Elemente der Analysis II

- Kapitel 3: Lineare Abbildungen und Gleichungssysteme
- Informationen zur Vorlesung:
  http://www.mathematik.uni-trier.de/∼wengenroth/

3.1 Beispiel

Sei $p = [p_1, \ldots, p_n]$ der Vektor der Preise von Gütern mit Nummern $1, \ldots, n$. Jedes $x = [x_1, \ldots, x_n]$ beschreibt einen "Warenkorb" mit x_k Einheiten (Stück, Liter, Gewicht, ...) des Gutes Nummer k. Dieser Warenkorb hat dann den Wert

$$f(x) = p_1x_1 + p_2x_2 + \ldots + p_nx_n = \langle p, x\rangle.$$

Wie schon in Satz 1.9 bemerkt, ist dann

$f(ax) = af(x)$ (der a-fache Warenkorb ist das a-fache wert) und

$f(x+y) = f(x)+f(y)$ (der Wert von zwei Körben ist die Summe der Werte).

(Beachte, dass auch dieses Modell eine Vereinfachung ist. Tatsächlich würde man wohl Rabatte aushandeln, wenn man viel einkauft.)

3.2 Definition

Eine Abbildung $f : \mathbb{R}^n \to \mathbb{R}^m$ heißt linear, wenn für alle $x, y \in \mathbb{R}^n$ und $a \in \mathbb{R}$ gelten:

$$f(ax) = af(x) \text{ und } f(x+y) = f(x) + f(y).$$

3.3 Satz

(a) $f : \mathbb{R}^n \to \mathbb{R}^m$ ist genau dann linear, wenn für alle $x, y \in \mathbb{R}^n$ und $a, b \in \mathbb{R}$ gilt $f(ax + by) = af(x) + bf(y)$.

(b) Ist $f : \mathbb{R}^n \to \mathbb{R}^m$ linear, so gilt für alle $x1, \dots, x^p \in \mathbb{R}^n$ und $a_1, \dots a_p \in \mathbb{R}$, dass $f\left(\sum_{j=1}^{p} a_j x^j\right) = \sum_{j=1}^{p} a_j f(x^j)$.
(Linearkombinationen)

(c) $f = [f_1, \dots, f_m] : \mathbb{R}^n \to \mathbb{R}^m$ ist genau dann linear, wenn alle Komponentenfunktionen $f_k : \mathbb{R}^n \to \mathbb{R}$ linear sind.

(d) Summen, Vielfache, Zusammensetzungen und Kompositionen linearer Abbildungen sind linear (natürlich, falls sie definiert sind).

Beweis

(a) $'' \Longrightarrow ''$ folgt aus $f(ax+by) = f(ax) + f(by) = af(x) + bf(y)$.
$'' \Longleftarrow ''$ folgt mit $b = 0$ bzw. $a = b = 1$.

(b) folgt induktiv aus (a)

(c) folgt aus der komponentenweisen Definition von Summen und Vielfachen von Vektoren.

(d) Die Linearität von $f + g$, $[f, g]$ und αf rechnet man direkt nach. Für $f : \mathbb{R}^h \to \mathbb{R}^m$ und $g : \mathbb{R}^m \to \mathbb{R}^p$ beide linear ist
$g(f(ax+by)) = g(af(x) + bf(y)) = ag(f(x)) + bg(f(y))$.

$\square$

3.4 Satz (Darstellungssatz)

Jede lineare Abbildung $f : \mathbb{R}^n \to \mathbb{R}$ ist von der Form $f(x) = \langle p, x \rangle$ mit $p = [f(e1), \ldots, f(e^n)] \in \mathbb{R}^n$ und den Einheitsvektoren $e1, \ldots, e^n$.

