Mathematik I

(für Informatiker, ET und IK)

Oliver Ernst

Professur Numerische Mathematik

Wintersemester 2015/16

Mathematik ! TU Chemnitz

Inhalt

1. Grundlagen
2. Lineare Algebra

Lineare Algebra

Mit der linearen Algebra lernen wir nun ein großes Teilgebiet der Mathematik kennen.

Dieses befasst sich unter anderem mit

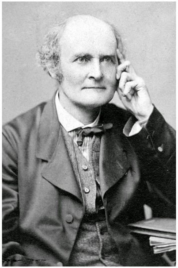
- Vektorräumen,
- linearen Abbildungen zwischen Vektorräumen,
- linearen Gleichungssystemen,
- Determinanten und Matrizen,
- Eigenwerten und -vektoren.

Insbesondere werden wir hier auch die Grundlagen für die mehrdimensionale Differential- und Integralrechnung legen.

Die Entwicklung der modernen linearen Algebra erfolgte vor allem in der Mitte des 19. Jahrhunderts, wenngleich erste Grundlagen bereits wesentlich früher bekannt waren. Wichtige Beiträge lieferten

- Gabriel Cramer (1704-1752, Schweizer Mathematiker),
- Sir William Rowan Hamilton (1805–1865, irischer Mathematiker),
- Herrmann Graßmann (1809–1877, deutscher Mathematiker),
- Arthur Cayley (1821–1895, englischer Mathematiker).



Inhalt

1. Grundlagen

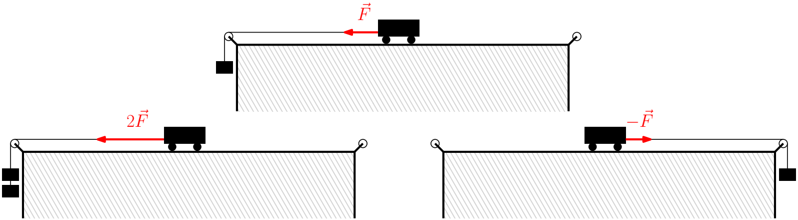
2. Lineare Algebra

2.1 Vektorräume
2.2 Matrizen und lineare Abbildungen
2.3 Lineare Gleichungssysteme
2.4 Determinanten
2.5 Invertierbare Matrizen
2.6 Orthogonalität, Skalarprodukt und Norm
2.7 Kreuz- und Spatprodukt
2.8 Elemente der analytischen Geometrie
2.9 Orthogonale Abbildungen
2.10 Eigenwerte und Eigenvektoren
2.11 Singulärwertzerlegung

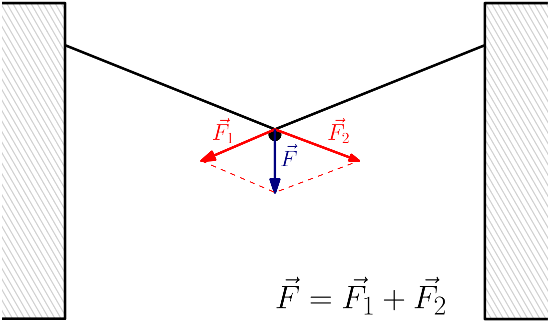
Physikalische Kräfte können nicht durch eine Zahl allein beschrieben werden: sie besitzen neben ihrem „Betrag“ auch eine Richtung.

Man beschreibt sie durch Vektoren. Wirken die Kräfte in einer Ebene, verwendet man „zweidimensionale“ Vektoren $\left[\begin{smallmatrix}\alpha\\ \beta\end{smallmatrix}\right]$.

Man kann Kräfte (Vektoren) mit einer Zahl λ multiplizieren: dabei wird die Richtung beibehalten oder (bei negativem λ) umgekehrt und der Betrag mit $|\lambda|$ multipliziert.



Desweiteren kann man Kräfte (Vektoren) addieren. Dies visualisiert man am sogenannten Kräfteparallelogramm.



Man beachte, dass sich die Beträge der Kräfte **nicht** einfach addieren. Es gelten aber auch für die Vektoraddition viele gewohnte Gesetzmäßigkeiten, die in die Definition des Vektorraums einfließen.

Um die algebraische Struktur des Vektorraums und damit den Vektorbegriff mathematisch exakt zu fassen, benötigen wir zunächst einen Körper $\mathbb{K}$.

Dies ist eine Menge, auf der zwei Operationen ($+$ und $\cdot$) definiert sind, die den Gesetzen genügen, die in Abschnitt 2.3, Folie 41, aufgelistet wurden.

In den meisten Fällen werden wir als Körper die reellen Zahlen wählen ($\mathbb{K} = \mathbb{R}$), mitunter auch die komplexen Zahlen ($\mathbb{K} = \mathbb{C}$).

Grundsätzlich könnte man aber jeden Körper wählen (z. B. $\mathbb{Q}$).

Definition 3.1 (Vektorraum)

Ein $\mathbb{K}$-Vektorraum $\mathscr{V} := (V; +, \cdot)$ besteht aus einer Menge $V \neq \emptyset$, deren Elemente Vektoren genannt werden, sowie zwei Operationen: einer (Vektor)addition $+\colon V \times V \to V$ und einer Skalarmultiplikation $\cdot : \mathbb{K} \times V \to V$.

Dabei müssen folgende Regeln gelten:

(1) $\boldsymbol{a} + (\boldsymbol{b} + \boldsymbol{c}) = (\boldsymbol{a} + \boldsymbol{b}) + \boldsymbol{c}$ für alle $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c} \in V$,

(2) es gibt einen Vektor $\mathbf{0}$ mit $\boldsymbol{a} + \mathbf{0} = \boldsymbol{a}$ für alle $\boldsymbol{a} \in V$,

(3) zu jedem $\boldsymbol{a} \in V$ gibt es ein $-\boldsymbol{a} \in V$ mit $\boldsymbol{a} + (-\boldsymbol{a}) = \mathbf{0}$,

(4) $\boldsymbol{a} + \boldsymbol{b} = \boldsymbol{b} + \boldsymbol{a}$ für alle $\boldsymbol{a}, \boldsymbol{b} \in V$,

(5) $(\lambda\mu) \cdot \boldsymbol{a} = \lambda \cdot (\mu \cdot \boldsymbol{a})$ für alle $\lambda, \mu \in \mathbb{K}$ und alle $\boldsymbol{a} \in V$,

(6) $(\lambda + \mu) \cdot \boldsymbol{a} = \lambda \cdot \boldsymbol{a} + \mu \cdot \boldsymbol{a}$ für alle $\lambda, \mu \in \mathbb{K}$ und alle $\boldsymbol{a} \in V$,

(7) $\lambda \cdot (\boldsymbol{a} + \boldsymbol{b}) = \lambda \cdot \boldsymbol{a} + \lambda \cdot \boldsymbol{b}$ für alle $\lambda \in \mathbb{K}$ und alle $\boldsymbol{a}, \boldsymbol{b} \in V$,

(8) $1 \cdot \boldsymbol{a} = \boldsymbol{a}$ für alle $\boldsymbol{a} \in V$.

- $+\colon V \times V \to V$ bedeutet, dass die Addition je zwei Vektoren $\boldsymbol{a}$ und $\boldsymbol{b}$ einen Vektor $\boldsymbol{a} + \boldsymbol{b}$ zuordnet
- $\cdot : \mathbb{K} \times V \to V$ bedeutet, dass die Skalarmultiplikation je einer Zahl λ und einem Vektor $\boldsymbol{a}$ einen Vektor $\lambda \cdot \boldsymbol{a}$ zuordnet.

Dies entspricht genau dem Charakter der am Beispiel physikalischer Kräfte diskutierten Operationen.

Auch wenn sich die Menge V und die algebraische Struktur $\mathscr{V} := (V; +, \cdot)$ prinzipiell unterscheiden, verwendet man statt des umständlichen $(V; +, \cdot)$ fast immer nur V als Bezeichnung des Vektorraums.

Das (momentan) wichtigste Beispiel für einen Vektorraum ist

$$\mathbb{K}^n := \left\{ \vec{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} : a_j \in \mathbb{K},\ j = 1, 2, \ldots, n \right\}$$

mit

$$\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} := \begin{bmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \vdots \\ a_n + b_n \end{bmatrix} \quad \text{und} \quad \lambda \cdot \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} := \begin{bmatrix} \lambda a_1 \\ \lambda a_2 \\ \vdots \\ \lambda a_n \end{bmatrix}$$

($\lambda \in \mathbb{K}$; $a_j, b_j \in \mathbb{K}$). Addition und Skalarmultiplikation sind also komponentenweise definiert.

Zwei Vektoren, $\vec{a} = [a_j]_{j=1}^n$ und $\vec{b} = [b_j]_{j=1}^n \in \mathbb{K}^n$, sind genau dann gleich, wenn $a_j = b_j$ für alle $j = 1, 2, \dots, n$ gilt.

Nullvektor und inverse Vektoren sind im $\mathbb{K}^n$ gegeben durch

$$\vec{0} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \text{bzw.} \quad -\vec{a} = \begin{bmatrix} -a_1 \\ -a_2 \\ \vdots \\ -a_n \end{bmatrix}, \quad \text{falls } \vec{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}.$$

Zu $\vec{a}, \vec{b} \in \mathbb{R}^2$ mit $\vec{a} = \left[\begin{smallmatrix} 3 \\ 4 \end{smallmatrix}\right]$ und $\vec{b} = \left[\begin{smallmatrix} -1 \\ 2 \end{smallmatrix}\right]$ berechne man $\vec{a} + 3\,\vec{b}$ und $\vec{a} - \vec{b}$.

Zu $\vec{a}, \vec{b} \in \mathbb{C}^2$ mit $\vec{a} = \left[\begin{smallmatrix} 3 \\ 4i \end{smallmatrix}\right]$ und $\vec{b} = \left[\begin{smallmatrix} -1+i \\ 2 \end{smallmatrix}\right]$ berechne man $\vec{a} + (1+i)\,\vec{b}$ und $\vec{a} - i\vec{b}$.

Verifizieren Sie für $V = \mathbb{R}^n$ einige der in Definition 3.1 genannten Beziehungen.

- Die Menge der Polynome bildet einen Vektorraum.
- Die Menge der Polynome vom maximalen Grad n bildet einen Vektorraum.
- Die Menge der stetigen reellen Funktionen bildet einen Vektorraum (Bezeichnung $C(\mathbb{R})$).
- Die Menge der k-mal stetig differenzierbaren reellen Funktionen bildet einen Vektorraum (Bezeichnung $C^k(\mathbb{R})$).

In jedem dieser Beispiele sind Addition und Skalarmultiplikation punktweise zu verstehen.

- Koordinatenvektoren sind **Spaltenvektoren**. Weil das oft zuviel Platz beansprucht, schreiben wir auch

$$\vec{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} =: [a_1,\, a_2,\, \ldots,\, a_n]^T\,.$$

  a_j heißt j-te Komponente von $\vec{a}$.
- Solange wir allgemeine Vektorräume betrachten, verwenden wir für Vektoren fette kleine lateinische Buchstaben ($\boldsymbol{a}, \boldsymbol{b}, \ldots$) und kleine griechische Buchstaben für Skalare.
- Für den Spezialfall $\mathbb{K}^n$, insbesondere für $\mathbb{R}^n$ verwenden wir für Vektoren die Schreibweise mit dem Pfeil ($\vec{a}, \vec{b}, \ldots$) und deren Komponenten die Schreibweise $a_j, b_j, \ldots \ (j = 1, \ldots, n)$.
- Der Punkt $\cdot$, der für die Skalarmultiplikation steht, wird meistens unterdrückt.

In Vektorräumen gelten weiterhin folgende Rechenregeln:

Satz 3.2

Sei V ein $\mathbb{K}$-Vektorraum. Dann gelten:

- $0\boldsymbol{v} = \lambda\mathbf{0} = \mathbf{0}$ *für alle* $\lambda \in \mathbb{K}$ *und alle* $\boldsymbol{v} \in V$*.*
- $(-\lambda)\boldsymbol{v} = \lambda(-\boldsymbol{v}) = -(\lambda\boldsymbol{v})$ *für alle* $\lambda \in \mathbb{K}$ *und alle* $\boldsymbol{v} \in V$*.*
- $(-\lambda)(-\boldsymbol{v}) = \lambda\boldsymbol{v}$ *für alle* $\lambda \in \mathbb{K}$ *und alle* $\boldsymbol{v} \in V$*.*

Dem Anwender dürften diese Regeln intuitiv klar sein; aber genaugenommen müssen sie aus Definition 3.1 hergeleitet werden.

Machen Sie sich für mindestens einen Punkt klar, wie das geschehen könnte.

Definition 3.3 (Unterraum)

Ist U eine nichtleere Teilmenge eines $\mathbb{K}$-Vektorraums V mit

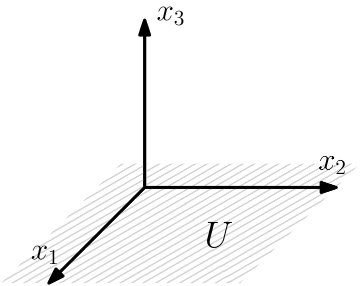
- $\boldsymbol{u} + \boldsymbol{v} \in U$ für alle $\boldsymbol{u}, \boldsymbol{v} \in U$ und
- $\lambda \boldsymbol{u} \in U$ für alle $\boldsymbol{u} \in U$ und alle $\lambda \in \mathbb{K}$,

dann nennt man U einen Unterraum von V.

Für die in Definition 3.3 genannten Punkte verwendet man auch zusammenfassend die Sprechweise: U ist abgeschlossen unter Addition und Skalarmultiplikation.

Natürlich ist U damit selbst wieder ein Vektorraum; daher verwendet man auch die Bezeichnung Unter(vektor-)raum.

- Jeder Vektorraum V enthält als triviale Unterräume den gesamten Raum, also V, und den Nullraum $\{\mathbf{0}\}$, der nur aus dem Nullvektor besteht.
- Die Menge $\{\vec{x} \in \mathbb{R}^3 : x_3 = 0\}$ bildet einen Unterraum U des $\mathbb{R}^3$.



Überzeugen Sie sich anhand der Definition, dass der zweite Punkt wahr ist.

Definition 3.4 (und Satz)

Sei V ein Vektorraum. Ein Vektor $\boldsymbol{y}$ der Form

$$\boldsymbol{y} = \sum_{j=1}^{k} \lambda_j \boldsymbol{x}_j \qquad (\lambda_j \in \mathbb{K},\ \boldsymbol{x}_j \in V,\ k \in \mathbb{N})$$

heißt Linearkombination der Vektoren $\boldsymbol{x}_1, \dots, \boldsymbol{x}_k$.

Ist $\emptyset \neq X \subseteq V$, so ist

$$\operatorname{span}(X) := \left\{ \sum_{j=1}^{k} \lambda_j \boldsymbol{x}_j : \lambda_j \in \mathbb{K}, \boldsymbol{x}_j \in X, k \in \mathbb{N} \right\}$$

ein Unterraum von V, genauer: der kleinste Unterraum von V, der X enthält. Man nennt $\operatorname{span}(X)$ die lineare Hülle von X oder den von X erzeugten Unterraum von V.

Die lineare Hülle $\operatorname{span}(X)$ ist also gerade die Menge aller Linearkombinationen von Vektoren aus X.

- $\operatorname{span}\{\boldsymbol{v}\} = \{\lambda \boldsymbol{v} \,:\, \lambda \in \mathbb{K}\}$,
- $\operatorname{span}\{\boldsymbol{v}, \boldsymbol{w}\} = \{\lambda \boldsymbol{v} + \mu \boldsymbol{w} \,:\, \lambda, \mu \in \mathbb{K}\}$
- Für $V = \mathbb{R}^3$ gilt

$$\operatorname{span}\left(\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}\right) = \{x \in \mathbb{R}^3 : x_3 = 0\},$$

aber auch

$$\operatorname{span}\left(\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \right\}\right) = \{x \in \mathbb{R}^3 : x_3 = 0\}.$$

Definition 3.5 (Lineare Unabhängigkeit)

Seien V ein $\mathbb{K}$-Vektorraum und $X \subseteq V$ eine Teilmenge von V. Die Vektoren aus X heißen linear unabhängig, wenn der Nullvektor nur trivial als Linearkombination von Vektoren aus X dargestellt werden kann; d. h. wenn aus

$$\sum_{j=1}^{k} \lambda_j \boldsymbol{x}_j = \mathbf{0} \quad (\text{mit } \boldsymbol{x}_j \in X \text{ und } \lambda_j \in \mathbb{K})$$

stets

$$\lambda_1 = \lambda_2 = \cdots = \lambda_k = 0$$

folgt. Vektoren, die nicht linear unabhängig sind, nennt man linear abhängig.

Seien V ein $\mathbb{K}$-Vektorraum und $X \subseteq V$ eine Teilmenge von V. Dann sind die Vektoren aus X genau dann linear unabhängig, wenn sich keiner der Vektoren aus X als Linearkombination der anderen schreiben lässt.

Dies ist wiederum äquivalent zur Forderung

$$\operatorname{span}(X \setminus \{\boldsymbol{x}\}) \subsetneq \operatorname{span}(X) \quad \text{für alle} \quad x \in X.$$

Machen Sie sich klar, dass es sich hierbei tatsächlich um eine äquivalente Charakterisierung handelt.

Wir werden später effiziente Möglichkeiten kennenlernen, Informationen über lineare Unabhängigkeit zu erhalten. Wir versuchen uns trotzdem bereits hier an folgender Aufgabe:

Für welche der folgenden Mengen $X_j \subset \mathbb{R}^2$ sind die Vektoren aus X_j linear unabhängig? Geben Sie jeweils eine schlüssige Begründung.

$$X_1 = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}, \qquad X_2 = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$$

$$X_3 = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right\}, \qquad X_4 = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 4 \\ 0 \end{bmatrix} \right\}$$

Definition 3.6

Ist V ein Vektorraum, so heißt eine Teilmenge $X \subseteq V$ ein Erzeugendensystem von V, wenn man jeden Vektor $\boldsymbol{v} \in V$ als Linearkombination von Vektoren aus X darstellen kann.

Ein Erzeugendensystem X von V, das aus linear unabhängigen Vektoren besteht, heißt Basis von V.

Erinnerung: „Als Linearkombination darstellbar" bedeutet, dass zu jedem $\boldsymbol{v} \in V$ Skalare $\lambda_1, \lambda_2, \dots, \lambda_k \in \mathbb{K}$ und Vektoren $\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_k \in X$ existieren mit

$$\boldsymbol{v} = \lambda_1 \boldsymbol{x}_1 + \lambda_2 \boldsymbol{x}_2 + \cdots + \lambda_k \boldsymbol{x}_k = \sum_{j=1}^{k} \lambda_j \boldsymbol{x}_j,$$

oder kurz, dass $V = \operatorname{span}(X)$ gilt.

Sei X eine Basis des $\mathbb{K}$-Vektorraums V. Dann gilt:

- Entfernt man aus X einen beliebigen Vektor $\boldsymbol{x}$, dann ist $X \setminus \{\boldsymbol{x}\}$ kein Erzeugendensystem von V.

  Mit anderen Worten: Eine Basis von V ist ein **minimales Erzeugendensystem** von V.

- Fügt man zu X einen Vektor $\boldsymbol{y}$ ($\boldsymbol{y} \notin X$) hinzu, dann sind die Vektoren aus $X \cup \{\boldsymbol{y}\}$ nicht mehr linear unabhängig.

  Mit anderen Worten: Eine Basis von V ist eine **maximale Menge linear unabhängiger Vektoren** aus V.

Ein Vektorraum V hat i.Allg. viele verschiedenen Basen, die aber alle dieselbe Anzahl von Elementen besitzen. Die Zahl der Vektoren, aus denen eine Basis von V besteht, heißt Dimension von V.

Schreibweise: $\dim(V)$.

- Für $V = \mathbb{R}^2$ gilt $\dim(V) = 2$, denn $\{\left[\begin{smallmatrix}1\\0\end{smallmatrix}\right], \left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]\}$ ist eine Basis des $\mathbb{R}^2$.
- Für den Vektorraum V_p der Polynome gilt $\dim(V_p) = \infty$. Eine Basis ist zum Beispiel gegeben durch $\{1, x, x^2, x^3, \ldots\}$. (Warum?)

Wir befassen uns hier (fast) nur mit **endlich-dimensionalen** Vektorräumen V (d. h. $\dim(V) = n < \infty$).

Welche der folgenden Mengen $X_j \subset \mathbb{R}^2$ sind Erzeugendensysteme bzw. Basen des $\mathbb{R}^2$?

$$X_1 = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}, \qquad X_2 = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$$

$$X_3 = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right\}, \qquad X_4 = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\}$$

Argumentieren Sie auch mit Hilfe der auf den letzten beiden Folien behandelten Eigenschaften von Basen.

Man charakterisiere die 1- und 2-dimensionalen Untervektorräume des $\mathbb{R}^3$ geometrisch.

Satz 3.7

Ist X eine Basis des $\mathbb{K}$-Vektorraums V, so lässt sich jeder Vektor $\boldsymbol{v} \in V$ auf ***eindeutige*** *Weise als Linearkombination der Basisvektoren darstellen.*

- Für endlichdimensionale Vektorräume und $X = \{\boldsymbol{x}_1, \dots, \boldsymbol{x}_n\}$ existieren somit zu jedem $v \in V$ eindeutig bestimmte Skalare $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ mit $\boldsymbol{v} = \sum_{j=1}^{n} \lambda_j \boldsymbol{x}_j$.
- Im allgemeinen Fall mit $X = \{\boldsymbol{x}_j : j \in J\}$ ist diese Beziehung zu ersetzen durch $\boldsymbol{v} = \sum_{j \in J} \lambda_j \boldsymbol{x}_j$, wobei jedoch nur endlich viele $\lambda_j \in \mathbb{K}$ ungleich 0 sein dürfen.

Man mache sich letzteres am Beispiel des Vektorraums der Polynome und der Exponentialfunktion (die natürlich kein Polynom ist) klar.

Der Vektor

$$\vec{e}_j := [0, \ldots, 0, \underbrace{1}_{j-\text{te-Komponente}}, 0, \ldots, 0]^T \in \mathbb{K}^n \quad (j = 1, 2, \ldots, n)$$

heißt j-ter (n-dimensionaler) Einheitsvektor des $\mathbb{K}^n$. Es gilt:

Satz 3.8

Mehr als n Vektoren aus $\mathbb{K}^n$ sind stets linear abhängig.

Im $\mathbb{K}^n$ sind k paarweise verschiedene Einheitsvektoren ($k \leq n$) immer linear unabhängig.

Insbesondere ist $\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\}$ eine Basis des $\mathbb{K}^n$, die sogenannte Standardbasis, und $\dim(\mathbb{K}^n) = n$.

Zwei Vektoren im $\mathbb{R}^2$ bilden genau dann eine Basis des $\mathbb{R}^2$, wenn sie nicht auf einer Geraden liegen. Drei Vektoren im $\mathbb{R}^3$ bilden genau dann eine Basis des $\mathbb{R}^3$, wenn sie nicht in einer Ebene liegen.

