1 Differentialrechnung in mehreren Variablen

1.1 Die Geometrie euklidischer Räume

Zur Erinnerung Die Elemente des $\mathbb{R}^n$ schreiben wir normalerweise als Zeilenvektoren:

$$\mathbf{x} = (x_1, \ldots, x_n).$$

Kommen Matrizen ins Spiel, so ist manchmal die Spalten-Schreibweise vorteilhafter:

$$[\mathbf{x}] := \mathbf{x}^\top = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

Ist $M_{n,m}(\mathbb{R})$ der Raum der Matrizen mit n Zeilen und m Spalten und $A \in M_{n,m}(\mathbb{R})$, so induziert A eine lineare Abbildung $f_A : \mathbb{R}^m \to \mathbb{R}^n$ durch $[f_A(\mathbf{x})] := A \cdot [\mathbf{x}]$, in Zeilenschreibweise also

$$f_A(\mathbf{x}) := (A \cdot \mathbf{x}^\top)^\top = \mathbf{x} \cdot A^\top.$$

Dabei steht der Punkt für die Matrizenmultiplikation.

Das euklidische Skalarprodukt zweier Vektoren $\mathbf{v}, \mathbf{w}$ wird durch

$$\mathbf{v} \bullet \mathbf{w} := \mathbf{v} \cdot \mathbf{w}^\top = \sum_{\nu=1}^{n} v_\nu w_\nu,$$

erklärt, die euklidische Norm eines Vektors $\mathbf{v}$ durch

$$\|\mathbf{v}\| := (\mathbf{v} \bullet \mathbf{v})^{1/2} = \sqrt{(v_1)^2 + \cdots + (v_n)^2}.$$

Neben dem $\mathbb{R}^n$ und dem Matrizenraum $M_{n,m}(\mathbb{R})$ werden wir auch mit anderen Vektorräumen zu tun haben, z.B. mit Unterräumen des $\mathbb{R}^n$. Sind V und W zwei beliebige $\mathbb{R}$-Vektorräume, so bezeichnen wir mit $L(V, W)$ den Raum der linearen Abbildungen von V nach W.

Wir werden uns in diesem Abschnitt mit der „Topologie“ des $\mathbb{R}^n$ und allgemeinerer Vektorräume beschäftigen. Das Wort „Topologie“ kommt vom griechischen „topos“ und bedeutet „Ort“, „Stelle“ oder „Raum“. Es geht also um die Wissenschaft vom Raum und der Lage der Dinge zueinander.

Definition

Eine ***Norm*** auf einem $\mathbb{R}$-Vektorraum V ist eine Funktion $N : V \to \mathbb{R}$ mit folgenden Eigenschaften:

1. $N(v) \geq 0$ für jedes $v \in V$, und $N(v) = 0 \iff v = 0$,
2. $N(\alpha\, v) = |\alpha| \cdot N(v)$ für $\alpha \in \mathbb{R}$ und $v \in V$,
3. $N(v + w) \leq N(v) + N(w)$ für $v, w \in V$ (Dreiecks-Ungleichung).

Ein ***normierter Vektorraum*** ist ein Vektorraum V, auf dem eine Norm gegeben ist.

1.1. Beispiele

A. Die *kanonische euklidische Norm* auf dem $\mathbb{R}^n$, gegeben durch $\|\mathbf{v}\| := \sqrt{\mathbf{v} \bullet \mathbf{v}}$, kennen wir schon.

B. Für $\mathbf{v} = (v_1, \ldots, v_n) \in \mathbb{R}^n$ sei die ***Maximumsnorm*** definiert durch

$$|\mathbf{v}| := \max_{i=1,\ldots,n} |v_i|.$$

Offensichtlich sind die Bedingungen (1) und (2) erfüllt, und es ist

$$|\mathbf{v} + \mathbf{w}| = \max_i |v_i + w_i| \leq \max_i (|v_i| + |w_i|) \leq \max_i |v_i| + \max_i |w_i| = |\mathbf{v}| + |\mathbf{w}|.$$

C. Ist $I = [a, b]$ ein abgeschlossenes Intervall, so ist der Raum $V := \mathscr{C}^0(I)$ der stetigen Funktionen auf I ein Beispiel für einen unendlich-dimensionalen Vektorraum. Durch

$$\|f\|_I := \sup\{|f(x)| \,:\, x \in I\}$$

wird eine Norm auf V eingeführt.

Ist V ein Vektorraum mit einer Norm N, so versteht man unter der (offenen) ***Kugel*** mit Radius r um x_0 (bezüglich der Norm N) die Menge

$$B_r(x_0) := \{x \in V \,:\, N(x - x_0) < r\}.$$

Die „Kugel“ bezüglich der Maximumsnorm, $\{\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n \,:\, |\mathbf{x}| < r\} = (-r, r) \times \ldots \times (-r, r)$, ist in Wirklichkeit ein Würfel.

Definition

Sei V ein Vektorraum mit Norm N. Eine Folge (x_ν) von Punkten in V ***konvergiert*** (bezüglich N) gegen einen Punkt x_0, falls es zu jedem $\varepsilon > 0$ ein ν_0 gibt, so dass für alle $\nu \geq \nu_0$ gilt: $N(x_\nu - x_0) < \varepsilon$. Man schreibt dann: $\lim\limits_{\nu \to \infty} x_\nu = x_0$.

Man kann auch sagen: (x_ν) konvergiert gegen x_0, falls $N(x_\nu - x_0)$ gegen 0 konvergiert. In $\mathbb{R}$ ergibt das den bereits bekannten Konvergenzbegriff. Genau wie dort folgt auch, dass der Grenzwert eindeutig bestimmt ist.

Ist $\mathbf{x}_\nu = (x_1^{(\nu)}, \dots, x_n^{(\nu)})$ eine Punktfolge im $\mathbb{R}^n$ und $\mathbf{x}_0 = (x_1^{(0)}, \dots, x_n^{(0)})$ ein fester Punkt, so ist

$$\|\mathbf{x}_\nu - \mathbf{x}_0\| = \sqrt{(x_1^{(\nu)} - x_1^{(0)})^2 + \cdots + (x_n^{(\nu)} - x_n^{(0)})^2}.$$

Die Zahlenfolge $\|\mathbf{x}_\nu - \mathbf{x}_0\|$ konvergiert offensichtlich genau dann gegen 0, wenn $|x_i^{(\nu)} - x_i^{(0)}|$ für jedes i gegen Null konvergiert. Also konvergiert die Punktfolge $(\mathbf{x}_\nu)$ genau dann bezüglich der kanonischen euklidischen Norm gegen $\mathbf{x}_0$, wenn die Komponenten $x_i^{(\nu)}$ jeweils gegen $x_i^{(0)}$ konvergieren.

Nicht alle Grenzwertsätze lassen sich übertragen, aber einige (wie etwa der über die Summen von Folgen) gelten sinngemäß wie in $\mathbb{R}$ und werden auch genauso bewiesen.

Wir haben nun das Problem, dass viele Aussagen von der benutzten Norm abhängen. Das ist sehr lästig, und deshalb führen wir den folgenden Begriff ein:

Zwei Normen N_1 und N_2 auf einem beliebigen Vektorraum V heißen ***äquivalent***, falls es Konstanten $c, c^* > 0$ gibt, so dass gilt:

$$c \cdot N_1(v) \le N_2(v) \le c^* \cdot N_1(v) \quad \text{für alle } v \in V.$$

Es handelt sich dabei wirklich um eine Äquivalenzrelation:

1. Offensichtlich ist jede Norm äquivalent zu sich selbst.

2. Ist $c \cdot N_1(v) \le N_2(v) \le c^* \cdot N_1(v)$ für alle $v \in V$, so ist auch

   $$\frac{1}{c^*} \cdot N_2(v) \le N_1(v) \le \frac{1}{c} \cdot N_2(v) \quad \text{für alle } v \in V.$$

   Das liefert die Symmetrie.

3. Ist $c \cdot N_1(v) \le N_2(v) \le c^* \cdot N_1(v)$ und $d \cdot N_2(v) \le N_3(v) \le d^* \cdot N_2(v)$ für alle $v \in V$, so ist

   $$(cd) \cdot N_1(v) \le N_3(v) \le (c^* d^*) \cdot N_1(v) \quad \text{für alle } v \in V.$$

   Also ist die Relation transitiv.

1.2. Beispiel

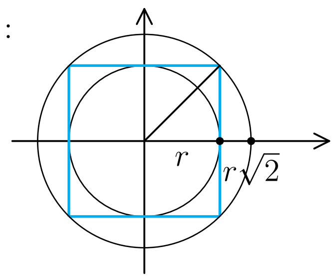
Wir betrachten die beiden Normen $N_1(\mathbf{x}) = \|\mathbf{x}\|$ und $N_2(\mathbf{x}) = |\mathbf{x}|$ auf dem $\mathbb{R}^n$. Dann ist

$$\begin{aligned} \|\mathbf{x}\|^2 &= |x_1|^2 + \cdots + |x_n|^2 \leq n \cdot (\max_i |x_i|)^2 \\ \text{und} \quad \max_i |x_i| &= \sqrt{(\max_i |x_i|)^2} \leq \sqrt{|x_1|^2 + \cdots + |x_n|^2}, \end{aligned}$$

also $\frac{1}{\sqrt{n}} \cdot \|\mathbf{x}\| \leq |\mathbf{x}| \leq \|\mathbf{x}\|$.

