Mathematik für InformatikerInnen 2

Frank-Olaf Schreyer

Universität des Saarlandes, SS 2020

Matrixnormen und Vektoriterationen

Die Themen heute sind

- Matrixnormen
- Eigenwertabschätzungen
- Beweisschritte zum Satz von Gerschgorin
- Vektoriteration

Heute werden wir verschiedene Techniken, um Abschätzungen für die Eigenwerte einer Matrix zu bekommen, besprechen.

Matrixnormen

Definition. Eine Matrixnorm ist eine Funktion

$$\| \quad \| : \mathbb{R}^{n\times n} \to \mathbb{R}$$

mit folgenden Eigenschaften:

1. $\|A\| \geqslant 0 \quad \forall A \in \mathbb{R}^{n\times n}$ und $\|A\| = 0 \Leftrightarrow A = 0$,
2. $\|\lambda A\| = |\lambda|\|A\| \quad \forall \lambda \in \mathbb{R}, \forall A \in \mathbb{R}^{n\times n}$,
3. $\|A + B\| \leqslant \|A\| + \|B\| \quad \forall A, B \in \mathbb{R}^{n\times n}$,
4. $\|A \cdot B\| \leqslant \|A\| \cdot \|B\| \quad \forall A, B \in \mathbb{R}^{n\times n}$ (Submultiplizität).

Beispiele. Sei $A = (a_{ik}) \in \mathbb{R}^{n\times n}$. Dann haben wir folgende Matrixnormen:

a) die **Gesamtnorm**

$$\|A\|_G = n \max_{i,j} |a_{ik}|,$$

b) die **Zeilensummennorm**

$$\|A\|_Z = \max_i \sum_{k=1}^{n} |a_{ik}|,$$

c) die **Spaltensummennorm**

$$\|A\|_S = \max_k \sum_{i=1}^{n} |a_{ik}|,$$

d) die **Frobeniusnorm**

$$\|A\|_F = \Big(\sum_{i,k=1}^{n} |a_{ik}|^2 \Big)^{\frac{1}{2}},$$

e) die **Spektalnorm**

$$\|A\|_2 = \sqrt{\lambda_{\max}(A^t A)},$$

wobei $\lambda_{\max}(A^t A)$ der größte Eigenwert von $A^t A$ ist.

Da Matrizen und Vektoren oft gemeinsam auftreten, sollten Matrixnormen und Vektornormen miteinander verträglich sein.

Definition. Eine Matrixnorm $\| \ \|_M$ heißt **verträglich** mit einer Vektornorm $\| \ \|_V$, falls

$$\|Ax\|_V \leqslant \|A\|_M \|x\|_V \quad \forall A \in \mathbb{R}^{n\times n}, \forall x \in \mathbb{R}^n$$

gilt.

Beispiele. Zu den p-Normen auf $\mathbb{R}^n$

$$\|x\|_p = \begin{cases} \left(\sum_{i=1}^n |x_i|^p\right)^{\frac{1}{p}} & \text{für } 1 \leqslant p < \infty, \\ \max_i |x_i| & \text{für } p = \infty \end{cases}$$

bestehen folgende Verträglichkeiten:

i) $\|A\|_G$ und $\|A\|_S$ sind mit der Betragssummennorm $\|x\|_1$ verträglich,
ii) $\|A\|_G, \|A\|_F$ und $\|A\|_2$ sind mit der euklidischen Norm $\|x\|_2$ verträglich,
iii) $\|A\|_G$ und $\|A\|_Z$ sind mit der Maximumnorm $\|x\|_\infty$ verträglich.

Beweis. Wir zeigen exemplarisch: $\|A\|_G$ und $\|x\|_\infty$ sind verträglich.

$$\begin{aligned}
\|Ax\|_\infty &= \max_i \{ | \sum_{k=1}^n a_{ik} x_k | \} \\
&\leqslant \max_i \{ \sum_{k=1}^n |a_{ik} x_k| \} \qquad (\Delta\text{-Ungleichung}) \\
&\leqslant \max \{ \sum_{k=1}^n \max_{r,s} |a_{rs}| \cdot \max_\ell |x_\ell| \} \\
&= n \max_{r,s} |a_{rs}| \cdot \max_\ell |x_\ell| = \|A\|_G \cdot \|x\|_\infty
\end{aligned}$$

□

Da zu einer gegebenen Vektornorm $\| \, \|_V$ oftmals viele verträgliche Matrixnormen $\| \, \|_M$ existieren, verwendet man in der Praxis oft gerne diejenige, für die die Abschätzung

$$\|Ax\|_V \leqslant \|A\|_M \cdot \|x\|_V$$

am schärfsten ist.

