

10. Eigenwertprobleme

10.1 Grundlagen

Zu einer vorgegebenen Matrix $A \in \mathbb{R}^{(n,n)} / \mathbb{C}^{(n,n)}$ werden die Zahlen $\lambda \in \mathbb{C}$ und Vektoren $v \in \mathbb{C}^n \setminus \{0\}$ gesucht, für die gilt:

$$Av = \lambda v \quad (10.1)$$

λ : Eigenwert, v : Eigenvektor

$$(10.1) \iff (A - \lambda I_n) v = 0, v \neq 0$$

$$\iff A - \lambda I_n \text{ singular}$$

$$\iff \det(A - \lambda I_n) = 0$$

Definition: $p_A(\lambda) := \det(A - \lambda I_n)$ (10.2)

ist ein Polynom vom Grad n , das charakteristische Polynom.

Nach obigem gilt also:

$$\lambda \text{ Eigenwert von } A \iff p_A(\lambda) = 0 \quad (10.3)$$

- Die Vielfachheit eines EW λ als Nullstelle von $p_A(\lambda)$ heißt algebraische Vielfachheit: $a(\lambda)$.
- Die Dimension des Eigenraumes $E_\lambda := \text{Kern}(A - \lambda I_n)$ heißt geometrische Vielfachheit: $g(\lambda)$.
- Zwei Matrizen $A, B \in \mathbb{C}^{(n,n)}$ heißen ähnlich, wenn es eine reguläre Matrix $S \in \mathbb{C}^{(n,n)}$ gibt mit $B = S^{-1}AS$.
Ähnliche Matrizen beschreiben die gleiche lineare Abbildung $l: \mathbb{C}^n \rightarrow \mathbb{C}^n$ bezügl. ev. verschiedener Basen von \mathbb{C}^n .
- Ähnliche Matrizen haben das gleiche Charakt. Polynom, also auch gleiche Eigenwerte. Auch die algebraischen und geometrischen Vielfachheiten stimmen überein!
- Für einen EW λ gilt stets:

$$1 \leq g(\lambda) \leq a(\lambda) \leq n$$

(10.4)

Beweis:

$$(i) \quad B = S^{-1}AS \quad \Rightarrow$$

$$\begin{aligned} p_B(\lambda) &= \det(S^{-1}AS - \lambda I_n) \\ &= \det[S^{-1}(A - \lambda I_n)S] \\ &= \det(S^{-1}) \cdot p_A(\lambda) \cdot \det(S) \\ &= p_A(\lambda). \end{aligned}$$

(ii) Ist λ_1 EW mit $g(\lambda_1) = k$, so gibt es eine Basis (v^1, \dots, v^k) von E_{λ_1} .

Erweitert man diese zur Basis (v^1, \dots, v^n) von \mathbb{C}^n , so folgt mit $Av^j = \lambda_1 v^j$, $j=1, \dots, k$ und $S := (v^1, \dots, v^n)$:

$$B := S^{-1}AS = \left(\begin{array}{c|c} \lambda_1 & * \\ \hline 0 & R \end{array} \right), \quad R \in \mathbb{C}^{(n-k, n-k)}$$

$$\begin{aligned} \Rightarrow p_A(\lambda) &= p_B(\lambda) \\ &= (\lambda - \lambda_1)^k \cdot \det(R - \lambda I_{n-k}) \end{aligned}$$

$$\Rightarrow a(\lambda) \geq g(\lambda_1) = k \quad \blacktriangle$$

Diagonalisierbarkeit: Ist $A \in \mathbb{C}^{(n,n)}$ zu einer Diagonalmatrix $\Lambda = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ ähnlich, so heißt A diagonalisierbar, also:

$$\Lambda = S^{-1}AS, \quad S = (v^1, \dots, v^n) \text{ regulär}$$

$$\Leftrightarrow S\Lambda = AS:$$

$$(v^1, \dots, v^n) \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = A(v^1, \dots, v^n)$$

$$\Leftrightarrow \forall i=1, \dots, n: \lambda_i v^i = Av^i$$

$$\Leftrightarrow \lambda_1, \dots, \lambda_n \text{ sind die EWe von } A \text{ und } (v^1, \dots, v^n) \text{ ist eine Basis aus EVen.}$$

Damit folgt:

$$A \text{ diagonalisierbar} \Leftrightarrow \forall \lambda: g(\lambda) = a(\lambda)$$

(10.5)

Beispiel (10.6)

$$\underline{\underline{a)}} \quad A = \begin{pmatrix} 3 & 2 & 0 \\ -3 & -2 & 0 \\ -3 & -3 & 1 \end{pmatrix}$$

$$P_A(\lambda) = \begin{vmatrix} 3-\lambda & 2 & 0 \\ -3 & -2-\lambda & 0 \\ -3 & -3 & 1-\lambda \end{vmatrix} = -\lambda(\lambda-1)^2$$

\Rightarrow Eigenwerte: $\lambda_1 = 0$, $\lambda_2 = 1$

Eigenräume:

$$\underline{\lambda_1 = 0}: \left(\begin{array}{ccc|c} 3 & 2 & 0 & 0 \\ -3 & -2 & 0 & 0 \\ -3 & -3 & 1 & 0 \end{array} \right) \Rightarrow v^1 = \begin{pmatrix} -2 \\ 3 \\ 3 \end{pmatrix}$$

$$\underline{\lambda_2 = 1}: \left(\begin{array}{ccc|c} 2 & 2 & 0 & 0 \\ -3 & -3 & 0 & 0 \\ -3 & -3 & 0 & 0 \end{array} \right) \Rightarrow v^2 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, v^3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

