

Inhaltsverzeichnis

I	Mathematik 2	2	6	Extremwertaufgaben	11
1	Anwendung der Differential- und Integralrechnung	2	6.1	Extremwertaufgaben ohne Nebenbedingungen	11
1.1	Taylorentwicklung	2	6.2	Ausgleichsrechnung (Least squares)	11
1.2	Landau-Symbole	2	6.3	Extremwertaufgaben unter Nebenbedingungen	11
1.2.1	Rechenregeln	2	7	Kurvenintegrale	12
1.3	Fixpunktiteration	2	7.1	Kurvenintegral eines Skalarfeldes	12
1.4	Newton-Verfahren	2	7.2	Kurvenintegral eines Vektorfeldes	12
1.5	Kurven in \mathbb{R}^n	3	7.3	Gradientenfelder	12
1.5.1	Singuläre Punkte	3			
1.5.2	Länge einer Kurve	3			
1.5.3	Umparametrisierung einer Kurve	3			
1.5.4	Krümmung einer Kurve	3			
2	Differentialrechnung in \mathbb{R}^n : Skalarfelder	3			
2.1	Topologie	3			
2.2	Folgen in \mathbb{R}^n	4			
2.3	Stetigkeit von Skalarfeldern	4			
2.4	Partielle Ableitung	4			
2.5	Gradient	4			
2.6	Totale (Fréchet-) Differenzierbarkeit	4			
2.7	Richtungsableitung	4			
2.8	Höhere partielle Ableitungen	4			
2.9	Hessematrix	4			
2.10	Taylorentwicklung für Skalarfelder	5			
3	Lineare Abbildungen	5			
3.1	Definitionen	5			
3.2	Dualraum	5			
3.3	Matrixdarstellung	5			
3.4	Transformationsmatrix	6			
3.5	Matrixnormen	6			
4	Differentialrechnung in \mathbb{R}^n : Vektorfelder	6			
4.1	Differenzierbarkeit von Vektorfeldern	6			
4.2	Jacobimatrix	7			
4.3	Rechenregeln	7			
4.4	Fréchet-Differenzierbarkeit	7			
4.5	Kettenregel	7			
4.6	Laplace-Operator	7			
4.7	Divergenz	7			
4.8	Rotation	7			
4.9	Besondere Zusammenhänge	7			
4.10	Orthogonale krummlinige Koordinaten	8			
4.11	Implizite Funktionen	8			
4.12	Mittelwertsatz	8			
4.12.1	Für Skalarfelder	8			
4.12.2	Für Vektorfelder	8			
5	Eigenwerte, Matrixfaktorisierungen	9			
5.1	Eigenwerte	9			
5.2	Eigenvektoren	9			
5.3	Diagonalisierung	9			
5.3.1	Spezielle Eigenschaften	9			
5.4	Ähnliche Matrizen	9			
5.5	Definitheit	10			
5.6	Spektralnorm	10			
5.7	Schurzerlegung	10			
5.8	Singulärwertzerlegung	10			

Teil I Mathematik 2

1 Anwendung der Differential- und Integralrechnung

1.1 Taylorentwicklung

Das m -te Taylorpolynom der Funktion f an der Stelle x_0 ist

$$T_m(x_0; x) = \sum_{k=0}^m \frac{f^{(k)}(x_0)}{k!} \cdot (x - x_0)^k$$

Restglied:

$$R_{m+1}(x_0; x) = f(x) - T_m(x_0; x)$$

$$R_{m+1}(x_0; x) = \frac{1}{m!} \int_{x_0}^x (x-t)^m f^{(m+1)}(t) dt$$

Außerdem gibt es für jedes $x \in I$ ein $\xi_x \in (x_0; x)$ oder $\xi_x \in (x, x_0)$ mit

$$R_{m+1}(x_0; x) = \frac{f^{(m+1)}(\xi_x)}{(m+1)!} (x - x_0)^{m+1}$$

Taylorformel für $h = x - x_0$:

$$f(x_0 + h) = \sum_{k=0}^m \frac{f^{(k)}(x_0)}{k!} h^k + R_{m+1}(x_0; x_0 + h)$$

mit

$$R_{m+1}(x_0; x_0 + h) = \frac{f^{(m+1)}(\xi_x)}{(m+1)!} h^{(m+1)}$$

Mit Landau-Notation gilt:

$$R_{m+1}(x_0; x_0 + h) = o(h^m) = O(h^{m+1})$$

Eine Taylorreihe konvergiert genau dann, wenn gilt:

$$\lim_{n \rightarrow \infty} R_n(x_0, x) = 0$$

bzw. falls es zwei Konstanten $A, B > 0$ gibt, sodass gilt:

$$|f^{(n)}(x)| \leq A \cdot B^n \quad \forall n \in \mathbb{N}, x \in I$$

1.2 Landau-Symbole

$$f(x) = o(g(x)) \text{ für } x \rightarrow a$$

gilt genau dann, wenn auch

$$\lim_{x \rightarrow a} \left| \frac{f(x)}{g(x)} \right| = 0$$

gilt.

$$f(x) = O(g(x)) \text{ für } x \rightarrow a$$

gilt genau dann, wenn in einer Umgebung $(a - \epsilon, a + \epsilon)$ für eine Konstante $C > 0$ gilt:

$$|f(x)| \leq C \cdot |g(x)| \text{ für } x \in (a - \epsilon, a + \epsilon)$$

bzw.

$$\limsup_{x \rightarrow a} \left| \frac{f(x)}{g(x)} \right| < \infty$$

1.2.1 Rechenregeln

$$f = o(g) \Rightarrow f = O(g)$$

$$f_1 = o(g), f_2 = o(g) \Rightarrow f_1 + f_2 = o(g)$$

$$f_1 = O(g), f_2 = O(g) \Rightarrow f_1 + f_2 = O(g)$$

$$f_1 = O(g_1), f_2 = O(g) \Rightarrow f_1 \cdot f_2 = O(g_1 \cdot g_2)$$

$$f_1 = O(g_1), f_2 = o(g_2) \Rightarrow f_1 \cdot f_2 = o(g_1 \cdot g_2)$$

1.3 Fixpunktiteration

Jede Nullstellengleichung $f(x) = 0$ kann man in der Form einer Fixpunktgleichung $g(x) = x$ darstellen, indem man $g(x) = f(x) + x$ setzt. Man kann versuchen, die unbekannte Lösung x^* mit $g(x^*) = x^*$ wie folgt zu approximieren:

Man betrachtet eine erste Näherung $x_0 \in I$ und definiert eine Folge x_n durch die rekursive Vorschrift:

$$x_{n+1} = g(x_n)$$

Ist die Folge konvergent, so ist der Grenzwert ein Fixpunkt von g .

Banach'scher Fixpunktsatz:

Sei $g : I = [a, b] \rightarrow I$ eine differenzierbare Funktion (Selbstabbildung des Intervalls I) und es gebe eine Konstante $0 \leq L \leq 1$ mit

$$|g'(x)| \leq L \text{ für alle } x \in I$$

Dann existiert genau eine Lösung x^* der Fixpunktgleichung $g(x) = x$ und die Iteration

$$x_0 \in I, x_{n+1} = g(x_n)$$

konvergiert gegen x^* . Außerdem gelten folgende Abschätzungen:

$$\underbrace{|x_n - x^*| \leq \frac{L^n}{1-L} |x_1 - x_0|}_{A\text{-Priori}} \quad \text{und} \quad \underbrace{|x_n - x^*| \leq \frac{L}{1-L} |x_n - x_{n-1}|}_{A\text{-Posteriori}}$$

1.4 Newton-Verfahren

Iterationsschritt:

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

Konvergenztest:

Sei $f : [a, b] \rightarrow \mathbb{R}$ eine zweimal stetig differenzierbare Funktion und $x^* \in (a, b)$ eine Nullstelle von f .

$$m = \min_{a \leq x \leq b} |f'(x)| > 0$$

$$M = \max_{a \leq x \leq b} |f''(x)|$$

Liegt der Startwert $x_0 \in I$ hinreichend nahe bei der Nullstelle, d.h.

$$|x_0 - x^*| \leq \frac{2m}{M}$$

dann konvergiert das Newton-Verfahren gegen die Nullstelle x^* und es gilt:

$$|x^* - x_n| \leq \frac{M}{2m} |x^* - x_{n-1}|^2$$

1.5 Kurven in \mathbb{R}^n

Sei $I = [a, b]$ ein Intervall. Eine stetige Funktion $k : I \rightarrow \mathbb{R}^n$ heißt Kurve im \mathbb{R}^n .

- Der Punkt $k(a) \in \mathbb{R}^n$ heißt Anfangspunkt und $k(b) \in \mathbb{R}^n$ Endpunkt der Kurve
- Gilt $k(a) = k(b)$, so heißt die Kurve geschlossen.

