1 Der n-dimensionale Zahlenraum $\mathbb{K}^n$

In diesem Kapitel werden als Vorbereitung der Differential- und Integralrechnung von Funktionen mehrerer Variabler zunächst die wichtigsten Eigenschaften des euklidischen Vektorraumes $\mathbb{K}^n$ zusammengestellt. Dabei steht $\mathbb{K}$ wieder für den Körper $\mathbb{R}$ der reellen Zahlen oder den Körper $\mathbb{C}$ der komplexen Zahlen. Die kanonischen Fälle bei geometrischen Anwendungen sind natürlich die Ortsdimensionen $n = 2$ und $n = 3$, oder bei Einbeziehung der Zeitvariablen auch $n = 4$. Bei der Diskretisierung von Differentialgleichungen treten aber auch beliebig hoch dimensionale Probleme auf, so dass die Betrachtung der allgemeinen Dimension $n \in \mathbb{N}$ keine bloße mathematische Spielerei ist.

1.1 Der euklidische Raum $\mathbb{K}^n$

Für $n \in \mathbb{N}$ bezeichnet $\mathbb{K}^n$ den Vektorrraum der n-Tupel $x = (x_1, \dots, x_n)$ mit Komponenten $x_i \in \mathbb{K}$, $i = 1, \dots, n$. Für diese sind Addition und skalare Multiplikation definiert:

$$x + y := (x_1 + y_1, \dots, x_n + y_n), \qquad \alpha x := (\alpha x_1, \dots, \alpha x_n), \quad \alpha \in \mathbb{K}.$$

Die Elemente $x \in \mathbb{K}^n$ werden je nach Interpretation als „Punkte“ oder „Vektoren“ angesprochen. Dabei kann man sich x als den Endpunkt eines Vektors vorstellen, der im Ursprung des gewählten kartesischen[1] Koordinatensystem angeheftet ist, und die Komponenten x_i als Koordinaten bezüglich dieses Koordinatensystems. In diesem Fall fassen wir Vektoren stets als „Spaltenvektoren“ auf und schreiben dafür im Rahmen des Matrizenkalküls auch $(x_1, \dots, x_n)^T$. Der „Nullvektor“ $(0, \dots, 0)^T$ wird ebenfalls kurz mit 0 bezeichnet. Wir bevorzugen diese koordinatenorientierte Darstellung wegen ihrer Vertrautheit; eine koordinatenfreie Beschreibung ist möglich aber weniger anschaulich.

Wir rekapitulieren einige Eigenschaften des $\mathbb{K}^n$, die im Folgenden verwendet werden. Ein System von m Vektoren $\{a^{(1)}, \dots, a^{(m)}\} \subset \mathbb{K}^n$ heißt „linear abhängig“, wenn es Skalare $\alpha_i \in \mathbb{K}$, $i = 1, \dots, m$, gibt, die nicht alle Null sind, so dass

$$\sum_{i=1}^{n} \alpha_i a^{(i)} = 0,$$

andernfalls „linear unabhängig“. Ein System linear unabhängiger Vektoren des $\mathbb{K}^n$ kann maximal n Elemente enthalten; ein solches „maximales“ System heißt „Basis“ des $\mathbb{K}^n$ und bestimmt mit seiner Mächtigkeit n die „Dimension“ des $\mathbb{K}^n$. Die natürliche Basis des $\mathbb{K}^n$ ist die „euklidische Basis“ („kartesische“) bestehend aus den Vektoren $e^{(i)} := (\delta_{i1}, \dots, \delta_{in})$, $i = 1, \dots, n$. Offenbar ist jedes $x \in \mathbb{K}^n$ darstellbar in der Form

$$x = \sum_{i=1}^{n} x_i e^{(i)},$$

d. h. als „Linearkombination“ der Basisvektoren $e^{(i)}$.

[1] René Descartes (1596–1650): Französischer Mathematiker und Philosoph („cogito ergo sum“); wirkte in Holland und zuletzt in Stockholm; erkannte als erster die enge Beziehung zwischen Geometrie und Arithmetik und begründete die analytische Geometrie.

Definition 1.1: *Sei V irgend ein Vektorraum über dem Körper $\mathbb{K}$. Eine Abbildung $\|\cdot\| : V \to \mathbb{R}$ heißt „Norm“ (auf V), wenn folgende Bedingungen erfüllt sind:*

(N1) Definitheit: $\|x\| \geq 0, \quad \|x\| = 0 \ \Leftrightarrow\ x = 0;$

(N2) Homogenität: $\|\alpha x\| = |\alpha|\,\|x\|, \ \ \alpha \in \mathbb{K};$

(N3) Dreiecksungleichung: $\|x+y\| \leq \|x\| + \|y\|.$

Das Paar $(V, \|\cdot\|)$ wird „normierter Raum“ genannt.

Bemerkung 1.1: Die Nichtnegativität $\|x\| \geq 0$ ist eigentlich bereits eine notwendige Konsequenz der anderen Annahmen. Mit (N2) folgt zunächst $0 = \|0\|$ und dann mit (N3) und (N2) $0 = \|x-x\| \leq \|x\| + \|-x\| = 2\|x\|$. Mit Hilfe von (N3) erhält man analog zum Absolutbetrag für eine beliebige Norm die wichtige Ungleichung

$$\big|\|x\| - \|y\|\big| \leq \|x-y\|, \quad x, y \in V. \tag{1.1.1}$$

Beispiel 1.1: Das bekannteste Beispiel einer Norm auf $\mathbb{K}^n$ ist die „euklidische Norm“

$$\|x\|_2 := \Big(\sum_{i=1}^n |x_i|^2\Big)^{1/2}.$$

Weitere Beispiele von gebräuchlichen Normen sind die „Maximumnorm“ (oder „l_∞-Norm) und die sog. „l_1-Norm“

$$\|x\|_\infty := \max_{i=1,\dots,n} |x_i|, \qquad \|x\|_1 := \sum_{i=1}^n |x_i|.$$

Die Normeigenschaft von $\|\cdot\|_\infty$ und $\|\cdot\|_1$ ergibt sich unmittelbar aus den entsprechenden Eigenschaften des Absolutbetrags. Die Normeigenschaft von $\|\cdot\|_2$ folgt aus seiner Beziehung zum euklidischen Skalarprodukt, was wir später noch genauer diskutieren werden. Quasi zwischen l_1-Norm und Maximumnorm liegen die sog. „l_p-Normen“ für $1 < p < \infty$:

$$\|x\|_p := \Big(\sum_{i=1}^n |x_i|^p\Big)^{1/p}.$$

Dass dies wirklich Normen sind, d. h. dass insbesondere die Dreiecksungleichung gilt, werden wir später noch sehen.

Mit Hilfe einer Norm $\|\cdot\|$ wird für Vektoren $x, y \in \mathbb{K}^n$ eine „Abstandsfunktion“ (oder „Metrik“) erklärt durch $d(x,y) := \|x-y\|$.

Definition 1.2: *Sei X irgend eine Menge. Eine Abbildung $d(\cdot,\cdot) : X \times X \to \mathbb{R}$ heißt „Metrik“ (auf X), wenn folgende Bedingungen erfüllt sind:*

(M1) Definitheit: $d(x,y) \geq 0, \quad d(x,y) = 0 \ \Leftrightarrow\ x = y;$

(M2) Symmetrie: $d(x,y) = d(y,x);$

(M3) Dreiecksungleichung: $d(x,y) \leq d(x,z) + d(z,y).$

Das Paar (X, d) wird „metrischer Raum“ genannt.

Bemerkung 1.2: Viele Aussagen dieses und des folgenden Kapitels, die nicht die Vektorraumstruktur benötigen, lassen sich auch ganz allgemein für metrische Räume formulieren. Die so gewonnene Theorie hat dann viele über den Rahmen des endlich dimensionalen $\mathbb{K}^n$ hinausgehende Anwendungen, z. B. auf unendlich dimensionale Funktionenräume wie den $C[a,b]$ und den $R[a,b]$ (siehe Kapitel 4 und Kapitel 7 des Bandes Analysis 1). Da aber die Behandlung wirklich interessanter Anwendungen in der Physik und in anderen Wissenschaften beträchtlichen Aufwand erfordert, wollen wir hier auf diese formale Allgemeinheit zugunsten größerer Anschaulichkeit verzichten.

Bemerkung 1.3: Die natürliche Verallgemeinerung des *endlich* dimensionalen Raumes $\mathbb{K}^n$ ist der *unendlich* dimensionale Folgenraum l_2 der quadratisch summierbaren Zahlenfolgen $x = (x_k)_{k\in\mathbb{N}},\ x_k \in \mathbb{K}$, d. h. der Folgen, für die $\sum_{k=1}^\infty |x_k|^2$ konvergiert, versehen mit der Norm

$$\|x\|_2 := \Big(\sum_{k=1}^\infty |x_k|^2 \Big)^{1/2}.$$

Der Nachweis der Vektorraum- und Normeigenschaften sei als Übungsaufgabe gestellt. Es ist zweckmäßig, sich die im Folgenden für den $\mathbb{K}^n$ formulierten Aussagen durch Überprüfung ihrer Gültigkeit im $(l_2, \|\cdot\|_2)$ klarzumachen.

Mit Hilfe einer Norm wird der „Abstand“ $d(x,x') := \|x-x'\|$ zweier Vektoren im $\mathbb{K}^n$ definiert. Damit lassen sich dann auch für Punktmengen des $\mathbb{K}^n$ die schon vom $\mathbb{K}^1$ bekannten topologischen Begriffe „offen“, „abgeschlossen“, „kompakt“, „Durchmesser“ und „Umgebung“ definieren. Im Folgenden verwenden wir hierzu zunächst die Maximumnorm $\|\cdot\|_\infty$, werden aber später sehen, dass dies unabhängig von der gewählten Norm ist. Für ein $a \in \mathbb{K}^n$ und $r > 0$ wird die Menge

$$K_r(a) := \{x \in \mathbb{K}^n :\ \|x-a\|_\infty < r\}$$

eine „Kugelumgebung“ mit Radius r genannt.

Definition 1.3: *Eine Folge von Vektoren $(x^{(k)})_{k\in\mathbb{N}}$ des $\mathbb{K}^n$ heißt:*

i) „beschränkt“, wenn alle ihre Elemente in einer Kugelumgebung $K_R(0)$ liegen;

ii) „Cauchy-Folge“, wenn es zu jedem $\varepsilon \in \mathbb{R}_+$ ein $N_\varepsilon \in \mathbb{N}$ gibt, so dass für alle $k,l \ge N$

$$\|x^{(k)} - x^{(l)}\|_\infty < \varepsilon;$$

iii) „konvergent“ gegen ein $x \in \mathbb{K}^n$, wenn

$$\|x^{(k)} - x\|_\infty \to 0 \quad (k \to \infty).$$

Für eine konvergente Folge $(x^{(k)})_{k\in\mathbb{N}}$ schreiben wir wieder $\lim_{k\to\infty} x^{(k)} = x$ oder auch kurz $x^{(k)} \to x\ (k \to \infty)$. Geometrisch bedeutet dies, dass jede Kugelumgebung $K_\varepsilon(x)$ von x fast alle (d. h. alle bis auf endlich viele) der Folgenelemente $x^{(k)}$ enthält. Die so definierte

Konvergenz $x^{(k)} \to x \ (k \to \infty)$ ist offenbar gleichbedeutend mit der komponentenweisen Konvergenz:

$$\|x^{(k)} - x\|_\infty \to 0 \ \ (k \to \infty) \quad \Leftrightarrow \quad x_i^{(k)} \to x_i \ \ (k \to \infty), \ i = 1, \dots, n.$$

Damit kann die Theorie der Konvergenz von Vektorfolgen in $\mathbb{K}^n$ auf die von Zahlenfolgen in $\mathbb{K}$ zurückgeführt werden. Wir gewinnen so zunächst das n-dimensionale Analogon des Cauchyschen Konvergenzkriteriums und des Satzes von Bolzano-Weierstraß.

Satz 1.1 (Satz von Cauchy und Satz von Bolzano-Weierstraß):
i) Jede Cauchy-Folge im $\mathbb{K}^n$ konvergiert, d. h.: Der normierte Raum $(\mathbb{K}^n, \|\cdot\|_\infty)$ ist vollständig. Ein vollständiger normierter Raum wird „Banach-Raum“ genannt.

ii) Jede beschränkte Folge in $\mathbb{K}^n$ besitzt eine konvergente Teilfolge.

