Fachbereich Mathematik
Prof. Dr. P. Spellucci
Wintersemester 2007/2008

TECHNISCHE UNIVERSITÄT DARMSTADT

14.1.2008

Numerik des Matrizeneigenwertproblems
Übung 11

Präsenzübung

Ü 31 Zeigen Sie, daß das LANCZOS-Verfahren angewandt auf die Matrix A mit dem Startvektor $x^{(0)}$ dieselbe Tridiagonalmatrix T liefert wie für die Matrix $Q^T AQ$ mit dem Startvektor $Q^T x^{(0)}$, falls Q unitär ist. Was folgt daraus für die theoretische Analyse des Verfahrens?

Ü 32 Beweis von Satz 1.8.2
Zeigen Sie:
Sei V_j eine orthonormierte Eigenvektormatrix von T_j (Tridiagonalmatrix aus dem Lanczos–Verfahren für die Matrix A):

$$V_j^T T_j V_j = diag(\Theta_1, \ldots, \Theta_j),$$

und Y_j sei definiert durch

$$Y_j = Q_j V_j = (y_1, \ldots, y_j).$$

Dann gilt

$$\|Ay_i - \Theta_i y_i\|_2 = |\beta_j|\ |v_{ji}|.$$

Hinweis: Fassen Sie die Rekursion des Lanczos-Verfahrens in Matrixschreibweise zusammen, um AQ_j zu berechnen.

Ü 33 Sei

$$A = \begin{pmatrix} 3 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

Berechnen Sie mit Startvektor $x_0 = (1, 1, 1)^T$ und dem Lanczos-Verfahren die symmetrische Tridiagonalmatrix T_3.

Hausübung

H 31 Zeigen Sie: Das Lanczos-Verfahren, angewandt auf die Matrix A und die Matrix $A - \sigma I$ bei beliebigem σ erzeugt die selbe Matrix Q_k, wenn der gleiche Startvektor benutzt wird.

H 32 Gegeben sei eine symmetrische Matrix $A \in \mathbb{R}^{n\times n}$ mit einem einfachen Eigenwert λ_1 und dem zugehörigen Eigenvektor r_1. Zeigen Sie, daß die tridiagonale Matrix T aus der LANCZOS-Zerlegung nach Abbruch mit $\beta_{k+1} = 0$ keinen Eigenwert gleich λ_1 hat, wenn der Startvektor $x^{(0)}$ senkrecht auf r_1 steht.
Hinweis: Verwenden Sie das Resultat der Aufgabe **Ü 31** und beachten Sie, daß A unitär diagonalisierbar ist.

H 33 Wegen des Verlustes der Orthogonalität beim Lanczos-Verfahren bei gerundetem Rechnen ist es üblich, die $q^{(i)}$–Vektoren (wenigstens partiell) zu orthogonalisieren. Wir betrachten hier den Fall der vollständigen Orthogonalisierung von $q^{(i)}$ bezüglich $q^{(1)}, \ldots, q^{(i-1)}$. Zeigen Sie, daß bei exakter Rechnung der folgende Algorithmus die gleiche Q_k–Matrix erzeugt wie der Standard-Lanczos-Algorithmus, daß aber unter Rundungseinfluss die von ihm erzeugten Vektoren bis auf Maschinengenauigkeit orthogonal bleiben. Benutzen Sie ohne Beweis, daß die Spalten von gerundet berechneten Produkten von Householdermatrizen numerisch bis auf Maschinengenauigkeit orthogonal bleiben. Im Folgenden bezeichnet H_k eine Householdermatrix, die durch die Transformationseigenschaft bezüglich eines spezifizierten Vektors festgelegt wird und die letzten $n - k$ Komponenten dieses Vektors in null überführt und $e_1, \ldots, e_{k-1}$ invariant lässt.

```
beta_0 = 0 ;  q_0 = 0 ;
q_1  ist der gegebene normierte Startvektor.
H_1 q_1 = e_1 ;

for k=1,...,n-1
   alpha_k = q_k'Aq_k ;
   r_{k+1}  = (A- alpha_{k}I) q_{k} - beta_{k-1} q_{k-1} ;
   w = (H_k * .... * H_1) r_{k+1} ;
   H_{k+1} w = ( w_{k,1},...,w{k,k},beta_k,0,...,0)' ;
   q_{k+1} = H_1*....*H_{k+1} e_{k+1};
end
alpha_n = q_n'Aq_n;
```