Die euklidische Norm und die Maximumsnorm sind demnach äquivalent.

$n = 2$:



Die Äquivalenz bedeutet, dass jede Kugel in der einen Norm eine Kugel bezüglich der anderen Norm enthält und in einer entsprechenden größeren enthalten ist. Deshalb hängen die folgenden Definitionen nicht davon ab, ob man die euklidische Norm oder die Maximumsnorm verwendet.

Definition

Ein Punkt $\mathbf{x}_0 \in \mathbb{R}^n$ heißt ***Häufungspunkt*** einer Folge von Punkten $\mathbf{x}_\nu$, falls in jeder Kugel um $\mathbf{x}_0$ unendlich viele Folgenglieder liegen.

Eine Menge $M \subset \mathbb{R}^n$ heißt ***beschränkt***, falls es ein $R > 0$ gibt, so dass M in der Kugel $B_R(0)$ enthalten ist. Eine Folge im $\mathbb{R}^n$ heißt *beschränkt*, wenn die Menge ihrer Glieder beschränkt ist. Auch hier gilt:

1.3. Satz von Bolzano-Weierstraß

Jede beschränkte Folge $(\mathbf{x}_\nu)$ im $\mathbb{R}^n$ besitzt eine konvergente Teilfolge.

BEWEIS: Es gibt ein $R > 0$, so dass alle $\mathbf{x}_\nu = (x_1^{(\nu)}, \ldots, x_n^{(\nu)})$ in $B_R(0)$ liegen. Aber dann liegen sie erst recht in $I^n = I \times \ldots \times I$, mit $I := [-R, R]$. Die Folge $x_1^{(\nu)}$ besitzt eine konvergente Teilfolge $x_1^{(\nu(i_1))}$ mit einem Grenzwert $x_1^{(0)} \in I$, die Folge $x_2^{(\nu(i_1))}$ besitzt eine konvergente Teilfolge $x_2^{(\nu(i_2))}$ mit einem Grenzwert $x_2^{(0)} \in I$, usw.

Schließlich erhält man eine Teilfolge $(\mathbf{x}_{\nu(i_n)})$ von $(\mathbf{x}_\nu)$, die gegen $\mathbf{x}_0 = (x_1^{(0)}, \ldots, x_n^{(0)})$ konvergiert. ∎

Die Äquivalenz von euklidischer und Maximumsnorm ist kein Zufall.

1.4. Satz

Je zwei Normen auf dem $\mathbb{R}^n$ sind äquivalent.

BEWEIS: Es reicht zu zeigen, dass eine beliebige Norm N äquivalent zur Maximumsnorm ist.

Jeder Vektor $\mathbf{x} \in \mathbb{R}^n$ besitzt eine eindeutige Darstellung $\mathbf{x} = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n$. Daraus folgt die Beziehung $N(\mathbf{x}) \le |x_1| \cdot N(\mathbf{e}_1) + \cdots + |x_n| \cdot N(\mathbf{e}_n) \le c^* \cdot |\mathbf{x}|$, mit $c^* := N(\mathbf{e}_1) + \cdots + N(\mathbf{e}_n)$.

Wir nehmen nun an, es gibt kein $c > 0$, so dass $c \cdot |\mathbf{x}| \le N(\mathbf{x})$ für alle $\mathbf{x}$ ist. Dann gibt es zu jedem $\nu \in \mathbb{N}$ ein $\mathbf{x}_\nu$ mit $|\mathbf{x}_\nu| > \nu \cdot N(\mathbf{x}_\nu)$, also

$$N(\mathbf{y}_\nu) < \frac{1}{\nu}\,, \text{ für } \mathbf{y}_\nu := \frac{\mathbf{x}_\nu}{|\mathbf{x}_\nu|}\,.$$

Weil $|\mathbf{y}_\nu| = 1$ ist, ist die Folge $(\mathbf{y}_\nu)$ beschränkt und besitzt eine konvergente Teilfolge. Ohne Beschränkung der Allgemeinheit können wir annehmen, dass $(\mathbf{y}_\nu)$ schon selbst gegen ein $\mathbf{y}_0$ (in der Maximumsnorm) konvergiert. Dann konvergiert aber $|\mathbf{y}_\nu - \mathbf{y}_0|$ gegen Null. Außerdem ist

$$\begin{aligned} N(\mathbf{y}_0) &= N(\mathbf{y}_0 - \mathbf{y}_\nu + \mathbf{y}_\nu) \le N(\mathbf{y}_0 - \mathbf{y}_\nu) + N(\mathbf{y}_\nu) \\ &\le c^* \cdot |\mathbf{y}_0 - \mathbf{y}_\nu| + \frac{1}{\nu}\,. \end{aligned}$$

Da die rechte Seite gegen Null konvergiert, ist $N(\mathbf{y}_0) = 0$, also $\mathbf{y}_0 = 0$. Aber andererseits ist

$$1 = |\mathbf{y}_\nu| = |\mathbf{y}_0 + (\mathbf{y}_\nu - \mathbf{y}_0)| \le |\mathbf{y}_0| + |\mathbf{y}_\nu - \mathbf{y}_0|\,.$$

Das ergibt einen Widerspruch. ■

Wir fassen noch einmal zusammen, was die Äquivalenz von Normen bedeutet:

1.5. Satz

Die Normen N_1 und N_2 auf dem Vektorraum V seien äquivalent. Dann folgt:

1. *Zu jedem Radius $r_1 > 0$ gibt es einen Radius $r_2 > 0$, so dass gilt:*

   $$\{x \in V \,:\, N_2(x) < r_2\} \subset \{x \in V \,:\, N_1(x) < r_1\},$$

   und umgekehrt.

2. *Konvergiert eine Folge (x_n) in V bezüglich N_1 gegen einen Grenzwert x_0, so konvergiert (x_n) auch bezüglich N_2 gegen x_0, und umgekehrt.*

Der Beweis ist offensichtlich.

Definition

Sei E ein reeller (oder komplexer) Vektorraum. Ein ***Skalarprodukt*** auf E ist eine Funktion, die je zwei Vektoren $v, w \in E$ eine reelle (bzw. komplexe) Zahl $(v \,|\, w)$ zuordnet und folgende Eigenschaften besitzt:

1. $(\mathbf{v}_1 + \mathbf{v}_2 \,|\, \mathbf{w}) = (\mathbf{v}_1 \,|\, \mathbf{w}) + (\mathbf{v}_2 \,|\, \mathbf{w})$ für $\mathbf{v}_1, \mathbf{v}_2, \mathbf{w} \in E$,
2. $(\alpha\,\mathbf{v} \,|\, \mathbf{w}) = \alpha \cdot (\mathbf{v} \,|\, \mathbf{w})$ für $\alpha \in \mathbb{R}$ (bzw. in $\mathbb{C}$) und $\mathbf{v}, \mathbf{w} \in E$,
3. $(\mathbf{w} \,|\, \mathbf{v}) = \overline{(\mathbf{v} \,|\, \mathbf{w})}$ für $\mathbf{v}, \mathbf{w} \in E$.
4. Ist $\mathbf{v} \neq \mathbf{0}$, so ist $(\mathbf{v} \,|\, \mathbf{v}) > 0$.

Unter einem ***euklidischen Raum*** (bzw. einem ***unitären Raum***) verstehen wir einen reellen (bzw. komplexen) Vektorraum mit einem Skalarprodukt.

Ein Skalarprodukt auf einem reellen Vektorraum ist also eine symmetrische Bilinearform, die zusätzlich ***positiv definit*** ist (Eigenschaft (4)). Im komplexen Fall nennt man eine Funktion mit den Eigenschaften (1) bis (3) eine ***hermitesche Form***. Auch sie wird zum Skalarprodukt, wenn sie positiv definit ist.

Im reellen Fall ist $(w \,|\, v) = (v \,|\, w)$ für $v, w \in E$. Im komplexen Fall ist $(v \,|\, v)$ stets reell und $(v \,|\, \alpha\, w) = \overline{\alpha} \cdot (v \,|\, w)$ für $\alpha \in \mathbb{C}$ und $v, w \in E$.

1.6. Beispiele

A. Das Skalarprodukt $(\mathbf{x}, \mathbf{y}) \mapsto \mathbf{x} \bullet \mathbf{y}$ bezeichnet man als das ***kanonische Skalarprodukt*** auf dem $\mathbb{R}^n$.

