2 Matrizen und Systeme Linearer Gleichungen

2.1 Matrizenalgebra

Ein Hauptziel der Vorlesung zur Linearen Algebra besteht darin, Aussagen über die Lösungsmenge linearer Gleichungssysteme zu machen. Etwa ob das Gleichungssystem

$$\begin{array}{rcrcrcr} 2x & - & y & + & z & = & 1 \\ x & + & y & - & z & = & -1 \\ 3x & - & 3y & + & 3z & = & 1 \end{array}$$

eine Lösung hat. Dabei ist es natürlich unerheblich, wie wir die Variablen nennen, sie sind in gewisser Weise nur Platzhalter. Wesentlich ist nur das System der Koeffizienten auf der linken Seite der Gleichungen, und die Werte auf der rechten Seite. Solche Systeme von Koeffizienten und rechten Seiten schreiben wir als rechteckige Systeme von Zahlen, sogenannte Matrizen. Mit Matrizen lässt sich hervorragend rechnen und auf diese Weise die Lösungen des Gleichungssystemes gestimmen.

In diesem ganzen Kapitel 2.1 sei R ein kommutativer Ring mit Einselement 1.

Definition 2.1.1: *Seien m und n positive ganze Zahlen. Eine* **Matrix** *A mit m* **Zeilen** *und n* **Spalten**, *oder kurz eine $m \times n$ Matrix,* **über** *R ist ein rechteckiges Schema von Elementen von R der Form*

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} .$$

Das Element a_{ij} in der i-ten Zeile und j-ten Spalte von A nennen wir oft auch das (i, j)-te Element von A. [1]
Die Menge der $m \times n$ Matrizen über R bezeichnen wir mit $\mathrm{Mat}_{m\times n}(R)$ oder – wenn R aus dem Kontext klar ist – einfach mit $\mathrm{Mat}_{m\times n}$.
Elemente $a \in \mathrm{Mat}_{1\times n}$ oder $b \in \mathrm{Mat}_{m\times 1}$ heissen **Vektoren**. *a ist ein* **Zeilenvektor**, *b ist ein* **Spaltenvektor**.
$m \times n$ heisst die **Dimension** *der Matrix A.* □

Oft bezeichnen wir eine Matrix A wie in der Definition einfach als

$$(a_{ij})_{m\times n} .$$

Beispiel 2.1.2: Wir betrachten Matrizen über $\mathbb{R}$.
(i) Die 3×3 Matrix

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

[1] Streng formal könnten wir eine $m\times n$ Matrix definieren als eine Funktion $A : \{1, \ldots, m\}\times\{1, \ldots, n\} \to R$. Das Element a_{ij} ist dann einfach $A(i, j)$. Wir werden aber im folgenden dennoch die eher informalen Begriffe aus Definition 2.1.1 verwenden.

kann auch geschrieben werden als $A = (a_{ij})_{3\times 3}$ mit $a_{ij} = i^j$.
(ii) Die variable 3×3 Matrix

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

kann auch geschrieben werden als $B = (b_{ij})_{3\times 3}$ mit

$$b_{ij} = \begin{cases} x & \text{falls } i \leq j \\ 0 & \text{sonst.} \end{cases}$$

(iii) Die $n \times n$ Matrix

$$C = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ e & 1 & 0 & \cdots & 0 \\ e^2 & e & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ e^{n-1} & e^{n-2} & e^{n-3} & \cdots & 1 \end{pmatrix}$$

kann auch geschrieben werden als $C = (c_{ij})_{n\times n}$ mit

$$c_{ij} = \begin{cases} e^{i-j} & \text{falls } i \geq j \\ 0 & \text{sonst.} \end{cases}$$

(iv) Das **Kroneckersymbol** δ_{ij} (für $i, j \in \mathbb{N}$) ist definiert als

$$\delta_{ij} = \begin{cases} 1 & \text{falls } i = j \\ 0 & \text{sonst.} \end{cases}$$

Die Matrix

$$I_m \;=\; \left(\delta_{ij}\right)_{m\times m} \;=\; \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \;\in\; \mathrm{Mat}_{m\times m}$$

heisst die $(m \times m)$-**Identitäts-** oder **Einheitsmatrix**. □

Bisher ist $\mathrm{Mat}_{m\times n}(R)$ nur eine Menge. Um sie zu einer Algebra zu machen, definieren wir Operationen $+$ und $\cdot$ für Matrizen und auch eine Skalarmultiplikation mit Elementen aus R. Elemente auf dem zugrundeliegenden Körper (hier R) heissen auch **Skalare**.

Definition 2.1.3: *Seien $A = (a_{ij})$ und $B = (b_{ij})$ in $\mathrm{Mat}_{m\times n}(R)$. Die* **Summe** *$A + B$ der Matrizen A und B ist die Matrix $C = (c_{ij}) \in \mathrm{Mat}_{m\times n}(R)$ mit*

$$\forall i, j: \; c_{ij} = a_{ij} + b_{ij} \; .$$ □

Beispiel 2.1.4: In $\mathrm{Mat}_{2\times 2}(\mathbb{Q})$ gilt etwa

$$\begin{pmatrix} -1 & 0 \\ 2 & -2 \end{pmatrix} + \begin{pmatrix} 1 & 2 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 2 \\ 1 & -2 \end{pmatrix} \; .$$ □

Satz 2.1.5: *In $\mathrm{Mat}_{m\times n}(R)$ gilt:*

(i) + *ist kommutativ, also* $\forall A, B: \ A + B = B + A$.

(ii) + *ist assoziativ, also* $\forall A, B, C: \ A + (B + C) = (A + B) + C$. □

Beweis: Wegen der Kommutativität und Assoziativität von "+" in R. □

Wegen Satz 2.1.5(ii) schreiben wir einfach $A+B+C$ statt $A+(B+C)$ oder $(A+B)+C$.

Satz und Definition 2.1.6: *Sei* $N = (n_{i,j}) \in \mathrm{Mat}_{m\times n}$ *mit* $n_{ij} = 0$ *für alle* i, j. *Dann gilt für alle* $m \times n$ *Matrizen* A*:* $A + N = A$. N *ist die einzige Matrix in* $\mathrm{Mat}_{m\times n}$ *mit dieser Eigenschaft.* N *heisst die* $m \times n$ **Nullmatrix**, *auch geschrieben als* $0_{m\times n}$ *oder einfach* 0. □

Satz und Definition 2.1.7: *In* $\mathrm{Mat}_{m\times n}$ *gibt es für jede Matrix* $A = (a_{ij})$ *eine eindeutig bestimmte Matrix* $B = (b_{ij})$, *sodass* $A + B = 0_{m\times n}$. *Für diese Matrix* B *gilt:* $\forall i, j: \ b_{ij} = -a_{i,j}$. B *heisst die* **additive Inverse** *von* A, *und wir schreiben sie auch als* $-A$. □

Ebenso wie bei Zahlen schreiben wir meist $A-B$ statt $A+(-B)$. Die binäre Operation "$-$" heisst **Subtraktion**.

Definition 2.1.8: *Sei* $A = (a_{ij})_{m\times n}$ *eine Matrix über* R *und* $\lambda \in R$. *Das* **Skalarprodukt** *von* A *und* λ, λA, *ist definiert wie folgt:*

$$\lambda A := (\lambda a_{ij})_{m\times n} \ .$$ □

Satz 2.1.9: *Seien* $A, B \in \mathrm{Mat}_{m\times n}$ *und* $\lambda, \mu \in R$. *Dann gilt:*

(i) $\lambda(A + B) = \lambda A + \lambda B$,

(ii) $(\lambda + \mu)A = \lambda A + \mu A$,

(iii) $\lambda(\mu A) = (\lambda\mu)A$,

(iv) $(-1)A = -A$,

(v) $0A = 0_{m\times n}$. □

Definition 2.1.10: *Seien* $A = (a_{ij}) \in \mathrm{Mat}_{m\times n}$ *und* $B = (b_{jk}) \in \mathrm{Mat}_{n\times p}$. *Das* **Produkt** $A \cdot B$ *von* A *und* B *ist die Matrix* $C = (c_{ik}) \in \mathrm{Mat}_{m\times p}$ *mit*

$$\forall i, k: \ c_{ik} = \sum_{j=1}^{n} a_{ij} \cdot b_{jk} \ .$$ □

Die Operation $\cdot$ *heisst* **Multiplikation**.

Ebenso wie bei Zahlen schreiben wir ein Produkt $A \cdot B$ oft einfach als AB. Man beachte, dass Matrizen A und B nur dann multipliziert werden können, wenn A soviele Spalten hat wie B Zeilen hat. Insbesondere folgt aus der Multiplizierbarkeit von A und B nicht, dass auch B und A multipliziert werden können.

Beispiel 2.1.11: Seien A und B die folgenden Matrizen über $\mathbb{R}$:

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 2 & 3 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 0 \\ 1 & 2 \\ 1 & 1 \end{pmatrix} \ .$$

A und B können multipliziert werden (A hat soviele Spalten wie B Zeilen hat), und wir erhalten dabei die folgende 2×2 Matrix:

$$AB = \begin{pmatrix} 0 \cdot 2 + 1 \cdot 1 + 0 \cdot 1 & 0 \cdot 0 + 1 \cdot 2 + 0 \cdot 1 \\ 2 \cdot 2 + 3 \cdot 1 + 1 \cdot 1 & 2 \cdot 0 + 3 \cdot 2 + 1 \cdot 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 8 & 7 \end{pmatrix} .$$

In diesem speziellen Fall können wir auch B und A (in dieser Reihenfolge) multiplizieren, erhalten dabei aber die folgende 3×3 Matrix:

$$BA = \begin{pmatrix} 2 \cdot 0 + 0 \cdot 2 & 2 \cdot 1 + 0 \cdot 3 & 2 \cdot 0 + 0 \cdot 1 \\ 1 \cdot 0 + 2 \cdot 2 & 1 \cdot 1 + 2 \cdot 3 & 1 \cdot 0 + 2 \cdot 1 \\ 1 \cdot 0 + 1 \cdot 2 & 1 \cdot 1 + 1 \cdot 3 & 1 \cdot 0 + 1 \cdot 1 \end{pmatrix} = \begin{pmatrix} 0 & 2 & 0 \\ 4 & 7 & 2 \\ 2 & 4 & 1 \end{pmatrix} . \qquad \square$$

Beispiel 2.1.12: Für die Matrizen

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

erhalten wir die Produkte $AB = 0_{2\times 2}$ und $BA = A$. $\square$

Satz 2.1.13: *Die Multiplikation von Matrizen ist assoziativ, falls sie überhaupt definiert ist. Also für $m, n, p, q \in \mathbb{N}^*$ gilt:*

$$\forall A \in \mathrm{Mat}_{m\times n}, B \in \mathrm{Mat}_{n\times p}, C \in \mathrm{Mat}_{p\times q} : \; A(BC) = (AB)C \; . \qquad \square$$

Beweis:

$$A(BC)_{ij} = \sum_{k=1}^{n} a_{ik}(BC)_{kj} = \sum_{k=1}^{n} a_{ik}(\sum_{t=1}^{p} b_{kt}c_{tj}) = \sum_{k=1}^{n}\sum_{t=1}^{p} a_{ik}b_{kt}c_{tj} =$$

$$\sum_{t=1}^{p}(\sum_{k=1}^{n} a_{ik}b_{kt})\, c_{tj} = \sum_{t=1}^{p}(AB)_{it}\, c_{tj} = ((AB)C)_{ij}. \qquad \square$$

Wegen Satz 2.1.13 schreiben wir einfach ABC statt $A(BC)$ oder $(AB)C$. Weiters schreiben wir A^n statt $AA \cdots A$ (n mal). Dabei ist $A^0 = I$ (das neutrale Element bzgl. der Matrizenmultiplikation, wie wir gleich sehen werden).