Eine Kurve $k : I \rightarrow \mathbb{R}^n$ heißt differenzierbar an der Stelle $t \in I$, falls alle Komponentenfunktionen $k_1, k_2, \dots, k_n : I \rightarrow \mathbb{R}$ an der Stelle t differenzierbar sind. Somit lautet die Ableitung von k an der Stelle t :

$$k'(t) = (k'_1(t), k'_2(t), \dots, k'_n(t))^T$$

Eine Kurve $k : I \rightarrow \mathbb{R}^n$ heißt C^1 -Kurve, falls die Funktion k stetig differenzierbar ist.

1.5.1 Singuläre Punkte

Singuläre Punkte einer Kurve sind diejenigen Stellen, an denen gilt

$$k'(t) = 0$$

1.5.2 Länge einer Kurve

$$L(k) = \int_a^b \|k'(t)\| dt$$

1.5.3 Umparametrisierung einer Kurve

Sei eine Kurve $k : I = [a, b] \rightarrow \mathbb{R}^n$ gegeben und sei $h : [c, d] \rightarrow [a, b]$ eine streng monoton wachsende Funktion mit $h(c) = a$ und $h(d) = b$. Wir definieren eine neue Kurve

$$\tilde{k} : [c, d] \rightarrow \mathbb{R}^n, \tilde{k}(\tau) = k(h(\tau))$$

Die Kurve \tilde{k} entsteht aus der Kurve k durch Umparametrisierung.

Eine Kurve $k : I \rightarrow \mathbb{R}^n$ heißt regulär, wenn gilt

$$k'(t) \neq 0 \text{ für alle } t \in I$$

Umparametrisierung nach der Bogenlänge:

Für eine reguläre Kurve $k : I \rightarrow \mathbb{R}^n$ definieren wir die Umparametrisierung der Kurve k nach der Bogenlänge. Dafür betrachten wir die Funktion $s : I = [a, b] \rightarrow \mathbb{R}$, die für jedes t die Bogenlänge des Kurvenstücks zwischen $k(a)$ und $k(t)$ beschreibt:

$$s(t) = \int_a^t \|k'(\tau)\| d\tau$$

Es gilt:

$$s'(t) = \|k'(t)\| > 0 \text{ für alle } t \in I$$

da die Kurve k regulär ist. Somit ist die Funktion $s : [a, b] \rightarrow [0, L(k)]$ streng monoton steigend und es gibt eine (streng monoton steigende) Umkehrfunktion $s^{-1} : [0, L(k)] \rightarrow [a, b]$.

Wir definieren die Umparametrisierung nach der Bogenlänge als

$$\begin{aligned} \tilde{k} : [0, L(k)] &\rightarrow \mathbb{R}^n, \tilde{k} = k(s^{-1}(\tau)) \\ &\Rightarrow \|\tilde{k}'(\tau)\| = 1 \end{aligned}$$

1.5.4 Krümmung einer Kurve

Sei $k : I \rightarrow \mathbb{R}^n$ eine C^1 -Kurve. Wir definieren den Tangenteinheitsvektor $T(t)$ als

$$T(t) = \frac{1}{\|k'(t)\|} k'(t)$$

Sei $k : I \rightarrow \mathbb{R}^n$ eine reguläre Kurve. Die Krümmung von k an der Stelle t ist gegeben als

$$\kappa(t) = \frac{1}{s'(t)} \|T'(t)\|$$

Sei $k : I \rightarrow \mathbb{R}^2$ eine Kurve mit $k(t) = (x(t), y(t))$, so gilt:

$$\kappa(t) = \frac{x'(t)y''(t) - y'(t)x''(t)}{[x'(t)^2 + y'(t)^2]^{\frac{3}{2}}}$$

Ist die Kurve k nach der Bogenlänge parametrisiert, so gilt $s'(t) = 1$ und

$$\tilde{\kappa} = x'(t)y''(t) - y'(t)x''(t)$$

2 Differentialrechnung in \mathbb{R}^n : Skalarfelder

Ein Skalarfeld ist eine Funktion $f : \Omega \rightarrow \mathbb{R}$, wobei $\Omega \subset \mathbb{R}^n$.

2.1 Topologie

Sei $\Omega \subset \mathbb{R}^n$

- Ein Punkt $x \in \Omega$ heißt **innerer Punkt** von Ω , falls es eine offene Kugel $B_\epsilon(x)$ (Umgebung) gibt, die ganz in Ω liegt, d.h. $B_\epsilon \subset \Omega$.
- Die **Menge aller inneren Punkte** von Ω heißt das Innere von Ω und wird mit $\text{int}(\Omega)$ bezeichnet.
- Eine Menge $\Omega \subset \mathbb{R}^n$ heißt **offen**, falls $\Omega = \text{int}(\Omega)$ gilt.
- Ein Punkt $x \in \mathbb{R}^n$ heißt **Randpunkt** von Ω , falls jede offene Kugel $B_r(x)$ sowohl Punkte aus Ω als auch Punkte, die nicht in Ω liegen, enthält
- Die **Menge aller Randpunkte** von Ω heißt der Rand von Ω und wird mit $\partial\Omega$ bezeichnet.
- Der **Abschluss** einer Menge Ω wird definiert als $\bar{\Omega} = \Omega + \partial\Omega$.
- Eine Menge $\Omega \subset \mathbb{R}^n$ heißt **abgeschlossen**, falls alle Randpunkte in der Menge Ω enthalten sind, d.h. $\partial\Omega \subset \Omega \Leftrightarrow \Omega = \bar{\Omega}$.
- Die Menge $\Omega \subset \mathbb{R}^n$ heißt **beschränkt**, wenn es eine Konstante $M > 0$ gibt, sodass $\|x\| \leq M$ für alle $x \in \Omega$ gilt.
- Eine Menge $\Omega \subset \mathbb{R}^n$ ist **kompakt**, wenn sie beschränkt und abgeschlossen ist.

Regeln:

- Sei $A \subset \mathbb{R}^n$ offen, dann ist $\mathbb{R}^n \setminus A$ abgeschlossen.
- Sei $A \subset \mathbb{R}^n$ abgeschlossen, dann ist $\mathbb{R}^n \setminus A$ offen.
- Seien $A, B \subset \mathbb{R}^n$ offen, dann sind auch $A \cup B$ und $A \cap B$ offen.
- Seien $A, B \subset \mathbb{R}^n$ abgeschlossen, dann sind auch $A \cup B$ und $A \cap B$ abgeschlossen.

2.2 Folgen in \mathbb{R}^n

Wir betrachten Folgen $\{x^k\} \subset \mathbb{R}^n$ analog zu den Folgen in \mathbb{R} oder in \mathbb{C} : x^1, x^2, x^3, \dots

Ein Folgenglied $x^k \in \mathbb{R}^n$ hat dann die Form

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

Ein Vektor $x \in \mathbb{R}^n$ heißt Grenzwert der Folge $\{x^k\} \subset \mathbb{R}^n$, wenn

$$\lim_{k \rightarrow \infty} \|x - x^k\| = 0$$

erfüllt ist. Man schreibt dann

$$\lim_{k \rightarrow \infty} x^k = x \text{ oder } x^k \rightarrow x, k \rightarrow \infty$$

2.3 Stetigkeit von Skalarfeldern

Sei $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ ein Skalarfeld und sei $y \in \Omega$.

f ist stetig an der Stelle y , wenn gilt:

$$\lim_{x \rightarrow y} f(x) = f(y)$$

Das bedeutet, dass für jede Folge $\{x^k\} \subset \Omega$ mit $x^k \rightarrow y$ die Folge der Funktionswerte $\{f(x^k)\} \subset \mathbb{R}$ gegen $f(y)$ konvergieren muss.

Sei $\Omega \subset \mathbb{R}^n$ eine kompakte Menge und sei $f : \Omega \rightarrow \mathbb{R}$ ein stetiges Skalarfeld, dann nimmt f auf Ω sein Minimum und sein Maximum an. D.h. es gibt Punkte $x_m, x_M \in \Omega$ mit

$$f(x_m) = \min_{x \in \Omega} f(x) \text{ und } f(x_M) = \max_{x \in \Omega} f(x)$$

2.4 Partielle Ableitung

Bei der Berechnung der partiellen Ableitung nach x_i werden alle anderen Variablen festgehalten, d.h. sie werden als Konstanten betrachtet:

$$\frac{\partial f}{\partial x_i}(x) = \lim_{h \rightarrow 0} \frac{f(x_1, \dots, x_{i-1}, x_i + h, x_{i+1}, \dots, x_n) - f(x_1, \dots, x_n)}{h}$$

Sei $\Omega \subset \mathbb{R}^n$ eine offene Menge, $f : \Omega \rightarrow \mathbb{R}$ ein Skalarfeld und $x \in \Omega$. Die Funktion f heißt partiell differenzierbar an der Stelle x , wenn alle partiellen Ableitungen an der Stelle x existieren. f ist stetig differenzierbar, falls alle partiellen Ableitungen stetige Skalarfelder definieren. Die Menge aller stetigen differenzierbaren Funktionen bezeichnet man mit $C^1(\Omega)$.