Beweis: Wegen 1.6 (a) ist $x = \sum_{k=1}^{n} x_k e^k$ für jedes $x = [x_1, \ldots, x_n]$. Mit 3.3 (b) (für $a_k = x_k$) folgt also

$$f(x) = \sum_{k=1}^{n} x_k f(e^k) = \sum_{k=1}^{n} x_k p_k = \langle x, p \rangle = \langle p, x \rangle.$$

$\square$

3.4 Satz (Darstellungssatz)

(a) Bemerkung: Eine lineare Abbildung $\mathbb{R}^n \to \mathbb{R}$ ist also durch die n Werte $f(e1), \ldots, f(e^n)$ der Standardbasis eindeutig bestimmt und damit sehr leicht auszurechnen. (Ist $z1, \ldots, z^n$ irgendeine Basis des $\mathbb{R}^n$, so ist f auch durch $f(z1), \ldots, f(z^n)$ eindeutig bestimmt, aber nur dann leicht auszurechnen, wenn man die Koeffizienten in $x = \sum_{k=1}^{n} a_k z^k$ effektiv bestimmen kann).

(b) Anschaulich $(n = 2)$ bedeutet 3.4, dass der Graph

$$\Gamma_f = \{[x_1, x_2, f(x_1, x_2)] : [x_1, x_2] \in \mathbb{R}2\}$$

die Ebene in $\mathbb{R}3$ ist die $[1, 0, f(1, 0)], [0, 1, f(0, 1)]$ und den Ursprung $[0, 0, 0] = [0, 0, f(0, 0)]$ enthält.

(c) Interpretation von $p_k = f(e^k)$: Dies ist der Faktor (Preis) mit dem die k-te Komponente x_k von $x = [x_1, \ldots, x_n]$ zum Funktionswert $f(x)$ beiträgt.

3.5 Beispiel

Ein Unternehmen produziere aus 3 Rohstoffen zwei Produkte.
Input: x_1 Menge der ersten, x_2 Menge des zweiten, x_3 Menge des dritten Rohstoffs.
Output: $f_1(x_1, x_2, x_3)$ produzierte Menge des ersten Produkts bei Einsatz x_1, x_2, x_3.
$f_2(x_1, x_2, x_3)$ produzierte Menge des zweiten Produkts bei Einsatz x_1, x_2, x_3.
Modell: $f = [f_1, f_2] : \mathbb{R}3 \to \mathbb{R}2$.

3.6 Matrizen

(a) Sei $f : \mathbb{R}^n \to \mathbb{R}^m$ eine lineare Abbildung mit Komponentenfunktionen $f_1, \ldots, f_m : \mathbb{R}^n \to \mathbb{R}$, das heißt $f_j(x)$ ist die j-te Komponente des Vektors $f(x) \in \mathbb{R}^m$. Die f_k sind wieder linear (wegen Satz 3.3(c)) und deshalb gibt es nach Satz 3.4 für jedes $j = 1, 2, \ldots, m$ einen Vektor $p^j = [p^j_1, p^j_2, \ldots, p^j_n] \in \mathbb{R}^n$ (nämlich $p^j_k = f_j(e^k)$), so dass

$$f_j(x) = p^j_1 x_1 + p^j_2 x_2 + \ldots + p^j_n x_n$$

(b) Die lineare Abbildung f ist also durch die $n\,m$ Zahlen $p1_1, \ldots, p1_n, p2_1, \ldots, p2_n, \ldots, p^m_1, \ldots, p^m_n$ eindeutig bestimmt und leicht auszurechnen. Es bietet sich an, diese Zahlen in eine Tabelle zu schreiben:

$$P = \begin{bmatrix} p1_1 & p1_2 & \ldots & p1_n \\ p2_1 & p2_2 & \ldots & p2_n \\ \vdots & & & \\ p^m_1 & p^m_2 & \ldots & p^m_n \end{bmatrix}$$

3.6 Matrizen

(b) Eine solche Tabelle heißt $m \times n$-Matrix.
Die Vektoren $p^j = [p^j_1, \dots, p^j_n] \in \mathbb{R}^n$ heißen Zeilen von P und die Spaltenvektoren $\begin{bmatrix} p1_k \\ \vdots \\ p^m_k \end{bmatrix}$ heißen Spalten von P.
Eine $m \times n$-Matrix hat also m Zeilen und n Spalten. Mit dem Symbol $\mathbb{R}^{m \times n}$ bezeichnen wir die Menge aller $m \times n$-Matrizen.