Beweis: i) Für eine Cauchy-Folge $(x^{(k)})_{k\in\mathbb{N}}$ sind wegen $|x_i| \le \|x\|_\infty$, $i = 1, \dots, n$, für $x \in \mathbb{K}^n$, auch die Komponentenfolgen $(x_i^{(k)})_{k\in\mathbb{N}}$, $i = 1, \dots, n$, Cauchy-Folgen in $\mathbb{K}$ und konvergieren jede für sich gegen Grenzwerte $x_i \in \mathbb{K}$. Der Vektor $x := (x_1, \dots, x_n) \in \mathbb{K}^n$ ist dann Limes der Folge $(x^{(k)})_{k\in\mathbb{N}}$ bzgl. der Maximumnormkonvergenz.

ii) Für eine beschränkte Folge $(x^{(k)})_{k\in\mathbb{N}}$ sind auch die Komponentenfolgen $(x_i^{(k)})_{k\in\mathbb{N}}$, $i = 1, \dots, n$, beschränkt. Durch sukzessive Anwendung des Satzes von Bolzano-Weierstraß in $\mathbb{K}$ erhalten wir zunächst eine konvergente Teilfolge $(x_1^{(k_{1j})})_{j\in\mathbb{N}}$ von $(x_1^{(k)})_{k\in\mathbb{N}}$ mit $x_1^{(k_{1j})} \to x_1 \ (j \to \infty)$, dann eine konvergente Teilfolge $(x_2^{(k_{2j})})_{j\in\mathbb{N}}$ von $(x_2^{(k_{1j})})_{j\in\mathbb{N}}$ mit $x_2^{(k_{2j})} \to x_2 \ (j \to \infty)$, u.s.w. Nach n Auswahlschritten haben wir schließlich eine Teilfolge $(x^{(k_{nj})})_{j\in\mathbb{N}}$ von $(x^{(k)})_{k\in\mathbb{N}}$, für die alle Komponentenfolgen $(x_i^{(k_{nj})})_{j\in\mathbb{N}}$, $i = 1, \dots, n$, konvergieren. Mit den Limiten $x_i \in \mathbb{K}$ setzen wir $x := (x_1, \dots, x_n) \in \mathbb{K}^n$ und erhalten die Konvergenz $x^{k_{nj}} \to x \ (j \to \infty)$. Q.E.D.

Der folgende wichtige Satz zeigt, dass auf dem $\mathbb{K}^n$ alle durch irgendwelche Normen definierten Konvergenzbegriffe äquivalent sind zur Konvergenz bezüglich der Maximumnorm, d. h. zur komponentenweisen Konvergenz.

Satz 1.2 (Äquivalenz von Normen): *Auf dem endlich dimensionalen Vektorraum $\mathbb{K}^n$ sind alle Normen äquivalent zur Maximumnorm, d. h.: Zu jeder Norm $\|\cdot\|$ gibt es positive Konstanten m, M, mit denen gilt:*

$$m \, \|x\|_\infty \le \|x\| \le M \, \|x\|_\infty \,, \quad x \in \mathbb{K}^n \,. \tag{1.1.2}$$

Beweis: Sei $\|\cdot\|$ irgendeine Norm. Für jeden Vektor $x = \sum_{k=1}^n x_k e^{(k)} \in \mathbb{K}^n$ gilt zunächst

$$\|x\| \le \sum_{k=1}^n |x_k| \, \|e^{(k)}\| \le M \, \|x\|_\infty \,, \quad M := \sum_{k=1}^n \|e^{(k)}\| \,.$$

Wir setzen:

$$S_1 := \{x \in \mathbb{K}^n : \ \|x\|_\infty = 1\}, \qquad m := \inf\{\|x\|, \ x \in S_1\} \ge 0.$$

Wir wollen zeigen, dass $m > 0$ ist, denn dann ergibt sich für $x \neq 0$ wegen $\|x\|_\infty^{-1} x \in S_1$ auch $m \leq \|x\|_\infty^{-1}\|x\|$ und folglich

$$0 < m\|x\|_\infty \leq \|x\|, \quad x \in \mathbb{K}^n.$$

Sei also angenommen, dass $m = 0$. Dann gibt es in S_1 eine Folge $(x^{(k)})_{k\in\mathbb{N}}$ mit $\|x^{(k)}\| \to 0 \ (k \to \infty)$. Da die Folge bzgl. der Maximumnorm beschränkt ist, gibt es nach dem Satz von Bolzano-Weierstraß eine Teilfolge, ebenfalls mit $x^{(k)}$ bezeichnet, welche bzgl. der Maximumnorm gegen ein $x \in \mathbb{K}^n$ konvergiert. Wegen

$$\big|1 - \|x\|_\infty\big| = \big|\|x^{(k)}\|_\infty - \|x\|_\infty\big| \leq \|x^{(k)} - x\|_\infty \to 0 \quad (k \to \infty)$$

ist auch $x \in S_1$. Andererseits gilt für alle $k \in \mathbb{N}$:

$$\|x\| \leq \|x - x^{(k)}\| + \|x^{(k)}\| \leq M\|x - x^{(k)}\|_\infty + \|x^{(k)}\|.$$

Für $k \to \infty$ folgt hieraus $\|x\| = 0$ und somit $x = 0$, im Widerspruch zu $x \in S_1$. Q.E.D.

Bemerkung 1.4: Für die beiden vorausgegangenen Sätze, den „Bolzano-Weierstraß“ sowie die „Normäquivalenz“, ist die *endliche* Dimension des $\mathbb{K}^n$ entscheidend. Beide Sätze gelten nicht in *unendlich* dimensionalen Räumen, wie z. B. dem $C[a,b]$, dem $R[a,b]$ und dem Folgenraum l_2.

Bemerkung 1.5: In vielen Anwendungen kommen Mengen von Paaren $\{x,y\}$ von Punkten $x, y \in \mathbb{K}^n$ vor. Diese liegen im sog. „Produktraum“ $V = \mathbb{K}^n \times \mathbb{K}^n$, den man mit der natürlichen Norm

$$\|\{x,y\}\| := \big(\|x\|^2 + \|y\|^2\big)^{1/2}.$$

versehen kann. Da dieser Raum mit dem $2n$-dimensionalen euklidischen Raum $\mathbb{K}^{2n}$ identifiziert werden kann, übertragen sich sinngemäß die bisherigen Aussagen für Mengen im $\mathbb{K}^n$ auch auf Mengen im $\mathbb{K}^n \times \mathbb{K}^n$. Dies lässt sich auf allgemeinere Konstruktionen von (endlich dimensionalen) Produkträumen übertragen, z. B.: $V = \mathbb{K}^{n_1} \times \cdots \times \mathbb{K}^{n_m}$.

1.2 Teilmengen des $\mathbb{K}^n$

Da im $\mathbb{K}^n$ alle Normen äquivalent sind, können wir für die folgenden Betrachtungen irgend eine Norm verwenden, die wir mit $\|\cdot\|$ bezeichnen.

Definition 1.4: *i) Eine Teilmenge $M \subset \mathbb{K}^n$ heißt „beschränkt“, wenn sie in einer Kugelumgebung $K_R(0)$ enthalten ist.*

ii) Eine Teilmenge $U \subset \mathbb{K}^n$ heißt „Umgebung“ des Punktes $a \in \mathbb{K}^n$, wenn sie eine Kugelumgebung $K_\varepsilon(a)$ von a enthält; letztere wird auch „ε-Umgebung“ von a genannt.

iii) Eine Menge $O \in \mathbb{K}^n$ heißt „offen“, wenn es zu jedem $a \in O$ eine Kugelumgebung $K_\varepsilon(a)$ gibt, die in O enthalten ist.

iv) Eine Menge $A \subset \mathbb{K}^n$ heißt „abgeschlossen“ , wenn ihr Komplement $A^c = \mathbb{K}^n \setminus A$ offen ist.

Beispiel 1.2: i) Die Kugel $K_r(a) = \{x \in \mathbb{K}^n \mid \|x-a\| < r\}$ ist Umgebung jedes ihrer Punkte $x \in K_r(a)$. Die Kugel $K_\rho(x)$ mit dem Radius $\rho := r - \|x-a\| > 0$ ist wegen

$$\|y-a\| \le \|y-x\| + \|x-a\| < r - \|x-a\| + \|x-a\| = r, \quad y \in K_\rho(x),$$

in $K_r(a)$ enthalten. Insbesondere ist also die Kugel $K_r(a)$ offen, d. h.: Kugelumgebungen sind stets *offene* Umgebungen.

ii) Die sog. *abgeschlossene* Kugel $\overline{K}_r(a) := \{x \in \mathbb{K}^n : \|x-a\| \le r\}$ ist im obigen Sinne abgeschlossen, denn für jeden Punkt $x \notin \overline{K}_r(a)$ liegt auch die Kugel $K_\varepsilon(x)$ mit Radius $\varepsilon = \|x-a\| - r$ wegen, $y \in K_\varepsilon(x)$,

$$\|y-a\| = \|(x-a)-(x-y)\| \ge \|x-a\| - \|x-y\| > \|x-a\| + r - \|x-a\| = r$$

außerhalb von $\overline{K}_r(a)$.

iii) Der $\mathbb{K}^n$ und die leere Menge $\emptyset$ sind zugleich offen und abgeschlossen.

iv) Die diskrete Menge $\{1/n,\, n \in \mathbb{N}\} \subset \mathbb{R}$ ist weder offen noch abgeschlossen.

v) Auf dem Unterraum $V_{n-1} := \{x \in \mathbb{K}^n : x = (x_1, \dots, x_{n-1}, 0)\}$ des $\mathbb{K}^n$ wird durch die Norm $\|\cdot\|$ von $\mathbb{K}^n$ eine eigene Norm induziert. Damit sind dann auch für Mengen in V_{n-1} die Begriffe „offen“ und „abgeschlossen“ (relativ zu V_{n-1}) definiert. Eine in V_{n-1} offene, nichtleere Menge ist dann aber bezogen auf den Oberraum $\mathbb{K}^n$ nicht offen.

Lemma 1.1: *Es gelten die folgenden Aussagen:*

i) Jede Obermenge einer Umgebung von a ist eine Umgebung von a.

ii) Der Durchschnitt zweier Umgebungen eines Punktes $a \in \mathbb{K}^n$ ist ebenfalls eine Umgebung von a.

iii) Zu je zwei verschiedenen Punkten $a, b \in \mathbb{K}^n$ existieren disjunkte Umgebungen (sog. Hausdorffsche[2] Trennungseigenschaft).

Beweis: i) Sei U Umgebung von a und $K_r(a) \subset U$ eine zugehörige Kugelumgebung. Dann ist diese auch in jeder Obermenge von U enthalten, so dass letztere auch Umgebung von a ist.

ii) Seien U_1 und U_2 Umgebungen von $a \in \mathbb{R}^n$. Es gibt also Kugelumgebungen $K_{r_1}(a) \subset U_1$ und $K_{r_2}(a) \subset U_2$. Die Kugelumgebung $K_r(a)$ mit $r := \min(r_1, r_2)$ ist dann im Schnitt $U_1 \cap U_2$ enthalten, d. h. dieser ist ebenfalls eine Umgebung von a.

iii) Für die Kugelumgebungen $K_r(a)$ und $K_r(b)$ mit Radius $r = \frac{1}{3}\|a-b\|$ gilt

$$\begin{aligned} x \in K_r(a) : \quad \|x-b\| &= \|x-a+a-b\| \ge \|a-b\| - \|x-a\| \ge \tfrac{2}{3}\|a-b\| = 2r, \\ x \in K_r(b) : \quad \|x-a\| &= \|x-b+b-a\| \ge \|a-b\| - \|x-b\| \ge \tfrac{2}{3}\|a-b\| = 2r, \end{aligned}$$

d. h. $K_r(a) \cap K_r(b) = \emptyset$. Q.E.D.

[2]Felix Hausdorff (1868–1942): Deutscher Mathematiker; Prof. in Bonn 1910–1932, Zwangspensionierung durch das nationalsozialistische Regime; grundlegende Beiträge zur Topologie und Mengenlehre.

Lemma 1.2: *Es gelten die folgenden Aussagen:*

i) Der Durchschnitt endlich vieler und die Vereinigung beliebig vieler offener Mengen ist offen.

ii) Die Vereinigung endlich vieler und der Durchschnitt beliebig vieler abgeschlossener Mengen ist abgeschlossen.

Diese Aussagen lassen sich nicht verschärfen, d. h.: Der Durchschnitt unendlich vieler offener Mengen ist nicht notwendig offen, und die Vereinigung unendlich vieler abgeschlossener Mengen ist nicht notwendig abgeschlossen.

Beweis: Übungsaufgabe. Q.E.D.

Lemma 1.3: *Eine Menge $A \subset \mathbb{K}^n$ ist genau dann abgeschlossen, wenn der Limes jeder konvergenten Folge von Punkten in A ebenfalls in A liegt.*

Beweis: i) Sei A abgeschlossen. Läge der Grenzwert a einer konvergenten Folge $(a^{(k)})_{k\in\mathbb{N}}$ mit $a^{(k)} \in A$ nicht in A, d. h.: a liegt in der offenen Menge $O := \mathbb{K}^n \setminus A$, so enthielte diese offene Menge als Umgebung von a fast alle $a^{(k)}$, im Widerspruch zur Voraussetzung $a^{(k)} \in A$.

ii) Sei nun der Limes jeder konvergenten Folge aus A ebenfalls in A. Angenommen A ist nicht abgeschlossen, d. h. $O := \mathbb{K}^n \setminus A$ ist nicht offen. Dann gibt es einen Punkt $a \in O$ derart, dass keine Kugelumgebung $K_\varepsilon(a)$ ganz in O liegt. Insbesondere enthält jede Kugel $K_{1/k}(a)$, $k \in \mathbb{N}$, einen Punkt $a^{(k)}$ mit $a^{(k)} \notin O$. Die Folge $(a^{(k)})_{k\in\mathbb{N}}$ liegt also in A und konvergiert wegen $\|a^{(k)} - a\| < 1/k$ gegen den Limes a, der aber nicht zu A gehört, im Widerspruch zur Voraussetzung. Q.E.D.