Inhalt

1. Grundlagen

2. Lineare Algebra

2.1 Vektorräume

2.2 Matrizen und lineare Abbildungen

2.3 Lineare Gleichungssysteme

2.4 Determinanten

2.5 Invertierbare Matrizen

2.6 Orthogonalität, Skalarprodukt und Norm

2.7 Kreuz- und Spatprodukt

2.8 Elemente der analytischen Geometrie

2.9 Orthogonale Abbildungen

2.10 Eigenwerte und Eigenvektoren

2.11 Singulärwertzerlegung

Eine $m \times n-$Matrix A ist ein rechteckiges Zahlenschema, in dem $m \cdot n$ reelle oder komplexe Einträge in m Zeilen und n Spalten angeordnet sind:

$$A = [a_{i,j}]_{1 \le i \le m, 1 \le j \le n} = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix}.$$

Die Zahl $a_{i,j}$, die in der i-ten Zeile und j-ten Spalte von A steht, heißt der (i,j)-te Eintrag von A.

Die Menge der reellen (komplexen) $m \times n-$Matrizen wird mit $\mathbb{R}^{m \times n}$ ($\mathbb{C}^{m \times n}$) bezeichnet. Um beide Fälle zu erfassen schreiben wir $\mathbb{K}^{m \times n}$.

Matrizen und lineare Abbildungen

Gleichheit von Matrizen, Vektoren als Spezialfall

Zwei Matrizen $A, B \in \mathbb{K}^{m\times n}$ mit $A = [a_{i,j}]_{1\le i\le m, 1\le j\le n}$ und $B = [b_{i,j}]_{1\le i\le m, 1\le j\le n}$ sind genau dann gleich, wenn

$$a_{i,j} = b_{i,j} \quad \text{für alle } i = 1, 2, \dots, m \text{ und alle } j = 1, 2, \dots, n.$$

Vektoren aus dem $\mathbb{K}^n$ (Spaltenvektoren) kann man als Matrizen aus $\mathbb{K}^{n\times 1}$ auffassen.

Zeilenvektoren lassen sich analog als Elemente von $\mathbb{K}^{1\times n}$ auffassen.

Dies wird insbesondere dann deutlich, wenn wir Addition und Skalarmultiplikation für Matrizen eingeführt haben.

Wie bei Vektoren werden Addition und Skalarmultiplikation komponentenweise erklärt.

Für

$$A = [a_{i,j}]_{1 \le i \le m, 1 \le j \le n}, B = [b_{i,j}]_{1 \le i \le m, 1 \le j \le n} \in \mathbb{K}^{m \times n} \text{ und } \lambda \in \mathbb{K}$$

definiert man

$$A + B := [a_{i,j} + b_{i,j}]_{1 \le i \le m, 1 \le j \le n} \in \mathbb{K}^{m \times n}$$

und

$$\lambda \cdot A := [\lambda a_{i,j}]_{1 \le i \le m, 1 \le j \le n} \in \mathbb{K}^{m \times n}.$$

Man beachte, dass die Addition nur für Matrizen gleicher Größe (gleiche Anzahl Zeilen und Spalten) definiert ist.

Satz 3.9

- $A + (B + C) = (A + B) + C$ *für alle* $A, B, C \in \mathbb{K}^{m \times n}$,
- *es gibt eine Matrix* $O \in \mathbb{K}^{m \times n}$, *die sogenannte* *Nullmatrix*, *mit* $A + O = A$ *für alle* $A \in \mathbb{K}^{m \times n}$,
- *zu jeder Matrix* $A \in \mathbb{K}^{m \times n}$ *gibt es eine Matrix* $-A \in \mathbb{K}^{m \times n}$ *mit* $A + (-A) = O$,
- $A + B = B + A$ *für alle* $A, B \in \mathbb{K}^{m \times n}$,
- $(\lambda\mu)A = \lambda(\mu A)$ *für alle* $\lambda, \mu \in \mathbb{K}$ *und alle* $A \in \mathbb{K}^{m \times n}$,
- $(\lambda + \mu)A = \lambda A + \mu A$ *für alle* $\lambda, \mu \in \mathbb{K}$ *und alle* $A \in \mathbb{K}^{m \times n}$,
- $\lambda(A + B) = \lambda A + \lambda B$ *für alle* $\lambda \in \mathbb{K}$ *und alle* $A, B \in \mathbb{K}^{m \times n}$,
- $1 \cdot A = A$ *für alle* $A \in \mathbb{K}^{m \times n}$.

Zusammenfassend: $(\mathbb{K}^{m \times n}; +, \cdot)$ ist ein $\mathbb{K}$-Vektorraum.

Naheliegenderweise enthält die Nullmatrix $O \in \mathbb{K}^{m\times n}$ als Einträge nur Nullen, d. h. $O = [0]_{1\le i\le m, 1\le j\le n}$.

Mit $A = [a_{i,j}] \in \mathbb{K}^{m\times n}$ gilt weiterhin $-A = [-a_{i,j}]$.

Für $A, B \in \mathbb{R}^{3\times 2}$ mit

$$A = \begin{bmatrix} 1 & 3 \\ 0 & -1 \\ 0 & 5 \end{bmatrix} \quad \text{und} \quad B = \begin{bmatrix} -2 & 4 \\ 5 & 1 \\ 0 & -2 \end{bmatrix}$$

berechne man $A + B$ und $A - 3B$.

Ist $A = [a_{i,j}] \in \mathbb{K}^{m \times n}$, dann heißt

$$A^T := [a_{j,i}] \in \mathbb{K}^{n \times m}$$

die Transponierte von A. In A^T sind also die Rollen der Zeilen und Spalten von A vertauscht.

Transponiert man einen Vektor $\boldsymbol{a} = [a_j] \in \mathbb{K}^n$, so ergibt sich ein Zeilenvektor $\boldsymbol{a}^T = [a_1, a_2, \dots, a_n]$.

Für eine komplexe Matrix $A = [a_{i,j}] \in \mathbb{C}^{m \times n}$ definiert man weiterhin die Konjugiert-Transponierte von A:

$$A^H := \overline{A}^T = [\overline{a}_{j,i}] \in \mathbb{C}^{n \times m}.$$

Hat A nur reelle Einträge, so sind A^T und A^H identisch.

Satz 3.10

Für $A, B \in \mathbb{K}^{m \times n}$ *und* $\lambda \in \mathbb{K}$ *gelten:*

- $\left(A^T\right)^T = A$ *und* $\left(A^H\right)^H = A$,
- $(\lambda A)^T = \lambda A^T$ *und* $(\lambda A)^H = \overline{\lambda} A^H$,
- $(A + B)^T = A^T + B^T$ *und* $(A + B)^H = A^H + B^H$.

Man bestimme die Transponierten und Konjugiert-Transponierten von

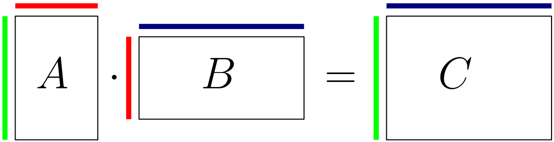
$$A = \begin{bmatrix} 1 & 3 \\ 0 & -1 \\ 0 & 5 \end{bmatrix} \quad \text{und} \quad B = \begin{bmatrix} 1+i & 3i \\ 0 & -1 \end{bmatrix}$$

Definition 3.11

Für $A = [a_{i,j}] \in \mathbb{K}^{m \times n}$ und $B = [b_{i,j}] \in \mathbb{K}^{n \times p}$ ist das Produkt $C = A \cdot B = [c_{i,j}] \in \mathbb{K}^{m \times p}$ definiert durch

$$c_{i,j} = \sum_{k=1}^{n} a_{i,k} b_{k,j} \qquad (i = 1, 2, \ldots, m \;\text{ und }\; j = 1, 2, \ldots, p).$$

Anmerkung: Das Produkt $C = AB$ ist nur dann erklärt, wenn A so viele Spalten wie B Zeilen hat. In diesem Fall übernimmt das Ergebnis C die Zeilenanzahl von A und die Spaltenanzahl von B, symbolisch:



Matrizen und lineare Abbildungen

Falk-Schema für Matrizenmultiplikation

$$
\begin{array}{ccccc|ccccc}
 & & & & & b_{1,1} & \cdots & b_{1,j} & \cdots & b_{1,p} \\
 & & & & & b_{2,1} & \cdots & b_{2,j} & \cdots & b_{2,p} \\
 & & & & & \vdots & & \vdots & & \vdots \\
 & & & & & \vdots & & \vdots & & \vdots \\
 & & & & & b_{n,1} & \cdots & b_{n,j} & \cdots & b_{n,p} \\
\hline
a_{1,1} & a_{1,2} & \cdots & \cdots & a_{1,n} & c_{1,1} & \cdots & c_{1,j} & \cdots & c_{1,p} \\
\vdots & & & & \vdots & \vdots & & \vdots & & \vdots \\
a_{i,1} & a_{i,2} & \cdots & \cdots & a_{i,n} & c_{i,1} & \cdots & c_{i,j} & \cdots & c_{i,p} \\
\vdots & & & & \vdots & \vdots & & \vdots & & \vdots \\
a_{m,1} & a_{m,2} & \cdots & \cdots & a_{m,n} & c_{m,1} & \cdots & c_{m,j} & \cdots & c_{m,p}
\end{array}
$$

Nur die grau markierte Zeile bzw. Spalte von A und B geht in die Berechnung von $c_{i,j}$ ein. Folgendes Vorgehen:

- Bilden Sie „von außen kommend“ Zahlenpärchen der Form $(a_{i,k}, b_{k,j})$.
- Multiplizieren Sie jeweils die beiden Zahlen und summieren Sie sämtliche Ergebnisse.

Berechnen Sie für

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 3 & -1 & 2 \end{bmatrix} \quad \text{und} \quad B = \begin{bmatrix} 1 & 2 \\ 0 & -2 \\ 2 & 2 \end{bmatrix}$$

die Produkte AB und BA und vergleichen Sie die Ergebnisse. Was können Sie aus dem Vergleich schließen?

- Auch wenn beide Produkte AB und BA definiert sind (was beispielsweise für $A, B \in \mathbb{K}^{n\times n}$ der Fall ist), gilt i. A. $AB \neq BA$.
- Aus $AB = O$ (Nullmatrix) folgt **keineswegs** $A = O$ oder $B = O$.
- Selbst aus $A^2 = AA = O$ folgt **nicht** $A = O$.

Berechnen Sie A^2 für

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

Finden Sie eine Matrix $B \in \mathbb{R}^{2\times 3}$, so dass $AB = O$.

Satz 3.12

Im Zusammenhang mit der Matrizenmultiplikation gelten folgende Rechenregeln:

- $(AB)C = A(BC)$ *für alle* $A \in \mathbb{K}^{m\times n}$, $B \in \mathbb{K}^{n\times p}$, $C \in \mathbb{K}^{p\times q}$,
- $A(B+C) = AB + AC$ *für alle* $A \in \mathbb{K}^{m\times n}$, $B, C \in \mathbb{K}^{n\times p}$,
- $(A+B)C = AC + BC$ *für alle* $A, B \in \mathbb{K}^{m\times n}$, $C \in \mathbb{K}^{n\times p}$,
- $\lambda(AB) = (\lambda A)B = A(\lambda B)$ *für alle* $\lambda \in \mathbb{K}$, $A \in \mathbb{K}^{m\times n}$, $B \in \mathbb{K}^{n\times p}$,
- $(AB)^T = B^T A^T$ *und* $(AB)^H = B^H A^H$ *für alle* $A \in \mathbb{K}^{m\times n}$, $B \in \mathbb{K}^{n\times p}$,

Satz 3.13

- *Für die m-dimensionale Einheitsmatrix*

$$I_m := \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & & 1 & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \in \mathbb{K}^{m\times m},$$

  gilt $I_m A = A$ für alle $A \in \mathbb{K}^{m\times n}$.

- *Für die n-dimensionale Einheitsmatrix $I_n \in \mathbb{K}^{n\times n}$ gilt $AI_n = A$ für alle $A \in \mathbb{K}^{m\times n}$,*

- *Für die wie gekennzeichnet dimensionierten Nullmatrizen gilt $AO_{n\times p} = O_{m\times p}$ und $O_{q\times m}A = O_{q\times n}$ für alle $A \in \mathbb{K}^{m\times n}$.*

Ein Spezialfall der Matrizenmultiplikation ist die Multiplikation von Matrix und Vektor.

Für $A = [a_{i,j}] \in \mathbb{K}^{m \times n}$ und $\vec{x} \in \mathbb{K}^n$ ist $\vec{y} = A\vec{x}$ definiert durch

$$y_i = \sum_{j=1}^{n} a_{i,j} x_j \qquad (i = 1, 2, \dots, m).$$

Somit ist $\vec{y} = A\vec{x} = \sum_{j=1}^{n} x_j \vec{a}_j$ eine Linearkombination der Spalten $\vec{a}_j$ von A:

$$\left[\begin{array}{c} \\ \vec{y} \\ \\ \end{array}\right] = \left[\begin{array}{c|c|c|c} & & & \\ \vec{a}_1 & \vec{a}_2 & \dots & \vec{a}_n \\ & & & \end{array}\right] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1 \left[\begin{array}{c} \\ \vec{a}_1 \\ \\ \end{array}\right] + x_2 \left[\begin{array}{c} \\ \vec{a}_2 \\ \\ \end{array}\right] + \dots + x_n \left[\begin{array}{c} \\ \vec{a}_n \\ \\ \end{array}\right].$$

Jede Matrix $A \in \mathbb{K}^{m\times n}$ induziert nun über die Matrix-Vektor- Multiplikation eine Abbildung (die wir wieder mit A bezeichnen):

$$A : \mathbb{K}^n \to \mathbb{K}^m, \qquad \vec{x} \mapsto \vec{y} = A\vec{x}.$$

Der Definitionsbereich von A ist $\mathbb{K}^n$ und der Wertebereich von A ist in $\mathbb{K}^m$ enthalten. Es werden Vektoren auf Vektoren abgebildet.

Die Abbildung A besitzt zwei bemerkenswerte Eigenschaften:

- $A(\vec{x}_1 + \vec{x}_2) = A\vec{x}_1 + A\vec{x}_2$ für alle $\vec{x}_1, \vec{x}_2 \in \mathbb{K}^n$,
- $A(\lambda\vec{x}) = \lambda(A\vec{x})$ für alle $\lambda \in \mathbb{K}$, $\vec{x} \in \mathbb{K}^n$.

Diese beiden Eigenschaften fasst man unter dem Begriff der Linearität zusammen.

Die Definition der Linearität für allgemeine Vektorräume lautet wie folgt:

Definition 3.14 (Lineare Abbildung)

Seien V und W Vektorräume über $\mathbb{K}$. Eine Abbildung $f : V \to W$ heißt linear, wenn

- $f(\boldsymbol{x} + \boldsymbol{y}) = f(\boldsymbol{x}) + f(\boldsymbol{y})$ für alle $\boldsymbol{x}, \boldsymbol{y} \in V$,
- $f(\lambda \boldsymbol{x}) = \lambda f(\boldsymbol{x})$ für alle $\lambda \in \mathbb{K}$, $\boldsymbol{x} \in V$.

Sind folgende Abbildungen linear im Sinne von Definition 3.14:

- $f : \mathbb{R} \to \mathbb{R}$, $f(x) = 3x$,
- $g : \mathbb{R} \to \mathbb{R}$, $g(x) = 3x + 42$,
- $h : \mathbb{R}^2 \to \mathbb{R}$, $h(\vec{x}) = [1, 0] \cdot \vec{x}$?

Auf Seite 140 hatten wir bereits bestätigt, dass Abbildungen der Form

$$A : \mathbb{K}^n \to \mathbb{K}^m, \qquad \vec{x} \mapsto \vec{y} = A\vec{x},$$

linear sind.

Es stellt sich nun heraus, dass sich jede lineare Abbildung von $\mathbb{K}^n$ nach $\mathbb{K}^m$ in dieser Form schreiben lässt.

Satz 3.15

Sei $f : \mathbb{K}^n \to \mathbb{K}^m$ eine lineare Abbildung. Dann existiert eine eindeutig bestimmte Matrix $A \in \mathbb{K}^{m\times n}$, so dass

$$f(\vec{x}) = A\vec{x} \qquad \text{für alle } \ \vec{x} \in \mathbb{K}^n.$$

In den Spalten $\vec{a}_j$ von A stehen dabei gerade die Bilder der Einheitsvektoren $\vec{e}_j$, d. h.

$$\vec{a}_j = A\vec{e}_j \qquad (j = 1, \ldots, n).$$

Beweisidee:

Machen Sie sich die Aussage von Satz 3.15 klar, indem Sie

- verifizieren, dass eine lineare Abbildung $f : \mathbb{K}^n \to \mathbb{K}^m$ durch Angabe aller Werte $f(\vec{e}_j),\ j = 1, \dots, n$, eindeutig bestimmt ist,
- die Beziehung $\vec{a}_j = A\vec{e}_j$ bestätigen,
- die beiden genannten Teilschritte zu einer Gesamtargumentation zusammenfügen.

Zu einer gegebenen Matrix $A \in \mathbb{K}^{m \times n}$ (bzw. der zugeordneten linearen Abbildung) heißt

- $\mathscr{N}(A) := \{\vec{x} \in \mathbb{K}^n : A\vec{x} = \vec{0}\} \subseteq \mathbb{K}^n$ Nullraum oder Kern von A.
- $\mathscr{R}(A) := \{\vec{y} = A\vec{x} \in \mathbb{K}^m : \vec{x} \in \mathbb{K}^n\} \subseteq \mathbb{K}^m$ das Bild von A (entspricht dem Wertebereich im bei uns gebrauchten Sinne).

Aufgrund der Linearität besitzen sowohl $\mathscr{N}(A)$ als auch $\mathscr{R}(A)$ eine spezielle Struktur:

Satz 3.16

Sei $A \in \mathbb{K}^{m \times n}$. $\mathscr{N}(A)$ ist ein Unterraum von $\mathbb{K}^n$. $\mathscr{R}(A)$ ist ein Unterraum von $\mathbb{K}^m$. Es gilt

$$\dim \mathscr{N}(A) + \dim \mathscr{R}(A) = n.$$

$\dim \mathscr{N}(A)$ *heißt Defekt von A und* $\operatorname{rang}(A) := \dim \mathscr{R}(A)$ *Rang von A.*

Illustrieren Sie die Aussagen von Satz 3.16 am Beispiel $A = [1, 1] \in \mathbb{R}^{1\times 2}$.

Verifizieren Sie, dass $\mathscr{N}(A)$ ein Unterraum von $\mathbb{K}^n$ ist.

Bezeichnen $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n \in \mathbb{K}^m$ die Spalten von $A \in \mathbb{K}^{m\times n}$, dann ist

$$\mathscr{R}(A) = \operatorname{span}\{\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n\}.$$

(Warum?) Insbesondere ist die maximale Anzahl linear unabhängiger Spalten von A gleich dem Rang von A.

Satz 3.17

Für $A \in \mathbb{K}^{m \times n}$ gilt

$$\operatorname{rang}(A) = \dim \mathscr{R}(A) = \dim \mathscr{R}\left(A^T\right).$$

Die maximale Anzahl linear unabhängiger Spalten von A ist also gleich der maximalen Anzahl linear unabhängiger Zeilen, und beide stimmen mit dem Rang von A überein.

Insbesondere gilt

$$\operatorname{rang}(A) \leq \min\{m, n\}.$$

Bestimmen Sie den Rang der Matrizen $A = [1, 1]$ und $B = \left[\begin{smallmatrix} 1 & 2 \\ 2 & 4 \end{smallmatrix}\right]$.

Anmerkung: Effiziente Verfahren zur Rangbestimmung werden wir im nächsten Abschnitt kennenlernen.

Es gibt aber Matrizen, bei denen man den Rang sofort ablesen kann. Ein Beispiel sind Matrizen in Trapezform

$$A = \left[\begin{array}{ccccc|ccc} a_{1,1} & a_{1,2} & a_{1,3} & \dots & a_{1,r} & a_{1,r+1} & \dots & a_{1,n} \\ 0 & a_{2,2} & a_{2,3} & \dots & a_{2,r} & a_{2,r+1} & \dots & a_{2,n} \\ 0 & 0 & a_{3,3} & \dots & a_{3,r} & a_{3,r+1} & \dots & a_{3,n} \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & a_{r,r} & a_{r,r+1} & \dots & a_{r,n} \\ \hline 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{array}\right] \in \mathbb{K}^{m \times n},$$

mit $a_{j,j} \neq 0\,(j = 1, \dots, r)$, wobei der untere, der rechte oder beide Teile entfallen können. Der Rang solcher Matrizen ist stets gleich r.

Machen Sie sich klar, warum das so ist.

Inhalt

1. Grundlagen

2. Lineare Algebra

2.1 Vektorräume

2.2 Matrizen und lineare Abbildungen

2.3 Lineare Gleichungssysteme

2.4 Determinanten

2.5 Invertierbare Matrizen

2.6 Orthogonalität, Skalarprodukt und Norm

2.7 Kreuz- und Spatprodukt

2.8 Elemente der analytischen Geometrie

2.9 Orthogonale Abbildungen

2.10 Eigenwerte und Eigenvektoren

2.11 Singulärwertzerlegung

Lineare Gleichungssysteme

Ein System aus Gleichungen der Form

$$\begin{array}{ccccccccc} a_{1,1}x_1 & + & a_{1,2}x_2 & + & \cdots & + & a_{1,n}x_n & = & b_1 \\ a_{2,1}x_1 & + & a_{2,2}x_2 & + & \cdots & + & a_{2,n}x_n & = & b_2 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ a_{m,1}x_1 & + & a_{m,2}x_2 & + & \cdots & + & a_{m,n}x_n & = & b_m \end{array} \tag{3.1}$$

heißt lineares Gleichungssystem (LGS), genauer: ein System von m linearen algebraischen Gleichungen in n Unbekannten.

Dabei sind die Koeffizienten $a_{i,j} \in \mathbb{K}$ sowie die Zahlen $b_i \in \mathbb{K}$ vorgegeben, während man die Zahlen $x_i \in \mathbb{K}$ zu bestimmen sucht.

Lineare Gleichungssysteme

Mit der Koeffizientenmatrix

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix} \in \mathbb{K}^{m\times n},$$

dem Vektor der Unbekannten

$$\vec{x} = [x_1, x_2, \dots x_n]^T \in \mathbb{K}^n$$

sowie der rechten Seite

$$\vec{b} = [b_1, b_2, \dots, b_m]^T \in \mathbb{K}^m$$

schreibt man (3.1) kürzer als:

$$A\vec{x} = \vec{b}. \tag{3.2}$$

Ein lineares Gleichungssystem $A\vec{x} = \vec{b}$ heißt homogen, wenn $\vec{b} = \vec{0}$ und inhomogen, wenn $\vec{b} \neq \vec{0}$.

Ein homogenes System besitzt immer (mindestens) eine Lösung, nämlich $\vec{x} = \vec{0}$. Die Lösungsmenge von $A\vec{x} = \vec{0}$ ist gerade der Kern $\mathscr{N}(A)$ von A.

Bei inhomogenen Systemen ist die Sache komplizierter. Wir beginnen mit folgendem Ergebnis:

Satz 3.18

Ist $A \in \mathbb{K}^{m\times n}$ *und ist* $\vec{b} \in \mathscr{R}(A)$, *d. h. es gibt (mindestens) ein* $\vec{x}_0 \in \mathbb{K}^n$ *mit* $A\vec{x}_0 = \vec{b}$, *dann gilt*

$$\begin{aligned} \{\vec{x} \in \mathbb{K}^n:\ A\vec{x} = \vec{b}\} &= \{\vec{x}_0\} + \mathscr{N}(A) \qquad (3.3) \\ &= \{\vec{x}_0 + \vec{y}:\ y \in \mathscr{N}(A)\}. \end{aligned}$$

Im Falle der Existenz einer Lösung erhält man also die Lösungsmenge des inhomogenen Systems durch Verschieben der Lösungsmenge des homogenen Systems zur gleichen Koeffizientenmatrix.

