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

TECHNISCHE UNIVERSITÄT DARMSTADT

21.1.2008

Numerik des Matrizeneigenwertproblems
Übung 12

Präsenzübung

Ü 34 Formulieren Sie das Wielandt-Verfahren mit Shift zur Bestimmung eines Eigenvektors eines allgemeinen Eigenwertproblems

$$Ax \;=\; \lambda Bx$$

mit reell symmetrischem A und reell symmetrischem positiv definitem B. Es soll dabei weder die Inverse von B noch die Cholesky-Zerlegung von B benutzt werden. Es tritt auch hier die Aufgabe einer mehrfachen Gleichungslösung mit einer festen Matrix auf, wenn der Shift festgehalten wird. Welches Problem entsteht bei dieser Gleichungslösung? Könnte man das Problem durch eine unitäre Vortransformation a la Tridiagonalisierung vereinfachen?
Hinweis: Arbeiten Sie zunächst formal mit der Choleskyzerlegung von B (Transformation auf ein Standard-Eigenwertproblem) und übersetzen Sie dann in die Originaldaten zurück.

Ü 35 Zeigen Sie: Ist A obere Hessenbergmatrix und B invertierbare obere Dreiecksmatrix, dann gilt:

1. AB^{-1} ist obere Hessenbergmatrix.
2. Die Elemente (2,1) von $AB^{-1} - \mu I$ und von $A - \mu B$ werden durch die gleiche Ähnlichkeitstransformation mit einer Givensmatrix in null überführt.

Ü 36 (*Transformation eines nichtlineares Eigenwertproblems*)
Gegeben sei das quadratische Eigenwertproblem

$$\left(A_2\lambda^2 + A_1\lambda + A_0\right) x = 0,$$

mit regulärer Matrix A_2.

a) Transformieren Sie dieses polynomiale Problem in ein gewöhnliches lineares Eigenwertproblem

$$C\tilde{x} = \lambda\tilde{x}.$$

Hinweis: Eliminieren Sie zunächst den Matrixkoeffizienten vor λ^2 und setzen Sie dann den neuen Eigenvektor an als $\tilde{x} = (\lambda x, x)^T$. Wie lautet die Matrix C ?

b) Wie würde die Transformation bei einem allgemeinen polynomialen Eigenwertproblem

$$(p(\lambda))\, x = \left(\sum_{i=0}^{m} A_i \lambda^i \right) x = 0$$

(A_m regulär) aussehen ?

Hausübung

H 34 Zeigen Sie: Sind A und B beliebige reelle $n \times n$ Matrizen, dann existieren (komplex) unitäre Matrizen Q und Z mit

$$Q^H AZ \;=\; R_A \quad \text{und} \quad Q^H BZ \;=\; R_B$$

wobei R_A und R_B obere Dreiecksmatrizen sind. Wie stellt sich damit die Lösung des allgemeinen Eignwertproblems

$$Ax \;=\; \lambda Bx$$

dar? Warum muss man u.U. Q und Z komplex wählen, auch wenn A und B reell sind? (Die Aussage gilt natürlich auch für komplexe A, B)
Hinweis: Approximieren Sie B durch eine Folge invertierbarer Matrizen B_k. Benutzen Sie die Existenz der `QR`-Zerlegung und der Schur-Normalform für beliebige Matrizen.

H 35 *Lanczos für nichtsymmetrisches* A Im Folgenden sei A eine reelle nichtsymmetrische $n \times n$ Matrix. Wir untersuchen die schrittweise Herstellung einer Transformation

$$P^T AQ \;=\; T \text{ mit tridiagonalem } T \text{ und } P^T \;=\; Q^{-1} \quad (*)$$

$$T \;=\; \begin{pmatrix} \alpha_1 & \gamma_1 & 0 & . & . & . & . \\ \beta_1 & \alpha_2 & \gamma_2 & . & . & . & . \\ 0 & \beta_2 & \alpha_3 & \gamma_3 & 0 & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & 0 & \beta_{n-1} & \alpha_n \end{pmatrix}$$

1. Schreiben Sie (*) um in
$$AQ \;=\; QT \;.$$
Wie lautet dann eine entsprechende Gleichung für A^T ?
2. Leiten Sie aus diesen beiden Relationen Rekursionen für die Spalten p_j von P und q_j von Q her, wobei Sie wie beim Lanczosverfahren mit Hilfsvektoren $q_0 = 0$ und $p_0 = 0$ arbeiten um eine einheitliche Darstellung zu haben. Sie beginnen mit gegebenen Vektoren p_1 und q_1 mit
$$p_1^T q_1 \;=\; 1 \;.$$
(Dies ersetzt die Normierung auf 1 im Lanczosverfahren für den symmetrischen Fall.) Was ergibt sich also für α_i?
3. Wie drückt sich die Biorthogonalität $p_k^T q_k \;=\; 1$ mittels β_k und γ_k aus?
4. Wann kann man β_k oder γ_k null setzen und was weiss man dann von den Eigenwerten der Untermatrix T_k, der $k \times k$ Hauptuntermatrix von T?
5. Wann bricht das Verfahren ohne Ergebnis zusammen?