B. Das ***kanonische „hermitesche“ Skalarprodukt*** auf dem $\mathbb{C}^n$ wird gegeben durch

$$<\mathbf{v} \,|\, \mathbf{w}> := \sum_{\nu=1}^{n} v_\nu \overline{w}_\nu \,.$$

C. Ist $f = g + \mathrm{i}\, h$ eine stetige komplexwertige Funktion über $[a, b] \subset \mathbb{R}$, so setzen wir

$$<f \,|\, g> := \int_a^b f(t)\overline{g(t)}\, dt \,.$$

Offensichtlich liefert das eine hermitesche Form auf dem Raum der stetigen komplexwertigen Funktionen auf $[a, b]$, und es ist

$$<f \,|\, f> = \int_a^b |f(t)|^2\, dt \geq 0 \text{ für alle } f.$$

Ist eine stetige Funktion $f : [a, b] \to \mathbb{C}$ nicht die Nullfunktion, so gibt es ein $t_0 \in [a, b]$ mit $f(t_0) \neq 0$, und wegen der Stetigkeit gibt es ein $\varepsilon > 0$, so dass $f(t) \neq 0$ auf $U_\varepsilon(t_0) \cap [a, b]$ ist. Dann nimmt $|f|$ auf $\{t \in [a, b] \,:\, |t - t_0| \leq \varepsilon/2\}$ ein Minimum $\delta > 0$ an, und dort ist $|f| \geq \delta$. Daraus folgt:

$$<f \,|\, f> = \int_a^b |f(t)|^2\, dt \geq \delta^2 \cdot \varepsilon/2 > 0 \,.$$

Also liegt sogar ein Skalarprodukt auf $\mathcal{C}^0([a, b]; \mathbb{C})$ vor.

1.7. Die Ungleichung von Cauchy-Schwarz

Sei E ein euklidischer oder unitärer Vektorraum. Dann gilt für $v, w \in E$:

$$|(v \,|\, w)|^2 \leq (v \,|\, v) \cdot (w \,|\, w)\,.$$

BEWEIS: Sei $a := (v \,|\, v)$, $b := (v \,|\, w)$ und $c := (w \,|\, w)$. Dann sind a, c reell und ≥ 0. Ist $v = w = 0$, so ist auch $a = b = c = 0$ und nichts mehr zu zeigen. Da die Aussage symmetrisch in v und w ist, können wir annehmen, dass $w \neq 0$ ist, also $c > 0$. Ist $\lambda \in \mathbb{C}$ beliebig, so gilt:

$$0 \leq (v + \lambda w \,|\, v + \lambda w) = a + \lambda\overline{b} + \overline{\lambda}b + |\lambda|^2 c\,.$$

Mit $\lambda = -b/c$ erhalten wir:

$$0 \leq a - \frac{b\overline{b}}{c} - \frac{b\overline{b}}{c} + \frac{b\overline{b}}{c^2} \cdot c = a - \frac{b\overline{b}}{c},$$

also $b\overline{b} \leq a \cdot c$. Und genau das war zu zeigen. ∎

Im Falle des euklidischen Skalarproduktes auf dem $\mathbb{R}^n$ liefert die Schwarz'sche Ungleichung

$$|\mathbf{v} \bullet \mathbf{w}| \leq \|\mathbf{v}\| \cdot \|\mathbf{w}\|,$$

dass $\left|\dfrac{\mathbf{v} \bullet \mathbf{w}}{\|\mathbf{v}\| \cdot \|\mathbf{w}\|}\right| \leq 1$ ist. Deshalb definiert man:

Der ***Winkel*** θ ***zwischen*** $\mathbf{v}$ ***und*** $\mathbf{w}$ ist gegeben durch $\cos\theta = \dfrac{\mathbf{v} \bullet \mathbf{w}}{\|\mathbf{v}\| \cdot \|\mathbf{w}\|}$.

Zwei Elemente $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ heißen ***orthogonal***, falls $\mathbf{v} \bullet \mathbf{w} = 0$ ist.

Der Begriff der Orthogonalität kann auf beliebige euklidische (oder unitäre) Vektorräume V verallgemeinert werden: Zwei Elemente $v, w \in V$ werden ***orthogonal*** genannt, falls $(v \,|\, w) = 0$ ist. Sind v, w orthogonal, so ist

$$(v + w \,|\, v + w) = (v \,|\, v) + (w \,|\, w) \quad \text{(Satz des Pythagoras)}\,.$$

1.8. Beispiel

Im Raum der stetigen Funktionen auf $I = [-\pi, \pi]$ ist

$$<e^{\mathrm{i}\,nt} \,|\, e^{\mathrm{i}\,mt}> = \int_{-\pi}^{\pi} e^{\mathrm{i}\,(n-m)t}\,dt = \begin{cases} 2\pi & \text{falls } n = m \\ 0 & \text{falls } n \neq m. \end{cases}$$

Für $n \neq m$ sind also die Funktionen $e^{\mathrm{i}\,nt}$ und $e^{\mathrm{i}\,mt}$ orthogonal zueinander. Die Funktionen

$$f_n(t) := \frac{1}{\sqrt{2\pi}} \cdot e^{\mathrm{i}\,nt}, \quad n \in \mathbb{Z},$$

bilden dann ein „Orthonormalsystem“ im Raum $\mathcal{C}^0([-\pi, \pi]; \mathbb{C})$, d.h. es ist $<f_n \,|\, f_m> = 0$ für $n \neq m$ und $<f_n \,|\, f_n> = 1$ für alle n.

Jedes Skalarprodukt liefert eine Norm, durch

$$N(\mathbf{v}) := (\mathbf{v} \,|\, \mathbf{v})^{1/2}.$$

Die Dreiecksungleichung folgt mit Hilfe der Cauchy-Schwarz'schen Ungleichung: Ist $z = x + \mathrm{i}\,y$ eine komplexe Zahl, so ist $z + \overline{z} = 2x = 2\,\mathrm{Re}(z)$. Daraus folgt:

$$\begin{aligned} N(\mathbf{a}+\mathbf{b})^2 = (\mathbf{a}+\mathbf{b} \,|\, \mathbf{a}+\mathbf{b}) &= (\mathbf{a} \,|\, \mathbf{a}) + 2\,\mathrm{Re}\,(\mathbf{a} \,|\, \mathbf{b}) + (\mathbf{b} \,|\, \mathbf{b}) \\ &\le N(\mathbf{a})^2 + 2 \cdot |(\mathbf{a} \,|\, \mathbf{b})| + N(\mathbf{b})^2 \\ &\le N(\mathbf{a})^2 + 2 \cdot N(\mathbf{a}) \cdot N(\mathbf{b}) + N(\mathbf{b})^2 \\ &= \big(N(\mathbf{a}) + N(\mathbf{b})\big)^2. \end{aligned}$$

Wurzelziehen auf beiden Seiten ergibt die gewünschte Dreiecksungleichung.

Nicht jede Norm wird mit Hilfe eines Skalarproduktes definiert. Ein typisches Beispiel für eine Norm, die nicht von einem Skalarprodukt kommt, ist die Supremumsnorm.

Definition

Sei X eine Menge. Unter einer ***Metrik*** auf X versteht man eine Abbildung $d : X \times X \to \mathbb{R}$ mit folgenden Eigenschaften:

1. $d(x, y) \ge 0$ und $= 0 \iff x = y$,
2. $d(x, y) = d(y, x)$,
3. $d(x, y) \le d(x, z) + d(z, y)$ (Dreiecksungleichung).

Jede Norm N auf einem Vektorraum V führt zu einer Metrik d_N auf V durch

$$d_N(\mathbf{v}, \mathbf{w}) := N(\mathbf{v} - \mathbf{w})\,.$$

Die Eigenschaften (1) und (2) einer Metrik sind für d_N offensichtlich erfüllt, und auch die Dreiecksungleichung folgt leicht:

$$\begin{aligned} d_N(\mathbf{x}, \mathbf{y}) &= N(\mathbf{x} - \mathbf{y}) = N\big((\mathbf{x} - \mathbf{z}) + (\mathbf{z} - \mathbf{y})\big) \\ &\le N(\mathbf{x} - \mathbf{z}) + N(\mathbf{z} - \mathbf{y}) \\ &= d_N(\mathbf{x}, \mathbf{z}) + d_N(\mathbf{z}, \mathbf{y}). \end{aligned}$$

Wir haben also folgende Abhängigkeit:

$$\boxed{\text{Skalarprodukt} \implies \text{Norm} \implies \text{Metrik.}}$$

Die Supremumsnorm ist ein typisches Beispiel für eine Norm, die nicht von einem Skalarprodukt kommt, und es gibt auch Metriken, die nicht von einer Norm kommen. Für eine Metrik braucht man nicht einmal einen Vektorraum.

Eine Menge mit einer Metrik wird als ***metrischer Raum*** bezeichnet. Schon mit Hilfe einer Metrik kann man Kugeln

$$B_r(x_0) := \{x \in X \,:\, d(x, x_0) < r\}$$

(und damit ε-Umgebungen) und dann in einem beliebigen metrischen Raum auch offene Mengen definieren.

Zur Erinnerung: Eine Menge $M \subset \mathbb{R}$ heißt *offen*, falls es zu jedem Element $x_0 \in M$ ein $\varepsilon > 0$ gibt, so dass

$$U_\varepsilon(x_0) := \{x \in \mathbb{R} \,:\, |x - x_0| < \varepsilon\}$$

ganz in M enthalten ist.