Machen Sie sich dies anhand des relativ trivialen Gleichungssystems

$$x_1 + x_2 = 7$$

graphisch klar. Verwenden Sie im Ansatz (3.3) aus Satz 3.18 auch verschiedene $\vec{x}_0$.

Verifizieren Sie die Beziehung (3.3).

Zur weiteren Analyse der Lösbarkeit von LGS beleuchten wir die Lösungssuche noch unter einem weiteren Aspekt.

Sind $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n$ die Spalten der Matrix $A \in \mathbb{K}^{m \times n}$, dann lässt sich $A\vec{x} = \vec{b}$ auch schreiben als (vgl. S. 139):

$$x_1\vec{a}_1 + x_2\vec{a}_2 + \cdots + x_n\vec{a}_n = \vec{b}$$

Die Suche nach einer Lösung von $A\vec{x} = \vec{b}$ ist also gleichbedeutend mit der Suche nach Linearkombinationen der Spalten $\vec{a}_j$, die die rechte Seite $\vec{b}$ ergeben.

Wir halten somit fest:

Satz 3.19 (Lösbarkeit linearer Gleichungssysteme)

Es seien $A \in \mathbb{K}^{m \times n}$ mit den Spalten $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n$ und $\vec{b} \in \mathbb{K}^m$.

Das lineare Gleichungssystem $A\vec{x} = \vec{b}$ ist genau dann lösbar, wenn

$$\vec{b} \in \mathscr{R}(A) = \operatorname{span}\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\}.$$

*Das lineare Gleichungssystem $A\vec{x} = \vec{b}$ ist genau dann **eindeutig** lösbar, wenn $\vec{b} \in \mathscr{R}(A)$ gilt, und die Spalten $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n$ linear unabhängig sind:*

$$A\vec{x} = \vec{b} \text{ ist eindeutig lösbar} \quad \Leftrightarrow \quad \vec{b} \in \mathscr{R}(A) \;\; \text{und} \;\; \operatorname{rang}(A) = n.$$

Zusammenfassung: Es können somit genau drei Fälle eintreten:

- $A\vec{x} = \vec{b}$ besitzt **keine** Lösung, d. h. $\vec{b} \notin \mathscr{R}(A)$ (kann nur bei inhomogenen LGS passieren).
- $A\vec{x} = \vec{b}$ besitzt **genau eine** Lösung, d. h. $\vec{b} \in \mathscr{R}(A)$ und $\operatorname{rang}(A) = n$.
  Dieser Fall kann nur für $m \geq n$ eintreten (mindestens so viele Gleichungen wie Unbekannte).
- $A\vec{x} = \vec{b}$ besitzt **unendlich viele** Lösungen, d. h. $\vec{b} \in \mathscr{R}(A)$ und $\operatorname{rang}(A) < n$.
  Mit einer beliebigen Lösung $\vec{x}_0$ erhält man als Lösungsmenge
  $\vec{x}_0 + \mathscr{N}(A) = \vec{x}_0 + \{$alle Lösungen von $A\vec{x} = \vec{0}\}$.

Lineare Gleichungssysteme mit mehr als einer, aber nur endlich vielen Lösungen existieren nicht.

Bislang wissen wir zwar über die Lösbarkeit eines LGS $A\vec{x} = \vec{b}$ Bescheid; können aber die Lösung selbst noch nicht konkret berechnen.

Für diese Aufgabe stehen verschiedenste Algorithmen zur Verfügung, so zum Beispiel die hier behandelten:

- Gauß-Algorithmus,
- Gauß-Jordan-Algorithmus – eine Erweiterung des Gauß-Algorithmus.

Desweiteren gibt es eine ganze Reihe numerischer Algorithmen, die Sie mittels Literatur oder in einer Numerik-Vorlesung erlernen können.

Die hier vorgestellten exakten Verfahren beruhen auf Umformungen in ein äquivalentes Gleichungssystem $\tilde{A}\vec{x} = \tilde{\vec{b}}$, für das man die Lösung relativ mühelos angeben kann.

Zunächst erzeugen wir aus der Koeffizientenmatrix $A \in \mathbb{K}^{m \times n}$ und der rechten Seite $b \in \mathbb{K}^m$ die erweiterte Koeffizientenmatrix $[A \,|\, \vec{b}]$. Diese enthält die vollständige Information über das betrachtete LGS.

Satz 3.20

Die folgenden elementaren Umformungen der erweiterten Koeffizientenmatrix verändern die Lösungsmenge eines linearen Gleichungssystems nicht:

- *Multiplikation einer Gleichung (Zeile) mit $\lambda \in \mathbb{K}$, $\lambda \neq 0$,*
- *Addition des Vielfachen einer Gleichung (Zeile) zu einer anderen Gleichung (Zeile),*
- *Vertauschung von zwei Gleichungen (Zeilen),*
- *Umnummerierung von zwei Unbekannten (Vertauschung von zwei Spalten). Hier muss man sich allerdings merken, welche Unbekannten man umnummeriert hat.*

Beim Gauß-Algorithmus wird das LGS $A\vec{x} = \vec{b}$ in ein äquivalentes System $\tilde{A}\vec{x} = \tilde{\vec{b}}$ mit einer Matrix $[\tilde{A} \,|\, \tilde{\vec{b}}]$ in Trapezform überführt:

$$\left[\begin{array}{cccccccc|c}
\tilde{a}_{1,1} & \tilde{a}_{1,2} & \tilde{a}_{1,3} & \dots & \tilde{a}_{1,r} & \tilde{a}_{1,r+1} & \dots & \tilde{a}_{1,n} & \tilde{b}_1 \\
0 & \tilde{a}_{2,2} & \tilde{a}_{2,3} & \dots & \tilde{a}_{2,r} & \tilde{a}_{2,r+1} & \dots & \tilde{a}_{2,n} & \tilde{b}_2 \\
0 & 0 & \tilde{a}_{3,3} & \dots & \tilde{a}_{3,r} & \tilde{a}_{3,r+1} & \dots & \tilde{a}_{3,n} & \tilde{b}_3 \\
\vdots & \vdots & \ddots & \ddots & \vdots & \vdots & & \vdots & \vdots \\
0 & 0 & 0 & \dots & \tilde{a}_{r,r} & \tilde{a}_{r,r+1} & \dots & \tilde{a}_{r,n} & \tilde{b}_r
\end{array}\right]$$

mit $\tilde{a}_{j,j} \neq 0 \, (j = 1, \dots, r)$.

Die Lösung solcher „gestaffelter“ Gleichungssysteme ist durch Rückwärtsauflösen leicht zu bestimmen.

$$
\left[\begin{array}{ccccccccc|c}
{\color{red}*} & * & * & \cdots & * & * & * & \cdots & * & * \\
0 & * & * & \cdots & * & * & * & \cdots & * & * \\
0 & * & * & \cdots & * & * & * & \cdots & * & * \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & & \vdots & \vdots \\
0 & * & * & \cdots & * & * & * & \cdots & * & * \\
0 & * & * & \cdots & * & * & * & \cdots & * & *
\end{array}\right]
$$

$$
\left[\begin{array}{ccccccccc|c}
{\color{red}*} & * & * & \cdots & * & * & * & \cdots & * & * \\
0 & {\color{red}*} & * & \cdots & * & * & * & \cdots & * & * \\
0 & 0 & * & \cdots & * & * & * & \cdots & * & * \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \\
0 & 0 & * & \cdots & * & * & * & \cdots & * & * \\
0 & 0 & * & \cdots & * & * & * & \cdots & * & *
\end{array}\right]
$$

$$\vdots$$

$$
\left[\begin{array}{ccccccccc|c}
{\color{red}*} & * & * & \cdots & * & * & * & \cdots & * & * \\
0 & {\color{red}*} & * & \cdots & * & * & * & \cdots & * & * \\
0 & 0 & {\color{red}*} & \cdots & * & * & * & \cdots & * & * \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & & \vdots & \vdots \\
0 & 0 & 0 & \cdots & {\color{red}*} & * & * & \cdots & * & * \\
0 & 0 & 0 & \cdots & 0 & {\color{red}*} & * & \cdots & * & *
\end{array}\right]
$$

roter Stern: Eintrag zwingend ungleich Null; schwarzer Stern: beliebiger Eintrag

Vor jedem Einzelschritt werden zunächst alle Zeilen der Form

$$0 \quad 0 \quad \ldots \quad 0 \;\big|\; 0$$

ersatzlos gestrichen.

Gibt es Zeilen der Form

$$0 \quad 0 \quad \ldots \quad 0 \;\big|\; \tilde{b}_j,$$

(mit $\tilde{b}_j \neq 0$) so wird der Algorithmus abgebrochen; das LGS hat dann keine Lösung.

Warum sind diese Schritte sinnvoll?

Schritt 1:

- Ist $a_{1,1} = 0$, so suche in der ersten Spalte nach einem Eintrag $a_{j,1} \neq 0$. Existiert ein solcher, tausche Zeilen 1 und j.
- Ist $a_{1,1} = 0$ und der erste Punkt war nicht erfolgreich, so existiert in der ersten Zeile ein Eintrag $a_{1,k} \neq 0$

  Tausche in diesem Fall Spalten 1 und k; merke wie die Unbekannten umnummeriert wurden,
- Ist jetzt $\tilde{a}_{1,1} \neq 0$, so erzeuge unterhalb von $\tilde{a}_{1,1}$ lauter Nullen durch Addition des jeweils $(-\frac{\tilde{a}_{i,1}}{\tilde{a}_{1,1}})$-fachen der ersten Zeile zur Zeile i.

Die erste Zeile und Spalte von $[\tilde{A} \,|\, \tilde{\vec{b}}]$ ist damit ermittelt und wird nicht mehr angefasst.

Schritt 2 etc.:

Nach Ausführung der Vorarbeiten (S. 160) wird nun mit der um die erste Zeile und Spalte reduzierten Matrix analog verfahren usw.

Im Falle der Existenz einer Lösung steht am Ende des Gauß-Algorithmus ein äquivalentes System der Form

$$\begin{bmatrix} \tilde{a}_{1,1} & \tilde{a}_{1,2} & \tilde{a}_{1,3} & \dots & \tilde{a}_{1,r} & \tilde{a}_{1,r+1} & \dots & \tilde{a}_{1,n} \\ 0 & \tilde{a}_{2,2} & \tilde{a}_{2,3} & \dots & \tilde{a}_{2,r} & \tilde{a}_{2,r+1} & \dots & \tilde{a}_{2,n} \\ 0 & 0 & \tilde{a}_{3,3} & \dots & \tilde{a}_{3,r} & \tilde{a}_{3,r+1} & \dots & \tilde{a}_{3,n} \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & \tilde{a}_{r,r} & \tilde{a}_{r,r+1} & \dots & \tilde{a}_{r,n} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_r \end{bmatrix}.$$

Die Variablen $x_{r+1}, \dots, x_n$ können als Parameter frei gewählt werden ($x_{r+1} = \lambda_{r+1}, \dots, x_n = \lambda_n$).

Die anderen Variablen $x_r, x_{r-1}, \dots, x_1$ lassen sich dann bestimmen, indem man die Gleichungen von unten her sukzessive auflöst.

Achtung: Umnummerierungen muss man natürlich berücksichtigen!

Trainieren Sie den Gauß-Algorithmus:

1. Bestimmen Sie die Lösungen von $A\vec{x} = \vec{b}$ und $A\vec{x} = \vec{0}$ für

$$A = \begin{bmatrix} 2 & -1 & -1 & 2 \\ 6 & -2 & 3 & -1 \\ -4 & 2 & 3 & -2 \\ 2 & 0 & 4 & -3 \end{bmatrix} \quad \text{und} \quad \vec{b} = \begin{bmatrix} 3 \\ -3 \\ -2 \\ -1 \end{bmatrix}$$

2. Bestimmen Sie die Lösungen von

$$\begin{aligned} 6x_1 + 4x_2 + 8x_3 + 17x_4 &= -20 \\ 3x_1 + 2x_2 + 5x_3 + 8x_4 &= -8 \\ 3x_1 + 2x_2 + 7x_3 + 7x_4 &= -4 \\ 2x_3 - x_4 &= 4 \end{aligned}$$

3. Wieviele Lösungen beisitzt das LGS $\left[\begin{smallmatrix} 1 & 2 \\ 2 & 4 \end{smallmatrix}\right] \left[\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right]$?

Die elementaren Umformungen von S. 157 lassen nicht nur die Lösung eines LGS unverändert.

Sie erhalten desweiteren natürlich auch den Rang der Koeffizientenmatrix. Es gilt also mit den bisherigen Bezeichnungen und den Erkenntnissen von S. 147:

$$\operatorname{rang}(A) = \operatorname{rang}(\tilde{A}) = r.$$

Somit steht uns mit dem Gauß-Verfahren auch ein effizientes Verfahren zur Rangbestimmung von Matrizen zur Verfügung.

Geben Sie die Ränge der Koeffizientenmatrizen aus den Beispielen von S. 163 an.

Beim Gauß-Jordan-Algorithmus wird das LGS $A\vec{x} = \vec{b}$, $A \in \mathbb{K}^{m \times n}$, ebenfalls in ein äquivalentes System mit Trapezform überführt.

Allerdings führt man die Umformungen noch weiter und erzeugt schließlich eine Matrix $[\tilde{A} \,|\, \tilde{\vec{b}}]$, in der $\tilde{A}$ folgende Blockstruktur besitzt:

$$\tilde{A} = [I_r \; R] = \begin{bmatrix} 1 & 0 & \dots & 0 & * & \dots & * \\ 0 & 1 & \dots & 0 & * & \dots & * \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \\ 0 & 0 & \dots & 1 & * & \dots & * \end{bmatrix}. \tag{3.4}$$

Dabei ist I_r die r-dimensionale Einheitsmatrix und $R \in \mathbb{K}^{r \times (n-r)}$ eine beliebige Restmatrix, die auch entfallen kann

Wir wollen hier nur den Fall betrachten, dass $A\vec{x} = \vec{b}$ eindeutig lösbar ist ($r = n$). Das Gleichungssystem wird dann in ein äquivalentes System $I_n\vec{x} = \tilde{\vec{b}}$ überführt, dessen Lösung $\vec{x} = \tilde{\vec{b}}$ direkt abgelesen werden kann.

(A quadratisch mit vollem Rang)

$$
\left[\begin{array}{cccccc|c}
1 & * & * & \dots & * & * & * \\
0 & * & * & \dots & * & * & * \\
0 & * & * & \dots & * & * & * \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \\
0 & * & * & \dots & * & * & * \\
0 & * & * & \dots & * & * & *
\end{array}\right]
\to
\left[\begin{array}{cccccc|c}
1 & 0 & * & \dots & * & * & * \\
0 & 1 & * & \dots & * & * & * \\
0 & 0 & * & \dots & * & * & * \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \\
0 & 0 & * & \dots & * & * & * \\
0 & 0 & * & \dots & * & * & *
\end{array}\right]
$$

$$
\to \cdots \to
\left[\begin{array}{cccccc|c}
1 & 0 & 0 & \dots & 0 & 0 & * \\
0 & 1 & 0 & \dots & 0 & 0 & * \\
0 & 0 & 1 & \dots & 0 & 0 & * \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \\
0 & 0 & 0 & \dots & 1 & 0 & * \\
0 & 0 & 0 & \dots & 0 & 1 & *
\end{array}\right]
$$

Bei Bedarf können auch mehrere rechte Seiten parallel mitgeführt werden.

- Ist $a_{k,k} = 0$, dann tausche die Zeilen (oder Spalten) wie beim Gauß-Algorithmus, so dass $\tilde{a}_{k,k} \neq 0$.
- Teile nun die k-te Zeile durch $\tilde{a}_{k,k}$, d. h. erzeuge eine 1 an Position (k, k).
- Erzeuge unter- und oberhalb der Position (k, k) in **allen** Zeilen Nullen durch Addition des jeweils $(-\tilde{a}_{i,k})$-fachen der k-ten Zeile zur Zeile i.

Wie beim Gauß-Algorithmus muss dabei über unterwegs vorgenommene Umnummerierungen von Variablen (Spaltenvertauschungen) sorgfältig buchgeführt werden.

Fehlt in (3.4) der Block R, so ist das Gleichungssystem $A\vec{x} = \vec{b}$ äquivalent zu $I_n\vec{x} = \tilde{\vec{b}}$. Es gibt eine eindeutige Lösung, die man direkt ablesen kann:

$$\vec{x} = \tilde{\vec{b}}.$$

Ist das Gleichungssystem äquivalent zu $[I_r \;\; R]\,\vec{x} = \tilde{\vec{b}}$ mit $r < n$, so teilt man $\vec{x}$ in zwei Teile $\vec{x}^1 = [x_1, \ldots, x_r]^T$ und $\vec{x}^2 = [x_{r+1}, \ldots, x_n]^T$.

Die Variablen $x_{r+1}, \ldots, x_n$ können als Parameter frei gewählt werden. Wegen

$$\tilde{\vec{b}} = \begin{bmatrix} I_r & R \end{bmatrix} \begin{bmatrix} \vec{x}^1 \\ \vec{x}^2 \end{bmatrix} = \vec{x}^1 + R\vec{x}^2$$

ist die Lösung von $A\vec{x} = \vec{b}$ gegeben durch

$$\mathscr{L} = \left\{ \vec{x} = \begin{bmatrix} \vec{x}^1 \\ \vec{x}^2 \end{bmatrix} : \; \vec{x}^2 \in \mathbb{R}^{n-r} \text{ beliebig und } \vec{x}^1 = \tilde{\vec{b}} - R\vec{x}^2 \right\}.$$

Ob man ein LGS mittels Gauß- oder Gauß-Jordan-Verfahren löst, ist letztlich Geschmackssache.

Beim Gauß-Jordan-Verfahren spart man das Rückwärtseinsetzen, muss aber eine höhere Anzahl elementarer Umformungen in Kauf nehmen.

Besonders vorteilhaft ist Gauß-Jordan jedoch dann, wenn es um die Lösung mehrerer LGS mit gleicher Koeffizientenmatrix $A \in \mathbb{R}^{n\times n}$, $\mathrm{rang}(A) = n$, und verschiedenen rechten Seiten geht:

$$A\vec{x} = \vec{b}_1, \quad A\vec{x} = \vec{b}_2, \quad \dots, \quad A\vec{x} = \vec{b}_\ell$$

Diese Aufgabe ist gleichbedeutend mit der Lösung der Matrixgleichung

$$AX = B,$$

wobei $X \in \mathbb{R}^{n\times\ell}$ gesucht ist, und $B \in \mathbb{R}^{n\times\ell}$ als Spalten gerade die rechten Seiten $\vec{b}_i$ enthält.

Mit Hilfe des Gauß-Jordan-Algorithmus bestimme man die Lösung der Matrixgleichung $AX = B$, wobei

$$A = \begin{bmatrix} 3 & 5 & 1 \\ 2 & 4 & 5 \\ 1 & 2 & 2 \end{bmatrix} \quad \text{und} \quad B = \begin{bmatrix} 6 & 13 \\ 17 & 10 \\ 7 & 5 \end{bmatrix}.$$

Inhalt

1. Grundlagen

2. Lineare Algebra

2.1 Vektorräume
2.2 Matrizen und lineare Abbildungen
2.3 Lineare Gleichungssysteme
2.4 Determinanten
2.5 Invertierbare Matrizen
2.6 Orthogonalität, Skalarprodukt und Norm
2.7 Kreuz- und Spatprodukt
2.8 Elemente der analytischen Geometrie
2.9 Orthogonale Abbildungen
2.10 Eigenwerte und Eigenvektoren
2.11 Singulärwertzerlegung

Determinanten

Die Determinante ist eine Abbildung, die einer quadratischen Matrix $A \in \mathbb{K}^{n \times n}$ eine Zahl zuordnet. Sie wird u. A. benutzt um

- die eindeutige Lösbarkeit von LGS,
- die lineare Unabhängigkeit von n Vektoren,
- die Veränderung von Flächen/Volumina bei Anwendung der Abbildung A

zu beschreiben.

Historisches zum Determinantenbegriff

Determinanten mit $n = 2$ wurden erstmals am Ende des 16. Jh. von Cardano, größere ca. 100 Jahre später von Leibniz behandelt. Ein moderner axiomatischer Ansatz (1864) geht auf Weierstraß zurück.

Wir werden den axiomatischen Aufbau hier umgehen und für die Fälle mit $n \leq 3$ explizite Formeln angeben. Für $n > 3$ definieren wir die Determinante induktiv.

$$n = 1: \qquad A = \begin{bmatrix} a_{1,1} \end{bmatrix}, \quad \det(A) := a_{1,1}.$$

$$n = 2: \qquad A = \begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix}, \quad \det(A) := a_{1,1}a_{2,2} - a_{1,2}a_{2,1}.$$

$$n = 3: \qquad A = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{bmatrix},$$

$$\begin{aligned} \det(A) := {} & +a_{1,1}a_{2,2}a_{3,3} + a_{1,2}a_{2,3}a_{3,1} + a_{1,3}a_{2,1}a_{3,2} \\ & -a_{1,3}a_{2,2}a_{3,1} - a_{1,1}a_{2,3}a_{3,2} - a_{1,2}a_{2,1}a_{3,3}. \end{aligned}$$

Im Falle $n = 3$ verwendet man zur Berechnung einer Determinante gerne folgendes Schema:

$$
\begin{array}{ccccc}
a_{1,1} & a_{1,2} & a_{1,3} & a_{1,1} & a_{1,2} \\
a_{2,1} & a_{2,2} & a_{2,3} & a_{2,1} & a_{2,2} \\
a_{3,1} & a_{3,2} & a_{3,3} & a_{3,1} & a_{3,2} \\
- & - & -,+ & + & +
\end{array}
$$

- Man schreibt 1. und 2. Spalte nochmals neben die Determinante.
- Entlang der Diagonalen ermittelt man die Produkte der Einträge und versieht die Ergebnisse mit den dargestellten Vorzeichen.
- Man summiert die 6 vorzeichenbehafteten Produkte.

* Pierre Frédéric Sarrus, 1798-1861, französischer Mathematiker

Für Determinanten mit $n \geq 4$ gehen wir induktiv vor, greifen also auf Determinanten geringerer Größe zurück.

Für $i, j \in \{1, 2, \ldots, n\}$ bezeichne $A_{i,j} \in \mathbb{K}^{(n-1)\times(n-1)}$ diejenige Matrix, die durch Streichen der Zeile i und der Spalte j aus A entsteht.

Sei j nun ein beliebiger Spaltenindex. Dann definieren wir

$$\det(A) := \sum_{i=1}^{n} a_{i,j} (-1)^{i+j} \det(A_{i,j}) \tag{3.5}$$

Diese Darstellung wird Laplace-Entwicklung nach der j-ten Spalte genannt.

