

VORLESUNGSBEILAGE ZUR MATHEMATIK FÜR ÖKONOMEN

Inhalt

Teil 1: Lineare Algebra

1. Vektoren
2. Matrizen
3. Lineare Gleichungssysteme
4. Gauß – Algorithmus
5. Inverse Matrizen
6. Determinanten
7. Lineare Optimierung

Teil 2: Analysis

8. Folgen und Reihen
9. Finanzmathematik
10. Globale Extrema für Funktionen einer Variablen
11. Differenzial, Wachstumsrate, Elastizität
12. Taylorentwicklung
13. Unbestimmte Ausdrücke: Regeln von l'Hospital
14. Newton-Verfahren
15. Lokale Extrema für Funktionen mit mehreren Variablen
16. Ableiten linearer und quadratischer Funktionen in Matrixschreibweise
17. Partielles und totales Differenzial, partielle Wachstumsrate, partielle Elastizität
18. Homogene Funktionen
19. Kettenregel, totale Ableitung
20. Ableiten impliziter Funktionen
21. Extrema unter Nebenbedingungen: Lagrange-Ansatz
22. Integralrechnung
23. Differenzialgleichungen

Formelsammlung zur Analysis

1. Potenzen und Logarithmen

- $a^x \cdot a^y = a^{x+y}$, $\frac{a^x}{a^y} = a^{x-y}$, $a^x \cdot b^x = (ab)^x$, $\frac{a^x}{b^x} = \left(\frac{a}{b}\right)^x$, $(a^x)^y = a^{x \cdot y} = (a^y)^x$,
 $a^{-x} = \frac{1}{a^x}$, $a^{\frac{1}{n}} = \sqrt[n]{a}$, $a^0 = 1$.
- $\log_a(a^x) = x = a^{\log_a(x)}$, $\log_a(x \cdot y) = \log_a(x) + \log_a(y)$, $\log_a\left(\frac{x}{y}\right) = \log_a(x) - \log_a(y)$,
 $x^y = a^{y \cdot \log_a(x)}$, $\log_a(x^y) = y \cdot \log_a(x)$. Speziell für $a = e = 2,718\dots$ und $\ln(x) = \log_e(x)$ gilt:
 $\ln(e^x) = x = e^{\ln(x)}$, also $\ln(e) = 1$, $\ln(1) = 0$, $\ln(x^y) = y \cdot \ln(x)$, $\log_a(x) = \frac{\ln(x)}{\ln(a)}$.

2. Finanzmathematik

m Jahre, k Zinszahlungen pro Jahr, $n = m \cdot k$ Zinsperioden, K_n Guthaben nach n Zinsperioden,
 p Nominalzinsfuß, $i = p \% = \frac{p}{100}$ Nominalzinssatz, $q = 1 + \frac{i}{k} = 1 + \frac{p}{k \cdot 100}$ Zinsfaktor

- Einmalige Einzahlung eines Betrages K und
 - k -malige Zinsgutschrift pro Jahr: $K_n = K \cdot q^n = K \cdot \left(1 + \frac{i}{k}\right)^{k \cdot m}$
 - stetige Verzinsung: $K_m = K \cdot e^{m \cdot i}$
- Einzahlung eines Betrages E zu Beginn jeder Zinsperiode: $K_n = E \cdot q \cdot \frac{q^n - 1}{q - 1}$

3. Ableitungsregeln

Funktion	Ableitung	Funktion	Ableitung
x^a	$a \cdot x^{a-1}$	$f(x) \cdot g(x)$	$f'(x) \cdot g(x) + f(x) \cdot g'(x)$
a^x	$a^x \cdot \ln(a)$	$\frac{f(x)}{g(x)}$	$\frac{f'(x) \cdot g(x) - f(x) \cdot g'(x)}{(g(x))^2}$
$\log_a(x)$	$\frac{1}{x \cdot \ln(a)}$	$f(g(x))$	$f'(g(x)) \cdot g'(x)$

4. Elastizität: $\epsilon_f(x) = x \cdot \frac{f'(x)}{f(x)} = x \cdot [\ln(|f(x)|)]'$

Funktion	Elastizität	Funktion	Elastizität
a	0	$f(x) \cdot g(x)$	$\epsilon_f(x) + \epsilon_g(x)$
x^a	a	$\frac{f(x)}{g(x)}$	$\epsilon_f(x) - \epsilon_g(x)$
e^x	x	$f(g(x))$	$\epsilon_f(g(x)) \cdot \epsilon_g(x)$
$\ln(x)$	$\frac{1}{\ln(x)}$		

5. Taylorpolynom von $f(x)$, entwickelt an der Stelle x_0 :

$$T_n(x) = f(x_0) + f'(x_0) \cdot (x - x_0) + \frac{1}{2} \cdot f''(x_0) \cdot (x - x_0)^2 + \frac{1}{3!} \cdot f'''(x_0) \cdot (x - x_0)^3 + \dots + \frac{1}{n!} \cdot f^{(n)}(x_0) \cdot (x - x_0)^n$$

6. Newton-Verfahren

Iterationsformel: $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$

7. Homogenität einer Funktion $f(x, y)$ vom Grade r : $f(\lambda x, \lambda y) = \lambda^r \cdot f(x, y) \quad \forall \lambda \in \mathbb{R}_+, x, y \in \mathbb{R}$

8. Ableitung einer impliziten Funktion $f(x, y) = 0$: $\frac{dy}{dx}(x, y) = -\frac{f_x(x, y)}{f_y(x, y)}$

9. Lokale Extrema einer Funktion $f(x, y)$ unter einer Nebenbedingung $g(x, y) = 0$

Lagrange-Funktion: $L(x, y, \lambda) = f(x, y) + \lambda \cdot g(x, y)$

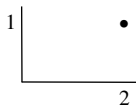
1. Vektoren

Skalar $\lambda \in \mathbb{R}$, z.B. $\lambda = 3$

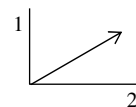
(Spalten) Vektor $\mathbf{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in \mathbb{R}^n$, z.B. $\mathbf{a} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \in \mathbb{R}^2$, $\mathbf{b} = \begin{pmatrix} 4 \\ -1 \\ 2 \end{pmatrix} \in \mathbb{R}^3$

Schreibweise: klein, fett (auch: \vec{a})

Graphisch, z.B. für $\mathbf{a} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$: Punkt



oder Pfeil



Zeilenvektor $\mathbf{a} = (a_1, \dots, a_n) \in \mathbb{R}^n$, z.B. $\mathbf{a} = (2, 1, -3) \in \mathbb{R}^3$

Transponieren \mathbf{a}' bzw. \mathbf{a}^T (\mathbf{a} transponiert)

\mathbf{a} Spaltenvektor $\Rightarrow \mathbf{a}'$ Zeilenvektor, \mathbf{b} Zeilenvektor $\Rightarrow \mathbf{b}'$ Spaltenvektor

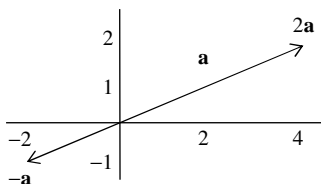
z.B. $\mathbf{a} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \Rightarrow \mathbf{a}' = (1, 2)$, $\mathbf{b} = (2, 1, 3) \Rightarrow \mathbf{b}' = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}$

Regel: $(\mathbf{a}')' = \mathbf{a}$

Spezielle Vektoren 0-Vektor: $\mathbf{0} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$, 1-Vektor: $\mathbf{1} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$, Einheitsvektor: $\mathbf{e}_i = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow i\text{-te Stelle}$

Skalar \cdot Vektor $\lambda \in \mathbb{R}$, $\mathbf{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in \mathbb{R}^n \Rightarrow \lambda \cdot \mathbf{a} = \begin{pmatrix} \lambda a_1 \\ \vdots \\ \lambda a_n \end{pmatrix}$, z.B. $3 \cdot \begin{pmatrix} 1 \\ -2 \end{pmatrix} = \begin{pmatrix} 3 \\ -6 \end{pmatrix}$, $2 \cdot (0, 1) = (0, 2)$

Graphisch, z.B. für $\mathbf{a} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$:



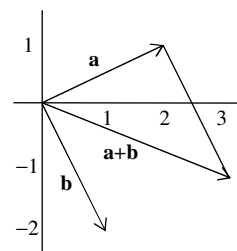
Die Menge der Punkte $\{ \lambda \mathbf{a} \mid \lambda \in \mathbb{R} \}$ stellt für einen Vektor $\mathbf{a} \in \mathbb{R}^n$ eine Gerade im \mathbb{R}^n dar, die durch den 0-Punkt geht.

Vektor \pm Vektor $\mathbf{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in \mathbb{R}^n$, $\mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \in \mathbb{R}^n \Rightarrow \mathbf{a} + \mathbf{b} = \begin{pmatrix} a_1 + b_1 \\ \vdots \\ a_n + b_n \end{pmatrix}$, $\mathbf{a} - \mathbf{b} = \begin{pmatrix} a_1 - b_1 \\ \vdots \\ a_n - b_n \end{pmatrix}$

z.B. $\begin{pmatrix} 1 \\ 2 \end{pmatrix} + \begin{pmatrix} 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 4 \\ 6 \end{pmatrix}$, $(3, 4) - (1, 2) = (2, 2)$

Graphisch, z.B. für $\mathbf{a} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ und $\mathbf{b} = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$, $\mathbf{a} + \mathbf{b} = \begin{pmatrix} 3 \\ -1 \end{pmatrix}$:

Nicht definiert sind z.B. $\begin{pmatrix} 1 \\ 2 \end{pmatrix} + (3, 4)$, $\begin{pmatrix} 1 \\ 2 \end{pmatrix} + \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$, $3 + \begin{pmatrix} 1 \\ 2 \end{pmatrix}$.



Zeilenvektor \cdot Spaltenvektor (Skalarprodukt bzw. inneres Produkt der beiden Spaltenvektoren $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$)

Für $\mathbf{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \in \mathbb{R}^n$ ergibt $\mathbf{a}' \mathbf{b} = \mathbf{b}' \mathbf{a}$ ein Skalar, definiert durch:

$$\mathbf{a}' \mathbf{b} = (a_1, \dots, a_n) \cdot \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = a_1 b_1 + \dots + a_n b_n = \sum_{i=1}^n a_i b_i = b_1 a_1 + \dots + b_n a_n = (b_1, \dots, b_n) \cdot \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \mathbf{b}' \mathbf{a}$$

Beispiele

$$(1) \quad (-1, 2, 0) \cdot \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} = -1 \cdot 2 + 2 \cdot 1 + 0 \cdot 3 = 0 = (2, 1, 3) \cdot \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix} = 2 \cdot (-1) + 1 \cdot 2 + 3 \cdot 0$$

$$(2) \quad (a_1, \dots, a_n) \cdot \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = a_1^2 + \dots + a_n^2 = \sum_{i=1}^n a_i^2 = \mathbf{a}' \mathbf{a}$$

$$(3) \quad (a_1, \dots, a_n) \cdot \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = a_1 \cdot 1 + \dots + a_n \cdot 1 = \sum_{i=1}^n a_i = \mathbf{a}' \mathbf{1} = \mathbf{1}' \mathbf{a}, \quad \mathbf{1} \in \mathbb{R}^n$$

$$(4) \quad (a_1, \dots, a_n) \cdot \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} = a_1 \cdot 0 + \dots + a_n \cdot 0 = 0 = \mathbf{a}' \mathbf{0}, \quad \mathbf{0} \in \mathbb{R}^n$$

$$(5) \quad (a_1, \dots, a_n) \cdot \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = \mathbf{a}' \mathbf{e}_i = a_i, \quad \mathbf{e}_i \in \mathbb{R}^n$$

$$(6) \quad (1, \dots, 1) \cdot \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \mathbf{1}' \mathbf{1} = n, \quad \mathbf{1} \in \mathbb{R}^n$$

Nicht definiert sind z.B. $\begin{pmatrix} 1 \\ 2 \end{pmatrix} \cdot \begin{pmatrix} 3 \\ 4 \end{pmatrix}$, $(2, 3) \cdot (4, 5)$, $(1, 2, 3) \cdot \begin{pmatrix} 4 \\ 5 \end{pmatrix}$.

Rechenregeln für Vektoren (folgen aus den Rechenregeln für Zahlen). Für $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^n$, $\lambda, \lambda_1, \lambda_2 \in \mathbb{R}$ gilt:

- (1) $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$
- (2) $\mathbf{a} + (\mathbf{b} + \mathbf{c}) = (\mathbf{a} + \mathbf{b}) + \mathbf{c}$
- (3) $(\lambda_1 + \lambda_2) \cdot \mathbf{a} = \lambda_1 \mathbf{a} + \lambda_2 \mathbf{a}$
- (4) $\lambda(\mathbf{a} + \mathbf{b}) = \lambda \mathbf{a} + \lambda \mathbf{b}$
- (5) $\lambda(\mathbf{a}' \mathbf{b}) = (\lambda \mathbf{a}') \mathbf{b} = \mathbf{a}' (\lambda \mathbf{b}) = \lambda \mathbf{a}' \mathbf{b}$
- (6) $(\mathbf{a} + \mathbf{b})' \mathbf{c} = \mathbf{a}' \mathbf{c} + \mathbf{b}' \mathbf{c}$

Linearform Lineare Funktion $f: \mathbb{R}^n \rightarrow \mathbb{R}$ mit $f(\mathbf{x}) = \mathbf{a}' \mathbf{x}$, $\mathbf{a}, \mathbf{x} \in \mathbb{R}^n$

z.B. für $\mathbf{a} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$, $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$: $f(\mathbf{x}) = \mathbf{a}' \mathbf{x} = (2, 3) \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 2x_1 + 3x_2$

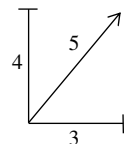
Lineare Gleichung mit n Variablen: $\mathbf{a}' \mathbf{x} = b$, $\mathbf{a}, \mathbf{x} \in \mathbb{R}^n$, $b \in \mathbb{R}$

z.B. für $\mathbf{a} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$, $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$, $b = 4$: $\mathbf{a}' \mathbf{x} = b \Leftrightarrow (1, 2, 3) \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 4 \Leftrightarrow x_1 + 2x_2 + 3x_3 = 4$

Betrag (Länge, euklidische Norm) eines Vektors $\mathbf{a} \in \mathbb{R}^n$: $|\mathbf{a}| = \sqrt{\mathbf{a}' \mathbf{a}} = \sqrt{a_1^2 + \dots + a_n^2}$

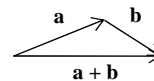
z.B. für $\mathbf{a} = \begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix}$: $|\mathbf{a}| = \sqrt{1^2 + 2^2 + (-2)^2} = \sqrt{9} = 3$

Graphisch: Länge des "Pfeils"; z.B. für $\mathbf{a} = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$ ist $|\mathbf{a}| = \sqrt{9 + 16} = 5$



Regeln für das Rechnen mit Beträgen

- (1) $|\mathbf{a}| = 0 \Leftrightarrow \mathbf{a} = \mathbf{0}$
- (2) $|\lambda \mathbf{a}| = |\lambda| \cdot |\mathbf{a}|$
- (3) $|\mathbf{a}'| = |\mathbf{a}|$
- (4) $|\mathbf{a} + \mathbf{b}| \leq |\mathbf{a}| + |\mathbf{b}|$ (Dreiecksungleichung)



Statistik (1) Arithmetischer Mittelwert $\bar{\mathbf{a}}$ eines Datenvektors $\mathbf{a} \in \mathbb{R}^n$: $\bar{\mathbf{a}} = \frac{1}{n} \mathbf{a}' \mathbf{1}$, (also auch $n \cdot \bar{\mathbf{a}} = \mathbf{a}' \mathbf{1}$),

(2) Zentrierter Vektor: $\mathbf{a}_Z = \mathbf{a} - \bar{\mathbf{a}} \cdot \mathbf{1}$,

(3) Normierter Vektor: $\mathbf{a}_N = \frac{1}{|\mathbf{a}|} \cdot \mathbf{a}$,

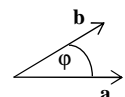
(4) Zentrierter und normierter Vektor: $\mathbf{a}_{ZN} = \frac{1}{|\mathbf{a}_Z|} \cdot \mathbf{a}_Z$; dabei ist $|\mathbf{a}_Z| = \sqrt{\mathbf{a}' \mathbf{a} - n \cdot \bar{\mathbf{a}}^2}$, denn

$$|\mathbf{a}_Z|^2 = (\mathbf{a} - \bar{\mathbf{a}} \cdot \mathbf{1})' (\mathbf{a} - \bar{\mathbf{a}} \cdot \mathbf{1}) = \mathbf{a}' \mathbf{a} - 2 \bar{\mathbf{a}} \cdot \mathbf{a}' \mathbf{1} + \bar{\mathbf{a}}^2 \cdot \mathbf{1}' \mathbf{1} = \mathbf{a}' \mathbf{a} - 2 \bar{\mathbf{a}} \cdot n \cdot \bar{\mathbf{a}} + \bar{\mathbf{a}}^2 \cdot n = \mathbf{a}' \mathbf{a} - n \cdot \bar{\mathbf{a}}^2;$$

(5) Standardisierter Vektor: $\mathbf{a}_S = \sqrt{n} \cdot \mathbf{a}_{ZN}$ (Mittelwert = 0, empirische Standardabweichung = 1)

Winkel φ_{ab} zwischen zwei Vektoren $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$;

graphisch :



Regel: $\cos(\varphi_{ab}) = \frac{\mathbf{a}' \mathbf{b}}{|\mathbf{a}| \cdot |\mathbf{b}|} = \frac{\mathbf{a}' \cdot \mathbf{b}}{|\mathbf{a}| \cdot |\mathbf{b}|} = \mathbf{a}_N' \mathbf{b}_N$, also $\varphi_{ab} = \cos^{-1}(\mathbf{a}_N' \mathbf{b}_N)$

Dieser Zusammenhang lässt sich relativ leicht mit Hilfe des Cosinus-Satzes beweisen.

φ_{ab} hängt offensichtlich nicht von der Länge der Vektoren \mathbf{a} und \mathbf{b} ab.

Wegen $0^\circ \leq \varphi_{ab} \leq 180^\circ$
 ist $1 \geq \cos(\varphi_{ab}) \geq -1$:

φ_{ab}	0°	90°	180°
$\cos(\varphi_{ab})$	1	0	-1

Bsp.: $\mathbf{a} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \Rightarrow \mathbf{a}'\mathbf{b} = 1$, $|\mathbf{a}| = \sqrt{2}$, $|\mathbf{b}| = 1$, $\cos(\varphi_{ab}) = \frac{1}{\sqrt{2}}$, $\varphi_{ab} = \cos^{-1}\left(\frac{1}{\sqrt{2}}\right) = 45^\circ$

Korrelation

In der Statistik basiert der gewöhnliche, empirische Korrelationskoeffizient r von Bravais-Pearson als Maß für die (lineare) Abhängigkeit zwischen zwei Beobachtungsreihen (Vektoren) \mathbf{a} und \mathbf{b} auf dieser Beziehung: $r(\mathbf{a}, \mathbf{b}) = \cos(\varphi(\mathbf{a}_Z, \mathbf{b}_Z)) = \mathbf{a}_{ZN}'\mathbf{b}_{ZN} = \frac{1}{n}\mathbf{a}_S'\mathbf{b}_S$.

Wie oben unter (4) erhält man $\mathbf{a}'_Z\mathbf{b}_Z = \mathbf{a}'\mathbf{b} - n\bar{a}\bar{b}$ und daraus die bekannte Formel

$$r(\mathbf{a}, \mathbf{b}) = \frac{\mathbf{a}'\mathbf{b} - n\bar{a}\bar{b}}{\sqrt{(\mathbf{a}'\mathbf{a} - n\bar{a}^2) \cdot (\mathbf{b}'\mathbf{b} - n\bar{b}^2)}} = \frac{n\mathbf{a}'\mathbf{b} - (\mathbf{a}'\mathbf{1}) \cdot (\mathbf{b}'\mathbf{1})}{\sqrt{(n\mathbf{a}'\mathbf{a} - (\mathbf{a}'\mathbf{1})^2) \cdot (n\mathbf{b}'\mathbf{b} - (\mathbf{b}'\mathbf{1})^2)}}$$

Orthogonalität

von $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$: $\mathbf{a} \perp \mathbf{b}$ (\mathbf{a} ist orthogonal zu \mathbf{b} , \mathbf{a} steht senkrecht auf \mathbf{b})

Es gilt: $\mathbf{a} \perp \mathbf{b} \Leftrightarrow \varphi_{ab} = 90^\circ \Leftrightarrow \mathbf{a}'\mathbf{b} = 0$, da $\cos(90^\circ) = 0$;

z.B. $\begin{pmatrix} 1 \\ 2 \end{pmatrix} \perp \begin{pmatrix} -2 \\ 1 \end{pmatrix}$, da $\begin{pmatrix} 1 \\ 2 \end{pmatrix}' \begin{pmatrix} -2 \\ 1 \end{pmatrix} = 0$.

Linearkombination (LK) von Vektoren

$\mathbf{b} = \lambda_1 \mathbf{a}_1 + \dots + \lambda_k \mathbf{a}_k$ ist eine LK der Vektoren $\mathbf{a}_1, \dots, \mathbf{a}_k \in \mathbb{R}^n$,

$\lambda_1, \dots, \lambda_k \in \mathbb{R}$ sind die linearen Gewichte (Koeffizienten).

Beispiele:

(1) Anwendung bei *linearen Gleichungssystemen* (LGSen)

$$\begin{cases} x + 2y = 3 \\ 2x + y = 3 \end{cases} \Leftrightarrow x \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} + y \cdot \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \end{pmatrix}$$

Das Lösen des LGS ist gleichbedeutend der Aufgabe, den Vektor $\begin{pmatrix} 3 \\ 3 \end{pmatrix}$ als LK der beiden Vektoren $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ und $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ darzustellen (hier möglich mit $x = y = 1$).

(2) *Parameterdarstellung einer Geraden* durch zwei Vektoren $\mathbf{a}_0, \mathbf{a}_1 \in \mathbb{R}^n$:

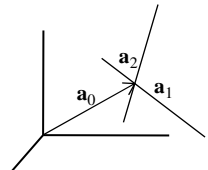
$\{\mathbf{b} \in \mathbb{R}^n \mid \mathbf{b} = \mathbf{a}_0 + \lambda \mathbf{a}_1, \lambda \in \mathbb{R}\}$; \mathbf{a}_0 : Ortsvektor, \mathbf{a}_1 : Richtungsvektor

Graphisch:  Die Parameterdarstellung ist nicht eindeutig.

(3) *Parameterdarstellung einer Ebenen* durch drei Vektoren $\mathbf{a}_0, \mathbf{a}_1, \mathbf{a}_2 \in \mathbb{R}^n$:

$\{\mathbf{b} \in \mathbb{R}^n \mid \mathbf{b} = \mathbf{a}_0 + \lambda_1 \mathbf{a}_1 + \lambda_2 \mathbf{a}_2, \lambda_1, \lambda_2 \in \mathbb{R}\}$,

\mathbf{a}_0 : Ortsvektor, $\mathbf{a}_1, \mathbf{a}_2$: Richtungsvektoren.

Graphisch: 

Konvexe Linearkombinationen: Eine LK ist konvex, wenn $\lambda_1, \dots, \lambda_k \geq 0$ und $\lambda_1 + \dots + \lambda_k = 1$.

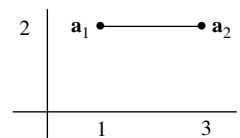
(1) *Verbindungsstrecke* zwischen zwei Punkten $\mathbf{a}_1, \mathbf{a}_2 \in \mathbb{R}^n$

$\{\mathbf{b} \in \mathbb{R}^n \mid \mathbf{b} = \lambda_1 \mathbf{a}_1 + \lambda_2 \mathbf{a}_2, \lambda_1, \lambda_2 \geq 0, \lambda_1 + \lambda_2 = 1\} =$

$\{\mathbf{b} \in \mathbb{R}^n \mid \mathbf{b} = \lambda \mathbf{a}_1 + (1-\lambda) \mathbf{a}_2, 0 \leq \lambda \leq 1\}$

(mit $\lambda_1 = \lambda, \lambda_2 = 1-\lambda$)

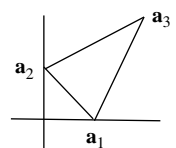
Graphisch, z.B. für $\mathbf{a}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, $\mathbf{a}_2 = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$:



(2) Dreieckige Fläche zwischen drei Punkten $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3 \in \mathbb{R}^n$

$\{\mathbf{b} \in \mathbb{R}^n \mid \mathbf{b} = \lambda_1 \mathbf{a}_1 + \lambda_2 \mathbf{a}_2 + \lambda_3 \mathbf{a}_3, \lambda_1, \lambda_2, \lambda_3 \geq 0, \lambda_1 + \lambda_2 + \lambda_3 = 1\}$

Graphisch, z.B. für $\mathbf{a}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\mathbf{a}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $\mathbf{a}_3 = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$:



(3) Die konvexe LK von k Vektoren erzeugt im \mathbb{R}^2 ein konvexes Vieleck, im \mathbb{R}^n einen konvexen Körper (Polyeder), aufgespannt durch die k Ecken.

Konvexe LK benötigt man für die Theorie des Simplex-Algorithmus (lineare Optimierung).

Lineare Abhängigkeit $\mathbf{a}_1, \dots, \mathbf{a}_k \in \mathbb{R}^n$ sind linear abhängig, wenn ein Vektor als LK der übrigen darstellbar ist;

z.B. $\begin{pmatrix} 5 \\ 8 \\ 11 \end{pmatrix} = 1 \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + 2 \cdot \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix}$, d.h. die Vektoren $\begin{pmatrix} 5 \\ 8 \\ 11 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix}$ sind linear abhängig.

Der Begriff der linearen Abhängigkeit findet Anwendung bei linearen Gleichungssystemen.

Das LGS $x \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} + y \cdot \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \end{pmatrix}$ beispielsweise kann nur dann eine Lösung besitzen, wenn der Vektor $\begin{pmatrix} 3 \\ 3 \end{pmatrix}$ linear abhängig ist von den Vektoren $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ und $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$.

Regeln

- (1) Zwei Vektoren $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ sind linear abhängig, wenn ein Vektor das Vielfache des anderen ist (die drei obigen Vektoren sind also paarweise linear unabhängig).
- (2) Mehr als n Vektoren $\mathbf{a}_i \in \mathbb{R}^n$ sind stets linear abhängig.
- (3) k von $\mathbf{0}$ verschiedene, paarweise orthogonale Vektoren sind stets linear unabhängig.
- (4) $\mathbf{a}_1, \dots, \mathbf{a}_k \in \mathbb{R}^n$ sind linear unabhängig genau dann, wenn aus $\lambda_1 \mathbf{a}_1 + \dots + \lambda_k \mathbf{a}_k = \mathbf{0}$ folgt: $\lambda_1 = \dots = \lambda_k = 0$ (der Nullvektor kann nur durch die *triviale* LK $\lambda_1 = \dots = \lambda_k = 0$ erzeugt werden).

Beispiele

- (1) $\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} -2 \\ -4 \end{pmatrix}$ sind linear abhängig, da $\begin{pmatrix} -2 \\ -4 \end{pmatrix} = -2 \begin{pmatrix} 1 \\ 2 \end{pmatrix}$; $\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ sind linear unabhängig.
- (2) $\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 3 \end{pmatrix}$ sind linear abhängig (3 Vektoren aus dem \mathbb{R}^2), aber paarweise lin. unabhängig.
- (3) $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$ sind linear unabhängig, da paarweise orthogonal.
- (4) $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ sind linear unabhängig, denn: $\lambda_1 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} + \lambda_3 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$
 $\Leftrightarrow \begin{pmatrix} 0 \\ 0 \\ \lambda_1 \end{pmatrix} + \begin{pmatrix} 0 \\ \lambda_2 \\ \lambda_2 \end{pmatrix} + \begin{pmatrix} \lambda_3 \\ \lambda_3 \\ \lambda_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{pmatrix} \lambda_3 \\ \lambda_2 + \lambda_3 \\ \lambda_1 + \lambda_2 + \lambda_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \lambda_3 = 0, \lambda_2 = 0, \lambda_1 = 0$

Spaltenvektor \cdot Zeilenvektor (dyadisches (äußeres) Produkt der beiden Vektoren $\mathbf{a} \in \mathbb{R}^m$ und $\mathbf{b} \in \mathbb{R}^n$)

$$\begin{aligned} \mathbf{a}\mathbf{b}' &= \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} \cdot (b_1, \dots, b_n) = \begin{pmatrix} a_1 b_1 & \dots & a_1 b_n \\ \vdots & & \vdots \\ a_m b_1 & \dots & a_m b_n \end{pmatrix} \in \mathbb{R}^{m \times n} \quad (\text{Matrix}) \\ &= (\mathbf{a}b_1, \dots, \mathbf{a}b_n) \quad (n \text{ Spaltenvektoren nebeneinander}) \\ &= \begin{pmatrix} a_1 \mathbf{b}' \\ \vdots \\ a_m \mathbf{b}' \end{pmatrix} \quad (m \text{ Zeilenvektoren untereinander}) \end{aligned}$$

z.B. $\begin{pmatrix} 1 \\ 2 \end{pmatrix} \cdot (2, 3, 4) = \begin{pmatrix} 2 & 3 & 4 \\ 4 & 6 & 8 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \cdot (2, 3) = \begin{pmatrix} 2 & 3 \\ 4 & 6 \\ 6 & 9 \end{pmatrix}$

Merken	Zeile \cdot Spalte = Skalar	Zeile \cdot Zeile : nicht definiert
	Spalte \cdot Zeile = Matrix	Spalte \cdot Spalte : nicht definiert

2. Matrizen

Matrix

$\mathbf{A} \in \mathbb{R}^{m \times n}$, m = Anzahl der Zeilen, n = Anzahl der Spalten

$$\begin{aligned} \mathbf{A} &= \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} = \begin{pmatrix} \mathbf{a}_{(1)} \\ \vdots \\ \mathbf{a}_{(m)} \end{pmatrix} \quad (m \text{ Zeilenvektoren } \mathbf{a}_{(i)} \in \mathbb{R}^n \text{ untereinander}) \\ &= (\mathbf{a}_1, \dots, \mathbf{a}_n) \quad (n \text{ Spaltenvektoren } \mathbf{a}_j \in \mathbb{R}^m \text{ nebeneinander}) \end{aligned}$$

$a_{ij} \in \mathbb{R}$ steht in Zeile i , Spalte j ;

$m \times n$: Ordnung der Matrix;

1×1 -Matrix: Skalar

Vektor: Matrix mit nur einer Spalte;

Zeilenvektor: Matrix mit nur einer Zeile

Transponieren \mathbf{A}' bzw. \mathbf{A}^T (\mathbf{A} transponiert)
 $\mathbf{A} \in \mathbb{R}^{m \times n} \Rightarrow \mathbf{A}' \in \mathbb{R}^{n \times m}$
 $\mathbf{A} = \begin{pmatrix} \mathbf{a}_{(1)} \\ \vdots \\ \mathbf{a}_{(m)} \end{pmatrix} \Rightarrow \mathbf{A}' = (\mathbf{a}'_{(1)}, \dots, \mathbf{a}'_{(m)}), \quad \mathbf{A} = (\mathbf{a}_1, \dots, \mathbf{a}_n) \Rightarrow \mathbf{A}' = \begin{pmatrix} \mathbf{a}'_1 \\ \vdots \\ \mathbf{a}'_n \end{pmatrix}$
z.B. $\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \in \mathbb{R}^{2 \times 3} \Rightarrow \mathbf{A}' = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix} \in \mathbb{R}^{3 \times 2}$

Regel: $(\mathbf{A}')' = \mathbf{A}$

Spezielle Matrizen Nullmatrix: $\mathbf{0} = \begin{pmatrix} 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \end{pmatrix}$

Quadratische Matrix: $\mathbf{A} \in \mathbb{R}^{n \times n}$ (Anzahl der Zeilen = Anzahl der Spalten);
Hauptdiagonale einer quadratischen Matrix: $a_{11}, a_{22}, \dots, a_{nn}$.

Symmetrische Matrix: Quadratische Matrix mit der Eigenschaft $\mathbf{A} = \mathbf{A}'$, also $a_{ij} = a_{ji}$,
d.h. \mathbf{A} ist spiegelsymmetrisch zur Hauptdiagonalen (z.B. $\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{pmatrix}$).

Einheitsmatrix: \mathbf{I}, \mathbf{I}_n bzw. \mathbf{E}_n , $\mathbf{I}_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$

(Matrix mit Einsen auf der Hauptdiagonalen und sonst nur Nullen)

Skalare Matrix: $\mathbf{S} = \lambda \mathbf{I}$, z.B. $\mathbf{S} = 3 \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix}$

Diagonalmatrix: $\mathbf{D} = \begin{pmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & d_n \end{pmatrix}$, z.B. $\mathbf{D} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ Nur die Hauptdiagonalelemente dürfen von Null verschieden sein

Dreiecksmatrix: $\mathbf{L} = \begin{pmatrix} l_{11} & 0 & \dots & 0 \\ l_{21} & l_{22} & & \vdots \\ \vdots & & \ddots & 0 \\ l_{n1} & \dots & l_{n,n-1} & l_{nn} \end{pmatrix}$, $\mathbf{R} = \begin{pmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ 0 & r_{22} & & \vdots \\ \vdots & & \ddots & r_{n-1,n} \\ 0 & \dots & 0 & r_{nn} \end{pmatrix}$

\mathbf{L} = Linke bzw. untere Dreiecksmatrix (oberhalb der Hauptdiagonalen 0)

\mathbf{R} = Rechte bzw. obere Dreiecksmatrix (unterhalb der Hauptdiagonalen 0)

Spur einer quadratischen Matrix \mathbf{A}
Spur(\mathbf{A}) = Summe der Hauptdiagonalelemente (engl.: trace, $\text{tr}(\mathbf{A}) = a_{11} + a_{22} + \dots + a_{nn}$)
z.B. für $\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{pmatrix}$: Spur(\mathbf{A}) = $\text{tr}(\mathbf{A}) = 1 + 4 + 6 = 11$

Skalar · Matrix $\lambda \in \mathbb{R}, \mathbf{A} \in \mathbb{R}^{m \times n} \Rightarrow \lambda \mathbf{A} = \begin{pmatrix} \lambda a_{11} & \dots & \lambda a_{1n} \\ \vdots & & \vdots \\ \lambda a_{m1} & \dots & \lambda a_{mn} \end{pmatrix}$, z.B. $2 \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 2 & 4 \\ 6 & 8 \end{pmatrix}$

Matrix \pm Matrix $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m \times n} \Rightarrow \mathbf{A} + \mathbf{B} = \begin{pmatrix} a_{11} + b_{11} & \dots & a_{1n} + b_{1n} \\ \vdots & & \vdots \\ a_{m1} + b_{m1} & \dots & a_{mn} + b_{mn} \end{pmatrix}$, $\mathbf{A} - \mathbf{B}$ analog

Nicht definiert sind z.B. $\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} + \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$, $3 + \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$.

Matrix · Matrix $\mathbf{A} \in \mathbb{R}^{m \times n}, \mathbf{B} \in \mathbb{R}^{n \times k} \Rightarrow \mathbf{A} \cdot \mathbf{B} = \mathbf{C} \in \mathbb{R}^{m \times k}$; also $\mathbf{A} \cdot \mathbf{B} = \mathbf{C}$. Beispiele:

(1) $\begin{pmatrix} 1 & 2 & 3 \\ 1 \times 3 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix} = 20$ $\begin{matrix} 1 \times 1 \\ 3 \times 1 \end{matrix}$ (2) $\begin{pmatrix} 1 \\ 2 \end{pmatrix} \cdot \begin{pmatrix} 2 & 3 & 4 \\ 1 \times 3 \end{pmatrix} = \begin{pmatrix} 2 & 3 & 4 \\ 4 & 6 & 8 \end{pmatrix}$ $\begin{matrix} 2 \times 1 \\ 1 \times 3 \\ 2 \times 3 \end{matrix}$

$$\begin{aligned}
(3) \quad & \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}_{2 \times 3} \cdot \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix}_{3 \times 1} = \begin{pmatrix} (1, 2, 3) \\ (3, 1, 2) \end{pmatrix}_{2 \times 3} \cdot \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix}_{3 \times 1} = \begin{pmatrix} (1, 2, 3) \cdot \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix} \\ (3, 1, 2) \cdot \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix} \end{pmatrix}_{2 \times 1} = \begin{pmatrix} 20 \\ 17 \end{pmatrix} \\
(4) \quad & (1, 2, 3)_{1 \times 3} \cdot \begin{pmatrix} 2 & 2 \\ 3 & 1 \\ 4 & 3 \end{pmatrix}_{3 \times 2} = (1, 2, 3)_{1 \times 3} \cdot \left(\begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} \right)_{3 \times 2} = \left((1, 2, 3) \cdot \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix}, (1, 2, 3) \cdot \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} \right)_{1 \times 2} = (20, 13) \\
(5) \quad & \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}_{2 \times 3} \cdot \begin{pmatrix} 2 & 2 \\ 3 & 1 \\ 4 & 3 \end{pmatrix}_{3 \times 2} = \begin{pmatrix} (1, 2, 3) \\ (3, 1, 2) \end{pmatrix}_{2 \times 3} \cdot \left(\begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} \right)_{3 \times 2} = \begin{pmatrix} (1, 2, 3) \cdot \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix}, (1, 2, 3) \cdot \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} \\ (3, 1, 2) \cdot \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix}, (3, 1, 2) \cdot \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} \end{pmatrix}_{2 \times 2} = \begin{pmatrix} 20 & 13 \\ 17 & 13 \end{pmatrix}
\end{aligned}$$

Verbal: Jede Zeile der linken Matrix **A** wird mit jeder Spalte der rechten Matrix **B** multipliziert.

Dazu muss gelten: Anzahl der Spalten von **A** = Anzahl der Zeilen von **B**.

Mit Hilfe des Schemas von Falk lässt sich die Matrixmultiplikation einfach darstellen:

Schema von Falk Beispiel $\mathbf{A} \cdot \mathbf{B} = \mathbf{C}$:

$$\begin{array}{ccc}
& & \begin{pmatrix} 2 & 3 \\ 3 & 2 \\ 4 & 1 \end{pmatrix} \mathbf{B} \\
\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \end{pmatrix} \mathbf{A} & & \begin{pmatrix} 20 & 10 \\ 29 & 16 \end{pmatrix} \mathbf{C}
\end{array}$$

Allgemein für $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{B} \in \mathbb{R}^{n \times k}$ besagt das Schema:

$$\begin{array}{ccc}
& & \begin{matrix} j\text{-te Spalte} \\ \downarrow \\ \begin{pmatrix} b_{11} & \cdots & b_{1j} & \cdots & b_{1n} \\ \vdots & & \vdots & & \vdots \\ b_{n1} & \cdots & b_{nj} & \cdots & b_{nk} \end{pmatrix} \mathbf{B} \end{matrix} \\
i\text{-te Zeile} \rightarrow \begin{pmatrix} a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \mathbf{A} & \begin{pmatrix} \vdots \\ \cdots & c_{ij} & \cdots \\ \vdots \end{pmatrix} \mathbf{C} & \begin{matrix} \text{d.h. für } i = 1, \dots, m \text{ und } j = 1, \dots, k \text{ ist} \\ c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \dots + a_{in}b_{nj} = \sum_{l=1}^n a_{il} b_{lj} \end{matrix}
\end{array}$$

Interpretation der Matrixmultiplikation $\mathbf{A} \cdot \mathbf{B}$, $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{B} \in \mathbb{R}^{n \times k}$ in Vektordarstellung (2 Varianten)

(1) Variante 1 entspricht dem Schema von Falk: **A** besteht aus Zeilen, **B** aus Spalten;

$$\mathbf{A} = \begin{pmatrix} \mathbf{a}_{(1)} \\ \vdots \\ \mathbf{a}_{(m)} \end{pmatrix}, \quad \mathbf{a}_{(i)} \in \mathbb{R}^n \text{ Zeilenvektoren, } \quad \mathbf{B} = (\mathbf{b}_1, \dots, \mathbf{b}_k), \quad \mathbf{b}_j \in \mathbb{R}^n \text{ Spaltenvektoren;}$$

$$\mathbf{A} \cdot \mathbf{B} = \begin{pmatrix} \mathbf{a}_{(1)} \\ \vdots \\ \mathbf{a}_{(m)} \end{pmatrix} \cdot (\mathbf{b}_1, \dots, \mathbf{b}_k) = \begin{pmatrix} \mathbf{a}_{(1)} \mathbf{b}_1 & \cdots & \mathbf{a}_{(1)} \mathbf{b}_k \\ \vdots & & \vdots \\ \mathbf{a}_{(m)} \mathbf{b}_1 & \cdots & \mathbf{a}_{(m)} \mathbf{b}_k \end{pmatrix}.$$

$\mathbf{A} \cdot \mathbf{B}$ ist ein äußeres Produkt von Vektoren, wobei die $\mathbf{a}_{(i)} \cdot \mathbf{b}_j$ innere Produkte sind.

