Mathematik für InformatikerInnen 2

Frank-Olaf Schreyer

Universität des Saarlandes, SS 2020

Ziele der heutigen Vorlesung:

- $\mathbb{R}^n$ als Vektorraum
- Euklidische Norm und Skalarprodukt
- Geraden und Hyperebenen
- Abstände und orthogonale Projektionen

Der $\mathbb{R}^n$ als Vektorraum

In der Physik wird ein Vektor mit dem Pfeil,

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3; \quad \vec{x} + \vec{y} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ x_3 + y_3 \end{pmatrix},$$

der vom Nullpunkt zum Punkt $(x_1, x_2, x_3) \in \mathbb{R}^3$ zeigt, identifiziert. Die Summe zweier Vektoren ist die Diagonale in dem Parallelogramm, welches von $\vec{x}$ und $\vec{y}$ aufgespannt wird.

Sind $\vec{x}$ und $\vec{y}$ zwei Kräfte, die auf ein Objekt wirken, dann beschreibt $\vec{x} + \vec{y}$ den Kraftvektor der gemeinsamen Wirkung.

In der Mathematik lassen wir den Pfeil über $\vec{x}$ weg und erlauben ein beliebiges n. Auf $\mathbb{R}^n$ haben wir die folgenden Verknüpfungen

$$\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n, \quad (x, y) \mapsto x + y \qquad \text{(Addition)}$$

und

$$\mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n, \quad (\lambda, x) \mapsto \lambda x \qquad \text{(Skalarmultiplikation).}$$

Dabei ist die Addition komponentenweise definiert

$$x + y = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}.$$

Die Skalarmultiplikation ist durch

$$\lambda x = \lambda \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{pmatrix}$$

gegeben. Diese Verknüpfungen geben $\mathbb{R}^n$ die Struktur eines $\mathbb{R}$-Vektorraums (im Sinne der Definition, die wir in einer der nächsten Vorlesungen besprechen werden).

Euklidische Struktur auf $\mathbb{R}^n$

Definition. Für Vektoren

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

ist das **Skalarprodukt** oder **dot-Produkt** durch

$$\langle x, y \rangle = \langle x | y \rangle = x \cdot y = \sum_{i=1}^{n} x_i y_i$$

definiert. Die **euklidische Norm** von x ist

$$\|x\| = \sqrt{\langle x, x \rangle} = \sqrt{x_1^2 + \ldots + x_n^2}.$$

Wir sagen x und y sind **senkrecht** zu einander, $x \perp y$, falls $\langle x, y \rangle = 0$. Nach dem Satz des Pythagoras können wir $\|x\|$ als die **Länge** des Vektors, oder den Abstand von $x \in \mathbb{R}^n$ zum Nullpunkt $0 \in \mathbb{R}^n$ interpretieren.

Eigenschaften des Skalarprodukts

Proposition. *Das folgende gilt:*

1. $\langle x+y,z\rangle = \langle x,z\rangle + \langle y,z\rangle$ *für alle* $x,y,z \in \mathbb{R}^n$ *(Linearität),*
2. $\langle \lambda x,y\rangle = \lambda\langle x,y\rangle$ *für alle* $x,y \in \mathbb{R}^n, \lambda \in \mathbb{R}$ *(Linearität),*
3. $\langle x,y\rangle = \langle y,x\rangle$ *für alle* $x,y \in \mathbb{R}^n$ *(Symmetrie),*
4. $\langle x,x\rangle \geq 0$ *und* $\langle x,x\rangle = 0$ *genau dann, wenn* $x = 0 \in \mathbb{R}^n$*,*
5. $\|x+y\|^2 = \|x\|^2 + 2\langle x,y\rangle + \|y\|^2$ *(Pythagoras),*
6. $\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2$ *(Parallelogramm-Gl.).*

Beweis.

1.) $\langle x+y,z\rangle = \sum(x_i+y_i)z_i = \sum x_i z_i + \sum y_i z_i = \langle x,z\rangle + \langle y,z\rangle$.

5.) $\|x+y\|^2 = \langle x+y, x+y\rangle \overset{(1\&3)}{=} \langle x,x\rangle + 2\langle x,y\rangle + \langle y,y\rangle$.

