Kapitel 1

Normen und Skalarprodukte

1.1 Normen

Definition (Norm). Sei V ein Vektorraum über $\mathbb{K}$. Eine Funktion

$$V \to \mathbb{R}, \quad v \mapsto \|v\|$$

heißt eine *Norm* auf V, wenn sie die nachfolgenden vier Eigenschaften erfüllt:

(1) *Nichtnegativität:* Für alle $v \in V$ gilt $\|v\| \geq 0$.

(2) *Definitheit:* Für alle $v \in V$ gilt $\|v\| = 0 \iff v = 0$.

(3) *Homogenität:* Für alle $v \in V$ und alle $\alpha \in \mathbb{K}$ gilt

$$\|\alpha v\| = |\alpha| \, \|v\|.$$

(4) *Dreiecksungleichung:* Für alle $v, w \in V$ gilt

$$\|v + w\| \leq \|v\| + \|w\|.$$

Es ist allgemein üblich, eine Norm $V \to \mathbb{R},\ v \mapsto \|v\|$ vereinfachend mit $\|\cdot\|$ zu bezeichnen. Da eine solche Norm eine Funktion von V nach $\mathbb{R}$ ist, kann man sie auch als

$$\|\cdot\| : V \to \mathbb{R}$$

in Funktionsschreibweise darstellen. Der Punkt deutet hierbei die Stelle der Variable an. Ist $v \in V$ ein beliebiger Vektor, so nennt man die reelle Zahl $\|v\|$ üblichweise die *Norm* von v. Der Begriff „Norm“ wird also sowohl als Bezeichnung für die Funktion $\|\cdot\|$ als auch für einzelne Funktionswerte von $\|\cdot\|$ verwendet. Nachfolgend geben wir einige wichtige Beispiele für Normen an.

Beispiele.

(a) Die *reelle Betragsfunktion* $\mathbb{R} \to \mathbb{R},\ a \mapsto |a|$ ist eine Norm auf $\mathbb{R}$.

(b) Die *komplexe Betragsfunktion* $\mathbb{C} \to \mathbb{R},\ z \mapsto |z|$ ist eine Norm auf $\mathbb{C}$.

(c) Für jede natürliche Zahl $n \in \mathbb{N}$ definiert man auf dem Vektorraum $\mathbb{R}^n$ die so genannte *euklidische Norm* $|\cdot| : \mathbb{R}^n \to \mathbb{R}$ durch

$$|x| := \left(\sum_{i=1}^{n} x_i^2\right)^{1/2} = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}$$

für alle $x = (x_1, x_2, \ldots, x_n)^{\mathrm{T}} \in \mathbb{R}^n$. Die euklidische Norm auf $\mathbb{R}^n$ wird gelegentlich auch als die *Standardnorm* auf $\mathbb{R}^n$ bezeichnet.

Man beachte, dass für $n = 1$ die euklidische Norm genau der Betragsfunktion auf $\mathbb{R} = \mathbb{R}^1$ entspricht, da $\sqrt{x^2} = |x|$ für alle $x \in \mathbb{R}$ gilt. Die euklidische Norm kann also als Verallgemeinerung der Betragsfunktion angesehen werden, weshalb wir auch Betragsstriche zur Kennzeichnung dieser Norm verwenden. Im Mathematikunterricht der Oberstufe wird die euklidische Norm eines Vektors im $\mathbb{R}^3$ als „der Betrag des Vektors“ eingeführt.

(d) Für jede natürliche Zahl $n \in \mathbb{N}$ definiert man auf dem Vektorraum $\mathbb{C}^n$ die so genannte *Standardnorm* $|\cdot| : \mathbb{C}^n \to \mathbb{R}$ durch

$$|x| := \left(\sum_{i=1}^{n} |x_i|^2\right)^{1/2} = \sqrt{|x_1|^2 + |x_2|^2 + \ldots + |x_n|^2}$$

für alle $x = (x_1, x_2, \ldots, x_n)^{\mathrm{T}} \in \mathbb{C}^n$.

Für $n = 1$ entspricht die Standardnorm auf $\mathbb{C}^n$ genau der komplexen Betragsfunktion.

(e) Für jede natürliche Zahl $n \in \mathbb{N}$ definiert man auf $\mathbb{K}^n$ die so genannte *Betragssummennorm* $\|\cdot\|_1 : \mathbb{K}^n \to \mathbb{R}$ durch

$$\|x\|_1 := \sum_{i=1}^{n} |x_i| = |x_1| + |x_2| + \ldots + |x_n|$$

für alle $x = (x_1, x_2, \ldots, x_n)^{\mathrm{T}} \in \mathbb{K}^n$. Die Betragssummennorm wird gelegentlich auch als *Einsnorm* bezeichnet.

(f) Für jede natürliche Zahl $n \in \mathbb{N}$ definiert man auf $\mathbb{K}^n$ die so genannte *Maximumnorm* $\|\cdot\|_\infty : \mathbb{K}^n \to \mathbb{R}$ durch

$$\|x\|_\infty := \max_{i=1,\ldots,n} |x_i| = \max\bigl\{|x_1|, |x_2|, \ldots, |x_n|\bigr\}$$

für alle $x = (x_1, x_2, \ldots, x_n)^{\mathrm{T}} \in \mathbb{K}^n$. Hierbei bezeichnet $\max\bigl\{|x_1|, |x_2|, \ldots, |x_n|\bigr\}$ das größte Element der Menge $\bigl\{|x_1|, |x_2|, \ldots, |x_n|\bigr\}$. Die Maximumnorm wird gelegentlich auch als *Tschebyschev–Norm* oder als *Unendlichnorm* bezeichnet.

(g) Für jede natürliche Zahl $n \in \mathbb{N}$ und jede reelle Zahl $p \geq 1$ definiert man auf dem Vektorraum $\mathbb{R}^n$ die so genannte *p-Norm* $\|\cdot\|_p : \mathbb{K}^n \to \mathbb{R}$ durch

$$\|x\|_p := \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p} = \sqrt[p]{|x_1|^p + |x_2|^p + \ldots + |x_n|^p}$$

für alle $x = (x_1, x_2, \ldots, x_n)^{\mathrm{T}} \in \mathbb{K}^n$.

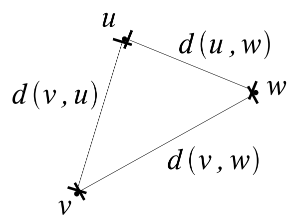

Abbildung 1.1: Die Bedeutung der Dreiecksungleichung: Der Abstand zwischen v und w ist stets kleiner oder gleich der Summe der Abstände zwischen v und u sowie zwischen u und w.

Für $p = 1$ entspricht die p-Norm genau der Betragssummennorm, und für $p = 2$ genau der euklidischen Norm auf $\mathbb{R}^n$ bzw. der Standardnorm auf $\mathbb{C}^n$. Die Betragssummennorm, die euklidische Norm auf $\mathbb{R}^n$ und die Standardnorm auf $\mathbb{C}^n$ sind also spezielle p-Normen.

(h) Sei $n \in \mathbb{N}$ eine natürliche Zahl, und sei $w = (w_1, w_2, \ldots, w_n)^{\mathrm{T}} \in \mathbb{R}^n$ ein Vektor mit positiven Komponenten, d.h. es gelte $w_i > 0$ für alle $i = 1, 2, \ldots, n$. Dann ist die Abbildung $\|\cdot\|_w : \mathbb{R}^n \to \mathbb{R}$, definiert durch

$$\|x\|_w := \left(\sum_{i=1}^{n} w_i x_i^2\right)^{1/2} = \sqrt{w_1 x_1^2 + w_2 x_2^2 + \ldots + w_n x_n^2}$$

für alle $x = (x_1, x_2, \ldots, x_n)^{\mathrm{T}} \in \mathbb{R}^n$, eine Norm auf $\mathbb{R}^n$.

Man beachte, dass für $w = (1, 1, \ldots, 1)^{\mathrm{T}}$ die Norm $\|\cdot\|_w$ genau der euklidischen Norm entspricht. $\diamondsuit$

Definition (normierter Raum). Sei V ein Vektorraum über $\mathbb{K}$, und sei $\|\cdot\|$ eine Norm auf V. Dann heißt das Paar $(V, \|\cdot\|)$ ein *normierter Raum* über $\mathbb{K}$.

Der Begriff „normierter Raum" ist ähnlich allgemein wie der Begriff „Gruppe" oder der Begriff „Vektorraum". Die Aussage, dass ein Paar $(V, \|\cdot\|)$ ein normierter Raum über $\mathbb{K}$ ist, bedeutet nicht mehr und nicht weniger als dass V ein Vektorraum über $\mathbb{K}$ ist, und dass $\|\cdot\|$ eine Norm auf V ist. Man sagt auch, dass der Vektorraum V mit der Norm $\|\cdot\|$ *versehen* und so zum normierten Raum $(V, \|\cdot\|)$ wird.