Numerik des Matrizeneigenwertproblems
Übung 12, Lösungsvorschlag

Präsenzübung

Ü 34 Formulieren Sie das Wielandt-Verfahren mit Shift zur Bestimmung eines Eigenvektors eines allgemeinen Eigenwertproblems

$$Ax = \lambda Bx$$

mit reell symmetrischem A und reell symmetrischem positiv definitem B. Es soll dabei weder die Inverse von B noch die Cholesky-Zerlegung von B benutzt werden. Es tritt auch hier die Aufgabe einer mehrfachen Gleichungslösung mit einer festen Matrix auf, wenn der Shift festgehalten wird. Welches Problem entsteht bei dieser Gleichungslösung? Könnte man das Problem durch eine unitäre Vortransformation a la Tridiagonalisierung vereinfachen?
Hinweis: Arbeiten Sie zunächst formal mit der Choleskyzerlegung von B (Transformation auf ein Standard-Eigenwertproblem) und übersetzen Sie dann in die Originaldaten zurück.

Sei

$$B = LL^T \quad \textit{Cholesky}$$

Dann ist

$$Ax = \lambda Bx \Leftrightarrow L^{-1}AL^{-T}L^Tx = \lambda L^Tx = Cy = \lambda y$$

mit

$$C = L^{-1}AL^{-T} \quad \textit{und} \quad y = L^Tx \ .$$

Das Wielandtverfahren für C lautet

$$(C - \mu I)y^{k+1} = y^k \ .$$

(Die Normierung unterbleibt hier zur Vereinfachung). In den Originaldaten ist das

$$(A - \mu B)x^{(k+1)} = Bx^{(k)} \ .$$

Wenn μ nicht mit einem Eigenwert des Problems übereinstimmt, ist die Matrix dieses Gleichungssystems zwar invertierbar, aber wenn μ im Inneren des Spektrums liegt, ist sie indefinit, sodaß zur Gleichungslösung entweder ein stabiler Algorithmus benutzt werden muss, der die Symmetrie nicht berücksichtigt (z.B. `QR`-Zerlegung) oder spezielle symmetrische Löser (z.B. Bunch-Parlett). Man kann das Problem auf 5-Bandform bringen, indem man in A und B spaltenweise vorgeht von unten nach oben, mittels Givensreflektoren: Z.B. in B Elemente $(n,1)$ und $(n-1,1)$ in null überführen, (und natürlich von rechts entsprechend $(1, n-1)$ und $(1, n)$), diese Transformation auch an A vornimmt, dann in A das Element $(n,1)$ bzw $(1,n)$, die gleiche Transformation auf B angewendet zerstört die erzeugten Nullen nicht usw. Man erhält dann B tridiagonal und A quindiagonal.

Ü 35 Zeigen Sie: Ist A obere Hessenbergmatrix und B invertierbare obere Dreiecksmatrix, dann gilt:

1. AB^{-1} ist obere Hessenbergmatrix.
2. Die Elemente (2,1) von $AB^{-1} - \mu I$ und von $A - \mu B$ werden durch die gleiche Ähnlichkeitstransformation mit einer Givensmatrix in null überführt.

1. *Sei* $S = R^{-1}$ *und* $C = AS$*. Da die Inverse einer oberen Dreiecksmatrix wieder eine obere Dreiecksmatrx ist, folgt mit*

$$c_{i,j} = \sum_{k=1}^{n} a_{i,k} s_{k,j}$$

und

$$a_{i,k} = 0 \text{ für } k < i-1 \quad \text{und} \quad s_{k,j} = 0 \text{ für } k > j$$

dass

$$c_{i,j} = \sum_{k=i-1}^{j} a_{i,k} s_{k,j} = 0 \text{ für } j < i-1$$

2. *Sei nun*

$$C = AB^{-1} - \mu I \ .$$

Die Überführung von $c_{2,1}$ *in Null durch eine Givensreflektion an den Zeilen und Spalten 1 und 2 von* C *wird geleistet durch die Matrix*

$$\Omega = \rho \begin{pmatrix} c_{1,1} & c_{2,1} \\ c_{2,1} & -c_{1,1} \end{pmatrix} \text{ mit } \rho = \frac{1}{\sqrt{c_{1,1}^2 + c_{2,1}^2}} \ .$$

