21. Euklidische und unitäre Räume

Wir wollen uns nun mit einem ganz anderen Thema beschäftigen, nämlich wie man Längen von Vektoren und Winkel zwischen zwei Vektoren berechnen (und überhaupt erst einmal definieren) kann. Zur Motivation betrachten wir dazu zunächst einmal den sehr einfachen Fall des Vektorraums $\mathbb{R}^2$, in dem sich diese beiden Fragen mit Hilfe von Elementargeometrie und Schulmathematik leicht beantworten lassen.

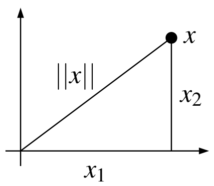
Beispiel 21.1 (Längen und Winkel in $\mathbb{R}^2$)**.** Wie ihr sicher aus der Schule wisst, ist das wesentliche Hilfsmittel für die Längen- und Winkelmessung in $\mathbb{R}^2$ das sogenannte *Skalarprodukt*

$$\langle x,y\rangle = x_1y_1 + x_2y_2 \in \mathbb{R} \qquad \text{für } x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},\ y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \in \mathbb{R}^2.$$

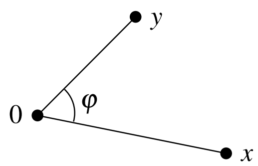
So ergibt sich z. B. wie im Bild unten links dargestellt aus dem Satz des Pythagoras, dass die Länge eines Vektors $x \in \mathbb{R}^2$ durch den Ausdruck

$$\sqrt{x_1^2 + x_2^2} = \sqrt{\langle x,x\rangle}$$

gegeben ist, den wir im Folgenden kurz als $||x||$ schreiben werden.



$||x|| = \sqrt{\langle x,x\rangle}$



$\frac{y}{||y||} = e^{i\varphi} \cdot \frac{x}{||x||}$

Wollen wir den Winkel φ zwischen zwei Vektoren $x,y \in \mathbb{R}^2\backslash\{0\}$ wie im Bild oben rechts berechnen, betrachten wir dazu am besten zunächst einmal die Vektoren $\frac{x}{||x||}$ und $\frac{y}{||y||}$, die in die gleiche Richtung wie x bzw. y zeigen, aber die Länge 1 haben. Fassen wir dann $x = x_1 + ix_2$ und $y = y_1 + iy_2$ als Elemente der komplexen Ebene $\mathbb{C} = \mathbb{R}^2$ auf, so folgt aus den Bemerkungen 6.5 und 9.11, dass

$$\frac{y}{||y||} = e^{i\varphi} \cdot \frac{x}{||x||}$$

ist, da sich die Winkel bei der komplexen Multiplikation addieren und $e^{i\varphi}$ eine Zahl mit Winkel φ und Betrag 1 ist. Einfache Umformungen in $\mathbb{C}$ ergeben nun wegen $||x||^2 = \bar{x}x$

$$\frac{\bar{x}y}{\bar{x}x} \cdot \frac{||x||}{||y||} = e^{i\varphi} \qquad \Rightarrow \qquad \frac{\bar{x}y}{||x|| \cdot ||y||} = \cos\varphi + i\sin\varphi.$$

Wegen $\bar{x}y = (x_1 - ix_2)(y_1 + iy_2) = x_1y_1 + x_2y_2 + i(x_1y_2 - x_2y_1)$ besagt der Realteil dieser Gleichung

$$\frac{x_1y_1 + x_2y_2}{||x|| \cdot ||y||} = \cos\varphi,$$

woraus wir mit der obigen Definition des Skalarprodukts folgern, dass

$$\varphi = \arccos \frac{\langle x,y\rangle}{||x|| \cdot ||y||}$$

ist (beachte hierbei, dass der Arkuskosinus nur Werte zwischen 0 und π zurückliefert und aufgrund der Symmetrien der Kosinusfunktion damit den *unorientierten* Winkel zwischen x und y ergibt).

Sowohl Längen als auch Winkel lassen sich damit durch das Skalarprodukt ausdrücken. Wenn wir diese beiden Konzepte auch in anderen Vektorräumen definieren wollen, sollten wir den Begriff des Skalarprodukts also auf beliebige Vektorräume verallgemeinern.

Das Problem dabei ist jedoch, dass die Formel $\langle x,y\rangle = x_1y_1+x_2y_2$ (oder eine entsprechend verallgemeinerte Version für höhere Dimensionen) explizit die Koordinaten der beiden Vektoren x und y benutzt. In einem allgemeinen Vektorraum gäbe es solche Koordinaten aber erst nach Wahl einer Basis — und die Formel würde natürlich auch unterschiedliche Ergebnisse liefern, wenn man die Koordinaten bezüglich verschiedener Basen nehmen würde. Wir schließen daraus, dass es in einem allgemeinen Vektorraum *kein natürlich definiertes* Skalarprodukt gibt, sondern dass ein Skalarprodukt eine *Zusatzstruktur* darstellt, die man zusätzlich zum Vektorraum erst einmal festlegen muss, bevor man mit konkreten Rechnungen anfangen kann.

Wir müssen diese Zusatzstruktur also zunächst erst einmal genauer definieren, d. h. konkret angeben, welche Eigenschaften ein Skalarprodukt haben soll. Klar ist, dass wir zwei Elementen eines K-Vektorraums V ein Element des zugrunde liegenden Körpers K zuordnen wollen, also formal eine Abbildung von $V\times V$ nach K betrachten müssen. Mit solchen Abbildungen wollen wir uns nun zunächst beschäftigen.

21.A Bilinearformen

Die erste wichtige Eigenschaft eines Skalarprodukts ist, dass es linear in beiden Vektoren ist. Derartige Abbildungen bezeichnet man als Bilinearformen.

Definition 21.2 (Bilinearformen)**.** Es sei V ein K-Vektorraum. Eine **Bilinearform** auf V ist eine Abbildung $b\colon V\times V\to K$, die in beiden Komponenten eine lineare Abbildung gemäß Definition 13.16 ist, d. h. für die für alle $x_1,x_2,x,y_1,y_2,y\in V$ und $\lambda\in K$ die Eigenschaften

$$\begin{aligned} b(x_1+x_2,y) &= b(x_1,y)+b(x_2,y),\\ b(\lambda x,y) &= \lambda\, b(x,y),\\ b(x,y_1+y_2) &= b(x,y_1)+b(x,y_2),\\ b(x,\lambda y) &= \lambda\, b(x,y) \end{aligned}$$

gelten. Wie man leicht nachprüft, ist die Menge aller Bilinearformen auf V ein Unterraum von $\mathrm{Abb}(V\times V,K)$ (siehe Beispiel 13.3 (d)) und damit ein K-Vektorraum. Wir bezeichnen ihn mit $\mathrm{BLF}(V)$.

Beispiel 21.3.

(a) Die Abbildung

$$b\colon \mathbb{R}^2\times\mathbb{R}^2\to\mathbb{R},\ \left(\begin{pmatrix}x_1\\x_2\end{pmatrix},\begin{pmatrix}y_1\\y_2\end{pmatrix}\right)\mapsto x_1y_1+x_1y_2+x_2y_1+4x_2y_2$$

ist offensichtlich eine Bilinearform: Hält man y_1 und y_2 fest, so ist der gegebene Ausdruck eine lineare Abbildung in x_1 und x_2, und umgekehrt. Hingegen ist

$$b\colon \mathbb{R}^2\times\mathbb{R}^2\to\mathbb{R},\ \left(\begin{pmatrix}x_1\\x_2\end{pmatrix},\begin{pmatrix}y_1\\y_2\end{pmatrix}\right)\mapsto x_1y_1+x_1+y_1$$

keine Bilinearform: Da lineare Abbildungen nach Bemerkung 13.18 (b) stets 0 auf 0 abbilden, kann b wegen

$$b\left(\begin{pmatrix}0\\0\end{pmatrix},\begin{pmatrix}1\\0\end{pmatrix}\right)=1$$

bei festgehaltener zweiter Komponente y nicht linear im ersten Argument x sein.

(b) Ist $A\in\mathrm{Mat}(n\times n,K)$ eine quadratische Matrix, so ist

$$b\colon K^n\times K^n\to K,\ b(x,y)=x^{\mathsf{T}}A\,y \qquad (*)$$

nach den Rechenregeln für Matrizen aus Lemma 16.7 eine Bilinearform — beachte, dass das Ergebnis hierbei als Produkt dreier Matrizen der Größen $1\times n$, $n\times n$ und $n\times 1$ eine

1×1-Matrix, also ein Element von K ist. Ist $A = (a_{i,j})_{i,j}$ und sind $x_1, \dots, x_n$ und $y_1, \dots, y_n$ die Koordinaten von x und y, so ist eine alternative Schreibweise für $(*)$ nach Definition 16.5

$$b(x,y) = (x_1 \ \cdots \ x_n) \cdot \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \sum_{i,j=1}^{n} x_i a_{i,j} y_j.$$

Wir wollen nun sehen, dass man in der Tat sogar jede mögliche Bilinearform auf K^n auf diese Art aus einer eindeutig bestimmten Matrix erhalten kann. Dies besagt der folgende Satz, der völlig analog zu Satz 16.11 über den Zusammenhang zwischen linearen Abbildungen und Matrizen ist, und der damit letztlich besagt, dass Bilinearformen auf K^n und $n \times n$-Matrizen über K „im Prinzip dasselbe" sind.

