

VII EIGENWERTAUFGABEN

S. BARTELS, 15.1.2014

VIII.A. Lokalisierung. Die Berechnung einzelner oder sämtlicher Eigenwerte einer Matrix und gegebenenfalls zugehöriger Eigenvektoren bezeichnet man als *Eigenwertaufgaben*. Im Allgemeinen ist es schwierig und ineffizient die Nullstellen eines charakteristischen Polynoms zu bestimmen, da schon die Auswertung des Polynoms mit hohem Aufwand verbunden ist.

Satz VIII.1. Sei $A \in \mathbb{R}^{n \times n}$ und $\lambda \in \mathbb{C}$ ein Eigenwert von A . Dann gilt

$$\lambda \in \bigcup_{i=1}^n K_i, \quad K_i = \left\{ z \in \mathbb{C} : |z - a_{ii}| \leq \sum_{j=1, \dots, n, j \neq i} |a_{ij}| \right\}.$$

Die Mengen K_i heißen Gerschgorin-Kreise.

Beweis. Es gelte $Ax = \lambda x$ für ein $x \in \mathbb{C}^n \setminus \{0\}$. Dann existiert ein i mit $|x_j| \leq |x_i|$ für alle $j = 1, 2, \dots, n$, und $x_i \neq 0$. Es gilt

$$\lambda x_i = (Ax)_i = \sum_{j=1}^n a_{ij} x_j$$

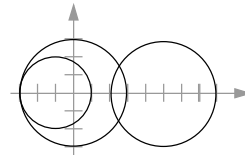
und nach Division durch $x_i \neq 0$ folgt

$$\lambda - a_{ii} = \sum_{j=1, \dots, n, j \neq i} a_{ij} \frac{x_j}{x_i}.$$

Die Dreiecksungleichung und $|x_j|/|x_i| \leq 1$ implizieren $\lambda \in K_i$ und damit die Behauptung. □

Beispiel VIII.2. Für die nachfolgende Matrix $A \in \mathbb{R}^{3 \times 3}$ ergeben sich die Gerschgorin-Kreise K_1, K_2, K_3 .

$$A = \begin{bmatrix} 5 & 1 & 2 \\ 1 & -1 & 1 \\ 2 & 1 & 0 \end{bmatrix}, \quad \begin{aligned} K_1 &= \{z \in \mathbb{C} : |z - 5| \leq 3\}, \\ K_2 &= \{z \in \mathbb{C} : |z + 1| \leq 2\}, \\ K_3 &= \{z \in \mathbb{C} : |z| \leq 3\}. \end{aligned}$$



Im Fall symmetrischer Matrizen lassen sich die Eigenwerte durch Extremwerte des sogenannten Rayleigh-Quotienten charakterisieren.

Satz VIII.3. Sei $A \in \mathbb{R}^{n \times n}$ symmetrisch. Für den maximalen beziehungsweise minimalen Eigenwert von A gilt

$$\lambda_{\min} = \min_{x \in \mathbb{R}^n \setminus \{0\}} \frac{x^\top Ax}{\|x\|_2^2}, \quad \lambda_{\max} = \max_{x \in \mathbb{R}^n \setminus \{0\}} \frac{x^\top Ax}{\|x\|_2^2}.$$

Beweis. Es sei $(v_1, v_2, \dots, v_n) \subset \mathbb{R}^n$ eine Orthonormalbasis des \mathbb{R}^n bestehend aus Eigenvektoren zu den Eigenwerten $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \in \mathbb{R}$ der Matrix A . Zu $x \in \mathbb{R}^n$ existieren $\alpha_1, \alpha_2, \dots, \alpha_n \in \mathbb{R}$, so dass $x = \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n$ und es gilt

$$Ax = \alpha_1 \lambda_1 v_1 + \alpha_2 \lambda_2 v_2 + \dots + \alpha_n \lambda_n v_n.$$