(c) Alternative (übliche) Schreibweise: $p^j_k = p_{j,k}$. Der erste/obere Index j ist die Zeilennummer und der zweite/untere Index k ist die Spaltennummer. Oft schreibt man $P = [p_{j,k}]_{\substack{j \in \{1,\dots,m\} \\ k \in \{1,\dots n\}}}$ oder nur $P = [p_{jk}]$, falls die Dimensionen n und m klar sind.

3.6 Matrizen

(d) Die Unterscheidung zwischen Zeilen und Spalten ist ab jetzt wesentlich, und wir fassen ab jetzt Vektoren immer als SPALTEN auf, das heißt $\mathbb{R}^n = \mathbb{R}^{n\times 1}$.

(e) Zusammenfassung:
Ist $f : \mathbb{R}^n \to \mathbb{R}^m$ eine lineare Abbildung und zugehöriger Matrix $P = [p_{jk}]$ (wobei $p_{j,k}$ die j-te Komponente von $f(e^k)$), so gilt für

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$$

$$f(x) = \begin{bmatrix} p_{1,1}x_1 + p_{1,2}x_2 + \ldots + p_{1,n}x_n \\ \vdots \\ p_{m,1}x_1 + p_{m,2}x_2 + \ldots + p_{m,n}x_n \end{bmatrix}$$

3.6 Matrizen

(f) Wir definieren deshalb für $P \in \mathbb{R}^{m \times n}$ und $x \in \mathbb{R}^n$ das Produkt

$$P \cdot x = \begin{bmatrix} p_{1,1}x_1 + p_{1,2}x_2 + \ldots + p_{1,n}x_n \\ \vdots \\ p_{m,1}x_1 + p_{m,2}x_2 + \ldots + p_{m,n}x_n \end{bmatrix} \in \mathbb{R}^m$$

(g) Jede lineare Abbildung $\mathbb{R}^n \to \mathbb{R}^m$ ist also von der Form $x \mapsto P \cdot x$.

3.7 Beispiel

In der Situation von Beispiel 3.5 sei konkret

$$\begin{aligned} f_1(x_1, x_2, x_3) &= 2x_1 + 4x_2 + x_3 \\ f_2(x_1, x_2, x_3) &= 6x_1 + x_2 + 2x_3 \end{aligned}$$

Dann hat man also $P = \begin{bmatrix} 2 & 4 & 1 \\ 6 & 1 & 2 \end{bmatrix}$. Das Unternehmen erhalte eine Anfrage, ob es y_1 Stück von Gut 1 und y_2 Stück von Gut 2 liefern könne.

Die Geschäftsführung frag sich dann: Gibt es Input $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, so dass $f(x) = y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$, also

$$\begin{array}{ll} (1) & 2x_1 + 4x_2 + x_3 = y_1 \\ (2) & 6x_1 + x_2 + 2x_3 = y_2 \end{array} \quad ?$$

3.7 Beispiel

Subtraktion des dreifachen der ersten Gleichung von der zweiten liefert, dass das Gleichungssystem äquivalent ist zu

$$\begin{array}{lrcl} (1) & 2x_1 + 4x_2 + x_3 & = & y_1 \\ (3) & -11x_2 - x_3 & = & y_2 - 3y_1. \end{array}$$

Dieses neue System kann man nun explizit auflösen:

- Wähle irgendein $x_3 \in \mathbb{R}$
- Definiere

$$\begin{array}{rcl} x_2 & = & \frac{1}{-11}(y_2 - 3y_1 + x_3) \\ x_1 & = & \frac{1}{2}(y_1 - x_3 - 4x_3) \end{array}$$

Es gibt also sehr viele Lösungen, und man kann die Anfrage positiv beantworten (und sich dann überlegen, welches x_3 zu einer Lösung mit weiteren günstigen Eigenschaften führt).