Ist nun X ein metrischer Raum mit Metrik d, so definiert man:

Definition

Eine Menge $M \subset X$ heißt ***offen***, falls es zu jedem Element $x_0 \in M$ ein $\varepsilon > 0$ gibt, so dass

$$U_\varepsilon(x_0) := \{x \in X \,:\, d(x, x_0) < \varepsilon\}$$

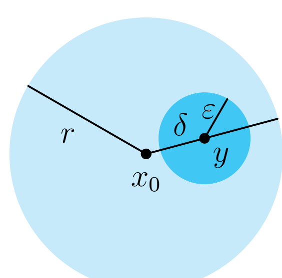
ganz in M enthalten ist.

Eine Menge $A \subset X$ heißt ***abgeschlossen***, falls $M := X \setminus A$ offen ist.

1.9. Satz

In einem metrischen Raum ist jede offene Kugel $B_r(x_0)$ eine offene Menge.

BEWEIS: Sei $y \in B_r(x_0)$. Wir suchen eine ε-Umgebung von y, die noch ganz in $B_r(x_0)$ enthalten ist. Dazu sei $\delta := \mathrm{dist}(y, x_0)$. Dann ist $0 \le \delta < r$. Man kann eine positive reelle Zahl $\varepsilon < r - \delta$ finden. Ist $x \in U_\varepsilon(y)$, also $d(x, y) < \varepsilon$, so ist $d(x, x_0) \le d(x, y) + d(y, x_0) < \varepsilon + \delta < (r - \delta) + \delta = r$.



Das zeigt, dass $U_\varepsilon(y) \subset B_r(x_0)$ ist. ∎

Die offenen Mengen eines metrischen Raumes X haben gewisse typische Eigenschaften:

- Die leere Menge besitzt kein Element, für das man etwas nachprüfen müsste. Deshalb ist sie offen. Und auch der ganze Raum X ist trivialerweise offen.

- Sind M_1 und M_2 offene Mengen und ist $x_0 \in M_1 \cap M_2$, so gibt es $\varepsilon_1, \varepsilon_2 > 0$, so dass $U_{\varepsilon_1}(x_0) \subset M_1$ und $U_{\varepsilon_2}(x_0) \subset M_2$ ist. Setzt man $\varepsilon := \min(\varepsilon_1, \varepsilon_2)$, so ist $U_\varepsilon(x_0) \subset M_1 \cap M_2$. Das zeigt, dass $M_1 \cap M_2$ offen ist.

  Der Durchschnitt von unendlich vielen offenen Mengen braucht nicht mehr offen zu sein! So ist z.B. der Durchschnitt aller offenen Nullumgebungen in $\mathbb{R}$ die Menge $\{0\}$.

- Ist $(M_\iota)_{\iota \in I}$ ein System von offenen Mengen und

  $$x_0 \in M := \bigcup_{\iota \in I} M_\iota = \{x \in X \,:\, \exists\, \iota \in I \text{ mit } x \in M_\iota\},$$

  so gibt es ein $\iota_0 \in I$ mit $x_0 \in M_{\iota_0}$ und daher ein $\varepsilon > 0$, so dass $U_\varepsilon(x_0) \subset M_{\iota_0} \subset M$ ist. Also ist M offen.

Das System aller offenen Mengen legt eine neue Art von Struktur auf unseren Raum.

Definition

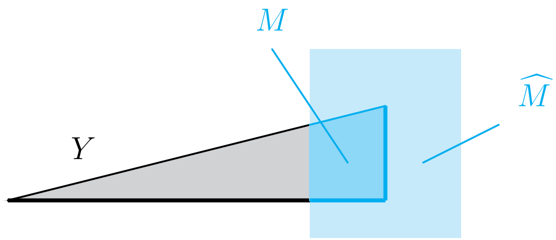
Sei X eine Menge. Eine ***Topologie*** auf X ist ein System $\mathscr{O}$ von Teilmengen von X mit folgenden Eigenschaften:

1. $\varnothing \in \mathscr{O}$ und $X \in \mathscr{O}$.
2. $M, N \in \mathscr{O} \implies M \cap N \in \mathscr{O}$.
3. Ist $(M_\iota)_{\iota \in I}$ eine Familie von Elementen aus $\mathscr{O}$, so gehört auch $\bigcup_{\iota \in I} M_\iota$ zu $\mathscr{O}$.

1.10. Beispiele

A. Das System aller offenen Mengen in einem metrischen Raum X bildet eine Topologie auf X.

B. Sei X mit einer Topologie versehen und $Y \subset X$ eine nicht leere Teilmenge. Wir nennen $M \subset Y$ ***(relativ) offen***, falls es eine offene Menge $\widehat{M} \subset X$ mit $M = X \cap \widehat{M}$ gibt.



- Weil $\varnothing = Y \cap \varnothing$ und $Y = Y \cap X$ ist, sind $\varnothing$ und Y relativ offen.
- Sind $M_1 = Y \cap \widehat{M}_1$ und $M_2 = Y \cap \widehat{M}_2$ relativ offen, so ist auch
$$M_1 \cap M_2 = (Y \cap \widehat{M}_1) \cap (Y \cap \widehat{M}_2) = Y \cap (\widehat{M}_1 \cap \widehat{M}_2)$$
relativ offen.
- Sei $(M_\iota)_{\iota \in I}$ ein System von relativ offenen Mengen in Y. Dann gibt es zu jedem $\iota \in I$ eine offene Menge $\widehat{M}_\iota$ in X mit $M_\iota = Y \cap \widehat{M}_\iota$. Aber dann ist auch
$$\bigcup_{\iota \in I} M_\iota = \bigcup_{\iota \in I} (Y \cap \widehat{M}_\iota) = Y \cap \bigcup_{\iota \in I} \widehat{M}_\iota$$
relativ offen.

Also bilden die relativ offenen Mengen in Y eine Topologie $\mathscr{O}_Y$ auf Y, die sogenannte ***Relativtopologie*** (oder ***induzierte Topologie***).

Klar ist:

1.11. Satz

Zwei äquivalente Normen auf einem Vektorraum E definieren die gleiche Topologie.

1.12. Folgerung

Jede Norm auf dem $\mathbb{R}^n$ induziert die gleiche Topologie.

BEWEIS: Klar! ∎

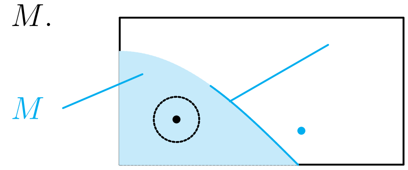
Jetzt können wir die Abhängigkeit noch um einen Begriff erweitern:

$$\boxed{\text{Skalarprodukt} \implies \text{Norm} \implies \text{Metrik} \implies \text{Topologie.}}$$

Es ist nicht schwer, eine Topologie anzugeben, die nicht von einer Metrik kommt. Das würde uns aber etwas zu weit vom Wege ablenken.

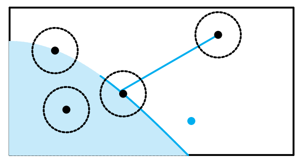
Sei jetzt wieder X ein metrischer Raum und $M \subset X$ eine Teilmenge. Ist $\mathbf{x}_0 \in X$ ein beliebiger Punkt, so kann man die möglichen Positionen von $\mathbf{x}_0$ gegenüber der Menge M mit Hilfe von Umgebungen beschreiben. Eine Teilmenge $U \subset X$ heißt eine ***Umgebung*** von $\mathbf{x}_0$, falls es eine offene Menge W mit $\mathbf{x}_0 \in W \subset U$ gibt. Dann heben wir die folgenden Situationen besonders hervor:

1. Es gibt eine Umgebung U von $\mathbf{x}_0$ in X, die ganz in M liegt. Dann heißt $\mathbf{x}_0$ ein ***innerer Punkt*** von M.



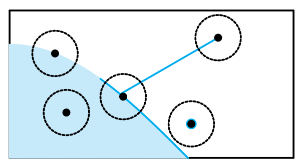
← innerer Punkt

2. Jede Umgebung von $\mathbf{x}_0$ enthält unendlich viele Punkte von M. Dann heißt $\mathbf{x}_0$ ein ***Häufungspunkt*** von M.



← Häufungspunkte

3. Jede Umgebung von $\mathbf{x}_0$ enthält wenigstens einen Punkt von M. Dann heißt $\mathbf{x}_0$ ein ***Berührungspunkt*** von M.



← Berührungspunkte

Jeder innere Punkt ist auch ein Häufungspunkt, aber die Umkehrung gilt i.a. nicht. Jeder Häufungspunkt ist ein Berührungspunkt, aber nicht nicht unbedingt umgekehrt. Ein Berührungspunkt $\mathbf{x}_0$, der kein Häufungspunkt ist, besitzt eine Umgebung U, so dass $U \cap M = \{\mathbf{x}_0\}$ ist. Dann nennt man $\mathbf{x}_0$ einen ***isolierten Punkt*** von M. Ist $\mathbf{x}_0$ nicht einmal ein Berührungspunkt von M, so gibt es eine Umgebung von $\mathbf{x}_0$, die keinen Punkt von M enthält.