\Rightarrow A diagonalisierbar, $\Lambda = S^{-1}AS = \begin{pmatrix} 0 & & \\ & 1 & \\ & & 1 \end{pmatrix}$,

$$S = \begin{pmatrix} -2 & -1 & 0 \\ 3 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix}.$$

b) $A = \begin{pmatrix} 8 & 8 & 4 \\ -1 & 2 & 1 \\ -2 & -4 & -2 \end{pmatrix}$

$$\Rightarrow P_A(\lambda) = -\lambda(\lambda-4)^2$$

\Rightarrow Eigenwerte $\lambda_1 = 0$, $\lambda_2 = 4$

Eigenraum E_{λ_2} : $\left(\begin{array}{ccc|c} 4 & 8 & 4 & 0 \\ -1 & -2 & 1 & 0 \\ -2 & -4 & -6 & 0 \end{array} \right)$

Matrix $A - \lambda_2 I$ hat Rang = 2, daher

$E_2 = \mathbb{C} \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}$ eindimensional, $g(\lambda_2) < a(\lambda_2)$

$\Rightarrow A$ ist nicht diagonalisierbar!

Warnung:

Die Berechnung der Eigenwerte als Nullstellen des charakt. Polynoms ist numerisch instabil !!

Grund: Kleine Änderungen in den Koeffizienten des charakt. Polynoms können zu großen Änderungen in den Nullstellen führen!

Vgl. hierzu das Beispiel (2.3)!

Definition (10.7)

$A \in \mathbb{C}^{(n,n)}$	A reell
$A^* := \bar{A}^T$	A^T
A <u>Hermitesch</u> $:\Leftrightarrow A = A^*$	A symmetrisch
A <u>unitär</u> $:\Leftrightarrow AA^* = A^*A = I$	A orthogonal
A <u>normal</u> $:\Leftrightarrow AA^* = A^*A$	$A^T A = A A^T$

Satz (10.8)

Ist $A \in \mathbb{C}^{(n,n)}$ eine Hermitesche Matrix, so sind alle EWe von A reell, die EVen zu verschiedenen EVen sind orthogonal bzgl. des Skalarproduktes $\langle u, v \rangle = u^* v$.

Beweis:

a) v EV zum EW $\lambda \in \mathbb{C} \Rightarrow$

$$\begin{cases} v^* A v = v^* (\lambda v) = \lambda \|v\|^2 \\ v^* A^* v = (A v)^* v = \bar{\lambda} \|v\|^2 \end{cases}$$

Aus $A = A^*$ folgt daher: $\lambda = \bar{\lambda}$, also $\lambda \in \mathbb{R}$.

b) gilt $A v = \lambda v$, $A w = \mu w$ mit $\lambda \neq \mu$ und $v \neq 0$, $w \neq 0$, so folgt

$$\begin{aligned} (\lambda - \mu) v^* w &= (\lambda v)^* w - v^* (\mu w) \\ &= (A v)^* w - v^* (A w) \\ &= v^* (A^* - A) w \\ &= 0, \end{aligned}$$

also $v^* w = 0$ \blacktriangle

Satz (10.9)

Jede Hermitesche Matrix $A \in \mathbb{C}^{(n,n)}$ ist diagonalisierbar. Die Transformationsmatrix S lässt sich unitär wählen.

Beweis: Wäre A nicht diagonalisierbar, so gäbe es (nach dem Satz über die Jordansche Normalform) einen Hauptvektor erster Stufe, d.h. es gäbe einen EW λ von A und einen Vektor $v \in \mathbb{C}^n$ mit

$$(A - \lambda I)^2 v = 0, \quad (A - \lambda I)v \neq 0 \quad (*)$$

Nun ist auch $(A - \lambda I)$ Hermitesch (da λ nach (10.8) reell ist). Damit

$$(A - \lambda I)^* (A - \lambda I) v = (A - \lambda I)^2 v = \vec{0}$$

$$\Rightarrow v^* (A - \lambda I)^* (A - \lambda I) v = 0$$

$$\Rightarrow \| (A - \lambda I) v \|^2 = 0$$

$$\Rightarrow (A - \lambda I) v = 0$$

im Widerspruch zu $(*)$ \blacktriangle

Kondition der EW aufgabe.

Sei $A \in \mathbb{C}^{(n,n)}$ vorgegeben. λ_1 sei EW von A mit zugehörigem EV $v^1 \in \mathbb{C}^n$.

$$\begin{aligned} \text{Wegen } \overline{p_A(\lambda)} &= \overline{\det(A - \lambda I)} \\ &= \det(\bar{A} - \bar{\lambda} I) \\ &= \det(A^* - \bar{\lambda} I) = p_{A^*}(\bar{\lambda}) \end{aligned}$$

ist dann auch $\bar{\lambda}_1$ EW von A^* . Ein zugehöriger EV sei $w^1 \in \mathbb{C}^n$.

Satz (10.10) Sei λ_1 einfacher EW von A .

a) Es ex. Umgebung U von A in $\mathbb{C}^{(n,n)}$ und eine C^1 -Fkt. $\lambda: U \rightarrow \mathbb{C}$ mit

$$\lambda(A) = \lambda_1, \quad \lambda(B) \text{ einfacher EW von } B,$$

$$\lambda'(A) \zeta = \frac{\langle w^1, \zeta v^1 \rangle}{\langle w^1, v^1 \rangle}$$

$$\|\lambda'(A)\|_2 = \frac{\|w^1\| \|v^1\|}{|\langle w^1, v^1 \rangle|} = \frac{1}{|\cos \angle(v^1, w^1)|}$$

b) Die Kondition von $A \mapsto \lambda_1$ lautet

$$K_{\text{abs}} = \frac{1}{|\cos \angle(v^1, w^1)|}, \quad K_{\text{rel}} = \frac{\|A\|_2}{|\lambda_1 \cdot \cos \angle(v^1, w^1)|}$$

Beweis: → Deuffhard, Hohmann,
Sätze 5.1, 5.2, Seiten 129-131.