Definition 1.5: *i) Ein Punkt $a \in \mathbb{K}^n$ heißt „Randpunkt“ einer Menge $M \subset \mathbb{K}^n$, wenn jede Umgebung von a Punkte sowohl aus M als auch aus dem Komplement $M^c := \mathbb{K}^n \setminus M$ enthält. Die Menge aller Randpunkte von M, der sog. „Rand“, wird mit ∂M bezeichnet. Aus Symmetriegründen gilt $\partial(M^c) = \partial M$. Jeder Randpunkt von M ist also sowohl Limes von Punktfolgen aus M als auch Limes von Punktfolgen aus M^c.*

ii) Für eine Menge $M \subset \mathbb{K}^n$ ist $M^o := M \setminus \partial M$ der sog. „offene Kern“ (oder auch das „Innere“) von M.

iii) Für eine Menge $M \subset \mathbb{K}^n$ ist $\overline{M} := M \cup \partial M$ die sog. „abgeschlossene Hülle“ (oder auch der „Abschluss“) von M.

iv) Für eine nichtleere, beschränkte Menge $M \subset \mathbb{K}^n$ ist der „Durchmesser“ $\operatorname{diam}(M)$ (bzgl. der Norm $\|\cdot\|$) definiert durch

$$\operatorname{diam}(M) := \sup\{\|x - y\|,\ x, y \in M\}.$$

Beispiel 1.3: *Der Rand der Kugel $K_r(a) \subset \mathbb{R}^n$ ist die „Sphäre“*

$$S = \{x \in \mathbb{R}^n :\ \|x - a\| = r\}.$$

Der Rand von $\mathbb{Q}^n$ in $\mathbb{R}^n$ ist der ganze $\mathbb{R}^n$. Der Rand von $\mathbb{R}^n$ ist leer.

Lemma 1.4: *Für jede Menge $M \subset \mathbb{K}^n$ gilt:*
i) Der Rand ∂M ist abgeschlossen.

ii) Die Menge $M^o = M \setminus \partial M$ ist offen. Jede offene Teilmenge $O \subset M$ ist in $M \setminus \partial M$ enthalten.

iii) Die Menge $\overline{M} = M \cup \partial M$ ist abgeschlossen. Jede abgeschlossene Menge A mit $M \subset A$ enthält $M \cup \partial M$.

Beweis: i) Wir zeigen, dass das Komplement ∂M^c offen ist. Sei $a \notin \partial M$. Dann gibt es nach Definition von ∂M eine Umgebung $K_r(a)$, welche entweder keine Punkte von M oder keine Punkte von M^c enthält. Dann ist aber auch $K_r(a) \cap \partial M = \emptyset$ und $(\partial M)^c$ ist somit offen.

ii) Der Beweis wird als Übungsaufgabe gestellt.

iii) Wir zeigen, dass $(M \cup \partial M)^c$ offen ist. Zu jedem Punkt $a \notin (M \cup \partial M)^c$ existiert eine Umgebung $K_r(a)$, welche keine Punkte von M enthält. Kein Punkt von $K_r(a)$ kann dann zu ∂M gehören, d. h. $K_r(a) \subset (M \cup \partial M)^c$. Also ist $(M \cup \partial M)^c$ offen. Sei A abgeschlossen mit $M \subset A$. Jeder Punkt $a \in \partial M$ ist Limes einer Folge von Punkten $a_n \in M$. Wegen $a_n \in M \subset A$ ist dann der Limes auch in A, d. h.: $\partial M \subset A$. Q.E.D.

Korollar 1.1: *Eine Menge $O \subset \mathbb{K}^n$ ist genau dann offen, wenn sie keinen ihrer Randpunkte enthält. Eine Menge $A \subset \mathbb{K}^n$ ist genau dann abgeschlossen, wenn sie alle ihre Randpunkte enthält.*

Beweis: Die Richtigkeit der Behauptungen ergibt sich unmittelbar mit Hilfe der Aussagen von Lemma 1.4. Die Beweisdetails werden als Übungsaufgabe gestellt. Q.E.D.

Definition 1.6: *Ein Punkt $x \in \mathbb{K}^n$ heißt „Häufungspunkt" einer Menge $M \subset \mathbb{K}^n$ wenn jede Umgebung von x mindestens einen Punkt aus $M \setminus \{x\}$ enthält. Die Menge der Häufungspunkte von M wird mit $\mathcal{H}(M)$ bezeichnet. Ein Punkt $x \in M \setminus \mathcal{H}(M)$ wird „isoliert" genannt.*

Satz 1.3: *i) Für jede Menge $M \subset \mathbb{K}^n$ gilt*

$$M \cup \mathcal{H}(M) = \overline{M}. \tag{1.2.3}$$

ii) Eine Menge $M \subset \mathbb{K}^n$ ist genau dann abgeschlossen, wenn sie alle ihre Häufungspunkte enthält.

Beweis: i) Sei $x \in \partial M$. Dann liegt in jeder $1/k$-Umgebung von x ein Punkt $x^{(k)} \in M$, d. h.: x ist Limes der Folge $(x^{(k)})_{k\in\mathbb{N}}$. Also ist x Häufungspunkt von M, und es gilt $M \cup \partial M = \overline{M} \subset M \cup \mathcal{H}(M)$. Weiter ist jedes $x \in \mathcal{H}(M)$ Limes einer Folge von Punkten aus M, d. h.: $x \notin \overline{M}^c$. Also ist $M \cup \mathcal{H}(M) \subset \overline{M}$. Dies impliziert die Richtigkeit der ersten Behauptung.

ii) Im Falle $\mathcal{H}(M) \subset M$ ist nach dem eben Gezeigten $M = \overline{M}$, d. h.: M ist abgschlossen. Ist andererseits M abgeschlossen, so ist $M = M \cup \partial M = \overline{M}$, d. h.: $\mathcal{H}(M) \subset M$. Q.E.D.

Definition 1.7 (Kompaktheit): *Eine Menge $M \subset \mathbb{K}^n$ heißt „kompakt" (bzw. „folgen-kompakt"), wenn jede Folge aus M eine konvergente Teilfolge mit Limes in M besitzt.*

Beispiel 1.4: Mit Hilfe des Satzes von Bolzano-Weiserstraß folgt, dass eine abgeschlossene Kugel $\overline{K}_r(a)$ kompakt ist. Ferner sind der Rand ∂M einer beschränkten Menge $M \subset \mathbb{K}^n$ und jede endliche Menge kompakt.

Satz 1.4 (Satz von der Kompaktheit): *Für eine Teilmenge $M \subset \mathbb{K}^n$ sind folgende Aussagen äquivalent:*

i) M ist folgen-kompakt.

ii) M ist beschränkt und abgeschlossen.

iii) Jede offene Überdeckung $\{O_\lambda,\ \lambda \in \Lambda\}$ von M, d. h.: $O_\lambda \subset \mathbb{K}^n$ offen und $M \subset \cup_{\lambda\in\Lambda} O_\lambda$ (Λ eine beliebige Indexmenge), enthält eine endliche Überdeckung, d. h.: $M \subset \cup_{i=1,\dots,m} O_i$. (Überdeckungseigenschaft von Heine[3] und Borel[4]).

Beweis: (i)⇒(ii): Die Menge $M \subset \mathbb{K}^n$ sei folgen-kompakt. Dann ist M auch abgeschlossen, denn jede konvergente Folge in M hat wegen der Kompaktheit eine konvergente Teilfolge mit Limes in M, d. h.: Auch der Limes der gesamten Folge liegt in M. Weiter muss M auch beschränkt sein, denn andernfalls gäbe es eine Folge $(x_n)_{n\in\mathbb{N}}$ in M mit $\|x_n\| \to \infty$, welche dann keine konvergente Teilfolge haben kann.

(ii)⇒(i): Ist $M \subset \mathbb{K}^n$ beschränkt und abgeschlossen, so besitzt sie nach dem Satz von Bolzano-Weiserstraß eine konvergente Teilfolge, deren Limes dann wegen der angenommenen Abgeschlossenheit von M ebenfalls in M liegt. Also ist M folgen-kompakt.

(iii)⇒(i): Die Menge $M \subset \mathbb{K}^n$ besitze die Überdeckungseigenschaft. Wir wollen zeigen, dass sie dann auch folgen-kompakt ist. Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge in M und $A := \{x_n, n \in \mathbb{N}\}$ (Gleiche Folgenelemente werden in der Menge A identifiziert.). Ist A endlich, so hat die Folge notwendig eine konstante (und damit konvergente) Teilfolge. Sei A also unendlich. Angenommen A hat *keine konvergente Teilfolge* mit Limes in M. Dann hat jeder Punkt $a \in M$ eine offene Umgebung $U(a)$, die nur endlich viele Punkte von A enthält. Diese Umgebungen $U(a)$ bilden nun eine offene Überdeckung von M, zu der es nach Annahme eine *endliche* Teilüberdeckung $\{U(a_k), k = 1,\dots,m\}$ gibt. Diese

[3] Eduard Heine (1821–1881): Deutscher Mathematiker; Professor in Halle; einer der wichtigsten Vertreter der „Weierstraßschen Schule" im 19. Jahrhundert; Beiträge zur Theorie der reellen Funktionen, Potentialtheorie und Theorie der Differentialgleichungen.

[4] Félix Éduard Justin Émile Borel (1871–1956): Französischer Mathematiker, u. a. Professor an der Universität Sorbonne in Paris; wichtige Beiträge zur Maßtheorie und zur Spieltheorie; war auch politisch aktiv (1925-1940 Marineminister) und während des Krieges Mitglied der Résistance.

„Teilüberdeckung“ kann dann auch nur endlich viele Punkte von A enthalten, d. h.: A ist endlich, im Widerspruch zur Annahme.

(ii)⇒(iii): Die Menge M sei beschränkt und abgeschlossen. Sei $\{O_\lambda,\, \lambda \in \Lambda\}$ eine (offene) Überdeckung von M, die keine endliche Überdeckung von M enthält. Als beschränkte Menge ist M in einem abgeschlossenen Würfel Q_0 mit Kantenlänge L enthalten. Wir zerlegen Q in 2^n Würfel der halben Kantenlänge. Dann gilt auch für mindestens einen dieser Teilwürfel Q_1, dass $M \cap Q_1$ nicht von endlich vielen der O_λ überdeckt wird. Durch rekursive Wiederholung dieses Verfahrens finden wir eine Folge abgeschlossener Würfel Q_k mit Kantenlänge $L_k = 2^{-k}L$, so dass $\ldots \subset Q_k \subset Q_{k-1} \subset \ldots \subset Q_0$, und keine der Mengen $M \cap Q_k$ wird durch endlich viele der O_λ überdeckt. Wir wählen nun in jeder der Mengen $M \cap Q_k$ einen Punkt x_k. Nach Konstruktion der Würfelfolge ist $(x_k)_{k\in\mathbb{N}}$ eine Cauchy-Folge und somit konvergent gegen einen Punkt $x \in M$. Dieser Limes liegt dann aber auch in einer der offenen Überdeckungsmengen O_λ. Diese muss dann auch fast alle der Würfel Q_k und damit insbesondere fast alle der Durchschnitte $M \cap Q_k$ enthalten. Dies ist ein Widerspruch zur Annahme, dass keiner dieser Durchschnitte von endlich vielen der O_λ überdeckt wird. Q.E.D.

Bemerkung 1.6: Die Charakterisierung kompakter Mengen durch die Überdeckungseigenschaft ist die Aussage des „Satzes von Heine-Borel“. Die entscheidende Voraussetzung für die Gültigkeit der Sätze von Bolzano-Weierstraß und Heine-Borel ist die *endliche* Dimension von $\mathbb{K}^n$. In *unendlich* dimensionalen Banach-Räumen, wie z. B. dem Raum $C[a,b]$ der stetigen Funktionen, ist dies nicht möglich. Hier werden zum Nachweis der Kompaktheit von beschränkten, abgeschlossenen Mengen noch zusätzliche Eigenschaften benötigt, wie z. B. die gleichgradige Stetigkeit beim „Auswahlsatz von Arzelà-Ascoli“.

Korollar 1.2: *Jede abgeschlossene Teilmenge einer kompakten Menge in $\mathbb{K}^n$ ist ebenfalls kompakt.*

Beweis: Sei $M \subset \mathbb{K}^n$ kompakt und $A \subset M$ abgeschlossen. Nach Satz 1.4 ist M, und damit auch A beschränkt. Also ist wieder nach Satz 1.4 A auch kompakt. Q.E.D.

