

Matrizennorm

Es seien $r, s \in \mathbb{N}$. Mit $M_{r,s}(\mathbb{R})$ bezeichnen wir die Menge der reellen $r \times s$ -Matrizen (also der linearen Abbildungen $\mathbb{R}^s \rightarrow \mathbb{R}^r$), und setze $M_s(\mathbb{R}) := M_{s,s}(\mathbb{R})$ (also die Menge der linearen Abbildungen $\mathbb{R}^s \rightarrow \mathbb{R}^s$). Analog könnte man die komplexen Matrizen als $M_{r,s}(\mathbb{C})$, bzw. $M_s(\mathbb{C})$ definieren. Wir werden zwar die Resultate (und Definitionen) hier nur für reelle Matrizen formulieren und beweisen (bzw. definieren), aber sie wären prinzipiell auch für komplexe Matrizen richtig (bei unverändertem Beweis).

Definition 1. Sei $A \in M_{r,s}(\mathbb{R})$. Dann heißt $\|A\| := \sup_{x \in \mathbb{R}^s \setminus \{0\}} \frac{|Ax|}{|x|}$ die Matrixnorm von A .

Wir wissen zunächst nicht, ob $\|A\|$ eine reelle Zahl ist. Wenn $\{\frac{|Ax|}{|x|} : x \in \mathbb{R}^s \setminus \{0\}\}$ nicht nach oben beschränkt wäre, dann setzen wir $\|A\| := +\infty$. Aus (2) in der folgenden Proposition 1 wird aber folgen, dass $\|A\|$ stets eine reelle Zahl ist.

Proposition 1. Sei $A \in M_{r,s}(\mathbb{R})$. Dann gelten folgende Eigenschaften.

- (1) $\|A\| = \sup_{\substack{x \in \mathbb{R}^s \\ |x|=1}} |Ax|$.
- (2) $\|A\| \in \mathbb{R}$.
- (3) Für alle $x \in \mathbb{R}^s$ gilt $|Ax| \leq \|A\||x|$.
- (4) $\|A\| \geq 0$.
- (5) Es gilt $\|A\| = 0$ genau dann, wenn $A = 0$.
- (6) Für jedes $\lambda \in \mathbb{R}$ gilt $\|\lambda A\| = |\lambda| \|A\|$.
- (7) Für alle $B \in M_{r,s}(\mathbb{R})$ gilt $\|A + B\| \leq \|A\| + \|B\|$.
- (8) Falls $B \in M_{p,r}(\mathbb{R})$, dann gilt $\|BA\| \leq \|B\| \|A\|$.

Beweis. (1) Sei $x \in \mathbb{R}^s$ mit $|x| = 1$. Wegen $|x| = 1$ ist $x \neq 0$, also $x \in \mathbb{R}^s \setminus \{0\}$. Dann ist $|Ax| = \frac{|Ax|}{|x|} \leq \|A\|$. Deshalb ist $\sup_{\substack{x \in \mathbb{R}^s \\ |x|=1}} |Ax| \leq \|A\|$.

Jetzt sei $x \in \mathbb{R}^s \setminus \{0\}$. Dann ist $|x| \neq 0$. Setze $y := \frac{1}{|x|}x$. Es ist $|y| = \left| \frac{1}{|x|}x \right| = \frac{1}{|x|}|x| = 1$. Somit ist $\frac{|Ax|}{|x|} = \frac{1}{|x|}|Ax| = \left| \frac{1}{|x|}Ax \right| = \left| A \left(\underbrace{\frac{1}{|x|}x}_{=y} \right) \right| = |Ay| \leq$

$\sup_{\substack{x \in \mathbb{R}^s \\ |x|=1}} |Ax|$. Daher ist $\|A\| \leq \sup_{\substack{x \in \mathbb{R}^s \\ |x|=1}} |Ax|$. Wegen $\sup_{\substack{x \in \mathbb{R}^s \\ |x|=1}} |Ax| \leq \|A\|$ ergibt sich daraus $\|A\| = \sup_{\substack{x \in \mathbb{R}^s \\ |x|=1}} |Ax|$.