Nun gilt

$$c_{1,1} = \frac{a_{1,1}}{b_{1,1}} - \mu \text{ und } c_{2,1} = \frac{a_{2,1}}{b_{1,1}} \ .$$

Die entsprechende Tranformation für $A - \mu B$ *lautet*

$$\begin{aligned} \Omega_1 &= \rho_1 \begin{pmatrix} a_{1,1} - \mu b_{1,1} & a_{2,1} - \mu b2,1 \\ a_{2,1} - \mu b2,1 & \mu b_{1,1} - a_{1,1} \end{pmatrix} \\ \rho_1 &= \frac{1}{\sqrt{(a_{1,1} - \mu b_{1,1})^2 + (a_{2,1} - \mu b_{2,1})^2}} \end{aligned}$$

Wegen $b_{2,1} = 0$ *liefert Multiplikation der Elemente* $c_{1,1}$ *und* $c_{2,1}$ *mit* $b_{1,1}$*, die sich in der Defintion von* Ω *wieder heraushebt, gerade die Daten zur Berechnung von* Ω_1

Bem.: Diese Tatsache macht man sich im `QZ`*-Algorithmus zu nutze. Das ist im Prinzip das* `QR`*-Verfahren für* C*, aber ausgeführt an den Originaldaten* $A,\ B$*.*

Ü 36 (*Transformation eines nichtlineares Eigenwertproblems*)

Gegeben sei das quadratische Eigenwertproblem

$$\left(A_2\lambda^2 + A_1\lambda + A_0\right) x = 0,$$

mit regulärer Matrix A_2.

a) Transformieren Sie dieses polynomiale Problem in ein gewöhnliches lineares Eigenwertproblem

$$C\tilde{x} = \lambda\tilde{x}.$$

Hinweis: Eliminieren Sie zunächst den Matrixkoeffizienten vor λ^2 und setzen Sie dann den neuen Eigenvektor an als $\tilde{x} = (\lambda x, x)^T$. Wie lautet die Matrix C ?

b) Wie würde die Transformation bei einem allgemeinen polynomialen Eigenwertproblem

$$(p(\lambda))\, x = \left(\sum_{i=0}^{m} A_i\lambda^i\right) x = 0$$

(A_m regulär) aussehen ?

a) *Nach dem Multiplizieren mit* A_2^{-1} *lautet das Eigenwertproblem*

$$\lambda^2 x + A_2^{-1}A_1(\lambda x) + A_2^{-1}A_0 x = 0.$$

Dies läst sich umschreiben in

$$\lambda^2 x = -A_2^{-1}A_1(\lambda x) - A_2^{-1}A_0 x.$$

Auf der rechten Seite der Gleichung steht das Matrix-Vektor-Produkt

$$\left(-A_2^{-1}A_1, -A_2^{-1}A_0\right)\tilde{x}$$

und links haben wir die erste Komponente von $\tilde{x}$ *multipliziert mit* λ. *Um daraus ein gewöhnliches Eigenwertproblem mit quadratischer Matrix zu machen müßte also die zweite Komponente von* $\lambda\tilde{x}$ *auf der rechten Seite ergänzt werden.* $(\lambda\tilde{x})_2$ *ist aber gerade* λx, *also die erste Komponente von* $\tilde{x}$. *Die Matrix* C *lautet deshalb*

$$C = \begin{pmatrix} -A_2^{-1}A_1 & -A_2^{-1}A_0 \\ I & 0 \end{pmatrix}$$

und das in λ *lineare Eigenwertproblem*

$$\begin{pmatrix} -A_2^{-1}A_1 & -A_2^{-1}A_0 \\ I & 0 \end{pmatrix}\begin{pmatrix} y \\ x \end{pmatrix} = \lambda\begin{pmatrix} y \\ x \end{pmatrix}.$$

Es gilt automatisch $y = \lambda x$.