Folgerung (10.11)

Für normale Matrizen (also auch für
Hermitesche und unitäre Matrizen) gilt
 $w^1 = v^1$ (bis auf Längennormierung)

⇒ Das EW-Problem ist gut konditioniert,
 $\kappa_{\text{abs}} = 1$.

Beweis:

$$\lambda_1 (A^* v^1) = A^* (\lambda_1 v^1) = A^* A v^1 = A (A^* v^1)$$

⇒ $A^* v^1$ EV von A zum EW λ_1 (oder $= 0$)

⇒ Einfachheit! $A^* v^1 = \mu v^1$, $\mu \in \mathbb{C}$

$$\begin{aligned} \Rightarrow \mu (v^{1*} v^1) &= v^{1*} A^* v^1 = (A v^1)^* v^1 \\ &= (\lambda_1 v^1)^* v^1 = \bar{\lambda}_1 (v^{1*} v^1) \end{aligned}$$

$$\Rightarrow \mu = \bar{\lambda}_1 \quad \blacktriangle$$

10.2. Eigenwertabschätzungen

Satz (10.12) (Hirsch)

Es bezeichne $\|\cdot\|$ eine Norm für $\mathbb{C}^{(n)}$ und die zugehörige Matrixnorm für $\mathbb{C}^{(n,n)}$.

Ist λ EW von $A \in \mathbb{C}^{(n,n)}$, so gilt

$$|\lambda| \leq \|A\|.$$

Beweis:

$$|\lambda| \|v\| = \|\lambda v\| = \|Av\| \leq \|A\| \|v\| \quad \blacktriangle$$

Beispiel

$$A = \begin{pmatrix} 1 & 0.1 & -0.1 \\ 0 & 2 & 0.4 \\ -0.2 & 0 & 3 \end{pmatrix}$$

Mit $\|\cdot\| = \|\cdot\|_\infty$ folgt:

$$|\lambda| \leq \max\{1.2, 2.4, 3.2\} = 3.2.$$

Bemerkung: Wie genau ist die Abschätzung (10.12)?

Es bezeichne

$$\sigma(A) := \{\lambda \in \mathbb{C} \mid \lambda \text{ EW von } A\} \quad (\text{Spektrum})$$

$$\rho(A) := \max_{\lambda \in \sigma(A)} |\lambda| \quad (\text{Spektralradius})$$

Dann lässt sich zeigen:

Zu jedem $\varepsilon > 0$ gibt es eine (von A abhängige) Matrixnorm $\|\cdot\|$ mit

$$\rho(A) \leq \|A\| \leq \rho(A) + \varepsilon.$$

Beweis: Stoer, Bulirsch: Satz (6.9.2).

Satz (10.13) (Gerschgorin)

Zu $A \in \mathbb{C}^{(n,n)}$ definiere man die Kreise

$$K_i := \left\{ \lambda \in \mathbb{C} : |\lambda - a_{ii}| \leq \sum_{\substack{k=1 \\ k \neq i}}^n |a_{ik}| \right\}$$

$(i=1, \dots, n)$

a) λ EW von $A \Rightarrow \lambda \in \bigcup_{i=1}^n K_i =: K$

b) Jede Zusammenhangskomponente $\bigcup_{j=1}^k K_{i_j}$ von K - bestehend aus k Kreisscheiben - enthält genau k EWe von A (der Vielfachheit nach gezählt).

Beweis: (nur a)

Ist v EV zum EW λ von A , so gilt für

irgendeine Matrix $B \in \mathbb{C}^{(n,n)}$:

$$(A - B)v = (\lambda I - B)v.$$

Ist nun λ kein EW von B , d.h. $(\lambda I - B)$ regulär, so folgt

$$(\lambda I - B)^{-1}(A - B)v = v$$

$$\Rightarrow \|(\lambda I - B)^{-1}(A - B)\| \geq 1. \quad (*)$$

Wähle nun $\|\cdot\| = \|\cdot\|_\infty$ und

$$B = \text{diag}(a_{11}, \dots, a_{nn})$$

$$(*) \Rightarrow \|(\lambda I - B)^{-1}(A - B)\|_\infty = \max_{1 \leq i \leq n} \sum_{\substack{j=1 \\ j \neq i}}^n \frac{|a_{ij}|}{|\lambda - a_{ii}|} \geq 1$$

$$\Rightarrow \exists i \in \{1, \dots, n\} : \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| \geq |\lambda - a_{ii}| \quad \blacktriangle$$

