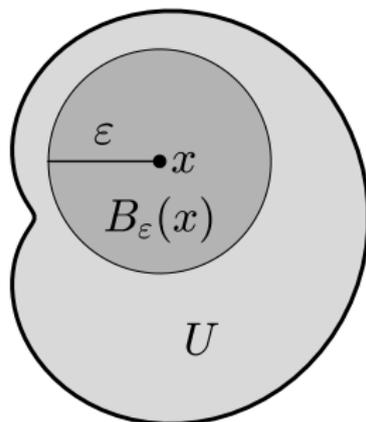


Umgebung

Als ε -Umgebung eines Punktes $x \in \mathbb{R}^n$ bezeichnet man die offene Kugel um x mit Radius ε :

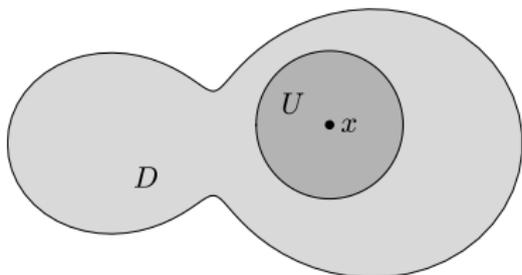
$$B_\varepsilon(x) = \{y \in \mathbb{R}^n : |y - x| < \varepsilon\}.$$



Allgemeiner ist eine Umgebung U von x eine Menge, die eine ε -Umgebung $B_\varepsilon(x)$ enthält.

Offene Menge

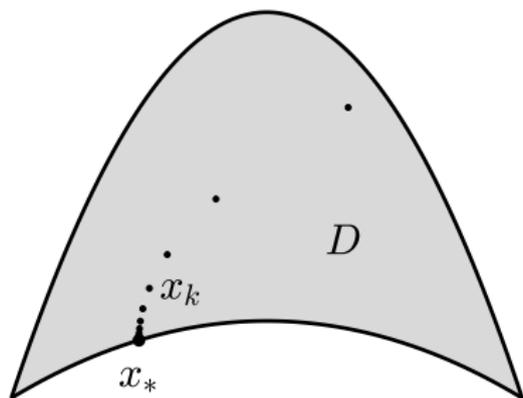
Eine Menge $D \subseteq \mathbb{R}^n$ heißt offen, wenn jeder Punkt $x \in D$ eine Umgebung U besitzt, die ganz in D enthalten ist. Damit enthält eine offene Menge keinen ihrer Randpunkte. Insbesondere sind \mathbb{R}^n und die leere Menge \emptyset offen.



Für eine beliebige Menge D bezeichnet $\overset{\circ}{D} \subseteq D$ das Innere von D , d.h. die Menge aller Punkte mit einer Umgebung in D .

Abgeschlossene Menge

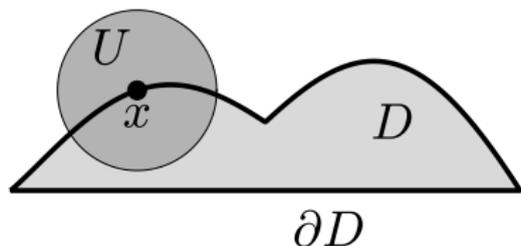
Eine Menge $D \subseteq \mathbb{R}^n$ heißt abgeschlossen, wenn die Grenzwerte x_* jeder konvergenten Folge von Punkten $x_k \in D$ in D liegen. Damit enthält eine abgeschlossene Menge jeden ihrer Randpunkte. Insbesondere sind \mathbb{R}^n und die leere Menge \emptyset abgeschlossen.



Für eine beliebige Menge D bezeichnet $\overline{D} \supseteq D$ den Abschluss von D , d.h. die Menge aller Grenzwerte von konvergenten Folgen in D .

Rand einer Menge

Der Rand ∂D einer Menge $D \subseteq \mathbb{R}^n$ besteht aus allen Punkten x , für die jede Umgebung U sowohl Punkte innerhalb als auch außerhalb von D enthält.



Alternativ kann der Rand als Differenz von Abschluss und Inneren,

$$\partial D = \bar{D} \setminus \overset{\circ}{D}$$

definiert werden.

Kompakte Menge

Eine beschränkte und abgeschlossene Menge $D \subseteq \mathbb{R}^n$ bezeichnet man als kompakt.

Äquivalent dazu sind folgende Charakterisierungen:

- Jede Folge in D besitzt eine konvergente Teilfolge mit Grenzwert in D .
 - Jede Überdeckung von D mit offenen Mengen besitzt eine endliche Teilüberdeckung.
-

Illustration von Mengeneigenschaften

Hälfte einer Kreisscheibe,

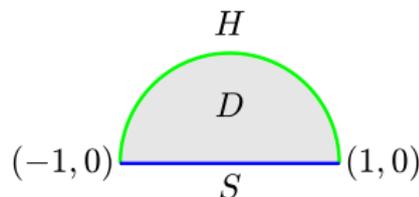
$$D : x^2 + y^2 < 1 \wedge y > 0,$$

mit Rand $\partial D = H \cup S$,

$H : x^2 + y^2 = 1 \wedge y \geq 0$ Halbkreis

$S : (-1, 1) \times \{0\}$ Geradensegment

- D offen: keine Randpunkte wegen strikter Ungleichungen
- H abgeschlossen: Relationen $=$, \geq und \leq bleiben bei Grenzwertbildung erhalten
- S weder offen noch abgeschlossen:
Produkt einer offenen und abgeschlossenen Menge
- \bar{D} , ∂D kompakt: abgeschlossen und beschränkt



Multivariate Funktion

Eine reelle Funktion f von mehreren Veränderlichen x_k ,

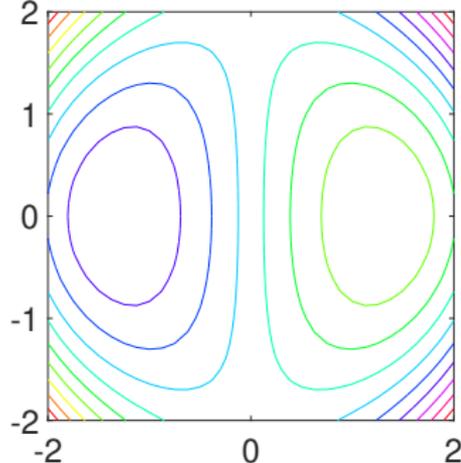
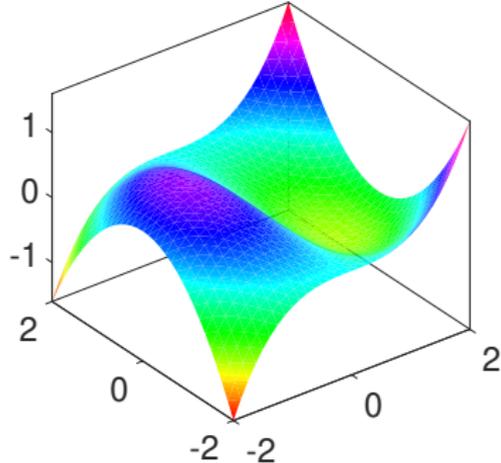
$$f : \mathbb{R}^n \supseteq D \rightarrow W \subseteq \mathbb{R}^m, \quad x \mapsto f(x),$$

ordnet einem Punkt/Vektor $x = (x_1, \dots, x_n)^t$ aus dem Definitionsbereich D einen Punkt/Vektor $f(x) = f(x_1, \dots, x_n) = (f_1(x), \dots, f_m(x))^t$ aus dem Wertebereich $W = f(D)$ zu.

Je nach der Dimension m unterscheidet man zwischen skalaren ($m = 1$) und vektorwertigen ($m > 1$) Funktionen. Ist $n = 1$ und f stetig, so parametrisiert f eine Kurve im \mathbb{R}^m .

Für $m, n \leq 3$ verwendet man meist keine Indizes und bezeichnet die Variablen mit x, y, z . Beispielsweise schreibt man für $m = n = 2$

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} f(x, y) \\ g(x, y) \end{pmatrix}.$$



Wie in der Abbildung illustriert ist, können zur Visualisierung skalarer Funktionen der Graph

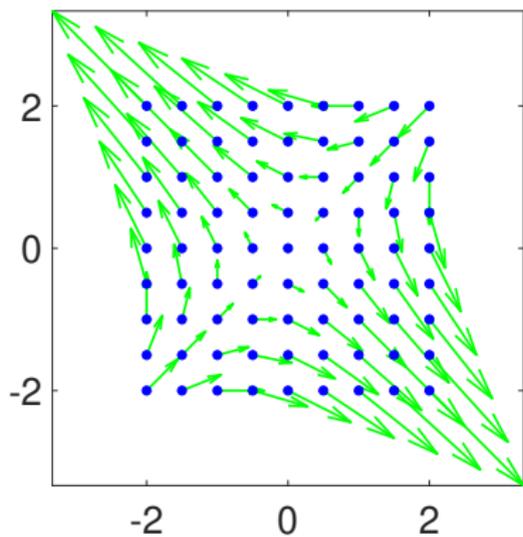
$$\{(x, f(x)) : x \in D\}$$

und die Niveaumengen

$$\{x \in D : f(x) = c\}$$

benutzt werden.

Eine vektorwertige Funktion lässt sich veranschaulichen, indem man die Vektoren $f(x)$ als Pfeile mit den Punkten x assoziiert.



Multivariate Polynome

Ein Polynom p in n Variablen x_1, \dots, x_n ist eine Linearkombination von Monomen:

$$p(x) = \sum_{\alpha} a_{\alpha} x^{\alpha}, \quad x^{\alpha} = x_1^{\alpha_1} \cdots x_n^{\alpha_n},$$

mit $\alpha_k \in \mathbb{N}_0$.

Je nach Summationsbereich unterscheidet man zwischen

- totalem Grad $\leq m$: $\sum \alpha = \alpha_1 + \cdots + \alpha_n \leq m$;
- maximalem Grad $\leq m$: $\max \alpha = \max_k \alpha_k \leq m$.

Für bivariate und trivariate Polynome bezeichnet man die Variablen meist mit x, y bzw. x, y, z . Beispielsweise bilden die Monome

$$(x, y) \mapsto x^j y^k, \quad j, k \geq 0, j + k \leq m$$

eine Basis für die bivariaten Polynome vom totalen Grad $\leq m$, und eine Basis für die trivariaten Polynome mit maximalen Grad $\leq m$ besteht aus den Monomen

$$(x, y, z) \mapsto x^j y^k z^\ell, \quad 0 \leq j, k, \ell \leq m.$$

Man bezeichnet ein n -variates Polynom p als homogen vom Grad k , wenn

$$p(sx) = s^k p(x) \quad \text{für } s \in \mathbb{R}.$$

Ein solches Polynom ist eine Linearkombination der Monome $x \mapsto x^\alpha$ mit $\sum \alpha = k$.

Die Dimensionen der drei n -variaten Polynomräume entsprechen den Anzahlen der relevanten Monome:

homogen vom Grad k	totaler Grad $\leq m$	maximaler Grad $\leq m$
$\binom{k+n-1}{n-1}$	$\binom{m+n}{n}$	$(m+1)^n$

Beweis

(i) Bivariate Polynome ($n = 2$):

Auflistung der homogenen Monome

$$k = 0 : \quad 1$$

$$k = 1 : \quad x, y$$

$$k = 2 : \quad x^2, xy, y^2$$

$$k = 3 : \quad x^3, x^2y, y^2x, y^3$$

...

