Mathematik I

Vorlesung 10 - Matrizen

Prof. Dr. Sandra Eisenreich

07. Dezember 2023

Hochschule Landshut

Motivation

Matrizen sind die Art und Weise wie man mit linearen Abbildungen oder großen Datenmengen beim Programmieren umgeht.

Ihre Wichtigkeit in der Informatik lässt sich am einfachsten dadurch veranschaulichen, dass das einzige was GPUs von CPUs unterscheidet die Fähigkeit ist, besonders effektiv Matrizen zu multiplizieren.

- Computergrafik besteht quasi nur aus Matrizenmultiplikationen. GPUs machen nichts anderes, als effizienter Matrizen miteinander zu multiplizieren. (Wie das funktioniert lernen wir indirekt im Rest der Vorlesung).
- KI besteht aus vielen vielen Matrizen. Auch Daten werden in Matrizenform gespeichert.
- sonst auch überall...

2×2-Matrizen

(2×2)-Matrizen

Definition

- Für Elemente $a_{11}, a_{12}, a_{21}, a_{22}$ in einem Körper K nennen wir $A := \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ eine (2×2)-**Matrix** mit Einträgen in K. Die Menge aller 2×2 Matrizen mit Einträgen in K bezeichnen wir mit $K^{2\times 2}$.
- Wir definieren ein **Matrix-Vektor Produkt** zwischen A und $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in K^2$ als

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} a_{11} \cdot x_1 + a_{12} \cdot x_2 \\ a_{21} \cdot x_1 + a_{22} \cdot x_2 \end{pmatrix}$$

- Die Spalten der Matrix $\begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}$ und $\begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}$ in $\mathbb{R}^2$ nennt man **Spaltenvektoren**, die Zeilen (a_{11}, a_{12}) und (a_{21}, a_{22}) **Zeilenvektoren**. Der erste Index heißt **Zeilenindex**, der zweite **Spaltenindex**.

lineare Abbildung $\Rightarrow$ Matrix

Jede lineare Abbildung $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ kann man schreiben als:

$$f\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} a_{11} \cdot x_1 + a_{12} \cdot x_2 \\ a_{21} \cdot x_1 + a_{22} \cdot x_2 \end{pmatrix} = A \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \qquad \text{für Zahlen } a_{11}, a_{12}, a_{21}, a_{22} \in \mathbb{R}.$$

also als Multiplikation mit einer Matrix A wie oben.

Definition

Man bezeichnet die Matrix A zu einer linearen Abbildung mit $\operatorname{Mat} f$ und nennt diese **darstellende Matrix** von f.

Matrix $\Rightarrow$ lineare Abbildung

Umgekehrt: Jede (2×2)-Matrix $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ stellt eine lineare Abbildung f dar:

$$f_A\colon \mathbb{R}^2 \to \mathbb{R}^2, \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto A \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} a_{11} \cdot x_1 + a_{12} \cdot x_2 \\ a_{21} \cdot x_1 + a_{22} \cdot x_2 \end{pmatrix}$$

Definition

Man schreibt die von der Matrix A definierte lineare Abbildung als f_A.

Lineare Abbildungen = Matrizen

Satz/Definition

Es gibt eine eins-zu-eins Korrespondenz zwischen

$$\left\{ \begin{array}{c} \textit{linearen Abbildungen} \\ f\colon \mathbb{R}^2 \to \mathbb{R}^2 \end{array} \right\} \quad \textit{und} \quad \left\{ \begin{array}{c} (2\times 2) - \textit{Matrizen} \\ A \end{array} \right\}$$

$$\begin{array}{ccc} f & \mapsto & \operatorname{Mat} f \\ f_A & \leftarrow\!\shortmid & A \end{array}$$

Das heißt: für jede lineare Abbildung f gibt es genau eine **darstellende Matrix** $\operatorname{Mat} f$*, und jede Matrix $A \in \mathbb{R}^{2\times 2}$ definiert genau eine lineare Abbildung $f_A\colon \mathbb{R}^2 \to \mathbb{R}^2$. Die Spaltenvektoren der Matrix A sind dann genau die Bilder der Standardbasisvektoren $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ und $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ unter f_A.*

Beispiele → Mitschrift.