6.) $\|x+y\|^2 + \|x-y\|^2 \overset{(5)}{=}$
$\|x\|^2 + 2\langle x,y\rangle + \|y\|^2 + \|x\|^2 + 2\langle x,-y\rangle + \|-y\|^2$
$\overset{(2\&3)}{=} \|x\|^2 + 2\langle x,y\rangle + \|y\|^2 + \|x\|^2 - 2\langle x,y\rangle + \|y\|^2 =$
$2(\|x\|^2 + \|y\|^2)$. $\square$

Cauchy-Schwarz'sche Ungleichung

Theorem. *Es gilt*

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

für alle $x, y \in \mathbb{R}^n$. *Ferner, für* $x \neq 0$ *gilt die Gleichheit* $|\langle x, y\rangle| = \|x\| \cdot \|y\|$ *genau dann, wenn* $y = \lambda x$ *für ein* $\lambda \in \mathbb{R}$ *gilt.*

Beweis. Wir dürfen $x \neq 0 \in \mathbb{R}^n$ annehmen. Setze

$$\mu := \langle x, x\rangle > 0 \text{ und } \varphi := -\langle x, y\rangle.$$

Dann gilt

$$\begin{aligned}
0 &\leq \langle \varphi x + \mu y, \varphi x + \mu y\rangle \\
&= \varphi^2\langle x, x\rangle + 2\varphi\mu\langle x, y\rangle + \mu^2\langle y, y\rangle \\
&= \mu \cdot \langle x, y\rangle^2 - 2\mu \cdot \langle x, y\rangle^2 + \mu \cdot \langle x, x\rangle\langle y, y\rangle \\
&= \mu \cdot \left(-\langle x, y\rangle^2 + \langle x, x\rangle\langle y, y\rangle\right).
\end{aligned}$$

Da $\mu > 0$, folgt

$$0 \leq -(\langle x, y\rangle)^2 + \|x\|^2 \cdot \|y\|^2 \quad \text{äquivalent} \quad |\langle x, y\rangle|^2 \leq \|x\|^2 \cdot \|y\|^2.$$

Die Monotonie der Quadratwurzel $\sqrt{\ }$ gibt die erste Behauptung.
Im Fall von Gleichheit gilt

$$0 = \langle \varphi x + \mu y, \varphi x + \mu y\rangle \;\Rightarrow\; \varphi x + \mu y = 0.$$

Also $y = \lambda x$ mit $\lambda = -\frac{\varphi}{\mu}$. $\square$
Der **Winkel** $\alpha \in [0, \pi]$ zwischen zwei von Null verschiedenen Vektoren x und y lässt sich also durch

$$\cos\alpha = \frac{\langle x, y\rangle}{\|x\| \cdot \|y\|} \quad \in [-1, 1]$$

definieren.

Eigenschaften der Norm

Proposition. *Für* $x, y \in \mathbb{R}^n$ *und* $\lambda \in \mathbb{R}$ *gilt:*

1. $\|x\| \geq 0$ *und* $\|x\| = 0$ *gilt genau dann, wenn* $x = 0 \in \mathbb{R}^n$,
2. $\|\lambda x\| = |\lambda| \cdot \|x\|$,
3. $\|x + y\| \leq \|x\| + \|y\|$ *(Dreiecksungleichung oder* Δ*-Ungl.)*

Beweis. Nur 3) braucht einen Beweis. Nach Definition gilt

$$\begin{aligned} \|x+y\|^2 &= \langle x+y, x+y\rangle \\ &= \|x\|^2 + 2\langle x, y\rangle + \|y\|^2. \end{aligned}$$

Cauchy–Schwarz gibt:

$$\|x+y\|^2 \leq \|x\|^2 + 2\|x\| \cdot \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2.$$

Die Behauptung folgt mit der Monotonie der Quadratwurzel. $\square$

Geraden

Eine **Gerade** $L \subseteq \mathbb{R}^n$ ist eine Teilmenge der Gestalt

$$L = \{p + \lambda v \mid \lambda \in \mathbb{R}\} =: p + \mathbb{R} \cdot v,$$

wobei $p \in L$ ein beliebiger **Aufpunkt** und $v \in \mathbb{R}^n \setminus \{0\}$ ein **Richtungsvektor** ist.