\rightsquigarrow Anzahl $k + 1 = \binom{k+2-1}{2-1}$ für Grad k

Dimension der bivariaten Polynome vom totalen Grad $\leq m$:

$$1 + 2 + \dots + (m + 1) = \frac{(m + 2)(m + 1)}{2} = \binom{m + 2}{2}$$

$(m + 1)^2$ Monome vom maximalen Grad $\leq m$

$$x^j y^k, \quad j, k \leq m$$

(ii) Homogene n -variate Polynome vom Grad k :
identifiziere den Exponent

$$(\alpha_1, \dots, \alpha_n), \alpha_1 + \dots + \alpha_n = k$$

eines relevanten Monoms mit einer strikt monotonen Folge

$$\beta_1 = \alpha_1 + 1$$

$$\beta_2 = \alpha_1 + \alpha_2 + 2$$

...

$$\beta_{n-1} = \alpha_1 + \dots + \alpha_{n-1} + (n-1)$$

aus $\{1, \dots, k+n-1\}$, d.h.

$$\alpha_j = \beta_j - \beta_{j-1} - 1$$

mit $\beta_0 = 0, \beta_n = k+n$

$\rightsquigarrow \binom{k+n-1}{n-1}$ Möglichkeiten

(iii) n -variate Monome vom totalen Grad $\leq m$:

$\iff (n+1)$ -variate homogene Monome vom Grad m :

$$x_1^{\alpha_1} \cdots x_n^{\alpha_n} \iff x_1^{\alpha_1} \cdots x_n^{\alpha_n} x_{n+1}^{m - \sum \alpha_i}$$

(Letzter Exponent liegt fest.)

\rightsquigarrow Dimension des Polynomraums

$$\binom{m + (n+1) - 1}{(n+1) - 1}$$

(iv) n -variate Polynome vom maximalen Grad $\leq m$:

Analog zum bivariaten Fall existieren $(m+1)^n$ Monome $x \mapsto x^\alpha$ mit $0 \leq \alpha_j \leq m$.

Verschiedene bi- und trivariate Polynome

bivariates Polynom mit totalen Grad ≤ 3

$$p(x, y) = a_{0,0} + (a_{1,0}x + a_{0,1}y) + (a_{2,0}x^2 + a_{1,1}xy + a_{0,2}y^2) \\ + (a_{3,0}x^3 + a_{2,1}x^2y + a_{1,2}xy^2 + a_{0,3}y^3)$$

bzw. $p(x_1, x_2) = \sum_{\alpha_1 + \alpha_2 \leq 3} a_{\alpha} x_1^{\alpha_1} x_2^{\alpha_2}$

spezielles homogenes bivariates Polynom vom Grad 5

$$(x, y) \mapsto p(x, y) = 7x^5 - 8x^3y^2 + 6xy^4$$

bzw. $p(x_1, x_2) = 7x^{(5,0)} - 8x^{(3,2)} + 6x^{(1,4)}$

trivariates Polynom mit maximalem Grad ≤ 1 in drei Variablen:

$$p(x, y, z) = a_{0,0,0} + (a_{1,0,0}x + a_{0,1,0}y + a_{0,0,1}z) \\ + (a_{1,1,0}xy + a_{1,0,1}xz + a_{0,1,1}yz) + a_{1,1,1}xyz$$

Stetigkeit multivariater Funktionen

Eine Funktion $f : \mathbb{R}^n \supseteq D \rightarrow \mathbb{R}^m$ ist in einem Punkt (a_1, \dots, a_n) ihres Definitionsbereichs D stetig, wenn

$$x \rightarrow a \quad \Longrightarrow \quad f(x) \rightarrow f(a),$$

d.h. für alle $\varepsilon > 0$ existiert $\delta(\varepsilon) > 0$, so dass

$$|x - a| < \delta \quad \Longrightarrow \quad |f(x) - f(a)| < \varepsilon.$$

Gilt dies für alle Punkte $a \in D$, so ist f stetig auf D .

Existiert der Grenzwert für einen Punkt a auf dem Rand ∂D des Definitionsbereichs, so lässt sich f in diesen Randpunkt stetig fortsetzen.

Stetigkeit ist verträglich mit den arithmetischen Operation, d.h. eine Summe, ein Produkt und ein Quotient stetiger Funktionen ist stetig. Bei der Bildung eines Quotienten muss lediglich vorausgesetzt werden, dass der Nenner keine Nullstelle hat. Desweiteren ist die Hintereinanderschaltung stetiger Funktionen stetig.

Beispiel

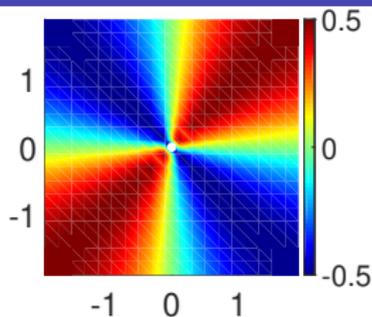
Unstetigkeit nicht-konstanter bivariater in Polarkoordinaten gegebener Funktionen f , die nur vom Winkel und nicht vom Radius abhängen

$f : (r, \varphi) \mapsto f(\varphi)$ mit $x = r \cos \varphi$,

$y = r \sin \varphi$, $r = \sqrt{x^2 + y^2}$

z.B.

$$f(x, y) = \frac{xy}{x^2 + y^2} = \cos \varphi \sin \varphi$$



Einschränkung auf Gerade durch den Ursprung, $g : y = mx$ bzw.

$g : \varphi = c \rightsquigarrow$ konstanter Funktionswert:

$$f(x, mx) = \frac{mx^2}{x^2 + (mx)^2} = \frac{m}{1 + m^2} = c_m \quad \text{bzw.} \quad f(\varphi) = \cos \varphi \sin \varphi = c_\varphi$$

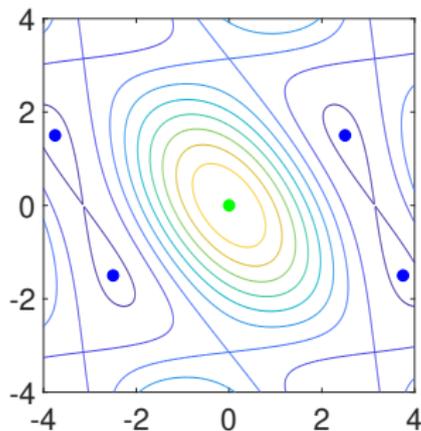
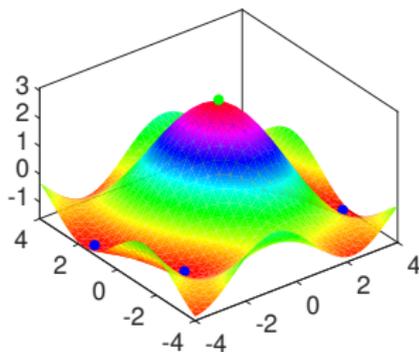
nicht stetig fortsetzbar für $(x, y) \rightarrow (0, 0)$, da verschiedener Grenzwert für jede Ursprungsgerade:

$$\lim_{x \rightarrow 0} f(x, mx) = \lim_{x \rightarrow 0} \frac{m}{1 + m^2} = \frac{m}{1 + m^2}$$

Extremwerte stetiger Funktionen

Eine stetige Funktion $f : \mathbb{R}^n \supseteq D \rightarrow \mathbb{R}$ besitzt auf einer kompakten Menge $K \subseteq D$ sowohl ein Minimum als auch ein Maximum.

Auf dem links abgebildeten Funktionsgraphen sind Minima und ein Maximum markiert. Bei den Niveaulinien erkennt man die Extremstellen durch geschlossene Kurven, die diese Punkte einschließen.



Äquivalenz von Normen

Jede Vektornorm $\|\cdot\|$ in \mathbb{R}^n ist zur euklidischen Norm $|\cdot|$ äquivalent, d.h. es gibt positive Konstanten c_k mit

$$c_1\|x\| \leq |x| \leq c_2\|x\|$$

für alle $x \in \mathbb{R}^n$.

Beweis

Die Ungleichungen sind skalierungsinvariant, d.h. invariant unter der Substitution $x \rightarrow sx$ mit $s \in \mathbb{R}$.

\rightsquigarrow betrachte nur Vektoren x auf der Einheitskugel

$$S : |x| = 1$$

$\|x\| \neq 0$ für $x \in S \implies$ Stetigkeit der Funktion $x \mapsto |x|/\|x\|$ (*)

S kompakt $\implies \exists$ positives Minimum c_1 und Maximum c_2

(*) Beweis der Stetigkeit der Norm mit Hilfe der Dreiecksungleichung:

$$\|x\| \leq \|x - y\| + \|y\|, \quad \|y\| \leq \|y - x\| + \|x\|$$

\implies (nach Umformung)

$$\pm(\|x\| - \|y\|) \leq \|x - y\| \quad \text{bzw.} \quad |\|x\| - \|y\|| \leq \|x - y\|,$$

d.h. $\|\cdot\|$ ist Lipschitz-stetig mit Konstante 1

Lipschitz-Stetigkeit

Eine Funktion f ist Lipschitz-stetig auf einer Menge D , wenn eine Lipschitz-Konstante c existiert, so dass

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

für alle $x, y \in D$.

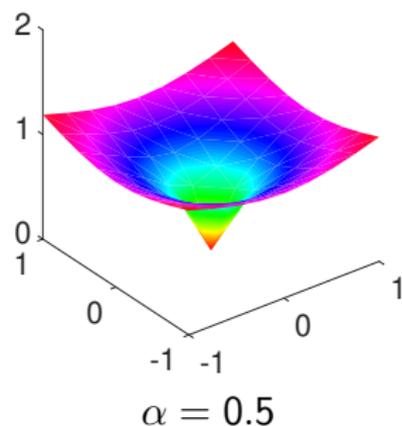
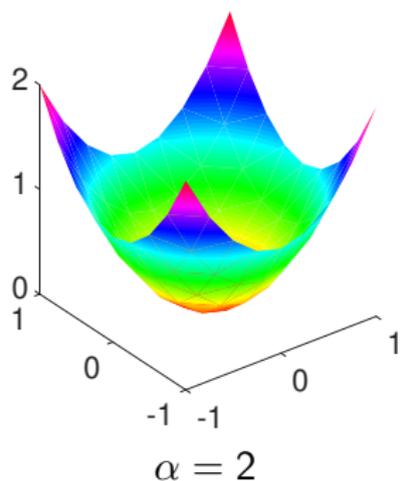
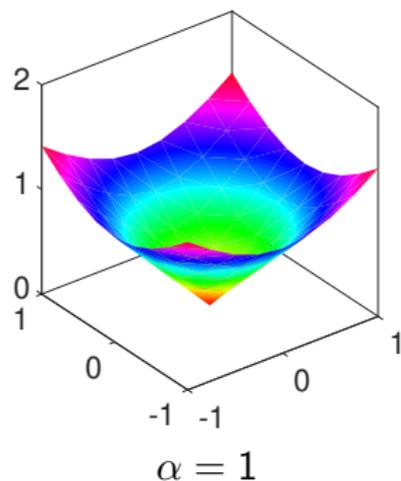
Ist f stetig differenzierbar und D konvex, so kann die Lipschitz-Konstante mit Hilfe der Norm der Jacobi-Matrix abgeschätzt werden:

$$c \leq \sup_{x \in D} \underbrace{\|f'(x)\|}_J$$

mit $\|J\| = \max_{\|z\|=1} \|Jz\|$.