3.7 Beispiel

Bemerkungen:

- In realistischen Beispielen hat man viel mehr Variablen $x_1, \dots, x_n$ und viel mehr Gleichungen.
- Die Menge der Lösungen $\{x \in \mathbb{R}^n : f(x) = y\}$ ist also eine Niveaumenge zu Niveau y.
- Weil man 3 Unbekannte und 2 Gleichungen hat, ist es nicht überraschend, dass es viele Lösungen gibt.
- Vorsicht: Ist $Q = \begin{bmatrix} 2 & 4 & 1 \\ 4 & 8 & 2 \end{bmatrix}$, so hat $\begin{array}{rcl} 2x_1 + 4x_2 + x_3 & = & 1 \\ 4x_1 + 8x_2 + 2x_3 & = & 1 \end{array}$ keine Lösung. (Die zweite Zeile von Q ist ein Vielfaches der ersten Zeile).

3.8 Lineare Gleichungssysteme (LGS)

(a) Für eine $m \times n$-Matrix $P = [p_{j,k}]$ und $y \in \mathbb{R}^m$ heißt die Gleichung

$$P \cdot x = y$$

ein LGS mit m Gleichungen und n Unbekannten oder $\text{LGS}_{m \times n}$.
Jedes $x \in \mathbb{R}^n$, das diese Gleichung erfüllt heißt eine Lösung des LGS.
$L(P, y) = \{x \in \mathbb{R}^n : P \cdot x = y\}$ heißt Lösungsmenge.

(b) $P \cdot x = y$ ist eine kurze und elegante Schreibweise für

$$\begin{array}{ccc}
p_{1,1}x_1 + p_{1,2}x_2 + \ldots + p_{1,n}x_n & = & y_1 \\
\vdots & \vdots & \\
p_{m,1}x_1 + p_{m,2}x_2 + \ldots + p_{m,n}x_n & = & y_m
\end{array}$$

3.8 Lineare Gleichungssysteme (LGS)

(c) LGS treten sehr oft mit riesigen Dimensionen n und m auf (Computertomographie: $n \approx m \approx 106$)

(d) Hat P die Spalten $p1, \ldots, p^n$, so ist $P \cdot x = x_1 p1 + \ldots, + x_n p^n$. Es gibt also genau dann eine Lösung von $P \cdot x = y$, wenn y eine Linearkombination der Spalten ist.

(e) Für $y = \mathrm{O} \in \mathbb{R}^m$ gibt es immer mindestens eine Lösung, nämlich $x = \mathrm{O} \in \mathbb{R}^n$. Oft gibt es aber auch $x \neq \mathrm{O}$ mit $P \cdot x = \mathrm{O}$.

3.9 Rekursive Algorithmen

(a) Der Gaußsche Algorithmus (Eliminationsverfahren) ist eine einfache Methode, wie man das Problem „Finde alle Lösungen des $\text{LGS}_{m\times n}\ P \cdot x = y$“ auf ein „kleineres“ $\text{LGS}_{(m-1)\times(n-1)}$ zurückführt. Das neue Problem wird mit der selben Methode auf ein noch kleineres $\text{LGS}_{(m-2)\times(n-2)}$ zurückgeführt, bis man schließlich entweder ein $\text{LGS}_{1,q}$ oder ein $\text{LGS}_{p,1}$ erhält (das man leicht lösen kann).
So eine Methode nennt man rekursiven Algorithmus.

(b) Türme von Hanoi

- Gegeben: Drei Stäbe, auf einem davon n Scheiben nach der (verschiedenen) Größe geordnet.
- Ziel: Stapel von einem Stab a auf anderen c bewegen.
- Erlaubte Züge: Oberste Scheibe eines Stabs auf einen anderen legen, so dass nie eine größere Scheibe über einer kleineren liegt.

Algorithmus:

- Verschiebe $(n-1)$ Scheiben von a (über c) nach b.
- Verschiebe oberste (einzige) Scheibe von a nach c.
- Verschiebe $(n-1)$ Scheiben von b (über a) nach c.