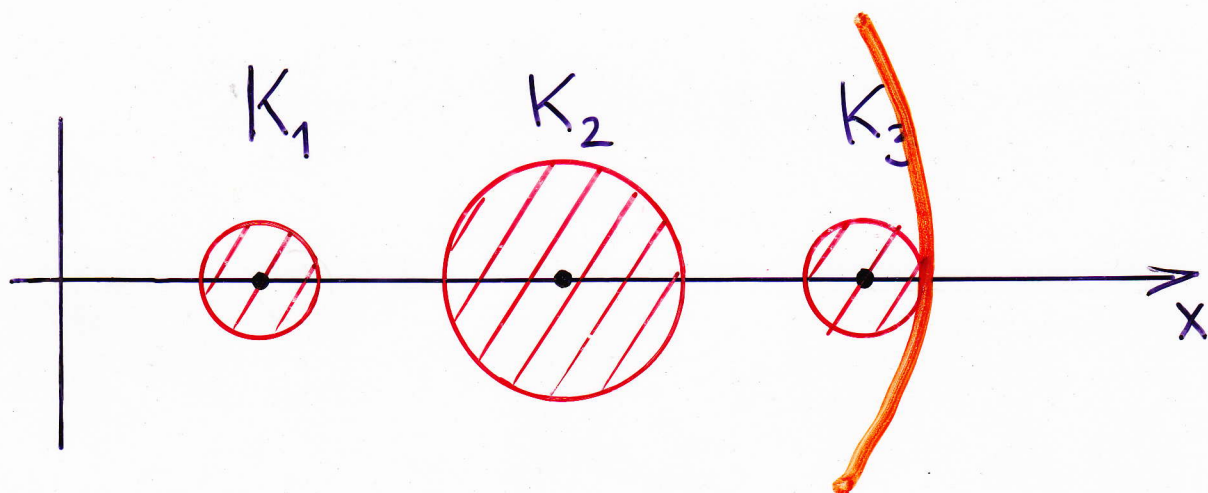
Beweis zu b): Stoer, Bulirsch : (6.9.5).

Beispiel:

$$A = \begin{pmatrix} 1 & 0.1 & -0.1 \\ 0 & 2 & 0.4 \\ -0.2 & 0 & 3 \end{pmatrix}$$

Die EWe liegen nach (10.13) in

$$\bar{K}_{0.2}(1) \cup \bar{K}_{0.4}(2) \cup \bar{K}_{0.2}(3)$$



Aus (10.13) b) folgt weiter, dass in jedem Kreis K_i genau ein Eigenwert liegt.

Insbesondere ist A damit diagonalisierbar.

MATLAB-Berechnung liefert:

$$\lambda_1 = 0.98615\dots$$

$$\lambda_2 = 2.00784\dots$$

$$\lambda_3 = 3.00600\dots$$

Reelle, symmetrische Matrizen

Im Folgenden gehen wir zum Eigenwertproblem für reelle, symmetrische Matrizen $A \in \mathbb{R}^{(n,n)}$ über. Nach (10.8) sind die EWe reell, es gelte:

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \quad (10.14)$$

Ferner ex. eine Orthonormalbasis (w^1, \dots, w^n) aus EVen, vgl. (10.9).

Definition (10.15)

Zu $x \in \mathbb{R}^n \setminus \{0\}$ heißt $R(x) := \frac{x^T A x}{x^T x}$
der Rayleigh-Quotient.

Satz (10.16) (Rayleighsches Prinzip)

a) $\forall x \in \mathbb{R}^n \setminus \{0\} : \lambda_n \leq R(x) \leq \lambda_1$

b) $\lambda_1 = \max_{x \neq 0} R(x), \quad \lambda_n = \min_{x \neq 0} R(x)$

c) $R(x) = \lambda_1 \Rightarrow x \text{ EV zu } \lambda_1$

$R(x) = \lambda_n \Rightarrow x \text{ EV zu } \lambda_n$

d) $\lambda_i = \min \{ R(x) : x \perp w^{i+1}, \dots, w^n, x \neq 0 \}$

$\lambda_i = \max \{ R(x) : x \perp w^1, \dots, w^{i-1}, x \neq 0 \}$

Anwendung: Ist x Näherung für einen EV von A , so ist $R(x)$ eine Näherung für den zugehörigen EW.

Beweis:

zu a): $x = \sum_{i=1}^n x_i w^i \Rightarrow$

$x^T x = \sum x_i^2, \quad Ax = \sum x_i \lambda_i w^i$

$$\Rightarrow x^T A x = \sum_i \lambda_i x_i^2$$

$$\Rightarrow R(x) = \frac{\sum \lambda_i x_i^2}{\sum x_i^2} \begin{cases} \leq \lambda_1 \\ \geq \lambda_n \end{cases}$$

zub): Einsetzen zeigt $R(w^j) = \lambda_j$. Damit

$$\lambda_n = R(w^n) \leq R(x) \leq R(w^1) = \lambda_1$$

zuc): $R(x) = \frac{\sum \lambda_i x_i^2}{\sum x_i^2} = \lambda_1$

$$\Rightarrow \forall i: (\lambda_i = \lambda_1 \vee x_i = 0)$$

$$\Rightarrow x = \sum_{j=1}^k x_j w^j \text{ mit } \lambda_1 = \lambda_2 = \dots = \lambda_k > \lambda_{k+1}$$

ist EV zu λ_1 .

zud): $x \perp w^{i+1}, \dots, w^n \Rightarrow x = \sum_{j=1}^i x_j w^j$

$$\Rightarrow R(x) = \frac{\sum_{j \leq i} \lambda_j x_j^2}{\sum_{j \leq i} x_j^2} \geq \lambda_i$$

und $R(w^i) = \lambda_i$

$$\Rightarrow \lambda_i = \min \{ R(x) : x \perp w^{i+1}, \dots, w^n, x \neq 0 \}$$

Satz (10.17) (Bogoljubow, Krylow)

Sei $\lambda \in \mathbb{R}$ eine Näherung für einen EW von A ,
 $x \in \mathbb{R}^n \setminus \{0\}$ eine Näherung für den zugehörigen
 EV. Dann gibt es einen EW λ_k mit

$$|\lambda - \lambda_k| \leq \sqrt{\frac{y^T y}{x^T x}}, \quad y := Ax - \lambda x$$