Definition

Die Menge M° der inneren Punkte von M heißt ***offener Kern*** von M, die Menge $\overline{M}$ der Berührungspunkte von M heißt ***abgeschlossene Hülle*** von M.

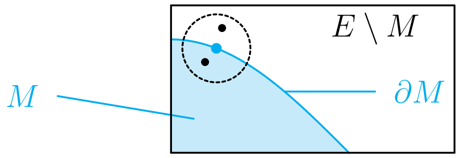
Die inneren Punkte von M gehören immer zu M. Für Häufungspunkte trifft das nicht unbedingt zu. So ist zum Beispiel 1 ein Häufungspunkt des offenen Intervalls $I := (0, 1)$, gehört aber nicht zu I. Isolierte Punkte einer Menge gehören immer zu der Menge dazu.

Die abgeschlossene Hülle $\overline{M}$ besteht aus den Häufungspunkten und den isolierten Punkten von M. Also ist $\overline{M}$ die Vereinigung von M mit allen Häufungspunkten von M.

Definition

Die Menge $\partial M := \overline{M} \setminus M^\circ$ heißt der ***Rand*** von M.

Ein Punkt $\mathbf{x}$ liegt also genau dann im Rand von M, wenn jede Umgebung von $\mathbf{x}$ sowohl M als auch $X \setminus M$ trifft. Es ist $M \cup \partial M = \overline{M}$ und $M \setminus \partial M = M^\circ$.



1.13. Beispiele

A. Ist $M = [a, b) \subset \mathbb{R}$, so ist $M^\circ = (a, b)$, $\overline{M} = [a, b]$ und $\partial M = \{a, b\}$.

B. Sei $M = [a, b] \cap \mathbb{Q}$. Dann ist $M^\circ = \varnothing$, $\overline{M} = [a, b]$ und $\partial M = [a, b]$.

C. Versieht man $\mathbb{Q}$ mit der von $\mathbb{R}$ induzierten Relativtopologie, so ist die Menge

$$M := \{x \in \mathbb{Q} \,:\, -\sqrt{2} < x < \sqrt{2}\}$$

relativ offen in $\mathbb{Q}$. Aber M ist auch relativ abgeschlossen in $\mathbb{Q}$, denn weil $\sqrt{2}$ keine rationale Zahl ist, ist das Komplement von M die Menge

$$\mathbb{Q} \setminus M = \{x \in \mathbb{Q} \,:\, |x| > \sqrt{2}\},$$

und die ist relativ offen. Also ist dann $\overline{M} = M^\circ = M$ und $\partial M = \varnothing$. Das gilt, wie gesagt, in der Relativtopologie. In $\mathbb{R}$ ist M weder offen, noch abgeschlossen. Dort ist $\overline{M} = [-\sqrt{2}, \sqrt{2}]$ und $M^\circ = \varnothing$, also $\partial M = \overline{M}$.

Wir müssen noch Ergebnisse über abgeschlossene Mengen (in einem metrischen Raum) herleiten.

1.14. Satz

Sind $A, B \subset X$ abgeschlossen, so sind auch $A \cap B$ und $A \cup B$ abgeschlossen.

Beweis: Der Satz folgt aus den Beziehungen

$$X \setminus (A \cap B) = (X \setminus A) \cup (X \setminus B) \;\text{ und }\; X \setminus (A \cup B) = (X \setminus A) \cap (X \setminus B).$$

∎

1.15. Abgeschlossenheitskriterium

Eine Teilmenge $M \subset X$ ist genau dann abgeschlossen, wenn gilt: Ist $(\mathbf{x}_\nu)$ eine Folge in M, die in X einen Grenzwert besitzt, so liegt dieser Grenzwert schon in M.

Beweis: 1) Sei M abgeschlossen und $(\mathbf{x}_\nu)$ eine Folge in M, die gegen einen Punkt $\mathbf{x}_0 \in X$ konvergiert. Ist die Menge $F := \{\mathbf{x}_\nu : \nu \in \mathbb{N}\}$ endlich, so muss $\mathbf{x}_0$ schon einer der Punkte $\mathbf{x}_\nu$ sein und deshalb in M liegen. Wir brauchen also nur den Fall zu betrachten, dass F unendlich ist. Wäre $\mathbf{x}_0$ ein Element von $X \setminus M$, so gäbe es ein $\varepsilon > 0$, so dass die ε-Umgebung von $\mathbf{x}_0$ auch noch in $X \setminus M$ liegt. Aber andererseits liegen fast alle Elemente von F (und damit unendlich viele Elemente von M) in $U_\varepsilon(\mathbf{x}_0)$. Das ergibt einen Widerspruch; $\mathbf{x}_0$ muss in M liegen.

2) M erfülle das Kriterium und $\mathbf{x}_0$ sei ein Punkt von $X \setminus M$. Wir nehmen an, in jeder $(1/n)$-Umgebung von $\mathbf{x}_0$ liegt ein Punkt $\mathbf{x}_n \in M$. Offensichtlich konvergiert dann die Folge $(\mathbf{x}_n)$ gegen $\mathbf{x}_0$, und $\mathbf{x}_0$ muss schon in M liegen. Das kann nicht sein! Also gibt es wenigstens ein $\varepsilon > 0$ mit $U_\varepsilon(\mathbf{x}_0) \subset X \setminus M$. So folgt, dass $X \setminus M$ offen und M selbst abgeschlossen ist. ∎

Häufungspunkte und Grenzwerte von Punktfolgen in einem metrischen Raum definiert man wie in normierten Vektorräumen:

Ein Punkt $\mathbf{x}_0 \in X$ ist ***Häufungspunkt*** einer Folge $(\mathbf{x}_n)$ in X, falls in jeder Umgebung von $\mathbf{x}_0$ unendlich viele Folgenglieder liegen. $\mathbf{x}_0$ ist ***Grenzwert*** der Folge $(\mathbf{x}_n)$, falls in jeder Umgebung von $\mathbf{x}_0$ fast alle Folgenglieder $\mathbf{x}_n$ liegen. Äquivalent dazu ist, dass $d(\mathbf{x}_\nu, \mathbf{x}_0)$ eine Nullfolge (in $\mathbb{R}$) ist.

1.16. Hausdorff'scher Trennungssatz

Sei X ein metrischer Raum. Sind $x, y \in X$ zwei Punkte mit $x \neq y$, so gibt es offene Umgebungen U von x und V von y, so dass $U \cap V = \varnothing$ ist.

Beweis: Wegen $x \neq y$ ist $r := d(x, y) > 0$. Nun sei $0 < \varepsilon < r/2$, $U = B_\varepsilon(x)$ und $V = B_\varepsilon(y)$. Gäbe es einen Punkt z in $U \cap V$, so wäre

$$d(x, y) \leq d(x, z) + d(z, y) < 2\varepsilon < r.$$

Das ist ein Widerspruch. ∎

In einem beliebigen topologischen Raum braucht der Trennungssatz nicht zu gelten.

1.17. Satz

In einem metrischen Raum ist der Grenzwert einer konvergenten Folge eindeutig bestimmt.

Beweis: Man argumentiert wie bei Folgen in $\mathbb{R}$ und benutzt den Trennungssatz.
∎

Zur Erinnerung Eine Menge $K \subset \mathbb{R}$ heißt *kompakt*, wenn jede unendliche Punktfolge in K eine Teilfolge besitzt, die gegen einen Punkt aus K konvergiert. Äquivalent dazu ist die Bedingung, dass jede Punktfolge in K wenigstens einen Häufungspunkt in K besitzt. Diese Eigenschaft wird in der Literatur meistens *„folgenkompakt"* genannt.

Wir haben auch gezeigt, dass eine Menge in $\mathbb{R}$ genau dann kompakt ist, wenn sie abgeschlossen und beschränkt ist.

Definition

Sei X ein metrischer Raum. Eine Teilmenge $K \subset X$ heißt ***kompakt***, falls jede Punktfolge in K eine Teilfolge besitzt, die gegen einen Punkt aus K konvergiert.

Sei X ein beliebiger metrischer Raum und $M \subset X$ eine beliebige Teilmenge. Eine ***offene Überdeckung*** von M ist ein System $(U_\iota)_{\iota \in I}$ von offenen Teilmengen von X mit

$$M \subset \bigcup_{\iota \in I} U_\iota \,.$$

Ist $I_0 = \{\iota_1, \ldots, \iota_N\} \subset I$ eine endliche Teilmenge und $M \subset U_{\iota_1} \cup \ldots \cup U_{\iota_N}$, so nennt man $\{U_{\iota_1}, \ldots, U_{\iota_N}\}$ eine ***endliche Teilüberdeckung***.

Definition

Eine Teilmenge $K \subset X$ besitzt die ***Überdeckungseigenschaft***, falls jede offene Überdeckung von K eine endliche Teilüberdeckung enthält.