Motivation: geometrische Transformationen

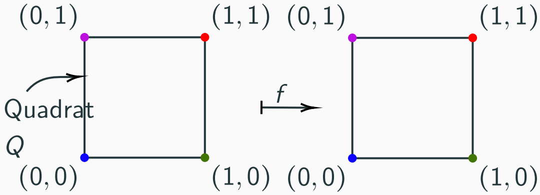
Wir wollten lineare Abbildugnen und Matrizen einführen um Drehungen oder andere geometrische Transformationen darstellen zu können. Ein Grafikprogramm führt in genau dieser Weise Skalierungen/Drehungen, etc. durch!

Wir betrachten nun eine Reihe von **geometrischen Transformationen in $\mathbb{R}^2$:**

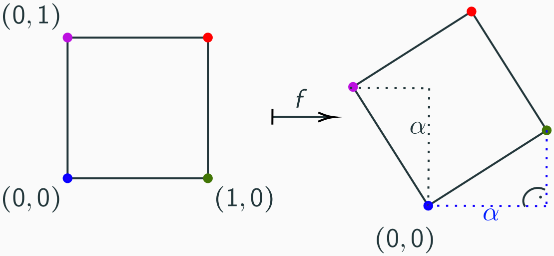
- die identische Abbildung, die alles "gleich lässt"
- Drehungen um einen Winkel α
- Streckungen um einen Faktor λ
- Spiegelungen an einer Achse
- Frage: ist eine Verschiebung auch eine lineare Abbildung?

Geometrische Transformationen: Identische Abbildung und Drehung

Identische Abbildung:

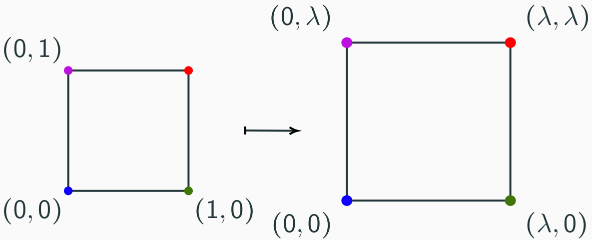


Drehung um α:

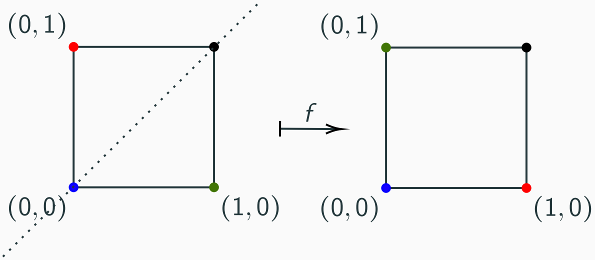


Streckung und Spiegelung

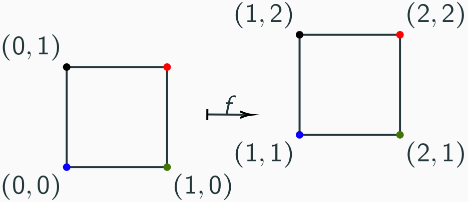
Streckung:



Spiegelung:



Verschiebung/Translation



Frage: Ist eine Verschiebung um $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ durch eine Matrix darstellerbar?

Antwort: ?

Hintereinanderausführung linearer Abbildungen im $\mathbb{R}^2$

Führt man zwei lineare Abbildungen hintereinander aus, z.B. eine Drehung f und dann eine Spiegelung g, so ist die resultierende Abbildung $g \circ f$ wieder linear.

Frage: Kann man die Matrix der Abbildung $g \circ f$ einfach aus den Matrizen der einzelnen Abbildungen g und f ausrechnen?

Wir betrachten zwei lineare Abbildungen $f : \mathbb{R}^2 \to \mathbb{R}^2$, und $g : \mathbb{R}^2 \to \mathbb{R}^2$ mit

$$A = \operatorname{Mat} f = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad B = \operatorname{Mat} f = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$

Wie sieht $\operatorname{Mat}(g \circ f)$ aus?