2.5 Gradient

Der Gradient an der Stelle x wird definiert als

$$\nabla f(x) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(x) \\ \frac{\partial f}{\partial x_2}(x) \\ \vdots \\ \frac{\partial f}{\partial x_n}(x) \end{pmatrix} \in \mathbb{R}^n$$

Rechenregeln:

$$\nabla(f + g)(x) = \nabla f(x) + \nabla g(x)$$

$$\nabla(\lambda f)(x) = \lambda \nabla f(x)$$

$$\nabla(fg)(x) = g(x)\nabla f(x) + f(x)\nabla g(x)$$

$$\nabla \left(\frac{f}{g} \right) (x) = \frac{g(x)\nabla f(x) - f(x)\nabla g(x)}{g(x)^2}$$

2.6 Totale (Fréchet-) Differenzierbarkeit

Sei $\Omega \subset \mathbb{R}^n$ eine offene Menge. Ein Skalarfeld $f : \Omega \rightarrow \mathbb{R}$ heißt total differenzierbar oder Fréchet-differenzierbar an der Stelle $x \in \Omega$, falls es einen Vektor $b \in \mathbb{R}^n$ gibt, sodass die Darstellung

$$f(x + d) = f(x) + b^T d + o(\|d\|)$$

für $d \in \mathbb{R}^n$ und $\|d\| \rightarrow 0$ gilt.

Ist das Skalarfeld an der Stelle x Fréchet-differenzierbar, so ist f auch partiell differenzierbar und $b = \nabla f(x)$ und f ist an dieser Stelle stetig.

Ein Skalarfeld ist Fréchet-differenzierbar in x_0 , falls dort alle partiellen Ableitungen existieren und das Skalarfeld stetig ist bzw. wenn gilt:

$$\lim_{d \rightarrow 0} \frac{f(x_0 + d) - f(x_0) - (\nabla f)^T d}{\|d\|} = 0$$

Fréchet-Differenzierbarkeit kann auch widerlegt werden, indem man eine Folge findet, die für $n \rightarrow \infty$ nicht gegen die Stelle x_0 konvergiert.

2.7 Richtungsableitung

Sei $\Omega \subset \mathbb{R}^n$ eine offene Menge und $f : \Omega \rightarrow \mathbb{R}$ ein Skalarfeld. Die Richtungsableitung im Punkt x in Richtung $d \in \mathbb{R}^n$ wird definiert als

$$\partial_d f(x) = \lim_{h \rightarrow 0} \frac{f(x + hd) - f(x)}{h}$$

Ist f Fréchet-differenzierbar im Punkt $x \in \Omega$, dann existiert für jede Richtung $d \in \mathbb{R}^n$ die Richtungsableitung $\partial_d f(x)$ und es gilt

$$\partial_d f(x) = \nabla f(x)^T d$$

2.8 Höhere partielle Ableitungen

$$\partial_{x_j} \partial_{x_i} f(x) = \lim_{h \rightarrow 0} \frac{\partial_{x_i} f(x + h e_j) - \partial_{x_i} f(x)}{h}$$

Wird zweimal nach x_i abgeleitet, so schreib man

$$\partial_{x_i} \partial_{x_i} f(x) = \frac{\partial^2}{\partial x_i^2} f(x) = \partial_{x_i}^2 f(x)$$

Die Menge aller zweimal stetig differenzierbaren Funktionen bezeichnet man mit $C^2(\Omega)$ usw.

Sei $f : \Omega \rightarrow \mathbb{R}$ und $f \in C^2(\Omega)$, dann gilt

$$\partial_i \partial_j f(x) = \partial_j \partial_i f(x)$$

für alle $1 \leq i, j \leq n$ und für alle $x \in \Omega$.

2.9 Hessematrix

Sei $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ ein zweimal partiell differenzierbares Skalarfeld, dann ist die Hesse-Matrix $H_f(x)$ definiert als

$$H_f(x) = \begin{pmatrix} \partial_{x_1} \partial_{x_1} f(x) & \dots & \partial_{x_n} \partial_{x_1} f(x) \\ \partial_{x_1} \partial_{x_2} f(x) & \dots & \partial_{x_n} \partial_{x_2} f(x) \\ \vdots & \ddots & \vdots \\ \partial_{x_1} \partial_{x_n} f(x) & \dots & \partial_{x_n} \partial_{x_n} f(x) \end{pmatrix}$$

Des Weiteren gilt:

$$H_f(x) = J_{\nabla f}(x)$$

2.10 Taylorentwicklung für Skalarfelder

Das Taylorpolynom m -ten Grades am Punkt a lautet:

$$\begin{aligned} T_m(a; a+h) &= f(a) + \nabla f(a)^T h + \frac{1}{2} h^T H_f(a) h + \dots \\ &= f(a) + \sum_{i=1}^n \partial_{x_i} f(a) h_i + \frac{1}{2!} \sum_{i,j=1}^n \partial_{x_i} \partial_{x_j} f(a) h_i h_j \\ &\quad + \frac{1}{3!} \sum_{i,j,k=1}^n \partial_{x_i} \partial_{x_j} \partial_{x_k} f(a) h_i h_j h_k + \dots \end{aligned}$$

mit $h_i = x_i - a_i$

Falls das Skalarfeld f bereits ein Polynom n -ten Grades oder kleiner ist, d.h. falls $f \in P_n$, dann folgt

$$T_n(a; x) = f(x)$$

Für das Restglied gilt:

$$\begin{aligned} R_{m+1}(a; a+h) &= \frac{1}{m!} \cdot \int_0^1 (1-t)^m g^{(m+1)}(t) dt \\ &= \frac{1}{(m+1)!} g^{(m+1)}(\xi) = O(\|h\|^{m+1}) \end{aligned}$$

mit $\xi \in (0, 1)$.

3 Lineare Abbildungen

3.1 Definitionen

- Eine **Abbildung (eine Funktion)** $f : X \rightarrow Y$ zwischen den Mengen X und Y ist eine Zuordnung, bei der jedem Element $x \in X$ ein Element $y = f(x) \in Y$ zugeordnet wird. Man schreibt:

$$f : X \rightarrow Y, x \rightarrow y = f(x)$$

- Das **Bild** der Funktion $f : X \rightarrow Y$ wird definiert als

$$Bild(f) = \{y \in Y \mid y = f(x), x \in X\}$$

- Eine Funktion heißt **surjektiv**, falls das Bild mit der Zielmenge Y übereinstimmt.
- Eine Funktion heißt **injektiv**, falls keine zwei verschiedenen Element aus X auf dasselbe Element aus Y abgebildet werden.
- Eine Funktion heißt **bijektiv**, falls sie injektiv und surjektiv ist. Eine bijektive Funktion besitzt eine Umkehrfunktion, die ebenfalls bijektiv ist.
- Eine Abbildung $f : V \rightarrow W$ zwischen zwei Vektorräumen V und W heißt **linear**, falls folgende Eigenschaft erfüllt ist:

$$f(\alpha x + y) = \alpha f(x) + f(y) \text{ für alle } x, y \in V \text{ und } \alpha \in \mathbb{R}$$

- Der **Kern** von f wird definiert als

$$Kern(f) = \{x \in V \mid f(x) = 0\}$$

- Der **Rang** einer linearen Abbildung f wird definiert als

$$Rang(f) = dim(Bild(f))$$

- Eine Funktion $f : V \rightarrow W$ zwischen zwei Vektorräumen V und W heißt **affinlinear**, wenn $f(x) = g(x) + a$ gilt, wobei $g : V \rightarrow W$ eine lineare Abbildung und $a \in W$ ist.
- Eine lineare Abbildung $f : V \rightarrow W$ ist genau dann surjektiv, wenn $Bild(f) = W$ und genau dann injektiv, wenn $Kern(f) = 0$.

- **Dimensionsformel:**

$$dim(Bild(f)) + dim(Kern(f)) = dim(V)$$

- Falls $f : V \rightarrow V : f$ ist injektiv $\Leftrightarrow f$ ist surjektiv

3.2 Dualraum

- Seien V und W zwei endlichdimensionale Vektorräume. Die Menge aller linearen Abbildungen von V nach W wird mit

$$\mathcal{L}(V, W) = \{f : V \rightarrow W \mid f \text{ ist linear}\}$$

bezeichnet.

- Sei V ein endlichdimensionaler Raum. Die Menge

$$V^* = \mathcal{L}(V, \mathbb{R}) \tag{1}$$

aller linearen Abbildungen von V nach \mathbb{R} nennt man Dualraum von V . Die linearen Abbildungen von V nach \mathbb{R} werden lineare Funktionale genannt.

Jedes lineare Funktional $f \in V^*$ kann man durch einen Zeilenvektor $w_f \in \mathbb{R}^{1 \times n}$ darstellen: $f(v) = w_f \cdot v$

- Sei V ein endlichdimensionaler Vektorraum. Der Vektorraum V^* ist auch endlichdimensional und es gilt

$$dim(V^*) = dim(V)$$

- Sei $\{v_1, v_2, \dots, v_n\} \subset V$ eine Basis von V . Dann bildet die Menge $\{v_1^*, v_2^*, \dots, v_n^*\} \subset V^*$ mit der Eigenschaft

$$v_i^*(v_j) = \delta_{ij} = \begin{cases} 1 & \text{falls } i = j \\ 0 & \text{sonst} \end{cases}$$

eine Basis von V^* . Diese Basis nennt man dual zu der Basis $\{v_1, v_2, \dots, v_n\}$.