3.10 Der Gauß-Algorithmus

Gesucht sind alle Lösungen $x \in \mathbb{R}^n$ das $LGS_{m \times n}$ $P \cdot x = y$ für gegebene Matrix $P \in \mathbb{R}^{m \times n}$ und $y \in \mathbb{R}^m$, also

$$\begin{array}{rcl} p_{1,1}x_1 + p_{1,2}x_2 + \ldots + p_{1,n}x_n & = & y_1 \\ \vdots & & \vdots \\ p_{m,1}x_1 + p_{m,2}x_2 + \ldots + p_{m,n}x_n & = & y_m. \end{array}$$

Lösungsstrategie: Elimination der Unbekannten x_n aus den ersten $(m-1)$ Gleichungen durch Addition geeigneter Vielfacher der letzten Gleichung zu den übrigen. Das geht, falls der Koeffizient $p_{m,n}$ ungleich 0 ist. Sonst Vertausche die Reihenfolge der Gleichungen.

3.10 Der Gauß-Algorithmus

I_n Vertausche gegebenenfalls die Zeilen, so dass $p_{m,n} \neq 0$. (Falls nicht möglich, gehe zu V_n).

II_n Für $j = 1, \ldots, m-1$ setze $\alpha_j = -\frac{p_{j,n}}{p_{m,n}}$ und addiere das α_j-fache der letzten Zeile zur j-ten Zeile. Dies liefert das $LGS_{m\times n}$ mit der selben Lösungsmenge

$$\left.\begin{array}{lcl} \tilde{p}_{1,1}x_1 + \ldots + \tilde{p}_{1,n-1}x_{n-1} & = & \tilde{y}_1 \\ \vdots & & \vdots \\ \tilde{p}_{m-1,1}x_1 + \ldots + \tilde{p}_{m-1,n-1}x_{n-1} & = & \tilde{y}_{n-1} \end{array}\right\} \tilde{P}\cdot x = \tilde{y}$$
$$p_{m,1}x_1 + \ldots + p_{m,n-1}x_{n-1} + p_{m,n}x_n = y_n$$

mit $\tilde{p}_{j,k} = p_{j,k} + \alpha_j p_{m,k}$ und $\tilde{y}_j = y_j + \alpha_j y_m$.

III_n Löse durch die Schritte I_{n-1} bis IV_{n-1} das kleinere $LGS_{m-1,n-1}$.

IV_n Dann löst $x = [x_1, \ldots, x_n]$ das LGS $P \cdot x = y$

$$\iff \begin{array}{l} \tilde{x} = [x_1, \ldots, x_{n-1}] \text{ löst } \tilde{P}\cdot\tilde{x} = \tilde{y} \text{ und} \\ x_n = \frac{1}{p_{m,n}}(y_n - p_{m,1}x_1 - p_{m,2}x_2 - \ldots - p_{m,n-1}x_{n-1}) \end{array}$$

3.10 Der Gauß-Algorithmus

V_n Falls $p_{1,n} = p_{2,n} = \ldots = p_{m,n} = 0$ ist Schritt I_n nicht durchführbar. In diesem Fall gilt: $x = [x_1, \ldots, x_n]$ löst $P \cdot x = y \iff$ $\tilde{x} = [x_1, \ldots, x_{n-1}]$ löst das $\mathrm{LGS}_{m,n-1}$ $\tilde{P}\tilde{x} = y$ mit

$$\tilde{P} = \begin{bmatrix} p_{1,1} & \cdots & p_{1,n-1} \\ \vdots & & \\ p_{m,1} & \cdots & p_{m,n-1} \end{bmatrix}$$ (und x_n ist beliebig).

3.11 Bemerkung zum Gauß-Algorithmus

(a) Die Schritte IV und V besagen, dass man für jede Lösung $\tilde{x}$ des kleineren System eine (IV) beziehungsweise unendlich viele (V) Lösungen des Ausgangssystems erhält.

(b) Es ist möglich, dass das kleinere System keine Lösung hat. Dann hat auch das Ausgangssystem keine Lösung, das System heißt dann inkonsistent.

(c) Durch Umbenennen der Variablen $x_1, \ldots, x_n$ und Vertauschen der Zeilen (Pivotieren) erhält man Varianten des Algorithmus. In der Literatur oft zuerst: x_1 aus Gleichungen 2 bis n eliminieren.