Hyperebenen

Eine **Hyperebene** $H \subseteq \mathbb{R}^n$ ist eine Teilmenge der Gestalt

$$H = \{x \in \mathbb{R}^n \mid a_1 x_1 + \cdots + a_n x_n = b\} = \{x \in \mathbb{R}^n \mid \langle a, x \rangle = b\},$$

wobei $a = (a_1, \ldots, a_n)^t \in \mathbb{R}^n \setminus \{0\}$ und $b \in \mathbb{R}$. Der Vektor a heißt **Normalenvektor**. Eine Hyperebene im $\mathbb{R}^3$ ist einfach eine **Ebene**. Für zwei Punkte $p, q \in H$ gilt für den Differenzvektor $v = p - q$

$$\langle a, v \rangle = \langle a, p - q \rangle = \langle a, p \rangle - \langle a, q \rangle = b - b = 0.$$

Also $a \perp p - q$. Daher der Name Normalenvektor.

Schnittpunkte

Sei $L \subseteq \mathbb{R}^n$ eine Gerade und $H \subseteq \mathbb{R}^n$ eine Hyperebene. Für die Schnittmenge $L \cap H$ gibt es drei Möglichkeiten:

1. $L \cap H = \{q\}$ besteht aus genau einem Punkt $q \in \mathbb{R}^n$,
2. $L \cap H = \emptyset$,
3. $L \subset H$.

Proposition. *Der Fall 1.) liegt genau dann vor, wenn der Richtungsvektor von L nicht senkrecht zum Normalenvektor a von H ist.*

Beweis. Setzen wir die Parametrisierung $L = \{p + \lambda v | \lambda \in \mathbb{R}\}$ der Geraden in Gleichung $\langle a, x\rangle = b$ von H ein, so erhalten wir mit

$$\langle a, p + \lambda v\rangle = b \Leftrightarrow \lambda\langle a, v\rangle = b - \langle a, p\rangle \qquad (*)$$

eine Gleichung für λ.

Ist a nicht senkrecht zu v, d.h. $\langle a, v\rangle \neq 0$, dann ist die einzige Lösung

$$\lambda = \frac{b - \langle a, p\rangle}{\langle a, v\rangle}, \text{ also } L \cap H \ni q = p + \frac{b - \langle a, p\rangle}{\langle a, v\rangle} v.$$

Ist $\langle a, v\rangle = 0$, also $a \perp v$, dann hat (*) nur dann eine Lösung, wenn $b - \langle a, p\rangle = 0 \Leftrightarrow p \in H \Rightarrow L \subseteq H$, da dann $\lambda \in \mathbb{R}$ beliebig gewählt werden kann. Dies entspricht dem Fall 3.)
Ist $\langle a, v\rangle = 0$ und $b \neq \langle a, p\rangle$, dann ist $L \cap H = \emptyset$, der Fall 2.). $\square$

Abstand zwischen Gerade und Punkt

Sei $L = \{p + \lambda v \mid \lambda \in \mathbb{R}\} \subseteq \mathbb{R}^n$ eine Gerade und $q \in \mathbb{R}^n$ ein weiterer Punkt. Dann ist für jeden Punkt $u \in L$ der Abstand $d(u, q) = \|u - q\|$.

Definition. Wir definieren den **Abstand von L zu q** durch

$$d(L, q) = \min_{u \in L} d(u, q).$$

Proposition. *Das Minimum $d(L, q)$ wird in genau einem Punkt u_q angenommen. Der Punkt u_q ist eindeutig bestimmt durch die Eigenschaft, dass $u_q - q$ senkrecht zu dem Richtungsvektor v steht.*

u_q heißt **Fußpunkt** des **Lots** von q auf L.

Beweis. Die Gleichung $\langle v, p + \lambda v - q \rangle = 0$ hat genau eine Lösung, nämlich

$$\lambda = \frac{\langle q - p, v \rangle}{\langle v, v \rangle},$$

da $\|v\|^2 \neq 0$.

Abstand zwischen Gerade und Punkt, 2

Also:

$$u_q = p + \frac{\langle q-p, v\rangle}{\langle v, v\rangle} \cdot v = p + \left\langle q-p, \frac{v}{\|v\|}\right\rangle \cdot \frac{v}{\|v\|}.$$