(2) Variante 2: **A** besteht aus Spalten, **B** aus Zeilen;

$$\mathbf{A} = (\mathbf{a}_1, \dots, \mathbf{a}_n), \quad \mathbf{a}_i \in \mathbb{R}^m \text{ Spaltenvektoren, } \quad \mathbf{B} = \begin{pmatrix} \mathbf{b}_{(1)} \\ \vdots \\ \mathbf{b}_{(n)} \end{pmatrix}, \quad \mathbf{b}_{(i)} \in \mathbb{R}^k \text{ Zeilenvektoren;}$$

$$\mathbf{A} \cdot \mathbf{B} = (\mathbf{a}_1, \dots, \mathbf{a}_n) \cdot \begin{pmatrix} \mathbf{b}_{(1)} \\ \vdots \\ \mathbf{b}_{(n)} \end{pmatrix} = \mathbf{a}_1 \mathbf{b}_{(1)} + \dots + \mathbf{a}_n \mathbf{b}_{(n)}.$$

$\mathbf{A} \cdot \mathbf{B}$ ist ein inneres Produkt von Vektoren, wobei die $\mathbf{a}_i \mathbf{b}_{(i)}$ äußere Produkte sind.

$$\mathbf{a}_i \mathbf{b}_{(i)} = \begin{pmatrix} a_{1i} \\ \vdots \\ a_{mi} \end{pmatrix} \cdot (b_{i1}, \dots, b_{ik}) = \begin{pmatrix} a_{1i} b_{i1} & \cdots & a_{1i} b_{ik} \\ \vdots & & \vdots \\ a_{mi} b_{i1} & \cdots & a_{mi} b_{ik} \end{pmatrix}$$

ist also eine $m \times k$ Matrix, d.h. $\mathbf{A} \cdot \mathbf{B}$ die Summe von n solchen Matrizen.

$$\begin{aligned} \text{Beispiel: } & \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} 2 & 2 \\ 3 & 1 \\ 4 & 3 \end{pmatrix} = \left(\begin{pmatrix} 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \end{pmatrix} \right) \cdot \begin{pmatrix} (2, 2) \\ (3, 1) \\ (4, 3) \end{pmatrix} = \\ & = \begin{pmatrix} 1 \\ 3 \end{pmatrix} \cdot (2, 2) + \begin{pmatrix} 2 \\ 1 \end{pmatrix} \cdot (3, 1) + \begin{pmatrix} 3 \\ 2 \end{pmatrix} \cdot (4, 3) = \begin{pmatrix} 2 & 2 \\ 6 & 6 \end{pmatrix} + \begin{pmatrix} 6 & 2 \\ 3 & 1 \end{pmatrix} + \begin{pmatrix} 12 & 9 \\ 8 & 6 \end{pmatrix} = \begin{pmatrix} 20 & 13 \\ 17 & 13 \end{pmatrix} \end{aligned}$$

Lineare Gleichungssysteme in Matrixschreibweise (Beispiel)

$$\begin{aligned} x_1 + 2x_2 = 5 \\ 2x_1 - x_2 = 0 \end{aligned} & \Leftrightarrow \begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 5 \\ 0 \end{pmatrix} & \text{(zeilenweise; Schreibweise 1)} \\ & \Leftrightarrow \begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 5 \\ 0 \end{pmatrix} & \text{(Matrixschreibweise)} \\ & \Leftrightarrow \mathbf{A} \cdot \mathbf{x} = \mathbf{b} & \text{(Matrixschreibweise)} \\ & \Leftrightarrow \begin{pmatrix} 1 \\ 2 \end{pmatrix} \cdot x_1 + \begin{pmatrix} 2 \\ -1 \end{pmatrix} \cdot x_2 = \begin{pmatrix} 5 \\ 0 \end{pmatrix} & \text{(spaltenweise; Schreibweise 2)} \end{aligned}$$

Funktionen

- (1) Lineare Funktion: $f(\mathbf{x}) = \mathbf{A} \cdot \mathbf{x}$ mit $\mathbf{A} \in \mathbb{R}^{m \times n}$ und $\mathbf{x} \in \mathbb{R}^n$ ist eine lineare Funktion vom \mathbb{R}^n in den \mathbb{R}^m ; beispielsweise für $\mathbf{A} = \begin{pmatrix} 1 & 2 & 0 \\ 0 & -2 & 1 \end{pmatrix}$ ordnet die Funktion f jedem Vektor $\mathbf{x} \in \mathbb{R}^3$ den Vektor $\mathbf{y} = \mathbf{A} \cdot \mathbf{x} = \begin{pmatrix} x_1 + 2x_2 \\ -2x_2 + x_3 \end{pmatrix}$ zu, z.B. $\mathbf{x} = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$ den Vektor $\mathbf{y} = \mathbf{A} \cdot \mathbf{x} = \begin{pmatrix} 7 \\ -3 \end{pmatrix}$.
- (2) Quadratische Funktion: $f(\mathbf{x}) = \mathbf{x}' \mathbf{A} \mathbf{x}$ mit $\mathbf{A} \in \mathbb{R}^{n \times n}$ und $\mathbf{x} \in \mathbb{R}^n$ ist eine quadratische Funktion von \mathbb{R}^n nach \mathbb{R} ; beispielsweise für $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ ordnet die Funktion f jedem Vektor $\mathbf{x} \in \mathbb{R}^2$ die Zahl $y = \mathbf{x}' \mathbf{A} \mathbf{x} = x_1^2 + 5x_1x_2 + 4x_2^2$ zu, z.B. $\mathbf{x} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ die Zahl $y = \mathbf{x}' \mathbf{A} \mathbf{x} = 18$.

Rechenregeln

für Matrizen (Voraussetzung: Die Matrizen "passen" zueinander)

- (1) $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$ (2) $(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C})$
(3) $(\lambda_1 + \lambda_2) \mathbf{A} = \lambda_1 \mathbf{A} + \lambda_2 \mathbf{A}$ (4) $\lambda(\mathbf{A} + \mathbf{B}) = \lambda \mathbf{A} + \lambda \mathbf{B}$
(5) $\lambda(\mathbf{A}\mathbf{B}) = (\lambda \mathbf{A})\mathbf{B} = \mathbf{A}(\lambda \mathbf{B})$ (6) $(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{A}\mathbf{C} + \mathbf{B}\mathbf{C}$
(7) $(\mathbf{A} + \mathbf{B})' \mathbf{C} = \mathbf{A}' \mathbf{C} + \mathbf{B}' \mathbf{C}$ (8) $(\mathbf{A}\mathbf{B})\mathbf{C} = \mathbf{A}(\mathbf{B}\mathbf{C})$
(9) $(\mathbf{A}\mathbf{B})' = \mathbf{B}' \mathbf{A}'$, also auch $(\mathbf{A}\mathbf{B}\mathbf{C})' = \mathbf{C}' \mathbf{B}' \mathbf{A}'$ etc.

(folgt aus der Schreibweise 2 der Matrix-Multiplikation unter Verwendung von $(\mathbf{a}\mathbf{b}')' = \mathbf{b}\mathbf{a}'$)

- (10) $\mathbf{A} \cdot \mathbf{0} = \mathbf{0}, \mathbf{0} \cdot \mathbf{A} = \mathbf{0}$ (11) $\mathbf{A} \cdot \mathbf{I} = \mathbf{A} = \mathbf{I} \cdot \mathbf{A}$, speziell $\mathbf{I} \cdot \mathbf{I} = \mathbf{I}$

- (12) $\mathbf{S} \cdot \mathbf{A} = (\lambda \mathbf{I}) \mathbf{A} = \lambda \mathbf{A} = \mathbf{A} \cdot \mathbf{S}$, \mathbf{S} skalare Matrix

$$(13) \mathbf{D} \cdot \mathbf{A} = \begin{pmatrix} d_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & d_n \end{pmatrix} \cdot \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix} = \begin{pmatrix} d_1 a_{11} & \cdots & d_1 a_{1m} \\ \vdots & & \vdots \\ d_n a_{n1} & \cdots & d_n a_{nm} \end{pmatrix} = \begin{pmatrix} d_1 \mathbf{a}_{(1)} \\ \vdots \\ d_n \mathbf{a}_{(n)} \end{pmatrix},$$

$\mathbf{D} \in \mathbb{R}^{n \times n}$ diagonal, $\mathbf{a}_{(i)}$ Zeilen von $\mathbf{A} \in \mathbb{R}^{n \times m}$,

$$\text{z.B. } \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 1 \cdot (1, 2) \\ 2 \cdot (3, 4) \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 6 & 8 \end{pmatrix}$$

$$\mathbf{A} \cdot \mathbf{D} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} d_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & d_n \end{pmatrix} = \begin{pmatrix} d_1 a_{11} & \cdots & d_n a_{1n} \\ \vdots & & \vdots \\ d_1 a_{m1} & \cdots & d_n a_{mn} \end{pmatrix} = (d_1 \mathbf{a}_1, \dots, d_n \mathbf{a}_n),$$

$\mathbf{D} \in \mathbb{R}^{n \times n}$ diagonal, \mathbf{a}_i Spalten von $\mathbf{A} \in \mathbb{R}^{m \times n}$,

$$\text{z.B. } \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} = \left(1 \cdot \begin{pmatrix} 1 \\ 3 \end{pmatrix}, 2 \cdot \begin{pmatrix} 2 \\ 4 \end{pmatrix} \right) = \begin{pmatrix} 1 & 4 \\ 3 & 8 \end{pmatrix}$$

$$\mathbf{A} \cdot \mathbf{D} \cdot \mathbf{B} = (d_1 \mathbf{a}_1, \dots, d_n \mathbf{a}_n) \cdot \begin{pmatrix} \mathbf{b}_{(1)} \\ \vdots \\ \mathbf{b}_{(n)} \end{pmatrix} = d_1 \mathbf{a}_1 \mathbf{b}_{(1)} + \dots + d_n \mathbf{a}_n \mathbf{b}_{(n)},$$

$\mathbf{D} \in \mathbb{R}^{n \times n}$ diagonal, \mathbf{a}_i Spalten von $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{b}_{(i)}$ Zeilen von $\mathbf{B} \in \mathbb{R}^{n \times k}$,

$$\begin{aligned} \text{z.B. } & \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \cdot \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \cdot \begin{pmatrix} 4 & 3 \\ 2 & 1 \end{pmatrix} = 2 \cdot \begin{pmatrix} 1 \\ 3 \end{pmatrix} \cdot (4, 3) + 3 \cdot \begin{pmatrix} 2 \\ 4 \end{pmatrix} \cdot (2, 1) \\ & = 2 \cdot \begin{pmatrix} 4 & 3 \\ 12 & 9 \end{pmatrix} + 3 \cdot \begin{pmatrix} 4 & 2 \\ 8 & 4 \end{pmatrix} = \begin{pmatrix} 20 & 12 \\ 48 & 30 \end{pmatrix} \end{aligned}$$

- (14) Für das Rechnen mit der Spur einer Matrix gilt:

(a) $\text{Spur}(\mathbf{A}) = \text{Spur}(\mathbf{A}')$, $\mathbf{A} \in \mathbb{R}^{n \times n}$

- (b) $\text{Spur}(\mathbf{A} \pm \mathbf{B}) = \text{Spur}(\mathbf{A}) \pm \text{Spur}(\mathbf{B}), \mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times n}$
- (c) $\text{Spur}(\mathbf{a}\mathbf{b}') = \text{Spur} \begin{pmatrix} a_1 b_1 & \cdots & a_1 b_n \\ \vdots & & \vdots \\ a_n b_1 & \cdots & a_n b_n \end{pmatrix} = a_1 b_1 + \cdots + a_n b_n = \mathbf{a}'\mathbf{b}, \mathbf{a}, \mathbf{b} \in \mathbb{R}^n$
- (d) $\text{Spur}(\mathbf{A} \cdot \mathbf{B}) = \text{Spur}(\mathbf{B} \cdot \mathbf{A}), \mathbf{A} \in \mathbb{R}^{n \times m}, \mathbf{B} \in \mathbb{R}^{m \times n}$, denn
 $\mathbf{A} \cdot \mathbf{B} = \mathbf{a}_1 \mathbf{b}_{(1)} + \dots + \mathbf{a}_n \mathbf{b}_{(n)}$ (siehe Multiplikation, Variante 2)
 $\Rightarrow \text{Spur}(\mathbf{A} \cdot \mathbf{B}) = \text{Spur}(\mathbf{a}_1 \mathbf{b}_{(1)} + \dots + \mathbf{a}_n \mathbf{b}_{(n)}) = \text{Spur}(\mathbf{a}_1 \mathbf{b}_{(1)}) + \dots + \text{Spur}(\mathbf{a}_n \mathbf{b}_{(n)})$ (nach (b))
 $= \mathbf{b}_{(1)} \mathbf{a}_1 + \dots + \mathbf{b}_{(n)} \mathbf{a}_n$ (nach (c), da $\mathbf{a}_i \mathbf{b}_{(i)}$ äußere, $\mathbf{b}_{(i)} \mathbf{a}_i$ innere Produkte)
 $= \text{Spur}(\mathbf{b}_{(1)} \mathbf{a}_1 + \dots + \mathbf{b}_{(n)} \mathbf{a}_n)$ (die $\mathbf{b}_{(i)} \mathbf{a}_i$ sind Skalare)
 $= \text{Spur}(\mathbf{B} \cdot \mathbf{A})$ (siehe Multiplikation, Variante 1)

Abweichend von den Rechenregeln für Zahlen gilt:

(1) $\mathbf{A} \cdot \mathbf{B} \neq \mathbf{B} \cdot \mathbf{A}$ i.a.; z.B. $\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 6 & 8 \end{pmatrix} \neq \begin{pmatrix} 1 & 4 \\ 3 & 8 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$

Wenn $\mathbf{A} \cdot \mathbf{B}$ definiert ist, muss $\mathbf{B} \cdot \mathbf{A}$ nicht einmal definiert sein; z.B. für $\mathbf{A} \in \mathbb{R}^{2 \times 3}, \mathbf{B} \in \mathbb{R}^{3 \times 4}$.

(2) $\mathbf{A} \cdot \mathbf{B} = \mathbf{0} \not\Rightarrow \mathbf{A} = \mathbf{0}$ oder $\mathbf{B} = \mathbf{0}$; z.B. $\begin{pmatrix} 1 & 2 \\ 0,5 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & -2 \\ -0,5 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$

- (3) $\mathbf{A}^2 = \mathbf{A} \cdot \mathbf{A}$ ist i.a. nicht definiert (nur für quadratische Matrizen);
aber $\mathbf{A}'\mathbf{A}$ und $\mathbf{A}\mathbf{A}'$ sind stets definiert und symmetrisch.

Definitheit

einer quadratischen Matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$

- (1) \mathbf{A} ist *positiv definit*, wenn $\mathbf{x}'\mathbf{A}\mathbf{x} > 0 \quad \forall \mathbf{x} \in \mathbb{R}^n, \mathbf{x} \neq \mathbf{0}$
- (2) \mathbf{A} ist *positiv semidefinit*, wenn $\mathbf{x}'\mathbf{A}\mathbf{x} \geq 0 \quad \forall \mathbf{x} \in \mathbb{R}^n, \mathbf{x} \neq \mathbf{0}$
- (3) \mathbf{A} ist *negativ definit*, wenn $\mathbf{x}'\mathbf{A}\mathbf{x} < 0 \quad \forall \mathbf{x} \in \mathbb{R}^n, \mathbf{x} \neq \mathbf{0}$
- (4) \mathbf{A} ist *negativ semidefinit*, wenn $\mathbf{x}'\mathbf{A}\mathbf{x} \leq 0 \quad \forall \mathbf{x} \in \mathbb{R}^n, \mathbf{x} \neq \mathbf{0}$
- (5) \mathbf{A} ist *indefinit*, wenn es Vektoren $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ gibt mit $\mathbf{x}'\mathbf{A}\mathbf{x} > 0$ und $\mathbf{y}'\mathbf{A}\mathbf{y} < 0$

Dabei ist $Q(\mathbf{x}) = \mathbf{x}'\mathbf{A}\mathbf{x}$ eine *quadratische Form* (quadratische Funktion von \mathbb{R}^n nach \mathbb{R}).

Beispiel: Matrizen der Form $\mathbf{M} = \mathbf{A}'\mathbf{A}$ oder $\mathbf{M} = \mathbf{A}\mathbf{A}'$ sind stets positiv semidefinit, denn
 $\mathbf{x}'\mathbf{M}\mathbf{x} = \mathbf{x}'\mathbf{A}'\mathbf{A}\mathbf{x} = (\mathbf{A}\mathbf{x})'(\mathbf{A}\mathbf{x}) = \mathbf{y}'\mathbf{y} = y_1^2 + \dots + y_n^2 \geq 0, \mathbf{y} = \mathbf{A}\mathbf{x}$.

Die Definitheit von Matrizen spielt eine wichtige Rolle in der Differenzialrechnung für Funktionen mit mehreren Variablen. Notwendige Bedingungen für das Vorliegen von Definitheit ergeben sich, wenn man für \mathbf{x} speziell die Einheitsvektoren einsetzt: \mathbf{A} kann nur dann positiv (negativ) definit sein, wenn alle Hauptdiagonalelemente positiv (negativ) sind. Tauchen auf der Hauptdiagonalen sowohl positive als auch negative Zahlen auf, so ist dies ein hinreichendes (aber nicht notwendiges) Kriterium für Indefinitheit. Genauere Kriterien zur Überprüfung einer Matrix auf Definitheit liefern später die Determinanten.

Idempotente Matrizen Eine quadratische Matrix \mathbf{A} heißt idempotent, wenn $\mathbf{A}^2 = \mathbf{A}$ ist. Beispiele:

(1) $\mathbf{A} = \begin{pmatrix} 4 & 2 \\ -6 & -3 \end{pmatrix}$ ist idempotent, denn $\mathbf{A}^2 = \begin{pmatrix} 4 & 2 \\ -6 & -3 \end{pmatrix} \cdot \begin{pmatrix} 4 & 2 \\ -6 & -3 \end{pmatrix} = \begin{pmatrix} 4 & 2 \\ -6 & -3 \end{pmatrix} = \mathbf{A}$.

(2) Für $\mathbf{v} \in \mathbb{R}^n$ mit $|\mathbf{v}| = 1$ ist $\mathbf{A} = \mathbf{v}\mathbf{v}'$ idempotent, denn
 $|\mathbf{v}| = \sqrt{\mathbf{v}'\mathbf{v}} = 1 \Rightarrow \mathbf{v}'\mathbf{v} = 1 \Rightarrow \mathbf{A}^2 = \mathbf{A} \cdot \mathbf{A} = (\mathbf{v}\mathbf{v}')(\mathbf{v}\mathbf{v}') = \mathbf{v}(\mathbf{v}'\mathbf{v})\mathbf{v}' = \mathbf{v}\mathbf{v}' = \mathbf{A}$.

Regel: (1) Ist \mathbf{A} idempotent, dann auch $\mathbf{I} - \mathbf{A}$, denn $(\mathbf{I} - \mathbf{A})(\mathbf{I} - \mathbf{A}) = \mathbf{I} - \mathbf{A} - \mathbf{A} + \mathbf{A}^2 = \mathbf{I} - \mathbf{A}$.

(2) Ist \mathbf{A} idempotent, dann auch \mathbf{A}' , denn $\mathbf{A}'\mathbf{A}' = (\mathbf{A}\mathbf{A})' = \mathbf{A}'$.

Idempotente Matrizen sind in der Statistik von Bedeutung.

Orthonormale Matrizen Eine quadratische Matrix \mathbf{A} ist orthonormal, wenn alle Spaltenvektoren \mathbf{a}_i aus \mathbf{A}

- (1) die Länge 1 besitzen, d.h. $|\mathbf{a}_i| = 1 \quad \forall i$, und
- (2) paarweise orthogonal zueinander sind, d.h. $\mathbf{a}_i' \mathbf{a}_j = 0 \quad \forall i \neq j$.

Beispiel: $\mathbf{A} = \frac{1}{5} \begin{pmatrix} 3 & -4 \\ 4 & 3 \end{pmatrix}$ ist orthonormal, denn für $\mathbf{a}_1 = \frac{1}{5} \begin{pmatrix} 3 \\ 4 \end{pmatrix}$ und $\mathbf{a}_2 = \frac{1}{5} \begin{pmatrix} -4 \\ 3 \end{pmatrix}$ erhält man
 $|\mathbf{a}_1| = 1 = |\mathbf{a}_2|$ sowie $\mathbf{a}_1' \mathbf{a}_2 = 0$.

Regel: Für eine orthonormale Matrix \mathbf{A} gilt: $\mathbf{A}'\mathbf{A} = \mathbf{I} = \mathbf{A}\mathbf{A}'$.

Begründung: Mit $\mathbf{A} = (\mathbf{a}_1, \dots, \mathbf{a}_n) \in \mathbb{R}^{n \times n}$, $\mathbf{a}_i \in \mathbb{R}^n$ Spaltenvektoren von \mathbf{A} , erhält man

$$\mathbf{A}'\mathbf{A} = \begin{pmatrix} \mathbf{a}_1' \\ \vdots \\ \mathbf{a}_n' \end{pmatrix} \cdot (\mathbf{a}_1, \dots, \mathbf{a}_n) = \begin{pmatrix} \mathbf{a}_1'\mathbf{a}_1 & \dots & \mathbf{a}_1'\mathbf{a}_n \\ \vdots & & \vdots \\ \mathbf{a}_n'\mathbf{a}_1 & \dots & \mathbf{a}_n'\mathbf{a}_n \end{pmatrix} = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} = \mathbf{I}.$$

Zum Nachweis von $\mathbf{A}\mathbf{A}' = \mathbf{I}$ siehe Regel (6) in Abschnitt 5.

Erfüllt \mathbf{A} zwar die Bedingungen (1) und (2), ist aber nicht quadratisch, dann gilt zwar weiterhin

$$\mathbf{A}'\mathbf{A} = \mathbf{I}, \text{ nicht aber } \mathbf{A}\mathbf{A}' = \mathbf{I} \quad (\text{z.B. f\u00fcr } \mathbf{A} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}).$$

Rang

einer Matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\text{rg}(\mathbf{A})$

Zeilenrang: Anzahl linear unabhängiger Zeilenvektoren von \mathbf{A}

Spaltenrang: Anzahl linear unabhängiger Spaltenvektoren von \mathbf{A}

Regeln: Für jede Matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ gilt:

- (1) Zeilenrang = Spaltenrang (= $\text{rg}(\mathbf{A})$)
- (2) $\text{rg}(\mathbf{A}) \leq \min(m, n)$
- (3) Der Rang einer Matrix ändert sich nicht bei Addition des Vielfachen einer Zeile (Spalte) zu einer anderen Zeile (Spalte)
- (4) $\text{rg}(\mathbf{A} \cdot \mathbf{B}) \leq \min(\text{rg}(\mathbf{A}), \text{rg}(\mathbf{B}))$, $\mathbf{B} \in \mathbb{R}^{n \times k}$; $\text{rg}(\mathbf{A}'\mathbf{A}) = \text{rg}(\mathbf{A}) = \text{rg}(\mathbf{A}\mathbf{A}')$

Der Rang ist nur bei kleinen (2×2 , 2×3 , 3×2) oder sehr einfachen Matrizen offensichtlich.

Ansonsten bestimmt man den Rang z.B. mittels Gauß-Algorithmus (nächster Abschnitt).

$$\text{rg} \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} = 2, \quad \text{rg} \begin{pmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 4 \end{pmatrix} = 2, \quad \text{rg} \begin{pmatrix} 1 & 2 & 3 \\ 3 & 4 & 5 \end{pmatrix} = 2, \quad \text{rg} \begin{pmatrix} 1 & -2 \\ 2 & -4 \end{pmatrix} = 1,$$

Beispiele:

$$\text{rg} \begin{pmatrix} 1 & 2 \\ 2 & 4 \\ 3 & 6 \end{pmatrix} = 1, \quad \text{rg} \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = 0, \quad \text{rg} \begin{pmatrix} 1 & 1 & 2 \\ 2 & 2 & 3 \\ 3 & 3 & 4 \end{pmatrix} = 2, \quad \text{rg} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = 3.$$

- \mathbf{A} besitzt *vollen Rang*, wenn $\text{rg}(\mathbf{A}) = \min(m, n)$
- \mathbf{A} ist *regulär*, wenn \mathbf{A} quadratisch ist und vollen Rang besitzt
- \mathbf{A} ist *singulär*, wenn \mathbf{A} quadratisch ist und nicht vollen Rang besitzt

Untersuchen einer Matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ auf vollen Rang

$m \geq n$ (weniger Spalten als Zeilen): Spaltenvektoren auf lineare Abhängigkeit überprüfen.

$$\text{Beispiel: } \mathbf{A} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}; \quad \lambda_1 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} + \lambda_3 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow \lambda_3 = 0, \lambda_2 = 0, \lambda_1 = 0,$$

d.h. die Spaltenvektoren sind linear unabhängig, also hat \mathbf{A} vollen Rang.

$$\text{In Matrixschreibweise: } \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

oder äquivalent: $\mathbf{A}\boldsymbol{\lambda} = \mathbf{0} \Rightarrow \boldsymbol{\lambda} = \mathbf{0}$; \mathbf{A} hat also vollen Rang, wenn das *homogene* LGS

$\mathbf{A}\boldsymbol{\lambda} = \mathbf{0}$ nur die *triviale Lösung* $\boldsymbol{\lambda} = \mathbf{0}$ besitzt.

$m < n$ (weniger Zeilen als Spalten): Zeilenvektoren auf lineare Abhängigkeit überprüfen.

\mathbf{A} besitzt vollen Rang, wenn gilt: $\mathbf{A}'\boldsymbol{\lambda} = \mathbf{0} \Rightarrow \boldsymbol{\lambda} = \mathbf{0}$.

3. Lineare Gleichungssysteme

Schreibweise eines linearen $m \times n$ Gleichungssystems (LGS) mit m Gleichungen und n Variablen

$$a_{11}x_1 + \dots + a_{1n}x_n = b_1$$

$$\vdots \quad \quad \quad \vdots \quad \quad \quad \vdots$$

$$a_{m1}x_1 + \dots + a_{mn}x_n = b_m$$

$$\Leftrightarrow \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} x_1 + \dots + \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} x_n = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

$$\Leftrightarrow \mathbf{a}_1 x_1 + \dots + \mathbf{a}_n x_n = \mathbf{b}, \quad \mathbf{a}_i, \mathbf{b} \in \mathbb{R}^m \text{ Spaltenvektoren}$$

$$\Leftrightarrow (\mathbf{a}_1, \dots, \mathbf{a}_n) \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \mathbf{b}$$

$$\Leftrightarrow \mathbf{A} \cdot \mathbf{x} = \mathbf{b}$$

Lösbarkeit

eines $m \times n$ LGS $\mathbf{A} \mathbf{x} = \mathbf{b}$

Sei $(\mathbf{A}, \mathbf{b}) = (\mathbf{a}_1, \dots, \mathbf{a}_n, \mathbf{b})$ die um den Vektor \mathbf{b} erweiterte Matrix \mathbf{A} ; offensichtlich gilt:

- (1) $\text{rg}(\mathbf{A}, \mathbf{b}) = \text{rg}(\mathbf{A})$, wenn \mathbf{b} linear abhängig ist von $\mathbf{a}_1, \dots, \mathbf{a}_n$
- (2) $\text{rg}(\mathbf{A}, \mathbf{b}) = \text{rg}(\mathbf{A}) + 1$, wenn \mathbf{b} linear unabhängig ist von $\mathbf{a}_1, \dots, \mathbf{a}_n$

Für die Lösung des LGS $\mathbf{A} \mathbf{x} = \mathbf{b}$ bedeutet dies:

- (1) Das LGS besitzt mindestens eine Lösung, wenn $\text{rg}(\mathbf{A}, \mathbf{b}) = \text{rg}(\mathbf{A})$, und zwar
 - (a) genau eine Lösung, wenn $\text{rg}(\mathbf{A}) = n$,
 - (b) ∞ viele Lösungen, wenn $\text{rg}(\mathbf{A}) < n$.
- (2) Das LGS besitzt keine Lösung, wenn $\text{rg}(\mathbf{A}, \mathbf{b}) = \text{rg}(\mathbf{A}) + 1$.

Die folgende Tabelle stellt dies noch etwas ausführlicher dar.

Anzahl der Lösungen eines LGS ($m \times n$) $\mathbf{A} \mathbf{x} = \mathbf{b}$

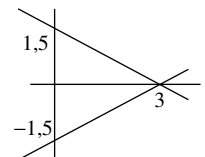
$m =$ Anzahl der Gleichungen $n =$ Anzahl der Variablen $=$ Anzahl der Spalten von \mathbf{A}	$\text{rg}(\mathbf{A}, \mathbf{b}) = \text{rg}(\mathbf{A}) + 1$	$\text{rg}(\mathbf{A}, \mathbf{b}) = \text{rg}(\mathbf{A})$	
		\mathbf{A} besitzt vollen Rang	\mathbf{A} besitzt nicht vollen Rang
$m < n$	0	∞	∞
$m \geq n$	0	1	∞

Insbesondere besagt die Tabelle:

- (1) Genau eine Lösung gibt es nur dann, wenn $m \geq n$ und $\text{rg}(\mathbf{A}, \mathbf{b}) = \text{rg}(\mathbf{A}) = \min(m, n) = n$ ist; dies ist insbesondere für jede reguläre Matrix \mathbf{A} erfüllt;
- (2) $m < n \Rightarrow$ es gibt keine oder ∞ viele Lösungen;
- (3) \mathbf{A} besitzt nicht vollen Rang \Rightarrow es gibt keine oder ∞ viele Lösungen.

Graphische Darstellung möglicher Lösungen eines LGS $\mathbf{A} \mathbf{x} = \mathbf{b}$ im \mathbb{R}^2 ($m = n = 2$); Beispiele:

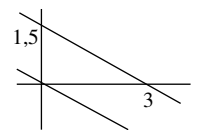
$$(1) \begin{cases} x_1 + 2x_2 = 3 \\ x_1 - 2x_2 = 3 \end{cases} \Leftrightarrow \begin{pmatrix} 1 & 2 \\ 1 & -2 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \end{pmatrix}; \text{ graphisch:}$$



Das LGS besitzt genau eine Lösung ($\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \end{pmatrix}$, im Schnittpunkt der Geraden), da die Matrix

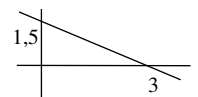
\mathbf{A} vollen Rang hat ($\text{rg}(\mathbf{A}) = 2$) und $\text{rg}(\mathbf{A}) = \text{rg}(\mathbf{A}, \mathbf{b})$, $(\mathbf{A}, \mathbf{b}) = \begin{pmatrix} 1 & 2 & 3 \\ 1 & -2 & 3 \end{pmatrix}$.

$$(2) \begin{cases} x_1 + 2x_2 = 3 \\ x_1 + 2x_2 = 0 \end{cases} \Leftrightarrow \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \end{pmatrix}; \text{ graphisch:}$$



Das LGS besitzt keine Lösung (die Geraden verlaufen parallel), da $\text{rg}(\mathbf{A}) = 1 \neq 2 = \text{rg}(\mathbf{A}, \mathbf{b})$.

$$(3) \begin{cases} x_1 + 2x_2 = 3 \\ 2x_1 + 4x_2 = 6 \end{cases} \Leftrightarrow \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 6 \end{pmatrix}; \text{ graphisch:}$$



Das LGS besitzt ∞ viele Lösungen (die Geraden fallen zusammen), da die Matrix \mathbf{A} nicht vollen Rang hat ($\text{rg}(\mathbf{A}) = 1$) und $\text{rg}(\mathbf{A}) = \text{rg}(\mathbf{A}, \mathbf{b})$.

4. Gauß¹-Algorithmus (GA)

- Ziel
- Lösen eines LGS ($m \times n$)
 - Bestimmen des Rangs einer $m \times n$ Matrix
 - Inverse einer $n \times n$ Matrix berechnen (siehe nächsten Abschnitt)

- Varianten des GA
- GA mit
 - vollständiger Elimination (1-stufig) (GA1)
 - teilweiser Elimination (2-stufig) (GA2)

Beispiel (1) (Das LGS besitzt genau eine Lösung)

$$\begin{aligned} x_1 + x_2 + 3x_3 &= 8 \\ 2x_1 + 2x_2 + 2x_3 &= 12 \\ 3x_1 + 2x_2 + 2x_3 &= 15 \end{aligned} \Leftrightarrow \begin{pmatrix} 1 & 1 & 3 \\ 2 & 2 & 2 \\ 3 & 2 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 8 \\ 12 \\ 15 \end{pmatrix}$$

In Tableauschreibweise:

Anfangstableau

Gewünschtes Endtableau (beim GA1)

x_1	x_2	x_3	r.S.
1	1	3	8
2	2	2	12
3	2	2	15

x_1	x_2	x_3	r.S.
1	0	0	c_1
0	1	0	c_2
0	0	1	c_3

also $x_1 = c_1, x_2 = c_2, x_3 = c_3$.

Vorgehen: Spaltenweises Erzeugen von Einheitsvektoren, indem das Vielfache einer Zeile zu einer anderen Zeile addiert wird.

(): Pivotelement² (muss 1 werden, der Rest der Spalte 0)

PZ: Pivotzeile (Zeile, die das Pivotelement enthält)

x_1	x_2	x_3	r.S.	
(1)	1	3	8	← PZ
2	2	2	12	−2·PZ
3	2	2	15	−3·PZ
1	1	3	8	
0	(0)	−4	−4	:(−4)
0	−1	−7	−9	·(−1)
1	1	3	8	− PZ
0	(1)	7	9	← PZ
0	0	1	1	
1	0	−4	−1	+4·PZ
0	1	7	9	−7·PZ
0	0	(1)	1	← PZ
1	0	0	3	
0	1	0	2	
0	0	1	1	

Tausche Zeile 2 mit Zeile 3

[*]

Also $x_1 = 3$
 $x_2 = 2$
 $x_3 = 1$

- Vorgehen beim GA2:
1. Stufe : obere Dreiecksmatrix erzeugen
 2. Stufe : LGS lösen durch Einsetzen von unten nach oben

- Im obigen Beispiel:
1. Stufe: beendet an der Stelle [*]
 2. Stufe: $x_3 = 1$ (Zeile 3)
 $x_2 = 9 - 7x_3 = 9 - 7 = 2$ (Zeile 2)
 $x_1 = 8 - x_2 - 3x_3 = 8 - 2 - 3 = 3$ (Zeile 1)

GA2 erfordert i.a. weniger Rechenaufwand als GA1.

Aus didaktischen Gründen wird im Weiteren dennoch GA1 verwandt.

¹ Carl Friedrich Gauß: 1777 - 1855

² Pivot (franz.): Drehpunkt

Beispiel (2)

(Das LGS besitzt ∞ viele Lösungen)

x_1	x_2	x_3	r.S.		
(1)	2	1	7	← PZ	
1	2	3	13	- PZ	
3	6	5	27	-3·PZ	
1	2	1	7		
0	(0)	2	6	: 2	Tausche Spalte 2
0	0	2	6	: 2	mit Spalte 3
x_1	x_3	x_2	r.S.		
1	1	2	7	- PZ	
0	(1)	0	3	← PZ	
0	1	0	3	- PZ	
1	0	2	4		
0	1	0	3		
0	0	0	0		← Zeile streichen (allgemeingültig)

Das LGS lautet ausgeschrieben:

$$\begin{aligned}
 x_1 + 2x_2 = 4 & \Leftrightarrow x_1 = 4 - 2x_2 & x_1 = 4 - 2 \cdot x_2 \\
 x_3 = 3 & \Leftrightarrow x_3 = 3 & x_2 = 0 + 1 \cdot x_2 \quad \leftarrow \text{(Gleichung eingefügt)} \\
 & & x_3 = 3 + 0 \cdot x_2
 \end{aligned}$$

Daraus ergibt sich die Lösungsmenge $L = \{ \mathbf{x} \in \mathbb{R}^3 \mid \mathbf{x} = \begin{pmatrix} 4 \\ 0 \\ 3 \end{pmatrix} + \lambda \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}, \lambda \in \mathbb{R} \}$,

die Parameterdarstellung einer Geraden im \mathbb{R}^3 .

Es gibt also ∞ viele Lösungen, da eine Variable (hier x_2) frei wählbar ist.

Beispiel (3)

(Das LGS besitzt keine Lösung)

x_1	x_2	x_3	r.S.		
(1)	0	2	3	← PZ	
3	2	2	7	-3·PZ	
2	2	0	5	-2·PZ	
1	0	2	3		
0	(2)	-4	-2	: 2	
0	2	-4	-1		
1	0	2	3		
0	(1)	-2	-1	← PZ	
0	2	-4	-1	-2·PZ	
1	0	2	3		
0	1	-2	-1		
0	0	0	1		← Gleichung nicht erfüllbar (widersprüchlich)

Das LGS besitzt keine Lösung.

Allgemeine Darstellung des GA

Ziel: Erzeugen einer möglichst großen Einheitsmatrix (beim GA1) bzw. Dreiecksmatrix (beim GA2) "unter" den Variablen.

Dazu wird das Anfangstableau

x_1	x_2	...	x_n	b
a_{11}	a_{12}	...	a_{1n}	b_1
\vdots	\vdots	\vdots	\vdots	\vdots
a_{m1}	a_{m2}	...	a_{mn}	b_m

so umgeformt, das folgendes Endtableau (beim GA1) entsteht:

x_1^*	x_2^*	...	x_r^*	x_{r+1}^*	...	x_n^*	\mathbf{b}^*
1	0	...	0	$a_{1,r+1}$...	a_{1n}^*	b_1^*
0	1	\ddots	\vdots	\vdots		\vdots	\vdots
\vdots	\ddots	\ddots	0	\vdots		\vdots	\vdots
0	...	0	1	$a_{r,r+1}$...	a_{rn}^*	b_r^*
0	0	0	...	0	b_{r+1}^*
\vdots			\vdots	\vdots		\vdots	\vdots
\vdots			\vdots	\vdots		\vdots	\vdots
0	0	0	...	0	b_m^*

Interpretation:

Die mit einem * indizierten Variablen deuten an, dass evtl. die Reihenfolge der x_i vertauscht wurde.

- $r = \text{Rang von } \mathbf{A}$: Die Spalten "unter" x_1^*, \dots, x_r^* sind linear unabhängig und erzeugen die übrigen Spalten.
- Ist einer der Werte $b_{r+1}^*, \dots, b_m^* \neq 0$, gibt es keine Lösung.
- Ansonsten können die Zeilen gestrichen werden. In diesem Fall gilt:

(a) Ist $r = n$, gibt es genau eine Lösung: $\begin{pmatrix} x_1^* \\ \vdots \\ x_n^* \end{pmatrix} = \begin{pmatrix} b_1^* \\ \vdots \\ b_n^* \end{pmatrix}$.

(b) Ist $r < n$, gibt es ∞ viele Lösungen: Die Variablen x_{r+1}^*, \dots, x_n^* sind frei wählbar und es

gilt $\begin{pmatrix} x_1^* \\ \vdots \\ x_r^* \end{pmatrix} = \begin{pmatrix} b_1^* \\ \vdots \\ b_r^* \end{pmatrix} - \mathbf{A}^* \cdot \begin{pmatrix} x_{r+1}^* \\ \vdots \\ x_n^* \end{pmatrix}$, $\mathbf{A}^* = \begin{pmatrix} a_{1,r+1}^* & \dots & a_{1n}^* \\ \vdots & & \vdots \\ a_{r,r+1}^* & \dots & a_{rn}^* \end{pmatrix}$.

Hilfsmittel für den *Simplex-Algorithmus* (lineare Optimierung)

- *Kanonische Form* eines LGS: Auf die Höchstzahl verschiedener Einheitsvektoren umgeformt (Endtableau GA1; Nullzeilen gestrichen)
- *Basisvariable (BV)*: x_1^*, \dots, x_r^*
- *Nichtbasisvariable (NBV)*: x_{r+1}^*, \dots, x_n^*
- *Basislösung*: $\begin{pmatrix} x_1^* \\ \vdots \\ x_r^* \end{pmatrix} = \begin{pmatrix} b_1^* \\ \vdots \\ b_r^* \end{pmatrix}$ und $\begin{pmatrix} x_{r+1}^* \\ \vdots \\ x_n^* \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$ (d.h. die NBV sind auf 0 gesetzt)
- *Basistausch*: Eine BV wird aus der Basis entfernt und dafür eine NBV neu in die Basis aufgenommen. Dazu erzeugt man in der Spalte der NBV den Einheitsvektor, der in der Spalte der BV steht (und dort verschwindet).

Beispiel

x_1	x_3	x_2	r.S.
1	0	1	3
0	1	2	4

Das LGS liegt in kanonischer Form vor. Basisvariablen: x_1, x_3 ; Nichtbasisvariable: x_2 ;

Basislösung: $\mathbf{x} = \begin{pmatrix} 3 \\ 0 \\ 4 \end{pmatrix}$. Tausche nun beispielsweise x_2 mit x_3 , d.h. in der x_2 -Spalte wird

der Einheitsvektor erzeugt, der in der x_3 -Spalte steht.

x_1	x_3	x_2	r.S.	
1	0	1	3	
0	1	(2)	4	: 2
1	0	1	3	- PZ
0	0,5	(1)	2	← PZ
1	-0,5	0	1	
0	0,5	1	2	
x_1	x_2	x_3	r.S.	Tausche Spalte 2 mit Spalte 3
1	0	-0,5	1	
0	1	0,5	2	

Basisvariablen: x_1, x_2 ; Nichtbasisvariable: x_3 ; Basislösung: $\mathbf{x} = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$.

Man könnte jetzt noch x_1 mit x_3 tauschen.

Lösen eines LGS für mehrere "rechte Seiten" gleichzeitig

Beispiel: LGS 1: $\mathbf{A}\mathbf{x} = \mathbf{b}$, LGS 2: $\mathbf{A}\mathbf{x} = \mathbf{c}$

$x_1 \dots x_n$	r.S. 1	r.S. 2	wird mittels GA1	$x_1^* \dots x_n^*$	r.S. 1	r.S. 2
\mathbf{A}	\mathbf{b}	\mathbf{c}	umgeformt zu	\mathbf{I}, \mathbf{A}^*	\mathbf{b}^*	\mathbf{c}^*

So erhält man in einem Durchgang gleichzeitig die Lösung für LGS 1 und LGS 2.

z.B. LGS 1: $x_1 + x_2 = 3$
 $2x_1 + 3x_2 = 8$ LGS 2: $x_1 + x_2 = 5$
 $2x_1 + 3x_2 = 13$

x_1	x_2	r.S. 1	r.S. 2	
(1)	1	3	5	← PZ
2	3	8	13	-2·PZ
1	1	3	5	- PZ
0	(1)	2	3	← PZ
1	0	1	2	
0	1	2	3	

Lösung von

LGS 1: $x_1 = 1$
 $x_2 = 2$,

LGS 2: $x_1 = 2$
 $x_2 = 3$.

5. Quadratische lineare Gleichungssysteme – Inverse einer quadratischen Matrix \mathbf{A}

Betrachtet sei das LGS $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$ mit $\mathbf{A} \in \mathbb{R}^{n \times n}$, $\mathbf{x}, \mathbf{b} \in \mathbb{R}^n$. Hat \mathbf{A} den vollen Rang und ist damit regulär, so ist auch $\text{rg}(\mathbf{A}, \mathbf{b}) = \text{rg}(\mathbf{A})$ und das LGS besitzt genau eine Lösung. In diesem Fall lässt sich die *Inverse* $\mathbf{A}^{-1} \in \mathbb{R}^{n \times n}$ von \mathbf{A} bestimmen. Sie ist *eindeutig* festgelegt durch die Beziehung $\boxed{\mathbf{A}^{-1}\mathbf{A} = \mathbf{I} = \mathbf{A}\mathbf{A}^{-1}}$. Durch Multiplikation des LGS von links mit \mathbf{A}^{-1} erhält man $\mathbf{A}^{-1}\mathbf{A}\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$,

also wegen $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$ und $\mathbf{I}\mathbf{x} = \mathbf{x}$ die Lösung des LGS: $\boxed{\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}}$.

Für die Funktion $f(\mathbf{x}) = \mathbf{A} \cdot \mathbf{x}$ gilt: $g(\mathbf{x}) = \mathbf{A}^{-1} \cdot \mathbf{x}$ ist die Umkehrfunktion zu f .

Inversion

einer regulären $n \times n$ Matrix \mathbf{A} mittels GA

Sei $\mathbf{A}^{-1} = (\tilde{\mathbf{a}}_1, \dots, \tilde{\mathbf{a}}_n)$ die Inverse von \mathbf{A} , d.h. \mathbf{A}^{-1} bestehe aus den Spaltenvektoren $\tilde{\mathbf{a}}_i \in \mathbb{R}^n$.

Mit $\mathbf{I} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ ist dann die Bedingung $\mathbf{A} \cdot \mathbf{A}^{-1} = \mathbf{I}$ äquivalent zu $\mathbf{A} \cdot (\tilde{\mathbf{a}}_1, \dots, \tilde{\mathbf{a}}_n) = (\mathbf{e}_1, \dots, \mathbf{e}_n)$.

Für die Spalten $\tilde{\mathbf{a}}_i$ muss daher gelten: $\mathbf{A} \cdot \tilde{\mathbf{a}}_1 = \mathbf{e}_1, \dots, \mathbf{A} \cdot \tilde{\mathbf{a}}_n = \mathbf{e}_n$,

d.h. die i -te Spalte $\tilde{\mathbf{a}}_i$ von \mathbf{A}^{-1} ergibt sich als Lösung des LGS $\mathbf{A} \cdot \tilde{\mathbf{a}}_i = \mathbf{e}_i$;

die Inversion von \mathbf{A} wird also durch das simultane Lösen dieser n Gleichungssysteme erreicht.