Satz und Definition 21.4 (Bilinearformen auf K^n und Matrizen)**.** *Für alle $n \in \mathbb{N}$ ist die Abbildung*

$$\mathrm{Mat}(n \times n, K) \to \mathrm{BLF}(K^n), \quad A \mapsto b_A \qquad \text{mit } b_A(x,y) := x^{\mathsf{T}} A\, y$$

ein K-Vektorraumisomorphismus mit Umkehrabbildung

$$\mathrm{BLF}(K^n) \to \mathrm{Mat}(n \times n, K), \quad b \mapsto A_b \qquad \text{mit } A_b := (b(e_i, e_j))_{i,j}.$$

*Man nennt A_b die **Gramsche Matrix** von b.*

Beweis. Nach Beispiel 21.3 (b) ist b_A wirklich eine Bilinearform auf K^n. Wir zeigen, dass die beiden angegebenen Abbildungen invers zueinander sind: Starten wir mit einer Matrix $A = (a_{i,j})_{i,j}$, bilden die zugehörige Bilinearform $b_A(x,y) = x^{\mathsf{T}} A\, y$ und dazu wieder die Gramsche Matrix, so erhalten wir

$$A_{b_A} = (b_A(e_i, e_j))_{i,j} = (e_i^{\mathsf{T}} A\, e_j)_{i,j} = (a_{i,j})_{i,j} = A.$$

Beginnen wir umgekehrt mit einer Bilinearform b, nehmen dazu die Gramsche Matrix A_b und bilden dazu wieder die zugehörige Bilinearform, so erhalten wir aufgrund der Bilinearität von b für alle $x, y \in K^n$

$$b_{A_b}(x,y) = x^{\mathsf{T}} A_b\, y = \sum_{i,j=1}^{n} x_i b(e_i, e_j) y_j = b\left(\sum_{i=1}^{n} x_i e_i, \sum_{j=1}^{n} y_j e_j\right) = b(x,y),$$

wobei $x_1, \dots, x_n$ und $y_1, \dots, y_n$ die Koordinaten von x bzw. y bezeichnen. Also sind die beiden gegebenen Abbildungen wirklich invers zueinander.

Schließlich ist die gegebene Abbildung $A \mapsto b_A$ auch linear, denn für $A, B \in \mathrm{Mat}(n \times n, K)$, $\lambda \in K$ und $x, y \in K^n$ gilt

$$b_{A+B}(x,y) = x^{\mathsf{T}}(A+B)\, y = x^{\mathsf{T}} A\, y + x^{\mathsf{T}} B\, y = b_A(x,y) + b_B(x,y)$$

$$\text{und} \qquad b_{\lambda A}(x,y) = x^{\mathsf{T}}(\lambda A)\, y = \lambda\, x^{\mathsf{T}} A\, y = \lambda\, b_A(x,y). \qquad \square$$

48

Beispiel 21.5.

(a) Zur Bilinearform

$$b \colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}, \quad \left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right) \mapsto x_1 y_1 + x_1 y_2 + x_2 y_1 + 4 x_2 y_2 \qquad (*)$$

aus Beispiel 21.3 (a) ist die zugehörige Gramsche Matrix

$$A_b = \begin{pmatrix} b(e_1, e_1) & b(e_1, e_2) \\ b(e_2, e_1) & b(e_2, e_2) \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix}.$$

Eine alternative Beschreibung von $A_b = (a_{i,j})_{i,j}$ ist offensichtlich, dass $a_{i,j}$ in einer Darstellung der Form $(*)$ von $b(x,y)$ genau der Koeffizient von $x_i y_j$ ist.

Umgekehrt können wir nun nach Satz 21.4 aus dieser Matrix auch die ursprüngliche Bilinearform durch die Formel

$$b(x,y) = (x_1 \ x_2) \cdot \begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = x_1 y_1 + x_1 y_2 + x_2 y_1 + 4 x_2 y_2$$

zurückgewinnen.

(b) Die Einheitsmatrix $E_n \in \operatorname{Mat}(n \times n, K)$ entspricht in der Korrespondenz aus Satz 21.4 genau der Bilinearform

$$b\colon K^n \times K^n \to K,\ (x,y) \mapsto x^{\mathsf{T}} y = x_1 y_1 + \cdots + x_n y_n,$$

die wir in Beispiel 21.1 im Fall $K = \mathbb{R}$ und $n = 2$ schon beim gewöhnlichen Skalarprodukt auf $\mathbb{R}^2$ gesehen haben.

Genau wie bei linearen Abbildungen in Satz 16.23 können wir unsere Korrespondenz zwischen Bilinearformen und Matrizen nun unmittelbar von K^n auf beliebige endlich-dimensionale Vektorräume erweitern, indem wir dort eine Basis wählen und mit den Koordinaten bezüglich dieser Basis arbeiten.

Folgerung 21.6 (Bilinearformen auf V und Matrizen)**.** *Es seien V ein endlich-dimensionaler K-Vektorraum sowie $B = (x_1, \ldots, x_n)$ eine Basis von V mit zugehöriger Koordinatenabbildung $\Phi_B\colon V \to K^n$ (siehe Definition 16.22). Dann ist die Abbildung*

$$\operatorname{Mat}(n \times n, K) \to \operatorname{BLF}(V), \quad A \mapsto b_A^B \qquad \text{mit } b_A^B(x,y) := \Phi_B(x)^{\mathsf{T}} A\, \Phi_B(y)$$

wieder ein K-Vektorraumisomorphismus mit Umkehrabbildung

$$\operatorname{BLF}(V) \to \operatorname{Mat}(n \times n, K), \quad b \mapsto A_b^B \qquad \text{mit } A_b^B := (b(x_i, x_j))_{i,j}.$$

*Wie oben nennt man A_b^B die **Gramsche Matrix** von b bezüglich der Basis B.*

Beweis. Die Abbildung

$$\operatorname{BLF}(K^n) \to \operatorname{BLF}(V), \quad b \mapsto \Big((x,y) \mapsto b(\Phi_B(x), \Phi_B(y)) \Big),$$

die einer Bilinearform b auf K^n die Bilinearform auf V zuordnet, bei der man einfach in b die Koordinatenvektoren der Vektoren aus V einsetzt, ist offensichtlich ein Isomorphismus mit Umkehrabbildung

$$\operatorname{BLF}(V) \to \operatorname{BLF}(K^n), \quad b \mapsto \Big((x,y) \mapsto b(\Phi_B^{-1}(x), \Phi_B^{-1}(y)) \Big).$$

Verketten wir den Isomorphismus aus Satz 21.4 mit dieser Abbildung, erhalten wir also wie behauptet einen Isomorphismus $\operatorname{Mat}(n \times n, K) \to \operatorname{BLF}(K^n) \to \operatorname{BLF}(V)$, der eine Matrix A auf die Bilinearform $(x,y) \mapsto b_A(\Phi_B(x), \Phi_B(y)) = \Phi_B(x)^{\mathsf{T}} A\, \Phi_B(y)$ abbildet, und dessen Umkehrung einer Bilinearform b auf V die Matrix $(b(\Phi_B^{-1}(e_i), \Phi_B^{-1}(e_j)))_{i,j} = (b(x_i, x_j))_{i,j}$ zuordnet. □

Natürlich hängt die Gramsche Matrix einer Bilinearform wie in Folgerung 21.6 von der gewählten Basis ab. Wie das folgende Lemma zeigt, ist die Transformationsformel bei einem Basiswechsel jedoch eine andere als für Endomorphismen (siehe Bemerkung 19.3 (b)).

Lemma 21.7 (Verhalten von Gramschen Matrizen unter Basiswechsel)**.** *Es seien b eine Bilinearform auf einem endlich erzeugten K-Vektorraum V sowie B und B' zwei Basen von V. Dann gilt für die Gramschen Matrizen von b bezüglich B und B'*

$$A_b^{B'} = T^{\mathsf{T}} A_b^B\, T,$$

wobei $T = A^{B',B}$ die Basiswechselmatrix aus Definition 16.28 ist.

Beweis. Es seien $B = (x_1, \ldots, x_n)$ und $B' = (y_1, \ldots, y_n)$ die gewählten Basen. Nach Definition 16.28 enthält die k-te Spalte von $T = (a_{i,j})_{i,j}$ für $k = 1, \ldots, n$ genau den Koordinatenvektor von y_k bezüglich B, d. h. es gilt

$$y_k = a_{1,k} x_1 + \cdots + a_{n,k} x_n.$$

Damit folgt für die Gramschen Matrizen mit der Formel aus Folgerung 21.6 sofort

$$A_b^{B'} = (b(y_k, y_l))_{k,l} = \left(b\left(\sum_{i=1}^n a_{i,k} x_i, \sum_{j=1}^n a_{j,l} x_j \right) \right)_{k,l} = \left(\sum_{i,j=1}^n a_{i,k}\, b(x_i, x_j)\, a_{j,l} \right)_{k,l} = T^{\mathsf{T}} A_b^B\, T. \qquad \square$$

Bemerkung 21.8 (Matrizen unter Basiswechsel)**.** Bisher hatten wir Matrizen nahezu ausschließlich zur Beschreibung von linearen Abbildungen benutzt. Nach Definition ist eine Matrix aber zunächst einmal nichts weiter als ein rechteckiges Zahlenschema, ohne Vorgabe einer Bedeutung dieser Zahlen. In der Tat haben wir nun gesehen, dass man Matrizen auch noch für ganz andere Dinge verwenden kann, nämlich z. B. zur Darstellung von Bilinearformen.

Ohne weitere Informationen ergibt es daher keinen Sinn zu fragen, wie sich eine Matrix unter einem Basiswechsel transformiert. Die Antwort auf diese Frage hängt nach Bemerkung 19.3 (b) und Lemma 21.7 davon ab, welche Bedeutung die Einträge in der Matrix haben:

Bei einem Basiswechsel mit zugehöriger Basiswechselmatrix T transformiert sich ...
... eine zu einem *Endomorphismus* gehörige Matrix A in die Matrix $T^{-1}AT$,
... eine zu einer *Bilinearform* gehörige Matrix A in die Matrix $T^{\mathsf{T}}AT$.

Wie bei linearen Abbildungen oder Endomorphismen könnten wir uns nun schließlich auch bei Bilinearformen wieder nach einer Normalform fragen: Wie können wir zu einer gegebenen Bilinearform $b \in \mathrm{BLF}(V)$ eine Basis von V so wählen, dass die zugehörige Gramsche Matrix A_b^B möglichst einfach wird? Wir wollen diese Frage hier allerdings nicht in dieser vollen Allgemeinheit beantworten, da wir im Folgenden hauptsächlich an Bilinearformen mit noch weiteren speziellen Eigenschaften interessiert sind. Diese Eigenschaften wollen wir jetzt einführen.

21.B Skalarprodukte

Wir kommen nun zu den in der Einleitung zu diesem Kapitel bereits angekündigten Skalarprodukten. Wir werden sie als Bilinearformen mit den folgenden beiden Eigenschaften definieren, die wir auch gleich wieder analog für Matrizen einführen wollen.