Ende Teil 1

Definition. Zu einer gegebenen Vektornorm $\| \ \|$ definieren wir die **zugehörige Matrixnorm** durch

$$\|A\| = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \max_{\|x\|=1} \|Ax\|.$$

Bemerkung. Dies ist in der Tat eine Matrixnorm.

Beispiele. Einige zugehörige Matrixnormen sind die folgenden:

Vektornorm	zugehörige Matrixnorm
Betragssummennorm $\|x\|_1$	Spaltensummennorm $\|A\|_S$
euklidische Norm $\|x\|_2$	Spektralnorm $\|A\|_2$
Maximumnorm $\|x\|_\infty$	Zeilensummennorm $\|A\|_Z$

Matrixnormen sind nützlich für Eigenwertabschätzungen.

Eigenwertabschätzungen mit Matrixnormen

Satz. *Ist λ ein Eigenwert von $A \in \mathbb{R}^{n \times n}$ und $\|A\|$ eine Matrixnorm, die mit einer Vektornorm $\|\ \|$ verträglich ist, dann gilt*

$$|\lambda| \leqslant \|A\|.$$

Beweis. Sei $x \in \mathbb{R}^n$ ein Eigenvektor zu λ. Dann gilt:

$$|\lambda| \cdot \|x\| = \|\lambda x\| = \|Ax\| \leqslant \|A\| \|x\|.$$

Da $\|x\| \neq 0$, folgt die Behauptung. □

Korollar.

$$|\lambda| \leqslant \max_i \{ \sum_{k=1}^{n} |a_{ik}| \} \leqslant n \cdot \max_{i,k} \{ |a_{ik}| \}.$$ □

Beispiel.

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

$$\begin{aligned}
\|A\|_G &= 3 \cdot \max_{i,k}\{|a_{ik}|\} = 3 \cdot 3 = 9 \\
\|A\|_Z &= \max\{1.2, 2.4, 3.2\} = 3.2 \\
\|A\|_S &= \max\{1.2, 2.1, 3.5\} = 3.5 \\
\|A\|_F &= (1^2 + (0.1)^2 + \qquad\qquad\qquad\qquad +3^2)^{\frac{1}{2}} = \sqrt{14.22} \approx 3.77
\end{aligned}$$

$\|A\|_Z = 3.2$ liefert die schärfste Abschätzung. Tatsächlich gilt

$$\lambda_1 \approx 3.0060, \lambda_2 \approx 2.0078, \lambda_3 \approx 0.9862.$$

Der Satz von Gerschgorin gibt:

Ende Teil 2

Beweisschritte zum Satz von Gerschgorin

Satz. *Sei* $A = (a_{jk}) \in \mathbb{C}^{n\times n}$ *eine quadratische komplexe Matrix. Es bezeichne*

$$D_j = \{z \in \mathbb{C} \mid |z - a_{jj}| \leqslant \sum_{k\neq j} |a_{jk}|\}.$$

Dann liegen alle Eigenwerte λ *von* A *in der Vereinigung* $\bigcup D_j$*, wobei jede zusammenhängende Komponente der Vereinigung genauso viele Eigenwerte enthält wie Kreise an deren Vereinigung beteiligt sind.*

Wir zeigen:
Jeder Eigenwert λ *liegt in wenigstens einem Gerschgorin-Kreis.*
Beweis. Sei $x \in \mathbb{C}^n$ ein Eigenvektor zu λ. Wir wählen x so, dass eine Komponente $x_j = 1$ und für alle Komponenten $|x_k| \leqslant 1$ gilt. Wir erreichen dies, indem wir x durch $\frac{1}{x_j}x$ ersetzen, wobei j ein Index mit $|x_j| = \max_k\{|x_k|\}$ ist. Dann gilt:

$$\lambda x_j = \sum_{k=1}^{n} a_{jk}x_k \Leftrightarrow \lambda - a_{jj} = \sum_{k\neq j} a_{jk}x_k$$

Die Dreiecksungleichung liefert

$$|\lambda - a_{jj}| \leqslant \sum_{k \neq j} |a_{jk}||x_k| \leqslant \sum_{k \neq j} |a_{jk}|,$$

d.h., λ liegt im j-ten Gerschgorin-Kreis. $\square$

Die Umkehrung, jeder Gerschgorin-Kreis enthält einen Eigenwert, ist im allgemeinen nicht wahr.