Beispiel: $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} 4 \\ 7 \end{pmatrix}$, $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$

Das LGS $\mathbf{A}\mathbf{x} = \mathbf{b}$ wird mittels GA1 gelöst und gleichzeitig die Inverse \mathbf{A}^{-1} berechnet.

x_1	x_2	r.S.	$\mathbf{I} \rightarrow \mathbf{A}^{-1}$		
(1)	2	4	1	0	\leftarrow PZ
2	3	7	0	1	$-2 \cdot$ PZ
1	2	4	1	0	
0	(-1)	-1	-2	1	$\cdot (-1)$
1	2	4	1	0	$-2 \cdot$ PZ
0	(1)	1	2	-1	\leftarrow PZ
1	0	2	-3	2	
0	1	1	2	-1	

Lösung des LGS: $x_1 = 2, x_2 = 1$; Inverse: $\mathbf{A}^{-1} = \begin{pmatrix} -3 & 2 \\ 2 & -1 \end{pmatrix}$

Probe: (1) $\mathbf{A} \mathbf{x} = \mathbf{b}$: $\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 7 \end{pmatrix}$
(2) $\mathbf{A} \mathbf{A}^{-1} = \mathbf{I}$: $\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} \cdot \begin{pmatrix} -3 & 2 \\ 2 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$
(3) $\mathbf{x} = \mathbf{A}^{-1} \mathbf{b}$: $\begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} -3 & 2 \\ 2 & -1 \end{pmatrix} \cdot \begin{pmatrix} 4 \\ 7 \end{pmatrix}$

Das Lösen eines LGS unter Verwendung der Beziehung $\mathbf{x} = \mathbf{A}^{-1} \mathbf{b}$ lohnt sich offensichtlich nicht, da die Berechnung von \mathbf{A}^{-1} zu aufwendig ist. Nur wenn ein LGS zukünftig mit verschiedenen, noch unbekanntem "rechten Seiten" gelöst werden soll, kann dieses Vorgehen sinnvoll sein.

Beachten: Die Inverse \mathbf{A}^{-1} ist nur für reguläre Matrizen \mathbf{A} definiert.

Regeln	für reguläre $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times n}$;	Begründung
(1)	$\mathbf{I}^{-1} = \mathbf{I}$	$\mathbf{I} \cdot \mathbf{I} = \mathbf{I}$
(2)	$(\mathbf{A}^{-1})^{-1} = \mathbf{A}$	
(3)	$(\mathbf{A}^{-1})' = (\mathbf{A}')^{-1}$	$(\mathbf{A}^{-1})' \mathbf{A}' = (\mathbf{A} \mathbf{A}^{-1})' = \mathbf{I}' = \mathbf{I}$
(4)	$(\mathbf{A} \mathbf{B})^{-1} = \mathbf{B}^{-1} \mathbf{A}^{-1}$	$(\mathbf{B}^{-1} \mathbf{A}^{-1}) (\mathbf{A} \mathbf{B}) = \mathbf{B}^{-1} (\mathbf{A}^{-1} \mathbf{A}) \mathbf{B} = \mathbf{B}^{-1} \mathbf{I} \mathbf{B} = \mathbf{B}^{-1} \mathbf{B} = \mathbf{I}$
(5)	$(\lambda \mathbf{A})^{-1} = \frac{1}{\lambda} \mathbf{A}^{-1}, \lambda \neq 0$	$(\frac{1}{\lambda} \mathbf{A}^{-1}) (\lambda \mathbf{A}) = \mathbf{A}^{-1} \mathbf{A} = \mathbf{I}$
(6)	Für eine orthonormale Matrix \mathbf{A} ist $\mathbf{A}^{-1} = \mathbf{A}'$.	Begründung: Bei orthonormalen Matrizen (siehe Abschnitt 2) gilt $\mathbf{A}' \mathbf{A} = \mathbf{I}$.
(7)	Für eine Diagonalmatrix \mathbf{D} mit den Diagonalelementen d_i ist auch \mathbf{D}^{-1} diagonal mit den Diagonalelementen $1/d_i$.	Begründung: $\mathbf{D} \mathbf{D}^{-1} = \mathbf{I}$.
(8)	Für Dreiecksmatrizen \mathbf{L} bzw. \mathbf{R} sind auch \mathbf{L}^{-1} bzw. \mathbf{R}^{-1} wieder Dreiecksmatrizen.	
(9)	Für eine symmetrische Matrix \mathbf{S} ist auch \mathbf{S}^{-1} symmetrisch.	

Anmerkung Nur reguläre Matrizen besitzen eine Inverse. Dies bedeutet insbesondere:

- (1) Wenn \mathbf{A} zwar quadratisch ist, aber nicht den vollen Rang besitzt, existiert keine Inverse.
- (2) Wenn \mathbf{A} nicht quadratisch ist, existiert keine Inverse.

Jedoch lässt sich für nicht quadratische Matrizen mit $m > n$ und vollem Rang eine Matrix \mathbf{A}^+ berechnen mit der Eigenschaft $\mathbf{A}^+ \mathbf{A} = \mathbf{I}$ (*Linksinverse* zu \mathbf{A}).

Für sie gilt aber nicht $\mathbf{A} \mathbf{A}^+ = \mathbf{I}$.

Ähnliche Matrizen $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times n}$ heißen ähnlich, wenn es eine reguläre Matrix \mathbf{M} gibt, so dass $\mathbf{A} = \mathbf{M}^{-1} \mathbf{B} \mathbf{M}$ ist.

Für ähnliche Matrizen \mathbf{A}, \mathbf{B} gilt: $\text{Spur}(\mathbf{A}) = \text{Spur}(\mathbf{B})$.

Begründung: $\text{Spur}(\mathbf{A}) = \text{Spur}(\mathbf{M}^{-1} \mathbf{B} \mathbf{M}) = \text{Spur}(\mathbf{B} \mathbf{M} \mathbf{M}^{-1}) = \text{Spur}(\mathbf{B} \mathbf{I}) = \text{Spur}(\mathbf{B})$.

6. Determinanten

- Ziel
- Lösen eines LGS ($n \times n$)
 - Inversion einer Matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$
 - Überprüfen einer symmetrischen Matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ auf Definitheit

Aufgrund ihrer numerischen Instabilität finden Determinanten in der Praxis für große Matrizen keine Verwendung.

Determinanten ordnen *quadratischen* Matrizen $\mathbf{A} \in \mathbb{R}^{n \times n}$ Zahlen zu. Schreibweise: $\det(\mathbf{A})$ bzw. $|\mathbf{A}|$
 $|\mathbf{A}| = a_{11}$ für $n = 1$

$$|\mathbf{A}| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11} a_{22} - a_{12} a_{21} \quad \text{für } n = 2$$

$$|\mathbf{A}| = \sum_{j=1}^n (-1)^{i+j} a_{ij} |\mathbf{A}_{ij}| \quad \text{für ein } i \in \{1, \dots, n\} \quad (\text{Entwicklung nach der } i\text{-ten Zeile})$$

$$= \sum_{i=1}^n (-1)^{i+j} a_{ij} |\mathbf{A}_{ij}| \quad \text{für ein } j \in \{1, \dots, n\} \quad (\text{Entwicklung nach der } j\text{-ten Spalte})$$

für beliebiges n (*Entwicklungssatz* von Laplace), wobei die Untermatrizen \mathbf{A}_{ij} der Ordnung $(n-1) \times (n-1)$ aus \mathbf{A} entstehen, indem man die i -te Zeile und die j -te Spalte streicht.

Graphisch $n = 2$: Für $\mathbf{a} = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$, $\mathbf{A} = \begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix}$ ist $|\det(\mathbf{A})| = |x_1 y_2 - x_2 y_1| = F$,

d.h. der Betrag der Determinante gibt den Flächeninhalt F des Parallelogramms an, das durch die beiden Vektoren \mathbf{a} und \mathbf{b} aufgespannt wird, denn:

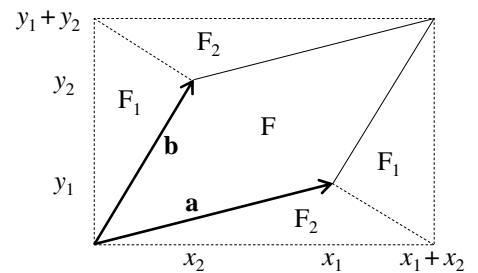
$$F = \text{Fläche des großen Rechtecks} - \text{Fläche der 4 Dreiecke}$$

$$= (x_1 + x_2) \cdot (y_1 + y_2) - 2 \cdot F_1 - 2 \cdot F_2$$

$$= (x_1 + x_2) \cdot (y_1 + y_2) - 2 \cdot \frac{1}{2} \cdot (y_1 + y_2) \cdot x_2 - 2 \cdot \frac{1}{2} \cdot (x_1 + x_2) \cdot y_1$$

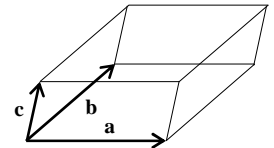
$$= x_1 y_1 + x_1 y_2 + x_2 y_1 + y_1 x_2 - y_2 x_2 - x_2 y_1 - x_1 y_1 - x_2 y_1$$

$$= x_1 y_2 - x_2 y_1$$



Sind die beiden Vektoren linear abhängig, so ist $F = 0$.

$n = 3$: $|\det(\mathbf{A})| = \text{Volumen des Parallelotops, aufgespannt durch die 3 Vektoren } \mathbf{a}, \mathbf{b}, \mathbf{c}$.



Beispiele

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

$$(2) \quad \begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} = \begin{array}{l} \text{(a) z.B. entwickelt nach der 1. Zeile} \\ \text{(b) z.B. entwickelt nach der 3. Spalte} \end{array}$$

$$(a) \quad (-1)^{1+1} \cdot 1 \cdot |\mathbf{A}_{11}| + (-1)^{1+2} \cdot 2 \cdot |\mathbf{A}_{12}| + (-1)^{1+3} \cdot 3 \cdot |\mathbf{A}_{13}| = 1 \cdot |\mathbf{A}_{11}| - 2 \cdot |\mathbf{A}_{12}| + 3 \cdot |\mathbf{A}_{13}| =$$

$$1 \cdot \begin{vmatrix} 5 & 6 \\ 8 & 9 \end{vmatrix} - 2 \cdot \begin{vmatrix} 4 & 6 \\ 7 & 9 \end{vmatrix} + 3 \cdot \begin{vmatrix} 4 & 5 \\ 7 & 8 \end{vmatrix} = 1 \cdot (45 - 48) - 2 \cdot (36 - 42) + 3 \cdot (32 - 35) = 0$$

$$(b) \quad (-1)^{1+3} \cdot 3 \cdot |\mathbf{A}_{13}| + (-1)^{2+3} \cdot 6 \cdot |\mathbf{A}_{23}| + (-1)^{3+3} \cdot 9 \cdot |\mathbf{A}_{33}| = 3 \cdot |\mathbf{A}_{13}| - 6 \cdot |\mathbf{A}_{23}| + 9 \cdot |\mathbf{A}_{33}| =$$

$$3 \cdot \begin{vmatrix} 4 & 5 \\ 7 & 8 \end{vmatrix} - 6 \cdot \begin{vmatrix} 1 & 2 \\ 7 & 8 \end{vmatrix} + 9 \cdot \begin{vmatrix} 1 & 2 \\ 4 & 5 \end{vmatrix} = 3 \cdot (32 - 35) - 6 \cdot (8 - 14) + 9 \cdot (5 - 8) = 0$$

$$(3) \quad \begin{vmatrix} 1 & 2 & 3 & 4 \\ 2 & 0 & 2 & 3 \\ 3 & 0 & 1 & 2 \\ 4 & 0 & 0 & 1 \end{vmatrix} = \begin{array}{l} \text{entwickelt} \\ \text{nach der} \\ \text{2. Spalte} \end{array} -2 \cdot \begin{vmatrix} 2 & 2 & 3 \\ 3 & 1 & 2 \\ 4 & 0 & 1 \end{vmatrix} = \begin{array}{l} \text{entwickelt} \\ \text{nach der} \\ \text{3. Zeile} \end{array} -2 \cdot \left(4 \cdot \begin{vmatrix} 2 & 3 \\ 1 & 2 \end{vmatrix} \right) + 1 \cdot \left(\begin{vmatrix} 2 & 2 \\ 3 & 1 \end{vmatrix} \right)$$

$$= -2 \cdot (4 \cdot (4 - 3) + 1 \cdot (2 - 6)) = 0$$

$$(4) \quad \begin{vmatrix} 4 & 0 & 0 & 0 \\ 3 & 3 & 0 & 0 \\ 2 & 2 & 2 & 0 \\ 1 & 1 & 1 & 1 \end{vmatrix} = \begin{array}{l} \text{entwickelt} \\ \text{nach der} \\ \text{1. Zeile} \end{array} 4 \cdot \begin{vmatrix} 3 & 0 & 0 \\ 2 & 2 & 0 \\ 1 & 1 & 1 \end{vmatrix} = \begin{array}{l} \text{entwickelt} \\ \text{nach der} \\ \text{1. Zeile} \end{array} 4 \cdot 3 \cdot \begin{vmatrix} 2 & 0 \\ 1 & 1 \end{vmatrix} = 4 \cdot 3 \cdot 2 \cdot 1 = 24$$

(Produkt der Hauptdiagonalelemente)

Rechenregeln für $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times n}$

- (1) \mathbf{A} regulär $\Leftrightarrow |\mathbf{A}| \neq 0$; \mathbf{A} singular $\Leftrightarrow |\mathbf{A}| = 0$ (vgl. oben, Beispiel (2) und (3))
 (2) $|\mathbf{A}'| = |\mathbf{A}|$

- (3) **A** Dreiecksmatrix $\Rightarrow |\mathbf{A}| = a_{11} \cdot a_{22} \cdot \dots \cdot a_{nn} = \prod_{i=1}^n a_{ii}$ (vgl. oben, Beispiel (4))
- (4) $|\mathbf{A}|$ ändert das Vorzeichen, wenn man zwei Zeilen (Spalten) vertauscht;
- (5) $|\mathbf{A}|$ ändert sich nicht, wenn man zu einer Zeile (Spalte) das Vielfache einer anderen Zeile (Spalte) addiert;
- (6) Multipliziert man eine Zeile (Spalte) mit einem Faktor, so ändert sich $|\mathbf{A}|$ um diesen Faktor;
speziell: $|\lambda \mathbf{A}| = \lambda^n \cdot |\mathbf{A}|$, $\lambda \in \mathbb{R}$
- (7) $|\mathbf{A}\mathbf{B}| = |\mathbf{A}| \cdot |\mathbf{B}| = |\mathbf{B}\mathbf{A}|$; aber: $|\mathbf{A} + \mathbf{B}| \neq |\mathbf{A}| + |\mathbf{B}|$ i.a.
- (8) **A** regulär $\Rightarrow |\mathbf{A}^{-1}| = \frac{1}{|\mathbf{A}|}$, denn $|\mathbf{A}| \cdot |\mathbf{A}^{-1}| = |\mathbf{A}\mathbf{A}^{-1}| = |\mathbf{I}| = 1$
- (9) Sind **A**, **B** zwei Matrizen, die sich höchstens in der k -ten Zeile unterscheiden und sonst gleich sind, also $\mathbf{a}_{(i)} = \mathbf{b}_{(i)} \quad \forall i \neq k$ und evtl. $\mathbf{a}_{(k)} \neq \mathbf{b}_{(k)}$, dann gilt: $|\mathbf{A}| + |\mathbf{B}| = |\mathbf{C}|$. Dabei ist die Matrix **C** wiederum bis auf die k -te Zeile mit **A** bzw. **B** identisch und die k -te Zeile von **C** ergibt sich als Summe der beiden Zeilen aus **A** und **B**, also $\mathbf{c}_{(i)} = \mathbf{a}_{(i)} = \mathbf{b}_{(i)} \quad \forall i \neq k$, $\mathbf{c}_{(k)} = \mathbf{a}_{(k)} + \mathbf{b}_{(k)}$.
- Die Aussage gilt analog auch für Spalten.
- (10) Für eine orthonormale Matrix **A** ist $|\mathbf{A}| = \pm 1$, denn aus Regel (7) und (2) folgt $|\mathbf{A}'\mathbf{A}| = |\mathbf{A}|^2$ und damit wegen $\mathbf{A}'\mathbf{A} = \mathbf{I}$ und $|\mathbf{I}| = 1$ das Resultat $|\mathbf{A}|^2 = 1$.
- (11) Für ähnliche Matrizen **A**, **B** gilt: $|\mathbf{A}| = |\mathbf{B}|$,
denn $|\mathbf{A}| = |\mathbf{M}^{-1} \mathbf{B} \mathbf{M}| = |\mathbf{M}^{-1}| \cdot |\mathbf{B}| \cdot |\mathbf{M}| = \frac{1}{|\mathbf{M}|} \cdot |\mathbf{B}| \cdot |\mathbf{M}| = |\mathbf{B}|$.

Beispiel zu (5):
$$\begin{vmatrix} 1 & 1 & 2 & 3 \\ 2 & 3 & 1 & 1 \\ 3 & 2 & 4 & 4 \\ 4 & 4 & 3 & 2 \end{vmatrix} \begin{array}{l} \text{Rechnung:} \\ 2. \text{ Zeile} - 2 \cdot 1. \text{ Zeile} \\ 3. \text{ Zeile} - 3 \cdot 1. \text{ Zeile} \\ 4. \text{ Zeile} - 4 \cdot 1. \text{ Zeile} \end{array} = \begin{vmatrix} 1 & 1 & 2 & 3 \\ 0 & 1 & -3 & -5 \\ 0 & -1 & -2 & -5 \\ 0 & 0 & -5 & -10 \end{vmatrix}$$

= entwickelt nach der 1. Spalte
$$1 \cdot \begin{vmatrix} 1 & -3 & -5 \\ -1 & -2 & -5 \\ 0 & -5 & -10 \end{vmatrix} \begin{array}{l} \text{Rechnung:} \\ 2. \text{ Zeile} + 1. \text{ Zeile} \end{array} = \begin{vmatrix} 1 & -3 & -5 \\ 0 & -5 & -10 \\ 0 & -5 & -10 \end{vmatrix} = \begin{vmatrix} -5 & -10 \\ -5 & -10 \end{vmatrix} = 0$$

Beispiel zu (9):
$$\begin{vmatrix} 2 & 3 \\ 1 & 2 \end{vmatrix} + \begin{vmatrix} 2 & 3 \\ 3 & 4 \end{vmatrix} = \begin{vmatrix} 2 & 3 \\ 4 & 6 \end{vmatrix}$$

Regel von Sarrus

(nur für 3x3 Matrizen)

$$|\mathbf{A}| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{array}{l} + + + - - - \\ a_{11} a_{22} a_{33} + a_{12} a_{23} a_{31} + a_{13} a_{21} a_{32} \\ - a_{13} a_{22} a_{31} - a_{11} a_{23} a_{32} - a_{12} a_{21} a_{33} \end{array}$$

Beispiel:
$$\begin{vmatrix} 3 & 1 & 2 \\ 3 & 3 & 5 \\ 6 & 8 & 14 \end{vmatrix} = \begin{array}{l} + + \pm - - \\ 3 \cdot 1 \cdot 14 + 1 \cdot 5 \cdot 6 + 2 \cdot 3 \cdot 8 - 2 \cdot 3 \cdot 6 - 3 \cdot 5 \cdot 8 - 1 \cdot 3 \cdot 14 = 6 \end{array}$$

Cramersche Regel

zum Lösen eines quadratischen LGS $\mathbf{A}\mathbf{x} = \mathbf{b}$

Für eine reguläre Matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ und $\mathbf{b} \in \mathbb{R}^n$ ergibt sich die Lösung des LGS $\mathbf{A}\mathbf{x} = \mathbf{b}$ aus

$x_j = \frac{|\mathbf{A}_j|}{|\mathbf{A}|}$, wobei \mathbf{A}_j aus **A** entsteht, indem man die j -te Spalte von **A** durch **b** ersetzt.

Beispiel: $\mathbf{A} = \begin{pmatrix} 3 & 1 & 2 \\ 3 & 3 & 5 \\ 6 & 8 & 14 \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$. Nach dem letzten Beispiel ist $|\mathbf{A}| = 6 \neq 0$, also **A** regulär.

Mit $|\mathbf{A}_1| = \begin{vmatrix} 2 & 1 & 2 \\ 1 & 3 & 5 \\ 0 & 8 & 14 \end{vmatrix} = 2 \cdot \begin{vmatrix} 3 & 5 \\ 8 & 14 \end{vmatrix} - 1 \cdot \begin{vmatrix} 1 & 2 \\ 8 & 14 \end{vmatrix} = 4 - (-2) = 6$

$|\mathbf{A}_2| = \begin{vmatrix} 3 & 2 & 2 \\ 3 & 1 & 5 \\ 6 & 0 & 14 \end{vmatrix} = -2 \cdot \begin{vmatrix} 3 & 5 \\ 6 & 14 \end{vmatrix} + 1 \cdot \begin{vmatrix} 3 & 2 \\ 6 & 14 \end{vmatrix} = -24 + 30 = 6$

$$|\mathbf{A}_3| = \left| \begin{pmatrix} 3 & 1 & 2 \\ 3 & 3 & 1 \\ 6 & 8 & 0 \end{pmatrix} \right| = 2 \cdot \left| \begin{pmatrix} 3 & 3 \\ 6 & 8 \end{pmatrix} \right| - 1 \cdot \left| \begin{pmatrix} 3 & 1 \\ 6 & 8 \end{pmatrix} \right| = 12 - 18 = -6$$

ergibt sich $x_1 = \frac{|\mathbf{A}_1|}{|\mathbf{A}|} = \frac{6}{6} = 1$, $x_2 = \frac{|\mathbf{A}_2|}{|\mathbf{A}|} = \frac{6}{6} = 1$, $x_3 = \frac{|\mathbf{A}_3|}{|\mathbf{A}|} = \frac{-6}{6} = -1$.

Die Cramersche Regel kann auch zur Inversion einer Matrix verwendet werden:

Inversion

einer regulären $n \times n$ Matrix \mathbf{A}

Für eine reguläre Matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ ist $\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|} \mathbf{A}'_{\text{adj}}$, wobei sich die Elemente c_{ij} der adjungierten Matrix \mathbf{A}'_{adj} errechnen durch $c_{ij} = (-1)^{i+j} |\mathbf{A}_{ij}|$.

Speziell für $n = 2$: Die Inverse von $\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ ist $\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$.

Beispiele

(1) $\mathbf{A} = \begin{pmatrix} 2 & 3 \\ 1 & 4 \end{pmatrix}$, $|\mathbf{A}| = 8 - 3 = 5 \neq 0$, also \mathbf{A} regulär mit $\mathbf{A}^{-1} = \frac{1}{5} \begin{pmatrix} 4 & -3 \\ -1 & 2 \end{pmatrix}$

(2) $\mathbf{A} = \begin{pmatrix} 3 & 1 & 2 \\ 3 & 3 & 5 \\ 6 & 8 & 14 \end{pmatrix}$, $|\mathbf{A}| = 6 \neq 0$ (s.o.), also \mathbf{A} regulär. Die Inversion erfolgt in 3 Schritten:

(a) Berechne die Determinanten aller Untermatrizen \mathbf{A}_{ij} und fasse sie wieder zu einer Matrix

$$\text{zusammen: } \begin{pmatrix} |\mathbf{A}_{11}| & |\mathbf{A}_{12}| & |\mathbf{A}_{13}| \\ |\mathbf{A}_{21}| & |\mathbf{A}_{22}| & |\mathbf{A}_{23}| \\ |\mathbf{A}_{31}| & |\mathbf{A}_{32}| & |\mathbf{A}_{33}| \end{pmatrix} = \begin{pmatrix} 2 & 12 & 6 \\ -2 & 30 & 18 \\ -1 & 9 & 6 \end{pmatrix}$$

(b) berücksichtige die Vorzeichen $(-1)^{i+j}$ und bestimme so die adjungierte Matrix

$$\mathbf{A}'_{\text{adj}} = \begin{pmatrix} 2 & -12 & 6 \\ -(-2) & 30 & -18 \\ -1 & -9 & 6 \end{pmatrix}$$

(c) transponiere \mathbf{A}'_{adj} und dividiere durch $|\mathbf{A}|$, also $\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|} \mathbf{A}'_{\text{adj}} = \frac{1}{6} \begin{pmatrix} 2 & 2 & -1 \\ -12 & 30 & -9 \\ 6 & -18 & 6 \end{pmatrix}$

Definitheit

Regeln für symmetrische Matrizen $\mathbf{A} \in \mathbb{R}^{n \times n}$

Bezeichne mit \mathbf{H}_i die $i \times i$ Hauptuntermatrizen von \mathbf{A} ,

also $\mathbf{H}_1 = a_{11}$, $\mathbf{H}_2 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, $\mathbf{H}_3 = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$, ..., $\mathbf{H}_n = \mathbf{A}$.

- (1) \mathbf{A} ist positiv definit genau dann, wenn $|\mathbf{H}_i| > 0$ für alle i ;
- (2) \mathbf{A} ist negativ definit genau dann, wenn $|\mathbf{H}_i|$ wechselndes Vorzeichen aufweist, und zwar $a_{11} < 0$, $|\mathbf{H}_2| > 0$, $|\mathbf{H}_3| < 0$, $|\mathbf{H}_4| > 0$ etc.;
- (3) \mathbf{A} ist indefinit, wenn in (1) bzw. (2) mindestens eine der Determinanten das "falsche" Vorzeichen besitzt, also z.B. $|\mathbf{H}_2| < 0$ ist.

Auf die etwas komplizierteren Kriterien zur Semidefinitheit soll hier nicht eingegangen werden.

Beispiele

(1) $\begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}$ ist positiv definit, da $1 > 0$ und $\left| \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix} \right| = 1 > 0$

(2) $\begin{pmatrix} -1 & 2 \\ 2 & -5 \end{pmatrix}$ ist negativ definit, da $-1 < 0$ und $\left| \begin{pmatrix} -1 & 2 \\ 2 & -5 \end{pmatrix} \right| = 1 > 0$

(3) $\begin{pmatrix} -1 & 2 \\ 2 & 5 \end{pmatrix}$ ist indefinit, da $\left| \begin{pmatrix} -1 & 2 \\ 2 & 5 \end{pmatrix} \right| = -9 < 0$

(4) $\begin{pmatrix} 2 & 1 & 0 \\ 1 & 4 & 1 \\ 0 & 1 & 2 \end{pmatrix}$ ist positiv definit, da $2 > 0$, $\left| \begin{pmatrix} 2 & 1 \\ 1 & 4 \end{pmatrix} \right| = 7 > 0$, $\left| \begin{pmatrix} 2 & 1 & 0 \\ 1 & 4 & 1 \\ 0 & 1 & 2 \end{pmatrix} \right| = 12 > 0$

$$(5) \begin{pmatrix} -2 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & -2 \end{pmatrix} \text{ ist negativ definit, da } -2 < 0, \left| \begin{pmatrix} -2 & 1 \\ 1 & -4 \end{pmatrix} \right| = 7 > 0, \left| \begin{pmatrix} -2 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & -2 \end{pmatrix} \right| = -12 < 0$$

$$(6) \begin{pmatrix} 1 & 2 & 0 \\ 2 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ ist indefinit, da } \left| \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} \right| = -1 < 0$$

Ist die Matrix \mathbf{A} nicht symmetrisch, so untersucht man statt dessen die symmetrische Matrix $\mathbf{B} = \frac{1}{2}(\mathbf{A} + \mathbf{A}')$, da \mathbf{B} die gleiche quadratische Form wie \mathbf{A} besitzt: $Q(\mathbf{x}) = \mathbf{x}'\mathbf{A}\mathbf{x} = \mathbf{x}'\mathbf{B}\mathbf{x}$.

Bsp.: Statt $\mathbf{A} = \begin{pmatrix} 1 & 4 \\ 0 & 5 \end{pmatrix}$ untersucht man $\mathbf{B} = \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}$ auf Definitheit.

7. Lineare Optimierung

Ziel Bestimmen von Minima bzw. Maxima einer linearen Zielfunktion $z: \mathbb{R}^n \rightarrow \mathbb{R}$ unter linearen Nebenbedingungen in Form von Ungleichungen

Beispiel eines linearen Optimierungsproblems (LOP)
 Ein Unternehmen produziert die beiden Produkte P_1, P_2 auf den Maschinen M_1, M_2, M_3 . Die dazu benötigten Produktionszeiten (in ZE / ME)³, die Kapazitäten der Maschinen (in ZE) und die Deckungsbeiträge DB (in GE / ME) sind in nebenstehender Tabelle zusammengefasst:

	P_1	P_2	Kapazität
M_1	1	–	10
M_2	–	1	6
M_3	1	2	16
DB	3	2	

Frage: Bei welchen Produktionsmengen x_1, x_2 für P_1, P_2 wird der DB maximal?

Lösung (1) Umsetzung der Daten in ein mathematisches Modell

$x_1 \leq 10$	Die Modellierung der Kapazitätsbeschränkungen führt zu einem linearen
$x_2 \leq 6$	Ungleichungssystem, auch als lineares Restriktionssystem (LRS) bezeichnet.
$x_1 + 2x_2 \leq 16$	Es enthält zusätzlich die so genannten Nichtnegativitätsbedingungen
$x_1 \geq 0$	(NNB). Sie stellen sicher, dass keine negativen Produktionsmengen zugelassen werden.
$x_2 \geq 0$	

Die zu maximierende, lineare Zielfunktion ergibt sich aus den beiden Deckungsbeiträgen.
 Zielfunktion: $z = 3x_1 + 2x_2 = \max!$

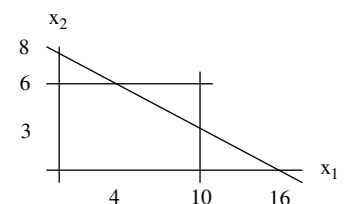
(2) Graphische Lösung (nur bei 2 Variablen sinnvoll)

(a) Zulässigen Bereich (ZB) ermitteln

Zunächst wird die Lösungsmenge des linearen Ungleichungssystems, der so genannte zulässige Bereich, graphisch bestimmt. Dazu zeichnet man die einzelnen Restriktionsgeraden, z.B. die Gerade $x_1 + 2x_2 = 16$, in ein Koordinatensystem ein.

Die Lösungsmenge der entsprechenden Ungleichung umfasst eine Hälfte der Koordinatenebene, in Fall der Ungleichung $x_1 + 2x_2 \leq 16$ die Hälfte links von bzw. unterhalb der Geraden, wie man durch Einsetzen des Punktes $(0, 0)$ überprüft.

Die Lösungsmenge des Gesamtsystems erhält man als Durchschnitt der einzelnen Lösungsmengen. Sie bildet als Durchschnitt von konvexen⁴ Mengen wiederum eine konvexe Menge, hier ein geschlossenes Fünfeck. Der ZB enthält also alle Mengenkombinationen (x_1, x_2) , die unter den gegebenen Kapazitätsbeschränkungen produziert werden könnten.



(b) Gerade $z = 0$ ($3x_1 + 2x_2 = 0$) einzeichnen

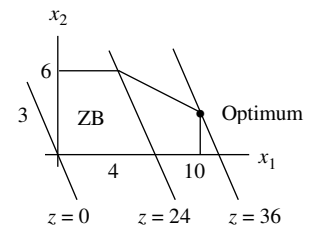
Für jede beliebige Zahl a verläuft die Gerade $z = a$ ($3x_1 + 2x_2 = a$) parallel zur Geraden

³ ZE = Zeiteinheit; ME = Mengeneinheit; GE = Geldeinheit

⁴ Eine Menge ist konvex, wenn mit je zwei Punkten der Menge auch jeweils die gesamte Verbindungsgerade zwischen den Punkten zur Menge gehört.

$z = 0$. Auf der Geraden $z = a$ liegen alle Mengenkombinationen, die einen Zielbeitrag (hier: Deckungsbeitrag) von a liefern.

- (c) Gerade $z = 0$ soweit parallel verschieben (in Richtung "Verbesserung der Zielfunktion"), bis sie den ZB nur noch in einem Punkt (evtl. einer Strecke) berührt. Dieser Punkt stellt das Optimum dar.



(3) Rechnerische Lösung mit Hilfe des *Simplex-Algorithmus* (SA)

(a) Grundlegende Aussagen zum SA; Prinzip des SA

- Das Optimum eines LOP liegt stets in einer Ecke des ZB.
Diese Aussage basiert letztendlich auf der Konvexität des ZB.
- Vorgehensweise des SA: Es werden systematisch die Ecken des ZB abgesucht. Dabei startet man üblicherweise im 0-Punkt.⁵
- Charakterisierung der Ecken des ZB
 - Jedes LRS lässt sich durch Einfügen von zusätzlichen Variablen in ein LGS überführen. Es kann so bestimmt werden, dass seine rechte Seite stets größer als Null ist. Das LGS muss dabei den NNB genügen.
 - Jede Basislösung dieses LGS, die die NNB erfüllt, entspricht einer Ecke des ZB und umgekehrt.
- Der Übergang von einer Ecke des ZB zu einer benachbarten Ecke geschieht durch einen Basistausch im LGS. Wie im Zusammenhang mit dem Gauß-Algorithmus besprochen führt man dazu einen Pivotschritt durch.

(b) Umformen eines LRS in ein LGS

Für jede Ungleichung wird eine zusätzliche Variable eingeführt.

Im obigen Beispiel ergibt sich folgende Umformung:

<u>LRS</u>	<u>LGS</u>
$x_1 \leq 10$	$x_1 + u_1 = 10$
$x_2 \leq 6$	$x_2 + u_2 = 6$
$x_1 + 2x_2 \leq 16$	$x_1 + 2x_2 + u_3 = 16$

Das LGS muss dabei die Nichtnegativitätsbedingungen $x_1, x_2, u_1, u_2, u_3 \geq 0$ erfüllen. Die Variablen u_i messen hier die freie Kapazität der Maschinen, die als Schlupf bezeichnet wird. Man nennt die u_i daher auch *Schlupfvariablen*.

Die Umformung ist immer so möglich, dass die Bedingungen $u_i \geq 0$ gelten müssen und die rechte Seite ≥ 0 ist, wie nachstehende Beispiele zeigen:

<u>LRS</u>	<u>LGS</u>	<u>NNB</u>
• $x \leq 1$	$x + u = 1,$	$u \geq 0$
• $x \geq 1$	$x - u = 1,$	$u \geq 0$
• $x \leq -1$	$-x \geq 1$	$-x - u = 1, \quad u \geq 0$
• $x \geq -1$	$-x \leq 1$	$-x + u = 1, \quad u \geq 0$

(c) Anfangstableau

Die Umsetzung eines LGS in ein Tableau ist vom Gauß-Algorithmus her bekannt.

Um auch die Zielfunktion im Tableau berücksichtigen zu können, formt man sie um:

$$z = 3x_1 + 2x_2 \quad \Leftrightarrow \quad z - 3x_1 - 2x_2 = 0$$

Insgesamt ergibt sich so ein LGS mit 4 Gleichungen und 6 Variablen. Es besitzt ∞ viele Lösungen und liegt in kanonischer Form vor. In Tableauschreibweise:

⁵ Falls dieser nicht zulässig ist, wird in einem ersten Schritt durch Einführen von zusätzlichen Hilfsvariablen mit Hilfe des SA eine zulässige Ecke ermittelt (2-Phasen Simplex-Algorithmus).

BV	x_1	x_2	u_1	u_2	u_3	z	r.S.
u_1	1	0	1	0	0	0	10
u_2	0	1	0	1	0	0	6
u_3	1	2	0	0	1	0	16
z	-3	-2	0	0	0	1	0

Basisvariable (BV) sind hier u_1, u_2, u_3, z , da "unter" ihnen die Einheitsvektoren stehen; sie können auch an der zusätzlichen, ersten Spalte abgelesen werden. Entsprechend sind x_1, x_2 die Nichtbasisvariablen (NBV).

Die Basislösung erhält man durch 0-Setzen der NBV: $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, $\begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \begin{pmatrix} 10 \\ 6 \\ 16 \end{pmatrix}$, $z = 0$

Sie entspricht dem 0-Punkt, dem Ausgangspunkt bei der Suche nach der optimalen Ecke im zulässigen Bereich.

(d) Basistausch: Pivotschritt

Fragen

Grafisch: Zu welcher benachbarten Ecke des zulässigen Bereiches soll gewechselt werden?

Inhaltlich:

- (1) Welches Produkt soll (zusätzlich) produziert werden?
- (2) Wie viel kann von diesem Produkt hergestellt werden, welche Maschine stellt also den Engpass dar?

Im Verlauf des Simplex-Algorithmus ist es keine Seltenheit, dass die gewählte Produktkonstellation sich als noch nicht optimal herausstellt. Die Fragen können dann auch lauten:

- (1) An welcher Maschine soll zunächst freie Kapazität geschaffen werden (um dann im nächsten Schritt ein anderes Produkt zu wählen)?
- (2) Auf welches der bisher gewählten Produkte soll dabei verzichtet werden?

Formal:

- (1) Welche bisherige NBV soll in die Basis aufgenommen werden, d.h. in welcher Spalte soll ein Einheitsvektor erzeugt werden?
- (2) Welcher Einheitsvektor soll dort erzeugt werden, d.h. welche bisherige BV entfällt dafür aus der Basis?

Antworten

- (1) Nehme eine Variable in die Basis auf, für die in der z -Zeile ein negativer Wert steht, denn dies liefert einen positiven Beitrag zur Zielfunktion. Im obigen Beispiel kann x_1 oder x_2 aufgenommen werden. Welche der Variablen zu einem schnelleren Erreichen des Optimums führt, kann i.a. nicht vorhergesagt werden. Üblicherweise wählt man den betragsmäßig größten Wert, hier x_1 , d.h. es soll zunächst das Produkt P_1 hergestellt werden. Im nachstehenden Tableau wird dies durch einen Pfeil verdeutlicht, der auf die x_1 -Spalte zeigt.

Im Folgenden wird die Spalte der aufzunehmenden Variablen als a -Spalte bezeichnet.

- (2) Welche Variable entfällt nun aus der Basis, d.h. welche der Kapazitätsbeschränkungen für P_1 ist am restriktivsten.

Dazu betrachtet man die Werte in der a -Spalte (hier x_1 -Spalte), die größer als 0 sind und bildet für sie zeilenweise die Quotienten: $\frac{\text{Wert der r.S.-Spalte}}{\text{Wert der } a\text{-Spalte}}$.

Die Ergebnisse werden in die entsprechende Zeile der Hilfsspalte θ (Theta)

eingetragen. Die z -Zeile bleibt bei der Bildung des Quotienten unberücksichtigt.

Bestimme nun das Minimum dieser Quotienten. Die Zeile, in der das Minimum steht (hier die u_1 -Zeile), heißt im Folgenden m -Zeile. Die Variable m (hier u_1) entfällt aus der Basis. Der Pfeil am rechten Rand des nachstehenden Tableaus verdeutlicht dies.

Gibt es mehrere gleich kleine Werte, so kann einer dieser Werte gewählt werden. Sind alle Werte in der a -Spalte ≤ 0 , so können keine Quotienten gebildet werden und das Optimierungsproblem besitzt keine Lösung.

Begründung für die obige Vorgehensweise (am Beispiel):

$$\text{Das LRS lautet in Vektorschreibweise: } x_1 \cdot \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + x_2 \cdot \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} \leq \begin{pmatrix} 10 \\ 6 \\ 16 \end{pmatrix}.$$

Da x_2 weiterhin NBV bleiben soll und somit gleich 0 ist, ergeben sich für x_1 die

$$\text{Bedingungen } x_1 \cdot \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \leq \begin{pmatrix} 10 \\ 6 \\ 16 \end{pmatrix}, \text{ also } x_1 \leq 10 \text{ und } x_1 \leq 16, \text{ insgesamt}$$

$x_1 \leq \min\{10/1, -, 16/1\}$. Genau diese Bedingung wird in der Hilfsspalte θ ermittelt. u_1 entfällt aus der Basis, da die Restriktion der Maschine 1 bei der Produktion von P_1 den Engpass darstellt.

Pivotschritt durchführen

Nachdem die beiden Fragen geklärt sind, kann der Basistausch (Pivotschritt) durchgeführt werden. Das Pivotelement liegt im Schnittpunkt von a -Spalte und m -Zeile. In diesem Fall bedeutet der Basistausch von u_1 mit x_1 , dass in der x_1 -Spalte derjenige Einheitsvektor erzeugt werden muss, der bislang in der u_1 -Spalte steht.

BV	x_1	x_2	u_1	u_2	u_3	z	r.S.	θ	
u_1	(1)	0	1	0	0	0	10	10/1	} min. $\left. \begin{array}{l} \leftarrow \text{PZ} \\ - \text{PZ} \\ + 3 \cdot \text{PZ} \end{array} \right\}$
u_2	0	1	0	1	0	0	6	-	
u_3	1	2	0	0	1	0	16	16/1	
z	-3	-2	0	0	0	1	0		

↑

Nach der Durchführung des Pivotschrittes erhält man folgendes Tableau:

BV	x_1	x_2	u_1	u_2	u_3	z	r.S.
x_1	1	0	1	0	0	0	10
u_2	0	1	0	1	0	0	6
u_3	0	2	-1	0	1	0	6
z	0	-2	3	0	0	1	30

Basisvariable: x_1, u_2, u_3, z ; Nichtbasisvariable: x_2, u_1

$$\text{Basislösung: } \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 10 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 6 \\ 6 \end{pmatrix}, \quad z = 30$$

Erneuter Basistausch, bis alle Werte der z -Zeile ≥ 0

Solange in der z -Zeile noch negative Werte stehen, kann das Zielkriterium durch Aufnahme der entsprechenden Variablen in die Basis verbessert werden. Sind alle Werte der z -Zeile ≥ 0 , so kann keine weitere Verbesserung erzielt werden und das Optimum ist erreicht.

Im Beispiel: Nehme x_2 in die Basis auf, eliminiere dafür u_3 aus der Basis.

BV	x_1	x_2	u_1	u_2	u_3	z	r.S.	θ	
x_1	1	0	1	0	0	0	10	-	} min. $\left. \begin{array}{l} - \text{PZ} \\ : 2 \leftarrow \text{PZ} \\ + 2 \cdot \text{PZ} \end{array} \right\}$
u_2	0	1	0	1	0	0	6	6/1	
u_3	0	(2)	-1	0	1	0	6	6/2	
z	0	-2	3	0	0	1	30		

↑

Nach der Durchführung des Pivotschrittes erhält man das Endtableau:

BV	x_1	x_2	u_1	u_2	u_3	z	r.S.
x_1	1	0	1	0	0	0	10
u_2	0	0	0,5	1	-0,5	0	3
x_2	0	1	-0,5	0	0,5	0	3
z	0	0	2	0	1	1	36

(e) Interpretation des Endtableaus

Das Optimum ist erreicht, da alle Werte der z -Zeile ≥ 0 .

Basisvariable: x_1, u_2, x_2, z ; Nichtbasisvariable: u_1, u_3 , d.h. $u_1 = 0, u_3 = 0$,

also Basislösung: $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 10 \\ 3 \end{pmatrix}$, $\begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \\ 0 \end{pmatrix}$, $z = 36$.

