

Numerisches Programmieren, Übungen

Musterlösung 12. Übungsblatt: Iterative Verfahren II

1) Kondition eines Eigenwertproblems

a) Finden Sie alle Eigenwerte $\lambda_i(\varepsilon)$ und Eigenvektoren $v_i(\varepsilon)$ der Matrix.

$$A(\varepsilon) = \begin{pmatrix} 1 + \varepsilon \cos(2/\varepsilon) & -\varepsilon \sin(2/\varepsilon) \\ -\varepsilon \sin(2/\varepsilon) & 1 - \varepsilon \cos(2/\varepsilon) \end{pmatrix}.$$

b) Wie verhalten sich $A(\varepsilon)$, $\lambda_i(\varepsilon)$, und $v_i(\varepsilon)$ für $\varepsilon \rightarrow 0$? Ist das Eigenwert- bzw. Eigenvektorproblem für kleine ε gut konditioniert?

Lösung: Eigenwerte: Löse $\det(A - \lambda_i I) = 0$:

Mit $c := \cos(2/\varepsilon)$ und $s := \sin(2/\varepsilon)$ gilt

$$\begin{aligned} 0 &\stackrel{!}{=} \det(A(\varepsilon) - \lambda I) = \begin{vmatrix} 1 + \varepsilon c - \lambda & -\varepsilon s \\ -\varepsilon s & 1 - \varepsilon c - \lambda \end{vmatrix} = \\ &= (1 + \varepsilon c - \lambda)(1 - \varepsilon c - \lambda) - \varepsilon^2 s^2 \end{aligned}$$

Mit der 3. binomischen Formel und der Eigenschaft $\sin^2 + \cos^2 = 1$ folgt:

$$\begin{aligned} 0 &\stackrel{!}{=} ((1 - \lambda)^2 - \varepsilon^2 c^2) - \varepsilon^2 s^2 \\ &= (1 - \lambda)^2 - \varepsilon^2 (c^2 + s^2) = (1 - \lambda)^2 - \varepsilon^2 \end{aligned}$$

und damit

$$\begin{aligned} (1 - \lambda)^2 &= \varepsilon^2 \\ \Rightarrow \lambda_1 &= 1 - \varepsilon \\ \lambda_2 &= 1 + \varepsilon \end{aligned}$$

Eigenvektoren: Löse $(A - \lambda_i I)v_i = 0$:

Als Standard-Methode zur Lösung dieses Systems wählen wir die Gauß-Elimination. Da λ ein Eigenwert ist, folgt automatisch die Nicht-Invertierbarkeit der Matrix $A - \lambda I$. Da nur ein (2x2)-System vorliegt, wird die 2. Zeile nach einem Gauß-Schritt immer zu einer Nullzeile. Also erfüllt zum Beispiel jede Lösung v_1 des Systems

$$(A - \lambda_1 I)v_1 = \begin{pmatrix} 1 + \varepsilon c - \lambda_1 & -\varepsilon s \\ -\varepsilon s & 1 - \varepsilon c - \lambda_1 \end{pmatrix} v_1 = \vec{0}$$

die Bedingung $(1+\varepsilon c-\lambda)v_1(1)+(-\varepsilon s)v_1(2)=0$. Beispiele für Eigenvektoren zu den Eigenwerten λ_1 und λ_2 sind:

$$\begin{aligned} v_1 &= \begin{pmatrix} \sin(2/\varepsilon) \\ 1 + \cos(2/\varepsilon) \end{pmatrix} \\ v_2 &= \begin{pmatrix} \sin(2/\varepsilon) \\ -1 + \cos(2/\varepsilon) \end{pmatrix} \end{aligned}$$

Grenzwerte für $\varepsilon \rightarrow 0$:

$$\begin{aligned} A(\varepsilon) &\xrightarrow{\varepsilon \rightarrow 0} I \\ \lambda_i(\varepsilon) &\xrightarrow{\varepsilon \rightarrow 0} 1 \\ v_i(\varepsilon) &\xrightarrow{\varepsilon \rightarrow 0} \text{ existiert nicht!} \end{aligned}$$

Kondition:

- Das Eigenwertproblem ist gut konditioniert:

$$\lambda_i(0) = 1 = \lim_{\varepsilon \rightarrow 0} \lambda_i(\varepsilon)$$

- Das Eigenvektorproblem ist nicht gut konditioniert: Bei einem gut konditionierten Problem ließen sich die Grenzwerte $\lim_{\varepsilon \rightarrow 0} v_i(\varepsilon)$ bilden, die dann mit der analytischen Lösung (den Eigenvektoren von $A(\varepsilon=0)$, sprich den beiden Einheitsvektoren e_i) übereinstimmen würden.