Diese Definition ist zwar nicht sehr anschaulich, aber recht praktisch.

1.18. Überdeckungssatz von Heine-Borel

Sei X ein metrischer Raum. Eine Teilmenge $K \subset X$ ist genau dann kompakt, wenn sie die Überdeckungseigenschaft besitzt.

Beweis: 1) Wir setzen voraus, dass K die Überdeckungseigenschaft besitzt und betrachten eine Folge $(\mathbf{x}_\nu)$ in K, so dass $F := \{\mathbf{x}_\nu \,:\, \nu \in \mathbb{N}\}$ eine unendliche Menge ist (sonst ist nichts zu zeigen).

Wenn F keinen Häufungspunkt in K besitzt, dann gibt es zu jedem $\mathbf{x} \in K$ eine offene Umgebung $U_\mathbf{x}$, so dass $U_\mathbf{x} \cap F$ endlich ist. Die Mengen $U_\mathbf{x}$ überdecken K. Wegen der Gültigkeit der Überdeckungseigenschaft gibt es endlich viele offene Mengen $U_1, \ldots, U_N$, die schon K überdecken und für die stets $U_i \cap F$ endlich ist. Das bedeutet, dass F endlich ist. Da haben wir unseren Widerspruch, K muss kompakt sein!

2) Ist umgekehrt K kompakt, so betrachten wir eine offene Überdeckung $(U_\iota)_{\iota\in I}$ von K und zeigen, dass es eine endliche Teilüberdeckung gibt.

a) Wir beweisen zunächst, dass es ein $\varepsilon > 0$ und zu jedem $\mathbf{x} \in K$ ein $\iota(\mathbf{x}) \in I$ mit $U_\varepsilon(\mathbf{x}) \subset U_{\iota(\mathbf{x})}$ gibt.

Gibt es nämlich dieses ε nicht, so finden wir zu jedem $\nu \in \mathbb{N}$ ein $\mathbf{x}_\nu \in K$, so dass $U_{1/\nu}(\mathbf{x}_\nu)$ in keinem U_ι enthalten ist. Weil K kompakt ist, können wir aus der Folge $(\mathbf{x}_\nu)$ eine Teilfolge $\mathbf{y}_i = \mathbf{x}_{\nu_i}$ auswählen, die gegen ein $\mathbf{y}_0 \in K$ konvergiert. Natürlich muss dieses $\mathbf{y}_0$ in einem Überdeckungselement U_{ι_0} liegen.

Wir wählen ein $r > 0$, so dass auch noch $U_r(\mathbf{y}_0) \subset U_{\iota_0}$ ist. Ist $i \in \mathbb{N}$ hinreichend groß, so ist $1/\nu_i < r/2$ und zugleich $\mathbf{y}_i \in U_{r/2}(\mathbf{y}_0)$. Für alle $\mathbf{x} \in U_{1/\nu_i}(\mathbf{y}_i)$ gilt dann

$$d(\mathbf{x}, \mathbf{y}_0) \leq d(\mathbf{x}, \mathbf{y}_i) + d(\mathbf{y}_i, \mathbf{y}_0) < \frac{1}{\nu_i} + \frac{r}{2} < r.$$

Also liegt $U_{1/\nu_i}(\mathbf{x}_{\nu_i}) = U_{1/\nu_i}(\mathbf{y}_i)$ in $U_r(\mathbf{y}_0) \subset U_{\iota_0}$. Das widerspricht der Konstruktion.

b) Wir können also annehmen, dass ein ε mit der gewünschten Eigenschaft existiert. Nun konstruieren wir eine neue Folge in K.

Zunächst wählen wir einen beliebigen Punkt $\mathbf{a}_1 \in K$. Wenn nicht schon $U_\varepsilon(\mathbf{a}_1)$ überdeckt, dann wählen wir einen Punkt $\mathbf{a}_2 \in K \setminus U_\varepsilon(\mathbf{a}_1)$. Wenn nicht schon $U_\varepsilon(\mathbf{a}_1)$ und $U_\varepsilon(\mathbf{a}_2)$ überdecken, dann wählen wir einen Punkt $\mathbf{a}_3 \in K \setminus (U_\varepsilon(\mathbf{a}_1) \cup U_\varepsilon(\mathbf{a}_2))$ und so weiter.

Würde die Folge der $\mathbf{a}_n$ nicht abbrechen, so müsste eine geeignete Teilfolge $\mathbf{z}_\mu = \mathbf{a}_{n_\mu}$ einen Grenzwert $\mathbf{z}_0 \in K$ besitzen. Für großes μ müsste dann $d(\mathbf{z}_\mu, \mathbf{z}_{\mu+1}) < \varepsilon$ sein, aber andererseits liegt $\mathbf{z}_{\mu+1}$ nach Konstruktion nicht in $U_\varepsilon(\mathbf{z}_\mu)$. Das ist ein Widerspruch, $(U_\iota)_{\iota\in I}$ enthält eben doch eine endliche Teilüberdeckung. Also besitzt K die Überdeckungseigenschaft. ∎

1.19. Satz

Im $\mathbb{R}^n$ ist eine Teilmenge K genau dann kompakt, wenn sie abgeschlossen und beschränkt ist.

Beweis: 1) Sei K abgeschlossen und beschränkt. Eine Folge in K ist dann auch beschränkt, und nach dem Satz von Bolzano-Weierstrass konvergiert eine Teilfolge gegen einen Punkt $\mathbf{x}_0$ des $\mathbb{R}^n$. Weil K abgeschlossen ist, muss $\mathbf{x}_0$ in K liegen.

2) Sei K kompakt. Da man K durch endlich viele Kugeln von festem Radius überdecken kann, ist K auch beschränkt. Sei nun $\mathbf{x}_0 \in \mathbb{R}^n \setminus K$. Zu jedem $\mathbf{x} \in K$ gibt es ein $\delta = \delta_x > 0$ und ein $\varepsilon = \varepsilon_x > 0$, so dass $U_\delta(x) \cap U_\varepsilon(x_0) = \varnothing$ ist. Endlich viele der δ-Umgebungen überdecken schon K. Der Durchschnitt der zugehörigen ε-Umgebungen liefert eine Umgebung von x_0, die ganz in $\mathbb{R}^n \setminus K$ liegt. Also ist $\mathbb{R}^n \setminus K$ offen und K selbst abgeschlossen. ∎

Hier kommt eine Anwendung:

1.20. Satz

Sei $M \subset \mathbb{R}^n$ offen und $K \subset M$ kompakt. Dann gibt es eine offene Menge $U \subset \mathbb{R}^n$, so dass $\overline{U}$ kompakt und $K \subset U \subset \overline{U} \subset M$ ist.

Beweis: Zu jedem $\mathbf{x} \in K$ gibt es eine offene Kugel $B_{\mathbf{x}}$ um $\mathbf{x}$, die noch ganz in M enthalten ist. Dann sei $B'_{\mathbf{x}}$ die Kugel mit dem halben Radius, so dass sogar $\overline{B'_{\mathbf{x}}} \subset M$ ist. Wegen der Kompaktheit von K gibt es endlich viele Punkte $\mathbf{x}_1, \ldots, \mathbf{x}_N$, so dass die Kugeln $B'_\nu := B'_{\mathbf{x}_\nu}$, $\nu = 1, \ldots, N$, schon ganz K überdecken. Dann ist $U = B'_1 \cup \ldots \cup B'_N$ offen, $\overline{U}$ kompakt, $K \subset U$ und $\overline{U} \subset M$. ■

Bemerkung: Man sagt, eine offene Menge U liegt ***relativ-kompakt*** in einer offenen Menge W, falls $\overline{U}$ kompakt und in W enthalten ist. Als Abkürzung dafür schreiben wir „$U \subset\subset W$“.

Wir kommen nun zum Begriff der stetigen Abbildung.

Definition

Sei $M \subset \mathbb{R}^n$. Eine Abbildung $\mathbf{f} : M \to \mathbb{R}^k$ heißt ***stetig in*** $\mathbf{x}_0 \in M$, falls es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass für alle $\mathbf{x} \in M$ gilt:

$$\text{Ist } \|\mathbf{x} - \mathbf{x}_0\| < \delta, \text{ so ist } \|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x}_0)\| < \varepsilon.$$

$\mathbf{f}$ heißt ***stetig auf*** M, falls $\mathbf{f}$ in jedem Punkt von M stetig ist.

Für Funktionen $f : I \to \mathbb{R}$ ist dies immer noch die alte Definition.

Eine Abbildung $\mathbf{f} : M \to \mathbb{R}^k$ ist genau dann stetig in $\mathbf{x}_0 \in M$, falls es zu jeder Umgebung U von $\mathbf{f}(\mathbf{x}_0)$ eine Umgebung V von $\mathbf{x}_0$ mit $\mathbf{f}(V \cap M) \subset U$ gibt.