3.3 Matrixdarstellung

Sei $f : V \rightarrow W$ eine lineare Abbildung, $B = \{b_1, b_2, \dots, b_n\}$ eine Basis von V und $C = \{c_1, c_2, \dots, c_m\}$ eine Basis von W , dann kann man die Bilder der Basisvektoren von V in der Basis von W darstellen:

$$f(b_j) = \sum_{i=1}^m a_{ij} c_i$$

Die Koeffizienten a_{ij} in der oberen Darstellung bilden die sogenannte Abbildungsmatrix $A \in \mathbb{R}^{m \times n}$, $A = (a_{ij})$ der Abbildung f bezüglich der Basen B von V und C von W . In der j -ten Spalte der Abbildungsmatrix stehen die

Koordinaten des Bildes $f(b_j)$ des j -ten Basiselements b_j der Basis C .

Falls C die Einheitsbasis ist, gilt

$$A = (f(b_1), f(b_2), \dots, f(b_n))$$

Hieraus kann A' bezüglich einer anderen Ausgangsbasis C' folgendermaßen berechnet werden:

$$(C'|A) \rightsquigarrow (I_n|A')$$

3.4 Transformationsmatrix

Sei $B = (b_1, b_2, \dots, b_n)$ die Ausgangsbasis des Raumes V und $B' = (b'_1, b'_2, \dots, b'_n)$ eine neue Basis. Die Vektoren der neuen Basis haben eine eindeutige Darstellung in der alten Basis:

$$b'_j = a_{1j}b_1 + a_{2j}b_2 + \dots + a_{nj}b_n$$

Man definiert die sog. Transformationsmatrix $T_{B'}^B \in \mathbb{R}^{n \times n}$ wie folgt:

$$T_{B'}^B = (a_{ij})$$

D.h. in der j -ten Spalte der Matrix $T_{B'}^B$ stehen die Koordinaten des j -ten Vektors b'_j der neuen Basis bezüglich der alten Basis B . Es gilt außerdem:

$$(T_{B'}^B)^{-1} = T_B^{B'}$$

Diese Transformationsmatrix kann folgendermaßen berechnet werden:

$$(B|B') \rightsquigarrow (I_n|T_{B'}^B)$$

bzw.

$$(B|B') \rightsquigarrow (T_B^{B'}|I_n)$$

Möchte man einen Vektor x in einen neuen Vektor x' bezüglich der neuen Basis umrechnen und umgekehrt, so gilt:

$$\begin{aligned} x' &= T_B^{B'} x \\ x &= T_{B'}^B x' \end{aligned}$$

Die Abbildungsmatrix A einer Funktion f bezüglich der Eingangsbasis $B = (b_1, b_2, \dots, b_n)$ und der Ausgangsbasis $C = (c_1, c_2, \dots, c_n)$ kann in eine Abbildungsmatrix \tilde{A} bezüglich der neuen Basen B' und C' wie folgt berechnet werden:

$$\tilde{A} = (T_{C'}^C)^{-1} A T_B^{B'}$$

3.5 Matrixnormen

Normen auf $\mathbb{R}^{m \times n}$ müssen folgende Bedingungen erfüllen:

- Definitheit:

$$\|A\| \geq 0 \text{ für alle } A \in \mathbb{R}^{m \times n} \text{ und}$$

$$\|A\| = 0 \Leftrightarrow A = 0$$

- Homogenität:

$$\|\alpha A\| = |\alpha| \cdot \|A\| \text{ für alle } A \in \mathbb{R}^{m \times n}, \alpha \in \mathbb{R}$$

- Dreiecksungleichung:

$$\|A + B\| \leq \|A\| + \|B\| \text{ für alle } A, B \in \mathbb{R}^{m \times n}$$

Frobenius-Norm:

$$\|A\|_F = \sqrt{\sum_{i=1}^m \sum_{j=1}^n a_{ij}^2}$$

Eine Norm auf $\mathbb{R}^{n \times n}$ heißt **submultiplikativ**, falls für alle $A, B \in \mathbb{R}^{n \times n}$ gilt:

$$\|AB\| \leq \|A\| \cdot \|B\|$$

Eine Matrixnorm heißt **verträglich** mit einer Vektornorm $\|\cdot\|_V$ auf \mathbb{R}^n falls gilt:

$$\|Ax\|_V \leq \|A\| \cdot \|x\|_V \text{ für alle } A \in \mathbb{R}^{n \times n}, x \in \mathbb{R}^n$$

Die von $\|\cdot\|$ induzierte Matrixnorm ist gegeben als

$$\|A\| = \sup_{x \in \mathbb{R}^n, x \neq 0} \frac{\|Ax\|_V}{\|x\|_V}$$

(natürliche Matrixnorm) mit den Eigenschaften:

- Diese Matrixnorm ist mit der entsprechenden Vektornorm verträglich.
- Diese Matrixnorm ist submultiplikativ
- Es gilt:

$$\|A\| = \sup_{x \in \mathbb{R}^n, \|x\|_V=1} \|Ax\|_V$$

Die von der **Maximumnorm** $\|\cdot\|_\infty$ induzierte Matrixnorm ist die "maximale Zeilensumme".

$$\|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|$$

Die von der **l_1 -Norm** $\|\cdot\|_1$ induzierte Matrixnorm ist die "maximale Spaltensumme":

$$\|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}|$$

4 Differentialrechnung in \mathbb{R}^n : Vektorfelder

Eine Funktion $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ heißt Vektorfeld. Ein Vektorfeld f besteht aus m Skalarfeldern, die man Komponentenfunktionen $f_i : \Omega \rightarrow \mathbb{R}$ nennt. Ein Punkt (Vektor) $x \in \Omega \subset \mathbb{R}^n$ wird auf einen Punkt $f(x) \in \mathbb{R}^m$ durch

$$f(x) = \begin{pmatrix} f_1(x) \\ f_2(x) \\ \dots \\ f_m(x) \end{pmatrix}$$

abgebildet. Die Begriffe des Grenzwertes und der Stetigkeit können Komponentenweise definiert werden.

4.1 Differenzierbarkeit von Vektorfeldern

Ein Vektorfeld $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ heißt **stetig** an der Stelle $x_0 \in \Omega$, falls alle Komponentenfunktionen $f_i : \Omega \rightarrow \mathbb{R}$ an der Stelle x_0 stetig sind.

Das Vektorfeld ist (partiell) an der Stelle x_0 **differenzierbar**, falls alle Komponentenfunktionen an der Stelle x_0 differenzierbar sind.

Sei $\Omega \in \mathbb{R}^n$ eine offene Menge. Ein Vektorfeld heißt **stetig differenzierbar** auf Ω , wenn alle Komponentenfunktionen auf Ω stetig differenzierbar sind. Man schreibt dann $f \in C^1(\Omega)$.

4.2 Jacobimatrix

Sei $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ ein Vektorfeld, das an der Stelle $x \in \Omega$ partiell differenzierbar ist. Die Jacobi-Matrix oder Funktionalmatrix $J_f(x) \in \mathbb{R}^{m \times n}$ wird definiert als:

$$J_f(x) = \begin{pmatrix} \partial_1 f_1(x) & \partial_2 f_1(x) & \dots & \partial_n f_1(x) \\ \partial_1 f_2(x) & \partial_2 f_2(x) & \dots & \partial_n f_2(x) \\ \vdots & \vdots & \ddots & \vdots \\ \partial_1 f_m(x) & \partial_2 f_m(x) & \dots & \partial_n f_m(x) \end{pmatrix}$$

Die Zeilen der Jacobi-Matrix sind also die transponierten Gradientenfunktionen:

$$J_f(x) = \begin{pmatrix} \nabla f_1(x)^T \\ \nabla f_2(x)^T \\ \vdots \\ \nabla f_m(x)^T \end{pmatrix}$$

Der i -te Spaltenvektor ist der Tangentialvektor zu der i -ten Koordinatenlinie.

4.3 Rechenregeln

- Seien $f, g : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ zwei differenzierbare Vektorfelder, dann ist die Linearkombination $\alpha f + \beta g$ mit $\alpha, \beta \in \mathbb{R}$ auch differenzierbar und es gilt:

$$J_{\alpha f + \beta g}(x) = \alpha J_f(x) + \beta J_g(x)$$

- Seien $f, g : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ zwei differenzierbare Vektorfelder, dann ist das Skalarfeld

$$h : \Omega \rightarrow \mathbb{R} \text{ mit } h(x) = f(x)^T g(x)$$

differenzierbar und es gilt:

$$J_h(x) = g(x)^T J_f(x) + f(x)^T J_g(x)$$

$$\nabla h(x) = J_h(x)^T = J_f(x)^T g(x) + J_g(x)^T f(x)$$

4.4 Fréchet-Differenzierbarkeit

Sei $\Omega \subset \mathbb{R}^n$ eine offene Menge. Ein Vektorfeld $f : \Omega \rightarrow \mathbb{R}^m$ heißt Fréchet-differenzierbar (oder total differenzierbar) an der Stelle $x \in \Omega$, falls es eine lineare Abbildung $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ gibt, sodass folgende Darstellung erfüllt ist:

$$f(x+h) = f(x) + T(h) + o(\|h\|)$$

für $\|h\| \rightarrow 0$.