2) Satz von Gerschgorin

a) Beweisen Sie den **Satz von Gerschgorin**:

Sei $A \in \mathbb{C}^{n \times n}$ eine quadratische Matrix mit Einträgen a_{ij} . Dann liegt jeder Eigenwert λ von A in mindestens einer der Kreisscheiben K_j , die durch

$$K_j := \left\{ z \in \mathbb{C} : |z - a_{jj}| \leq \sum_{k=1, k \neq j}^n |a_{jk}| \right\}, \quad j = 1, \dots, n,$$

definiert sind.

Hinweis: Betrachten Sie die j te Komponente der Eigenwertgleichung $Ax = \lambda x$, wobei x_j der maximale Eintrag von x ist.

b) Zeichnen Sie für die Matrix

$$A = \begin{pmatrix} 7 & 2 & 1 & 1 \\ 2 & 5 & 0 & 0 \\ 1 & 0 & 4 & 0 \\ 1 & 0 & 0 & 3 \end{pmatrix}$$

die Gerschgorin-Kreise und finden Sie eine obere Schranke für die Kondition $\kappa_2(A)$. Hinweis: Benutzen Sie die Euklidische Norm.

Lösung:

a) Sei x ein Eigenvektor zum Eigenwert λ und x_j der betragsmäßig größte Eintrag davon.

$$\begin{aligned}\sum_{k=1}^n a_{jk}x_k &= \lambda x_j \\ \lambda x_j - a_{jj}x_j &= \sum_{k=1, k \neq j}^n a_{jk}x_k \\ \lambda - a_{jj} &= \sum_{k=1, k \neq j}^n a_{jk} \frac{x_k}{x_j} \\ |\lambda - a_{jj}| &\leq \sum_{k=1, k \neq j}^n |a_{jk}| \frac{|x_k|}{|x_j|} \\ &\leq \sum_{k=1, k \neq j}^n |a_{jk}|\end{aligned}$$

Bemerkung: Division durch x_j immer möglich: $x_j \neq 0$, da sonst x der Nullvektor und damit kein Eigenvektor wäre. Zusätzlich sei erwähnt, dass sich die Gerschgorin-Kreise analog zu oben spaltenweise aufbauen lassen, da A und A^\top die selben Eigenwerte besitzen.

ii) Gerschgorinkreise für A :

$$\begin{aligned}|\lambda - 7| &\leq 4 \\ |\lambda - 5| &\leq 2 \\ |\lambda - 4| &\leq 1 \\ |\lambda - 3| &\leq 1\end{aligned}$$

Damit liegen offensichtlich alle Eigenwerte $\lambda_1 \approx 8.5, \lambda_2 \approx 4.4, \lambda_3 \approx 3.6$ und $\lambda_4 \approx 2.5$ in mindestens einem der Gerschgorinkreise. Weiter können wir folgern:

A ist symmetrisch $\Rightarrow \lambda \in \mathbb{R}$

$\Rightarrow \lambda \in [3, 11] \cup [3, 7] \cup [3, 5] \cup [2, 4] = [2, 11]$

$\Rightarrow A$ ist positiv definit ($\lambda > 0$), A^{-1} existiert.

Abschätzung der Kondition von A bezüglich der Spektralnorm $\|\cdot\|_2$:

$$\begin{aligned}\kappa_2(A) &= \|A\|_2 \cdot \|A^{-1}\|_2 = \sqrt{\lambda_{\max}(A^T A)} \cdot \sqrt{\lambda_{\max}(A^{-T} A^{-1})} \\ &\stackrel{\text{Symmetrie}}{=} |\lambda_{\max}(A)| \cdot |\lambda_{\max}(A^{-1})| = \frac{\lambda_{\max}(A)}{\lambda_{\min}(A)} \leq \frac{11}{2} = 5.5.\end{aligned}$$

3) Rayleigh Quotient

- a) Sei $x \in \mathbb{R}^n$ Eigenvektor einer Matrix $A \in \mathbb{R}^{n \times n}$. Zeigen Sie, dass die zugehörigen Eigenwerte λ mit Hilfe des Rayleigh Quotienten berechnet werden können:

$$\lambda = \frac{x^T A x}{x^T x}.$$

- b) Sei λ_{\min} und λ_{\max} der kleinste bzw. größte Eigenwert einer symmetrischen Matrix $A \in \mathbb{R}^{n \times n}$. Zeigen Sie, dass folgende Zusammenhänge gelten:

$$\lambda_{\min} = \min_{x \neq 0} \frac{x^T A x}{x^T x}, \quad \lambda_{\max} = \max_{x \neq 0} \frac{x^T A x}{x^T x}.$$

