
Kapitel 5

Lineare Algebra

- Grundbegriffe
 - Das Skalarprodukt
 - Matrizen
 - Die Determinante
 - Lineare Gleichungssysteme
 - Die Inverse einer Matrix
 - Eigenwerte und Eigenvektoren
 - Anwendungen
-

Skalare und Vektoren

Ein Skalar ist einfach eine reelle Zahl (in der entsprechenden Maßeinheit), wie etwa die Temperatur (z.B. $25^{\circ}C$).

Bei Vektoren kommt noch eine Richtung hinzu: Beispielsweise hat eine durch Vektoren repräsentierte Geschwindigkeit nicht nur einen Wert (z.B. $3m/s$), sondern auch eine Bewegungsrichtung.

Die einzige Ausnahme: Der Nullvektor hat keine Richtung.

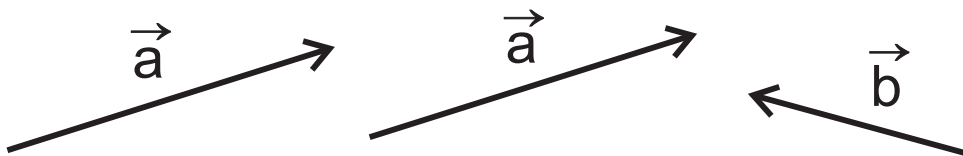
Vektoren veranschaulicht man sich gewöhnlich als Verschiebungspfeile:

Definition

Unter einem Vektor versteht man eine gerichtete Strecke. Man bezeichnet Vektoren mit \vec{a} , \vec{b} , ... Zwei Vektoren heißen gleich, wenn sie sich durch Parallelverschiebung ineinander überführen lassen.

Veranschaulichung von Vektoren

Bei Vektoren kommt es also nur auf Richtung und Länge an — der Anfangspunkt ist egal.



In vielen Anwendungen hat man es mit ebenen oder mit räumlichen Vektoren zu tun:

Dabei identifizieren wir den \mathbb{R}^2 mit den Punkten der Ebene und ordnen jedem Punkt $(x_1, x_2) \in \mathbb{R}^2$ einen (zweidimensionalen) Vektor \vec{x} zu:

Bei beliebigem Anfangspunkt gehe man x_1 Einheiten nach rechts (bei *negativem* x_1 entsprechend nach links) in x -Richtung und x_2 Einheiten in y -Richtung eines *Kartesischen Koordinatensystems*.

Vektoren im \mathbb{R}^n

Man beachte, dass bei Vektoren das Zahlenpaar üblicherweise als Spalte (Spaltenvektor) geschrieben wird:

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Will man Vektoren als Zeilen (Zeilenvektoren) schreiben, so benutzt man den *transponierten Vektor*:

$$\vec{x} = (x_1, x_2)^T.$$

Analoges gilt für räumliche Vektoren des \mathbb{R}^3 . Man kann sogar ganz allgemein definieren:

Definition

Einen Vektor \vec{a} des \mathbb{R}^n stellt man dar als

$$\vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ \dots \\ a_n \end{pmatrix} = (a_1, a_2, \dots, a_n)^T$$

mit $a_i \in \mathbb{R}$ für alle $i = 1, \dots, n$. Der Vektor $(0, 0, \dots, 0)^T$ heißt Nullvektor.

Vektoraddition und Skalarmultiplikation im \mathbb{R}^n

Vektoren kann man bekanntermaßen addieren und mit einem Skalar multiplizieren:

Definition

Jeweils zwei Vektoren $\vec{a} = (a_1, a_2, \dots, a_n)^T$ und $\vec{b} = (b_1, b_2, \dots, b_n)^T$ des \mathbb{R}^n kann man (komponentenweise) addieren:

$$\vec{a} + \vec{b} := (a_1 + b_1, a_2 + b_2, \dots, a_n + b_n)^T$$

bzw. einen Vektor \vec{a} mit einem Skalar $\lambda \in \mathbb{R}$ multiplizieren:

$$\lambda \vec{a} := (\lambda a_1, \lambda a_2, \dots, \lambda a_n)^T.$$

Kommutativgesetz der Vektoraddition

Für die beiden Operationen

„Vektoraddition“ und „Skalarmultiplikation“

gelten nun einige einfache Rechengesetze, etwa

$$\vec{a} + \vec{b} = \vec{b} + \vec{a} \quad \text{für alle } \vec{a}, \vec{b} \in \mathbb{R}^n.$$

Dies ist ganz einfach der Fall, weil jeweils komponentenweise für die reellen Zahlen das Kommutativgesetz gilt:

$$a_i + b_i = b_i + a_i.$$

Veranschaulichen kann man sich dieses Rechengesetz am so genannten Kräfteparallelogramm.

Übung

Welche der folgenden Rechengesetze gelten für Vektoren?

(Dabei seien \vec{a} und \vec{b} Vektoren des \mathbb{R}^n und λ und μ Skalare aus \mathbb{R} .)

a) $\vec{a} + \lambda = \lambda + \vec{a},$

b) $\lambda(\vec{a} + \vec{b}) = \lambda\vec{a} + \lambda\vec{b},$

c) $\vec{a} \cdot \vec{b} = \vec{b} \cdot \vec{a},$

d) $(\lambda + \mu)\vec{a} = \lambda\vec{a} + \mu\vec{a},$

e) $\vec{a} + \vec{0} = \vec{a},$

f) $0\vec{a} = \vec{0}.$

Lösung

Es gelten die Rechengesetze b), d), e) und f).

Im Ausdruck a) werden verbotenerweise Vektoren und Skalare addiert, was überhaupt nicht definiert ist.

Im Ausdruck c) werden Vektoren multipliziert (und nicht ein Skalar mit einem Vektor), was erst später als Skalarprodukt definiert wird.

\mathbb{R}^n mit Vektoraddition ist kommutative Gruppe

Für die Vektoraddition gelten Gesetze, die wir im Zusammenhang mit Gruppen bereits kennen gelernt haben:

- So ist die Summe zweier Vektoren wiederum ein Vektor;
- es gibt ein *neutrales Element*, den Nullvektor $\vec{0}$, mit $\vec{a} + \vec{0} = \vec{a}$;
- zu jedem Vektor \vec{a} gibt es einen *inversen Vektor* $-\vec{a}$ (mit den Komponenten $-a_i$ anstelle von a_i).

Außerdem gelten das Kommutativgesetz und das Distributivgesetz.

Derartige Mengen (hier \mathbb{R}^n) mit einer entsprechenden Verknüpfung (hier die Vektoraddition) heißen kommutative (oder abelsche) Gruppen.

Vektorraum

Nimmt man nun noch die Skalarmultiplikation hinzu mit den unten aufgeführten Gesetzen, so spricht man von einem Vektorraum:

Definition

Eine Menge V bildet einen Vektorraum über \mathbb{R} , wenn folgende Axiome gelten:

- a) Die Menge V mit der Vektoraddition $+$, also $(V, +)$, ist eine abelsche Gruppe.**
- b) Zwischen einem Skalar $\lambda \in \mathbb{R}$ und einem Vektor $\vec{a} \in V$ ist eindeutig ein Produkt $\lambda\vec{a} \in V$ erklärt. Dabei gelten für $\lambda, \mu \in \mathbb{R}$ und $\vec{a}, \vec{b} \in V$ die folgenden Rechengesetze:**
 - $\lambda(\vec{a} + \vec{b}) = \lambda\vec{a} + \lambda\vec{b}$,
 - $(\lambda\mu)\vec{a} = \lambda(\mu\vec{a})$,
 - $(\lambda + \mu)\vec{a} = \lambda\vec{a} + \mu\vec{a}$,
 - $1\vec{a} = \vec{a}$.

Beispiel

Wir betrachten die Menge aller reellwertigen Funktionen auf dem Intervall $[0, 1]$.

Offenbar kann man zwei Funktionen f und g addieren und die Summenfunktion $f + g$ ist definiert durch

$$(f + g)(x) := f(x) + g(x)$$

für alle $x \in [0, 1]$.

Ähnlich funktioniert die Multiplikation einer Funktion f mit einer reellen Zahl λ :

$$(\lambda f)(x) := \lambda \cdot f(x)$$

für alle $x \in [0, 1]$.

Mit diesen beiden Operationen ist die Menge aller reellwertigen Funktionen auf dem Intervall $[0, 1]$ ein Vektorraum.

Übung

Wie lauten in vorangegangenem Beispiel der Nullvektor und wie der zum Vektor

$$f = \sin x$$

inverse Vektor?

Lösung

Der Nullvektor ist hier ganz einfach die Funktion

$$f \equiv 0,$$

die jedem $x \in [0, 1]$ den Funktionswert 0 zuordnet.

Der zu $f(x) = \sin x$ inverse Vektor ist

$$-f(x) = -\sin x,$$

denn

$$f(x) + (-f)(x) = \sin x + (-\sin x) = 0,$$

also gleich der Funktion, die konstant den Wert 0 ergibt.

(Von der Umkehrfunktion $\arcsin x$ sprechen wir in einem ganz anderen Zusammenhang!)

Linearkombination

Die Begriffe „Linearkombination“, „lineare Abhängigkeit“ bzw. „lineare Unabhängigkeit“ sowie „Basis“ und „Dimension“ lassen sich ganz allgemein für beliebige Vektorräume erklären.

Im \mathbb{R}^2 beispielsweise können wir aus den beiden Vektoren $\vec{a} = (1, 4)^T$ und $\vec{b} = (-2, 5)^T$ die *Linearkombination* erzeugen:

$$2\vec{a} - 0.5\vec{b} = (3, 5.5)^T.$$

Definition

Einen Vektor \vec{b} der Form

$$\vec{b} = \lambda_1 \vec{a}_1 + \lambda_2 \vec{a}_2 + \dots + \lambda_n \vec{a}_n$$

mit $\lambda_i \in \mathbb{R}$ für $i = 1, \dots, n$ nennt man eine Linearkombination der Vektoren

$$\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n.$$

Übung

Stellen Sie, falls möglich, den (Zeilen-)Vektor

$$(-1, 5)$$

jeweils als Linearkombination der folgenden Vektoren dar:

- a) $(1, 0), (0, 1)$;
- b) $(1, 2), (-4, -1)$;
- c) $(1, 2), (2, 4)$;
- d) $(-1, 5), (2, -10)$;
- e) $(1, 2)$;
- f) $(2, -10)$;
- g) $(1, 0), (0, 1), (1, 1)$.

In welchen Fällen ist dies evtl. sogar auf mehrere Arten möglich?

Lösung

- a) *Es gilt:* $(-1, 5) = (-1) \cdot (1, 0) + 5 \cdot (0, 1)$.
- b) *Hier ist:* $(-1, 5) = 3 \cdot (1, 2) + 1 \cdot (-4, -1)$.
- c) *Es ist unmöglich, $(-1, 5)$ als Linearkombination von $(1, 2)$ und $(2, 4)$ darzustellen.*
- d) *Es gilt z.B. $(-1, 5) = 1 \cdot (-1, 5) + 0 \cdot (2, -10)$ oder auch $(-1, 5) = 7 \cdot (-1, 5) + 3 \cdot (2, -10)$.*
- e) *Es ist unmöglich, $(-1, 5)$ als Linearkombination von $(1, 2)$ darzustellen.*
- f) *Natürlich ist $(-1, 5) = (-0.5) \cdot (2, -10)$.*
- g) *Hier ist z.B. $(-1, 5) = (-1) \cdot (1, 0) + 5 \cdot (0, 1) + 0 \cdot (1, 1)$ oder $(-1, 5) = 2 \cdot (1, 0) + 8 \cdot (0, 1) + (-3) \cdot (1, 1)$.*

In den Fällen d) und g) gibt es jeweils sogar unendlich viele Möglichkeiten, $(-1, 5)$ als Linearkombination der angegebenen Vektoren zu schreiben.

Lineare Unabhängigkeit, Lineare Abhängigkeit

Definition

Die Vektoren $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n$ heißen **linear unabhängig**, wenn aus der Gleichung

$$\lambda_1 \vec{a}_1 + \lambda_2 \vec{a}_2 + \dots + \lambda_n \vec{a}_n = \vec{0} \quad (*)$$

folgt, dass alle Koeffizienten $\lambda_1, \lambda_2, \dots, \lambda_n$ gleich Null sind:

$$\lambda_1 = \lambda_2 = \dots = \lambda_n = 0.$$

Gibt es hingegen Koeffizienten $\lambda_1, \lambda_2, \dots, \lambda_n$, die nicht alle gleich 0 sind, für die aber (*) erfüllt ist, so heißen die Vektoren $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n$ **linear abhängig**.

Übung

Welche der angegebenen Vektoren sind linear unabhängig?

- a) $(1, 0), (0, 1)$;
- b) $(1, 2), (-4, -1)$;
- c) $(1, 2), (2, 4)$;
- d) $(-1, 5), (2, -10)$;
- e) $(1, 2)$;
- f) $(2, -10)$;
- g) $(1, 0), (0, 1), (1, 1)$.

Lösung

Die Vektoren unter a), b), e) und f) sind linear unabhängig.

Alle anderen sind jeweils linear abhängig.

Bemerkung:

Die Darstellung des (Zeilen-) Vektors $(-1, 5)$ als Linearkombination der beiden Vektoren

$$(1, 0), (0, 1)$$

war am einfachsten, denn hier musste man überhaupt nicht rechnen.

Basis und Basisvektoren

Es lässt sich allgemein für Vektoren des \mathbb{R}^2 zeigen:

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = a_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + a_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = a_1 \vec{e}_1 + a_2 \vec{e}_2.$$

Die Vektoren $\vec{e}_1 := (1, 0)^T$ und $\vec{e}_2 := (0, 1)^T$ nennt man *Basisvektoren* des \mathbb{R}^2 .

Aber die Vektoren $(1, 2)$ und $(-4, -1)$ bilden ebenfalls eine so genannte *Basis* des Vektorraums \mathbb{R}^2 , denn auch hier lässt sich jeder Vektor des \mathbb{R}^2 eindeutig als Linearkombination dieser beiden Vektoren schreiben.