In der Regel gibt es eine Vielzahl unterschiedlicher Normen, mit denen ein gegebener Vektorraum versehen werden kann. Man betrachte dazu etwa die zuvor genannten Beispiele. Für manche Vektorräume ist es jedoch in gewisser Weise „natürlich", diese mit einer ganz bestimmten Norm zu versehen. Das gilt auch für die Vektorräume $\mathbb{R}^n$, $n \in \mathbb{N}$, für die man die folgende Vereinbarung trifft.

Vereinbarung. Ist nichts gegenteiliges ausgesagt, so ist für alle $n \in \mathbb{N}$ der Vektorraum $\mathbb{R}^n$ mit der euklidischen Norm versehen.

Wir wenden uns nun einigen Begriffen zu, die man in einem normierten Raum definieren kann. Ist $(V, \|\cdot\|)$ ein normierter Raum über $\mathbb{K}$, so kann man auf V den *Abstand* $d(v, w)$ zwischen je zwei Vektoren $v, w \in V$ gemäß

$$d(v, w) := \|w - v\|$$

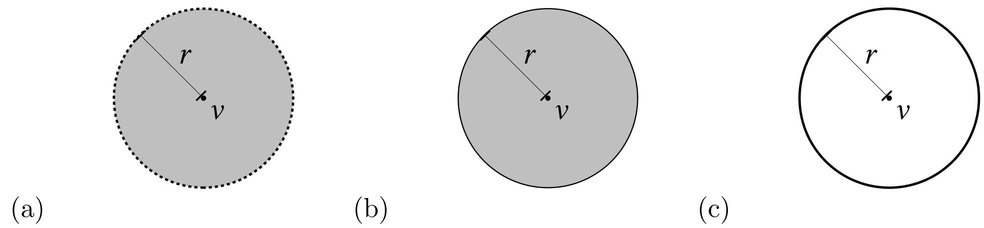

Abbildung 1.2: (a) Die offene Kugel $B_r(v)$ in $\mathbb{R}^2$ bezüglich der euklidischen Norm. (b) Die abgeschlossene Kugel $\overline{B_r(v)}$ in $\mathbb{R}^2$ bezüglich der euklidischen Norm. (c) Die Sphäre $S_r(v)$ in $\mathbb{R}^2$ bezüglich der euklidischen Norm.

definieren. Aufgrund der Nichtnegativität der Norm $\|\cdot\|$, ist der Abstand zwischen zwei beliebigen Vektoren stets größer oder gleich Null. Aufgrund der Definitheit von $\|\cdot\|$ beträgt der Abstand zwischen zwei Vektoren genau dann Null, wenn die beiden Vektoren identisch sind. Die Homogenität von $\|\cdot\|$ garantiert außerdem, dass $d(v,w) = d(w,v)$ für alle $v, w \in V$ gilt. Die Dreiecksungleichung impliziert schließlich, dass

$$d(v,w) \leq d(v,u) + d(u,w)$$

für alle Vektoren $u, v, w \in V$ gilt. Anschaulich bedeutet dies, dass die Länge der Verbindungsstrecke zwischen den „Punkten" v und w stets kleiner oder gleich groß ist wie die Summe der Längen zweier Verbindungsstrecken, welche den Punkt v bzw. den Punkt w mit einem beliebigen dritten Punkt u verbinden. Die Verbindungsstreck zwischen v und w gibt also den kürzesten Weg zwischen beiden Punkten an.

Man beachte, dass die hier diskutierten Eigenschaften des Abstandes zwischen zwei Vektoren oder Punkten unabhängig von der Norm gelten, durch die der Abstand definiert ist. Die Norm bestimmt lediglich die Größe des Abstandes zwischen zwei gegebenen Vektoren. Es zeigt sich, dass in den Vektorräumen $\mathbb{R}^1$, $\mathbb{R}^2$ und $\mathbb{R}^3$ der Abstand, welchen man für die euklidische Norm erhält, genau dem geometrischen Abstand entspricht. Dies ist mit ein Grund dafür, weshalb man zu gegebenem $n \in \mathbb{N}$ den Vektorraum $\mathbb{R}^n$ standardmäßig mit der euklidischen Norm versieht.

Betrachtet man zu einem vorgegebenen Punkt $x \in \mathbb{R}^2$ sowie zu einer vorgegebenen Zahl $r > 0$ die Menge aller Punkte, deren Abstand (bezüglich der euklidischen Norm) zum Punkt x kleiner oder gleich r ist, so stellt man fest, dass diese Menge eine Kreisscheibe mit Mittelpunkt x und Radius r ist. Im Vektorraum $\mathbb{R}^3$ ist eine solche Menge eine Kugel mit Mittelpunkt x und Radius r. Man kann also Kreisscheiben und Kugeln anhand von Abständen zwischen Punkten in $\mathbb{R}^2$ und $\mathbb{R}^3$ definieren. In gleicher Weise kann man auch in beliebigen normierten Räumen so genannte *Kugeln* definieren.

Definition (offene Kugel). Sei $(V, \|\cdot\|)$ ein normierter Raum. Zu jedem Vektor $v \in V$ und jeder positiven Zahl $r > 0$ definiert man die Menge

$$B_r(v) := \big\{ w \in V \;\big|\; \|w - v\| < r \big\}.$$

Diese Menge heißt die *offene Kugel* in V mit Mittelpunkt v und Radius r.

Definition (abgeschlossene Kugel). Sei $(V, \|\cdot\|)$ ein normierter Raum. Zu jedem Vektor $v \in V$ und jeder positiven Zahl $r > 0$ definiert man die Menge

$$\overline{B_r(v)} := \big\{ w \in V \;\big|\; \|w - v\| \leq r \big\}.$$

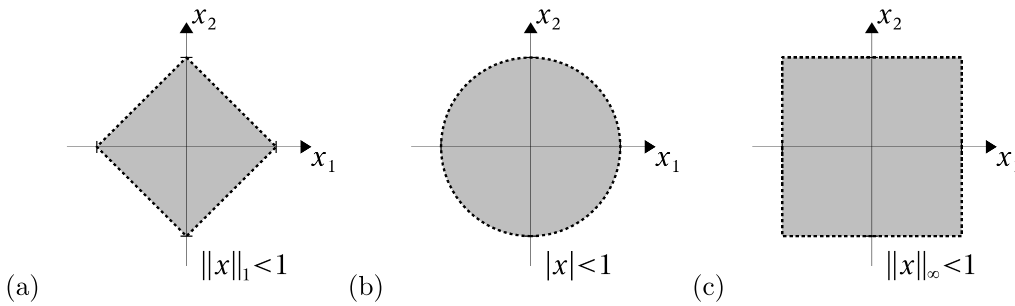

Abbildung 1.3: Offene Einheitskugeln um den Ursprung in $\mathbb{R}^2$ bezüglich (a) der Betragssummennorm $\|\cdot\|_1$, (b) der euklidische Norm $|\cdot|$ und (c) der Maximumnorm $\|\cdot\|_\infty$.

Diese Menge heißt die *abgeschlossene Kugel* in V mit Mittelpunkt v und Radius r.

Die Menge aller Punkte im Vektorraum $\mathbb{R}^2$, welche denselben Abstand $r > 0$ (bezüglich der euklidischen Norm) zu einem vorgegebenen Punkt x haben, ist bekanntlich ein Kreis mit Radius r und Mittelpunkt x. Im Vektorraum $\mathbb{R}^3$ ist bildet eine solche Menge die Oberfläche einer Kugel. Eine solche Menge wird auch *Sphäre* genannt. Solche Sphären kann man auch in allgemeinen normierten Räumen definieren.

Definition (Sphäre). Sei $(V, \|\cdot\|)$ ein normierter Raum. Zu jedem Vektor $v \in V$ und jeder positiven Zahl $r > 0$ definiert man die Menge

$$S_r(v) := \big\{ w \in V \;\big|\; \|w - v\| = r \big\}.$$

Man nennt diese Menge die *Sphäre* in V mit Mittelpunkt v und Radius r.

Offenbar gilt

$$\begin{aligned} B_r(v) &= \overline{B_r(v)} \setminus S_r(v), \\ \overline{B_r(v)} &= B_r(v) \cup S_r(v), \\ S_r(v) &= \overline{B_r(v)} \setminus B_r(v) \end{aligned}$$

für alle positiven Radien $r > 0$ und für alle Vektoren v eines normierten Raums.

Von besonderer Bedeutung sind oft Kugeln und Sphären mit dem Radius 1. Solche Kugeln und Sphären werden als *Einheitskugeln* bzw. als *Einheitssphären* bezeichnet. Ist der Mittelpunkt einer Kugel oder einer Sphäre der Nullvektor, so spricht man von einer Kugel bzw. einer Sphäre *um den Ursprung*.