Numerik des Matrizeneigenwertproblems
Übung 11, Lösungsvorschlag

Präsenzübung

Ü 31 Zeigen Sie, daß das Lanczos-Verfahren angewandt auf die Matrix A mit dem Startvektor $x^{(0)}$ dieselbe Tridiagonalmatrix T liefert wie für die Matrix Q^TAQ mit dem Startvektor $Q^Tx^{(0)}$, falls Q unitär ist. Was folgt daraus für die theoretische Analyse des Verfahrens?

Das Lanczos-Verfahren für die Matrix und dem Startvektor $x^{(0)}$ lautet:

Gegeben der Startvektor $x^{(0)} \neq 0$. Dann setze $\beta_1 = \|x^{(0)}\|$, $r^1 = x^{(0)}$ und $q^0 = 0$.

Setze $j = 1$ und berechne:

Schritt 1 : $q^j = \frac{r^j}{\beta_j}$.

Schritt 2 : $u^j = Aq^j - \beta_j q^{j-1}$.

Schritt 3 : $\alpha_j = (u^j)^T q^j$.

Schritt 4 : $r^{j+1} = u^j - \alpha_j q^j$.

Schritt 5 : $\beta_{j+1} = \|r^{j+1}\|$.

Schritt 6 : *Falls kein Abbruch wegen $\beta_{j+1} = 0$ erfolgt, setze $j = j + 1$ und gehe zu Schritt 1.*

Es ist zu zeigen, daß der Lanczos-Algorithmus für die Matrix Q^TAQ und den Startvektor $Q^Tx^{(0)}$ die gleichen Skalare α_j und β_j liefert.

Die Behauptung wird induktiv gezeigt. (Die Vektoren und Skalare ohne Tilde bezeichnen die entsprechenden Werte aus dem obigen Algorithmus.)

Gegeben der Startvektor $Q^Tx^{(0)} \neq 0$. Dann gilt $\tilde{\beta}_1 = \|Q^Tx^{(0)}\| = \|x^{(0)}\| = \beta_1$, $\tilde{r}^1 = Q^Tx^{(0)} = Q^Tr^1$ und $\tilde{q}^0 = 0$.

Im ersten Schritt ergibt sich

Schritt 1 : $\tilde{q}^1 = \frac{\tilde{r}^1}{\tilde{\beta}_1} = Q^T\frac{r^1}{\beta_1} = Q^Tq^1$.

Schritt 2 : $\tilde{u}^1 = Q^TAQ\tilde{q}^1 = Q^TAQQ^Tq^1 = Q^TAq^1 = Q^Tu^1$.

Schritt 3 : $\tilde{\alpha}_1 = (\tilde{u}^1)^T\tilde{q}^1 = (u^1)^TQQ^Tq^1 = (u^1)^Tq^1 = \alpha_1$.

Schritt 4 : $\tilde{r}^2 = \tilde{u}^1 - \tilde{\alpha}_1\tilde{q}^1 = Q^Tu^1 - \alpha_1Q^Tq^1 = Q^T(u^1 - \alpha_1q^1) = Q^Tr^2$.

Schritt 5 : $\tilde{\beta}_2 = \|\tilde{r}^2\| = \|Q^Tr^2\| = \|r^2\| = \beta_2$.

Im zweiten Schritt tritt erstmals eine Drei-Term-Rekursion auf

Schritt 1 : $\tilde{q}^2 = \frac{\tilde{r}^2}{\tilde{\beta}_2} = Q^T\frac{r^2}{\beta_2} = Q^Tq^2$.