Definition

Die Vektoren $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n$ eines Vektorraums bilden eine Basis, wenn sie linear unabhängig sind und wenn sich jeder Vektor \vec{x} des Vektorraums (eindeutig) als Linearkombination dieser Basisvektoren mit geeigneten $\lambda_i \in \mathbb{R}$ darstellen lässt:

$$\vec{x} = \lambda_1 \vec{a}_1 + \lambda_2 \vec{a}_2 + \dots + \lambda_n \vec{a}_n.$$

Übung

Welche der angegebenen Vektoren bilden eine Basis des \mathbb{R}^2 ?

- a) $(1, 0), (0, 1)$;
- b) $(1, 2), (-4, -1)$;
- c) $(1, 2), (2, 4)$;
- d) $(-1, 5), (2, -10)$;
- e) $(1, 2)$;
- f) $(2, -10)$;
- g) $(1, 0), (0, 1), (1, 1)$.

Lösung

Nur die Vektoren in a) und b) bilden jeweils eine Basis.

Die Vektoren in c), d) und g) sind nicht linear unabhängig.

Die Vektoren in e) und f) sind zwar linear unabhängig, aber nicht jeder Vektor des \mathbb{R}^2 lässt sich als Linearkombination (hier: als Vielfaches) der angegebenen Vektoren darstellen.

Dimension eines Vektorraumes

Mit Hilfe der Anzahl der Basisvektoren (die bei allen Basen identisch ist) lässt sich mathematisch der Begriff der Dimension definieren.

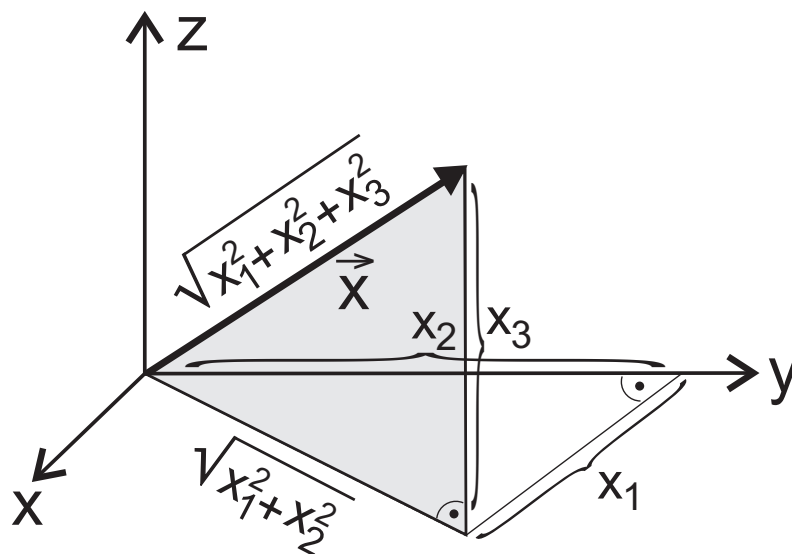
Natürlich gilt dann, dass die Dimension des \mathbb{R}^2 gleich 2, die Dimension des \mathbb{R}^3 gleich 3 und die Dimension des \mathbb{R}^n gleich n ist.

Definition

Die (endliche) Anzahl n der Vektoren in einer Basis eines Vektorraums ist immer gleich. Man sagt, dass der Vektorraum die Dimension n hat.

Länge eines Vektors

Wie man die Länge eines Vektors berechnet, lässt sich anschaulich sehr gut erläutern an Vektoren $\vec{x} = (x_1, x_2, x_3)^T \in \mathbb{R}^3$:



Nach dem Lehrsatz von Pythagoras hat zunächst die senkrechte Projektion von \vec{x} in die xy -Ebene die Länge $\sqrt{x_1^2 + x_2^2}$ (rechtwinkliges Dreieck!).

Eine weitere Anwendung des Pythagoräischen Lehrsatzes auf das grau unterlegte (rechtwinklige) Dreieck liefert dann die Vektorlänge

$$\sqrt{\left(\sqrt{x_1^2 + x_2^2}\right)^2 + x_3^2} = \sqrt{x_1^2 + x_2^2 + x_3^2}.$$

Betrag bzw. Norm eines Vektors

Formal gelten diese Überlegungen auch im \mathbb{R}^n , so dass man definiert:

Definition

Die Länge eines Vektors

$$\vec{x} = (x_1, x_2, \dots, x_n)^T \in \mathbb{R}^n$$

nennt man **Betrag oder Norm** von \vec{x} . Sie ist gegeben durch

$$\|\vec{x}\| := \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}.$$

Vektoren mit $\|\vec{x}\| = 1$ heißen Einheitsvektoren.

Übung

- a) *Bestimmen Sie alle Vektoren, die den Betrag 0 haben.*
- b) *Welche Norm hat der Vektor $\vec{x} = (2, 4, 4)^T$? Bestimmen Sie einen Einheitsvektor, der dieselbe Richtung wie \vec{x} hat.*

Lösung

a) Es ist $\|\vec{x}\| \stackrel{!}{=} 0$ äquivalent zu

$$\sqrt{x_1^2 + x_2^2 + \dots + x_n^2} \stackrel{!}{=} 0,$$

also zu $x_1 = x_2 = \dots = x_n = 0$. Der einzige Vektor mit Länge 0 ist also der Nullvektor.

b) Es ist

$$\|\vec{x}\| = \sqrt{2^2 + 4^2 + 4^2} = \sqrt{36} = 6.$$

Aus jedem Vektor $\vec{x} \neq \vec{0}$ lässt sich mittels

$$\vec{y} = \vec{x} / \|\vec{x}\|$$

ein Einheitsvektor konstruieren. Hier gilt

$$\vec{y} = \frac{1}{6}(2, 4, 4)^T = \left(\frac{1}{3}, \frac{2}{3}, \frac{2}{3}\right)^T.$$

Skalarprodukt, Inneres Produkt

Man kann nun jeweils die entsprechenden Komponenten a_i, b_i zweier Vektoren $\vec{a}, \vec{b} \in \mathbb{R}^n$ miteinander multiplizieren und anschließend aufsummieren und erhält einen Skalar:

Definition

Für zwei Vektoren $\vec{a} = (a_1, a_2, \dots, a_n)^T$ und $\vec{b} = (b_1, b_2, \dots, b_n)^T$ ist das Skalarprodukt bzw. innere Produkt von \vec{a} und \vec{b} , bezeichnet mit

$$\vec{a} \cdot \vec{b} \quad \text{oder} \quad \vec{a}^T \vec{b},$$

definiert als die reelle Zahl

$$\vec{a} \cdot \vec{b} := a_1 b_1 + a_2 b_2 + \dots + a_n b_n.$$

Beispiel

Wir berechnen das Skalarprodukt der Vektoren

$$\vec{a} = (\sqrt{12}, 1, 6)^T, \vec{b} = (0, 1, 1)^T :$$

Es ergibt sich zu:

$$\vec{a} \cdot \vec{b} = \sqrt{12} \cdot 0 + 1 \cdot 1 + 6 \cdot 1 = 7.$$

Rechenregeln für Skalarprodukte

Wichtige Rechenregeln für das Skalarprodukt ergeben sich unmittelbar aus dessen Definition:

Definition

Für Vektoren $\vec{a}, \vec{b}, \vec{c} \in \mathbb{R}^n$ und Skalare $\lambda \in \mathbb{R}$ gilt:

- a) $\vec{a} \cdot \vec{a} \geq 0, \quad \vec{a} \cdot \vec{a} = 0 \iff \vec{a} = \vec{0},$
- b) $\|\vec{a}\| = \sqrt{\vec{a} \cdot \vec{a}},$
- c) $\vec{a} \cdot \vec{b} = \vec{b} \cdot \vec{a},$
- d) $(\lambda\vec{a}) \cdot \vec{b} = \vec{a} \cdot (\lambda\vec{b}) = \lambda(\vec{a} \cdot \vec{b}),$
- e) $\vec{a} \cdot (\vec{b} + \vec{c}) = \vec{a} \cdot \vec{b} + \vec{a} \cdot \vec{c}.$

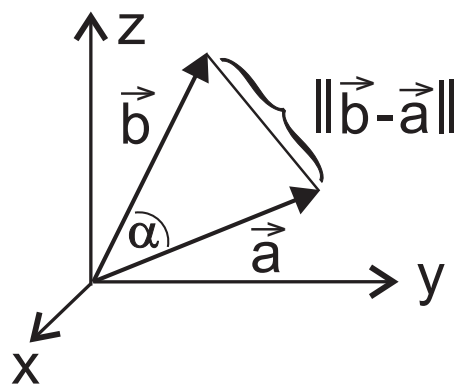
Beispiel

Für zwei Vektoren $\vec{a}, \vec{b} \in \mathbb{R}^n$ berechnen wir $\|\vec{b} - \vec{a}\|^2$.
Unter Anwendung obiger Rechenregeln b), c) und e) ergibt sich:

$$\begin{aligned}\|\vec{b} - \vec{a}\|^2 &= (\vec{b} - \vec{a}) \cdot (\vec{b} - \vec{a}) \\ &= \vec{b} \cdot (\vec{b} - \vec{a}) - \vec{a} \cdot (\vec{b} - \vec{a}) \\ &= \vec{b} \cdot \vec{b} - \vec{b} \cdot \vec{a} - \vec{a} \cdot \vec{b} + \vec{a} \cdot \vec{a} \\ &= \|\vec{a}\|^2 + \|\vec{b}\|^2 - 2\vec{a} \cdot \vec{b}.\end{aligned}$$

Winkel zwischen zwei Vektoren

Mit Hilfe des Skalarproduktes kann man auch Winkel α zwischen zwei Vektoren \vec{a} und \vec{b} bestimmen:



Das durch die beiden Vektoren \vec{a} und \vec{b} aufgespannte Dreieck hat die Seitenlängen $\|\vec{a}\|$, $\|\vec{b}\|$ und $\|\vec{b} - \vec{a}\|$. Den Winkel α kann man mit dem *Kosinussatz für schiefwinklige Dreiecke* berechnen:

$$\|\vec{b} - \vec{a}\|^2 = \|\vec{a}\|^2 + \|\vec{b}\|^2 - 2\|\vec{a}\|\|\vec{b}\| \cos \alpha.$$

Wegen $\|\vec{b} - \vec{a}\|^2 = \|\vec{a}\|^2 + \|\vec{b}\|^2 - 2\vec{a} \cdot \vec{b}$ ist obige Gleichung aber äquivalent zu

$$\|\vec{a}\|^2 + \|\vec{b}\|^2 - 2\vec{a} \cdot \vec{b} = \|\vec{a}\|^2 + \|\vec{b}\|^2 - 2\|\vec{a}\|\|\vec{b}\| \cos \alpha.$$

Winkel zwischen zwei Vektoren

$$\|\vec{a}\|^2 + \|\vec{b}\|^2 - 2\vec{a} \cdot \vec{b} = \|\vec{a}\|^2 + \|\vec{b}\|^2 - 2\|\vec{a}\|\|\vec{b}\|\cos\alpha.$$

Kürzen gemeinsamer Terme und Auflösen der Gleichung nach α liefert:

Für den Winkel α mit $0 \leq \alpha \leq \pi$ zwischen zwei Vektoren $\vec{a} \neq \vec{0}$, $\vec{b} \neq \vec{0}$ gilt

$$\vec{a} \cdot \vec{b} = \|\vec{a}\|\|\vec{b}\|\cos\alpha$$

bzw.

$$\alpha = \arccos\left(\frac{\vec{a} \cdot \vec{b}}{\|\vec{a}\|\|\vec{b}\|}\right).$$

Übung

Berechnen Sie den Winkel α zwischen den Vektoren

$$\vec{a} = (\sqrt{12}, 1, 6)^T \quad \text{und} \quad \vec{b} = (0, 1, 1)^T.$$

Lösung

Es ist $\vec{a} \cdot \vec{b} = 7$ und

$$\|\vec{a}\| = \sqrt{12 + 1 + 36} = 7,$$

$$\|\vec{b}\| = \sqrt{0 + 1 + 1} = \sqrt{2}.$$

Damit gilt

$$\cos \alpha = \frac{7}{7 \cdot \sqrt{2}},$$

also

$$\alpha = \arccos\left(\frac{1}{\sqrt{2}}\right) = \frac{\pi}{4}.$$

Die beiden Vektoren bilden daher einen 45° -Winkel.

Bemerkung:

Man bezeichnet Vektoren, die aufeinander senkrecht stehen (im Zeichen $\vec{a} \perp \vec{b}$), also einen 90° -Winkel bilden, als *orthogonale Vektoren*. Da für $\alpha \in [0, \pi]$ gilt: $\cos \alpha = 0 \iff \alpha = \frac{\pi}{2}$, liefert das Skalarprodukt ein einfaches Kriterium für die Orthogonalität von Vektoren:

$$\vec{a} \perp \vec{b} \iff \vec{a} \cdot \vec{b} = 0.$$

Grundlegende Definitionen

Wir können uns zwei Vektorräume — etwa $\mathbb{R}^n, \mathbb{R}^m$ — vorgeben und einem Vektor \vec{x} des ersten Raumes einen Vektor \vec{z} des zweiten zuordnen.

Interessant sind in diesem Zusammenhang spezielle Funktionen, so genannte *lineare Abbildungen*, die zusätzlich bestimmte (lineare) Eigenschaften erfüllen und üblicherweise mit φ bezeichnet werden:

Definition

Eine Abbildung $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}^m$ heißt lineare Abbildung, wenn für alle $\vec{x}, \vec{y} \in \mathbb{R}^n$ und $c \in \mathbb{R}$ gilt:

$$\varphi(\vec{x} + \vec{y}) = \varphi(\vec{x}) + \varphi(\vec{y}), \quad \varphi(c\vec{x}) = c\varphi(\vec{x}).$$

Beispiel

Setzt man $\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ und $\vec{z} = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$, dann ist durch

$$z_1 = 3x_1 + x_2 + 5x_3$$

$$z_2 = -2x_1 + 8x_3$$

offensichtlich eine lineare Abbildung $\varphi : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ gegeben. Jedem Vektor $\vec{x} \in \mathbb{R}^3$ wird ein Vektor $\vec{z} \in \mathbb{R}^2$ zugeordnet.

Im Prinzip ist die Abbildung durch die Gleichungskoeffizienten definiert. Daher kann man diese auch beschreiben, indem man die Koeffizienten zu einem Schema zusammenfasst:

$$\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} 3 & 1 & 5 \\ -2 & 0 & 8 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

Das Koeffizientenschema $\begin{pmatrix} 3 & 1 & 5 \\ -2 & 0 & 8 \end{pmatrix}$ nennt man Matrix.