Satz 2.1.14: *Die Matrizenmultiplikation ist distributiv bzgl. der Matrizenaddition. Also wenn die folgenden Summen und Produkte definiert sind, dann gilt*

$$A(B + C) = AB + AC, \quad (A + B)C = AC + BC \; . \qquad \square$$

Beweis:

$$[A(B+C)]_{ij} = \sum_{k=1}^{n} a_{ik}(B+C)_{kj} = \sum_{k=1}^{n} a_{ik}(b_{kj}+c_{kj}) = \sum_{k=1}^{n} a_{ik}b_{kj} + \sum_{k=1}^{n} a_{ik}c_{kj} = [AB+AC]_{ij} \; . \square$$

Die Skalarmultiplikation kann in ein Matrizenprodukt hineingezogen werden.

Satz 2.1.15: *Ist das Produkt AB definiert, dann gilt für alle $\lambda \in R$:*

$$\lambda(AB) = (\lambda A)B = A(\lambda B) \ . \qquad \square$$

Definition 2.1.16: *Eine Matrix A heisst* **quadratisch**, *wenn sie gleich viele Zeilen wie Spalten hat, also wenn $A \in \mathrm{Mat}_{n\times n}$ für ein $n \in \mathbb{N}^*$.* $\square$

Satz und Definition 2.1.17: *Sei $I_n = (\delta_{i,j}) \in \mathrm{Mat}_{n\times n}$ wie in Beispiel 2.1.2(iv). Dann gilt für alle $n \times n$ Matrizen A: $A \cdot I_n = I_n \cdot A = A$. I_n ist die einzige Matrix in $\mathrm{Mat}_{n\times n}$ mit dieser Eigenschaft. I_n heisst die $(n \times n)$-***Einheitsmatrix** *oder* **Identitätsmatrix**. *Ist n aus dem Kontext klar, so schreiben wir statt I_n auch einfach I.* $\square$

Aus den obigen Sätzen können wir sehen, dass $\mathrm{Mat}_{n\times n}(R)$ ein Ring mit Einselement ist. Dieser Ring ist aber nicht kommutativ und er enthält Nullteiler (vgl. Beispiel 2.1.12).

Satz 2.1.18: *Die Menge der quadratischen Matrizen $\mathrm{Mat}_{n\times n}$ mit den Operationen "+" und "·" bildet einen Ring. Dieser Ring ist nicht kommutativ und enthält Nullteiler.*

Definition 2.1.19: *Wir sagen, dass die beiden Matrizen A und B* **kommutieren**, *wenn gilt $AB = BA$.*

Beispiel 2.1.20: Nicht nur ist $\mathrm{Mat}_{n\times n}$ nicht-kommutativ und enthält Nullteiler. Es treten auch noch andere (etwa in Zahlenbereichen) ungewöhnliche Eigenschaften auf. Ist etwa R ein Körper, so gilt für jeden Skalar $\lambda \in R^*$

$$\begin{pmatrix} 0 & \lambda \\ \lambda^{-1} & 0 \end{pmatrix} \begin{pmatrix} 0 & \lambda \\ \lambda^{-1} & 0 \end{pmatrix} = I_2 \ .$$

I_2 hat also unendlich viele Quadratwurzeln, falls R ein unendlicher Körper ist. $\square$

In I_n steht an jeder Position in der Diagonale 1, und an jeder Position ausserhalb der Diagonale 0. Solche speziellen Matrizen, welche nur in der Diagonale von 0 verschieden sind, spielen in der Linearen Algebra eine wichtige Rolle.

Definition 2.1.21: *Eine quadratische Matrix $D = (d_{ij})_{n\times n}$ heisst* **Diagonalmatrix**, *wenn $d_{ij} = 0$ für alle $i \neq j$.* $\square$

Definition 2.1.22: *Ist A eine $m \times n$ Matrix, dann ist die* **Transponierte** *A^T von A die $n \times m$ Matrix, deren (i, j)-tes Element das (j, i)-te Element von A ist. Ist also*

$$A = (a_{ij})_{m\times n} \ ,$$

dann ist

$$A^T = (a_{ji})_{n\times m} \ . \qquad \square$$

Satz 2.1.23: *Wenn die jeweiligen Summen und Produkte definiert sind, dann gilt*

(i) $(A^T)^T = A, \quad (A + B)^T = A^T + B^T, \quad (\lambda A)^T = \lambda A^T, \quad (AB)^T = B^T A^T$

(ii) $(A(B + C))^T = B^T A^T + C^T A^T$

(iii) $(A^n)^T = (A^T)^n$ *für alle* $n \in \mathbb{N}$.

Beweis: ad (i): das einzig Interessante ist $(AB)^T = B^T A^T$. Sei $A = (a_{ij})_{m\times n}, B = (b_{kl})_{n\times\ell}$.

$$(AB)_{rs} = \sum_{t=1}^{n} a_{rt} b_{ts}, \quad \text{also} \quad [(AB)^T]_{uv} = \sum_{t=1}^{n} a_{vt} b_{tu} \ .$$

$$[B^T A^T]_{uv} = \sum_{t=1}^{n} (B^T)_{ut} (A^T)_{tv} = \sum_{t=1}^{n} b_{tu} a_{vt} \ .$$

ad (ii) und (iii): folgt inmittelbar aus (i). □

Definition 2.1.24: *Eine quadratische Matrix A ist* **symmetrisch**, *wenn $A = A^T$.*
Eine quadratische Matrix A ist **schiefsymmetrisch**, *wenn $A = -A^T$.*

Satz 2.1.25: *Für jede quadratische Matrix A gilt:*

(i) $A + A^T$ ist symmetrisch,

(ii) $A - A^T$ ist schiefsymmetrisch,

(iii) A kann auf eindeutige Weise ausgedrückt werden als Summe einer symmetrischen und einer schiefsymmetrischen Matrix.

Beweis: (i) $(A + A^T)^T = A^T + A$ wegen Satz 2.1.23.
(iii) $A = \frac{1}{2}(A + A^T) + \frac{1}{2}(A - A^T)$.
Zur Eindeutigkeit: Seien $U, \tilde{U}$ symmetrisch, $V, \tilde{V}$ schiefsymmetrisch, mit

$$U + V \ = \ A \ = \ \tilde{U} + \tilde{V} \ .$$

Dann ist $U - \tilde{U}$ symmetrisch und $\tilde{V} - V$ schiefsymmetrisch, und es gilt $U - \tilde{U} = \tilde{V} - V$. Das geht nur für die Nullmatrix. □

Rotationsmatrizen

Wir betrachten die Rotation des Koordinatensystems um den Winkel ϑ gegen den Uhrzeigersinn. Sei $P = (x, y)$ ein Punkt im ersten Quadranten der Ebene. Sei r der Abstand des Punktes P vom Ursprung, und sei α der von der Verbindungslinie $\overline{OP}$ und der x-Achse eingeschlossene Winkel.

Wir berechnen die Koordinaten (x', y') des Punktes P im neuen (um den Winkel ϑ gedrehten) Koordinatensystem. Es gilt

$$x = r\cos\alpha, \quad y = r\sin\alpha \ .$$

Im neuen Koordinatensystem haben wir

$$\begin{aligned} x' &= r\cos(\alpha - \vartheta) = r\cos\alpha\cos\vartheta + r\sin\alpha\sin\vartheta &&= x\cos\vartheta + y\sin\vartheta, \\ y' &= r\sin(\alpha - \vartheta) = r\sin\alpha\cos\vartheta - r\cos\alpha\sin\vartheta &&= y\cos\vartheta - x\sin\vartheta. \end{aligned}$$

Diese Gleichungen, welche die Koordinaten x', y' im neuen Koordinatensystem ausdrücken in Termini der Koordinaten x, y des alten Koordinatensystems und des Winkels ϑ, können mittels Matrizenschreibweise geschrieben werden als

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\vartheta & \sin\vartheta \\ -\sin\vartheta & \cos\vartheta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \ .$$

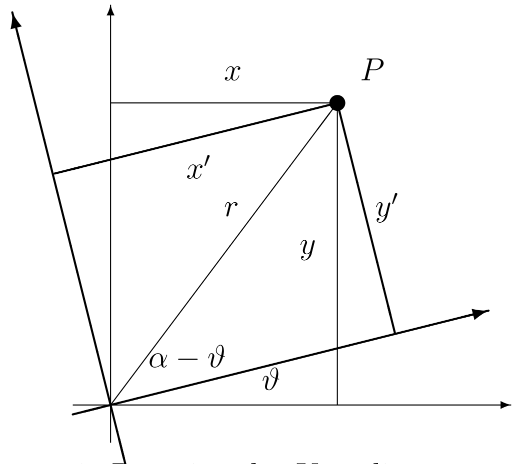

Figure 1: Rotation des Koordinatensystems um ϑ

Die 2×2 Matrix

$$R_\vartheta = \begin{pmatrix} \cos\vartheta & \sin\vartheta \\ -\sin\vartheta & \cos\vartheta \end{pmatrix}$$

heisst **Rotationsmatrix** um den Winkel ϑ. Sie hat die Eigenschaft

$$R_\vartheta R_\vartheta^T = I_2 = R_\vartheta^T R_\vartheta \ .$$

Damit ist sie ein Beispiel für eine sogenannte orthogonale Matrix.