Stetigkeit der radialen Funktion

$$f(x) = r^\alpha, \quad r = |x| = \sqrt{x_1^2 + \dots + x_n^2}$$



(i) $\alpha = 1$:

f ist global Lipschitz-stetig, denn

$$|r - r'| = ||x| - |x'|| \leq |x - x'|,$$

d.h. $c = 1$ ist Lipschitz-Konstante für $D = \mathbb{R}^n$

(ii) $\alpha > 1$:

f ist Lipschitz-stetig auf beschränkten Mengen D

Mittelwertsatz \implies

$$|r^\alpha - (r')^\alpha| = \alpha s^{\alpha-1} |r - r'|$$

mit s zwischen r und r'

\rightsquigarrow Lipschitz-Konstante $c = \max_{x \in D} \alpha |x|^{\alpha-1}$

(iii) $0 < \alpha < 1$:

f ist stetig aber nicht Lipschitz-stetig in einer Umgebung von $(0, 0)$, denn

$$|r^\alpha - 0^\alpha| \leq c |r - 0|$$

ist für $r \rightarrow 0$ nicht erfüllbar

(iv) $\alpha < 0$:

f ist unstetig bei 0, denn $\lim_{r \rightarrow 0} r^\alpha = \infty$

Konvergenz von Vektoren

Eine Folge von Vektoren $x_k \in \mathbb{R}^n$ konvergiert gegen einen Vektor x_* ,

$$\lim_{k \rightarrow \infty} x_k = x_* \quad \text{bzw.} \quad x_k \rightarrow x_* \text{ f\"ur } k \rightarrow \infty,$$

wenn f\"ur alle $\varepsilon > 0$ ein Index k_ε existiert mit

$$|x_k - x_*| < \varepsilon \quad \text{f\"ur } k > k_\varepsilon.$$

Mit anderen Worten enth\"alt jede ε -Umgebung

$$B_\varepsilon(x_*) = \{y : |y - x_*| < \varepsilon\}$$

alle bis auf endlich viele Folgeelemente.

\"Aquivalent zur Konvergenz von (x_k) ist die Konvergenz aller Komponenten der Folge, d.h. es k\"onnen eindimensionale Konvergenzkriterien herangezogen werden.

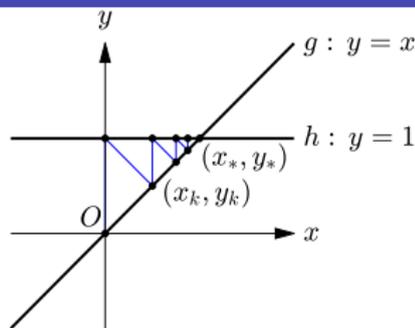
Beispiel

alternierende Projektion auf zwei Geraden

\rightsquigarrow Konvergenz der Folge $(x_k, y_k)^t$ gegen den Schnittpunkt $(x_*, y_*)^t$

Illustration für die abgebildeten Geraden

$$\begin{pmatrix} x_k \\ y_k \end{pmatrix} : \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix}, \\ \begin{pmatrix} 1/2 \\ 1 \end{pmatrix}, \begin{pmatrix} 3/4 \\ 3/4 \end{pmatrix}, \dots \rightarrow \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$



Konvergenz der ersten Komponente:

$$x_{2k} = x_{2k+1} = 1 - (1/2)^k \rightarrow 1,$$

denn

$$\left| 1 - 1 + (1/2)^k \right| = (1/2)^k < \varepsilon \quad \text{für } k > k_\varepsilon = -\lg \varepsilon$$

Cauchy-Kriterium für Vektoren

Eine Folge von Vektoren $x_k \in \mathbb{R}^n$ konvergiert genau dann, wenn sie eine Cauchy-Folge ist, d.h. wenn für alle $\varepsilon > 0$ ein k_ε existiert mit

$$|x_\ell - x_k| < \varepsilon \quad \text{für } \ell, k > k_\varepsilon.$$

Dies ist gleichbedeutend damit, dass jede der n Komponenten der Folge x_1, x_2, \dots konvergiert.

Aufgrund der Äquivalenz von Normen im \mathbb{R}^n kann anstelle der Euklidischen Norm auch jede andere Vektornorm verwendet werden.

Eine oft einfach zu verifizierende hinreichende Bedingung für das Cauchy-Kriterium ist geometrische Konvergenz. In diesem Fall gilt

$$|x_{k+1} - x_k| \leq c|x_k - x_{k-1}|$$

mit einer Konstanten $c < 1$.

Beweis

Geometrische Konvergenz \implies Cauchy-Kriterium

Dreiecksungleichung, Abschätzung für benachbarte Folgeelemente

\implies

$$\begin{aligned} |x_\ell - x_k| &= |x_\ell - x_{\ell-1} + x_{\ell-1} - x_{\ell-2} + \cdots + x_{k+1} - x_k| \\ &\leq |x_\ell - x_{\ell-1}| + |x_{\ell-1} - x_{\ell-2}| + \cdots + |x_{k+1} - x_k| \\ &\leq c\lambda^{\ell-1} + c\lambda^{\ell-2} + \cdots + c\lambda^k \\ &= c\lambda^k(1 + \lambda + \lambda^2 + \cdots) \leq \frac{c}{1-\lambda}\lambda^k = c'\lambda^k < \varepsilon \end{aligned}$$

für $\ell > k > k_\varepsilon = \ln(\varepsilon/c')/\ln \lambda$

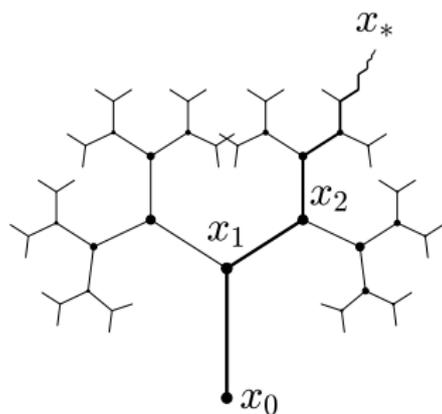
Beispiel

binärer Baum:

neue Segmente jeweils um den Faktor λ ($0 < \lambda < 1$) verkürzt und in einem Winkel von ϑ angefügt

geometrische Konvergenz sukzessiver Verzweigungspunkte x_k :

$$|x_{k+1} - x_k| = c\lambda^k$$



Cauchy-Folge, denn

$$\begin{aligned} |x_\ell - x_k| &\leq |x_\ell - x_{\ell-1}| + |x_{\ell-1} - x_{\ell-2}| + \dots + |x_{k+1} - x_k| \\ &\leq c\lambda^{\ell-1} + c\lambda^{\ell-2} + \dots + c\lambda^k \\ &= c\lambda^k (1 + \lambda + \lambda^2 + \dots) \leq \frac{c}{1-\lambda} \lambda^k = c' \lambda^k < \varepsilon \end{aligned}$$

für $\ell > k > k_\varepsilon = \ln(\varepsilon/c') / \ln \lambda$

Kontrahierende Abbildung

Eine Abbildung

$$g : D \rightarrow D, \quad D \subseteq \mathbb{R}^n,$$

ist kontrahierend, wenn in einer geeigneten Norm

$$\|g(x) - g(y)\| \leq c \|x - y\|, \quad x, y \in D,$$

mit $c < 1$ gilt.

Die Konstante c wird als Kontraktionskonstante von g bezeichnet. Sie kann für eine konvexe Menge D mit Hilfe der Jacobi-Matrix durch

$$c \leq \sup_{x \in D} \|g'(x)\|$$

abgeschätzt werden mit der Matrixnorm $\|J\| = \max_{\|x\|=1} \|Jx\|$.

Beispiel

Richardson-Iteration zur Lösung eines linearen Gleichungssystems $Ax = b$ für eine symmetrische positiv definite Matrix: A

$$x \mapsto g(x) = x - \omega(Ax - b)$$

kontrahierend, falls ω genügend klein gewählt wird

Begründung:

$$g(x) - g(y) = Q(x - y), \quad Q = E - \omega A$$

Eigenwerte von Q :

$$\varrho_k = 1 - \omega \lambda_k$$

mit $\lambda_k > 0$ den Eigenwerten von A

$$\omega = 1 / \max_k \lambda_k \quad \implies \quad \varrho_k \geq 0 \text{ und}$$

$$c = \|Q\| = \max_k \varrho_k = 1 - (\min_k \lambda_k) / (\max_k \lambda_k) < 1$$

bei Verwendung der der euklidischen Norm zugeordneten Matrixnorm $\| \cdot \|$

Banachscher Fixpunktsatz

Ist g eine kontrahierende Abbildung, die eine nicht leere, abgeschlossene Menge $D \subset \mathbb{R}^n$ in sich abbildet, d.h. gilt

- $D = \overline{D} \neq \emptyset$
- $x \in D \implies g(x) \in D$
- $\|g(x) - g(y)\| \leq c\|x - y\| \quad \forall x, y \in D$ mit $c < 1$,

dann besitzt g einen eindeutigen Fixpunkt $x_* = g(x_*) \in D$.

Ausgehend von einem beliebigen Punkt $x_0 \in D$ kann x_* durch die Iterationsfolge

$$x_0, x_1 = g(x_0), x_2 = g(x_1), \dots$$

approximiert werden.

Für den Fehler gilt

$$\|x_* - x_\ell\| \leq \frac{c^\ell}{1 - c} \|x_1 - x_0\|$$

d.h. die Iterationsfolge konvergiert für jeden Startpunkt linear.

Der Fixpunktsatz gilt allgemeiner in vollständigen metrischen Räumen. Da die Translationsinvarianz und Homogenität der Norm nicht benötigt wird, kann man $\|x - y\|$ durch eine allgemeine Abstandsfunktion $d(x, y)$ ersetzen.

Beweis

(i) $g(D) \subseteq D \implies x_\ell \in D$ für alle $\ell > 0$

(ii) Kontraktionsbedingung \implies

$$\|x_{\ell+1} - x_\ell\| = \|g(x_\ell) - g(x_{\ell-1})\| \leq c \|x_\ell - x_{\ell-1}\|$$

Iteration \rightsquigarrow

$$\|x_{\ell+1} - x_\ell\| \leq c^\ell \|x_1 - x_0\|$$

(iii) Dreiecksungleichung \implies

$$\begin{aligned} \|x_k - x_\ell\| &\leq \|x_k - x_{k-1}\| + \|x_{k-1} - x_{k-2}\| + \dots + \|x_{\ell+1} - x_\ell\| \\ &\leq (c^{k-1} + \dots + c^\ell) \|x_1 - x_0\| = \frac{c^k - c^\ell}{c-1} \|x_1 - x_0\| \\ &\leq \frac{c^\ell}{1-c} \|x_1 - x_0\| \end{aligned}$$

\rightsquigarrow Cauchy-Konvergenz der Folge x_0, x_1, \dots gegen einen Grenzwert x_*

(iv) Kontraktionsbedingung \implies

$$\|g(x_*) - x_*\| \leq \|g(x_*) - g(x_k)\| + \|g(x_k) - x_*\| \leq c \|x_* - x_k\| + \|x_{k+1} - x_*\|$$

Grenzwert $k \rightarrow \infty \implies x_*$ Fixpunkt

(v) x_* eindeutig, da

$$\|\tilde{x}_* - x_*\| = \|g(\tilde{x}_*) - g(x_*)\| \leq c \|\tilde{x}_* - x_*\|$$

mit $c < 1$

(vi) Abschätzung für den Fehler \longleftarrow Bilden des Grenzwerts für $k \rightarrow \infty$ in der Ungleichung (iii) für $\|x_k - x_\ell\|$

Gestörtes lineares System:

$$Ax + \varepsilon f(x) = b$$

mit einer quadratischen invertierbaren Matrix A und einer Lipschitz-stetigen Funktion f (Konstante c_f)

Lösung mit Hilfe der Iteration

$$x \rightarrow g(x) = A^{-1}(b - \varepsilon f(x))$$

Prüfe die Voraussetzungen des Banachschen Fixpunktsatzes für die abgeschlossene Kugel um $p = A^{-1}b$ mit Radius r ,

$$D = \{y : \|y - p\| \leq r\}, \quad p = A^{-1}b$$

(i) $g(D) \subset D$:

für $x \in D$

$$\|g(x) - p\| = \varepsilon \|A^{-1}f(x)\| \leq \varepsilon \|A^{-1}\| \max_{y \in D} \|f(y)\|$$

$g(x) \in D$ (Abstand zu $p \leq r$), falls

$$\varepsilon \leq \frac{r}{\|A^{-1}\| \max_{y \in D} \|f(y)\|}$$

(ii) Kontraktionsbedingung:

$$\|g(x) - g(y)\| = \varepsilon \|A^{-1}(f(x) - f(y))\| \leq \underbrace{\varepsilon \|A^{-1}\| c_f}_c \|x - y\|$$

mit $c < 1$, falls

$$\varepsilon < \frac{1}{\|A^{-1}\| c_f}$$

Beide Bedingungen, (i) und (ii), sind für hinreichend kleines ε erfüllt.