Lineare Abbildung

Allgemein ist eine lineare Abbildung $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}^m$ gegeben durch

$$\begin{aligned} z_1 &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ z_2 &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \dots & \quad \dots \quad \dots \quad \dots \quad \dots \\ z_m &= a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n. \end{aligned}$$

Jedem Vektor $\vec{x} \in \mathbb{R}^n$ wird ein Vektor $\vec{z} \in \mathbb{R}^m$ zugeordnet. In Matrix-Schreibweise lauten die Gleichungen dann:

$$\begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_m \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

Matrix

Definition

Ein rechteckiges Zahlenschema aus m Zeilen und n Spalten nennt man eine Matrix vom Typ (m, n) :

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1k} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2k} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ik} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mk} & \cdots & a_{mn} \end{pmatrix}.$$

Die Zahlen a_{ik} heißen Elemente der Matrix. Das Element a_{ik} steht in der i -ten Zeile und k -ten Spalte. Daher heißt i Zeilenindex und k Spaltenindex.

Matrix

Die Gleichungen

$$\begin{aligned} z_1 &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ z_2 &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \dots & \quad \dots \quad \dots \quad \dots \quad \dots \\ z_m &= a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{aligned}$$

können jetzt abkürzend geschrieben werden als

$$\vec{z} = A\vec{x}.$$

Die i -te Komponente z_i des Vektors \vec{z} ergibt sich immer als *Skalarprodukt* aus der i -ten Matrixzeile und dem Vektor \vec{x} :

$$z_i = (a_{i1}, a_{i2}, \dots, a_{in}) \cdot \vec{x}.$$

Übung

Wie lauten die Gleichungen von $\vec{z} = A\vec{x}$ mit

$$A = \begin{pmatrix} 2 & 0 & 4 & 7 \\ -5 & 1 & 3 & 0 \end{pmatrix}$$

ausgeschrieben?

Welches Bild \vec{z} ergibt sich für $\vec{x}^T = (1, 2, 3, 4)$?

Lösung

Der $(2, 4)$ -Matrix A entnimmt man, dass

$$a_{11} = 2, a_{12} = 0, a_{13} = 4, a_{14} = 7$$

(1. Zeile) und

$$a_{21} = -5, a_{22} = 1, a_{23} = 3, a_{24} = 0$$

(2. Zeile) ist.

Somit lauten die Gleichungen:

$$z_1 = 2x_1 + 0x_2 + 4x_3 + 7x_4$$

$$z_2 = -5x_1 + 1x_2 + 3x_3 + 0x_4$$

Konkret ergibt sich

$$z_1 = 2 \cdot 1 + 0 \cdot 2 + 4 \cdot 3 + 7 \cdot 4 = 42$$

$$z_2 = -5 \cdot 1 + 1 \cdot 2 + 3 \cdot 3 + 0 \cdot 4 = 6.$$

Matrizen — Weitere Begriffe

Matrizen notiert man üblicherweise mit großen Buchstaben: A, B, C, \dots . Möchte man auch den Typ auführen, so schreibt man kurz $A_{(m,n)}$ für eine (m, n) -Matrix. Gebräuchlich ist auch die Schreibweise

$$(a_{ik}), (b_{ik}), (c_{ik}), \dots,$$

wenn man notieren möchte, wie das allgemeine Element der jeweiligen Matrix in Position (i, k) definiert ist.

Wichtige Begriffe bzw. Sonderfälle:

- Eine Matrix A mit gleich vielen Zeilen und Spalten, d.h. eine (m, m) -Matrix, nennt man *quadratisch*. Ihre Elemente $a_{11}, a_{22}, \dots, a_{mm}$ bilden die so genannte *Hauptdiagonale*. Ihren Typ, üblicherweise *Ordnung* genannt, notiert man abkürzend zu A_m .
- Eine quadratische (m, m) -Matrix D_m , bei der alle Elemente außerhalb der Hauptdiagonalen verschwinden ($d_{ik} = 0$ für $i \neq k$), heißt *Diagonalmatrix*. Abkürzend schreibt man auch $D_m = \text{diag}(d_{11}, d_{22}, \dots, d_{mm})$.

Matrizen — Weitere Begriffe

- Eine quadratische Matrix, die nur auf und oberhalb bzw. unterhalb der Hauptdiagonalen von Null verschiedene Elemente haben darf, heißt *obere Dreiecksmatrix* bzw. *untere Dreiecksmatrix*.
- Eine quadratische (m, m) -Matrix, die auf der Hauptdiagonalen nur „1“, sonst „0“ stehen hat, nennt man *Einheitsmatrix* der Ordnung m . Üblicherweise bezeichnet man sie mit I_m oder I (I für Identität).
- Eine (m, n) -Matrix, deren Elemente alle 0 sind, heißt *Nullmatrix*, bezeichnet mit 0 bzw. $0_{(m, n)}$.
- Ein Spezialfall ist die $(m, 1)$ -Matrix, sie besteht nur aus einer Spalte und ist unser üblicher *Spaltenvektor*. Eine $(1, n)$ -Matrix besteht dagegen nur aus einer Zeile und wird *Zeilenvektor* genannt.

Beispiel

$$D_4 = \begin{pmatrix} 7 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 5 \end{pmatrix}, \quad I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

$$O_{(3,2)} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$$

$D_4 = \text{diag}(7, 2, 0, 5)$ ist eine Diagonalmatrix der Ordnung 4.

I_3 ist die Einheitsmatrix der Ordnung 3.

$O_{(3,2)}$ ist die (3, 2)-Nullmatrix.

Operationen und Rechenregeln für Matrizen

Zunächst halten wir fest, dass zwei Matrizen A, B genau dann *gleich* sind (im Zeichen $A = B$), wenn sie vom gleichen Typ sind *und* elementweise übereinstimmen ($a_{ik} = b_{ik}$ für alle i, k).

Im Folgenden werden wir nun die wichtigsten Rechenregeln für Matrizen auführen.

Skalarmultiplikation

Eine Matrix A wird mit einem Skalar λ multipliziert, indem man *alle* Elemente von A mit λ multipliziert:

$$\lambda A = \lambda \cdot (a_{ik}) = (\lambda \cdot a_{ik}).$$

Beispiel

$$A = \begin{pmatrix} -1 & 2 & 3 \\ 4 & 5 & 0 \end{pmatrix} \implies 3A = \begin{pmatrix} -3 & 6 & 9 \\ 12 & 15 & 0 \end{pmatrix}$$

Matrixaddition/-subtraktion

Zwei Matrizen $A = (a_{ik})$ und $B = (b_{ik})$ des *gleichen Typs* werden addiert bzw. subtrahiert, indem man ihre entsprechenden Elemente addiert bzw. subtrahiert:

$$A \pm B = (a_{ik}) \pm (b_{ik}) = (a_{ik} \pm b_{ik}).$$

Beispiel

$$\begin{aligned} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} + \begin{pmatrix} -1 & 5 \\ 6 & -7 \end{pmatrix} &= \begin{pmatrix} 1 - 1 & 2 + 5 \\ 3 + 6 & 4 - 7 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 7 \\ 9 & -3 \end{pmatrix} \end{aligned}$$

Transponierte einer Matrix

Vertauscht man in einer Matrix A Zeilen mit Spalten, so entsteht die Transponierte von A : A^T . Für die Elemente von $A = (a_{ik})$ und $A^T = (a_{ik}^T)$ gilt

$$a_{ik}^T = a_{ki} \text{ für alle } i \text{ und } k.$$

Beispiel

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \implies A^T = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{pmatrix}$$

Für transponierte Matrizen folgen unmittelbar aus der Definition die Rechengesetze:

$$(A + B)^T = A^T + B^T \quad \text{und} \quad (A^T)^T = A.$$

In vielen Anwendungen treten übrigens so genannte *symmetrische Matrizen* auf, bei denen die Elemente spiegelsymmetrisch zur Hauptdiagonalen angeordnet sind, d.h. es gilt

$$a_{ik} = a_{ki} \text{ für alle } i \text{ und } k \text{ bzw. } A^T = A.$$

Übung

Vereinfachen Sie den Ausdruck

$$(A^T + B)^T - A.$$

Lösung

$$\begin{aligned}(A^T + B)^T - A &= (A^T)^T + B^T - A \\ &= A + B^T - A = B^T.\end{aligned}$$

Die Matrizenmultiplikation

Die Multiplikation zweier Matrizen A und B wird so definiert, dass sie der Hintereinanderschaltung der zugehörigen Abbildungen entspricht. Wir betrachten hierzu zunächst ein Beispiel:

Gegeben seien die zwei Abbildungen

$$\vec{y} = B\vec{x},$$

$$\text{d.h.} \quad \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

und

$$\vec{z} = A\vec{y},$$

$$\text{d.h.} \quad \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.$$

Gesucht ist nun die zusammengesetzte Abbildung $\vec{z} = C\vec{x}$, die \vec{z} direkt in Abhängigkeit von \vec{x} darstellt.

Die Matrizenmultiplikation

Dazu berechnen wir

$$\begin{aligned}z_1 &= a_{11}y_1 + a_{12}y_2 \\ &= a_{11}(b_{11}x_1 + b_{12}x_2 + b_{13}x_3) + a_{12}(b_{21}x_1 + b_{22}x_2 + b_{23}x_3) \\ &= \underbrace{(a_{11}b_{11} + a_{12}b_{21})}_{=: c_{11}}x_1 + \underbrace{(a_{11}b_{12} + a_{12}b_{22})}_{=: c_{12}}x_2 \\ &\quad + \underbrace{(a_{11}b_{13} + a_{12}b_{23})}_{=: c_{13}}x_3\end{aligned}$$

und

$$\begin{aligned}z_2 &= a_{21}y_1 + a_{22}y_2 \\ &= a_{21}(b_{11}x_1 + b_{12}x_2 + b_{13}x_3) + a_{22}(b_{21}x_1 + b_{22}x_2 + b_{23}x_3) \\ &= \underbrace{(a_{21}b_{11} + a_{22}b_{21})}_{=: c_{21}}x_1 + \underbrace{(a_{21}b_{12} + a_{22}b_{22})}_{=: c_{22}}x_2 \\ &\quad + \underbrace{(a_{21}b_{13} + a_{22}b_{23})}_{=: c_{23}}x_3\end{aligned}$$

Die Matrizenmultiplikation

Man erkennt, dass sich die c_{ik} jeweils als *Skalarprodukt* der i -ten Zeile von A und k -ten Spalte von B ergeben!

Es gilt einerseits

$$\vec{z} = C\vec{x} \quad \text{mit} \quad C = \begin{pmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \end{pmatrix},$$

andererseits aber

$$\vec{z} = A\vec{y} = AB\vec{x}.$$

Man definiert daher C als Produkt:

$$C = A \cdot B.$$

Definition

Für zwei Matrizen A und B ist das Produkt $A \cdot B$ genau dann definiert, wenn die Spaltenzahl von A gleich der Zeilenzahl von B ist. Es gilt dann

$$A_{(m,n)} \cdot B_{(n,s)} = C_{(m,s)}$$

Die Elemente c_{ik} ($i = 1, \dots, m; k = 1, \dots, s$) von C sind definiert als Skalarprodukte der i -ten Zeile von A und der k -ten Spalte von B :

$$c_{ik} = a_{i1}b_{1k} + a_{i2}b_{2k} + \dots + a_{in}b_{nk}.$$

Die Matrizenmultiplikation – Typcheck

Ob ein Produkt $A \cdot B$ definiert ist, lässt sich leicht überprüfen, wenn man den Typ notiert :

$$A_{(m,n)} \cdot B_{(r,s)} = C_{(m,s)}.$$

Die inneren Elemente n, r der beiden „Typ-Paare“ müssen gleich sein: $n = r$. In diesem Fall kann man den Typ der Produktmatrix ablesen: Er entspricht den beiden äußeren Elementen, d.h. ergibt sich zu (m, s) .

Die Berechnung des Elementes c_{ik} der Produktmatrix lässt sich einfach durchführen, wenn man die Matrizen nebeneinander schreibt und das Skalarprodukt aus der i -ten Zeile von A mit der k -ten Spalte von B bildet.

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \dots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} b_{11} & \dots & b_{1k} & \dots & b_{1s} \\ \vdots & & \vdots & & \vdots \\ b_{n1} & \dots & b_{nk} & \dots & b_{ns} \end{pmatrix} = \begin{pmatrix} c_{11} & \dots & c_{1k} & \dots & c_{1s} \\ \vdots & & \vdots & & \vdots \\ c_{i1} & \dots & c_{ik} & \dots & c_{is} \\ \vdots & & \vdots & & \vdots \\ c_{m1} & \dots & c_{mk} & \dots & c_{ms} \end{pmatrix}$$

Übung

Berechnen Sie das Produkt

$$C = A \cdot B$$

der Matrizen

$$A_{(2,3)} = \begin{pmatrix} 0 & 1 & 2 \\ 3 & 4 & 0 \end{pmatrix}$$

und

$$B_{(3,2)} = \begin{pmatrix} 5 & 6 \\ -1 & 7 \\ 0 & -8 \end{pmatrix}.$$

Ist auch das Produkt $B \cdot A$ definiert?

Von welchem Typ ist es?

Lösung**Das Produkt**

$$C = A_{(2,3)} \cdot B_{(3,2)}$$

ist wegen der Gleichheit der inneren Elemente ($3 = 3$) definiert. Es hat den (an den äußeren Elementen abzulesenden) Typ $(2, 2)$ und ergibt sich durch folgende Skalarproduktbildungen (Falk-Schemas):

				5				6
				-1				7
				0				-8
0	1	2		0 · 5 - 1 · 1 + 2 · 0 = -1				0 · 6 + 1 · 7 - 2 · 8 = -9
3	4	0		3 · 5 - 4 · 1 + 0 · 0 = 11				3 · 6 + 4 · 7 - 0 · 8 = 46

Die Produktmatrix C lautet also $C = \begin{pmatrix} -1 & -9 \\ 11 & 46 \end{pmatrix}$.

Das Produkt $B_{(3,2)} \cdot A_{(2,3)}$ ist wegen $2 = 2$ (innere Elemente) ebenfalls definiert und vom Typ $(3, 3)$ (äußere Elemente).