Man sollte sich klar machen, dass die geometrische Gestalt einer Kugel oder einer Sphäre stets von der Norm abhängt, mit der ein Vektorraum versehen wurde. So sind beispielsweise alle in Abbildung 1.3 skizzierten Mengen offene Kugeln in $\mathbb{R}^2$, jede jedoch bezüglich einer anderen Norm auf $\mathbb{R}^2$. Man erkennt, dass die nur die offene Kugel bezüglich der euklidischen Norm die geometrische Gestalt einer Kreisscheibe besitzt.

Am Ende dieses Abschnitts geben wir noch ein nützliches Resultat an, welches als *umgekehrte Dreieckungleichung* bekannt ist.

Satz 1.1 (Umgekehrte Dreiecksungleichung). Sei $(V, \|\cdot\|)$ ein normierter Raum. Dann gilt

$$\big| \|v\| - \|w\| \big| \leq \|v - w\|$$

für alle Vektoren $v, w \in V$.

Da bekanntlich $x \leq |x|$ für alle $x \in \mathbb{R}$ gilt, folgt aus der umgekehrten Dreiecksungleichung auch, dass für alle Vektoren v und w eines mit einer Norm $\|\cdot\|$ versehenden Vektorraums die Ungleichung $\|v\| - \|w\| \leq \|v - w\|$ gilt. Ersetzt man w durch $-w$, so erhält man außerdem die Ungleichung $\|v\| - \|w\| \leq \|v + w\|$. Da jede Norm die Dreiecksungleichung erfüllt, erhält man somit die Ungleichungskette

$$\|v\| - \|w\| \leq \|v \pm w\| \leq \|v\| + \|w\|$$

für alle Vektoren $v, w \in V$.

Übungsaufgaben

1. Berechnen Sie die euklidische Norm, die Maximumnorm und die Betragssummennorm der folgenden Vektoren:

$$a := \begin{pmatrix} 2 \\ -1 \\ 5 \end{pmatrix}, \quad b := \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad c := \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad d := \begin{pmatrix} 0 \\ -4 \\ 3 \end{pmatrix}.$$

2. Skizzieren Sie die Einheitssphären $S_1(0)$ in $\mathbb{R}^2$ bezüglich der Betragssummennorm, der euklidischen Norm und der Maximumnorm.

3. Sei V ein reeller Vektorraum, und sei $\|\cdot\|$ eine Norm auf V. Zeigen Sie, dass die nachfolgenden Gleichung und Ungleichungen für alle Vektoren $u, v, w \in V$ und alle nichtnegativen Zahlen $\alpha, \beta \geq 0$ gelten.
   - $\|-v\| = \|v\|$.
   - $\|v - w\| \leq \|v - u\| + \|u - w\|$.
   - $\|\alpha v + \beta w\| \leq \alpha\|v\| + \beta\|w\|$.

4. Berechnen Sie (evtl. unter Zuhilfenahme eines Rechners) die p-Norm des Vektors

$$x = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

   für $p = 1$, $p = 2$, $p = 4$, $p = 8$ und $p = 16$, sowie dessen Maximumnorm. Was fällt Ihnen auf?

5. Sei $(V, \|\cdot\|)$ ein normierter Raum. Zeigen Sie, dass für alle Vektoren $v, w \in V$ und alle positiven Zahlen $r, s > 0$ die folgenden Aussagen gelten:
   - $\|v - w\| \leq s - r \implies B_r(v) \subseteq B_s(w)$.
   - $\|v - w\| \geq r + s \implies B_r(v) \cap B_s(w) = \emptyset$.

   Hierbei bezeichnen $B_r(v)$ und $B_s(w)$ die offenen Kugeln in V mit den Mittelpunkten v bzw. w und den Radien r bzw. s.

6. Weisen Sie nach, dass die Betragsfunktion eine Norm auf $\mathbb{R}$ ist.

7. Weisen Sie nach, dass die Betragssummennorm für alle $n \in \mathbb{N}$ eine Norm auf $\mathbb{R}^n$ ist.

1.2 Skalarprodukte

Definition (Skalarprodukt, inneres Produkt). Sei V ein reeller Vektorraum. Eine Funktion
$$V \times V \to \mathbb{R}, \quad (v, w) \mapsto \langle v\,, w \rangle$$
heißt ein *Skalarprodukt* oder ein *inneres Produkt* auf V, wenn die nachfolgenden vier Eigenschaften gelten:

(1) *Bilinearität:* Für alle $v, v_1, v_2, w, w_1, w_2 \in V$ und alle $\alpha, \beta \in \mathbb{R}$ gilt
$$\begin{aligned} \langle v_1 + v_2\,, w \rangle &= \langle v_1\,, w \rangle + \langle v_2\,, w \rangle, \\ \langle \alpha v\,, w \rangle &= \alpha \langle v\,, w \rangle, \\ \langle v\,, w_1 + w_2 \rangle &= \langle v\,, w_1 \rangle + \langle v\,, w_2 \rangle, \\ \langle v\,, \beta w \rangle &= \beta \langle v\,, w \rangle. \end{aligned}$$

(2) *Nichtnegativität:* Für alle $v \in V$ gilt $\langle v\,, v \rangle \geq 0$.

(3) *Definitheit:* Für alle $v \in V$ gilt $\langle v\,, v \rangle = 0 \iff v = 0$.

(4) *Symmetrie:* Für alle $v, w \in V$ gilt $\langle v\,, w \rangle = \langle w\,, v \rangle$.

Gemäß obiger Definition ist ein Skalarprodukt auf einem reellen Vektorraum eine Funktion, welche von zwei Variablen (oder Argumenten) abhängt, und welche *bilinear*, *nichtnegativ*, *definit* und *symmetrisch* ist. „Bilinear“ bedeutet hierbei, dass die Funktion linear bezüglich der ersten wie auch bezüglich der zweiten Variable ist. „Symmetrisch“ bedeutet, dass sich der Wert eines Skalarprodukts nicht ändert, wenn man beide Variablen vertauscht.

Es ist allgemein üblich, ein Skalarprodukt $V \times V \to \mathbb{R}$, $(v, w) \mapsto \langle v\,, w \rangle$ vereinfachend mit $\langle \cdot\,, \cdot \rangle$ zu bezeichnen. Da ein solches Skalarprodukt eine Funktion von $V \times V$ nach $\mathbb{R}$ ist, kann man es auch in Funktionsschreibweise gemäß
$$\langle \cdot\,, \cdot \rangle : V \times V \to \mathbb{R}$$
darstellen. Die Punkte deuten hierbei die Stellen an, an denen die Variablen des Skalarprodukts stehen. Nachfolgend geben wir einige wichtige Beispiele für Skalarprodukte an.

Beispiele.

(a) Die gewöhnliche *Multiplikation* $\mathbb{R} \times \mathbb{R} \to \mathbb{R}$, $(a, b) \mapsto ab$ ist ein Skalarprodukt auf $\mathbb{R}$.

(b) Für alle $n \in \mathbb{N}$ definiert man auf dem Vektorraum $\mathbb{R}^n$ das so genannte *euklidische Skalarprodukt* $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $(x, y) \mapsto x \cdot y$ durch
$$x \cdot y := \sum_{i=1}^{n} x_i y_i = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$$
für alle $x = (x_1, x_2, \ldots, x_n)^{\mathrm{T}} \in \mathbb{R}^n$ und alle $y = (y_1, y_2, \ldots, y_n)^{\mathrm{T}} \in \mathbb{R}^n$. Das euklidische Skalarprodukt auf $\mathbb{R}^n$ wird gelegentlich auch als das *Standardskalarprodukt* auf $\mathbb{R}^n$ bezeichnet.

Man beachte, dass für $n = 1$ das euklidische Skalarprodukt genau der gewöhnlichen Multiplikation auf $\mathbb{R} = \mathbb{R}^1$ entspricht. Das euklidische Skalarprodukt kann also als

Verallgemeinerung der gewöhnlichen Multiplikation angesehen werden. Daher verwenden wir auch den Malpunkt $\cdot$ zur Kennzeichnung dieses Skalarprodukts. Im Mathematikunterricht der Oberstufe wird das euklidische Skalarprodukt auf $\mathbb{R}^3$ als „das Skalarprodukt" eingeführt. Tatsächlich ist es aber eines von vielen Skalarprodukten, die man auf dem Vektorraum $\mathbb{R}^3$ definieren kann.

(c) Sei $n \in \mathbb{N}$ vorgegeben, und sei $w = (w_1, w_2, \dots, w_n)^{\mathrm{T}} \in \mathbb{R}^n$ ein Vektor mit positiven Komponenten, d.h. es gelte $w_i > 0$ für alle $i = 1, 2, \dots, n$. Dann ist die Abbildung $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R},\ (x, y) \mapsto \langle x\,,\, y\rangle_w$, welche durch

$$\langle x\,,\, y\rangle_w := \sum_{i=1}^{n} w_i x_i y_i$$

für alle $x = (x_1, x_2, \dots, x_n)^{\mathrm{T}} \in \mathbb{R}^n$ und alle $y = (y_1, y_2, \dots, y_n)^{\mathrm{T}} \in \mathbb{R}^n$ definiert ist, ein Skalarprodukt auf $\mathbb{R}^n$.