1.3 Geometrie des $\mathbb{K}^n$

Das Betreiben von „Geometrie“ auf dem $\mathbb{K}^n$, oder allgemeiner auf einem beliebigen Vektorraum, bedeutet zunächst das Formulieren der uns aus der Elementargeometrie der Ebene wohl bekannten Begriffsbildungen und Zusammenhänge in einer abstrakteren Sprache, z. B. die Eigenschaft „orthogonal“ für Vektoren und der Begriff des „Winkels“ zwischen Vektoren. Dazu dient das auf dem $\mathbb{K}^n$ definierte sog. „euklidische Skalarprodukt“:

$$(x,y)_2 := \sum_{i=1}^{n} x_i \overline{y}_i.$$

Definition 1.8: *Sei V irgendein Vektorraum über dem Körper $\mathbb{K}$. Eine Abbildung $(\cdot,\cdot): V\times V\to\mathbb{K}$ heißt „Skalarprodukt“, wenn folgende Bedingungen erfüllt sind:*

(S1) Linearität: $(\alpha x_1+\beta x_2,y)=\alpha(x_1,y)+\beta(x_2,y),\quad \alpha,\beta\in\mathbb{K}$;

(S2) Symmetrie: $(x,y)=\overline{(y,x)}$;

(S3) Definitheit: $(x,x)\in\mathbb{R},\quad (x,x)\ge 0,\qquad (x,x)=0\ \Rightarrow\ x=0.$

Bemerkung 1.7: *i) Verzichtet man in der obigen Definition des Skalarprodukts auf die strenge „Definitheit“, d. h. wird nur $(x,x)\in\mathbb{R},\ (x,x)\ge 0$ verlangt, so spricht man von einem „Semi-Skalarprodukt“.*

ii) Aus den Eigenschaften (S1) (Linearität im ersten Argument) und (S2) (Symmetrie) folgt auch die Linearität im zweiten Argument und damit die volle „Bilinearität“ des Skalarprodukts als eine „Sesquilinearfom“ (im Komplexen) bzw. eine „Bilinearform“ (im Reellen).

iii) Die Eigenschaft (S1) (Linearität) kann bei ihrem Nachweis auch in die beiden Bestandteile „Additivität“, $(x_1+x_2,y)=(x_1,y)+(x_2,y)$, und „Homogenität“, $(\alpha x,y)=\alpha(x,y),\ \alpha\in\mathbb{K}$, aufgespalten werden.

Lemma 1.5: *Für ein Skalarprodukt $(\cdot,\cdot)$ auf einem Vektorraum V über $\mathbb{K}$ gilt die „Schwarzsche Ungleichung“*

$$|(x,y)|^2\le (x,x)\,(y,y),\quad x,y\in V. \tag{1.3.4}$$

Beweis: Da die Behauptung für $y=0$ offensichtlich richtig ist, können wir o.B.d.A. $y\ne 0$ annehmen. Für beliebiges $\alpha\in\mathbb{K}$ ist

$$0\le (x+\alpha y,x+\alpha y)=(x,x)+\alpha(y,x)+\overline{\alpha}(x,y)+\alpha\overline{\alpha}(y,y).$$

Mit $\alpha:=-(x,y)(y,y)^{-1}$ impliziert dies

$$\begin{aligned}0&\le (x,x)-(x,y)(y,y)^{-1}(y,x)-\overline{(x,y)}(y,y)^{-1}(x,y)+(x,y)\overline{(x,y)}(y,y)^{-1}\\&=(x,x)-|(x,y)|^2(y,y)^{-1}\end{aligned}$$

bzw. $0\le (x,x)(y,y)-|(x,y)|^2$. Dies zeigt die Richtigkeit der Behauptung. Q.E.D.

Korollar 1.3: *a) Von einem Skalarprodukt $(\cdot,\cdot)$ auf einem Vektorraum V über $\mathbb{K}$ wird durch*

$$\|x\|:=(x,x)^{1/2},\quad x\in V,$$

eine Norm erzeugt. Ist der so entstehende normierte Raum $(V,\|\cdot\|)$ vollständig, so heißt das Paar $(V,(\cdot,\cdot))$ „Hilbert-Raum“.

b) Das euklidische Skalarprodukt $(\cdot,\cdot)_2$ auf dem $\mathbb{K}^n$ erzeugt durch

$$\|x\|_2:=(x,x)_2^{1/2}$$

eine Norm, die sog. „euklidische Norm“. Damit ist dann $(\mathbb{K}^n,(\cdot,\cdot)_2)$ ein Hilbert-Raum.

Beweis: Die Normeigenschaften (N1) (Definitheit) und (N2) (Homogenität) sind offensichtlich gegeben. Es bleibt die Dreiecksungleichung (N3) zu zeigen. Mit Hilfe der Schwarzschen Ungleichung erhalten wir

$$\begin{aligned}\|x+y\|^2 &= (x+y,x+y) = (x,x)+(x,y)+(y,x)+(y,y) \\ &\le \|x\|^2 + 2|(x,y)| + \|y\|^2 \le \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\| + ||y||)^2,\end{aligned}$$

was zu zeigen war. Q.E.D.

Die Schwarzsche Ungleichung für das euklidische Skaarprodukt wird durch die im Folgenden bewiesene „Höldersche[5] Ungleichung" verallgemeinert. Als Vorbereitung stellen wir einen einfachen Spezialfall einer ganzen Klasse von Ungleichungen, den sog. „Youngschen[6] Ungleichungen" bereit.

Lemma 1.6 (Youngsche Ungleichung): *Für $p,q\in\mathbb{R}$ mit $1<p,q<\infty$ und $1/p+1/q=1$ gilt die sog. „Youngsche Ungleichung"*

$$|xy| \le \frac{|x|^p}{p} + \frac{|y|^q}{q}, \quad x,y\in\mathbb{K}. \tag{1.3.5}$$

Beweis: Da die Logarithmus-Funktion $\ln(x)$ auf $\mathbb{R}_+$ wegen $\ln''(x) = -1/x^2 < 0$ konkav ist, gilt für $x,y\in\mathbb{K}$:

$$\ln\left(\tfrac{1}{p}|x|^p + \tfrac{1}{q}|y|^q\right) \ge \tfrac{1}{p}\ln(|x|^p) + \tfrac{1}{q}\ln(|y|^q) = \ln(|x|) + \ln(|y|).$$

Wegen der Monotonie der Exponentialfunktion e^x folgt weiter für $x,y\in\mathbb{K}$:

$$\tfrac{1}{p}|x|^p + \tfrac{1}{q}|y|^q \ge \exp\big(\ln(|x|)+\ln(|y|)\big) = \exp\big(\ln(|x|)\big)\exp\big(\ln(|y|)\big) = |x||y| = |xy|,$$

was zu beweisen war. Q.E.D.

Lemma 1.7 (Höldersche Ungleichung): *Für das euklidische Skalarprodukt gilt für beliebige $p,q\in\mathbb{R}$ mit $1<p,q<\infty$ und $1/p+1/q=1$ die sog. „Höldersche Ungleichung"*

$$|(x,y)_2| \le \|x\|_p\,\|y\|_q, \quad x,y\in\mathbb{K}^n. \tag{1.3.6}$$

Diese Ungleichung gilt auch im Grenzfall $p=1,\ q=\infty$.

[5] Ludwig Otto Hölder (1859–1937): Deutscher Mathematiker; Prof. in Tübingen; Beiträge zunächst zur Theorie der Fourier-Reihen und später vor allem zur Gruppentheorie; fand 1884 die nach ihm benannte Ungleichung.

[6] William Henry Young (1863–1942): Englischer Mathematiker; lehrte an verschiedenen Universitäten weltweit, u.a. in Calcutta, Liverpool und Wales; Beiträge zur Differential- und Intergralrechnung, Topologischen Mengentheorie und Geometrie.

Beweis: Für $x = 0$ oder $y = 0$ ist die behauptete Ungleichung trivialerweise richtig. Sei also o.B.d.A. $\|x\|_p \neq 0$ und $\|y\|_q \neq 0$. Zunächst gilt

$$\frac{|(x,y)_2|}{\|x\|_p\|y\|_q} = \frac{1}{\|x\|_p\|y\|_q}\Big|\sum_{i=1}^n x_i\bar{y}_i\Big| \leq \sum_{i=1}^n \frac{|x_i||y_i|}{\|x\|_p\|y\|_q}.$$

Mit Hilfe der Youngschen Ungleichung folgt weiter

$$\begin{aligned}\frac{|(x,y)_2|}{\|x\|_p\|y\|_q} &\leq \sum_{i=1}^n \Big\{\frac{|x_i|^p}{p\|x\|_p^p} + \frac{|y_i|^q}{q\|y\|_q^q}\Big\} \\ &= \frac{1}{p\|x\|_p^p}\sum_{i=1}^n |x_i|^p + \frac{1}{q\|y\|_q^q}\sum_{i=1}^n |y_i|^q = \frac{1}{p} + \frac{1}{q} = 1.\end{aligned}$$

Dies impliziert die Behauptung. Q.E.D.

Als Folgerung aus der Hölderschen Ungleichung gewinnen wir die sog. „Minkowskische[7] Ungleichung", welche gerade die Dreiecksungleichung für die l_p-Norm ist.

Lemma 1.8 (Minkowskische Ungleichung): *Für beliebiges $p \in \mathbb{R}$ mit $1 \leq p < \infty$ sowie für $p = \infty$ gilt die sog. „Minkowskische Ungleichung"*

$$\|x+y\|_p \leq \|x\|_p + \|y\|_p, \quad x, y \in \mathbb{K}^n. \tag{1.3.7}$$

Beweis: Für $p = 1$ und $p = \infty$ ergibt sich die behauptete Abschätzung unmittelbar aus der Dreiecksungleichung für reelle Zahlen:

$$\begin{aligned}\|x+y\|_1 &= \sum_{i=1}^n |x_i + y_i| \leq \sum_{i=1}^n |x_i| + \sum_{i=1}^n |y_i| = \|x\|_1 + \|y\|_1, \\ \|x+y\|_\infty &= \max_{1\leq i\leq n} |x_i + y_i| \leq \max_{1\leq i\leq n} |x_i| + \max_{1\leq i\leq n} |y_i| = \|x\|_\infty + \|y\|_\infty.\end{aligned}$$

Sei nun $1 < p < \infty$ und q definiert durch $1/p + 1/q = 1$, d. h. $q = p/(p-1)$. Wir setzen

$$\xi_i := |x_i + y_i|^{p-1}, \quad i = 1, \dots, n, \quad \xi := (\xi_i)_{i=1}^n.$$

Damit gilt zunächst

$$\|x+y\|_p^p = \sum_{i=1}^n |x_i + y_i||x_i + y_i|^{p-1} \leq \sum_{i=1}^n |x_i|\xi_i + \sum_{i=1}^n |y_i|\xi_i$$

und weiter mit Hilfe der Hölderschen Ungleichung

$$\|x+y\|_p^p \leq \|x\|_p\|\xi\|_q + \|y\|_p\|\xi\|_q = (\|x\|_p + \|y\|_p)\|\xi\|_q.$$

[7] Hermann Minkowski (1864–1909): Russisch-deutscher Mathematiker; Prof. in Göttingen; verschiedene Beiträge zur reinen Mathematik; „erfand" das nicht-euklidische, 4-dimensionale Raum-Zeit-Kontinuum (Minkowski-Raum) zur Beschreibung der Einsteinschen Relativitätstheorie.

Bei Beachtung von $q = p/(p-1)$ folgt

$$\|\xi\|_q^q = \sum_{i=1}^n |\xi_i|^q = \sum_{i=1}^n |x_i + y_i|^p = \|x+y\|_p^p.$$

und damit

$$\|x+y\|_p^p \leq (\|x\|_p + \|y\|_p)\|x+y\|_p^{p/q} = (\|x\|_p + \|y\|_p)\|x+y\|_p^{p-1}.$$

Dies impliziert offenbar die Behauptung. Q.E.D.

Mit Hilfe des euklidischen Skalarprodukts lässt sich der geometrische Begriff „orthogonal“ definieren. Zwei Vektoren $x, y \in \mathbb{K}^n$ heißen „orthogonal“ („$x \perp y$“), wenn

$$(x,y)_2 = 0.$$

Für orthogonale Vektoren gilt der „Satz des Pythagoras“ (Übungsaufgabe):

$$\|x+y\|_2^2 = \|x\|_2^2 + \|y\|_2^2, \quad x, y \in \mathbb{K}^n,\ x \perp y. \tag{1.3.8}$$

Ein Satz von Vektoren $\{a^{(1)}, \dots, a^{(m)}\}$, $a^{(i)} \neq 0$, des $\mathbb{K}^n$, welche paarweise orthogonal sind, d. h. $(a^{(k)}, a^{(l)}) = 0$ für $k \neq l$, ist linear unabhängig. Denn für $\sum_{k=1}^m c_k a^{(k)} = 0$ folgt durch Skalarproduktbildung mit $a^{(l)}$, $l = 1, \dots, m$:

$$0 = \sum_{k=1}^m c_k (a^{(k)}, a^{(l)}) = c_l (a^{(l)}, a^{(l)}),$$

und folglich $c_l = 0$.

Definition 1.9: *Ein Satz von Vektoren $\{a^{(1)}, \dots, a^{(m)}\}$, $a^{(i)} \neq 0$, des $\mathbb{K}^n$ welche paarweise orthogonal sind, $(a^{(k)}, a^{(l)}) = 0$, $k \neq l$, heißt „Orthogonalsystem“ bzw. im Fall $m = n$ „Orthogonalbasis“. Gilt $(a^{(k)}, a^{(k)}) = 1$, so spricht man von einem „Orthonormalsystem“ bzw. einer „Orthonormalbasis“.*

Beispiel 1.5: Die euklidische Basis $\{e^{(1)}, \dots, e^{(n)}\}$ ist offenbar eine Orthonormalbasis des $\mathbb{R}^n$. Es gibt aber noch andere, die man etwa durch Drehung und Spiegelung erhält.