Die Orthonormalität $v_i \cdot v_j = \delta_{ij}$, $1 \leq i, j \leq n$, der Vektoren v_1, v_2, \dots, v_n impliziert

$$\begin{aligned} x^\top x &= \left(\sum_{i=1}^n \alpha_i v_i \right) \cdot \left(\sum_{j=1}^n \alpha_j v_j \right) = \sum_{i,j=1}^n \alpha_i \alpha_j v_i \cdot v_j = \sum_{i=1}^n \alpha_i^2, \\ x^\top Ax &= \left(\sum_{i=1}^n \alpha_i v_i \right) \cdot \left(\sum_{j=1}^n \alpha_j Av_j \right) = \sum_{i,j=1}^n \alpha_i \lambda_j \alpha_j v_i \cdot v_j = \sum_{i=1}^n \lambda_i \alpha_i^2. \end{aligned}$$

Daraus folgt

$$x^\top Ax \geq \lambda_n \sum_{i=1}^n \alpha_i^2 = \lambda_n \|x\|_2^2 = \lambda_{\min} \|x\|_2^2,$$

wobei Gleichheit für $x = v_n$ gilt. Die Aussage für $\lambda_1 = \lambda_{\max}$ folgt analog. \square

VIII.B. Konditionierung. Eine Matrix $A \in \mathbb{R}^{n \times n}$ ist komplex diagonalisierbar, wenn es eine reguläre Matrix $V \in \mathbb{C}^{n \times n}$ und eine Diagonalmatrix $D \in \mathbb{C}^{n \times n}$ gibt, so dass $A = VDV^{-1}$ gilt. In dieser Situation lässt sich folgendes Resultat von Bauer und Fike beweisen.

Satz VIII.4. Sei $A \in \mathbb{R}^{n \times n}$ komplex diagonalisierbar mit $A = VDV^{-1}$, sei $E \in \mathbb{R}^{n \times n}$ und sei $\tilde{\lambda} \in \mathbb{C}$ ein Eigenwert von $A + E$. Dann existiert ein komplexer Eigenwert von A , so dass

$$|\tilde{\lambda} - \lambda| \leq \text{cond}_2(V) \|E\|_2,$$

wobei die Operatornorm und Konditionszahl in naheliegender Weise für komplexe Matrizen verallgemeinert seien.

Beweis. Siehe Vorlesung. \square

Bemerkungen VIII.5. (i) Nicht jede Matrix ist komplex diagonalisierbar, jedoch ist jede Matrix komplex trigonalisierbar.

(ii) Normale Matrizen, das heißt Matrizen mit der Eigenschaft $AA^\top = A^\top A$, sind komplex diagonalisierbar mit unitärer Transformationsmatrix V , das heißt V erfüllt $\bar{V}^\top V = I_n$.

Korollar VIII.6. Sei $A \in \mathbb{R}^{n \times n}$ normal, $E \in \mathbb{R}^{n \times n}$ und $\tilde{\lambda}$ ein Eigenwert von $A + E$. Dann existiert ein Eigenwert von A mit

$$|\lambda - \tilde{\lambda}| \leq \|E\|_2$$

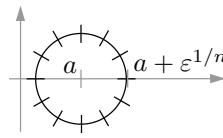
Beweis. Da A diagonalisierbar ist mit einer unitären Matrix $V \in \mathbb{C}^{n \times n}$ und $\text{cond}_2(V) = 1$ gilt, folgt die Abschätzung aus dem vorangegangenen Satz. \square

Für normale, insbesondere symmetrische Matrizen ist die Bestimmung der Eigenwerte damit ein gut konditioniertes Problem. Im Allgemeinen ist dies nicht der Fall.

Beispiel VIII.7. Ist $p(t) = t^n + a_{n-1}t^{n-1} + \dots + a_1t + a_0$ ein beliebiges Polynom, so gilt $p(t) = (-1)^n \det(A - tI_n)$ mit der Frobenius-Begleitmatrix

$$A = \begin{bmatrix} 0 & & & & -a_0 \\ 1 & 0 & & & -a_1 \\ & \ddots & \ddots & & \vdots \\ & & & 1 & 0 \\ & & & & 1 & -a_{n-1} \end{bmatrix},$$