Definition 2.1.26: *Eine $n \times n$ Matrix A heisst* **orthogonal**, *wenn*

$$AA^T = I_n = A^T A \ .$$

Nehmen wir nun an, dass wir zwei Rotationen hintereinander ausführen wollen. Zunächst werden die Koordinaten (x, y) transformiert in (x', y') durch Rotation um den Winkel ϑ, und anschliessend wird (x', y') transformiert in (x'', y'') durch Rotation um φ. Es ergibt sich also

$$\begin{aligned} \begin{pmatrix} x'' \\ y'' \end{pmatrix} &= \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} \\ &= \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix} \begin{pmatrix} \cos\vartheta & \sin\vartheta \\ -\sin\vartheta & \cos\vartheta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos(\varphi+\vartheta) & \sin(\varphi+\vartheta) \\ -\sin(\varphi+\vartheta) & \cos(\varphi+\vartheta) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \ . \end{aligned}$$

Zwischenrechnung Matrizenprodukt (letzte Gleichung):

$$\begin{pmatrix} \underbrace{\cos\varphi\cos\vartheta - \sin\varphi\sin\vartheta}_{\cos(\varphi+\vartheta)} & \underbrace{\cos\varphi\sin\vartheta + \cos\vartheta\sin\varphi}_{\sin(\varphi+\vartheta)} \\ \underbrace{-\sin\varphi\cos\vartheta - \cos\varphi\sin\vartheta}_{-\sin(\varphi+\vartheta)} & \underbrace{-\sin\varphi\sin\vartheta + \cos\vartheta\sin\varphi}_{\cos(\varphi+\vartheta)} \end{pmatrix}$$

Die Hintereinanderausführung von Rotationen kann also beschrieben werden durch das Produkt der betreffenden Rotationsmatrizen. Dabei spielt die Reihenfolge keine Rolle, Rotationsmatrizen kommutieren also wechselseitig, d.h.

$$R_\vartheta R_\varphi = R_{\vartheta+\varphi} = R_\varphi R_\vartheta \ ,$$

wie sich leicht mittels der trigonometrischen Additionstheoreme nachweisen lässt.

Beispiel 2.1.27: Wir betrachten etwa die durch die Gleichung

$$x^2 - y^2 = 1$$

beschriebene Hyperbel. Wenn wir diese Hyperbel um $\pi/4$ (also 45°) rotieren, wie sieht dann die Gleichung der rotierten Hyperbel aus? Dazu beobachten wir zunächst, dass die Rotation der Hyperbel um $\pi/4$ gleichbedeutend ist mit der Rotation des Koordinatensystems um $-\pi/4$, es gilt also

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = R_{-\pi/4} \begin{pmatrix} x \\ y \end{pmatrix} \ .$$

Wenn wir nun unter Ausnutzung der Beziehung

$$R_{\pi/4} R_{-\pi/4} = R_{\pi/4} R_{-\pi/4} = R_0 = I_2$$

die obige Gleichung mit $R_{\pi/4}$ multiplizieren, so erhalten wir

$$\begin{pmatrix} x \\ y \end{pmatrix} = R_{\pi/4} \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} \ .$$

also

$$x = \frac{1}{\sqrt{2}} x' + \frac{1}{\sqrt{2}} y' \ , \quad y = -\frac{1}{\sqrt{2}} x' + \frac{1}{\sqrt{2}} y' \ .$$

Die Gleichung $x^2 - y^2 = 1$ wird also transformiert in die neue Gleichung

$$\left(\frac{1}{\sqrt{2}} x' + \frac{1}{\sqrt{2}} y' \right)^2 - \left(-\frac{1}{\sqrt{2}} x' + \frac{1}{\sqrt{2}} y' \right)^2 = 1 \ ,$$

also in

$$2x'y' = 1 \ . \qquad \square$$

2.2 Systeme linearer Gleichungen

Wir wenden uns nun dem Problem der Lösung linearer Gleichungssysteme zu.

Beispiel 2.2.1: Wir betrachten etwa das folgende System linearer Gleichungen:

$$\begin{array}{rcrcrcrl} & & y & + & 2z & = & 1 & \quad (1) \\ x & - & 2y & + & z & = & 0 & \quad (2) \\ & & 3y & - & 4z & = & 23 & \quad (3) \end{array}$$

Multiplizieren wir die Gleichung (1) mit 3 und subtrahieren die neue Gleichung von (3), so erhalten wir

$$-10z = 20 \ .$$

Daraus sehen wir, dass $z = -2$ sein muss. Aus Gleichung (1) folgt damit $y = 5$ und weiters aus Gleichung (2) $x = 2y - z = 12$.

Definition 2.2.2: *Wir betrachten das folgende* **System von** m **linearen Gleichungen** *(***SLG***)* **in** n **Variablen über dem Körper** K*:*

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + \cdots + & a_{1n}x_n & = & b_1 \\ & & & & & \vdots & \\ a_{m1}x_1 & + & a_{m2}x_2 & + \cdots + & a_{mn}x_n & = & b_m \end{array}$$

wobei $a_{ij}, b_i \in K$ *für alle* i, j*. Wir schreiben dieses lineare Gleichungssystem auch als*

$$A \cdot x = b \ ,$$

wobei

$$A = (a_{ij})_{m \times n}, \quad x = (x_1, \ldots, x_n)^T, \quad b = (b_1, \ldots, b_m)^T \ .$$

Die $m \times n$ *Matrix* A *heisst die* **Koeffizientenmatrix** *des Systems, der Vektor* b *heisst die* **rechte Seite** *des Systems. Das Gleichungssystem heisst* **homogen***, wenn die rechte Seite der Nullvektor ist; ansonsten heisst das Gleichungssystem* **inhomogen***.*
Fügen wir zur Koeffizientenmatrix A *die rechte Seite* b *als weitere Spalte hinzu, so erhalten wir die* **erweiterte Matrix** $A|b \in \mathrm{Mat}_{m \times (n+1)}$ *des Gleichungssystems.*
Ein Vektor $a = (a_1, \ldots, a_n)^T \in \mathrm{Mat}_{n \times 1}(K)$ *heisst* **Lösungsvektor** *oder einfach* **Lösung** *dieses Gleichungssystems, wenn gilt*

$$Aa = b \ .$$

Ein Lösungsvektor heisst **nichttrivial***, wenn er verschieden ist vom Nullvektor.* □

Im folgenden wollen wir eine systematische Methode erarbeiten, um zu entscheiden, ob ein gegebenes Gleichungssystem überhaupt Lösungen hat, und gegebenenfalls alle Lösungen zu bestimmen. Diese Methode heisst das **Gaußsche Eliminationsverfahren**, benannt nach Carl Friedrich Gauß (1777–1855).

Definition 2.2.3: *Sei A eine Matrix. Jede der folgenden Operationen (und nur diese) ist eine* **elementare Zeilenoperation** *auf A:*

(1) Vertauschung zweier Zeilen;

(2) Multiplikation einer Zeile mit einem von 0 verschiedenen Skalar;

(3) Addition einer Zeile zu einer anderen Zeile.

Geht die Matrix B aus A hervor durch eine endliche Folge von elementaren Zeilenoperationen, so nennen wir B **zeilenäquivalent** *zu A.*
Analog kann man natürlich auch **elementare Spaltenoperationen** *definieren. Geht die Matrix B aus A hervor durch eine endliche Folge von elementaren Spaltenoperationen, so nennen wir B* **spaltenäquivalent** *zu A.* □

Elementare Zeilenoperationen können umgekehrt werden. Die Zeilenäquivalenz ist also eine Äquivalenzrelation auf Matrizen gleicher Dimension.

Satz 2.2.4: *Wenn die erweiterte Matrix $A'|b'$ aus $A|b$ durch eine elementare Zeilenoperation hervorgeht, so haben die linearen Gleichungssysteme*

$$Ax = b \quad \text{und} \quad A'x = b'$$

dieselben Lösungen.

Beweis: Seien $a_i = (a_{i1}, \ldots, a_{in})$, $1 \leq i \leq m$ die Zeilen von A, bzw. $a'_i = (a'_{i1}, \ldots, a'_{in})$, $1 \leq i \leq m$ die Zeilen von A'. Wegen der Symmetrie genügt es zu zeigen, dass jede Lösung von $Ax = b$ auch Lösung von $A'x = b'$ ist.
Fall (i): $a'_j = a_i, a'_i = a_j$, $b'_j = b_i, b'_i = b_j$ und $a'_k = a_k, b'_k = b_k$ für alle $k \neq i, j$. Offensichtlich ist jede Lösung von $Ax = b$ auch Lösung von $A'x = b'$.
Fall (ii): $a'_i = \lambda a_i, b'_i = \lambda b_i$ und $a'_k = a_k, b'_k = b_k$ für alle $k \neq i$. Offensichtlich ist jede Lösung von $Ax = b$ auch Lösung von $A'x = b'$.
Fall (iii): $a'_j|b'_j = a_j|b_j + a_i|b_i$ und $a'_k|b'_k = a_k|b_k$ für alle $k \neq j$. Offensichtlich ist jede Lösung von $Ax = b$ auch Lösung von $A'x = b'$. □

Die elementaren Zeilenoperationen haben eine Entsprechung in termini von Matrizenmultiplikation.

Definition 2.2.5: *Eine $n \times n$* **Elementarmatrix** *ist eine Matrix, welche aus I_n hervorgeht durch eine der elementaren Zeilenoperationen. Insbesondere bezeichnen wir (für fixes n)*

(1) mit $V_{i,j}$ die $n \times n$ Matrix, welche aus I_n hervorgeht durch Vertauschen der i-ten und der j-ten Zeile,

(2) mit $M_{i,\lambda}$ die $n \times n$ Matrix, welche aus I_n hervorgeht durch Multiplikation der i-ten Zeile mit dem von 0 verschiedenen Skalar λ,

(3) mit $A_{i,j}$ die $n \times n$ Matrix, welche aus I_n hervorgeht durch Addition der i-ten Zeile zur j-ten Zeile. □

Beispiel 2.2.6: Typische 3×3 Elementarmatrizen sind etwa

$$V_{1,2} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad M_{2,3} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad A_{1,3} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

Was passiert nun, wenn wir diese Elementarmatrizen von links etwa mit der Matrix

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

multiplizieren?

$$V_{1,2} \cdot A = \begin{pmatrix} 3 & 4 \\ 1 & 2 \\ 5 & 6 \end{pmatrix} \qquad \text{also die Zeilen 1 und 2 werden vertauscht}$$

$$M_{2,3} \cdot A = \begin{pmatrix} 1 & 2 \\ 9 & 12 \\ 5 & 6 \end{pmatrix} \qquad \text{also Zeile 2 wird mit 3 multipliziert}$$

$$A_{1,3} \cdot A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 6 & 8 \end{pmatrix} \qquad \text{also Zeile 1 wird zu Zeile 3 addiert}$$

Satz 2.2.7: *Sei A eine Matrix mit m Zeilen. Dann können die elementaren Zeilenoperationen auf A durch Multiplikation von links mit $m \times m$ Elementarmatrizen erreicht werden.*

(1) $V_{i,j}A$ ist die Matrix, welche aus A hervorgeht durch Vertauschen der i-ten und der j-ten Zeile.

(2) $M_{i,\lambda}A$ ist die Matrix, welche aus A hervorgeht durch Multiplikation der i-ten Zeile mit dem Skalar λ.

(3) $A_{i,j}A$ ist die Matrix, welche aus A hervorgeht durch Addition der i-ten Zeile zur j-ten Zeile.