Interpretation:

- Optimale Produktionsmengen: $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 10 \\ 3 \end{pmatrix}$ ME
- Kapazitätsauslastung: $u_2 = 3$, d.h. an Maschine M_2 sind noch 3 ZE freie Kapazität; Maschine M_1 und M_3 sind ausgelastet.
- Maximal erreichbarer Deckungsbeitrag: $DB_{\text{opt}} = 36$ GE
- Weitere Interpretation des Endtableaus (Spalten der NBV; hier u_1, u_3)

In Vektorschreibweise lautet das Endtableau:

$$x_1 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + x_2 \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + u_1 \cdot \begin{pmatrix} 1 \\ 0,5 \\ -0,5 \end{pmatrix} + u_2 \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + u_3 \cdot \begin{pmatrix} 0 \\ -0,5 \\ 0,5 \end{pmatrix} = \begin{pmatrix} 10 \\ 3 \\ 3 \end{pmatrix}$$

$$\text{oder äquivalent} \quad \begin{pmatrix} x_1 \\ u_2 \\ x_2 \end{pmatrix} = \begin{pmatrix} 10 \\ 3 \\ 3 \end{pmatrix} - u_1 \cdot \begin{pmatrix} 1 \\ 0,5 \\ -0,5 \end{pmatrix} - u_3 \cdot \begin{pmatrix} 0 \\ -0,5 \\ 0,5 \end{pmatrix}.$$

Setzt man z.B. $u_1 = -1$, d.h. man vergrößert die Ausgangskapazität der Maschine M_1 um eine ZE (hier von 10 auf 11) und lässt $u_3 = 0$ weiterhin, dann ergibt sich

$$\begin{pmatrix} x_1 \\ u_2 \\ x_2 \end{pmatrix} = \begin{pmatrix} 10 \\ 3 \\ 3 \end{pmatrix} + 1 \cdot \begin{pmatrix} 1 \\ 0,5 \\ -0,5 \end{pmatrix} = \begin{pmatrix} 11 \\ 3,5 \\ 2,5 \end{pmatrix}.$$

x_1 steigt also um 1 ME, x_2 sinkt um 0,5 ME und die freie Kapazität an M_2 steigt um 0,5 ZE. Analog können durch $u_1 = 1$ die Auswirkungen einer Kapazitätssenkung von M_1 um eine Zeiteinheit analysiert werden. Man interpretiert die Werte in den Spalten der NBV daher als Anpassungs- bzw. Verdrängungskoeffizienten.

Der maximale Deckungsbeitrag *steigt* im Fall $u_1 = -1$ um 2 GE, wie im Schnittpunkt von z -Zeile und u_1 -Spalte abgelesen werden kann. Die Werte der z -Zeile werden daher als Grenzgewinne gedeutet.

Analoge Aussagen können auch für u_3 abgeleitet werden.

Verringert man hingegen die Kapazität von M_2 um bis zu 3 ZE, so hat dies keine Auswirkungen auf die optimale Produktion, da M_2 noch nicht ausgelastet ist (s.o.).

Häufig kommt es bei größeren Optimierungsproblemen vor, dass im Optimum bestimmte Produkte nicht hergestellt werden also z.B. $x_n = 0$ ist. Setzt man in diesem Falle $x_n = 1$, erzwingt also die Produktion einer ME von x_n , so gibt die Spalte unter x_n (mit vertauschtem Vorzeichen) an, welche Auswirkungen dies hat. Der Wert in der z -Zeile misst dann die Opportunitätskosten der Produktion von x_n .

(f) Anmerkungen

- (1) Da die z -Spalte immer konstant bleibt, lässt man sie weg.
- (2) Man kann auch die Spalten der Basisvariablen weglassen, da sie stets die Einheitsvektoren enthalten und diese Information bereits in der Spalte "BV" enthalten ist.

- (3) Soll eine Zielfunktion minimiert werden, so kehrt man alle Vorzeichen in der Zielfunktion um und maximiert sie dann.
- (4) Bei nichtlinearer Zielfunktion und / oder Ungleichungen kann der Simplex-Algorithmus nicht mehr angewandt werden. Im Falle *konvexer* Zielfunktion und konvexen Ungleichungen kann der *Satz von Kuhn-Tucker* weiterhelfen. Liegen die Nebenbedingungen in *Gleichungsform* vor, so verwendet man den *Lagrange-Ansatz*, der in der Analysis vorgestellt wird. Zielfunktion und Nebenbedingungen dürfen dabei beliebige Form aufweisen.

8. Folgen und Reihen

Folge

Abbildung von \mathbb{N} nach \mathbb{R} . Schreibweise: $(a_n)_{n \in \mathbb{N}}$

Folgen bilden die Grundlage der Differenzial- und Integralrechnung. Begriffe wie Stetigkeit, Differenzierbarkeit, Ableitung und Integral werden mit Hilfe von Folgen definiert.

Eigenschaften von Folgen wie Monotonie und Beschränktheit definieren sich direkt aus den entsprechenden Eigenschaften von Funktionen.

Wesentlich bei der Betrachtung von Folgen ist ihr Verhalten für steigendes n ($n \rightarrow \infty$), also die Frage nach der Konvergenz einer Folge. Vor der Präzisierung des Begriffs einige Beispiele.

Beispiele

(1) $a_n = \frac{1}{n}$: $a_1 = 1$, $a_2 = \frac{1}{2}$, $a_3 = \frac{1}{3}$, ... Grafisch:

in Analogie zur Darstellung der Funktion $f(x) = \frac{1}{x}$.

a_n ist streng monoton fallend, beschränkt und konvergiert gegen 0.

(2) $a_n = 1 + \frac{1}{n}$: $a_1 = 1 + 1 = 2$, $a_2 = 1 + \frac{1}{2} = \frac{3}{2}$, $a_3 = 1 + \frac{1}{3} = \frac{4}{3}$, ...

a_n ist streng monoton fallend, beschränkt und konvergiert gegen 1.

(3) $a_n = (1 + \frac{1}{n})^n$: $a_1 = 2$, $a_2 = \frac{9}{4} = 2,25$, $a_3 = \frac{64}{27} = 2,37$, ..., $a_{100} = (\frac{101}{100})^{100} = 2,70$, ...

a_n ist streng monoton steigend, beschränkt und konvergiert gegen die Zahl $e = 2,718...$, wie später mit Hilfe der Differenzialrechnung (Regeln von l'Hospital) nachgewiesen wird.

(4) $a_n = (-1)^n \cdot \frac{1}{n}$: $a_1 = -1$, $a_2 = \frac{1}{2}$, $a_3 = -\frac{1}{3}$, ...

a_n ist eine beschränkte, *alternierende* Folge (wechselt stets das Vorzeichen, also nicht monoton) und konvergiert gegen 0.

(5) $a_n = (-1)^n \cdot (1 + \frac{1}{n})$: $a_1 = -2$, $a_2 = \frac{3}{2}$, $a_3 = -\frac{4}{3}$, ...

a_n ist eine beschränkte, alternierende Folge und nicht konvergent. Die beiden Werte 1 bzw. -1, denen sich die Folge für $n \rightarrow \infty$ nähert, sind *Häufungspunkte* der Folge.

(6) $a_n = n$: $a_1 = 1$, $a_2 = 2$, $a_3 = 3$, ...

a_n ist streng monoton steigend und nicht beschränkt, strebt also gegen ∞ für $n \rightarrow \infty$.

Man spricht in diesem Fall von *uneigentlicher Konvergenz*, da ∞ keine Zahl ist.

(7) *Arithmetische Folge*: $a_n = a_{n-1} + d$, $d \in \mathbb{R}$ (rekursive Darstellung);

d.h. $a_2 = a_1 + d$, $a_3 = a_2 + d = a_1 + 2d$, ..., $a_n = a_1 + (n-1) \cdot d$ (explizite Darstellung);

z.B. $a_1 = 1$, $d = 2 \Rightarrow a_2 = 3$, $a_3 = 5$, $a_4 = 7$, ... (ungerade Zahlen);

strebt gegen ∞ bzw. $-\infty$ für $n \rightarrow \infty$, falls $d \neq 0$.

(8) *Geometrische Folge*: $a_n = a_{n-1} \cdot q$, $q \in \mathbb{R}$ (rekursive Darstellung);

d.h. $a_2 = a_1 \cdot q$, $a_3 = a_2 \cdot q = a_1 \cdot q^2$, ..., $a_n = a_1 \cdot q^{n-1}$ (explizite Darstellung);

konvergiert gegen 0, falls $|q| < 1$.

Konvergenz einer Folge $(a_n)_{n \in \mathbb{N}}$ gegen einen Grenzwert a
 Zu jeder (beliebig kleinen) Zahl $\varepsilon > 0$ existiert ein Index n_ε , so dass $\forall n \geq n_\varepsilon$ gilt: $|a_n - a| < \varepsilon$.
 Anders formuliert: Zu jeder ε -Umgebung von a , $U_\varepsilon(a)$, existiert ein Zeitpunkt n_ε , ab dem alle weiteren Folgenglieder in dieser Umgebung liegen.
 Dabei ist $U_\varepsilon(a) = \{x \in \mathbb{R} \mid |x-a| < \varepsilon\} = (a-\varepsilon, a+\varepsilon)$.
 Schreibweise: $a_n \xrightarrow{n \rightarrow \infty} a$ oder $\lim_{n \rightarrow \infty} a_n = a$. Nicht konvergente Folgen heißen *divergent*.

Nullfolge Folge mit Grenzwert 0.
 Folgen, die gegen ∞ (bzw. $-\infty$) streben, deren Folgenglieder also ab einem Zeitpunkt jenseits einer beliebig groß (bzw. klein) gewählten Zahl liegen, nennt man *uneigentlich konvergent*.

Regel Jede monotone und beschränkte Folge ist konvergent.
 Begründung: Der Grenzwert ist die kleinste obere Schranke (bei monoton steigenden Folgen) bzw. die größte untere Schranke (bei monoton fallenden Folgen).

Regeln Für konvergente Folgen $(a_n)_{n \in \mathbb{N}}$ und $(b_n)_{n \in \mathbb{N}}$ mit $\lim_{n \rightarrow \infty} a_n = a$, $\lim_{n \rightarrow \infty} b_n = b$ gilt:
 (1) $\lim_{n \rightarrow \infty} (a_n \pm b_n) = a \pm b$ (2) $\lim_{n \rightarrow \infty} (a_n \cdot b_n) = a \cdot b$ (3) $\lim_{n \rightarrow \infty} (a_n : b_n) = a : b$, $b_n \neq 0$

Beispiel: $a_n = \frac{3n^2 + 2n + 1}{4n^2 + 5n + 6} = \frac{3 + \frac{2}{n} + \frac{1}{n^2}}{4 + \frac{5}{n} + \frac{6}{n^2}} \xrightarrow{n \rightarrow \infty} \frac{3}{4}$ (nach den Regeln (1) und (3)).

Reihe Folge von Teilsummen. Schreibweise: $(s_n)_{n \in \mathbb{N}}$
 Eine Reihe entsteht aus einer Folge $(a_n)_{n \in \mathbb{N}}$ durch Aufsummieren: $s_n = a_1 + \dots + a_n = \sum_{i=1}^n a_i$

Folge $(a_n)_{n \in \mathbb{N}}$	Reihe $(s_n)_{n \in \mathbb{N}}$
a_1	$s_1 = a_1$
a_2	$s_2 = a_1 + a_2$
a_3	$s_3 = a_1 + a_2 + a_3$
\vdots	\vdots
a_n	$s_n = a_1 + a_2 + \dots + a_n$

Regel Wenn $(s_n)_{n \in \mathbb{N}}$ konvergiert, muss $(a_n)_{n \in \mathbb{N}}$ notwendigerweise eine Nullfolge bilden.
 Oder äquivalent: Wenn $(a_n)_{n \in \mathbb{N}}$ keine Nullfolge ist, kann auch $(s_n)_{n \in \mathbb{N}}$ nicht konvergieren.
 Aber: Wenn $(a_n)_{n \in \mathbb{N}}$ eine Nullfolge ist, muss $(s_n)_{n \in \mathbb{N}}$ nicht unbedingt konvergieren.

Beispiele (1) $a_n = 1 + \frac{1}{n} \xrightarrow{n \rightarrow \infty} 1 \neq 0$ (keine Nullfolge) $\Rightarrow s_n$ kann nicht konvergieren
 (2) $a_n = \frac{1}{n} \xrightarrow{n \rightarrow \infty} 0$ (Nullfolge), aber s_n (harmonische Reihe) konvergiert trotzdem nicht, denn für

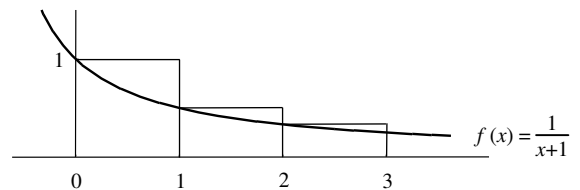
$n = 2^m$ (also $m = \log_2(n)$) erhält man

$$s_n = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \dots + \frac{1}{8} + \frac{1}{9} + \dots + \frac{1}{16} + \frac{1}{17} + \dots + \frac{1}{32} + \dots + \frac{1}{2^{m-1}+1} + \dots + \frac{1}{2^m}$$

$$> 1 + \frac{1}{2} + 2 \cdot \frac{1}{4} + 4 \cdot \frac{1}{8} + 8 \cdot \frac{1}{16} + 16 \cdot \frac{1}{32} + \dots + 2^{m-1} \cdot \frac{1}{2^m}$$

$$= 1 + \frac{1}{2} \cdot m = 1 + \frac{1}{2} \cdot \log_2(n) \xrightarrow{n \rightarrow \infty} \infty.$$

Alternativ lässt sich die Reihe grafisch veranschaulichen als Fläche von Rechtecken mit der Breite 1 und der Höhe $\frac{1}{n}$. Die Fläche dieser Rechtecke



ist größer als die Fläche zwischen x -Achse und dem Grafen der Funktion $f(x) = \frac{1}{x+1}$, die sich mit Hilfe der Integralrechnung einfach bestimmen lässt (siehe Abschnitt Integrale):

$$\int_0^n f(x) dx = [\ln(x+1)]_0^n = \ln(n+1) - \ln(1) = \ln(n+1).$$

Somit erhält man $s_n > \ln(n+1) \xrightarrow{n \rightarrow \infty} \infty$.

(3) $a_n = \frac{1}{n^2} \xrightarrow{n \rightarrow \infty} 0$ und s_n konvergiert, denn wegen $\frac{1}{n^2} < \frac{1}{n \cdot (n-1)} = \frac{1}{n-1} - \frac{1}{n}$ ist

$$s_n = 1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \dots + \frac{1}{n^2} < 1 + \frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} + \frac{1}{3 \cdot 4} + \dots + \frac{1}{n \cdot (n-1)}$$

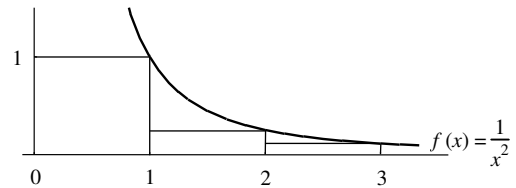
$$= 1 + (1 - \frac{1}{2}) + (\frac{1}{2} - \frac{1}{3}) + (\frac{1}{3} - \frac{1}{4}) + \dots + (\frac{1}{n-1} - \frac{1}{n}) = 2 - \frac{1}{n} \xrightarrow{n \rightarrow \infty} 2.$$

s_n ist also beschränkt durch 2 (und monoton steigend) und somit konvergent. Der Grenzwert beträgt $\frac{\pi^2}{6} = 1,64\dots$, was hier nicht nachgewiesen werden soll. Der Grund für die Konvergenz dieser Reihe im Gegensatz zur harmonischen Reihe liegt in der höheren Geschwindigkeit, mit der die Folge a_n gegen 0 strebt.

Alternativ kann auch hier die Beschränktheit von s_n mit Hilfe von Integralen leicht nachgewiesen werden, denn die Fläche der Rechtecke ist

in diesem Fall kleiner als die Fläche zwischen x -Achse und dem Grafen der

Funktion $f(x) = \frac{1}{x^2}$, d.h. es gilt:



$$s_n < 1 + \int_1^n f(x) dx = 1 + \left[-\frac{1}{x} \right]_1^n = 1 + \left(-\frac{1}{n} - (-1) \right) = 2 - \frac{1}{n} \xrightarrow{n \rightarrow \infty} 2.$$

(4) $a_n = \frac{1}{n!} \xrightarrow{n \rightarrow \infty} 0$ und $s_n \xrightarrow{n \rightarrow \infty} e - 1$

(5) $a_n = (-1)^{n+1} \cdot \frac{1}{n} \xrightarrow{n \rightarrow \infty} 0$ und $s_n \xrightarrow{n \rightarrow \infty} \ln(2)$

(6) *Arithmetische Reihe:* $s_n = n \cdot a_1 + \frac{n \cdot (n-1)}{2} \cdot d \xrightarrow{n \rightarrow \infty} \pm \infty$, falls $d \neq 0$

(7) *Geometrische Reihe:* $s_n = a_1 \cdot \frac{q^n - 1}{q - 1} \xrightarrow{n \rightarrow \infty} \frac{1}{1 - q}$, falls $|q| < 1$

Die Darstellung der arithmetischen und geometrischen Reihe ergibt sich dabei wie folgt:

n	arithm. Folge	arithm. Reihe	geom. Folge	geom. Reihe
1	a_1	a_1	a_1	a_1
2	$a_1 + d$	$2 a_1 + d$	$a_1 \cdot q$	$a_1 + a_1 \cdot q = a_1 \cdot (1+q)$
3	$a_1 + 2d$	$3 a_1 + d + 2d$	$a_1 \cdot q^2$	$a_1 \cdot (1+q+q^2)$
\vdots	\vdots	\vdots	\vdots	\vdots
n	$a_1 + (n-1)d$	$n a_1 + d(1+2+\dots+(n-1))$	$a_1 \cdot q^{n-1}$	$a_1 \cdot (1+q+q^2+\dots+q^{n-1})$

Dabei ist

$$1 + 2 + \dots + (n-1) + (n-1) + (n-2) + \dots + 1 = \frac{n \cdot (n-1)}{2} + \frac{n \cdot (n-1)}{2} = n \cdot (n-1),$$

also $2(1+2+\dots+(n-1)) = n \cdot (n-1)$ und somit $1 + 2 + \dots + (n-1) = \frac{n \cdot (n-1)}{2}$.

Analog ist

$$T - \frac{T \cdot q}{q - 1} = \frac{1 + q + q^2 + \dots + q^{n-1} - (q + q^2 + \dots + q^{n-1} + q^n)}{q - 1} = \frac{1 - q^n}{q - 1},$$

also $T \cdot (1 - q) = 1 - q^n$ und somit $T = 1 + q + q^2 + \dots + q^{n-1} = \frac{1 - q^n}{1 - q} = \frac{q^n - 1}{q - 1}$.

9. Finanzmathematik (Zinseszinsrechnung)

Beispiel (1)

Lineares bzw. progressives Wachstum

Umsatz eines Unternehmens 2018: 2,1 Mio. €.

Geplante jährliche Steigerung alternativ (1) linear: 0,1 Mio. € oder

(2) progressiv: 4% gegenüber dem Vorjahr.

Gesucht: (a) Umsatz im Jahre 2024, (b) Summe der Umsätze bis einschl. 2024.

Lösung: Die Umsatzentwicklung in beiden Szenarien ergibt sich aus der nachstehenden Tabelle.

Jahr	(1) lineares Wachstum		(2) progressives Wachstum	
	(a) Umsatz	(b) Summe	(a) Umsatz	(b) Summe
2018	2,1	2,1	2,1	2,1
2019	2,2	4,3	$2,1 \cdot 1,04 = 2,184$	4,284
2020	2,3	6,6	$2,184 \cdot 1,04 = 2,271$	6,555
⋮	⋮	⋮	⋮	⋮
2024	2,7	16,8	2,657	16,586

Das lineare Wachstum kann mittels arithmetischer Folge bzw. Reihe beschrieben werden, das progressive Wachstum mittels geometrischer Folge bzw. Reihe:

zu (1): $a_1 = 2,1$, $d = 0,1$, $n = 7$ (2018 zählt schon mit)

$$(a) a_n = a_1 + (n-1)d, \text{ also hier } a_7 = 2,1 + 6 \cdot 0,1 = 2,7;$$

$$(b) s_n = n \cdot a_1 + \frac{n \cdot (n-1)}{2} \cdot d, \text{ also hier } s_7 = 7 \cdot 2,1 + \frac{7 \cdot 6}{2} \cdot 0,1 = 14,7 + 2,1 = 16,8;$$

zu (2): $a_1 = 2,1$, $q = 1,04$, $n = 7$

$$(a) a_n = a_1 \cdot q^{n-1}, \text{ also hier } a_7 = 2,1 \cdot 1,04^6 = 2,1 \cdot 1,265 = 2,657;$$

$$(b) s_n = a_1 \cdot \frac{q^n - 1}{q - 1}, \text{ also hier } s_7 = 2,1 \cdot \frac{1,04^7 - 1}{1,04 - 1} = 2,1 \cdot 7,898 = 16,586.$$

Analog kann auch die lineare bzw. degressive Abschreibung dargestellt werden.

Beispiel (2)

Zinseszinsrechnung mit ein- bzw. mehrmaliger Einzahlung

Auf ein Konto wird alternativ eingezahlt (1) ein einmaliger Betrag K oder (2) ein jährlicher Betrag E .

Das Guthaben wird mit p % verzinst und die Zinsen dem Konto nach jedem Jahr gutgeschrieben. Gesucht ist das Guthaben K_n nach n Jahren.

Lösung: Das Guthaben in beiden Szenarien ergibt sich aus der nachstehenden Tabelle.

Dabei bezeichnet $q = 1 + p\% = 1 + \frac{p}{100}$ den Zinsfaktor.

Jahre	(1) Guthaben K_n bei einmaliger Einzahlung	(2) Guthaben K_n bei jährlicher Einzahlung
1	$K \cdot q$	$E \cdot q$
2	$K \cdot q^2$	$E \cdot q^2 + E \cdot q = E \cdot q \cdot (q+1)$
3	$K \cdot q^3$	$E \cdot q^3 + E \cdot q^2 + E \cdot q = E \cdot q \cdot (q^2 + q + 1)$
⋮	⋮	⋮
n	$K \cdot q^n$	$E \cdot q \cdot (1 + q + q^2 + \dots + q^{n-1}) = E \cdot q \cdot \frac{q^n - 1}{q - 1}$

Die Entwicklung des Guthabens lässt sich also mit Hilfe der geometrischen Folge bzw. Reihe beschreiben. Die hier abgeleiteten Formeln

$$(1) K_n = K \cdot q^n \quad (\text{bei einmaliger Einzahlung}),$$

$$(2) K_n = E \cdot q \cdot \frac{q^n - 1}{q - 1} \quad (\text{bei jährlicher Einzahlung})$$

reduzieren sich um den Faktor q , wenn die Einzahlung nicht wie hier unterstellt vorschüssig, also zu Beginn des Jahres erfolgt, sondern nachschüssig, d.h. am Ende des Jahres.

Häufig müssen die Formeln nach einer anderen Variablen umgestellt werden, wie im folgenden

Beispiel: Auf einen Bausparvertrag mit 2,5 % iger Verzinsung sollen jeweils zu Beginn eines Jahres 2.000 € eingezahlt werden. Die Zinsen werden dem Konto am Ende jeden Jahres gutgeschrieben.

Wann sind 20.000 € angespart ?

Lösung: Hier ist die Formel (2) anzuwenden. Gegeben sind $K_n = 20.000$, $E = 2.000$ und

$q = 1 + \frac{2,5}{100} = 1,025$. Gesucht wird n . Dazu wird die Formel nach n umgestellt:

$$K_n = E \cdot q \cdot \frac{q^n - 1}{q - 1} \Leftrightarrow \frac{K_n}{E} \cdot \frac{q - 1}{q} + 1 = q^n \Leftrightarrow n = \frac{\ln\left(\frac{K_n}{E} \cdot \frac{q - 1}{q} + 1\right)}{\ln(q)}$$

Für das Beispiel erhält man so $n = \frac{\ln\left(\frac{20.000}{2.000} \cdot \frac{0,025}{1,025} + 1\right)}{\ln(1,025)} = 8,84$, also nach 9 Jahren.

Beispiel (3)

Unterjährige Zinszahlung; *stetige (momentane)* Verzinsung

Zu Beginn eines Jahres werden $K = 100.000$ € auf ein Konto mit 5 % iger Verzinsung eingezahlt.

Gesucht: Kapital nach einem Jahr bei (1) jährlicher, (2) vierteljährlicher, (3) monatlicher, (4) täglicher und (5) stetiger (momentaner) Zinszahlung.

Lösung: Das Guthaben in den 5 Anlageformen ergibt sich aus der nachstehenden Übersicht.

<u>Zinszahlung</u>	<u>Kapital nach einem Jahr</u>
(1) jährlich	$100.000 \cdot \left(1 + \frac{5}{100}\right) = 105.000$
(2) vierteljährlich	$100.000 \cdot \left(1 + \frac{5}{100} \cdot \frac{1}{4}\right)^4 = 105.094,53$
(3) monatlich	$100.000 \cdot \left(1 + \frac{5}{100} \cdot \frac{1}{12}\right)^{12} = 105.116,19$
(4) täglich	$100.000 \cdot \left(1 + \frac{5}{100} \cdot \frac{1}{360}\right)^{360} = 105.126,75$
(5) stetig (momentan)	$100.000 \cdot \lim_{k \rightarrow \infty} \left(1 + \frac{5}{100} \cdot \frac{1}{k}\right)^k = 100.000 \cdot e^{5/100} = 105.127,11$

Der Grenzwert in der letzten Zeile basiert auf der Aussage $\lim_{k \rightarrow \infty} \left(1 + x \cdot \frac{1}{k}\right)^k = e^x$, die später mit Hilfe der Regeln von l' Hospital einfach nachgewiesen werden kann.

Stetiges Wachstum spielt eine wichtige Rolle in Natur, Technik und Ökonomie.

Die Ergebnisse der vorausgegangenen Beispiele seien noch einmal zusammengefasst:

Bezeichnungen	m : Anzahl der Jahre	k : Anzahl der Zinszahlungen pro Jahr
	$n = m \cdot k$: Anzahl der Zinsperioden insgesamt	K_n : Guthaben nach n Zinsperioden
	p : Zinsfuß, $i = p \% = \frac{p}{100}$: Zinssatz (nominal)	$q = 1 + \frac{i}{k} = 1 + \frac{p}{k \cdot 100}$: Zinsfaktor

Einmalige Einzahlung eines Betrages K und

k -malige Zinsgutschrift pro Jahr: $K_n = K \cdot q^n = K \cdot \left(1 + \frac{i}{k}\right)^{k \cdot m}$

stetige Verzinsung: $K_m = \lim_{k \rightarrow \infty} K \cdot \left(1 + \frac{i}{k}\right)^{k \cdot m} = K \cdot e^{m \cdot i}$

Mehrmalige Einzahlung eines Betrages E

vorschüssig, d.h. zu Beginn jeder Zinsperiode:⁶ $K_n = E \cdot q \cdot \frac{q^n - 1}{q - 1}$

nachschüssig, d.h. am Ende jeder Zinsperiode: $K_n = E \cdot \frac{q^n - 1}{q - 1}$

10. Globale Extrema von Funktionen einer Variablen

Relevanter Bereich Menge der x -Werte, die für das *aktuelle Problem* relevant sind.

In der Regel ein Intervall oder die reellen Zahlen insgesamt.

Definitionsbereich D_f einer Funktion f : Menge aller x -Werte, für die $f(x)$ mathematisch erklärt ist.

Definitionslücken treten z.B. an Stellen auf, an denen durch 0 geteilt würde oder die Quadratwurzel aus einer negativen Zahl zu ziehen wäre.

Beachten: Definitionsbereich und relevanter Bereich sind zwei unterschiedliche Dinge.

Lokale Extrema oder relative Extrema: Punkte mit dem größten (Maximum) bzw. kleinsten (Minimum)

Funktionswert von allen x -Werten in einer (kleinen) Umgebung.

⁶ Wegen $q \cdot \frac{q^n - 1}{q - 1} = \frac{q}{q - 1} \cdot (q^n - 1)$, $\frac{q}{q - 1} = \frac{1 + i/k}{i/k} = 1 + \frac{k}{i}$, gilt auch $K_n = E \cdot \left(1 + \frac{k}{i}\right) \cdot (q^n - 1)$.

Lokale Extrema können vorliegen an

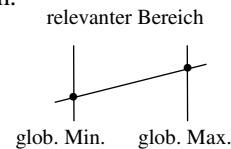
- (1) Stellen mit *Steigung* 0, z.B. ein lokales Maximum \cap oder ein lokales Minimum \cup ,
- (2) *Sprungstellen* der Funktion, z.B. ein lokales Maximum \swarrow oder ein lokales Minimum \searrow ,
- (3) *Knickstellen* der Funktion, z.B. ein lokales Maximum \wedge oder ein lokales Minimum \vee .

Das Auffinden von Sprung- bzw. Knickstellen, also Stellen, an denen die Funktion nicht *stetig* bzw. nicht *differenzierbar* ist, wird als erstes besprochen. Anschließend wird dargestellt, wie die Steigung einer Funktion mit Hilfe der 1. *Ableitung* bestimmt werden kann.

Globale Extrema oder absolute Extrema: Punkte mit dem größten (Maximum) bzw. kleinsten (Minimum) Funktionswert von allen x -Werten im relevanten Bereich.

"Kandidaten" für globale Extrema:

- (1) Lokale Extrema,
- (2) Ränder des relevanten Bereiches, z.B.



Lokale bzw. globale Extrema müssen nicht eindeutig bestimmt sein, d.h. es kann z.B. mehrere globale Maxima geben, die alle einen gleich hohen Funktionswert aufweisen. Eindeutig bestimmte Extrema werden als *strenge* (strikte) Extrema bezeichnet.

Stetigkeit f ist *stetig* in $x_0 \in \mathbb{D}_f$, wenn für jede Folge $(x_n)_{n \in \mathbb{N}}$ mit $\lim_{n \rightarrow \infty} x_n = x_0$ gilt: $\lim_{n \rightarrow \infty} f(x_n) = f(x_0)$.

Ist f an einer Stelle x_0 nicht definiert, also $x_0 \notin \mathbb{D}_f$, so ist f dort weder stetig noch unstetig.

Eine *Sprungstelle* liegt vor, wenn die Funktion dort nicht stetig ist.

f ist *stetig in einer Menge* M , wenn f an jeder Stelle $x \in M$ stetig ist.

Die nachstehenden Regeln ergeben sich sofort aus den entsprechenden Regeln für Folgen.

- Regeln**
- (1) Sind f und g stetig in x_0 , dann auch $f + g$, $f - g$, $f \cdot g$ und f / g , $g(x_0) \neq 0$.
 - (2) Ist f stetig in x_0 und g stetig in $y_0 = f(x_0)$, dann ist $g \circ f$ stetig in x_0 .

- Beispiele**
- (1) Die Funktion $f(x) = x$ ist stetig in \mathbb{R} , denn $\forall x_0 \in \mathbb{R}$ gilt: $x_n \xrightarrow{n \rightarrow \infty} x_0 \Rightarrow f(x_n) = x_n \xrightarrow{n \rightarrow \infty} x_0 = f(x_0)$.
 - (2) Wenn $f(x) = x$ stetig ist, dann auch $f(x) + f(x) = 2x$, $f(x) \cdot f(x) = x^2$, $x^2 + 2x$, ...
 \Rightarrow Alle Polynome sind stetig in \mathbb{R} .
 - (3) Auch die Funktionen x^a , a^x , $\log_a(x)$ sowie die trigonometrischen Funktionen sind überall in ihrem Definitionsbereich stetig.

Da auch die Verknüpfung all dieser Funktionen wiederum stetig ist, erhält man insgesamt:

Alle in der Ökonomie relevanten Funktionen⁷ sind überall in ihrem Definitionsbereich stetig.

Ausnahme: Stückweise definierte Funktionen können dort, wo sie zusammengesetzt sind ("Nahtstellen") evtl. unstetig sein. Nur hier muss eine Überprüfung auf Stetigkeit erfolgen.

Nahtstelle $x_0 \in \mathbb{D}_f$ ist eine *Nahtstelle* von f , wenn die Funktion in einer Umgebung von x_0 nicht einheitlich

$$\text{definiert ist, also } f(x) = \begin{cases} f_l(x) & x < x_0 \\ f(x_0) & \text{für } x = x_0 \\ f_r(x) & x > x_0 \end{cases}$$

x_0 gehört dabei praktisch immer entweder zum linken Funktionsast f_l oder zum rechten

$$\text{Funktionsast } f_r, \text{ also } f(x) = \begin{cases} f_l(x) & \text{für } x \leq x_0 \\ f_r(x) & \text{für } x > x_0 \end{cases} \quad \text{oder} \quad f(x) = \begin{cases} f_l(x) & \text{für } x < x_0 \\ f_r(x) & \text{für } x \geq x_0 \end{cases}$$

Wir beschränken uns daher im Weiteren auf diese beiden Fälle.

In der Praxis treten stückweise definierte Funktionen beispielsweise auf bei Angabe des Preises p eines Produktes in Abhängigkeit von der Bestellmenge x , z.B. $p(x) = \begin{cases} 5 \text{ € / Stück} & \text{für } x < 100 \\ 4 \text{ € / Stück} & \text{für } x \geq 100 \end{cases}$.

- Regel** f ist stetig an einer Nahtstelle $x_0 \in \mathbb{D}_f$, wenn $f_l(x_0) = f_r(x_0)$.
 Sind die Funktionswerte nicht gleich, so ist f unstetig in x_0 .

⁷ Für die Ökonomie irrelevant ist beispielsweise die Funktion $f(x) = 0$ für rationales x , $f(x) = 1$ für irrationales x , die überall in \mathbb{R} definiert, jedoch an keiner Stelle stetig ist.

Beispiele

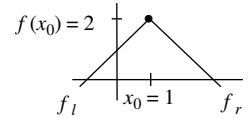
$$(1) f(x) = \begin{cases} f_l(x) = x + 1 & \text{für } x \leq 1 \\ f_r(x) = -x + 3 & \text{für } x > 1 \end{cases} \quad \text{bzw.} \quad f(x) = \begin{cases} f_l(x) = x + 1 & \text{für } x < 1 \\ f_r(x) = -x + 3 & \text{für } x \geq 1 \end{cases} .$$

Die Nahtstelle $x_0 = 1$ gehört bei der ersten Funktion zum linken

Funktionsast f_l , bei der zweiten Funktion zum rechten Funktionsast f_r .

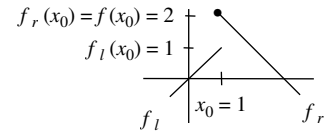
In beiden Fällen ist f stetig an der Nahtstelle $x_0 = 1$ (und somit $\forall x \in \mathbb{R}$),

da $f_l(x_0) = 2 = f_r(x_0)$.



$$(2) f(x) = \begin{cases} f_l(x) = x & \text{für } x < 1 \\ f_r(x) = -x + 3 & \text{für } x \geq 1 \end{cases}$$

$f_l(x_0) = 1, f_r(x_0) = 2$, also $f_l(x_0) \neq f_r(x_0)$, d.h. f ist unstetig in x_0 .



Sonderfall

Ist $f_l(x_0)$ mathematisch nicht erklärt ($x_0 \notin \mathbb{D}_{f_l}$), so muss statt dessen der Grenzwert $\lim_{\substack{x \rightarrow x_0 \\ x < x_0}} f_l(x)$ berechnet und wie oben bei der Überprüfung auf Stetigkeit verwendet werden.

Analog: Wenn $x_0 \notin \mathbb{D}_{f_r}$, dann $\lim_{\substack{x \rightarrow x_0 \\ x > x_0}} f_r(x)$ verwenden.

Ist einer der Grenzwerte gleich ∞ oder $-\infty$, so ist f unstetig in x_0 .

Beispiele

$$(1) f(x) = \begin{cases} f_l(x) = x + 1 & \text{für } x \leq 0 \\ f_r(x) = \frac{e^x - 1}{x} & \text{für } x > 0 \end{cases} . \text{ An der Nahtstelle } x_0 = 0 \text{ ist } f_l(x_0) = f(x_0) = 1;$$

$f_r(x_0)$ ist nicht definiert ($x_0 \notin \mathbb{D}_{f_r}$). Berechne daher $\lim_{\substack{x \rightarrow x_0 \\ x > x_0}} f_r(x) = \lim_{\substack{x \rightarrow 0 \\ x > 0}} \frac{e^x - 1}{x} = 1$.

(Berechnung des Grenzwertes: siehe Abschnitt 13 (Regeln von l'Hospital), Bsp. (2)).

f ist somit stetig an der Nahtstelle x_0 .

$$(2) f(x) = \begin{cases} f_l(x) = 0 & \text{für } x \leq 0 \\ f_r(x) = 1/x & \text{für } x > 0 \end{cases} . \text{ An der Nahtstelle } x_0 = 0 \text{ ist } f_l(x_0) = f(x_0) = 0;$$

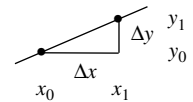
$f_r(x_0)$ ist nicht definiert ($x_0 \notin \mathbb{D}_{f_r}$). Berechne daher $\lim_{\substack{x \rightarrow x_0 \\ x > x_0}} f_r(x) = \lim_{\substack{x \rightarrow 0 \\ x > 0}} \frac{1}{x} = \infty$.

f ist somit unstetig an der Nahtstelle x_0 .

Steigung

- **Steigung s einer Geraden:** Verhältnis von vertikaler zu horizontaler Änderung zwischen 2

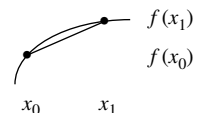
Punkten (x_0, y_0) und (x_1, y_1) auf der Geraden: $s = \frac{y_1 - y_0}{x_1 - x_0} = \frac{\Delta y}{\Delta x}$



- **Sekante:** Verbindungsgerade zwischen 2 Punkten $(x_0, f(x_0))$ und $(x_1, f(x_1))$ auf dem Grafen einer Funktion f .

- **Sekantensteigung:** Steigung der Sekanten: $s_{01} = \frac{f(x_1) - f(x_0)}{x_1 - x_0}$

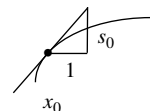
(durchschnittliche Steigung der Funktion f im Intervall $[x_0, x_1]$).



- **Steigung von f an einer Stelle x_0 :** $s_0 = \lim_{n \rightarrow \infty} \frac{f(x_n) - f(x_0)}{x_n - x_0} = \lim_{n \rightarrow \infty} s_{0n}$, wobei $(x_n)_{n \in \mathbb{N}}$ eine Folge

von Zahlen ist, die gegen x_0 konvergiert. Die Steigung in einem Punkt ist also als Grenzwert einer Folge von Sekantensteigungen definiert, wobei der Abstand zwischen den beiden Stellen x_0 und x_n gegen 0 geht. s_0 gibt die

- **Steigung der Tangenten** an, die durch den Punkt $(x_0, f(x_0))$ an den Funktionsgraphen gelegt wird. Zur konkreten Berechnung der Steigung von f an einer Stelle x_0 verwenden wir später den Funktionswert der



- 1. **Ableitung an der Stelle x_0 :** $s_0 = f'(x_0)$.

Bemerkung zur Schreibweise: Jede Folge $(x_n)_{n \in \mathbb{N}}$, die gegen x_0 konvergiert, lässt sich schreiben als $x_n = x_0 + h_n$, wobei $(h_n)_{n \in \mathbb{N}}$ eine Folge ist, die gegen 0 konvergiert. Wegen $x_n - x_0 = h_n$ erhält man damit die häufig verwendete, alternative Darstellung für die Definition der Steigung

$$s_0 = \lim_{n \rightarrow \infty} \frac{f(x_0 + h_n) - f(x_0)}{h_n} \quad \text{oder kurz} \quad s_0 = \lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h} .$$

Beispiel

Betrachte die Funktion $f(x) = x^2$ an der Stelle $x_0 = 1$ mit $f(x_0) = x_0^2 = 1$.

- Die Steigung der Sekante zwischen den beiden Punkten $(x_0, f(x_0))$ und $(x_1, f(x_1))$ beträgt dann $s_{01} = \frac{f(x_1) - f(x_0)}{x_1 - x_0} = \frac{x_1^2 - 1}{x_1 - 1} = \frac{(x_1 - 1) \cdot (x_1 + 1)}{x_1 - 1} = x_1 + 1$, also z.B. für $x_1 = 2$ ist $s_{01} = 3$.
- Sei $(x_n)_{n \in \mathbb{N}}$ nun eine Folge, die gegen $x_0 = 1$ konvergiert. Für die entsprechende Folge von Sekantensteigungen s_{0n} gilt $s_{0n} = \frac{f(x_n) - f(x_0)}{x_n - x_0} = \frac{x_n^2 - 1}{x_n - 1} = \frac{(x_n - 1) \cdot (x_n + 1)}{x_n - 1} = x_n + 1 \xrightarrow{n \rightarrow \infty} 1 + 1 = 2$, d.h. die Steigung von f an der Stelle $x_0 = 1$ ist $s_0 = \lim_{n \rightarrow \infty} s_{0n} = 2$.

Exemplarisch für die Folge $x_n = 1 + \frac{1}{n}$ ergibt sich so die nachstehende Tabelle:

n	1	2	3	4	... $\rightarrow \infty$
x_n	2	1,5	1,33	1,25	... $\rightarrow x_0 = 1$
s_{0n}	3	2,5	2,33	2,25	... $\rightarrow s_0 = 2$

- Abschließend bezeichne x_0 nun eine beliebige Zahl. Dann gilt analog zur obigen Rechnung:

$$s_{0n} = \frac{f(x_n) - f(x_0)}{x_n - x_0} = \frac{x_n^2 - x_0^2}{x_n - x_0} = \frac{(x_n - x_0) \cdot (x_n + x_0)}{x_n - x_0} = x_n + x_0 \xrightarrow{n \rightarrow \infty} 2x_0, \text{ falls } x_n \xrightarrow{n \rightarrow \infty} x_0.$$

Damit ist die Steigung der Funktion f an einer beliebigen Stelle x_0 gleich $s_0 = 2x_0$:

x_0	-2	-1	0	1	2	3
s_0	-4	-2	0	2	4	6

Die so ermittelte "Steigungsfunktion" bezeichnet man mit f' (hier $f'(x) = 2x$) und nennt sie die 1. Ableitung von f (siehe nächste Seite).

Differenzierbarkeit f ist differenzierbar (diff.bar) in x_0 , wenn die Steigung dort eindeutig bestimmt werden kann, also von der Wahl der Folge $(x_n)_{n \in \mathbb{N}}$ nicht abhängt.

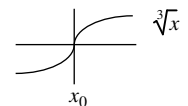
Eine *Knickstelle* liegt vor, wenn die Funktion dort stetig, aber nicht diff.bar ist.

f ist differenzierbar in einer Menge M , wenn f an jeder Stelle $x \in M$ diff.bar ist.

Beispiele

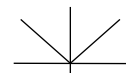
- Die Funktion $f(x) = x^2$ ist $\forall x \in \mathbb{R}$ diff.bar, da die Steigung $\forall x \in \mathbb{R}$ eindeutig bestimmt werden kann (siehe Beispiel oben). Gleiches gilt für alle Polynome.
- Die Funktion $f(x) = \sqrt[3]{x}$ ist an der Stelle $x_0 = 0$ stetig, aber nicht diff.bar.

Die Steigung dort ist ∞ , denn $s_{0n} = \frac{f(x_n) - f(x_0)}{x_n - x_0} = \frac{\sqrt[3]{x_n} - 0}{x_n - 0} \xrightarrow{n \rightarrow \infty} \infty$,



falls $x_n \xrightarrow{n \rightarrow \infty} 0$. Man könnte hier von "uneigentlicher Diff.barkeit" sprechen; optisch liegt kein Knick vor.

- Die Funktion $f(x) = |x| = \begin{cases} -x & \text{für } x < 0 \\ x & \text{für } x \geq 0 \end{cases}$



ist an der Stelle $x_0 = 0$ stetig, aber nicht diff.bar, denn für die Folge $x_n = -1/n$ ist stets $s_{0n} = -1$, also auch die Steigung $s_0 = \lim_{n \rightarrow \infty} s_{0n} = -1$, während man für die Folge $x_n = 1/n$ die

Sekantensteigungen $s_{0n} = 1$, also auch $s_0 = 1$ erhält. Die Steigung an der Stelle $x_0 = 0$ ist daher nicht eindeutig bestimmt, somit f nicht diff.bar in x_0 .