insbesondere entsprechen die komplexen Nullstellen von p den komplexen Eigenwerten von A . Für $a \in \mathbb{R} \setminus \{0\}$ und $\varepsilon > 0$ hat das Polynom $p_0(t) = (t-a)^n$ die n -fache Nullstelle $\lambda = a$, während das Polynom $p_\varepsilon(t) = (t-a)^{n-\varepsilon}$ die Nullstellen $\lambda_k = a - \varepsilon^{1/n} e^{i2\pi k/n}$, $k = 1, 2, \dots, n$, besitzt. Die Polynome p_0 und p_ε unterscheiden sich nur im konstanten Koeffizienten und für die Differenz $A - A_\varepsilon$ der zugehörigen Begleitmatrizen gilt $\|A - A_\varepsilon\|_\ell = \varepsilon$ für $\ell \in \{1, 2, \infty\}$. Es gilt $|\lambda - \lambda_k| = \varepsilon^{1/n}$ und für die relativen Fehler folgt

$$\begin{aligned} \frac{|\lambda - \lambda_k|}{|\lambda|} &= \frac{\varepsilon^{1/n} \|A\|_\ell \|A - A_\varepsilon\|_\ell}{|a| \|A\|_\ell \varepsilon} \\ &= \frac{\varepsilon^{1/n} \|A\|_\ell \|A - A_\varepsilon\|_\ell}{\varepsilon |a| \|A\|_\ell}. \end{aligned}$$


Der Faktor $\varepsilon^{(1-n)/n}$ ist unbeschränkt für $\varepsilon \rightarrow 0$, sofern $n > 1$ gilt.

VIII.C. Potenzmethode. Sei $A \in \mathbb{R}^{n \times n}$ reell diagonalisierbar mit Eigenwerten $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{R}$ und zugehörigen linear unabhängigen Eigenvektoren $v_1, v_2, \dots, v_n \in \mathbb{R}^n$, für die $\|v_i\|_2 = 1$, $i = 1, 2, \dots, n$, gelte. Für jedes $x \in \mathbb{R}^n$ mit

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

ergibt die k -malige Anwendung von A auf v

$$A^k x = A^{k-1} \left(\sum_{i=1}^n \lambda_i \alpha_i v_i \right) = \dots = \sum_{i=1}^n \lambda_i^k \alpha_i v_i.$$

Ist λ_1 der betragsmäßig größte Eigenwert, so folgt für k hinreichend groß, dass

$$A^k x \approx \alpha_1 \lambda_1^k v_1.$$

Wir betrachten die Normen $\|A^k x\|_2$ sowie $\|A^{k+1} x\|_2$ und bilden deren Quotient, so dass wegen $\|v_1\|_2 = 1$ folgt

$$\frac{\|A^{k+1} x\|_2}{\|A^k x\|_2} \approx |\lambda_1|.$$

Mit dieser Beobachtung kann man den dominierenden Eigenwert einer Matrix bestimmen.

Algorithmus VIII.8 (von Mises-Potenzmethode). Seien $A \in \mathbb{R}^{n \times n}$, $x \in \mathbb{R}^n \setminus \{0\}$ und $\varepsilon_{stop} > 0$. Setze $x_0 = x/\|x\|_2$, $\mu_0 = 0$ und $k = 0$.

- (1) Berechne $\tilde{x}_{k+1} = Ax_k$, $\mu_{k+1} = \|\tilde{x}_{k+1}\|_2$ und $x_{k+1} = \tilde{x}_{k+1}/\|\tilde{x}_{k+1}\|_2$.
- (2) Stoppe, falls $|\mu_{k+1} - \mu_k| \leq \varepsilon_{stop}$ gilt; erhöhe $k \rightarrow k + 1$ und wiederhole Schritt (1) andernfalls.

Bemerkung VIII.9. Induktiv zeigt man, dass $x_k = A^k x / \|A^k x\|_2$ gilt. Um ein Verlassen des Bereichs darstellbarer Zahlen zu vermeiden, muss in jedem Iterationsschritt eine Normierung durchgeführt werden.

Wir zeigen, dass die Folge $(x_k)_{k \in \mathbb{N}}$ einen normierten Eigenvektor zum betragsmäßig maximalen Eigenwert von A approximiert, dessen Betrag durch die Folge $(\mu_k)_{k \in \mathbb{N}}$ angenähert wird.