Definition 21.9 (Symmetrie und positive Definitheit)**.** Es seien $b \in \mathrm{BLF}(V)$ eine Bilinearform auf einem K-Vektorraum V und $A \in \mathrm{Mat}(n \times n, K)$.

(a) Die Bilinearform b heißt **symmetrisch**, wenn $b(x,y) = b(y,x)$ für alle $x, y \in V$.

Die Matrix A heißt **symmetrisch**, wenn $A^{\mathsf{T}} = A$.

(b) Es sei nun zusätzlich $K = \mathbb{R}$.

Die Bilinearform b heißt dann **positiv definit**, wenn $b(x,x) > 0$ für alle $x \in V$ mit $x \neq 0$.

Die Matrix A heißt dann **positiv definit**, wenn $x^{\mathsf{T}}Ax > 0$ für alle $x \in \mathbb{R}^n$ mit $x \neq 0$.

Bemerkung 21.10.

(a) Die Bedingung der positiven Definitheit lässt sich offensichtlich nur für einen geordneten Körper (siehe Kapitel 4.B) formulieren. Für uns ist hierbei eigentlich nur der Fall $K = \mathbb{R}$ interessant. Wir werden die Bedingung der positiven Definitheit in Konstruktion 21.18 aber noch etwas abändern, so dass sie dann auch im Fall $K = \mathbb{C}$ anwendbar ist.

(b) Da eine Bilinearform b linear in jedem Eintrag ist, gilt natürlich stets $b(0,0) = 0$. Eine positiv definite Bilinearform auf einem $\mathbb{R}$-Vektorraum V erfüllt damit also immer $b(x,x) \geq 0$ für alle $x \in V$. Diese Bedingung, die wir später für Skalarprodukte fordern werden, wird uns dann sicherstellen, dass wir aus $b(x,x)$ die Wurzel ziehen und so die Länge von x definieren können.

In manchen Fällen (siehe Satz 26.20) benötigt man allerdings auch noch die folgenden zur positiven Definitheit analogen Bedingungen: Eine Bilinearform $b \in \mathrm{BLF}(V)$ auf einem reellen Vektorraum V heißt ...

- ... **negativ definit**, wenn $b(x,x) < 0$ für alle $x \in V$ mit $x \neq 0$;
- ... **positiv semidefinit**, wenn $b(x,x) \geq 0$ für alle $x \in V$;
- ... **negativ semidefinit**, wenn $b(x,x) \leq 0$ für alle $x \in V$;

- ... **indefinit**, wenn sie weder positiv noch negativ semidefinit ist, also wenn es $x,y \in V$ gibt mit $b(x,x) > 0$ und $b(y,y) < 0$.

Entsprechende Eigenschaften definiert man natürlich auch für reelle quadratische Matrizen.

Als Erstes wollen wir nun die wohl erwartete Aussage zeigen, dass sich die in Definition 21.9 eingeführten Begriffe für Bilinearformen und Matrizen entsprechen.

Lemma 21.11 (Symmetrie und positive Definitheit bei Bilinearformen und Matrizen)**.** *Es seien b eine Bilinearform auf einem endlich-dimensionalen K-Vektorraum V, B eine Basis von V, und A_b^B wie in Folgerung 21.6 die zugehörige Gramsche Matrix. Dann gilt:*

(a) *Die Bilinearform b ist genau dann symmetrisch, wenn die Matrix A_b^B symmetrisch ist.*

(b) *Im Fall $K = \mathbb{R}$ ist b genau dann positiv definit, wenn A_b^B positiv definit ist.*

Beweis. Es sei $B = (x_1, \dots, x_n)$.

(a) „$\Rightarrow$“: Ist b symmetrisch, so folgt natürlich sofort

$$(A_b^B)^\mathsf{T} = (b(x_j,x_i))_{i,j} = (b(x_i,x_j))_{i,j} = A_b^B.$$

„$\Leftarrow$“: Ist umgekehrt A_b^B symmetrisch, so gilt nach Folgerung 21.6 für alle $x,y \in V$

$$b(x,y) = \Phi_B(x)^\mathsf{T} A_b^B \Phi_B(y) \overset{(*)}{=} \Phi_B(y)^\mathsf{T} (A_b^B)^\mathsf{T} \Phi_B(x) = \Phi_B(y)^\mathsf{T} A_b^B \Phi_B(x) = b(y,x),$$

wobei wir in $(*)$ gemäß Lemma 16.7 (d) die transponierte 1×1-Matrix gebildet haben.

(b) Da die Koordinatenabbildung $\Phi_B \colon V \to \mathbb{R}^n$ ein Isomorphismus ist, ist A_b^B genau dann positiv definit, wenn $\Phi_B(x)^\mathsf{T} A_b^B \Phi_B(x) > 0$, nach Folgerung 21.6 also genau dann wenn $b(x,x) > 0$ für alle $x \in V$ mit $x \neq 0$ gilt. □

Bemerkung 21.12 (Invarianz von Symmetrie und positiver Definitheit)**.** Sind $A, A' \in \operatorname{Mat}(n \times n, K)$ zwei quadratische Matrizen mit $A' = T^\mathsf{T} A T$ für ein $T \in \operatorname{GL}(n,K)$, so besagt Lemma 21.11 insbesondere, dass A' genau dann symmetrisch (bzw. im Fall $K = \mathbb{R}$ positiv definit) ist, wenn dies für A gilt: A' und A beschreiben nach Lemma 21.7 nämlich die gleiche Bilinearform bezüglich zweier evtl. verschiedener Basen, und nach Lemma 21.11 hängt es nur von dieser Bilinearform (aber eben nicht von der gewählten Basis) ab, ob die Matrix symmetrisch bzw. positiv definit ist.

Mit Hilfe der eingeführten Konzepte können wir nun Skalarprodukte auf reellen Vektorräumen definieren.

Definition 21.13 (Skalarprodukte)**.** Es sei V ein $\mathbb{R}$-Vektorraum. Ein **Skalarprodukt** auf V ist eine positiv definite, symmetrische Bilinearform $b \colon V \times V \to \mathbb{R}$. Ein $\mathbb{R}$-Vektorraum V zusammen mit einem Skalarprodukt heißt ein **euklidischer Raum**.

Für $x,y \in V$ schreiben wir statt $b(x,y)$ dann auch $\langle x,y \rangle$. Die (wegen der positiven Definitheit existierende) Zahl

$$||x|| := \sqrt{\langle x,x \rangle} \quad \in \mathbb{R}_{\geq 0}$$

heißt in Verallgemeinerung von Beispiel 21.1 die **Norm** oder **Länge** von x (bezüglich b). Man nennt einen Vektor $x \in V$ **normiert** (bezüglich b), falls $||x|| = 1$.

Bemerkung 21.14. Ist V ein endlich erzeugter $\mathbb{R}$-Vektorraum und $B = (x_1, \dots, x_n)$ eine Basis von V, so lässt sich ein Skalarprodukt auf V nach Lemma 21.11 also genau durch eine positiv definite, symmetrische Matrix $A \in \operatorname{Mat}(n \times n, \mathbb{R})$ beschreiben bzw. definieren (nämlich durch die Gramsche Matrix des Skalarprodukts bezüglich der Basis B).

Beispiel 21.15.

(a) Für $V = \mathbb{R}^n$ ist

$$\langle x,y \rangle := x^\mathsf{T} y = \sum_{i=1}^n x_i y_i$$

(wobei $x_1,\dots,x_n$ und $y_1,\dots,y_n$ die Koordinaten von x bzw. y bezeichnen) ein Skalarprodukt: Die Bilinearität und Symmetrie sind offensichtlich, und die positive Definitheit ergibt sich sofort daraus, dass für alle $x \in \mathbb{R}^n$ mit $x \neq 0$ natürlich

$$\langle x,x\rangle = \sum_{i=1}^{n} x_i^2 > 0$$

gilt. Es ist die direkte Verallgemeinerung von Beispiel 21.1 und wird als das **Standardskalarprodukt** auf $\mathbb{R}^n$ bezeichnet. Bezüglich der Standardbasis ist die Gramsche Matrix dieses Skalarprodukts offensichtlich die Einheitsmatrix (siehe Beispiel 21.5 (b)).

(b) Die in Beispiel 21.5 (a) bereits betrachtete reelle Matrix $A = \begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix}$ ist positiv definit: Für alle $x \in \mathbb{R}^2$ ist zunächst

$$x^{\mathsf{T}} A\, x = x_1^2 + x_1x_2 + x_2x_1 + 4x_2^2 = (x_1+x_2)^2 + 3x_2^2 \geq 0,$$

und die Gleichheit kann hier nur gelten für $x_1 + x_2 = x_2 = 0$, also für $x = 0$. Da A auch symmetrisch ist, ist die durch diese Matrix (bezüglich der Standardbasis) definierte Bilinearform

$$\langle x,y\rangle = x^{\mathsf{T}} A\, y = x_1y_1 + x_1y_2 + x_2y_1 + 4x_2y_2$$

also ein Skalarprodukt auf $\mathbb{R}^2$.