Schritt 2 : $\tilde{u}^2 = Q^TAQ\tilde{q}^2 - \tilde{\beta}_2\tilde{q}^1 = Q^TAQQ^Tq^2 - \beta_2Q^Tq^1 = Q^T(Aq^2 - \beta_2q^1) = Q^Tu^2$.

Schritt 3 : $\tilde{\alpha}_2 = (\tilde{u}^2)^T\tilde{q}^2 = (u^2)^TQQ^Tq^2 = (u^2)^Tq^2 = \alpha_2$.

Schritt 4 : $\tilde{r}^3 = \tilde{u}^2 - \tilde{\alpha}_2 \tilde{q}^2 = Q^T u^2 - \alpha_2 Q^T q^2 = Q^T(u^2 - \alpha_2 q^2) = Q^T r^3.$

Schritt 5 : $\tilde{\beta}_3 = \|\tilde{r}^3\| = \|Q^T r^3\| = \|r^3\| = \beta_3.$

Alle weiteren Schritte folgen analog. Man kann also zur theoretischen Analyse des Verfahrens voraussetzen, daß A bereits diagonal ist.

Ü 32 Beweis von Satz 1.8.2

Zeigen Sie:

Sei V_j eine orthonormierte Eigenvektormatrix von T_j (Tridiagonalmatrix aus dem Lanczos–Verfahren für die Matrix A):

$$V_j^T T_j V_j = diag(\Theta_1, \ldots, \Theta_j),$$

und Y_j sei definiert durch

$$Y_j = Q_j V_j = (y_1, \ldots, y_j).$$

Dann gilt

$$\|Ay_i - \Theta_i y_i\|_2 = |\beta_j| \; |v_{ji}|.$$

Hinweis: Fassen Sie die Rekursion des Lanczos-Verfahrens in Matrixschreibweise zusammen, um AQ_j zu berechnen.

Die Rekursion lautet in Matrixschreibweise zusammengefasst und nach AQ_j aufgelöst:

$$AQ_j = Q_j T_j + \beta_j q_{j+1} e_j^T.$$

Damit zur berechnende Größe:

$$\begin{aligned}
\|AQ_j v_i - \Theta_i y_i\|_2 &= \|Q_j T_j v_i + \beta_j q_{j+1} e_j^T v_i - \Theta_i y_i\|_2 \\
&= \|Q_j V_j diag(\Theta_k) V_j^T v_i + \beta_j q_{j+1} v_{ji} - \Theta_i y_i\|_2 \\
&= \|Y_j \Theta_i e_i + \beta_j q_{j+1} v_{ji} - \Theta_i y_i\|_2 \\
&= |\beta_j| \; |v_{ji}|.
\end{aligned}$$

Verwendet wurde noch, dass die q_i normiert sind.

Ü 33 Sei

$$A = \begin{pmatrix} 3 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

Berechnen Sie mit Startvektor $x_0 = (1, 1, 1)^T$ und dem Lanczos-Verfahren die symmetrische Tridiagonalmatrix T_3.

Initialisierung:

$$\begin{aligned} q^{(0)} &= 0 \\ \beta_0 &= 1 \\ q^{(1)} &= \frac{x^{(0)}}{||x^{(0)}||_2} = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \end{aligned}$$

Für $i = 1$ erhält man somit

$$\begin{aligned} \alpha_1 &= {q^{(1)}}^T A q^{(1)} = \frac{8}{3} \\ r^{(2)} &= Aq^{(1)} - \alpha_1 q^{(1)} - \beta_0 q^{(0)} = \frac{1}{3\sqrt{3}} \begin{pmatrix} 4 \\ -2 \\ -2 \end{pmatrix} \\ \beta_1 &= ||r^{(2)}||_2 = \frac{2\sqrt{2}}{3} \\ q^{(2)} &= \frac{r^{(2)}}{\beta_1} = \frac{1}{\sqrt{6}} \begin{pmatrix} 2 \\ -1 \\ -1 \end{pmatrix} \end{aligned}$$

