II OPERATORNORM UND KONDITIONSZAHL

S. BARTELS, 30.10.2013

II.A. **Vektornormen.** Um die Begriffe der Konditionierung und Stabilität präzisieren zu können, müssen Abstände zwischen Punkten im $\mathbb{R}^n$ beziehungsweise Längen von Vektoren gemessen werden können.

Definition II.1. *Eine* Norm *auf* $\mathbb{R}^n$ *ist eine Abbildung* $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}$ *mit den folgenden Eigenschaften:*
(i) $\|x\| = 0 \implies x = 0$ *für alle* $x \in \mathbb{R}^n$ *(Definitheit);*
(ii) $\|x+y\| \le \|x\| + \|y\|$ *für alle* $x, y \in \mathbb{R}^n$ *(Dreiecksungleichung);*
(iii) $\|\lambda x\| = |\lambda|\|x\|$ *für alle* $\lambda \in \mathbb{R}$ *und* $x \in \mathbb{R}^n$ *(Homogenität).*

Beispiel II.2. *Die* ℓ^p*-Normen sind für* $1 \le p \le \infty$ *und* $x = [x_1, \dots, x_n]^\top \in \mathbb{R}^n$ *definiert durch*

$$\|x\|_p = \begin{cases} \Big(\sum_{j=1}^n |x_j|^p\Big)^{1/p}, & p < \infty, \\ \max_{j=1,\dots,n} |x_j|, & p = \infty. \end{cases}$$

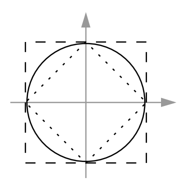
Die Norm $\|\cdot\|_2$ *heißt* euklidische Norm *und erfüllt* $\|x\|_2^2 = x \cdot x = x^\top x$.

Bemerkungen II.3. *(i) Die* ℓ^p*-Normen sind äquivalent in dem Sinne, dass für alle* $1 \le p, q \le \infty$ *eine Konstante* $C_{pq} > 0$ *existiert, so dass für alle* $x \in \mathbb{R}^n$ *gilt*

$$C_{pq}^{-1}\|x\|_p \le \|x\|_q \le C_{pq}\|x\|_p.$$

Die Konstante C_{pq} *ist abhängig von* p, q *und* n.
(ii) Die ℓ^p*-Normen unterscheiden sich durch ihre Niveaumengen*

$$N_p(1) = \{x \in \mathbb{R}^n : \|x\|_p = 1\}.$$



II.B. **Matrixnormen.** Wir identifizieren im Folgenden stets Matrizen $A \in \mathbb{R}^{m\times n}$ mit linearen Abbildungen $A : \mathbb{R}^n \to \mathbb{R}^m$, wobei dabei die jeweiligen kanonischen Basen gewählt seien. Eine lineare Abbildung wird auch als linearer Operator bezeichnet.

Definition II.4. *Für Normen* $\|\cdot\|_{\mathbb{R}^m}$ *und* $\|\cdot\|_{\mathbb{R}^n}$ *auf* $\mathbb{R}^m$ *beziehungsweise* $\mathbb{R}^n$ *ist die* (induzierte) Operatornorm *für alle* $A \in \mathbb{R}^{m\times n}$ *definiert durch*

$$\|A\|_{op} = \sup_{x\in\mathbb{R}^n, \|x\|_{\mathbb{R}^n}=1} \|Ax\|_{\mathbb{R}^m}.$$

Die Operatornorm misst, wie stark Niveaumengen verformt werden.

Beispiel II.5. *Durch eine symmetrische Matrix $A \in \mathbb{R}^{2\times 2}$ wird die kreisförmige Niveaumenge $N_2(1)$ auf eine Ellipse abgebildet, die im Kreis mit Radius $\|A\|_2$ enthalten ist.*

Die Operatornorm definiert eine Norm mit folgenden Eigenschaften.