Beweis:

ad (i): i-te Zeile von $V_{i,j} \cdot A$: $(0 \ldots 0\,1\,0 \ldots 0)A = (a_{j1} \ldots a_{jn})$
$\uparrow$ j-te Stelle

j-te Zeile von $V_{i,j} \cdot A$: $(0 \ldots 0\,1\,0 \ldots 0)A = (a_{i1} \ldots a_{in})$
$\uparrow$ i-te Stelle

k-te Zeile $(k \neq i, j)$ von $V_{i,j} \cdot A$: $(0 \ldots 0\,1\,0 \ldots 0)A = (a_{k1} \ldots a_{kn})$
$\uparrow$ k-te Stelle

ad (ii): i-te Zeile von $M_{i,\lambda} \cdot A$: $(0 \ldots 0\,\lambda\,0 \ldots 0)A = (\lambda a_{i1} \ldots \lambda a_{in})$
$\uparrow$ i-te Stelle

k-te Zeile $(k \neq i)$ von $M_{i,\lambda} \cdot A$: $(0 \ldots 0\,1\,0 \ldots 0)A = (a_{k1} \ldots a_{kn})$
$\uparrow$ k-te Stelle

ad (iii): für $i = j$ siehe (ii), $\lambda = 2$. Sonst

j-te Zeile von $A_{i,j} \cdot A$: $(0 \ldots 0\,1\,0 \ldots 0\,1\,0 \ldots 0)A = (a_{i1} + a_{j1} \ldots a_{in} + a_{jn})$
$\uparrow i \quad \uparrow j$

k-te Zeile $(k \neq j)$ von $A_{i,j} \cdot A$: $(0 \ldots 0\,1\,0 \ldots 0)A = (a_{k1} \ldots a_{kn})$
$\uparrow$ k-te Stelle $\square$

Satz 2.2.8: *Sei $A \in \mathrm{Mat}_{m\times n}$ und $b \in \mathrm{Mat}_{m\times 1}$. Sei E eine $m \times m$ Elementarmatrix, und sei $(A'|b') := E(A|b)$. Dann haben die Gleichungssysteme*

$$Ax = b \quad \text{und} \quad A'x = b'$$

dieselben Lösungen.

Beweis: wegen Satz 2.2.4 und Satz 2.2.7. □

Wir wollen nun mittels einer Folge von elementaren Zeilenumformungen (bzw. eine Folge von Multiplikationen mit Elementarmatrizen, vgl. Satz 2.2.8) ein gegebenes lineares Gleichungssystem

$$Ax = b$$

umformen (ohne dabei die Lösungsmenge zu ändern) in ein äquivalentes Gleichungssystem

$$Bx = c\ ,$$

sodass im transformierten System die Lösungen einfach abgelesen werden können.

Definition 2.2.9: *Unter einer* **Matrix in Zeilenstaffelform** *verstehen wir eine Matrix der Form*

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}\ ,$$

zu welcher es einen Index k gibt, $0 \le k \le m$, sodass

- $\forall l > k, 1 \le p \le n :\ a_{lp} = 0$
  (alle Zeilen nach der k-ten sind Nullzeilen);
- *für jeden Zeilenindex i, $1 \le i \le k$, gibt es einen Spaltenindex n_i, $1 \le n_i \le n$, sodass*

  $$a_{ij} = 0 \text{ für } j < n_i \text{ und } a_{in_i} \neq 0\ ;$$

- *die n_i nehmen strikt monoton zu, also*

  $$\forall 1 \le i < k :\ n_i < n_{i+1}\ .$$

Wir nennen die Elemente a_{in_i} die **Staffelelemente** *der Matrix.*

Beispiel 2.2.10: Die folgende Matrix ist in Zeilenstaffelform:

$$A = \begin{pmatrix} 1 & 1 & 3 & 0 & 2 & 9 & 0 & 1 \\ 0 & 0 & 0 & 4 & 0 & 3 & 5 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Der zugehörige Index k ist 3. Dabei sind $n_1 = 1, n_2 = 4, n_3 = 6$. Die Staffelelemente sind $a_{11} = 1, a_{24} = 4, a_{36} = 1$.
Die $m \times n$ Nullmatrix ist in Zeilenstaffelform. Der zugehörige Index k ist 0. □

Satz 2.2.11: *Mittels elementarer Zeilenoperationen kann jede Matrix A in eine Matrix in Zeilenstaffelform transformiert werden.*

Beweis: Ist A die Nullmatrix, so ist A offensichtlich in Zeilenstaffelform ($k = 0$).
Sei also nun A eine von der Nullmatrix verschiedene $m \times n$ Matrix. Wir betrachten die erste Spalte von A (von links her), welche nicht der Nullvektor ist. Durch eine geeignete Vertauschung von Zeilen erreichen wir, dass in dieser Spalte das erste Element verschieden von 0 ist. Wir erhalten also eine Matrix der Form

$$B = \begin{pmatrix} 0 & \cdots & 0 & b_{11} & b_{12} & \cdots & b_{1k} \\ 0 & \cdots & 0 & b_{21} & b_{22} & \cdots & b_{2k} \\ & \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & b_{m1} & b_{m2} & \cdots & b_{mk} \end{pmatrix} ,$$

in welcher gilt $b_{11} \neq 0$.
Nun subtrahieren wir, für $i = 2, 3, \ldots, n$, die 1. Zeile b_{i1}/b_{11} mal von der i-ten Zeile. Durch diese Kombination von Elementaroperationen wird die Matrix B transformiert in eine Matrix der Form

$$C = \begin{pmatrix} 0 & \cdots & 0 & b_{11} & b_{12} & \cdots & b_{1k} \\ 0 & \cdots & 0 & 0 & c_{22} & \cdots & c_{2k} \\ & \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & c_{m2} & \cdots & c_{mk} \end{pmatrix} .$$

Wenn wir nun mit der $(m-1) \times (k-1)$ Matrix (c_{ij}) rekursiv so verfahren wie mit der Ausgangsmatrix A, so erhalten wir schliesslich das gewünschte Ergebnis. □

Aus dem Beweis von Satz 2.2.11 können wir einen Algorithmus (ein Rechenverfahren) extrahieren, um aus einer beliebigen Matrix eine äquivalente Matrix in Zeilenstaffelform herzustellen. Dieses Verfahren heisst das **Gaußsche Eliminationsverfahren**. Es ist das wichtigste Verfahren zur Lösung linearer Gleichungssysteme.

Zu einer Matrix A gibt es i.a. nicht "die" äquivalente Matrix in Zeilenstaffelform. Das sieht man etwa daran, dass man im ersten Schritt des Gaußschen Eliminationsverfahrens i.a. auswählen kann, welche Zeile man zur 1. Zeile macht. Im folgenden führen wir mit der Hermite-Matrix eine Form ein, die eindeutig bestimmt ist.

Definition 2.2.12: *Eine $m \times n$ Matrix ist eine* **Hermite-Matrix** *bzw. ist eine* **Matrix in reduzierter Zeilenstaffelform**, *wenn*

- *A eine Matrix in Zeilenstaffelform ist,*
- *alle Staffelelemente (vgl. Def. 2.2.9) 1 sind, und*
- *alle Elemente oberhalb von Staffelelementen (in derselben Spalte) 0 sind.*

Beispiel 2.2.13: Durch elementare Zeilenoperationen können wir die Matrix A aus Beispiel 3.10 transformieren in die Hermite-Matrix

$$B = \begin{pmatrix} 1 & 1 & 3 & 0 & 2 & 0 & -9 & -8 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1/2 & -1/2 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} .$$

Satz 2.2.14: *Mittels elementarer Zeilenumformungen lässt sich jede Matrix A transformieren in eine äquivalente Hermite-Matrix.*

Beweis: Sei Z die zu A äquivalente Matrix in Zeilenstaffelform (nach Satz 2.2.11). Dividiere jede von der Nullzeile verschiedene Zeile in Z durch das Staffelelement dieser Zeile; somit sind dann alle Staffelelemente 1. Durch geeignete Subtraktion von Vielfachen dieser Zeilen eliminiert man alle Einträge oberhalb der Staffelelemente. □

Diese Hermite-Form ist genau das, was wir bei der Lösung von Systemen linearer Gleichungen brauchen. Um aber geeignet über diese Lösungen sprechen zu können, führen wir zunächst noch einige Begriffe ein.

Definition 2.2.15: *Seien $a_1 = (a_{11}, \ldots, a_{1n}),\ \ldots\ , a_m = (a_{m1}, \ldots, a_{mn})$ Zeilenvektoren, und seien $\lambda_1, \ldots, \lambda_m \in K$ Skalare. Dann ist der Zeilenvektor*

$$a = \sum_{i=1}^{m} \lambda_i a_i$$

eine **Linearkombination** *von $a_1, \ldots, a_m$.*
Auf analoge Weise definieren wir Linearkombinationen von Spaltenvektoren.

Beispiel 2.2.16: Der Zeilenvektor

$$(2, -3, 1) = 2(1, 0, 0) - 3(0, 1, 0) + 1(0, 0, 1)$$

ist eine Linearkombination der Zeilen der Einheitsmatrix I_3. Offensichtlich ist jeder Zeilenvektor der Länge 3 eine Linearkombination der Zeilen von I_3.
Ebenso ist jeder Spaltenvektor der Länge 3 eine Linearkombination der Spalten von I_3. □

Definition 2.2.17: *Die Zeilenvektoren $a_1, \ldots, a_m$ der Länge n heissen* **linear abhängig***, wenn es Skalare $\lambda_1, \ldots, \lambda_m \in K$ gibt, von denen mindestens einer von 0 verschieden ist, sodass sich der Nullvektor der Länge n mittels dieser Skalare als Linearkombination der a_i schreiben lässt; also*

$$0_n = \sum_{i=1}^{m} \lambda_i a_i \quad .$$

Ist das nicht der Fall, so heissen die Vektoren $a_1, \ldots, a_n$ **linear unabhängig***.*
Auf analoge Weise definieren wir lineare Abhängigkeit bzw. Unabhängigkeit für Spaltenvektoren.

Beispiel 2.2.18: Die Zeilen bzw. Spalten der Einheitsmatrix I_n sind linear unabhängig. Die Spalten der Matrix

$$\begin{pmatrix} 1 & 2 & 0 & 3 \\ 0 & 1 & 1 & 2 \\ 1 & 0 & 1 & 2 \end{pmatrix}$$

sind linear abhängig, da die 4. Spalte die Summe der ersten drei Spalten ist. Die ersten 3 Spalten aber sind linear unabhängig. □

Ist einer der Vektoren $a_i, 1 \leq i \leq p$, der Nullvektor, dann sind die a_i offensichtlich linear abhängig.
Es ist sinnvoll für eine Summe über einen leeren Indexbereich den Wert 0 festzusetzen

(weil 0 das neutrale Element bzgl. der Addition ist). So ist also der Nullvektor die Linearkombination von 0 Vektoren. Man beachte das im folgenden Satz.
Die leere Menge von Vektoren ist linear unabhängig. Für $I = \emptyset$ gibt es nämlich offensichtlich keinen Skalar $\lambda_i, i \in I$, sodass mindestens ein λ_i verschieden von 0 ist.

