_m \in [v_1, \dots, v_m] = \mathcal{K}_m$.

Z.z.: $\mathcal{K}_m = \mathcal{K}_{m+1} = \dots = \mathcal{K}_n$

Wir zeigen die Behauptung mit Induktion.

IA: $\mathcal{K}_m = \mathcal{K}_{m+1}$ ist erfüllt gemäß obiger Rechnung

IV: Es gelte $\mathcal{K}_j = \mathcal{K}_{j+1}$

IS: Z.z.: $\mathcal{K}_{j+1} = \mathcal{K}_{j+2}$

$$\mathcal{K}_{j+2} = [\mathcal{K}_{j+1}, A^{j+1}r_0] = [\mathcal{K}_{j+1}, AA^j r_0] = [\mathcal{K}_{j+1}, A\mathcal{K}_{j+1}] \stackrel{\text{IV}}{=} [\mathcal{K}_{j+1}, A\mathcal{K}_j] = [\mathcal{K}_{j+1}, \mathcal{K}_{j+1}] = \mathcal{K}_{j+1}$$



6. Beispiel

a)

Sei $\Re(A) \geq \gamma > 0$.

Z.z.: Der 're-started' GMRES(m) Algorithmus konvergiert für beliebiges $m \geq 1$.

Beweis: Für das Residuum im m -ten Schritt des ursprünglichen GMRES Algorithmus gilt (10.15)

$$\|r_m\|_2 \leq \sqrt{1 - \frac{\gamma^2}{\|A\|_2^2}} \|r_{m-1}\|_2.$$

Die Konstante der Abschätzung ist < 1 und unabhängig von der Anzahl der bereits durchgeführten Schritte, damit gilt nach m Schritten (und beliebig vielen Restarts dazwischen)

$$\|r_m\|_2 \leq \left(\sqrt{1 - \frac{\gamma^2}{\|A\|_2^2}} \right)^m \|r_0\|_2.$$

Da die Konstante für $m \rightarrow \infty$ gegen 0 geht, konvergiert das Verfahren.

Anmerkung: Die Abschätzung ist also sowohl für das ursprüngliche als auch für das restarted Verfahren gleich, allerdings ist bei zweiterem nicht gesichert, dass das Verfahren nach höchstens n (= Größe von A) Schritten die exakte Lösung liefert. Es ist nicht einmal gesichert, dass je die exakte Lösung erreicht wird, nur dass man der Lösung beliebig nahe kommt.



b)

Seien A, W SPD und es gelte

$$aW \leq A \leq bW.$$

Z.z.:

$$\kappa_\sigma(W^{-1}A) := \frac{\lambda_{\max}(W^{-1}A)}{\lambda_{\min}(W^{-1}A)} \leq \frac{b}{a}$$

Beweis:

$$\begin{aligned} \sigma(W^{-1}A) &= \sigma(W^{\frac{1}{2}}W^{-1}AW^{-\frac{1}{2}}) \\ &= \sigma(W^{-\frac{1}{2}}AW^{-\frac{1}{2}}) \end{aligned}$$

Die letzte Matrix ist offensichtlich bezüglich $(\cdot, \cdot)_2$ symmetrisch, über das die größten und kleinsten Eigenwerte gefunden werden können:

$$\begin{aligned} (W^{-\frac{1}{2}}A \underbrace{W^{-\frac{1}{2}}x}_{=:y}, x) &= (Ay, W^{-\frac{1}{2}}x) = (Ay, y) \geq a(Wy, y) = a(WW^{-\frac{1}{2}}x, W^{-\frac{1}{2}}x) = a(x, x) \\ &\leq b(Wy, y) = b(WW^{-\frac{1}{2}}x, W^{-\frac{1}{2}}x) = b(x, x) \end{aligned}$$

Also ist $\lambda_{\max} \leq b$ und $\lambda_{\min} \geq a$.

Z.z.:

$$\kappa_\sigma(W^{-1}A) = \kappa_A(W^{-1}A)$$

Beweis: Es gilt

$$(W^{-1}Ax, x)_A = (AW^{-1}Ax, x) = (x, AW^{-1}Ax) = (Ax, W^{-1}Ax) = (x, W^{-1}Ax)_A$$

also ist $W^{-1}A$ selbstadjungiert bezüglich $(\cdot, \cdot)_A$ und die Konditionszahl κ_A wird über die Eigenwerte bestimmt.