Ist das Vektorfeld $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ an der Stelle $x \in \Omega$ Fréchet-differenzierbar, so folgt für die lineare Abbildung T die Darstellung

$$T(h) = J_f(x)h$$

Das Vektorfeld ist genau dann Fréchet-differenzierbar, wenn alle Komponentenfunktionen Fréchet-differenzierbar sind.

Für die Fréchet-Ableitung schreibt man oft $Df(x)$ und meint die lineare Abbildung $Df(x) : \mathbb{R}^n \rightarrow \mathbb{R}^m$. Die Jacobi-Matrix ist die Abbildungsmatrix dieser Abbildung bezüglich der Standardbasis. Es gilt:

$$Df(x)(h) = J_f(x)h \text{ (totale Ableitung)}$$

f ist auf ganz Ω Fréchet-differenzierbar, falls gilt

$$\Omega \subset \mathbb{R}^n \text{ offen und } f \in C^1(\Omega)$$

4.5 Kettenregel

Seien $\Omega \subset \mathbb{R}^n, D \subset \mathbb{R}^m$ offene Mengen und $f : \Omega \rightarrow \mathbb{R}^m, g : D \rightarrow \mathbb{R}^l$ zwei Vektorfelder. Seien das Vektorfeld f an der Stelle $x \in \Omega$ und das Vektorfeld g an der Stelle $y = f(x) \in D$ Fréchet-differenzierbar, dann ist auch die Komposition $h = g \circ f$ an der Stelle x Fréchet-differenzierbar und es gilt:

$$Dh(x) = Dg(f(x)) \circ Df(x)$$

$$J_h(x) = J_g(f(x))J_f(x)$$

Für die partiellen Ableitungen folgt:

$$\partial_j h_i(x) = \sum_{k=1}^m \partial_k g_i(f(x)) \partial_j f_k(x)$$

4.6 Laplace-Operator

Für ein zweimal stetig differenzierbares Skalarfeld $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ definiert man den Laplace-Operator

$$\Delta f(x) = \sum_{i=1}^n \partial_i^2 f(x)$$

4.7 Divergenz

Für ein Vektorfeld $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ definiert man die Divergenz

$$\nabla \cdot f(x) = \operatorname{div}(f(x)) = \sum_{i=1}^n \partial_i f_i(x)$$

4.8 Rotation

Für ein Vektorfeld $f : \Omega \subset \mathbb{R}^3 \rightarrow \mathbb{R}^3$ definiert man die Rotation

$$\nabla \times f(x) = \operatorname{rot}(f(x)) = \begin{pmatrix} \partial_2 f_3(x) - \partial_3 f_2(x) \\ \partial_3 f_1(x) - \partial_1 f_3(x) \\ \partial_1 f_2(x) - \partial_2 f_1(x) \end{pmatrix}$$

4.9 Besondere Zusammenhänge

- Sei $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ ein zweimal stetig differenzierbares Skalarfeld, dann gilt:

$$\nabla \nabla f(x) = \Delta f(x)$$

- Für ein zweimal stetig differenzierbares Skalarfeld $f : \Omega \subset \mathbb{R}^3 \rightarrow \mathbb{R}$ gilt:

$$\operatorname{rot}(\nabla f) = 0$$

- Für ein zweimal stetig differenzierbares Vektorfeld $g : \Omega \subset \mathbb{R}^3 \rightarrow \mathbb{R}^3$ gilt:

$$\operatorname{div}(\operatorname{rot}(g)) = 0$$

4.10 Orthogonale krummlinige Koordinaten

Siehe Formelsammlung (z.B. Springer)

Eine Koordinatentransformation (für orthogonale Koordinaten) kann auch mittels Matrizenmultiplikation durchgeführt werden.

Seien die Vektoren $\tilde{g}(x), \hat{g}(x) \in \mathbb{R}^3$ mit

$\tilde{g}(x)$: neue Eingangsbasis, kartesische Ausgangsbasis

$\hat{g}(x)$: neue Ein- und Ausgangsbasis

Dann gilt:

$$\hat{g}(x) = B^T \cdot \tilde{g}(x)$$

mit den Matrizen

- Zylinderkoordinaten (r, φ, z)

$$B = \begin{pmatrix} \underbrace{\cos(\varphi)}_{e_r} & \underbrace{-\sin(\varphi)}_{e_\varphi} & \underbrace{0}_{e_z} \\ \underbrace{\sin(\varphi)}_{e_r} & \underbrace{\cos(\varphi)}_{e_\varphi} & \underbrace{0}_{e_z} \\ \underbrace{0}_{e_r} & \underbrace{0}_{e_\varphi} & \underbrace{1}_{e_z} \end{pmatrix}$$

- Kugelkoordinaten (r, φ, θ)

$$B = \begin{pmatrix} \underbrace{\cos(\varphi) \sin(\theta)}_{e_r} & \underbrace{-\sin(\varphi)}_{e_\varphi} & \underbrace{\cos(\varphi) \cos(\theta)}_{e_\theta} \\ \underbrace{\sin(\varphi) \sin(\theta)}_{e_r} & \underbrace{\cos(\varphi)}_{e_\varphi} & \underbrace{\sin(\varphi) \cos(\theta)}_{e_\theta} \\ \underbrace{\cos(\theta)}_{e_r} & \underbrace{0}_{e_\varphi} & \underbrace{-\sin(\theta)}_{e_\theta} \end{pmatrix}$$

4.11 Implizite Funktionen

Eine implizite Funktion ist gegeben durch $f(x, y, \dots) = 0$.

Sei $D \in \mathbb{R}^2$ offen und $f : D \rightarrow \mathbb{R}$ stetig differenzierbar. Sei der Punkt $(x_0, y_0) \in D$ derart gegeben, dass

$$f(x_0; y_0) = 0 \text{ und } \partial_y f(x_0; y_0) \neq 0$$

(Funktion auflösbar nach x , falls $\partial_x f \neq 0$)

Dann gibt es zwei offene Intervalle $I \subset \mathbb{R}$ mit Mittelpunkt x_0 und $K \subset \mathbb{R}$ mit Mittelpunkt y_0 , sodass gilt

- $\mathbb{R} = I \times K \subset D$ und $\partial_y f(x, y) \neq 0$ für alle $(x, y) \in \mathbb{R}$
- Durch $f(x, y) = 0$ ist auf I eindeutig eine differenzierbare Funktion $g : I \rightarrow K$ implizit definiert mit

$$f(x, g(x)) = 0 \text{ für alle } x \in I$$

Die Ableitung von g ist gegeben als

$$y' = g'(x) = -\frac{\partial_x f(x, g(x))}{\partial_y f(x, g(x))} \text{ für alle } x \in I$$

Die zweite Ableitung von g ist gegeben als

$$y'' = g''(x) = -\left[\frac{\partial_x^2 f(x, g(x)) + 2\partial_y \partial_x f(x, g(x))g'(x)}{\partial_y f(x, g(x))} + \frac{\partial_y^2 f(x, g(x))(g'(x))^2}{\partial_y f(x, g(x))^2} \right]$$

Sei $f(x, y, z) = 0$ mit $z = g(x, y)$. Dann gilt:

$$\nabla g(x, y) = \begin{pmatrix} -\frac{\partial_x f(x, y, z)}{\partial_z f(x, y, z)} \\ -\frac{\partial_y f(x, y, z)}{\partial_z f(x, y, z)} \end{pmatrix}$$

Außerdem gilt:

Der Tangentialvektor steht senkrecht zum Gradienten.

Sei $f : \mathbb{R}^{k+m} \rightarrow \mathbb{R}^m$ ein stetig differenzierbares **Vektorfeld** und $f(z) = 0, z \in \mathbb{R}^{k+m}$.

Man kann jeden Vektor $z \in \mathbb{R}^{k+m}$ als ein Paar $z = (x, y)$ mit $x \in \mathbb{R}^k$ und $y \in \mathbb{R}^m$ schreiben. In Analogie zu dem eindimensionalen Fall sucht man ein Vektorfeld $g : \mathbb{R}^k \rightarrow \mathbb{R}^m$ mit $f(x, g(x)) = 0$.

Die Bedingung für die Existenz einer solchen implizit definierten Funktion in einer Umgebung der Stelle $z_0 = (x_0, y_0)$ mit $f(z_0) = 0$ ist die folgende: Die Teilmatrix $J_{f,y}(z_0) \in \mathbb{R}^{m \times m}$ der Jacobi-Matrix $J_f(z_0) \in \mathbb{R}^{m \times (m+k)}$ bestehend aus

$$\partial_j f_i(z), 1 \leq i \leq m, k+1 \leq j \leq k+m$$

muss invertierbar sein.