1.21. Verkettungen von stetigen Abbildungen sind stetig

Sei $M \subset \mathbb{R}^n$ und $N \subset \mathbb{R}^m$. Ist $\mathbf{f} : M \to \mathbb{R}^m$ stetig in $\mathbf{x}_0 \in M$, $\mathbf{f}(M) \subset N$ und $\mathbf{g} : N \to \mathbb{R}^k$ stetig in $\mathbf{y}_0 := \mathbf{f}(\mathbf{x}_0) \in N$, so ist auch $\mathbf{g} \circ \mathbf{f} : M \to \mathbb{R}^k$ stetig in $\mathbf{x}_0$.

Beweis: Sei $\mathbf{z}_0 := \mathbf{g}(\mathbf{y}_0) = (\mathbf{g} \circ \mathbf{f})(\mathbf{x}_0)$ und $W = W(\mathbf{z}_0) \subset \mathbb{R}^k$ eine Umgebung. Dann gibt es eine Umgebung $V = V(\mathbf{y}_0) \subset \mathbb{R}^m$ mit $\mathbf{g}(V \cap N) \subset W$, sowie eine Umgebung $U = U(\mathbf{x}_0) \subset \mathbb{R}^n$ mit $\mathbf{f}(U \cap M) \subset V \cap N$. Es folgt, dass $(\mathbf{g} \circ \mathbf{f})(U \cap M) \subset W$ ist, also $\mathbf{g} \circ \mathbf{f}$ stetig in $\mathbf{x}_0$. ■

1.22. Beispiele

A. Die *identische Abbildung* $\mathrm{id} : \mathbb{R}^n \to \mathbb{R}^n$ mit $\mathrm{id}(\mathbf{x}) := \mathbf{x}$ ist stetig, denn für jede Umgebung $U = U(\mathrm{id}(\mathbf{x}_0))$ ist U auch eine Umgebung von $\mathbf{x}_0$ und $\mathrm{id}(U) = U$.

B. Sei $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ eine *lineare Abbildung*. Einfachstes Beispiel ist eine lineare Funktion $f : \mathbb{R} \to \mathbb{R}$, gegeben durch $f(x) = ax$ mit einem festen Faktor a.

Sind $\mathbf{e}_1, \dots, \mathbf{e}_n$ die Einheitsvektoren im $\mathbb{R}^n$, so kann jeder Vektor $\mathbf{x} \in \mathbb{R}^n$ in der Form $\mathbf{x} = (x_1, \dots, x_n) = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n$ geschrieben werden. Für eine lineare Abbildung $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ ist dann

$$\mathbf{f}(\mathbf{x}) = x_1\mathbf{f}(\mathbf{e}_1) + \cdots + x_n\mathbf{f}(\mathbf{e}_1),$$

und wir erhalten die Abschätzung

$$\begin{aligned} \|\mathbf{f}(\mathbf{x})\| &= \Big\|\sum_{i=1}^{n} x_i \cdot \mathbf{f}(\mathbf{e}_i)\Big\| \\ &\le \sum_{i=1}^{n} |x_i| \cdot \|\mathbf{f}(\mathbf{e}_i)\| \qquad \text{(Dreiecksungleichung)} \\ &\le C \cdot \max_i |x_i| \qquad \Big(\text{mit } C := \sum_{i=1}^{n} \|\mathbf{f}(\mathbf{e}_i)\|\Big) \\ &\le C \cdot \|\mathbf{x}\|. \end{aligned}$$

Dabei ist C eine nur von $\mathbf{f}$ abhängige Konstante, und es wurde die Ungleichung $|x_i| = \sqrt{x_i^2} \le \sqrt{x_1^2 + \cdots + x_n^2} = \|\mathbf{x}\|$ benutzt.

Aus der gewonnenen Ungleichung und der Linearität von $\mathbf{f}$ ergibt sich

$$\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})\| = \|\mathbf{f}(\mathbf{x} - \mathbf{y})\| \le C \cdot \|\mathbf{x} - \mathbf{y}\|.$$

Daraus folgt, dass $\mathbf{f}$ überall stetig ist. Ist nämlich $\mathbf{x}_0 \in \mathbb{R}^n$ und $\varepsilon > 0$, so kann man $\delta := \varepsilon / C$ setzen. Ist dann $\|\mathbf{x} - \mathbf{x}_0\| < \delta$, so ist $\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x}_0)\| \le C \cdot \|\mathbf{x} - \mathbf{x}_0\| < C \cdot \delta = \varepsilon$.

C. Sei $M \subset \mathbb{R}^n$ und $\mathbf{x}_0 \in M$. Eine Abbildung $\mathbf{f} = (f_1, \dots, f_m) : M \to \mathbb{R}^m$ ist genau dann stetig in $\mathbf{x}_0$, wenn alle Komponenten-Funktionen $f_i : M \to \mathbb{R}$ stetig in $\mathbf{x}_0$ sind. In der einen Richtung folgt das aus der Gleichung $f_i = \mathrm{pr}_i \circ \mathbf{f}$ und dem Satz über die Stetigkeit der Verkettung stetiger Funktionen. Zum Beweis der anderen Richtung sei ein $\varepsilon > 0$ vorgegeben. Dann gibt es zu jedem i ein $\delta_i > 0$, so dass $|f_i(\mathbf{x}) - f_i(\mathbf{x}_0)| < \varepsilon/\sqrt{m}$ für alle $\mathbf{x} \in U_{\delta_i}(\mathbf{x}_0) \cap M$ ist. Wir setzen $\delta := \min(\delta_1, \dots, \delta_m)$. Ist $\mathbf{x} \in U_\delta(\mathbf{x}_0) \cap M$, so ist

$$\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x}_0)\| = \sqrt{\sum_{i=1}^{m} |f_i(\mathbf{x}) - f_i(\mathbf{x}_0)|^2} < \sqrt{m\varepsilon^2/m} = \varepsilon.$$

Man darf nun allerdings nicht dem Trugschluss unterliegen, eine Funktion von mehreren Veränderlichen sei schon stetig, wenn sie in jeder einzelnen Variablen stetig ist. Sei etwa

$$f(x,y) := \begin{cases} \dfrac{2xy}{x^2+y^2} & \text{für } (x,y) \neq (0,0), \\ 0 & \text{für } (x,y) = (0,0). \end{cases}$$

Diese Funktion ist überall definiert, und die Funktionen $f(x,0) \equiv 0$ und $f(0,y) \equiv 0$ sind im Nullpunkt stetig. Dennoch ist f selbst dort nicht stetig, denn für $x \neq 0$ ist $f(x,x) \equiv 1$.

1.23. Folgenkriterium

Sei $M \subset \mathbb{R}^n$. Eine Abbildung $\mathbf{f} : M \to \mathbb{R}^m$ ist genau dann stetig in $\mathbf{x}_0 \in M$, wenn für jede Folge $(\mathbf{x}_\nu)$ in M mit $\mathbf{x}_\nu \neq \mathbf{x}_0$ und $\lim\limits_{\nu\to\infty} \mathbf{x}_\nu = \mathbf{x}_0$ gilt: $\lim\limits_{\nu\to\infty} \mathbf{f}(\mathbf{x}_\nu) = \mathbf{f}(\mathbf{x}_0)$.

Beweis: Wie in einer Veränderlichen. ∎

1.24. Offenheit von Ungleichungen

Sei $M \subset \mathbb{R}^n$ offen und $f : M \to \mathbb{R}$ eine stetige Funktion. Dann ist auch die Menge $P := \{\mathbf{x} \in M \,:\, f(\mathbf{x}) > 0\}$ offen.

Beweis: Wie in einer Veränderlichen. ∎

1.25. Stetige Bilder kompakter Mengen sind kompakt

Sei $K \subset \mathbb{R}^n$ kompakt und $\mathbf{f} : K \to \mathbb{R}^m$ eine stetige Abbildung. Dann ist auch $\mathbf{f}(K)$ kompakt.

Beweis: Sei $(\mathbf{y}_\nu)$ eine Folge von Punkten in $\mathbf{f}(K)$. Dann gibt es zu jedem ν einen Punkt $\mathbf{x}_\nu \in K$ mit $\mathbf{f}(\mathbf{x}_\nu) = \mathbf{y}_\nu$. Weil K kompakt ist, besitzt die Folge $(\mathbf{x}_\nu)$ eine konvergente Teilfolge $(\mathbf{x}_{\nu_i})$, ihr Grenzwert in K sei mit $\mathbf{x}_0$ bezeichnet. Wegen der Stetigkeit von $\mathbf{f}$ konvergiert $(\mathbf{y}_{\nu_i})$ gegen $\mathbf{f}(\mathbf{x}_0)$, und dieser Punkt liegt in $\mathbf{f}(K)$. ∎

1.26. Folgerung

Sei $K \subset \mathbb{R}^n$ kompakt. Dann nimmt jede stetige Funktion $f : K \to \mathbb{R}$ auf K ihr Maximum und ihr Minimum an.