(d) Lösen von $\mathrm{LGS}_{1,n}$ $p_1x_1 + \ldots + p_nx_n = y$ ist einfach:
Wähle k mit $p_k \neq 0$ und beliebige $x_\ell \in \mathbb{R}$ für $\ell \neq k$ und löse nach x_k auf.

(e) Lösen von $\mathrm{LGS}_{m,1}$
$$\begin{aligned} p_1x_1 &= y_1 \\ &\vdots \\ p_nx_n &= y_n \end{aligned}$$
ist auch einfach: Löse eine Gleichung und überprüfe die anderen.

3.12 Beispiel a

$$\begin{aligned} x_1 + 0x_2 + 2x_3 &= -1 \\ 2x_1 + x_2 + x_3 &= 0 \\ 2x_1 + x_2 + \tfrac{1}{2}x_3 &= 2 \end{aligned}$$

I_3 Keine Vertauschung nötig.

II_3 $\alpha_1 = -\frac{2}{1/2} = -4, \alpha_2 = -\frac{1}{1/2} = -2$. Das neue LGS ist

$$\left.\begin{aligned} -7x_1 - 4x_2 &= -9 \\ -2x_1 - x_2 &= -4 \end{aligned}\right\} \iff \tilde{P}x = \tilde{y}$$
$$2x_1 + x_2 + 1/2x_3 = 2$$

III_3 Lösen des $LGS_{2\times 2}$ in II_3: I_2 Wieder keine Vertauschung nötig.

II_2 $\alpha_1 = -\frac{-4}{-1} = -4$. Das neue LGS ist

$$\begin{aligned} (-7-4(-2))x_1 &= -9-4(-4) \iff x_1 = 7 \\ -2x_1 - x_2 &= -4 \end{aligned}$$

III_2 Das LGS in II_2 hat Lösung $x_1 = 7$.

IV_2 Das LGS in II_3 hat Lösung $x_1 = 7, x_2 = -10$.

IV_3 Das $LGS_{3\times 3}$ $P \cdot x = y$ hat Lösung $x_1 = 7, x_2 = -10, x_3 = 4$

3.12 Beispiel

(b) Gesucht sind alle Lösungen des $\mathrm{LGS}_{3\times 3}$

$$\begin{aligned} x_1 + x_2 - x_3 &= 0 \\ x_1 - x_2 + 2x_3 &= 3 \\ 3x_1 + x_2 &= 1 \end{aligned}$$

I_3. $p_{3,1} = 0$. Vertausche 1. und 3. Zeile. Neues $\mathrm{LGS}_{3\times 3}$.

$$\begin{aligned} 3x_1 + x_2 &= 1 \\ x_1 - x_2 + 2x_3 &= 3 \\ x_1 + x_2 - x_3 &= 0 \end{aligned}$$

II_3. $\alpha_1 = -\frac{0}{-1} = 0, \alpha_2 = -\frac{2}{-1} = 2$. Neues $\mathrm{LGS}_{3\times 3}$

$$\left.\begin{aligned} 3x_1 + x_2 = 1 \\ 3x_1 + x_2 = 3 \\ x_1 + x_2 - x_3 = 0 \end{aligned}\right\} \iff \tilde{P}x = \tilde{y}$$

3.12 Beispiel

III_3 Löse $\text{LGS}_{2\times 2}$ in II_3.

- $\text{I}_{2\cdot}$ Keine Vertauschung nötig.
- $\text{II}_{2\cdot}$ $\alpha_1 = 1$. Das neue $\text{LGS}_{2\times 2}$ ist

$$\left.\begin{array}{rl} 0 \times x_1 \quad &= -2 \\ 3x_1 + x_2 &= \quad 3. \end{array}\right\} \quad \tilde{P} \cdot x = \tilde{y}$$

- III_2 Das $\text{LGS}_{1\times 1}$ in II_2 hat keine Lösung.
- IV_2 Das $\text{LGS}_{2\times 2}$ in II_3 hat keine Lösung.