Lemma 1.9: *Sei $\{a^{(i)}, i = 1, \dots, n\}$ eine Orthonormalbasis des $\mathbb{K}^n$. Dann besitzt jeder Vektor $x \in \mathbb{K}^n$ eine Darstellung der Form (in Analogie zur „Fourier-Entwicklung“)*

$$x = \sum_{i=1}^n (x, a^{(i)})_2\, a^{(i)}, \quad x \in \mathbb{K}^n, \tag{1.3.9}$$

und es gilt die „Vollständigkeitsrelation“ (auch „Parsevalsche[8] Gleichung“ genannt)

$$\|x\|_2^2 = \sum_{i=1}^n |(x, a^{(i)})_2|^2. \tag{1.3.10}$$

[8]Marc-Antoine Parseval des Chênes (1755–1836): Französischer Mathematiker; Arbeiten über partielle Differentialgleichungen der Physik (nur fünf mathematische Publikationen); bekannt durch die nach ihm benannte Gleichung, die er aber ohne Beweis und Bezug zu Fourier-Reihen angegeben hat.

Beweis: Aus der Darstellung $x = \sum_{j=1}^n \alpha_j a^{(j)}$ folgt durch Produktbildung mit $a^{(i)}$:

$$(x, a^{(i)})_2 = \sum_{j=1}^n \alpha_j (a^{(j)}, a^{(i)})_2 = \alpha_i, \quad i = 1, \dots, n,$$

und somit die Darstellung (1.3.9). Ferner gilt:

$$\|x\|_2^2 = (x, x)_2 = \sum_{i,j=1}^n (x, a^{(i)})_2 \overline{(x, a^{(j)})_2}\, (a^{(i)}, a^{(j)})_2 = \sum_{i=1}^n |(x, a^{(i)})_2|^2,$$

was zu beweisen war. Q.E.D.

Bemerkung 1.8: *Die Aussage von Lemma 1.9 gilt sinngemäß auch in unendlich dimensionalen Skalarprodukträumen mit „vollständigen“ Orthonormalsystemen (Verallgemeinerung des Basisbegriffs); ein Beispiel ist der in der Fourier-Analyse verwendete Raum* $R[0, 2\pi]$ *mit den normierten trigonometrischen Funktionen als vollständigem Orthogonormalsystem.*

Der folgende Gram-Schmidt-Algorithmus erlaubt es, aus einer beliebigen Basis des $\mathbb{K}^n$ eine Orthonormalbasis zu konstruieren. Die gilt auch in beliebigen Vektorräumen mit Skalarprodukt.

Satz 1.5 (Gram-Schmidt-Verfahren): *Sei* $\{a^{(1)}, \dots, a^{(n)}\}$ *eine Basis des* $\mathbb{K}^n$. *Dann erhält man durch das sog. „Gram[9]-Schmidtsche[10] Orthogonalisierungsverfahren“,*

$$\begin{aligned} &b^{(1)} := \|a^{(1)}\|_2^{-1} a^{(1)}, \\ &\tilde{b}^{(k)} := a^{(k)} - \sum_{j=1}^{k-1} (a^{(k)}, b^{(j)})_2 b^{(j)}, \quad b^{(k)} := \|\tilde{b}^{(k)}\|_2^{-1} \tilde{b}^{(k)}, \quad k = 2, \dots, n, \end{aligned} \tag{1.3.11}$$

eine Orthonormalbasis $\{b^{(1)}, \dots, b^{(n)}\}$.

Beweis: Wir zeigen zunächst, dass der Konstruktionsprozeß für die $b^{(k)}$ nicht mit $k < n$ abbrechen kann. Die Vektoren $b^{(k)}$ sind gemäß Konstruktion Linearkombinationen der $a^{(1)}, \dots, a^{(k)}$. Wäre nun für ein $k \le n$

$$a^{(k)} - \sum_{j=1}^{k-1} (a^{(k)}, b^{(j)})_2 b^{(j)} = 0,$$

[9] Jørgen Pedersen Gram (1850–1916): Dänischer Mathematiker, Mitarbeiter und später Eigentümer einer Versicherungsgesellschaft, Beiträge zur Algebra (Invariantentheorie), Wahrscheinlichkeitstheorie, Numerik und Forstwissenschaft; das u. a. nach ihm benannte Orthogonalisierungsverfahren geht aber wohl auf Laplace zurück und wurde bereits von Cauchy 1836 verwendet.

[10] Erhard Schmidt (1876-1959): Deutscher Mathematiker, Prof. in Berlin, Gründer des dortigen Instituts für Angewandte Mathematik 1920, nach dem Krieg Direktor des Mathematischen Instituts der Akademie der Wissenschaften der DDR; Beiträge zur Theorie der Integralgleichungen und der Hilbert-Räume sowie später zur Topologie.

so müssten die Vektoren $\{a^{(1)},\dots,a^{(k)}\}$ linear abhängig sein, im Widerspruch zur Annahme, dass $\{a^{(1)},\dots,a^{(n)}\}$ eine Basis ist. Wir zeigen nun durch Induktion, dass der Gram-Schmidt-Prozess tatsächlich eine Orthonormalbasis erzeugt. Offenbar ist $\|b^{(1)}\|_2 = 1$. Sei nun $\{b^{(1)},\dots,b^{(k)}\}$ für $k \le n$ bereits als Orthonormalsystem nachgewiesen. Dann gilt für $l = 1,\dots,k$:

$$(b^{(k+1)}, b^{(l)})_2 = (a^{(k+1)}, b^{(l)})_2 - \sum_{j=1}^{k} (a^{(k+1)}, b^{(j)})_2 \underbrace{(b^{(j)}, b^{(l)})_2}_{=\,\delta_{jl}} = 0$$

und $\|b^{(k+1)}\|_2 = 1$, d. h.: $\{b^{(1)},\dots,b^{(k+1)}\}$ ist ebenfalls ein Orthonormalsystem. Q.E.D.

Für einen Punkt $x \in \mathbb{K}^n$ ist die sog. „orthogonale Projektion“ $P_W x \in W$ (geometrisch der „Lotfußpunkt“) auf einen Unterraum $W \subset \mathbb{K}^n$ anschaulich charakterisiert durch die Eigenschaft

$$\|x - P_W x\|_2 = \min_{y \in W} \|x - y\|_2. \tag{1.3.12}$$

Diese „Bestapproximationseigenschaft“ ist äquivalent zu den Beziehungen

$$(P_W x, y)_2 = (x, y)_2 \quad \forall y \in W, \tag{1.3.13}$$

aus denen sich $P_W x$ berechnen ließe.

1.4 Lineare Abbildungen auf dem $\mathbb{K}^n$

Wir betrachten nun Abbildungen des n-dimensionalen Raumes $\mathbb{K}^n$ in den m-dimensionalen Raum $\mathbb{K}^m$, wobei nicht notwendig $m = n$ sein muss. Der Spezialfall $m = n$ spielt aber eine wichtige Rolle. Eine Abbildung $\varphi : \mathbb{K}^n \to \mathbb{K}^m$ heißt „linear“, wenn für $x, y \in \mathbb{K}^n$ und $\alpha, \beta \in \mathbb{K}$ gilt:

$$\varphi(\alpha x + \beta y) = \alpha\varphi(x) + \beta\varphi(y). \tag{1.4.14}$$

Die Wirkung einer linearen Abbildung auf Vektoren $x \in \mathbb{K}^n$ lässt sich auf unterschiedliche Weise beschreiben. Es genügt offenbar, die Wirkung auf die Elemente einer Basis, z. B. einer kartesischen Basis $\{e^{(i)}, i = 1,\dots,n\}$, anzugeben:

$$x = \sum_{i=1}^{n} x_i e^{(i)} \quad \to \quad \varphi(x) = \varphi\Big(\sum_{i=1}^{n} x_i e^{(i)}\Big) = \sum_{i=1}^{n} x_i \varphi(e^{(i)}).$$

Dabei wird dem Punkt $x \in \mathbb{K}^n$ (eineindeutig) der „Koordinatenvektor“ $\hat{x} = (x_i)_{i=1}^n$ zugeordnet. Stellt man auch die Bilder $\varphi(x)$ bzgl. einer kartesischen Basis des $\mathbb{K}^m$ dar,

$$\varphi(x) = \sum_{j=1}^{m} \varphi_j(x) e^{(j)} = \sum_{j=1}^{m} \Big(\sum_{i=1}^{n} \underbrace{\varphi_j(e^{(i)})}_{=:\, a_{ji}} x_i\Big) e^{(j)},$$

mit dem Koordinatenvektor $\hat{\varphi}(x) = (\varphi_j(x))_{j=1}^m$, so erhält man das in Matrixform angeordnete Zahlenschema

$$\begin{pmatrix} \varphi_1(e^{(1)}) & \cdots & \varphi_1(e^{(n)}) \\ \vdots & \ddots & \vdots \\ \varphi_m(e^{(1)}) & \cdots & \varphi_m(e^{(n)}) \end{pmatrix} =: \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} = A \in \mathbb{K}^{m\times n}.$$

Damit gilt dann nach den üblichen Regeln der Matrix-Vektor-Multiplikation:

$$\varphi_j(x) = (A\hat{x})_j := \sum_{i=1}^n a_{ji}x_i, \quad j = 1,\dots,m.$$

Die lineare Abbildung $\varphi : \mathbb{K}^n \to \mathbb{K}^m$ lässt sich also bzgl. der fest gewählten kartesischen Basen von $\mathbb{K}^n$ und $\mathbb{K}^m$ (eindeutig) durch die Matrix $A \in \mathbb{K}^{m\times n}$ beschreiben:

$$\hat{\varphi}(x) = A\hat{x}, \quad x \in \mathbb{K}^n. \tag{1.4.15}$$

Im Folgenden werden wir zur Vereinfachung der Schreibweise meist den Punkt x mit seiner speziellen kartesischen Koordinatendarstellung $\hat{x}$ identifizieren.

Wir folgen hier der Konvention, dass bei der Bezeichnung $\mathbb{K}^{m\times n}$ der erste Parameter m zum Bildraum $\mathbb{K}^m$, d. h. zur Anzahl der Zeilen der Matrix, und der zweite n zum Urbildraum $\mathbb{K}^n$, d. h. zur Anzahl der Spalten, korrespondiert. Entsprechend bedeutet bei Matrixelementen a_{ij} der erste Index die Zeilennummer und der zweite die Spaltennummer. Wir betonen nochmals, dass dies nur eine von vielen möglichen konkreten Darstellungen einer linearen Abbildung in $\mathbb{K}^n$ ist. In diesem Sinne definiert z. B. jede quadratische Matrix $A \in \mathbb{K}^{n\times n}$ eine lineare Abbildung in $\mathbb{K}^n$. Eine quadratische Matrix $A \in \mathbb{K}^{n\times n}$ ist „regulär", wenn die zugehörige lineare Abbildung injektiv und surjektiv, d. h. bijektiv, ist.

Lemma 1.10: *Für $A = (a_{ij})_{i,j=1}^n \in \mathbb{K}^{n\times n}$ sind die folgenden Aussagen äquivalent:*

i) A ist regulär.

ii) $Ax = b$ ist für jedes $b \in \mathbb{K}^n$ eindeutig lösbar (Bijektivität).

iii) $Ax = 0$ ist nur durch $x = 0$ lösbar (Injektivität).

iv) $Ax = b$ ist für jedes $b \in \mathbb{K}^n$ lösbar (Surjektivität).

v) $\mathrm{Rang}(A) = n$.

vi) $\det(A) \neq 0$.

vii) Alle Eigenwerte $\lambda \in \mathbb{C}$ von A sind ungleich Null.

viii) Die (komplex) Transponierte $\bar{A}^T$ ist regulär.

Die Begriffe „Rang" $\mathrm{Rang}(A)$, „Determinante" $\det(A)$, „Transponierte" $\bar{A}^T$ sowie „Eigenwert" λ einer Matrix A werden als bekannt vorausgesetzt ($\Rightarrow$ Lineare Algebra) und werden im Folgenden nur bei Bedarf näher diskutiert.

Zwei Matrizen $A, A' \in \mathbb{K}^{n\times n}$ sind identisch, d. h. $a_{ij} = a'_{ij}$ $(i,j = 1,\dots,n)$, genau dann wenn

$$Ax = A'x \quad \forall x \in \mathbb{K}^n.$$

Zwei Matrizen $A, A' \in \mathbb{K}^{n\times n}$ heißen „ähnlich“ , wenn es eine reguläre Matrix $T \in \mathbb{K}^{n\times n}$ gibt, so dass gilt:

$$A' = T^{-1}AT.$$

Der Übergang $A \to A'$ wird „Ähnlichkeitstransformation“ genannt. Aus dem Determinantensatz $\det(AB) = \det(A)\det(B)$ folgt $\det(T^{-1}) = \det(T)^{-1}$ und weiter

$$\begin{aligned}\det(A' - zI) &= \det(T^{-1}AT - zT^{-1}T) = \det(T^{-1}(A - zI)T)\\ &= \det(T^{-1})\det(A - zI)\det(T) = \det(A - zI),\end{aligned}$$

für $z \in \mathbb{C}$, wobei $I = (\delta_{ij})_{i,j=1}^n$ die sog. „Einheitsmatrix“ ist. Hieraus entnehmen wir, dass ähnliche Matrizen dieselben Eigenwerte (Nullstellen ihrer charakteristischen Polynome) haben; sie haben aber i. Allg. unterschiedliche Eigenvektoren.