Man sieht: Wenn f in x_0 stetig ist, so muss f dort nicht notwendig auch diff.bar sein.

Umgekehrt zeigt man leicht die folgende Regel.

Regel

Wenn f in x_0 diff.bar ist, dann ist f dort auch stetig.

Oder äquivalent: Wenn f nicht stetig ist in x_0 , dann auch nicht diff.bar.

Beispiel (4): Die Funktion $f(x) = \begin{cases} x & \text{für } x \leq 0 \\ x+1 & \text{für } x > 0 \end{cases}$



ist an der Stelle $x_0 = 0$ nicht stetig und daher dort nicht diff.bar.

Ähnlich wie im Fall der Stetigkeit zeigt sich, dass die in der Praxis relevanten Funktionen "fast überall" in ihrem Definitionsbereich diff.bar sind. "Fast überall" bedeutet, dass einige Funktionen an bestimmten Stellen die Steigung ∞ oder $-\infty$ besitzen (siehe Beispiel (2)).

Ansonsten können Knickstellen allenfalls an den Nahtstellen stückweise definierter Funktionen auftreten (siehe Beispiel (3)). Die praktische Überprüfung auf Diff.barkeit bzw. die Berechnung der Steigung erfolgt nun nicht mit Hilfe von Folgen, sondern durch Ableitungen.

Ableitungsfunktionen Zu jeder Funktion f lässt sich mittels der weiter unten aufgeführten Ableitungsregeln deren erste *Ableitungsfunktion* (kurz: 1. *Ableitung*) f' angeben. f' besitzt die Eigenschaft, dass ihr Funktionswert an einer Stelle x_0 , $f'(x_0)$, die Steigung von f an der Stelle x_0 misst. Eine andere Schreibweise für $f'(x)$, auf die wir im Zusammenhang mit dem Differenzial noch einmal zurückkommen, ist $\frac{df}{dx}(x)$.

Durch erneutes Anwenden der Ableitungsregeln erhält man aus f' die 2. *Ableitung* f'' , daraus die 3. *Ableitung* f''' und schließlich allgemein die n -te *Ableitung* $f^{(n)}$.

Existiert $f^{(n)}(x_0)$, so nennt man f *n-fach diff.bar* an der Stelle x_0 .

Im Bereich der Wirtschaftswissenschaft bezeichnet man die 1. Ableitung in der Regel als *Grenzfunktion* (z.B. Grenzkosten, Grenzerlöse etc.) oder als *Marginalfunktion* (z.B. marginale Konsumquote).

Regel f ist diff.bar an einer Nahtstelle $x_0 \in D_f$, wenn f dort stetig ist und zusätzlich gilt $f_l'(x_0) = f_r'(x_0)$. Ist f unstetig in x_0 oder gilt $f_l'(x_0) \neq f_r'(x_0)$, so ist f in x_0 nicht diff.bar. f_l' bzw. f_r' bezeichnen die 1. Ableitung des linken bzw. rechten Funktionsastes, d.h. $f_l'(x_0)$ bzw. $f_r'(x_0)$ geben die linksseitige bzw. rechtsseitige Steigung von f an der Stelle x_0 an.

- Beispiele**
- $f(x) = \begin{cases} f_l(x) = x^2 & \text{für } x \leq 1 \\ f_r(x) = 2x - 1 & \text{für } x > 1 \end{cases}$ ist stetig an der Nahtstelle $x_0 = 1$, da $f_l(x_0) = f(x_0) = 1 = f_r(x_0)$. Mit $f_l'(x) = 2x$ und $f_r'(x) = 2$ erhält man $f_l'(x_0) = 2 = f_r'(x_0)$, d.h. f ist diff.bar an der Stelle x_0 (und damit $\forall x \in \mathbb{R}$).
 - $f(x) = |x|$ (siehe oben Beispiel (3)) mit den Funktionsästen $f_l(x) = -x$ und $f_r(x) = x$ ist stetig an der Nahtstelle $x_0 = 0$, da $f_l(x_0) = 0 = f_r(x_0) = f(x_0)$. Wegen $f_l'(x) = -1$ und $f_r'(x) = 1$ ist $f_l'(x_0) = -1 \neq 1 = f_r'(x_0)$, d.h. f ist nicht diff.bar an der Stelle x_0 .
 - Im obigen Beispiel (4) mit $f_l(x) = x$ und $f_r(x) = x + 1$ ist f nicht stetig an der Nahtstelle $x_0 = 0$ und somit dort auch nicht diff.bar, obwohl $f_l'(x) = 1$ und $f_r'(x) = 1$, also $f_l'(x_0) = 1 = f_r'(x_0)$ (links- und rechtsseitige Steigung stimmen zwar überein, aber in x_0 liegt eine Sprungstelle vor).

Sonderfall Ist $f_l'(x_0)$ mathematisch nicht erklärt, so muss statt dessen der Grenzwert $\lim_{\substack{x \rightarrow x_0 \\ x < x_0}} \frac{f_l(x) - f_l(x_0)}{x - x_0} = \lim_{\substack{x \rightarrow x_0 \\ x < x_0}} f_l'(x)$ berechnet und zur Überprüfung auf Diff.barkeit verwendet werden.

Analog: Wenn $x_0 \notin D_{f_r'}$, dann $\lim_{\substack{x \rightarrow x_0 \\ x > x_0}} \frac{f_r(x) - f_r(x_0)}{x - x_0} = \lim_{\substack{x \rightarrow x_0 \\ x > x_0}} f_r'(x)$ verwenden.

Ist einer der Grenzwerte gleich ∞ oder $-\infty$, so ist f nicht diff.bar in x_0 .

- Beispiele**
- $f(x) = \begin{cases} f_l(x) = 0 & \text{für } x \leq 0 \\ f_r(x) = x^2 \cdot (2 \ln(x) - 1) & \text{für } x > 0 \end{cases}$ besitzt die Ableitung $f_r'(x) = 4x \cdot \ln(x)$ (siehe Ableitungsregeln). f_r und f_r' sind an der Stelle x_0 nicht definiert, da $\ln(0)$ nicht erklärt ist. Mit Hilfe der später zu besprechenden Regeln von l'Hospital zeigt man:
 $-\lim_{\substack{x \rightarrow 0 \\ x > 0}} f_r(x) = 0 = f_l(0)$, also f stetig in x_0 , $-\lim_{\substack{x \rightarrow 0 \\ x > 0}} f_r'(x) = 0 = f_l'(0)$, also f diff.bar in x_0 .
 - $f(x) = \begin{cases} f_l(x) = 0 & \text{für } x \leq 0 \\ f_r(x) = \sqrt{x} & \text{für } x > 0 \end{cases}$ ist stetig an der Nahtstelle $x_0 = 0$ und besitzt die Ableitungsfunktion $f_r'(x) = \frac{1}{2\sqrt{x}}$ (siehe Ableitungsregeln). f_r' ist an der Stelle x_0 nicht definiert, der zu berechnende Grenzwerte gleich ∞ ; f ist daher nicht diff.bar in x_0 .

Ableitungsregeln (auf den Beweis der Regeln wird verzichtet)

Funktion	1. Ableitung	Funktion	1. Ableitung	Funktion	1. Ableitung
a	0	$\sin(x)$	$\cos(x)$	$f \pm g$	$f' \pm g'$
x^a	$a \cdot x^{a-1}$	$\cos(x)$	$-\sin(x)$	$f \cdot g$	$f' \cdot g + f \cdot g'$
a^x	$a^x \cdot \ln(a)$	$\tan(x)$	$\frac{1}{\cos^2(x)}$	$\frac{f}{g}$	$\frac{f' \cdot g - f \cdot g'}{g^2}$
$\log_a(x)$	$\frac{1}{x \cdot \ln(a)}$	$\cot(x)$	$\frac{-1}{\sin^2(x)}$	$f(g)$	$f'(g) \cdot g'$

Beispiele

Funktion	1. Ableitung	Funktion	1. Ableitung
$x^0 = 1$	0	$\log_2(x)$	$1 / (x \cdot \ln(2))$
$x^1 = x$	1	$\log_e(x) = \ln(x)$	$1/x$, da $\ln(e) = 1$
x^2	$2x$	$x^2 - x + 3$	$2x - 1$
x^3	$3x^2$	$4x^3 - 5x^2 + 6x - 7$	$12x^2 - 10x + 6$
x^4	$4x^3$	$(x^2 - 1) \cdot (3x + 5)$	$2x \cdot (3x + 5) + (x^2 - 1) \cdot 3$
$x^{-1} = 1/x$	$-x^{-2} = -1/x^2$	$x^3 \cdot \ln(x)$	$3x^2 \cdot \ln(x) + x^3 \cdot 1/x$
$x^{-2} = 1/x^2$	$-2x^{-3} = -2/x^3$	$\frac{2x+1}{3x-2}$	$\frac{2 \cdot (3x-2) - (2x+1) \cdot 3}{(3x-2)^2}$
$x^{-3} = 1/x^3$	$-3x^{-4} = -3/x^4$	$\frac{e^x}{\sin(x)}$	$\frac{e^x \cdot \sin(x) - e^x \cdot \cos(x)}{\sin^2(x)}$
$x^{1/2} = \sqrt{x}$	$1/2 \cdot x^{-1/2} = 1/(2\sqrt{x})$	$(x^2 - 1)^3$	$3 \cdot (x^2 - 1)^2 \cdot 2x$
$x^{1/3} = \sqrt[3]{x}$	$1/3 \cdot x^{-2/3} = 1/(3\sqrt[3]{x^2})$	$\sqrt{x^4 + 1} = (x^4 + 1)^{1/2}$	$\frac{1}{2} \cdot (x^4 + 1)^{-1/2} \cdot 4x^3$
$x^{3/4} = \sqrt[4]{x^3}$	$3/4 \cdot x^{-1/4} = 3/(4\sqrt[4]{x})$	$\frac{1}{(x^3 - 1)^2} = (x^3 - 1)^{-2}$	$-2(x^3 - 1)^{-3} \cdot 3x^2$
$x^{-3/2} = 1/\sqrt{x^3}$	$-3/2 \cdot x^{-5/2} = -3/(2\sqrt{x^5})$		
2^x	$2^x \cdot \ln(2)$		
e^x	$e^x \cdot \ln(e) = e^x$		

Funktion	1. Ableitung	Funktion	1. Ableitung
$[f(x)]^a$	$a \cdot [f(x)]^{a-1} \cdot f'(x)$	$\ln(x^4 + 1)$	$\frac{1}{x^4 + 1} \cdot 4x^3$
e^{3x+1}	$e^{3x+1} \cdot 3$	$\log_{10}(3x^2 + 4)$	$\frac{1}{(3x^2 + 4) \cdot \ln(10)} \cdot 6x$
2^{x^3-1}	$2^{x^3-1} \cdot \ln(2) \cdot 3x^2$	$\log_a(f(x))$	$\frac{1}{f(x) \cdot \ln(a)} \cdot f'(x)$
$a^{f(x)}$	$a^{f(x)} \cdot \ln(a) \cdot f'(x)$	$\ln(f(x))$	$f'(x) / f(x)$

Produkt-, Quotienten- und Kettenregel können auch mehrfach hintereinander angewandt und miteinander kombiniert werden, z.B.

Funktion	1. Ableitung
$f(x) \cdot g(x) \cdot h(x)$	$f'(x) \cdot g(x) \cdot h(x) + f(x) \cdot g'(x) \cdot h(x) + f(x) \cdot g(x) \cdot h'(x)$
$\frac{x^3 \cdot \ln(x)}{x^2 - 1}$	$\frac{(3x^2 \cdot \ln(x) + x^3 \cdot \frac{1}{x}) \cdot (x^2 - 1) - x^3 \cdot \ln(x) \cdot 2x}{(x^2 - 1)^2}$
$\sin(\ln(x^2 + 1))$	$\cos(\ln(x^2 + 1)) \cdot \frac{1}{x^2 + 1} \cdot 2x$

Besteht eine Funktion f aus Produkten und/ oder Quotienten von Teilfunktionen, so bietet sich das logarithmische Ableiten an:

Logarithmisches Ableiten Leitet man eine logarithmierte Funktion ab, so erhält man (siehe oben) $[\ln(|f(x)|)]' = \frac{f'(x)}{f(x)}$,

oder anders geschrieben: $f'(x) = f(x) \cdot [\ln(|f(x)|)]'$.

Beispiel: Gesucht ist die 1. Ableitung der Funktion $f(x) = \frac{x^2 \cdot e^x}{(x-1)^4 \cdot \sqrt[5]{(x^3-1)^2}}$.

Lösung: Zunächst wird $\ln(|f(x)|)$ bestimmt. Unter Beachtung der Rechenregeln

$$\ln(x \cdot y) = \ln(x) + \ln(y), \quad \ln(x/y) = \ln(x) - \ln(y), \quad \ln(x^y) = y \cdot \ln(x), \quad \ln(e^x) = x$$

$$\text{erhält man } \ln(|f(x)|) = \ln(x^2 \cdot e^x) - \ln((x-1)^4 \cdot \sqrt[5]{(x^3-1)^2}) =$$

$$\ln(x^2) + \ln(e^x) - [\ln(x-1)^4 + \ln(\sqrt[5]{(x^3-1)^2})] = 2 \ln(|x|) + x - 4 \ln(|x-1|) - \frac{2}{5} \cdot \ln(|x^3-1|).$$

Die Ableitung von diesem Ausdruck ist $[\ln(|f(x)|)]' = \frac{2}{x} + 1 - \frac{4}{x-1} - \frac{2}{5} \cdot \frac{1}{x^3-1} \cdot 3x^2$ und daher

$$f'(x) = f(x) \cdot [\ln(|f(x)|)]' = \frac{x^2 \cdot e^x}{(x-1)^4 \cdot \sqrt[5]{(x^3-1)^2}} \cdot \left[\frac{2}{x} + 1 - \frac{4}{x-1} - \frac{6}{5} \cdot \frac{x^2}{x^3-1} \right].$$

Die Ableitung dieser Funktion mittels Produkt-, Quotienten- und Kettenregel gestaltet sich erheblich aufwendiger und unübersichtlicher.

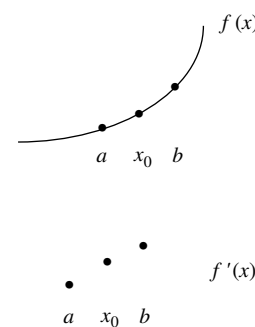
- Funktionsverlauf**
- Krümmung:** *konvex* = linksgekrümmt (z.B. das Symbol \cup);
konkav = rechtsgekrümmt (z.B. das Symbol \cap).
 - Wendestelle:** x_0 ist Wendestelle von f , wenn f die Krümmung an der Stelle x_0 wechselt (von konkav nach konvex oder umgekehrt).
 - Sattelstelle:** Wendestelle mit Steigung 0.



Regel Charakterisierung des Funktionsverlaufs mittels Ableitungen
Voraussetzung: f ist hinreichend oft diff.bar im betrachteten Intervall I.

- **Steigung** von f
 - (1) $f'(x) > 0 \quad \forall x \in I \Rightarrow f$ streng monoton steigend in I;
 - (2) $f'(x) < 0 \quad \forall x \in I \Rightarrow f$ streng monoton fallend in I.
- **Krümmung** von f
 - (1) $f''(x) > 0 \quad \forall x \in I \Rightarrow f$ streng konvex in I;
 - (2) $f''(x) < 0 \quad \forall x \in I \Rightarrow f$ streng konkav in I.
- **Wendestellen** von f
 - (1) x_0 Wendestelle von $f \Rightarrow f''(x_0) = 0$;
oder äquivalent: $f''(x_0) \neq 0 \Rightarrow x_0$ keine Wendestelle von f ;
 - (2) $f''(x_0) = 0$ und $f'''(x_0) \neq 0 \Rightarrow x_0$ Wendestelle von f .

Begründung des Zusammenhangs zwischen f'' und der Krümmung von f .
Betrachte eine Funktion f , die in der Umgebung einer Stelle x_0 konvex ist. Konvexität bedeutet, dass die Funktion von links nach rechts einen zunehmend steileren Verlauf aufweist. Links von x_0 , z.B. an der Stelle a , ist die Steigung geringer als rechts von x_0 , z.B. in b . Für die Funktionswerte $f'(x)$, die diese Steigung angeben, muss also gelten $f'(a) < f'(x_0) < f'(b)$. Die Funktion f' weist daher ebenfalls eine steigende Tendenz auf in der Umgebung von x_0 . Diese positive Steigung von f' an der Stelle x_0 wiederum wird gemessen durch $f''(x_0)$, d.h. $f''(x_0) > 0$.



Die Argumentationskette zur Begründung der oben angegebenen Regel verläuft jetzt natürlich genau andersherum:

- $f''(x_0) > 0 \Rightarrow f'$ besitzt an der Stelle x_0 eine positive Steigung, d.h. $f'(a) < f'(x_0) < f'(b)$.
- $\Rightarrow f$ verläuft links von x_0 flacher als rechts von x_0 , d.h. f ist konvex in einer Umgebung von x_0 .
- Analog: $f''(x_0) < 0 \Rightarrow f'$ besitzt an der Stelle x_0 eine negative Steigung, d.h. $f'(a) > f'(x_0) > f'(b)$.
- $\Rightarrow f$ verläuft links von x_0 steiler als rechts von x_0 , d.h. f ist konkav in einer Umgebung von x_0 .

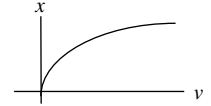
- Beispiele**
- (1) $f(x) = x^2, f'(x) = 2x, f''(x) = 2$, also
 - f streng monoton fallend $\forall x < 0$, da $f'(x) < 0 \quad \forall x < 0$;
 - f streng monoton steigend $\forall x > 0$, da $f'(x) > 0 \quad \forall x > 0$;
 - f streng konvex $\forall x \in \mathbb{R}$, da $f''(x) > 0 \quad \forall x \in \mathbb{R}$.
 - (2) $f(x) = -x^2, f'(x) = -2x, f''(x) = -2$, also
 - f streng monoton steigend $\forall x < 0$, da $f'(x) > 0 \quad \forall x < 0$;
 - f streng monoton fallend $\forall x > 0$, da $f'(x) < 0 \quad \forall x > 0$;
 - f streng konkav $\forall x \in \mathbb{R}$, da $f''(x) < 0 \quad \forall x \in \mathbb{R}$.
 - (3) $f(x) = x^3, f'(x) = 3x^2, f''(x) = 6x, f'''(x) = 6$, also
 - f streng monoton steigend $\forall x \in \mathbb{R}$, da $f'(x) > 0 \quad \forall x \in \mathbb{R}$;

f streng konkav $\forall x < 0$, da $f''(x) < 0 \forall x < 0$,
 f streng konvex $\forall x > 0$, da $f''(x) > 0 \forall x > 0$;
 $f'(x) = 0 \Leftrightarrow x_0 = 0, f''(x_0) = 0, f'''(x_0) = 6 \neq 0 \Rightarrow x_0$ Sattelstelle von f .

- (4) Sowohl in der Betriebs- als auch in der Volkswirtschaftstheorie verwendet man Ableitungen, um Funktionen mit bestimmten Eigenschaften zu konstruieren.

Beispiel: Konstruktion einer neoklassischen Produktionsfunktion $x(v)$, wobei x den Output, v den Input bezeichnet. x soll folgende Eigenschaften aufweisen:

- (a) $x(0) = 0$,
 (b) x positiv $\forall v > 0$, also $x(v) > 0 \forall v > 0$,
 (c) x monoton steigend $\forall v > 0$, also $x'(v) > 0 \forall v > 0$
 (d) x degressiv steigend (konkav) $\forall v > 0$, also $x''(v) < 0 \forall v > 0$.



Als möglicher Funktionstyp kommt eine Potenzfunktion der Form $x(v) = a \cdot v^b$ in Frage.

An die Koeffizienten a, b müssen dann die folgenden Bedingungen gestellt werden:

- (a) $x(0) = a \cdot 0^b = 0$ ist für jedes beliebige a, b erfüllt,
 (b) $x(v) = a \cdot v^b > 0 \forall v > 0$ gilt nur für $a > 0$, da $v^b > 0$ für beliebige b ,
 (c) $x'(v) = a \cdot b \cdot v^{b-1} > 0 \forall v > 0$ gilt nur für $b > 0$, da $a > 0$ und $v^{b-1} > 0$,
 (d) $x''(v) = a \cdot b \cdot (b-1) \cdot v^{b-2} < 0 \forall v > 0$ gilt nur für $b-1 < 0$, also $b < 1$, da $a \cdot b \cdot v^{b-2} > 0$.
 Insgesamt erfüllt daher die Funktion $x(v) = a \cdot v^b$ mit $a > 0$ und $0 < b < 1$ alle Bedingungen.

Die Kriterien für das Vorliegen von Extrema sind nun sofort einsichtig:

- (1) Ist die Steigung links von einer Stelle x_0 negativ und rechts von x_0 positiv $\searrow \swarrow$, so liegt in x_0 ein Minimum vor; umgekehrt liegt in x_0 ein Maximum vor, wenn die Steigung links von x_0 positiv und rechts von x_0 negativ ist $\swarrow \searrow$.
 (2) Ist die Steigung an einer Stelle x_0 gleich 0 (d.h. $f'(x_0) = 0$), so handelt es sich um ein Minimum, wenn die Funktion dort konvex verläuft \cup (d.h. $f''(x_0) > 0$), Maximum, wenn die Funktion dort konkav verläuft \cap (d.h. $f''(x_0) < 0$).

Regel

Lokale und globale Extrema

Voraussetzung: Die verwendeten Ableitungen an der Stelle x_0 existieren, d.h. die Funktionen sind entsprechend oft diff.bar.

- (1) *Lokale Extrema an Nahtstellen* x_0 , an denen die Funktion f nicht differenzierbar ist
 (a) $f'_l(x_0) < 0$ und $f'_r(x_0) > 0 \Rightarrow$ strenges lokales Minimum an der Stelle x_0 ;
 (b) $f'_l(x_0) > 0$ und $f'_r(x_0) < 0 \Rightarrow$ strenges lokales Maximum an der Stelle x_0 .⁸
 (2) *Lokale Extrema* an differenzierbaren Stellen x_0
 (a) $f'(x_0) = 0$ und $f''(x_0) > 0 \Rightarrow$ strenges lokales Minimum an der Stelle x_0 ;
 (b) $f'(x_0) = 0$ und $f''(x_0) < 0 \Rightarrow$ strenges lokales Maximum an der Stelle x_0 ;
 (c) $f'(x_0) = 0$ und $f''(x_0) = 0$ und $f'''(x_0) \neq 0 \Rightarrow$ Sattelstelle in x_0 ;
 (d) $f'(x_0) \neq 0 \Rightarrow$ kein lokales Extremum in x_0 .
 (3) *Globale Extrema* in einem Intervall $I, x_0 \in I$
 • Monotonie
 (a) $f'(x) < 0 \forall x < x_0$ und $f'(x) > 0 \forall x > x_0 \Rightarrow x_0$ strenges globales Minimum in I ;
 (b) $f'(x) > 0 \forall x < x_0$ und $f'(x) < 0 \forall x > x_0 \Rightarrow x_0$ strenges globales Maximum in I .⁹

⁸ Die Aussagen (a) und (b) sind falsch, wenn f in x_0 unstetig ist. Man sieht dies z.B. am Bildchen: $\swarrow \searrow$ Die Nahtstelle x_0 (Punkt) gehört zum rechten Funktionsast. Daher liegt in x_0 kein lokales Maximum vor (links von x_0 gibt es noch höhere Funktionswerte). In der ökonomischen Praxis würde man trotzdem von einem Maximum an der Stelle x_0 sprechen, da aufgrund der in der Praxis vorhandenen Meßgenauigkeit x_0 als zum "günstigeren" Funktionsast zugehörig angesehen werden darf (z.B. Briefporto beim Gewicht bis 20g / über 20g).

⁹ siehe vorherige Fußnote

- Krümmung
 - (a) $f'(x_0) = 0$ und $f''(x) > 0 \quad \forall x \in I \Rightarrow x_0$ strenges globales Minimum in I ;
 - (b) $f'(x_0) = 0$ und $f''(x) < 0 \quad \forall x \in I \Rightarrow x_0$ strenges globales Maximum in I .

Beispiel $f(x) = x^2$ bzw. $f(x) = -x^2$ besitzen in $x_0 = 0$ nicht nur ein lokales, sondern sogar ein globales Minimum bzw. Maximum $\forall x \in \mathbb{R}$ (siehe oben, Beispiel (1) bzw. (2)).

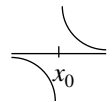
Schema zum Bestimmen globaler Extrema

(1) f auf Definitionslücken im relevanten Bereich hin untersuchen

Falls eine Definitionslücke x_0 im relevanten Bereich (der jeweiligen Teilfunktion) existiert:

$$\lim_{\substack{x \rightarrow x_0 \\ x < x_0}} f(x) \quad \text{und} \quad \lim_{\substack{x \rightarrow x_0 \\ x > x_0}} f(x) \quad \text{ermitteln und} \quad (x_0, \lim_{\substack{x \rightarrow x_0 \\ x < x_0}} f(x)), \quad (x_0, \lim_{\substack{x \rightarrow x_0 \\ x > x_0}} f(x))$$

als "Kandidaten" merken (siehe auch die Bemerkung unten).



(2) Randpunkte des relevanten Bereiches untersuchen

Fall (a): Relevanter Bereich = Intervall $[x_L; x_R]$

$(x_L, f(x_L)), (x_R, f(x_R))$ als "Kandidaten" merken.

Fall (b): Relevanter Bereich = \mathbb{R} ("Randpunkte": $-\infty, +\infty$)

$$\lim_{x \rightarrow -\infty} f(x) \quad \text{und} \quad \lim_{x \rightarrow \infty} f(x) \quad \text{berechnen und} \quad (-\infty, \lim_{x \rightarrow -\infty} f(x)), \quad (+\infty, \lim_{x \rightarrow \infty} f(x))$$

als "Kandidaten" merken.

Fall (c): Relevanter Bereich = $[x_L; \infty)$ oder $(-\infty; x_R]$: Analog zu (a) bzw. (b).

(3) Bei stückweise definierten Funktionen: "Nahtstellen" untersuchen

Fall (a) $f_l(x_0) = f_r(x_0)$ (also f stetig an der Nahtstelle x_0 mit $f(x_0) = f_l(x_0) = f_r(x_0)$):

$(x_0, f(x_0))$ als "Kandidat" merken.

Fall (b) $f_l(x_0) \neq f_r(x_0)$ (also f nicht stetig an der Nahtstelle x_0):

$(x_0, f_l(x_0)), (x_0, f_r(x_0))$ als "Kandidaten" merken.

Dabei ist $\lim_{\substack{x \rightarrow x_0 \\ x < x_0}} f_l(x)$ statt $f_l(x_0)$ bzw. $\lim_{\substack{x \rightarrow x_0 \\ x > x_0}} f_r(x)$ statt $f_r(x_0)$ zu verwenden, wenn $f_l(x_0)$

bzw. $f_r(x_0)$ mathematisch nicht erklärt ist.

(4) Stellen mit Steigung 0 im relevanten Bereich bestimmen

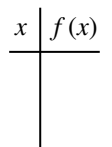
f' ermitteln und gleich 0 setzen (bei stückweise definierten Funktionen für jede Teilfunktion);

wenn die gefundenen Nullstellen x_i der 1. Ableitung im relevanten Bereich (der jeweiligen Teilfunktion) liegen: $(x_i, f(x_i))$ als "Kandidaten" merken.

(5) Zusammenfassung: Auflisten aller "Kandidaten" aus (1) – (4)

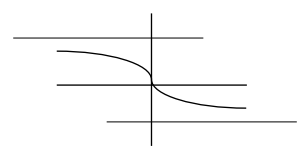
Sofern existent, erhält man das globale

- Maximum als Punkt mit dem größten $f(x)$ -Wert aller "Kandidaten",
- Minimum als Punkt mit dem kleinsten $f(x)$ -Wert aller "Kandidaten".



Bemerkung:

- Ist der kleinste $f(x)$ -Wert gleich $-\infty$, so existiert kein globales Minimum; ist der größte $f(x)$ -Wert gleich ∞ , so existiert kein globales Maximum.
- Wird der kleinste $f(x)$ -Wert für $x \rightarrow -\infty$ bzw. $x \rightarrow \infty$ (also nie) erreicht, so existiert kein globales Minimum; wird der größte $f(x)$ -Wert für $x \rightarrow -\infty$ bzw. $x \rightarrow \infty$ (also nie) erreicht, so existiert kein globales Maximum.
- Da Nahtstellen auch als Randpunkte des relevanten Bereiches von zwei Teilfunktionen zu sehen sind, könnte man (2) und (3) zu einem Unterpunkt zusammenfassen.



Beispiele zur Anwendung dieses Schemas finden sich in der Aufgabensammlung.

11. Differenzial, Wachstumsrate, Elastizität

Ziel Näherungsweise Berechnen der Änderung eines Funktionswertes $f(x_0)$ bei Variation von x_0 .

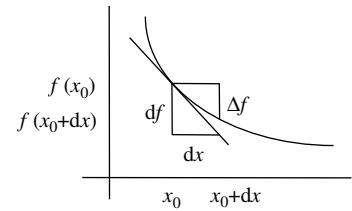
Anwendung Die Preiselastizität der Nachfrage beispielsweise gibt näherungsweise an, um wie viel % sich die Nachfrage ändert bei Variation des aktuellen Preises p_0 um 1%.

Idee Der Graf einer Funktion f lässt sich in einer (kleinen) Umgebung eines Punktes $(x_0, f(x_0))$ relativ gut durch die Tangente an die Kurve annähern.

Differenzial Variiert man x_0 um dx Einheiten, so ändert sich $f(x_0)$ um

$$\Delta f = f(x_0+dx) - f(x_0) \quad \text{Einheiten.}$$

$T(x)$ bezeichne die Tangente an den Funktionsgraphen durch den Punkt $(x_0, f(x_0))$, also $T(x_0) = f(x_0)$ und $f'(x_0) = \frac{df}{dx}$.



Ändert man dann wieder x_0 um dx Einheiten, so ändert sich der Funktionswert auf der Tangenten um

$$df = T(x_0+dx) - f(x_0) = f'(x_0) \cdot dx \quad \text{Einheiten.}$$

Für betragsmäßig kleine Variationen dx stimmen Δf und df gut überein.

df und dx bezeichnet man als *Differenziale*, $f' = \frac{df}{dx}$ daher auch als *Differenzialquotient*.

Beispiel: Für $f(x) = x^2$, $x_0 = 3$, $dx = 0,2$ ist $f'(x) = 2x$ und

$$\Delta f = f(x_0+dx) - f(x_0) = f(3,2) - f(3) = 10,24 - 9 = 1,24 \quad (\text{exakt})$$

$$\approx df = f'(x_0) \cdot dx = f'(3) \cdot 0,2 = 6 \cdot 0,2 = 1,2 \quad (\text{angenähert}).$$

Wachstumsrate Variiert man x_0 um dx Einheiten, so beträgt die *relative* Änderung von $f(x_0)$: $\frac{\Delta f}{f(x_0)}$.

Dies entspricht für betragsmäßig kleine dx näherungsweise der relativen Änderung auf der Tangenten

$$\frac{df}{f(x_0)} = \frac{f'(x_0)}{f(x_0)} \cdot dx = r_f(x_0) \cdot dx.$$

Die Funktion $r_f = f' / f$ bezeichnet man als *Wachstumsrate* von f .

Beispiel: Für $f(x) = x^2$, $x_0 = 3$, $dx = 0,2$ ist $f'(x) = 2x$, $r_f(x) = \frac{f'(x)}{f(x)} = \frac{2x}{x^2} = \frac{2}{x}$ und

$$\frac{\Delta f}{f(x_0)} = \frac{1,24}{9} = 0,1378 = 0,1378 \cdot 100 \% = 13,78 \% \quad (\text{exakt})$$

$$\approx \frac{df}{f(x_0)} = r_f(x_0) \cdot dx = \frac{2}{3} \cdot 0,2 = 0,1333 = 13,33 \% \quad (\text{angenähert}).$$

Beispiel: Die Wachstumsrate der Funktion $f(x) = e^x$ ist $r_f(x) = e^x / e^x = 1$.

Die e -Funktion drückt also konstantes Wachstum aus, wie bereits im Zusammenhang mit der stetigen Verzinsung dargelegt.

Elastizität Variiert man x_0 um s Prozent, so ändert sich $f(x_0)$ *relativ* um $\frac{f(x_0+x_0 \cdot s\%) - f(x_0)}{f(x_0)}$.

Die Variation von x_0 um s % entspricht einer Änderung von x_0 um $dx = x_0 \cdot s$ Einheiten.

Für betragsmäßig kleine Variationen von s % ergibt sich wie oben die Annäherung

$$\frac{f'(x_0)}{f(x_0)} \cdot x_0 \cdot s \% = r_f(x_0) \cdot x_0 \cdot s \% = \varepsilon_f(x_0) \cdot s \%.$$

Die Funktion ε_f mit $\varepsilon_f(x) = x \cdot \frac{f'(x)}{f(x)}$ bezeichnet man als *Elastizität* von f .

Speziell für $s = 1$ erhält man: Die Elastizität an einer Stelle x_0 gibt näherungsweise an, um wie viel % sich $f(x_0)$ ändert, wenn x_0 um 1% variiert.

Beispiel: Für $f(x) = x^2$, $x_0 = 3$, $s = 5$ ist $f'(x) = 2x$, $\varepsilon_f(x) = x \cdot \frac{f'(x)}{f(x)} = x \cdot \frac{2x}{x^2} = 2$ und

$$\frac{f(x_0+x_0 \cdot s\%) - f(x_0)}{f(x_0)} = \frac{f(3+3 \cdot 5\%) - f(3)}{f(3)} = \frac{f(3,15) - 9}{9} = 0,1025 = 10,25 \% \quad (\text{exakt})$$

$$\approx \varepsilon_f(x_0) \cdot s \% = 2 \cdot 5 \% = 10 \% \quad (\text{angenähert}).$$

Die Bedeutung der Elastizität als Maß für Empfindlichkeit einer Funktion auf Variation des Arguments beruht nicht zuletzt auch auf der einfachen Berechenbarkeit von $\varepsilon_f(x)$. Wegen

$[\ln(|f(x)|)]' = \frac{f'(x)}{f(x)}$ erhält man $\varepsilon_f(x) = x \cdot [\ln(|f(x)|)]'$. Die Berechnung der Elastizität mittels

dieser Formel bietet sich an, wenn f ein Produkt bzw. Quotient von Teilfunktionen ist. Außerdem erhält man hieraus sofort die Produkt- bzw. Quotientenregel für Elastizitäten in folgender Tabelle.

Rechenregeln für Elastizitäten (ergeben sich sofort aus den Ableitungsregeln)

Funktion	Elastizität	Funktion	Elastizität	Funktion	Elastizität
a	0	$f \cdot g$	$\epsilon_f + \epsilon_g$	$a \cdot f(x)$	$\epsilon_f(x)$
x^a	a	f / g	$\epsilon_f - \epsilon_g$	$[f(x)]^a$	$a \cdot \epsilon_f(x)$
e^x	x	$f(g)$	$\epsilon_f(g) \cdot \epsilon_g$	$e^{f(x)}$	$x \cdot f'(x)$
$\ln(x)$	$\frac{1}{\ln(x)}$			$\ln(f(x))$	$\frac{\epsilon_f(x)}{\ln(f(x))}$

Beispiel: Die Elastizität der Funktion $f(x) = \frac{e^{-x}}{\sqrt{x}} = \frac{e^{-x}}{x^{1/2}}$ ist gleich der Elastizität des Zählers ($= -x$) minus der Elastizität des Nenners ($= \frac{1}{2}$), also $\epsilon_f(x) = -x - \frac{1}{2}$.

Fehlerrechnung Beispiel: Ein runder Deckel mit dem Radius $x = 5$ cm soll hergestellt werden. Die Fläche des Deckels, sprich der Materialverbrauch, beträgt $M(x) = x^2 \cdot \pi$.

- (1) Um wie viel cm^2 kann der Materialverbrauch pro Deckel ca. schwanken, wenn der Radius um ± 1 mm variiert?
- (2) Um wie viel % kann der Materialverbrauch pro Deckel ca. schwanken, wenn der Radius um ± 1 mm variiert?
- (3) Um wie viel % kann der Materialverbrauch pro Deckel ca. schwanken, wenn der Radius um $\pm 1\%$ variiert?

Lösung: Die Fragen können unter Verwendung von Differenzial, Wachstumsrate und Elastizität beantwortet werden.

- (1) Mit $x_0 = 5$ cm, $dx = 0,1$ cm und $M'(x) = 2x \cdot \pi$ ergibt sich eine Schwankungsbreite von ca. $M'(x_0) \cdot dx = 10 \cdot \pi \cdot 0,1 = \pi \text{ cm}^2$.
- (2) Wegen $r_M(x) = \frac{2x \cdot \pi}{x^2 \cdot \pi} = \frac{2}{x}$ ergibt sich eine Schwankungsbreite von ca. $r_M(x_0) \cdot dx = \frac{2}{5} \cdot 0,1 = 0,04 = 4 \%$.
- (3) Mit $s = 1$ und $\epsilon_M(x) = 2$ ergibt sich eine Schwankungsbreite von ca. $\epsilon_M(x_0) \cdot s \% = 2 \cdot 1 \% = 2 \%$.

12. Taylorentwicklung einer Funktion

Ziel Approximation einer Funktion f in einer Umgebung eines Punktes $(x_0, f(x_0))$ durch ein Polynom T_n n -ten Grades.

Anwendung Im letzten Abschnitt wurde bereits das Taylorpolynom 1. Grades, nämlich die Tangente T_1 , als Annäherung an eine Funktion verwendet. Mit Polynomen höherer Ordnung, also quadratischen, kubischen etc. Funktionen wird die Güte der Annäherung weiter verbessert. Viele Funktionen sind als Polynom der Ordnung ∞ darstellbar (in eine Taylorreihe zu entwickeln).

Taylorpolynom Ist eine Funktion f n -fach diff.bar an der Stelle x_0 , dann nennt man die Funktion T_n mit

$$T_n(x) = f(x_0) + f'(x_0) \cdot (x - x_0) + \frac{1}{2} \cdot f''(x_0) \cdot (x - x_0)^2 + \frac{1}{3!} \cdot f'''(x_0) \cdot (x - x_0)^3 + \dots + \frac{1}{n!} \cdot f^{(n)}(x_0) \cdot (x - x_0)^n$$

$$= \sum_{k=0}^n \frac{1}{k!} \cdot f^{(k)}(x_0) \cdot (x - x_0)^k, \quad f^{(0)}(x_0) = f(x_0), \quad n! = 1 \cdot 2 \cdot 3 \cdot \dots \cdot n, \quad 0! = 1,$$

das Taylorpolynom n -ten Grades von f , entwickelt an der Stelle x_0 .

Also ist $T_1(x) = f(x_0) + f'(x_0) \cdot (x - x_0)$ eine lineare Funktion (Gerade),

$T_2(x) = T_1(x) + \frac{1}{2} \cdot f''(x_0) \cdot (x - x_0)^2$ eine quadratische Funktion (Parabel),

$T_3(x) = T_2(x) + \frac{1}{3!} \cdot f'''(x_0) \cdot (x - x_0)^3$ eine kubische Funktion etc..

Bemerkung: • Je höher der Grad des Polynoms, umso besser stimmen $T_n(x)$ und $f(x)$ überein.
• Je näher ein Wert x an x_0 liegt, umso besser stimmen $T_n(x)$ und $f(x)$ überein.

Eigenschaften der Funktionen T_n : Taylorpolynome sind so konstruiert, dass sie folgende Eigenschaften besitzen:
 $T_n(x_0) = f(x_0)$: Alle Taylorpolynome gehen durch den Punkt $(x_0, f(x_0))$;
 $T_n'(x_0) = f'(x_0)$: Sie besitzen dort die gleiche Steigung wie f ;
 $T_n''(x_0) = f''(x_0)$: Ab dem Taylorpolynom 2. Grades besitzen sie dort die gleiche Krümmung wie f ;
 $T_n'''(x_0) = f'''(x_0); \dots; T_n^{(n)}(x_0) = f^{(n)}(x_0)$.
 An der Stelle x_0 stimmen also die Ableitungen von T_n und f bis zum Grade n überein.
 Durch Ableiten von T_n und Einsetzen von x_0 für x kann man diese Eigenschaften einfach nachweisen.

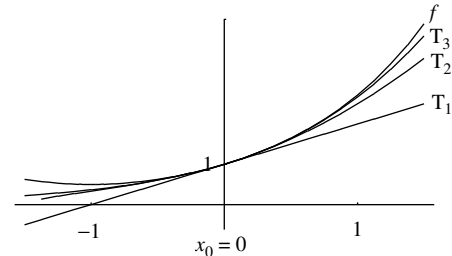
Vorgehen beim Berechnen von T_n

- Ableitungen $f'(x), f''(x), \dots, f^{(n)}(x)$ ermitteln;
- Funktionswerte an der Stelle x_0 bestimmen: $f(x_0), f'(x_0), f''(x_0), \dots, f^{(n)}(x_0)$;
- x_0 und Funktionswerte in die Formel einsetzen.

Beachten: T_n ist ein Polynom und nicht zu verwechseln mit dem Funktionswert $T_n(x_0) = f(x_0)$.

Beispiel Gesucht ist das Taylorpolynom n -ten Grades der Funktion $f(x) = e^x$, entwickelt an der Stelle $x_0 = 0$.
 Lösung: Die Ableitung von f ergibt wiederum $f'(x) = e^x$ und damit auch allgemein $f^{(n)}(x) = e^x$.
 An der Stelle x_0 erhält man so die Funktionswerte $f(x_0) = e^0 = 1 = f'(x_0) = \dots = f^{(n)}(x_0)$.
 Eingesetzt in die Formel: $T_n(x) = 1 + 1 \cdot (x - 0) + \frac{1}{2} \cdot 1 \cdot (x - 0)^2 + \frac{1}{3!} \cdot 1 \cdot (x - 0)^3 + \dots + \frac{1}{n!} \cdot 1 \cdot (x - 0)^n$,
 also $T_n(x) = 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 + \dots + \frac{1}{n!}x^n$. Speziell ist $T_1(x) = 1 + x$, $T_2(x) = 1 + x + \frac{1}{2}x^2$, etc..

x	-1	-0,5	$x_0 = 0$	0,5	1
$f(x)$	0,37	0,607	1	1,649	2,72
$T_1(x)$	0	0,5	1	1,5	2
$T_2(x)$	0,5	0,625	1	1,625	2,5
$T_3(x)$	0,38	0,604	1	1,646	2,67



Restgliedabschätzung: Den Fehler $R_n(x) = f(x) - T_n(x)$, der bei Annäherung von $f(x)$ durch $T_n(x)$ entsteht, nennt man das *Restglied* der Taylorentwicklung. Für die Funktion R_n gilt die auf Lagrange zurückgehende Aussage: Zu jedem x gibt es eine Zahl z zwischen x und x_0 , so dass $R_n(x) = \frac{1}{(n+1)!} \cdot f^{(n+1)}(z) \cdot (x - x_0)^{n+1}$.
 Mit Hilfe dieser Darstellung gewinnt man eine Abschätzung über den maximalen Fehler bei der Approximation von f durch T_n auf einem Intervall.
 Ist die Funktion $f \infty$ oft diff.bar in x_0 und geht der Fehler $R_n(x)$ gegen 0 (für $n \rightarrow \infty$) für alle x aus einem Intervall I , dann lässt sich f an der Stelle x_0 in eine Taylorreihe entwickeln und es gilt $\forall x \in I$

$$f(x) = \sum_{k=0}^{\infty} \frac{1}{k!} \cdot f^{(k)}(x_0) \cdot (x - x_0)^k.$$

Beispiel Wie oben sei die Funktion $f(x) = e^x$ in der Umgebung von $x_0 = 0$ durch das Taylorpolynom $T_n(x) = 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 + \dots + \frac{1}{n!}x^n$ angenähert.