Satz der Umkehrabbildung:

Ist ein Vektorfeld $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ umkehrbar, dann gilt:

$$f^{-1}(f(x)) = x \text{ für alle } x \in \Omega$$

Sind f und f^{-1} stetig differenzierbar, so folgt aus der Kettenregel:

$$J_{f^{-1}}(f(x))J_f(x) = I$$

Sei $\Omega \subset \mathbb{R}^n$ offen und sei $f : \Omega \rightarrow \mathbb{R}^n$ ein stetig differenzierbares Vektorfeld. Sei $x_0 \in \Omega, y_0 = f(x_0)$ und sei die Jacobi-Matrix $J_f(x_0)$ invertierbar. Dann gibt es eine Umgebung $U \subset \Omega$ von x_0 und eine Umgebung $V \subset f(\Omega)$ von y_0 , sodass f die Menge U auf die Menge V bijektiv abbildet. Die Umkehrfunktion $f^{-1} : V \rightarrow U$ ist stetig differenzierbar und es gilt:

$$J_{f^{-1}}(f(x)) = (J_f(x))^{-1} \text{ für alle } x \in U$$

4.12 Mittelwertsatz

4.12.1 Für Skalarfelder

Sei $\Omega \subset \mathbb{R}^n$ eine offene Menge, $f : \Omega \rightarrow \mathbb{R}$ ein stetig differenzierbares Skalarfeld und $x, y \in \Omega$. Es gelte zusätzlich $[x, y] \subset \Omega$.

Dann gibt es eine Zwischenstelle $\xi \in [x, y]$, sodass

$$f(y) - f(x) = \nabla f(\xi)^T (y - x)$$

gilt. Außerdem hat man die folgende Abschätzung

$$|f(y) - f(x)| \leq C \|y - x\|$$

mit

$$C = \max_{z \in [x, y]} \|\nabla f(z)\|$$

4.12.2 Für Vektorfelder

Der Mittelwertsatz ist nicht direkt auf Vektorfelder übertragbar.

⇒ Ungleichungsvariante des Mittelwertsatzes:

Sei $\Omega \subset \mathbb{R}^n$ eine offene Menge und $f : \Omega \rightarrow \mathbb{R}^m$ ein stetig differenzierbares Vektorfeld. Seien $x, y \in \Omega$ und $[x, y] \subset \Omega$, dann gilt:

$$\|f(y) - f(x)\| \leq C \|y - x\|$$

wobei die Konstante C gegeben als

$$C = \max_{z \in [x, y]} \|J_f(z)\|$$

mit der Jacobi-Matrix $J_f(x)$ und der Matrizenorm

$$\|A\| = \sup_{v \in \mathbb{R}^n, v \neq 0} \frac{\|Av\|}{\|v\|}$$

5 Eigenwerte, Matrixfaktorisierungen

Sei $f : V \rightarrow V$ eine lineare Selbstabbildung des Vektorraumes V . Um die Struktur dieser Abbildung zu studieren, sucht man oft nach einer Basis $B = (v_1, v_2, \dots, v_n)$ von V derart, dass die entsprechende Abbildungsmatrix eine einfache Form besitzt, z.B. eine Diagonalmatrix ist.

5.1 Eigenwerte

Eine reelle Zahl λ heißt Eigenwert der Matrix $A \in \mathbb{R}^{n \times n}$, wenn es einen Vektor $x \in \mathbb{R}^n$ gibt mit

$$Ax = \lambda x \text{ und } x \neq 0$$

Entsprechend definiert man auch komplexe Eigenwerte $\lambda \in \mathbb{C}$ für Matrizen in $\mathbb{C}^{n \times n}$.

Berechnung:

Eine Zahl $\lambda \in \mathbb{R}$ (bzw. $\lambda \in \mathbb{C}$) ist genau dann ein Eigenwert einer $(n \times n)$ -Matrix A , wenn das sog. charakteristische Polynom gleich Null ist, d.h.

$$\det(A - \lambda I_n) \stackrel{!}{=} 0$$

Die Vielfachheit k_i der Nullstelle λ_i heißt **algebraische Vielfachheit** des Eigenwertes λ_i .

Seien $\lambda_1, \lambda_2, \dots, \lambda_n$ alle Eigenwerte der Matrix A mit den algebraischen Vielfachheiten k_i , dann gilt:

$$\det(A) = \lambda_1^{k_1} \cdot \lambda_2^{k_2} \cdot \dots \cdot \lambda_n^{k_n}$$

5.2 Eigenvektoren

Die Eigenvektoren ergeben sich als nichttriviale Lösungen des homogenen LGS

$$(A - \lambda I_n)x = 0$$

Die Lösungsmenge dieses LGS ist ein Unterraum $V_\lambda = \text{Kern}(A - \lambda I_n)$. Alle Vektoren x mit $x \neq 0$ in diesem Eigenraum V_λ sind Eigenvektoren zum Eigenwert λ .

Die Dimension des Eigenraumes V_λ nennt man **geometrische Vielfachheit** des Eigenwertes λ . Es gilt:

$$1 \leq \dim(V_{\lambda_i}) \leq k_i$$

Die Eigenvektoren w_1, w_2, \dots, w_i zu paarweise verschiedenen Eigenwerten $\lambda_1, \lambda_2, \dots, \lambda_i$ sind linear unabhängig.

5.3 Diagonalisierung

- Eine lineare Abbildung $f : V \rightarrow V$ heißt diagonalisierbar, falls es eine Basis gibt, bezüglich derer die Abbildungsmatrix von f eine Diagonalmatrix ist.

Sei die Abbildung $f : V \rightarrow V$ diagonalisierbar,

(w_1, w_2, \dots, w_n) die entsprechende Basis und D die Abbildungsmatrix von f in der Form

$$D = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{pmatrix}$$

mit nicht notwendigerweise verschiedenen Zahlen λ_i . Dann gilt

$$f(w_i) = \lambda_i w_i$$

und somit ist λ_i ein Eigenwert und w_i ein Eigenvektor zum Eigenwert λ_i .

- Eine Matrix heißt diagonalisierbar, falls es eine invertierbare Matrix $T \in \mathbb{R}^{n \times n}$ und eine Diagonalmatrix $D \in \mathbb{R}^{n \times n}$ gibt mit der Eigenschaft

$$T^{-1}AT = D$$

Die Matrix D besteht dann aus den Eigenwerten der Matrix A und die Spalten der Matrix T bilden eine Basis aus linear unabhängigen Eigenvektoren der Matrix A .

- Besitzt eine Matrix $A \in \mathbb{R}^{n \times n}$ n verschiedene Eigenwerte, so ist diese Matrix diagonalisierbar.
- Eine Matrix $A \in \mathbb{R}^{n \times n}$ (bzw. $A \in \mathbb{C}^{n \times n}$) ist genau dann diagonalisierbar, wenn folgende zwei Bedingungen erfüllt sind:

- 1) Das charakteristische Polynom \mathcal{X}_A zerfällt in Linearfaktoren

$$\mathcal{X}_A(t) = (\lambda_1 - t)^{k_1} \cdot (\lambda_2 - t)^{k_2} \cdot \dots \cdot (\lambda_r - t)^{k_r}$$

mit $\lambda_i \in \mathbb{R}$ (bzw. $\lambda_i \in \mathbb{C}$).

- 2) Die algebraischen Vielfachheiten aller Eigenwerte stimmen mit den geometrischen Vielfachheiten überein, d.h.

$$k_i = \dim(V_{\lambda_i}), \quad i = 1, 2, \dots, r$$

5.3.1 Spezielle Eigenschaften

- A und A^{-1} besitzen die gleichen Eigenvektoren, die Eigenwerte sind invers zueinander
- Eine symmetrische Matrix $A \in \mathbb{R}^{n \times n}$ hat nur reelle Eigenwerte und es existiert eine Orthonormalbasis aus Eigenvektoren.

Eigenvektoren zu unterschiedlichen Eigenwerten sind bei symmetrischen Matrizen stets orthogonal zueinander. Falls die Eigenvektoren des gleichen Eigenwertes nicht orthogonal sind, muss das Gram-Schmidt-Verfahren angewendet werden.

5.4 Ähnliche Matrizen

Zwei $(n \times n)$ -Matrizen A und B heißen ähnlich, wenn es eine invertierbare Matrix T gibt, mit der Eigenschaft

$$T^{-1}AT = B$$

Ähnliche Matrizen haben dieselben Eigenwerte mit identischen algebraischen und geometrischen Vielfachheiten (Umkehrung gilt nicht!) aber nicht notwendigerweise die gleichen Eigenvektoren. Daraus folgt, dass sie

- 1) den gleichen Rang
- 2) die gleiche Determinante
- 3) die gleiche Spur
- 4) das gleiche charakteristische Polynom

haben.

Eine diagonalisierbare Matrix ist immer zu einer Diagonalmatrix ähnlich.

5.5 Definitheit

Für eine symmetrische Matrix $A \in \mathbb{R}^{n \times n}$ gilt:

- Falls alle Eigenwerte $> 0 \Rightarrow$ Positiv definit
- Falls alle Eigenwerte $\geq 0 \Rightarrow$ Positiv semidefinit
- Falls alle Eigenwerte $< 0 \Rightarrow$ Negativ definit
- Falls alle Eigenwerte $\leq 0 \Rightarrow$ Negativ semidefinit
- Falls positive und negative Eigenwerte existieren \Rightarrow Indefinit

Die Matrix $A^T A$ ist positiv semidefinit.