Wir betrachten nun den Vektorraum der $m\times n$-Matrizen $A \in \mathbb{K}^{m\times n}$. Offenbar kann dieser mit dem Vektorraum der mn-Vektoren identifiziert werden. Somit übertragen sich alle Aussagen für Vektornormen auch auf Normen für Matrizen. Insbesondere sind alle Normen für $m\times n$-Matrizen äquivalent und die Konvergenz von Folgen von Matrizen ist die komponentenweise Konvergenz:

$$A^{(k)} \to A \ (k\to\infty) \quad \Longleftrightarrow \quad a_{ij}^{(k)} \to a_{ij}\ (k\to\infty), \quad i = 1,\dots,m\,, j = 1,\dots,n\,.$$

Für eine beliebige Vektornorm $\|\cdot\|$ auf $\mathbb{K}^n$ wird für Matrizen $A \in \mathbb{K}^{n\times n}$ durch

$$\|A\| := \sup_{x\in\mathbb{K}^n\setminus\{0\}} \frac{\|Ax\|}{\|x\|} = \sup_{x\in\mathbb{K}^n,\|x\|=1} \|Ax\|$$

eine Norm erklärt (Übungsaufgabe). Diese heißt die von $\|\cdot\|$ erzeugte „natürliche Matrixnorm“ und wird meist, wenn Missverständnisse ausgeschlossen sind, genauso wie die erzeugende Vektornorm bezeichnet. Für natürliche Matrixnormen gilt notwendig $\|I\| = 1$. Eine solche natürliche Matrixnorm ist mit der erzeugenden Vektornorm „verträglich“, d. h.:

$$\|Ax\| \le \|A\|\,\|x\|\,, \quad x \in \mathbb{K}^n,\ A \in \mathbb{K}^{n\times n}\,. \tag{1.4.16}$$

Ferner ist sie „submultiplikativ“:

$$\|AB\| \le \|A\|\,\|B\|\,, \quad A, B \in \mathbb{K}^{n\times n}. \tag{1.4.17}$$

Eine submultiplikative Matrixnorm wird oft auch als „Matrizennorm“ bezeichnet. Wir werden im Folgenden diese subtile Unterscheidung der Normbegriffe aber nicht verwenden.

Nicht jede Matrixnorm ist auch „natürlich“; z. B. sieht man leicht mit Hilfe der Schwarzschen Ungleichung, dass die Quadratsummennorm (sog. „Frobenius[11]-Norm“)

$$\|A\|_F := \Big(\sum_{j,k=1}^n |a_{jk}|^2\Big)^{1/2}$$

[11] Ferdinand Georg Frobenius (1849–1917): Deutscher Mathematiker; Prof. in Zürich und Berlin; bed. Beiträge zur Theorie der Differentialgleichungen, zu Determinanten und Matrizen sowie zur Gruppentheorie.

zwar mit der euklidischen Norm verträglich und submultiplikativ ist, aber wegen $\|I\|_F = \sqrt{n}$ (für $n \geq 2$) keine *natürliche* Matrixnorm sein kann.

Lemma 1.11 (Natürliche Matrixnormen): *Die natürlichen Matrixnormen zur Maximumnorm $\|\cdot\|_\infty$ und zur l_1-Norm $\|\cdot\|_1$ sind die sog. „Maximale-Zeilensummen-Norm“*

$$\|A\|_\infty := \max_{1\leq i\leq n} \sum_{j=1}^n |a_{ij}| \tag{1.4.18}$$

bzw. die „Maximale-Spaltensummen-Norm“

$$\|A\|_1 := \max_{1\leq j\leq n} \sum_{i=1}^n |a_{ij}|\,. \tag{1.4.19}$$

Beweis: i) Offenbar ist die maximale Zeilensumme $\|\cdot\|_\infty$ eine Matrizennorm. Wegen

$$\|Ax\|_\infty = \max_{1\leq i\leq n} |\sum_{j=1}^n a_{ij}x_j| \leq \max_{1\leq i\leq n} \sum_{j=1}^n |a_{ij}| \max_{1\leq j\leq n} |x_j| = \|A\|_\infty \|x\|_\infty$$

ist sie verträglich mit $\|\cdot\|_\infty$. Im Falle $\|A\|_\infty = 0$ ist $A = 0$, d. h. trivialerweise

$$\|A\|_\infty = \sup_{\|x\|_\infty=1} \|Ax\|_\infty\,.$$

Sei also $\|A\|_\infty > 0$ und $m \in \{1,\dots,n\}$ ein Index mit der Eigenschaft

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

Wir setzen für $j = 1,\dots,n$: $z_j \equiv |a_{mj}|/a_{mj}$ für $a_{mj} \neq 0$ und $z_j \equiv 0$ sonst, d. h.: $z = (z_j)_{j=1}^n \in \mathbb{K}^n$, $\|z\|_\infty = 1$. Für $v := Az$ gilt dann

$$v_m = \sum_{j=1}^n a_{mj}z_j = \sum_{j=1}^n |a_{mj}| = \|A\|_\infty\,.$$

Folglich ist

$$\|A\|_\infty = v_m \leq \|v\|_\infty = \|Az\|_\infty \leq \sup_{\|y\|_\infty=1} \|Ay\|_\infty\,.$$

ii) Der Beweis für die l_1-Norm verläuft analog und sei als Übungsaufgabe gestellt. Q.E.D.

Definition 1.10: *i) Die „Eigenwerte“ $\lambda \in \mathbb{K}$ einer Matrix $A \in \mathbb{K}^{n\times n}$ sind definiert als die Nullstellen ihres charakteristischen Polynoms $p(\lambda) = \det(A - \lambda I)$. Folglich existieren genau n (ihrer Vielfachheit als Nullstelle, „algebraische Vielfachheit“, entsprechend oft gezählte) Eigenwerte λ.*

ii) Die Eigenwerte einer Matrix bilden deren „Spektrum“ $\sigma(A)$.

iii) Zu jedem $\lambda \in \sigma(A)$ *existiert ein* Eigenvektor $w \in \mathbb{K}^n \setminus \{0\}$*, so dass*

$$Aw = \lambda w.$$

Die Eigenvektoren zu einem Eigenwert $\lambda \in \sigma(A)$ *bilden einen Vektorraum, den „Eigenraum“ zu* λ*, dessen Dimension ist die sog. „geometrische“ Vielfachheit von* λ.

Die Matrix $A \in \mathbb{K}^{n\times n}$ heißt „hermitesch“, wenn gilt:

$$A = \bar{A}^T \quad \text{bzw.} \quad a_{ij} = \overline{a_{ji}},\ i,j = 1,\dots,n\,.$$

Reelle hermitesche Matrizen werden „symmetrisch“ genannt. Der Begriff der Symmetrie ist eng verknüpft mit dem des Skalarprodukts. Mit dem euklidischen Skalarprodukt gilt:

$$A = \bar{A}^T \quad \Leftrightarrow \quad (Ax,y)_2 = (x,Ay)_2, \quad x,y \in \mathbb{K}^n.$$

Lemma 1.12: *i) Die geometrische Vielfachheit eines Eigenwerts ist stets kleiner oder gleich seiner algebraischen Vielfachheit. Für hermitesche/symmetrische Matrizen, oder allgemeiner für „normale“ Matrizen (d. h.:* $\bar{A}^T A = A\bar{A}^T$*), sind sie gleich.*

ii) Eine hermitesche/symmetrische Matrix oder allgemeiner eine normale Matrix ist „diagonalisierbar“, d. h. ähnlich zu einer Diagonalmatrix. Dies ist äquivalent zur Existenz einer zugehörigen Basis von Eigenvektoren.

iii) Für hermitische/symmetrische Matrizen sind alle Eigenwerte reell. Eigenvektoren zu unterschiedlichen Eigenwerten sind orthogonal zueinander, und es existiert eine Orthonormalbasis aus Eigenvektoren.

Beweis: Übungsaufgabe (s. Lineare Algebra). Q.E.D.

Sei nun $\|\cdot\|$ eine beliebige Vektornorm und $\|\cdot\|$ eine damit verträgliche Matrizennorm. Mit einem normierten Eigenvektor, $\|w\| = 1$, zum Eigenwert λ gilt dann:

$$|\lambda| = |\lambda|\,\|w\| = \|\lambda w\| = \|Aw\| \ \le\ \|A\|\,\|w\| = \|A\|, \tag{1.4.20}$$

d. h. alle Eigenwerte von A liegen in einer Kreisscheibe in $\mathbb{C}$ mit Mittelpunkt Null und Radius $\|A\|$. Speziell mit $\|A\|_\infty$ erhält man die Abschätzung

$$\max_{\lambda\in\sigma(A)} |\lambda| \le \|A\|_\infty = \max_{1\le i\le n} \sum_{j=1}^{n} |a_{ij}|.$$

Eine Matrix $A \in \mathbb{K}^{n\times n}$ heißt „positiv definit“, wenn gilt:

$$(Ax,x)_2 \in \mathbb{R}\,, \quad (Ax,x)_2 > 0 \quad \forall x \in \mathbb{K}^n \setminus \{0\}.$$

Eine hermitesche Matrix ist genau dann positiv definit, wenn alle ihre (reellen) Eigenwerte positiv sind. Im folgenden werden wir in Verbindung mit der Eigenschaft *positiv definit* stets auch die Eigenschaft *hermitesch* (bzw. *symmetrisch* im Reellen) einer Matrix annehmen. Dies ist im Komplexen automatisch gegeben, im Rellen aber eine zusätzliche Bedingung (Übungsaufgabe).

Die von der euklidischen Vektornorm erzeugte natürliche Matrizennorm heißt die „Spektralnorm" und wird mit $\|\cdot\|_2$ bezeichnet. Diese Bezeichnung ist durch das folgende Resultat gerechtfertigt:

Lemma 1.13 (Spektralnorm): *Für eine beliebige Matrix $A \in \mathbb{K}^{n\times n}$ ist die Matrix $\overline{A}^T A \in \mathbb{K}^{n\times n}$ stets hermitesch und positiv semi-definit. Für die Spektralnorm von A gilt:*

$$\|A\|_2 = \max\{|\lambda|^{1/2},\ \lambda \in \sigma(\bar{A}^T A)\}. \tag{1.4.21}$$

Ist A hermitesch (bzw. symmetrisch), so gilt:

$$\|A\|_2 = \max\{|\lambda|,\ \lambda \in \sigma(A)\}. \tag{1.4.22}$$

Beweis: Wir geben den Beweis nur für den Fall, dass A hermitesch ist. Der Beweis für den allgemeinen Fall wird als Übungsaufgabe gestellt. Seien $\lambda_i \in \sigma(A)$ die n, ihrer Vielfachheiten entsprechend oft gezählten (reellen) Eigenwerte von A und $\{w^{(i)},\ i = 1,\dots,n\}$ eine zugehörige Orthonormalbasis von Eigenvektoren, so dass $Aw^{(i)} = \lambda_i w^{(i)}$. Aufgrund der Eigenwertschranke (1.4.20) gilt zunächs $|\lambda_{\max}| \le \|A\|_2$. Ferner ist

$$\begin{aligned}\|Ax\|_2^2 &= \sum_{i,j=1}^n (x,w^{(i)})_2\overline{(x,w^{(j)})_2}(Aw^{(i)},Aw^{(j)})_2 = \sum_{i,j=1}^n (x,w^{(i)})_2\overline{(x,w^{(j)})_2}\lambda_i\overline{\lambda_i}(w^{(i)},w^{(j)})_2 \\ &= \sum_{i=1}^n |\lambda_i|^2|(x,w^{(i)})_2|^2 \le |\lambda_{\max}|^2\|x\|_2^2,\end{aligned}$$

und folglich

$$\|A\|_2 = \sup_{x\in\mathbb{K}^n, x\neq 0}\frac{\|Ax\|_2}{\|x\|_2} \le |\lambda_{\max}| \sup_{x\in\mathbb{K}^n, x\neq 0}\frac{\|x\|_2}{\|x\|_2} \le |\lambda_{\max}|.$$

Q.E.D.

Definition 1.11: *Eine Matrix $Q \in \mathbb{K}^{m\times n}$ heißt „orthonormal", wenn ihre Spaltenvektoren ein Orthonormalsystem im $\mathbb{K}^m$ bilden. Im Fall $n = m$ heißt eine solche Matrix „unitär".*

Lemma 1.14: *Eine unitäre Matrix $Q \in \mathbb{K}^{n\times n}$ ist regulär und ihre Inverse ist $Q^{-1} = \overline{Q}^T$. Ferner gelten die Beziehungen*

$$(Qx,Qy)_2 = (x,y)_2, \quad x,y \in \mathbb{K}^n, \tag{1.4.23}$$

$$\|Qx\|_2 = \|x\|_2, \quad x \in \mathbb{K}^n, \tag{1.4.24}$$

d. h. euklidisches Skalarprodukt und euklidische Norm von Vektoren sind invariant unter einer unitären Transformation. Dies impliziert insbesondere, dass $\|Q\|_2 = \|Q^{-1}\|_2 = 1$.