Rechenregeln für Matrixmultiplikation

Auch bei der Multiplikation von Matrizen gelten viele, von den reellen Zahlen her bekannte, Rechengesetze:

- Assoziativgesetz: $(AB)C = A(BC)$,
- Distributivgesetze: $A(B + C) = AB + AC$,
 $(A + B)C = AC + BC$,
- Speziell gilt: $AI = IA = A$ (I : Einheitsmatrix),
- $(\lambda A)B = A(\lambda B) = \lambda(AB)$,
- $(AB)^T = B^T A^T$.

Übung

a) Gegeben seien die Matrizen

$$A_{(2,3)} = \begin{pmatrix} 0 & 1 & 2 \\ 3 & 4 & 0 \end{pmatrix},$$

$$B_{(3,2)} = \begin{pmatrix} 5 & 6 \\ -1 & 7 \\ 0 & -8 \end{pmatrix}$$

und $C = AB$.

Berechnen Sie möglichst einfach das Produkt ABC .

b) Vereinfachen Sie den Ausdruck

$$(C + I)^T D^T - (DC)^T.$$

Lösung

a) *Es ist nach Assoziativgesetz*

$$\begin{aligned} ABC &= (AB)C = CC \\ &= \begin{pmatrix} -1 \cdot (-1) - 9 \cdot 11 & -1 \cdot (-9) - 9 \cdot 46 \\ 11 \cdot (-1) + 46 \cdot 11 & 11 \cdot (-9) + 46 \cdot 46 \end{pmatrix} \\ &= \begin{pmatrix} -98 & -405 \\ 495 & 2017 \end{pmatrix}. \end{aligned}$$

b) *Es ist*

$$\begin{aligned} (C + I)^T D^T - (DC)^T &= C^T D^T + I^T D^T - C^T D^T \\ &= ID^T = D^T. \end{aligned}$$

Fehlendes Kommutativgesetz

Im Gegensatz zur kommutativen Multiplikation von reellen Zahlen ist bei der Multiplikation von Matrizen die Reihenfolge der Faktoren wichtig: Das *Kommutativgesetz* gilt also *nicht*!

So existiert beispielsweise das Produkt

$$A_{(m,n)} \cdot B_{(n,s)} = C_{(m,s)}.$$

Jedoch existiert das vertauschte Produkt

$$B_{(n,s)} \cdot A_{(m,n)}$$

nur im Spezialfall $s = m$, da die Spaltenzahl von B und die Zeilenzahl von A übereinstimmen müssen. In diesem Fall ist aber

$$A_{(m,n)} \cdot B_{(n,m)} = (AB)_{(m,m)},$$

$$B_{(n,m)} \cdot A_{(m,n)} = (BA)_{(n,n)}.$$

Jetzt existieren zwar die beiden Produkte AB und BA , sie können aber nur dann identisch sein, wenn $m = n$ gilt.

Fehlendes Kommutativgesetz

Aber auch in letzterem Fall gilt im Allg. $AB \neq BA$, wie das folgende Beispiel zeigt: Aus den Matrizen

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

erhält man die Produkte

$$AB = \begin{pmatrix} 1 & -1 \\ 0 & -1 \end{pmatrix}, \quad BA = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}.$$

Es gilt also tatsächlich

$$AB \neq BA.$$

Dies ist auch der Grund, warum Matrizen eines bestimmten Typs *keinen Körper* bilden: Bezüglich der Matrizenaddition liegt zwar eine kommutative Gruppe vor, die Nullmatrix ist neutrales Element, invers zu A ist die Matrix $-A$).

Aber bezüglich der Multiplikation kann keine abelsche Gruppe vorliegen, da eben das Kommutativgesetz nicht für alle Matrizen erfüllt ist. Weil die Multiplikation assoziativ ist und die Distributivgesetze gelten, liegt jedoch ein *Ring* vor.

Warnung vor Fehler

Beim Rechnen mit Matrizen sei abschließend vor einem weiteren Fehler gewarnt: Aus der reellen Analysis kennt man die Aussage:

„Ein Produkt ist genau dann Null, wenn mindestens einer der beiden Faktoren Null ist“.

Diese Aussage gilt für Matrizenprodukte *nicht*, wie das nachfolgende Beispiel zeigt: Mit

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

folgt offensichtlich

$$AB = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = 0$$

(Nullmatrix). D.h. aus $AB = 0$ folgt im Allg. eben *nicht* $A = 0$ oder $B = 0$.

Rang einer Matrix

In der Lösungstheorie linearer Gleichungssysteme ist ein weiterer Begriff im Zusammenhang mit Matrizen wichtig:

Definition

Die Maximalzahl linear unabhängiger Spalten einer Matrix A heißt Spaltenrang von A , die Maximalzahl linear unabhängiger Zeilen heißt Zeilenrang von A . Da immer „Zeilenrang = Spaltenrang“ gilt, spricht man vom Rang der Matrix schlechthin:

$$\text{Rang von } A := \text{Rg}(A).$$

Die obige Feststellung „Zeilenrang = Spaltenrang“ lässt sich natürlich mathematisch beweisen, was wir hier aber nicht nachvollziehen wollen.

Beispiel

Die Matrix

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}$$

hat die Spalten

$$\vec{a}_1^T = (1, 1, 1), \quad \vec{a}_2^T = (2, 2, 2), \quad \vec{a}_3^T = (3, 3, 3).$$

Offensichtlich besteht die Menge $\{\vec{a}_1, \vec{a}_2, \vec{a}_3\}$ lediglich aus einem linear unabhängigen Vektor, also ist $\text{Rg}(A) = 1$.

Dagegen gilt, dass alle Spalten der Matrix

$$B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

linear unabhängig sind, also ist $\text{Rg}(B) = 3$.

Nichtsinguläre bzw. reguläre Matrix

Speziell für quadratische Matrizen ist eine weitere Definition wichtig:

Definition

Eine quadratische (n, n) -Matrix A heißt nichtsingulär oder regulär, falls

$$\text{Rg}(A) = n$$

gilt. Ist $\text{Rg}(A) < n$, wird sie singulär genannt.

Bei einer nichtsingulären Matrix sind also alle n Spalten (und damit auch Zeilen) linear unabhängig.

Die Determinante

Eine quadratische $(1, 1)$ -Matrix A besteht nur aus einem einzigen Element a_{11} . Dieses ist gleichzeitig auch der Wert der Determinante von A .

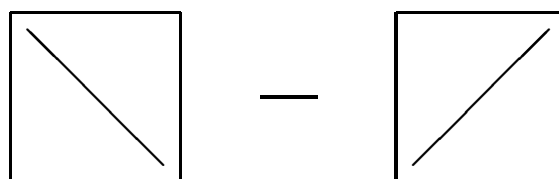
Definition

Ist $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ eine $(2, 2)$ -Matrix, dann heißt

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12}$$

zweireihige Determinante von A .

Statt die vielen Indices in obiger Formel auswendig zu lernen, empfiehlt sich das Merken der Berechnungsregel in folgender Symbolik:



Beispiel

Für die Determinante der Matrix

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

gilt:

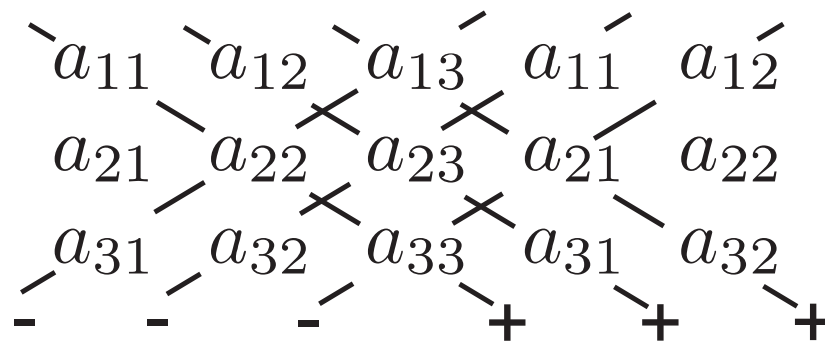
$$\det(A) = \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = 1 \cdot 4 - 2 \cdot 3 = -2.$$

Regel von Sarrus

Auch die Berechnung von dreireihigen Determinanten für (3, 3)-Matrizen lässt sich ähnlich einfach mit der so genannten *Regel von Sarrus* durchführen:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{matrix} a_{11}a_{22}a_{33} & + & a_{12}a_{23}a_{31} & + & a_{13}a_{21}a_{32} \\ -a_{31}a_{22}a_{13} & - & a_{32}a_{23}a_{11} & - & a_{33}a_{21}a_{12} \end{matrix}$$

Diese Formel lässt sich schematisiert sehr leicht merken und anwenden:



Übung

Berechnen Sie die 3-reihige Determinante

$$\det(A) = \begin{vmatrix} 2 & 9 & 5 \\ 2 & -3 & 4 \\ 1 & 2 & 2 \end{vmatrix}.$$

Lösung

Nach obiger Vorschrift erhalten wir das folgende Rechenschema:

$$\left| \begin{array}{ccc|cc} 2 & 9 & 5 & 2 & 9 \\ 2 & -3 & 4 & 2 & -3 \\ 1 & 2 & 2 & 1 & 2 \end{array} \right| .$$

Damit ergibt sich:

$$\det(A) = 2 \cdot (-3) \cdot 2 + 9 \cdot 4 \cdot 1 + 5 \cdot 2 \cdot 2 - 1 \cdot (-3) \cdot 5 - 2 \cdot 4 \cdot 2 - 2 \cdot 2 \cdot 9 = 7.$$

Determinante und Rang

Man beachte, dass für n -reihige Determinanten mit $n > 3$ eine entsprechende Regel *nicht* mehr gilt. Diese lassen sich aber mit dem so genannten *Laplace'schen Entwicklungssatz* berechnen.

Ohne Beweis weisen wir noch auf folgenden wichtigen Zusammenhang hin:

Für eine (n, n) -Matrix A gilt folgende Äquivalenz:

$$\det(A) \neq 0 \iff \text{Rg}(A) = n$$

Das Gauß'sche Eliminationsverfahren

Wir betrachten ein (m, n) -System von m linearen Gleichungen mit n Unbekannten ($m < n$ stets!):

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots & \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m. \end{aligned}$$

Mit der Koeffizientenmatrix $A = (a_{ik})$ ($i = 1, \dots, m$, $k = 1, \dots, n$) und den Vektoren $\vec{x}^T = (x_1, \dots, x_n)$, $\vec{b}^T = (b_1, \dots, b_m)$ lautet das System in Matrixschreibweise $A\vec{x} = \vec{b}$.

Definition

Ein lineares Gleichungssystem

$$A\vec{x} = \vec{b}$$

heißt homogen, wenn $\vec{b} = \vec{0}$. Andernfalls nennt man es inhomogen. Ist $\vec{b} \neq \vec{0}$, so heißt $A\vec{x} = \vec{0}$ das zugehörige homogene System.

Lösungsmenge $L(A, \vec{b})$, Erweiterte Koeffizientenmatrix

Die Lösungsmenge

$$L(A, \vec{b}) := \{\vec{x} \in \mathbb{R}^n \mid A\vec{x} = \vec{b}\}$$

des Systems $A\vec{x} = \vec{b}$ lässt sich nun mit dem *Gauß'schen Eliminationsverfahren* ermitteln, das die so genannte *erweiterte Koeffizientenmatrix* benutzt:

$$(A|\vec{b}) = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array} \right).$$

Das Verfahren arbeitet mit *elementaren Zeilenumformungen* an der erweiterten Koeffizientenmatrix, welche die Lösungsmenge des Systems offenbar nicht ändern:

- Vertauschung zweier Zeilen,
- Addition des λ -fachen einer Zeile zu einer anderen Zeile,
- Multiplikation einer Zeile mit einer Zahl $\lambda \neq 0$.

Zeilenstufenform

Die Zeilenumformungen werden nun benutzt, um die Koeffizientenmatrix in folgende so genannte *Zeilenstufenform* $(\bar{A}, \vec{\bar{b}})$ (siehe Abb.) zu bringen:

$$(\bar{A}, \vec{\bar{b}}) = \left(\begin{array}{cccc|cccc} & & & & \bar{b}_1 & & & & & \\ & & & & \bar{b}_2 & & & & & \\ & & & & \vdots & & & & & \\ & & & & \vdots & & & & & \\ & & & & \bar{b}_r & & & & & \\ & & & & \bar{b}_{r+1} & & & & & \\ & & & & \vdots & & & & & \\ & & & & \bar{b}_m & & & & & \end{array} \right) \begin{array}{l} \left. \vphantom{\begin{array}{c} \bar{b}_1 \\ \bar{b}_2 \\ \vdots \\ \bar{b}_r \end{array}} \right\} r \\ \left. \vphantom{\begin{array}{c} \bar{b}_{r+1} \\ \vdots \\ \bar{b}_m \end{array}} \right\} m-r \end{array}$$

In dieser Form müssen alle Einträge, die mit „*“ gekennzeichnet sind, ungleich Null sein. Man nennt diese *Pivotelemente*, die Zeile entsprechend *Pivotzeile*.

Unterhalb der skizzierten „Stufenlinie“ dürfen in \bar{A} nur Nullen stehen. Der durch die Umformungen ebenfalls geänderte Vektor $\vec{\bar{b}}$ kann beliebige Komponenten haben.

Eliminationsfaktor

Um nun beispielsweise in der k -ten Spalte unterhalb des Pivots — bezeichnen wir es mit p — Nullen zu erzeugen, müssen wir die entsprechenden Elemente der darunter liegenden Zeilen mittels Addition des λ -fachen (so genannter Eliminationsfaktor) der Pivotzeile zur jeweiligen Zeile zu Null machen. Sind

$$(0, \dots, 0, p, \dots)$$

die Pivotzeile und

$$(0, \dots, 0, a, \dots)$$

eine Zeile, in der das Element a zu Null werden muss, dann ergibt sich der *Eliminationsfaktor* λ durch die Forderung

$$a + \lambda p \stackrel{!}{=} 0, \text{ also zu } \lambda = -\frac{a}{p}.$$

Ist die Zeilenstufenform erreicht, so können nun im Falle der Lösbarkeit des Systems durch „Rückwärtsauflösen“ die entsprechenden Variablenwerte ermittelt werden.