Satz 2.2.19: *Seien $a_1, \ldots, a_p$ Zeilenvektoren gleicher Länge, $p \geq 1$. Folgende Aussagen sind äquivalent (FAÄ):*

(i) Die Vektoren $a_1, \ldots, a_p$ sind linear abhängig.

(ii) Einer der Vektoren a_i kann ausgedrückt werden als Linearkombination der anderen.

(iii) Es gibt einen Vektor, der durch (mindestens) 2 verschiedene Linearkombinationen von $a_1, \ldots, a_p$ ausgedrückt werden kann.

Beweis: (i) $\Longrightarrow$ (ii): Seien nicht alle $\lambda_i, 1 \leq \lambda \leq p$, gleich 0, und

$$0 \;=\; \sum_{i=1}^{p} \lambda_i a_i \;.$$

O.B.d.A. sei $\lambda_1 \neq 0$. Dann gilt

$$a_1 \;=\; \sum_{i=2}^{p} (-\lambda_i/\lambda_1) a_i \;.$$

(ii) $\Longrightarrow$ (iii): o.B.d.A. sei

$$a_1 = \lambda_2 a_2 + \cdots + \lambda_p a_p \;.$$

Dann lässt sich der Nullvektor auf 2 verschiedene Weisen ausdrücken:

$$0a_1 + \cdots + 0a_p = 0 = -1a_1 + \lambda_2 a_2 + \cdots + \lambda_p a_p \;.$$

(iii) $\Longrightarrow$ (i): Ergeben zwei verschiedene Linearkombinationen denselben Vektor, also

$$\sum_{i=1}^{p} \lambda_i a_i \;=\; \sum_{i=1}^{p} \mu_i a_i \;,$$

so lässt sich der Nullvektor nichttrivial schreiben als

$$0 = \sum_{i=1}^{p} (\lambda_i - \mu_i) a_i \;,$$

die Vektoren $a_1, \ldots, a_p$ sind also linear abhängig. □

Korollar: *Die Zeilen (bzw. Spalten) einer Matrix sind linear abhängig, g.d.w. eine aus den anderen durch elementare Zeilenoperationen (Spaltenoperationen) erzeugt werden kann.*

Beweis: Jede Linearkombination der Zeilen (bzw. Spalten) entspricht einer Folge von elementaren Zeilenoperationen (Spaltenoperationen). □

Definition 2.2.20: *Der* **Zeilenrang** *einer Matrix A ist die maximale Anzahl linear unabhängiger Zeilen der Matrix A. Wir schreiben dafür* $\mathrm{rang}_z(A)$.

Beispiel 2.2.21: Der Zeilenrang der Nullmatrix $0_{m\times n}$ ist 0.
Der Zeilenrang der Identitätsmatrix I_n ist n.
Der Zeilenrang der Matrix

$$\begin{pmatrix} 1 & 2 & 0 & 0 \\ 2 & 1 & -1 & 1 \\ 5 & 4 & -2 & 2 \end{pmatrix}$$

ist 2. Die 3 Zeilen sind linear abhängig, aber die 1. und 3. Zeile sind linear unabhängig.
□

Satz 2.2.22: *Elementare Zeilenoperationen verändern den Zeilenrang einer Matrix nicht. Sind also A und B zeilenäquivalent, so gilt* $\mathrm{rang}_z(A) = \mathrm{rang}_z(B)$.

Beweis: Offensichtlich hat die Vertauschung zweier Zeilen keinen Einfluss auf die maximale Anzahl linear unabhängiger Zeilen.
Ist nun etwa die k-te Zeile eine Linearkombination von p anderen Zeilen, so ist auch das λ-fache ($\lambda \neq 0$) der k-ten Zeile so eine Linearkombination; und umgekehrt.
Schliesslich nehmen wir an, dass wir die i-te Zeile zur j-ten Zeile addieren, und so die neue j-te Zeile erhalten. Also

$$A_j^* \ = \ A_i + A_j \ .$$

Aus

$$\Big(\sum_{k=1, k\neq j}^{m} \lambda_k A_k \Big) \ + \ \lambda_j A_j^* \ = \ \Big(\sum_{k=1, k\neq i}^{m} \lambda_k A_k \Big) \ + \ (\lambda_i + \lambda_j) A_i$$

sehen wir, dass $A_1, \ldots, A_j, \ldots, A_p$ linear unabhängig sind g.d.w. $A_1, \ldots, A_j^*, \ldots, A_p$ linear unabhängig sind. Die Addition einer Zeile zu einer anderen Zeile hat also keinen Einfluss auf den Zeilenrang einer Matrix. □

Satz 2.2.23: *Jede Matrix kann mittels elementarer Zeilenoperationen in eine eindeutig bestimmte Hermite-Matrix transformiert werden.*
Zwei Matrizen A und B sind also zeilenäquivalent g.d.w. sie dieselbe zugehörige Hermite-Matrix haben.

Beweis: Es genügt zu zeigen:

"Sind A, B zwei zeilenäquivalente ($A \sim B$) Hermite-Matrizen, dann $A = B$"

Seien also $A \sim B$ zwei Hermite-Matrizen der Dimension $m \times n$.
Die Nullmatrix ist nur zu sich selbst zeilenäquivalent. Ist also A oder B die Nullmatrix, dann gilt $A = B$.
Für den Fall "$A, B \neq 0$" verfahren wir mittels Induktion über n:
$n = 1$: in diesem Fall gibt es nur eine Hermite-Matrix, nämlich

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

Induktionsannahme: "*Sind $A \sim B$ Hermite-Matrizen der Dimension $m \times (n-1)$, so sind $A = B$*"
Induktionsschritt $n-1 \to n$: Seien nun $A \sim B$ Hermite-Matrizen der Dimension $m \times n$, also

$$A = (a_{ij})_{m \times n}, \quad B = (b_{ij})_{m \times n} \ .$$

Für ein Produkt von Elementarmatrizen F gilt also

$$B = F \cdot A \ .$$

Seien $\hat{A}, \hat{B}$ die Matrizen, welche aus den ersten $n-1$ Spalten von A bzw. B bestehen,

$$\hat{A} = (a_{ij})_{m \times (n-1)}, \quad \hat{B} = (b_{ij})_{m \times (n-1)} \ .$$

$\hat{A} \sim \hat{B}$ sind Hermite-Matrizen mit $n-1$ Spalten, also wegen der Induktionshypothese gilt

$$\hat{A} = \hat{B} \ .$$

Es bleibt also zu zeigen, dass auch die n-ten Spalten von A und B übereinstimmen.
Der Zeilenrang einer Hermite-Matrix ist offenbar die Anzahl der von 0 verschiedenen Zeilen. Wegen $A \sim B$ ist $\operatorname{rang}_z(A) = \operatorname{rang}_z(B)$ (Satz 3.22), also A und B haben dieselbe Anzahl von Staffelelementen, etwa

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

Dann ist

$$\operatorname{rang}_z(\hat{A}) = \operatorname{rang}_z(\hat{B}) = \begin{cases} r-1 & (1) \\ \qquad \text{oder} & \\ r & (2) \ . \end{cases}$$

Im Fall (1) gibt es in den letzten Spalten von A und B nur eine 1,

$$\begin{pmatrix} a_{in} \\ \vdots \\ a_{rn} \\ \vdots \\ a_{mn} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} b_{in} \\ \vdots \\ b_{rn} \\ \vdots \\ b_{mn} \end{pmatrix} , \quad \text{also} \quad A = B \ .$$

Im Fall (2): wegen $B = FA$ lässt sich jede Zeile von B schreiben als Linearkombination von Zeilen von A:

$$(b_{i1} \dots b_{in}) \ = \ \sum_{k=1}^{r} \lambda_k (a_{k1} \dots a_{kn}) \quad \text{für} \ \ 1 \leq i \leq r \ .$$

Also gilt das auch für die Zeilen von $\hat{A} = \hat{B}$:

$$(a_{i1} \dots a_{i\,n-1}) = (b_{i1} \dots b_{i\,n-1}) \ = \ \sum_{k=1}^{r} \lambda_k (a_{k1} \dots a_{k\,n-1}) \quad \text{für} \ \ 1 \leq i \leq r \ .$$

Die ersten r Zeilen von $\hat{A}$ sind aber linear unabhängig, also

$$\lambda_i = 1 \quad \text{und} \quad \lambda_j = 0 \text{ für } j \neq i \ .$$

Somit ist $(b_{i1} \dots b_{in}) = (a_{i1} \dots a_{in})$, insbesondere $a_{in} = b_{in}$ für alle $1 \leq i \leq m$, also $A = B$.
□

Korollar: *Der Zeilenrang einer Matrix ist die Anzahl der vom Nullvektor verschiedenen Zeilen in einer Zeilenstaffelform der Matrix.*

Wir untersuchen nun den Effekt von elementaren Spaltenoperationen.

Satz 2.2.24: *Eine elementare Spaltenoperation auf einer $m \times n$ Matrix A kann erreicht werden durch Multiplikation von rechts mit einer geeigneten Elementarmatrix. Dabei multipliziert man genau mit derjenigen Elementarmatrix, welche dieser Spaltenoperation entspricht.*

Beweis: analog zu Satz 2.2.7. □

Definition 2.2.25: *Der* **Spaltenrang** *einer Matrix A ist die maximale Anzahl linear unabhängiger Spalten der Matrix A. Wir schreiben dafür $\operatorname{rang}_s(A)$.*

Satz 2.2.26: *Elementare Spaltenoperationen verändern den Spaltenrang einer Matrix nicht. Sind also A und B spaltenäquivalent, so gilt $\operatorname{rang}_s(A) = \operatorname{rang}_s(B)$.*

Beweis: Analog zu Satz 2.2.22. □

Satz 3.27: *Zeilen- und Spaltenrang einer Matrix A sind invariant sowohl unter elementaren Zeilen- als auch Spaltenoperationen.*

Beweis: Elementare Spaltenoperation haben keinen Einfluss auf die lineare Unabhängigkeit von Zeilen. Das ist offensichtlich für Vertauschung und Verlambdafachung von Spalten. Es gilt auch für die Addition einer j-ten Spalte zur k-ten Spalte:

$$\sum_{i=1}^{r} \lambda_i (a_{i1} \dots a_{ij} \dots a_{ik} \dots a_{in}) \;=\; 0 \quad \Longleftrightarrow \quad \sum_{i=1}^{r} \lambda_i (a_{i1} \dots a_{ij} \dots a_{ik} + a_{ij} \dots a_{in}) \;=\; 0 \; .$$

Ebenso haben elementare Zeilenoperationen keinen Einfluss auf die lineare Unabhängigkeit von Spalten. □