Beispiel. Die Matrix $A = \begin{pmatrix} 0 & 1 \\ 4 & 0 \end{pmatrix}$ hat die Eigenwerte $\lambda = \pm 2$.

Beide sind in dem Gerschgorin-Kreis $\{|z| \leqslant 1\}$ nicht enthalten.

Der restliche Beweis beruht darauf, dass die Nullstellen eines Polynoms stetig von den Koeffizienten abhängen.
Wir betrachten die 1-Parameterfamilie von Matrizen

$$B(s) = (1-s)D + sA \quad \text{für } s \in [0,1],$$

wobei D die Diagonalmatrix mit den gleichen Diagonaleinträgen wie A ist. Die Mittelpunkte der Gerschgorin-Kreise von $B(s)$ sind die gleichen wie die von A, die Radien wachsen mit s.

Sei nun K eine Zusammenhangskomponente von $\bigcup_j D_j$, und D_k ein Gerschgorin-Kreis mit $D_k \cap K = \varnothing$. Dann hat D_k zu K einen positiven Abstand.

Wir behaupten nun:

Jedes $B(s)$ hat mit algebraischen Vielfachheiten gezählt genau ℓ Eigenwerte in K, wobei ℓ die Anzahl der Kreise bezeichnet, die an K beteiligt sind.

Dies ist klar für $B(0)$. Hätte nun $A = B(1)$ mehr oder weniger Eigenwerte in K, dann müsste zwischendurch ein Eigenwert $\lambda_k(s)$ von $B(s)$ die Komponente wechseln.

Das ist nicht möglich, da die $\lambda_k(s)$ stetig von s abhängig gewählt werden können. Für einen Zwischenwert $\tilde{s}$ müsste dann ein $\lambda_k(\tilde{s})$ außerhalb der Gerschgorin-Kreise sein, was dem ersten Teil des Beweises widerspricht. $\square$

Bemerkung. Dass die $\lambda_k(s)$ für $k = 1, \dots, n$ mit Anfangswerten $\lambda_k(0) = a_{kk}$ stetig gewählt werden können, ist vielleicht nicht so offensichtlich. Eigenwerte $\lambda_k(s)$ könnten zum Beispiel zusammenlaufen. Im Prinzip ist dies eine Aussage über die Geometrie der Nullstellenmenge C des Polynoms

$$f(s,t) = \det((1-s)D + sA - tE) \in \mathbb{C}[s,t].$$

Man nennt die **Nullstellenmenge**

$$C = \{(w,z) \in \mathbb{C}^2 \mid f(w,z) = 0\}$$

eines Polynoms $f(s,t) \in \mathbb{C}[s,t]$ eine **ebene algebraische Kurve**. Wir sagen eben, weil sie in der komplexen Ebene $\mathbb{C}^2$ enthalten ist, und sprechen von einer Kurve, weil sie endlich über der s-Achse liegt, was ein 1-dimensionaler $\mathbb{C}$-Vektorraum ist.

Beispiel. Die Nullstellenmenge des Polynoms $f(s,t) = t^2 - s$ hat für jeden Punkt $w \in \mathbb{C}\setminus\{0\}$ zwei Lösungen. Für reelle $w > 0$ sind dies $z = \pm\sqrt{w}$. Für reelle $w < 0$ sind es $z = \pm i\sqrt{|w|}$, ein Paar von konjugiert komplexen Lösungen. In Polarkoordinaten sehen wir:

Über dem Kreis $\{w \in \mathbb{C} \mid |w| = r\}$ ist die Lösungsmenge durch

$$\{\begin{pmatrix} w \\ z \end{pmatrix} = \begin{pmatrix} r(\cos 2\alpha + i\sin 2\alpha) \\ \sqrt{r}(\cos\alpha + i\sin\alpha) \end{pmatrix} \in \mathbb{C}^2 \mid \alpha \in [0, 2\pi]\}$$

parametrisiert.

Bemerkung. Wegen solchen Effekten können die

$$\lambda_k : [0,1] \to \mathbb{C}, \quad s \mapsto \lambda_k(s)$$

nur stetig, aber nicht differenzierbar gewählt werden.