- Natürlich müsste man zunächst zeigen, dass die Darstellung in (3.5) von der gewählten Spalte unabhängig ist.
- Bei Berechnungen mit Hilfe der Laplace-Entwicklung ist es häufig zweckmäßig, Spalten (oder Zeilen, siehe später) mit möglichst vielen Nulleinträgen zu wählen.
- Die Laplace-Entwicklung (3.5) gilt auch für Determinanten mit $n = 2$ oder 3.
- In der Literatur wird die Determinante häufig anders aufgebaut. Formel (3.5) ist dann ein Bestandteil des **Laplaceschen Entwicklungssatzes**.

Man verwendet für Deteminanten mit $n \geq 2$ auch folgende verkürzende Schreibweise:

$$\begin{vmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,n} \end{vmatrix} := \det\left(\begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,n} \end{bmatrix} \right).$$

Berechnen Sie folgende Determinanten:

$$\begin{vmatrix} 1 & 2 \\ 2 & 4 \end{vmatrix}, \begin{vmatrix} 1 & 0 & 2 \\ 3 & -1 & 1 \\ 3 & 1 & 5 \end{vmatrix}, \begin{vmatrix} 1 & 42 & 23 \\ 0 & 17 & -110 \\ 0 & 0 & 3 \end{vmatrix}, \begin{vmatrix} 1 & 0 & 1 & 3 \\ -1 & 2 & 0 & 0 \\ 0 & 2 & -1 & 0 \\ 1 & -2 & 0 & -1 \end{vmatrix}.$$

Satz 3.21

Für $A = [a_{i,j}] \in \mathbb{K}^{n\times n}$ *und* $\lambda \in \mathbb{K}$ *gelten:*

- *Multipliziert man eine Spalte von* A *mit* λ*, so multipliziert sich auch die Determinante mit* λ*:*

$$\begin{vmatrix} a_{1,1} & \cdots & \lambda a_{1,j} & \cdots & a_{1,n} \\ a_{2,1} & \cdots & \lambda a_{2,j} & \cdots & a_{2,n} \\ \vdots & & \vdots & & \vdots \\ a_{n,1} & \cdots & \lambda a_{n,j} & \cdots & a_{n,n} \end{vmatrix} = \lambda \cdot \begin{vmatrix} a_{1,1} & \cdots & a_{1,j} & \cdots & a_{1,n} \\ a_{2,1} & \cdots & a_{2,j} & \cdots & a_{2,n} \\ \vdots & & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,j} & \cdots & a_{n,n} \end{vmatrix}$$

- *Die Determinante einer Matrix* A *ändert sich nicht, wenn man zu einer Spalte das Vielfache einer* ***anderen*** *Spalte addiert.*
- *Vertauscht man in einer quadratischen Matrix zwei verschiedene Spalten, so multipliziert sich die Determinate mit* -1*.*

Satz 3.22

Seien $A, B \in \mathbb{K}^{n \times n}$, $A = [a_{i,j}]$ *und* $\lambda \in \mathbb{K}$. *Dann gelten:*

- $\det(\lambda A) = \lambda^n \det(A)$.
- $\det(A^T) = \det(A)$.
- $\det(AB) = \det(A)\det(B)$.
- *Ist* A *eine Dreiecksmatrix (d. h.* $a_{i,j} = 0$ *für alle* $i < j$ *oder alle* $i > j$*), dann gilt*

  $$\det(A) = \prod_{i=1}^{n} a_{i,i},$$

  d. h. die Determinante ist das Produkt der (Haupt-) Diagonalelemente.

- Wegen (3.5) und Satz 3.22, Punkt 2, lässt sich eine Determinante auch nach der i-ten Zeile entwickeln:

$$\det(A) = \sum_{j=1}^{n} a_{i,j} (-1)^{i+j} \det(A_{i,j})$$

- Elementare Zeilen- und Spaltenumformungen wie im Gauß- Algorithmus können die Determinante der Koeffizientenmatrix ändern.

  Die Eigenschaft, gleich oder ungleich Null zu sein, bleibt dabei aber immer erhalten.

- Die Determinante einer Dreiecksmatrix ist genau dann Null, wenn für eines der Diagonalelemente $a_{i,i} = 0$ gilt.

Kombiniert man die letzten beiden Punkte mit den Kenntnissen über den Endzustand der Matrix $\tilde{A}$ beim Gauß-Algorithmus, so ergibt sich weiterhin:

- Ist $\operatorname{rang}(A) = n$, so ist $\det(A) \neq 0$.
- Ist $\operatorname{rang}(A) < n$, so ist $\det(A) = 0$.

Dabei ist natürlich immer $A \in \mathbb{K}^{n \times n}$ anzusetzen.

Das bedeutet wiederum, dass man die eindeutige Lösbarkeit eines linearen Gleichungssystems leicht mit der Determinante verifizieren kann.

Bestimmen Sie alle $\alpha \in \mathbb{R}$, für die das LGS $A\vec{x} = \vec{b}$ mit

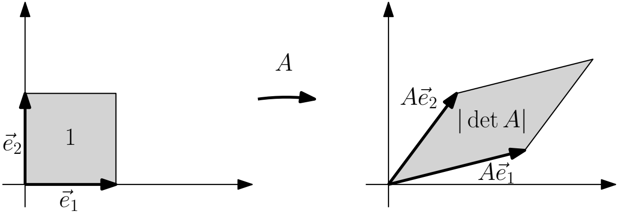
$$A = \begin{bmatrix} \alpha & 0 & 0 \\ 0 & 2\alpha & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

und beliebigem $\vec{b} \in \mathbb{R}^3$ eine eindeutige Lösung besitzt

Sei $A \in \mathbb{R}^{n\times n}$ und $\mathrm{vol}(S)$ das n-dimensionale Volumen* einer geeigneten** Punktmenge S. Dann ist das Volumen des Bildes $f(S)$ unter $f(\vec{x}) = A\vec{x}$ gegeben durch

$$\mathrm{vol}(f(S)) = |\det(A)| \cdot \mathrm{vol}(S).$$

Determinanten können also als Flächen- oder Volumenverzerrungsfaktor interpretiert werden:



* in 2D ist das eine Fläche, in 3D das gewohnte Volumen
** “Geeignet“ lässt sich im Rahmen der Maßtheorie sauber definieren. Gängige geometrische Objekte machen i. d. R. kein Probleme.

Inhalt

1. Grundlagen

2. Lineare Algebra

2.1 Vektorräume
2.2 Matrizen und lineare Abbildungen
2.3 Lineare Gleichungssysteme
2.4 Determinanten
2.5 Invertierbare Matrizen
2.6 Orthogonalität, Skalarprodukt und Norm
2.7 Kreuz- und Spatprodukt
2.8 Elemente der analytischen Geometrie
2.9 Orthogonale Abbildungen
2.10 Eigenwerte und Eigenvektoren
2.11 Singulärwertzerlegung

Invertierbare Matrizen

Definition 3.23

Eine quadratische Matrix $A \in \mathbb{K}^{n\times n}$ heißt invertierbar oder regulär, wenn es eine Matrix $B \in \mathbb{K}^{n\times n}$ gibt mit

$$AB = BA = I_n. \tag{3.6}$$

Die Matrix B ist dann eindeutig durch A bestimmt, wird die Inverse von A genannt und mit A^{-1} bezeichnet.

Warum ist $B = A^{-1}$ eindeutig bestimmt? Sie finden die Antwort leicht, wenn Sie (3.6) als Menge von LGS der Form $A\vec{x} = \vec{e}_i$ lesen.

Die Inverse einer invertierbaren 2×2–Matrix

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix}$$

kann man explizit angeben:

$$A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} a_{2,2} & -a_{1,2} \\ -a_{2,1} & a_{1,1} \end{bmatrix} = \frac{1}{a_{1,1}a_{2,2} - a_{1,2}a_{2,1}} \begin{bmatrix} a_{2,2} & -a_{1,2} \\ -a_{2,1} & a_{1,1} \end{bmatrix}.$$

Für $n \geq 3$ existieren zwar auch Formeln, diese sind jedoch sperrig und kaum in Gebrauch.

Bestätigen Sie obige Formel durch Nachrechnen.

Im allgemeinen Fall muss man die Matrixgleichung $AX = I_n$ lösen, oder eben sämtliche LGSe

$$A\vec{x} = \vec{e}_i \quad (i = 1, \dots, n).$$

Für diese Aufgabe bietet sich der auf Folie 165 ff behandelte Gauß-Jordan-Algorithmus an.

$$\begin{array}{c|c} A & I_n \\ \hline \vdots & \vdots \\ \hline I_n & A^{-1} \end{array}$$

Man berechne die Inversen von

$$A = \begin{bmatrix} 1 & 23 \\ 0 & 42 \end{bmatrix} \quad \text{und} \quad B = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 3 & 0 \\ 3 & 4 & 1 \end{bmatrix}.$$

Satz 3.24

Sei $A \in \mathbb{K}^{n\times n}$. Dann sind folgende Aussagen äquivalent:

- *A ist invertierbar.*
- $\det(A) \neq 0$.
- *Es gilt $\operatorname{rang}(A) = n$, d. h. die Spalten (Zeilen) von A bilden eine Basis des $\mathbb{K}^n$.*
- *Das homogene System $A\vec{x} = \vec{0}$ besitzt nur die triviale Lösung $\vec{x} = \vec{0}$.*
- *Für jede rechte Seite $\vec{b} \in \mathbb{K}^n$ besitzt das System $A\vec{x} = \vec{b}$ genau eine Lösung, nämlich*
$$\vec{x} = A^{-1}\vec{b}.$$

Satz 3.25

Seien $A, B \in \mathbb{K}^{n \times n}$ *invertierbar. Dann gelten:*

- A^T *und* A^H *sind invertierbar, und es gilt*
$$(A^T)^{-1} = (A^{-1})^T \quad \textit{bzw.} \quad (A^H)^{-1} = (A^{-1})^H.$$
- A^{-1} *ist invertierbar, und es gilt*
$$(A^{-1})^{-1} = A.$$
- AB *ist invertierbar, und es gilt*
$$(AB)^{-1} = B^{-1}A^{-1}.$$

Inhalt

1. Grundlagen

2. Lineare Algebra

2.1 Vektorräume
2.2 Matrizen und lineare Abbildungen
2.3 Lineare Gleichungssysteme
2.4 Determinanten
2.5 Invertierbare Matrizen
2.6 Orthogonalität, Skalarprodukt und Norm
2.7 Kreuz- und Spatprodukt
2.8 Elemente der analytischen Geometrie
2.9 Orthogonale Abbildungen
2.10 Eigenwerte und Eigenvektoren
2.11 Singulärwertzerlegung

Orthogonalität, Skalarprodukt und Norm

- Visualisieren wir zwei Vektoren $\vec{x}, \vec{y}$ aus $\mathbb{R}^2$ oder $\mathbb{R}^3$ als Pfeile mit Startpunkt in $\vec{0}$, so haben wir eine recht konkrete Vorstellung, was mit „senkrecht" oder dem „Winkel zwischen $\vec{x}$ und $\vec{y}$" gemeint ist.
- Die bisherigen Überlegungen blenden diese Anschauung aber noch völlig aus – es fehlen mathematische Begriffe zur Beschreibung solcher Eigenschaften.
- Um diesen Mangel zu beheben, benötigen wir den Begriff des Innen- oder Skalarprodukts.
- Wir geben zunächst die allgemeine Form, ziehen uns aber dann im wesentlichen auf $\mathbb{R}^n$ (manchmal $\mathbb{C}^n$) zurück.

Definition 3.26

Sei V ein $\mathbb{K}$-Vektorraum, wobei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Dann heißt eine Abbildung

$$\langle \cdot, \cdot \rangle : V \times V \to \mathbb{K}$$

Innenprodukt oder Skalarprodukt, wenn für alle $\boldsymbol{x}, \boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{y} \in V$ und $\lambda \in \mathbb{K}$ gilt:

- $\langle \boldsymbol{x}, \boldsymbol{x} \rangle \geq 0$, wobei $\langle \boldsymbol{x}, \boldsymbol{x} \rangle = 0 \Leftrightarrow \boldsymbol{x} = \boldsymbol{0}$ („positiv definit"),
- $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{y}, \boldsymbol{x} \rangle$, falls $\mathbb{K} = \mathbb{R}$ („symmetrisch"),
  $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \overline{\langle \boldsymbol{y}, \boldsymbol{x} \rangle}$, falls $\mathbb{K} = \mathbb{C}$ („hermitesch"),
- $\langle \lambda \boldsymbol{x}_1 + \boldsymbol{x}_2, \boldsymbol{y} \rangle = \lambda \langle \boldsymbol{x}_1, \boldsymbol{y} \rangle + \langle \boldsymbol{x}_2, \boldsymbol{y} \rangle$ („linear im ersten Argument").

Wir nennen $\boldsymbol{x}, \boldsymbol{y} \in V$ orthogonal ($\boldsymbol{x} \perp \boldsymbol{y}$), wenn $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = 0$ gilt.

Hinweis: Denken Sie am besten bereits hier an das Standard-Skalarprodukt $\langle \vec{x}, \vec{y} \rangle = \sum_{i=1}^{n} x_i y_i$ im $\mathbb{R}^n$.

Hat man einmal ein Skalarprodukt festgelegt, so ist damit immer eine „Längenmessung“ für Vektoren verbunden:

Satz 3.27 (und Definition)

Sei V ein $\mathbb{K}$-Vektorraum mit $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$ und $\langle \cdot, \cdot \rangle$ ein Skalarprodukt auf V. Dann gilt Cauchy-Schwarzsche-Ungleichung

$$|\langle \boldsymbol{x}, \boldsymbol{y} \rangle|^2 \leq \langle \boldsymbol{x}, \boldsymbol{x} \rangle \langle \boldsymbol{y}, \boldsymbol{y} \rangle, \qquad (\boldsymbol{x}, \boldsymbol{y} \in V) \tag{CSU}$$

mit Gleichheit genau dann, wenn $\boldsymbol{x}$ und $\boldsymbol{y}$ linear abhängig sind. Die durch

$$\|\boldsymbol{x}\| := \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle} \qquad (\boldsymbol{x} \in V)$$

definierte Abbildung heißt die von diesem Skalarprodukt erzeugte Norm. Jede Norm erfüllt für alle $\boldsymbol{x}, \boldsymbol{y} \in V, \lambda \in \mathbb{K}$ folgende Beziehungen:

- $\|\boldsymbol{x}\| \geq 0$, *wobei* $\|\boldsymbol{x}\| = 0 \Leftrightarrow \boldsymbol{x} = \boldsymbol{0}$ *(„positiv definit“),*
- $\|\lambda \boldsymbol{x}\| = |\lambda| \|\boldsymbol{x}\|$ *(„homogen“),*
- $\|\boldsymbol{x} + \boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$ *(„Dreiecksungleichung“).*

Das (für uns) bei weitem wichtigste Skalarprodukt ist das Euklidische Skalarprodukt. Auf $\mathbb{R}^n$ ist es definiert durch

$$\langle \vec{x}, \vec{y} \rangle := \vec{y}^T \vec{x} = \sum_{i=1}^{n} x_i y_i \qquad (\vec{x}, \vec{y} \in \mathbb{R}^n).$$

Die Entsprechung in $\mathbb{C}^n$ lautet

$$\langle \vec{x}, \vec{y} \rangle := \vec{y}^H \vec{x} = \sum_{i=1}^{n} x_i \bar{y}_i \qquad (\vec{x}, \vec{y} \in \mathbb{C}^n),$$

so dass wir nun z. B. auch über Orthogonalität von Vektoren aus $\mathbb{C}^n$ entscheiden können.

Anmerkung: Man kann auch andere Skalarprodukte auf $\mathbb{K}^n$ definieren. Diese modellieren dann aber nicht unbedingt unsere klassischen Vorstellungen von Winkelmessung und Orthogonalität.

Prüfen Sie folgende Vektoren des $\mathbb{R}^3$ auf paarweise Orthogonalität:

$$\vec{a} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \quad \vec{c} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Bestimmen Sie das Skalarprodukt der Vektoren $\vec{x} = (1, i)^T$ und $\vec{y} = (1 + i, i)^T$ in $\mathbb{C}^2$. Lassen sich diese beiden Vektoren mit unseren gewohnten Vorstellungen visualisieren?

Bestätigen Sie die in Definition 3.26 genannten Punkte für einen der Fälle $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$.

Das Euklidische Skalarprodukt erzeugt die Euklidische Norm

$$\|\vec{x}\|_2 := \|\vec{x}\| := \sqrt{\sum_{i=1}^{n} |x_i|^2} \quad (\vec{x} \in \mathbb{R} \text{ bzw. } \vec{x} \in \mathbb{C}).$$

In $\mathbb{R}^n$ kann man die Betragsstriche weglassen, d. h. $\|\vec{x}\| = \sqrt{\sum_{i=1}^{n} x_i^2}$.

Nach dem Satz des Pythagoras ist dies gerade die Länge des $\vec{x}$ zugeordneten Vektorpfeils.

Machen Sie sich das am Beispiel des $\mathbb{R}^2$ anhand einer Skizze klar.

Berechnen Sie die Norm der Vektoren $\vec{a}, \vec{b}$ und $\vec{c}$ von Seite 194.

Anmerkung: In der Literatur findet man auch das Symbol $|\vec{x}|$ und die Bezeichnung „Betrag“ für die Euklidische Norm auf $\mathbb{R}^n$.

In $\mathbb{R}^2$ oder $\mathbb{R}^3$ lässt sich mittels elementarer Geometrie zeigen, dass für den von zwei Vektoren $\vec{x}$ und $\vec{y}$ eingeschlossenen Winkel $\phi \in [0, \pi)$ gilt:

$$\cos\phi = \frac{\vec{x}^T \vec{y}}{\|\vec{x}\| \, \|\vec{y}\|}. \tag{3.7}$$

Anmerkung: Mittels

$$\cos\phi = \frac{\langle \boldsymbol{x}, \boldsymbol{y} \rangle}{\|\boldsymbol{x}\| \|\boldsymbol{y}\|}$$

lassen sich auch Winkel in beliebigen Vektorräumen mit Skalarprodukt definieren*, dies nutzt man praktisch aber selten.

Die Wohldefiniertheit wird dabei durch die allgemeingültige Cauchy-Schwarz-Ungleichung

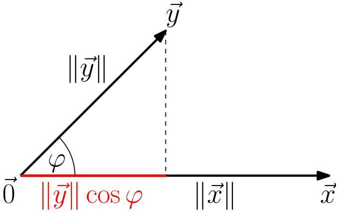
$$|\langle \boldsymbol{x}, \boldsymbol{y} \rangle| \le \|\boldsymbol{x}\| \, \|\boldsymbol{y}\| \qquad (\boldsymbol{x}, \boldsymbol{y} \in V)$$

gesichert. Machen Sie sich klar, welche Bedingung es eigentlich zu sichern gilt.

Aus Gleichung (3.7) erhält man in $\mathbb{R}^n$ desweiteren folgende Darstellung des Euklidischen Skalarprodukts:

$$\vec{x}^T \vec{y} = \|\vec{x}\| \, \|\vec{y}\| \, \cos\phi$$

Interpretiert man diese Formel an folgender Skizze,



so sieht man, dass das Skalarprodukt gleich der (vorzeichenbehafteten) Länge der orthogonalen Projektion von $\vec{y}$ auf $\vec{x}$, multipliziert mit der Länge von $\vec{x}$ ist.

Definition 3.28

Eine Basis $B = \{\vec{b}_1, \vec{b}_2, \dots, \vec{b}_n\}$ des $\mathbb{R}^n$ heißt Orthonormalbasis (ONB), wenn

$$\vec{b}_j^T \vec{b}_i = \begin{cases} 1, & \text{für } i = j, \\ 0, & \text{für } i \neq j. \end{cases}$$

Bei einer Orthonormalbasis stehen die Basisvektoren also paarweise aufeinander senkrecht und haben allesamt die Länge 1.

Ein prominentes Beispiel für eine Orthonormalbasis des $\mathbb{R}^n$ ist die Standardbasis $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$.

Aus den Vektoren $\vec{a}, \vec{b}, \vec{c}$ auf S. 194 kann man durch Normieren eine ONB des $\mathbb{R}^3$ erhalten. Geben Sie diese an.

Die Komponenten eines Vektors bezüglich einer Orthonormalbasis lassen sich besonders leicht über Orthogonalprojektionen berechnen:

Satz 3.29

Ist $B = \{\vec{b}_1, \vec{b}_2, \dots, \vec{b}_n\}$ eine Orthonormalbasis des $\mathbb{R}^n$, dann besitzt jeder Vektor $\vec{x} \in \mathbb{R}^n$ die eindeutige Darstellung

$$\vec{x} = \sum_{j=1}^{n} \left(\vec{b}_j^T \vec{x}\right) \vec{b}_j.$$

Machen Sie sich die Gültigkeit der Aussage für $\mathbb{R}^2$ anhand der Skizze auf S. 197 klar. Bestätigen Sie sie für den Fall der Standardbasis auf $\mathbb{R}^2$ auch rechnerisch.

Machen Sie sich klar, warum zwei von Null verschiedene orthogonale Vektoren immer linear unabhängig sind.

Aus einem Satz Vektoren $\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_n$ kann man durch folgendes Verfahren eine Orthonormalbasis des Unterraumes $\operatorname{span}\{\boldsymbol{a}_1, \dots, \boldsymbol{a}_n\}$ konstruieren.

$$
\begin{aligned}
\boldsymbol{b}_1 &= \frac{\boldsymbol{a}_1}{\|\boldsymbol{a}_1\|}, & & \\
\tilde{\boldsymbol{b}}_2 &= \boldsymbol{a}_2 - \langle \boldsymbol{a}_2, \boldsymbol{b}_1\rangle \boldsymbol{b}_1, & \boldsymbol{b}_2 &= \frac{\tilde{\boldsymbol{b}}_2}{\|\tilde{\boldsymbol{b}}_2\|}, \\
\tilde{\boldsymbol{b}}_3 &= \boldsymbol{a}_3 - \langle \boldsymbol{a}_3, \boldsymbol{b}_2\rangle \boldsymbol{b}_2 - \langle \boldsymbol{a}_3, \boldsymbol{b}_1\rangle \boldsymbol{b}_1, & \boldsymbol{b}_3 &= \frac{\tilde{\boldsymbol{b}}_3}{\|\tilde{\boldsymbol{b}}_3\|}, \\
&\vdots & & \\
\tilde{\boldsymbol{b}}_n &= \boldsymbol{a}_n - \sum_{k=1}^{n-1} \langle \boldsymbol{a}_n, \boldsymbol{b}_k\rangle \boldsymbol{b}_k, & \boldsymbol{b}_n &= \frac{\tilde{\boldsymbol{b}}_n}{\|\tilde{\boldsymbol{b}}_n\|}.
\end{aligned}
$$

Im Fall dass die Vektoren $\boldsymbol{a}_1, \dots, \boldsymbol{a}_n$ nur einen Unterraum der Dimension $m < n$ aufspannen gilt $\boldsymbol{b}_{m+1} = \dots = \boldsymbol{b}_n = \boldsymbol{0}$.

In der Quantenmechanik beschreibt man den Zustand von Teilchen (in einer Raumdimension) mittels einer komplexwertigen Wellen- oder Zustandsfunktion

$$\phi : \mathbb{R} \to \mathbb{C}.$$

Diese Funktion ist selbst schwer interpretierbar, allerdings liefert

$$\int_a^b |\phi(x)|^2 \, \mathrm{d}x$$

die Wahrscheinlichkeit, dass sich das Teilchen zwischen a und b aufhält.