Für $i = 2$

$$\begin{aligned} \alpha_2 &= {q^{(2)}}^T A q^{(2)} = \frac{11}{6} \\ r^{(3)} &= Aq^{(2)} - \alpha_2 q^{(2)} - \beta_1 q^{(1)} = \frac{1}{2\sqrt{6}} \begin{pmatrix} 0 \\ -3 \\ 3 \end{pmatrix} \\ \beta_2 &= ||r^{(3)}||_2 = \frac{\sqrt{3}}{2} \\ q^{(3)} &= \frac{r^{(3)}}{\beta_2} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \end{aligned}$$

Für $i = 3$

$$\begin{aligned} \alpha_3 &= {q^{(3)}}^T A q^{(3)} = \frac{3}{2} \\ r^{(4)} &= Aq^{(3)} - \alpha_3 q^{(3)} - \beta_2 q^{(2)} = \frac{1}{2\sqrt{2}} \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\ \beta_3 &= ||r^{(4)}||_2 = 0 \end{aligned}$$

Man hat also

$$T_3 = \begin{pmatrix} \frac{8}{3} & \frac{2\sqrt{2}}{3} & 0 \\ \frac{2\sqrt{2}}{3} & \frac{11}{6} & \frac{\sqrt{3}}{2} \\ 0 & \frac{\sqrt{3}}{2} & \frac{3}{2} \end{pmatrix}$$

Hausübung

H 31 Zeigen Sie: Das Lanczos-Verfahren, angewandt auf die Matrix A und die Matrix $A - \sigma I$ bei beliebigem σ erzeugt die selbe Matrix Q_k, wenn der gleiche Startvektor benutzt wird.

Die Spalten von Q_k entstehen durch Normierung der Vektoren r_{k+1} und die β_k sind die Normierungsfaktoren. Es genügt also zu zeigen, daß die gleiche Vektorfolge r_j entsteht. Natürlich ist der Startvektor $q^{(1)}$ in beiden Fällen gleich. Wir erhalten formal zwei Folgen

$$\begin{aligned} r^{(k+1)} &= Aq^{(k)} - \alpha_k q^{(k)} - \beta_{k-1} q^{(k-1)} \\ \hat{r}^{(k+1)} &= (A - \sigma I)\hat{q}^{(k)} - \hat{\alpha}_k \hat{q}^{(k)} - \hat{\beta}_{k-1} \hat{q}^{(k-1)} \end{aligned}$$

mit

$$\alpha_k = R(q^k, A) \quad \textit{und} \quad \hat{\alpha}_k = R(\hat{q}^{(k)}, A - \sigma I)$$

Wegen des Induktionsanfanges $q^{(1)} = \hat{q}^{(1)}$ und

$$R(x, A) = R(x, A - \sigma I) + \sigma$$

folgt jedoch

$$r^{(2)} = \hat{r}^{(2)} \quad \textit{somit} \quad \hat{\beta}_1 = \beta_1$$

d.h.

$$q^{(2)} = \hat{q}^{(2)}$$

und induktive Fortsetzung dieses Arguments liefert die Behauptung.

H 32 Gegeben sei eine symmetrische Matrix $A \in \mathbb{R}^{n \times n}$ mit einem einfachen Eigenwert λ_1 und dem zugehörigen Eigenvektor r_1. Zeigen Sie, daß die tridiagonale Matrix T aus der Lanczos-Zerlegung nach Abbruch mit $\beta_{k+1} = 0$ keinen Eigenwert gleich λ_1 hat, wenn der Startvektor $x^{(0)}$ senkrecht auf r_1 steht.
Hinweis: Verwenden Sie das Resultat der Aufgabe **Ü 31** und beachten Sie, daß A unitär diagonalisierbar ist.

Aus dem Lanczos-Verfahren mit dem entsprechenden Startvektor ergibt sich die Tridiagonalmatrix T. Nach Aufgabe Ü31 liefert der selbe Algorithmus für $R^T AR$ ebenfalls diese Matrix T, falls der Startvektor $R^T x^{(0)}$ gewählt wurde und R unitär ist.