Für $w = (1, 1, \dots, 1)^{\mathrm{T}}$ entspricht das Skalarprodukt $\langle \cdot\,,\, \cdot\rangle_w$ genau dem euklidischen Skalarprodukt.

(d) Die Abbildung $\langle \cdot\,,\, \cdot\rangle_{\#} : \mathbb{R}^2 \to \mathbb{R}^2 \to \mathbb{R}$, welche durch

$$\langle x\,,\, y\rangle_{\#} := 2x_1y_1 - x_1y_2 - x_2y_1 + 2x_2y_2$$

für alle $x = (x_1, x_2)^{\mathrm{T}} \in \mathbb{R}^2$ und alle $y = (y_1, y_2)^{\mathrm{T}} \in \mathbb{R}^2$ definiert ist, ist ein Skalarprodukt auf $\mathbb{R}^2$. $\lozenge$

Definition (Innenproduktraum). Sei V ein reeller Vektorraum und sei $\langle \cdot\,,\, \cdot\rangle$ ein Skalarprodukt auf V. Dann heißt das Paar $(V, \langle \cdot\,,\, \cdot\rangle)$ ein reeller *Innenproduktraum*.

Die Aussage, dass ein Paar $(V, \langle \cdot\,,\, \cdot\rangle)$ ein reeller Innenproduktraum ist, bedeutet nicht mehr und nicht weniger als dass V ein Vektorraum über $\mathbb{R}$ ist, und dass $\langle \cdot\,,\, \cdot\rangle$ ein Skalarprodukt auf V ist. Man sagt auch, dass der Vektorraum V mit dem Skalarprodukt $\langle \cdot\,,\, \cdot\rangle$ *versehen* und so zum Innenproduktraum $(V, \langle \cdot\,,\, \cdot\rangle)$ wird. Ähnlich wie bei den normierten Räumen, ist es üblich gewisse Vektorräume mit ganz bestimmten Normen standardmäßig zu versehen. Insbesondere trifft man die nachfolgende Vereinbarung.

Vereinbarung. Ist nichts gegenteiliges ausgesagt, so ist für alle $n \in \mathbb{N}$ der Vektorraum $\mathbb{R}^n$ mit dem euklidischen Skalarprodukt versehen.

Abschließend erwähnen wir noch eine wichtige Eigenschaft transponierter Matrizen in Bezug auf das euklidische Skalarprodukt.

Satz 1.2. Seien $m, n \in \mathbb{N}$ zwei natürliche Zahlen, sei $A \in \mathbb{R}^{m\times n}$ eine beliebige Matrix, und sei $A^{\mathrm{T}} \in \mathbb{R}^{n\times m}$ die Transponierte zu A. Dann gilt

$$x \cdot Ay = (A^{\mathrm{T}}x) \cdot y$$

für alle $x \in \mathbb{R}^m$ und alle $y \in \mathbb{R}^n$.

Übungsaufgaben

1. Geben seien die Vektoren

$$a := \begin{pmatrix} -2 \\ -5 \\ 3 \end{pmatrix}, \quad b := \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}, \quad c := \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}.$$

   Berechnen Sie $a \cdot b$, $a \cdot c$ und $b \cdot c$, sowie $a \cdot a$, $b \cdot b$ und $c \cdot c$.

2. Sei V ein reeller Vektorraum, und sei $\langle \cdot \, , \cdot \rangle$ ein Skalarprodukt auf V. Weisen Sie nach, dass die nachfolgenden Rechenregeln für alle $v, w, x, y \in V$ und alle $\alpha, \beta \in \mathbb{R}$ gelten:

   - $\langle \alpha v + \beta w \, , x \rangle = \alpha \langle v \, , x \rangle + \beta \langle w \, , x \rangle$.
   - $\langle v \, , \alpha x + \beta y \rangle = \alpha \langle v \, , x \rangle + \beta \langle v \, , y \rangle$.
   - $\langle v + w \, , x + y \rangle = \langle v \, , x \rangle + \langle v \, , y \rangle + \langle w \, , x \rangle + \langle w \, , y \rangle$.
   - $\langle -v \, , -w \rangle = \langle v \, , w \rangle$.
   - $\langle -v \, , w \rangle = -\langle v \, , w \rangle = \langle v \, , -w \rangle$.
   - $\langle v \, , 0 \rangle = \langle 0 \, , w \rangle = 0$.

1.3 Induzierte Normen

Bislang wurden Innenprodukträume und normierte Räume separat eingeführt. In diesem Abschnitt wird gezeigt, dass jeder Innenproduktraum auch als normierter Raum aufgefasst werden kann.

Satz und Definition 1.3 (induzierte Norm). Sei $(V, \langle \cdot\,,\,\cdot \rangle)$ ein reeller Innenproduktraum. Dann kann man V mit einer Norm $V \to \mathbb{R},\ v \mapsto \|v\|$ versehen, welche durch

$$\|v\| := \sqrt{\langle v\,,\,v \rangle}$$

für alle $v \in V$ definiert ist. Diese Norm wird als die vom Skalarprodukt $\langle \cdot\,,\,\cdot \rangle$ *induzierte Norm* auf V bezeichnet.

Da jedes Skalarprodukt auf einem Vektorraum gemäß Satz und Definition 1.3 auch eine Norm auf dem Vektorraum induziert, ist jeder Innenproduktraum in natürlicher Weise auch ein normierter Raum. Man versieht einen Innenproduktraum nämlich standardmäßig mit der Norm, die vom jeweiligen Skalarprodukt induziert wird.

Beispiele.

(a) Die reelle Betragsfunktion wird als Norm auf $\mathbb{R}$ von der gewöhnlichen Multiplikation auf $\mathbb{R}$ induziert, da bekanntlich $\sqrt{x^2} = |x|$ für alle $x \in \mathbb{R}$ gilt.

(b) Die euklidische Norm wird vom euklidischen Skalarprodukt induziert, d.h. es gilt

$$|x| = \sqrt{x \cdot x}$$

für alle $x \in \mathbb{R}^n$ und alle $n \in \mathbb{N}$, wie man leicht nachrechnet.

(c) Für jedes fest gewählte $n \in \mathbb{N}$ und jeden fest gewählten Vektor $w \in \mathbb{R}^n$ mit positiven Komponenten, wird die Norm $\|\cdot\|_w$ (siehe Beispiel (f) auf Seite 10) vom Skalarprodukt $\langle \cdot\,,\,\cdot \rangle_w$ (Beispiel (c) auf Seite 15) induziert. $\diamondsuit$

Für Normen, die von Skalarprodukten induziert werden, gelten eine Reihe wichtiger Gleichungen und Ungleichungen, die im nachfolgenden Satz zusammengetragen wurden.

Satz 1.4. Sei $(V, \langle \cdot\,,\cdot \rangle)$ ein Innenproduktraum über $\mathbb{R}$, und sei $\|\cdot\|$ die vom Skalarprodukt $\langle \cdot\,,\,\cdot \rangle$ induzierte Norm auf V. Dann gelten die folgenden Resultate:

(1) *Binomische Formeln:* Für alle $v, w \in V$ gilt

$$\begin{aligned}
\|v + w\|^2 &= \|v\|^2 + 2\langle v\,,\,w \rangle + \|w\|^2,\\
\|v - w\|^2 &= \|v\|^2 - 2\langle v\,,\,w \rangle + \|w\|^2,\\
\langle v - w\,,\,v + w \rangle &= \|v\|^2 - \|w\|^2.
\end{aligned}$$

(2) Für alle $v, w \in V$ gilt

$$|\langle v\,,\,w \rangle| \le \frac{1}{2}\|v\|^2 + \frac{1}{2}\|w\|^2.$$

(3) *Cauchy–Schwarzsche Ungleichung:* Für alle $v, w \in V$ gilt

$$|\langle v\,,\, w\rangle| \leq \|v\|\,\|w\|.$$

(4) *Dreiecksungleichung:* Für alle $v, w \in V$ gilt

$$\|v + w\| \leq \|v\| + \|w\|.$$

(5) *Parallelogrammgleichung:* Für alle $v, w \in V$ gilt

$$\|v + w\|^2 + \|v - w\|^2 = 2\|v\|^2 + 2\|w\|^2.$$

(6) Für alle $v, w \in V$ gilt

$$\langle v\,,\, w\rangle = \frac{1}{4}\big(\|v + w\|^2 - \|v - w\|^2\big).$$

Die Cauchy–Schwarzsche Ungleichung und die Dreiecksungleichung sind wichtige Hilfsmittel der Analysis. Hinsichtlich der Frage, ob eine gegebene Norm auf einem Vektorraum von einem Skalarprodukt induziert wird, kommt der Parallelogrammgleichung kommt eine besondere Bedeutung zu. Man kann nämlich zeigen, dass eine Norm genau dann von einem Skalarprodukt induziert wird, wenn sie die Parallelogrammgleichung erfüllt.