Beweis: Sei (w^1, \dots, w^n) wieder eine ONB aus
 EVen von A . Es folgt:

$$x = \sum_{i=1}^n x_i w^i, \quad Ax = \sum_{i=1}^n x_i \lambda_i w^i$$

$$\Rightarrow y = Ax - \lambda x = \sum_{i=1}^n (\lambda_i - \lambda) x_i w^i$$

$$\Rightarrow y^T y = \sum_{i=1}^n (\lambda_i - \lambda)^2 x_i^2 \geq (\lambda_k - \lambda)^2 \sum_{i=1}^n x_i^2$$

$$\text{mit } |\lambda_k - \lambda| = \min_i |\lambda_i - \lambda|$$

$$\Rightarrow (\lambda_k - \lambda)^2 \leq \frac{y^T y}{x^T x} \quad \blacktriangle$$

Beispiel:

$$A = \begin{pmatrix} 1 & 0.1 & -0.1 \\ 0 & 2 & 0.4 \\ -0.2 & 0 & 3 \end{pmatrix}$$

Zur Einschließung des Eigenwertes $\lambda \approx 2$
 setzen wir $x := (0, 1, 0)^T$.

$$\Rightarrow Ax = \begin{pmatrix} 0.1 \\ 2 \\ 0 \end{pmatrix}; \quad y = Ax - \lambda x = \begin{pmatrix} 0.1 \\ 0 \\ 0 \end{pmatrix}$$

$$\Rightarrow |\lambda_k - 2| \leq \sqrt{\frac{y^T y}{x^T x}} = \sqrt{0.01} = \underline{0.1}.$$

Verbesserung der Näherungen:

$$\tilde{x} = Ax = \begin{pmatrix} 0.1 \\ 2 \\ 0 \end{pmatrix}$$

$$\tilde{\lambda} = R(\tilde{x}) = \frac{\tilde{x}^T A \tilde{x}}{\tilde{x}^T \tilde{x}} \doteq 2.00249$$

$$\tilde{y} = A\tilde{x} - \tilde{\lambda}\tilde{x} \doteq \begin{pmatrix} 0.09975 \\ -0.004987 \\ -0.02 \end{pmatrix}$$

$$\Rightarrow |\lambda_2 - 2.00249| \leq \sqrt{\frac{\tilde{y}^T \tilde{y}}{\tilde{x}^T \tilde{x}}} \doteq \underline{0.0508}.$$

Bemerkung (10.18)

Zu einem Polynom $p(x) = \sum_0^n a_j x^j$ liefert

$$A = \begin{pmatrix} 0 & & & -x_0 \\ 1 & & & \vdots \\ & \ddots & & 0 \\ & & 0 & -x_{n-1} \\ & & 1 & \end{pmatrix}, \quad x_i := \frac{a_i}{a_n} \quad (\text{Frobenius})$$

eine Matrix mit $p_A(\lambda) = \frac{1}{a_n} (-1)^n p(\lambda)$.

Mit Hilfe der Eigenwertabschätzungen lassen sich also auch brauchbare Abschätzungen für die Nullstellen eines Polynoms gewinnen!

10.3 Vektoriteration

auch: von Mises - Verfahren

nach Richard von Mises (1883 - 1953)

Es sei $A \in \mathbb{R}^{(n,n)}$ wieder eine symmetrische Matrix und für die (reellen) Eigenwerte gelte

$$|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_n|. \quad (10.19)$$

Insbesondere sei also der betragsgrößte EW λ_1 einfach!

Ausgangspunkt ist die Iteration

$x^{(0)} \in \mathbb{R}^n$ Startvektor ($\neq 0$)

$$\begin{array}{l} \text{für } i = 0, 1, 2, \dots \\ | \quad x^{(i+1)} = A x^{(i)} \end{array} \quad (10.20)$$

Ist (v_1, \dots, v_n) eine Basis aus EV zu A (wobei v_i EV zum EW λ_i) und gilt

$$x^{(0)} = \sum_{j=1}^n \alpha_j v_j, \quad \alpha_1 = \left\langle x^{(0)}, \frac{v_1}{\|v_1\|} \right\rangle \neq 0,$$

so folgt für die Iterierten:

$$\begin{aligned} x^{(i)} &= A^i x^{(0)} = \sum_{j=1}^n \alpha_j \lambda_j^i v_j = \\ &= \alpha_1 \lambda_1^i \left\{ v_1 + \frac{\alpha_2}{\alpha_1} \left(\frac{\lambda_2}{\lambda_1}\right)^i v_2 + \dots + \frac{\alpha_n}{\alpha_1} \left(\frac{\lambda_n}{\lambda_1}\right)^i v_n \right\} \end{aligned}$$

Wegen $|\frac{\lambda_j}{\lambda_1}| < 1$ für $j=2, \dots, n$, konvergiert die Klammer $\{ \}$ und damit auch die Richtung von $x^{(i)}$ gegen v_1 ($i \rightarrow \infty$).

Ferner: Die Kvgz. ist umso schneller, je kleiner die Quotienten $|\lambda_j/\lambda_1|$ sind.

Zur Vermeidung von Exponenten über- oder unterlauf wird $x^{(i)}$ zumeist normiert:

$$x^{(i+1)} = Ax^{(i)} / \|Ax^{(i)}\|. \quad (10.21)$$

Für große i gilt dann $Ax^{(i)} \approx \lambda_1 x^{(i)}$, so dass sich auch λ_1 numerisch berechnen lässt.