Jeder andere Punkt $u \in L$ hat einen größeren Abstand, da
$\|u-q\|^2 \overset{\text{Pythagoras}}{=} \|u_q - q\|^2 + \|u - u_q\|^2 \geq \|u_q - q\|^2$, also:

$$d(L, q) = \|u_q - q\|. \qquad \square$$

Abstand zwischen Hyperebene und Punkt

Definition. Sei $H = \{x \mid \langle a, x\rangle = b\} \subseteq \mathbb{R}^n$ eine Hyperebene und $q \in \mathbb{R}^n$ ein Punkt. Dann definieren wir

$$d(H, q) := \min_{u \in H} d(u, q).$$

Proposition. *Das Minimum $d(H, q)$ wird in genau einem Punkt $u_q \in H$ angenommen. u_q ist durch die Bedingung, dass die Differenz $u_q - q$ ein skalares Vielfaches des Normalenvektors a ist, eindeutig bestimmt.*

Beweis. Jeder andere Punkt $u \in H$ hat größeren Abstand zu q, nach Pythagoras. Um u_q auszurechnen, betrachten wir die Gerade $L = \{q + \lambda a \mid \lambda \in \mathbb{R}\} \subseteq \mathbb{R}^n$ und bestimmen $L \cap H$:
$\langle q + \lambda a, a\rangle = b$ liefert $\lambda = \frac{b - \langle q, a\rangle}{\langle a, a\rangle}$, also

$$u_q = q + \frac{b - \langle a, q\rangle}{\langle a, a\rangle} \cdot a. \qquad \square$$

Abstand zwischen Hyperebene und Punkt, 2

Der Abstand ist somit:

$$d(H,q) = \left\| \frac{b - \langle a, q\rangle}{\langle a, a\rangle} \cdot a \right\| = \frac{|b - \langle a, q\rangle|}{\langle a, a\rangle} \cdot \|a\| = \left| \frac{b}{\|a\|} - \left\langle \frac{a}{\|a\|}, q \right\rangle \right| .$$

Die Abbildung

$$\mathbb{R}^n \to H, \quad q \to u_q$$

heißt **orthogonale Projektion** auf die Hyperebene H.

Abstand zwischen Hyperebene und Punkt, 3

Wählen wir den Normalenvektor **normiert**, d.h. von der Länge 1, dann gilt:

$$d(H, q) = |b - \langle a, q \rangle|.$$

In diesem Fall lässt sich $|b|$ als Abstand $d(H, 0)$ von H zum Nullpunkt interpretieren. Auch das Vorzeichen von b hat eine Interpretation:

$$\begin{array}{lll} b > 0 & \Leftrightarrow & \text{0 liegt in dem Halbraum } \{x \mid \langle a, x \rangle < b\} \\ & \Leftrightarrow & \text{Der Normalenvektor, auf } H \text{ angetragen, zeigt} \\ & & \text{in den Halbraum, der 0 nicht enthält.} \end{array}$$

Abstand windschiefer Geraden

Definition. Seien $L_1 = \{p_1 + \lambda v_1 \mid \lambda \in \mathbb{R}\}$ und $L_2 = \{p_2 + \lambda v_2 \mid \lambda \in \mathbb{R}\}$ zwei Geraden im $\mathbb{R}^n$. L_1 und L_2 heißen **parallel**, wenn $v_1 = \lambda v_2$ für ein $\lambda \in \mathbb{R}$, das heißt, wenn die Richtungsvektoren bis auf den Skalarfaktor übereinstimmen. L_1 und L_2 heißen **windschief**, wenn gilt:

1. L_1 und L_2 sind nicht parallel,
2. $L_1 \cap L_2 = \emptyset$.

$$d(L_1, L_2) := \min_{x \in L_1, y \in L_2} d(x, y)$$

nennen wir den **Abstand von L_1 zu L_2**

Abstand windschiefer Geraden, 2

Proposition. *Es seien L_1 und L_2 zwei windschiefe Geraden mit Richtungsvektoren v_1 bzw. v_2. Dann wird das Minimum $d(L_1, L_2)$ in genau einem Paar von Punkten $(\tilde{x}, \tilde{y}) \in L_1 \times L_2$ angenommen. $(\tilde{x}, \tilde{y})$ ist durch die Bedingung, dass $\tilde{x} - \tilde{y}$ senkrecht zu v_1 und v_2 steht, eindeutig bestimmt.*

Beweis. $(\tilde{x}, \tilde{y})$ erfülle die Bedingung. Für jedes andere Paar $(x, y) \in L_1 \times L_2$ gilt:

$$(x, y) = (\tilde{x} + \lambda_1 v_1,\ \tilde{y} + \lambda_2 v_2) \text{ für gewisse } \lambda_1, \lambda_2 \in \mathbb{R}.$$

