Kapitel 2: Lineare Algebra

2.1 Der $\mathbb{R}^n$: Vektorraum- und euklidische Struktur

Def.: $$\mathbb{R}^n := \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} : x_1, \dots, x_n \in \mathbb{R} \right\}$$

= Menge aller geordneten n-Tupel der Form $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, sogenannter Spaltenvektoren. Die Einträge x_k nennen wir die Komponenten des Vektors $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$.

"Geordnet" bedeutet, dass durch Vertauschung zweier Komponenten im allgemeinen ein anderer Vektor entsteht, z. B. ist $\begin{pmatrix} 2 \\ 1 \end{pmatrix} \neq \begin{pmatrix} 1 \\ 2 \end{pmatrix}$.

Bem.: • $\mathbb{R}^n = \underbrace{\mathbb{R} \times \mathbb{R} \times \dots \times \mathbb{R}}_{n\text{-mal}}$, das n-fache kartesische Produkt von $\mathbb{R}$ mit sich selbst.

• bekannt: $n = 1$: $\mathbb{R}^1 = \mathbb{R}$, geometrisch interpretierbar als eine Gerade ("die Zahlengerade"),

• vermutlich ebenso bekannt

$$\mathbb{R}^2 = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} : x, y \in \mathbb{R} \right\},$$ geometrisch interpretierbar als Ebene,

$$\mathbb{R}^3 = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} : x, y, z \in \mathbb{R} \right\},$$ der dreidimensionale Raum.

Die Identifikation von Punkten im Raum bzw. in der Ebene mit Zahlentupeln erlaubt das Rechnen mit geometrischen Objekten.

Besonders ausgezeichnete Elemente des $\mathbb{R}^n$ sind die sog. "kanonischen Basisvektoren" oder "kanonischen Einheitsvektoren" e_k, definiert durch

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots \quad e_k = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \begin{matrix} \\ \\ \\ \leftarrow k\text{-te Stelle} \\ \\ \\ \\ \end{matrix} .$$

Welche Rechenoperationen lassen sich für die Elemente des $\mathbb{R}^n$ in naheliegender Weise erklären?

Def.: Für $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$, $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n$ und $\lambda \in \mathbb{R}$

definieren wir

$$\underbrace{x + y}_{+ \text{ in } \mathbb{R}^n \times \mathbb{R}^n} := \begin{pmatrix} x_1 + y_1 \\ \vdots \\ \underbrace{x_n + y_n}_{+ \text{ in } \mathbb{R} \times \mathbb{R}} \end{pmatrix} \quad (\text{"Vektoraddition"})$$

und

$$\underbrace{\lambda x}_{\cdot \text{ in } \mathbb{R} \times \mathbb{R}^n} = \begin{pmatrix} \lambda x_1 \\ \vdots \\ \underbrace{\lambda x_n}_{\cdot \text{ in } \mathbb{R} \times \mathbb{R}} \end{pmatrix} \quad (\text{"Multiplikation mit einem Skalar"})$$

Bem.: Da $\lambda = -1$ hier zugelassen ist, ist mit dieser Definition auch die Subtraktion zweier Vektoren erklärt. Hingegen ist es i. allgemeinen nicht möglich, den Quotienten aus zwei Vektoren zu bilden.

Die Vektoraddition und skalare Multiplikation genügen den folgenden

Rechenregeln:

(V1) $x+y=y+x$ (Kommutativität)

(V2) $x+(y+z)=(x+y)+z$ (Assoziativgesetz)

(V3) $x+0=x$ für $0=\begin{pmatrix}0\\ \vdots\\ 0\end{pmatrix}$ (Nullvektor, neutrales Element)

(V4) Zu $x=\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}\in\mathbb{R}^n$ existiert $-x=\begin{pmatrix}-x_1\\ \vdots\\ -x_n\end{pmatrix}\in\mathbb{R}^n$,

so daß $x+(-x)=0$ (negatives Element).