Beweis: Der Beweis wird wie im Falle einer Veränderlichen geführt. ∎

Die Zahlenbereiche $\mathbb{R}$ und $\mathbb{C}$ sind nicht nur $\mathbb{R}$-Vektorräume, sie besitzen zusätzlich eine multiplikative Struktur, so dass $|x \cdot y| = |x| \cdot |y|$ ist. Es gibt auch andere Vektorräume mit einer multiplikativen Struktur, z.B. den Raum $M_n(\mathbb{R})$ der quadratischen Matrizen mit der Matrizenmultiplikation.

Definition

Unter einer $\mathbb{R}$-***Algebra*** (bzw. $\mathbb{C}$-***Algebra***) versteht man einen $\mathbb{R}$-Vektorraum (bzw. $\mathbb{C}$-Vektorraum) E mit einer zusätzlichen Multiplikation, so dass für $\mathbf{u}, \mathbf{v}, \mathbf{w} \in E$ gilt:

1. $\mathbf{u} \cdot (\mathbf{v} \cdot \mathbf{w}) = (\mathbf{u} \cdot \mathbf{v}) \cdot \mathbf{w}$ (Assoziativgesetz),
2. $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$ und $(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w}$ (Distributivgesetze).
3. $\alpha(\mathbf{v} \cdot \mathbf{w}) = (\alpha\mathbf{v}) \cdot \mathbf{w} = \mathbf{v} \cdot (\alpha\mathbf{w})$ für $\alpha \in \mathbb{R}$ (bzw. $\in \mathbb{C}$).

Ist E normiert und $\|\mathbf{v} \cdot \mathbf{w}\| \le \|\mathbf{v}\| \cdot \|\mathbf{w}\|$ für alle $\mathbf{v}, \mathbf{w} \in E$, so nennt man E eine ***normierte Algebra***.

Offensichtlich ist der Raum $\mathscr{C}^0\big([a,b]\big)$ der stetigen Funktionen auf einem abgeschlossenen Intervall ein Beispiel für eine normierte Algebra.

Wir wollen zeigen, dass auch $M_n(\mathbb{R})$ eine normierte $\mathbb{R}$-Algebra ist. Dazu müssen wir noch die Norm einer Matrix einführen.

Für eine Matrix

$$A = \left(a_{ij} \;\middle|\; \begin{array}{l} i = 1, \ldots, n \\ j = 1, \ldots, n \end{array} \right) \in M_n(\mathbb{R})$$

setzen wir

$$\|A\| := \sqrt{\sum_{i,j} a_{ij}^2}.$$

Das ist nichts anderes, als die gewöhnliche euklidische Norm von A in $M_n(\mathbb{R}) \cong \mathbb{R}^{n^2}$. Nun gilt:

1. $\|A + B\| \le \|A\| + \|B\|$.
2. $\|\lambda \cdot A\| = |\lambda| \cdot \|A\|$ für $\lambda \in \mathbb{R}$.
3. $\|A\| = 0 \iff A = 0$.
4. Sind $A, B \in M_n(\mathbb{R})$, so ist $\|A \cdot B\| \le \|A\| \cdot \|B\|$.

   Diese Aussage muss noch bewiesen werden. $A \cdot B$ hat an der Stelle (i,j) den Eintrag

$$\sum_{k=1}^{n} a_{ik} b_{kj} = \mathbf{z}_i(A) \bullet \mathbf{s}_j(B),$$

wenn man mit $\mathbf{z}_i(A) \in \mathbb{R}^n$ die i-te Zeile von A und mit $\mathbf{s}_j(B) \in \mathbb{R}^n$ die j-te Spalte von B bezeichnet. Mit Hilfe der Schwarz'schen Ungleichung folgt dann:

$$\begin{aligned}
\|A \cdot B\|^2 &= \sum_{i,j} (\mathbf{z}_i(A) \bullet \mathbf{s}_j(B))^2 \\
&\le \sum_{i,j} \|\mathbf{z}_i(A)\|^2 \cdot \|\mathbf{s}_j(B)\|^2 \\
&= \left(\sum_i \|\mathbf{z}_i(A)\|^2 \right) \cdot \left(\sum_j \|\mathbf{s}_j(B)\|^2 \right) \\
&= \|A\|^2 \cdot \|B\|^2.
\end{aligned}$$

Also ist $M_n(\mathbb{R})$ eine normierte Algebra. Eng verwandt mit dem Raum der Matrizen ist der Raum $L(\mathbb{R}^n, \mathbb{R}^n)$ der linearen Abbildungen von $\mathbb{R}^n$ nach $\mathbb{R}^n$, mit der Verknüpfung von Abbildungen als Multiplikation. Hier bietet sich eine andere Norm an, die genauso auch für Matrizen verwendet werden kann.

Definition

Sei $f : \mathbb{R}^n \to \mathbb{R}^n$ linear. Dann nennen wir

$$\|f\|_{\text{op}} := \sup\{\|f(\mathbf{x})\| \ : \ \|\mathbf{x}\| \le 1\}$$

die ***Operator-Norm*** von f.

1.27. Satz

Die Operator-Norm ist eine Norm, und für $\mathbf{x} \in \mathbb{R}^n$ ist $\|f(\mathbf{x})\| \le \|f\|_{\text{op}} \cdot \|\mathbf{x}\|$. Außerdem ist $\|f \circ g\|_{\text{op}} \le \|f\|_{\text{op}} \cdot \|g\|_{\text{op}}$.

BEWEIS: 1) Offensichtlich ist stets $\|f\|_{\text{op}} \ge 0$ und $\|f\|_{\text{op}} = 0 \iff f = 0$.

2) Für $\alpha \in \mathbb{R}$ ist $\|\alpha f\|_{\text{op}} = \sup\limits_{\|\mathbf{x}\| \le 1} \|(\alpha f)(\mathbf{x})\| = |\alpha| \cdot \sup\limits_{\|\mathbf{x}\| \le 1} \|f(\mathbf{x})\| = |\alpha| \cdot \|f\|_{\text{op}}$.

3) Es ist $\|f + g\|_{\text{op}} = \sup\limits_{\|\mathbf{x}\| \le 1} \|(f + g)(\mathbf{x})\| \le \sup\limits_{\|\mathbf{x}\| \le 1} \big(\|f(\mathbf{x})\| + \|g(\mathbf{x})\|\big) \le \|f\|_{\text{op}} + \|g\|_{\text{op}}$.

4) Ist $\mathbf{x} \ne \mathbf{0}$ ein beliebiger Vektor, so ist

$$\frac{\|f(\mathbf{x})\|}{\|\mathbf{x}\|} = \|f\Big(\frac{\mathbf{x}}{\|\mathbf{x}\|}\Big)\| \le \|f\|_{\text{op}}, \text{ also } \|f(\mathbf{x})\| \le \|f\|_{\text{op}} \cdot \|\mathbf{x}\|.$$

5) Es ist $\|f \circ g\|_{\text{op}} = \sup\limits_{\|\mathbf{x}\| \le 1} \|f(g(\mathbf{x}))\| \le \sup\limits_{\|\mathbf{x}\| \le 1} \|f\|_{\text{op}} \cdot \|g(\mathbf{x})\| = \|f\|_{\text{op}} \cdot \|g\|_{\text{op}}$. ∎

Jede Matrix $A \in M_n(\mathbb{R})$ definiert eine lineare Abbildung $f_A : \mathbb{R}^n \to \mathbb{R}^n$ durch $f_A(\mathbf{x}) := \mathbf{x} \cdot A^\top$. Daher kann man $\|A\|_{\text{op}} := \|f_A\|_{\text{op}}$ setzen. Sei nun $|A|$ die Maximumnorm von A. Dann gilt:

1.28. Satz

Für die Normen auf $M_n(\mathbb{R})$ gilt $\quad |A| \le \|A\|_{\text{op}} \le \|A\|$.

Beweis: 1) Es ist $f_A(\mathbf{e}_j) = \mathbf{s}_j(A)$ und daher

$$|a_{ij}| \le \sqrt{a_{1j}^2 + \cdots + a_{nj}^2} = \|f_A(\mathbf{e}_j)\| \le \|f_A\|_{\text{op}}.$$

Daraus folgt die Ungleichung $|A| \le \|A\|_{\text{op}}$.

2) Bezeichnen wir weiterhin die i-te Zeile von A mit $\mathbf{z}_i(A)$, so sind die Komponenten von $A \cdot \mathbf{x}^\top$ die Skalarprodukte $\mathbf{z}_i(A) \bullet \mathbf{x}$. Daraus folgt:

$$\begin{aligned}
\|A\|_{\text{op}}^2 &= \sup_{\|\mathbf{x}\| \le 1} \|A \cdot \mathbf{x}^\top\|^2 = \sup_{\|\mathbf{x}\| \le 1} \sum_{i=1}^n (\mathbf{z}_i(A) \bullet \mathbf{x})^2 \\
&\le \sup_{\|\mathbf{x}\| \le 1} \sum_{i=1}^n \|\mathbf{z}_i(A)\|^2 \cdot \|\mathbf{x}\|^2 \qquad \text{(Cauchy-Schwarz)} \\
&\le \sum_{i=1}^n \|\mathbf{z}_i(A)\|^2 = \|A\|.
\end{aligned}$$

■