Partielle Ableitungen

Die partielle Ableitung $\partial_k f$ einer Funktion $f : \mathbb{R}^n \ni D \rightarrow \mathbb{R}^m$ nach der k -ten Variablen x_k ist die Ableitung der univariaten Funktion

$$x_k \mapsto f(x_1, \dots, x_k, \dots, x_n),$$

bei der die Variablen x_j , $j \neq k$, als Konstanten betrachtet werden. Man schreibt auch

$$\partial_k f = f_{x_k} = \frac{\partial f}{\partial x_k},$$

bzw. wenn x, y, z als Variablen verwendet werden, $\partial_x f = f_x = \partial f / \partial x$, $\partial_y f = f_y$, etc.. Gemäß der Definition der univariaten Ableitung gilt

$$\partial_k f(x) = \lim_{h \rightarrow 0} \frac{f(\dots, x_k + h, \dots) - f(\dots, x_k, \dots)}{h}.$$

Partielle Ableitungen sind sowohl für skalare ($m = 1$) als auch für vektorwertige ($m > 1$) reelle Funktionen definiert. In beiden Fällen bleibt der Funktionstyp beim partiellen Ableiten erhalten. Definitionsgemäß gilt

$$\partial_k \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix} = \begin{pmatrix} \partial_k f_1 \\ \vdots \\ \partial_k f_n \end{pmatrix},$$

d.h. die partielle Ableitung wird simultan in den Komponenten gebildet.

Die übliche Konvention ist, sowohl für die Variable x als auch für die Funktion f Spaltenvektoren zu verwenden. Dies ist wichtig bei der Definition der totalen Ableitung bzw. einer linearen Approximation von f .

Partielle Ableitungen für verschiedene Funktionstypen

(i) Skalare Funktionen:

skalare partielle Ableitungen

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto f(x, y) = xy^2 \quad \rightsquigarrow$$

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

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mapsto f(x_1, x_2, x_3) = x_1^3 x_2 + x_1 x_3^2 \quad \rightsquigarrow$$

$$\partial_1 f(x) = 3x_1^2 x_2 + x_3^2, \quad \frac{f(x)}{\partial x_2} = x_1^3, \quad f_{x_3}(x_1, x_2, x_3) = 2x_1 x_3$$

(Verwendung der verschiedenen alternativen Notationen)

(ii) Vektorwertige Funktion:

gleiche Komponentenzahl der partiellen Ableitungen

$$\begin{pmatrix} r \\ t \end{pmatrix} \mapsto f(r, t) = \begin{pmatrix} r \cos t \\ r \sin t \end{pmatrix} \rightsquigarrow$$

$$f_r(r, t) = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}, \quad f_t(r, t) = \begin{pmatrix} -r \sin t \\ r \cos t \end{pmatrix}$$

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mapsto \begin{pmatrix} f_1(x_1, x_2, x_3) \\ f_2(x_1, x_2, x_3) \end{pmatrix} = \begin{pmatrix} x_2^3 x_3 \\ x_1^2 + x_1 x_3^2 \end{pmatrix} \rightsquigarrow$$

$$\begin{aligned} \partial_1 f(x) &= \begin{pmatrix} \partial_1 f_1(x) \\ \partial_1 f_2(x) \end{pmatrix} = \begin{pmatrix} 0 \\ 2x_1 + x_3^2 \end{pmatrix} \\ \frac{\partial f(x_1, x_2, x_3)}{\partial x_2} &= \begin{pmatrix} 3x_2^2 x_3 \\ 0 \end{pmatrix} \\ \partial_3 f(x_1, x_2, x_3) &= \begin{pmatrix} x_2^3 \\ 2x_1 x_3 \end{pmatrix} \end{aligned}$$

Mehrfache partielle Ableitungen

Zweifache (hintereinander ausgeführte) partielle Ableitungen werden mit

$$\partial_j \partial_k f = f_{x_k x_j} = \frac{\partial^2 f}{\partial x_j \partial x_k}$$

bezeichnet. Analog schreibt man $\partial_j \partial_k \partial_\ell \dots f$ für partielle Ableitungen höherer Ordnung.

Sind die partiellen Ableitungen stetig, so spielt die Reihenfolge der Differentiation keine Rolle. In diesem (Normal)fall kann man alternativ die Multiindex-Notation

$$\partial^\alpha f = \partial_1^{\alpha_1} \dots \partial_n^{\alpha_n} f, \quad \alpha = (\alpha_1, \dots, \alpha_n),$$

verwenden, wobei der Index $\alpha_k \in \mathbb{N}_0$ die Anzahl der partiellen Ableitungen nach der k -ten Variablen bezeichnet. Die Summe $|\alpha| = \alpha_1 + \dots + \alpha_n$ ist die Ordnung der partiellen Ableitung. Beispielsweise ist für eine glatte Funktion f von 3 Variablen $\partial^{(2,1,3)} f = \partial_1^2 \partial_2 \partial_3^3 f$ eine partielle Ableitung der Ordnung $|(2, 1, 3)| = 2 + 1 + 3 = 6$.

Partielle Ableitungen höherer Ordnung verschiedener Funktionstypen

(i) $f(x, y) = x/y$:

$$f_x = 1/y, \quad f_y = -x/y^2, \quad f_{xx} = 0, \quad f_{yy} = 2x/y^3, \quad f_{xy} = f_{yx} = -1/y^2$$

(ii) $f(x_1, x_2) = \begin{pmatrix} \cos(x_1 x_2) \\ \sin(x_1 + x_2) \end{pmatrix}$:

$$\partial_1 f = \begin{pmatrix} -x_2 \sin(x_1 x_2) \\ \cos(x_1 + x_2) \end{pmatrix}, \quad \partial_2 \partial_1 f = \begin{pmatrix} -x_1 x_2 \cos(x_1 x_2) \\ -\sin(x_1 + x_2) \end{pmatrix}$$

(iii) $f(x_1, x_2, x_3) = \exp(x_1 + 2x_2 + 3x_3)$:

$$\partial_1 f = f, \quad \partial_2 f = 2 \exp(x_1 + 2x_2 + 3x_3) = 2f, \quad \partial_3 f = 3f \quad \rightsquigarrow$$

$$\partial^{(3,1,2)} f = \partial_1^3 \partial_2 \partial_3^2 f = 1^3 \cdot 2 \cdot 3^2 f = 18f$$

Beispiel

Partielle Ableitungen der Funktion

$$f(x_1, \dots, x_n) = \exp(i\omega^t x) = \exp(i(\omega_1 x_1 + \dots + \omega_n x_n))$$

(ebene Welle)

Kettenregel \rightsquigarrow

$$\begin{aligned}\partial_k f(x) &= i\omega_k \exp(i\omega^t x) \\ \partial_\ell \partial_k f(x) &= (i\omega_\ell)(i\omega_k) \exp(i\omega^t x)\end{aligned}$$

und somit

$$\partial^\alpha f(x) = (i\omega)^\alpha \exp(i\omega^t x) = i^{|\alpha|} \omega^\alpha \exp(i\omega^t x)$$

mit $\alpha = (\alpha_1, \dots, \alpha_n)$ und $\omega^\alpha = \omega_1^{\alpha_1} \dots \omega_n^{\alpha_n}$

z.B. ($n = 2$):

$$\partial^{(3,4)} f(x) = \underbrace{i^7}_{-i} \omega_1^3 \omega_2^4 \exp(i(\omega_1 x_1 + \omega_2 x_2))$$

Partielle Ableitungen von multivariaten Polynomen

Die partielle Ableitung

$$\partial^\alpha x^\beta = \partial_1^{\alpha_1} \dots \partial_n^{\alpha_n} \left(x_1^{\beta_1} \dots x_n^{\beta_n} \right)$$

eines Monoms ist nur dann ungleich Null, wenn

$\alpha \leq \beta \iff \alpha_k \leq \beta_k \forall k$, und in diesem Fall gleich

$$c x^{\beta-\alpha}, \quad c = \prod_{k=1}^n \beta_k (\beta_k - 1) \dots (\beta_k - \alpha_k + 1) = \beta! / \alpha!$$

mit $(j, k, \dots)! = j! k! \dots$. Insbesondere ist

$$\partial^\alpha x^\beta |_{x=0} = \alpha! \delta_{\alpha, \beta}.$$

Gilt für ein Polynom $p(x) = \sum_{\beta} c_{\beta} x^{\beta}$

$$\partial^\alpha p \equiv 0 \quad \forall \alpha \text{ mit } |\alpha| = \alpha_1 + \dots + \alpha_n = m,$$

so hat p totalen Grad $< m$, d.h. $c_{\beta} = 0$ für $|\beta| \geq m$.

Beweis

zu zeigen:

$$p = \sum_{\beta} c_{\beta} x^{\beta}, \partial^{\alpha} p \equiv 0 \quad \forall \alpha \text{ mit } |\alpha| = m \quad \implies \quad c_{\beta} = 0 \quad \forall \beta \text{ mit } |\beta| \geq m$$

Widerspruchsannahme: $c_{\beta} \neq 0$ für ein β mit $|\beta| \geq m$

wähle α mit $|\alpha| = m$ und $\alpha \leq \beta \quad \rightsquigarrow$

$$\partial^{\alpha} p(x) = \sum_{\beta' \geq \alpha} c_{\beta'} (\beta'! / \alpha!) x^{\beta' - \alpha} \neq 0,$$

da mindestens der Summand mit $\beta' = \beta$ nicht verschwindet im Gegensatz zu der Annahme, dass alle partiellen Ableitungen der Ordnung m identisch Null sind

Vertauschbarkeit partieller Ableitungen

Sind die ersten und zweiten partiellen Ableitungen einer Funktion $f : D \ni \mathbb{R}^n \rightarrow \mathbb{R}$ stetig, so gilt

$$\partial_j \partial_k f = \partial_k \partial_j f .$$

Für hinreichend glatte Funktionen ist also die Reihenfolge partieller Ableitungen vertauschbar. Insbesondere rechtfertigt dies die Multiindex-Schreibweise.

Angewandt auf die Komponenten gilt die Aussage ebenfalls für vektorwertige Funktionen $f : D \ni \mathbb{R}^n \rightarrow \mathbb{R}^m$.