Frage: Wie groß kann der Fehler $R_n(x)$ maximal für x -Werte aus dem Intervall $[-1; 1]$ werden?
 Antwort: Es ist $R_n(x) = \frac{1}{(n+1)!} f^{(n+1)}(z) \cdot (x - x_0)^{n+1} = \frac{1}{(n+1)!} \cdot e^z \cdot x^{n+1}$ für ein z zwischen 0 und x .
 Für x -Werte aus dem Intervall $[-1; 1]$ liegt dann auch z in diesem Intervall. Daher ist stets $|e^z| \leq e$.
 Außerdem ist $|x^{n+1}| \leq 1 \quad \forall x \in [-1; 1]$. Somit erhält man insgesamt: $|R_n(x)| \leq \frac{e}{(n+1)!} \quad \forall x \in [-1; 1]$.
 Beispielsweise für das Taylorpolynom 5. Grades gilt dann: $|e^x - T_5(x)| \leq \frac{e}{6!} = 0,0038 \quad \forall x \in [-1; 1]$.

Da f unendlich oft diff.bar ist in x_0 und $R_n(x) \xrightarrow{n \rightarrow \infty} 0 \quad \forall x$ aus einem beliebigen Intervall I , lässt sich f an der Stelle x_0 für alle $x \in \mathbb{R}$ in eine Taylorreihe entwickeln, d.h. es gilt $e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} \quad \forall x \in \mathbb{R}$.

13. Unbestimmte Ausdrücke: Die Regeln von l'Hospital

Ziel Ermitteln des Verhaltens einer Funktion $\frac{f}{g}$

- an einer Stelle x_0 mit $f(x_0) = 0 = g(x_0)$ bzw. $\lim_{x \rightarrow x_0} f(x) = \pm\infty$ und $\lim_{x \rightarrow x_0} g(x) = \pm\infty$,
- für $x \rightarrow \pm\infty$ mit $\lim_{x \rightarrow \pm\infty} f(x) = 0 = \lim_{x \rightarrow \pm\infty} g(x)$ bzw. $\lim_{x \rightarrow \pm\infty} f(x) = \pm\infty$ und $\lim_{x \rightarrow \pm\infty} g(x) = \pm\infty$.

Ausdrücke der Form " $\frac{0}{0}$ " bzw. " $\frac{\pm\infty}{\pm\infty}$ " heißen *unbestimmt* ($\pm\infty$ steht für $+\infty$ oder $-\infty$).

Regeln von l'Hospital: An einer Stelle x_0 sei $f(x_0) = 0 = g(x_0)$.
Existiert dann $\lim_{x \rightarrow x_0} \frac{f'(x)}{g'(x)}$, so ist $\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = \lim_{x \rightarrow x_0} \frac{f'(x)}{g'(x)}$.

Die Regel gilt analog auch für die übrigen genannten Fälle.

Herleitung (Idee) Für x -Werte in der Nähe von x_0 gilt nach der Taylorformel die Annäherung $f(x) \approx T_1(x) = f(x_0) + f'(x_0) \cdot (x - x_0) = f'(x_0) \cdot (x - x_0)$, da $f(x_0) = 0$ nach Voraussetzung.
Analog ist $g(x) \approx g'(x_0) \cdot (x - x_0)$ und damit $\frac{f(x)}{g(x)} \approx \frac{f'(x_0) \cdot (x - x_0)}{g'(x_0) \cdot (x - x_0)} = \frac{f'(x_0)}{g'(x_0)}$.

Beispiele

- (1) $\lim_{x \rightarrow 1} \frac{x^2 - 1}{x - 1} = \lim_{x \rightarrow 1} \frac{2x}{1} = 2$ (Fall " $\frac{0}{0}$ ")
- (2) $\lim_{x \rightarrow 0} \frac{e^x - 1}{x} = \lim_{x \rightarrow 0} \frac{e^x}{1} = 1$ (Fall " $\frac{0}{0}$ ")
- (3) $\lim_{x \rightarrow 0} \frac{\ln(x)}{1/x} = \lim_{x \rightarrow 0} \frac{1/x}{-1/x^2} = \lim_{x \rightarrow 0} -x = 0$ (Fall " $\frac{-\infty}{\infty}$ ")
- (4) $\lim_{x \rightarrow 1} \frac{x^3 - 3x + 2}{x^2 - 2x + 1} = \lim_{x \rightarrow 1} \frac{3x^2 - 3}{2x - 2} = \lim_{x \rightarrow 1} \frac{6x}{2} = 3$ (zwei mal der Fall " $\frac{0}{0}$ ")
- (5) $\lim_{x \rightarrow \infty} \frac{x + 1}{x^2 + 1} = \lim_{x \rightarrow \infty} \frac{1}{2x} = 0$ (Fall " $\frac{\infty}{\infty}$ ")

Rückführung der Fälle (1) " $0 \cdot \pm\infty$ ", (2) " $1^{\pm\infty}$ ", (3) " 0^0 ", (4) " $\pm\infty^0$ " auf die Fälle " $\frac{0}{0}$ " bzw. " $\frac{\pm\infty}{\pm\infty}$ ".

(1) Sei $f(x_0) = 0$ und $\lim_{x \rightarrow x_0} g(x) = \pm\infty$, also $\lim_{x \rightarrow x_0} f(x) \cdot g(x)$ ein Ausdruck der Form " $0 \cdot \pm\infty$ ".

Wegen $f(x) \cdot g(x) = \frac{f(x)}{\frac{1}{g(x)}} = \frac{g(x)}{\frac{1}{f(x)}}$ und $\lim_{x \rightarrow x_0} \frac{1}{g(x)} = 0$, $\lim_{x \rightarrow x_0} \frac{1}{f(x)} = \pm\infty$, lässt sich der

Ausdruck " $0 \cdot \pm\infty$ " zurückführen auf den Ausdruck " $\frac{0}{0}$ " bzw. " $\frac{\pm\infty}{\pm\infty}$ ".

(2) - (4): Bei diesen Fällen verwendet man die Transformation $f(x)^{g(x)} = e^{g(x) \cdot \ln(f(x))}$. Der Grenzwert des Exponenten $g(x) \cdot \ln(f(x))$ kann unter Verwendung von (1) bestimmt werden.

Beispiele

- (1) $\lim_{x \rightarrow 0} x \cdot \ln(x) = \lim_{x \rightarrow 0} \frac{\ln(x)}{1/x} = 0$ (der Fall " $0 \cdot -\infty$ " wird transformiert zu " $\frac{-\infty}{\infty}$ "; s.o. Bsp. (3)).
- (2) $\lim_{x \rightarrow \infty} \left(1 + \frac{a}{x}\right)^x = \lim_{x \rightarrow \infty} e^{x \cdot \ln(1 + a/x)}$ und $\lim_{x \rightarrow \infty} x \cdot \ln(1 + \frac{a}{x}) = \lim_{x \rightarrow \infty} \frac{\ln(1 + a/x)}{1/x} = \lim_{x \rightarrow \infty} \frac{\frac{1}{1+a/x} \cdot \frac{-a}{x^2}}{-1/x^2}$
 $= \lim_{x \rightarrow \infty} \frac{a}{1 + a/x} = a$, also $\lim_{x \rightarrow \infty} \left(1 + \frac{a}{x}\right)^x = e^a$ (der Fall " 1^∞ " wird transformiert zu " $e^{\infty \cdot 0}$ " und der Exponent " $\infty \cdot 0$ " dann zu " $\frac{0}{0}$ ", so dass l'Hospital angewandt werden kann).
- (3) $\lim_{x \rightarrow 0} x^x = \lim_{x \rightarrow 0} e^{x \cdot \ln(x)}$ und $\lim_{x \rightarrow 0} x \cdot \ln(x) = 0$ (siehe (1)), also $\lim_{x \rightarrow 0} x^x = e^0 = 1$
 (der Fall " 0^0 " wird transformiert zu " $e^{0 \cdot -\infty}$ " und der Exponent " $0 \cdot -\infty$ " dann zu " $\frac{-\infty}{\infty}$ ").
- (4) $\lim_{x \rightarrow \infty} \sqrt[x]{x} = \lim_{x \rightarrow \infty} x^{1/x} = \lim_{x \rightarrow \infty} e^{1/x \cdot \ln(x)}$ und $\lim_{x \rightarrow \infty} \frac{\ln(x)}{x} = \lim_{x \rightarrow \infty} \frac{1/x}{1} = 0$, also $\lim_{x \rightarrow \infty} \sqrt[x]{x} = e^0 = 1$
 (der Fall " ∞^0 " wird transformiert zu " $e^{0 \cdot \infty}$ " und der Exponent " $0 \cdot \infty$ " dann zu " $\frac{\infty}{\infty}$ ").

Grenzwerte von Folgen Im Abschnitt "Folgen" wurde bereits angedeutet, dass sich die Grenzwerte vieler Folgen bestimmen lassen über die Grenzwerte der entsprechenden Funktionen. Aus den vorangegangenen Beispielen (1) - (4) ergibt sich sofort:

$$(1) \lim_{n \rightarrow \infty} 1/n \cdot \ln(1/n) = 0 \quad (\text{mit } x = 1/n); \quad (2) \lim_{n \rightarrow \infty} (1 + a/n)^n = e^a \quad (\text{mit } x = n);$$

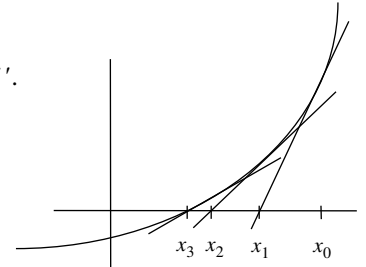
$$(3) \lim_{n \rightarrow \infty} (1/n)^{1/n} = \lim_{n \rightarrow \infty} \sqrt[n]{1/n} = 1 \quad (\text{mit } x = 1/n); \quad (4) \lim_{n \rightarrow \infty} \sqrt[n]{n} = 1 \quad (\text{mit } x = n).$$

14. Newton-Verfahren

Ziel Numerisches Ermitteln einer Nullstelle von f unter Verwendung von f' .

Prinzip Berechne iterativ (nacheinander) neue Annäherungen an die Nullstelle, wobei

- der Startwert x_0 "fast" beliebig gewählt werden kann (vgl. Bemerkung unten),
- die erste Annäherung x_1 errechnet wird als Schnittpunkt der Tangente an die Kurve durch $(x_0, f(x_0))$ mit der x -Achse,
- die zweite Annäherung x_2 bestimmt wird als Schnittpunkt der Tangente an die Kurve durch $(x_1, f(x_1))$ mit der x -Achse,
- weitere Versuchspunkte $(x_2, f(x_2)), (x_3, f(x_3)), \dots, (x_n, f(x_n))$ analog ermittelt werden,
- das Verfahren solange fortgesetzt wird, bis mit einer vorgegebenen Genauigkeit gilt: $|f(x_n)| \approx 0$ und $x_n \approx x_{n-1}$.



Herleitung der Iterationsformel: Für die Tangente an die Kurve durch den Punkt $(x_0, f(x_0))$ erhält man aus der Taylorformel die Darstellung $T_1(x) = f(x_0) + f'(x_0) \cdot (x - x_0)$. Den Schnittpunkt dieser Tangente mit der x -Achse, also die Nullstelle von $T_1(x)$, berechnet man durch

$$f(x_0) + f'(x_0) \cdot (x - x_0) = 0 \Leftrightarrow x - x_0 = -\frac{f(x_0)}{f'(x_0)} \Leftrightarrow x = x_0 - \frac{f(x_0)}{f'(x_0)}, \text{ falls } f'(x_0) \neq 0.$$

Die erste Annäherung ist also $x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$. Analog ergeben sich die weiteren Approximationen

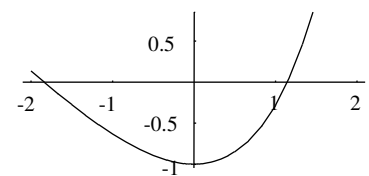
$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}, \quad x_3 = x_2 - \frac{f(x_2)}{f'(x_2)}, \quad \text{und damit die allgemeine Iterationsformel.}$$

Allgemeine Iterationsformel: $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$.

Beispiel Die Nullstellen der Funktion $f(x) = e^x - x - 2$ können analytisch nicht ermittelt werden.

Wie die Wertetabelle zeigt, muss zwischen -2 und -1 und zwischen 1 und 2 eine Nullstelle liegen.

x	-2	-1	0	1	2
$f(x)$	$0,14$	$-0,63$	-1	$-0,28$	$3,39$



Mit $f'(x) = e^x - 1$ und dem Startwert $x_0 = 1$ folgt

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} = 1 - \frac{e^1 - 1 - 2}{e^1 - 1} = 1 - \frac{-0,282}{1,718} = 1,164$$

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)} = 1,164 - \frac{e^{1,164} - 1,164 - 2}{e^{1,164} - 1} = 1,1464$$

$$x_3 = x_2 - \frac{f(x_2)}{f'(x_2)} = 1,1464 - \frac{e^{1,1464} - 1,1464 - 2}{e^{1,1464} - 1} = 1,146193259$$

$$x_4 = x_3 - \frac{f(x_3)}{f'(x_3)} = 1,146193221 = x_5.$$

Nach 4 Iterationen ist die erste Nullstelle bis auf 9 Nachkommastellen (Genauigkeit des Taschenrechners) bestimmt. Analog erhält man mit dem Startwert $x_0 = -2$ die zweite Nullstelle durch $x_1 = -1,843$, $x_2 = -1,841406066$, $x_3 = -1,841405660 = x_4$.

Probleme können auftreten, wenn während der Iteration eine Stelle x_n erreicht wird mit $f'(x_n) = 0$ bzw. $f'(x_n) \approx 0$ (vgl. oben).
 Beispiel: Startet man bei der obigen Funktion mit $x_0 = 0$, so ergibt sich $f'(x_0) = 0$, d.h. x_1 kann nicht bestimmt werden. Beginnt man statt dessen mit $x_0 = 0,01$, so erhält man $f'(x_0) = 0,001$ und daraus $x_1 = 99,5$, $x_2 = 98,5$, ..., d.h. das Newton-Verfahren konvergiert sehr langsam.

15. Lokale Extrema für Funktionen mit mehreren Variablen

Grafische Darstellung von Funktionen mit mehreren Variablen (siehe Übersicht)

Funktionen f mit zwei Variablen lassen sich grafisch als dreidimensionales "Gebirge" darstellen, indem jedem Punkt der x - y -Ebene der entsprechende Funktionswert $f(x, y)$ als "Höhe" zugeordnet wird.

Bei Funktionen mit mehr als zwei Variablen (also $f: \mathbb{R}^m \rightarrow \mathbb{R}$ mit $f(\mathbf{x}) \in \mathbb{R}$, $\mathbf{x} \in \mathbb{R}^m$, $m > 2$) versagt zwar die Möglichkeit der einfachen grafischen Darstellung, jedoch können alle Begriffe und Konzepte problemlos dorthin übertragen werden.

Lokale Extrema

$\mathbf{x}_0 \in \mathbb{R}^m$ ist ein (strenges)

lokales Minimum, wenn es ein $\varepsilon > 0$ gibt, so dass $f(\mathbf{x}_0) \leq (<) f(\mathbf{x}) \quad \forall \mathbf{x} \in U_\varepsilon(\mathbf{x}_0)$;

lokales Maximum, wenn es ein $\varepsilon > 0$ gibt, so dass $f(\mathbf{x}_0) \geq (>) f(\mathbf{x}) \quad \forall \mathbf{x} \in U_\varepsilon(\mathbf{x}_0)$;

dabei ist $U_\varepsilon(\mathbf{x}_0) = \{ \mathbf{x} \in \mathbb{R}^m \mid |\mathbf{x} - \mathbf{x}_0| < \varepsilon \}$ eine ε -Umgebung von $\mathbf{x}_0 \in \mathbb{R}^m$.

Im Gegensatz zu den Funktionen einer Variablen wird hier auf das Bestimmen globaler Extrema verzichtet, da die Vorgehensweise im mehrdimensionalen Fall erheblich aufwendiger ist. Das Ermitteln lokaler Extrema verläuft bei Funktionen mehrerer Variablen parallel zum einfachen Fall.

Konvergenz

einer Folge $(\mathbf{x}_n)_{n \in \mathbb{N}} \in \mathbb{R}^m$ gegen $\mathbf{x}_0 \in \mathbb{R}^m$: $\mathbf{x}_n = \begin{pmatrix} x_{1(n)} \\ \vdots \\ x_{m(n)} \end{pmatrix}$ konvergiert gegen $\mathbf{x}_0 = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}$ für $n \rightarrow \infty$,

wenn alle Komponenten einzeln konvergieren, also $\lim_{n \rightarrow \infty} x_{1(n)} = x_1, \dots, \lim_{n \rightarrow \infty} x_{m(n)} = x_m$;

z.B. $\begin{pmatrix} 1+1/n \\ 2-1/n \end{pmatrix} \rightarrow \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ für $n \rightarrow \infty$.

Stetigkeit

einer Funktion $f: \mathbb{R}^m \rightarrow \mathbb{R}$ an der Stelle $\mathbf{x}_0 \in \mathbb{R}^m$

f ist stetig in $\mathbf{x}_0 \in \mathbb{R}^m$, wenn für jede gegen \mathbf{x}_0 konvergente Folge $(\mathbf{x}_n)_{n \in \mathbb{N}} \in \mathbb{R}^m$ auch die entsprechende Folge der Funktionswerte $f(\mathbf{x}_n)$ gegen $f(\mathbf{x}_0)$ konvergiert.

Kurz: $\lim_{n \rightarrow \infty} \mathbf{x}_n = \mathbf{x}_0 \Rightarrow \lim_{n \rightarrow \infty} f(\mathbf{x}_n) = f(\mathbf{x}_0)$.

Wie bei den Funktionen einer Variablen sind auch im mehrdimensionalen Fall alle in der Praxis relevanten Funktionen überall in ihrem Definitionsbereich stetig, mit Ausnahme von stückweise definierten Funktionen, die an ihre Nahtstellen unstetig sein können.

Beispiel: Die Funktion $f(x, y) = \frac{x \cdot y}{x^2 + y^2}$ ist definiert $\forall (x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\}$ und dort auch stetig.

An der Definitionslücke $(0, 0)$ kann die Funktion nicht stetig ergänzt werden, d.h. es kann kein Funktionswert $f(0, 0)$ gefunden werden, so dass die Funktion dort stetig wird, denn:

Die Folgen $(1/n, 1/n)$, $(-1/n, 1/n)$, $(0, 1/n)$ beispielsweise konvergieren alle gegen $(0, 0)$. Die entsprechenden Folgen der Funktionswerte streben aber nicht gegen einen einheitlichen Grenzwert, da

$$f(1/n, 1/n) = \frac{1/n^2}{1/n^2 + 1/n^2} = \frac{1}{2}, \quad f(-1/n, 1/n) = \frac{-1/n^2}{1/n^2 + 1/n^2} = -\frac{1}{2}, \quad f(0, 1/n) = \frac{0}{0 + 1/n^2} = 0 \quad \forall n \in \mathbb{N}.$$

Partielle Ableitungen 1. Ordnung von f an der Stelle $\mathbf{x}_0 \in \mathbb{R}^m$:

$$\frac{\partial f}{\partial x_1}(\mathbf{x}_0) = \lim_{h \rightarrow 0} \frac{f(\mathbf{x}_0 + \mathbf{e}_1 \cdot h) - f(\mathbf{x}_0)}{h}, \dots, \frac{\partial f}{\partial x_m}(\mathbf{x}_0) = \lim_{h \rightarrow 0} \frac{f(\mathbf{x}_0 + \mathbf{e}_m \cdot h) - f(\mathbf{x}_0)}{h}.$$

Dabei bezeichnet \mathbf{e}_i den i -ten Einheitsvektor im \mathbb{R}^m .

Das Symbol " ∂ " wird statt des Buchstabens "d" verwendet, um anzudeuten, dass es sich um eine Funktion mehrerer Variablen handelt, die nach einer bestimmten Variablen partiell abgeleitet wird.

Statt $\frac{\partial f}{\partial x_i}(\mathbf{x}_0)$ schreibt man auch kurz $f_i(\mathbf{x}_0)$.

Existieren alle m Grenzwerte, so nennt man $f(\mathbf{x})$ *partiell differenzierbar* in \mathbf{x}_0 .

Zur *konkreten Berechnung* der partiellen Ableitungen verwendet man die bekannten Ableitungsregeln für Funktionen einer Variablen. Die partielle Ableitungsfunktion f_i z.B. entsteht durch Ableiten von f nach der Variablen x_i , indem alle übrigen Variablen als Konstante angesehen werden.

Beispiele

Auf die folgenden Beispiele wird im Weiteren immer wieder zurückgegriffen.

$$(1) f(x, y) = x - y, \quad \frac{\partial f}{\partial x}(x, y) = f_x(x, y) = 1, \quad \frac{\partial f}{\partial y}(x, y) = f_y(x, y) = -1$$

$$(2) f(x, y) = -x^2 - y^2 + 4, \quad \frac{\partial f}{\partial x}(x, y) = f_x(x, y) = -2x, \quad \frac{\partial f}{\partial y}(x, y) = f_y(x, y) = -2y$$

$$(3) f(x, y) = -x^2 - y^2 - xy + 4, \quad \frac{\partial f}{\partial x}(x, y) = f_x(x, y) = -2x - y, \quad \frac{\partial f}{\partial y}(x, y) = f_y(x, y) = -2y - x$$

$$(4) f(x, y) = -x^2 - y^2 - 2xy + 4, \quad \frac{\partial f}{\partial x}(x, y) = f_x(x, y) = -2x - 2y, \quad \frac{\partial f}{\partial y}(x, y) = f_y(x, y) = -2y - 2x$$

$$(5) f(x, y) = -x^2 - y^2 - 4xy + 4, \quad \frac{\partial f}{\partial x}(x, y) = f_x(x, y) = -2x - 4y, \quad \frac{\partial f}{\partial y}(x, y) = f_y(x, y) = -2y - 4x$$

$$(6) f(x, y) = (x^2 - 1) \cdot (y^2 - 1), \quad \frac{\partial f}{\partial x}(x, y) = f_x(x, y) = 2x \cdot (y^2 - 1), \quad \frac{\partial f}{\partial y}(x, y) = f_y(x, y) = 2y \cdot (x^2 - 1)$$

Um $f_x(x_0, y_0)$ *grafisch* zu deuten, durchschneidet man das von f erzeugte "Gebirge" im Punkt $(x_0, y_0, f(x_0, y_0))$ parallel zur x -Achse. $f_x(x_0, y_0)$ misst dann die Steigung der Schnittfunktion in diesem Punkt. $f_y(x_0, y_0)$ gibt analog die Steigung der Schnittfunktion parallel zur y -Achse an.

Gradient

Den Vektor der partiellen Ableitungsfunktionen 1. Ordnung bezeichnet man als Gradienten von f .

Schreibweise: $\begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix}$, grad_f , $\frac{df}{d\mathbf{x}}$ oder ∇f .

Für die vorstehenden Beispiele (1) bis (6) sind die Gradienten

$$(1) \text{grad}_f(x, y) = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad (2) \text{grad}_f(x, y) = \begin{pmatrix} -2x \\ -2y \end{pmatrix}, \quad (3) \text{grad}_f(x, y) = \begin{pmatrix} -2x-y \\ -2y-x \end{pmatrix},$$

$$(4) \text{grad}_f(x, y) = \begin{pmatrix} -2x-2y \\ -2y-2x \end{pmatrix}, \quad (5) \text{grad}_f(x, y) = \begin{pmatrix} -2x-4y \\ -2y-4x \end{pmatrix}, \quad (6) \text{grad}_f(x, y) = \begin{pmatrix} 2x \cdot (y^2-1) \\ 2y \cdot (x^2-1) \end{pmatrix}.$$

Der Gradient ist also eine Funktion von \mathbb{R}^m nach \mathbb{R}^m .

Für jeden Punkt \mathbf{x}_0 gibt $\text{grad}_f(\mathbf{x}_0)$ die Richtung des "steilsten Anstiegs" von f in diesem Punkt an.

Für den Punkt $(x_0, y_0) = (2, 2)$ und den Gradienten $\text{grad}_f(x, y) = \begin{pmatrix} -2x-y \\ -2y-x \end{pmatrix}$ aus Beispiel (3) erhält

man $\text{grad}_f(2, 2) = \begin{pmatrix} -6 \\ -6 \end{pmatrix}$. Man muss also von (x_0, y_0) aus zunächst ein Stück in Richtung $\begin{pmatrix} -6 \\ -6 \end{pmatrix}$

gehen, um den kürzesten Weg zum Gipfel einzuschlagen. Der Gradient besagt i.a. weder, dass sich das nächste lokale Maximum genau in der angegebenen Richtung befindet, noch wie weit es bis dorthin ist. Trotzdem wird diese Eigenschaft des Gradienten verwendet, um auf numerischem Wege lokale Extrema zu bestimmen (Gradientenverfahren).

Kritische Punkte

$\mathbf{x}_0 \in \mathbb{R}^m$ ist ein *kritischer Punkt* von f , wenn alle partiellen Ableitungen an der Stelle \mathbf{x}_0 gleich 0 sind, also $\text{grad}_f(\mathbf{x}_0) = \mathbf{0}$ ist. Kritische Punkte sind "Kandidaten" für lokale Extrema.

Beispiele

Für die obigen Beispiele (2), (3) und (5) ergibt sich stets der kritischen Punkt $(x_0, y_0) = (0, 0)$, z.B. in

$$(3) \text{ durch } \begin{pmatrix} -2x-y \\ -2y-x \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{matrix} -2x-y=0 \\ -x-2y=0 \end{matrix} \Leftrightarrow (x_0, y_0) = (0, 0) \quad (\text{Cramersche Regel oder Gauß-Algorithmus}).$$

Die Funktion (4) besitzt ∞ viele kritische Punkte (x, y) , nämlich die mit $x = -y$.

Im Beispiel (6) erhält man 5 kritische Punkte: $(0, 0), (1, 1), (1, -1), (-1, 1), (-1, -1)$. Das Gleichungssystem ist in diesem Fall nicht mehr linear wie in (2) - (5).

Das Auffinden kritischer Punkte in nichtlinearen Gleichungssystemen ist häufig analytisch nicht möglich, z. B. für $f(x, y) = x^6y + xy^6 + x^2 + y^2$ mit $f_x(x, y) = 6x^5y + y^6 + 2x$, $f_y(x, y) = x^6 + 6xy^5 + 2y$.

Regel *Notwendiges Kriterium zum Bestimmen lokale Extrema* mittels Ableitungen

Sei f partiell diff.bar in \mathbf{x}_0 . Dann gilt:

- (1) \mathbf{x}_0 ist ein lokales Extremum von $f \Rightarrow \mathbf{x}_0$ ist ein kritischer Punkt von f ;
 oder äquivalent: \mathbf{x}_0 ist kein kritischer Punkt von $f \Rightarrow \mathbf{x}_0$ ist kein lokales Extremum von f ;
- (2) \mathbf{x}_0 ist ein kritischer Punkt von $f \Rightarrow \mathbf{x}_0$ ist ein lokales Extremum oder ein Sattelpunkt von f .

Sattelpunkte sind kritische Punkte, an denen kein lokales Extremum vorliegt. Der Name legt die Verbindung zum Pferdesattel bzw. Gebirgssattel nahe. Anschaulich handelt es sich dabei in der Regel um Punkte, die aus einer Richtung betrachtet lokale Minima sind, aus einer anderen Richtung gesehen jedoch lokale Maxima.

Funktionalmatrix (*Jacobimatrix*): Matrix der partiellen Ableitungen 1. Ordnung einer Funktion \mathbf{g} von \mathbb{R}^m nach \mathbb{R}^n mit $\mathbf{g}(\mathbf{x}) = \begin{pmatrix} g_1(\mathbf{x}) \\ \vdots \\ g_n(\mathbf{x}) \end{pmatrix}, \mathbf{x} \in \mathbb{R}^m$.

Wie oben erwähnt, gehört beispielsweise der Gradient zu diesem Typ von Funktionen. Jede Teilfunktion g_i kann nach allen m Variablen partiell abgeleitet werden. Die insgesamt $n \cdot m$ Ableitungen fasst man zur Jacobimatrix zusammen:

$$\mathbf{J}_g(\mathbf{x}) = \frac{d\mathbf{g}}{d\mathbf{x}}(\mathbf{x}) = \begin{pmatrix} \frac{\partial g_1}{\partial x_1}(\mathbf{x}) & \cdots & \frac{\partial g_1}{\partial x_m}(\mathbf{x}) \\ \vdots & & \vdots \\ \frac{\partial g_n}{\partial x_1}(\mathbf{x}) & \cdots & \frac{\partial g_n}{\partial x_m}(\mathbf{x}) \end{pmatrix}$$

Beispiel: Für $\mathbf{g}(x, y) = \begin{pmatrix} g_1(x, y) = x^2 + y^2 \\ g_2(x, y) = 3xy + 4 \end{pmatrix}$ ist $\mathbf{J}_g(x, y) = \begin{pmatrix} 2x & 2y \\ 3y & 3x \end{pmatrix}$.

Partielle Ableitungen 2. Ordnung: Um festzustellen, ob es sich bei einem kritischen Punkt um ein lokales Minimum, Maximum oder um einen Sattelpunkt handelt, müssen die 2. Ableitungen herangezogen werden. Dazu wird jede der m Ableitungsfunktionen 1. Ordnung nochmals nach jeder einzelnen Variablen partiell abgeleitet, so dass insgesamt m^2 Ableitungsfunktionen 2. Ordnung entstehen:

$$\begin{matrix} \frac{\partial^2 f}{\partial x_1^2}, & \frac{\partial^2 f}{\partial x_1 \partial x_2}, & \cdots, & \frac{\partial^2 f}{\partial x_1 \partial x_m} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}, & \frac{\partial^2 f}{\partial x_2^2}, & \cdots, & \frac{\partial^2 f}{\partial x_2 \partial x_m} \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_m \partial x_1}, & \frac{\partial^2 f}{\partial x_m \partial x_2}, & \cdots, & \frac{\partial^2 f}{\partial x_m^2} \end{matrix}$$

Die Ableitungsfunktion 2. Ordnung, $\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial f_i}{\partial x_j}$, in Kurzschreibweise f_{ij} , entsteht also durch nochmaliges Ableiten der 1. Ableitungsfunktion f_i nach der Variablen x_j .

Beispielsweise erhält man für die Funktion (3) aus $f_x(x, y) = -2x - y$ und $f_y(x, y) = -2y - x$ die Ableitungen 2. Ordnung

$$\begin{aligned} \frac{\partial^2 f}{\partial x^2}(x, y) &= \frac{\partial f_x}{\partial x}(x, y) = f_{xx}(x, y) = -2, & \frac{\partial^2 f}{\partial x \partial y}(x, y) &= \frac{\partial f_x}{\partial y}(x, y) = f_{xy}(x, y) = -1, \\ \frac{\partial^2 f}{\partial y \partial x}(x, y) &= \frac{\partial f_y}{\partial x}(x, y) = f_{yx}(x, y) = -1, & \frac{\partial^2 f}{\partial y^2}(x, y) &= \frac{\partial f_y}{\partial y}(x, y) = f_{yy}(x, y) = -2. \end{aligned}$$

Man fasst die Ableitungen 2. Ordnung zusammen zur

Hesse-Matrix $\mathbf{H}_f = \frac{d^2 f}{d\mathbf{x}^2} = \begin{pmatrix} f_{11} & f_{12} & \cdots & f_{1m} \\ f_{21} & f_{22} & \cdots & f_{2m} \\ \vdots & \vdots & \cdots & \vdots \\ f_{m1} & f_{m2} & \cdots & f_{mm} \end{pmatrix}$ von f .

Für die oben angeführten Beispiele (1) bis (6) ergeben sich folgende Hesse-Matrizen:

$$\begin{aligned} (1) \mathbf{H}_f(x, y) &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, & (2) \mathbf{H}_f(x, y) &= \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}, & (3) \mathbf{H}_f(x, y) &= \begin{pmatrix} -2 & -1 \\ -1 & -2 \end{pmatrix}, \\ (4) \mathbf{H}_f(x, y) &= \begin{pmatrix} -2 & -2 \\ -2 & -2 \end{pmatrix}, & (5) \mathbf{H}_f(x, y) &= \begin{pmatrix} -2 & -4 \\ -4 & -2 \end{pmatrix}, & (6) \mathbf{H}_f(x, y) &= \begin{pmatrix} 2(y^2-1) & 4xy \\ 4xy & 2(x^2-1) \end{pmatrix}. \end{aligned}$$

Die Hesse-Matrix ist eine Funktion von \mathbb{R}^m nach $\mathbb{R}^{m \times m}$ und kann als Jacobimatrix des Gradienten aufgefasst werden.

Existieren alle partiellen Ableitungen 2. Ordnung in einem Punkt \mathbf{x}_0 , so nennt man f dort *2-fach partiell differenzierbar*. Für die gemischten partiellen Ableitungen f_{ij} mit $i \neq j$ gilt der folgende Satz.

Satz von Schwarz Besitzt f stetige partielle Ableitungen 2. Ordnung im Punkt \mathbf{x}_0 , so ist $f_{ij}(\mathbf{x}_0) = f_{ji}(\mathbf{x}_0)$, d.h. die Hesse-Matrix ist in der Praxis symmetrisch (siehe obige Beispiele).

Krümmung von f Wie bei Funktionen einer Variablen ermöglichen die Ableitungen 2. Ordnung Aussagen über die Krümmung der Funktion und damit auch über das Vorhandensein lokaler Extrema. Der Funktionswert $f_{11}(\mathbf{x}_0)$ z.B. misst die Krümmung der Schnittfunktion, die parallel zur x_1 -Achse verläuft, im Punkt \mathbf{x}_0 .

Regel: $M \subset \mathbb{R}^m$ sei eine konvexe Menge¹⁰ und f 2-fach partiell diff.bar in M . Dann gilt:

(1) $\mathbf{H}_f(\mathbf{x})$ positiv definit $\forall \mathbf{x} \in M \Rightarrow f$ streng konvex in M ;

(2) $\mathbf{H}_f(\mathbf{x})$ negativ definit $\forall \mathbf{x} \in M \Rightarrow f$ streng konkav in M .

Die Überprüfung einer Matrix auf Definitheit in einem Punkt \mathbf{x} erfolgt durch Berechnung der Determinanten aller Hauptuntermatrizen:

Fall (1): Alle Determinanten sind positiv $\Rightarrow \mathbf{H}_f(\mathbf{x})$ ist positiv definit;

Fall (2): die Determinanten weisen wechselndes Vorzeichen auf, beginnend mit "-"
 $\Rightarrow \mathbf{H}_f(\mathbf{x})$ ist negativ definit;

Fall (3): eine der Determinanten in (1) oder (2) besitzt ein "falsches" Vorzeichen
 $\Rightarrow \mathbf{H}_f(\mathbf{x})$ ist indefinit.

Die Hesse-Matrizen der oben angeführten Beispiele (2) und (3) sind negativ definit $\forall (x, y) \in \mathbb{R}^2$, d.h. die beiden Funktionen $f(x, y) = -x^2 - y^2 + 4$, $f(x, y) = -x^2 - y^2 - xy + 4$ sind konkav im \mathbb{R}^2 .

Regel *Hinreichendes Kriterium für das Vorliegen lokaler Extrema*

$\mathbf{x}_0 \in \mathbb{R}^m$ sei ein kritischer Punkt von f mit Hesse-Matrix $\mathbf{H}_f(\mathbf{x}_0)$. Dann gilt:

(1) $\mathbf{H}_f(\mathbf{x}_0)$ positiv definit (also f streng konvex in einer Umgebung von \mathbf{x}_0)
 $\Rightarrow \mathbf{x}_0$ ist ein strenges lokales Minimum von f ;

(2) $\mathbf{H}_f(\mathbf{x}_0)$ negativ definit (also f streng konkav in einer Umgebung von \mathbf{x}_0)
 $\Rightarrow \mathbf{x}_0$ ist ein strenges lokales Maximum von f ;

(3) $\mathbf{H}_f(\mathbf{x}_0)$ indefinit $\Rightarrow \mathbf{x}_0$ ist ein Sattelpunkt.

Bei semidefiniter Hesse-Matrix kann ohne zusätzliche Information keine Aussage gemacht werden, denn die Funktion $f(x, y) = x^2 + y^3$ besitzt bei semidefiniter Hesse-Matrix in $(0, 0)$ einen Sattelpunkt, während $f(x, y) = x^4 + y^4$ bei ebenfalls semidefiniter Hesse-Matrix in $(0, 0)$ ein Minimum hat.

Beispiele Die Hesse-Matrizen der Beispiele (2) und (3) sind negativ definit $\forall (x, y) \in \mathbb{R}^2$, also auch an ihrem jeweiligen kritischen Punkt $(0, 0)$. Dort liegt in beiden Fällen ein strenges lokales Maximum vor. Im Beispiel (4) ist die Hesse-Matrix negativ semidefinit $\forall (x, y) \in \mathbb{R}^2$ und damit zunächst keine Aussage möglich. Aus der grafischen Darstellung geht hervor, dass es sich bei den ∞ vielen kritischen Punkten um (gleich hohe) lokale Maxima handelt. Es existiert aber kein strenges lokales Maximum. Die Funktion (5) besitzt eine indefinite Hesse-Matrix $\forall (x, y) \in \mathbb{R}^2$ und daher einen Sattelpunkt in $(0, 0)$.

Beachten: Es reicht nicht aus, nur die beiden partiellen Ableitungen f_{xx} und f_{yy} am kritischen Punkt auf ihr Vorzeichen hin zu untersuchen (wie man in Analogie zum Fall einer Variablen evtl. vermutet).

Bei der Funktion (6) erhält man für die 5 kritischen Punkte folgende Hesse-Matrizen:

$\mathbf{H}_f(0, 0) = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}$ (negativ definit), $\mathbf{H}_f(1, 1) = \begin{pmatrix} 0 & 4 \\ 4 & 0 \end{pmatrix} = \mathbf{H}_f(-1, -1)$ (indefinit),

$\mathbf{H}_f(1, -1) = \begin{pmatrix} 0 & -4 \\ -4 & 0 \end{pmatrix} = \mathbf{H}_f(-1, 1)$ (indefinit). Im Punkt $(0, 0)$ liegt daher ein strenges lokales

Maximum vor, bei allen anderen Punkten handelt es sich um Sattelpunkte.

Regel *Globale Extrema bei konvexen oder konkaven Funktionen*

$\mathbf{x}_0 \in M \subset \mathbb{R}^m$ sei ein kritischer Punkt der 2-fach diff.baren Funktion f in einer konvexen Menge M .

Dann gilt:

(1) $\mathbf{H}_f(\mathbf{x})$ positiv definit $\forall \mathbf{x} \in M$ (also f streng konvex in M)

$\Rightarrow \mathbf{x}_0$ ist ein strenges globales Minimum von f in M ;

(2) $\mathbf{H}_f(\mathbf{x})$ negativ definit $\forall \mathbf{x} \in M$ (also f streng konkav in M)

¹⁰d.h. mit je zwei Punkten aus M liegt auch die gesamte Verbindungsstrecke zwischen den beiden Punkten in M

$\Rightarrow \mathbf{x}_0$ ist ein strenges *globales Maximum* von f in M .

Bei den kritischen Punkten der Beispiele (2) und (3) handelt es sich daher nicht nur um lokale, sondern sogar um globale Maxima.

Zusammenfassung Vorgehen beim *Bestimmen lokaler Extrema* für Funktionen mehrerer Variablen

- Funktion nach jeder vorkommenden Variable einzeln partiell ableiten;
- die Ableitungsfunktionen alle gleichzeitig Null setzen; so entsteht ein Gleichungssystem;
- Lösungsmenge des Gleichungssystems bestimmen (kritische Punkte);
- jede Ableitungsfunktion 1. Ordnung nochmals nach jeder Variablen partiell ableiten; dabei kann man sich Arbeit ersparen, wenn man berücksichtigt, dass $f_{ij}(\mathbf{x}) = f_{ji}(\mathbf{x})$;
- Funktionswerte aller Ableitungen 2. Ordnung an einem kritischen Punkt \mathbf{x}_0 bestimmen;
- Funktionswerte in Matrixform zusammenfassen (Hesse-Matrix $\mathbf{H}_f(\mathbf{x}_0)$);
- Hesse-Matrix auf Definitheit überprüfen:
 $\mathbf{H}_f(\mathbf{x}_0)$ positiv definit $\Rightarrow \mathbf{x}_0$ strenges lokales Minimum;
 $\mathbf{H}_f(\mathbf{x}_0)$ negativ definit $\Rightarrow \mathbf{x}_0$ strenges lokales Maximum;
 $\mathbf{H}_f(\mathbf{x}_0)$ indefinit $\Rightarrow \mathbf{x}_0$ Sattelpunkt.

16. Symbolisches Ableiten von Linearformen und quadratischen Formen

Linearform Lineare Funktion $f: \mathbb{R}^n \rightarrow \mathbb{R}$ mit $f(\mathbf{x}) = \mathbf{a}'\mathbf{x} = a_1x_1 + a_2x_2 + \dots + a_nx_n$, $\mathbf{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$, $\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$.

Die partielle Ableitung der Funktion f nach der Variablen x_i ergibt a_i . Hieraus folgt sofort die *Regel*: Eine Linearform $f(\mathbf{x}) = \mathbf{a}'\mathbf{x}$ besitzt den Gradienten $\text{grad}_f(\mathbf{x}) = \mathbf{a}$ und die Hesse-Matrix $\mathbf{H}_f(\mathbf{x}) = \mathbf{0}$.

Beispiel: $f(x_1, x_2, x_3) = 3x_1 + x_2 + 2x_3 = (3, 1, 2) \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \mathbf{a}'\mathbf{x}$ mit $\mathbf{a} = \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix}$, $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$.

Dann ist $\text{grad}_f(\mathbf{x}) = \mathbf{a} = \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix}$, $\mathbf{H}_f(\mathbf{x}) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$.

Quadratische Form Quadratische Funktion $f: \mathbb{R}^n \rightarrow \mathbb{R}$ mit

$$f(\mathbf{x}) = \mathbf{x}'\mathbf{M}\mathbf{x} = m_{11}x_1^2 + (m_{12} + m_{21})x_1x_2 + \dots + (m_{ij} + m_{ji})x_ix_j + \dots + m_{nn}x_n^2,$$

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n, \quad \mathbf{M} = \begin{pmatrix} m_{11} & \dots & m_{1n} \\ \vdots & & \vdots \\ m_{m1} & \dots & m_{mn} \end{pmatrix} \in \mathbb{R}^{n \times n}.$$

Leitet man f partiell nach jeder Variablen x_i ab und fasst die Ergebnisse wieder in Matrixform zusammen, so gelangt man zur

Regel: Eine quadratische Form $f(\mathbf{x}) = \mathbf{x}'\mathbf{M}\mathbf{x}$ besitzt den Gradienten $\text{grad}_f(\mathbf{x}) = (\mathbf{M} + \mathbf{M}') \cdot \mathbf{x}$ und die Hesse-Matrix $\mathbf{H}_f(\mathbf{x}) = \mathbf{M} + \mathbf{M}'$.