Die Regeln (V1) bis (V4) werden häufig zu der Aussage zusammengefaßt, daß $(\mathbb{R}^n,+)$ eine "kommutative Gruppe" ist (man sagt auch "abelsche Gruppe"). Desweiteren gelten für die skalare Multiplikation

(V5) $\lambda(\mu x)=(\lambda\mu)x$

(V6) $1\cdot x = x$

und ferner die Distributivgesetze

(V7) $\lambda(x+y)=\lambda x+\lambda y$

(V8) $(\lambda+\mu)x=\lambda x+\mu x$

Diese Rechenregeln gelten für alle $x,y,z\in\mathbb{R}^n$ und $\lambda,\mu\in\mathbb{R}$. Sie sind leicht aus den entsprechenden Eigenschaften der reellen Zahlen herzuleiten. Alle weiteren

Rechenregeln für $+$ und $\cdot$ im $\mathbb{R}^n$ (z.B. die Eindeutigkeit des Nullelements und des Negativen) lassen sich hieraus herleiten. Daher nennt man (V1) bis (V8) auch Axiome, genauer: Vektorraumaxiome. Eine Menge, die diesen Forderungen genügt, nennt man einen $\mathbb{R}$-Vektorraum:

<u>Def.</u> (Vektorraum): Ein Tripel $(V, +, \cdot)$ bestehend aus

- einer Menge $V \neq \emptyset$,
- einer inneren Verknüpfung $+ : V \times V \to V$,

  $(x, y) \mapsto x + y$ und
- einer skalaren Multiplikation $\cdot : \mathbb{R} \times V \to V$,

  $(\lambda, x) \mapsto \lambda x$

heißt ein <u>Vektorraum</u> über $\mathbb{R}$ (kurz: $\mathbb{R}$-Vektorraum), falls die Rechenregeln (V1) bis (V8) gelten. Die Elemente eines Vektorraums nennen wir <u>Vektoren</u>.

Wie festgestellt, ist $(\mathbb{R}^n, +, \cdot)$ mit komponentenweiser Addition und skalarer Multiplikation ein $\mathbb{R}$-VR. Ein weiteres Bsp. ist das folgende.

Bsp.: $V := \{f : \mathbb{R} \to \mathbb{R},\ f \text{ Abbildung}\}$

Hierbei werden Addition und skalare Multiplikation punktweise erklärt, d.h. man setzt

$$(f+g)(x) := f(x)+g(x) \quad \text{und} \quad (\lambda f)(x) := \lambda \cdot f(x),$$

wobei auf der rechten Seite + und · in $\mathbb{R}$ gemeint sind.

Weitere Beispiele werden durch Untervektorräume geliefert:

Def. Eine Teilmenge $U \neq \emptyset$ eines $\mathbb{R}$-Vektorraums V heißt ein <u>Untervektorraum</u>, falls gilt

$$x, y \in U \text{ und } \lambda, \mu \in \mathbb{R} \Rightarrow \lambda x + \mu y \in U.$$

Bsp. 1: Die Untervektorräume des $\mathbb{R}^2$ sind (nach Dimension geordnet):

(a) $\{0\}$ ("Nullvektorraum"),

(b) für jedes $x \in \mathbb{R}^2 \setminus \{0\}$ ist

$$G_x := \{\lambda x : \lambda \in \mathbb{R}\}$$

ein Untervektorraum von $\mathbb{R}^2$. Es handelt sich dabei um eine Gerade durch den Nullpunkt, der auch Ursprung genannt wird.

(c) $\mathbb{R}^2$ (Konvention: stets wird V selbst auch als Untervektorraum bezeichnet!)

Bem.: Geraden, die nicht durch den Nullpunkt gehen, sind keine Untervektorräume. In diesem Fall spricht man von affinen Teilräumen, das sind sozusagen "verschobene Untervektorräume".

Bsp. 2.: Untervektorräume des $\mathbb{R}^3$

(a) $\{0\}$,

(b) $G_x := \{\lambda x : \lambda \in \mathbb{R}\}$, dabei $x \in \mathbb{R}^3 \setminus \{0\}$ fest,

(c) $E_{x,y} := \{\lambda x + \mu y : \lambda, \mu \in \mathbb{R}\}$, dabei $x, y \in \mathbb{R}^3 \setminus \{0\}$

fest und nicht Vielfache voneinander.

Dies sind die Ebenen durch den Nullpunkt.

(d) $\mathbb{R}^3$

Bsp. 3: Ein Beispiel für einen UVR des Vektorraums

$$V = \{f : \mathbb{R} \to \mathbb{R}\}$$

aller Funktionen f von $\mathbb{R}$ nach $\mathbb{R}$ wird gegeben durch die Polynomabbildungen

$$P := \{p : \mathbb{R} \to \mathbb{R} : \exists n \in \mathbb{N},\ a_0, \ldots, a_n \in \mathbb{R}, \text{ so daß } p(x) = \sum_{k=0}^{n} a_k x^k\}.$$

Zwei weitere wichtige Beispiele von Untervektorräumen sind eng verknüpft mit dem Begriff der linearen Abbildung.