Wegen der unitären Diagonalisierbarkeit der Matrix A sei R die unitäre Matrix der Eigenvektoren von A, so daß

$$\tilde{A} = R^T AR = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n).$$

Der neue Startvektor $R^T x^{(0)}$ hat dann die Gestalt

$$R^T x^{(0)} = (0, *, \cdots, *)^T,$$

da $e_1^T R^T x^{(0)} = r_1^T x^{(0)} = 0$.
Es ist leicht zu sehen, daß damit alle LANCZOS-*Vektoren q^i die selbe Gestalt haben, da $\tilde{A}$ eine Diagonalmatrix ist.*

Nach dem Abbruch des Verfahren wegen $\beta_{\nu+1} = 0$ erhalten wir

$$\tilde{A}Q = QT.$$

Dabei ist $Q \in \mathbb{R}^{n\times\nu}$ und $T \in \mathbb{R}^{\nu\times\nu}$ ist die selbe Tridiagonalmatrix, die aus A mit Startvektor $x^{(0)}$ entstanden wäre. Diagonalisiert man T so ergibt sich

$$\tilde{A}Q = QT = QVDV^T$$

mit unitärem V. Setzt man $QV = \tilde{Q}$ so hat diese Matrix, ebenso wie Q, die Gestalt

$$\tilde{Q} = \begin{pmatrix} 0 & \cdots & 0 \\ \hline * & \cdots & * \\ \vdots & & \vdots \\ * & \cdots & * \end{pmatrix}.$$

Angenommen λ_1 sei doch Eigenwert von D, dann ergibt sich aber

$$\tilde{A}(\tilde{Q})_\mu = \lambda_1(\tilde{Q})_\mu$$

für ein $\mu \in \{1,\ldots,\nu\}$. Damit ist aber λ_1 ein Eigenwert von $\tilde{A}$ mit dem Eigenvektor $\tilde{Q}_\mu$. Da λ_1 aber das erste Diagonalelement von $\tilde{A}$ ist, muß e_1 auch ein Eigenvektor sein. Dies ist ein Widerspruch zu der Voraussetzung, daß λ_1 ein einfacher Eigenvektor ist.

H 33 Wegen des Verlustes der Orthogonalität beim Lanczos-Verfahren bei gerundetem Rechnen ist es üblich, die $q^{(i)}$–Vektoren (wenigstens partiell) zu orthogonalisieren. Wir betrachten hier den Fall der vollständigen Orthogonalisierung von $q^{(i)}$ bezüglich $q^{(1)},\ldots,q^{(i-1)}$. Zeigen Sie, daß bei exakter Rechnung der folgende Algorithmus die gleiche Q_k–Matrix erzeugt wie der Standard-Lanczos-Algorithmus, daß aber unter Rundungseinfluss die von ihm erzeugten Vektoren bis auf Maschinengenauigkeit orthogonal bleiben. Benutzen Sie ohne Beweis, daß die Spalten von gerundet berechneten Produkten von Householdermatrizen numerisch bis auf Maschinengenauigkeit orthogonal bleiben. Im Folgenden bezeichnet H_k eine Householdermatrix, die durch die Transformationseigenschaft bezüglich eines spezifizierten Vektors festgelegt wird und die letzten $n-k$ Komponenten dieses Vektors in null überführt und $e_1,\ldots,e_{k-1}$ invariant lässt.

```
beta_0 = 0 ;  q_0 = 0 ;
q_1  ist der gegebene normierte Startvektor.
H_1 q_1 = e_1 ;

for k=1,...,n-1
```

```
    alpha_k = q_k'Aq_k ;
    r_{k+1}  = (A- alpha_{k}I) q_{k} - beta_{k-1} q_{k-1} ;
    w = (H_k * .... * H_1) r_{k+1} ;
    H_{k+1} w = ( w_{k,1},...,w{k,k},beta_k,0,...,0)' ;
    q_{k+1} = H_1*....*H_{k+1} e_{k+1};
end
alpha_n = q_n'Aq_n;
```