Da sich das Teilchen an irgendeiner Stelle auf der Zahlengeraden befinden muss, ist es sinnvoll,

$$\int_{-\infty}^{\infty} |\phi(x)|^2 \, \mathrm{d}x = 1$$

zu fordern.

Als zugrundeliegende Räume verwendet man daher sogenannte L^2-Räume*. Grob gesprochen sind das Räume von Funktionen mit

$$\|f\|_2 := \left(\int_{-\infty}^{\infty} |f(x)|^2 \, \mathrm{d}x \right)^{\frac{1}{2}} < \infty.$$

Die so definierte L^2-Norm lässt sich mit folgendem (komplexen) Skalarprodukt erzeugen:

$$\langle f, \, g \rangle = \int_{-\infty}^{\infty} f(x) \, \overline{g(x)} \, \mathrm{d}x.$$

Eine reelle Version verwendet man häufig bei der theoretischen und numerischen Behandlung partieller Differentialgleichungen.

* Das "L" steht für den hier benötigten erweiterten Integralbegriff – das „Lebesgue-Integral" (Henri Lebesgue, 1875-1941).

Mit Hilfe einer Norm lässt sich schließlich ein Abstandsbegriff (Metrik) in V einführen. Der Abstand zweier Vektoren $\boldsymbol{x}, \boldsymbol{y} \in V$ wird durch

$$d(\boldsymbol{x}, \boldsymbol{y}) = \|\boldsymbol{x} - \boldsymbol{y}\| \tag{3.8}$$

erklärt.

Für alle $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in V$ gelten:

- $d(\boldsymbol{x}, \boldsymbol{x}) = 0$, $d(\boldsymbol{x}, \boldsymbol{y}) > 0$, falls $\boldsymbol{x} \neq \boldsymbol{y}$ (Definitheit),
- $d(\boldsymbol{x}, \boldsymbol{y}) = d(\boldsymbol{y}, \boldsymbol{x})$ (Symmetrie),
- $d(\boldsymbol{x}, \boldsymbol{y}) \leq d(\boldsymbol{x}, \boldsymbol{z}) + d(\boldsymbol{z}, \boldsymbol{y})$ (Dreiecksungleichung),
- $d(\boldsymbol{x} + \boldsymbol{z}, \boldsymbol{y} + \boldsymbol{z}) = d(\boldsymbol{x}, \boldsymbol{y})$ (Translationsinvarianz).

Setzt man wieder die Euklid-Norm auf $\mathbb{R}^n$ an, so ergibt sich aus (3.8) die gewohnte Abstandsformel für zwei Punkte $\vec{x}, \vec{y} \in \mathbb{R}^n$:

$$d(\vec{x}, \vec{y}) = \|\vec{x} - \vec{y}\| = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}.$$

Mit welchem klassischen Satz deckt sich diese Formel?

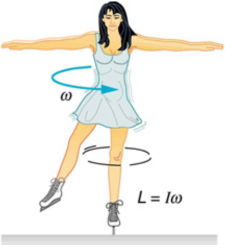
Berechnen Sie den Abstand der Punkte $(5, 2)^T$ und $(1, -1)^T$ in $\mathbb{R}^2$.

1 Grundlagen

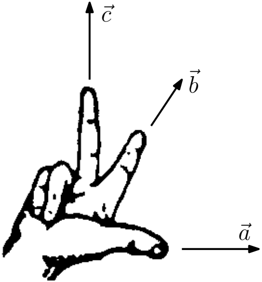
2 Lineare Algebra
2.1 Vektorräume
2.2 Matrizen und lineare Abbildungen
2.3 Lineare Gleichungssysteme
2.4 Determinanten
2.5 Invertierbare Matrizen
2.6 Orthogonalität, Skalarprodukt und Norm
2.7 Kreuz- und Spatprodukt
2.8 Elemente der analytischen Geometrie
2.9 Orthogonale Abbildungen
2.10 Eigenwerte und Eigenvektoren
2.11 Singulärwertzerlegung

Bei vielen physikalischen Anwendungen benötigt man eine weitere Operation, das sogenannte Vektor- oder Kreuzprodukt. Beispiele sind

- das Drehmoment $\vec{M} = \vec{r} \times \vec{F}$ zum Abstandsvektor $\vec{r}$ und zur Kraft $\vec{F}$,
- die Lorentzkraft $\vec{F_L} = q(\vec{v} \times \vec{B})$, welche auf ein Teilchen mit Ladung q und Geschwindigkeit $\vec{v}$ im Magnetfeld $\vec{B}$ wirkt;
- die Corioliskraft $\vec{F_C} = -2m(\vec{\omega} \times \vec{v})$, welche auf einen Körper der Masse m wirkt, welcher sich mit Geschwindigkeit $\vec{v}$ relativ zu einem mit Winkelgeschwindigkeit $\vec{\omega}$ rotierenden Bezugssystem bewegt;




Wir gehen bei der Definition des Kreuzprodukts von den gewünschten geometrischen Eigenschaften aus und leiten dann die Rechenregeln sowie die Formel zur Berechnung her. Sämtliche Vektoren in diesem Kapitel sind als Elemente des $\mathbb{R}^3$ aufzufassen.
Drei linear unabhängige Vektoren $\vec{a}, \vec{b}, \vec{c} \in \mathbb{R}^3$ bilden (in dieser Reihenfolge) ein Rechtssystem, wenn sich — von der Spitze von $\vec{c}$ aus gesehen — $\vec{a}$ durch Drehung um einen Winkel $\phi \in [0, \pi)$ im Gegenuhrzeigersinn in die gleiche Richtung wie $\vec{b}$ bringen lässt.



Ob ein Rechtssystem vorliegt, kann mit der Rechten-Hand-Regel (Korkenzieherregel) entschieden werden:
Analog kann man ein Linkssystem definieren.

Definition 3.30

Seien $\vec{a}, \vec{b} \in \mathbb{R}^3$ linear unabhängig. Der Vektor

$$\vec{c} = \vec{a} \times \vec{b} \ \in \mathbb{R}^3$$

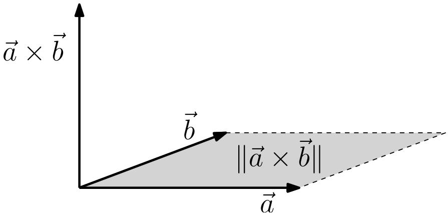
heißt Kreuzprodukt (auch Vektorprodukt oder äußeres Produkt) von $\vec{a}$ und $\vec{b}$, wenn

(1) $\vec{c} \perp \vec{a}$ und $\vec{c} \perp \vec{b}$ (d. h. $\vec{c} \perp \operatorname{span}\{\vec{a}, \vec{b}\}$),

(2) $\|\vec{c}\| = \|\vec{a}\| \, \|\vec{b}\| \, |\sin \measuredangle(\vec{a}, \vec{b})|$,

(3) $\vec{a}, \vec{b}, \vec{c}$ bilden in dieser Reihenfolge ein Rechtssystem.

Sind $\vec{a}$ und $\vec{b}$ linear abhängig (was die Fälle $\vec{a} = \vec{0}$ oder $\vec{b} = \vec{0}$ mit einschließt), so definieren wir $\vec{a} \times \vec{b} := \vec{0}$.



- Punkt (1) legt die Gerade fest, auf der $\vec{a} \times \vec{b}$ liegt. Die Richtung von $\vec{a} \times \vec{b}$ ist damit bis auf das Vorzeichen eindeutig bestimmt.
- Punkt (2) legt die Länge von $\vec{a} \times \vec{b}$ fest. Diese stimmt mit der Fläche des von $\vec{a}$ und $\vec{b}$ aufgespannten Parallelogramms überein.
- Punkt (3) liefert die Entscheidung, welcher der verbliebenen zwei Vektoren (Richtungen) zu verwenden ist.

Allein aufgrund der geometrischen Definition ergeben sich folgende Rechenregeln für Kreuzprodukte:

Satz 3.31

Seien $\vec{a}, \vec{b}, \vec{c} \in \mathbb{R}^3$ *und* $\lambda \in \mathbb{R}$. *Dann gelten:*

- $\vec{b} \times \vec{a} = -(\vec{a} \times \vec{b})$,
- $(\vec{a} + \vec{b}) \times \vec{c} = \vec{a} \times \vec{c} + \vec{b} \times \vec{c}$,
- $\lambda(\vec{a} \times \vec{b}) = (\lambda\vec{a}) \times \vec{b} = \vec{a} \times (\lambda\vec{b})$.

Illustrieren Sie zumindest den ersten und dritten Punkt anhand geeigneter Skizzen.

Satz 3.32

Für $\vec{a}, \vec{b} \in \mathbb{R}^3$ *gilt*

$$\vec{a} \times \vec{b} = \begin{bmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{bmatrix}. \tag{3.9}$$

Beweisidee:

- Für die Einheitsvektoren des $\mathbb{R}^3$ gilt offenbar $\vec{e}_1 \times \vec{e}_2 = \vec{e}_3$, $\vec{e}_2 \times \vec{e}_3 = \vec{e}_1$ und $\vec{e}_3 \times \vec{e}_1 = \vec{e}_2$.
- Schreibe $\vec{a} = a_1\vec{e}_1 + a_2\vec{e}_2 + a_3\vec{e}_3$ bzw. $\vec{b} = b_1\vec{e}_1 + b_2\vec{e}_2 + b_3\vec{e}_3$ und wende Satz 3.31 an.

Man führe die Rechnung zum zweitgenannten Punkt aus.

Formel (3.9) merkt man sich am besten mit Hilfe der formalen 3×3-Determinante

$$\vec{a} \times \vec{b} = \det\left(\begin{bmatrix} \vec{e}_1 & \vec{e}_2 & \vec{e}_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix}\right),$$

die man mit der Regel von Sarrus auswertet.

Berechnen Sie auf diese Weise das Kreuzprodukt

$$\begin{bmatrix} 1 \\ 2 \\ -4 \end{bmatrix} \times \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}.$$

Mitunter ist noch eine Kombination von Kreuz- und Skalarprodukt in Gebrauch. Unter dem Spatprodukt (auch gemischten Produkt) dreier Vektoren $\vec{a}, \vec{b}, \vec{c} \in \mathbb{R}^3$ versteht man die reelle Zahl

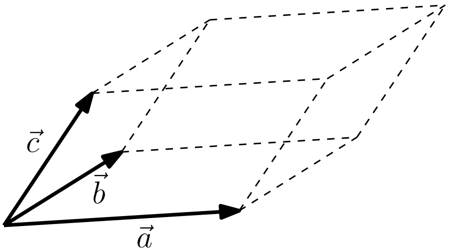
$$[\vec{a}, \vec{b}, \vec{c}] := \vec{a}^T(\vec{b} \times \vec{c}).$$

Eine einfache Rechnung zeigt, dass sich das Spatprodukt als gewöhnliche Determinante interpretieren lässt:

$$[\vec{a}, \vec{b}, \vec{c}] = \det\left(\begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix}\right).$$

Beim Spatprodukt handelt es sich also eigentlich um nichts Neues.

Wie im Abschnitt Determinanten erörtert lässt sich damit der Betrag des Spatprodukts als Volumen des von $\vec{a}, \vec{b}$ und $\vec{c}$ aufgespannten Parallelepipeds interpretieren.



Der Begriff **Spat** steht synonym für Parallelepiped; dies begründet die Namensgebung.

Satz 3.33

Seien $\vec{a}, \vec{b}, \vec{c} \in \mathbb{R}^3$*. Dann gelten:*

- *Sind* $\vec{a}, \vec{b}, \vec{c}$ *linear abhängig, so ist* $[\vec{a}, \vec{b}, \vec{c}] = 0$*.*
- *Sind* $\vec{a}, \vec{b}, \vec{c}$ *linear unabhängig, und bilden sie in dieser Reihenfolge ein Rechtssystem, so ist* $[\vec{a}, \vec{b}, \vec{c}] > 0$*.*
- *Sind* $\vec{a}, \vec{b}, \vec{c}$ *linear unabhängig, und bilden sie in dieser Reihenfolge ein Linkssystem, so ist* $[\vec{a}, \vec{b}, \vec{c}] < 0$*.*
- $[\vec{b}, \vec{a}, \vec{c}] = [\vec{c}, \vec{b}, \vec{a}] = [\vec{a}, \vec{c}, \vec{b}] = -[\vec{a}, \vec{b}, \vec{c}]$*.*
- $[\vec{a}, \vec{b}, \vec{c}] = [\vec{c}, \vec{a}, \vec{b}] = [\vec{b}, \vec{c}, \vec{a}]$*.*

Inhalt

❶ Grundlagen

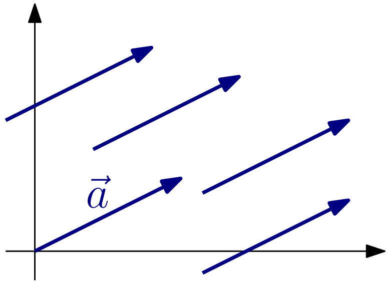
❷ Lineare Algebra

2.1 Vektorräume
2.2 Matrizen und lineare Abbildungen
2.3 Lineare Gleichungssysteme
2.4 Determinanten
2.5 Invertierbare Matrizen
2.6 Orthogonalität, Skalarprodukt und Norm
2.7 Kreuz- und Spatprodukt
2.8 Elemente der analytischen Geometrie
2.9 Orthogonale Abbildungen
2.10 Eigenwerte und Eigenvektoren
2.11 Singulärwertzerlegung

Vektoren aus $\mathbb{R}^2$ bzw. $\mathbb{R}^3$ lassen sich anschaulich auf verschiedene Weisen interpretieren:

- als Punkt der Anschauungsebene bzw. des Anschauungsraums,
- als Pfeil vom Koordinatenursprung zu eben diesem Punkt („Ortsvektor“), also einem Objekt mit Richtung und Länge.

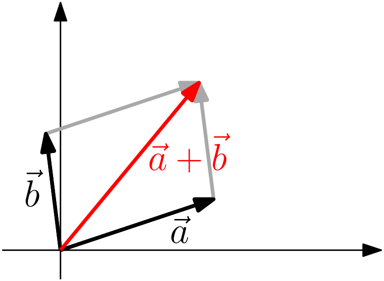
Manchmal ist es zudem anschaulicher, die so entstandenen Pfeile an bestimmte Stellen in der Ebene bzw. im Raum zu verschieben (z. B. wenn eine Kraft $\vec{F}$ an einem bestimmten Punkt „angreift“).

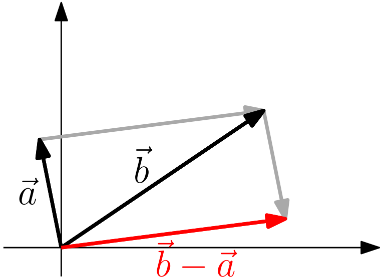


verschiedene Repräsentanten
des Vektors $\vec{a}$.

Vektoren des $\mathbb{R}^n$ lassen sich also mit sogenannten Pfeilklassen assoziieren. Jede dieser Pfeilklassen korrespondiert wiederum mit einer speziellen Parallelverschiebung.

Wie zweckmäßig solche Verschiebungen des Ortsvektors sind, wird an der geometrischen Interpretation der Addition und Subtraktion von Vektoren in deutlich:





In welchem Zusammenhang haben Sie diese Skizzen schon einmal gesehen? Was ist der Grund für diese Analogie?

In der analytischen Geometrie werden Punkte und Vektoren mitunter auch als verschiedene Objekte betrachtet.

Punkte werden dann meist durch Großbuchstaben gekennzeichnet (z. B. O für den Ursprung) und Vektoren häufig in der Form $\overrightarrow{AB}$ über ihre Anfangs- und Endpunkte. Es gilt dann:

- Jeder Punkt P korrespondiert mit seinem Ortsvektor $\vec{p} = \overrightarrow{OP}$.
- Sind $\vec{a} = \overrightarrow{OA}$ und $\vec{b} = \overrightarrow{OB}$ die Ortsvektoren zu zwei Punkten A und B, so gilt

$$\overrightarrow{AB} = \vec{b} - \vec{a}$$

Wir werden bei unserer bisherigen Notation bleiben und $\vec{p}$ statt $\overrightarrow{OP}$ bzw. $\vec{b} - \vec{a}$ statt $\overrightarrow{AB}$ schreiben.

In diesem Abschnitt betrachten wir nur noch Objekte im $\mathbb{R}^3$. Einige Ergebnisse haben ihre Entsprechungen im $\mathbb{R}^2$, die klar ersichtlich sind.

Konkret wollen wir uns mit den Lagebeziehungen von Punkt, Gerade und Ebene befassen.

Der Stoff gehört zum Standardrepertoire an Gymnasien und wird dort auch in der nötigen Tiefe behandelt. Daher sollen diese Inhalte nur überblicksartig dargestellt werden.

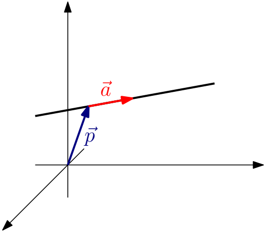
Eine Gerade, die durch den Punkt $\vec{p}$ und parallel zur Richtung $\vec{a} \neq \vec{0}$ verläuft, besitzt die Parameterdarstellung

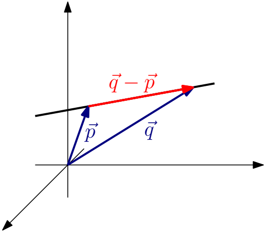
$$\vec{x} = \vec{p} + \lambda\,\vec{a}, \quad \lambda \in \mathbb{R}.$$

Diese Form wird auch Punkt-Richtungs-Form genannt.

Alternativ ist eine Gerade durch zwei verschiedene Punkte $\vec{p}$ und $\vec{q}$ eindeutig festgelegt. Man erhält die Parameterdarstellung über die Zwei-Punkte-Form

$$\vec{x} = \vec{p} + \lambda\,(\vec{q} - \vec{p}), \quad \lambda \in \mathbb{R}.$$





Analog ist eine Ebene durch einen Punkt $\vec{p}$ und zwei linear unabhängige Richtungsvektoren $\vec{a}$ und $\vec{b}$ eindeutig festgelegt. Sie besitzt also die Parameterform

$$\vec{x} = \vec{p} + \lambda\,\vec{a} + \mu\,\vec{b}, \qquad \lambda, \mu \in \mathbb{R}.$$

Diese Form wird auch Punkt-Richtungs-Form der Ebene genannt.

Alternativ kann man zur eindeutigen Festlegung drei Punkte $\vec{p}, \vec{q}$ und $\vec{r}$ verwenden. Man erhält die Parameterdarstellung über die Drei-Punkte-Form

$$\vec{x} = \vec{p} + \lambda\,(\vec{q} - \vec{p}) + \mu\,(\vec{r} - \vec{p}), \qquad \lambda, \mu \in \mathbb{R}.$$

Jeder Vektor $\vec{n} \neq 0$, der senkrecht auf der Ebene steht, heißt Normalenvektor der Ebene.

Auch durch Vorgabe eines Punktes $\vec{p}$ und eines Normelenvektors $\vec{n}$ ist eine Ebene eindeutig festgelegt. Es entsteht die Normalenform

$$\begin{aligned} &\vec{n}^T(\vec{x} - \vec{p}) = 0 \quad \text{bzw.} \\ &n_1 x_1 + n_2 x_2 + n_3 x_3 = c \ (\text{mit} \ \ c = \vec{n}^T \vec{p}). \end{aligned} \tag{3.10}$$

Ist $\vec{n}$ auf Länge Eins normiert, so ist der Abstand des Punktes $\vec{x} = (x_1, x_2, x_3)$ von der Ebene gegeben durch (Hessesche Normalform)

$$d = |n_1 x_1 + n_2 x_2 + n_3 x_3 - c|.$$

Wie kann man einen Normalenvektor zur Ebene $\vec{x} = \vec{p} + \lambda\,\vec{a} + \mu\,\vec{b}$ berechnen?

Interpretieren Sie Formel (3.10) im Kontext orthogonaler Projektionen auf den Normalenvektor.

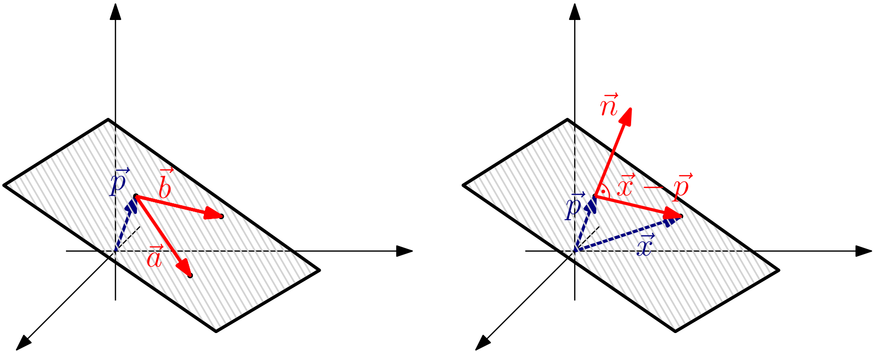
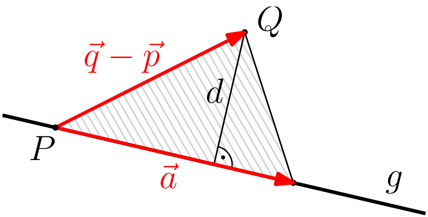


Illustration von Punkt-Richtungsform (links) und Normalenform (rechts).

Ob ein Punkt $\vec{q}$ auf einer Geraden $g : \vec{x} = \vec{p} + \lambda\,\vec{a}$ liegt oder nicht, kann man direkt mit der Parameterdarstellung prüfen. Alternativ berechnet man einfach den Abstand

$$d(\vec{q}, g) = \frac{\|\vec{a} \times (\vec{q} - \vec{p})\|}{\|\vec{a}\|}. \tag{3.11}$$

Bestätigen Sie die Formel. Berechnen Sie dazu den Flächeninhalt des von $\vec{a}$ und $\vec{q} - \vec{p}$ aufgespannten Dreiecks auf zwei verschiedene Weisen.



Für zwei Geraden

$$g_1 : \vec{x} = \vec{p} + \lambda\,\vec{a} \qquad \text{und} \qquad g_2 : \vec{x} = \vec{q} + \mu\,\vec{b}$$

im $\mathbb{R}^3$ liegt genau eine der drei folgenden Situationen vor:

- sie sind parallel (Spezialfall der Gleichheit einbezogen),
- sie schneiden sich in genau einem Punkt,
- sie sind windschief, d. h. sie sind weder parallel noch schneiden sie sich.

Parallelität lässt sich am leichtesten erkennen: In diesem Fall sind $\vec{a}$ und $\vec{b}$ linear abhängig und unterscheiden sich nur um einen skalaren Faktor.

Im Falle der Parallelität kann man Formel (3.11) zur Abstandsbestimmung nutzen (warum?) und erhält

$$d(g_1, g_2) = \frac{\|\vec{a} \times (\vec{q} - \vec{p})\|}{\|\vec{a}\|} = \frac{\|\vec{b} \times (\vec{p} - \vec{q})\|}{\|\vec{b}\|}.$$

Für die restlichen Fälle erzeugt man durch Gleichsetzen das lineare Gleichungssystem

$$\vec{p} + \lambda\vec{a} = \vec{q} + \mu\vec{b} \tag{3.12}$$

(drei Gleichungen für die zwei Unbekannten λ und μ).