b) Im Falle eines allgemeinen Polynoms wäre

$$\tilde{x} = \begin{pmatrix} \lambda^{m-1}x \\ \lambda^{m-2}x \\ \vdots \\ x \end{pmatrix}$$

und die Matrix C lautet

$$C = \begin{pmatrix} -A_m^{-1}A_{m-1} & -A_m^{-1}A_{m-2} & \cdots & & \cdots & -A_m^{-1}A_0 \\ I & 0 & \cdots & & \cdots & 0 \\ 0 & I & 0 & \cdots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & & \vdots \\ 0 & \cdots & \cdots & 0 & I & 0 \end{pmatrix}.$$

Hausübung

H 34 Zeigen Sie: Sind A und B beliebige reelle $n \times n$ Matrizen, dann existieren (komplex) unitäre Matrizen Q und Z mit

$$Q^H AZ = R_A \quad \text{und} \quad Q^H BZ = R_B$$

wobei R_A und R_B obere Dreiecksmatrizen sind. Wie stellt sich damit die Lösung des allgemeinen Eignwertproblems

$$Ax = \lambda Bx$$

dar? Warum muss man u.U. Q und Z komplex wählen, auch wenn A und B reell sind? (Die Aussage gilt natürlich auch für komplexe A, B)
Hinweis: Approximieren Sie B durch eine Folge invertierbarer Matrizen B_k. Benutzen Sie die Existenz der `QR`-Zerlegung und der Schur-Normalform für beliebige Matrizen.

$$Ax = \lambda B_k x \Leftrightarrow AB_k^{-1} y = \lambda y \quad \text{mit } y = B_k x$$

Nach dem Satz von Schur gilt

$$\exists\ Q_k \text{ unitär } : Q_k^H AB_k^{-1} Q_k = R_k \text{ obere Dreiecksmatrix}$$

Aber für beliebiges unitäres Z_k gilt

$$Q_k^H AB_k^{-1} Q_k = Q_k^H AZ_k Z_k^H B_k^{-1} Q_k \ .$$

Wähle nun Z_k so, daß

$$Z_k^H B_k^{-1} Q_k = R_{B_k} \text{ obere Dreiecksmatrix.}$$

Dann ist

$$\begin{aligned} Q_k^H AZ_k &= R_k R_{B_k}^{-1} \text{ obere Dreiecksmatrix und} \\ Q_k^H B_k Z_k &= R_{B_k}^{-1} \text{ obere Dreiecksmatrix.} \end{aligned}$$

Als Folgen unitärer Matrizen haben $\{Q_k\}$ und $\{Z_k\}$ Häufungswerte Q und Z und B_k hat den Grenzwert B. Also existieren auch Häufungswerte für die rechts stehenden Dreiecksmatrizen(-produkte) und damit die Behauptung. Weil AB_k^{-1} komplexe Eigenwerte haben kann, ist auch die Schurnormalform komplex zu verstehen, also werden die Matrizen Q und Z u.U. komplex unitär sein. Das allgemeine Eigenwertproblem wird dann zu

$$R_A y = \lambda R_B y \ , \ y \neq 0$$

Gilt nun $(R_B)_{i,i} \neq 0$, dann ist offenbar $\frac{(R_A)_{i,i}}{(R_B)_{i,i}}$ ein Eigenwert, im Falle $(R_B)_{i,i} = 0$ und $(R_A)_{i,i} \neq 0$ fehlt ein Eigenwert, während im Falle $(R_B)_{i,i} = 0 = (R_A)_{i,i}$ jeder beliebige komplexe Wert Eigenwert ist.

H 35 *Lanczos für nichtsymmetrisches* A Im Folgenden sei A eine reelle nichtsymmetrische $n \times n$ Matrix. Wir untersuchen die schrittweise Herstellung einer Transformation

$$P^T AQ \;=\; T \text{ mit tridiagonalem } T \text{ und } P^T \;=\; Q^{-1} \quad (*)$$

$$T \;=\; \begin{pmatrix} \alpha_1 & \gamma_1 & 0 & . & . & . & . \\ \beta_1 & \alpha_2 & \gamma_2 & . & . & . & . \\ 0 & \beta_2 & \alpha_3 & \gamma_3 & 0 & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & 0 & \beta_{n-1} & \alpha_n \end{pmatrix}$$

1. Schreiben Sie (*) um in
$$AQ \;=\; QT \;.$$
Wie lautet dann eine entsprechende Gleichung für A^T ?
2. Leiten Sie aus diesen beiden Relationen Rekursionen für die Spalten p_j von P und q_j von Q her, wobei Sie wie beim Lanczosverfahren mit Hilfsvektoren $q_0 = 0$ und $p_0 = 0$ arbeiten um eine einheitliche Darstellung zu haben. Sie beginnen mit gegebenen Vektoren p_1 und q_1 mit
$$p_1^T q_1 \;=\; 1 \;.$$
(Dies ersetzt die Normierung auf 1 im Lanczosverfahren für den symmetrischen Fall.) Was ergibt sich also für α_i?
3. Wie drückt sich die Biorthogonalität $p_k^T q_k \;=\; 1$ mittels β_k und γ_k aus?
4. Wann kann man β_k oder γ_k null setzen und was weiss man dann von den Eigenwerten der Untermatrix T_k, der $k \times k$ Hauptuntermatrix von T?
5. Wann bricht das Verfahren ohne Ergebnis zusammen?