Lemma II.6. *Zu fixierten Normen $\|\cdot\|$ auf $\mathbb{R}^n$ beziehungsweise $\mathbb{R}^m$ sei $\|\cdot\|_{op}$ die induzierte Operatornorm auf $\mathbb{R}^{m\times n}$. Dann gilt:*
(i) $\|\cdot\|_{op}$ definiert eine Norm auf $\mathbb{R}^{m\times n}$;
(ii) $\|A\|_{op} = \sup_{x\in\mathbb{R}^n, \|x\|=1} \|Ax\| = \inf\{C > 0 : \forall x \in \mathbb{R}^n \; \|Ax\| \le C\|x\|\}$;
(iii) für $A \neq 0$ und $x \in \mathbb{R}^n$ mit $\|x\| \le 1$ und $\|Ax\| = \|A\|_{op}$ folgt $\|x\| = 1$;
(iv) das Infimum und das Supremum in (ii) werden angenommen.

Beweis. Übung. □

Bemerkung II.7. *Aus (ii) folgt $\|Ax\| \le \|A\|_{op}\|x\|$ für alle $x \in \mathbb{R}^n$.*

Für einige ℓ^p-Normen lassen sich die induzierten Operatornormen explizit angeben. Die Einträge einer Matrix $A \in \mathbb{R}^{m\times n}$ seien mit a_{ij}, $1 \le i \le m$, $1 \le j \le n$, bezeichnet.

Beispiel II.8. *(i) Die ℓ^1-Norm auf $\mathbb{R}^m$ und $\mathbb{R}^n$ induziert die* Spaltensummennorm

$$\|A\|_1 = \max_{j=1,\dots,n} \sum_{i=1}^{m} |a_{ij}|.$$

(ii) Die ℓ^∞-Norm auf $\mathbb{R}^m$ und $\mathbb{R}^n$ induziert die Zeilensummennorm

$$\|A\|_\infty = \max_{i=1,\dots,m} \sum_{j=1}^{n} |a_{ij}|.$$

(iii) Die ℓ^2-Norm auf $\mathbb{R}^m$ und $\mathbb{R}^n$ induziert die Spektralnorm

$$\|A\|_2 = \sqrt{\varrho(A^\top A)} = \big(\max\{|\lambda| : \lambda \textit{ ist Eigenwert von } A^\top A\}\big)^{1/2}.$$

Die Zahl $\varrho(A^\top A)$ heißt Spektralradius *von $A^\top A$.*

Einige weitere Eigenschaften der Operatornorm sind die folgenden.

Lemma II.9. *Seien Normen auf $\mathbb{R}^\ell$, $\mathbb{R}^m$ und $\mathbb{R}^n$ fixiert und die induzierten Operatornormen mit $\|\cdot\|$ bezeichnet.*
(i) Für $A \in \mathbb{R}^{\ell\times m}$ und $B \in \mathbb{R}^{m\times n}$ gilt $\|AB\| \le \|A\|\|B\|$.
(ii) Die Einheitsmatrix $I_n \in \mathbb{R}^{n\times n}$ erfüllt $\|I_n\| = 1$.
(iii) Jede induzierte Operatornorm auf $\mathbb{R}^{n\times n}$ erfüllt $\|A\|_{op} \ge |\lambda|$ für alle symmetrischen Matrizen $A \in \mathbb{R}^{n\times n}$ und jeden Eigenwert λ von A.

Beweis. Nach dem vorigen Lemma gilt $\|ABx\| \le \|A\|\|Bx\| \le \|A\|\|B\|\|x\|$ und dies impliziert $\|AB\| \le \|A\|\|B\|$. Die anderen Aussagen folgen unmittelbar aus der Definition der Operatornorm. □

Die euklidische Norm lässt sich in naheliegender Weise auf $\mathbb{R}^{m\times n}$ definieren, ist aber keine induzierte Operatornorm.

Beispiel II.10. *Die Frobenius-Norm einer Matrix $A \in \mathbb{R}^{m\times n}$ ist definiert durch $\|A\|_{\mathcal{F}} = \left(\sum_{i=1}^m \sum_{j=1}^n a_{ij}^2\right)^{1/2}$. Sie ist keine induzierte Operatornorm für $n > 1$, da $\|I_n\|_{\mathcal{F}} = \sqrt{n}$ gilt. Auch die skalierte Frobeniusnorm $n^{-1/2}\|A\|_{\mathcal{F}}$ ist keine induzierte Operatornorm, denn diese verletzt die Eigenschaft $\|A\|_{op} \geq |\lambda|$ für jeden Eigenwert λ von A.*

II.C. **Konditionszahl.** Mit Hilfe des Begriffs der Operatornorm lässt sich die Konditionierung eines linearen Gleichungssystems präzisieren.