Def.: Es seien V und W $\mathbb{R}$-Vektorräume. Eine Abbildung $F: V \to W$ heißt <u>linear</u>, falls für alle $x, y \in V$ und $\lambda, \mu \in \mathbb{R}$ gilt

$$F(\lambda x + \mu y) = \lambda F(x) + \mu F(y).$$

Bsp. 4. Die linearen Abbildungen $F: \mathbb{R} \to \mathbb{R}$ sind alle der Form $F(x) = a x$, dabei $a \in \mathbb{R}$ fest.

Für $b \neq 0$ ist die Abbildung

$$\varphi: \mathbb{R} \to \mathbb{R}, \quad x \mapsto \varphi(x) = ax + b$$

<u>keine</u> lineare Abbildung im Sinne der obigen Definition, da $\varphi(0) = b \neq 0$. Man nennt φ eine affin-lineare Abbildung.

Bsp. 5: Die linearen Abbildungen $F: \mathbb{R}^2 \to \mathbb{R}^2$ haben alle die Gestalt

$$F\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix},$$

dabei sind $a, b, c, d \in \mathbb{R}$ feste Zahlen.

Die Nullstellenmenge einer linearen Abbildung nennt man ihren Kern. Genauer:

Def.: Es seien V, W $\mathbb{R}$-Vektorräume und $F: V \to W$ linear. Dann heißt

$$\mathrm{Ker}(F) = \{x \in V : F(x) = 0\}$$

der <u>Kern</u> von F.

<u>Bem.</u> (1) F ist injektiv $\Leftrightarrow$ $\mathrm{Ker}(F) = \{0\}$

(2) $\mathrm{Ker}(F)$ ist ein UVR von V.

Begründung: (1) Für lineare Abbildungen ist stets $F(0) = 0$. Ist F injektiv, gilt $F(x) \neq 0$ $\forall x \in V \setminus \{0\}$. Dies zeigt "$\Rightarrow$". Umgekehrt sei $\mathrm{Ker}(F) = 0$. Dann folgt aus $F(x) = F(y)$, daß $F(x-y) = F(x) - F(y) = 0$, also $x - y = 0$ bzw. $x = y$. D.h. F ist injektiv.

(2) Sind $x, y \in \mathrm{Ker}(F)$ und $\lambda, \mu \in \mathbb{R}$, so folgt

$F(\lambda x + \mu y) = \lambda F(x) + \mu F(y) = 0$, d.h.

$\lambda x + \mu y \in \mathrm{Ker}(F)$.

<u>Def</u>: Es seien V, W $\mathbb{R}$-Vektorräume und $F: V \to W$ linear. Dann heißt

$$R(F) = \{F(x) : x \in V\}$$

das Bild von V unter F (oder kurz: das Bild von F).

(R für range)

Bem.: (1) $R(F)$ ist ein UVR von W

(2) F ist surjektiv genau dann, wenn $R(F) = W$.

Begründung von (1): $\lambda F(x) + \mu F(y) = F(\lambda x + \mu y)$.

Die Begriffe Kern und Bild sollen an einem konkreten Beispiel veranschaulicht werden:

<u>Bsp.</u> **6**: Es sei $F: \mathbb{R}^2 \to \mathbb{R}^2$ definiert durch

$$F\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x+y \\ -3x-3y \end{pmatrix}.$$ Dann ist F linear.

Es ist $F\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \iff \begin{pmatrix} x+y \\ -3x-3y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \iff x = -y$

Damit gilt:

$$\operatorname{Ker}(F) = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : x+y=0 \right\} = \left\{ \lambda \begin{pmatrix} 1 \\ -1 \end{pmatrix} : \lambda \in \mathbb{R} \right\},$$

also handelt es sich um eine Gerade durch den Nullpunkt. Auch bei $R(F)$ handelt es sich um eine Ursprungsgerade, denn wir haben:

$$\begin{pmatrix} u \\ v \end{pmatrix} \in R(F) \iff \exists\, x, y \in \mathbb{R} \text{ so daß } \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} x+y \\ -3(x+y) \end{pmatrix}$$

$$\iff v = -3u, \text{ also}$$

$$R(F) = \left\{ \begin{pmatrix} u \\ v \end{pmatrix} \in \mathbb{R}^2 : v = -3u \right\} = \left\{ \lambda \begin{pmatrix} 1 \\ -3 \end{pmatrix} : \lambda \in \mathbb{R} \right\}.$$

(Skizzen!)