(c) Es sei $V = C^0([0,1])$ der reelle Vektorraum der stetigen Funktionen auf dem Intervall $[0,1]$. Wir behaupten, dass dann die Abbildung

$$\langle f,g\rangle := \int_0^1 f(x)g(x)\,dx$$

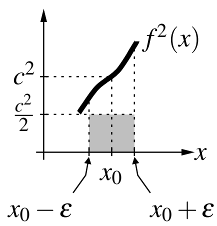
ein Skalarprodukt auf V definiert. Die Bilinearität ergibt sich hierbei einfach aus den Eigenschaften des Integrals, z. B. ist für $f_1, f_2, g \in V$

$$\begin{aligned}\langle f_1+f_2,g\rangle &= \int_0^1 (f_1(x)+f_2(x))g(x)\,dx = \int_0^1 f_1(x)g(x)\,dx + \int_0^1 f_2(x)g(x)\,dx \\ &= \langle f_1,g\rangle + \langle f_2,g\rangle\,.\end{aligned}$$

Die Symmetrie ist natürlich offensichtlich. Für die positive Definitheit bemerken wir zunächst, dass die Funktion f^2 für jedes $f \in V$ überall nicht-negativ ist und damit schon einmal

$$\langle f,f\rangle = \int_0^1 f(x)^2\,dx \geq 0$$

gilt. Ist nun zusätzlich $f \neq 0$ nicht die Nullfunktion, so gibt es also ein $x_0 \in [0,1]$ mit $c := f(x_0) \neq 0$. Da f und damit auch f^2 stetig ist, gibt es dann aber wie im Bild rechts ein $\varepsilon > 0$, so dass $f(x)^2 \geq \frac{c^2}{2}$ für alle $x \in [0,1]$ mit $|x - x_0| < \varepsilon$ ist. Unter dem Graphen von f^2 liegt also sicher ein positives Flächenstück (der Breite 2ε und Höhe $\frac{c^2}{2}$, falls x_0 nicht zu weit am Rand des Intervalls liegt), und somit ist wie gewünscht $\langle f,f\rangle = \int_0^1 f(x)^2\,dx > 0$.



Beachte, dass hierfür die Stetigkeit der betrachteten Funktionen entscheidend ist: Wäre z. B. die nicht stetige Funktion

$$f\colon [0,1] \to \mathbb{R},\ x \mapsto \begin{cases} 0 & \text{für } x \neq \frac{1}{2}, \\ 1 & \text{für } x = \frac{1}{2} \end{cases}$$

zugelassen, so wäre hier zwar $f \neq 0$, aber trotzdem $\langle f,f\rangle = \int_0^1 f(x)^2\,dx = 0$. Wir hätten in diesem Fall also kein Skalarprodukt (siehe auch Aufgabe 21.23 (c)).

(d) Die Einschränkung eines Skalarprodukts auf einen Untervektorraum ist offensichtlich wieder ein Skalarprodukt. So können wir z. B. das Skalarprodukt aus (c) genauso auch auf den Räumen aller differenzierbaren Funktionen oder aller Polynomfunktionen betrachten.

Aufgabe 21.16. Es sei $n \in \mathbb{N}_{>0}$. Zeige, dass die folgenden Abbildungen b Skalarprodukte auf dem reellen Vektorraum V sind:

(a) $V = \mathbb{R}^n$, $b(x,y) = x^\mathsf{T} A\, y$ mit $A = \begin{pmatrix} 2 & 1 & & & \\ 1 & 2 & 1 & & \huge 0 \\ & \ddots & \ddots & \ddots & \\ \huge 0 & & 1 & 2 & 1 \\ & & & 1 & 2 \end{pmatrix}$;

(b) $V = \mathrm{Mat}(n \times n, \mathbb{R})$, $b(A,B) = \mathrm{Spur}(A^\mathsf{T} B)$.

Aufgabe 21.17 (Duale Vektorräume)**.** Es sei V ein K-Vektorraum. Man nennt $V^* := \mathrm{Hom}(V,K)$ den *Dualraum* zu V.

(a) Ist V ein euklidischer Vektorraum, so zeige man, dass

$$\Phi \colon V \to V^*, \quad x \mapsto \varphi_x \qquad \text{mit } \varphi_x(y) := \langle x,y \rangle$$

eine injektive lineare Abbildung ist.

(b) Zeige, dass die Abbildung Φ aus (a) sogar ein Isomorphismus ist, wenn V endlich erzeugt ist.

(c) Im Fall des euklidischen Raumes $V = C^0([0,1])$ aller stetigen Funktionen auf dem Intervall $[0,1]$ mit dem Skalarprodukt $\langle f,g \rangle = \int_0^1 f(x)g(x)\,dx$ aus Beispiel 21.15 (c) zeige man, dass die Abbildung Φ aus (a) nicht surjektiv und damit kein Isomorphismus ist.

49

Im Rest dieses Kapitels wollen wir nun die grundlegenden Eigenschaften von euklidischen Räumen untersuchen. Da es in der Praxis öfters einmal vorkommt, werden wir den Begriff des Skalarprodukts aber zunächst noch auf den Fall von komplexen Vektorräumen erweitern.

Konstruktion 21.18 (Skalarprodukte im komplexen Fall)**.** Wollen wir die Definition 21.13 auf einen komplexen Vektorraum übertragen, so haben wir das Problem, dass die Bedingung der positiven Definitheit über $\mathbb{C}$ zunächst einmal keinen Sinn ergibt, da $\mathbb{C}$ kein geordneter Körper ist. Dies hat zur Folge, dass wir die Norm eines Vektors nicht mehr wie gewohnt definieren können: Würden wir wie beim Standardskalarprodukt im Reellen auch in $\mathbb{C}^n$ die Formeln

$$\langle x,y \rangle = \sum_{i=1}^n x_i y_i \quad \text{und} \quad ||x|| = \sqrt{\langle x,x \rangle} = \sqrt{x_1^2 + \cdots + x_n^2}$$

verwenden, so müssten wir hier die Wurzel aus einer im Allgemeinen komplexen Zahl $x_1^2 + \cdots + x_n^2$ bilden — was nicht eindeutig möglich ist und auch nicht wie gewünscht zu einer nicht-negativen reellen Zahl als Länge eines Vektors führen würde. Die Lösung dieses Problems besteht darin, im Skalarprodukt grundsätzlich jede Koordinate des *ersten* (aber nicht des zweiten) Eintrags komplex zu konjugieren, so dass wir z. B. für das Standardskalarprodukt die Formeln

$$\langle x,y \rangle = \sum_{i=1}^n \overline{x_i}\, y_i \quad \text{und damit} \quad ||x|| = \sqrt{\langle x,x \rangle} = \sqrt{\overline{x_1}\, x_1 + \cdots + \overline{x_n}\, x_n} = \sqrt{|x_1|^2 + \cdots + |x_n|^2}$$

erhalten, die wieder zu einer reellen, nicht-negativen Länge eines Vektors führen.

Mit dem Hintergrund dieser Idee sind die entsprechenden Abänderungen für beliebige Skalarprodukte auf komplexen Vektorräumen relativ offensichtlich. Wir werden die sich daraus ergebenden Definitionen und Resultate im Folgenden kurz auflisten; die Beweise dieser Aussagen sind völlig analog zu denen im reellen Fall.

Es sei also V ein $\mathbb{C}$-Vektorraum. Eine **Sesquilinearform** auf V ist eine Abbildung $s\colon V \times V \to \mathbb{C}$, so dass für alle $x_1, x_2, x, y_1, y_2, y \in V$ und $\lambda \in \mathbb{C}$ die Eigenschaften

$$\begin{aligned} s(x_1 + x_2, y) &= s(x_1, y) + s(x_2, y), \\ s(\lambda x, y) &= \overline{\lambda}\, s(x, y), \\ s(x, y_1 + y_2) &= s(x, y_1) + s(x, y_2), \\ s(x, \lambda y) &= \lambda\, s(x, y) \end{aligned}$$

gelten. Der einzige Unterschied zu Bilinearformen besteht also in der komplexen Konjugation des Skalars in der zweiten Zeile oben — und in der Tat kommt der Begriff „Sesquilinearform" aus dem Lateinischen und bedeutet „eineinhalbfach lineare Form".

Die Sesquilinearformen auf V bilden einen $\mathbb{C}$-Vektorraum, den wir mit $\mathrm{SLF}(V)$ bezeichnen wollen. Ist V endlich-dimensional und $B = (x_1, \dots, x_n)$ eine Basis von V, so ist $\mathrm{SLF}(V)$ wie in Folgerung 21.6 isomorph zu $\mathrm{Mat}(n \times n, \mathbb{C})$ über die beiden zueinander inversen Isomorphismen

$$\begin{array}{llll} & \mathrm{Mat}(n \times n, \mathbb{C}) \to \mathrm{SLF}(V), & A \mapsto s_A^B & \text{mit } s_A^B(x,y) := \overline{\Phi_B(x)}^{\mathsf{T}} A\, \Phi_B(y) \\ \text{und} & \mathrm{SLF}(V) \to \mathrm{Mat}(n \times n, \mathbb{C}), & s \mapsto A_s^B & \text{mit } A_s^B := (s(x_i, x_j))_{i,j}, \end{array}$$

wobei der Querstrich über $\Phi_B(x) \in \mathbb{C}^n$ bedeutet, dass jede Komponente des Vektors komplex konjugiert wird. Man bezeichnet A_s^B wieder als die *Gramsche Matrix* von s. Ist B' eine weitere Basis von V, so transformieren sich die Gramschen Matrizen analog zu Lemma 21.7 gemäß

$$A_s^{B'} = \overline{T}^{\mathsf{T}} A_s^B\, T \quad \text{mit } T = A^{B',B},$$

wobei $T = A^{B',B}$ die übliche Basiswechselmatrix ist.

Eine Sesquilinearform s auf einem $\mathbb{C}$-Vektorraum V heißt **hermitesch**, wenn $s(y,x) = \overline{s(x,y)}$ für alle $x, y \in V$. Ist dies der Fall, so erhalten wir daraus insbesondere mit $y = x$

$$s(x,x) = \overline{s(x,x)}, \quad \text{also} \quad s(x,x) \in \mathbb{R}$$

für alle $x \in V$. Wir können daher fragen, ob diese reelle Zahl immer nicht-negativ ist, und nennen eine hermitesche Sesquilinearform s analog zum reellen Fall **positiv definit**, wenn $s(x,x) > 0$ für alle $x \in V$ mit $x \neq 0$.

Entsprechend heißt eine Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{C})$ **hermitesch**, wenn $\overline{A}^{\mathsf{T}} = A$. In diesem Fall ist $\overline{x}^{\mathsf{T}} A x \in \mathbb{R}$ für alle $x \in \mathbb{C}^n$, denn es ist

$$\overline{\overline{x}^{\mathsf{T}} A x} = x^{\mathsf{T}} \overline{A}\, \overline{x} = (x^{\mathsf{T}} \overline{A}\, \overline{x})^{\mathsf{T}} = \overline{x}^{\mathsf{T}} \overline{A}^{\mathsf{T}} x = \overline{x}^{\mathsf{T}} A x$$

nach Lemma 16.7 (d). Gilt sogar $\overline{x}^{\mathsf{T}} A x > 0$ für alle $x \in \mathbb{C}^n$ mit $x \neq 0$, so heißt A **positiv definit**. Wie in Lemma 21.11 ist eine Sesquilinearform auf einem endlich-dimensionalen Vektorraum genau dann hermitesch, bzw. eine hermitesche Sesquilinearform genau dann positiv definit, wenn ihre Gramsche Matrix zu einer beliebigen Basis diese Eigenschaft besitzt. Die anderen Definitheitsbegriffe aus Bemerkung 21.10 (b) definiert man natürlich sowohl für hermitesche Sesquilinearformen als auch für hermitesche Matrizen analog.

