Otto-von-Guericke-Universität Magdeburg
Max-Planck-Institut für Dynamik komplexer technischer Systeme
Dr. Patrick Kürschner, Janna Puderbach

Wintersemester 2017/18

Einführung Numerische Linear Algebra – 6. Übung

Problem 1 (Beweis b) Satz IX.5)

Sei $A = A^T \in \mathbb{R}^{n\times n}$ mit $Ax_j = \lambda_j x_j$ wobei $\lambda_1 \geq ... \geq \lambda_n$ and $x_j^T x_k = \delta_{jk}$. Weiterhin sei $T_k = T_k^T$ die Tridiagonalmatrix nach k Lanczos Schritten, dann gilt

$$\lambda_n < \theta_k < \lambda_n + \frac{(\lambda_1 - \lambda_n)\tan^2\phi_n}{(\tau_{k-1}(1+2\rho_n))^2}$$

wobei $\rho_n := \frac{\lambda_{n-1}-\lambda_n}{\lambda_1-\lambda_{n-1}}$ und $\phi_n := \arccos|v_1^T x_n|$ und $\tau_{k-1}(x)$ das Tschebycheff Polynom vom Grad $k-1$.

Problem 2 (Erweiterung Satz IX.5)

Unter Benutzung der selben Notation wie Satz IX.5 kann man zeigen, dass für die Ritzwerte $\theta_i = \lambda_i(T_k)$, $i = 1, \dots, k$ gilt

$$\lambda_i \geq \theta_i \geq \lambda_i + (\lambda_1 - \lambda_n)\left(\frac{\kappa_i \tan\phi_i}{\tau_{k-i}(1+2\rho_i)}\right)^2,$$

wobei $\rho_i := \frac{\lambda_i-\lambda_{i+1}}{\lambda_{i+1}-\lambda_n}$, $\kappa_i := \prod_{j=1}^{i-1}\frac{\theta_j-\lambda_n}{\theta_j-\lambda_i}$, und $\phi_i := \arccos|v_1^T x_i|$.
Erklären Sie, warum diese Abschätzung eine langsamere Konvergenz zu λ_i für $1 < i < n$ impliziert.

Problem 3 (Krylov-SVD Verfahren)

a) Leiten Sie einen Lanczos-artigen Algorithmus zur Berechnung von Singulärwerten und - vektoren einer großen, sparsen Matrix $A \in \mathbb{R}^{n\times m}$ her (Annahme $\operatorname{rank}(A) = m$). Ausgangspunkt dafür ist die spaltenweise Entwicklung der Golub-Kahan-Bidiagonalisierung von A wie in der Vorlesung (Kap. IX.5.1):

$$V^T A Q = \begin{bmatrix} \ddots & \ddots \end{bmatrix} = \begin{bmatrix} \alpha_1 & \beta_1 & & \\ 0 & \ddots & \ddots & \\ & \ddots & \alpha_{m-1} & \beta_{m-1} \\ & & 0 & \alpha_m \\ 0 & \dots & \dots & 0 \end{bmatrix}, \quad V \in \mathbb{R}^{n\times n},\ Q \in \mathbb{R}^{m\times m} \quad \text{orthogonal.}$$

Vervollständigen Sie dazu Algorithmus X.4.

Hinweis: Gehen Sie analog zur Herleitung von Lanczos (Alg. X.1) aus der Tridiagonalisierung, Arnoldi (Alg. X.2) aus der Hessenbergreduktion, bzw. Bi-Lanczos aus einer unsymmetrischen Tridiagonalisierung (Übung 2) vor.

b) Im k-ten Schritt des Algorithmus hat man $V_k \in \mathbb{R}^{n\times k}$, $Q_k \in \mathbb{R}^{m\times k}$ mit orthonormalen Spalten konstruiert. Welche Räume werden durch V_k, Q_k aufgespannt?

c) Es ist möglich, dass das Verfahren (ähnlich wie Lanczos/Arnoldi) vorzeitig abbricht. Welche Informationen hat man dann gewonnen?

Algorithmus X.4 Golub-Kahan-Lanczos-Verfahren für SVD (vervollständigen Sie)

INPUT: $A \in \mathbb{R}^{n\times m}$, $n \geq m$ Startvektor $q \in \mathbb{R}^m$.
OUTPUT: Approximation an Singulärwert und ggf.-vektoren von A

$\beta_0 = 0$, $q_1 := q/\|q\|$, $q_0 = 0$.
for $k = 1, 2, \ldots$ **do**
  $v_k =$
  $\alpha_k =$
  $v_k = v_k/\alpha_k$
  $q_{k+1} =$
  $\beta_k =$
  $q_{k+1} = q_{k+1}/\beta_k$
  Löse SVD: $S_k\Sigma_k T_k^T = B_k = \begin{bmatrix} \diagdown\!\!\diagdown \end{bmatrix} = \begin{bmatrix} \alpha_1 & \beta_1 & & \\ 0 & \ddots & \ddots & \\ & \ddots & \alpha_{k-1} & \beta_{k-1} \\ & & 0 & \alpha_k \end{bmatrix}$,
  $\Sigma_k = \mathrm{diag}(\sigma_1, \ldots, \sigma_k)$, $S_k = [s_1, \ldots, s_k]$, $T_k = [t_1, \ldots, t_k]$ orthonormal.
  Test auf Konvergenz, z.B. mit Residuen __________________.
end for

Problem 4 (Details Arnoldi-Algorithmus mit impliziten Neustarts)

Sei $A \in \mathbb{C}^{n\times n}$ und $v_1 \in \mathbb{C}^n$, so dass der Arnoldi-Algorithmus für A mit Startvektor v_1 in den ersten m Schritten nicht abbricht, d.h. es gibt $V_m = [v_1, \ldots, v_m] \in \mathbb{C}^{n\times m}$ isometrisch und eine Hessenbergmatrix $H_m \in \mathbb{C}^{m\times m}$, so dass

$$AV_m = V_m H_m + h_{m+1,m} v_{m+1} e_m^T \quad \text{mit } h_{m+1,m} \neq 0.$$

Ferner seien U und R die Matrizen, die man nach ℓ Schritten des QR-Algorithmus mit Shifts $\nu_1, \ldots, \nu_\ell$ für H_m erhält (vgl. Vorlesung Kap. VIII), sowie $p(t) = (t - \nu_1) \cdots (t - \nu_\ell)$, $1 \leq \ell \leq m$. Zeigen Sie durch Induktion nach ℓ:

a) $p(H_m) = UR$.

b) $p(A)V_m = V_m p(H_m) + F_m$, wobei

$$F_m = \begin{array}{c c} & \begin{array}{cc} m-\ell & \ell \end{array} \\ n & \begin{bmatrix} 0 & \tilde{F}_m \end{bmatrix} \end{array} \quad \text{und} \quad V_m^* F_m = \begin{array}{c c} & \begin{array}{cc} m-\ell+1 & \ell-1 \end{array} \\ n & \begin{bmatrix} 0 & \hat{F}_m \end{bmatrix} \end{array}.$$