5.6 Spektralnorm

Es gilt:

$$|\lambda| \leq \|A\|$$

$$\|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}| \quad (\text{max. Zeilensumme})$$

$$\|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}| \quad (\text{max. Spaltensumme})$$

Sei A eine symmetrische Matrix, dann gilt:

$$\|A\|_2 = |\lambda_{\max}| = \{|\lambda| \mid \lambda \in \mathbb{R} \text{ ist Eigenwert von } A\}$$

Für beliebige $\mathbb{R}^{n \times n}$ -Matrizen gilt:

$$\|A\|_2 = \max\{\sqrt{|\mu|} \mid \mu \in \mathbb{R} \text{ ist Eigenwert von } A^T A\}$$

5.7 Schurzerlegung

Gegeben sei $A \in \mathbb{R}^{n \times n}$ mit dem charakteristischen Polynom $\mathcal{X}_A(t)$, das in Linearfaktoren zerfällt. Dann gibt es eine orthogonale Matrix $U \in \mathbb{R}^{n \times n}$, sodass

$$R = U^T A U$$

mit einer oberen Dreiecksmatrix R gilt (falls A symmetrisch $\Rightarrow R$ ist Diagonalmatrix).

Wenn das charakteristische Polynom nicht über \mathbb{R} zerfällt, dann über \mathbb{C} mit einer Matrix $U \in \mathbb{C}^{n \times n}$, sodass $\bar{U}^T U = I_n$ (=unitär) und für eine obere Dreiecksmatrix $R \in \mathbb{C}^{n \times n}$ $R = \bar{U}^T A U$ gilt.

Berechnung der Schurzerlegung:

Sei $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ die lineare Abbildung, die mit der Matrix A assoziiert wird, d.h. $f(x) = Ax$.

- 1) Bestimme einen Eigenwert λ_1 von A und einen normierten Eigenvektor v_1 zu λ_1 .

- 2) Wähle $n - 1$ Vektoren w_2, w_3, \dots, w_n , sodass die Vektoren $v_1, w_2, w_3, \dots, w_n$ eine Orthonormalbasis bilden. Diese Vektoren bilden die Spalten einer orthogonalen Matrix V_1

$$V_1 = (v_1, w_2, w_3, \dots, w_n)$$

- 3) Aufgrund der Bedingung $f(v_1) = Av_1 = \lambda v_1$ hat die Abbildungsmatrix der Abbildung f in der neuen Basis $v_1, w_2, w_3, \dots, w_n$ die Form

$$\tilde{A} = V_1^T A V_1 = \begin{pmatrix} \lambda_1 & * \\ 0 & A_1 \end{pmatrix}$$

mit einer Matrix $A_1 \in \mathbb{R}^{(n-1) \times (n-1)}$.

- 4) Man wiederholt die vorherigen Schritte für die Matrizen A_1, A_2, \dots, A_{n-2} solange, bis man alle Matrizen V_1, V_2, \dots, V_{n-1} erhalten hat.
- 5) Berechne die Matrizen $\hat{V}_2, \hat{V}_3, \dots, \hat{V}_{n-1} \in \mathbb{R}^{n \times n}$ wie folgt:

$$\hat{V}_i = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & V_i \end{pmatrix}$$

- 6) Berechne $Q = V_1 \cdot \hat{V}_2 \cdot \hat{V}_3 \cdot \dots \cdot \hat{V}_{n-1}$

- 7) Berechne

$$R = Q^T A Q = \begin{pmatrix} \lambda_1 & * & \dots & * \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \dots & 0 & \lambda_n \end{pmatrix}$$

5.8 Singulärwertzerlegung

Sei $A \in \mathbb{R}^{m \times n}$ eine Matrix. Dann gibt es zwei orthogonale Matrizen $U \in \mathbb{R}^{m \times m}, V \in \mathbb{R}^{n \times n}$ und eine Diagonalmatrix $\Sigma \in \mathbb{R}^{m \times n}$ der Form

$$\Sigma = \begin{pmatrix} \sigma_1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \sigma_m & 0 & \dots & 0 \end{pmatrix} \quad \text{falls } m \leq n$$

$$\Sigma = \begin{pmatrix} \sigma_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \sigma_n \\ 0 & \dots & 0 \end{pmatrix} \quad \text{falls } m \geq n$$

mit

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_k \geq 0, \quad k = \min(m, n)$$

die die folgende Darstellung erlauben:

$$A = U \Sigma V^T$$

Die Zahlen $\sigma_1, \sigma_2, \dots, \sigma_k$ heißen Singulärwerte von A und die Darstellung $A = U \Sigma V^T$ heißt Singulärwertzerlegung (SVD).

Berechnung der Singulärwertzerlegung:

1) Man bestimmt die Eigenwerte λ_l der Matrix $A^T A \in \mathbb{R}^{n \times n}$ und sortiert sie

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0$$

und berechnet die zugehörigen normierten Eigenvektoren v_l .

2) Die Matrix Σ mit den Singulärwerten ergibt sich durch die Gleichung

$$\sigma_l = \sqrt{\lambda_l}, \quad l = 1, 2, \dots, k = \min(m, n)$$

3) Die Matrix V ergibt sich als

$$V = (v_1, v_2, \dots, v_n)$$

4) Die Vektoren u_1, u_2, \dots, u_m werden durch die Gleichung

$$u_l = \frac{1}{\sigma_l} A v_l$$

bestimmt.

Falls $n < m$ oder falls $\sigma_l = 0$, dann müssen die Vektoren mit Hilfe des Gram-Schmidt-Verfahrens zu einer Orthonormalbasis von \mathbb{R}^m ergänzt werden.

Anschließend erhält man die Matrix U :

$$U = (u_1, u_2, \dots, u_m)$$

Für die Zerlegung von A^T gilt:

$$\Sigma_{A^T} = (\Sigma_A)^T, \quad V_{A^T} = U_A, \quad U_{A^T} = V_A$$

6 Extremwertaufgaben

6.1 Extremwertaufgaben ohne Nebenbedingungen

Sei $\Omega \subset \mathbb{R}^n$, $f : \Omega \rightarrow \mathbb{R}$ ein Skalarfeld. Ein Punkt $\hat{x} \in \Omega$ heißt lokales Minimum von f , wenn es eine (offene) Umgebung $B_\epsilon(\hat{x})$ gibt, sodass

$$f(\hat{x}) \leq f(x) \text{ für alle } x \in B_\epsilon(\hat{x}) \cap \Omega$$

gilt.

Sei \hat{x} ein **lokales Minimum** und es gelte zusätzlich

$$f(\hat{x}) < f(x) \text{ für alle } x \in B_\epsilon(\hat{x}) \cap \Omega, \quad x \neq \hat{x}$$

in einer Umgebung $B_\epsilon(\hat{x})$, so spricht man von einem **strikten lokalen Minimum**.

Analog dazu wird das (strikte) lokale Maximum definiert.

Notwendige Bedingung erster Ordnung:

Sei $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ partiell differenzierbar und sei $\hat{x} \in \text{int}(\Omega)$ ein lokales Extremum, dann gilt

$$\nabla f(\hat{x}) = 0$$

Notwendige Bedingung zweiter Ordnung:

Sei $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ ein zweimal stetig differenzierbares Skalarfeld und sei $\hat{x} \in \text{int}(\Omega)$ ein lokales Minimum, dann gilt $\nabla f(\hat{x}) = 0$ und die Hessematrix $H_f(\hat{x})$ ist an der Stelle \hat{x} positiv semidefinit.

Dementsprechend gilt für ein Maximum, dass die Hessematrix $H_f(\hat{x})$ an der Stelle \hat{x} negativ semidefinit ist.

Für strikte lokale Extrema gelten die Kriterien positiv bzw. negativ definit.

Vorgehen zur Bestimmung von Extrema:

1) Suche alle stationären Punkte $\nabla f(\hat{x}) = 0$

2) Bestimme die Hessematrix $H_f(\hat{x})$

- Falls $H_f(\hat{x})$ positiv definit \Rightarrow Minimum
- Falls $H_f(\hat{x})$ negativ definit \Rightarrow Maximum
- Falls $H_f(\hat{x})$ indefinit \Rightarrow Sattelpunkt
- Falls $H_f(\hat{x})$ semidefinit \Rightarrow keine Aussage möglich

6.2 Ausgleichsrechnung (Least squares)

Man betrachte eine Funktion

$$b = f(t) = x_0 + x_1 t + \dots + x_n t^n$$

mit unbekanntem Koeffizienten x_0, x_1, \dots, x_n . Es sind m Paare (t_i, b_i) z.B. aus Messungen gegeben. Man kann die Koeffizienten aufgrund von Messfehlern meist nicht durch Gleichungssysteme finden.