Beweis

Variablen x_ℓ , $\ell \neq j, k$, irrelevant für die partielle Ableitungen ∂_j , ∂_k
 \rightsquigarrow betrachte o.B.d.A. eine bivariate Funktion $f(x, y)$

bezeichne mit

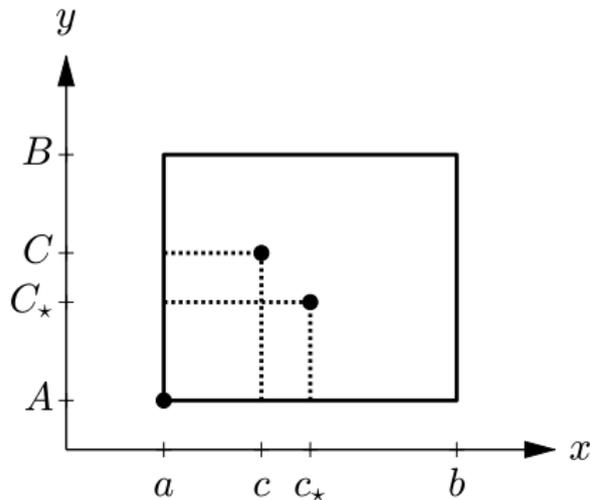
$$g(x) = f(x, B) - f(x, A),$$

$$h(y) = f(b, y) - f(a, y)$$

die Differenzen in y - bzw. x -
Richtung und berechne

$$Q = f(b, B) - f(b, A) - f(a, B) + f(a, A)$$

mit Hilfe des eindimensionalen
Mittelwertsatzes (MWS) auf zwei
verschiedene Arten



$$\begin{aligned}
Q &= [f(b, B) - f(b, A)] - [f(a, B) - f(a, A)] \\
&= g(b) - g(a) \\
&\stackrel{\text{MWS}}{=} (b - a)g_x(c) \\
&= (b - a)[f_x(c, B) - f_x(c, A)] \\
&\stackrel{\text{MWS}}{=} (b - a)(B - A)f_{xy}(c, C)
\end{aligned}$$

für ein $c \in (a, b)$ und $C \in (A, B)$

$$\begin{aligned}
Q &= [f(b, B) - f(a, B)] - [f(b, A) - f(a, A)] \\
&= h(B) - h(A) \\
&\stackrel{\text{MWS}}{=} (B - A)h_y(C^*) \\
&= (B - A)[f_y(b, C^*) - f_y(a, C^*)] \\
&\stackrel{\text{MWS}}{=} (B - A)(b - a)f_{yx}(c^*, C^*)
\end{aligned}$$

für ein $c^* \in (a, b)$ und $C^* \in (A, B)$

Gleichsetzen \implies

$$f_{xy}(c, C) = f_{yx}(c^*, C^*)$$

Verkleinerung des Rechtecks durch Grenzübergang ($b \rightarrow a, B \rightarrow A$) \implies

$$f_{xy}(a, A) = f_{yx}(a, A)$$

aufgrund der Stetigkeit der gemischten zweiten Ableitungen

Totale Ableitung und Jacobi-Matrix

Eine Funktion $f : D \ni \mathbb{R}^n \rightarrow \mathbb{R}^m$ ist in einem Punkt x differenzierbar, wenn

$$f(x+h) = f(x) + f'(x)h + o(|h|)$$

für $|(h_1, \dots, h_n)| \rightarrow 0$.

Die Ableitung f' ist die Jacobi-Matrix der partiellen Ableitungen:

$$f' = (\partial_1 f, \dots, \partial_n f) = \begin{pmatrix} \partial_1 f_1 & \cdots & \partial_n f_1 \\ \vdots & & \vdots \\ \partial_1 f_m & \cdots & \partial_n f_m \end{pmatrix}$$

und gemäß den Regeln des Matrix/Vektor-Kalküls ist

$$f'(x)h = \partial_1 f(x)h_1 + \cdots + \partial_n f(x)h_n.$$

Hinreichend für die Existenz der Ableitung $f'(x)$ ist die Stetigkeit der partiellen Ableitungen in einer Umgebung von x .

Alternative gebräuchliche Schreibweisen sind

$$f' = Jf = \frac{\partial(f_1, \dots, f_m)}{\partial(x_1, \dots, x_n)}.$$

Für eine skalare Funktion ($m = 1$) bezeichnet man die Ableitung als Gradient,

$$(\partial_1 f, \dots, \partial_n f) = f' = (\text{grad } f)^t = \begin{pmatrix} \partial_1 f \\ \vdots \\ \partial_n f \end{pmatrix}^t.$$

Dabei ist zu beachten, dass die $n \times 1$ -Jacobi-Matrix ein Zeilen- und der Gradient ein Spaltenvektor ist; deshalb ist die Transposition t notwendig.

Für die Parametrisierung einer Kurve $t \mapsto f(t) = (f_1(t), \dots, f_m(t))^t$ ($n = 1$) bezeichnet man den m -Vektor $f'(t)$ als Tangentenvektor.

Um die lineare Approximation kleiner Änderungen ($|h| \rightarrow 0$) hervorzuheben, schreibt man

$$df = \frac{\partial f}{\partial x_1} dx_1 + \dots + \frac{\partial f}{\partial x_n} dx_n$$

mit sogenannten Differentialen df und dx_k .

Beweis

betrachte bivariate Funktionen (analoge Argumentation im multivariaten Fall)

(i) zeige: Existenz von $f' \implies f' = (\partial_1 f, \dots, \partial_n f)$

Für eine bivariate Funktion $f(x, y)$ und $h = (s, 0)^t$ gilt

$$\begin{aligned} f(x + s, y) &= f(x, y) + f'(x, y) (s, 0)^t + o(|(s, 0)|) \\ &= f(x, y) + (Jf)_1 s + o(|s|), \end{aligned}$$

wobei $(Jf)_1$ die erste Spalte von Jf bezeichnet.

Division durch s und Bilden des Grenzwerts für $s \rightarrow 0 \rightsquigarrow$

$$(Jf)_1 = \lim_{s \rightarrow 0} \frac{f(x + s, y) - f(x, y)}{s} + \lim_{s \rightarrow 0} \frac{o(|s|)}{s} = \partial_1 f(x, y) + 0$$

analoge Argumentation für $(Jf)_2$

(ii) zeige: Stetigkeit von $\partial_1 f, \dots, \partial_n f \implies$ Existenz von f'

betrachte eine skalare Funktion $f(x, y)$

Mittelwertsatz \implies

$$\begin{aligned} f(x + s, y + t) &= [f(x + s, y) - f(x, y)] + [f(x + s, y + t) - f(x + s, y)] \\ &\quad + f(x, y) \\ &= s f_x(\xi, y) + t f_y(x + s, \eta) + f(x, y) \end{aligned}$$

mit $\xi \in (x, x + s)$, $\eta \in (y, y + t)$

Stetigkeit von f_x und $f_y \implies$

$$f_x(\xi, y) = f_x(x, y) + o(|s|), \quad f_y(x + s, \eta) = f_y(x, y) + o(|(s, t)|)$$

für $\underbrace{|(s, t)|}_h \rightarrow 0$ und damit die Differenzierbarkeit der Funktion f :

$$f(x + s, y + t) = f(x, y) + s f_x(x, y) + t f_y(x, y) + o(|h|)$$

separates Betrachten der Komponenten \rightsquigarrow vektorwertiger Fall

Differenzierbarkeit der Funktion

$$f(x, y) = \frac{xy}{x^2 + y^2}$$

(i) Unstetigkeit im Ursprung:

$$\lim_{x \rightarrow 0} f(x, x) = \frac{x^2}{x^2 + x^2} = \frac{1}{2} \neq -\frac{1}{2} = \lim_{x \rightarrow 0} f(x, -x)$$

$\implies f$ nicht stetig und damit auch nicht differenzierbar bei $(0, 0)$

(ii) Existenz der partiellen Ableitungen:

$$f(x, 0) = f(0, y) = 0$$

\implies

$$f_x(0, 0) = f_y(0, 0) = 0$$

(iii) Unstetigkeit beider partieller Ableitungen im Ursprung:
bestimme den Grenzwert der partiellen Ableitung

$$f_x(x, y) = \frac{y^3 - x^2 y}{(x^2 + y^2)^2}$$

für $(x, y) \rightarrow (0, 0)$ entlang verschiedener Kurven:

$$y = x \quad \rightsquigarrow \quad \lim_{x \rightarrow 0} f_x(x, x) = \lim_{x \rightarrow 0} \frac{0}{(x^2 + x^2)^2} = 0$$

und

$$y = x^2 \quad \rightsquigarrow \quad \lim_{x \rightarrow 0} f_x(x, x^2) = \lim_{x \rightarrow 0} \frac{x^6 - x^4}{(x^2 + x^4)^2} = \lim_{x \rightarrow 0} \frac{x^2 - 1}{(1 + x^2)^2} = -1$$

Die Unstetigkeit der partiellen Ableitung f_y im Ursprung folgt analog.

Beispiel \implies

Die Existenz der partiellen Ableitungen ist nicht ausreichend für
Differenzierbarkeit.

Beispiel

Überprüfung der Definition der Ableitung/Jacobi-Matrix für

$$f(x, y) = \begin{pmatrix} x + 2y \\ xy \\ 3x^2 + y^2 \end{pmatrix}, \quad f'(x, y) = (f_x(x, y), f_y(x, y)) = \begin{pmatrix} 1 & 2 \\ y & x \\ 6x & 2y \end{pmatrix}$$

berechne den Fehler $R = f(x + s, y + t) - f(x, y) - f'(x, y)(s, t)^t$

$$f(x + s, y + t) = \begin{pmatrix} (x + s) + 2(y + t) \\ (x + s)(y + t) \\ 3(x + s)^2 + (y + t)^2 \end{pmatrix}, \quad f'(x, y) \begin{pmatrix} s \\ t \end{pmatrix} = \begin{pmatrix} s + 2t \\ ys + xt \\ 6xs + 2yt \end{pmatrix}$$

$$\implies R = (0, st, 3s^2 + t^2)^t = o(\|(s, t)\|),$$

d.h. jede Komponente strebt für $(s, t) \rightarrow (0, 0)$ schneller gegen 0 als $|h| = \|(s, t)\|$

Jacobi-Matrizen verschiedener Dimension

(i) Gradient der skalaren Funktion $(x, y)^t \mapsto r(x, y) = \sqrt{x^2 + y^2}$:
 $r_x = (1/2)(x^2 + y^2)^{-1/2}(2x) = x/r$, Symmetrie $\implies r_y = y/r$ und
folglich

$$r' = (r_x, r_y) = (x/r, y/r) = (\text{grad } r)^t$$

(ii) Tangentenvektor, der durch

$$c(t) = \begin{pmatrix} \cos t \\ \sin t \\ t \end{pmatrix}$$

parametrisierten Kurve (Schraubenlinie):

$$c'(t) = \begin{pmatrix} -\sin t \\ \cos t \\ 1 \end{pmatrix}$$

(iii) Jacobi-Matrix, der durch

$$\begin{pmatrix} \vartheta \\ \varphi \end{pmatrix} \xrightarrow{P} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \sin \vartheta \cos \varphi \\ \sin \vartheta \sin \varphi \\ \cos \vartheta \end{pmatrix}$$

definierten Parametrisierung der Einheitssphäre:

$$P' = \frac{\partial(x, y, z)}{\partial(r, \vartheta, \varphi)} = \begin{pmatrix} x_{\vartheta} & x_{\varphi} \\ y_{\vartheta} & y_{\varphi} \\ z_{\vartheta} & z_{\varphi} \end{pmatrix} = \begin{pmatrix} \cos \vartheta \cos \varphi & -\sin \vartheta \sin \varphi \\ \cos \vartheta \sin \varphi & \sin \vartheta \cos \varphi \\ -\sin \vartheta & 0 \end{pmatrix}$$

Multivariate Kettenregel

Für die Hintereinanderschaltung

$$h = g \circ f : x \mapsto y = f(x) \mapsto z = g(y) = h(x),$$

stetig differenzierbarer Funktionen $f : \mathbb{R}^n \supseteq U \rightarrow \mathbb{R}^\ell$ und $g : \mathbb{R}^\ell \supseteq V \rightarrow \mathbb{R}^m$ mit $f(U) \subseteq V$ gilt

$$h'(x) = \underbrace{g'(y)}_{m \times \ell} \underbrace{f'(x)}_{\ell \times n},$$

d.h. die $m \times n$ -Jacobi-Matrix von h ist das Produkt der Jacobi-Matrizen von f und g . Die einzelnen Einträge von h' ergeben sich durch Matrixmultiplikation:

$$\frac{\partial h_i}{\partial x_k} = \sum_{j=1}^{\ell} \frac{\partial g_i}{\partial y_j} \frac{\partial f_j}{\partial x_k}, \quad i = 1, \dots, m, k = 1, \dots, n.$$

Diese Identität vereinfacht sich, wenn eine oder zwei der Dimensionen gleich eins sind. Beispielsweise hat eine multivariate Funktion g entlang einer Kurve mit Parametrisierung f ,

$$x \mapsto h(x) = g(f_1(x), \dots, f_\ell(x)),$$

die Ableitung

$$\frac{dh}{dx} = \partial_1 g(f(x)) f_1'(x) + \dots + \partial_\ell g(f(x)) f_\ell'(x) = (\text{grad } g)^t|_{f(x)} f'(x),$$

d.h. $h'(x)$ ist das Skalarprodukt aus Gradient von g und Tangentenvektor von f .