Auf reellen Innenprodukträumen kann man in der folgenden Weise Winkel zwischen zwei Vektoren definieren.

Definition (Winkel). Sei $(V, \langle\,\cdot\,,\,\cdot\,\rangle)$ ein Innenproduktraum über $\mathbb{R}$, und sei $\|\cdot\|$ die von $\langle\,\cdot\,,\,\cdot\,\rangle$ induzierte Norm auf V. Dann definiert man zu je zwei Vektoren $v, w \in V \setminus \{0\}$ die Zahl $\angle(v, w) \in [0, \pi]$ durch

$$\angle(v, w) := \arccos\left(\frac{\langle v\,,\, w\rangle}{\|v\|\,\|w\|}\right).$$

Die Zahl $\angle(v, w)$ heißt der *Winkel* zwischen v und w bezüglich $\langle\,\cdot\,,\,\cdot\,\rangle$.

Es muss betont werden, dass der Winkel zwischen zwei gegebenen Vektoren eines reellen Vektorraums V von dem Skalarprodukt abhängt, mit dem V versehen ist. Man betrachte beispielsweise die Vektoren

$$x := \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad y := \begin{pmatrix} 1 \\ \sqrt{3} \end{pmatrix}$$

des $\mathbb{R}^2$. Der Winkel zwischen beiden Vektoren bezüglich dem euklidischen Skalarprodukt beträgt

$$\angle(x, y) = \arccos\left(\frac{x \cdot y}{|x|\,|y|}\right) = \arccos\left(\frac{1}{2}\right) = \frac{\pi}{3} \quad (= 60^\circ).$$

Bezüglich dem Skalarprodukt $\langle\,\cdot\,,\,\cdot\,\rangle_w$ auf $\mathbb{R}^2$ mit $w := (3, 1)^{\mathrm{T}}$ (siehe Beispiel (c) auf Seite 15) hingegen, gilt

$$\angle(x, y) = \arccos\left(\frac{\langle x\,,\, y\rangle_w}{\|x\|_w\,\|y\|_w}\right) = \arccos\left(\frac{\sqrt{2}}{2}\right) = \frac{\pi}{4} \quad (= 45^\circ).$$

Man stellt fest, dass der Winkel zwischen je zwei Vektoren des $\mathbb{R}^2$ bzw. des $\mathbb{R}^3$ bezüglich euklidischen Skalarprodukt genau dem geometrischen Winkel entspricht, den man beispielsweise mit Hilfe eines Geodreiecks messen kann. Dies ist mit ein Grund dafür, weshalb man für jede natürliche Zahl $n \in \mathbb{N}$ den Vektorraum $\mathbb{R}^n$ standardmäßig mit dem euklidische Skalarprodukt versieht.

Übungsaufgaben

1. Zeigen Sie mit Hilfe der Cauchy–Schwarzschen Ungleichung, dass

$$(x_1\,y_1 + x_2\,y_2 + \cdots + x_n\,y_n)^2 \le \left(x_1^2 + x_2^2 + \cdots + x_n^2\right)\left(y_1^2 + y_2^2 + \cdots + y_n^2\right)$$

   für beliebige reelle Zahlen $x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_n \in \mathbb{R}$ gilt.

2. Geben Sie die Norm $\|\cdot\|_\# : \mathbb{R}^2 \to \mathbb{R}$ an, die vom Skalarprodukt $\langle \cdot\,,\,\cdot \rangle_\#$ induziert wird (siehe Beispiel (d) auf Seite 15). Berechnen Sie anschließend $\|a\|_\#$, $\|b\|_\#$ und $\|c\|_\#$, wobei

$$a := \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \quad b := \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad c := \frac{1}{2}\begin{pmatrix} \sqrt{2} \\ \sqrt{2} \end{pmatrix}.$$

3. Berechnen Sie für die Vektoren

$$a := \begin{pmatrix} 4 \\ 0 \end{pmatrix}, \quad b := \begin{pmatrix} \sqrt{2} \\ \sqrt{2} \end{pmatrix}, \quad c := \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

   die Winkel $\angle(a, a)$, $\angle(a, b)$, $\angle(-a, b)$, $\angle(a, c)$ und $\angle(b, c)$. Der Vektorraum $\mathbb{R}^2$ sei hierbei mit dem euklidischen Skalarprodukt versehen.

4. Sei $(V, \langle \cdot\,,\,\cdot \rangle)$ ein Innenproduktraum über $\mathbb{R}$. Zeigen Sie, dass $\angle(\alpha v, \beta w) = \angle(v, w)$ für alle Vektoren $v, w \in V \setminus \{0\}$ und alle positiven reellen Zahlen $\alpha, \beta > 0$ gilt.

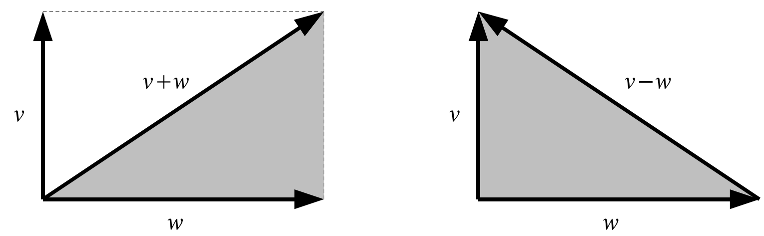

Abbildung 1.4: Die Vektoren v, w und $v + w$ bzw. $v - w$ bilden die Kanten eines rechtwinkligen Dreiecks. Für die entsprechenden Kantenlängen

1.4 Orthonormalbasen

Ein wichtiger Begriff, den man in Innenprodukträumen definieren kann, ist der Begriff der Orthogonalität.

Definition (Orthogonalität). Sei $(V, \langle \cdot \, , \, \cdot \rangle)$ ein Innenproduktraum über $\mathbb{R}$. Man sagt dass zwei Vektoren $v, w \in V$ zueinander *orthogonal* sind, wenn

$$\langle v \, , w \rangle = 0$$

gilt. In diesem Fall schreibt man $v \perp w$.

Bekanntlich gilt $\cos(\pi/2) = 0$. Daher sind zwei von Null verschiedene Vektoren w und v eines reellen Vektorraums V genau dann orthogonal zueinander bezüglich einem Skalarprodukt $\langle \cdot \, , \, \cdot \rangle$ auf V, wenn

$$\angle(v, w) = \arccos\left(\frac{\langle v \, , w \rangle}{\|v\| \, \|w\|} \right) = \frac{\pi}{2} \quad (= 90^\circ)$$

gilt, wobei $\| \cdot \|$ die von $\langle \cdot \, , \, \cdot \rangle$ induzierte Norm auf V bezeichnet. Die Vektoren sind also genau dann zueinander orthogonal, wenn sie im rechten Winkel zueinander stehen. Man muss sich jedoch immer wieder klar machen, dass der Winkel zwischen zwei Vekoren ebenso wie die Eigenschaft der Orthogonalität von dem Skalarprodukt abhängig sind, mit dem ein reeller Vektorraum versehen ist. Bezüglich dem euklidischen Skalarprodukt gilt beispielsweise, dass zwei von Null verschiedene Vektoren des $\mathbb{R}^2$ bzw. des $\mathbb{R}^3$ genau dann zueinander orthogonal sind, wenn sie im geometrischen Sinne senkrecht zueinander stehen. Versieht man die Räume $\mathbb{R}^2$ bzw. $\mathbb{R}^3$ mit anderen Skalarprodukten, gilt diese Aussage im allgemeinen jedoch nicht.

Ein wichtiges Resultat, welches für orthogonale Vektoren gilt, ist der berühmte *Satz des Pythagoras.* Man betrachte dazu auch die Abbildung 1.4.

Satz 1.5 (Pythagoras). Sei $(V, \langle \cdot \, , \cdot \rangle)$ ein Innenproduktraum über $\mathbb{R}$, und sei $\| \cdot \|$ die von $\langle \cdot \, , \, \cdot \rangle$ induzierte Norm. Seien außerdem $v, w \in V$ zwei zueinander orthogonale Vektoren. Dann gilt

$$\|v + w\|^2 = \|v\|^2 + \|w\|^2 = \|v - w\|^2.$$

Mit dem Satz des Pythagoras kann man insbesondere das folgende Lemma beweisen.

Lemma 1.6. Sei $(V, \langle \cdot\ ,\ \cdot \rangle)$ ein Innenproduktraum über $\mathbb{R}$, und seien $v_1, v_2, \ldots, v_n \in V \setminus \{0\}$ von Null verschiedene Vektoren, die paarweise zueinander orthogonal sind, d.h. es gelte $\langle v_i\ ,\ v_j \rangle = 0$ für alle $i, j \in \{1, 2, \ldots, n\}$ mit $i \neq j$. Dann sind diese Vektoren linear unabhängig.