Ein **Skalarprodukt** auf einem komplexen Vektorraum V ist nun eine positiv definite, hermitesche Sesquilinearform s auf V, die wir dann typischerweise als $\langle x, y \rangle := s(x,y)$ schreiben. Ein $\mathbb{C}$-Vektorraum zusammen mit einem Skalarprodukt heißt ein **unitärer Raum**. Für $x \in V$ nennen wir in diesem Fall wieder

$$||x|| := \sqrt{\langle x, x \rangle} \quad \in \mathbb{R}_{\geq 0}$$

die *Norm* (oder *Länge*) von x.

Das **Standardskalarprodukt** auf $\mathbb{C}^n$ ist dasjenige, das bezüglich der Standardbasis der Einheitsmatrix entspricht, also

$$\langle x, y \rangle := \overline{x}^{\mathsf{T}} y = \overline{x_1}\, y_1 + \cdots + \overline{x_n}\, y_n,$$

wobei $x_1,\dots,x_n$ und $y_1,\dots,y_n$ die Koordinaten von x bzw. y sind. Beachte, dass die Norm eines Vektors $x \in \mathbb{C}^n$ in diesem Fall dann

$$||x|| = \sqrt{\langle x,x\rangle} = \sqrt{|x_1|^2+\cdots+|x_n|^2} = \sqrt{(\operatorname{Re} x_1)^2+(\operatorname{Im} x_1)^2+\cdots+(\operatorname{Re} x_n)^2+(\operatorname{Im} x_n)^2}$$

und damit gleich der Norm dieses Vektors in $\mathbb{R}^{2n}$ bezüglich des Standardskalarprodukts ist.

Im Folgenden wollen wir den reellen und komplexen Fall in der Regel zusammen behandeln. Wie in der Analysis schreiben wir daher wieder $\mathbb{K}$ für einen der Körper $\mathbb{R}$ oder $\mathbb{C}$, und sprechen von einem *Vektorraum mit Skalarprodukt*, wenn wir einen euklidischen bzw. unitären Raum meinen. Die Formeln werden wir dabei wie in Konstruktion 21.18 mit der komplexen Konjugation schreiben, so dass sie dann für $\mathbb{K}=\mathbb{R}$ und $\mathbb{K}=\mathbb{C}$ gleichermaßen gelten — im reellen Fall ist die komplexe Konjugation dann zwar unnötig, schadet aber natürlich auch nicht.

Als erstes Resultat über Skalarprodukte wollen wir nun eine sehr wichtige Ungleichung beweisen.

Satz 21.19 (**Cauchy-Schwarz-Ungleichung**). *In jedem Vektorraum V mit Skalarprodukt gilt*

$$|\langle x,y\rangle| \le ||x||\cdot||y||$$

für alle $x,y \in V$, wobei die Gleichheit genau dann gilt, wenn x und y linear abhängig sind.

Beweis. Für $y=0$ ist die Aussage des Satzes offensichtlich, denn dann sind beide Seiten gleich Null und die Vektoren linear abhängig. Andernfalls setzen wir $\lambda := \frac{\langle y,x\rangle}{\langle y,y\rangle}$ (und damit $\overline{\lambda} = \frac{\langle x,y\rangle}{\langle y,y\rangle}$ wegen der Symmetrie bzw. Hermitizität des Skalarprodukts) und folgern aus der positiven Definitheit, dass

$$\begin{aligned}
0 &\le \langle x-\lambda y, x-\lambda y\rangle && (*)\\
&= \langle x,x\rangle - \lambda\langle x,y\rangle - \overline{\lambda}\langle y,x\rangle + \lambda\overline{\lambda}\langle y,y\rangle\\
&= \langle x,x\rangle - \frac{\langle y,x\rangle\langle x,y\rangle}{\langle y,y\rangle} - \frac{\langle x,y\rangle\langle y,x\rangle}{\langle y,y\rangle} + \frac{\langle x,y\rangle\langle y,x\rangle}{\langle y,y\rangle}\\
&= \langle x,x\rangle - \frac{\langle x,y\rangle\langle y,x\rangle}{\langle y,y\rangle}\\
&= ||x||^2 - \frac{|\langle x,y\rangle|^2}{||y||^2}
\end{aligned}$$

und damit $|\langle x,y\rangle|^2 \le ||x||^2\cdot||y||^2$ gilt. Wurzelziehen liefert nun die behauptete Ungleichung.

Gilt in dieser Ungleichung sogar die Gleichheit, gilt also die Gleichheit in $(*)$, so ergibt sich aus der positiven Definitheit des Skalarprodukts sofort $x-\lambda y=0$, d. h. x und y sind linear abhängig. Sind umgekehrt x und y linear abhängig, gilt also $x=\mu y$ für ein $\mu\in\mathbb{K}$, so ist $\langle y,x\rangle = \langle y,\mu y\rangle = \mu\langle y,y\rangle$ und damit $\mu=\lambda$, so dass in $(*)$ oben und damit auch in der Cauchy-Schwarz-Ungleichung sogar die Gleichheit gilt. □

Mit der Cauchy-Schwarz-Ungleichung können wir die folgenden Eigenschaften der Norm herleiten:

Satz 21.20 (Eigenschaften der Norm). *In jedem Vektorraum V mit Skalarprodukt gilt*

(a) $||\lambda x|| = |\lambda|\cdot||x||$ *für alle $\lambda\in\mathbb{K}$ und $x\in V$;*

(b) $||x||>0$ *für alle $x\in V\backslash\{0\}$;*

(c) $||x+y|| \le ||x||+||y||$ *für alle $x,y\in V$ (**Dreiecksungleichung**).*

Beweis.

(a) Es ist $||\lambda x|| = \sqrt{\langle\lambda x,\lambda x\rangle} = \sqrt{\overline{\lambda}\lambda\,\langle x,x\rangle} = \sqrt{|\lambda|^2\langle x,x\rangle} = |\lambda|\cdot||x||$.

(b) folgt sofort aus der positiven Definitheit des Skalarprodukts.

(c) Es gilt

$$\begin{aligned}
||x+y||^2 &= \langle x+y, x+y\rangle \\
&= \langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle + \langle y,y\rangle \\
&= ||x||^2 + \langle x,y\rangle + \overline{\langle x,y\rangle} + ||y||^2 \\
&= ||x||^2 + 2\,\mathrm{Re}\,\langle x,y\rangle + ||y||^2 && \text{(wegen } \mathrm{Re}\,z = \tfrac{1}{2}(z+\overline{z}) \text{ für alle } z \in \mathbb{C}) \\
&\le ||x||^2 + 2\,|\langle x,y\rangle| + ||y||^2 && \text{(wegen } \mathrm{Re}\,z \le |z| \text{ für alle } z \in \mathbb{C}) \\
&\le ||x||^2 + 2\,||x||\cdot||y|| + ||y||^2 && \text{(Satz 21.19)} \\
&= (||x|| + ||y||)^2,
\end{aligned}$$

woraus durch Wurzelziehen die Behauptung folgt. □

Bemerkung 21.21 (Normen in der linearen Algebra und Analysis)**.** In der Analysis werden wir später normierte Vektorräume definieren als reelle oder komplexe Vektorräume mit einer reellwertigen „Normabbildung“ $||\cdot||$, die die drei Eigenschaften aus Satz 21.20 erfüllt (siehe Definition 23.1). In diesem Sinne sind Vektorräume mit Skalarprodukt also immer normierte Vektorräume aus der Sicht der Analysis. Wir werden allerdings sehen, dass es auch sehr viele Normen gibt, die nicht von einem Skalarprodukt kommen, z. B. in $\mathbb{R}^2$ die Summennorm $||x|| = |x_1| + |x_2|$ oder die Maximumsnorm $||x|| = \max\{|x_1|, |x_2|\}$.

Wir kommen nun zum Winkel zwischen zwei Vektoren. Man definiert ihn in der Regel nur im reellen Fall und orientiert sich dabei an der geometrischen Deutung aus Beispiel 21.1 im Fall des Standardskalarprodukts.

Konstruktion 21.22 (Winkel)**.** Es seien V ein euklidischer Raum und $x,y \in V\backslash\{0\}$. Nach der Cauchy-Schwarz-Ungleichung aus Satz 21.19 ist dann

$$\frac{|\langle x,y\rangle|}{||x||\cdot||y||} \le 1, \qquad \text{also} \qquad -1 \le \frac{\langle x,y\rangle}{||x||\cdot||y||} \le 1.$$

Die Zahl

$$\varphi = \arccos \frac{\langle x,y\rangle}{||x||\cdot||y||} \quad \in [0,\pi]$$

ist daher wohldefiniert; wir nennen sie in Analogie zu Beispiel 21.1 den (unorientierten) **Winkel** zwischen x und y.

Aufgabe 21.23 (Skalarprodukte aus positiv semidefiniten Formen)**.** Zu einer symmetrischen Bilinearform b auf einem reellen Vektorraum V sei $U_b = \{x \in V : b(x,x) = 0\}$. Man zeige:

(a) U_b ist im Allgemeinen kein Unterraum von V.

(b) Ist b jedoch positiv semidefinit, so ist U_b ein Unterraum, und $\overline{b}(\overline{x},\overline{y}) := b(x,y)$ ist ein wohldefiniertes Skalarprodukt auf V/U_b.

(c) Was ergibt sich aus dieser Konstruktion, wenn V der Vektorraum aller stückweise stetigen Funktionen auf einem Intervall $[a,b]$ und $b(f,g) = \int_a^b f(x)g(x)\,dx$ ist? Wie kann man sich in diesem Fall die Elemente von U_b und V/U_b anschaulich vorstellen?