Satz VIII.10. Es gelte $|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_n| \geq 0$ und es sei $x = \sum_{i=1}^n \alpha_i v_i$ mit den normierten Eigenvektoren v_1, v_2, \dots, v_n von A . Gilt $\alpha_1 \neq 0$ so folgt mit $q = |\lambda_2|/|\lambda_1| < 1$ und $k \geq K$, dass

$$\left| \|Ax_k\|_2 - |\lambda_1| \right| \leq 4\|A\|_2 C q^k$$

mit einer von k unabhängigen Konstanten $C \geq 0$.

Beweis. Siehe Vorlesung. □

Bemerkungen VIII.11. (i) In jedem Schritt der Iteration verringert sich der Approximationsfehler um den Faktor $q < 1$.

(ii) Die letzte Ungleichung im Beweis zeigt, dass $\lambda_1 < 0$ genau dann gilt, wenn die Vorzeichen von $(x_k)_{k=1,2,\dots}$ alternieren, und dass die Folge $(x_k)_{k=1,2,\dots}$ bis auf den Faktor $\text{sign } \lambda_1^k$ gegen einen Eigenvektor konvergiert.

(iii) Die Voraussetzung $\alpha_1 \neq 0$ muss im konkreten Fall sichergestellt werden. Aufgrund von Rundungsfehlern kann dies zwar angenommen werden, allerdings kann dann die Konstante $C \sim 1/|\alpha_1|$ sehr groß werden.

Im Fall symmetrischer Matrizen lässt sich eine bessere Konvergenzaussage beweisen, die impliziert, dass der Fehler in jedem Schritt um den Faktor q^2 verringert wird.

Satz VIII.12. Ist $A \in \mathbb{R}^{n \times n}$ symmetrisch, so gilt unter den Voraussetzungen des vorigen Satzes

$$|\lambda_1 - x_k^\top A x_k| \leq 2\|A\|_2 C^2 q^{2k}.$$

Beweis. Siehe Vorlesung. □

Bemerkungen VIII.13. (i) Gilt $0 < |\lambda_n| < |\lambda_{n-1}| \leq \dots \leq |\lambda_1|$, so liefert die Potenzmethode mit A^{-1} statt A eine Approximation von $|\lambda_n|^{-1}$. Dies wird als inverse Iteration bezeichnet.

(ii) Wendet man die Potenzmethode auf die Matrix $(A - \mu I_n)^{-1}$ an, so konvergiert sie unter geeigneten Voraussetzungen gegen den Eigenwert, der

am nächsten bei μ liegt.

(iii) Der dominierende Eigenwert kann mehrfach auftreten, das heißt die Voraussetzung der Sätze lässt sich auf $|\lambda_1| = \dots = |\lambda_p| > |\lambda_{p+1}| \geq \dots \geq |\lambda_n| \geq 0$ abschwächen.

VIII.D. QR-Verfahren. Das QR-Verfahren berechnet unter geeigneten Voraussetzungen Approximationen sämtlicher Eigenwerte einer Matrix.

Algorithmus VIII.14 (QR-Verfahren). Sei $A \in \mathbb{R}^{n \times n}$. Setze $A_0 = A$ und $k = 0$.

- (1) Bestimme die QR-Zerlegung $A_k = Q_k R_k$ und setze $A_{k+1} = R_k Q_k$.
- (2) Erhöhe $k \rightarrow k + 1$ und fahre fort mit (1).

Das Verfahren lässt sich als verallgemeinerte Potenzmethode interpretieren und die Iterierten sind ähnlich zueinander.

Lemma VIII.15. Es gilt

$$\begin{aligned} A_{k+1} &= Q_k^\top A_k Q_k = (Q_0 \dots Q_k)^\top A (Q_0 \dots Q_k), \\ A^{k+1} &= (Q_0 \dots Q_k) (R_k \dots R_0). \end{aligned}$$

Beweis. Siehe Vorlesung. □

Mit Hilfe dieses Lemmas und einer Stabilitätseigenschaft der QR-Zerlegung lässt sich das folgende Resultat beweisen.