Als nächstes betrachten wir Basen von endlichdimensionalen Vektorräumen, die aus zueinander orthogonalen Vektoren bestehen.

Definition (Orthogonalbasis). Sei U ein Untervektorraum eines reellen Vektorraums V, welcher mit einem Skalarprodukt $\langle \cdot\ ,\ \cdot \rangle$ versehen ist. Eine Basis $\{p_1, p_2, \ldots, p_m\}$ von U heißt *Orthogonalbasis* bezüglich $\langle \cdot\ ,\ \cdot \rangle$, wenn $\langle p_i\ ,\ p_j \rangle = 0$ für alle $i, j \in \{1, 2, \ldots, m\}$ mit $i \neq j$ gilt, d.h. wenn die Basisvektoren paarweise orthogonal zueinander sind.

Im Rest dieses Abschnitts wenden wir uns eine speziellen Klasse von Orthogonalbasen, den so genannten *Orthonormalbasen*, zu.

Definition (Orthonormalbasis). Sei U ein Untervektorraum eines reellen Vektorraums V, welcher mit einem Skalarprodukt $\langle \cdot\ ,\ \cdot \rangle$ versehen ist. Eine Basis $\{q_1, q_2, \ldots, q_m\}$ von U heißt *Orthonormalbasis* (abgekürzt *ONB*) bezüglich $\langle \cdot\ ,\ \cdot \rangle$, wenn

$$\langle q_i\ ,\ q_j \rangle = \begin{cases} 1 & \text{falls } i = j, \\ 0 & \text{falls } i \neq j \end{cases} \qquad \text{für alle } i, j = 1, 2, \ldots, m$$

gilt.

Man vergegenwärtige sich noch einmal, dass für alle Vektoren v eines reellen Vektorraums V, welcher mit einem Skalarprodukt $\langle \cdot\ ,\ \cdot \rangle$ versehen ist, die Gleichung

$$\langle v\ ,\ v \rangle = \|v\|^2$$

gilt, wobei $\|\cdot\|$ die von $\langle \cdot\ ,\ \cdot \rangle$ induzierte Norm auf V bezeichnet. Daher ist die Basis $\{q_1, q_2, \ldots, q_m\}$ eines Untervektorraums genau dann eine Orthonormalbasis, wenn sie eine *Orthogonal*basis ist, und wenn alle Basisvektoren bezüglich der induzierten Norm *auf Eins normiert* sind, d.h. wenn

$$\|q_i\| = 1 \qquad \text{für alle } i = 1, 2, \ldots, m$$

gilt. Nachfolgend geben wir einige Beispiele für Orthonormalbasen endlichdimensionale Vektorräume über $\mathbb{R}$ an.

Beispiele.

(a) Für jede natürliche Zahl $n \in \mathbb{N}$ ist die so genannte *Standardbasis* $\{e^{(1)}, e^{(2)}, \ldots, e^{(n)}\}$ des $\mathbb{R}^n$, welche durch

$$e^{(1)} := \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad e^{(2)} := \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \ldots, \quad e^{(n)} := \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

gegeben ist, eine Orthonormalbasis bezüglich dem euklidischen Skalarprodukt.

(b) Die Vektoren

$$v^{(1)} := \frac{\sqrt{2}}{2}\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad v^{(2)} := \frac{\sqrt{3}}{3}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad v^{(3)} := \frac{\sqrt{6}}{6}\begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$$

bilden eine Orthonormalbasis des $\mathbb{R}^3$ bezüglich dem euklidischen Skalarprodukt.

Für Orthonormalbasen gelten eine Reihe wichtiger Resultate, welche im folgenden Satz zusammengefasst sind.

Satz 1.7. Sei $(V, \langle \cdot\,,\, \cdot \rangle)$ ein endlichdimensionaler, reeller Innenproduktraum, und sei $\|\cdot\|$ die von $\langle \cdot\,,\, \cdot \rangle$ induzierte Norm. Sei außerdem $\{q_1, q_2, \ldots, q_n\}$ eine Orthonormalbasis von V. Dann gelten die folgenden Aussagen.

(1) *Orthogonalentwicklung:* Für alle $v \in V$ gilt

$$v = \langle q_1\,,\, v\rangle q_1 + \langle q_2\,,\, v\rangle q_2 + \cdots + \langle q_n\,,\, v\rangle q_n.$$

(2) *Parsevalsche Gleichung:* Für alle $v \in V$ gilt

$$\|v\|^2 = |\langle q_1\,,\, v\rangle|^2 + |\langle q_2\,,\, v\rangle|^2 + \cdots + |\langle q_n\,,\, v\rangle|^2.$$

(3) *Äquidistanz:* Für alle $i, j \in \{1, 2, \ldots, n\}$ gilt

$$\|q_i - q_j\| = \begin{cases} 0 & \text{falls } i = j, \\ \sqrt{2} & \text{falls } i \neq j. \end{cases}$$

Abschließend soll hier noch diskutiert werden, wie man zu einem gegebenen Vektorraum eine Orthonormalbasis konstruieren kann. Wichtigstes Hilfsmittel einer solchen Konstruktion ist das so genannte *Gram–Schmidtsche Orthogonalsierungsverfahren*, welches durch das nachfolgende Lemma motiviert ist.

Lemma 1.8. Sei $(V, \langle \cdot\,,\, \cdot \rangle)$ ein reeller Innenproduktraum und $\{p_1, p_2, \ldots, p_m\} \subseteq V \setminus \{0\}$ eine Menge von Vektoren, die paarweise zueinander orthogonal sind. Sei außerdem $v \in V$ ein beliebiger Vektor. Dann ist der Vektor

$$p := v - \frac{\langle p_1\,,\, v\rangle}{\langle p_1\,,\, p_1\rangle} p_1 - \frac{\langle p_2\,,\, v\rangle}{\langle p_2\,,\, p_2\rangle} p_2 - \cdots - \frac{\langle p_m\,,\, v\rangle}{\langle p_m\,,\, p_m\rangle} p_m$$

zu jedem Vektor der Menge $\{p_1, p_2, \ldots, p_m\}$ orthogonal.

Algorithmus (Gram–Schmidtsches Orthogonalisierungsverfahren).
Sei $(V, \langle \cdot\,,\, \cdot \rangle)$ ein reeller Innenproduktraum, und seien $v_1, v_2, \ldots, v_m \in V$ linear unabhängige Vektoren. Berechnet man die Vektoren $p_1, p_2, \ldots, p_m \in V$ gemäß

$$p_1 := v_1,$$

$$p_2 := v_2 - \frac{\langle p_1\,,\, v_2\rangle}{\langle p_1\,,\, p_1\rangle} p_1,$$

$$p_3 := v_3 - \frac{\langle p_1\,,v_3\rangle}{\langle p_1\,,p_1\rangle}\,p_1 - \frac{\langle p_2\,,v_3\rangle}{\langle p_2\,,p_2\rangle}\,p_2,$$
$$\vdots$$
$$p_m := v_m - \frac{\langle p_1\,,v_m\rangle}{\langle p_1\,,p_1\rangle}\,p_1 - \frac{\langle p_2\,,v_m\rangle}{\langle p_2\,,p_2\rangle}\,p_2 - \cdots - \frac{\langle p_{m-1}\,,v_m\rangle}{\langle p_{m-1}\,,p_{m-1}\rangle}\,p_{m-1},$$

dann ist $\{p_1, p_2, \ldots, p_m\}$ eine Orthogonalbasis von $\operatorname{span}\{v_1, v_2, \ldots, v_m\}$.

Mit Hilfe des Gram–Schmidtschen Orthogonalisierungsverfahrens kann man aus einer gegebenen Basis $\{v_1, v_2, \ldots, v_m\}$ eines Vektorraums eine Orthogonalbasis $\{p_1, p_2, \ldots, p_m\}$ desselben Vektorraums konstruieren. Normiert man die Vektoren dieser Orthogonalbasis auf Eins, so erhält man eine Orthonormalbasis des Vektorraums.

Beispiel. Gegeben seien die linear unabhängigen Vektoren

$$v^{(1)} := \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix}, \quad v^{(2)} := \begin{pmatrix} -1 \\ 2 \\ 2 \end{pmatrix}.$$

Gesucht ist eine Orthonormalbasis $\{q^{(1)}, q^{(2)}\}$ von $\operatorname{span}\{v^{(1)}, v^{(2)}\}$ bezüglich des euklidischen Skalarprodukts.