**Iterative Lösung großer Gleichungssysteme,
Übung 4, Aufgabe 3
Lukas Kogler**

Algorithm 1 Block Arnoldi

```

1: procedure BLOCK_ARNOLDI( $A, p$ )
2:   Wähle eine orthogonale Matrix  $V_1 \in \mathbb{R}^{n \times p}$ 
3:   for  $j = 1 \dots m$  do
4:     Berechne  $H_{ij} := V_i^T AV_j$  für  $i = 1 \dots j$ 
5:      $W_j := AV_j - \sum_{i=1}^j V_i H_{ij}$ 
6:     Berechne die reduziertet QR-Zerlegung  $W_j = V_{j+1} H_{j+1,j}$ 
7:   end for
8: end procedure

```

Außerdem wird folgende Notation benutzt:

- Die aneinandergereihten V_i :

$$U_m := [V_1, V_2, \dots, V_m]$$

- Die letzten p Spalten der m -dimensionalen Einheitsmatrix I_m :

$$E_m := [0, I_p] \cdot I_m \in \mathbb{R}^{m \times p}$$

- Die zusammengefügteten H_{ij} :

$$H_m = (H_{i,j})_{1 \leq i, j \leq m}, \quad H_{i,j} = 0 \quad i > j + 1$$

Zu beachten ist, dass H_m eine obere Block-Dreiecks-Matrix aus Blöcken der Größe p ist, wobei eine zusätzliche untere "Block-Diagonale" besetzt ist, wobei die "oberen" Blöcke aus der Orthogonalisierung (Zeile 4 im Algorithmus) und die "unteren" Blöcke aus der QR-Zerlegung (Zeile 6) stammen, also selbst obere Dreiecks-Matrizen sind (weshalb H_m einfach obere Dreiecks-Matrix mit p zusätzlichen besetzten Diagonalen ist):

$$H_m = \begin{pmatrix} H_{1,1} & H_{1,2} & \dots & \dots & H_{1,m} \\ H_{2,1} & H_{2,2} & & & H_{2,m} \\ 0 & H_{3,2} & \ddots & & \vdots \\ \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & H_{m-1,m} \\ 0 & \dots & \dots & 0 & H_{m,m-1} & H_{m,m} \end{pmatrix}$$

Zu zeigen ist, dass algorithmus (1) eine Block-Verallgemeinerung des Arnoldi-Verfahrens ist, das heißt, dass einerseits die Spalten von U_m eine Orthogonalbasis darstellen, und dass außerdem folgende Verallgemeinerung von Formel (9.7) aus dem Skriptum gilt:

$$AU_m = U_m H_m + V_{m+1} H_{m+1,m} E_m^T$$

U_m ist orthogonal:

Zuerst wird gezeigt, dass:

$$V_i^T V_j = \begin{cases} I_p & i = j \\ 0 & \text{else} \end{cases}$$

Jedes V_j ist eine orthogonale Matrix (da sie aus einer QR-Zerlegung gewonnen wurde), also $V_j^T V_j = I_p$. Statt $V_i^T V_j = 0$ für $i \neq j$ zu zeigen, reicht es auch, $V_i^T W_j = 0$ für $i \leq j$ zu zeigen, da $V_i^T V_{j+1} = V_i^T W_j H_{j+1,j}^{-1}$ und $H_{j+1,j}$ als obere Dreiecks-Matrix vollen Rang hat und da man durch Transponieren den Fall $i > j$ bekommt. Dies lässt sich mit einem Induktionsargument zeigen:

z.Z: $V_i^T W_j = 0 \quad i \leq j$

IA:

$$V_1^T W_1 = V_1^T A V_1 - V_1 H_{1,1} = V_1^T A V_1 - \underbrace{V_1^T V_1}_{=I_p} V_1^T A V_1 = 0 \quad \text{☺}$$