Satz VIII.16. Sei $A \in \mathbb{R}^{n \times n}$ diagonalisierbar mit $A = V D V^{-1}$ derart, dass für die Eigenwerte $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{R}$ gilt

$$|\lambda_1| > |\lambda_2| > \dots > |\lambda_n| > 0$$

und die Inverse der Matrix V eine LU-Zerlegung besitzt. Dann gilt

$$\|\text{diag}(A_k) - \text{diag}(D)\|_2 \leq C q^k$$

mit $q = \max_{1 \leq i < j \leq n} |\lambda_j|/|\lambda_i|$ und einer Konstanten $C \geq 0$.

Bemerkungen VIII.17. (i) In der Praxis beobachtet man Konvergenz unter deutlich schwächeren Voraussetzungen an A .

(ii) Im Allgemeinen führt ein Schritt im QR-Verfahren auf einen Aufwand der Ordnung $\mathcal{O}(n^3)$. Wird A zunächst mittels Householder-Transformationen auf eine sogenannte Hessenberg-Matrix

$$\widehat{A} = H^\top A H = \begin{bmatrix} \widehat{a}_{11} & \dots & & & \widehat{a}_{1n} \\ \widehat{a}_{21} & \widehat{a}_{22} & \dots & & \widehat{a}_{2n} \\ & \widehat{a}_{32} & \dots & & \widehat{a}_{3n} \\ & & \ddots & & \vdots \\ & & & \widehat{a}_{n,n-1} & \widehat{a}_{nn} \end{bmatrix}$$

transformiert, das heißt gilt $\widehat{a}_{ij} = 0$ für $i > j + 1$, so lässt sich die QR-Zerlegung mit Givens-Rotationen in $\mathcal{O}(n^2)$ Schritten bestimmen.

VIII.E. Jacobi-Verfahren. Der Satz über die Gerschgorin-Kreise zeigt, dass die Diagonaleinträge einer Matrix Approximationen der Eigenwerte definieren, die besonders gut sind, wenn die Nichtdiagonalelemente klein sind. Im Jacobi-Verfahren werden diese Einträge einer symmetrischen Matrix sukzessive mit Ähnlichkeitstransformationen verringert.

Definition VIII.18. Für $A \in \mathbb{R}^{n \times n}$ sei

$$\mathcal{N}(A) = \|A\|_{\mathcal{F}}^2 - \sum_{i=1}^n a_{ii}^2 = \sum_{1 \leq i, j \leq n, i \neq j} a_{ij}^2.$$

Offensichtlich ist A eine Diagonalmatrix genau dann, wenn $\mathcal{N}(A) = 0$ gilt. Allgemeiner lässt sich zeigen, dass zu jedem Diagonaleintrag a_{jj} , $1 \leq j \leq n$, ein Eigenwert λ mit der Eigenschaft $|\lambda - a_{jj}| \leq \sqrt{\mathcal{N}(A)}$ existiert. Aus dem Gerschgorinschen Kreisesatz erhält man die schwächere Aussage, dass zu jedem Eigenwert λ von A ein Diagonaleintrag a_{jj} existiert, so dass $|\lambda - a_{jj}| \leq (n-1)^{1/2} \sqrt{\mathcal{N}(A)}$ gilt.

Definition VIII.19. Für $c, s \in \mathbb{R}$ mit $c^2 + s^2 = 1$ und $1 \leq p, q \leq n$ wird eine Givens-Rotation $G_{pq} \in O(n)$ definiert durch

$$(G_{pq})_{ij} = \begin{cases} 1, & i = j, i \neq p, \\ 1, & i = j, i \neq q, \\ c, & i = p, j = p, \\ c, & i = q, j = q, \\ s, & i = q, j = p, \\ -s, & i = p, j = q, \\ 0 & \text{sonst.} \end{cases} \quad G_{pq} = \begin{bmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & c & & s & & & & \\ & & & \ddots & & & & & \\ & & & & -s & & c & & \\ & & & & & & & \ddots & \\ & & & & & & & & 1 \end{bmatrix}$$

Im folgenden Lemma wird verwendet, dass die Frobenius-Norm invariant ist unter orthogonalen Transformationen, das heißt dass $\|Q^T M\|_{\mathcal{F}} = \|MQ\|_{\mathcal{F}} = \|M\|_{\mathcal{F}}$ für alle $M \in \mathbb{R}^{n \times n}$ und $Q \in O(n)$ gilt. Dies folgt zum Beispiel aus $\|M\|_{\mathcal{F}}^2 = \text{tr}(M^T M)$ und $\|M\|_{\mathcal{F}} = \|M^T\|_{\mathcal{F}}$.