Lösung:

Man sucht den Koeffizientenvektor $x = (x_0, x_1, \dots, x_n)^T \in \mathbb{R}^n$ für den die sogenannten Residuen $r_i(x) = b_i - x_0 - x_1 t_i - \dots - x_n t_i^n$ möglichst klein sind. Man formuliert dies über den Ansatz der kleinsten Quadrate:

$$\text{Minimiere } \sum_{i=1}^m (b_i - x_0 - x_1 t_i - \dots - x_n t_i^n)$$

Fasst man die Residuen $r_i(x)$ zu einem Vektor $r(x) \in \mathbb{R}^m$ zusammen, so muss man folgendes Problem lösen:

$$\text{Minimiert } \|r(x)\|^2, \quad r(x) = b - Ax$$

mit

$$A = \begin{pmatrix} 1 & t_1 & t_1^2 & \dots & t_1^n \\ 1 & t_2 & t_2^2 & \dots & t_2^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & t_m & t_m^2 & \dots & t_m^n \end{pmatrix}$$

Ist \hat{x} ein Minimum $\|r(x)\|^2$, so muss \hat{x} die sogenannte Normalengleichung lösen:

$$A^T A \hat{x} = A^T b$$

6.3 Extremwertaufgaben unter Nebenbedingungen

Aufgaben vom Typ: Minimiere $f(x)$ unter der Nebenbedingung $g(x) = 0$ mit $f, g : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ heißen Optimierungsaufgaben unter Nebenbedingungen.

Die Menge

$$\Omega_{ad} = \{x \in \Omega \mid g(x) = 0\}$$

wird zulässige Menge genannt.

Ein Punkt $\hat{x} \in \Omega_{ad}$ heißt lokale Lösung dieser Optimierungsaufgabe, wenn es eine Umgebung $B_\epsilon(\hat{x})$ gibt, sodass gilt

$$f(\hat{x}) \leq f(x) \text{ für alle } x \in B_\epsilon(\hat{x}) \cap \Omega_{ad}$$

Für die lokale Lösung \hat{x} ist die Bedingung $\nabla f(\hat{x}) = 0$ im Allgemeinen nicht erfüllt!

Entsprechend kann man Maximierungsaufgaben mit Nebenbedingungen betrachten.

Notwendige Bedingung erster Ordnung:

Seien $f, g : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ zwei Fréchet-differenzierbare Skalarfelder und $\hat{x} \in \text{int}(\Omega)$ sei ein lokales Minimum der Optimierungsaufgabe

Minimiere $f(x)$ unter der Nebenbedingung $g(x) = 0$.

Gilt zusätzlich $\nabla g(\hat{x}) \neq 0$, dann gibt es einen Lagrange-Multiplikator $\hat{\lambda} \in \mathbb{R}$ mit

$$\nabla f(\hat{x}) + \hat{\lambda} \nabla g(\hat{x}) = 0$$

Lösung:

1) Formuliere die Lagrange-Funktion

$$L(x, \lambda) = f(x) + \lambda g(x)$$

2) Suche die stationären Punkte von L :

$$\nabla L(x, \lambda) = \begin{pmatrix} \nabla_x L(x, \lambda) \\ \partial_\lambda L(x, \lambda) \end{pmatrix} = \begin{pmatrix} \nabla f(x) + \lambda \nabla g(x) \\ g(x) \end{pmatrix} \stackrel{!}{=} 0$$

3) Setze die stationären Punkte in $f(x)$ ein und suche den höchsten Wert.

Ist die Nebenbedingung $g(x)$ ein Vektorfeld, so wird λ ebenfalls zum Vektor und es gilt

$$L(x, \lambda) = f(x) + \lambda^T g(x)$$

7 Kurvenintegrale

7.1 Kurvenintegral eines Skalarfeldes

Sei $D \subset \mathbb{R}^n$ offen, $w : I = [a, b] \rightarrow D$ eine differenzierbare Kurve und $f : D \rightarrow \mathbb{R}$ ein stetiges Skalarfeld, dann heißt

$$\int_w f ds = \int_a^b f(w(t)) \|w'(t)\| dt$$

das **Kurvenintegral** von f entlang von w .

Die **Länge einer Kurve** ist gleich dem Kurvenintegral für $f = 1$ entlang der Kurve:

$$L(w) = \int_w ds = \int_a^b \|w'(t)\| dt$$

Ist eine Kurve stückweise differenzierbar, so kann man über die einzelnen Kurvenstücke integrieren und danach aufaddieren.

Rechenregel (Linearität):

$$\int_w (\alpha f + \beta g) ds = \alpha \int_w f ds + \beta \int_w g ds$$

7.2 Kurvenintegral eines Vektorfeldes

Sei $E \subset \mathbb{R}^n$ offen, $w : I = [a, b] \rightarrow D$ eine differenzierbare Kurve und $f : D \rightarrow \mathbb{R}^n$ ein stetiges Vektorfeld, dann heißt

$$\int_w f dx = \int_a^b f(w(t)) \cdot w'(t) dt = \int_a^b f(w(t)) T(t) \|w'(t)\| dt$$

das **Kurvenintegral** von f entlang von w .

Mit dem Tangenteneinheitsvektor

$$T(t) = \frac{1}{\|w'(t)\|} w'(t)$$

gilt

$$\int_w f dx = \int_w (f \cdot T) ds$$

Ist die Kurve geschlossen, d.h. $w(a) = w(b)$, so verwendet man oftmals die Bezeichnung

$$\oint_w f dx$$

Für das Kurvenintegral des Vektorfeldes gilt die gleiche Linearität wie für das Kurvenintegral des Skalarfeldes.

7.3 Gradientenfelder

Eine Menge $D \subset \mathbb{R}^n$ heißt **wegzusammenhängend**, falls es zu je zwei Punkten x, y aus D eine Kurve $w : [a, b] \rightarrow D$ mit $w(a) = x$ und $w(b) = y$ gibt.

Eine wegzusammenhängende Menge $D \subset \mathbb{R}^n$ heißt **einfach zusammenhängend**, wenn jede geschlossene doppelpunktfreie Kurve in D stetig auf einen Punkt in D zusammengezogen werden kann, ohne dass D verlassen wird. Sei $C \in \mathbb{R}^n$ eine offene und wegzusammenhängende Menge. Ein stetiges Vektorfeld $f : D \rightarrow \mathbb{R}^n$ heißt **konservativ**, **Potentialfeld** oder **Gradientenfeld**, wenn es ein Skalarfeld $F : D \rightarrow \mathbb{R}$ gibt mit

$$f(x) = \nabla F(x) \text{ für alle } x \in D$$

In diesem Fall heißt F **Stammfunktion** und $U = -F$ eine **Potentialfunktion** von f .

Ist ein Vektorfeld konservativ und kennt man die Stammfunktion, so kann man direkt das entsprechende Kurvenintegral bestimmen:

$$\int_w f dx = F(w(b)) - F(w(a))$$

Sei $D \subset \mathbb{R}^n$ eine wegzusammenhängende offene Menge und sei $f : D \rightarrow \mathbb{R}^n$ ein stetiges Vektorfeld, dann sind folgende Aussagen äquivalent:

- f ist ein Gradientenfeld
- Für alle stetig differenzierbaren Kurven w in D hängt das Kurvenintegral $\int_w f dx$ nur vom Anfangs- und Endpunkt von w ab.

- Für alle geschlossenen stetig differenzierbaren Kurven w in D gilt

$$\oint_w f dx = 0$$

Überprüfung auf Gradientenfeld:

Sei $D \subset \mathbb{R}^n$ eine offene, einfach zusammenhängende Menge und sei $f : D \rightarrow \mathbb{R}^n$ ein stetig differenzierbares Vektorfeld. Dieses Vektorfeld ist genau dann ein Gradientenfeld, wenn die sogenannte **Integrabilitätsbedingung**

$$J_f(x) = J_f(x)^T \text{ für alle } x \in D$$

bzw.

$$\partial_{x_i} f_j(x) = \partial_{x_j} f_i(x) \text{ für alle } 1 \leq i, j \leq n$$

erfüllt ist.

Für dreidimensionale Gebiete:

Sei $D \subset \mathbb{R}^n$ eine offene, einfach zusammenhängende Menge und sei $f : D \rightarrow \mathbb{R}^3$ ein stetig differenzierbares Vektorfeld. f ist genau dann ein Gradientenfeld, wenn gilt:

$$\text{rot}(f(x)) = 0 \text{ für alle } x \in D$$

Berechnung der Stammfunktion:

- 1) Man wähle einen festen Punkt $x_0 \in D$
- 2) Zu jedem weiteren Punkt $y \in D$ wählt man eine Kurve $w_y : [a, b] \rightarrow D$ mit $w_y(a) = x_0$ und $w_y(b) = y$
- 3) Man berechnet die Stammfunktion durch

$$F(y) = \int_{w_y} f dx$$

Diese Formelsammlung ist eine überarbeitete und erweiterte Version der "Formelsammlung Mathematik 2 für Elektroingenieure" von Sebastian Wagner.

Lizenz: CC BY-NC-SA 3.0

<http://creativecommons.org/licenses/by-nc-sa/3.0/de/>