Beispiel

Das Verfahren sei an folgendem linearen Gleichungssystem verdeutlicht:

$$3x_1 - 3x_2 + 6x_3 = 9$$

$$2x_1 \quad \quad + 3x_3 = 6$$

$$x_1 + x_2 + 2x_3 = 4$$

Die erweiterte Koeffizientenmatrix dieses Systems schreiben wir als Tableau, d.h. ohne die runden Klammern, auf:

$$\begin{array}{cccc|c} (1) & \textcircled{3} & -3 & 6 & 9 \\ (2) & 2 & 0 & 3 & 6 \\ (3) & 1 & 1 & 2 & 4. \end{array}$$

Im 1. Schritt ist das Pivotelement die „eingekreiste“ 3 in der 1. Spalte. Darunter müssen nun zwei Nullen erzeugt werden.

Da die Pivotzeile die Form $(3, -3, 6, 9)$ hat und die darunterliegende Zeile $(2, 0, 3, 6)$ lautet, bestimmt sich der erste Eliminationsfaktor aus $2 + \lambda \cdot 3 \stackrel{!}{=} 0$ zu $\lambda = -\frac{2}{3}$, der zweite analog zu $\lambda = -\frac{1}{3}$.

Beispiel — Fortsetzung

Bezeichnen wir mit z_i die Zeile (i) des Tableaus, so sind die elementaren Umformungen $z_{2'} = z_2 - \frac{2}{3}z_1$ und $z_{3'} = z_3 - \frac{1}{3}z_1$ (jeweils elementweise!) durchzuführen. Dies ergibt ein neues Tableau, bei dem im 2. Schritt nun in der zweiten Spalte unterhalb des neuen Pivotelements 2 Nullen erzeugt werden müssen. Hierzu wird mit der *Eliminationszeile* ($2'$) die Umformung $z_{3''} = z_{3'} - z_{2'}$ ausgeführt.

$$\begin{array}{ccc|c}
 (1') & 3 & -3 & 6 & 9 \\
 (2') & 0 & \textcircled{2} & -1 & 0 \\
 (3') & 0 & 2 & 0 & 1
 \end{array}
 \xrightarrow{\text{2. Schritt}}
 \begin{array}{ccc|c}
 (1'') & 3 & -3 & 6 & 9 \\
 (2'') & 0 & 2 & -1 & 0 \\
 (3'') & 0 & 0 & 1 & 1.
 \end{array}$$

Jetzt liegt ein *gestaffeltes System* vor. Die Lösung kann bei solchen Systemen immer durch „Rückwärtsauflösen“ aus den Gleichungen ermittelt werden: $x_3 = 1$,

$$x_2 = \frac{1}{2}(0 + x_3) = \frac{1}{2}, \quad x_1 = \frac{1}{3}(9 + 3x_2 - 6x_3) = \frac{3}{2}.$$

Übung

Wenden Sie das Gauß'sche Verfahren auf folgendes System an:

$$3x_1 - 3x_2 + 6x_3 = 9$$

$$2x_1 \quad \quad + 3x_3 = 6$$

$$x_1 + x_2 + x_3 = 4$$

Lösung

Das System entspricht bis auf eine Änderung in der dritten Gleichung ($a_{33} = 2$ wird zu $a_{33} = 1$) dem Gleichungssystem des vorherigen Beispiels. Mit denselben elementaren Umformungen wie oben erhält man daher die Tableaufolge:

$$\begin{array}{ccc|c} (1) & 3 & -3 & 6 & 9 \\ (2) & 2 & 0 & 3 & 6 \\ (3) & 1 & 1 & 1 & 4 \end{array} \rightarrow \begin{array}{ccc|c} (1') & 3 & -3 & 6 & 9 \\ (2') & 0 & 2 & -1 & 0 \\ (3') & 0 & 2 & -1 & 1 \end{array} \rightarrow \begin{array}{ccc|c} (1'') & 3 & -3 & 6 & 9 \\ (2'') & 0 & 2 & -1 & 0 \\ (3'') & 0 & 0 & 0 & 1. \end{array}$$

Der letzten Zeile ($3''$) des Endtableaus entspricht nun die Gleichung

$$0 \cdot x_1 + 0 \cdot x_2 + 0 \cdot x_3 = 1.$$

Dies ist offensichtlich ein Widerspruch. Somit hat das System keine Lösung.

Unlösbares System/ Freie Parameter

Die *Unlösbarkeit* eines inhomogenen Gleichungssystems erkennt man also daran, dass es in der Zeilenstufenform mindestens ein

$$b_i \neq 0 \quad \text{mit} \quad (r + 1) \leq i \leq m$$

gibt, bei dem die restliche (linke) Zeile aus lauter Nullen besteht.

Jetzt fehlt uns nur noch der Fall unendlich vieler Lösungen mit frei wählbaren Unbekannten, die man dann *freie Parameter* nennt.

Beispiel

Wir betrachten das System der letzten Übung, ändern aber die rechte Seite $\vec{b}^T = (9, 6, 4)$ ab in $\vec{b}^T = (9, 7, 4)$. Analoge Zeilenumformungen liefern dann die Tableaufolge

$$\begin{array}{l} (1) \quad 3 \quad -3 \quad 6 \quad | \quad 9 \\ (2) \quad 2 \quad 0 \quad 3 \quad | \quad 7 \\ (3) \quad 1 \quad 1 \quad 1 \quad | \quad 4 \end{array} \rightarrow \begin{array}{l} (1') \quad 3 \quad -3 \quad 6 \quad | \quad 9 \\ (2') \quad 0 \quad 2 \quad -1 \quad | \quad 1 \\ (3') \quad 0 \quad 2 \quad -1 \quad | \quad 1 \end{array} \rightarrow \begin{array}{l} (1'') \quad 3 \quad -3 \quad 6 \quad | \quad 9 \\ (2'') \quad 0 \quad 2 \quad -1 \quad | \quad 1 \\ (3'') \quad 0 \quad 0 \quad 0 \quad | \quad 0. \end{array}$$

Letzte Zeile (3''): $0 \cdot x_1 + 0 \cdot x_2 + 0 \cdot x_3 = 0$, offensichtlich stets erfüllt. Damit reduziert sich das System auf zwei Gleichungen für drei Unbekannte. Wir setzen $x_3 = t$ mit $t \in \mathbb{R}$ beliebig.

Wieder ergeben sich die restlichen Unbekannten durch „Rückwärtsauflösen“ zu $x_2 = \frac{1}{2}(1 + x_3) = \frac{1}{2}(1 + t)$, $x_1 = \frac{1}{3}(9 + 3x_2 - 6x_3) = \frac{1}{2}(7 - 3t)$. Mit $\vec{u} = (\frac{7}{2}, \frac{1}{2}, 0)^T$ und $\vec{v}^T = (-\frac{3}{2}, \frac{1}{2}, 1)$ lässt sich die Lösungsmenge auch in Parameterform zu $\vec{x} = \vec{u} + t \cdot \vec{v}$ angeben.

Freie Parameter

Im Allgemeinen erkennt man an der Zeilenstufenform wie viele Parameter frei gewählt werden können:

Ist r die Anzahl der nicht aus lauter Nullen bestehenden Zeilen, so sind $n - r$ Unbekannte frei wählbar. Diese fungieren dann als Parameter und die Lösungsmenge kann in *Parameterform* angegeben werden.

Nicht immer sind die Parameter beliebig wählbar: Man kann aber stets die Variablen nehmen, bei denen in den zugehörigen Spalten ein *horizontaler* Verlauf der „Stufen“ beginnt bzw. fortgesetzt wird.

Gauß'sches Eliminationsverfahren

Das Gauß'sche Eliminationsverfahren zur Lösung von $A_{(m,n)}\vec{x} = \vec{b}$ besteht aus folgenden Schritten:

- a) Man erstelle die erweiterte Koeffizientenmatrix $(A | \vec{b})$ in Tableauform.
- b) Man bringe die Matrix A mittels elementarer Zeilenumformungen auf „Zeilenstufenform“, wobei auch die Spalte \vec{b} mit umgeformt werden muss. Ergebnis: $(\bar{A} | \vec{\bar{b}})$.
- c) Aus $(\bar{A}, \vec{\bar{b}})$ ermittle man die Anzahl r der von Null verschiedenen Zeilen von \bar{A} und stelle durch Überprüfung von $\vec{\bar{b}}$ fest, ob Lösungen existieren.
- d) Falls ja ($r = m$ oder $r < m$ und $\bar{b}_i = 0$ für alle $r + 1 \leq i \leq m$), ermittle man durch „Rückwärtsauflösen“ die Lösung. Diese hat immer $n - r$ frei wählbare Parameter.

Rangbestimmung mittels Gauß-Verfahren

Die Anwendung des Gauß'schen Eliminationsverfahrens auf die Matrix A liefert eine Matrix \bar{A} in „Zeilenstufenform“. Offensichtlich sind die ersten r Zeilen von \bar{A} linear unabhängig.

Die dabei benutzten elementaren Zeilenumformungen ändern aber nicht die lineare Ab- bzw. Unabhängigkeit der Ausgangszeilen (aus A).

Man kann den Rang der Matrix A also direkt am Endtableau des Gauß-Verfahrens ablesen:

Ist r die Anzahl der von Null verschiedenen Zeilen von \bar{A} im Endtableau des Gauß-Verfahrens, dann gilt:

$$\mathbf{Rg}(A) = r.$$

Lösungstheorie mittels Rangbegriff

Betrachtet man nun die erweiterte Koeffizientenmatrix $(\bar{A} | \vec{\bar{b}})$, so unterscheidet sich deren Rang von $\text{Rg}(\bar{A})$ genau dann, wenn $r < m$ und mindestens ein $\bar{b}_i \neq 0$ mit $r + 1 \leq i \leq m$ existiert, das System also unlösbar ist. Da aber $\text{Rg}(A) = \text{Rg}(\bar{A})$ und $\text{Rg}((A | \vec{b})) = \text{Rg}((\bar{A} | \vec{\bar{b}}))$ gilt, können wir festhalten:

Ein lineares (m, n) -Gleichungssystem $A\vec{x} = \vec{b}$ ist genau dann lösbar, wenn der Rang $r = \text{Rg}(A)$ der Koeffizientenmatrix A mit dem Rang der erweiterten Koeffizientenmatrix $(A | \vec{b})$ übereinstimmt, d.h. wenn gilt

$$\text{Rg}(A) = \text{Rg}((A | \vec{b})).$$

Die Lösung enthält dann $n - r$ freie Parameter.

Lösungsstruktur inhomogenes/zugehöriges homogenes System

Ein homogenes Gleichungssystem $A\vec{x} = \vec{0}$ besitzt wegen $A\vec{0} = \vec{0}$ stets die so genannte *triviale Lösung* $\vec{x} = \vec{0}$, ist also immer lösbar. Dieser Sachverhalt folgt übrigens auch aus der obigen Lösbarkeitsbedingung, es gilt nämlich

$$\operatorname{Rg}(A) = \operatorname{Rg}\left((A|\vec{0})\right)$$

in jedem Fall.

Das zu einem inhomogenen (m, n) -System $A\vec{x} = \vec{b}$ mit $\operatorname{Rg}(A) = r$ gehörende homogene System $A\vec{x} = \vec{0}$ ist also stets lösbar: die Lösungsmenge $L(A, \vec{0}) \neq \emptyset$ enthält $n - r$ freie Parameter.

Wir nehmen nun an, dass $A\vec{x} = \vec{b}$ lösbar ist. Ist dann \vec{x}_{IH} eine beliebige spezielle Lösung des inhomogenen Systems und $\vec{x}_H \in L(A, \vec{0})$, so gilt:

$$A(\vec{x}_{IH} + \vec{x}_H) = A\vec{x}_{IH} + A\vec{x}_H = \vec{b} + \vec{0} = \vec{b}.$$

Lösungsstruktur inhomogenes/zugehöriges homogenes System

Es ist also $\vec{x}_{IH} + \vec{x}_H$ eine Lösung des inhomogenen Systems. Die Menge

$$\{\vec{x}_{IH} + \vec{x}_H \mid \vec{x}_H \in L(A, \vec{0})\}$$

hat aber ebenfalls $n - r$ freie Parameter, stellt also die gesamte Lösungsmenge des inhomogenen Systems dar.

Wir halten fest:

Die allgemeine Lösung eines lösbaren inhomogenen Gleichungssystems $A\vec{x} = \vec{b}$ erhält man durch Addition einer beliebigen speziellen Lösung \vec{x}_{IH} des inhomogenen Systems und der allgemeinen Lösung des zugehörigen homogenen Systems $A\vec{x} = \vec{0}$:

$$L(A, \vec{b}) = \vec{x}_{IH} + L(A, \vec{0}).$$

Beispiel

Zum inhomogenen Gleichungssystem des vorangegangenen Beispiels gehört die spezielle Lösung $\vec{x} = (2, 1, 1)^T$ (für $t = 1$).

Das zugehörige homogene System lässt sich mittels Gauß-Verfahren und analogen Zeilenumformungen lösen:

$$\begin{array}{l} (1) \quad 3 \quad -3 \quad 6 \quad | \quad 0 \\ (2) \quad 2 \quad 0 \quad 3 \quad | \quad 0 \\ (3) \quad 1 \quad 1 \quad 1 \quad | \quad 0 \end{array} \rightarrow \begin{array}{l} (1') \quad 3 \quad -3 \quad 6 \quad | \quad 0 \\ (2') \quad 0 \quad 2 \quad -1 \quad | \quad 0 \\ (3') \quad 0 \quad 2 \quad -1 \quad | \quad 0 \end{array} \rightarrow \begin{array}{l} (1'') \quad 3 \quad -3 \quad 6 \quad | \quad 0 \\ (2'') \quad 0 \quad 2 \quad -1 \quad | \quad 0 \\ (3'') \quad 0 \quad 0 \quad 0 \quad | \quad 0. \end{array}$$

„Rückwärtsauflösen“ liefert $L(A, \vec{0}) = \{t \cdot (-\frac{3}{2}, \frac{1}{2}, 1)^T \mid t \in \mathbb{R}\}$, falls man $x_3 = t$ setzt. Die triviale Lösung $\vec{0}$ ist für $t = 0$ dabei. Die allgemeine Lösung des inhomogenen Systems erhält man zu

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} + t \begin{pmatrix} -3/2 \\ 1/2 \\ 1 \end{pmatrix}, \quad t \in \mathbb{R}.$$

Beispiel — Fortsetzung

Die allgemeine Lösung des inhomogenen Systems erhält man zu

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} + t \begin{pmatrix} -3/2 \\ 1/2 \\ 1 \end{pmatrix}, \quad t \in \mathbb{R}.$$

Wählt man hier $t = -1$, so erhält man die spezielle Lösung $\vec{u} = (\frac{7}{2}, \frac{1}{2}, 0)^T$.