Hinweis: Wegen

$$\lambda_{\min} = \min_{x \neq 0} \frac{x^T A x}{x^T x} = \min_{x \neq 0} \left(\frac{1}{\|x\|_2} x \right)^T A \left(\frac{1}{\|x\|_2} x \right) = \min_{\|y\|_2=1} y^T A y$$

ist es ausreichend, folgende Zusammenhänge zu zeigen:

$$\lambda_{\min} = \min_{\|x\|_2=1} x^T A x \quad \text{and} \quad \lambda_{\max} = \max_{\|x\|_2=1} x^T A x.$$

Lösung:

- a)

$$\begin{aligned} Ax &= \lambda x \\ x^T A x &= \lambda x^T x \\ \lambda &= \frac{x^T A x}{x^T x} \end{aligned}$$

- b) A ist symmetrisch mit reellen Einträgen. \Rightarrow Es existiert eine orthonormale Basis bestehend aus den Eigenvektoren $\{v_i; i = 1, \dots, n\}$ von A .

Sei $x \in \mathbb{R}^n$ ein beliebiger Vektor mit Einheitslänge. Wir können x schreiben als

$$x = \sum_{i=1}^n \alpha_i v_i$$

wobei gilt $\alpha_i \in \mathbb{R}$, $i = 1, \dots, n$.

Wegen der Orthonormalität der v_i erhalten wir

$$x^T x = \left(\sum_{i=1}^n \alpha_i v_i \right)^T \left(\sum_{i=1}^n \alpha_i v_i \right) \stackrel{\text{Orthogonalität}}{=} \sum_{i=1}^n \alpha_i^2 \|v_i\|_2^2 \stackrel{\text{Orthonormalität}}{=} \sum_{i=1}^n \alpha_i^2$$

und mit $\|x\|_2 = 1$, ergibt sich für die α_i die Bedingung

$$\sum_{i=1}^n \alpha_i^2 = 1. \tag{1}$$

Analog können wir auch $x^T Ax$ berechnen

$$x^T Ax = \left(\sum_{i=1}^n \alpha_i v_i \right)^T A \left(\sum_{i=1}^n \alpha_i v_i \right) = \left(\sum_{i=1}^n \alpha_i v_i \right)^T \left(\sum_{i=1}^n \lambda_i \alpha_i v_i \right)$$

$$\stackrel{\text{Orthogonalität}}{=} \sum_{i=1}^n \lambda_i \alpha_i^2 \|v_i\|_2^2 \stackrel{\text{Orthonormalität}}{=} \sum_{i=1}^n \lambda_i \alpha_i^2 \rightarrow \min!$$

Mit Gleichung (1) ergibt sich, dass $x^T Ax$ genau dann minimiert wird, wenn $|\alpha_i| = 1$ gilt, für einen Index $i = i^*$ mit $\lambda_{i^*} = \lambda_{\min}$, sowie $\alpha_i = 0$ für $i \neq i^*$. Damit wird $x^T Ax$ minimiert für folgenden Vektor x

$$x = \sum_{i=1}^n \alpha_i v_i = v_{i^*},$$

und wir erhalten

$$\min_{\|x\|_2=1} x^T Ax = v_{i^*}^T A v_{i^*} \stackrel{(i)}{=} \lambda_{i^*} = \lambda_{\min}$$

Nach demselben Schema lässt sich

$$\lambda_{\max} = \max_{\|x\|_2=1} x^T Ax$$

zeigen.

4) Iterationsverfahren

- a) Berechnen Sie analytisch alle Eigenwerte und Eigenvektoren der Matrix $A = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$.
- b) Führen Sie zwei Iteration der direkten Vektoriteration (power iteration) für A
 - a) ohne Shift,
 - b) mit Shift $\mu = 1.5$ und
 - c) mit Shift $\mu = 3.5$

durch und berechnen Sie die zugehörige Eigenwertapproximation. Benutzen Sie dazu $x_0 = (1, 0)^T$ als Startvektor. Gegen welchen Eigenwert konvergiert die Iteration? Wie ist die Konvergenzrate für den jeweiligen Fall?

- c) Für welchen Shift μ konvergiert die direkte Vektoriteration (power iteration) gegen den ersten bzw. zweiten Eigenwert von A ? Was passiert, falls $\mu = 3$ gewählt wird?