Satz 2.2.28: *Zeilenrang und Spaltenrang einer Matrix sind gleich. Für eine beliebige Matrix A gilt also:*

$$\operatorname{rang}_z(A) = \operatorname{rang}_s(A) \; .$$

Beweis: Sei A eine beliebige aber fix gewählte $m \times n$ Matrix.
Wir transformieren A mittels Zeilenoperationen zur Hermite-Matrix $H(A)$. Wegen Satz 3.27 gilt

$$\operatorname{rang}_z(A) = \operatorname{rang}_z(H(A)) =: r, \quad \operatorname{rang}_s(A) = \operatorname{rang}_s(H(A)) \; .$$

Mittels Spaltenoperationen transformieren wir $H(A)$ in eine Matrix der Form

$$\begin{pmatrix} I_r & * \\ 0_{m-r,r} & 0_{m-r,n-r} \end{pmatrix} ,$$

und weiter in

$$B := \begin{pmatrix} I_r & 0_{r,n-r} \\ 0_{m-r,r} & 0_{m-r,n-r} \end{pmatrix} \; .$$

Wir wissen

$$\operatorname{rang}_z(A) = \operatorname{rang}_z(B) = r \quad \text{und} \quad \operatorname{rang}_s(A) = \operatorname{rang}_s(B) = r \ .$$

Also $\operatorname{rang}_z(A) = \operatorname{rang}_s(A)$. □

Definition 2.2.29: *Aufgrund des Satzes 2.2.28 sprechen wir einfach vom* **Rang** *einer Matrix A, also von der maximalen Anzahl linear unabhängiger Zeilen bzw. Spalten von A.*

Korollar zu Satz 2.2.28: *Für jede Matrix A gilt:* $\operatorname{rang}(A) = \operatorname{rang}(A^T)$.

Aus dem Beweis des Satzes 2.2.28 sieht man, dass sich jede Matrix A in eine Matrix der Form

$$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$$

transformieren lässt mittels Multiplikation von links und rechts durch jeweils ein Produkt von Elementarmatrizen. Also

$$PAQ = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \ . \tag{1}$$

Dabei ist r der Rang der Matrix A. Die Matrizen P und Q in (1) können systematisch berechnet werden, indem man auf das Schema

$$\begin{matrix} I_n & \\ A & I_m \end{matrix}$$

elementare Zeilen- und Spaltenoperationen anwendet. Ist N die Normalform von A, so erhält man schliesslich

$$\begin{matrix} Q & \\ N & P \end{matrix} \ .$$

Die Matrizen P und Q sind jedoch im allgemeinen nicht eindeutig bestimmt.

Definition 2.2.30: *Die Matrix auf der rechten Seite der Gleichung (1) heisst die* **Normalform** *von A. Zwei $m \times n$ Matrizen A und B heissen* **äquivalent***, wenn sie dieselbe Normalform haben, bzw. wenn* $\operatorname{rang}(A) = \operatorname{rang}(B)$. □

Zwei Matrizen sind also äquivalent, wenn sie durch elementare Zeilen- und Spaltenumformungen ineinander transformiert werden können

Beispiel 2.2.31: Wir bestimmen die Normalform und die zugehörigen Transformationsmatrizen für die Matrix

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

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

1. und 2. Zeile vertauschen und eliminieren

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

2. Zeile mal $-1/3$ und eliminieren

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

2. und 3. Spalte vertauschen und eliminieren

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

1. Spalte von 3. Spalte subtrahieren

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

Also somit sehen wir

$$\frac{1}{3} \cdot \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & 0 \\ -3 & 3 & 3 \end{pmatrix} \cdot \begin{pmatrix} 2 & 2 & 1 \\ 1 & 1 & 2 \\ 1 & 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

und wir haben die Normalform von A bestimmt. □

Satz 2.2.32: *Sei A eine $m \times n$ Matrix, b ein Spaltenvektor der Länge m.*

(i) Das homogene Gleichungssystem

$$Ax = 0$$

besitzt eine nichttriviale Lösung g.d.w. $\mathrm{rang}(A) < n$.

(ii) Das inhomogene Gleichungssystem

$$Ax = b$$

besitzt eine Lösung g.d.w. $\mathrm{rang}(A) = \mathrm{rang}(A|b)$.

Beweis: (i): bezeichne a_i die i-te Spalte von A, also $A = (a_1 \dots a_n)$. $x = (x_1 \dots x_n)^T$ ist Lösung g.d.w. $x_1 a_1 + \dots + x_n a_n = 0$. Also es gibt eine nichttriviale Lösung g.d.w. die Spalten $a_1, \dots, a_n$ linear abhängig sind g.d.w. $\mathrm{rang}(A) < n$.
(ii): $Ax = b$ heisst $x_1 a_1 + \dots + x_n a_n = b$. Also es gibt eine Lösung g.d.w. b linear abhängig ist von den Spalten $a_1, \dots, a_n$ g.d.w. $\mathrm{rang}(A) = \mathrm{rang}(A|b)$. □

Definition 2.2.33: *Ein System linearer Gleichungen (SLG) $Ax = b$ heisst* **konsistent** *g.d.w. es eine Lösung hat. Sonst heisst es* **inkonsistent**.

Satz 2.2.34: *Sei $Ax = b$ ein konsistentes SLG von m Gleichungen in n Variablen. Sei $r = \mathrm{rang}(A) = \mathrm{rang}(A|b)$. Seien $n_i, 1 \le i \le r$ die Staffelindizes von $H(A)$ (vgl. Def. 3.9). Dann kann den $n-r$ Variablen x_j, mit $j \notin \{n_i | 1 \le i \le r\}$, ein beliebiger Wert zugeordnet werden, und das SLG in den übrigen Variablen gelöst werden.*

Beweis: Wir transformieren die erweiterte Matrix $A|b$ (oder im Fall eines homogenen Systems einfach A) mittels elementarer Zeilenoperationen in die äquivalente Hermite-Matrix. Wir erhalten also eine Matrix der Form $H(A)|c$, welche genau r von 0 verschiedene Zeilen hat, wenn $\mathrm{rang}(A) = r$. Das neue SLG

$$H(A)x = c$$

ist äquivalent zum Ausgangssystem, und es erlaubt uns offenbar, den $n - r$ Variablen x_j, mit $j \notin \{n_i | 1 \le i \le r\}$, einen beliebigen Wert zuzuordnen und diese Teillösung zu einer vollen Lösung zu erweitern. □

Beispiel 2.2.35: Wir betrachten das SLG

$$\begin{array}{rcrcrcl} x & + & y & + & z & = & 1 \\ 2x & - & y & + & 2z & = & 1 \\ x & + & 2y & + & z & = & \alpha \end{array}$$

und wollen untersuchen, für welche Werte von α es eine Lösung hat (also konsistent ist).

Dazu bestimmen wir zunächst den Rang der erweiterten Matrix:

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 2 & -1 & 2 & 1 \\ 1 & 2 & 1 & \alpha \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & -3 & 0 & -1 \\ 0 & 1 & 0 & \alpha - 1 \end{pmatrix}$$

$$\longrightarrow \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & \frac{1}{3} \\ 0 & 1 & 0 & \alpha - 1 \end{pmatrix}$$

$$\longrightarrow \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & \frac{1}{3} \\ 0 & 0 & 0 & \alpha - \frac{4}{3} \end{pmatrix} \quad .$$

Der Rang der Koeffizientenmatrix und der Rang der erweiterten Matrix sind also gleich g.d.w. $\alpha = \frac{4}{3}$. Ist das nicht der Fall, so gibt es keine Lösung und das SLG ist inkonsistent. Ist jedoch $\alpha = \frac{4}{3}$, so ist der Rang 2, und die Hermite-Form der erweiterten Matrix ist

$$\begin{pmatrix} 1 & 0 & 1 & \frac{2}{3} \\ 0 & 1 & 0 & \frac{1}{3} \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad .$$

Wir können also $3 - 2 = 1$ Variablen einen beliebigen Wert zuweisen (etwa der Variablen z) und das SLG

$$\begin{array}{ccccccc} x & & & + & z & = & \frac{2}{3} \\ & & y & & & = & \frac{1}{3} \end{array}$$

nach den übrigen Variablen lösen. Dabei erhalten wir die Lösung

$$(\frac{2}{3} - z,\ \frac{1}{3},\ z)^T$$

für beliebig gewähltes z. □

Satz 2.2.36: *Sei A eine $m \times n$ Matrix und b ein Spaltenvektor der Länge m.*

(i) Seien a und a' Lösungen des homogenen SLG $Ax = 0$. Dann ist auch $\lambda a + \mu a'$ eine Lösung, für beliebige $\lambda, \mu \in K$. (Ist der Grundkörper K unendlich, so gibt es also in diesem Fall unendlich viele Lösungen.)

(ii) Sei a Lösung des homogenen SLG, und sei c Lösung des inhomogenen SLG $Ax = b$. Dann ist auch $a + c$ Lösung des inhomogenen SLG.

(iii) Seien c, d Lösungen des inhomogenen SLG. Dann ist $c - d$ Lösung des homogenen SLG.

Beweis:
(i): $A(\lambda a + \mu a') = \lambda Aa + \mu Aa' = 0$.
(ii): $A(a + c) = Aa + Ac = 0 + b = b$.
(iii): $A(c - d) = Ac - Ad = b - b = 0$. □

Korollar: *Sei $A \cdot x = b$ ein konsistentes SLG.*
Sei L_h die Lösungsmenge des homogenen SLG $Ax = 0$.
Sei L_i die Lösungsmenge des inhomogenen SLG $Ax = b$.
Sei $c \in L_i$.
Dann ist $L_i = \{a + c \mid a \in L_h\}$ und $L_h = \{d - c \mid d \in L_i\}$.

Beweis: Offensichtlich gilt wegen Satz(ii) für alle $a \in L_h$: $a + c \in L_i$.
Sei andererseits $d \in L_i$. Dann ist wegen Satz(iii) $d - c \in L_h$, und es gilt $d = (d - c) + c$. □

Bemerkung 2.2.37: Wir stellen einige Beziehungen zur Analytischen Geometrie her. Die Lösungen einer linearen Gleichung

$$ax + by + cz = d$$

hängen von 2 frei wählbaren Parametern ab und bilden eine Ebene im Raum.
Eine Gerade im Raum erhält man als Schnitt von 2 Ebenen, also durch 2 lineare Gleichungen.