Matrizenprodukt

Definition

Seien $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in \mathbb{R}^{2\times 2}$ und $B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \in \mathbb{R}^{2\times 2}$ zwei Matrizen. Dann definiert man ihr **Matrizenprodukt** als

$$B \cdot A = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \cdot \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} := \begin{pmatrix} b_{11}a_{11} + b_{12}a_{21} & b_{11}a_{12} + b_{12}a_{22} \\ b_{21}a_{11} + b_{22}a_{21} & b_{21}a_{12} + a_{22}b_{22} \end{pmatrix}$$

Ist $A = \operatorname{Mat} f$ und $B = \operatorname{Mat} g$ für lineare Abbildungen f, g, so ist $B \cdot A = \operatorname{Mat}(g \circ f)$

Merkregel

Wie kann man sich das merken?

$$\begin{array}{cc} & \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \\ \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} & \begin{pmatrix} b_{11}a_{11} + b_{12}a_{21} & b_{11}a_{12} + b_{12}a_{22} \\ b_{21}a_{11} + b_{22}a_{21} & b_{21}a_{12} + b_{22}a_{22} \end{pmatrix} \end{array}$$

Um das Element im Eintrag ij der neuen Matrix $B \cdot A$ zu berechnen, werden nacheinander die Elemente der i-ten Zeile von B mit denen der j-ten Spalte von A multipliziert und dann aufaddiert.

Beispiele → Mitschrift

Definition

Für Elemente a_{ij} in einem Körper K ($i = 1, \ldots, m$ und $j = 1, \ldots, n$) nennen wir

$$A := \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

als $(m \times n)$-**Matrix** mit Einträgen in K. Die Menge aller $m \times n$ Matrizen mit Einträgen in K bezeichnen wir mit $K^{m \times n}$. Wir definieren ein Matrix-Vektor Produkt zwischen A und dem Vektor $(x_1, \ldots, x_n) \in K^n$ als

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11} \cdot x_1 + a_{12} \cdot x_2 + \ldots + a_{1n} \cdot x_n \\ \vdots \\ a_{m1} \cdot x_1 + a_{m2} \cdot x_2 + \ldots + a_{mn} \cdot x_n \end{pmatrix}$$

Zeilen und Spalten

Definition

- Der Vektor $\begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} \in K^m$ heißt j**-ter Spaltenvektor** von A.
- Der Vektor $\begin{pmatrix} a_{i1} \\ a_{i2} \\ \vdots \\ a_{in} \end{pmatrix} \in K^n$ heißt i**-ter Zeilenvektor** von A.
- für jeden Eintrag a_{ij} von A heißt i der **Zeilenindex** und j der **Spaltenindex**
- Man schreibt auch $A = (a_{ij})_{i=1,\ldots,m,j=1,\ldots,n}$

Wie im Fall von (2×2)-Matrizen liefert jede Matrix $A \in K^{m \times n}$ eine lineare Abbildung

$$f \colon K^m \to K^n; x \mapsto A \cdot x;$$

Umgekehrt hat jede lineare Abbildung $f \colon K^m \to K^n$ die Form

$$f\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11} \cdot x_1 + a_{12} \cdot x_2 + \ldots + a_{1n} \cdot x_n \\ \vdots \\ a_{m1} \cdot x_1 + a_{m2} \cdot x_2 + \ldots + a_{mn} \cdot x_n \end{pmatrix}$$

Satz/Definition

Es gibt eine eins-zu-eins Korrespondenz zwischen

$$\begin{array}{ccc} \left\{ \begin{array}{c} \textit{linearen Abbildungen} \\ f \colon K^m \to K^n \end{array} \right\} & \textit{und} & \left\{ \begin{array}{c} (m \times n) - \textit{Matrizen} \\ A \end{array} \right\} \\ f & \mapsto & \operatorname{Mat} f \\ f_A & \leftarrow\!\shortmid & A \end{array}$$