Hat (3.12) genau eine Lösung (λ^*, μ^*), so schneiden sich die Geraden im Punkt $\vec{p} + \lambda^*\vec{a}$ (identisch mit $\vec{q} + \mu^*\vec{b}$). Der Schnittwinkel ist gegeben durch

$$\cos\phi = \frac{\vec{a}^T\vec{b}}{\|\vec{a}\|\|\vec{b}\|}$$

Hat (3.12) unendlich viele Lösungen, so sind die Geraden g_1 und g_2 identisch. (Das kann man aber im Rahmen des Tests auf Parallelität bereits entscheiden.)

Hat (3.12) keine Lösung, so sind die Geraden g_1 und g_2 windschief oder aber parallel und verschieden (s. o.).

Liegt der windschiefe Fall vor, erhält man den Abstand der Geraden mittels

$$d(g_1, g_2) = \frac{\left|[\vec{a}, \vec{b}, \vec{q} - \vec{p}]\right|}{\|\vec{a} \times \vec{b}\|}. \tag{3.13}$$

Formel (3.13) lässt sich in Analogie zu (3.11) bestätigen, indem man das Volumen eines geeigneten Körpers auf zwei verschiedene Weisen berechnet. Um welchen Körper handelt es sich?

Ob ein Punkt $\vec{q}$ zur Ebene

$$E: \vec{n}^T(\vec{x}-\vec{p})=0$$

gehört, lässt sich durch Einsetzen prüfen. (Gilt $\vec{n}^T(\vec{q}-\vec{p})=0$?)

Allgemein berechnet sich der Abstand von $\vec{q}$ zu E gemäß

$$d(\vec{q},E)=\frac{|\vec{n}^T(\vec{q}-\vec{p})|}{\|\vec{n}\|}. \tag{3.14}$$

Welcher Ansatz liegt Formel (3.14) zugrunde?

Eine Gerade und eine Ebene sind im $\mathbb{R}^3$ entweder

- parallel (beinhaltet den Fall, dass die Gerade in der Ebene liegt),
- oder sie schneiden sich in genau einem Punkt.

Die Gerade $g : \vec{x} = \vec{q} + \lambda\vec{a}$ und die Ebene $E : \vec{n}^T(\vec{x} - \vec{p}) = 0$ sind genau dann parallel, wenn $\vec{a}$ und $\vec{n}$ orthogonal sind, d. h. wenn

$$\vec{n}^T\vec{a} = 0.$$

In diesem Fall ist der Abstand Gerade-Ebene nach (3.14) gegeben durch

$$d = \frac{|\vec{n}^T(\vec{q} - \vec{p})|}{\|\vec{n}\|}.$$

Sind die Ebene und die Gerade nicht parallel ($\vec{n}^T\vec{a} \neq 0$), dann ist

$$\vec{q} + \frac{\vec{n}^T(\vec{p} - \vec{q})}{\vec{n}^T\vec{a}}\,\vec{a}$$

ihr Schnittpunkt. Der Schnittwinkel ϕ erfüllt

$$\sin(\phi) = \frac{|\vec{n}^T\vec{a}|}{\|\vec{n}\|\|\vec{a}\|}.$$

Leiten Sie beide Formeln her.

Zwei Ebenen im $\mathbb{R}^3$ sind

- entweder parallel (Spezialfall: identisch)
- oder schneiden sich entlang einer Geraden.

Zwei Ebenen

$$E_1 : \vec{n}_1^T(\vec{x} - \vec{p}_1) = 0 \qquad \text{und} \qquad E_2 : \vec{n}_2^T(\vec{x} - \vec{p}_2) = 0,$$

sind genau dann parallel, wenn $\vec{n}_1$ und $\vec{n}_2$ linear abhängig (also bis auf Vielfache gleich) sind. In diesem Fall ist der Abstand der Ebenen

$$d(E_1, E_2) = \frac{|\vec{n}_1^T(\vec{p}_2 - \vec{p}_1)|}{\|\vec{n}_1\|} = \frac{|\vec{n}_2^T(\vec{p}_1 - \vec{p}_2)|}{\|\vec{n}_2\|}.$$

Schneiden sich die Ebenen entlang einer Geraden $\vec{x} = \vec{p} + \lambda\vec{a}$ (äquivalent zu $\vec{n}_1 \times \vec{n}_2 \neq \vec{0}$), dann ist

$$\vec{a} = \vec{n}_1 \times \vec{n}_2,$$

und $\vec{p}$ ist (jede) Lösung des linearen Gleichungssystems

$$\begin{aligned} \vec{n}_1^T(\vec{p} - \vec{p}_1) &= 0, \\ \vec{n}_2^T(\vec{p} - \vec{p}_2) &= 0 \end{aligned}$$

(zwei Gleichungen für drei Unbekannte, nämlich die drei Komponenten von $\vec{p}$). Der Schnittwinkel ϕ erfüllt

$$\cos\phi = \frac{\vec{n}_1^T\vec{n}_2}{\|\vec{n}_1\|\|\vec{n}_2\|}.$$

Inhalt

1. Grundlagen

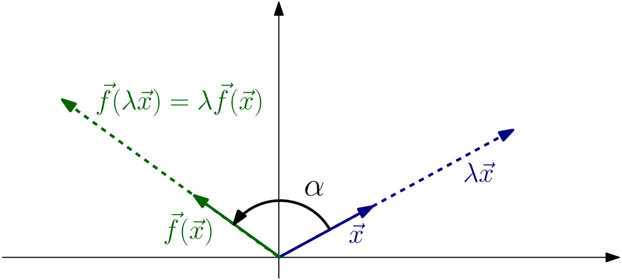
2. Lineare Algebra

2.1 Vektorräume
2.2 Matrizen und lineare Abbildungen
2.3 Lineare Gleichungssysteme
2.4 Determinanten
2.5 Invertierbare Matrizen
2.6 Orthogonalität, Skalarprodukt und Norm
2.7 Kreuz- und Spatprodukt
2.8 Elemente der analytischen Geometrie
2.9 Orthogonale Abbildungen
2.10 Eigenwerte und Eigenvektoren
2.11 Singulärwertzerlegung

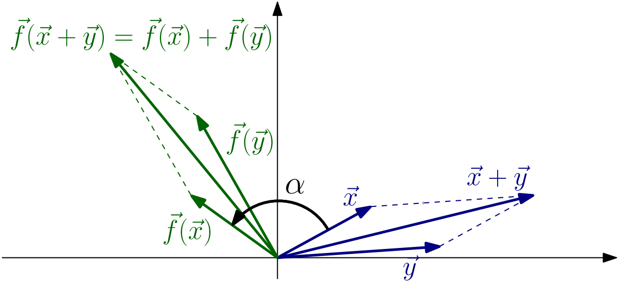
Orthogonale Abbildungen

- Eine große Klasse linearer Abbildungen sind die orthogonalen Abbildungen.
- Sie spielen insbesondere bei geometrischen Transformationen eine Rolle und zeichnen sich durch die Eigenschaft aus, Längen und Winkel unverändert zu lassen.
- So lassen sich zum Beispiel Drehungen und Spiegelungen mit orthogonalen Abbildungen mathematisch beschreiben.

Drehungen des $\mathbb{R}^2$ um einen Winkel α und Mittelpunkt in $\vec{0}$ sind zunächst lineare Abbildungen, wie an folgenden Skizzen deutlich wird:



$\vec{f}(\lambda\vec{x}) = \lambda\vec{f}(\vec{x})$
$\lambda\vec{x}$
α
$\vec{f}(\vec{x})$
$\vec{x}$



$\vec{f}(\vec{x}+\vec{y}) = \vec{f}(\vec{x}) + \vec{f}(\vec{y})$
$\vec{f}(\vec{y})$
α
$\vec{x}$
$\vec{x}+\vec{y}$
$\vec{f}(\vec{x})$
$\vec{y}$

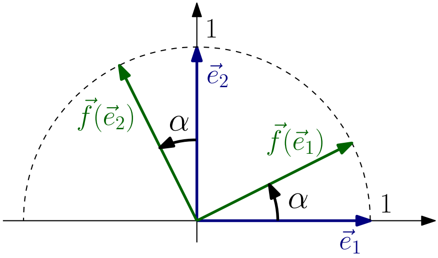
Zeichnen Sie eine analoge Skizze für Spiegelung des $\mathbb{R}^2$ an einer Geraden durch $\vec{0}$.

Natürlich lässt sich dieses geometrische Argument analog auf Drehungen und auf Spiegelungen an Ebenen im $\mathbb{R}^3$ anwenden.

Sowohl Drehungen um den Ursprung als auch Spiegelungen an einer Geraden (Ebene) durch den Ursprung können also als Matrix-Vektor-Multiplikationen beschrieben werden.

Wir begeben uns auf die Suche nach den Abbildungsmatrizen zu diesen Abbildungen.

Die Spalten der gesuchten Abbildungsmatrix sind gerade die Bilder der Einheitsvektoren unter der Drehung.



Es gilt

$$\vec{f}(\vec{e}_1) = \begin{bmatrix} \cos\alpha \\ \sin\alpha \end{bmatrix}$$

und

$$\vec{f}(\vec{e}_2) = \begin{bmatrix} -\sin\alpha \\ \cos\alpha \end{bmatrix}.$$

Damit ist die gesuchte Abbildungsmatrix

$$D_\alpha = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}.$$

Die Matrix D_α ist invertierbar, denn

$$\det D_\alpha = \cos^2\alpha + \sin^2\alpha = 1.$$

Ihre Inverse realisiert gerade die Drehung um $-\alpha$, d. h.

$$D_\alpha^{-1} = D_{-\alpha} = \begin{bmatrix} \cos(-\alpha) & -\sin(-\alpha) \\ \sin(-\alpha) & \cos(-\alpha) \end{bmatrix} = \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix}$$

Für D_α gilt also die bemerkenswerte Beziehung

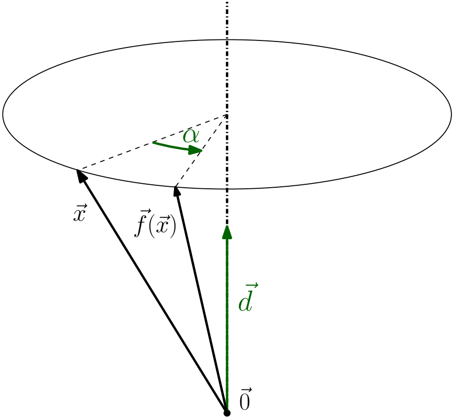
$$D_\alpha^{-1} = D_\alpha^T.$$

Berechnen Sie Drehmatrix für eine Drehung um den Ursprung mit $\alpha = 30°$. Geben Sie das Bild des Vektors $[2, 3]^T$ an.

Wiederholen Sie bei Bedarf die Begriffe Inverse, Invertierbarkeit und Determinante aus Kapiteln 7.4 und 7.5.

Drehungen im $\mathbb{R}^3$ werden durch eine Drehachse (die den Ursprung enthält) und einen Drehwinkel α festgelegt.

Die Drehachse wird dabei durch einen Vektor $\vec{d}$ festgelegt, der in die positive Achsenrichtung zeigt.



Orthogonale Abbildungen

Drehungen im Raum

Besonders einfach wird die Angabe der Drehmatrizen, wenn man die Einheitsvektoren (also die Koordinatenachsen) als Drehachsen verwendet:

$$D_{x,\alpha} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{bmatrix}, \quad D_{y,\alpha} = \begin{bmatrix} \cos\alpha & 0 & \sin\alpha \\ 0 & 1 & 0 \\ -\sin\alpha & 0 & \cos\alpha \end{bmatrix},$$

$$D_{z,\alpha} = \begin{bmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Jede beliebige Drehung im $\mathbb{R}^3$ lässt sich als Komposition von Drehungen um die Koordinatenachsen schreiben.

Physiker sprechen daher häufig von drei möglichen Freiheitsgraden der Rotation.

Machen Sie sich an einem der obigen Beispiele klar, dass die angegebene Matrix die gewünschte Transformation realisiert. Berechnen Sie dazu das Produkt $D\vec{x}$ für die gewählte Drehmatrix D.

Natürlich kann man auch im allgemeinen Fall die Drehmatrix angeben. Bei vorgegebenem Achsvektor $\vec{d}$ (mit $\|\vec{d}\| = 1$) und Winkel α lautet diese

$$\begin{bmatrix} \cos\alpha + (1-\cos\alpha)d_1^2 & (1-\cos\alpha)d_1 d_2 - d_3\sin\alpha & (1-\cos\alpha)d_1 d_3 + d_2\sin\alpha \\ (1-\cos\alpha)d_1 d_2 + d_3\sin\alpha & \cos\alpha + (1-\cos\alpha)d_2^2 & (1-\cos\alpha)d_2 d_3 - d_1\sin\alpha \\ (1-\cos\alpha)d_1 d_3 - d_2\sin\alpha & (1-\cos\alpha)d_2 d_3 + d_1\sin\alpha & \cos\alpha + (1-\cos\alpha)d_3^2 \end{bmatrix}$$

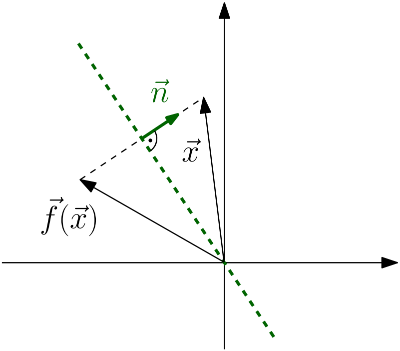
Eine explizite Abbildungsvorschrift ist gegeben durch

$$\vec{f}(\vec{x}) = \cos\alpha\,\vec{x} + (1-\cos\alpha)(\vec{x}^T\vec{d})\,\vec{d} + \sin\alpha(\vec{d}\times\vec{x}).$$

Dies alles schaut man bei Bedarf aber am besten in der Literatur nach.

Wir betrachten zunächst Spiegelungen im $\mathbb{R}^2$ an einer Geraden durch $\vec{0}$ senkrecht zum Vektor $\vec{n}$. Dabei sei $\vec{n}$ auf Länge 1 normiert ($\|\vec{n}\| = 1$).

Wir lesen das Spiegelbild $\vec{f}(x)$ von $\vec{x}$ aus folgender Skizze ab:



$$\vec{f}(x) = \vec{x} - 2(\vec{n}^T\vec{x})\vec{n}$$

Beachten Sie dabei, dass wegen der Normierung von $\vec{n}$ die Länge der Projektion von $\vec{x}$ auf $\vec{n}$ gerade $\vec{n}^T\vec{x}$ ist.

Es gilt also

$$\begin{aligned}\vec{f}(x) &= \vec{x} - 2(\vec{n}^T\vec{x})\vec{n} = \vec{x} - 2\vec{n}(\vec{n}^T\vec{x})\\ &= \vec{x} - 2(\vec{n}\vec{n}^T)\vec{x} = (I - 2\vec{n}\vec{n}^T)\vec{x},\end{aligned}$$

d. h. die Spiegelung wird durch Multiplikation mit der Matrix

$$S_{\vec{n}} = I - 2\vec{n}\vec{n}^T = \begin{bmatrix} 1-2n_1^2 & -2n_1n_2 \\ -2n_1n_2 & 1-2n_2^2 \end{bmatrix} = \begin{bmatrix} n_2^2-n_1^2 & -2n_1n_2 \\ -2n_1n_2 & n_1^2-n_2^2 \end{bmatrix}$$

beschrieben (beachte $n_1^2 + n_2^2 = 1$).

Geben Sie die Spiegelungsmatrix für die Spiegelung an der Geraden $x_2 = -x_1$ an. Multiplizieren Sie diese Matrix mit einem Vektor $\vec{x}$. Kann man das Ergebnis auch rein geometrisch begründen?

Bei Spiegelungen im $\mathbb{R}^3$ verwendet man statt der Spiegelachse eine Spiegelebene durch den Ursprung, welche ganz analog durch den Normalenvektor $\vec{n}$ festgelegt ist ($\|\vec{n}\| = 1$).

Die Spiegelungsmatrix besitzt jetzt drei Zeilen und Spalten, allerdings die gleiche Struktur:

$$S_{\vec{n}} = I - 2\vec{n}\vec{n}^T.$$

Zeichnen Sie eine geeignete Skizze, in welcher die Analogie sichtbar wird. Wie lautet die Matrix $S_{\vec{n}}$ in ausgeschriebener Form?

Drehungen und Spiegelungen gehören zur Klasse der orthogonalen linearen Abbildungen, die sich auch für höhere Dimensionen erklären lassen:

Definition 3.34

Eine Matrix $U \in \mathbb{R}^{n \times n}$ heißt orthogonal, wenn sie das Innenprodukt nicht verändert, d. h. wenn

$$(U\boldsymbol{y})^T(U\boldsymbol{x}) = \boldsymbol{y}^T\boldsymbol{x} \quad \text{für alle } \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n.$$

Insbesondere ist eine orthogonale Matrix

- längenerhaltend, d. h. $\|U\boldsymbol{x}\| = \|\boldsymbol{x}\|$ für alle $\boldsymbol{x}$ und
- winkeltreu , d. h. $\measuredangle(U\boldsymbol{x}, U\boldsymbol{y}) = \measuredangle(\boldsymbol{x}, \boldsymbol{y})$ für alle $\boldsymbol{x}, \boldsymbol{y}$.

Satz 3.35 (Charakterisierung orthogonaler Matrizen)

Sei $U \in \mathbb{R}^{n\times n}$. Die folgenden Aussagen sind äquivalent:

- *U ist orthogonal.*
- *Die Spalten (Zeilen) von U bilden eine Orthonormalbasis des $\mathbb{R}^n$.*
- *U ist invertierbar mit $U^{-1} = U^T$.*

Machen Sie sich klar, warum aus Punkt 1 Punkt 2 und daraus wiederum Punkt 3 folgt.

Bestätigen Sie mit Punkt 3, dass die 2D-Spiegelungsmatrix $S = I - 2\vec{n}\vec{n}^T$ orthogonal ist.

Die Entsprechung zu orthogonalen Matrizen im $\mathbb{C}^n$ sind unitäre Matrizen. Auch hier verwendet man das unveränderte Skalarprodukt zur Definition.

Dabei muss natürlich statt $\boldsymbol{y}^T\boldsymbol{x}$ immer das komplexe Skalarprodukt $\boldsymbol{y}^H\boldsymbol{x}$ verwendet werden.

Die Aussagen von Satz 3.35 gelten dann analog – die Beziehung im letzten Punkt lautet dabei

$$U^{-1} = U^H.$$

Unitäre Abbildungen werden Ihnen möglicherweise in der Quantenmechanik begegnen.

Inhalt

1. Grundlagen

2. Lineare Algebra

2.1 Vektorräume
2.2 Matrizen und lineare Abbildungen
2.3 Lineare Gleichungssysteme
2.4 Determinanten
2.5 Invertierbare Matrizen
2.6 Orthogonalität, Skalarprodukt und Norm
2.7 Kreuz- und Spatprodukt
2.8 Elemente der analytischen Geometrie
2.9 Orthogonale Abbildungen
2.10 Eigenwerte und Eigenvektoren
2.11 Singulärwertzerlegung

Wir betrachten ein Flugzeug, das auf einer holprigen Piste landet.

Modell des Flugzeugs:

- drei Massen: m_1 (Rumpf und Motor); m_2, m_3 (Flügel),
- drei Steifigkeiten (k_1, k_2, k_3) für die „federnde“ Verbindung der Teile.

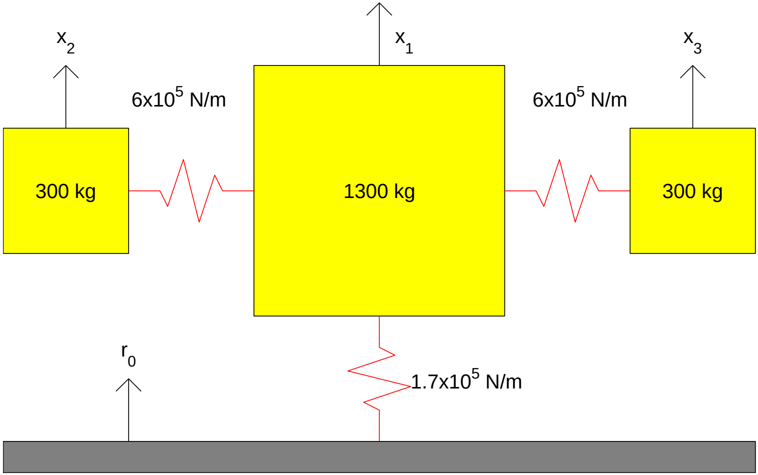
Die Piste wird durch eine Sinuskurve

$$r(t) = r_0 \sin(\omega_0 t)$$

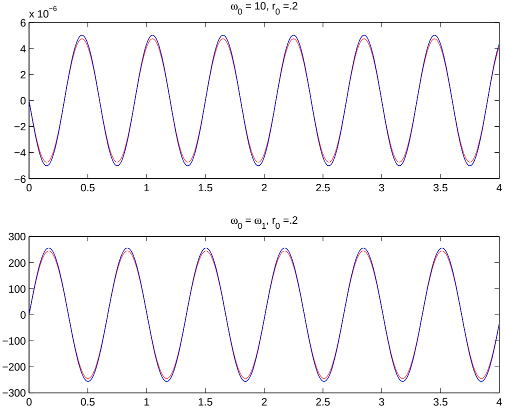
modelliert. Der Rumpf ist dann einer externen Kraft

$$f_1(t) = k_1 r_0 \sin(\omega_0 t)$$

ausgesetzt. Die Frequenz ω_0 hängt von der Landungsgeschwindigkeit v ab.


x2
x1
x3
6x10^5 N/m
6x10^5 N/m
300 kg
1300 kg
300 kg
r0
1.7x10^5 N/m

Die mathematische Analyse des Beispiels wird erst am Ende des Semesters gelingen. Wir zeigen hier aber schon die Lösungen $x_{1,2,3}(t)$ [in m] über t [in s] für $v = 120\,$km/h und $v = 108\,$km/h:



Beachten Sie, dass sich die Amplituden bei den verschiedenen Landungsgeschwindigkeiten um 7 Größenordnungen(!!!) unterscheiden.

Im zweiten Fall tritt ein Resonanzphänomen auf: Wenn die Anregungsfrequenz ω_0 (nahezu) mit einer der Eigenfrequenzen $\omega_{1,2,3}$ des Flugzeugs übereinstimmt, kommt es zu gefährlich großen Oszillationen.

Die Nichtbeachtung von Eigenfrequenzen und Resonanz kann z. B. bei Brücken oder Hochhäusern katastrophale Auswirkungen haben:

Tacoma Narrows Bridge (WA, 1940).
Bild: Prelinger Archives

Weitere Beispiele:

Broughton suspension bridge, Manchester 1831;
Millennium footbridge, London 2000.

Um z.B. Resonanzphänomene zu analysieren, benötigt man die folgenden Begriffe:

Definition 3.36

Sei $A \in \mathbb{R}^{n \times n}$ (oder $\mathbb{C}^{n \times n}$). Eine Zahl $\lambda \in \mathbb{C}$ heißt Eigenwert von A, wenn es einen Vektor $\boldsymbol{v} \in \mathbb{C}^n,\ \boldsymbol{v} \neq \boldsymbol{0}$, gibt, so dass

$$A\boldsymbol{v} = \lambda \boldsymbol{v}. \tag{3.15}$$

Jeder Vektor $\boldsymbol{v} \in \mathbb{C}^n \setminus \{\boldsymbol{0}\}$, der (3.15) erfüllt, heißt Eigenvektor von A zum Eigenwert λ.