Beweis: i) Wir zeigen zunächst, dass $\overline{Q}^T$ die Inverse von Q ist. Seien mit $q_i \in \mathbb{K}^n$ die Spaltenvektoren von Q. Für diese gilt $(q_i, q_j)_2 = q_i^T \overline{q}_j = \delta_{ij}$. Damit folgt:

$$\overline{Q}^T Q = \begin{pmatrix} \overline{q}_1^T q_1 & \dots & \overline{q}_1^T q_n \\ \vdots & \ddots & \vdots \\ \overline{q}_n^T q_1 & \dots & \overline{q}_n^T q_n \end{pmatrix} = \begin{pmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \end{pmatrix} = I.$$

ii) Mit Hilfe von (i) ergibt sich

$$(Qx, Qy)_2 = (x, \overline{Q}^T Qy)_2 = (x, y)_2,$$

und somit auch $\|Qx\|_2 = (Qx, Qx)_2^{1/2} = \|x\|_2$. Damit folgt dann

$$\|Q\|_2 = \sum_{x \in \mathbb{K}^n, x \neq 0} \frac{\|Qx\|_2}{\|x\|_2} = \sum_{x \in \mathbb{K}^n, x \neq 0} \frac{\|x\|_2}{\|x\|_2} = 1,$$

sowie

$$\|Q^{-1}\|_2 = \sum_{x \in \mathbb{K}^n, x \neq 0} \frac{\|Q^{-1}x\|_2}{\|x\|_2} = \sum_{y \in \mathbb{K}^n, y \neq 0} \frac{\|y\|_2}{\|Qy\|_2} = \sum_{y \in \mathbb{K}^n, y \neq 0} \frac{\|y\|_2}{\|y\|_2} = 1.$$

Q.E.D.

Beispiel 1.6: Die Matrix

$$Q_\theta^{(ij)} = \begin{array}{c} \begin{array}{ccccc} & i & & j & \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & \cos(\theta) & 0 & -\sin(\theta) & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & \sin(\theta) & 0 & \cos(\theta) & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array} \begin{array}{c} \\ \\ i \\ \\ j \\ \\ \end{array}$$

beschreibt eine Drehung in der (x_i, x_j)-Ebene um den Ursprung $x = 0$ mit dem Drehwinkel $\theta \in [0, 2\pi)$. Sie ist offenbar unitär.

Lemma 1.15: *Zu jeder regulären Matrix $A \in \mathbb{R}^{n \times n}$ existiert eine multiplikative Zerlegung*

$$A = Q_1 D Q_2 \tag{1.4.25}$$

mit einer Diagonalmatrix $D = \operatorname{diag}(\mu_1, \dots, \mu_n)$ mit Zahlen $\mu_i > 0$ und zwei orthonormalen Matrizen $Q_1, Q_2 \in \mathbb{R}^{n \times n}$.

Beweis: Die Matrix AA^T ist symmetrisch und positiv-definit (Übungsaufgabe). Daher gibt es eine orthonormale Matrix $Q \in \mathbb{R}^{n\times n}$, so dass $Q^TA^TAQ = D_1 = \operatorname{diag}(\mu_1, \dots, \mu_n)$ mit den Eigenwerten $\mu_i > 0$ von A^TA. Wir setzen $D := \operatorname{diag}(\lambda_1, \dots, \lambda_n)$ mit $\lambda_i := \sqrt{\mu_i}$. Dann ist

$$I = D^{-1}D_1D^{-1} = D^{-1}Q^TA^TAQD^{-1}.$$

Für $Q_2 := AQD^{-1}$ ist $Q_2^T = D^{-1}Q^TA^T$; also $Q_2^TQ_2 = D^{-1}Q^TA^TAQD^{-1} = D^{-1}D_1D^{-1} = I$, d. h. Q_2 ist orthonormal. Ferner gilt $A = (AQD^{-1})DQ^T$, was den Beweis vervollständigt. Q.E.D.

Bemerkung 1.9: Die Schwarzsche Ungleichung (1.3.4) erlaubt die Definition eines „Winkels" zwischen zwei Vektoren. Zu jeder Zahl $\alpha \in [-1, 1]$ gibt es genau ein $\theta \in [0, \pi]$ mit $\alpha = \cos(\theta)$. Für $x, y \in \mathbb{K}^n \setminus \{0\}$ wird also durch

$$\cos(\theta) = \frac{(x, y)_2}{\|x\|_2\|y\|_2}$$

ein $\theta \in [0, \pi]$ eindeutig festgelegt. Dies ist dann der „Winkel" zwischen den Vektoren x und y. Diese Definition ist verträglich mit der üblichen Definition des Winkels in der Ebene, was man wie folgt sieht: Die Beziehung (1.4.23) besagt, dass das euklidische Skalarprodukt zweier Vektoren im $\mathbb{K}^n$ invariant gegenüber Drehungen ist. Durch eine Drehung Q im $\mathbb{R}^n$ lässt sich erreichen, dass $Qx, Qy \in \operatorname{span}\{e^{(1)}, e^{(2)}\}$ liegt und $Qx = \|x\|_2e^{(1)}$.

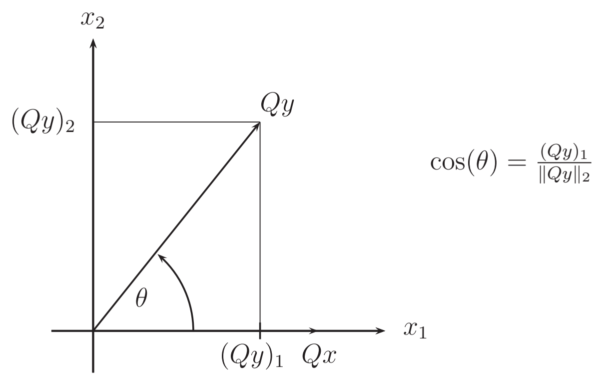

Abbildung 1.1: *Winkel zwischen zwei Vektoren* $x = \|x\|e^{(1)}$ *und* y *im* $\mathbb{R}^2$.

Dann ist

$$\begin{aligned}(x, y)_2 = (Qx, Qy)_2 = \|x\|_2(e^{(1)}, Qy)_2 = \|x\|_2(Qy)_1 = \|x\|_2\|Qy\|_2\cos(\theta)\\ = \|x\|_2\|y\|_2\cos(\theta),\end{aligned}$$

d. h.: Bei θ handelt es sich tatsächlich um den elementargeometrischen Winkel zwischen den beiden Vektoren.

Bemerkung 1.10: Skalarprodukte sind wichtig zum Studium der geometrischen Eigenschaften des $\mathbb{K}^n$ sowie der Spektraleigenschaften von linearen Abbildungen bzw. Matrizen. Daher stellt sich die Frage nach der allgemeinen Gestalt von solchen Skalarprodukten. Man zeigt leicht (Übungsaufgabe), dass sich jedes Skalarprodukt $(\cdot,\cdot)$ auf dem $\mathbb{K}^n$ mit dem euklidischen Skalarprodukt $(\cdot,\cdot)_2$ und einer geeigneten hermiteschen, positiv definiten Matrix $A \in \mathbb{K}^{n\times n}$ in der Form

$$(x,y) = (Ax,y)_2, \quad x,y \in \mathbb{K}^n,$$

darstellen lässt.

Der folgende Hilfssatz liefert ein nützliches Kriterium für die Regularität von „kleinen“ Störungen der Einheitsmatrix.

Lemma 1.16 (Störungssatz): *Sei $\|\cdot\|$ eine beliebige natürliche Matrixnorm auf $\mathbb{K}^{n\times n}$. Die Störmatrix $B \in \mathbb{K}^{n\times n}$ habe die Norm $\|B\| < 1$. Dann ist die Matrix $I+B$ regulär, und es gilt*

$$\|(I+B)^{-1}\| \le \frac{1}{1-\|B\|}\,. \tag{1.4.26}$$

Beweis: Für alle $x \in \mathbb{K}^n$ gilt

$$\|(I+B)x\| \ge \|x\| - \|Bx\| \ge (1-\|B\|)\|x\|\,.$$

Wegen $1-\|B\| > 0$ ist also $I+B$ injektiv und folglich regulär. Mit der Abschätzung

$$\begin{aligned} 1 = \|I\| &= \|(I+B)(I+B)^{-1}\| = \|(I+B)^{-1} + B(I+B)^{-1}\| \\ &\ge \|(I+B)^{-1}\| - \|B\|\,\|(I+B)^{-1}\| = \|(I+B)^{-1}\|(1-\|B\|) > 0\,. \end{aligned}$$

erhält man die behauptete Ungleichung. Q.E.D.

Korollar 1.4: *Sei $A \in \mathbb{K}^{n\times n}$ regulär und $\tilde{A} \in \mathbb{K}^{n\times n}$ mit*

$$\|\tilde{A} - A\| < \frac{1}{\|A^{-1}\|}.$$

Dann ist auch $\tilde{A}$ regulär.

Beweis: Es ist $\tilde{A} = A + \tilde{A} - A = A(I + A^{-1}(\tilde{A}-A))$. Wegen

$$\|A^{-1}(\tilde{A}-A)\| \le \|A^{-1}\|\,\|\tilde{A}-A\| < 1$$

ist nach Lemma 1.16 die Matrix $I + A^{-1}(\tilde{A}-A)$ regulär. Dann ist auch das Produkt $A(I+A^{-1}(\tilde{A}-A))$ regulär, woraus wiederum die Regularität von $\tilde{A}$ folgt. Q.E.D.

1.5 Übungen

Übung 1.1: Die Einheitssphäre bzgl. einer Norm $\|\cdot\|$ auf dem Vektorraum $\mathbb{R}^n$ ist definiert durch

$$S := \{x \in \mathbb{R}^n \mid \|x\| = 1\}.$$

Man skizziere die durch die l_1-Norm, euklidische Norm und die l_∞-Norm erzeugten Einheitssphären im $\mathbb{R}^2$. Wie lauten die Lösungen für die folgenden „gewichteten"Normen:

a) gewichtete l_1-Norm: $\|x\|_{1,\omega} := |x_1| + 2|x_2|$,

b) gewichtete l_2-Norm: $\|x\|_{2,\omega} := \big(|x_1|^2 + 2|x_2|^2\big)^{1/2}$,

c) gewichtete l_∞-Norm: $\|x\|_{\infty,\omega} := \max\big\{|x_1|, 2|x_2|\big\}$.

Übung 1.2: Man zeige:

a) Durchschnitt endlich vieler und Vereinigung beliebig vieler offener Mengen sind offen.

b) Vereinigung endlich vieler und Durchschnitt beliebig vieler abgeschlossener Mengen sind abgeschlossen.

c) Man zeige durch Gegenbeispiele, dass weitergehende Verallgemeinerungen dieser Aussagen nicht richtig sind, d. h. dass

- der Durchschnitt *unendlich* vieler offener Mengen nicht offen sein muss, und

- die Vereinigung *unendlich* vieler abgeschlossenen Mengen nicht abgeschlossen sein muss.

Übung 1.3: Welche der folgenen Gleichungen für Mengen $A \subset \mathbb{K}^n$ sind richtig, welche sind falsch?

$$\begin{array}{llll} a) & (\overline{A})^o = A^o, & b) & \overline{A^o} = \overline{A}, \\ c) & A^o \cap B^o = (A \cap B)^o, & d) & A^o \cup B^o = (A \cup B)^o, \\ e) & \overline{A \cap B} = \overline{A} \cap \overline{B}, & f) & \overline{A \cup B} = \overline{A} \cup \overline{B}. \end{array}$$

(Hinweis: Man mache sich die Aussagen anhand einfacher Beispiele mit Mengen im $\mathbb{R}^1$ oder $\mathbb{R}^2$ klar.)

Übung 1.4: Man betrachte den $n-1$-dimensionalen Unterraum

$$V_{n-1} := \{x \in \mathbb{K}^n \mid x = (x_1, \ldots, x_{n-1}, 0)\}$$

des $\mathbb{K}^n$. Dieser kann wieder als ein eigenständiger normierter Raum aufgefasst und auf offensichtliche Weise mit dem Vektorraum $\mathbb{K}^{n-1}$ identifiziert werden. Sei $O \subset \mathbb{K}^n$ eine offene Menge.

a) Ist $O \cap V_{n-1}$ aufgefasst als Teilmenge im $\mathbb{K}^n$ offen?

b) Ist $O \cap V_{n-1}$ aufgefasst als Teilmenge im $\mathbb{K}^{n-1}$ offen?

Übung 1.5: Sei l_2 die Menge der Folgen $x = (x_k)_{k\in\mathbb{N}}$ reeller Zahlen, welche „quadratisch summierbar“ sind, d. h. $\sum_{i=1}^\infty x_i^2 < \infty$. Man zeige:

i) Die Menge l_2 ist mit der natürlichen Addition $x + y = (x_i + y_i)_{i\in\mathbb{N}}$ und skalaren Multiplikation $\alpha x = (\alpha x_i)_{i\in\mathbb{N}}$ von Folgen ein Vektorraum.

ii) Auf l_2 sind durch

$$(x, y)_2 := \sum_{i=1}^{\infty} x_i y_i, \qquad \|x\|_2 := \Big(\sum_{i=1}^{\infty} x_i^2\Big)^{1/2},$$

ein Skalarprodukt mit zugehöriger Norm definiert.

iii) Der normierte Raum $(l_2, \|\cdot\|_2)$ ist vollständig, d. h. ein Banach-Raum.