(2) Definiere $C := \{x \in \mathbb{R}^s : |x| = 1\}$. Diese Menge ist beschränkt (weil ja $|x| \leq 1$ für alle $x \in C$). Da $x \mapsto |x|$ stetig ist, $\{1\}$ in \mathbb{R} abgeschlossen ist und C das Urbild von $\{1\}$ unter dieser Funktion ist, ist C als stetiges Urbild einer abgeschlossenen Menge abgeschlossen. Weil C beschränkt und abgeschlossen ist, ist C kompakt. Die Funktion $x \mapsto |Ax|$ ist stetig. Nach dem Satz vom Minimum und Maximum besitzt daher diese Funktion ein Maximum x_0 auf C . Wegen (1) ist $\|A\| = \sup_{\substack{x \in \mathbb{R}^s \\ |x|=1}} |Ax| = |Ax_0| \in \mathbb{R}$.

(3) Für $x = 0$ ergibt sich wegen $|x| = 0$ und $Ax = 0$, dass $|Ax| = |0| = 0 = \|A\|0 = \|A\||x|$. Sei jetzt $x \neq 0$. Dann ist $\frac{|Ax|}{|x|} \leq \|A\|$. Somit ist $|Ax| \leq \|A\||x|$.

(4) Es ist wegen (1) $\|A\| = \sup_{\substack{x \in \mathbb{R}^s \\ |x|=1}} |Ax| \geq 0$.

(5) Wegen (1) ist $\|0\| = \sup_{\substack{x \in \mathbb{R}^s \\ |x|=1}} |0x| = 0$. Für $A \neq 0$ gibt es ein $x_0 \in \mathbb{R}^s \setminus \{0\}$ mit $Ax_0 \neq 0$. Deshalb ist $\|A\| \geq \frac{|Ax_0|}{|x_0|} > 0$.

(6) Für $x \in \mathbb{R}^s$ mit $|x| = 1$ gilt $|(\lambda A)x| = |\lambda(Ax)| = |\lambda||Ax|$. Daher gilt $\| \lambda A \| = \sup_{\substack{x \in \mathbb{R}^s \\ |x|=1}} |(\lambda A)x| = |\lambda| \sup_{\substack{x \in \mathbb{R}^s \\ |x|=1}} |Ax| = |\lambda| \|A\|$.

(7) Sei $x \in \mathbb{R}^s$ mit $|x| = 1$. Es ist

$$\underbrace{|(A+B)x|}_{=Ax+Bx} \leq \underbrace{|Ax|}_{\leq \|A\| \text{ wegen (1)}} + \underbrace{|Bx|}_{\leq \|B\| \text{ wegen (1)}} \leq \|A\| + \|B\|.$$

Nach (1) folgt daraus, dass $\|A+B\| \leq \|A\| + \|B\|$.

(8) Wieder sei $x \in \mathbb{R}^s$ mit $|x| = 1$. Dann ist $|(AB)x| = \underbrace{|A(Bx)|}_{\leq \|A\| \text{ wegen (3)}} \leq \|A\| \underbrace{|Bx|}_{\leq \|B\| \text{ wegen (3)}} = \|A\| \|B\| |x| = \|A\| \|B\|$. Wegen (1) ergibt sich daraus, dass $\|AB\| \leq \|A\| \|B\|$. \square

Bemerkung. Falls die Eigenschaften (4)–(7) aus Proposition 1 erfüllt sind, dann spricht man von einer Norm. Ist zusätzlich auch die Eigenschaft (8) aus Proposition 1 erfüllt, dann spricht man von einer Matrixnorm oder Operatornorm.