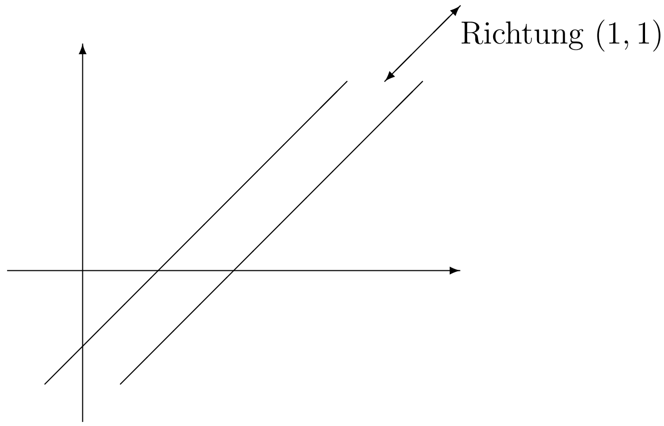

Fig.: parallele Geraden

Sind die beiden Vektoren v und w Lösungen eines SLG, so sind alle Punkte auf der Geraden durch v und w, also alle Vektoren

$$v + \lambda(w - v) \ ,$$

Lösungen dieses SLG.
Zwei Gerade in der Ebene haben einen Schnittpunkt, es sei denn sie sind parallel. Sind die beiden Geraden parallel, so haben wir ein SLG der Form

$$(1) \quad Ax = b$$

von 2 Gleichungen in 2 Variablen, wobei $\text{rang}(A) = 1$ und $\text{rang}(A|b) = \text{rang}(A| - b) = 2$.

Beispiel 2.2.38: Wir betrachten das SLG aus Beispiel 2.2.1. Die Hermite-Form der erweiterten Matrix ist

$$\begin{pmatrix} 1 & 0 & 0 & \vdots & 12 \\ 0 & 1 & 0 & \vdots & 5 \\ 0 & 0 & 1 & \vdots & -2 \end{pmatrix} .$$

Bei dieser Transformation bleiben die Lösungen invariant. Aus der Hermite-Form ist die Lösung $(12, 5, -2)^T$ unmittelbar ablesbar. $\square$

2.3 Invertierbare Matrizen

In diesem Kapitel befassen wir uns mit der Frage, welche Matrizen eine multiplikative Inverse haben und wie wir im gegebenen Fall diese Inverse bestimmen können. Die Matrizen sind alle über demselben Körper K.

Definition 2.3.1: *Sei A eine $m \times n$ Matrix. Eine $n \times m$ Matrix X heisst* **Linksinverse** *von A wenn gilt $XA = I_n$.*
Eine $n \times m$ Matrix Y heisst **Rechtsinverse** *von A wenn gilt $AY = I_m$.*

Beispiel 2.3.2: (i) Die Matrix

$$\begin{pmatrix} 1 & 1 \\ 4 & 3 \\ 3 & 4 \\ 0 & 0 \end{pmatrix}$$

hat unendlich viele Linksinverse der Form

$$\begin{pmatrix} -3 & 1 & 0 & a \\ -3 & 0 & 1 & b \end{pmatrix}$$

aber keine Rechtsinverse.
(ii) Die Matrix

$$\begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}$$

hat die eindeutig bestimmte Links- und Rechtsinverse

$$\begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix}$$

Satz 2.3.3: *Sei A eine $m \times n$ Matrix. Dann gilt:*

(i) A hat eine Rechtsinverse g.d.w. $\operatorname{rang}(A) = m$;

(ii) A hat eine Linksinverse g.d.w. $\operatorname{rang}(A) = n$.

Beweis:

(i) "$\Rightarrow$": Sei Y eine Rechtsinverse zu A, also $AY = I_m$.
Bezeichne $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)^T$ (1 an der i-ten Stelle) die i-te Spalte von I_m, und bezeichne y_i die i-te Spalte von Y. Also

$$Ay_i = e_i \quad \text{für alle } i \in \{1, \ldots, m\} \ .$$

Für jedes i ist dieses SLG konsistent, also $\operatorname{rang}(A) = \operatorname{rang}(A|e_i)$. Jedes e_i ist linear abhängig von den Spalten von A. Ausserdem sind die Vektoren $e_1, \ldots, e_m$ linear unabhängig. Somit gilt

$$\begin{aligned} \operatorname{rang}(A) &= \operatorname{rang}(A|e_1) \\ &= \operatorname{rang}(A|e_1|e_2) \\ &\ \vdots \\ &= \operatorname{rang}(A|e_1|\cdots|e_m) = \operatorname{rang}(A|I_m) = m \ . \end{aligned}$$

"$\Leftarrow$": Sei $\operatorname{rang}(A) = m$. Für die Hermite-Matrix $H(A^T)$ gilt

$$\operatorname{rang}(H(A^T)) = \operatorname{rang}(A^T) = \operatorname{rang}(A) = m \ ,$$

also für ein Produkt E von Elementarmatrizen $EA^T = H(A^T)$ und

$$H(A^T) = \begin{pmatrix} I_m \\ 0_{n-m,m} \end{pmatrix} \ ,$$

bzw.

$$AE^T = \begin{pmatrix} I_m \vdots 0_{m,n-m} \end{pmatrix} \ .$$

Sei Y die Matrix bestehend aus den ersten m Spalten von E^T. Y ist offensichtlich eine Rechtsinverse von A.

(ii) Wir sehen das aus folgenden Äquivalenzen:

$$\begin{aligned} XA = I_n \quad &\Longleftrightarrow \quad A^T X^T = I_n \\ &\Longleftrightarrow \text{(i)} \quad \operatorname{rang}(A) = \operatorname{rang}(A^T) = n \ . \end{aligned}$$

□

Satz 2.3.4: *Hat die Matrix A sowohl eine Linksinverse X als auch eine Rechtsinverse Y, dann gilt:*

(i) A ist quadratisch;

(ii) $X = Y$.

Wenn die Matrix A also sowohl eine Linksinverse als auch eine Rechtsinverse hat, so ist diese eindeutig bestimmt.

Beweis:
(i) Wegen Satz 2.3.3 gilt: $n = \operatorname{rang}(A) = m$.
(ii) $XA = I_n = AY$, somit

$$X = XI_n = X(AY) = (XA)Y = I_nY = Y \ .$$

□

Satz 2.3.5: *Sei A eine quadratische $n \times n$ Matrix. FAÄ:*

(i) A hat eine Linksinverse;

(ii) A hat eine Rechtsinverse;

(iii) $\operatorname{rang}(A) = n$;

(iv) die Hermite-Form von A ist I_n;

(v) A ist ein Produkt von Elementarmatrizen.

Beweis: Die Äquivalenz von (i), (ii) und (iii) ist eine unmittelbare Konsequenz von Satz 2.3.3.
"(iii) $\iff$ (iv)": hat A den Rang n, so hat $H(A)$ n vom Nullvektor verschiedene Zeilen, also n Staffelelemente 1. Das geht nur für $H(A) = I_n$. Umgekehrt haben wir

$$\operatorname{rang}(A) = \operatorname{rang}(H(A)) = \operatorname{rang}(I_n) = n \ .$$

"(i) $\Rightarrow$ (v)": Habe also A eine Linksinverse X. Das heisst aber auch, dass X eine Rechtsinverse hat, nämlich A. Wegen (ii) $\Rightarrow$ (iv) gibt es also Elementarmatrizen $E_1, \ldots, E_q$ sodass

$$E_q \cdots E_1 X = I_n \ .$$

Somit haben wir auch

$$A = I_n A = (E_q \cdots E_1 X)A = E_q \cdots E_1(XA) = E_q \cdots E_1 \ ,$$

d.h. A ist ein Produkt von Elementarmatrizen.
"(iii) $\Leftarrow$ (v)": Sei A ein Produkt von Elementarmatrizen, also $A = E_1 \cdots E_p$. Somit gilt

$$A = E_1 \cdots E_p I_n \ .$$

Die Identitätsmatrix I_n hat Rang n, dieser wird durch Elementaroperationen nicht verändert (Satz 2.2.27), also hat A den Rang n. □

Definition 2.3.6: *Aus den obigen Sätzen sehen wir, dass eine einseitige Inverse einer quadratischen Matrix A automatisch eine beidseitige Inverse ist. Wir sprechen also nur von der* **Inversen** *von A und schreiben diese als A^{-1}. Die Inverse ist, wenn sie existiert, eindeutig (Satz 2.3.4).*

Hat die quadratische Matrix A eine Inverse, so heisst A **invertierbar** *bzw.* **regulär**. *Hat A keine Inverse, so heisst A auch* **singulär**.

Korollar zu Satz 2.3.5: *Sei A eine quadratische Matrix.*

(i) Ist A invertierbar, dann ist auch A^{-1} invertierbar. Weiters gilt $\left(A^{-1}\right)^{-1} = A$.

(ii) Jedes Produkt von Elementarmatrizen ist invertierbar. B ist also zeilenäquivalent zu A g.d.w. es eine invertierbare Matrix E gibt sodass $B = EA$.

Beweis: (i) Offensichtlich ist A die Inverse zu A^{-1}.
(ii) Wegen der Äquivalenz von (v) und (i),(ii) in Satz 2.3.5. □

Aus Satz 2.3.5 können wir sofort eine Methode herleiten um die Invertierbarkeit der quadratischen $n \times n$ Matrix A zu entscheiden, und gegebenenfalls A^{-1} zu berechnen. Dazu gehen wir von der Matrix $A|I_n$ aus und wenden elementare Zeilenoperationen an, um den linken Teil in I_n zu transformieren. Ist das nicht möglich, so ist A nicht invertierbar. Gelingt es aber, so haben wir im rechten Teil automatisch die Inverse A^{-1} stehen. Dieser Prozess sieht wie folgt aus:

$$A|I_n \to E_1 A|E_1 \to E_2 E_1 A|E_2 E_1 \to \cdots \ .$$

In jedem Schritt dieses Prozesses haben wir also Matrizen der Form

$$S|Q = E_i \cdots E_1 A|E_i \cdots E_1$$

vorliegen, für welche gilt $QA = S$. Ist A invertierbar, so erreichen wir schliesslich die Endkonfiguration

$$I_n | E_p \cdots E_1 \ ,$$

sodass $E_p \cdots E_1 A = I_n$, also

$$A^{-1} = E_p \cdots E_1 \ .$$

Beispiel 2.3.7: Wir wollen den oben beschriebenen Prozess anwenden auf die Matrix

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

Dabei erhalten wir die folgenden Konfigurationen:

$$\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 1 & 3 & 4 & 0 & 1 & 0 \\ 1 & 4 & 4 & 0 & 0 & 1 \end{array} \quad \to \quad \begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & 1 & -1 & 1 & 0 \\ 0 & 2 & 1 & -1 & 0 & 1 \end{array}$$

$$\to \quad \begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & 1 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 & -2 & 1 \end{array}$$

$$\to \quad \begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 1 \\ 0 & 0 & -1 & 1 & -2 & 1 \end{array}$$

$$\to \quad \begin{array}{ccc|ccc} 1 & 2 & 0 & 4 & -6 & 3 \\ 0 & 1 & 0 & 0 & -1 & 1 \\ 0 & 0 & -1 & 1 & -2 & 1 \end{array}$$

$$\to \quad \begin{array}{ccc|ccc} 1 & 0 & 0 & 4 & -4 & 1 \\ 0 & 1 & 0 & 0 & -1 & 1 \\ 0 & 0 & 1 & -1 & 2 & -1 \end{array}$$

Die Matrix A ist also invertierbar, und die Inverse ist

$$\begin{pmatrix} 4 & -4 & 1 \\ 0 & -1 & 1 \\ -1 & 2 & -1 \end{pmatrix} \ . \qquad \square$$

Sind die beiden $n \times n$ Matrizen A und B invertierbar, so ist i.a. die Summe $A + B$ nicht invertierbar. Das sieht man etwa am Beispiel $A = I_n, B = -I_n$. Das Produkt AB ist aber sehr wohl invertierbar.