Aufgabe 21.24. Es sei $f\colon V \to V$ ein Endomorphismus eines $\mathbb{K}$-Vektorraums mit Skalarprodukt, so dass $\langle f(x), x\rangle = 0$ für alle $x \in V$. Man zeige:

(a) Ist V ein unitärer Raum (also $\mathbb{K} = \mathbb{C}$), so ist f die Nullabbildung.

(b) Ist V ein euklidischer Raum (also $\mathbb{K} = \mathbb{R}$), so ist f nicht notwendig die Nullabbildung.

21.C Orthogonalität

Der mit Abstand wichtigste Fall der Winkeldefinition ist derjenige, in dem die beiden betrachteten Vektoren „senkrecht aufeinander stehen", also dieser Winkel gleich $\frac{\pi}{2}$ und nach Konstruktion 21.22 damit das Skalarprodukt gleich 0 ist. Im Gegensatz zu allgemeinen Winkeln können wir diesen Spezialfall auch wieder sowohl für reelle als auch für komplexe Vektorräume definieren.

Definition 21.25 (Orthogonale Vektoren)**.** Es sei V ein Vektorraum mit Skalarprodukt.

(a) Zwei Vektoren $x, y \in V$ heißen **orthogonal** bzw. **senkrecht** zueinander (in Zeichen: $x \perp y$), wenn $\langle x, y \rangle = 0$ gilt. Für $x \in V$ und einen Unterraum $U \leq V$ schreiben wir kurz $x \perp U$, falls $x \perp y$ für alle $y \in U$.

(b) Eine Familie $B = (x_i)_{i \in I}$ in V heißt **orthogonal**, wenn $x_i \perp x_j$ für alle $i, j \in I$ mit $i \neq j$ gilt, also wenn die gegebenen Vektoren paarweise zueinander senkrecht sind. Gilt zusätzlich $||x_i|| = 1$ für alle $i \in I$, so nennt man B **orthonormal**.

(c) Eine **Orthonormalbasis** von V ist eine orthonormale Familie, die gleichzeitig eine Basis von V ist.

Beispiel 21.26.

(a) Nach Definition ist der Nullvektor orthogonal zu jedem anderen Vektor.

(b) Die Standardbasis von $\mathbb{R}^n$ bzw. $\mathbb{C}^n$ ist natürlich eine Orthonormalbasis bezüglich des Standardskalarprodukts.

Bemerkung 21.27. Es sei U ein Unterraum eines Vektorraums V mit Skalarprodukt. Sind dann $x \in V$ und $(x_1, \dots, x_n)$ eine Basis von U, so genügt es für die Bedingung $x \perp U$ zu überprüfen, dass $x \perp x_i$ für alle $i = 1, \dots, n$: In diesem Fall gilt nämlich $\langle x, x_i \rangle = 0$ für alle $i = 1, \dots, n$, also auch

$$\langle x, \lambda_1 x_1 + \cdots + \lambda_n x_n \rangle = \lambda_1 \langle x, x_1 \rangle + \cdots + \lambda_n \langle x, x_n \rangle = 0$$

für alle $\lambda_1, \dots, \lambda_n \in \mathbb{K}$, und damit $\langle x, y \rangle = 0$ für alle $y \in U$.

Lemma 21.28. *In einem Vektorraum mit Skalarprodukt ist jede orthogonale Familie, die nicht den Nullvektor enthält, linear unabhängig.*

Beweis. Es sei $B = (x_i)_{i \in I}$ orthogonal mit $x_i \neq 0$ für alle $i \in I$. Weiterhin sei $\sum_{i \in I} \lambda_i x_i = 0$ eine Linearkombination des Nullvektors (wobei nur endlich viele $\lambda_i \in \mathbb{K}$ ungleich 0 sind, siehe Definition 14.1 (b)). Bilden wir dann das Skalarprodukt dieser Gleichung mit einem x_k für $k \in I$, so folgt

$$0 = \left\langle x_k, \sum_{i \in I} \lambda_i x_i \right\rangle = \sum_{i \in I} \lambda_i \langle x_k, x_i \rangle = \lambda_k ||x_k||^2,$$

weil B orthogonal ist. Da weiterhin nach Voraussetzung $x_k \neq 0$ und damit aufgrund der positiven Definitheit $||x_k|| \neq 0$ ist, folgt $\lambda_k = 0$. Dies gilt aber für alle k, d. h. die ursprüngliche Linearkombination ist trivial, und B ist damit linear unabhängig. □

Wir wollen nun sehen, dass Orthonormalbasen in jedem (endlich-dimensionalen) Vektorraum mit Skalarprodukt existieren und zu besonders schönen Eigenschaften führen. In der Tat gilt sogar noch mehr, nämlich eine „Basisergänzungseigenschaft" analog zu Folgerung 14.16 für Orthonormalbasen. Um dies zu beweisen, brauchen wir die folgende Konstruktion der orthogonalen Projektion auf einen Unterraum.

Lemma und Definition 21.29 (Orthogonale Projektionen)**.** *Es sei U ein endlich-dimensionaler Unterraum eines Vektorraums V mit Skalarprodukt. Ferner sei $(x_1, \dots, x_n)$ eine Orthonormalbasis von U (wir werden in Satz 21.31 noch sehen, dass eine solche Orthonormalbasis immer existiert).*

Dann gibt es genau eine lineare Abbildung $f\colon V \to U$ mit $x - f(x) \perp U$ für alle $x \in V$, nämlich

$$f(x) = \sum_{i=1}^{n} \langle x_i, x \rangle \cdot x_i.$$

Man nennt sie die ***orthogonale Projektion*** von V *auf* U.

Beweis. Für ein zunächst festes $x \in V$ seien $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ die Koordinaten des Vektors $f(x) \in U$ bezüglich der gegebenen Basis $(x_1, \dots, x_n)$, also $f(x) = \lambda_1 x_1 + \cdots + \lambda_n x_n$. Dann gilt

$$\begin{aligned} x - f(x) \perp U &\overset{21.27}{\Leftrightarrow} \langle x_i, x - f(x) \rangle = 0 \text{ für alle } i = 1, \dots, n \\ &\Leftrightarrow \langle x_i, x - \lambda_1 x_1 - \cdots - \lambda_n x_n \rangle = 0 \text{ für alle } i = 1, \dots, n \\ &\overset{(*)}{\Leftrightarrow} \langle x_i, x \rangle - \lambda_i = 0 \text{ für alle } i = 1, \dots, n \\ &\Leftrightarrow f(x) = \langle x_1, x \rangle\, x_1 + \cdots + \langle x_n, x \rangle\, x_n, \end{aligned}$$

wobei wir in $(*)$ wieder die Orthonormalität der Basis verwendet haben. Da diese Vorschrift offensichtlich linear in x ist, ist das Lemma damit gezeigt. □ 50

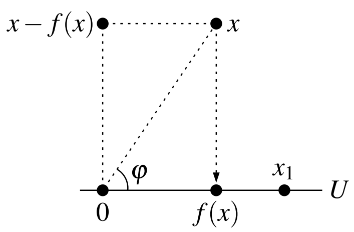
Bemerkung 21.30 (Geometrische Deutung orthogonaler Projektionen)**.** Orthogonale Projektionen haben eine einfache geometrische Deutung, die im Bild unten rechts für einen eindimensionalen Unterraum $U = \operatorname{Lin}(x_1)$ von $V = \mathbb{R}^2$ mit $||x_1|| = 1$ dargestellt ist.

Die Abbildung konstruiert in diesem Fall zu einem $x \in V$ das Lot auf den Unterraum U; der so entstehende Punkt ist die orthogonale Projektion $f(x)$. Nach Konstruktion 21.22 ist dann nämlich $f(x) = \lambda\, x_1$ mit

$$\lambda = ||x|| \cos\varphi = ||x|| \cdot \frac{\langle x_1, x \rangle}{||x_1|| \cdot ||x||} = \langle x_1, x \rangle,$$

und damit wie in Lemma 21.29

$$f(x) = \langle x_1, x \rangle\, x_1.$$



Beachte, dass die Differenz $x - f(x)$ dann wie in Lemma 21.29 senkrecht zu x_1 ist. Wir können also einen Vektor konstruieren, der auf (allen Vektoren von) U senkrecht steht, indem wir von einem beliebigen Vektor $x \notin U$ seine orthogonale Projektion auf U abziehen. Dies ist die Grundidee des folgenden Verfahrens zur Bestimmung von Orthonormalbasen.

Satz 21.31 (Gram-Schmidtsches Orthonormalisierungsverfahren). *Es seien V ein endlich-dimensionaler Vektorraum mit Skalarprodukt und $U \leq V$ ein Untervektorraum. Dann lässt sich jede Orthonormalbasis von U zu einer Orthonormalbasis von V ergänzen.*

Insbesondere besitzt also jeder endlich-dimensionale Vektorraum mit Skalarprodukt eine Orthonormalbasis.

Beweis. Der Beweis dieses Satzes ist konstruktiv und erlaubt daher auch eine einfache (iterative) Konstruktion solcher Orthonormalbasen.

Es sei $(x_1, \dots, x_k)$ die gegebene Orthonormalbasis von U. Ist bereits $U = V$, so sind wir natürlich fertig. Ansonsten führen wir die folgenden Schritte aus:

- Wähle einen beliebigen Vektor $x \in V \backslash U$.
- Wir subtrahieren von x die orthogonale Projektion von x auf U und erhalten so nach Lemma 21.29 den auf $x_1, \dots, x_k$ senkrecht stehenden Vektor

  $$y_{k+1} := x - \sum_{i=1}^{k} \langle x_i, x \rangle\, x_i.$$

  Wegen $x \notin U = \operatorname{Lin}(x_1, \dots, x_k)$ ist dies nicht der Nullvektor, und damit ist die Familie $(x_1, \dots, x_k, y_{k+1})$ nach Lemma 21.28 linear unabhängig.

- Normieren wir y_{k+1} nun zu
$$x_{k+1} := \frac{y_{k+1}}{||y_{k+1}||},$$
so ist $(x_1,\dots,x_{k+1})$ also eine Orthonormalbasis eines Unterraums $U' \supsetneq U$.