IV_3 Das $\text{LGS}_{3\times 3} P \cdot x = y$ in I_3 hat keine Lösung.

3.13 Simultanes Lösen von LGS

- Gegeben seien eine $m \times n$ Matrix P und s Vektoren $y1, \ldots, y^s \in \mathbb{R}^m$. Gesucht sind Lösungen $x1, \ldots, x^s \in \mathbb{R}^n$ von $P \cdot x^\ell = y^\ell$ für alle $\ell = 1, \ldots, s$.
- Eine Möglichkeit: s mal Gauß-Algorithmus durchführen.
- Bessere Variante: Die Matrix $\tilde{P}$ im Schritt II hängt nicht von der rechten Seite ab, also sind alle kleineren Probleme wieder von der Form $\tilde{P} \cdot x^\ell = \tilde{y}^\ell$.
- Nur IV hängt von y^ℓ ab.
- Schreibt man $x1, \ldots, x^s$ als Spalten einer Matrix $Q \in \mathbb{R}^{n \times s}$ (also $q_{j,\ell} = x_j^\ell$) und $y1, \ldots, y1$ als Spalten einer Matrix $R \in \mathbb{R}^{m \times s}$ (also $r_{j,\ell} = y_j^\ell$), so ist

$$\begin{aligned} r_{j,\ell} &= y_j^\ell = [P \cdot x^\ell]_j = p_{j,1} x_1^\ell + \ldots + p_{j,n} x_n \\ &= p_{j,1} q_{1,\ell} + \ldots + p_{j,n} q_{n,\ell} \end{aligned}$$

3.14 Matrix-Multiplikation

(a) Für $m, n, q \in \mathbb{N}$ und Matrizen $P \in \mathbb{R}^{m\times n}$ und $Q^{n\times s}$ definieren wir das Produkt $R = P \cdot Q \in \mathbb{R}^{m\times s}$ durch

$$[P \cdot Q]_{j,\ell} = \sum_{k=1}^{n} p_{j,k} q_{k,\ell} = p_{j,1} q_{1,\ell} + \ldots + p_{j,n} q_{n,\ell}.$$

(b) Das Element in der j-ten Zeile und ℓ-ten Spalte von $P \cdot Q$ ist also das Skalarprodukt der j-ten Zeile von P und der ℓ-ten Spalte von Q. Außerdem ist die ℓ-te Spalte von $P \cdot Q$ gleich dem Produkt von P mit der ℓ-ten Spalte von Q.

(c) Die j-te Zeile von $P \cdot Q$ ist gleich dem Produkt der j-ten Zeile $p_j \in \mathbb{R}^{1\times n}$ von P mit $Q \in \mathbb{R}^{n\times m}$.

3.14 Matrix-Multiplikation

(d) Seien $f : \mathbb{R}^n \to \mathbb{R}^m$ linear, also $f(x) = P \cdot x$, wobei $P \in \mathbb{R}^{m\times n}$ die Spalten $f(e1), \dots, f(e^n)$ hat, sowie $g : \mathbb{R}^q \to \mathbb{R}^n$ linear, also $g(y) = Q \cdot y$, wobei $Q \in \mathbb{R}^{n\times q}$ die Spalten $g(e1), \dots, g(e^q)$ hat. Dann ist $f \circ g : \mathbb{R}^q \to \mathbb{R}^m$ linear und

$$(f \circ g)(e^\ell) = f(g(e^\ell)) = P \cdot g(e^\ell) = P \cdot \ell\text{-te Spalte von } Q.$$

Also: $f \circ g(y) = (P \cdot Q) \cdot y$.

3.15 Satz

(a) $A \in \mathbb{R}^{m \times n}, B \in \mathbb{R}^{n \times q}, C \in \mathbb{R}^{q \times r}$
$\Longrightarrow (A \cdot B) \cdot C = A \cdot (B \cdot C)$ Assoziativität

(b) $A \in \mathbb{R}^{m \times n}, B, C \in \mathbb{R}^{n \times q} \Longrightarrow A \cdot (B + C) = A \cdot B + A \cdot C$
Distributivität (wobei die Summe von Matrizen gleicher Dimension komponentenweise erklärt ist).