Wir betrachten jetzt $(M_{r,s}(\mathbb{R}), \|\cdot\|)$, also der Raum der reellen $r \times s$ -Matrizen mit der Metrik $d(A, B) := \|A - B\|$. Nachdem man $M_{r,s}(\mathbb{R})$ mit \mathbb{R}^{rs} identifizieren kann, haben wir auf dieser Menge bereits eine andere Metrik, nämlich diejenige, die vom Betrag gegeben ist ($d(A, B) := |A - B|$). Diese beiden Metriken wollen wir jetzt vergleichen. Nachdem wir in Proposition 2 zeigen werden, dass es positive Konstanten α, β gibt mit $\alpha|A| \leq \|A\| \leq \beta|A|$, sind die offenen und abgeschlossenen Mengen bezüglich beider Metriken dieselben. Ebensovienig unterscheiden sich kompakte und zusammenhängende Mengen, und auch Konvergenz bedeutet dasselbe. Nachdem wir wissen, dass $(\mathbb{R}^{rs}, |\cdot|)$ vollständig ist, ist auch $(M_{r,s}(\mathbb{R}), \|\cdot\|)$ vollständig.

Proposition 2. *Sei $A \in M_{r,s}(\mathbb{R})$. Dann gilt $\frac{1}{\sqrt{rs}}|A| \leq \|A\| \leq |A|$. Insbesondere ist $(M_{r,s}(\mathbb{R}), \|\cdot\|)$ vollständig.*

Beweis. Es sei $A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,s} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,s} \\ \vdots & \vdots & \ddots & \vdots \\ a_{r,1} & a_{r,2} & \cdots & a_{r,s} \end{pmatrix}$ und $x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_s \end{pmatrix}$ so, dass $|x| = 1$.

Dann ist $Ax = \begin{pmatrix} \sum_{j=1}^s a_{1,j}x_j \\ \sum_{j=1}^s a_{2,j}x_j \\ \vdots \\ \sum_{j=1}^s a_{r,j}x_j \end{pmatrix}$. Für ein festes $k \in \{1, 2, \dots, r\}$ ist wegen der

Cauchy-Schwarz'schen Ungleichung $\left(\sum_{j=1}^s a_{k,j}x_j\right)^2 \leq \left(\sum_{j=1}^s a_{k,j}^2\right) \underbrace{|x|^2}_{=1} = \sum_{j=1}^s a_{k,j}^2$. Deshalb ist

$$|Ax| = \sqrt{\sum_{k=1}^r \left(\sum_{j=1}^s a_{k,j}x_j\right)^2} \leq \sqrt{\sum_{k=1}^r \sum_{j=1}^s a_{k,j}^2} = |A|.$$

Wegen (1) aus Proposition 1 ist $\|A\| = \sup_{\substack{x \in \mathbb{R}^s \\ |x|=1}} \underbrace{|Ax|}_{\leq |A|} \leq |A|$.

Wähle $j \in \{1, 2, \dots, s\}$ und $k \in \{1, 2, \dots, r\}$. Definiere v_j als denjenigen Vektor in \mathbb{R}^s , der in der j -ten Komponente 1 und in den anderen Komponenten 0 hat. Offensichtlich ist $|v_j| = 1$ und $Av_j = \begin{pmatrix} a_{1,j} \\ a_{2,j} \\ \vdots \\ a_{r,j} \end{pmatrix}$. Der Betrag von Av_j ist größer oder gleich dem Betrag seiner k -ten Komponente, also $|Av_j| \geq |a_{k,j}|$. Nach (1) aus Proposition 1 ist $|a_{k,j}| \leq |Av_j| \leq \|A\|$ und somit $a_{k,j}^2 \leq \|A\|^2$. Daher ist $|A|^2 = \sum_{k=1}^r \sum_{j=1}^s \underbrace{a_{k,j}^2}_{\leq \|A\|^2} \leq rs\|A\|^2$. Daraus ergibt

sich $\frac{1}{\sqrt{rs}}|A| \leq \|A\|$. □

Unser nächstes Resultat ist der Majorantentest für Reihen von Matrizen. Eigentlich formulieren wir nur das Analogon dazu, dass jede absolut konvergente Reihe konvergiert. Aus dem gewöhnlichen Majorantentest ergibt sich aber daraus auch der Majorantentest für Matrizenreihen.