IS:

$$\begin{aligned} V_i^T W_j &= V_i^T A V_j - \sum_{l=1}^j V_i^T V_l H_{l,j} \\ &= V_i^T A V_j - \sum_{l=1}^j \underbrace{V_i^T V_l}_{=\delta^{il} I_p} V_l^T A V_j \\ &= V_i^T A V_j - V_i^T A V_j = 0 \quad \text{☺} \end{aligned}$$

Nun gilt (wobei die Indices für $p \times p$ Unter-blöcke stehen):

$$U_m^T U_m = (V_i^T V_j)_{ij} = \delta^{ij} I_p = I_{m \cdot p}$$

U_m ist also orthogonal! ☺

Die Identität:

$$\begin{aligned}
 AV_j &= (I_n + \sum_{i=1}^m V_i V_i^T - \sum_{i=1}^m V_i V_i^T) AV_j = \\
 &= \sum_{i=1}^m V_i V_i^T AV_j + (I_n - \sum_{i=1}^m V_i V_i^T) AV_j = \\
 &= \sum_{i=1}^m V_i H_{ij} + \underbrace{(I_n - \sum_{i=1}^m V_i^T V_i)}_{=W_m} AV_j = \\
 &= \sum_{i=1}^m V_i H_{ij} + V_{j+1} H_{j+1,j} = \\
 &= U_{m+1} H_{:,j}^{(m+1)}
 \end{aligned}$$

Dabei steht $H_{:,j}^{(m+1)}$ für die j Spalten von H_{m+1} , die zum j -ten Block gehören, und der Index $m+1$ aus Bequemlichkeit oben geschrieben wurde.

$$H_{:,j}^{(m+1)} = \begin{pmatrix} H_{1,j} \\ \vdots \\ H_{j,j} \\ H_{j+1,j} \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Nun fügen wir dieses Resultat noch zusammen, um die gewünschte Identität für AU_m zu erhalten:

$$\begin{aligned}
 AU_m &= [AV_1, AV_2, \dots, AV_m] = \\
 &= [U_{m+1} H_{:,1}^{(m+1)}, U_{m+1} H_{:,2}^{(m+1)}, \dots, U_{m+1} H_{:,m}^{(m+1)}] = \\
 &= U_{m+1} H_{1\dots m}^{(m+1)}
 \end{aligned}$$

$H_{1\dots m}^{(m+1)}$ steht (in konsistenter Weise) für die ersten m Spalten von H_{m+1} .

$$\begin{aligned}
 U_{m+1} H_{1\dots m}^{(m+1)} &= (U_m \mid V_{m+1}) \cdot \begin{pmatrix} H_m \\ 0 \dots 0 \ H_{m+1,m} \end{pmatrix} = \\
 &= (U_m \mid V_{m+1}) \cdot \begin{pmatrix} H_m \\ H_{m+1,m} E_m^T \end{pmatrix} = \\
 &= U_m H_m + V_{m+1} H_{m+1,m} E_m^T
 \end{aligned}$$

■

4. Beispiel

c)

Für die Poisson-Matrix

$$P = \begin{pmatrix} 2 & -1 & & \\ -1 & & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & 2 \end{pmatrix}$$

ist der maximale Eigenwert gegeben durch $\lambda_{max} = 4 \sin^2\left(\frac{n\pi}{2(n-1)}\right) \approx 3.99999$. Für den Realteil der Matrix A gilt:

$$\Re A = \frac{1}{2} \begin{pmatrix} 2 & 1 & & \\ 1 & & \ddots & \\ & \ddots & \ddots & 1 \\ & & 1 & 2 \end{pmatrix} = \frac{1}{2}(-P + 4I)$$

Damit folgt:

$$\lambda_{min}(\Re A) = \frac{1}{2}(-\lambda_{max}(P) + 4) \approx 5 \cdot 10^{-6}$$

Wird die Diagonale auf 2 gesetzt, folgt analog:

$$\Re A = \frac{1}{2} \begin{pmatrix} 4 & 1 & & \\ 1 & & \ddots & \\ & \ddots & \ddots & 1 \\ & & 1 & 4 \end{pmatrix} = \frac{1}{2}(-P + 6I)$$

Und damit $\lambda_{min}(\Re A) = \frac{1}{2}(-\lambda_{max}(P) + 6) \approx 1.000004$.