- *Orthogonalisierung* (Gram–Schmidtsches Orthogonalisierungsverfahren):

$$p^{(1)} := v^{(1)} = \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix},$$
$$p^{(2)} := v^{(2)} - \frac{p^{(1)} \cdot v^{(2)}}{p^{(1)} \cdot p^{(1)}}\,p^{(1)} = \begin{pmatrix} -1 \\ 2 \\ 2 \end{pmatrix} - \frac{(-2)}{6}\begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1/3 \\ 7/3 \\ 5/3 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} -1 \\ 7 \\ 5 \end{pmatrix}.$$

- *Normierung:*

$$q^{(1)} := \frac{1}{|p^{(1)}|}\,p^{(1)} = \frac{\sqrt{6}}{6}\begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix},$$
$$q^{(2)} := \frac{1}{|p^{(2)}|}\,p^{(2)} = \frac{\sqrt{3}}{5} \cdot \frac{1}{3}\begin{pmatrix} -1 \\ 7 \\ 5 \end{pmatrix} = \frac{\sqrt{3}}{15}\begin{pmatrix} -1 \\ 7 \\ 5 \end{pmatrix}.$$

$\Diamond$

Zum Abschluss dieses Abschnitts wenden wird uns nun noch einer besonderen Klasse von Funktionen zu, die eine besondere Rolle in vielen Teilbereichen der Mathematik spielen.

Definition (Funktional). Sei V ein Vektorraum über $\mathbb{K}$. Eine Funktion $f : V \to \mathbb{K}$ heißt ein *Funktional* auf V.

Ein Funktional ist also eine Funktion, welche die Vektoren eines Vektorraums auf Elemente des Körpers abbildet, über dem der Vektorraum definiert ist. Für lineare Funktionale auf den Vektorräumen $\mathbb{R}^n$ gilt ein wichtiger Satz, der als *Rieszscher Darstellungssatz* bekannt ist.

Satz 1.9 (Rieszscher Darstellungssatz). Sei $n \in \mathbb{N}$ eine natürliche Zahl, $f : \mathbb{R}^n \to \mathbb{R}$ ein lineares Funktional auf $\mathbb{R}^n$ und $\langle \cdot\, , \cdot \rangle$ ein Skalarprodukt auf $\mathbb{R}^n$. Dann existiert ein eindeutig bestimmter Vektor $v_f \in \mathbb{R}^n$, so dass

$$f(x) = \langle v_f\, , x \rangle$$

für alle $x \in \mathbb{R}^n$ gilt.

Der Rieszsche Darstellungssatz besagt, dass jedes lineare Funktional auf $\mathbb{R}^n$ durch ein beliebig gewähltes Skalarprodukt dargestellt werden kann. Insbesondere existiert zu jedem linearen Funktional $f : \mathbb{R}^n \to \mathbb{R}$ ein eindeutig bestimmter Vektor $v_f \in V$, so dass $f(x) = v_f \cdot x$ für alle $x \in \mathbb{R}^n$ gilt.

Übungsaufgaben

1. Bestimmen Sie die Orthogonalentwicklungen der Vektoren

$$x := \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad y := \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad z := \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$$

   bezüglich der Orthormalbasis $\{q^{(1)}, q^{(2)}, q^{(3)}\}$ des $\mathbb{R}^3$, welche durch

$$q^{(1)} := \frac{\sqrt{3}}{3} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad q^{(2)} := \frac{\sqrt{2}}{2} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad q^{(3)} := \frac{\sqrt{6}}{6} \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$$

   gegeben ist.

2. Bestimmen Sie zwei Orthonormalbasen des $\mathbb{R}^3$ bezüglich dem euklidischen Skalarprodukt, indem Sie die folgenden Basen unter Verwendung des Gram–Schmidtschen Orthogonalisierungsverfahrens orthonormalsieren:

$$\left\{ \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 5 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} \right\}, \quad \left\{ \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \right\}.$$

3. Bestimmen Sie eine Orthonormalbasis des $\mathbb{R}^2$ bezüglich $\langle \cdot\, , \cdot \rangle_\#$ (siehe Beispiel (d) auf Seite 15), indem Sie die Standardbasis des $\mathbb{R}^2$ entsprechend orthonormalisieren.

4. Sei $n \in \mathbb{N}$ eine natürliche Zahl. Für jedes $j = 1, 2, \ldots, n$ sei das lineare Funktional $p_j : \mathbb{R}^n \to \mathbb{R}$ durch $p_j(x) = x_j$ für alle $x = (x_1, x_2, \ldots, x_n)^{\mathrm{T}} \in \mathbb{R}^n$ definiert. Geben Sie für jedes $j = 1, 2, \ldots, n$ einen Vektor $v_{p_j} \in \mathbb{R}^n$ an, so dass $p_j(x) = v_{p_j} \cdot x$ für alle $x \in \mathbb{R}^n$ gilt.

1.5 Definitheit symmetrischer Matrizen

Definition (symmetrische Matrix). Zu einer gegebenen natürlichen Zahl $n \in \mathbb{N}$ heißt eine Matrix $A \in \mathbb{R}^{n\times n}$ *symmetrisch*, wenn $A^{\mathrm{T}} = A$ gilt. Die Menge aller symmetrischen $n \times n$-Matrizen mit reellen Matrixkomponenten wird gelegentlich mit $\mathbb{R}^{n\times n}_{\mathrm{sym}}$ bezeichnet.

Beispiele.

(a) Die folgenden Matrizen sind symmetrisch:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}, \quad \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{pmatrix}.$$

(b) Die folgenden Matrizen sind nicht symmetrisch:

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 0 \\ 3 & 1 & 2 \\ 0 & 2 & 1 \end{pmatrix}.$$

Ist $n \in \mathbb{N}$ eine natürliche Zahl und $A \in \mathbb{R}^{n\times n}_{\mathrm{sym}}$ eine symmetrische Matrix, so gilt nach Satz 1.2 insbesondere

$$x \cdot Ay = (Ax) \cdot y$$

für alle Vektoren $x, y \in \mathbb{R}^n$. Der folgende Satz macht weiterhin Aussagen darüber, welche Matrixoperationen die Symmetrie einer Matrix erhalten.

Satz 1.10. Sei $n \in \mathbb{N}$ eine beliebige natürliche Zahl. Dann gelten für alle symmetrischen Matrizen $A, B \in \mathbb{R}^{n\times n}_{\mathrm{sym}}$ und für alle reellen Zahlen $\alpha \in \mathbb{R}$ die folgenden Aussagen

(1) $A + B \in \mathbb{R}^{n\times n}_{\mathrm{sym}}$.

(2) $\alpha A \in \mathbb{R}^{n\times n}_{\mathrm{sym}}$.

(3) $A^{-1} \in \mathbb{R}^{n\times n}_{\mathrm{sym}}$, falls A regulär ist.

(4) $A^k \in \mathbb{R}^{n\times n}_{\mathrm{sym}}$ für alle $k \in \mathbb{N}_0$.

Die Aussagen in Teil (1) und (2) von Satz 1.10 implizieren, dass die Menge $\mathbb{R}^{n\times n}_{\mathrm{sym}}$ ein Untervektorraum von $\mathbb{R}^{n\times n}$ ist. Die Menge $\mathbb{R}^{n\times n}_{\mathrm{sym}}$ ist jedoch keine Gruppe bezüglich der Matrixmultiplikation, da das Produkt zweier symmetrischer Matrizen im allgemeinen keine symmetrische Matrix ist. Man betrachte etwa die beiden Matrix $A, B \in \mathbb{R}^{2\times 2}_{\mathrm{sym}}$, welche durch

$$A := \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad B := \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

gegeben sind. Für das Produkt der beiden Matrizen gilt

$$AB = \begin{pmatrix} 0 & 0 \\ 1 & 2 \end{pmatrix} \neq \begin{pmatrix} 0 & 1 \\ 0 & 2 \end{pmatrix} = (AB)^{\mathrm{T}},$$

was $AB \notin \mathbb{R}^{2\times 2}_{\mathrm{sym}}$ impliziert.

Als nächstes führen wir das Konzept der *positiven bzw. negativen Definitheit* für quadratische Matrizen ein. Dieses spielt eine wichtige Rolle bei der Identifikation lokaler Extrema von Funktionalen auf $\mathbb{R}^n$ (siehe Abschnitt 6.6).

Definition (positiv definite Matrix). Sei $n \in \mathbb{N}$ eine natürliche Zahl. Eine quadratische Matrix $A \in \mathbb{R}^{n\times n}$ heißt *positiv definit*, wenn $x \cdot Ax > 0$ für alle $x \in \mathbb{R}^n \setminus \{0\}$ gilt.

Definition (negativ definite Matrix). Sei $n \in \mathbb{N}$ eine beliebige natürliche Zahl. Eine quadratische Matrix $A \in \mathbb{R}^{n\times n}$ heißt *negativ definit*, wenn $-A$ positiv definit ist, d.h. wenn $x \cdot Ax < 0$ für alle $x \in \mathbb{R}^n \setminus \{0\}$ gilt.