Achtung: Auch wenn wir nur reelle Matrizen betrachten: bei Eigenwerten und Eigenvektoren lässt man immer auch komplexe Zahlen zu!

Eigenwerte und Eigenvektoren
Definition

Die Richtung eines Eigenvektors wird durch die lineare Abbildung A nicht verändert – der Eigenvektor wird lediglich gestreckt.

Das rechte Bild entsteht z. B. aus dem linken durch eine „Scherung“ der Leinwand.

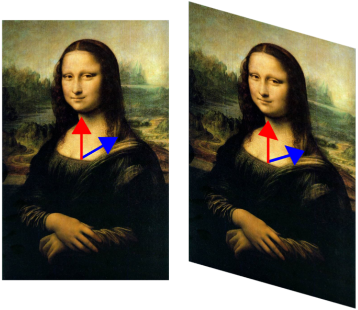


Bild: Wikimedia Commons

Der rote Vektor bleibt unverändert und ist damit ein Eigenvektor zum Eigenwert 1. Der blaue Vektor ändert hingegen seine Richtung und ist daher kein Eigenvektor der Scherungsabbildung.

Eigenwerte und Eigenvektoren

Welche reellen Eigenwerte besitzt eine Drehung im Raum um die x_1-Achse, und welcher reelle Eigenvektor kommt in Frage?

Wie verhält es sich mit einer Spiegelung in der Ebene an einer Geraden durch 0 und senkrecht zu $\vec{n}$?

Argumentieren Sie rein geometrisch!

Bestätigen Sie, dass $\lambda = 1$ Eigenwert von $A = \left[\begin{smallmatrix} 5 & -8 \\ -1 & 3 \end{smallmatrix}\right]$ mit zugehörigem Eigenvektor $\boldsymbol{v} = [2, 1]^T$ ist.

Zeigen Sie, dass jedes komplexe Vielfache des Vektors $[1, i]^T$ ein Eigenvektor von $B = \left[\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right]$ zum Eigenwert $\lambda = i$ ist.

Formulieren Sie eine allgemeingültige Aussage und bestätigen Sie diese durch Einsetzen in (3.15).

Anmerkung: Bislang wissen wir weder, wie man EW und EV berechnet, noch ob in den Beispielen alle EW und EV erfasst wurden.

Wir benutzen zur Herleitung der Formel für die Eigenwertberechnung folgende Äquivalenzkette:

$\lambda \in \mathbb{C}$ ist Eigenwert von $A \in \mathbb{R}^{n\times n}$.

- $\Leftrightarrow$ Es gibt einen Vektor $\boldsymbol{v} \neq \mathbf{0}$ mit $A\boldsymbol{v} = \lambda \boldsymbol{v}$.
- $\Leftrightarrow$ Ein Vektor $\boldsymbol{v} \neq \mathbf{0}$ löst das homogene LGS $(A - \lambda I)\boldsymbol{v} = \mathbf{0}$.
- $\Leftrightarrow$ $\det(A - \lambda I) = 0$.

Die Eigenwerte sind also gerade die Nullstellen der Funktion

$$c_A(\lambda) := \det(A - \lambda I).$$

Diese Funktion ist ein Polynom vom Grad n und wird das charakteristische Polynom von A genannt.

Satz 3.37

Für $A \in \mathbb{R}^{n \times n}$ ist das charakteristische Polynom $c_A(\lambda) = \det(A - \lambda I)$ ein Polynom vom exakten Grad n mit reellen Koeffizienten und Höchstkoeffizienten 1 oder -1.

Die Eigenwerte von A sind genau die Nullstellen von c_A, also die Lösungen der Gleichung

$$c_A(\lambda) = \det(A - \lambda I) \stackrel{!}{=} 0. \tag{3.16}$$

Berechnen Sie sämtliche Eigenwerte der Matrizen

$$A = \begin{bmatrix} 5 & -8 \\ -1 & 3 \end{bmatrix} \quad \text{und} \quad B = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

(vgl. 2. und 3. Kasten auf S. 257).

Sei λ Eigenwert der Matrix A. Die zugehörigen Eigenvektoren von A sind die (nicht-trivialen) Lösungen des homogenen Gleichungssystems

$$(A - \lambda I)\boldsymbol{v} = \boldsymbol{0}. \tag{3.17}$$

Zusammen mit $\boldsymbol{0}$ bilden sie einen Unterraum des $\mathbb{C}^n$, den sogenannten Eigenraum von A zum Eigenwert λ,

$$\begin{aligned} \text{Eig}(A, \lambda) &:= \{\boldsymbol{x} \in \mathbb{C}^n \,:\, A\boldsymbol{x} = \lambda\boldsymbol{x}\} \\ &= \{\boldsymbol{x} \in \mathbb{C}^n \,:\, (A - \lambda I)\boldsymbol{x} = \boldsymbol{0}\} \\ &= \mathscr{N}(A - \lambda I) \end{aligned}$$

Berechnen Sie sämtliche Eigenvektoren zu den Eigenwerten der Matrizen A und B von Seite 259.

Definition 3.38

Zwei Matrizen $A, B \in \mathbb{C}^{n\times n}$ heißen ähnlich, wenn es eine invertierbare Matrix $V \in \mathbb{C}^{n\times n}$ gibt, so dass

$$A = V^{-1}BV. \tag{3.18}$$

Bemerkungen:

1. Gleichung (3.18) ist äquivalent zu

$$VA = BV \quad \text{und} \quad B = VAV^{-1}.$$

2. Ähnlichkeit ist eine Äquivalenzrelation.

Ähnliche Matrizen haben das gleiche charakteristische Polynom, wie folgende Rechnung zeigt:

$$\begin{aligned} c_{V^{-1}BV}(\lambda) &= \det(V^{-1}BV - \lambda I) \\ &= \det(V^{-1}(B - \lambda I)V) \\ &= (\det V)^{-1} \det(B - \lambda I) \det V \\ &= c_B(\lambda). \end{aligned}$$

Wir fassen zusammen:

Satz 3.39

Ähnliche Matrizen haben das gleiche charakteristische Polynom und damit die gleichen Eigenwerte.

Die Eigenvektoren zu einem Eigenwert λ sind dabei allerdings verschieden: $\vec{x}$ ist genau dann Eigenvektor von $A = V^{-1}BV$, wenn $V\boldsymbol{x}$ Eigenvektor von B ist.

Um den Ähnlichkeitsbegriff vollständig zu verstehen, muss man sich mit Basistransformationen auseinandersetzen. Wir gehen dabei von der Darstellung $B = VAV^{-1}$ von S. 261 aus.

Zu einem gegebenen Vektor $\boldsymbol{x}$ lässt sich $V\boldsymbol{x}$ als Linearkombination der Spalten $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$ von V schreiben:

$$V\boldsymbol{x} = x_1\boldsymbol{v}_1 + x_2\boldsymbol{v}_2 + \ldots + x_n\boldsymbol{v}_n.$$

Fasst man $\vec{x}$ als Koordinatenvektor bezüglich der Basis $\mathscr{B}_V = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ auf, so ist $V\boldsymbol{x}$ gerade der zugehörige Koordinatenvektor bezüglich der Standardbasis.

Man sagt daher auch, V stellt eine Basistransformation von der Basis $\mathscr{B}_V$ in die Standardbasis dar.

Die Abbildung V^{-1} macht die Transformation rückgängig und stellt somit eine Basistransformation von der Standardbasis nach $\mathscr{B}_V$ dar.

Verifizieren Sie beide Aussagen am Beispiel der Basis $\mathscr{B} = \{\frac{1}{\sqrt{2}}\left[\begin{smallmatrix}1\\1\end{smallmatrix}\right], \frac{1}{\sqrt{2}}\left[\begin{smallmatrix}1\\-1\end{smallmatrix}\right]\}$ und des Vektors $\begin{bmatrix}\sqrt{2}\\0\end{bmatrix} = \left[\begin{smallmatrix}1\\1\end{smallmatrix}\right]_{\mathscr{B}}$.

In $B\boldsymbol{x} = VAV^{-1}\boldsymbol{x}$ kann man die rechte Seite nun von rechts nach links wie folgt lesen:

- Stelle $\boldsymbol{x}$ als Koordinatenvektor bezüglich der Basis $\mathscr{B}_V = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ dar (d. h. Multiplikation mit V^{-1}).
- Führe die durch B beschriebene lineare Abbildung in den Koordinaten bezüglich $\mathscr{B}_V$ aus (d. h. Multiplikation mit A).
- Transformiere das Ergebnis wieder zurück in die Standardkoordinaten (Multiplikation mit V).

Wir haben somit erkannt:

Satz 3.40

Ähnliche Matrizen stellen die gleiche lineare Abbildung bezüglich verschiedener Basen dar.

Folglich besitzen ähnliche Matrizen

- dieselbe Determinante,
- denselben Rang und
- denselben Defekt.

Auch die Aussage von Satz 3.39 wird mit dieser Erkenntnis noch ein Stück verständlicher.

Definition 3.41

Eine Matrix $A \in \mathbb{C}^{n\times n}$ heißt diagonalisierbar, wenn A ähnlich zu einer Diagonalmatrix $D \in \mathbb{C}^{n\times n}$ ist, d. h. wenn es eine invertierbare Matrix $V \in \mathbb{C}^{n\times n}$ gibt mit

$$D = V^{-1}AV.$$

Satz 3.42

Eine Matrix $A \in \mathbb{C}^{n\times n}$ ist genau dann diagonalisierbar, wenn es n linear unabhängige Eigenvektoren von A – also eine Basis aus Eigenvektoren von A – gibt.

Sei $A \in \mathbb{C}^{n \times n}$ eine Matrix, für eine Basis des $\mathbb{C}^n$ aus Eigenvektoren von A existiert (vgl. Satz 3.42).

Dann kann die Abbildung A bezüglich dieser Basis nur durch Streckungen der Basisvektoren ausgedrückt werden (Multiplikation mit einer Diagonalmatrix).

Beispiel Die Matrix $A = \left[\begin{smallmatrix} 5 & -8 \\ -1 & 3 \end{smallmatrix}\right]$ besitzt die Eigenwerte $\lambda_1 = 1$ und $\lambda_2 = 7$ mit den Eigenvektoren $\boldsymbol{v}_1 = [2, 1]^T$ und $\boldsymbol{v}_2 = [-4, 1]^T$ (vgl. S. 259). Es gilt die Darstellung

$$A = VDV^{-1} = \begin{bmatrix} 2 & -4 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 7 \end{bmatrix} \begin{bmatrix} 2 & -4 \\ 1 & 1 \end{bmatrix}^{-1}.$$

In der Diagonalmatrix D stehen die Eigenwerte und in den Spalten von V die Eigenvektoren.

Wir widmen uns nun der Frage, wie groß die Dimension des von den Eigenvektoren einer Matrix $A \in \mathbb{C}^{n \times n}$ aufgespannten Unterraums ist.

Insbesondere wollen wir wissen, wann es eine Basis des $\mathbb{C}^n$ aus Eigenvektoren von A gibt.

Wir beginnen mit folgendem Satz:

Satz 3.43

Gehören die Eigenvektoren $\boldsymbol{v}_1, \dots, \boldsymbol{v}_r$ zu ***verschiedenen*** *Eigenwerten $\lambda_1, \dots, \lambda_r$ der $n \times n$-Matrix A, dann sind sie linear unabhängig.*

Machen Sie sich dies zumindest für den Fall $r = 2$ klar.

Wenn es n verschiedene Eigenwerte zur Matrix $A \in \mathbb{C}^{n \times n}$ gibt, existiert also eine Basis des $\mathbb{C}^n$, die nur aus Eigenvektoren von A besteht.

Definition 3.44

Die algebraische Vielfachheit $\nu_{\text{alg}}(\lambda)$ eines Eigenwerts λ von A ist seine Vielfachheit als Nullstelle des charakteristischen Polynoms c_A von A.

Die geometrische Vielfachheit $\nu_{\text{geom}}(\lambda)$ eines Eigenwerts λ von A ist die Dimension des zugehörigen Eigenraums $\dim(\text{Eig}(A, \lambda))$.

Anmerkung zur Berechnung Um die algebraischen Vielfachheiten zu bestimmen, muss man nur das charakteristische Polynom c_A kennen. Um die geometrische Vielfachheiten zu berechnen, reicht die Kenntnis von c_A allein nicht aus.

Nach dem Fundamentalsatz der Algebra summieren sich die algebraischen Vielfachheiten der Eigenwerte einer Matrix $A \in \mathbb{C}^{n \times n}$ zu n.

Die Summe der geometrischen Vielfachheiten kann dagegen kleiner sein. Es gilt folgender Satz:

Satz 3.45

Sei λ Eigenwert der Matrix $A \in \mathbb{C}^{n \times n}$ mit algebraischer Vielfachheit $\nu_{\mathsf{alg}}(\lambda)$ und geometrischer Vielfachheit $\nu_{\mathsf{geom}}(\lambda)$. Dann gilt:

$$1 \le \nu_{\mathsf{geom}}(\lambda) \le \nu_{\mathsf{alg}}(\lambda) \le n.$$

Machen Sie sich am Beispiel der Matrix $\left[\begin{smallmatrix} 2 & 1 \\ 0 & 2 \end{smallmatrix}\right]$ klar, dass die geometrische Vielfachheit eines Eigenwerts tatsächlich kleiner sein kann als die algebraische.

Beim Beispiel im Kasten auf S. 270 handelt es sich um einen sogenannten Jordan-Block, d. h. eine Matrix der Form

$$\begin{bmatrix} \lambda & 1 & & \\ & \lambda & 1 & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix} \in \mathbb{R}^{n\times n}.$$

Dieser besitzt nur den Eigenwert λ mit geometrischer Vielfachheit 1 und algebraischer Vielfachheit n.

Bestätigen Sie diese Aussagen. Für die geometrische Vielfachheit nutzen Sie am besten die Beziehung

$$\nu_{\mathsf{geom}}(\lambda) = \dim(\mathsf{Eig}(A,\lambda)) = n - \mathsf{rang}(A - \lambda I).$$

Diagonalisierbarkeit lässt sich nun mittels Vielfachheiten auch folgendermaßen charakterisieren:

Satz 3.46

Eine Matrix $A \in \mathbb{C}^{n\times n}$ ist genau dann diagonalisierbar, wenn die algebraische und geometrische Vielfachheit für jeden Eigenwert übereinstimmen.

Sind $\lambda_1, \dots, \lambda_n$ die Eigenwerte und $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$ zugeordnete Eigenvektoren, die eine Basis des $\mathbb{C}^n$ bilden, so gilt die Darstellung

$$A = VDV^{-1}.$$

Dabei gilt $D = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$, und $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$ sind in dieser Reihenfolge die Spalten von V.

Die Matrix $A = \left[\begin{smallmatrix} 5 & -8 \\ -1 & 3 \end{smallmatrix}\right]$ besitzt die Eigenwerte $\lambda_1 = 1$ und $\lambda_2 = 7$ mit den Eigenvektoren $\boldsymbol{v}_1 = [2, 1]^T$ und $\boldsymbol{v}_2 = [-4, 1]^T$ (vgl. S. 259 f.). Sie bilden eine Basis des $\mathbb{C}^2$ (und in diesem Falle auch eine Basis des $\mathbb{R}^2$).

Das charakteristische Polynom einer reellen 2×2-Matrix A besitzt die Struktur

$$c_A(\lambda) = \lambda^2 + p\lambda + q \qquad (\text{mit } p, q \in \mathbb{R}).$$

In Kombination mit Satz 3.45 ergeben sich verschiedene Möglichkeiten für die Eigenwerte. Es existieren

- entweder zwei verschiedene reelle Eigenwerte λ_1 und λ_2, die geometrische und algebraische Vielfachheit 1 haben. Zu jedem Eigenwert gibt es reelle Eigenvektoren.
- oder zwei verschiedene konjugiert-komplexe Eigenwerte λ_1 und λ_2, die geometrische und algebraische Vielfachheit 1 haben. Zu jedem Eigenwert gibt es komplexe Eigenvektoren.

(b. w.)

- oder nur einen Eigenwert λ, der dann zwangsläufig reell ist.
  Er besitzt die algebraische Vielfachheit 2 und
  - entweder die geometrische Vielfachheit 2, d. h. jeder Vektor aus $\mathbb{C}^2$ ist Eigenvektor. A besitzt dann die Form
    $$A = \lambda I = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}$$
  - oder die geometrische Vielfachheit 1.

Ordnen Sie die Beispiele von S. 259 und S. 270 den entsprechenden Fällen zu.

Führen Sie eine ähnliche Analyse für den 3×3-Fall durch. (Auf die Unterscheidung zwischen reellen und komplexen Eigenwerten können Sie dabei verzichten.)

Gilt $\nu_{\text{geom}}(\lambda) < \nu_{\text{alg}}(\lambda)$ für ein Eigenwert λ einer Matrix $A \in \mathbb{C}^{n\times n}$, so ist A nicht diagonalisierbar. Folgende „fast diagonale“ Form existiert aber im allgemeinen Fall:

Satz 3.47 (Jordansche Normalform)

Jede Matrix $A \in \mathbb{C}^{n \times n}$ ist ähnlich zu einer Blockdiagonalmatrix $J = \operatorname{diag}(J_1, \ldots, J_s)$ mit Diagonalblöcken der Form

$$J_k = \begin{bmatrix} \lambda_k & 1 & & \\ & \lambda_k & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_k \end{bmatrix}, \qquad \textit{d.h.} \qquad A = VJV^{-1} \tag{3.19}$$

mit Eigenwerten λ_k von A. Zu jedem Eigenwert von A existieren ein oder mehrere solcher Jordan-Blöcke. Ist n_j die Dimension des größten zum Eigenwert λ_j von A gehörenden Jordan-Blocks, so heißt $m_A(\lambda) = \prod_j (\lambda - \lambda_j)^{n_j}$ das Minimalpolynom von A, welches das charakteristische Polynom von A teilt. Für die Spaltenvektoren $\boldsymbol{v}$ von V im Indexbereich eines Jordanblocks zum Eigenwert λ_j gilt

$$(A - \lambda_j I)^\ell \boldsymbol{v} = \mathbf{0} \qquad \textit{für ein } \ell \textit{ mit } 1 \le \ell \le n_j.$$

Bemerkungen:

(1) Die Spalten der Ähnlichkeitstransformation V in (3.19) heißen Haupvektoren. Zu jeden Eigenwert λ von A existieren $\nu_{\mathsf{alg}}(\lambda)$ linear unabhängige Hauptvektoren. Da Hauptvektoren zu verschiedenen Eigenwerten linear unabhängig sind existiert somit eine Basis des $\mathbb{C}^n$ aus Hauptvektoren von A.

(2) Im Fall $\nu_{\mathsf{geom}}(\lambda) = \nu_{\mathsf{alg}}(\lambda)$ sind alle Hauptvektoren von A zum Eigenwert λ auch Eigenvektoren.

(3) Es gilt $m_A(A) = O$ und m_A ist das Polynom niedrigsten Grades mit dieser Eigenschaft. Somit gilt auch $c_A(A) = O$, was auch als Satz von Cayley-Hamilton bekannt ist.

Satz 3.48

Sei $A \in \mathbb{R}^{n\times n}$. Dann gelten:

- *A und A^T besitzen dasselbe charakteristische Polynom, also dieselben Eigenwerte (mit i. A. verschiedenen Eigenräumen).*
- *Besitzt A den Eigenvektor $\boldsymbol{x}$ zum Eigenwert λ, dann besitzen*

  $$\alpha A,\ A^m,\ A+\beta I_n,\ p(A) = \alpha_m A^m + \cdots + \alpha_1 A + \alpha_0 I_n$$

  denselben Eigenvektor $\boldsymbol{x}$, allerdings zum Eigenwert

  $$\alpha\lambda,\ \lambda^m,\ \lambda+\beta,\ p(\lambda) = \alpha_m \lambda^m + \cdots + \alpha_1\lambda + \alpha_0.$$
- *A ist genau dann invertierbar, wenn alle Eigenwerte von A von 0 verschieden sind. Ist dann λ ein Eigenwert von A mit Eigenvektor $\boldsymbol{x}$, so ist λ^{-1} ein Eigenwert von A^{-1} mit demselben Eigenvektor $\boldsymbol{x}$.*

Verifizieren Sie einige dieser Aussagen.

Weiterhin kann man Eigenwerte sehr einfach bestimmen, wenn A bestimmte strukturelle Eigenschaften besitzt:

Satz 3.49

Ist A eine (untere oder obere) Dreiecksmatrix, so sind die Hauptdiagonaleinträge von A genau die Eigenwerte von A.

Dies trifft insbesondere dann zu, wenn A eine Diagonalmatrix ist. In diesem Fall sind die Einheitsvektoren zugehörige Eigenvektoren.

Erinnerung/Bemerkung: Eine Diagonalmatrix enthält nur auf der Hauptdiagonalen Einträge ungleich Null. Sind $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ die Einträge entlang der Diagonalen, so schreibt man $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$.

Bei Dreiecksmatrizen sind alle Elemente ober- bzw. unterhalb der Hauptdiagonalen gleich Null.

Eine Matrix $A \in \mathbb{R}^{n \times n}$ heißt symmetrisch, wenn $A = A^T$ gilt, d. h. die Einträge symmetrisch zur Hauptdiagonalen liegen.

Symmetrische Matrizen haben bemerkenswerte Eigenschaften:

Satz 3.50

Ist $A \in \mathbb{R}^{n \times n}$ symmetrisch, so besitzt A nur reelle Eigenwerte. Es gibt eine Orthonormalbasis des $\mathbb{R}^n$, die aus Eigenvektoren von A besteht.

Es gibt also eine Diagonalmatrix $D \in \mathbb{R}^{n \times n}$ und eine orthogonale Matrix $U \in \mathbb{R}^{n \times n}$ (d. h. $U^T = U^{-1}$), so dass

$$A = UDU^T.$$

Insbesondere ist A diagonalisierbar.

Für orthogonale Matrizen gibt es ein ähnliches Ergebnis. Beachten Sie aber, dass die Eigenwerte hier i. A. komplex sind!

Satz 3.51

Ist $A \in \mathbb{R}^{n\times n}$ orthogonal, so besitzt A nur Eigenwerte mit Betrag 1. Es gibt eine Orthonormalbasis des $\mathbb{C}^n$, die aus Eigenvektoren von A besteht.

Es gibt also eine Diagonalmatrix $D \in \mathbb{C}^{n\times n}$ und eine unitäre Matrix $U \in \mathbb{C}^{n\times n}$ (d. h. $U^H = U^{-1}$), so dass

$$A = UDU^H.$$

Insbesondere ist A diagonalisierbar.

Bemerkung: Die Klasse der Matrizen, welche eine Basis aus orthonormalen Eigenvektoren besitzen, ist größer als die der symmetrischen (Hermiteschen) Matrizen. Genau trifft dies zu für normale Matrizen. Diese sind charakterisiert durch die Eigenschaft

$$AA^T = A^T A, \quad \text{bzw.} \quad AA^H = A^H A.$$

Bestimmen Sie die Eigenwerte und Eigenvektoren der Matrizen

$$A = \begin{bmatrix} 8 & 0 & 0 \\ 0 & 1 & -3 \\ 0 & -3 & 1 \end{bmatrix} \quad \text{und} \quad D_\alpha = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}$$

Geben Sie in beiden Fällen eine Darstellung der Form VDV^{-1} an. Erkennen Sie die Matrix B von S. 259 f. wieder?