Die Lösungsmenge kann also auch in der Form

$$\vec{x} = \vec{u} + t \cdot \vec{v}$$

geschrieben werden.

Dies sind lediglich Darstellungen ein und derselben Geraden in Parameterform. Geometrisch entsprechen dem System drei Ebenen, die eine gemeinsame Schnittgerade besitzen.

Definition und Rechenregeln

Ist A eine reguläre (n, n) -Matrix, dann gilt per definitionem $\text{Rg}(A) = n$. Es stellt sich nun die Frage, ob es eine Matrix X gibt, für die gilt:

$$AX = I.$$

Bezeichnen wir mit \vec{e}_i die i -te Spalte der Einheitsmatrix $I = (\vec{e}_1, \dots, \vec{e}_n)$, dann sind wegen $\text{Rg}(A) = n$ die folgenden Gleichungssysteme *eindeutig lösbar*:

$$A\vec{x}_1 = \vec{e}_1, A\vec{x}_2 = \vec{e}_2, \dots, A\vec{x}_n = \vec{e}_n.$$

Wir können daher eine Matrix X definieren, deren Spalten den n *eindeutigen* Lösungen $\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n$ dieser Systeme entsprechen:

$$X := (\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n).$$

Offensichtlich gilt dann

$$\begin{aligned} AX &= (A\vec{x}_1, A\vec{x}_2, \dots, A\vec{x}_n) \\ &= (\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n) = I. \end{aligned}$$

Diese Konstruktion ist für jede beliebige reguläre Matrix möglich.

Inverse A^{-1} einer regulären Matrix A

Definition

Zu jeder regulären Matrix A existiert genau eine Matrix X , für die

$$AX = I$$

gilt. Man nennt X zu A invers oder die zu A inverse Matrix und schreibt $X = A^{-1}$. Es gilt damit stets

$$AA^{-1} = A^{-1}A = I.$$

Die Regularität ist dabei für die Existenz einer solchen Matrix notwendige Voraussetzung.

Rechenregeln für inverse Matrizen

Auch A^{-1} ist wieder regulär und es gelten folgende Rechenregeln, die wir nicht beweisen wollen:

Für den Umgang mit Inversen sind folgende Rechenregeln wichtig:

- $(A^{-1})^{-1} = A,$
- $(A^{-1})^T = (A^T)^{-1},$
- $(AB)^{-1} = B^{-1}A^{-1},$
- $(\lambda A)^{-1} = \frac{1}{\lambda}A^{-1} \ (\lambda \neq 0).$

Beispiel

Wir können jetzt die Lösung eines linearen (n, n) -Systems

$$A\vec{x} = \vec{b}$$

mit regulärer Matrix A mittels der Inversen berechnen. Das System ist nämlich äquivalent zu

$$A^{-1}A\vec{x} = A^{-1}\vec{b},$$

woraus wegen $A^{-1}A = I$ sofort folgt:

$$\vec{x} = A^{-1}\vec{b}.$$

Kennt man also die Inverse A^{-1} , so lässt sich die Lösung des Systems sofort angeben. Dies ist von Vorteil, wenn für verschiedene rechte Seiten \vec{b} Lösungen gesucht sind.

Bei einer rechten Seite bedenke man, dass die praktische Berechnung von A^{-1} wesentlich aufwendiger ist als die einmalige Durchführung des Gauß-Verfahrens.

Übung

Gegeben seien die regulären Matrizen A, B . Vereinfachen Sie den Ausdruck

$$\left(2AB^{-1}\right)^{-1} \left(B^{-1}A^T\right)^T$$

unter der Annahme, dass B symmetrisch ist, soweit wie möglich.

Lösung

Für den ersten Faktor gilt

$$\left(2AB^{-1}\right)^{-1} = \frac{1}{2} \left(B^{-1}\right)^{-1} A^{-1} = \frac{1}{2}BA^{-1}.$$

Der zweite Faktor vereinfacht sich wegen der Symmetrie von B zu

$$\begin{aligned} \left(B^{-1}A^T\right)^T &= \left(A^T\right)^T \left(B^{-1}\right)^T \\ &= A\left(B^T\right)^{-1} = AB^{-1}. \end{aligned}$$

Insgesamt ergibt sich also

$$\begin{aligned} \left(2AB^{-1}\right)^{-1} \left(B^{-1}A^T\right)^T &= \frac{1}{2}B\left(A^{-1}A\right)B^{-1} \\ &= \frac{1}{2}BB^{-1} \\ &= \frac{1}{2}I. \end{aligned}$$

Das Gauß-Jordan-Verfahren

Zur Bestimmung der Inversen gibt es ein numerisches Verfahren, das *Gauß-Jordan-Verfahren*. Dieses lässt sich am besten anhand eines Beispiels erläutern.

Beispiel

Wir geben uns nun die reguläre Matrix

$$A = \begin{pmatrix} 3 & -3 & 6 \\ 2 & 0 & 3 \\ 1 & 1 & 2 \end{pmatrix}$$

vor. Wie bisherige Überlegungen zeigen, müssen zur Bestimmung der Inversen die 3 Gleichungssysteme

$$A\vec{x}_i = \vec{e}_i, \quad i = 1, 2, 3$$

gelöst werden.

Am geringsten ist der Rechenaufwand, wenn man alle Systeme simultan mit dem Gauß-Verfahren löst.

Beispiel — Fortsetzung

Man schreibt in das Starttableau auf die rechte Seite alle drei Vektoren \vec{e}_i , also die Matrix I . Auf diese wendet man gleichzeitig die benötigten elementaren Umformungen an, um A auf *obere Dreiecksgestalt* zu bringen:

$$\begin{array}{l}
 (1) \quad 3 \quad -3 \quad 6 \quad | \quad 1 \quad 0 \quad 0 \\
 (2) \quad 2 \quad 0 \quad 3 \quad | \quad 0 \quad 1 \quad 0 \\
 (3) \quad 1 \quad 1 \quad 2 \quad | \quad 0 \quad 0 \quad 1
 \end{array}
 \longrightarrow
 \begin{array}{l}
 (1') \quad 3 \quad -3 \quad 6 \quad | \quad 1 \quad 0 \quad 0 \\
 (2') \quad 0 \quad 2 \quad -1 \quad | \quad -\frac{2}{3} \quad 1 \quad 0 \\
 (3') \quad 0 \quad 0 \quad 1 \quad | \quad \frac{1}{3} \quad -1 \quad 1
 \end{array}$$

Drei Lösungen effizient bestimmen: Obere Dreiecksmatrix auf der linken Tableauseite mittels elementarer Umformungen in die Einheitsmatrix überführen. Hierzu erzeugen wir — zunächst in der letzten Spalte der Dreiecksmatrix — oberhalb der Hauptdiagonalen Nullen ($z_{1''} = z_{1'} - 6z_{3'}$, $z_{2''} = z_{2'} + z_{3'}$), danach in der mittleren Spalte ($z_{1'''} = z_{1''} + \frac{3}{2}z_{2''}$).

Beispiel — Fortsetzung

$$\begin{array}{l}
 (1'') \quad 3 \quad -3 \quad 0 \quad | \quad -1 \quad 6 \quad -6 \\
 (2'') \quad 0 \quad 2 \quad 0 \quad | \quad -\frac{1}{3} \quad 0 \quad 1 \\
 (3'') \quad 0 \quad 0 \quad 1 \quad | \quad \frac{1}{3} \quad -1 \quad 1
 \end{array}
 \longrightarrow
 \begin{array}{l}
 3 \quad 0 \quad 0 \quad | \quad -\frac{3}{2} \quad 6 \quad -\frac{9}{2} \\
 0 \quad 2 \quad 0 \quad | \quad -\frac{1}{3} \quad 0 \quad 1 \\
 0 \quad 0 \quad 1 \quad | \quad \frac{1}{3} \quad -1 \quad 1
 \end{array}$$

Abschließend müssen wir lediglich alle Zeilen durch das entsprechende Diagonalelement (dies ist die dritte elementare Umformung!) dividieren und erhalten:

$$\begin{array}{l}
 1 \quad 0 \quad 0 \quad | \quad -\frac{1}{2} \quad 2 \quad -\frac{3}{2} \\
 0 \quad 1 \quad 0 \quad | \quad -\frac{1}{6} \quad 0 \quad \frac{1}{2} \\
 0 \quad 0 \quad 1 \quad | \quad \frac{1}{3} \quad -1 \quad 1
 \end{array}$$

Die Lösungen der drei Systeme können jetzt abgelesen werden:

Die 1. Spalte auf der rechten Seite ist \vec{x}_1 , die 2. Spalte ist \vec{x}_2 und die 3. Spalte entspricht \vec{x}_3 .

Beispiel — Fortsetzung

D.h. die zu A inverse Matrix A^{-1} ergibt sich zu

$$A^{-1} = \begin{pmatrix} -\frac{1}{2} & 2 & -\frac{3}{2} \\ -\frac{1}{6} & 0 & \frac{1}{2} \\ \frac{1}{3} & -1 & 1 \end{pmatrix}.$$

Um Rechenfehler auszuschließen, empfiehlt sich abschließend eine Probe: $AA^{-1} = I$.

Das Gauß-Jordan-Verfahren

Das Gauß-Jordan-Verfahren zur Bestimmung der Inversen A^{-1} einer regulären (n, n) -Matrix A lautet:

- a) Bilde ein Tableau, bestehend aus der Matrix A (linke Seite) und der Einheitsmatrix $I = I_n$ (rechte Seite).
- b) Führe A mittels Gauß-Verfahren in eine obere Dreiecksmatrix über.
- c) Wende auf beide Seiten elementare Umformungen (beginnend mit der letzten Spalte der linken Tableauseite) an, so dass aus der Dreiecksmatrix eine Diagonalmatrix wird.
- d) Dividiere alle Elemente jeder Zeile des Tableaus durch das entsprechende Diagonalelement, so dass aus der Diagonalmatrix die Einheitsmatrix wird.

Die rechte Tableauseite entspricht nun der gesuchten Inversen A^{-1} .

Übung

Welche Lösungen hat das System $A\vec{x} = \vec{b}$ mit der Matrix

$$A = \begin{pmatrix} 3 & -3 & 6 \\ 2 & 0 & 3 \\ 1 & 1 & 2 \end{pmatrix}$$

für die rechten Seiten

a) $\vec{b} = \vec{0}$,

b) $\vec{b} = (1, 1, 1)^T$?

Lösung

Da wir die Inverse A^{-1} im vorangegangenen Beispiel bereits zu

$$A^{-1} = \begin{pmatrix} -\frac{1}{2} & 2 & -\frac{3}{2} \\ -\frac{1}{6} & 0 & \frac{1}{2} \\ \frac{1}{3} & -1 & 1 \end{pmatrix}$$

berechnet haben, können wir die Lösung des Systems als $\vec{x} = A^{-1}\vec{b}$ schreiben. Damit ist:

a) $\vec{x} = A^{-1}\vec{0} = \vec{0}$.

Die triviale Lösung ist also die einzige Lösung des homogenen Systems.

b) $\vec{x} = A^{-1} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1/3 \\ 1/3 \end{pmatrix}$.

Lineare Abbildungen, Ausgangsvektor und Bildvektor

Lineare Abbildungen werden im Allgemeinen durch Matrizen beschrieben.

Wenn man eine Matrix mit einem Vektor multipliziert, so erhält man wiederum einen Vektor, der aber in den meisten Fällen auf den ersten Blick gar nichts mit dem Ausgangsvektor gemeinsam hat:

$$\begin{pmatrix} 1 & 4 \\ 1 & -2 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 9 \\ -3 \end{pmatrix}.$$

Lineare Abbildungen, Ausgangsvektor und Bildvektor

In anderen sehr speziellen Fällen ist der Bildvektor ein Vielfaches des Ausgangsvektors:

$$\begin{pmatrix} 1 & 4 \\ 1 & -2 \end{pmatrix} \cdot \begin{pmatrix} 4 \\ 1 \end{pmatrix} = \begin{pmatrix} 8 \\ 2 \end{pmatrix} = 2 \cdot \begin{pmatrix} 4 \\ 1 \end{pmatrix}.$$

Trivialerweise wird der Nullvektor unter einer linearen Abbildung immer auf sich selbst abgebildet:

$$\begin{pmatrix} 1 & 4 \\ 1 & -2 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

(Das Studium des Nullvektors ist also völlig uninteressant; wir werden ihn bei den folgenden Betrachtungen weglassen.)

Lineare Abbildungen, Ausgangsvektor und Bildvektor

Wir werden im Folgenden eine Methode vorstellen, wie man Vektoren identifiziert, die unter einer linearen Abbildung auf Vielfache von sich selbst überführt werden:

$$A \cdot \vec{x} = \lambda \cdot \vec{x}$$

$$A \cdot \vec{x} - \lambda \cdot \vec{x} = \vec{0}$$

$$A \cdot \vec{x} - \lambda I \cdot \vec{x} = \vec{0}$$

$$(A - \lambda I) \cdot \vec{x} = \vec{0}$$

$$\vec{x} \neq \vec{0} \Rightarrow \det(A - \lambda I) \stackrel{!}{=} 0$$

Gesucht sind also Vektoren $\vec{x} \neq \vec{0}$, die durch die lineare Abbildung/Matrix A auf das λ -fache ihrer selbst abgebildet werden.

Definitionen

Ein Vektor $\vec{x} \neq \vec{0}$, der bei Anwendung der Matrix A auf sein λ -faches übergeht, heißt Eigenvektor von A zum Eigenwert λ .

Eigenwerte λ sind dabei die Lösungen des so genannten charakteristischen Polynoms

$$\det(A - \lambda I) = 0.$$

Eigenvektoren \vec{x} zum Eigenwert λ sind die nicht-trivialen Lösungen des linearen Gleichungssystems

$$(A - \lambda I) \cdot \vec{x} = \vec{0}.$$

Bemerkungen

Das charakteristische Polynom ist bei Vorliegen einer (n, n) -Matrix A ein Polynom n -ten Grades. Nach dem Hauptsatz der Algebra hat ein Polynom n -ten Grades n (möglicherweise komplexe, evtl. auch zusammenfallende) Lösungen.