Ist jetzt $U' = V$, so sind wir fertig. Ansonsten setzen wir das obige Verfahren iterativ mit dem neuen Unterraum $U' = \operatorname{Lin}(x_1,\dots,x_{k+1})$ fort, bis wir genügend Vektoren gefunden haben. $\square$

Beispiel 21.32. Wir wollen eine Orthonormalbasis von $\mathbb{R}^2$ für das Skalarprodukt
$$\langle x,y\rangle = x^{\mathsf{T}} \begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix} y$$
aus Beispiel 21.15 (b) bestimmen. Beachte zunächst, dass die Standardbasis (e_1,e_2) *keine* Orthonormalbasis ist, denn es ist z. B.
$$\langle e_1,e_2\rangle = (\,1\;\;0\,) \begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 1 \neq 0.$$
Wir wenden also das Gram-Schmidtsche Orthonormalisierungsverfahren aus Satz 21.31 an. Den ersten Vektor können wir dabei (ungleich 0) beliebig wählen, z. B. $y_1 = e_1$. Wegen $\langle e_1,e_1\rangle = 1$, also $||e_1|| = 1$, erhalten wir nach Normieren mit den Notationen wie im Beweis des Satzes
$$x_1 = \frac{y_1}{||y_1||} = \frac{e_1}{1} = e_1.$$
Für den zweiten Vektor starten wir mit einem beliebigen Element von $\mathbb{R}^2 \setminus \operatorname{Lin} x_1$, z. B. $x = e_2$, und subtrahieren von ihm die orthogonale Projektion auf $\operatorname{Lin} e_1$:
$$y_2 = e_2 - \langle x_1,e_2\rangle\, x_1 = e_2 - \langle e_1,e_2\rangle\, e_1 = e_2 - 1\,e_1 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$
Wegen
$$||y_2||^2 = \langle y_2,y_2\rangle = (\,-1\;\;1\,) \begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \end{pmatrix} = 3$$
erhalten wir nach Normieren also die Orthonormalbasis
$$\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{3}} \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right)$$
von $\mathbb{R}^2$ bezüglich des gegebenen Skalarprodukts.

Bemerkung 21.33 (Gram-Schmidt als Normalformensatz)**.** Es sei V ein endlich erzeugter Vektorraum mit Skalarprodukt $\langle x,y\rangle = b(x,y)$. Ist dann $B = (x_1,\dots,x_n)$ gemäß Satz 21.31 eine Orthonormalbasis von V, so ist die Gramsche Matrix von b bezüglich B nach Folgerung 21.6 gerade
$$A_b^B = \big(\langle x_i,x_j\rangle\big)_{i,j} = E_n,$$
denn es ist ja $\langle x_i,x_j\rangle$ gleich 1 für $i = j$ und 0 für $i \neq j$. In Analogie zu den Normalformaussagen aus Satz 16.43 und Folgerung 20.15 kann man die Existenz von Orthonormalbasen damit auch so auffassen, dass es zu jedem Skalarprodukt eine Basis gibt, bezüglich der die Gramsche Matrix die Einheitsmatrix ist, also diese sehr einfache Form hat.

Da sich Gramsche Matrizen unter Basiswechsel wie in Lemma 21.7 bzw. Konstruktion 21.18 verhalten, ist die entsprechende Aussage in Matrixform gemäß Lemma 21.11 also, dass es zu jeder positiv definiten, symmetrischen (im Fall $\mathbb{K} = \mathbb{R}$) bzw. hermiteschen (im Fall $\mathbb{K} = \mathbb{C}$) Matrix $A \in \operatorname{Mat}(n \times n, \mathbb{K})$ eine invertierbare Matrix $T \in \operatorname{GL}(n,\mathbb{K})$ gibt mit $\overline{T}^{\mathsf{T}} A\, T = E$: Man muss in die Spalten von T einfach eine Orthonormalbasis von $\mathbb{K}^n$ bezüglich des Skalarprodukts $\langle x,y\rangle = \overline{x}^{\mathsf{T}} A\, y$ schreiben. Für das Skalarprodukt und die Orthonormalbasis aus Beispiel 21.32 ergibt sich also z. B.
$$T^{\mathsf{T}} A\, T = E \quad \text{mit} \quad A = \begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix} \quad \text{und} \quad T = \begin{pmatrix} 1 & -\frac{1}{\sqrt{3}} \\ 0 & \frac{1}{\sqrt{3}} \end{pmatrix},$$
was man durch direkte Berechnung des Matrixprodukts auch sofort bestätigen kann.

Eine Verallgemeinerung dieser Aussage auf nicht notwendig positiv definite symmetrische Bilinearformen bzw. hermitesche Sesquilinearformen werden wir in Satz 22.38 bzw. Bemerkung 22.39 kennenlernen.

Folgerung 21.34 (Determinante hermitescher und positiv definiter Matrizen)**.** *Für jede hermitesche Matrix* $A \in \mathrm{Mat}(n \times n, \mathbb{C})$ *gilt:*

(a) $\det A \in \mathbb{R}$.

(b) *Ist A zusätzlich positiv definit, so ist sogar* $\det A \in \mathbb{R}_{>0}$.

Beweis.

(a) Ist $\overline{A}^{\mathsf{T}} = A$, so folgt $\overline{\det A} = \det \overline{A} = \det \overline{A}^{\mathsf{T}} = \det A$ und damit $\det A \in \mathbb{R}$.

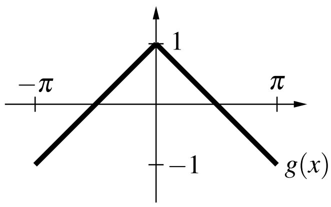
(b) Nach Bemerkung 21.33 gibt es eine invertierbare Matrix T mit $\overline{T}^{\mathsf{T}} A T = E$. Damit ist

$$1 = \det E = \det \overline{T}^{\mathsf{T}} \cdot \det A \cdot \det T = \det A \cdot \det \overline{T} \cdot \det T = \det A \cdot |\det T|^2,$$

woraus $\det A > 0$ folgt. □

Aufgabe 21.35. Es sei U ein endlich-dimensionaler Unterraum eines Vektorraums V mit Skalarprodukt. Zeige, dass die orthogonale Projektion eines Vektors $x \in V$ auf U der eindeutig bestimmte Punkt $y \in U$ ist, für den der Abstand $||y - x||$ von x zu y minimal ist.

Aufgabe 21.36. Es sei $V = C^0([-\pi, \pi])$ der reelle Vektorraum aller stetigen Funktionen auf dem Intervall $[-\pi, \pi]$ mit dem Skalarprodukt $\langle f, g \rangle := \int_{-\pi}^{\pi} f(x) g(x)\, dx$ wie in Beispiel 21.15 (c). Wir betrachten darin das Element $g \in V$ definiert durch $g(x) := 1 - \frac{2}{\pi} \cdot |x|$.



Für $n \in \mathbb{N}_{>0}$ seien weiterhin $f_n \in V$ mit $f_n(x) := \frac{1}{\sqrt{\pi}} \cdot \cos(nx)$, und $U_n := \mathrm{Lin}(f_1, \dots, f_n) \leq V$.

(a) Zeige, dass $(f_1, \dots, f_n)$ für alle n eine Orthonormalbasis von U_n ist.

(b) Berechne für alle $n \in \mathbb{N}_{>0}$ die orthogonale Projektion g_n von g auf U_n (die nach Aufgabe 21.35 also g in U_n am besten approximiert).

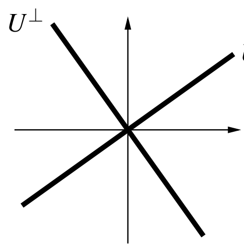
(c) Zeichne die Funktionen g_n für kleine n mit einem Computerprogramm und vergleiche sie mit der ursprünglichen Funktion g.

Wir wollen nun noch zwei Anwendungen von Orthonormalbasen betrachten. Die erste betrifft Komplemente von Unterräumen in endlich erzeugten Vektorräumen: Wir wissen ja nach Beispiel 15.9 und Satz 15.10 bereits, dass solche Komplemente zwar immer existieren, aber in der Regel nicht eindeutig bestimmt sind. Falls im zugrundeliegenden Vektorraum aber ein Skalarprodukt gegeben ist, wollen wir jetzt sehen, dass es zu jedem Unterraum stets ein besonderes Komplement gibt — nämlich das *orthogonale Komplement* — das immer eindeutig bestimmt ist.

Definition 21.37 (Orthogonales Komplement)**.** Es seien V ein Vektorraum mit Skalarprodukt und $U \leq V$ ein Unterraum von V. Dann nennen wir die Menge

$$\begin{aligned} U^{\perp} &:= \{x \in V : x \perp U\} \\ &= \{x \in V : \langle x, y \rangle = 0 \text{ für alle } y \in U\} \end{aligned}$$

das **orthogonale Komplement** von U. (Das Bild rechts illustriert dies für den Vektorraum $\mathbb{R}^2$ mit dem Standardskalarprodukt.)



Satz 21.38. *Es seien V ein endlich-dimensionaler Vektorraum mit Skalarprodukt und* $U \leq V$ *ein Unterraum. Dann ist das orthogonale Komplement* $U^{\perp}$ *ein Komplement von U im Sinne von Definition 15.7, d. h. es gilt* $V = U \oplus U^{\perp}$.