Lemma VIII.20. Ist $A \in \mathbb{R}^{n \times n}$ symmetrisch und G_{pq} eine beliebige Givens-Rotation, so gilt für $B = G_{pq}^T A G_{pq}$, dass

$$\mathcal{N}(B) = \mathcal{N}(A) - 2(a_{pq}^2 - b_{pq}^2),$$

wobei $b_{pq} = cs(a_{qq} - a_{pp}) + (c^2 - s^2)a_{pq}$.

Beweis. Siehe Vorlesung. □

Kann die Givens-Rotation G_{pq} so gewählt werden, dass $b_{pq} = 0$ gilt, so ergibt sich eine Verringerung der Nichtdiagonaleinträge. Durch Betrachtung von $c = \cos(\alpha)$, $s = \pm \sin(\alpha)$ und $D = \cos(2\alpha)$ erhält man folgendes Resultat.

Lemma VIII.21. *Ist $a_{pq} \neq 0$ und G_{pq} definiert durch $c = \sqrt{(1+D)/2}$ und $s = -\text{sign}(a_{pq})\sqrt{(1-D)/2}$ mit*

$$D = \frac{a_{pp} - a_{qq}}{((a_{pp} - a_{qq})^2 + 4a_{pq}^2)^{1/2}} \in [-1, 1]$$

so gilt $b_{pq} = 0$.

Beweis. Siehe Übung. □

Um eine größtmögliche Reduktion von $\mathcal{N}(A)$ zu erzielen, ist es naheliegend, das betragsmäßig größte Nichtdiagonalelement von A auszuwählen.

Satz VIII.22. *Ist a_{pq} das betragsmäßig größte Nichtdiagonalelement von A , so gilt mit der im vorigen Lemma definierten Givens-Rotation G_{pq} für die Matrix $B = G_{pq}^\top A G_{pq}$ und mit $\varepsilon_n = 2/(n(n-1))$, dass*

$$\mathcal{N}(B) \leq (1 - \varepsilon_n)\mathcal{N}(A),$$

Beweis. Siehe Vorlesung. □

Aus dem Satz folgt die Konvergenz des folgenden Verfahrens.

Algorithmus VIII.23 (Jacobi-Verfahren). *Sei $A \in \mathbb{R}^{n \times n}$ symmetrisch. Setze $A_0 = A$ und $k = 0$.*

- (1) *Seien p, q die Indizes des betragsmäßig größten Nichtdiagonalelements von A_k und wähle die Givens-Rotation G_{pq} so, dass für $A_{k+1} = G_{pq}^\top A_k G_{pq}$ der Eintrag $(A_{k+1})_{pq}$ verschwindet.*
- (2) *Erhöhe $k \rightarrow k + 1$ und fahre fort mit (1).*

Bemerkungen VIII.24. (i) *Im Allgemeinen werden $\mathcal{O}(n^2 \log(1/\delta))$ viele Iterationen benötigt, um $\mathcal{N}(A_k) \leq \delta$ zu garantieren.*

(ii) *Ein bereits zu Null transformierter Eintrag kann im Laufe des Verfahrens wieder von Null abweichen.*

(iii) *Das Verfahren konstruiert eine Faktorisierung $A = G D G^\top$ mit einer orthogonalen Matrix G und einer näherungsweise Diagonalmatrix D , insbesondere sind damit die Spaltenvektoren von G näherungsweise Eigenvektoren von A .*

(iv) *Da die Suche nach dem maximalen Nichtdiagonalelement sehr aufwändig ist, arbeitet man in der Praxis eher sukzessive alle Nichtdiagonalelemente ab und wiederholt dies so oft, bis $\mathcal{N}(A_k)$ hinreichend klein ist. Dieses Vorgehen bezeichnet man als zyklisches Jacobi-Verfahren.*