Ende Teil 3

Vektoriteration

Die Vektoriteration ist eine Methode, mit der man häufig den betragsmäßig größten Eigenwert approximativ betimmen kann. Die Idee ist folgende:

Sei $A \in \mathbb{R}^{n \times n}$ eine (nicht notwendig symmetrische) Matrix. Für einen beliebigen normierten Startvektor

$$w_0 \in \mathbb{R}^n$$

berechnen wir rekursiv

$$w_{k+1} = \frac{1}{\|Aw_k\|} Aw_k$$

Unsere Hoffnung ist, das

$$v = \lim_{k \to \infty} w_k$$

existiert und ein Eigenvektor von A ist.

Satz. *Sei* $A \in \mathbb{R}^{n\times n}$ *eine symmetrische Matrix und seien* $\lambda_1, \dots, \lambda_n$ *deren Eigenwerte. Sei* $\lambda_{\max} = \max_j\{|\lambda_j|\} > 0$ *und*

$$W = \mathrm{Eig}(A, \lambda_{\max}) \oplus \mathrm{Eig}(A, -\lambda_{\max})$$

die Summe der Eigenräume der betragsmäßig größten Eigenwerte. Für einen (normierten) Startvektor

$$w_0 \notin W^\perp \subsetneq \mathbb{R}^n,$$

konvergiert die Teilfolge (w_{2k}) *der rekursiv definierten Folge*

$$(w_k) \text{ mit } w_{k+1} = \frac{1}{\|Aw_k\|} Aw_k$$

gegen einen Eigenvektor v *von* A^2 *zum Eigenwert* $\lambda_{\max}^2$*. Ist* $-\lambda_{\max}$ *kein Eigenwert, dann konvergiert schon die Folge* (w_k) *gegen einen Eigenvektor* v *von* A *zum Eigenwert* $\lambda_{\max}$*.*

Beweis. Sei $\mathcal{B} = \{v_1, \ldots, v_n\}$ eine Orthonormalbasis aus Eigenvektoren von A. Der Startvektor w_0 ist eine Linearkombination von diesen, etwa

$$w_0 = \sum_{j=1}^{n} \mu_j v_j \quad \text{mit } \sum_j \mu_j^2 = 1,$$

da w_0 normiert ist.

$$Aw_0 = \sum_{j=1}^{n} \mu_j \lambda_j v_j \quad \text{und} \quad w_1 = \frac{1}{\|Aw_0\|} Aw_0 = \sum_{j=1}^{n} \mu_j^{(1)} v_j.$$

Wir nehmen nun an, dass $|\lambda_1| = \lambda_{\max}$ und $\mu_1 \neq 0$ gilt. Dann folgt:

$$\frac{\mu_j^{(1)}}{\mu_1^{(1)}} = \frac{\lambda_j}{\lambda_1} \cdot \frac{\mu_j}{\mu_1} \quad \text{und} \quad \frac{\mu_j^{(k)}}{\mu_1^{(k)}} = \left(\frac{\lambda_j}{\lambda_1}\right)^k \cdot \frac{\mu_j}{\mu_1}.$$

Für j mit $|\lambda_j| < \lambda_{\max}$ konvergiert der relative Anteil von w_k in Richtung v_j also gegen Null.

Für gerades $k = 2\ell$ bleibt für j mit $|\lambda_j| = \lambda_{\max}$ der relative Anteil

$$\frac{\mu_j^{(2\ell)}}{\mu_1^{(2\ell)}} = \frac{\mu_j}{\mu_1}$$

unverändert. Es folgt, die Folge $(w_{2\ell})$ konvergiert gegen

$$v = \frac{1}{\|\tilde{w}\|}\tilde{w}, \text{ wobei } \quad \tilde{w} = \sum_{j:|\lambda_j|=\lambda_{\max}} \mu_j v_j$$

die Projektion von w_0 auf $W = \mathrm{Eig}(A, \lambda_{\max}) \oplus \mathrm{Eig}(A, -\lambda_{\max})$ bezeichnet. Sind alle λ_j mit $|\lambda_j| = \lambda_{\max}$ positiv, dann konvergiert schon (w_k) gegen v. $\square$

Beispiel. Für $A = \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{pmatrix}$ gibt die Vektoriteration Folgen $(w_0, w_1, w_2, \ldots, w_{10}, \ldots, w_{20}, \ldots)$, die konvergieren.