Wegen

$$e_k^T e_j \;=\; 0 \quad \text{für } j \;<\; k$$

gilt auch

$$(q^{(k)})^T q^{(j)} \;=\; e_k^T H_k \cdot \ldots \cdot H_1 H_1 \cdot \ldots \cdot H_j e_j \;=\; 0$$

für $j \;<\; k$, *weil*

$$H_k \cdot \ldots \cdot H_{j+1} e_j \;=\; e_j \;.$$

Wir zeigen induktiv, daß die hier bestimmten q^j *mit denen aus dem Standard-Lanczos-Verfahren, die wir mit* $\hat{q}^{(j)}$ *bezeichnen, übereinstimmen. Nach Annahme ist*

$$q^{(1)} \;=\; \hat{q}^{(1)} \;.$$

Somit ist auch der Vektor $r^{(2)}$ *in beiden Verfahren der gleiche. Nun ist einerseits*

$$\hat{q}^{(2)} \;=\; r^{(2)}/||r^{(2)}|| \text{ und } \beta_1 \;= ||r^{(2)}||$$

und andererseits

$$q^{(2)} \;=\; H_1 H_2 e_2$$

wobei

$$H_2 H_1 r^{(2)} \;= (w_{2,1}, \beta_1, 0, \ldots, 0)^T$$

ist. Also ist

$$\begin{aligned} r^{(2)} &= H_1 H_2 (w_{2,1} e_1 + \beta_1 e_2) \\ &= H_1 H_2 (w_{2,1} e_1 + \beta_1 H_2 H_1 q^{(2)}) \\ &= H_1 w_{2,1} e_1 + \beta_1 q^{(2)} \\ &= w_{2,1} q^{(1)} + \beta_1 q^{(2)} \;. \end{aligned}$$

Wegen $(r^{(2)})^T \hat{q}^{(1)} \;=\; 0$ *und der Orthogonalität der* $q^{(j)}$ *und des Induktionsanfanges ist also (mit Multiplikation der letzten Gleichung mit* $\hat{q}^{(1)}$ *)*

$$w_{2,1} \;=\; 0$$

und daher auch

$$q^{(2)} \;=\; \hat{q}^{(2)} \;.$$

Sei nun die Behauptung gültig für $j = 1, \ldots, k$. *Dann gilt wie zuvor daß*

$$r^{k+1} = \hat{r}^{(k+1)} .$$

Somit

$$\begin{aligned}
(\prod_{j=1}^{k+1} H_j)^T r^{(k+1)} &= (w_{k+1,1}, \ldots, w_{k+1,k}, \beta_k, 0, \ldots, 0)^T \\
\hat{q}^{(k+1)} &= r^{k+1}/||r^{k+1}|| \ , \ \beta_k = ||r^{k+1}|| \\
q^{(k+1)} &= (\prod_{j=1}^{k+1} H_j) e_{k+1} \\
&= \Big(\prod_{j=1}^{k+1} H_j\Big)\Big((\prod_{j=1}^{k+1} H_j)^T r^{(k+1)} - \sum_{j=1}^{k} w_{k+1,j} e_j\Big)/\beta_k \\
&= r^{(k+1)}/\beta_k - \Big(\sum_{j=1}^{k} w_{k+1,j} (\prod_{l=1}^{k+1} H_l) e_j\Big)/\beta_k \\
&= r^{(k+1)}/\beta_k - \Big(\sum_{j=1}^{k} w_{k+1,j} (\prod_{l=1}^{j} H_l) e_j\Big)/\beta_k \\
&= r^{(k+1)}/\beta_k - \Big(\sum_{j=1}^{k} w_{k+1,j} q_j\Big)/\beta_k
\end{aligned}$$

und wegen der Induktionsannahme und der bereits bewiesenen Orthogonalität der $q^{(j)}$ *untereinander folgt wieder*

$$w_{k+1,l} = 0 \quad \text{für } l = 1 \, , \ldots , \ k$$

und damit die Behauptung. Da die $q^{(j)}$ *hier als Spalten aus Produkten von Householdermatrizen bestehen, bleiben sie bis auf Maschinengenauigkeit orthogonal.*