Für Funktionen von zwei oder drei Veränderlichen werden oft statt der Index-Schreibweise verschiedene Buchstaben für die Variablen verwendet. Beispielsweise ist die Jacobi-Matrix der Funktion

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} p(u(x, y), v(x, y)) \\ q(u(x, y), v(x, y)) \end{pmatrix}$$

gemäß der Kettenregel das Matrixprodukt

$$\begin{pmatrix} p_u & p_v \\ q_u & q_v \end{pmatrix} \Big|_{(u(x,y), v(x,y))} \begin{pmatrix} u_x & u_y \\ v_x & v_y \end{pmatrix} \Big|_{(x,y)}.$$

Beweis

Definition der Ableitung und Jacobi-Matrix:

$$\varphi(x + \Delta x) = \varphi(x) + \varphi'(x)\Delta x + o(|\Delta x|)$$

Existenz der Ableitungen $f'(x)$ und $g'(y)$ \implies

$$g(f(x+\Delta x)) = g(\underbrace{f(x) + f'(x)\Delta x + o(|\Delta x|)}_{\Delta y}) = g(y) + [g'(y) \Delta y] + o(|\Delta y|)$$

\rightsquigarrow Formel für die Jacobi-Matrix von $h = g \circ f$, da

$$[g'(y) \Delta y] = \underbrace{g'(y)f'(x)}_{h'(x)} \Delta x + o(|\Delta x|)$$

und $|\Delta y| = O(|\Delta x|)$

Beispiel

Jacobi-Matrix der Hintereinanderschaltung der Funktionen

$$y = f(x) = \begin{pmatrix} x_3 \sin x_1 \\ e^{x_2}/x_3 \end{pmatrix}, \quad g(y) = \begin{pmatrix} y_1 \\ y_1 + \ln y_2 \\ 0 \\ y_2 \cos y_1 \end{pmatrix}$$

für $x = p = (\pi, 0, 1)^t$

$$f'(x)|_{x=p} = \begin{pmatrix} x_3 \cos x_1 & 0 & \sin x_1 \\ 0 & e^{x_2}/x_3 & -e^{x_2}/x_3^2 \end{pmatrix} \Big|_{x=p} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & -1 \end{pmatrix}$$

$$g'(y)|_{y=f(p)} = \begin{pmatrix} 1 & 0 \\ 1 & 1/y_2 \\ 0 & 0 \\ -y_2 \sin y_1 & \cos y_1 \end{pmatrix} \Big|_{y=(0,1)} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}$$

Kettenregel $\rightsquigarrow h'(\pi, 0, 1) = g'(0, 1) f'(\pi, 0, 1)$ und nach Einsetzen

$$h'(\pi, 0, 1) = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & -1 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 \\ -1 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & 1 & -1 \end{pmatrix}$$

Beispiel

Kettenregel für eine skalare Funktion $f(x, y)$ entlang einer Kurve
 $t \mapsto (x(t), y(t))^t$

Spezialisierung der allgemeinen Formel \rightsquigarrow

$$\frac{d}{dt}f(x(t), y(t)) = f_x(x(t), y(t))x'(t) + f_y(x(t), y(t))y'(t)$$

(Skalarprodukt von Gradient und Tangentenvektor)

Anwendung im konkreten Fall

$$f(x, y) = xy^2, \quad x = \cos t, \quad y = \sin t \text{ (Kreis)}$$

Gradient: $(f_x, f_y)^t = (y^2, 2xy)^t$, Auswertung entlang der Kurve \rightsquigarrow
 $(\sin^2 t, 2 \cos t \sin t)^t$

Tangentenvektor: $(x', y')^t = (-\sin t, \cos t)^t$

Einsetzen in die Kettenregel \rightsquigarrow

$$\begin{aligned}\frac{d}{dt}f(x(t), y(t)) &= (\sin^2 t, 2 \cos t \sin t) \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix} \\ &= -\sin^3 t + 2 \cos^2 t \sin t = 2 \sin t - 3 \sin^3 t\end{aligned}$$

Vergleich mit der direkten Berechnung:

$$f(x(t), y(t)) = \cos t \sin^2 t \quad \implies$$

$$\frac{d}{dt}f(x(t), y(t)) = -\sin t(\sin^2 t) + \cos t(2 \sin t \cos t) \quad \checkmark$$

Beispiel

Berechnung des Gradienten der Funktion $h = g \circ f$ für

$$f(x, y) = \begin{pmatrix} x + y \\ x - y \\ x^2 + y^2 - 1 \end{pmatrix}, \quad g(u, v, w) = u^2 + v^2 + w^2$$

Jacobi-Matrix von f

$$f' = (f_x, f_y) = \begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 2x & 2y \end{pmatrix}$$

Gradient von g

$$(\text{grad } g)^t = g' = (2u, 2v, 2w) = 2(x + y, x - y, x^2 + y^2 - 1)$$

Kettenregel, $h'(x, y) = g'(f(x, y))f'(x, y) \implies$

$$\underbrace{(\text{grad } h)^t}_{h'} = \underbrace{(\text{grad } g)^t}_{g'} f' = 2 \begin{pmatrix} x + y + x - y + 2x(x^2 + y^2 - 1) \\ x + y - x + y + 2y(x^2 + y^2 - 1) \end{pmatrix}^t$$

und nach Vereinfachung

$$\text{grad } h = 4(x^2 + y^2) \begin{pmatrix} x \\ y \end{pmatrix}$$

Transformation von Gradienten bei affiner Abbildung eines Referenzdreiecks

Parametrisierung eines allgemeinen Dreiecks D mit Eckpunkten $a = (a_1, a_2)^t$, b , c , ausgehend von dem Referenzdreieck $D_* : x_1 + x_2 \leq 1, x_k \geq 0$

$$y = p(x) = a + (b - a)x_1 + (c - a)x_2$$

Jacobi-Matrix

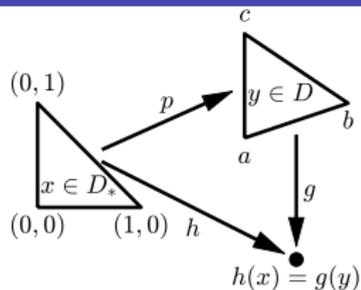
$$p'(x) = \frac{\partial(y_1, y_2)}{\partial(x_1, x_2)} = (b - a, c - a) = \begin{pmatrix} b_1 - a_1 & c_1 - a_1 \\ b_2 - a_2 & c_2 - a_2 \end{pmatrix}$$

Kettenregel \implies

$$(\text{grad } h(x))^t = (\text{grad } g(y))^t p'(x), \quad h(x) = g(p(x))$$

für skalare Funktionen g und h

Anwendung: Aufstellen von Steifigkeitsmatrizen für Finite Elemente



Richtungsableitung

Die Ableitung einer Funktion $f : \mathbb{R}^n \ni D \rightarrow \mathbb{R}$ in Richtung eines Vektors $v = (v_1, \dots, v_n)^t$ im Punkt x ist die Steigung der univariaten Funktion $t \mapsto f(x + tv)$ an der Stelle $t = 0$:

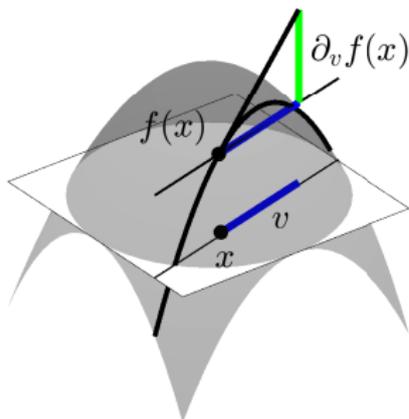
$$\partial_v f(x_1, \dots, x_n) = \lim_{t \rightarrow 0} \frac{f(x + tv) - f(x)}{t} = \left(\frac{d}{dt} f(x + tv) \right)_{t=0}.$$

Speziell ist $\partial_{e_k} f$ mit e_k dem k -ten Einheitsvektor die partielle Ableitung bzgl. der k -ten Koordinate.

Aufgrund der Kettenregel gilt

$$\partial_v f(x) = (\text{grad } f(x))^t v = \partial_1 f(x)v_1 + \dots + \partial_n f(x)v_n.$$

Die lokale Änderung von f ist somit maximal (minimal) für $v = s \text{ grad } f(x)$ mit $s > 0$ ($s < 0$).



Die Richtungsableitung kann allgemeiner auch für eine vektorwertige Funktion $f : \mathbb{R}^n \ni D \rightarrow \mathbb{R}^m$ definiert werden. Bei der Berechnung ist dann der Gradient durch die Jacobi-Matrix zu ersetzen:

$$\partial_v f(x) = f'(x) v.$$

Beispiel

Berechnung der Richtungsableitung der Funktion

$$f(x, y) = x^2 y^3$$

im Punkt $(x, y) = (2, -1)$

Gradient

$$\text{grad } f(x, y)|_{(2, -1)} = \left(\begin{array}{c} 2xy^3 \\ 3x^2y^2 \end{array} \right) \Big|_{(2, -1)} = \left(\begin{array}{c} -4 \\ 12 \end{array} \right)$$

↪ Richtungsableitung

$$\partial_v f(2, -1) = (-4, 12) \left(\begin{array}{c} v_1 \\ v_2 \end{array} \right) = -4v_1 + 12v_2$$

maximal für

$$v = s \left(\begin{array}{c} -4 \\ 12 \end{array} \right) \quad \text{mit } s > 0,$$

also z.B. für $v = (-1, 3)$

Näherung (lineare Taylor-Approximation) für den größten lokalen Anstieg von f :

$$\begin{aligned} f(2-t, -1+3t) &\approx f(2, -1) + t \partial_v f(2, -1) \\ &= -4 + t(-4, 12) \begin{pmatrix} -1 \\ 3 \end{pmatrix} \\ &= -4 + 40t \end{aligned}$$

für kleine Werte von t

Tangente

Der Tangentenvektor einer mit einer stetig differenzierbaren Funktion

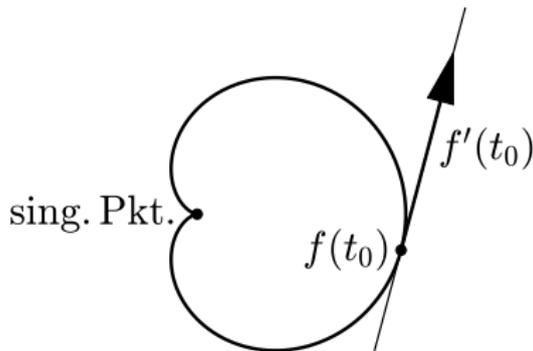
$$f : t \mapsto (f_1(t), \dots, f_n(t))^t$$

parametrisierten Kurve im Punkt $f(t_0)$ ist die Ableitung $f'(t_0)$, falls mindestens eine der Komponenten $f'_k(t_0)$ ungleich Null ist. Die Tangente ist in diesem Normalfall die durch

$$f(t_0) + f'(t_0)(t - t_0), \quad t \in \mathbb{R},$$

parametrisierte Gerade.