Die Vektorraumstruktur des $\mathbb{R}^n$ erlaubt die Addition und skalare Multiplikation von Vektoren. Sie liefert den geeigneten Rahmen zur Beschreibung linearer Abbildungen. Sie ist hingegen nicht ausreichend zur Messung von Abständen und Winkeln zwischen Vektoren. Dazu dient ein weiteres Strukturelement, das Skalarprodukt.

Def.: Für $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ und $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n$ heißt

$$\langle x, y \rangle := x_1 y_1 + \dots + x_n y_n = \sum_{k=1}^{n} x_k y_k$$

das Skalarprodukt von x und y.

Weitere Schreibweisen: $\langle x, y \rangle = x \cdot y = x^T y$.

Im letzten Ausdruck ist x^T der "Zeilenvektor" $x^T = (x_1, \dots, x_n)$. T steht allgemein für "transponiert", d.h. aus einer Zeile wird eine Spalte und umgekehrt. Ausgeschrieben haben wir also

$$x^T y = (x_1, \dots, x_n) \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix},$$ und wir multiplizieren "Zeile mal Spalte".

Das Skalarprodukt wird manchmal auch als inneres Produkt bezeichnet.

Satz 1 (Eigenschaften des Skalarprodukts): Es seien $x, y, z \in \mathbb{R}^n$ und $\lambda \in \mathbb{R}$. Dann gelten: (2.11)

$$\left.\begin{array}{ll} (S1) & \langle x+y, z\rangle = \langle x, z\rangle + \langle y, z\rangle \\ & \langle \lambda x, y\rangle = \lambda \langle x, y\rangle \end{array}\right\} \text{„Linearität in der ersten Komponente"}$$

$$(S2) \quad \langle x, y\rangle = \langle y, x\rangle \qquad \text{„Symmetrie"}$$

$$(S3) \quad \langle x, x\rangle > 0 \quad \forall x \in \mathbb{R}^n \setminus \{0\} \qquad \text{„Definitheit"}$$

Folgerungen:

(1) Linearität in der zweiten Komponente:

$$\langle x, y+z\rangle = \langle x, y\rangle + \langle x, z\rangle, \quad \langle x, \lambda y\rangle = \lambda \langle x, y\rangle.$$

Folgt aus (S1) und (S2).

(2) $\langle x, x\rangle \geq 0$ mit Gleichheit nur für $x = 0$.

Ergibt sich aus (S3) und $\langle 0, 0\rangle = 0$, letzteres gilt nach (S1) mit $\lambda = 0$.

In ähnlicher Weise, wie wir beim Begriff des Vektorraums vom $\mathbb{R}^n$ auf einen allgemeinen Vektorraum V abstrahiert haben, geschieht dies auch beim Begriff des Skalarprodukts:

Def.: Es sei V ein $\mathbb{R}$-Vektorraum und

$$\langle\ ,\ \rangle : V \times V \longrightarrow \mathbb{R}$$

eine Abbildung, die den Eigenschaften (S1) - (S3) genügt. Dann heißt $\langle\ ,\ \rangle$ ein Skalarprodukt und das Paar $(V, \langle\ ,\ \rangle)$ ein euklidischer Vektorraum.

Bsp. 7: $V := \{ f : [a,b] \to \mathbb{R} \}$ und U der UVR aller stetigen Funktionen in V, also

$$U := \{ f \in V : f \text{ ist stetig} \}.$$

Dann wird durch

$$\langle f, g \rangle = \int_a^b f(x)\, g(x)\, dx$$

ein Skalarprodukt auf U definiert. Das Paar $(U, \langle\ \rangle)$ bildet also einen euklidischen Vektorraum.

Hat man ein Skalarprodukt zur Verfügung, werden Längen- und Abstandsmessungen möglich.

Def.: Es sei $(V, \langle , \rangle)$ ein euklidischer Vektorraum.