3.16 Beispiel

(a) Für $A \in \mathbb{R}^{m \times n}$ und $B \in \mathbb{R}^{n \times q}$ mit $q \neq m$ ist kein Produkt B mit A definiert!

(b) Selbst wenn $A \cdot B$ und $B \cdot A$ definiert sind, ist sehr oft $A \cdot B \neq B \cdot A$.

Etwa $A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$:

$A \cdot B = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}$ aber $B \cdot A = \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix}$

3.17 Inverse Matrizen

(a) Für $n \in \mathbb{N}$ heißt $E_n = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & 0 & 0 \\ \vdots & & 1 & \vdots \\ 0 & \dots & & 1 \end{bmatrix} \in \mathbb{R}^{n \times n}$

die $n \times n$-dimensionale Einheitsmatrix (nur auf der Diagonalen Einsen sonst Nullen).

Für $A \in \mathbb{R}^{m \times n}$ gilt dann $A \cdot E_n = A$ und $E_m \cdot A = A$.

(b) Eine quadratische Matrix $A \in \mathbb{R}^{n \times n}$ heißt invertierbar, wenn es $B \in \mathbb{R}^{n \times n}$ gibt mit $A \cdot B = E_n$.

(c) Satz: $A \cdot B = E_n \Longrightarrow B \cdot A = E_n$.

3.17 Inverse Matrizen

(d) Folgerung: $A \cdot B = E_n$ und $A \cdot C = E_n \Longrightarrow B = C$.
Beweis:

(1) $A \cdot (B - C) = A \cdot B - A \cdot C = E_n - E_n = 0$
(2) $B - C = E_n \cdot (B - C) = (B \cdot A) \cdot (B - C) = B \cdot (A \cdot (B - C)) = B \cdot 0 = 0$.
(3) $B = C$.

(e) Ist also $A \in \mathbb{R}^{n \times n}$ invertierbar, so gibt es genau eine Matrix $B \in \mathbb{R}^{n \times n}$ mit $B \cdot A = E_n$. Diese Matrix heißt Inverse von A, und wir schreiben dann $B = A^{-1}$.

(f) $A = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$ ist invertierbar mit $A^{-1} = \begin{bmatrix} 1 & -\frac{2}{3} \\ 0 & 1/3 \end{bmatrix}$.

(g) $M = \begin{bmatrix} 1 & 2 \\ 4 & 8 \end{bmatrix}$ ist nicht invertierbar.
$x = \begin{bmatrix} 2 \\ -1 \end{bmatrix} \Longrightarrow M \cdot x = 0$. Wäre M invertierbar, folgte $x = M^{-1} \cdot M \cdot x = 0$, was nicht der Fall ist.

3.18 Satz

Sei $A \in \mathbb{R}^{n\times n}$ eine quadratische Matrix.

(a) Ist A invertierbar, so besitzt das $\text{LGS}_{n\times n}$ $A \cdot x = y$ für jedes $y \in \mathbb{R}^n$ genau eine Lösung, nämlich $x = A^{-1} \cdot y$.

(b) Sind alle n $\text{LGS}_{n\times n}$ $A \cdot x = e^k$ (k-ter Einheitsvektor) lösbar mit Lösung $x^k \in \mathbb{R}^n$, so ist A invertierbar und $x1, \ldots x^n$ sind die Spalten von A^{-1}.

3.18 Satz

Beweis.

(a) $x = A^{-1} \cdot y$ ist eine Lösung, weil $A \cdot (A^{-1} \cdot y) = (A \cdot A^{-1}) \cdot y = E_n \cdot y = y$. Ist $z \in \mathbb{R}^n$ irgendeine Lösung, so folgt aus Satz 3.17 (c) $z = A^{-1} \cdot A \cdot z = A^{-1} \cdot y$.

(b) Nach Definition der Matrixmultiplikation sind $A \cdot x^k = e^k$ die Spalten von $A \cdot [x1, \dots, x^n]$. Andererseits hat E_n gerade die Spalten $e1, \dots, e^n$. $\square$