Übung 1.6: Sei l_1 die Menge der „absolut summierbaren“ Folgen $x = (x_k)_{k\in\mathbb{N}}$ reeller oder komplexer Zahlen, d. h. die Menge aller Folgen mit der Eigenschaft

$$\sum_{i=1}^{\infty} |x_i| = \lim_{n\to\infty} \sum_{i=1}^{n} |x_i| < \infty.$$

Man zeige:

a) Die Menge l_1 ist mit der natürlichen Addition $x + y := (x_i + y_i)_{i\in\mathbb{N}}$ und skalaren Multiplikation $\alpha x := (\alpha x_i)_{i\in\mathbb{N}}$ von Folgen ein Vektorraum. Was ist dessen Dimension?

b) Auf l_1 ist eine Norm definiert durch

$$\|x\|_1 := \sum_{i=1}^{\infty} |x_i|.$$

c) Der normierte Raum $(l_1, \|\cdot\|_1)$ ist vollständig, d. h. ein Banach-Raum.

Übung 1.7: Der „Rand“ ∂M einer Menge $M \subset \mathbb{K}^n$ ist definiert durch

$$\partial M := \{x \in \mathbb{K}^n \,|\, \text{Jede Umgebung } K_r(x) \text{ enthält Punkte aus } M \text{ und } M^c.\}.$$

Man zeige:

a) Eine Menge $O \subset \mathbb{K}^n$ ist genau dann offen, wenn sie keinen ihrer Randpunkte enthält.

b) Eine Menge $A \subset \mathbb{K}^n$ ist genau dann abgeschlossen, wenn sie alle ihre Randpunkte enthält.

Übung 1.8: Man beweise oder widerlege:

a) Sei $d(\cdot,\cdot)\colon \mathbb{K}^n \times \mathbb{K}^n \to \mathbb{R}$ eine Metrik, zu der es eine stetige Funktion $\rho\colon K^n \to \mathbb{R}$ gibt mit $d(x, y) = \rho(x - y)$ für alle $x, y \in \mathbb{K}^n$. Dann ist $\rho(\cdot)$ eine Norm.

b) Eine Teilmenge $O \subset \mathbb{K}^n$ ist genau dann offen, wenn sie keinen ihrer Randpunkte enthält.

c) Der Rand ∂M einer Teilmenge $M \subset \mathbb{K}^n$ ist abgeschlossen.

d) Für Teilmengen $M \subset \mathbb{K}^n$ gilt $(\overline{M})^\circ = \overline{(M^\circ)}$.

e) Für Mengen $A, B \subset \mathbb{K}^n$ ist $A^\circ \cup B^\circ \neq (A \cup B)^\circ$.

f) Für Mengen $A, B \subset \mathbb{K}^n$ ist $A^\circ \cup B^\circ = (A \cup B)^\circ$.

Übung 1.9: Man rekapituliere für Teilmengen des $\mathbb{K}^n$ die Begriffe „offener Kern“, „Abschluss“ und „Rand“. Man bestimme den offenen Kern, den Abschluss und den Rand für die folgenden Mengen im $\mathbb{R}^n$:

a) $M := \{x \in \mathbb{R}^n : \|x\|_\infty < 1,\, x_i \in \mathbb{Q},\, i = 1, \dots, n\}$;

b) $M := \{x \in \mathbb{R}^n : \|x\|_2 \leq 1,\, x_1 = 0\}$.

c) $M := \{x \in \mathbb{R}^n : f(x) < 1\}$ mit $f(x) := \begin{cases} 1 & x \in (-1, 1)^n, \\ 0 & \text{sonst.} \end{cases}$

d) $M := \{x \in \mathbb{R}^n : f(x) \leq 1\}$ mit $g(x) := \frac{3}{2} - f(x)$.

Übung 1.10: Der Satz über die Normäquivalenz und die Sätze von Bolzano-Weierstraß und Heine-Borel wurden im Text nur für den $\mathbb{K}^n$ (bzw. allgemein für endlich dimensionale Vektorräume) bewiesen.

a) In unendlich-dimensionalen normierten Räumen gilt der Satz von der Normäquivalenz nicht immer. Man mache sich dies durch Konstruktion eines Gegenbeispiels im Raum $C[0,1]$ der stetigen Funktionen (mit der Maximumnorm) klar.

b) In unendlich dimensionalen normierten Räumen gelten die Sätze von Bolzano-Weierstraß und Heine-Borel nicht immer. Man mache sich dies durch Konstruktion eines Gegenbeispiels im Folgenraum l_2 klar.

Übung 1.11: Sei $(\cdot, \cdot)$ irgendein Skalarprodukt mit zugehöriger Norm $\|\cdot\|$ auf einem *reellen* Vektorraum V (z. B. dem $\mathbb{R}^n$). Man beweise für Punkte $x, x' \in V$ die folgenden Aussagen:

a) $x = x' \quad \Leftrightarrow \quad (x, y) = (x', y) \quad \forall y \in \mathbb{R}^n$;

b) $(x, x') = 0 \quad \Leftrightarrow \quad \|x + x'\|^2 = \|x\|^2 + \|x'\|^2$ („Satz von Pythagoras“).

c) $\|x + x'\|^2 + \|x - x'\|^2 = 2\|x\|^2 + 2\|x'\|^2$. („Parallelogrammidentität“).

Gelten diese Aussagen auch für *komplexe* Vektorräume (z. B. dem $\mathbb{C}^n$)?
(Bemerkung: Man mache sich diese Aussagen auch durch geometrische Überlegungen für das euklidischen Skalarprodukt auf dem $\mathbb{R}^2$ klar.)

Übung 1.12: Sei $a(\cdot,\cdot) : V \times V \to \mathbb{C}$ eine Sesquilinearform auf einem komplexen Vektorraum V, d. h. für beliebige $x, y \in V$ und $\alpha, \beta \in \mathbb{C}$ gilt:

$$a(\alpha x_1 + \beta x_2, y) = \alpha a(x_1, y) + \beta a(x_2, y),$$
$$a(x, \alpha y_1 + \beta y_2) = \overline{\alpha} a(x, y_1) + \overline{\beta} a(x, y_2).$$

a) Man zeige, dass $a(\cdot,\cdot)$ im Falle $a(x,x) \in \mathbb{R}$ notwendig hermitesch ist:

$$a(x,y) = \overline{a(y,x)}, \quad x, y \in V.$$

b) Gilt die entsprechende Aussage auch für Bilinearformen auf reellen Vektorräumen, wenn man zusätzlich noch die Definitheit $a(x,x) \geq 0$ fordert, d. h.: Folgt in diesem Fall aus ihrer Definitheit notwendig auch ihre Symmetrie?

Übung 1.13: a) Sei $(V, \|\cdot\|)$ ein *reeller* normierter Raum, dessen Norm $\|\cdot\|$ die „Parallelogrammidentität“ erfüllt:

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2, \quad x, y \in V.$$

Man zeige, dass dann durch

$$(x,y) := \tfrac{1}{4}\|x+y\|^2 - \tfrac{1}{4}\|x-y\|^2, \quad x, y \in V,$$

auf V ein Skalarprodukt definiert ist, durch welches die gegebene Norm erzeugt wird. (Hinweis: Der Nachweis der verschiedenen Skalarprodukteigenschaften ist von sehr unterschiedlicher Schwierigkeit. Die Aussage gilt auch für *komplexe* Vektorräume, allerdings mit einer entsprechend angepassten Definition des Skalarprodukts $(\cdot,\cdot)$.)

b) Man zeige, dass für $p \in [1,\infty) \cup \{\infty\} \setminus \{2\}$ die l_p-Normen auf dem $\mathbb{R}^n$ mit $n > 1$,

$$\|x\|_p := \Big(\sum_{i=1}^n |x_i|^p \Big)^{1/p}, \quad x \in \mathbb{R}^n,$$

nicht von Skalarprodukten erzeugt werden.

Übung 1.14: a) Man zeige, dass für jede Norm $\|\cdot\|$ auf $\mathbb{K}^n$ durch

$$\|A\| := \sup_{x \in \mathbb{K}^n \setminus \{0\}} \frac{\|Ax\|}{\|x\|} = \sup_{\substack{x \in \mathbb{K}^n \\ \|x\|=1}} \|Ax\| = \max_{\substack{x \in \mathbb{K}^n \\ \|x\|=1}} \|Ax\|, \quad A \in \mathbb{K}^{n\times n},$$

eine Matrixnorm auf $\mathbb{K}^{n\times n}$ definiert ist. Diese wird als die von der Vektornorm $\|\cdot\|$ erzeugte „natürliche Matrixnorm“ bezeichnet.

b) Man zeige, dass die sog. „Frobenius-Norm“

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

zwar mit der euklidischen Vektornorm verträglich und submultiplikativ (und damit sogar eine „Matrizennorm“) ist, jedoch nicht von einer Vektornorm erzeugt wird.

Übung 1.15: Sei $A \in \mathbb{K}^{n\times n}$ eine hermitesche Matrix. Man zeige:

i) Alle Eigenwerte von A sind reell.

ii) Eigenvektoren zu unterschiedlichen Eigenwerten sind orthogonal zueinander.

iii) Es gibt eine Orthonormalbasis aus Eigenvektoren von A.

Übung 1.16: Sei $A \in \mathbb{K}^{n\times n}$ beliebig. Man zeige:

i) Die Matrix $\bar{A}^T A$ ist hermitesch und positiv semi-definit. Für reguläres A ist $\bar{A}^T A$ sogar positiv definit.

ii) Für die Spektralnorm von A gilt:

$$\|A\|_2 = \max\{|\lambda|^{1/2},\ \lambda \in \sigma(\bar{A}^T A)\}.$$

Übung 1.17: a) Man verifiziere, dass die beiden (diagonalisierbaren) 2×2-Matrizen

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

nicht kommutieren, d. h.: $AB \neq BA$.

b) Man zeige, dass je zwei diagonalisierbare Matrizen $A, B \in \mathbb{K}^{n\times n}$ kommutieren, d. h.:

$$AB = BA,$$

wenn sie eine gemeinsame Basis von Eigenvektoren besitzen.
Zusatz für Ehrgeizige: Man zeige, dass letztere Bedingung auch notwendig für das Kommutieren ist, d. h.: Sind A, B diagonalisierbar mit $AB = BA$, dann besitzen sie eine gemeinsame Basis von Eigenvektoren.

(Hinweis: Die Existenz einer gemeinsamen Basis von Eigenvektoren bedeutet, dass die beiden Matrizen A, B durch dieselbe Ähnlichkeitstransformation simultan auf Diagonalgestalt gebracht werden können:

$$T^{-1}AT = \Lambda_A, \qquad T^{-1}BT = \Lambda_B.$$

Dabei sind die gemeinsamen Eigenvektoren die Spaltenvektoren der Transformationsmatrix $T \in \mathbb{K}^{n\times n}$ und die Diagonalmatrizen $\Lambda_A = \mathrm{diag}(\lambda_i(A))_{i=1}^n$ und $\Lambda_B = \mathrm{diag}(\lambda_i(B))_{i=1}^n$ enthalten gerade die Eigenwerte von A bzw. B.)

Übung 1.18: a) Man zeige, dass die Menge M der regulären Matrizen in $\mathbb{K}^{n\times n}$,

$$M := \{A \in \mathbb{K}^{n\times n} \,|\, A \text{ regulär}\} \subset \mathbb{K}^{n\times n},$$

bzgl. jeder Matrixnorm offen ist.

b) Für eine Matrix $A \in \mathbb{K}^{n\times n}$ ist auf ihrer „Resolventenmenge“

$$\mathrm{Res}(A) := \{z \in \mathbb{C} \,|\, A - zI \in \mathbb{K}^{n\times n} \text{ regulär}\} \subset \mathbb{C}$$

die „Resolvente“ $R(z) := (A - zI)^{-1}$ definiert. Als Komplement des Spektrums $\sigma(A)$ ist die Resolventenmenge offen. Man zeige, dass die Resolvente $R(z) := (A - zI)^{-1}$ auf $\mathrm{Res}(A)$ stetig und auf jeder kompakten Teilmenge von $\mathrm{Res}(A)$ gleichmäßig Lipschitz-stetig ist.

Übung 1.19: Für eine Matrix $A \in \mathbb{K}^{n\times n}$ ist die Exponentialmatrix $e^A \in \mathbb{K}^{n\times n}$ formal durch die folgende Potenzreihe definiert:

$$e^A := \sum_{k=0}^{\infty} \frac{A^k}{k!} := \lim_{n\to\infty} \sum_{k=0}^{n} \frac{A^k}{k!}.$$

a) Man zeige, dass diese Reihe bzgl. jeder beliebigen Matrixnorm konvergiert, d. h. dass die Folge der Partialsummen bzgl. jeder solchen Norm einen Limes hat, der dann eindeutig bestimmt ist und mit e^A bezeichnet wird. (Hinweis: Cauchysches Konvergenzkriterium)

b) Man zeige für diagonalisierbare Matrizen $A, B \in \mathbb{K}^{n\times n}$, welche eine gemeinsame Basis von Eigenvektoren besitzen, die Beziehung

$$e^{A+B} = e^A e^B.$$

(Hinweis: Man beachte den Hinweis zu Aufgabe 1.17, die Beziehung für die Eigenwerte $\lambda_i(A+B) = \lambda_i(A) + \lambda_i(B)$ und die bekannte Beziehung $e^{a+b} = e^a e^b$ für Zahlen $a, b \in \mathbb{K}$. Im Allgemeinen ist für nicht kommutierende Matrizen $e^{A+B} \neq e^A e^B$.)