Ist $f'(t_0)$ der Nullvektor, so ist die Parametrisierung bei t_0 singular. Ein Tangentenvektor kann, muss aber nicht existieren, denn die Tangentenrichtung kann sich im Punkt $f(t_0)$ abrupt ändern.



Beispiel

Tangentenvektor und Tangente für die Schraubenlinie

$$C : t \mapsto f(t) = (\cos t, \sin t, t)^t$$

(i) Tangentenvektor:

$$f'(t) = (-\sin t, \cos t, 1)^t$$

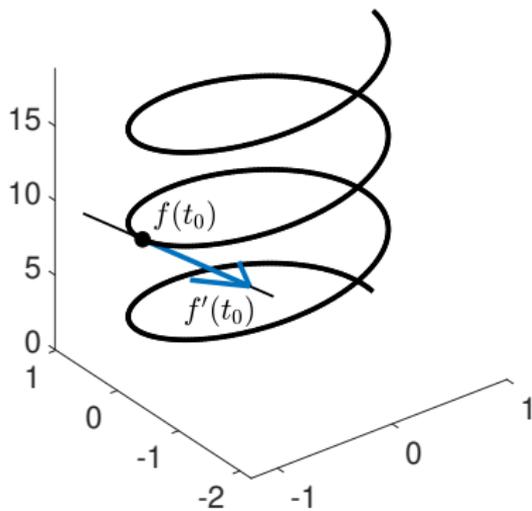
(ii) Tangente im Punkt $f(3\pi)$:

$$f(3\pi) = (-1, 0, 3\pi)^t$$

$$f'(3\pi) = (0, -1, 1)^t$$

↪ Parametrisierung

$$t \mapsto \begin{pmatrix} -1 \\ 0 \\ 3\pi \end{pmatrix} + \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} (t - 3\pi)$$



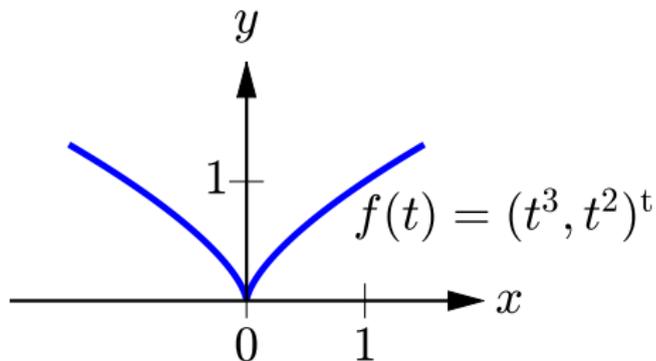
Beispiel

Stetige und un stetige Änderung des Tangentenvektors an einem singulären Kurvenpunkt

(i) $t \mapsto f(t) = (t^3, t^2)^t$:

abrupte Richtungsänderung für $t_0 = 0$ von $(0, -1)^t$ nach $(0, 1)^t$

möglich, da $f'(0) = (0, 0)^t$



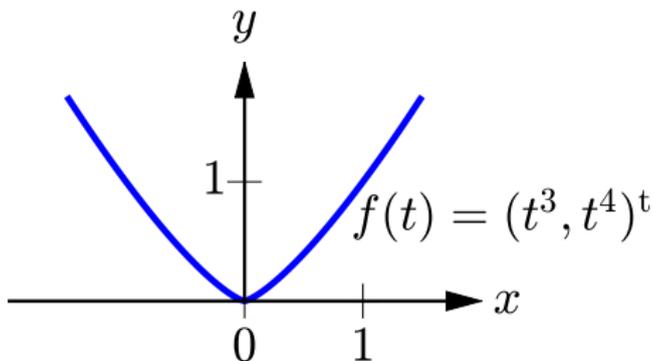
$$(ii) t \mapsto g(t) = (t^3, t^4)^t$$

stetige Tangentenänderung trotz $f'(0) = (0, 0)^t$

Umparametrisierung

$$s = t^3, \quad f(t) = g(s) = \begin{pmatrix} s \\ s^{4/3} \end{pmatrix}$$

$$\implies g'(s) = (1, 4s^{1/3}/3)^t \text{ stetig bei } s = 0$$



Tangentialebene

Die Tangentialebene im Punkt $p = (p_1, \dots, p_n)^t$ einer durch

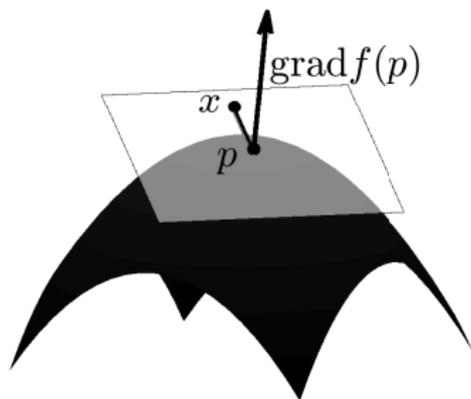
$$S : f(x_1, \dots, x_n) = c$$

implizit definierten Fläche besitzt die Darstellung

$$E : 0 = (\text{grad } f(p))^t (x - p) = \sum_{k=1}^n \partial_k f(p)(x_k - p_k),$$

falls mindestens eine der Komponenten $\partial_k f(p)$ des Gradienten ungleich Null ist. Der Normalenvektor von E ist also parallel zu $\text{grad } f$.

Ist $\text{grad } f(p) = (0, \dots, 0)^t$, so muss eine Tangentialebene im Punkt p nicht existieren. Beispielsweise kann die Fläche eine Kante oder Spitze haben.



Für den Graph einer Funktion $x \mapsto x_n = g(x_1, \dots, x_{n-1})$ ist

$$E : x_n - g(q) = \sum_{k=1}^{n-1} \partial_k g(q)(x_k - q_k)$$

die Gleichung der Tangentialebene im Punkt $(q_1, \dots, q_{n-1}, g(q))^t$. Die partielle Ableitung $\partial_k g(q)$ entspricht somit der Steigung der Tangentialebene in Richtung der k -ten Koordinatenachse. Die Normale der Tangentialebene ist parallel zu $(-\partial_1 g(q), \dots, -\partial_{n-1} g(q), 1)^t$.

Wird eine Fläche $S \subset \mathbb{R}^n$ durch eine Parametrisierung beschrieben,

$$S : (s_1, \dots, s_{n-1})^t \mapsto (h_1(s), \dots, h_n(s))^t,$$

so spannen die partiellen Ableitungen $\partial_k h(s^*)$ die Tangentialebene E im Punkt $p = h(s^*)$ auf, d.h.

$$E : p + \sum_{k=1}^{n-1} s_k \partial_k h(p), \quad s_k \in \mathbb{R}.$$

Beweis

(i) Implizite Darstellung einer Fläche durch $f(x_1, \dots, x_n) = c$:

Definition der totalen Ableitung $f' = (\text{grad } f)^t \implies$

$$f(x) = f(p) + \text{grad}f(p)^t(x - p) + o(|x - p|)$$

Vernachlässigung des Terms $o(|x - p|)$, $f(x) = f(p) = c \rightsquigarrow$

Gleichung der Tangentialebene

(ii) Darstellung einer Fläche als Funktionsgraph $y = g(x_1, \dots, x_{n-1})$:

$$\Delta y = \sum_{i=1}^{n-1} \partial_i g(q)(\Delta x_i) + o(|\Delta x|)$$

mit $\Delta y = y - g(q)$ und $\Delta x = x - q$

Vernachlässigung des Restgliedes \rightsquigarrow Darstellung der Tangentialebene

Tangentialebenen für den Kegel

$$K : f(x, y, z) = x^2 + y^2 - z^2 = 0$$

$\text{grad } f(x, y, z) = (2x, 2y, -2z)^t \rightsquigarrow$ Tangentialebene im Punkt (x_0, y_0, z_0)

$$E : 2x_0(x - x_0) + 2y_0(y - y_0) - 2z_0(z - z_0) = 0$$

Tangentialebene im Punkt $(x_0, y_0, z_0) = (3, 4, 5)$

$$E : 0 = 6(x - 3) + 8(y - 4) - 10(z - 5) = 6x + 8y - 10z$$

(Jede) Tangentialebene enthält den Ursprung.

$\text{grad } f(x_0, y_0, z_0) = (0, 0, 0)^t$ für $(x_0, y_0, z_0) = (0, 0, 0)$

\rightsquigarrow keine Tangentialebene an der Spitze des Kegels

Beispiel

Tangentialebenen für den den Funktionsgraph von

$$g(x) = |x - p|^{-1} - |x + p|^{-1}, \quad x = (x_1, x_2)$$

(Potential eines Dipols mit Ladungen in den Punkten $\pm p = \pm(p_1, p_2)$)

$$\partial_k (x_1^2 + x_2^2)^{-1/2} = (-1/2)(x_1^2 + x_2^2)^{-3/2}(2x_k) \rightsquigarrow$$

$$\partial_k g(x) = \frac{x_k + p_k}{|x + p|^3} - \frac{x_k - p_k}{|x - p|^3}, \quad k = 1, 2$$

Gleichung der Tangentialebene E im Punkt $x = q = (q_1, q_2)$

$$\begin{aligned} x_3 - g(q) &= \sum_{k=1}^2 \underbrace{\left(\frac{q_k + p_k}{|q + p|^3} - \frac{q_k - p_k}{|q - p|^3} \right)}_{\partial_k g(q)} (x_k - q_k) \\ &= \frac{(q + p)((x_1, x_2) - q)^t}{|q + p|^3} - \frac{(q - p)((x_1, x_2) - q)^t}{|q - p|^3} \end{aligned}$$

z.B. für $q = (0, 0)$

$$g(q) = 0, \quad E : x_3 = 2 \frac{p_1 x_1 + p_2 x_2}{|p|^3}$$

Tangentialebene verläuft durch den Ursprung und enthält die Gerade senkrecht zu p

keine Tangentialebenen in den Singularitäten ($x = \pm p$)

Tangentialebene des Hyperboloids

$$H : (\varphi, z) \mapsto h(\varphi, z) = (\sqrt{1+z^2} \cos \varphi, \sqrt{1+z^2} \sin \varphi, z)^t$$

im Punkt $p = (1, -1, 1) = h(-\pi/4, 1)^t$

partielle Ableitungen der Parametrisierung

$$h_\varphi(\varphi, z) = \begin{pmatrix} -\sqrt{1+z^2} \sin \varphi \\ \sqrt{1+z^2} \cos \varphi \\ 0 \end{pmatrix}, \quad h_z(\varphi, z) = \begin{pmatrix} \frac{z}{\sqrt{1+z^2}} \cos \varphi \\ \frac{z}{\sqrt{1+z^2}} \sin \varphi \\ 1 \end{pmatrix}$$

Auswertung im Berührungspunkt

$$h_\varphi(-\pi/4, 1) = \begin{pmatrix} -\sqrt{2}(-1/\sqrt{2}) \\ \sqrt{2}(1/\sqrt{2}) \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad h_z(-\pi/4, 1) = \begin{pmatrix} 1/2 \\ -1/2 \\ 1 \end{pmatrix}$$

↪ parametrische Darstellung der Tangentialebene

$$E: \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} + s \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 1/2 \\ -1/2 \\ 1 \end{pmatrix}, \quad s, t \in \mathbb{R}$$

Vektorprodukt der aufspannenden Vektoren ↪ Normale

$$\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \times \begin{pmatrix} 1/2 \\ 1/2 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}$$

und der impliziten Darstellung

$$E: 0 = (-1, -1, 1) \begin{pmatrix} x - 1 \\ y - (-1) \\ z - 1 \end{pmatrix} \quad \text{bzw.} \quad x + y + z = -1$$

Multivariate Taylor-Approximation

Eine in einer Umgebung D eines Punktes $a = (a_1, \dots, a_m)$ $(n + 1)$ -mal stetig differenzierbare Funktion $f : D \rightarrow \mathbb{R}$ von m Veränderlichen x_k kann durch ein Taylor-Polynom vom totalen Grad $\leq n$ approximiert werden:

$$f(x) = \sum_{|\alpha| \leq n} \frac{1}{\alpha!} \partial^\alpha f(a) (x - a)^\alpha + R$$

mit $\alpha! = \alpha_1! \cdots \alpha_m!$, $\partial_\alpha = \partial_1^{\alpha_1} \cdots \partial_m^{\alpha_m}$, $y^\alpha = y_1^{\alpha_1} \cdots y_m^{\alpha_m}$ und dem Restglied

$$R = \sum_{|\alpha|=n+1} \frac{1}{\alpha!} \partial^\alpha f(u) (x - a)^\alpha, \quad u = a + \theta(x - a),$$

für ein $\theta \in [0, 1]$.