Proposition 3. *Es sei $(A_n)_{n \in \mathbb{N}}$ eine Folge in $M_{r,s}(\mathbb{R})$. Falls $\sum_{n=1}^{\infty} \|A_n\|$ konvergiert, dann konvergiert $\sum_{n=1}^{\infty} A_n$.*

Beweis. Für $n \in \mathbb{N}$ setze $S_n := \sum_{k=1}^n A_k$ und $s_n := \sum_{k=1}^n \|A_k\|$. Weil $\sum_{n=1}^{\infty} \|A_n\|$ konvergiert ist $(s_n)_{n \in \mathbb{N}}$ eine Cauchyfolge. Sei $\varepsilon > 0$. Dann gibt es ein N , sodass $|s_n - s_m| < \varepsilon$ für alle $n, m \geq N$. Seien $n, m \geq N$. Ohne Beschränkung der Allgemeinheit können wir $m < n$ annehmen. Es ist $S_n - S_m = \sum_{k=1}^n A_k - \sum_{k=1}^m A_k = \sum_{k=m+1}^n A_k$ und analog $s_n - s_m = \sum_{k=m+1}^n \|A_k\|$. Wegen (7) aus Proposition 1 gilt $\|\sum_{k=m+1}^n A_k\| \leq \sum_{k=m+1}^n \|A_k\|$ und daher

$$\| \underbrace{S_n - S_m}_{=\sum_{k=m+1}^n A_k} \| = \|\sum_{k=m+1}^n A_k\| \leq \underbrace{\sum_{k=m+1}^n \|A_k\|}_{=s_n - s_m} = s_n - s_m \leq |s_n - s_m| < \varepsilon. \text{ Also}$$

ist $(S_n)_{n \in \mathbb{N}}$ eine Cauchyfolge. Da $M_{r,s}(\mathbb{R})$ vollständig ist konvergiert $(S_n)_{n \in \mathbb{N}}$, und dieser Grenzwert ist ja nach der Definition der Reihe $\sum_{n=1}^{\infty} A_n$. \square

In unendlichdimensionalen Räumen gibt es lineare Abbildungen A und B sodass $B \circ A = \text{id}$, aber A nicht invertierbar ist. Das nächste Resultat zeigt, dass solche Beispiele im Endlichdimensionalen nicht möglich sind.

Proposition 4. *Sei $A \in M_s(\mathbb{R})$. Falls es ein $B \in M_s(\mathbb{R})$ mit $BA = \text{id}$ gibt, dann ist A invertierbar und es gilt $B = A^{-1}$.*

Beweis. Angenommen $\text{rg } A < s$. Dann wäre $\dim \text{im } A = \text{rg } A < s$ und daher $\dim \text{im } BA \leq \dim \text{im } A < s$. Wegen $BA = \text{id}$ ist aber $\dim \text{im } BA = s$, wodurch sich ein Widerspruch ergibt. Es muss daher $\dim \text{im } A = \text{rg } A = s$ gelten. Wegen $\dim \text{im } A = s$ ist A surjektiv. Aus $s = \underbrace{\dim \text{im } A}_{=s} + \dim \ker A$ er-

halten wir $\dim \ker A = 0$, und somit ist A injektiv. Es ist also A bijektiv, und deshalb besitzt A eine Inverse A^{-1} . Weiters ist $B = B \underbrace{\text{id}}_{=AA^{-1}} = \underbrace{BA}_{=\text{id}} A^{-1} =$

$\text{id } A^{-1} = A^{-1}$. \square

Unser nächstes Resultat zeigt, dass jede Matrix, die in der Matrixnorm um weniger als 1 von der Einheitsmatrix abweicht, invertierbar ist. Die Reihendarstellung von A^{-1} wird die *von Neumann'sche Reihe* genannt.