Beispiel:

$$A = \begin{pmatrix} 1 & 0.1 & -0.1 \\ 0 & 2 & 0.4 \\ -0.2 & 0 & 3 \end{pmatrix}$$

$$x^{(0)} = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix}$$

$$x^{(25)} = (-2.7895 \cdot 10^{-2}, 3.6932 \cdot 10^{-1}, 9.2889 \cdot 10^{-1})$$

$$x^{(30)} = (-2.7894 \cdot 10^{-2}, 3.6933 \cdot 10^{-1}, 9.2888 \cdot 10^{-1})$$

$$Ax^{(30)} / x^{(30)} = (3.005998, 3.006011, 3.0060058)$$

Inverse Vektoriteration :

Zur Konvergenzbeschleunigung wird die Vektoriteration auf die Matrix $B = (A - \mu I)^{-1}$ angewendet.

In jedem Iterationsschritt ist dann ein lineares Gleichungssystem mit **universeller** Koeffizientenmatrix zu lösen!

Dabei ist μ eine Näherung für den gesuchten EW, etwa λ_k (einfach!).

$B = (A - \mu I)^{-1}$ hat die EWe $\tilde{\lambda}_j = \frac{1}{\lambda_j - \mu}$.

Das Konvergenzverhalten wird also durch die Quotienten

$$\left| \frac{\tilde{\lambda}_j}{\tilde{\lambda}_k} \right| = \left| \frac{\lambda_k - \mu}{\lambda_j - \mu} \right| \quad (j \neq k) \quad \underline{(10.22)}$$

bestimmt.

Iteration Für $i = 0, 1, 2, \dots$

$$\text{Löse } (A - \mu I) y = x^{(i)} \quad \underline{(10.23)}$$

$$\underline{x^{(i+1)} := y / \|y\|}$$

Beispiel:

$$A = \begin{pmatrix} 1 & 0.1 & -0.1 \\ 0 & 2 & 0.4 \\ -0.2 & 0 & 3 \end{pmatrix}, \quad \mu = 3$$

$$x^{(0)} = (0, 0, 1)$$

$$x^{(1)} = (-2.7843 \cdot 10^{-2}, 3.7125 \cdot 10^{-1}, 9.2812 \cdot 10^{-1})$$

$$x^{(2)} = (-2.7894 \cdot 10^{-2}, 3.6932 \cdot 10^{-1}, 9.2888 \cdot 10^{-1})$$

$$x^{(3)} = (-2.7893 \cdot 10^{-2}, 3.6933 \cdot 10^{-1}, 9.2888 \cdot 10^{-1})$$

$$Ax^{(3)} / x^{(3)} = \underline{\underline{(3.0060060, 3.0060056, 3.0060058)}}$$

Bemerkung:

Für gute Näherungen $\mu \approx \lambda_k$ ist die Matrix $B = A - \mu I$ numerisch singular.

Es treten dabei jedoch keine numerischen Probleme auf, da lediglich die Richtung von x zu bestimmen ist. Man kann ein kleines Pivotelement durch ϵ ersetzen. (Dabei muss der EW λ_k einfach sein!)

10.4 Reduktion auf Tridiagonalform

$A \in \mathbb{R}^{(n,n)}$: symmetrische Matrix

Zur Beschleunigung der numerischen EW-Verfahren ist es sinnvoll, die Matrix A zunächst durch orthogonale Ähnlichkeits-transformationen auf Tridiagonalform zu bringen. Hierzu lassen sich sowohl Givens wie Householder-Transformationen verwenden.

Wir beschreiben das Vorgehen mittels Householder-Transformationen:

$$\begin{pmatrix} * & * \text{---} * \\ * & * \text{---} * \\ \vdots & \vdots & \vdots \\ * & * \text{---} * \end{pmatrix} \xrightarrow{Q_1 A} \begin{pmatrix} * & * \text{---} * \\ \alpha_2 & * \text{---} * \\ 0 & \vdots & \vdots \\ 0 & * \text{---} * \end{pmatrix}$$

Hierbei wurden nur die Zeilen 2-n verändert.

$$\xrightarrow{Q_1 A Q_1^T} \begin{pmatrix} * & \alpha_2 & 0 \text{---} 0 \\ \alpha_2 & * & \text{---} * \\ 0 & \vdots & \vdots \\ 0 & * & \text{---} * \end{pmatrix}$$

Nach $(n-1)$ solcher Transformationen hat man

$$\tilde{A} = Q A Q^T = \begin{pmatrix} \delta_1 & \alpha_2 & & 0 \\ & \alpha_2 & & \\ & & \ddots & \\ 0 & & & \alpha_n & \delta_n \\ & & & & \alpha_n & \delta_n \end{pmatrix} \quad (10.24)$$

mit $Q := Q_{n-1} \cdot Q_{n-2} \cdot \dots \cdot Q_1$.

10.5 Eigenwerte für symmetr. Tridiagonalmatrizen

$A \in \mathbb{R}^{(n,n)}$ sei eine symmetrische Tridiagonalmatrix:

$$A = \begin{pmatrix} \delta_1 & \alpha_2 & & 0 \\ & \alpha_2 & & \\ & & \ddots & \\ 0 & & & \alpha_n & \delta_n \\ & & & & \alpha_n & \delta_n \end{pmatrix}$$

Die EWe von A werden als Nullstellen des charakt. Polynoms mittels Bisektion od. Newton-Verfahren berechnet.

Hierzu benutzt man, dass sich $p_A(\lambda)$ durch eine Dreiterm-Rekursion berechnen lässt.

Satz (10.25)

Zu vorgegebenem $\lambda \in \mathbb{R}$ wird berechnet

$$p_0 := 1, \quad p_1 := \delta_1 - \lambda$$

für $k=2, \dots, n$

$$p_k := (\delta_k - \lambda) p_{k-1} - \alpha_k^2 p_{k-2}$$

Damit gilt: $p_A(\lambda) = p_n$.