Theorem II.11. *Sei $\|\cdot\|$ eine Operatornorm auf $\mathbb{R}^{n\times n}$. Sei $A \in \mathbb{R}^{n\times n}$ regulär und seien $x, \widetilde{x}, b, \widetilde{b} \in \mathbb{R}^n$ so, dass*

$$Ax = b, \quad A\widetilde{x} = \widetilde{b}.$$

Dann gilt

$$\frac{\|x-\widetilde{x}\|}{\|x\|} \leq \|A\|\|A^{-1}\|\frac{\|b-\widetilde{b}\|}{\|b\|}$$

Beweis. Es gilt $\|x-\widetilde{x}\| = \|A^{-1}(b-\widetilde{b})\| \leq \|A^{-1}\|\|b-\widetilde{b}\|$ sowie $\|b\| = \|Ax\| \leq \|A\|\|x\|$ beziehungsweise $\|x\| \geq \|A\|^{-1}\|b\|$. Damit folgt

$$\frac{\|x-\widetilde{x}\|}{\|x\|} \leq \frac{\|A^{-1}\|\|b-\widetilde{b}\|}{\|x\|} \leq \frac{\|A^{-1}\|\|b-\widetilde{b}\|}{\|A\|^{-1}\|b\|},$$

also die behauptete Abschätzung. □

Das Produkt $\|A\|\|A^{-1}\|$ kontrolliert die Verstärkung des relativen Fehlers beim Lösen eines linearen Gleichungssystems.

Definition II.12. *Die* Konditionszahl *einer regulären Matrix $A \in \mathbb{R}^{n\times n}$ bezüglich der durch die Norm $\|\cdot\|$ auf $\mathbb{R}^n$ induzierten Operatornorm ist definiert durch*

$$\mathrm{cond}_{\|\cdot\|}(A) = \|A\|\|A^{-1}\|.$$

Im Fall einer ℓ^p-Norm schreiben wir cond_p statt $\mathrm{cond}_{\|\cdot\|_p}$.

Bemerkungen II.13. *(i) Die Konditionszahl einer Matrix ist stets nach unten durch 1 beschränkt, denn für jede Operatornorm gilt $1 = \|AA^{-1}\| \leq \|A\|\|A^{-1}\| = \mathrm{cond}_{\|\cdot\|}(A)$.*
(ii) Ist A symmetrisch mit Eigenwerten $\lambda_1, \ldots, \lambda_n$, so gilt

$$\mathrm{cond}_2(A) = \frac{\max_{j=1,\ldots,n}|\lambda_j|}{\min_{k=1,\ldots,n}|\lambda_k|}.$$

Wir betrachten die Konditionszahl der Matrix aus dem eingehenden Beispiel **??**, in dem Störungen der rechten Seite große Fehler verursachten.

Beispiel II.14. *Die Matrix* $A = \begin{bmatrix} 1 & 1 \\ 1 & 1+\varepsilon \end{bmatrix}$ *hat die Eigenwerte* $\lambda_{1,2} = 1+\varepsilon/2 \pm (1+\varepsilon^2/4)^{1/2}$*. Mit der Taylor-Approximation* $(1+x)^{1/2} \approx 1 + x/2$ *folgt, dass für kleine Zahlen* ε *gilt* $\lambda_1 \approx 2 + \varepsilon/2$ *und* $\lambda_2 \approx \varepsilon/2$*. Damit folgt* $\mathrm{cond}_2(A) \approx 4\varepsilon^{-1}$*, was das empfindliche Verhalten zugehöriger Gleichungssysteme gegenüber Störungen erklärt.*

Geometrisch interpretiert misst die Konditionszahl die durch die lineare Abbildung A definierte Verzerrung, ist jedoch unabängig von uniformen Skalierungen.

Beispiel II.15. *Für eine symmetrische Matrix* $A \in \mathbb{R}^{2\times 2}$ *beschreibt* $\mathrm{cond}_2(A)$ *das Verhältnis der Radien der Ellipse* $A(N_2(1))$*.*