Im allgemeinen ist es nicht leicht zu erkennen, ob eine Matrix positiv bzw. negativ definit ist. Für symmetrische Matrizen kann man jedoch ein hinreichendes Kriterium für die positive Definitheit wie auch für die negative Definitheit angeben, welches auf den so genannten *Hauptminoren* der Matrix beruht.

Definition (Hauptminor). Sei $n \in \mathbb{N}$ eine natürliche Zahl und $A \in \mathbb{R}^{n\times n}$ eine quadratische Matrix. Für jede natürliche Zahl $k \in \{1, 2, \ldots, n\}$ definiert man die reelle Zahl $A_{[k]} \in \mathbb{R}$ durch

$$A_{[k]} := \det \begin{pmatrix} A_{11} & A_{12} & \ldots & A_{1k} \\ A_{21} & A_{22} & \ldots & A_{2k} \\ \vdots & \vdots & & \vdots \\ A_{k1} & A_{k2} & \ldots & A_{kk} \end{pmatrix}.$$

Die Zahl $A_{[k]}$ wird der k-te *Hauptminor* von A genannt.

Offenbar gilt $A_{[1]} = A_{11}$ und $A_{[n]} = \det A$ für jede natürliche Zahl $n \in \mathbb{N}$ und jede quadratische Matrix $A \in \mathbb{R}^{n\times n}$. Alle übrigen Hauptminoren berechnet man, indem man die Determinante einer quadratischen „Teilmatrix“ von A berechnet, welche die Matrixkomponente A_{11} enthält. Man betrachte dazu die nachfolgenden Beispiele.

Beispiele.

(a) Die beiden Hauptminoren der Matrix

$$A := \begin{pmatrix} -4 & 1 \\ -2 & -1 \end{pmatrix}$$

lauten

$$A_{[1]} = \det \begin{pmatrix} -4 \end{pmatrix} = -4, \quad A_{[2]} = \det \begin{pmatrix} -4 & 1 \\ -2 & -1 \end{pmatrix} = 6.$$

(b) Die drei Hauptminoren der Matrix

$$B := \begin{pmatrix} 2 & 1 & 0 \\ -2 & 3 & 1 \\ -1 & 0 & 2 \end{pmatrix}$$

sind durch

$$B_{[1]} = \det \begin{pmatrix} 2 \end{pmatrix} = 2, \quad B_{[2]} = \det \begin{pmatrix} 2 & 1 \\ -2 & 3 \end{pmatrix} = 8, \quad B_{[3]} = \det \begin{pmatrix} 2 & 1 & 0 \\ -2 & 3 & 1 \\ -1 & 0 & 2 \end{pmatrix} = 15$$

gegeben. ◊

Es sollte noch erwähnt werden, dass in der mathematischen Fachliteratur keine einheitliche Bezeichnungsweise für die Hauptminoren einer Matrix existiert. Auch die hier gewählte Schreibweise $A_{[k]}$ ist keinesfalls allgemein üblich.

Als nächstes zeigen wir, wie man mit Hilfe von Hauptminoren die positive bzw. die negative Definitheit einer symmetrischen Matrix nachweisen kann.

Satz 1.11 (**Hauptminorenkriterium**). Sei $n \in \mathbb{N}$ eine natürliche Zahl und $A \in \mathbb{R}^{n\times n}_{\mathrm{sym}}$ eine symmetrische Matrix. Dann gelten die folgenden Aussagen.

(1) A ist genau dann positiv definit, wenn alle Hauptminoren von A positiv sind, d.h. wenn
$$\operatorname{sgn}(A_{[k]}) = 1$$
für alle $k = 1, 2, \ldots, n$ gilt.

(2) A ist genau dann negativ definit, wenn die Vorzeichen der Hauptminoren von A mit Minus beginnend alternieren, d.h. wenn
$$\operatorname{sgn}(A_{[k]}) = (-1)^k$$
für alle $k = 1, 2, \ldots, n$ gilt.

In den nachfolgenden Beispielen wird gezeigt, wie man das Hauptminorenkriterium anwendet.

Beispiele.

(a) Für die Hauptminoren der symmetrischen Matrix
$$A := \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}$$
gilt $A_{[1]} = 2 > 0$, $A_{[2]} = 3 > 0$ und $A_{[3]} = 4 > 0$. Daher ist A positiv definit.

(b) Für die Hauptminoren der symmetrischen Matrix
$$B := \begin{pmatrix} -4 & 0 & 1 \\ 0 & -2 & 0 \\ 1 & 0 & -3 \end{pmatrix}$$
gilt $B_{[1]} = -4 < 0$, $B_{[2]} = 8 > 0$ und $B_{[3]} = -22 < 0$. Daher ist B negativ definit.

(c) Für die Hauptminoren der symmetrischen Matrix
$$C := \begin{pmatrix} 3 & 2 & 3 \\ 2 & -2 & 0 \\ 3 & 0 & 1 \end{pmatrix}$$
gilt $C_{[1]} = 3 > 0$, $C_{[2]} = -10 < 0$ und $C_{[3]} = 8 > 0$. Die Matrix C ist daher weder positiv definit noch negativ definit. Tatsächlich gilt $e^{(1)} \cdot Ce^{(1)} = 3 > 0$ und $e^{(2)} \cdot Ce^{(2)} = -2 < 0$, wobei $e^{(1)} = (1, 0, 0)^{\mathrm{T}}$ und $e^{(2)} = (0, 1, 0)^{\mathrm{T}}$.

Der Vollständigkeit halber geben wir zum Abschluss diese Abschnitts noch das folgende Resultat an, welches es einem erlaubt, auch nicht-symmetrische Matrizen mit dem Hauptminorenkriterium auf positive oder negative Definitheit hin zu untersuchen.

Lemma 1.12. Eine quadratische Matrix $A \in \mathbb{R}^{n\times n}$ ist genau dann positiv definit bzw. negativ definit, wenn die symmetrische Matrix $A + A^{\mathrm{T}}$ positiv definit bzw. negativ definit ist.

Übungsaufgaben

1. Sei $n \in \mathbb{N}$ eine natürliche Zahl, und sei $A \in \mathbb{R}^{n\times n}$ eine quadratische Matrix. Zeigen Sie, dass die Matrizen $A^{\mathrm{T}}A$ und AA^{T} symmetrisch sind.

2. Zeigen Sie, dass für jede natürliche Zahl $n \in \mathbb{N}$ eine Matrix $A \in \mathbb{R}^{n\times n}$ genau dann symmetrisch ist, wenn
$$A = \frac{1}{2}\big(A + A^{\mathrm{T}}\big)$$
gilt.

3. Untersuchen Sie die folgenden symmetrischen Matrizen auf positive oder negative Definitheit.
$$A := \begin{pmatrix} -4 & 2 & 0 \\ 2 & -3 & -2 \\ 0 & -2 & -1 \end{pmatrix}, \quad B := \begin{pmatrix} 1 & 1 & 2 \\ 1 & 3 & -1 \\ 2 & -1 & 5 \end{pmatrix}, \quad C := \begin{pmatrix} 4 & -1 & -1 \\ -1 & 4 & -1 \\ -1 & -1 & 4 \end{pmatrix}.$$

4. Sei $n \in \mathbb{N}$ eine natürliche Zahl, und sei $A \in \mathbb{R}^{n\times n}$ eine symmetrische und positiv definite Matrix. Zeigen Sie, dass dann die Funktion $\langle \cdot\,,\cdot \rangle_A : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, welche durch $\langle x\,, y\rangle_A := x \cdot Ay$ für alle $x, y \in \mathbb{R}^n$ definiert ist, ein Skalarprodukt auf $\mathbb{R}^n$ ist.

Lernzielkontrolle

Nach dem Durcharbeiten dieses Kapitels sollten Sie ...

... die definierenden Eigenschaften von Normen und Skalarprodukten kennen.

... wissen, wie *offene Kugeln*, *abgeschlossene Kugeln* und *Sphären* in einem normierten Raum definiert sind.

... wissen, auf welche Weise ein Skalarprodukt eine Norm induziert.

... wichtige Gleichungen und Ungleichungen für Normen und Skalarprodukte kennen. Sie sollten insbesondere die *Dreiecksungleichung*, die *umgekehrte Dreiecksungleichung* und die *Cauchy–Schwarzsche Ungleichung* kennen.

... das *euklidische Skalarprodukt*, die *euklidische Norm*, die *Betragssummennorm* und die *Maximumnorm* kennen.

... wissen, dass die euklidische Norm vom euklidischen Skalarprodukt induziert wird.

... wissen, was eine *Orthonormalbasis* ist.

... das *Gram–Schmidtsche Orthogonalisierungsverfahren* anwenden können.

... wissen, was eine *positiv definite* und was eine *negativ definite*, *symmetrische* Matrix ist.

... die *Hauptminoren* einer quadratischen Matrix berechnen können.

... das *Hauptminorenkriterium* für die positive bzw. die negative Definitheit einer symmetrischen Matrix kennen.