Proposition 5. *Sei $A \in M_s(\mathbb{R})$. Falls $\|A - \text{id}\| < 1$ gilt, dann ist A invertierbar. Weiters gelten $\|A^{-1}\| \leq \frac{1}{1 - \|A - \text{id}\|}$ und $A^{-1} = \sum_{n=0}^{\infty} (\text{id} - A)^n$.*

Beweis. Sei $n \in \mathbb{N}_0$. Wegen Eigenschaft (8) aus Proposition 1 erhalten wir $\|(\text{id} - A)^n\| \leq \|\text{id} - A\|^n$. Nachdem $\|\text{id} - A\| < 1$ gilt, konvergiert die geometrische Reihe $\sum_{n=0}^{\infty} \|\text{id} - A\|^n$ und es gilt $\sum_{n=0}^{\infty} \|\text{id} - A\|^n = \frac{1}{1 - \|\text{id} - A\|}$. Wegen des Majorantentests konvergiert $\sum_{n=0}^{\infty} \|(\text{id} - A)^n\|$. Daher konvergiert $\sum_{n=0}^{\infty} (\text{id} - A)^n$ nach Proposition 3.

Setze $B := \sum_{n=0}^{\infty} (\text{id} - A)^n$. Es ist

$$\|B\| \leq \sum_{n=0}^{\infty} \|(\text{id} - A)^n\| \leq \sum_{n=0}^{\infty} \|\text{id} - A\|^n = \frac{1}{1 - \|\text{id} - A\|} .$$

Dann gilt

$$\begin{aligned} BA &= \sum_{n=0}^{\infty} (\text{id} - A)^n \underbrace{A}_{=A - \text{id} + \text{id} = -(\text{id} - A) + \text{id}} = \\ &= - \underbrace{\sum_{n=0}^{\infty} (\text{id} - A)^{n+1}}_{= \sum_{n=1}^{\infty} (\text{id} - A)^n} + \sum_{n=0}^{\infty} (\text{id} - A)^n = (\text{id} - A)^0 = \text{id} . \end{aligned}$$

Nach Proposition 4 ist A invertierbar und $A^{-1} = B$, womit das Resultat gezeigt ist. \square

Wenn eine Matrix genügend nahe bei einer invertierbaren Matrix liegt, dann ist sie invertierbar. Das zeigt unser nächstes Resultat.

Proposition 6. *Seien $A, B \in M_s(\mathbb{R})$. Weiters sei A invertierbar und es gelte $\|B - A\| < \frac{1}{\|A^{-1}\|}$. Dann ist B invertierbar und es gelten $\|B^{-1}\| \leq \frac{\|A^{-1}\|}{1 - \|A^{-1}\| \|B - A\|}$ und*

$$\|B^{-1} - A^{-1}\| \leq \|A^{-1}\| \|B^{-1}\| \|B - A\| \leq \frac{\|A^{-1}\|^2}{1 - \|A^{-1}\| \|B - A\|} \|B - A\| .$$

Beweis. Es gilt

$$\|A^{-1}B - \underbrace{\text{id}}_{=A^{-1}A}\| = \|A^{-1}(B - A)\| \underset{\text{wegen (8) aus Proposition 1}}{\leq} \|A^{-1}\| \|B - A\| < 1 .$$

Nach Proposition 5 ist $A^{-1}B$ invertierbar und $\|(A^{-1}B)^{-1}\| \leq \frac{1}{1 - \|A^{-1}B - \text{id}\|}$. Aus $\|A^{-1}B - \text{id}\| \leq \|A^{-1}\| \|B - A\|$ ergibt sich daraus $\|(A^{-1}B)^{-1}\| \leq$

$\frac{1}{1 - \|A^{-1}\| \|B - A\|}$. Setze $C := (A^{-1}B)^{-1}A^{-1}$. Dann gilt $CB = (A^{-1}B)^{-1}A^{-1}B = \text{id}$. Wegen Proposition 4 ist B invertierbar und $B^{-1} = C$. Es ist

$$\begin{aligned} \left\| \underbrace{B^{-1}}_{=C=(A^{-1}B)^{-1}A^{-1}} \right\| &= \|(A^{-1}B)^{-1}A^{-1}\| \stackrel{\text{wegen (8) aus Proposition 1}}{\leq} \\ &\leq \underbrace{\|(A^{-1}B)^{-1}\|}_{\leq \frac{1}{1 - \|A^{-1}\| \|B - A\|}} \|A^{-1}\| \leq \frac{\|A^{-1}\|}{1 - \|A^{-1}\| \|B - A\|}. \end{aligned}$$