Hat man einen Eigenwert λ gefunden, so erhält man wegen $\det(A - \lambda I) = 0$ auch immer zumindest eine nicht-triviale Lösung des Gleichungssystems $(A - \lambda I) \cdot \vec{x} = \vec{0}$. Dies bedeutet, dass man zu jedem Eigenwert mindestens einen Eigenvektor findet.

Weitere Bemerkungen

Wenn man einen Eigenvektor \vec{x} zu einem Eigenwert λ der Matrix A gefunden hat (also $A \cdot \vec{x} = \lambda \cdot \vec{x}$), so sind selbstverständlich alle Vielfache dieses Eigenvektors $a \cdot \vec{x}$ ebenfalls Eigenvektoren zum Eigenwert λ der Matrix A wegen

$$A \cdot (a\vec{x}) = aA\vec{x} = a\lambda\vec{x} = \lambda \cdot (a\vec{x}).$$

Bei mehrfachen Eigenwerten kann es mehrere linear unabhängige Eigenvektoren zu einem Eigenwert geben oder auch nur einen einzigen.

Beispiel

Wir ermitteln die Eigenwerte zur Matrix

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

$$\begin{aligned} \det(A - \lambda I) &= \det \begin{pmatrix} 1 - \lambda & 4 \\ 1 & -2 - \lambda \end{pmatrix} \\ &= (1 - \lambda) \cdot (-2 - \lambda) - 4 \\ &= \lambda^2 + \lambda - 6 = 0 \end{aligned}$$

Es folgt:

$$\lambda^2 + \lambda - 6 = 0 \quad \Rightarrow \quad \lambda_1 = 2, \lambda_2 = -3.$$

Die Eigenwerte zur Matrix A sind also:

$$\lambda_1 = 2 \text{ und } \lambda_2 = -3.$$

Wir suchen zunächst die Eigenvektoren \vec{x} zum Eigenwert $\lambda_1 = 2$:

Wegen $(A - \lambda_1 I) \cdot \vec{x} = \vec{0}$ folgt:

$$\begin{pmatrix} 1 - 2 & 4 \\ 1 & -2 - 2 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

und damit das lineare Gleichungssystem:

$$-x_1 + 4x_2 = 0.$$

Lösungen dieses linearen Gleichungssystems sind:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = a \cdot \begin{pmatrix} 4 \\ 1 \end{pmatrix}, \quad a \in \mathbb{R}.$$

Die Probe liefert:

$$\begin{pmatrix} 1 & 4 \\ 1 & -2 \end{pmatrix} \cdot \begin{pmatrix} 4 \\ 1 \end{pmatrix} = \begin{pmatrix} 8 \\ 2 \end{pmatrix} = 2 \cdot \begin{pmatrix} 4 \\ 1 \end{pmatrix}.$$

Wir suchen nun die Eigenvektoren \vec{x} zum Eigenwert $\lambda_2 = -3$:

Wegen $(A - \lambda_2 I) \cdot \vec{x} = \vec{0}$ folgt:

$$\begin{pmatrix} 1 - (-3) & 4 \\ 1 & -2 - (-3) \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

und damit das lineare Gleichungssystem:

$$x_1 + x_2 = 0.$$

Lösungen dieses linearen Gleichungssystems sind:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = a \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad a \in \mathbb{R}.$$

Die Probe liefert:

$$\begin{pmatrix} 1 & 4 \\ 1 & -2 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -3 \\ 3 \end{pmatrix} = (-3) \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Insgesamt:

Die Matrix

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

besitzt zwei Eigenwerte: $\lambda_1 = 2$ und $\lambda_2 = -3$.

Zum Eigenwert $\lambda_1 = 2$ gehören die Eigenvektoren

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = a \cdot \begin{pmatrix} 4 \\ 1 \end{pmatrix}, \quad a \in \mathbb{R} \setminus \{0\}.$$

Zum Eigenwert $\lambda_2 = -3$ gehören die Eigenvektoren

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = a \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad a \in \mathbb{R} \setminus \{0\}.$$

Übung

Ermitteln Sie alle Eigenwerte und Eigenvektoren zur Matrix

$$A = \begin{pmatrix} 3 & 1 \\ -1 & 5 \end{pmatrix}.$$

Lösung

$$\begin{aligned}\det(A - \lambda I) &= \det \begin{pmatrix} 3 - \lambda & 1 \\ -1 & 5 - \lambda \end{pmatrix} \\ &= (3 - \lambda) \cdot (5 - \lambda) + 1 \\ &= \lambda^2 - 8\lambda + 16 = (\lambda - 4)^2\end{aligned}$$

Es folgt: $(\lambda - 4)^2 \stackrel{!}{=} 0 \Rightarrow \lambda_1 = \lambda_2 = 4$

Die Eigenvektoren berechnen sich über:

$$\begin{pmatrix} 3 - 4 & 1 \\ -1 & 5 - 4 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow -x_1 + x_2 = 0$$

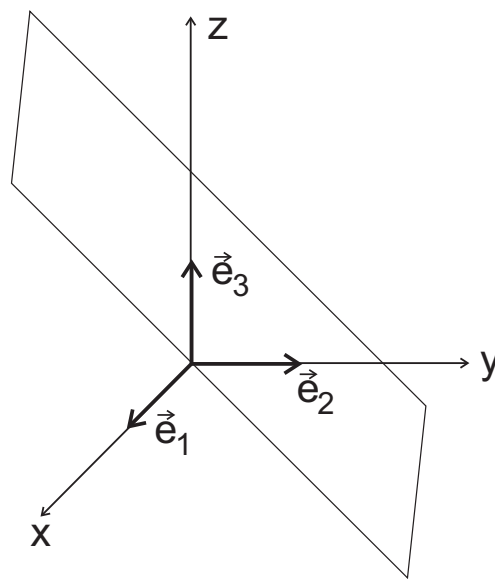
Eigenvektoren sind also:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = a \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad a \in \mathbb{R} \setminus \{0\}.$$

Zum doppelten Eigenwert $\lambda = 4$ gibt es also nur einen linear unabhängigen Eigenvektor $(x_1, x_2)^T = (1, 1)^T$ und seine Vielfachen.

Beispiele im \mathbb{R}^3

Wir betrachten ein Beispiel aus dem dreidimensionalen Raum, nämlich die Spiegelung an der Ebene $x = y$:



Diese lineare Abbildung wird durch die folgende Matrix wiedergegeben:

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Beachte: Durch die Spiegelung gehen die Einheitsvektoren in folgende Bildvektoren über: $\vec{e}_1 \mapsto \vec{e}_2$, $\vec{e}_2 \mapsto \vec{e}_1$, $\vec{e}_3 \mapsto \vec{e}_3$.

Es gibt auch bei der Spiegelung an der Ebene $x = y$ Vektoren, die auf Vielfache ihrer selbst abgebildet werden:

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = 1 \cdot \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} -1 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \\ 0 \end{pmatrix} = (-1) \cdot \begin{pmatrix} -1 \\ -1 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = 1 \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Auch hier liegen also Eigenwerte und Eigenvektoren vor, nämlich:

Eigenwert $\lambda = 1$: Eigenvektoren $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$

Eigenwert $\lambda = -1$: Eigenvektor $\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$

Die Rechnung erfolgt analog.

$$\det(A - \lambda I) = \det \begin{pmatrix} 0 - \lambda & 1 & 0 \\ 1 & 0 - \lambda & 0 \\ 0 & 0 & 1 - \lambda \end{pmatrix}$$

$$= (-\lambda) \cdot (-\lambda) \cdot (1 - \lambda) - 1 \cdot 1 \cdot (1 - \lambda)$$

$$= (1 - \lambda) \cdot (\lambda^2 - 1)$$

Es folgt:

$$\det(A - \lambda I) \stackrel{!}{=} 0 \quad \Rightarrow \quad \lambda_1 = \lambda_2 = 1, \lambda_3 = -1$$

Die Eigenvektoren zum Eigenwert $\lambda_1 = \lambda_2 = 1$ berechnen sich über:

$$\begin{pmatrix} -1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

und damit aus dem linearen Gleichungssystem:

$$-x_1 + x_2 = 0.$$

Lösungen sind hier die linear unabhängigen Eigenvektoren:

$$\vec{x}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \quad \text{bzw.} \quad \vec{x}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Die Eigenvektoren zum Eigenwert $\lambda_3 = -1$ berechnen sich über:

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

und damit aus dem linearen Gleichungssystem:

$$x_1 + x_2 = 0, \quad 2x_3 = 0.$$

Lösung ist hier der Eigenvektor:

$$\vec{x}_3 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$$

Übung

Ermitteln Sie alle Eigenwerte und Eigenvektoren zur Matrix

$$A = \begin{pmatrix} 2 & -2 & -2 \\ -2 & 5 & -1 \\ -2 & -1 & 5 \end{pmatrix}.$$

Lösung

Das charakteristische Polynom der Matrix

$$A = \begin{pmatrix} 2 & -2 & -2 \\ -2 & 5 & -1 \\ -2 & -1 & 5 \end{pmatrix}.$$

kann mit der Regel von Sarrus berechnet werden.

Es ergibt sich

$$\begin{aligned} \det(A - \lambda I) &= -\lambda^3 + 12\lambda^2 - 36\lambda \\ &= -\lambda \cdot (\lambda - 6)^2 = 0. \end{aligned}$$

Die Eigenwerte lauten $\lambda_1 = 0$ und $\lambda_2 = \lambda_3 = 6$.

Die Eigenvektoren zum Eigenwert $\lambda_1 = 0$ sind Lösungen des linearen Gleichungssystems

$$\begin{pmatrix} 2 & -2 & -2 \\ -2 & 5 & -1 \\ -2 & -1 & 5 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Um dieses lineare Gleichungssystem zu lösen, muss man zunächst die Matrix mittels Gauss-Algorithmus auf obere Dreiecksgestalt überführen. Man erhält:

$$\begin{array}{ccc} 2 & -2 & -2 \\ -2 & 5 & -1 \\ -2 & -1 & 5 \\ \hline 2 & -2 & -2 \\ 0 & 3 & -3 \\ 0 & -3 & 3 \\ \hline 2 & -2 & -2 \\ 0 & 3 & -3 \\ 0 & 0 & 0 \end{array}$$

Für die Eigenvektoren zum Eigenwert $\lambda = 0$ erhält man also das lineare Gleichungssystem:

$$\begin{aligned}x_1 - x_2 - x_3 &= 0 \\x_2 - x_3 &= 0.\end{aligned}$$

Lösung ist der Eigenvektor:

$$\vec{x}_1 = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}.$$

Die Eigenvektoren zum Eigenwert $\lambda_2 = \lambda_3 = 6$ sind Lösungen des linearen Gleichungssystems

$$\begin{pmatrix} -4 & -2 & -2 \\ -2 & -1 & -1 \\ -2 & -1 & -1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

d.h. von

$$-2x_1 - x_2 - x_3 = 0.$$

Lösungen sind hier die Eigenvektoren:

$$\vec{x}_2 = \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix} \quad \text{und} \quad \vec{x}_3 = \begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix}.$$

Beobachtung

Im letzten Beispiel und in der letzten Übung haben wir Eigenwerte und Eigenvektoren von symmetrischen Matrizen berechnet. Alle Eigenwerte waren reell. Gab es mehrfache Eigenwerte, so existierten auch entsprechend viele linear unabhängige Eigenvektoren.

Außerdem besaßen die Eigenvektoren eine interessante Eigenschaft: Eigenvektoren zu verschiedenen Eigenwerten stehen sogar senkrecht aufeinander.

Wir prüfen dies bei der letzten Übung nach:

$$(2, 1, 1) \cdot (-1, 2, 0)^T = 0$$

und

$$(2, 1, 1) \cdot (-1, 0, 2)^T = 0.$$

Dies ist kein Zufall!

Eigenwerte und Eigenvektoren einer symmetrischen Matrix

Für die Eigenwerte und Eigenvektoren einer symmetrischen (n, n) -Matrix gilt insbesondere:

- **Alle Eigenwerte sind reell.**
- **Es gibt insgesamt genau n linear unabhängige Eigenvektoren.**
- **Eigenvektoren, die zu verschiedenen Eigenvektoren gehören, stehen senkrecht aufeinander.**

Orthonormalbasis

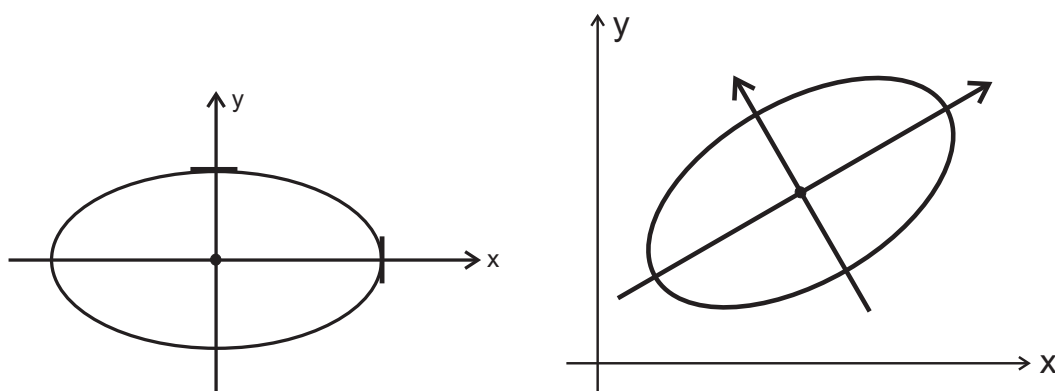
Linear unabhängige Eigenvektoren, die zum gleichen Eigenwert gehören, kann man „orthogonalisieren“ (also so wählen, dass sie orthogonal zueinander sind). Normiert man alle Eigenvektoren noch auf Länge 1, so erhält man insgesamt bei symmetrischen (n, n) -Matrizen eine so genannte Orthonormalbasis des \mathbb{R}^n bestehend aus Eigenvektoren.

Gekoppelte Schwingungen

In vielen physikalisch-technischen Anwendungen kommen symmetrische Matrizen vor. Ein Beispiel sind gekoppelte Schwingungen, die durch Systeme von Differentialgleichungen beschrieben werden. Für die auftretenden symmetrischen Matrizen lassen sich Eigenwerte und Eigenvektoren bestimmen, welche die so genannten Normalschwingungen des gekoppelten Systems beschreiben.