$$\begin{pmatrix} .5 \\ .5 \\ .5 \\ .5 \end{pmatrix}, \begin{pmatrix} .707107 \\ 0 \\ 0 \\ .707107 \end{pmatrix}, \begin{pmatrix} .632456 \\ -.316228 \\ -.316228 \\ .632456 \end{pmatrix}, \begin{pmatrix} .601501 \\ -.371748 \\ -.371748 \\ .601501 \end{pmatrix}, \begin{pmatrix} .601501 \\ -.371748 \\ -.371748 \\ .601501 \end{pmatrix}.$$

Für den ersten Startwert $w_0 = (0.5\, 0.5\, 0.5\, 0.5)^t$ konvergiert die Folge gegen einen Eigenvektor zum Eigenwert $\lambda_2 \approx 2.61803$, welcher allerdings nicht $\lambda_{\max}$ ist.

Die zweite Folge mit Startwert $w_0 = (0.2 \ -0.4 \ 0.4 \ -0.8)^t$ konvergiert gegen einen Eigenvektor zum Eigenwert $\lambda_1 = \lambda_{\max} \approx 3.61803$:

$$\begin{pmatrix} .2 \\ -.4 \\ .4 \\ -.8 \end{pmatrix}, \begin{pmatrix} .245718 \\ -.430007 \\ .614295 \\ -.614295 \end{pmatrix}, \begin{pmatrix} .261991 \\ -.48905 \\ .646245 \\ -.523982 \end{pmatrix}, \begin{pmatrix} .361695 \\ -.595219 \\ .607599 \\ -.381726 \end{pmatrix}, \begin{pmatrix} .371354 \\ -.601257 \\ .601744 \\ -.372142 \end{pmatrix}.$$

In der Tat, das charakteristische Polynom

$$\chi_A(t) = \left(t^2 - 5\,t + 5\right)\left(t^2 - 3\,t + 1\right)$$

faktorisiert, und die Eigenwerte sind

$$\lambda_1 = \frac{5+\sqrt{5}}{2}, \lambda_2 = \frac{3+\sqrt{5}}{2}, \lambda_3 = \frac{5-\sqrt{5}}{2}, \lambda_2 = \frac{3-\sqrt{5}}{2}.$$

Bemerkung. In der Situation vom Satz kann man den Anteil des Grenzwerts in $\mathrm{Eig}(A, \lambda_{\max})$ und $\mathrm{Eig}(A, -\lambda_{\max})$ trennen, indem man die Grenzwert v_+ und v_- der Folgen $(\frac{1}{2}(w_{2k+1} + w_{2k}))$ und $(\frac{1}{2}(w_{2k+1} - w_{2k}))$ betrachtet.

Beispiel.

$$A = \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & -2 & -1 \\ 0 & 0 & -1 & -2 \end{pmatrix}$$

hat charakteristisches Polynom $\chi_A(t) = t^4 - 11\,t^2 + 13$ und Eigenwerte

$$\lambda_j = \pm\sqrt{\frac{11 \pm \sqrt{69}}{2}},$$

also $\lambda_{\max} \approx 3.10698$. Mit Startwert $w_0 = (0.5, -0.5, 0.5, 0.5)^t$ nehmen w_{18}, w_{19}, w_{20} die folgenden Werte an.

Auf 6 Stellen genau sind dies

$$\begin{pmatrix} .525742 \\ -.464986 \\ .586499 \\ .404230 \end{pmatrix}, \begin{pmatrix} .488086 \\ -.657299 \\ -.357982 \\ -.448976 \end{pmatrix}, \begin{pmatrix} .525742 \\ -.464986 \\ .586499 \\ .404230 \end{pmatrix}.$$

Es folgt:

$$\tilde{v}_+ = \frac{1}{2}(w_{20} + w_{19}) \text{ und } \tilde{v}_- = \frac{1}{2}(w_{20} - w_{19})$$

sind Approximationen von Eigenvektor v_+ und v_- zum Eigenwert $\lambda_{\max}$ bzw. $-\lambda_{\max}$ von A. Es gilt:

$$A\tilde{v}_+ - \tilde{\lambda}_{\max}\tilde{v}_+ = \begin{pmatrix} 6.76046 \cdot 10^{-9} \\ -7.66281 \cdot 10^{-9} \\ 6.33198 \cdot 10^{-10} \\ 7.27885 \cdot 10^{-10} \end{pmatrix}$$

für

$$\tilde{\lambda}_{\max} = \frac{1}{4}\sum_{j=1}^{4}(A\tilde{v}_+)_i/(\tilde{v}_+)_i.$$

Ende Teil 4