Die durch die Gleichungen

$$\frac{x^2}{4} + \frac{y^2}{9} = 1 \qquad \text{bzw.} \qquad y = 4(x+1)^2 + 2$$

festgelegten Punktmengen erkennt man leicht als Ellipse symmetrisch zu den Koordinatenachsen mit Halbachsen 2 und 3 bzw. nach oben offene Parabel mit Scheitelpunkt $(x_S, y_S) = (-1, 2)$.

Dies ist weniger offensichtlich bei

$$\frac{x^2}{9} + 4xy - \frac{y^2}{4} + x - 2y = 4.$$

Die Hauptachsentransformation bietet eine Variablensubstitution, mit Hilfe derer eine Klassifikation quadratischer Funktionen einfach möglich ist.

Jede quadratische Funktion $\phi : \mathbb{R}^n \to \mathbb{R}$ besitzt die allgemeine Form

$$\phi(\boldsymbol{x}) = \boldsymbol{x}^T A \boldsymbol{x} + \boldsymbol{b}^T \boldsymbol{x} + c \tag{3.20}$$

mit einer symmetrischen Matrix $A \in \mathbb{R}^{n \times n}$, einem Vektor $\boldsymbol{b}$ und einer Zahl $c \in \mathbb{R}$.

Beispiel: $n = 2$

$$\phi(x_1, x_2) = a_{1,1} x_1^2 + a_{1,2} x_1 x_2 + a_{2,1} x_2 x_1 + a_{2,2} x_2^2 + b_1 x_1 + b_2 x_2 + c$$

Die Aufgabe der Hauptachsentransformation besteht darin, durch einen geeigneten Basiswechsel im $\mathbb{R}^n$ die quadratische Funktion in eine einfache Form zu bringen, an der man leicht ihren Typ ablesen kann.

Dies ist auch hilfreich bei der Klassifikation von Quadriken (auch Kegelschnitte genannt), also den Punktmengen

$$Q = \{\boldsymbol{x} \in \mathbb{R}^n : \phi(\boldsymbol{x}) = 0\}.$$

Da die Matrix A ohne Beschränkung der Allgemeinheit als symmetrisch angenommen werden kann besitzt diese eine Basis aus orthonormalen Eigenvektoren, d.h. eine orthogonale Matrix Q mit

$$Q^T AQ = D = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$$

mit reellen Eigenwerten $\lambda_1, \dots, \lambda_n$. Der Hauptteil $x^T Ax$ einer quadratischen Funktion (3.20) heißt auch quadratische Form. Die orthogonalen Eigenrichtungen der Matrix A heißen Hauptachsen der quadratischen Funktion.

Für die neue Variable $\boldsymbol{y} := Q^T \boldsymbol{x}$ gilt dann $\boldsymbol{x} = Q\boldsymbol{y}$ und, mit dem neuen Vektor $\boldsymbol{d} := Q^T \boldsymbol{b}$ erhalten wir

$$\begin{aligned}
\phi(\boldsymbol{x}) &= \boldsymbol{x}^T A\boldsymbol{x} + \boldsymbol{b}^T \boldsymbol{x} + c = (Q\boldsymbol{y})^T A(Q\boldsymbol{y}) + \boldsymbol{b}^T (Q\boldsymbol{y}) + c \\
&= \boldsymbol{y}^T Q^T AQ\boldsymbol{y} + \boldsymbol{b}^T Q\boldsymbol{y} + c = \boldsymbol{y}^T D\boldsymbol{y} + \boldsymbol{d}^T \boldsymbol{y} + c \\
&= \sum_{j=1}^{n} \lambda_j y_j^2 + \sum_{j=1}^{n} d_j y_j + c.
\end{aligned}$$

Als letzter Schritt folgt nun quadratisches Ergänzen in jeder Variablen y_j:

$$\lambda_j y_j^2 + d_j y_j = \lambda_j \left(y_j + \frac{d_j}{2\lambda_j} \right)^2 - \frac{d_j^2}{4\lambda_j}.$$

Mit einer weiteren Substitution $z_j = y_j + \frac{d_j}{2\lambda_j}$, $j = 1, \dots, n$ erhalten wir

$$\phi(\boldsymbol{x}) = \sum_{j=1}^{n} \lambda_j z_j^2 + d = \lambda_1 z_1^2 + \cdots + \lambda_n z_n^2 + d.$$

mit

$$d := c - \sum_{j=1}^{n} \frac{d_j^2}{4\lambda_j}.$$

Führen Sie die Hauptachsentransformation durch für die Quadrik

$$2x^2 - y^2 + 4xy - 2x + y - 6 = 0.$$

Alle Eigenwerte von A ungleich Null	
$\frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 = 0$	Ellipse mit Halbachsen a, b
$\frac{x^2}{a^2} + \frac{y^2}{b^2} + 1 = 0$	leere Menge
$\frac{x^2}{a^2} - \frac{y^2}{b^2} - 1 = 0$	Hyperbel
$x^2 + a^2y^2 = 0$	Punkt $\{(0,0)\}$
$x^2 - a^2y^2 = 0$	Geradenpaar $y = \pm\lvert a\rvert x$
Ein Eigenwert von A gleich Null	
$x^2 - 2py = 0$	Parabel
$x^2 - a^2 = 0$	paralleles Geradenpaar
$x^2 + a^2 = 0$	leere Menge
$x^2 = 0$	Gerade $x = 0$ (y-Achse)

Inhalt

1. Grundlagen

2. Lineare Algebra

2.1 Vektorräume
2.2 Matrizen und lineare Abbildungen
2.3 Lineare Gleichungssysteme
2.4 Determinanten
2.5 Invertierbare Matrizen
2.6 Orthogonalität, Skalarprodukt und Norm
2.7 Kreuz- und Spatprodukt
2.8 Elemente der analytischen Geometrie
2.9 Orthogonale Abbildungen
2.10 Eigenwerte und Eigenvektoren
2.11 Singulärwertzerlegung

Singulärwertzerlegung

Auch wenn sie zum Verständnis wichtig sind weisen die bisher betrachteten Zerlegungen bzw. Normalformen für Matrizen einige Schönheitsfehler auf:

- Nicht alle Matrizen sind diagonalisierbar, und die Jordansche Normalform lässt sich bei Rundungsfehlern nicht stabil berechnen.
- Auch bei diagonalisierbaren Matrizen können Eigenvektoren zu verschiedenen Eigenwerten nahezu linear abhängig sein, sodass auch hier sich Rundungsfehler stark bemerkbar machen können.
- In vielen Anwendungen beschreiben Matrizen Daten oder Signale und sind nicht notwengig quadratisch, d.h. es sind auch Zerlegungen für rechteckige Matrizen nötig.

Die Singulärwertzerlegung stellt die für die Praxis wichtigste Matrixzerlegung dar, mit Anwendungen u.a. in der Datenanalyse, Maschinenlernen, Computer-Vision, digitalen Signalverarbeitung, insbesondere Signal-Rauschentrennung.

Satz 3.52 (Singulärwertzerlegung)

Sei $A \in \mathbb{R}^{m\times n}$ eine Matrix vom Rang r. Dann gibt es orthogonale Matrizen $U \in \mathbb{R}^{m\times m}$ und $V \in \mathbb{R}^{n\times n}$ sowie eine „Diagonalmatrix“

$$\Sigma = \begin{bmatrix} \Sigma_r & O \\ O & O \end{bmatrix} \in \mathbb{R}^{m\times n} \qquad \textit{mit} \qquad \Sigma_r = \operatorname{diag}(\sigma_1, \sigma_2, \dots, \sigma_r) \in \mathbb{R}^{r\times r}$$

und $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$, so dass A die Zerlegung

$$A = U\Sigma V^\top \tag{SVD}$$

besitzt.

Die Darstellung (SVD) heißt Singulärwertzerlegung von A (*engl.* singular value decomposition).
Die positiven Zahlen $\sigma_1, \dots, \sigma_r$ nennt man die Singulärwerte von A.
Schreibt man $U = [\boldsymbol{u}_1, \boldsymbol{u}_2, \dots, \boldsymbol{u}_m]$ und $V = [\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n]$, so heißen $\boldsymbol{u}_i \in \mathbb{R}^m$ bzw. $\boldsymbol{v}_i \in \mathbb{R}^n$ zugehörige linke bzw. rechte Singulärvektoren.

(1) Darstellung von A als Summe von r Rang-1-Matrizen:

$$A = U\Sigma V^\top = [\boldsymbol{u}_1, \boldsymbol{u}_2, \dots, \boldsymbol{u}_r]\, \Sigma_r\, [\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_r]^\top = \sum_{i=1}^{r} \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i^\top$$

(2) Es gelten:

$$A\boldsymbol{v}_i = \begin{cases} \sigma_i \boldsymbol{u}_i & i = 1, 2, \dots, r, \\ \boldsymbol{0} & i = r+1, \dots, n \end{cases}$$

und

$$A^\top \boldsymbol{u}_i = \begin{cases} \sigma_i \boldsymbol{v}_i & i = 1, 2, \dots, r, \\ \boldsymbol{0} & i = r+1, \dots, m. \end{cases}$$

(3)

$\{\boldsymbol{u}_1, \dots, \boldsymbol{u}_r\}$	ist eine ON-Basis von	$\mathscr{R}(A)$.
$\{\boldsymbol{u}_{r+1}, \dots, \boldsymbol{u}_m\}$	ist eine ON-Basis von	$\mathscr{N}(A^\top) = \mathscr{R}(A)^\perp$.
$\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_r\}$	ist eine ON-Basis von	$\mathscr{R}(A^\top) = \mathscr{N}(A)^\perp$.
$\{\boldsymbol{v}_{r+1}, \dots, \boldsymbol{v}_n\}$	ist eine ON-Basis von	$\mathscr{N}(A)$.

(4) $\sigma_1^2, \dots, \sigma_r^2$ sind die von Null verschiedenen Eigenwerte von $A^\top A$ bzw. $AA^\top$:

$$A^\top A = V\Sigma^\top \Sigma V^\top = V \begin{bmatrix} \Sigma_r^2 & O \\ O & O \end{bmatrix} V^\top,$$
$$AA^\top = U\Sigma\Sigma^\top U^\top = U \begin{bmatrix} \Sigma_r^2 & O \\ O & O \end{bmatrix} U^\top.$$

Insbesondere sind die Singulärwerte $\sigma_1, \dots, \sigma_r$ durch A eindeutig festgelegt. Die rechten Singulärvektoren $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$ bilden eine ON-Basis des $\mathbb{R}^n$ aus Eigenvektoren von $A^\top A$:

$$A^\top A \boldsymbol{v}_i = \begin{cases} \sigma_i^2 \boldsymbol{v}_i & i = 1, 2, \dots, r, \\ \mathbf{0} & i = r+1, \dots, n. \end{cases}$$

Die linken Singulärvektoren $\boldsymbol{u}_1, \dots, \boldsymbol{u}_m$ bilden eine ON-Basis des $\mathbb{R}^m$ aus Eigenvektoren von $AA^\top$:

$$AA^\top \boldsymbol{u}_i = \begin{cases} \sigma_i^2 \boldsymbol{u}_i & i = 1, 2, \dots, r, \\ \mathbf{0} & i = r+1, \dots, m. \end{cases}$$

(5) Ist $A = A^\top \in \mathbb{R}^{n \times n}$ mit von Null verschiedenen Eigenwerten $\lambda_1, \dots, \lambda_r$, $|\lambda_1| \geq \cdots \geq |\lambda_r| > 0$, dann sind $\sigma_i = |\lambda_i|$ die Singulärwerte von A.

(6) Das Bild der (n-dimensionalen) Einheitskugel unter A ist ein Ellipsoid (im $\mathbb{R}^m$) mit Mittelpunkt $\mathbf{0}$ und Halbachsen $\sigma_i \boldsymbol{u}_i$ ($\sigma_i := 0$ für $i > r$).

(7) Für $A \in \mathbb{R}^{m \times n}$ gilt $\|A\|_2 = \sigma_1$.
Ist $A \in \mathbb{R}^{n \times n}$ invertierbar, gilt außerdem $\|A^{-1}\|_2 = \sigma_n^{-1}$.

(8) Analoge Aussagen gelten für komplexe Matrizen $A = U \Sigma V^H$ (U, V unitär). (In (5) ist 'symmetrisch' durch 'normal' zu ersetzen.)

Die lineare Ausgleichsrechnung führt auf Minimierungsprobleme der Form

$$\|\boldsymbol{b} - A\boldsymbol{x}\|_2 \to \min_{\boldsymbol{x}\in\mathbb{R}^n} \tag{3.21}$$

zu gegebener Matrix $A \in \mathbb{R}^{m\times n}$ $(m \geq n)$ und Vektor $\boldsymbol{b} \in \mathbb{R}^m$.

Satz 3.53

Es sei $A \in \mathbb{R}^{m\times n}$ eine Matrix vom Rang r mit Singulärwertzerlegung $A = U\Sigma V^\top = U \left[\begin{smallmatrix} \Sigma_r & O \\ O & O \end{smallmatrix}\right] V^\top$. Dann löst

$$\boldsymbol{x}^\star = V \begin{bmatrix} \Sigma_r^{-1} & O \\ O & O \end{bmatrix} U^\top \boldsymbol{b}$$

die lineare Ausgleichsaufgabe (3.21). *Darüberhinaus ist $\boldsymbol{x}^\star$ die eindeutig bestimmte Lösung von* (3.21) *mit minimaler Euklid-Norm.*

Mit Hilfe der Singulärwertzerlegung einer Matrix A lassen sich optimale Approximationen an A durch Matrizen niedrigen Rangs konstruieren. Hierzu ist ein Abstandsbegriff für Matrizen erforderlich. Die in Satz 3.27 eingeführten Normen für Vektoren lassen sich auch auf Matrizen verallgemeinern:

Definition 3.54

Eine Matrixnorm ist eine Abbildung

$$\|\cdot\| : \mathbb{K}^{n\times n} \to \mathbb{R}_0^+, \qquad A \mapsto \|A\|, \qquad \text{die}$$

- positiv definit, d.h. $\|A\| > 0 \;\; \forall\, A \in \mathbb{K}^{n\times n}$, $A \neq O$, sowie
- homogen, d.h. $\|\alpha A\| = |\alpha|\|A\| \;\; \forall\, \alpha \in \mathbb{K}$, $A \in \mathbb{K}^{n\times n}$, ist und der
- Dreiecksungleichung, d.h. $\|A + B\| \leq \|A\| + \|B\| \;\; \forall\, A, B \in \mathbb{K}^{n\times n}$, genügt. Zusätzlich soll sie
- submultiplikativ sein, d.h. $\|AB\| \leq \|A\|\|B\| \;\; \forall\, A, B \in \mathbb{K}^{n\times n}$.

Beispiel: Frobenius oder Schur-Norm ($A = [a_{i,j}]_{1 \le i,j \le n} \in \mathbb{K}^{n \times n}$)

$$\|A\|_F := \left(\sum_{i,j=1}^{n} |a_{i,j}|^2 \right)^{1/2}.$$

Definition 3.55

Jede Vektornorm $\|\cdot\|_V$ im $\mathbb{K}^n$ induziert durch

$$\|A\|_M := \max_{\|\boldsymbol{x}\|_V = 1} \|A\boldsymbol{x}\|_V = \max_{\boldsymbol{x} \neq 0} \frac{\|A\boldsymbol{x}\|_V}{\|\boldsymbol{x}\|_V}$$

eine Matrixnorm in $\mathbb{K}^{n \times n}$, die von $\|\cdot\|_V$ induzierte Matrixnorm.

Es ist üblich, für $\|\cdot\|_V$ und $\|\cdot\|_M$ das gleiche Symbol zu verwenden:

- $\|\cdot\|_1$ induziert die Spaltensummennorm $\|A\|_1 = \max_{1 \le j \le n} \sum_{i=1}^{n} |a_{i,j}|$.
- $\|\cdot\|_2$ induziert die Spektralnorm $\|A\|_2 = \sqrt{\lambda_{\max}(A^H A)}$.
- $\|\cdot\|_\infty$ induziert die Zeilensummennorm $\|A\|_\infty = \max_{1 \le i \le n} \sum_{j=1}^{n} |a_{i,j}|$.

Definition 3.56

Eine Vektornorm $\|\cdot\|_V$ und eine Matrixnorm $\|\cdot\|_M$ sind miteinander verträglich (oder passen zueinander), wenn

$$\|A\boldsymbol{x}\|_V \le \|A\|_M \|\boldsymbol{x}\|_V \quad \forall\, \boldsymbol{x} \in \mathbb{K}^n,\ A \in \mathbb{K}^{n\times n}$$

gilt.

- Bezeichnet $\|\cdot\|_M$ die von der Vektornorm $\|\cdot\|_V$ induzierte Matrixnorm, so sind $\|\cdot\|_V$ und $\|\cdot\|_M$ miteinander verträglich.
- $\|\cdot\|_M$ ist die kleinste Matrixnorm, die mit $\|\cdot\|_V$ verträglich ist: Für alle $A \in \mathbb{K}^{n\times n}$ gilt

$$\|A\|_M = \min\{\|A\| \;:\; \|\cdot\| \text{ ist mit } \|\cdot\|_V \text{ verträglich}\}.$$

- Die Euklidsche Vektornorm $\|\cdot\|_2$ ist mit der Frobenius-Norm $\|\cdot\|_F$ verträglich, was $\|A\|_2 \le \|A\|_F$ für alle $A \in \mathbb{K}^{n\times n}$ impliziert.

Satz 3.57 (Schmidt, 1907; Eckart & Young, 1936; Mirsky, 1960)

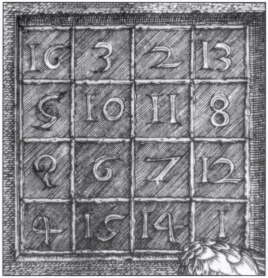
Für eine Matrix $A \in \mathbb{R}^{m \times n}$ vom Rang r mit Singulärwertzerlegung $A = U \Sigma V^\top$ besitzt die Approximationsaufgabe

$$\min\{\|A - B\|_2 \,:\, B \in \mathbb{R}^{m \times n} \text{ und } \mathrm{Rang}(B) \le k\}$$

für $k < r$ die Lösung

$$A_k := \sum_{i=1}^{k} \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i^\top \qquad \text{mit} \qquad \|A - A_k\|_2 = \sigma_{k+1}.$$

Digitale Fotografien kann man als Matrizen aus Helligkeits- und Farbwerten auffassen. Wir betrachten als Beispiel einen Ausschnitt aus Albrecht Dürers Kupferstich Melancholie I (1514), der ein magisches Quadrat enthält.



- Die Bildinformation ist in einer Pixelmatrix A der Dimension 359×371 gespeichert, deren Einträge – ganze Zahlen zwischen 1 und 64 – verschiedene Graustufen repräsentieren.
- Wir approximieren A durch Matrizen niedrigen Rangs k (vgl. Satz 3.57). Im technischen Berechnungssystem MATLAB geschieht dies bspw. durch die Anweisungen

```
load detail.mat;
[U,S,V] = svd(A);
A_k = U(:,1:k)*S(1:k,1:k)*V(:,1:k)';
image(X_k), colormap('gray'), axis('image'), axis('off')
```

- Zur Speicherung von A_k sind im Wesentlichen

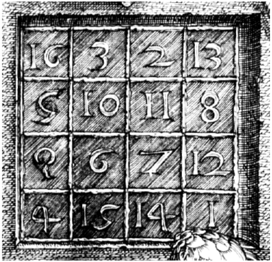
$$k(m+n) = 730k$$

Zahlen erforderlich; die volle Matrix A erfordert

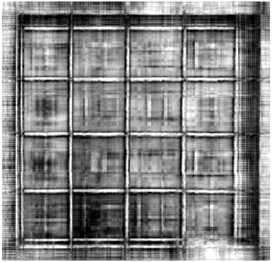
$$mn = 359 \cdot 371 = 133\,189$$

.

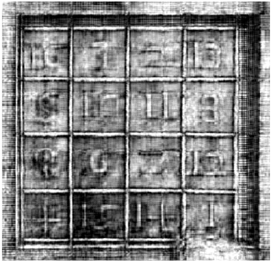
Original



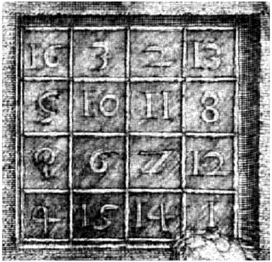
k=10



k=20



k=40



Als Maße für die Qualität der Kompression kann man den relativen Fehler

$$\frac{\|A - A_k\|_2}{\|A\|_2} = \frac{\sigma_{k+1}}{\sigma_1}$$

sowie die Kompressionrate

$$\frac{\text{Speicherbedarf von } A_k}{\text{Speicherbedarf von } A} = \frac{k(m+n)}{mn}$$

heranziehen. In diesem Beispiel:

k	Relativer Fehler σ_{k+1}/σ_1	Kompressionsrate
10	0.0666	0.055
20	0.0528	0.110
40	0.0382	0.219

Ziele erreicht?

Sie sollten nun (bzw. nach Abschluss der Übungen/Selbststudium):

- wissen, was die Begriffe Vektorraum, Unterraum, Basis und lineare Unabhängigkeit bedeuten,
- diese Vorstellungen insbesondere im Fall des $\mathbb{K}^n$ sicher anwenden können,
- sicher mit Matrizen rechnen können und den Zusammenhang zwischen Matrizen und linearen Abbildungen kennen,
- Dimension von Kern und Bild (Rang) einer Matrix (linearen Abbildung) sicher bestimmen können,
- lineare Gleichungssysteme sicher von Hand lösen können und sicher über die Anzahl der Lösungen entscheiden können,
- Determinanten moderater Größe sicher berechnen können und anhand der Ergebnisse über die eindeutige Lösbarkeit von LGS entscheiden können,
- wissen, was man unter Invertierbarkeit von Matrizen versteht und welche Beziehungen zur eindeutigen Lösbarkeit von LGS bestehen,
- die Inverse für kleinere Matrizen sicher berechnen können,

Ziele erreicht?

(Fortsetzung)

- über Skalarprodukte, Orthogonalität, Längen- und Winkelmessung bescheidwissen (vor allem im Fall $\mathbb{K}^n$),
- Kreuz- und Spatprodukt sicher berechnen können und über deren geometrische Interpretation bescheidwissen,
- Lagebeziehungen von Punkten, Geraden und Ebenen im Raum sicher analysieren können.
- über die Zusammenhänge zwischen Matrizen und linearen Abbildungen bescheidwissen,
- einfache geometrische Transformationen wie Drehungen und Spiegelungen mit Hilfe orthogonaler Matrizen beschreiben können,
- wissen, was orthogonale Matrizen charakterisiert,
- die Begriffe Eigenwert, Eigenvektor und Vielfachheit tiefgreifend verstanden haben,
- Eigenwerte und Eigenvektoren sicher berechnen können, ggf. auch unter Beachtung der Matrixstruktur,
- wissen, was es mit Basen von Eigenvektoren und Diagonalisierbarkeit auf sich hat.