Beweis: Entwicklung von

$$p_k(\lambda) = \begin{vmatrix} (\delta_1 - \lambda) & \alpha_2 & & 0 \\ \alpha_2 & & & \\ & & & \alpha_k \\ 0 & & \alpha_k & (\delta_k - \lambda) \end{vmatrix}$$

nach der letzten Zeile! ▲

Die Besonderheit der in Satz (10.25) konstruierten Folge (p_0, \dots, p_n) besteht darin, dass sie eine so genannte Sturmsche Kette bildet.

Hieran lässt sich ablesen, wieviele Eigenwerte links von λ liegen.

Satz (10.26) Es gelte $\alpha_j \neq 0$, $j=2, \dots, n$, d.h. die Matrix A zerfällt nicht. Dann gelten:

- a) Alle EWe sind einfach $\lambda_1 > \lambda_2 > \dots > \lambda_n$.
- b) Für $\lambda \in \mathbb{R}$ gibt die Anzahl $w(\lambda)$ der Vorzeichenwechsel in der Folge p_0, \dots, p_n von (10.25) an, wieviele Eigenwerte $\lambda_i < \lambda$ sind.

Beweis: \rightarrow Stoer, Abschnitt 5.6.

Der Satz (10.26) wird dazu benutzt, gezielt den m -ten EW λ_m in der Nummerierung

$$\lambda_1 > \lambda_2 > \dots > \lambda_n \quad (10.26)$$

Dazu sei $\alpha_j \neq 0$ ($\forall j=2, \dots, n$) und $[\lambda_n, \lambda_0]$ ein Intervall, das den EW λ_m garantiert enthält (vgl. (10.12) bzw. (10.13)).

Zur Realisierung des Bisektionsverfahrens dividiert man die Rekursion (10.25) durch p_{k-1} und setzt $q_k := p_k / p_{k-1}$.

Man erhält dann etwa folgenden Algorithmus:

Algorithmus (10.27)

$\alpha_1 := 0$, eps : Maschinengenauigkeit

für iter = 1, 2, ...

$l := 0$, $q := 1$, $\lambda := (\lambda_u + \lambda_o)/2$,

falls $|\lambda_o - \lambda_u| \leq |\lambda| \cdot \text{eps}$: Stopp!

für $k = 1, \dots, n$

falls ($q = 0$) $q := |\alpha_k| / \text{eps}$

$q := \delta_k - \lambda - \alpha_k^2 / q$

falls ($q < 0$) $l := l + 1$

falls ($l < m$) $\lambda_u := \lambda$

sonst $\lambda_o := \lambda$

Zur Anwendung von Newton-Techniken benötigt

man neben $p_A(\lambda)$ auch die Ableitung $p'_A(\lambda)$.

Hierzu wird die Rekursion (10.25) differenziert:

$p_0 := 1$, $p'_0 := 0$, $p_1 := \delta_1 - \lambda$, $p'_1 := -1$,

für $k = 2, \dots, n$

$p_k := (\delta_k - \lambda) p_{k-1} - \alpha_k^2 p_{k-2}$

$p'_k := -p_{k-1} + (\delta_k - \lambda) p'_{k-1} - \alpha_k^2 p'_{k-2}$

(10.28)

10.6 QR-Algorithmus

Die gebräuchlichste – und inzwischen sehr ausgefeilte – Methode zur Berechnung von Eigenwerten und Eigenvektoren ist der so genannte QR-Algorithmus. Er geht zurück auf J.G. Francis (1961/62).

Wir beschränken uns wieder auf den Fall einer symmetrischen, reellen Tridiagonalmatrix.

In der einfachsten Form lautet der QR-Algorithmus dann:

$$\begin{array}{l}
 A_1 := A \\
 \text{für } k=1,2,\dots \\
 \quad \left| \begin{array}{l}
 \text{Berechne QR-Zerleg. } A_k := Q_k R_k \\
 A_{k+1} := R_k Q_k
 \end{array} \right.
 \end{array}
 \quad (10.29)$$

Satz (10.30)

- a) Alle Iterierten A_k sind ähnlich zu A ,
symmetrisch und tridiagonal.
- b) Mit $|\lambda_1| > |\lambda_2| > \dots > |\lambda_n| > 0$ gilt:

$$\lim_{k \rightarrow \infty} Q_k = I_n, \quad \lim_{k \rightarrow \infty} A_k = \text{diag}(\lambda_1, \dots, \lambda_n)$$

$$\text{und } |a_{i,i-1}^{(k)}| \leq C \left| \frac{\lambda_i}{\lambda_{i-1}} \right|^k, \quad i=2, \dots, n$$

Beweis:

zu a): $Q_1 A_2 Q_1^T = Q_1 (R_1 Q_1) Q_1^T = Q_1 R_1 = A_1$

$\Rightarrow A_2 = Q_1^T A_1 Q_1$ ist symmetrisch und ähnlich zu A_1 .

Zur Tridiagonalgestalt:

Man sieht mittels Gauß-Elimination, dass die Inverse einer regulären oberen Dreiecksmatrix wieder eine obere Dreiecksmatrix ist.