Beispiel: $f(\mathbf{x}) = \mathbf{x}'\mathbf{M}\mathbf{x} = (x_1, x_2, x_3) \cdot \begin{pmatrix} 1 & 2 & 2 \\ 1 & 2 & 3 \\ 2 & 2 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_1^2 + 2x_2^2 + x_3^2 + 3x_1x_2 + 4x_1x_3 + 5x_2x_3$

mit $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$, $\mathbf{M} = \begin{pmatrix} 1 & 2 & 2 \\ 1 & 2 & 3 \\ 2 & 2 & 1 \end{pmatrix}$. Dann ist $\text{grad}_f(\mathbf{x}) = (\mathbf{M} + \mathbf{M}') \cdot \mathbf{x} = \begin{pmatrix} 2 & 3 & 4 \\ 3 & 4 & 5 \\ 4 & 5 & 2 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} =$

$\begin{pmatrix} 2x_1 + 3x_2 + 4x_3 \\ 3x_1 + 4x_2 + 5x_3 \\ 4x_1 + 5x_2 + 2x_3 \end{pmatrix}$ und $\mathbf{H}_f(\mathbf{x}) = \mathbf{M} + \mathbf{M}' = \begin{pmatrix} 2 & 3 & 4 \\ 3 & 4 & 5 \\ 4 & 5 & 2 \end{pmatrix}$.

Eine Anwendung dieser Regeln findet man in den Aufgaben 17.3 (a), 17.4 und 17.5 der Aufgabensammlung sowie im folgenden Exkurs.

Exkurs

Lineare Regression

Gegebene Daten: Vektoren $\mathbf{a}_1, \dots, \mathbf{a}_m, \mathbf{b} \in \mathbb{R}^n$,

z.B. $m+1$ Zeitreihen ökonomischer Größen zu jeweils n Zeitpunkten.

Fragestellung: Wie gut lässt sich die Größe \mathbf{b} vorhersagen, wenn man $\mathbf{a}_1, \dots, \mathbf{a}_m$ kennt?
 Genauer: Welcher lineare Zusammenhang besteht zwischen der Größe \mathbf{b} und den Größen $\mathbf{a}_1, \dots, \mathbf{a}_m$?

Lösung: Suche lineare Gewichte x_1, \dots, x_m , so dass $|\mathbf{a}_1 x_1 + \mathbf{a}_2 x_2 + \dots + \mathbf{a}_m x_m - \mathbf{b}| = \min!$, d.h. der Vektor der Fehler bei der Annäherung von \mathbf{b} durch eine Linearkombination der \mathbf{a}_i soll betragsmäßig minimal werden.

Mit den Bezeichnungen $\mathbf{A} = (\mathbf{a}_1, \dots, \mathbf{a}_m)$, $\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}$, $\mathbf{u} = \mathbf{A}\mathbf{x} - \mathbf{b}$ lautet die

Zielfunktion $|\mathbf{u}| = \min!$, also $\sqrt{\mathbf{u}'\mathbf{u}} = \min!$. Dies ist genau dann der Fall, wenn $\mathbf{u}'\mathbf{u}$ minimal wird. Somit erhält man die zu minimierende Zielfunktion

$$f(\mathbf{x}) = \mathbf{u}'\mathbf{u} = (\mathbf{A}\mathbf{x} - \mathbf{b})'(\mathbf{A}\mathbf{x} - \mathbf{b}) = \mathbf{x}'\mathbf{A}'\mathbf{A}\mathbf{x} - 2\mathbf{b}'\mathbf{A}\mathbf{x} + \mathbf{b}'\mathbf{b},$$

eine Kombination der quadratischen Form $\mathbf{x}'\mathbf{A}'\mathbf{A}\mathbf{x}$ (mit der symmetrischen, positiv semidefiniten Matrix $\mathbf{M} = \mathbf{A}'\mathbf{A} = \mathbf{M}'$), der Linearform $2\mathbf{b}'\mathbf{A}\mathbf{x}$ (mit dem Zeilenvektor $\mathbf{a}' = 2\mathbf{b}'\mathbf{A}$, also $\mathbf{a} = 2\mathbf{A}'\mathbf{b}$) und der Konstanten $\mathbf{b}'\mathbf{b}$.

Nach den obigen Ableitungsregeln ergibt sich dann $\text{grad}_f(\mathbf{x}) = 2\mathbf{A}'\mathbf{A}\mathbf{x} - 2\mathbf{A}'\mathbf{b}$ und $\mathbf{H}_f(\mathbf{x}) = 2\mathbf{A}'\mathbf{A}$. Setzt man den Gradienten gleich $\mathbf{0}$, so erhält man einen kritischen Punkt \mathbf{x}_0 von f als Lösung des linearen $m \times m$ Gleichungssystems $\mathbf{A}'\mathbf{A}\mathbf{x} = \mathbf{A}'\mathbf{b}$. Sind die Vektoren \mathbf{a}_i linear unabhängig (und $n \geq m$), dann ist $\mathbf{A}'\mathbf{A}$ regulär, \mathbf{x}_0 somit eindeutig bestimmt und darstellbar durch $\mathbf{x}_0 = (\mathbf{A}'\mathbf{A})^{-1} \mathbf{A}'\mathbf{b}$. Außerdem ist $\mathbf{A}'\mathbf{A}$ in diesem Falle positiv definit und \mathbf{x}_0 daher das globale Minimum der Zielfunktion. Als ein Maß zur Beurteilung der Güte der Anpassung dient der quadrierte Korrelationskoeffizient zwischen \mathbf{b} und dem Vektor $\mathbf{A}\mathbf{x}_0$.

17. Partielles und totales Differenzial, partielle Wachstumsrate, partielle Elastizität

Partielles Differenzial df_i : Sei f eine Funktion von \mathbb{R}^m nach \mathbb{R} , $\mathbf{x}_0 = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}$ ein Punkt im \mathbb{R}^m , $d\mathbf{x} = \begin{pmatrix} d_1 \\ \vdots \\ d_m \end{pmatrix} \in \mathbb{R}^m$ ein

Vektor von (kleinen) Änderungen von \mathbf{x}_0 . Variiert man \mathbf{x}_0 nur in der i -ten Komponente um d_i und lässt alle übrigen Komponenten konstant (ceteris paribus Betrachtung), dann ändert sich der Funktionswert um $f(\mathbf{x}_0 + \mathbf{e}_i \cdot d_i) - f(\mathbf{x}_0)$.

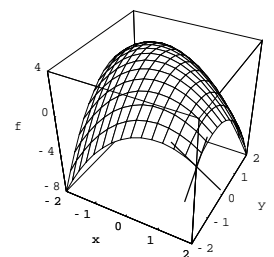
Dieser Wert ist für kleine Variationen d_i näherungsweise gleich dem *partiellen Differenzial*

$$df_i = f_i(\mathbf{x}_0) \cdot d_i.$$

Das partielle Differenzial df_i gibt die Änderung des Funktionswertes auf der Tangenten T_i an, $T_i(\mathbf{x}_0 + \mathbf{e}_i \cdot d_i) = f(\mathbf{x}_0) + f_i(\mathbf{x}_0) \cdot d_i$.

Tangente T_i bezeichnet die Tangente durch den Punkt $(\mathbf{x}_0, f(\mathbf{x}_0))$ an die Schnittfunktion, die durch das von f erzeugte Gebirge parallel zur x_i -Achse verläuft.

Bsp. (Grafik): $f(x, y) = -x^2 - y^2 - xy + 4$, Tangenten durch $(2, -2, 0)$.



Partielle Wachstumsrate r_{f_i} : Analog wie im Fall der Funktionen einer Variablen definiert und interpretiert man die partielle Wachstumsrate $r_{f_i}(\mathbf{x}) = \frac{f_i(\mathbf{x})}{f(\mathbf{x})}$ sowie die

Partielle Elastizität $\varepsilon_{f_i}(\mathbf{x}) = x_i \cdot r_{f_i}(\mathbf{x}) = x_i \cdot \frac{f_i(\mathbf{x})}{f(\mathbf{x})}$.

Die partielle Elastizität in einem Punkt \mathbf{x}_0 , $\varepsilon_{f_i}(\mathbf{x}_0)$, gibt näherungsweise an, um wie viel % sich $f(\mathbf{x}_0)$ ändert, wenn \mathbf{x}_0 in der i -ten Komponente um 1 % variiert.

Die bekannten Rechenregeln für Elastizitäten bleiben auch hier weiter erhalten.

Beispiel Betrachte die Funktion $f(x, y) = 2x^{1/3}y^{2/3}$ an der Stelle $(x_0, y_0) = (64, 8)$ mit $f(x_0, y_0) = 2 \cdot 4 \cdot 4 = 32$, $f_x(x, y) = \frac{2}{3}x^{-2/3}y^{2/3}$, $f_x(x_0, y_0) = \frac{2}{3} \cdot \frac{1}{16} \cdot 4 = \frac{1}{6}$, $f_y(x, y) = \frac{4}{3}x^{1/3}y^{-1/3}$, $f_y(x_0, y_0) = \frac{4}{3} \cdot 4 \cdot \frac{1}{2} = \frac{8}{3}$,

$$r_{f_x}(x, y) = \frac{f_x(x, y)}{f(x, y)} = \frac{2/3 \cdot x^{-2/3} \cdot y^{2/3}}{2x^{1/3}y^{2/3}} = \frac{1}{3x}, \quad r_{f_x}(x_0, y_0) = \frac{1}{3 \cdot 64}, \quad r_{f_y}(x, y) = \frac{f_y(x, y)}{f(x, y)} = \frac{4/3 \cdot x^{1/3} \cdot y^{-1/3}}{2x^{1/3}y^{2/3}} = \frac{2}{3y},$$

$$r_{f_y}(x_0, y_0) = \frac{2}{3 \cdot 8} = \frac{1}{12}, \quad \epsilon_{f_x}(x, y) = x \cdot r_{f_x}(x, y) = \frac{1}{3} = \epsilon_{f_x}(x_0, y_0), \quad \epsilon_{f_y}(x, y) = y \cdot r_{f_y}(x, y) = \frac{2}{3} = \epsilon_{f_y}(x_0, y_0).$$

- Variiert man x_0 um beispielsweise $dx = 2$ (und lässt y_0 konstant), so ändert sich $f(x_0, y_0)$ um

$$f(x_0 + dx, y_0) - f(x_0, y_0) = 32,33 - 32 = 0,330 \text{ Einheiten} \quad (\text{exakt}).$$

Dieser Wert entspricht näherungsweise dem partiellen Differenzial

$$df_x = f_x(x_0, y_0) \cdot dx = \frac{1}{6} \cdot 2 = 0,333 \text{ Einheiten} \quad (\text{angenähert}).$$

Variiert man analog y_0 um beispielsweise $dy = 1$ (und lässt x_0 konstant), so ändert sich $f(x_0, y_0)$

um $f(x_0, y_0 + dy) - f(x_0, y_0) = 34,61 - 32 = 2,61$ Einheiten (exakt).

Dieser Wert entspricht näherungsweise dem partiellen Differenzial

$$df_y = f_y(x_0, y_0) \cdot dy = \frac{8}{3} \cdot 1 = 2,66 \text{ Einheiten} \quad (\text{angenähert}).$$

- Die relative (prozentuale) Änderung von $f(x_0, y_0)$ beträgt hier bei Variation von

$$x_0 \text{ um } dx = 2 : \frac{0,33}{32} = 0,0103 = 1,03 \% \quad (\text{exakt}),$$

$$y_0 \text{ um } dy = 1 : \frac{2,61}{32} = 0,0817 = 8,17 \% \quad (\text{exakt}).$$

Diese Werte entsprechen näherungsweise den partiellen Wachstumsraten

$$r_{f_x}(x_0, y_0) \cdot dx = \frac{1}{3 \cdot 64} \cdot 2 = \frac{1}{96} = 0,0104 = 1,04 \% \quad (\text{angenähert}),$$

$$r_{f_y}(x_0, y_0) \cdot dy = \frac{1}{12} \cdot 1 = 0,0833 = 8,33 \% \quad (\text{angenähert}).$$

Die prozentuale Änderung von $f(x_0, y_0)$ beträgt beispielsweise bei Variationen von

$$x_0 \text{ um } 5\% : \frac{f(67,2; 8) - 32}{32} = 0,0164 = 1,64 \% \quad (\text{exakt}),$$

$$y_0 \text{ um } 10\% : \frac{f(64; 8,8) - 32}{32} = 0,0656 = 6,56 \% \quad (\text{exakt}).$$

Diese Werte entsprechen näherungsweise den partiellen Elastizitäten

$$\epsilon_{f_x}(x_0, y_0) \cdot s_x \% = \frac{1}{3} \cdot 5 \% = 1,67 \% \quad (\text{angenähert}),$$

$$\epsilon_{f_y}(x_0, y_0) \cdot s_y \% = \frac{2}{3} \cdot 10 \% = 6,67 \% \quad (\text{angenähert}).$$

Totales Differenzial (vollständiges Differenzial)

Variiert man (unter Verwendung obiger Schreibweise) nun \mathbf{x}_0 um $d\mathbf{x}$, also simultan mehr als eine Komponente, so ändert sich $f(\mathbf{x}_0)$ näherungsweise um

$$df = df_1 + \dots + df_m = f_1(\mathbf{x}_0) \cdot d_1 + \dots + f_m(\mathbf{x}_0) \cdot d_m = \begin{pmatrix} f_1(\mathbf{x}_0) \\ \vdots \\ f_m(\mathbf{x}_0) \end{pmatrix} \begin{pmatrix} d_1 \\ \vdots \\ d_m \end{pmatrix} = (\text{grad}_f(\mathbf{x}_0))' \cdot d\mathbf{x}.$$

Bei der oben angeführten Funktion $f(x, y) = 2x^{1/3}y^{2/3}$ führt die simultane Variation von $x_0 = 64$ um $dx = 2$ und $y_0 = 8$ um $dy = 1$ zu einer Änderung des Funktionswertes von $f(x_0, y_0) = 32$ um

$$f(x_0 + dx, y_0 + dy) - f(x_0, y_0) = 34,97 - 32 = 2,97 \text{ Einheiten} \quad (\text{exakt}).$$

Dieser Wert entspricht näherungsweise dem totalen Differenzial

$$df = df_x + df_y = \frac{1}{3} + \frac{8}{3} = 3 \text{ Einheiten} \quad (\text{angenähert}).$$

Das totale Differenzial df als Summe der partiellen Differenziale gibt die Änderung des Funktionswertes $f(\mathbf{x}_0)$ auf der Tangential(hyper)ebenen T an (siehe nachstehende Grafik).

Tangential(hyper)ebene T : T wird durch die Tangenten T_i aufgespannt und berührt das von f erzeugte Gebirge im Punkt

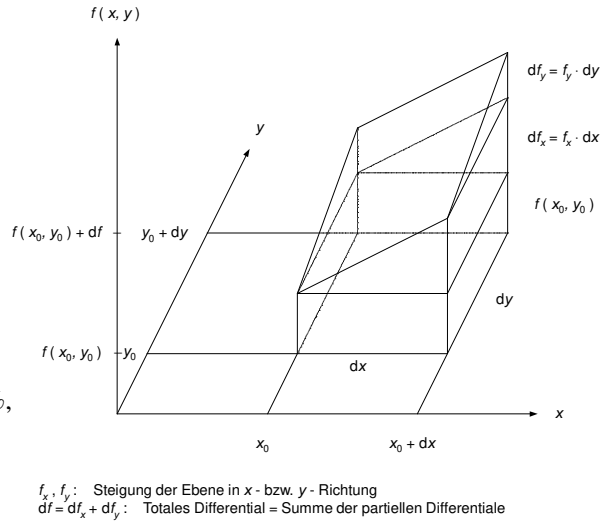
$(\mathbf{x}_0, f(\mathbf{x}_0))$. Für sie gilt: $T(\mathbf{x}_0 + d\mathbf{x}) = f(\mathbf{x}_0) + df = f(\mathbf{x}_0) + (\text{grad}_f(\mathbf{x}_0))' \cdot d\mathbf{x}$,

also mit $\mathbf{x} = \mathbf{x}_0 + d\mathbf{x} \Leftrightarrow d\mathbf{x} = \mathbf{x} - \mathbf{x}_0$: $T(\mathbf{x}) = f(\mathbf{x}_0) + (\text{grad}_f(\mathbf{x}_0))' \cdot (\mathbf{x} - \mathbf{x}_0)$.

Beispielsweise für die Funktion $f(x, y) = -x^2 - y^2 - xy + 4$ erhält man für den Punkt $(2, -2, 0)$:

$$\text{grad}_f(x, y) = \begin{pmatrix} -2x - y \\ -2y - x \end{pmatrix}, \quad \text{grad}_f(2, -2) = \begin{pmatrix} -2 \\ 2 \end{pmatrix}, \quad T(x, y) = 0 + (-2, 2) \cdot \begin{pmatrix} x - 2 \\ y + 2 \end{pmatrix} = -2x + 2y + 8.$$

Warum die Änderung des Funktionswertes auf der Tangentialebene gerade die Summe der Änderungen auf den Tangenten ergibt, erkennt man an nebenstehender Grafik.



Totale Elastizität

Auch wenn der Begriff nicht etabliert ist, könnte man ihn analog zum totalen Differential einführen. Variiert man speziell \mathbf{x}_0 in allen Komponenten simultan um $s\%$, von \mathbf{x}_0 auf $\mathbf{x}_0 + \mathbf{x}_0 \cdot s\%$, also um $d\mathbf{x} = \mathbf{x}_0 \cdot s\%$, so ändert sich $f(\mathbf{x}_0)$ näherungsweise um

$$\frac{df}{f(\mathbf{x}_0)} = \frac{(\text{grad}_f(\mathbf{x}_0))' d\mathbf{x}}{f(\mathbf{x}_0)} = \frac{(\text{grad}_f(\mathbf{x}_0))' \mathbf{x}_0 \cdot s\%}{f(\mathbf{x}_0)} = \frac{f_1(\mathbf{x}_0) \cdot x_1 + \dots + f_m(\mathbf{x}_0) \cdot x_m}{f(\mathbf{x}_0)} \cdot s\% = (\varepsilon_{f_1}(\mathbf{x}) + \dots + \varepsilon_{f_m}(\mathbf{x})) \cdot s\% . \quad (*)$$

Taylorentwicklung

von f : Die Tangential(hyper)ebene T stellt gleichzeitig das Taylorpolynom 1. Grades dar (vgl. die Analogie zur Darstellung $T_1(x) = f(x_0) + f'(x_0) \cdot (x - x_0)$ bei Funktionen einer Variablen).

Das Taylorpolynom 2. Grades von f , entwickelt an der Stelle \mathbf{x}_0 , ist

$$T_2(\mathbf{x}) = f(\mathbf{x}_0) + (\text{grad}_f(\mathbf{x}_0))'(\mathbf{x} - \mathbf{x}_0) + \frac{1}{2} \cdot (\mathbf{x} - \mathbf{x}_0)' \mathbf{H}_f(\mathbf{x}_0) (\mathbf{x} - \mathbf{x}_0).$$

Beispiel

Für $f(x, y) = 2x^{1/3}y^{2/3}$, $\text{grad}_f(x, y) = \begin{pmatrix} \frac{2}{3}x^{-2/3}y^{2/3} \\ \frac{4}{3}x^{1/3}y^{-1/3} \end{pmatrix}$, $\mathbf{H}_f(x, y) = \frac{4}{9} \begin{pmatrix} -x^{-5/3}y^{2/3} & x^{-2/3}y^{-1/3} \\ x^{-2/3}y^{-1/3} & -x^{1/3}y^{-4/3} \end{pmatrix}$ erhält

man für den Punkt $(64, 8, 32)$: $\text{grad}_f(64, 8) = \begin{pmatrix} 1/6 \\ 8/3 \end{pmatrix}$, $\mathbf{H}_f(x_0, y_0) = \frac{-1}{9} \begin{pmatrix} 1/64 & -1/8 \\ -1/8 & 1 \end{pmatrix}$,

Tangentialebene $T_1(x, y) = 32 + \begin{pmatrix} 1/6 \\ 8/3 \end{pmatrix} \cdot \begin{pmatrix} x - 64 \\ y - 8 \end{pmatrix} = \frac{1}{6}x + \frac{8}{3}y$,

Taylorpolynom 2. Grades $T_2(x, y) = \frac{1}{6}x + \frac{8}{3}y - \frac{1}{2 \cdot 9} (x - 64, y - 8) \cdot \begin{pmatrix} 1/64 & -1/8 \\ -1/8 & 1 \end{pmatrix} \cdot \begin{pmatrix} x - 64 \\ y - 8 \end{pmatrix}$
 $= \frac{1}{6}x + \frac{8}{3}y - \frac{1}{18} \left[\frac{1}{64}(x - 64)^2 - \frac{2}{8}(x - 64) \cdot (y - 8) + (y - 8)^2 \right]$
 $= \frac{1}{6}x + \frac{8}{3}y - \frac{1}{18} \left[\frac{1}{8}(x - 64) - (y - 8) \right]^2 = \frac{1}{6}x + \frac{8}{3}y - \frac{1}{18} \left[\frac{1}{8}x - y \right]^2$.

Auf Taylorpolynome höherer Ordnung soll hier nicht eingegangen werden.

18. Homogene Funktionen

Homogenität

Eine Funktion $f: \mathbb{R}^m \rightarrow \mathbb{R}$ heißt *homogen vom Grade r* , wenn

$$f(\lambda \cdot \mathbf{x}) = \lambda^r \cdot f(\mathbf{x}) \quad \forall \lambda \in \mathbb{R}_+, \mathbf{x} \in \mathbb{R}^m \neq \mathbf{0}.$$
 Für $r = 1$ nennt man f *linear homogen*.

Beispiel

(1) Die Funktion $f(x, y) = 2x^{1/3}y^{2/3}$ ist linear homogen, da

$$f(\lambda x, \lambda y) = 2(\lambda x)^{1/3}(\lambda y)^{2/3} = 2\lambda^{1/3}x^{1/3}\lambda^{2/3}y^{2/3} = 2\lambda^{1/3}\lambda^{2/3}x^{1/3}y^{2/3} = \lambda \cdot 2x^{1/3}y^{2/3} = \lambda \cdot f(x, y).$$

(2) Die Funktion $f(x, y) = xy$ ist homogen vom Grade 2, da $f(\lambda x, \lambda y) = \lambda x \cdot \lambda y = \lambda^2 xy = \lambda^2 \cdot f(x, y)$.

(3) Die Funktion $f(x, y) = x + y^2$ ist nicht homogen, da $f(\lambda x, \lambda y) = \lambda x + (\lambda y)^2 \neq \lambda^r \cdot (x + y^2)$.

Prozentuale Änderung von $f(\mathbf{x}_0)$ bei Variation von \mathbf{x}_0 um jeweils $s\%$ (in allen Komponenten)

Ändert man $\mathbf{x}_0 \in \mathbb{R}^m$ in allen m Komponenten um jeweils $s\%$, von \mathbf{x}_0 auf $\mathbf{x}_0 + \mathbf{x}_0 \cdot s\% = \mathbf{x}_0 \cdot (1 + s\%)$,

so führt dies mit $\lambda = 1 + s\%$ unter Verwendung der Homogenität zu

$$\frac{f(\lambda \cdot \mathbf{x}_0) - f(\mathbf{x}_0)}{f(\mathbf{x}_0)} = \frac{\lambda^r \cdot f(\mathbf{x}_0) - f(\mathbf{x}_0)}{f(\mathbf{x}_0)} = \frac{(\lambda^r - 1) f(\mathbf{x}_0)}{f(\mathbf{x}_0)} = \lambda^r - 1 = (1 + s\%)^r - 1.$$

Speziell für linear homogene Funktionen beträgt also auch die Änderung des Funktionswertes $s\%$.

Für $r = 2$ ist $(1 + s\%)^2 - 1 = (1 + s\% - 1)(1 + s\% + 1) = s\% (2 + s\%) = (2s + \frac{s^2}{100})\%$,

z.B. für $s = 3$: $(6 + \frac{9}{100})\% = 6,09\%$.

Andererseits gilt nach (*) die Annäherung $\frac{f(\lambda \cdot \mathbf{x}_0) - f(\mathbf{x}_0)}{f(\mathbf{x}_0)} \approx (\varepsilon_{f_1}(\mathbf{x}) + \dots + \varepsilon_{f_m}(\mathbf{x})) \cdot s\%$, also

$$(1 + s\%)^r - 1 \approx (\varepsilon_{f_1}(\mathbf{x}) + \dots + \varepsilon_{f_m}(\mathbf{x})) \cdot s\% \quad \text{und mit } s\% = \lambda - 1: \quad \varepsilon_{f_1}(\mathbf{x}) + \dots + \varepsilon_{f_m}(\mathbf{x}) \approx \frac{\lambda^r - 1}{\lambda - 1}.$$

Für $s \rightarrow 0$, also $\lambda \rightarrow 1$, erhält man mit Hilfe der Regel von l'Hospital (Fall 0/0) folgende Aussage:

Regel

(Eulersche Homogenitätsrelation): Für eine vom Grade r homogene Funktion $f: \mathbb{R}^m \rightarrow \mathbb{R}$ und

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \text{ gilt} \quad r \cdot f(\mathbf{x}) = x_1 \cdot f_1(\mathbf{x}) + \dots + x_m \cdot f_m(\mathbf{x}) = \mathbf{x}' \cdot \text{grad}_f(\mathbf{x}).$$

Teilt man die Gleichung durch $f(\mathbf{x})$, so erhält man $r = x_1 \cdot \frac{f_1(\mathbf{x})}{f(\mathbf{x})} + \dots + x_m \cdot \frac{f_m(\mathbf{x})}{f(\mathbf{x})} = \varepsilon_{f_1}(\mathbf{x}) + \dots + \varepsilon_{f_m}(\mathbf{x})$.

Für homogene Funktionen ergibt also die Summe der partiellen Elastizitäten den Homogenitätsgrad.

Beispielsweise für die obige Funktion $f(x, y) = 2x^{1/3}y^{2/3}$ ist $\varepsilon_{f_x}(x, y) = \frac{1}{3}$ und $\varepsilon_{f_y}(x, y) = \frac{2}{3}$, also

$$\varepsilon_{f_x}(x, y) + \varepsilon_{f_y}(x, y) = \frac{1}{3} + \frac{2}{3} = 1, \text{ d.h. die Funktion ist linear homogen (s.o.).}$$

Näherungsweise kann die Änderung des Funktionswertes $f(\mathbf{x}_0)$ bei Variation von \mathbf{x}_0 in allen m

Komponenten um jeweils $s\%$ bestimmt werden (siehe (*)) durch $(\varepsilon_{f_1}(\mathbf{x}) + \dots + \varepsilon_{f_m}(\mathbf{x})) \cdot s\% = r \cdot s\%$.

Variiert man beispielsweise bei der vom Grade 2 homogenen Funktion $f(x, y) = x \cdot y$ einen beliebigen Punkt (x_0, y_0) um 3% in beiden Komponenten, so ändert sich der Funktionswert $f(x_0, y_0)$ um $6,09\%$ exakt (s.o.) und näherungsweise um $2 \cdot 3\% = 6\%$.

Im Falle linear homogener Funktionen stimmen Annäherung und exakter Wert überein.

19. Kettenregel, totale Ableitung

Für Funktionen einer Variablen besagt die Kettenregel $[f(g(x))]' = f'(g(x)) \cdot g'(x)$ oder in anderer Schreibweise $\frac{df}{dx} = \frac{df}{dg} \cdot \frac{dg}{dx}$. Analoge Aussagen erhält man auch für Funktionen mehrerer Variablen.

Regeln

(1) Sei $f: \mathbb{R}^m \rightarrow \mathbb{R}$ eine Funktion des Variablenvektors $\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}$, wobei die Komponenten x_i

wiederum von einer Variablen t abhängen, also $x_i = x_i(t)$. Die Funktion f hängt im Grunde nur von dieser Variablen t ab. Die totale Ableitung von f nach t errechnet sich dann aus

$$f'(t) = \frac{df}{dt} = \frac{\partial f}{\partial x_1} \cdot \frac{dx_1}{dt} + \dots + \frac{\partial f}{\partial x_m} \cdot \frac{dx_m}{dt} = f_1 \cdot x_1' + \dots + f_m \cdot x_m'$$

$$\text{In Vektorschreibweise: } \frac{df}{dt} = \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_m} \right) \cdot \begin{pmatrix} \frac{dx_1}{dt} \\ \vdots \\ \frac{dx_m}{dt} \end{pmatrix} = (\text{grad}_f(\mathbf{x}))' \cdot \mathbf{J}_x(t) = \frac{df}{d\mathbf{x}} \cdot \frac{d\mathbf{x}}{dt}.$$

(2) Hängt nun jede Komponente x_i nicht nur von einer, sondern von K Variablen u_1, \dots, u_K ab, so

ergibt sich die totale partielle Ableitung von f nach u_k aus $\frac{\partial f}{\partial u_k} = \frac{\partial f}{\partial x_1} \cdot \frac{\partial x_1}{\partial u_k} + \dots + \frac{\partial f}{\partial x_m} \cdot \frac{\partial x_m}{\partial u_k}$.

$$\text{In Vektorschreibweise: } \frac{\partial f}{\partial u_k} = \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_m} \right) \cdot \begin{pmatrix} \frac{\partial x_1}{\partial u_k} \\ \vdots \\ \frac{\partial x_m}{\partial u_k} \end{pmatrix} = (\text{grad}_f(\mathbf{x}))' \cdot \begin{pmatrix} \frac{\partial x_1}{\partial u_k} \\ \vdots \\ \frac{\partial x_m}{\partial u_k} \end{pmatrix}.$$

Man fasst die totalen partiellen Ableitungen von f nach u_1, \dots, u_K zu einem Zeilenvektor

$$\text{zusammen und erhält } \frac{df}{d\mathbf{u}} = \left(\frac{\partial f}{\partial u_1}, \dots, \frac{\partial f}{\partial u_K} \right) = (\text{grad}_f(\mathbf{x}))' \cdot \mathbf{J}_x(\mathbf{u}) = \frac{df}{d\mathbf{x}} \cdot \frac{d\mathbf{x}}{d\mathbf{u}}.$$

(3) Betrachtet man schließlich eine Funktion $f: \mathbb{R}^m \rightarrow \mathbb{R}^n$,

$$\text{also } \mathbf{f}(\mathbf{x}) = \begin{pmatrix} f_1(\mathbf{x}) \\ \vdots \\ f_n(\mathbf{x}) \end{pmatrix}, \text{ mit } \mathbf{x} = \begin{pmatrix} x_1(\mathbf{u}) \\ \vdots \\ x_m(\mathbf{u}) \end{pmatrix}, \quad \mathbf{u} = \begin{pmatrix} u_1 \\ \vdots \\ u_K \end{pmatrix},$$

$$\text{so erhält man die Funktionalmatrix } \frac{d\mathbf{f}}{d\mathbf{u}} \text{ durch } \frac{d\mathbf{f}}{d\mathbf{u}} = \mathbf{J}_f(\mathbf{u}) = \mathbf{J}_f(\mathbf{x}) \cdot \mathbf{J}_x(\mathbf{u}) = \frac{d\mathbf{f}}{d\mathbf{x}} \cdot \frac{d\mathbf{x}}{d\mathbf{u}}.$$

Beispiele

(1) $f(x, y) = (2x + y^2)^3$ mit $x(t) = e^t$, $y(t) = \ln(t)$. Die Funktion f hängt prinzipiell nur von t ab. Mit $f_x = 3(2x + y^2)^2 \cdot 2 = 6(2x + y^2)^2$, $f_y = 3(2x + y^2)^2 \cdot 2y = 6y(2x + y^2)^2$, $x'(t) = e^t$, $y'(t) = \frac{1}{t}$ erhält man die totale Ableitung von f nach t :

$$f'(t) = \frac{df}{dt}(t) = \frac{\partial f}{\partial x} \cdot \frac{dx}{dt} + \frac{\partial f}{\partial y} \cdot \frac{dy}{dt} = f_x \cdot x' + f_y \cdot y' = 6(2x + y^2)^2 \cdot e^t + 6y(2x + y^2)^2 \cdot \frac{1}{t}.$$

Auch die 2. Ableitung kann analog bestimmt werden: $f''(t) = \frac{d^2f}{dt^2}(t) = \frac{\partial f'}{\partial x} \cdot \frac{dx}{dt} + \frac{\partial f'}{\partial y} \cdot \frac{dy}{dt} + \frac{\partial f'}{\partial t}$

(2) $f(x, y, z) = x^2 + 2yz$ mit $x(s, t) = s - t$, $y(s, t) = s \cdot t$, $z(s, t) = s^2 + t^2$. Die Funktion f hängt also von s und t ab. Mit $f_x = 2x$, $f_y = 2z$, $f_z = 2y$, $x_s = 1$, $x_t = -1$, $y_s = t$, $y_t = s$, $z_s = 2s$, $z_t = 2t$ erhält man die totalen partiellen Ableitungen:

$$\frac{\partial f}{\partial s} = \frac{\partial f}{\partial x} \cdot \frac{\partial x}{\partial s} + \frac{\partial f}{\partial y} \cdot \frac{\partial y}{\partial s} + \frac{\partial f}{\partial z} \cdot \frac{\partial z}{\partial s} = f_x \cdot x_s + f_y \cdot y_s + f_z \cdot z_s = 2x \cdot 1 + 2z \cdot t + 2y \cdot 2s,$$

$$\frac{\partial f}{\partial t} = \frac{\partial f}{\partial x} \cdot \frac{\partial x}{\partial t} + \frac{\partial f}{\partial y} \cdot \frac{\partial y}{\partial t} + \frac{\partial f}{\partial z} \cdot \frac{\partial z}{\partial t} = f_x \cdot x_t + f_y \cdot y_t + f_z \cdot z_t = 2x \cdot (-1) + 2z \cdot s + 2y \cdot 2t.$$

In Matrixschreibweise: $\text{grad}_f(x, y, z) = \begin{pmatrix} 2x \\ 2z \\ 2y \end{pmatrix}$, $\mathbf{J}(s, t) = \begin{pmatrix} 1 & -1 \\ t & s \\ 2s & 2t \end{pmatrix}$, also $\begin{pmatrix} \frac{\partial f}{\partial s} \\ \frac{\partial f}{\partial t} \end{pmatrix} = \text{grad}_f \cdot \mathbf{J}$.

20. Ableiten impliziter Funktionen

Implizite Funktion der Variablen x_1, \dots, x_m : Funktionaler Zusammenhang zwischen den Variablen, der durch eine Gleichung gegeben ist, aber evtl. nicht explizit dargestellt werden kann.

Beispiel: (1) $2x + y = 3$ (explizite Darstellung: $y = -2x + 3$),

(2) $x^5y^2 - xy^5 + x^2 - y = 1$ (nicht explizit anzugeben, da nicht auflösbar nach y).

Die Lösungsmenge einer solchen Gleichung stellt eine Kurve im \mathbb{R}^m (hier im \mathbb{R}^2) dar.

Dazu betrachtet man die Funktion

$$f(x, y) = x^5y^2 - xy^5 + x^2 - y - 1$$

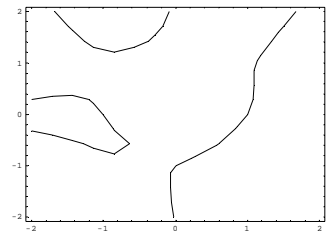
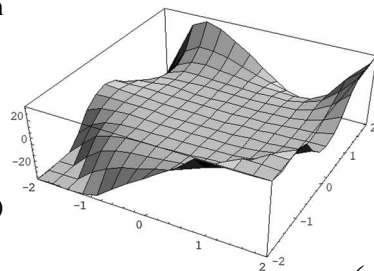
und den Schnitt durch dieses

„Gebirge“ auf der Höhe 0

(siehe Grafiken).

Die Punkte $(0, -1)$, $(-1, 0)$, $(1, 0)$

gehören z.B. zur Lösungsmenge.



Ziel

Bestimmen der Steigung $\frac{\partial x_i}{\partial x_j}(\mathbf{x}_0)$ in einem beliebigen Punkt $\mathbf{x}_0 = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}$ der Kurve.

Regel

Satz über implizite Funktionen: Eine Gleichung mit m Variablen sei durch die Funktion $f(\mathbf{x}) = 0$ gegeben. In einer Umgebung eines Punktes \mathbf{x}_0 mit $f(\mathbf{x}_0) = 0$ sei f stetig und besitze dort stetige partielle Ableitungen $f_j(\mathbf{x})$, $j = 1, \dots, m$. Für eine Variable x_i sei $f_i(\mathbf{x}_0) \neq 0$.

Dann existiert in einer Umgebung von \mathbf{x}_0 eine eindeutig bestimmte implizite Funktion

$$x_i = g(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_m) \quad \text{mit} \quad \frac{\partial g}{\partial x_j}(\mathbf{x}_0) = \frac{\partial x_i}{\partial x_j}(\mathbf{x}_0) = -\frac{f_j(\mathbf{x}_0)}{f_i(\mathbf{x}_0)}.$$

Speziell für eine Gleichung mit 2 Variablen, gegeben durch $f(x, y) = 0$, existiert eine Funktion

$$y = g(x) \quad \text{mit der Steigung} \quad \frac{dy}{dx}(x_0, y_0) = -\frac{f_x(x_0, y_0)}{f_y(x_0, y_0)}.$$

Begründung

der Ableitungsregel. Wegen $f(\mathbf{x}) = 0$ ist auch die totale partielle Ableitung von f nach einer Variablen x_j gleich 0. Diese totale partielle Ableitung

erhält man mit Hilfe der Kettenregel: $\frac{\partial f}{\partial x_j} + \frac{\partial f}{\partial x_i} \cdot \frac{\partial x_i}{\partial x_j}$.

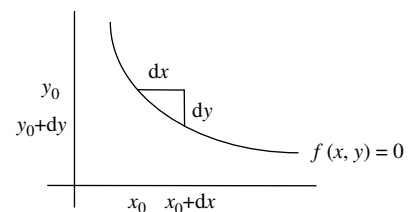
0-Setzen und Umstellen führt dann zu obiger Regel.

Im Fall von 2 Variablen ist auch die grafische Begründung

hilfreich zum Verständnis (siehe Grafik): Betrachte den

Schnitt durch die Funktion $f(x, y)$ auf der Höhe 0 (Höhenlinie $f(x, y) = 0$) und 2 Punkte

(x_0, y_0) und $(x_0 + dx, y_0 + dy)$ auf dieser Kurve. Wegen $f(x_0, y_0) = 0 = f(x_0 + dx, y_0 + dy)$ ist auch



$f(x_0 + dx, y_0 + dy) - f(x_0, y_0) = 0$. Andererseits ist die linke Seite näherungsweise gleich dem totalen Differential df , also $df = f_x(x_0, y_0) \cdot dx + f_y(x_0, y_0) \cdot dy \approx 0 \Leftrightarrow \frac{dy}{dx} \approx -\frac{f_x(x_0, y_0)}{f_y(x_0, y_0)}$.

Für $dx \rightarrow 0$ erhält man die Aussage.

Bemerkung Der Satz über implizite Funktionen gilt nicht nur für $f(x) = 0$, sondern auch für $f(x) = c$, $c \in \mathbb{R}$ (Begründung wie oben mit $f(x, y) = c$ statt $f(x, y) = 0$).

Beispiel (1) $2x + y = 3 \Leftrightarrow f(x, y) = 2x + y - 3 = 0$, $f_x(x, y) = 2$, $f_y(x, y) = 1$,
 also $\frac{dy}{dx}(x, y) = -\frac{2}{1} = -2 \quad \forall (x, y) \in \mathbb{R}^2$. Dies erkennt man auch sofort an der Darstellung $y = -2x + 3$ (Gerade mit Steigung -2).

(2) $x^5y^2 - xy^5 + x^2 - y = 1 \Leftrightarrow f(x, y) = x^5y^2 - xy^5 + x^2 - y - 1 = 0$, $f_x(x, y) = 5x^4y^2 - y^5 + 2x$,
 $f_y(x, y) = 2x^5y - 5xy^4 - 1$, also $\frac{dy}{dx}(x, y) = -\frac{5x^4y^2 - y^5 + 2x}{2x^5y - 5xy^4 - 1}$.

Beispielsweise im Punkt $(0, -1)$ ist die Steigung $\frac{dy}{dx}(0, -1) = -\frac{-(-1)}{-1} = 1$.

Anwendung Betrachtet sei die Produktionsfunktion $f(x, y) = 2x^{1/3}y^{2/3}$ mit den beiden substitutionalen Produktionsfaktoren x und y . Die aktuellen Einsatzmengen betragen $(x_0, y_0) = (8, 64)$ und führen zu einem Output von $f(x_0, y_0) = 2 \cdot 2 \cdot 16 = 64$.

Frage: Wie viele Einheiten von y können (näherungsweise) durch den zusätzlichen Einsatz einer Einheit von x substituiert (ersetzt) werden, wenn die Produktion konstant bleiben soll?

Lösung: Verwende als Annäherung für das gesuchte Austauschverhältnis die Steigung der Isoquanten $2x^{1/3}y^{2/3} = 64$ im Punkt $(8, 64)$. Mit $f_x(x, y) = \frac{2}{3}x^{-2/3}y^{2/3}$, $f_y(x, y) = \frac{4}{3}x^{1/3}y^{-1/3}$, $\frac{f_x(x, y)}{f_y(x, y)} = \frac{1}{2} \cdot \frac{y}{x}$

erhält man $\frac{dy}{dx}(8, 64) = -\frac{1}{2} \cdot \frac{64}{8} = -4$.

Die Interpretation dieses Ausdrucks ist einfach zu merken, wenn man dafür schreibt $dy = -4 dx$.

Setzt man $dx = 1$, also 1 Einheit von x zusätzlich ein, so ist $dy = -4$, d.h. es können dafür 4 Einheiten von y eingespart werden (die Schreibweise ist vom Differenzial her bekannt).

Der positive Wert des Austauschverhältnisses, also $-\frac{dy}{dx} = 4$, wird im Bereich der

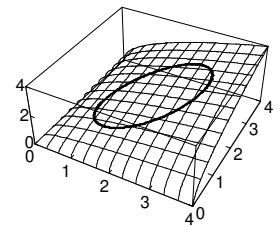
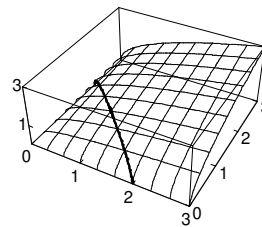
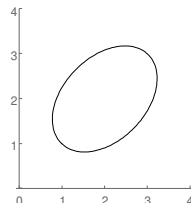
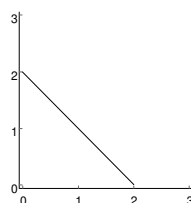
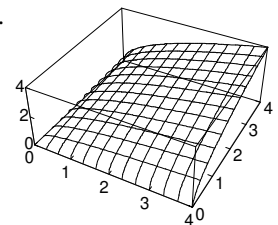
Wirtschaftswissenschaften als *Grenzrate der Substitution* bezeichnet.

21. Extrema unter Nebenbedingungen: Lagrange-Ansatz

Ziel Bestimmen lokaler Extrema einer Funktion $f: \mathbb{R}^m \rightarrow \mathbb{R}$ unter Beachtung von Nebenbedingungen, die als Gleichungen vorliegen.

Beispiel Gesucht sind lokale Extrema der Funktion $f(x, y) = \sqrt{x \cdot y} = x^{1/2} \cdot y^{1/2}$ unter Beachtung einer Nebenbedingung wie (a) $x + y = 2$ oder (b) $-x^2 - y^2 + 2x + 2y + xy = 3$. Der Graf der Funktion f stellt ein 3-dimensionales Gebirge dar (Grafik 1), die Lösungsmengen der Nebenbedingungen „Kurven“ in der x - y -Ebene, (a) eine Gerade (Grafik 2), (b) eine Ellipse (Grafik 3).