Bemerkungen

- Im Computer Algebra System Sage kann man Matrizen wie folgt eingeben
  $A = \text{matrix}\big[[1,2,3],[4,5,6]\big] \left(= \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \right)$
- Die Multiplikation zwischen einer Matrix $A = (a_{ij})_{ij}$ und einem Vektor ist ein komponentenweises Produkt der i-ten Zeiten von A und dem Vektor:

$$\text{z.B.} \quad \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 2 & -1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 \cdot x_1 + 2x_2 + 3x_3 + 4x_4 \\ 0 \cdot x_1 + 2x_2 + (-1) \cdot x_3 + 1 \cdot x_4 \\ 0 \cdot x_1 + 0 \cdot x_2 + 0x_3 + 1 \cdot x_4 \end{pmatrix} = \begin{pmatrix} x_1 + 2x_2 + 3x_3 + 4x_4 \\ 2x_2 - x_3 + x_4 \\ x_4 \end{pmatrix}$$

Matrizenmultiplikation

Satz

Seien $f : K^n \to K^m$ und $g : K^m \to K^k$ lineare Abbildungen und $A = Mat\ f = (a_{ij})_{ij}$ bzw $B = Mat\ g = (b_{rs})_{rs}$ die zugehörigen Matrizen. Dann gilt für die zu $g \circ f : K^n \to K^k$ zugehörige Matrix $C = Matg \circ f = (c_{ij})_{ij}$ (das **Produkt von** *B* **und** *A) dass:*

$$
\begin{aligned}
C = Mat(g \circ f) &= B \cdot A \\
&= \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1m} \\ \vdots & & & \\ b_{i1} & b_{i2} & \cdots & b_{im} \\ & \vdots & & \\ b_{k1} & \cdots & \cdots & b_{km} \end{pmatrix} \cdot \begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & & & \\ a_{21} & \cdots & a_{2j} & \cdots & a_{2n} \\ & \vdots & & & \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{pmatrix} \\
&:= \begin{array}{cc} \begin{pmatrix} \ddots & & \\ & b_{ii}a_{ij} + b_{i2}a_{2j} + \ldots + b_{im}b_{mj} & \\ & & \ddots \end{pmatrix} & \leftarrow \quad ite\ zeile \\ \uparrow & \\ j\text{-te spalte} & \end{array}
\end{aligned}
$$

Beispiele → Mitschrift

Komplexität von Matritzenmultiplikation

In der Informatik ist es immer wichtig, zu wissen, wie viele Rechenoperationen ein Algorithmus benötigt, um die Laufzeit abschätzen zu können. Gerade für Matrizenmultiplikationen (GPUs...) ist das wichtig.

Beachte: Will man zwei Matrizen aus $K^{n \times n}$ multiplizieren, so braucht man

- n^3 Additionen und
- n^3 Multiplikationen

in K. Das geht besser: Es gibt Verfahren, die das in etwa $n^{2.3\ldots}$ Operation können (schwierig!).

Offene Frage: Geht das besser, z.B. $n^2 \log n$ oder $n^{2+\varepsilon}$?

Weitere Beispiele → Mitschrift

Matrizen = K-Vektorraum

Wir betrachten noch weitere Operation auf Matrizen:

Definition

Seien $A = (a_{ij})$ und $B = (b_{ij})$ Matrizen in $K^{m \times n}$, und sei $\lambda \in K$. Dann definieren wir:

$A + B = (c_{ij}) := (a_{ij} + b_{ij})$ und $\lambda \cdot A = (d_{ij}) := (\lambda \cdot a_{ij})$

Beachte: Durch die oben eingeführte Addition und Skalarmultiplikation wird $K^{m \times n}$ zu einen K-Vektorraum der Dimension $m \cdot n$, d.h. $K^{m \times n} \cong K^{m \cdot n}$

Beispiele → Mitschrift

Rechenregeln für Matrizen

Satz

Seien A, B, C Matrizen passender Größe (d.h. die Operation müssen ausführbar sein), dann gilt:

- $A + B = B + A$
- $(A + B) \cdot C = A \cdot C + B \cdot C$
- $A \cdot (B + C) = A \cdot B + A \cdot C$
- $A \cdot (B \cdot C) = (A \cdot B) \cdot C$

(Im Allgemeinen ist aber $A \cdot B \neq B \cdot A$).