Dann heißt für $x \in V$

$$|x| := \langle x, x \rangle^{\frac{1}{2}}$$

die <u>euklidische Norm</u> von x.

Geometrische Interpretation:

$|x|$ = Länge des Vektors x = Abstand zum Nullpunkt,

allgemeiner: $|x-y|$ = Abstand der Vektoren x und y.

Konkretisierung: Für $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ und $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n$

ist die euklidische Norm gegeben durch

$$|x| = \left(\sum_{k=1}^{n} x_k^2 \right)^{\frac{1}{2}}$$

bzw. der Abstand von x und y durch

$$|x-y| = \left(\sum_{k=1}^{n} (x_k - y_k)^2 \right)^{\frac{1}{2}}.$$

Satz (Eigenschaften der euklidischen Norm): Es seien $(V, \langle , \rangle)$ ein euklidischer Vektorraum, $\lambda \in \mathbb{R}$ und $x, y \in V$. Dann gelten:

(N1) $|x| > 0$ für alle $x \in V \setminus \{0\}$ (Definitheit)

(N2) $|\lambda x| = |\lambda| \, |x|$ (Homogenität)

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

Skizze!

Die Eigenschaften (N1)–(N3) kann man – teilweise sogar recht leicht – aus den Eigenschaften (S1)–(S3) herleiten. Für (N3) erfordert dies die Verwendung der Cauchy-Schwarzschen Ungleichung:

Lemma: Es seien $(V, \langle\,,\rangle)$ ein euklidischer Vektorraum und $x, y \in V$. Dann gilt

$$|\langle x, y\rangle| \leq |x|\,|y|.$$

Herleitung für den Fall $V = \mathbb{R}^n$:

(i) Aus $0 \leq (a-b)^2 = a^2 - 2ab + b^2$ folgt für $a, b \in \mathbb{R}$, dass $a \cdot b \leq \frac{1}{2}(a^2 + b^2)$.

(ii) Sind $x, y \in \mathbb{R}^n$ mit $|x| = |y| = 1$, hat man

$$|\langle x, y\rangle| = \Big|\sum_{k=1}^{n} x_k y_k\Big| \leq \sum_{k=1}^{n} |x_k y_k| \leq \frac{1}{2}\sum_{k=1}^{n} x_k^2 + y_k^2$$

$$= \frac{1}{2}(1+1) = 1 = |x|\,|y|$$

(iii) Sind $x, y \in \mathbb{R}^n$ beliebig, so ist n.z.z. für $x = 0$ oder $y = 0$.

Andernfalls: $|\langle x, y\rangle| = \left|\left\langle \frac{x}{|x|}, \frac{y}{|y|}\right\rangle\right| |x|\,|y| \underset{(ii)}{\leq} |x|\,|y|$. $\square$

Damit: Beweis der Dreiecksungleichung:

$$|x+y|^2 = \langle x+y, x+y\rangle = \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle$$

$$= |x|^2 + 2\langle x, y\rangle + |y|^2 \underset{C.S.}{\leq} |x|^2 + 2|x|\,|y| + |y|^2$$

$$= (|x| + |y|)^2. \quad \text{Jetzt: } \sqrt{}. \Rightarrow \text{(N3)} \qquad \square$$

Abschließende Bem. zu Norm und Skalarprodukt:

(1) Nicht jede Norm stammt von einem Skalarprodukt, Bsp.:

$$\|x\|_p = \left(\sum_{k=1}^{n} |x_k|^p\right)^{\frac{1}{p}} \qquad (x \in \mathbb{R}^n,\ p \geq 1,\ p \neq 2)$$

$$\|x\|_\infty = \max\{|x_k| : 1 \leq k \leq n\} \qquad (x \in \mathbb{R}^n)$$

(2) Mit Hilfe des Skalarprodukts kann man auch den von zwei Vektoren x und y eingeschlossenen Winkel messen (bzw. definieren). Dazu setzt man für den Winkel α zwischen x und y

$$\cos(\alpha) := \frac{\langle x, y\rangle}{|x||y|},$$

was wir hier als <u>Definition</u> auffassen.

Speziell: x und y heißen senkrecht bzw. orthogonal zueinander, wenn ihr Skalarprodukt verschwindet. In Zeichen:

$$x \perp y \ :\Longleftrightarrow\ \langle x, y\rangle = 0$$

Insbes. ist dann $0 \perp x \quad \forall x \in V$.