Hauptachsentransformation

Ein anderes Beispiel wäre die Hauptachsentransformation. Eine Ellipse um den Nullpunkt ist einfach zu beschreiben. Eine Ellipse irgendwo in der Ebene ist schwieriger zu beschreiben, es sei denn, man kennt die beiden orthogonalen Hauptachsen (Eigenvektoren der zugehörigen Matrix) und verwendet ein dieses angepasstes Koordinatensystem:



Der Hamming–Abstand

Die Länge $\|\vec{x}\|$ eines Vektors $\vec{x} \in \mathbb{R}^n$ (auch Norm genannt) und davon ausgehend der Abstand $\|\vec{x} - \vec{y}\|$ zweier Vektoren \vec{x} und \vec{y} wurden bereits eingeführt.

Als Norm hatten wir dabei die übliche „Euklidische Norm“ gewählt:

$$\|\vec{x}\| = \|(x_1, x_2, \dots, x_n)^T\| := \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}.$$

Die wichtigsten Rechenregeln für Normen lauten:

- a) $\|\vec{x}\| \geq 0$; $\|\vec{x}\| = 0$ genau dann, wenn $\vec{x} = \vec{0}$,
- b) $\|\lambda\vec{x}\| = |\lambda| \cdot \|\vec{x}\|$ für alle $\lambda \in \mathbb{R}$,
- c) $\|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|$.

Wenn man nun unter einer Norm einfach eine Vorschrift versteht, die jedem Vektor \vec{x} eine reelle Zahl $\|\vec{x}\|$ zuordnet, so dass die genannten drei Rechenregeln erfüllt sind, gibt es plötzlich auch noch weitere Kandidaten.

Normen und Metriken

Ein Beispiel wäre die (zugegebenermaßen etwas ungewöhnliche) Norm

$$\|\vec{x}\|_{\infty} := \max_{i=1,\dots,n} |x_i|.$$

Hier hätte etwa der Vektor $(-1, 0, 2)^T$ die uns zunächst gänzlich unvertraute „Länge“

$$\|(-1, 0, 2)^T\|_{\infty} = \max\{|-1|, |0|, |2|\} = 2$$

und *nicht* die für uns gebräuchliche Länge

$$\|(-1, 0, 2)^T\| = \sqrt{(-1)^2 + 0^2 + 2^2} = \sqrt{5}.$$

Ähnlich zur Norm kann man beim Abstand zweier Vektoren $\|\vec{x} - \vec{y}\|$, auch *Metrik*

$$d(\vec{x}, \vec{y})$$

genannt, vorgehen. Wiederum stellt man die wichtigsten Rechengesetze zusammen und fasst nun als Metrik *jedes* $d(\vec{x}, \vec{y})$ auf, welches sie erfüllt.

Metriken

Die Rechenregeln für Metriken (oder Abstände) lauten:

- a) $d(\vec{x}, \vec{y}) \geq 0$; $d(\vec{x}, \vec{y}) = 0 \iff \vec{x} = \vec{y}$,
- b) $d(\vec{x}, \vec{y}) = d(\vec{y}, \vec{x})$,
- c) $d(\vec{x}, \vec{y}) \leq d(\vec{x}, \vec{z}) + d(\vec{z}, \vec{y})$ für beliebiges \vec{z} .

Wir wollen nun eine spezielle Metrik aus der *Codierungstheorie*, den so genannten *Hamming–Abstand*, kennen lernen.

Bekanntlich kann es bei der Übertragung von Daten zu Fehlern aufgrund von (zufälligen) Störungen kommen.

Betrachten wir ein Beispiel: Nehmen wir an, ein Sender (Quelle) codiert eine Nachricht, die Nachricht wird übermittelt und dabei evtl. gestört. Der Empfänger muss nun die erhaltene Nachricht decodieren.

Codewörter

Nehmen wir gleichzeitig an, dass die zulässigen *Codewörter* aus allen 3-Tupeln über den Binärzahlen $\{0, 1\}$ bestehen, dass also die Codewörter

000, 001, 010, 011, 100, 101, 110 und 111 erlaubt sind. Wenn nun das Codewort 000 versandt wurde, aber aufgrund einer Störung beim Empfänger das Wort 001 ankam, so wird dieser *nicht* erkennen können, dass eine Störung vorliegt, denn 001 ist ja auch ein zulässiges Codewort.

Dies ist anders, wenn etwa nur die Teilmenge

000, 011, 101 und 110

obiger Codewörter zulässig wäre. Würde man nun das Signal 001 empfangen, so könnte man sofort erkennen, dass es *kein* zulässiges Codewort ist. Man könnte die empfangene Nachricht aber nicht korrigieren: Selbst wenn man davon ausgeht, dass nur ein einziges Bit gestört wurde, gibt es doch mehrere Möglichkeiten. Das Ausgangssignal könnte 000, 011 oder 101 sein, während bei 110 ganze drei Fehler gegenüber 001 aufgetreten wären.

Der Hamming–Abstand

Wir können hier ganz naheliegend einen Abstand (Metrik) zwischen Codewörtern definieren: Der Hamming-Abstand zweier Codewörter

$$\vec{x} = (x_1, x_2, \dots, x_n)^T \quad \text{und} \quad \vec{y} = (y_1, y_2, \dots, y_n)^T$$

über den Binärzahlen $\{0, 1\}$ ist die Anzahl der Stellen, an denen sich \vec{x} und \vec{y} unterscheiden:

$$d(\vec{x}, \vec{y}) := |\{i \mid x_i \neq y_i, i = 1, \dots, n\}|.$$

Z.B. gilt:

$$d(000, 011) = 2, \quad d(010, 011) = 1,$$

$$d(000, 111) = 3, \quad d(010, 010) = 0.$$

Beim ersten Code mit den acht Codewörtern betrug der Hamming–Abstand $d(\vec{x}, \vec{y})$ für alle $\vec{x} \neq \vec{y}$ immer 1. Dagegen haben die vier Codewörtern 000, 011, 101, 110 für $\vec{x} \neq \vec{y}$ immer den Hamming-Abstand $d(\vec{x}, \vec{y}) = 2$. Man erkennt hier zwar Fehler, kann sie aber nicht korrigieren.

Fehlererkennung und -korrektur

Die Fähigkeit zur Fehlerkorrektur wäre nur in Codes mit größerem Hamming-Abstand der Codewörter von einander der Fall. Dabei kommt es auf die *Minimaldistanz* des Codes an, d.h. auf das Minimum aller Hamming–Abstände zwischen Codewörtern.

In der Informatik zeigt man:

Ein Code ist *t-fehlererkennend* (er erkennt also, dass t Fehler bei der Übertragung aufgetreten sind), wenn die Minimaldistanz größer oder gleich $t + 1$ ist.

Ein Code ist *t-fehlerkorrigierend*, wenn die Minimaldistanz größer oder gleich $2t + 1$ ist.

In unserem Beispiel war der zweite Code bestehend aus den vier Codewörtern 1-fehlererkennend, da seine Minimaldistanz gleich 2 war. Allerdings war er nicht 1-fehlerkorrigierend, denn dazu bedarf es einer Minimaldistanz von mindestens $2 \cdot 1 + 1 = 3$.

Computer-Tomographie und lineare Gleichungssysteme

Im Jahre 1973 entstand eine Technik, welche die durch Organüberlagerungen verursachten Schwächen herkömmlicher Röntgenbilder beseitigte: die *Computer-Tomographie*, abgekürzt CT.

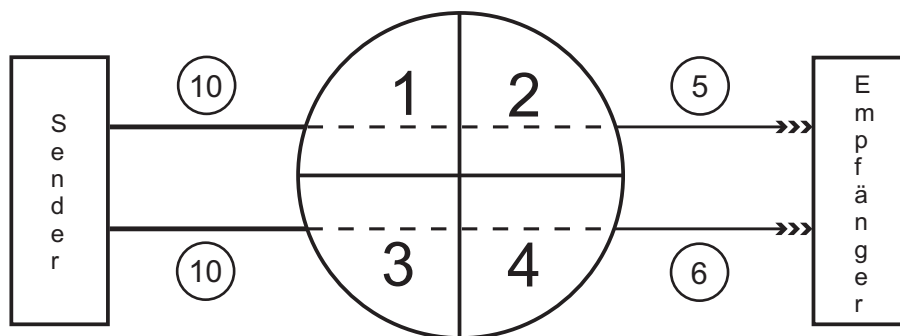
Die CT beruht auf Röntgenstrahlung. CT-Bilder geben einen Querschnitt durch den menschlichen Körper wieder. Die Bilder entstehen aufgrund von Messungen mit Hilfe von Computern.



Ein Röntgenstrahl wird von einem Sender ausgestrahlt und durchquert die vorgegebene Hirnschicht. Nach Verlassen des Körpers trifft er auf einen Strahlenempfänger, der misst, wie stark der Strahl jetzt noch ist.

Die Messvorrichtung von oben

Tatsächlich sendet die Strahlenquelle aber nicht nur einen Strahl, sondern viele parallele Strahlen aus. Der Empfänger misst demzufolge für jeden parallelen Strahl die Stärke seiner Abschwächung.



Wir gehen nun davon aus, dass alle Strahlen den Sender mit einer Stärke von 10 Einheiten verlassen. Die sukzessive Abschwächung des ersten Strahls lässt sich durch folgende Gleichung beschreiben:

$$10 - x_1 - x_2 = 5 \iff x_1 + x_2 = 5.$$

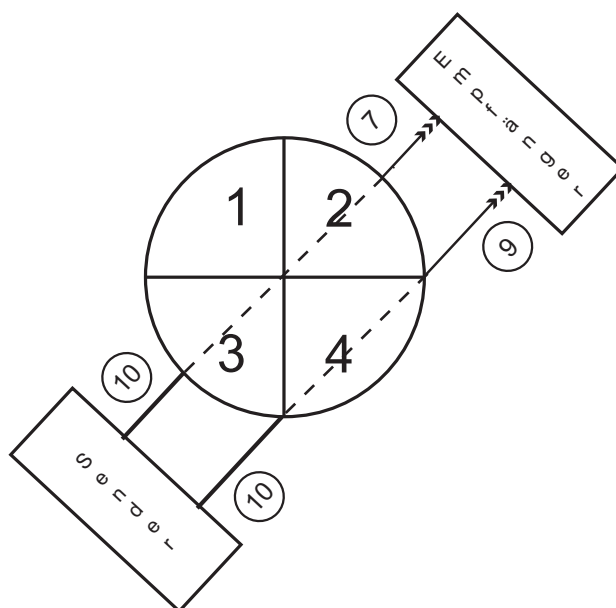
Analog wird der zweite Strahl beim Durchqueren der Hirnteile 3 und 4 um x_3 und x_4 Einheiten von 10 auf 6 Einheiten abgeschwächt. Daraus ergibt sich die Gleichung

$$10 - x_3 - x_4 = 6 \iff x_3 + x_4 = 4.$$

Erneute Messung durch Drehung

Die beiden Messungen ergeben ein Gleichungssystem mit 2 Gleichungen, aber 4 Unbekannten x_i , $i = 1, \dots, 4$. Die Unbekannten aus diesem System lassen sich nicht eindeutig bestimmen.

Deshalb wird die Messvorrichtung gedreht. Jetzt kann eine neue Messung durchgeführt werden:



Der obere Strahl führt nun auf die Gleichung

$$10 - x_3 - x_2 = 7 \iff x_2 + x_3 = 3.$$

Der untere Strahl liefert die Gleichung

$$10 - x_4 = 9 \iff x_4 = 1.$$

Das resultierende Gleichungssystem

Insgesamt ergibt sich damit das folgende Gleichungssystem:

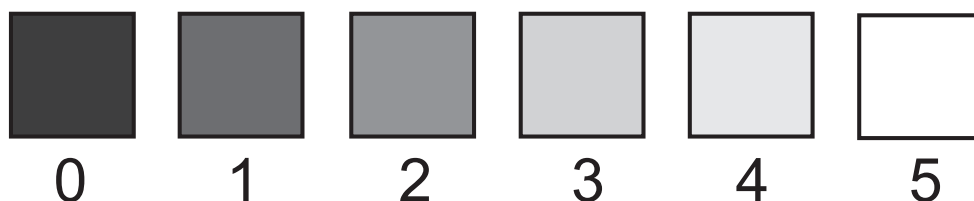
$$\begin{aligned}x_1 + x_2 &= 5 \\x_3 + x_4 &= 4 \\x_2 + x_3 &= 3 \\x_4 &= 1.\end{aligned}$$

Dieses System hat genau eine Lösung:

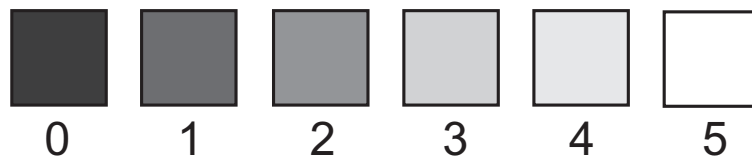
$$x_1 = 5, x_2 = 0, x_3 = 3 \quad \text{und} \quad x_4 = 1.$$

Diese besagt, dass Teil 1 den Strahl um 5 Einheiten, Teil 2 um 0 Einheiten usw. abschwächt.

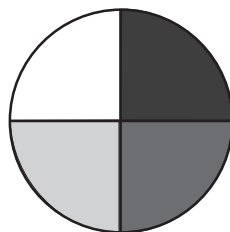
Zum CT-Bild kommt man jetzt, wenn man diese Lösungszahlen mittels einer *Grayoutskala* umsetzt:



Entstehen des CT-Bildes



Teil 1 unseres vereinfachten CT-Bildes wird also mit dem Grauton Nummer 5, Teil 2 mit dem Ton Nummer 0 etc. eingefärbt. So entsteht folgendes CT-Bild:



Um medizinisch verwertbare Bilder zu erhalten, müssen natürlich sehr viele Messungen durchgeführt werden. Richtige CT-Bilder bestehen aus Tausenden von Quadraten, die in unterschiedlichen Grautönen gefärbt sind. Jedem Quadrat entspricht eine Unbekannte x_i .

Die Konstruktion eines solchen Bildes bedingt daher die Lösung großer linearer Gleichungssysteme mit Tausenden von Gleichungen und Unbekannten.