Beweis. Nach Satz 21.31 können wir eine Orthonormalbasis $(x_1,\dots,x_k)$ von U finden und zu einer Orthonormalbasis $(x_1,\dots,x_n)$ von V ergänzen. Nach Bemerkung 15.11 ist $U' := \operatorname{Lin}(x_{k+1},\dots,x_n)$ dann ein Komplement von U. Es genügt also zu zeigen, dass $U^\perp = U'$ gilt: Für einen Vektor $x \in V$ mit der Darstellung $x = \lambda_1 x_1 + \cdots + \lambda_n x_n$ gilt

$$\begin{aligned} x \in U^\perp &\overset{21.27}{\Leftrightarrow} \langle x_i, \lambda_1 x_1 + \cdots + \lambda_n x_n \rangle = 0 \text{ für alle } i = 1,\dots,k \\ &\Leftrightarrow \lambda_i \,||x_i||^2 = 0 \text{ für alle } i = 1,\dots,k \\ &\Leftrightarrow \lambda_i = 0 \text{ für alle } i = 1,\dots,k \\ &\Leftrightarrow x \in U'. \end{aligned}$$

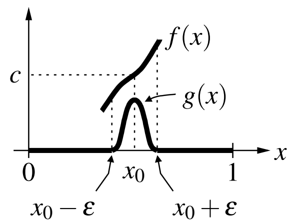
□

Bemerkung 21.39. Der Beweis von Satz 21.38 zeigt auch, wie man das orthogonale Komplement eines Unterraums U in V berechnen kann: Man ergänzt eine Orthonormalbasis $(x_1,\dots,x_k)$ von U zu einer Orthonormalbasis $(x_1,\dots,x_n)$ von V; die dabei neu hinzu genommenen Vektoren $(x_{k+1},\dots,x_n)$ bilden dann eine Orthonormalbasis von $U^\perp$.

Da in diesem Fall natürlich auch $x_1,\dots,x_k$ die Vektoren $x_{k+1},\dots,x_n$ zu einer Orthonormalbasis von V ergänzen, folgt aus dieser Beschreibung auch sofort, dass $(U^\perp)^\perp = U$.

Beispiel 21.40 (Orthogonale Komplemente in unendlich-dimensionalen Vektorräumen)**.** Als „abschreckendes Beispiel" dafür, dass Satz 21.38 nicht so selbstverständlich ist, wie er vielleicht scheint, wollen wir kurz zeigen, dass dieser Satz für unendlich-dimensionale Vektorräume im Allgemeinen *falsch* ist, dass das orthogonale Komplement dann also nicht unbedingt immer ein Komplement im Sinne von Definition 15.7 ist. Dazu sei $V = C^0([0,1])$ der Vektorraum aller stetigen Funktionen auf dem Intervall $[0,1]$ mit dem in Beispiel 21.15 (c) betrachteten Skalarprodukt $\langle f,g\rangle = \int_0^1 f(x)g(x)\,dx$. Wir betrachten hierin den Unterraum $U \subset V$ aller *differenzierbaren* Funktionen; offensichtlich ist $U \subsetneq V$. Wir werden aber zeigen, dass $U^\perp = \{0\}$ ist, so dass also insbesondere $U + U^\perp = U \neq V$, d. h. $U^\perp$ kein Komplement von U ist.

Dazu sei $f \in V\backslash\{0\}$ beliebig; wir werden zeigen, dass $f \notin U^\perp$. Wegen $f \neq 0$ gibt es ein $x_0 \in [0,1]$ mit $f(x_0) =: c \neq 0$. Wir können ohne Einschränkung annehmen, dass $c > 0$ ist. Da f stetig ist, gibt es ein $\varepsilon > 0$, so dass $f(x) \geq \frac{c}{2}$ für alle x mit $|x - x_0| < \varepsilon$ ist. Es sei nun $g \in U$ eine differenzierbare Funktion wie im Bild. Dann ist offensichtlich $\langle f,g\rangle = \int_0^1 f(x)g(x)\,dx \geq \frac{c}{2}\int_0^1 g(x)\,dx > 0$, denn die Funktion g ist überall nicht negativ und hat außerdem eine nicht verschwindende Fläche unter ihrem Graphen.



Wir haben damit also ein $g \in U$ mit $\langle f,g\rangle \neq 0$ gefunden. Damit ist $f \notin U^\perp$. Da f beliebig war, folgt also wie behauptet $U^\perp = \{0\}$.

Aufgabe 21.41. Es seien $n \in \mathbb{N}$ und V_n der Vektorraum aller reellen Polynome vom Grad höchstens n.

(a) Für welche $m \in \mathbb{N}$ definiert $\langle f,g\rangle := \sum_{i=0}^{m} f(i)\,g(i)$ ein Skalarprodukt auf V_n?

(b) Berechne für dieses Skalarprodukt eine Orthonormalbasis von V_2 im Fall $m = n = 2$.

Die zweite Anwendung der Existenz von Orthonormalbasen aus Satz 21.31 ist das folgende einfache Kriterium für die positive Definitheit einer Matrix.

Satz 21.42 (Hurwitz-Kriterium). *Es sei $A = (a_{i,j})_{i,j} \in \operatorname{Mat}(n \times n, \mathbb{K})$ eine symmetrische (im Fall $\mathbb{K} = \mathbb{R}$) bzw. hermitesche (im Fall $\mathbb{K} = \mathbb{C}$) Matrix. Für $k = 1,\dots,n$ bezeichnen wir mit $A_k = (a_{i,j})_{i,j=1,\dots,k} \in \operatorname{Mat}(k \times k, \mathbb{K})$ die Matrizen, die man erhält, wenn man von A nur die ersten k Zeilen und Spalten betrachtet. (Beachte, dass ihre Determinanten nach Folgerung 21.34 (a) in jedem Fall reell sind.)*

Dann ist A genau dann positiv definit, wenn $\det A_k > 0$ für alle $k = 1,\dots,n$.

Beweis. Es sei b die zu A gehörige symmetrische Bilinearform bzw. hermitesche Sesquilinearform auf $\mathbb{K}^n$, also $b(x,y) = \overline{x}^\mathsf{T} A y$ für alle $x, y \in \mathbb{K}^n$.

„$\Rightarrow$“ Ist A positiv definit und b damit ein Skalarprodukt, so ist nach Beispiel 21.15 (d) auch die Einschränkung von b auf den Unterraum $\mathrm{Lin}(e_1, \dots, e_k)$ ein Skalarprodukt. Die zugehörige Gramsche Matrix $(b(e_i, e_j))_{i,j=1,\dots,k} = A_k$ ist also ebenfalls positiv definit und hat damit nach Folgerung 21.34 (b) eine positive Determinante.

„$\Leftarrow$“ Wir zeigen die Aussage mit Induktion über n, der Induktionsanfang für $n = 1$ ist trivial.

Für den Induktionsschritt $n \to n+1$ bemerken wir zunächst, dass nach Annahme insbesondere $\det A_k > 0$ für $k = 1, \dots, n$ gilt. Damit ist A_n nach Induktionsvoraussetzung positiv definit, und wir können nach Satz 21.31 eine Orthonormalbasis $(x_1, \dots, x_n)$ für das zugehörige Skalarprodukt finden, das gerade die Einschränkung von b auf $\mathrm{Lin}(e_1, \dots, e_n)$ ist. Bilden wir nun wie im Gram-Schmidtschen Orthonormalisierungsverfahren in Satz 21.31 den Vektor

$$x_{n+1} := e_{n+1} - \sum_{i=1}^{n} b(x_i, e_{n+1})\, x_i \qquad \notin \mathrm{Lin}(x_1, \dots, x_n),$$

so gilt wieder

$$b(x_j, x_{n+1}) = b(x_j, e_{n+1}) - \sum_{i=1}^{n} b(x_i, e_{n+1}) b(x_j, x_i) = b(x_j, e_{n+1}) - b(x_j, e_{n+1}) = 0$$

für alle $j = 1, \dots, n$, da $(x_1, \dots, x_n)$ orthonormal ist. Damit hat die Gramsche Matrix von b bezüglich der Basis $B = (x_1, \dots, x_{n+1})$ die Form

$$A_b^B = \overline{T}^\mathsf{T} A\, T = (b(x_i, x_j))_{i,j} = \left(\begin{array}{c|c} E_n & 0 \\ \hline 0 & b(x_{n+1}, x_{n+1}) \end{array} \right),$$

wobei wir wie üblich $T = (\, x_1 \mid \cdots \mid x_{n+1} \,)$ gesetzt haben. Insbesondere folgt damit

$$b(x_{n+1}, x_{n+1}) = \det A_b^B = \det\left(\overline{T}^\mathsf{T} A\, T\right) = \overline{\det T} \cdot \det A \cdot \det T = |\det T|^2 \cdot \det A > 0.$$

Ist nun $x = \lambda_1 x_1 + \dots + \lambda_{n+1} x_{n+1} \neq 0$ beliebig, so ergibt sich daraus

$$b(x,x) = \sum_{i,j=1}^{n+1} \overline{\lambda_i} \lambda_j b(x_i, x_j) = |\lambda_1|^2 + \dots + |\lambda_n|^2 + \underbrace{b(x_{n+1}, x_{n+1})}_{>0} |\lambda_{n+1}|^2 > 0.$$

Also ist b und damit auch A positiv definit. □

Bemerkung 21.43. Die Determinanten von quadratischen Teilmatrizen einer gegebenen Matrix A werden oft auch *Minoren* von A genannt, die in Satz 21.42 auftretenden Determinanten $\det A_k$ bezeichnet man als *Hauptminoren* von A. Das Hurwitz-Kriterium ist daher auch unter dem Namen *Hauptminorenkriterium* bekannt.

51

Beispiel 21.44. Wir betrachten noch ein letztes Mal die symmetrische reelle Matrix

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

aus Beispiel 21.15 (b), die wir dort schon durch eine direkte Rechnung als positiv definit erkannt haben. Mit dem Hurwitz-Kriterium könnten wir dies aber nun auch noch einfacher zeigen, indem wir (mit den Notationen von Satz 21.42) die beiden Determinanten

$$\det A_1 = \det(1) = 1 > 0 \qquad \text{und} \qquad \det A_2 = \det A = 3 > 0$$

berechnen und sehen, dass sie beide positiv sind.

Aufgabe 21.45. Für ein $\lambda \in \mathbb{R}$ sei $A = \begin{pmatrix} \lambda & 1 & 2 \\ 1 & 2 & 3 \\ 2 & 3 & 7 \end{pmatrix}$.

(a) Für welche λ ist A positiv definit? Für welche λ ist A negativ definit?

(b) Wir betrachten nun $V = \mathbb{R}^3$ als euklidischen Vektorraum mit dem Skalarprodukt b_A für den Wert $\lambda = 1$. Berechne zu $U = \mathrm{Lin}(e_1) \leq V$ eine Orthonormalbasis des orthogonalen Komplements $U^\perp$.