Mit dieser Notation gilt

$$\begin{aligned} \|x - y\|^2 &= \|\tilde{x} + \lambda_1 v_1 - \tilde{y} - \lambda_2 v_2\|^2 \\ &= \|\tilde{x} - \tilde{y} + \lambda_1 v_1 - \lambda_2 v_2\|^2 \\ &= \|\tilde{x} - \tilde{y}\|^2 + \|\lambda_1 v_1 - \lambda_2 v_2\|^2 \end{aligned}$$

nach Pythagoras, da $\tilde{x} - \tilde{y}$ nach Voraussetzung zu jeder Linearkombination $\lambda_1 v_1 - \lambda_2 v_2$ senkrecht steht.

Abstand windschiefer Geraden, 3

In der Tat,

$$\langle \tilde{x} - \tilde{y},\ \lambda_1 v_1 - \lambda_2 v_2 \rangle = \lambda_1 \langle \tilde{x} - \tilde{y},\ v_1 \rangle - \lambda_2 \langle \tilde{x} - \tilde{y},\ v_2 \rangle = 0.$$

Insgesamt folgt: $\|x - y\|^2 \geq \|\tilde{x} - \tilde{y}\|^2$ (also auch $d(x, y) \geq d(\tilde{x}, \tilde{y})$), und Gleichheit gilt genau dann, wenn

$$\lambda_1 v_1 - \lambda_2 v_2 = 0 \Leftrightarrow \lambda_1 = \lambda_2 = 0,$$

da v_1 und v_2 keine skalaren Vielfachen voneinander sind. Wir haben somit:

$$d(L_1, L_2) = \|\tilde{x} - \tilde{y}\|.$$

Abstand windschiefer Geraden, 4

Es bleibt zu zeigen, dass die angegebene Bedingung $\tilde{x}$ und $\tilde{y}$ eindeutig bestimmt. Wir schreiben: $\tilde{x} = p_1 + \lambda_1 v_1$, $\tilde{y} = p_2 + \lambda_2 v_2$ für gewisse $\lambda_1, \lambda_2 \in \mathbb{R}$ und $p_1, p_2, v_1, v_2 \in \mathbb{R}^n$. Wegen der Bedingung gilt nun:

$$\langle \tilde{x} - \tilde{y}, v_1 \rangle = 0, \quad \langle \tilde{x} - \tilde{y}, v_2 \rangle = 0,$$

also:

$$\langle p_1 - p_2 + \lambda_1 v_1 - \lambda_2 v_2, v_1 \rangle = 0, \quad \langle p_1 - p_2 + \lambda_1 v_1 - \lambda_2 v_2, v_2 \rangle = 0.$$

Dies liefert ein **lineares Gleichungssystem** für λ_1 und λ_2:

$$\lambda_1 \cdot \|v_1\|^2 - \lambda_2 \langle v_2, v_1 \rangle = \langle p_2 - p_1, v_1 \rangle,$$
$$\lambda_1 \cdot \langle v_1, v_2 \rangle - \lambda_2 \|v_2\|^2 = \langle p_2 - p_1, v_2 \rangle.$$

Abstand windschiefer Geraden, 5

In **Matrixschreibweise** erhalten wir

$$\begin{pmatrix} \|v_1\|^2 & -\langle v_2, v_1\rangle \\ \langle v_1, v_2\rangle & -\|v_2\|^2 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} \langle p_2 - p_1, v_1\rangle \\ \langle p_2 - p_1, v_2\rangle \end{pmatrix}.$$

Wir werden sehen, dass diese Gleichung genau dann eine eindeutig bestimmte Lösung $(\lambda_1, \lambda_2)^t \in \mathbb{R}^2$ hat, wenn die **Determinante** der 2×2 Matrix

$$\det \begin{pmatrix} \|v_1\|^2 & -\langle v_2, v_1\rangle \\ \langle v_1, v_2\rangle & -\|v_2\|^2 \end{pmatrix} = -\|v_1\|^2 \cdot \|v_2\|^2 + \langle v_1, v_2\rangle^2$$

ungleich Null ist. Nach der Cauchy–Schwarz'schen Ungleichung ist aber $|\langle v_1, v_2\rangle| \leq \|v_1\| \cdot \|v_2\|$, und es gilt $<$, da v_1 und v_2 nicht skalare Vielfache voneinander sind. $\square$