Weiters ist

$$\begin{aligned} \left\| \underbrace{B^{-1}}_{=C=(A^{-1}B)^{-1}A^{-1}} - A^{-1} \right\| &= \|((A^{-1}B)^{-1} - \text{id})A^{-1}\| \stackrel{\text{wegen (8) aus Proposition 1}}{\leq} \\ &\leq \underbrace{\|(A^{-1}B)^{-1} - \text{id}\|}_{=B^{-1}A} \|A^{-1}\| = \|A^{-1}\| \|B^{-1}(A - B)\| \stackrel{\text{wegen (8) aus Proposition 1}}{\leq} \\ &\|A^{-1}\| \|B^{-1}\| \|B - A\| \stackrel{\text{da } \|B^{-1}\| \leq \frac{\|A^{-1}\|}{1 - \|A^{-1}\| \|B - A\|}}{\leq} \frac{\|A^{-1}\|^2}{1 - \|A^{-1}\| \|B - A\|} \|B - A\|, \end{aligned}$$

womit das behauptete Resultat bewiesen ist. \square

Die Menge der invertierbaren reellen $s \times s$ -Matrizen bezeichnet man mit $\text{GL}(s, \mathbb{R})$ (und analog dazu die Menge der invertierbaren komplexen $s \times s$ -Matrizen mit $\text{GL}(s, \mathbb{C})$). Aus der linearen Algebra weiß man, dass A genau dann invertierbar ist, wenn $\det A \neq 0$. Wir werden jetzt zeigen, dass die Menge der invertierbaren Matrizen offen ist, und $A \mapsto A^{-1}$ stetig ist. Offensichtlich ist für jede invertierbare Matrix A auch A^{-1} invertierbar, also $A^{-1} \in \text{GL}(s, \mathbb{R})$.

Proposition 7. *Die Menge $\text{GL}(s, \mathbb{R})$ ist offen in $M_s(\mathbb{R})$. Weiters ist die von $\text{GL}(s, \mathbb{R})$ nach $M_s(\mathbb{R})$ führende Abbildung $A \mapsto A^{-1}$ stetig.*

Beweis. Sei A invertierbar und sei $\varepsilon > 0$. Setze $\delta := \min \left\{ \frac{1}{2\|A^{-1}\|}, \frac{\varepsilon}{2\|A^{-1}\|^2} \right\}$. Es sei $B \in M_s(\mathbb{R})$ mit $\|B - A\| < \delta$. Weil $\|B - A\| < \delta \leq \frac{1}{2\|A^{-1}\|} \leq \frac{1}{\|A^{-1}\|}$ folgt aus Proposition 6, dass B invertierbar ist. Damit ist gezeigt, dass $\text{GL}(s, \mathbb{R})$ offen in $M_s(\mathbb{R})$ ist. Nachdem $\|B - A\| < \delta \leq \frac{1}{2\|A^{-1}\|}$ ist $\|A^{-1}\| \|B - A\| \leq \frac{1}{2}$, also $1 - \|A^{-1}\| \|B - A\| \geq \frac{1}{2}$, und deshalb $\frac{\|A^{-1}\|^2}{1 - \|A^{-1}\| \|B - A\|} \leq 2\|A^{-1}\|^2$. Nach Proposition 6 und weil $\|B - A\| < \delta \leq \frac{\varepsilon}{2\|A^{-1}\|^2}$ erhalten wir

$$\|B^{-1} - A^{-1}\| \leq \underbrace{\frac{\|A^{-1}\|^2}{1 - \|A^{-1}\| \|B - A\|}}_{\leq 2\|A^{-1}\|^2} \underbrace{\|B - A\|}_{< \frac{\varepsilon}{2\|A^{-1}\|^2}} < 2\|A^{-1}\|^2 \frac{\varepsilon}{2\|A^{-1}\|^2} = \varepsilon.$$

Somit ist $A \mapsto A^{-1}$ stetig. \square