Matrizen = Ring, aber kein Körper

Die obigen Ergebnisse zeigen, dass die Menge $K^{n\times n}$ (also quadratische Matrizen) zusammen mit der Addition und Multiplikation einen Ring bilden. Die neutralen Elemente für die Addition bzw. Multiplikation sind hier

$$0 = \begin{pmatrix} 0 & 0\cdots & 0 \\ \vdots & & \\ 0 & \cdots & 0 \end{pmatrix} \in K^{n\times n} \text{ bzw. } E_n = \begin{pmatrix} 1 & \cdots & 0 \\ 0 & & \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix} \in K^{n\times n}$$

Das Inverse einer Matrix $A = (a_{ij})$ bzgl. der Addition ist einfach: $-A = (-a_{ij})$. für die Multiplikation kann man hingegen nicht immer ein Inverses bestimmen; wie das folgende Beispiel zeigt:

$$\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \cdot \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \Leftrightarrow \begin{matrix} b_{11} + 2b_{21} = 1 & b_{12} + 2b_{22} = 0 \\ 2b_{11} + 4b_{21} = 0 & 2b_{12} + 4b_{22} = 0 \end{matrix}$$

Die Gleichungen widersprechen sich! $\Rightarrow$ keine Lösung!

Invertierbare Matrizen

Abhilfe verschafft die Einschränkung auf alle Matrizen in $K^{n\times n}$, für die es ein Inverses bzgl. der Multiplikation gibt.

Definition

Für $n \in \mathbb{N}$ bezeichnen wir mit

$\mathrm{GL}(n) := \{A \in K^{n\times n} | \exists = A^{-1}K^{n\times n} \text{ mit } A^{-1} \cdot A = E_n\}$ ("general linear")

Die Menge aller Matrizen $A \in K^{n\times n}$, fur die ein Inverses $A^{-1} \in K^{n\times n}$ hinsichtlich der Multiplikation existiert. Wir nennen solche Matrizen **invertierbar**.

Satz

Die Menge $\mathrm{GL}(n)$ *mit der Matrizenmultiplikation ist eine Gruppe.*

Matrizen für beliebige Basen

Bisher: Die Spalten sind die Bilder der Standardbasisvektoren. Was, wenn wir andere Basen hernehmen ("vor und nach" der Abbildung)?

Betrachte die lineare Abbildung $f: \mathbb{R}^2 \to \mathbb{R}^2, \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} 2x - y \\ x + y \end{pmatrix}.$

$$\Rightarrow \mathrm{Mat}_{E_2,E_2}(f) = \begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix}$$

.

Wählt man hingegen die Basen $B := \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$ für den Vektorraum "vor" der Abbildung und $C := \left\{ \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$ für den Vektorraum "nach" der Abbildung, dann müssen wir genau wie zuvor erst die Bilder der Basisvektoren von B unter f berechnen, und diese dann ausdrücken als Linearkombination der Basis C. Die Vorfaktoren sind dann die Einträge der "neuen" Basis:

$\to$ Mitschrift $\Rightarrow \mathrm{Mat}_{B,C}(f) = \begin{pmatrix} 2 & 1 \\ -4 & 0 \end{pmatrix}$

Matrix bezüglich beliebiger Basen

Genauso kann man jede lineare Abb. $f : U \to V$ bezüglich fester Basen B und C von U bzw. V als eindeutige Matrix $(a_{ij})_{ij}$ schreiben: Der j-te Spaltenvektor von A entspricht genau dem Koordinatenvektor des Bildes $f(b_j)$ von b_j, d.h.

$$f(b_j) = \begin{pmatrix} a_{j1} \\ \vdots \\ a_{jn} \end{pmatrix}_C .$$

Wir schreiben für die f beschreibende Matrix A:

$$A = \mathrm{Mat}_{B,C}(f).$$

Beachte, dass die Matrix A nicht nur von f, sondern auch von den Basen B und C abhängt.