Jedem Punkt der Kurve wird der entsprechende Funktionswert von f zugeordnet (Grafik 4 und Grafik 5). Das Ziel besteht nun darin, von allen Punkten auf den Kurven denjenigen mit dem größten bzw. kleinsten Funktionswert zu finden. Man könnte sich z.B. vorstellen, dass der höchste Punkt auf einem Weg durch ein Gebirge gesucht wird.



Alternative Lösungsmöglichkeiten: *Variablensubstitution* oder *Lagrange-Ansatz*

Variablensubstitution Nur anwendbar, falls die Nebenbedingung(en) nach einer Variablen aufzulösen ist.

Beispiel (a): $f(x, y) = \sqrt{x \cdot y}$, Nebenbedingung $x + y = 2$.

- Löse die Nebenbedingung nach einer Variablen auf, z.B. nach y : $y = 2 - x$.
- Setze diese Gleichung in die Funktion f ein: $f(x, y) = \sqrt{x \cdot y} = \sqrt{x \cdot (2 - x)} = (2x - x^2)^{1/2} =: h(x)$.
- Bestimme wie üblich die lokalen Extrema der Funktion h :

$$h'(x) = \frac{1}{2}(2x - x^2)^{-1/2} \cdot (2 - 2x) = (2x - x^2)^{-1/2} \cdot (1 - x) = 0 \Leftrightarrow x = 1 \Rightarrow y = 1.$$

$$h''(x) = -\frac{1}{2}(2x - x^2)^{-3/2} \cdot (2 - 2x) \cdot (1 - x) + (2x - x^2)^{-1/2} \cdot (-1), \quad h''(1) = -1 < 0.$$

$x = 1$ ist also ein lokales Maximum von h und damit auch $(x, y) = (1, 1)$ ein lokales Maximum (hier auch globales Maximum) von f unter Beachtung der Nebenbedingung $x + y = 2$. Da f nur für $x, y \geq 0$ definiert ist, wird das Minimum von f unter der Nebenbedingung an den Randpunkten $(2, 0)$ bzw. $(0, 2)$ erreicht.

Lagrange-Ansatz Beispiel (b): $f(x, y) = \sqrt{x \cdot y}$, Nebenbedingung $-x^2 - y^2 + 2x + 2y + xy = 3$.

Die Nebenbedingung kann hier nur aufwändig nach x oder y umgestellt werden.

Schreibt man die Nebenbedingung als implizite Funktion $g(x, y) = -x^2 - y^2 + 2x + 2y + xy - 3 = 0$, so besagt der Satz über implizite Funktionen, dass in jedem Punkt (x_0, y_0) der Ellipse eine Funktion

$$y(x) \text{ existiert mit der Steigung } \frac{dy}{dx}(x_0, y_0) = -\frac{g_x(x_0, y_0)}{g_y(x_0, y_0)}.$$

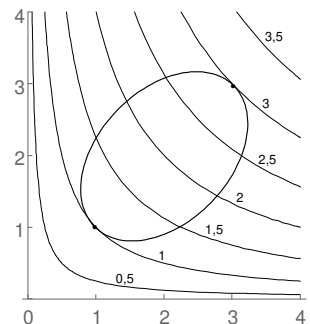
Setzt man in die Funktion f für y diese Funktion $y(x)$ ein und leitet sie ab, so erhält man die totale Ableitung $\frac{df}{dx} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} \cdot \frac{dy}{dx} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} \cdot \left(-\frac{g_x}{g_y}\right) = f_x - f_y \cdot \frac{g_x}{g_y}$. Setzt man die Ableitung gleich 0, so

erhält man die notwendige Bedingung für das Vorliegen eines lokalen Extremums im Punkt (x_0, y_0) :

$$-\frac{f_x(x_0, y_0)}{f_y(x_0, y_0)} = -\frac{g_x(x_0, y_0)}{g_y(x_0, y_0)} \quad (1).$$

Die Bedingung lässt sich auch grafisch gut herleiten (Grafik):

Außer der Ellipse (Nebenbedingung) sind einige Höhenlinien eingezeichnet, also Linien von Punkten mit gleichem Funktionswert $f(x, y) = c$ (hier c zwischen 0,5 und 3,5). Man erkennt unschwer, dass im Punkt $(3, 3)$ die Höhenlinie 3 und damit der höchste Punkt auf der Ellipse erreicht wird (lokales Maximum von f unter der Nebenbedingung), während im Punkt $(1, 1)$ der tiefste Punkt auf



der Höhenlinie 1 liegt (lokales Minimum von f unter der Nebenbedingung). Lokalen Extrema unter Nebenbedingungen findet man somit an Stellen, an denen der Graf der Nebenbedingung eine Höhenlinie berührt. An diesen Punkten muss die Steigung von Graf und Höhenlinie gleich sein (sonst würden sie sich schneiden). Die Steigung in einem beliebigen Punkt (x_0, y_0) einer Höhenlinie

(= implizite Funktion $f(x, y) = c$) ist

$$\frac{dy}{dx}(x_0, y_0) = -\frac{f_x(x_0, y_0)}{f_y(x_0, y_0)}, \quad \text{die der Nebenbedingung } \frac{dy}{dx}(x_0, y_0) = -\frac{g_x(x_0, y_0)}{g_y(x_0, y_0)}.$$

Gleichsetzen führt zu Bedingung (1).

$$\text{Die Bedingung (1) formt man leicht um zu } -\frac{f_x(x_0, y_0)}{g_x(x_0, y_0)} = -\frac{f_y(x_0, y_0)}{g_y(x_0, y_0)} \quad (2).$$

Der Lagrange-Ansatz beruht nun darauf, eine neue Funktion (die Lagrange-Funktion) so zu konstruieren, dass sie beim 0-Setzen ihrer partiellen Ableitungen genau zur Bedingung (2) führt.

Die Vorgehensweise ist dann wie folgt:

- Schreibe die Nebenbedingung um zu $g(x, y) = 0$ (siehe oben).
- Definiere die Lagrange-Funktion L durch $L(x, y, \lambda) = f(x, y) + \lambda \cdot g(x, y)$, hier: $L(x, y, \lambda) = x^{1/2} \cdot y^{1/2} + \lambda \cdot (-x^2 - y^2 + 2x + 2y + xy - 3)$.

Die zusätzliche Variable λ bezeichnet man als Lagrangeschen Multiplikator.

Ebenfalls möglich ist auch die Definition $L(x, y, \lambda) = f(x, y) - \lambda \cdot g(x, y)$.

- Ermittle die partiellen Ableitungen 1. Ordnung von L nach x und y :

$$L_x(x, y, \lambda) = f_x + \lambda \cdot g_x = \frac{1}{2}x^{-1/2} \cdot y^{1/2} + \lambda \cdot (-2x + 2 + y),$$

$$L_y(x, y, \lambda) = f_y + \lambda \cdot g_y = \frac{1}{2}x^{1/2} \cdot y^{-1/2} + \lambda \cdot (-2y + 2 + x).$$

Die partielle Ableitung nach λ muss nicht gebildet werden, denn es gilt stets $L_\lambda(x, y, \lambda) = g(x, y)$.

- Bestimme auf folgende Weise die kritischen Punkte von L :

- (1) Die beiden partiellen Ableitungen gleich 0 setzen und jeweils nach λ auflösen:

$$L_x(x, y, \lambda) = 0 \Leftrightarrow f_x + \lambda \cdot g_x = 0 \Rightarrow \lambda = -\frac{f_x}{g_x}, \quad \text{hier } \lambda = -\frac{1/2 \cdot x^{-1/2} \cdot y^{1/2}}{-2x + 2 + y}, \quad (*)$$

$$L_y(x, y, \lambda) = 0 \Leftrightarrow f_y + \lambda \cdot g_y = 0 \Rightarrow \lambda = -\frac{f_y}{g_y}, \quad \text{hier } \lambda = -\frac{1/2 \cdot x^{1/2} \cdot y^{-1/2}}{-2y + 2 + x}. \quad (**)$$

- (2) Die beiden Ausdrücke gleichsetzen (führt zur Bedingung (2)) und nach x oder y auflösen:

$$-\frac{f_x}{g_x} = -\frac{f_y}{g_y} \Rightarrow f_x \cdot g_y = f_y \cdot g_x, \quad \text{hier } x^{-1/2} \cdot y^{1/2} \cdot (-2y + 2 + x) = x^{1/2} \cdot y^{-1/2} \cdot (-2x + 2 + y)$$

$$\Rightarrow y \cdot (-2y + 2 + x) = x \cdot (-2x + 2 + y) \Leftrightarrow -2y^2 + 2y = -2x^2 + 2x \Leftrightarrow y^2 - y = x^2 - x$$

$$\Leftrightarrow y^2 - y + \frac{1}{4} = x^2 - x + \frac{1}{4} \Leftrightarrow (y - \frac{1}{2})^2 = (x - \frac{1}{2})^2 \Leftrightarrow |y - \frac{1}{2}| = |x - \frac{1}{2}|$$

$$\Leftrightarrow y = x \quad \text{oder} \quad y = 1 - x. \quad (***)$$

- (3) Den gefundenen Zusammenhang in die Nebenbedingung einsetzen:

$$\text{Fall (i) } y = x: \quad -x^2 - x^2 + 2x + 2x + x^2 = 3 \Leftrightarrow x^2 - 4x + 3 = (x - 1) \cdot (x - 3) = 0$$

$$\Leftrightarrow x_0 = 1 \quad \text{oder} \quad x_1 = 3 \Rightarrow y_0 = 1, \quad y_1 = 3;$$

$$\text{Fall (ii) } y = 1 - x: \quad -x^2 - (1 - x)^2 + 2x + 2(1 - x) + x(1 - x) = 3$$

$$\Leftrightarrow -3x^2 + 3x = 2 \Leftrightarrow x^2 - x + \frac{2}{3} = 0 \quad (\text{keine Lösung in } \mathbb{R});$$

insgesamt also 2 Punkte $(x_0, y_0) = (1, 1)$ und $(x_1, y_1) = (3, 3)$.

- (4) Die zugehörigen Werte für λ aus der Bedingung (*) oder (**) berechnen:

$$\lambda_0 = -\frac{1}{2}, \quad \lambda_1 = \frac{1}{2}, \quad \text{also kritische Punkte } (1, 1, -\frac{1}{2}) \quad \text{und} \quad (3, 3, \frac{1}{2}).$$

Um zu ermitteln, ob es sich um ein Minimum oder Maximum handelt, benötigt man die totale

Ableitung 2. Ordnung $\frac{d^2 f}{dx^2}$. Man rechnet nach (etwas aufwändig), dass an einem kritischen Punkt

(x, y, λ) gilt: $\frac{d^2 f}{dx^2}(x, y, \lambda) > 0 \Leftrightarrow |\mathbf{H}_L(x, y, \lambda)| < 0$. Dabei bezeichnet \mathbf{H}_L die Hesse-Matrix

der Lagrange-Funktion. Für $|\mathbf{H}_L(x, y, \lambda)| < 0$ liegt also ein lokales Minimum vor,

für $|\mathbf{H}_L(x, y, \lambda)| > 0$ ein lokales Maximum.

- Bestimme die symmetrische Hesse-Matrix $\mathbf{H}_L(x, y, \lambda)$

$$\mathbf{H}_L(x, y, \lambda) = \begin{pmatrix} L_{xx}(x, y, \lambda) & L_{xy}(x, y, \lambda) & L_{x\lambda}(x, y, \lambda) \\ L_{yx}(x, y, \lambda) & L_{yy}(x, y, \lambda) & L_{y\lambda}(x, y, \lambda) \\ L_{\lambda x}(x, y, \lambda) & L_{\lambda y}(x, y, \lambda) & L_{\lambda\lambda}(x, y, \lambda) \end{pmatrix} = \begin{pmatrix} L_{xx}(x, y, \lambda) & L_{xy}(x, y, \lambda) & g_x(x, y) \\ L_{yx}(x, y, \lambda) & L_{yy}(x, y, \lambda) & g_y(x, y) \\ g_x(x, y) & g_y(x, y) & 0 \end{pmatrix};$$

es gilt stets $L_{x\lambda}(x, y, \lambda) = g_x(x, y)$, $L_{y\lambda}(x, y, \lambda) = g_y(x, y)$ und $L_{\lambda\lambda}(x, y, \lambda) = 0$.

In diesem Beispiel ist $g_x(x, y) = -2x + 2 + y$, $g_y(x, y) = -2y + 2 + x$ (siehe oben),

$$L_{xx}(x, y, \lambda) = -\frac{1}{4}x^{-3/2} \cdot y^{1/2} - 2\lambda, \quad L_{xy}(x, y, \lambda) = \frac{1}{4}x^{-1/2} \cdot y^{-1/2} + \lambda, \quad L_{yy}(x, y, \lambda) = -\frac{1}{4}x^{1/2} \cdot y^{-3/2} - 2\lambda.$$

- Berechne die Determinante der Hesse-Matrix an den kritischen Punkten; wegen $x_i = y_i$ ist hier

$$L_{xx} = -\frac{1}{4x_i} - 2\lambda_i = L_{yy}, \quad L_{xy} = \frac{1}{4x_i} + \lambda_i, \quad g_x = 2 - x_i = g_y,$$

$$|\mathbf{H}_L(1, 1, -\frac{1}{2})| = \left| \begin{pmatrix} 3/4 & -1/4 & 1 \\ -1/4 & 3/4 & 1 \\ 1 & 1 & 0 \end{pmatrix} \right| = -2 < 0, \quad |\mathbf{H}_L(3, 3, \frac{1}{2})| = \left| \begin{pmatrix} -13/12 & 7/12 & -1 \\ 7/12 & -13/12 & -1 \\ -1 & -1 & 0 \end{pmatrix} \right| = \frac{10}{3} > 0$$

\Rightarrow im Punkt $(1, 1)$ liegt ein lokales Minimum von f unter der Nebenbedingung $g = 0$ vor,

im Punkt $(3, 3)$ ein lokales Maximum.

Bemerkung: Man rechnet schnell nach, dass

$$|\mathbf{H}_L(x, y, \lambda)| = 2g_x(x, y) \cdot g_y(x, y) \cdot L_{xy}(x, y, \lambda) - g_x^2(x, y) \cdot L_{yy}(x, y, \lambda) - g_y^2(x, y) \cdot L_{xx}(x, y, \lambda).$$

Merken

- (1) Während bei Funktionen ohne Nebenbedingung die Hesse-Matrix auf Definitheit überprüft wird, muss beim Lagrange-Ansatz nur ihre Determinante auf das Vorzeichen hin untersucht werden.

- (2) Für Funktionen mit mehr als 2 Variablen und / oder mehr als einer Nebenbedingung gestaltet sich das Überprüfen der hinreichenden Kriterien beim Lagrange-Ansatz erheblich aufwendiger, wie die folgende Regel zeigt.

Regel Es seien die Funktionen $f: \mathbb{R}^n \rightarrow \mathbb{R}$ und $g: \mathbb{R}^n \rightarrow \mathbb{R}^m$ gegeben mit $m < n$ und $g(x) = \begin{pmatrix} g_1(x) \\ \vdots \\ g_m(x) \end{pmatrix}$,

$x \in \mathbb{R}^n$. Dann gilt:

- (1) *Notwendiges Kriterium*

Notwendig für das Vorliegen eines lokalen Extremums von f unter Berücksichtigung der

Nebenbedingungen $g_1(x) = 0, \dots, g_m(x) = 0$ ist $\text{grad}_L(x, \lambda) = \mathbf{0}$, $\lambda = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_m \end{pmatrix}$, wobei

$L(x, \lambda) = f(x) + \lambda_1 g_1(x) + \dots + \lambda_m g_m(x)$ als *Lagrange-Funktion* bezeichnet wird.

Die λ_i nennt man *Lagrangesche Multiplikatoren*.

- (2) *Hinreichendes Kriterium*

(x_0, λ_0) sei ein kritischer Punkt der Lagrange-Funktion, $\lambda_0 = \begin{pmatrix} \lambda_1^0 \\ \vdots \\ \lambda_m^0 \end{pmatrix}$.

f besitzt dann in x_0 unter den Nebenbedingungen $g(x_0) = \mathbf{0}$ ein lokales

- *Maximum*, wenn die Determinanten der "geränderten" Hesse-Matrix $\begin{vmatrix} \mathbf{H}_L^{ii}(x_0, \lambda_0) & (\mathbf{J}_g^{mi}(x_0))' \\ \mathbf{J}_g^{mi}(x_0) & \mathbf{0} \end{vmatrix}$

für $i = m+1, \dots, n$ wechselndes Vorzeichen aufweisen, beginnend mit $(-1)^{m+1}$,

- *Minimum*, wenn die Determinanten alle das Vorzeichen $(-1)^m$ besitzen.

Dabei ist $\mathbf{H}_L(x_0, \lambda_0) = \mathbf{H}_f(x_0) + \lambda_1^0 \mathbf{H}_{g_1}(x_0) + \dots + \lambda_m^0 \mathbf{H}_{g_m}(x_0)$ die Matrix der partiellen

Ableitungen 2. Ordnung der Lagrange-Funktion nach den Variablen x_1, \dots, x_n im Punkt

(x_0, λ_0) , $\mathbf{J}_g(x_0)$ die Jacobi-Matrix von $g(x)$ im Punkt x_0 und \mathbf{M}^{kl} der Teil einer Matrix \mathbf{M} ,

welcher aus den ersten k Zeilen und l Spalten besteht.

Beispiele

- (1) Bei einer Funktion mit zwei Variablen und einer Nebenbedingung wie im obigen Beispiel ist $n = 2$, $m = 1$ und daher beim hinreichenden Kriterium nur für $i = 2$ die Determinante der "geränderten" Hesse-Matrix in den kritischen Punkten zu bestimmen, also z.B. für (x_0, y_0, λ_0) die Determinante

$$\begin{vmatrix} \mathbf{H}_L^{22}(x_0, y_0, \lambda_0) & (\mathbf{J}_g^{12}(x_0, y_0))' \\ \mathbf{J}_g^{12}(x_0, y_0) & 0 \end{vmatrix}. \text{ Dabei ist } \mathbf{H}_L^{22}(x, y, \lambda) = \begin{pmatrix} L_{xx}(x, y, \lambda) & L_{xy}(x, y, \lambda) \\ L_{yx}(x, y, \lambda) & L_{yy}(x, y, \lambda) \end{pmatrix} \text{ und} \\ (\mathbf{J}_g^{12}(x, y))' = \text{grad}_g(x, y) = \begin{pmatrix} g_x(x, y) \\ g_y(x, y) \end{pmatrix}, \text{ also } \begin{pmatrix} \mathbf{H}_L^{22}(x, y, \lambda) & (\mathbf{J}_g^{12}(x, y))' \\ \mathbf{J}_g^{12}(x, y) & 0 \end{pmatrix} = \begin{pmatrix} \begin{pmatrix} L_{xx} & L_{xy} \\ L_{yx} & L_{yy} \end{pmatrix} & \begin{pmatrix} g_x \\ g_y \end{pmatrix} \\ \begin{pmatrix} g_x & g_y \end{pmatrix} & 0 \end{pmatrix}.$$

Im Falle einer Funktion mit 2 Variablen und einer Nebenbedingung stimmt also die zu untersuchende "geränderte" Hesse-Matrix mit der "normalen" Hesse-Matrix der Lagrange-Funktion überein.

- (2) Bei einer Funktion $f(x, y, z)$ mit drei Variablen und einer Nebenbedingung $g(x, y, z) = 0$ ist $n = 3$, $m = 1$ und daher sind beim hinreichenden Kriterium für $i = m+1, \dots, n$, also für die beiden Werte $i = 2$ und $i = 3$ "geränderte" Hesse-Matrizen zu berechnen.

Die Lagrange-Funktion lautet $L(x, y, z, \lambda) = f(x, y, z) + \lambda \cdot g(x, y, z)$.

$$\text{Ihre Hesse-Matrix ist } \mathbf{H}_L = \begin{pmatrix} L_{xx} & L_{xy} & L_{xz} & L_{x\lambda} \\ L_{yx} & L_{yy} & L_{yz} & L_{y\lambda} \\ L_{zx} & L_{zy} & L_{zz} & L_{z\lambda} \\ L_{\lambda x} & L_{\lambda y} & L_{\lambda z} & L_{\lambda\lambda} \end{pmatrix} = \begin{pmatrix} L_{xx} & L_{xy} & L_{xz} & g_x \\ L_{yx} & L_{yy} & L_{yz} & g_y \\ L_{zx} & L_{zy} & L_{zz} & g_z \\ g_x & g_y & g_z & 0 \end{pmatrix} = \begin{pmatrix} \mathbf{H}_L^{33} & (\mathbf{J}_g^{13})' \\ \mathbf{J}_g^{13} & 0 \end{pmatrix}$$

mit der Jacobi-Matrix $\mathbf{J}_g = (g_x, g_y, g_z) = \mathbf{J}_g^{13}$. Die Matrix \mathbf{H}_L entspricht der für $i = 3$ zu

berechnenden "geränderten" Hesse-Matrix. Für $i = 2$ ergibt sich die Matrix

$$\begin{pmatrix} \mathbf{H}_L^{22} & (\mathbf{J}_g^{12})' \\ \mathbf{J}_g^{12} & 0 \end{pmatrix} = \begin{pmatrix} L_{xx} & L_{xy} & g_x \\ L_{yx} & L_{yy} & g_y \\ g_x & g_y & 0 \end{pmatrix}. \text{ Die Regel besagt dann: Besitzen die Determinanten beider}$$

Matrizen an einem kritischen Punkt ein negatives Vorzeichen ($(-1)^m = (-1)^1 = -1$), so liegt dort ein

lokales Minimum vor; ist die Determinante der kleineren Matrix ($i = 2$) positiv und die der größeren Matrix ($i = 3$) negativ, so handelt es sich dort um ein lokales Maximum.

Abgrenzung zu anderen Verfahren: Liegen die Nebenbedingungen in Ungleichungsform vor, so verwendet man bei

- linearer Zielfunktion und linearen Nebenbedingungen Verfahren aus der *linearen Optimierung*,
- konvexer Zielfunktion und konvexen Nebenbedingungen den *Satz von Kuhn-Tucker*.

22. Integralrechnung

Ziel (1) Umkehren des Differenzierens (*unbestimmtes Integral*)
 (2) Flächen- bzw. Volumensberechnung (*bestimmtes Integral*)

Unbestimmtes Integral: $\int f(x) dx = \{ F(x) \mid F'(x) = f(x) \} =$ Menge aller *Stammfunktionen* F von f .

Anmerkung:

- (1) Zwei Stammfunktionen der gleichen Funktion f unterscheiden sich höchstens um eine additive Konstante (die beim Ableiten wegfällt).
- (2) Nicht jede Funktion besitzt eine elementare Stammfunktion, z.B. $f(x) = e^{x^2}$ nicht. Die Stammfunktion kann aber durch eine unendliche Reihe angegeben werden, indem man f z.B. in eine Taylorreihe entwickelt und die Summanden der Reihe einzeln integriert.

Aus den Ableitungsregeln erhält man sofort die folgenden Regeln zum Ermitteln unbestimmter Integrale. Die Regeln können durch Ableiten leicht überprüft werden.

Regeln F sei im Folgenden eine beliebige Stammfunktion von f , G sei eine beliebige Stammfunktion von g und $c \in \mathbb{R}$ eine beliebige additive Konstante.

Funktion	Unbestimmtes Integral	Funktion	Unbestimmtes Integral
a	$ax + c$	$\sin(x)$	$-\cos(x) + c$
$x^a, a \neq -1$	$\frac{1}{a+1} \cdot x^{a+1} + c$	$\cos(x)$	$\sin(x) + c$
$x^{-1} = 1/x$	$\ln(x) + c$	$f \pm g$	$F \pm G + c$
a^x	$\frac{1}{\ln(a)} \cdot a^x + c$	$a \cdot f$	$a \cdot F + c$
$(a > 0, a \neq 1)$		$f(ax + b)$	$\frac{1}{a} \cdot F(ax + b) + c$
$\log_a(x)$	$x \cdot (\log_a(x) - \frac{1}{\ln(a)}) + c$	$f(g) \cdot g'$	$F(g) + c$
$(a > 0, a \neq 1)$			

Beispiele (die additiven Konstanten $c, c \in \mathbb{R}$, werden zur Erhöhung der Übersichtlichkeit weggelassen).

Funktion	Unbestimmtes Integral	Funktion	Unbestimmtes Integral
$x = x^1$	$\frac{1}{2}x^2$	$\frac{1}{\sqrt{2x+3}} = (2x+3)^{-1/2}$	$\frac{1}{2} \cdot 2 \cdot (2x+3)^{1/2} = \sqrt{2x+3}$
x^2	$\frac{1}{3}x^3$	$(x^3+2x^2+4)^4 \cdot (3x^2+4x)$	$\frac{1}{5}(x^3+2x^2+4)^5$
$\sqrt{x} = x^{1/2}$	$\frac{2}{3}x^{3/2}$	$g(x) = x^3+2x^2+4,$	$F(g(x))$
$\frac{1}{x^2} = x^{-2}$	$-x^{-1} = -\frac{1}{x}$	$g'(x) = 3x^2+4x$	
e^x	e^x	$f(x) = x^4, F(x) = \frac{1}{5}x^5$	
$\ln(x)$	$x \cdot (\ln(x) - 1)$	$\frac{2x}{x^2+1}$	$\ln(x^2+1) = \ln(x^2+1)$
$2x - 3x^2$	$2 \cdot \frac{1}{2}x^2 - 3 \cdot \frac{1}{3}x^3 = x^2 - x^3$	$g(x) = x^2+1, g'(x) = 2x$	$F(g(x))$
e^{2x+3}	$\frac{1}{2}e^{2x+3}$	$f(x) = \frac{1}{x}, F(x) = \ln(x)$	
$\frac{1}{2x+3}$	$\frac{1}{2}\ln(2x+3)$		

Durch Ableiten überprüft man leicht die folgenden Regeln:

Regel

Partielle Integration (Produktregel)

Für das Produkt von zwei Funktionen $f \cdot g$ gilt:

$$\int f \cdot g \, dx = F \cdot g - \int F \cdot g' \, dx = f \cdot G - \int f' \cdot G \, dx.$$

Beispiele:

(1) $\int \ln(x) \, dx = \int 1 \cdot \ln(x) \, dx$, also $f(x) = 1$, $g(x) = \ln(x)$, $F(x) = x$, $g'(x) = \frac{1}{x}$.

Damit ist $\int \ln(x) \, dx = x \cdot \ln(x) - \int x \cdot \frac{1}{x} \, dx = x \cdot \ln(x) - \int 1 \, dx = x \cdot \ln(x) - x$.

(2) $\int x \cdot e^x \, dx$, also $f(x) = x$, $g(x) = e^x$, $G(x) = e^x$, $f'(x) = 1$.

Damit ist $\int x \cdot e^x \, dx = x \cdot e^x - \int 1 \cdot e^x \, dx = x \cdot e^x - e^x = (x-1) \cdot e^x$.

Regel

Integration durch Substitution (Kettenregel; eine Version wurde bereits oben dargestellt)

$$\int f(x) \, dx = \int f(g(t)) \cdot g'(t) \, dt. \quad \text{Dabei ist } x = g(t), \quad t = g^{-1}(x), \quad g'(t) = \frac{dx}{dt}.$$

Beispiel: Zu berechnen ist $\int f(x) \, dx$ mit $f(x) = \frac{1}{\sqrt{1-x^2}}$.

– Ersetze (substituiere) x durch eine geeignete Funktion von t , hier $x = g(t) = \sin(t)$.¹¹

– Bilde die 1. Ableitung von g , hier $g'(t) = \cos(t)$.

– Berechne das Integral $\int f(g(t)) \cdot g'(t) \, dt$, hier $\int f(g(t)) \cdot g'(t) \, dt = \int \frac{1}{\sqrt{1-\sin^2(t)}} \cdot \cos(t) \, dt$.

Wegen $\sin^2(t) + \cos^2(t) = 1$, also $\sqrt{1-\sin^2(t)} = \cos(t)$, erhält man

$$\int f(g(t)) \cdot g'(t) \, dt = \int 1 \, dt = t.$$

– Ersetze t umgekehrt wieder durch x , hier $t = g^{-1}(x) = \arcsin(x)$, und erhalte so das Ergebnis

$$\int f(x) \, dx = \arcsin(x).$$

Bei komplizierteren Integralen verwendet man *Integraltafeln* oder entsprechende Software.

Bestimmtes Integral Sei f stetig in einem Intervall $[a, b]$. Dann ist das bestimmte Integral $\int_a^b f(x) \, dx$ definiert durch

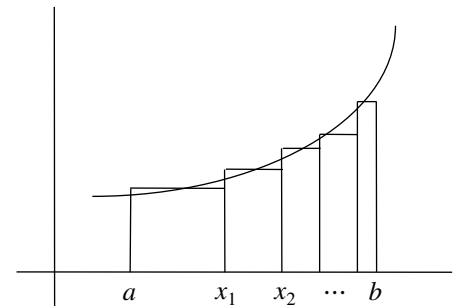
$$\int_a^b f(x) \, dx = \lim_{n \rightarrow \infty} \sum_{i=1}^n f(z_i) \cdot (x_i - x_{i-1}), \text{ wobei } (x_i)_{i \in \mathbb{N}} \text{ eine Zerlegungsfolge von } [a, b] \text{ ist mit der}$$

Eigenschaft $a = x_0 < x_1 < \dots < x_n = b$ und

$$\max_{i=1, \dots, n} (x_i - x_{i-1}) \xrightarrow{n \rightarrow \infty} 0; \quad z_i \in [x_{i-1}, x_i] \text{ ein beliebiger}$$

Punkt. Das bestimmte Integral misst die *Fläche* zwischen dem Grafen von f und der x -Achse im Bereich von a bis b .

Die nachstehenden Regeln ergeben sich direkt aus den Rechenregeln für Folgen.



Regeln

f, g seien stetig im Intervall $[a, b]$. Dann gilt:

(1) $\int_a^a f(x) \, dx = 0$

(2) $\int_b^a f(x) \, dx = - \int_a^b f(x) \, dx$

(3) $\int_a^b f(x) \, dx = \int_a^c f(x) \, dx + \int_c^b f(x) \, dx$ für beliebiges $c \in [a, b]$

(4) $\int_a^b f(x) \, dx \geq 0$, falls $f(x) \geq 0 \, \forall x \in [a, b]$; $\int_a^b f(x) \, dx \leq 0$, falls $f(x) \leq 0 \, \forall x \in [a, b]$

(5) $\int_a^b r \cdot f(x) \, dx = r \cdot \int_a^b f(x) \, dx$ für beliebiges $r \in \mathbb{R}$

(6) $\int_a^b (f(x) \pm g(x)) \, dx = \int_a^b f(x) \, dx \pm \int_a^b g(x) \, dx$

Die nächste Regel klärt nun den Zusammenhang zwischen unbestimmtem und bestimmtem Integral.

¹¹Es erfordert viel Übung zu erkennen, wie substituiert werden soll, denn nicht jede Substitution führt zum Ziel.

Regel

Hauptsatz der Integralrechnung

f sei stetig im Intervall $I = [a, b]$, F eine beliebige Stammfunktion von f in I .

Dann ist $\int_a^b f(x) dx = F(b) - F(a)$. (Statt $F(b) - F(a)$ schreibt man auch $[F(x)]_a^b$.)

Beispiel: $\int_1^2 3x^2 dx = [x^3]_1^2 = 2^3 - 1^3 = 8 - 1 = 7$.

Herleitung

der Regel: Für $c \in [a, b]$ definiert man die *Integralfunktion*

$I(c) = \int_a^c f(x) dx$. Für einen kleinen Wert h ist dann

$$I(c+h) = \int_a^{c+h} f(x) dx = \int_a^c f(x) dx + \int_c^{c+h} f(x) dx = I(c) + \int_c^{c+h} f(x) dx.$$

$\int_c^{c+h} f(x) dx$ misst eine Fläche, die für kleine Werte von h angenähert werden kann durch ein

Rechteck der Breite h und der Höhe $f(c)$, also der Fläche $h \cdot f(c)$. Insgesamt erhält man so

$I(c+h) \approx I(c) + h \cdot f(c)$. Umgestellt nach $f(c)$ ergibt sich $f(c) \approx \frac{I(c+h) - I(c)}{h}$ und damit

für $h \rightarrow 0$: $f(c) = \lim_{h \rightarrow 0} \frac{I(c+h) - I(c)}{h} = I'(c)$. Die Integralfunktion I ist also eine Stammfunktion

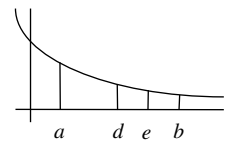
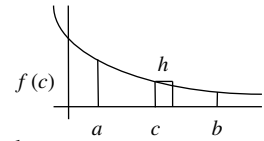
von f . Da 2 Stammfunktionen von f sich höchstens um eine additive Konstante unterscheiden, gibt es zu jeder Stammfunktion F von f eine Konstante r , so dass

$I(x) = F(x) + r$. Betrachtet man nun zwei beliebige Zahlen

$d, e \in [a, b]$, $d < e$, so erhält man die Fläche zwischen Kurve und x -Achse über dem Intervall $[d, e]$, indem man von der Fläche über dem Intervall $[a, e]$ die Fläche über dem Intervall $[a, d]$ abzieht:

$$\int_d^e f(x) dx = \int_a^e f(x) dx - \int_a^d f(x) dx = I(e) - I(d) = (F(e) + r) - (F(d) + r) = F(e) - F(d)$$

für eine beliebige Stammfunktion F von f . Mit $d = a$ und $e = b$ ergibt sich daraus die Aussage.



Uneigentliche Integrale: Als uneigentliche Integrale bezeichnet man bestimmte Integrale, bei dem eine (oder auch beide) Integrationsgrenzen ∞ bzw. $-\infty$ sind oder der Funktionswert $f(x)$ an einer (oder beiden)

Integrationsgrenzen ∞ bzw. $-\infty$ ist. Beispiele:

$$(1) \int_0^\infty e^{-x} dx = \lim_{b \rightarrow \infty} \int_0^b e^{-x} dx = \lim_{b \rightarrow \infty} [-e^{-x}]_0^b = \lim_{b \rightarrow \infty} (-e^{-b}) - (-e^0) = 0 - (-1) = 1.$$

$$(2) \int_{-\infty}^\infty e^{-x^2/2} dx = \lim_{a \rightarrow -\infty} \int_a^0 e^{-x^2/2} dx + \lim_{b \rightarrow \infty} \int_0^b e^{-x^2/2} dx.$$

Obwohl $f(x) = e^{-x^2/2}$ keine elementare Stammfunktion besitzt, kann gezeigt werden, dass beide Grenzwerte gleich $\sqrt{\pi/2}$ sind, also das uneigentliche Integral den Wert $2 \cdot \sqrt{\pi/2} = \sqrt{2\pi}$ aufweist. Die Funktion $f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ ist die Dichtefunktion der Standardnormalverteilung.

$$(3) \int_0^1 \frac{1}{\sqrt{x}} dx = \lim_{a \rightarrow 0} \int_a^1 \frac{1}{\sqrt{x}} dx = \lim_{a \rightarrow 0} [2 \cdot \sqrt{x}]_a^1 = 2 \cdot \sqrt{1} - 2 \cdot \sqrt{0} = 2.$$

Mehrfache (bestimmte) Integrale

Ziel: Volumen von Körpern berechnen.

Beispielsweise bestimmt $\int_a^b \int_c^d f(x, y) dy dx$ das Volumen

zwischen x - y -Ebene und dem Grafen der Funktion f über der rechteckigen Fläche a, b, c, d .

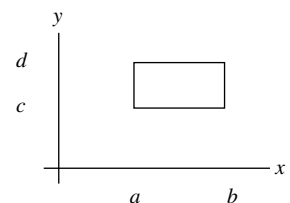
$$\text{Beispiel: } \int_0^1 \int_0^1 -x^2 - y^2 + 4 dy dx = \int_0^1 \left(\int_0^1 -x^2 - y^2 + 4 dy \right) dx.$$

Man berechnet nun zunächst das innere Integral, indem man die Funktion hier bezüglich y integriert und x dabei wie eine Konstante behandelt.

$$\int_0^1 -x^2 - y^2 + 4 dy = \left[-x^2 y - \frac{1}{3} y^3 + 4y \right]_0^1 = (-x^2 \cdot 1 - \frac{1}{3} \cdot 1^3 + 4 \cdot 1) - (-x^2 \cdot 0 - \frac{1}{3} \cdot 0^3 + 4 \cdot 0) = -x^2 + \frac{11}{3}.$$

Dieses Ergebnis des inneren Integrals stellt die zu integrierende Funktion für das äußere Integral dar,

$$\text{also } \int_0^1 -x^2 + \frac{11}{3} dx = \left[-\frac{1}{3} x^3 + \frac{11}{3} x \right]_0^1 = -\frac{1}{3} + \frac{11}{3} = \frac{10}{3}.$$



Regel: Beim Vertauschen der Integrationsreihenfolge müssen auch die Integrationsgrenzen vertauscht werden, d.h. es gilt $\int_a^b \int_c^d f(x, y) dy dx = \int_c^d \int_a^b f(x, y) dx dy$.

23. Differenzialgleichungen

Definition Eine *Differenzialgleichung* (DGL) ist eine Gleichung, die neben einer Funktion auch mindestens eine ihrer Ableitungen enthält. Man unterscheidet DGLen nach den Kriterien

- *gewöhnlich* (bei Funktionen 1 Variablen) oder *partiell* (bei Funktionen mehrerer Variablen),
- *linear* (bei linearer Verknüpfung von Funktion und Ableitung(en)) oder *nichtlinear*,
- *homogen* (ohne Absolutglied) oder *inhomogen*,
- *n-ter Ordnung* (bei Vorliegen von Ableitungen bis maximal der Ordnung n).

Beispiele:

- (1) $a \cdot f(x) + b \cdot f'(x) + c = 0$ ist eine gewöhnliche, lineare, inhomogene DGL 1. Ordnung.
- (2) $f(x) + e^x \cdot f'(x) - f''(x) = 0$ ist eine gewöhnliche, homogene, nichtlineare DGL 2. Ordnung.
- (3) $(f(x, y))^2 + f_y(x, y) + f_{xx}(x, y) = 0$ ist eine partielle, nichtlineare, homogene DGL 2. Ordnung.

Ziel Auffinden von Funktionen, welche der DGL genügen. Oftmals existiert keine entsprechende elementare Funktion, so dass nur eine numerische bzw. grafische Lösung möglich ist.

Anwendung DGLen werden insbesondere häufig zur Beschreibung von Zusammenhängen im Zeitablauf, also zur Modellierung dynamischer Prozesse verwendet (beispielsweise in der Wachstumstheorie). Nur für den einfachen Fall einer gewöhnlichen DGL 1.Ordnung soll die Lösung angegeben werden.

Regel Für eine gewöhnliche DGL 1. Ordnung gilt

$$(1) f'(x) = f(x) \cdot g(x) \Leftrightarrow |f(x)| = e^{\int g(x) dx}$$

$$(2) f'(x) + a \cdot f(x) = b \Leftrightarrow f(x) = \frac{b}{a} + c \cdot e^{-ax}, c \in \mathbb{R}$$

Herleitung: Die Richtung " \Leftarrow " folgt sofort durch Differenzieren jeweils der rechten Gleichung, denn zu (1): $f'(x) = (e^{\int g(x) dx})' = e^{\int g(x) dx} \cdot g(x) = f(x) \cdot g(x)$ und zu (2): $f'(x) = -ac e^{-ax} = -a(c \cdot e^{-ax} + \frac{b}{a} - \frac{b}{a}) = -a(c \cdot e^{-ax} + \frac{b}{a}) + b = -a \cdot f(x) + b$.

Die umgekehrte Richtung " \Rightarrow " ergibt sich wie folgt:

$$\text{zu (1): } f'(x) = f(x) \cdot g(x) \Leftrightarrow \frac{f'(x)}{f(x)} = g(x), f(x) \neq 0 \Leftrightarrow [\ln(|f(x)|)]' = g(x).$$

Integration beider Seiten führt zu $\ln(|f(x)|) = \int g(x) dx$ und die Operation "e hoch" schließlich zum Ergebnis $|f(x)| = e^{\int g(x) dx}$.

zu (2): $f'(x) + a \cdot f(x) = b \Leftrightarrow f'(x) \cdot e^{ax} + a \cdot f(x) \cdot e^{ax} = b \cdot e^{ax} \Leftrightarrow [f(x) \cdot e^{ax}]' = b \cdot e^{ax}$.

Durch Integration ergibt sich $f(x) \cdot e^{ax} = \int b \cdot e^{ax} dx = \frac{b}{a} \cdot e^{ax} + c, c \in \mathbb{R}$. Division durch e^{ax} führt schließlich zu $f(x) = \frac{b}{a} + \frac{c}{e^{ax}} = \frac{b}{a} + c \cdot e^{-ax}, c \in \mathbb{R}$.

Beispiele (1) Gesucht sind Funktionen, welche die konstante Elastizität $\epsilon_f(x) = a$ aufweisen.

$$\epsilon_f(x) = x \cdot \frac{f'(x)}{f(x)} = a \Leftrightarrow f'(x) = f(x) \cdot \frac{a}{x}, x \neq 0 \Leftrightarrow f'(x) = f(x) \cdot g(x) \text{ mit } g(x) = \frac{a}{x}.$$

Nach der Regel (1) ist dann $|f(x)| = e^{\int \frac{a}{x} dx} = e^{a \cdot \ln(|x|) + c} = e^{a \cdot \ln(|x|)} \cdot e^c = |x|^a \cdot e^c = b \cdot |x|^a, b \in \mathbb{R}_+,$ also insgesamt $f(x) = b \cdot |x|^a, b \neq 0$.

(2) Der Marktpreis p eines Produktes sei von der Zeit t abhängig, gemessen vom Zeitpunkt $t = 0$ der Produkteinführung an. Die Angebots- und die Nachfragefunktion seien gegeben durch $x_A(p(t)) = 3 + p(t) + 2 p'(t), x_N(p(t)) = 5 - p(t) + p'(t)$.

$p'(t)$ modelliert die Erwartungshaltung von Anbietern und Nachfragern im Zeitablauf. Werden steigende Preise erwartet, also $p'(t) > 0$, so erhöhen die Anbieter die Produktion, während sich die Nachfrager zeitig bevorraten; entsprechend senken die Anbieter bei erwarteten Preisrückgängen, $p'(t) < 0$, ihre Produktion, während sich die Nachfrager abwartend verhalten.

Gesucht ist die Funktion $p(t)$, bei der stets Marktgleichgewicht herrscht, d.h. $x_A(p(t)) = x_N(p(t))$ ist, unter der zusätzlichen Bedingung, dass der Preis zur Produkteinführung 2 € betrug.

Lösung: $x_A(p(t)) = x_N(p(t)) \Leftrightarrow 3 + p(t) + 2p'(t) = 5 - p(t) + p'(t) \Leftrightarrow p'(t) + 2p(t) = 2.$

Nach der Regel (2) ergibt sich für $a = 2$ und $b = 2$: $p(t) = \frac{2}{2} + c \cdot e^{-2t} = 1 + c \cdot e^{-2t}, c \in \mathbb{R}.$

Die zusätzliche Bedingung $p(0) = 2$ besagt dann $p(0) = 1 + c \cdot e^0 = 1 + c = 2$, also $c = 1.$

Die Funktion des Gleichgewichtspreises lautet daher $p(t) = 1 + e^{-2t}.$