Lösung:

- a) Bestimmung der Eigenwerte:

$$\begin{aligned} 0 &\stackrel{!}{=} \det(A - \lambda I) \\ &= (3 - \lambda)^2 - 1 \\ &= \lambda^2 - 6\lambda + 8 \\ \Rightarrow \lambda_i &= 3 \pm 1 \end{aligned}$$

$$\lambda_1 = 4, \lambda_2 = 2$$

Bestimmung der Eigenvektoren durch Lösung der homogenen Gleichungssysteme

$$(A - \lambda_i I) v_i = 0$$

$$v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

b) Mit Algorithmus aus der Vorlesung:

a)

$$w_0 = Ax_0 = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$$

$$\lambda^{(0)} = x_0^T w_0 = (1, 0) \cdot \begin{pmatrix} 3 \\ 1 \end{pmatrix} = 3$$

$$x_1 = \frac{w_0}{\|w_0\|_2} = \frac{1}{\sqrt{10}} \begin{pmatrix} 3 \\ 1 \end{pmatrix}$$

$$w_1 = Ax_1 = \frac{1}{\sqrt{10}} \begin{pmatrix} 10 \\ 6 \end{pmatrix}$$

$$\lambda^{(1)} = x_1^T w_1 = \frac{36}{10} = 3.6$$

$$x_2 = \frac{w_1}{\|w_1\|_2}$$

$$\lambda^{(k)} \xrightarrow{k \rightarrow \infty} 4$$

Konvergenzrate:

$$\frac{\lambda_2}{\lambda_1} = \frac{2}{4} = \frac{1}{2}$$

b)

$$w_0 = (A - \mu I)x_0 = \begin{pmatrix} \frac{3}{2} \\ 1 \end{pmatrix}$$

$$\tilde{\lambda}^{(0)} = x_0^T w_0 = \frac{3}{2}$$

$$x_1 = \frac{w_0}{\|w_0\|_2} = \frac{1}{\sqrt{13}} \begin{pmatrix} 3 \\ 2 \end{pmatrix}$$

$$w_1 = (A - \mu I)x_1 = \frac{1}{\sqrt{13}} \begin{pmatrix} \frac{13}{2} \\ 6 \end{pmatrix}$$

$$\tilde{\lambda}^{(1)} = x_1^T w_1 = \frac{1}{13} (3, 2) \cdot \left(\frac{13}{2}, 6\right)^T = \frac{63}{26} \approx 2.42$$

$$\lambda^{(1)} = \tilde{\lambda}^{(1)} + \mu \approx 3.92$$

$$x_2 = \frac{w_1}{\|w_1\|_2}$$

$$\lambda^{(k)} \xrightarrow{k \rightarrow \infty} 4$$

Konvergenzrate:

$$\frac{\lambda_2 - \mu}{\lambda_1 - \mu} = \frac{0.5}{2.5} = \frac{1}{5}$$

c)

$$\begin{aligned}
(A - \mu I) &= \begin{pmatrix} -0.5 & 1 \\ 1 & -0.5 \end{pmatrix} \\
w_0 &= (A - \mu I)x_0 = \begin{pmatrix} -1 \\ 2 \end{pmatrix} \\
\tilde{\lambda}^{(0)} &= x_0^T w_0 = \frac{-1}{2} \\
x_1 &= \frac{w_0}{\|w_0\|_2} = \frac{1}{\sqrt{5}} \begin{pmatrix} -1 \\ 2 \end{pmatrix} \\
w_1 &= (A - \mu I)x_1 = \begin{pmatrix} -0.5 & 1 \\ 1 & -0.5 \end{pmatrix} \cdot \frac{1}{\sqrt{5}} \begin{pmatrix} -1 \\ 2 \end{pmatrix} = \frac{1}{\sqrt{5}} \begin{pmatrix} \frac{5}{2} \\ -2 \end{pmatrix} \\
\tilde{\lambda}^{(1)} &= x_1^T w_1 = \frac{1}{5} \left(\frac{-5}{2} - 4 \right) = \frac{-13}{10} = -1.3 \\
\lambda^{(1)} &= \tilde{\lambda}^{(1)} + \mu = 2.2
\end{aligned}$$

$$\lambda^{(k)} \xrightarrow{k \rightarrow \infty} 2$$

Konvergenzrate (kleinstes mit Shift durch größtes mit Shift):

$$\frac{\lambda_1 - \mu}{\lambda_2 - \mu} = \frac{0.5}{-1.5} = \frac{-1}{3}.$$

Die Konvergenzrate wird (durch Beträge) natürlich eigentlich immer positiv angegeben

c)

$$\begin{aligned}
\mu < 3 &\Rightarrow \lambda^{(k)} \xrightarrow{k \rightarrow \infty} 4 \\
\mu > 3 &\Rightarrow \lambda^{(k)} \xrightarrow{k \rightarrow \infty} 2
\end{aligned}$$

Für $\mu = 3$ erhalten wir:

$$\begin{aligned}
x_0 &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\
w_1 &= (A - \mu I)x_0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\
x_1 &= \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\
w_2 &= (A - \mu I)x_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\
x_2 &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} = x_0 \Rightarrow \text{periodisch, keine Konvergenz!}
\end{aligned}$$