1.
$$P^T AQ \;=\; T \;\Leftrightarrow\; Q^T A^T P \;=\; T^T \;\Leftrightarrow\; A^T P \;=\; PT^T \;.$$

2. Betrachtet man nun die $k-$*te Spalte von* AQ *und* $A^T P$ *dann gilt*

$$\begin{aligned} AQe_k &= Aq_k \;=\; QTe_k \;=\; q_{k-1}\gamma_{k-1} + q_k\alpha_k + q_{k+1}\beta_k \\ A^T Pe_k &= A^T p_k \;=\; PT^T e_k \;=\; p_{k-1}\beta_{k-1} + p_k\alpha_k + p_{k+1}\gamma_k \end{aligned}$$

und mit der Setzung

$$\beta_0 \;=\; 0 \;,\; \gamma_0 \;=\; 0 \;,\; p_0 \;=\; 0 \;,\; q_0 \;=\; 0$$

und gegebenen p_1, q_1 *kann man nun zunächst rekursiv auflösen:*

$$\begin{aligned} r_{k+1} &\overset{def}{=} \beta_k q_{k+1} \;=\; (A - \alpha_k I)q_k - \gamma_{k-1}q_{k-1} \\ s_{k+1} &\overset{def}{=} \gamma_k p_{k+1} \;=\; (A^T - \alpha_k I)p_k - \beta_{k-1}p_{k-1} \;. \end{aligned}$$

Die Biorthogonalität bedeutet

$$1 = p_{k+1}^T q_{k+1} = s_{k+1}^T r_{k+1} = \frac{1}{\beta_k \gamma_k} .$$

Es dürfen also weder s_{k+1}, r_{k+1} *noch das Skalarprodukt* $s_{k+1}^T r_{k+1}$ *null sein. Aus der Biorthogonalität folgert noch keine Normierungseindeutigkeit für die Vektoren. Üblich ist die Normierung*

$$||q_k|| = 1 \; \forall k$$

also

$$\begin{aligned} \beta_k &= ||r_{k+1}|| \\ \gamma_k &= \frac{s_{k+1}^T r_{k+1}}{\beta_k} . \end{aligned}$$

3. *Ist* T_k *die* $k \times k$ *Hauptuntermatrix von* T, *bedeuten* Q_k *bzw* P_k *jeweils die ersten* k *Spalten von* Q *bzw.* P, *dann lautet die Rekursion in kompakter Form*

$$\begin{aligned} AQ_k &= Q_k T_k + r_{k+1} e_k^T \\ A^T P_k &= P_k T_k^T + s_{k+1} e_k^T . \end{aligned}$$

Ist z.B. $r_{k+1} = 0$, *dann kann man* β_k *auch zu null annehmen. Ist dann* T_k *diagonalähnlich (dies ist hier eine echte Zusatzvoraussetzung!) dann wird mt der Spektralzerlegung von* T_k

$$\begin{aligned} T_k &= V_k^{-1} \Theta_k V_k \Rightarrow \\ AQ_k V_k^{-1} &= Q_k V_k^{-1} \Theta_k \end{aligned}$$

d.h. man hat ein Subsystem von Eigenvektoren und Eigenwerten von A *bestimmt und analoges gilt im Falle* $s_{k+1} = 0$ *mit* $\gamma_k = 0$ *für* A^T.

4. *Wenn* $r_{k+1} \neq 0$ *und* $s_{k+1} \neq 0$ *aber* $r_{k+1}^T s_{k+1} = 0$, *dann ist das Verfahren nicht fortsetzbar, aber man hat keine Information über das Spektrum von* A *erhalten. Dies kann durchaus eintreten.*
Bem.: es gibt eine sogenannte "look ahead" Technik für diesen Fall (man versucht dann die Faktorisierung mittels Blockmatrizen statt Skalaren in T *fortzusetzen). Details siehe Spezialliteratur.*