Daher folgt:

$$\begin{aligned} A_2 &= R_1 Q_1 = R_1 A_1 R_1^{-1} \\ &= \begin{pmatrix} * & \dots & * \\ & \ddots & \\ 0 & & * \end{pmatrix} \begin{pmatrix} * & * & 0 \\ * & \ddots & \\ 0 & & * \end{pmatrix} \begin{pmatrix} * & \dots & * \\ & \ddots & \\ 0 & & * \end{pmatrix} \\ &= \begin{pmatrix} * & \dots & * \\ & \ddots & \\ 0 & & * \end{pmatrix} \begin{pmatrix} * & \dots & * \\ * & \ddots & \\ 0 & & * \end{pmatrix} = \begin{pmatrix} * & \dots & * \\ * & \ddots & \\ 0 & & * \end{pmatrix} \end{aligned}$$

$\Rightarrow A_2$ obere Hessenberg-Matrix

$\Rightarrow A_2$ symmetrisch! A_2 Tridiagonalmatrix.

Zu b) (nur grober Beweisablauf, technische Details weglassen)

$$\textcircled{i} \quad A^k = (Q_1 \dots Q_k) (R_k \dots R_1) \quad (1)$$

$$\begin{aligned} \text{denn: } A_{k+1} &= R_k Q_k = Q_k^T A_k Q_k = \dots \\ &= Q_k^T \dots Q_1^T A Q_1 \dots Q_k \end{aligned}$$

$$\Rightarrow Q_1 \dots Q_k A_{k+1} = A Q_1 \dots Q_k \quad (2)$$

$$\Rightarrow [Q_1 \dots Q_k R_k \dots R_1] = Q_1 \dots Q_{k-1} A_k R_{k-1} \dots R_1$$

$$\stackrel{(2)}{=} A [Q_1 \dots Q_{k-1} R_{k-1} \dots R_1]$$

$$\stackrel{\cdot}{=} A^k \Rightarrow (1).$$

$$\textcircled{ii} \quad A = U^T \Lambda U, \quad U \text{ orthogonal,} \\ \Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$$

$$\Rightarrow A^k = U^T \Lambda^k U$$

LR-Zerleg.
von U

$$= U^T \Lambda^k L R$$

$$= U^T \underbrace{(\Lambda^k L \Lambda^{-k})}_{\text{LR-Zerleg. von } U} \Lambda^k R$$

$$= (I + E_k), \quad E_k \rightarrow 0 \quad (k \rightarrow \infty)$$

$$\textcircled{\text{iii}} \quad \text{QR-Zerlegung von } I + E_k = \tilde{Q}_k \tilde{R}_k \\ \Rightarrow \tilde{Q}_k \rightarrow I, \tilde{R}_k \rightarrow I \quad (k \rightarrow \infty)$$

$$A^k = (U^T \tilde{Q}_k) (\tilde{R}_k \Lambda^k R)$$


ist eine weitere QR-Zerlegung

$$\begin{aligned} \Rightarrow \left. \begin{array}{l} \text{"Eindeutigkeit"} \\ \text{und (1)} \end{array} \right\} \begin{array}{l} U^T \tilde{Q}_k \approx Q_1 \dots Q_k \\ \tilde{R}_k \Lambda^k R \approx R_k \dots R_1 \end{array} \quad (3) \end{aligned}$$

$$\begin{aligned} \textcircled{\text{iv}} \quad Q_k &= (Q_1 \dots Q_{k-1})^T (Q_1 \dots Q_k) \\ &\stackrel{(3)}{\approx} \tilde{Q}_{k-1}^T U U^T \tilde{Q}_k \\ &= \tilde{Q}_{k-1}^T \tilde{Q}_k \rightarrow I \quad (k \rightarrow \infty) \end{aligned}$$

$$\begin{aligned} R_k &= (R_k \dots R_1) (R_{k-1} \dots R_1)^{-1} \\ &\stackrel{(3)}{\approx} (\tilde{R}_k \Lambda^k R) (\tilde{R}_{k-1} \Lambda^{k-1} R)^{-1} \\ &\rightarrow \Lambda \quad (k \rightarrow \infty) \end{aligned}$$

Damit folgt auch $\lim_{k \rightarrow \infty} A_k = \lim_{k \rightarrow \infty} Q_k R_k = \Lambda$

$\textcircled{\text{v}}$ Die Konvergenzgeschwindigkeit ergibt sich aus der von E_k : $\sim \left(\frac{\lambda_j}{\lambda_{j-1}}\right)^k$ 

Bemerkung (10.31)

Wegen $A_{k+1} = (Q_1 \dots Q_k)^T A (Q_1 \dots Q_k)$ konvergiert die Folge der orthogonalen Matrizen $V_k := Q_1 \dots Q_k$ gegen eine orthogonale Matrix $V = (v_1, \dots, v_n)$, die die Eiven von A als Spaltenvektoren enthält.

Shift - Techniken

Analog zur inversen Vektoriteration lässt sich der QR-Algorithmus durch Verschiebung des Spektrums erheblich beschleunigen.

Der Algorithmus lautet dann:

$$A_1 := A \quad (10.32)$$

für $k = 1, 2, \dots$

	Bestimme Verschiebung $\sigma_k \in \mathbb{R}$
	QR-Zerlegung $A_k - \sigma_k I = Q_k R_k$
	$A_{k+1} = R_k Q_k + \sigma_k I$

σ_k wird so gewählt, dass das Spektrum von $A_k - \sigma_k I$ nahe bei Null liegt.

Schritt 2:

Anwendung des QR-Algorithmus auf $B^T B$, allerdings in einer Form, die nur auf B operiert (um eine schlechte Kondition von $B^T B$ zu vermeiden).

Man erhält eine Folge von Bidiagonal =
matrizen B_k mit

$$B_k^T B_k \rightarrow \text{diag}(\sigma_1^2, \dots, \sigma_n^2) \quad (k \rightarrow \infty)$$

und

$$B_k \rightarrow \text{diag}(\sigma_1, \dots, \sigma_n) \quad (k \rightarrow \infty)$$