Satz 2.3.8: *Seien A und B zwei $n \times n$ Matrizen. Sind sowohl A als auch B invertierbar, dann ist auch AB invertierbar und es gilt*

$$(AB)^{-1} = B^{-1} A^{-1} \ .$$

Beweis: Offensichtlich gilt

$$ABB^{-1}A^{-1} = AI_nA^{-1} = AA^{-1} = I_n \ .$$

Also ist $B^{-1}A^{-1}$ eine Rechtsinverse zu AB. Wegen Satz 2.3.5 hat A auch eine Linksinverse, also eine eindeutig bestimmt Inverse, diese ist $B^{-1}A^{-1}$. □

Korollar: *Ist A invertierbar, dann ist auch A^m invertierbar für jede natürliche Zahl m (man beachte, dass $A^0 = I_n$). Weiters gilt*

$$(A^m)^{-1} = (A^{-1})^m \ .$$

Beweis: Induktion über m. Die Behauptung ist trivial für $m = 0$. Für $m > 0$ haben wir

$$(A^{-1})^m = A^{-1}(A^{-1})^{m-1} = A^{-1}(A^{m-1})^{-1} = (A^{m-1}A)^{-1} = (A^m)^{-1} \ . \qquad \square$$

Satz 2.3.9: *Ist A invertierbar, dann ist auch die Transponierte A^T invertierbar. Weiters gilt*

$$(A^T)^{-1} = (A^{-1})^T \ .$$

Beweis: Es gilt

$$I_n = I_n^T = (AA^{-1})^T = (A^{-1})^T A^T \ .$$

Also ist $(A^{-1})^T$ eine Linkinverse, und somit eine Inverse, von A^T. □

Wir erinnern daran, dass die $n \times n$ Matrix A **orthogonal** ist, wenn gilt

$$AA^T = I_n = A^T A \ .$$

Aus Satz 2.3.5 sehen wir, dass nur eine dieser Gleichungen notwendig ist. Eine orthogonale Matrix ist also eine invertierbare Matrix deren Inverse die Transponierte ist.

Satz 2.3.10: *Sei $A \cdot x = b$ ein SLG von n Gleichungen in n Variablen. A ist also eine $n \times n$ Matrix, und b ein Spaltenvektor der Länge n. Ist A invertierbar, so ist $A^{-1} \cdot b$ der eindeutig bestimmte Lösungsvektor des Gleichungssystems.*

Beweis: Sei x ein Lösungsvektor. Dann gilt

$$x = I_n x = A^{-1}Ax = A^{-1}b \ . \qquad \square$$

Beispiel 2.3.11: Wir betrachten das SLG aus Beispiel 2.2.1 und 2.2.38.

$$A = \begin{pmatrix} 0 & 1 & 2 \\ 1 & -2 & 1 \\ 0 & 3 & -4 \end{pmatrix} \quad , \quad A^{-1} = \begin{pmatrix} \frac{1}{2} & 1 & \frac{1}{2} \\ \frac{2}{5} & 0 & \frac{1}{5} \\ \frac{3}{10} & 0 & \frac{-1}{10} \end{pmatrix}$$

und der Lösungsvektor ist

$$A^{-1} \cdot b = \begin{pmatrix} 12 \\ 5 \\ -2 \end{pmatrix} . \square$$

Beispiel 2.3.12: Eine wichtige Anwendung finden inverse Matrizen im **mehrdimensionalen Newton-Verfahren** (benannt nach Sir Isaac Newton, 1642–1727) zur Approximation von Nullstellen von Funktionen.

Im (eindimensionalen) Newton-Verfahren geht es darum, zu einer einmal stetig differenzierbaren reellen Funktion, also

$$f : \mathbb{R} \longrightarrow \mathbb{R} \quad \in \quad C^1(\mathbb{R}),$$

eine Nullstelle $x \in \mathbb{R}$ approximativ zu finden, also $f(x) = 0$.
Dazu startet man "in der Nähe" der Nullstelle etwa bei $x^{(0)}$, und bestimmt eine Folge von Approximationswerten $x^{(0)}, x^{(1)}, x^{(2)}, \ldots$. Um den jeweils nächsten Approximationswert $x^{(m+1)}$ zu bestimmen, betrachtet man die Tangente an f bei $x^{(m)}$:

$$t(x) \;=\; f(x^{(m)}) + f'(x^{(m)}) \cdot (x - x^{(m)}) \;,$$

und schneidet diese mit der x-Achse. $x^{(m+1)}$ ist dieser Schnittpunkt:

$$x^{(m+1)} \;:=\; x^{(m)} - \frac{f(x^{(m)})}{f'(x^{(m)})} \;.$$

In der Analysis beweist man, dass dieses Verfahren lokal konvergiert, d.h. wenn der Startpunkt $x^{(0)}$ genügend nahe an einer Nullstelle liegt, so wird tatsächlich diese Nullstelle approximiert als Limes der Folge $\left(x^{(m)}\right)_{m\in\mathbb{N}}$. Verfahren zur Wurzelberechnung (etwa $\sqrt{2}$) sind Spezialfälle dieses Newtonschen Näherungsverfahrens.

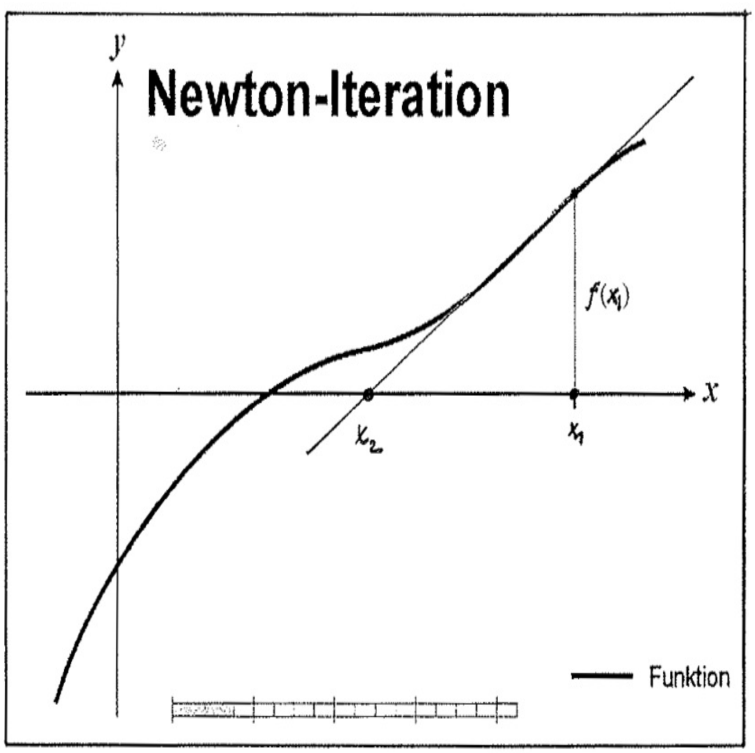

Figure 2: Newton-Verfahren

Im mehrdimensionalen Newton-Verfahren betrachtet man nun eine Funktion

$$f \;:\; \begin{array}{ccc} \mathbb{R}^n & \longrightarrow & \mathbb{R}^n \\ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} & \mapsto & \begin{pmatrix} f_1(x_1, \ldots, x_n) \\ \vdots \\ f_n(x_1, \ldots, x_n) \end{pmatrix} \end{array} ,$$

wobei alle Komponenten f_i differenzierbar sind nach allen Variablen, und die partiellen Ableitungen $\partial f_i / \partial x_j$ stetig sind.

Man startet nun wiederum "in der Nähe" einer Nullstelle, etwa mit $x^{(0)}$. Die weiteren Näherungswerte werden bestimmt als

$$x^{(m+1)} \;:=\; x^{(m)} \;-\; J_f^{-1}(x^{(m)}) \cdot f(x^{(m)}) \;.$$

Dabei ist

$$J_f(x) \;:=\; \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{pmatrix}$$

die **Jakobi-Matrix** von f. Das mehrdimensionale Newton-Verfahren funktioniert also, wenn die Jakobi-Matrix invertierbar ist.

Wir demonstrieren das am Beispiel

$$f \;:\; \begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ \begin{pmatrix} x \\ y \end{pmatrix} & \mapsto & \begin{pmatrix} xy + 2x \\ y^2 + y + x \end{pmatrix} \end{array} .$$

Für eine Nullstelle $(x, y)^T$ muss gelten $y^2 + y + x = x(y + 2) = 0$; daraus sieht man leicht, dass etwa $(0, -1)^T$ eine Nullstelle ist. Wir wollen sehen, wie wir diese Nullstelle mittels des Newton-Verfahrens approximieren können.

$$J_f(x, y) \;=\; \begin{pmatrix} y + 2 & x \\ 1 & 2y + 1 \end{pmatrix} ,$$

$$J_f^{-1}(x, y) \;=\; \frac{1}{2y^2 + 5y - x + 2} \cdot \begin{pmatrix} 2y + 1 & -x \\ -1 & y + 2 \end{pmatrix} .$$

Wir führen 2 Näherungsschritte durch:

$$x^{(0)} \;:=\; \begin{pmatrix} 0,5 \\ -0,8 \end{pmatrix}$$

$$f(x^{(0)}) \;=\; \begin{pmatrix} 0,60 \\ 0,34 \end{pmatrix}$$

$$x^{(1)} \;=\; x^{(0)} - J_f^{-1}(x^{(0)}) \cdot f(x^{(0)}) \;=\; \begin{pmatrix} 0,0655737705 \\ -0,9573770495 \end{pmatrix}$$

$$f(x^{(1)}) \;=\; \begin{pmatrix} 0,06836871807 \\ 0,0247675359 \end{pmatrix}$$

$$x^{(2)} \;=\; x^{(1)} - J_f^{-1}(x^{(1)}) \cdot f(x^{(1)}) \;=\; \begin{pmatrix} 0,002625108412 \\ -0,9991162739 \end{pmatrix}$$

$$f(x^{(2)}) \;=\; \begin{pmatrix} 0,002627428289 \\ 0,001742163312 \end{pmatrix}$$

Wir sehen also, dass bereits nach 2 Näherungsschritten die Nullstelle ziemlich gut approximiert wird. □