Für $x \rightarrow a$ strebt der Fehler mit der Ordnung $n + 1$ gegen Null:

$$R = O(|x - a|^{n+1}).$$

Für Funktionen von zwei oder drei Variablen werden meist anstelle von (x_1, \dots) die Bezeichnungen (x, y) bzw. (x, y, z) verwendet. Beispielsweise hat das quadratische Taylor-Polynom für eine bivariate Funktion in dieser Notation die Form

$$p(x, y) = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) + \frac{1}{2}f_{xx}(x_0, y_0)(x - x_0)^2 + f_{xy}(x_0, y_0)(x - x_0)(y - y_0) + \frac{1}{2}f_{yy}(x_0, y_0)(y - y_0)^2$$

und für eine trivariate Funktion

$$p(x, y, z) = f(x_0, y_0, z_0) + \underbrace{\dots}_{8 \text{ Terme}} + \frac{1}{2}f_{zz}(x_0, y_0, z_0)(z - z_0)^2.$$

Beweis

Zurückführung auf den univariaten Fall (o.B.d.A. $a = 0$):

setze

$$f(x_1, \dots, x_m) = f(tx_1, \dots, tx_m)|_{t=1} = g(t)|_{t=1}$$

Taylor-Entwicklung der univariaten Funktion $g(t)$ im Nullpunkt

$$g(t) = g(0) + g'(0)t + \dots + \frac{1}{n!}g^{(n)}(0)t^n + R$$

mit

$$R = \frac{1}{(n+1)!}g^{(n+1)}(\theta t)t^{n+1}$$

für ein $\theta \in [0, 1]$

Kettenregel \implies

$$g(0) = f(0, \dots, 0)$$

$$g'(0) = \sum_j \partial_j f(tx_1, \dots, tx_m) x_j \Big|_{t=0} = \sum_j (\partial_j f(0)) x_j$$

$$g''(0) = \sum_i \sum_j (\partial_i \partial_j f(0)) x_i x_j$$

\vdots

m^k Terme bei k -ter Ableitung

Zusammenfassen gleicher partieller Ableitungen \rightsquigarrow Koeffizienten der Entwicklung

z.B. für $m = 2$, $k = 5$, Zusammenfassen von

$$\partial_1 \partial_1 \partial_1 \partial_2 \partial_2, \partial_1 \partial_1 \partial_2 \partial_1 \partial_2, \partial_1 \partial_1 \partial_2 \partial_2 \partial_1, \dots$$

$$\partial^\alpha, \alpha = (3, 2) \rightsquigarrow \binom{5}{3} = 5! / (3! 2!) \text{ Terme}$$

allgemein

$$\begin{aligned} & \binom{k}{\alpha_1} \cdot \binom{k - \alpha_1}{\alpha_2} \cdots \binom{k - \alpha_1 - \dots - \alpha_{m-1}}{\alpha_m} \\ &= \frac{k!}{\alpha_1!(k - \alpha_1)!} \frac{(k - \alpha_1)!}{\alpha_2!(k - \alpha_1 - \alpha_2)!} \frac{(k - \alpha_1 - \alpha_2)!}{\alpha_3!(k - \alpha_1 - \alpha_2 - \alpha_3)!} \cdots \\ &= k! / (\alpha_1! \dots \alpha_m!) \end{aligned}$$

Terme, wenn nach der ν -ten Komponente jeweils α_ν -mal abgeleitet wird
Einsetzen der Ableitungen in die Funktion $g(t)$, Kürzen des Faktors $k!$

\rightsquigarrow Entwicklung von f

Taylor-Entwicklung einer Funktion von zwei Variablen

$$\begin{aligned} f(x, y) = & f + f_x(x - x_0) + f_y(y - y_0) \\ & + \frac{f_{xx}}{2}(x - x_0)^2 + f_{xy}(x - x_0)(y - y_0) + \frac{f_{yy}}{2}(y - y_0)^2 \\ & + \frac{f_{xxx}}{6}(x - x_0)^3 + \frac{f_{xxy}}{2}(x - x_0)^2(y - y_0) \\ & + \frac{f_{xyy}}{2}(x - x_0)(y - y_0)^2 + \frac{f_{yyy}}{6}(y - y_0)^3 + R, \end{aligned}$$

wobei f und sämtliche partielle Ableitungen im Punkt (x_0, y_0) ausgewertet werden

Konkreter Fall: Entwickeln von

$$f(x, y) = \sin(x - \omega y)$$

im Punkt $(x_0, y_0) = (0, 0)$

$\sin^{(1)} = \cos, \sin^{(2)} = -\sin, \sin^{(3)} = -\cos, \dots$ und

$$\partial_x^\alpha \partial_y^\beta f(0,0) = \sin^{(\alpha+\beta)}(0) (-\omega)^\beta$$

\rightsquigarrow Approximation

$$\sin(x - \omega y) = x - \omega y - \frac{1}{6} \underbrace{(x^3 - 3\omega x^2 y + 3\omega^2 x y^2 - \omega^3 y^3)}_{(x-\omega y)^3} + R$$

mit dem Restglied

$$R = \frac{1}{4!} (f_{xxxx} x^4 + 4f_{xxx y} x^3 y + 6f_{xx yy} x^2 y^2 + 4f_{xy yy} x y^3 + f_{yyyy} y^4)$$

und Auswertung der Ableitungen an der Stelle $(\theta x, \theta y)$ für ein $\theta \in [0, 1]$

$$f_{xxxx} = \sin(\theta x - \omega \theta y), f_{xxx y} = \sin(\theta x - \omega \theta y)(-\omega), \dots \rightsquigarrow$$

$$\begin{aligned} R &= \frac{\sin(\theta x - \omega \theta y)}{4!} (x^4 + 4x^3(-\omega)y + 6x^2\omega^2 y^2 + 4x(-\omega^3)y^3 + \omega^4 y^4) \\ &= \frac{1}{4!} \sin(\theta(x - \omega y))(x - \omega y)^4 \end{aligned}$$

Alternative Berechnung des Taylor-Polynoms durch Einsetzen von $t = x - \omega y$ in die eindimensionale Reihendarstellung der Sinusfunktion:

$$\sin(t) = t - \frac{t^3}{3!} + \dots$$

$$t = x - \omega y \quad \rightsquigarrow$$

$$f(x, y) = (x - \omega y) - \frac{1}{3!}(x - \omega y)^3 + \frac{1}{4!} \sin(\theta t)(x - \omega y)^4$$

für ein $\theta \in [0, 1]$ (univariates Restglied)

vierte Ableitung des Sinus am Entwicklungspunkt Null \rightsquigarrow kleineres Restglied

$$R = \frac{1}{5!} \cos(\theta t)(x - \omega y)^5$$

Taylor-Darstellung des Polynoms

$$f(x, y, z) = z^2 - xy$$

an der Stelle $(0, -2, 1)$

von Null verschiedene Ableitungen

$$f_x = -y, f_y = -x, f_z = 2z, f_{xy} = -1, f_{zz} = 2$$

Auswerten am Entwicklungspunkt \rightsquigarrow 1 + 5 Terme

$$\begin{aligned} z^2 - xy &= f + f_x x + f_y (y + 2) + f_z (z - 1) + f_{xy} x (y + 2) + \frac{1}{2} f_{zz} (z - 1)^2 \\ &= 1 + 2x + (-0)(y + 2) + 2(z - 1) + (-1)x(y + 2) + \frac{1}{2} 2(z - 1)^2 \end{aligned}$$

alternativ: Entwicklung durch Umformung

Substitution

$$y + 2 = \eta, \quad z - 1 = \zeta$$

\rightsquigarrow

$$\begin{aligned} f(x, y, z) &= z^2 - xy \\ &= (\zeta + 1)^2 - x(\eta - 2) = \zeta^2 + 2\zeta + 1 - x\eta + 2x \end{aligned}$$

Hesse-Matrix

Die quadratische Taylor-Approximation einer Funktion $f : \mathbb{R}^n \supseteq D \rightarrow \mathbb{R}$ in einem Punkt $a = (a_1, \dots, a_n)^t \in D$ lässt sich in der Form

$$f(x_1, \dots, x_n) = f(a) + (\text{grad } f(a))^t (x - a) + \frac{1}{2}(x - a)^t H f(a)(x - a) + O(\|x - a\|^3)$$

schreiben, wobei $\text{grad } f(a) = (\partial_1 f(a), \dots, \partial_n f(a))^t$ der Gradient im Punkt a ist und die symmetrische Hesse-Matrix

$$H f(a) = \begin{pmatrix} \partial_1 \partial_1 f(a) & \cdots & \partial_1 \partial_n f(a) \\ \vdots & & \vdots \\ \partial_n \partial_1 f(a) & \cdots & \partial_n \partial_n f(a) \end{pmatrix}$$

die zweiten Ableitungen enthält.

Bei zwei oder drei Veränderlichen werden die Variablen meist mit (x, y) bzw. (x, y, z) bezeichnet. In dieser Notation hat der quadratische Term der Taylor-Approximation für eine bivariate Funktion im Punkt (x_0, y_0) die Form

$$\frac{1}{2}(x - x_0, y - y_0) \underbrace{\begin{pmatrix} f_{xx}(x_0, y_0) & f_{xy}(x_0, y_0) \\ f_{yx}(x_0, y_0) & f_{yy}(x_0, y_0) \end{pmatrix}}_{Hf(x_0, y_0)} \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix}.$$

Beweis

Umschreiben der quadratischen Terme der Taylor-Approximation für eine bivariate Funktion ($n = 2$) \rightsquigarrow

$$\begin{aligned} f(x, y) &= f + f_x(x - x_0) + f_y(y - y_0) \\ &\quad + \frac{1}{2!} (f_{xx}(x - x_0)^2 + 2f_{xy}(x - x_0)(y - y_0) + f_{yy}(y - y_0)^2) \\ &\quad + R \\ &= f + f_x(x - x_0) + f_y(y - y_0) \\ &\quad + \frac{1}{2}(x - x_0, y - y_0) \begin{pmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{pmatrix} \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix} + R, \end{aligned}$$

wobei f und sämtliche partiellen Ableitungen im Punkt (x_0, y_0) ausgewertet werden

analoger Beweis im allgemeinen Fall ($n \geq 2$)

Beispiel

Quadratisches Taylor-Polynom der Funktion

$$f(x, y) = \ln(x + 1/y)$$

im Punkt $(0, 1)$

partielle Ableitungen

$$f_x = \frac{1}{x + 1/y}, \quad f_y = \frac{1}{x + 1/y} \left(-\frac{1}{y^2} \right) = -\frac{1}{xy^2 + y},$$
$$f_{xx} = -\frac{1}{(x + 1/y)^2}, \quad f_{xy} = \frac{1}{(xy + 1)^2}, \quad f_{yy} = \frac{2xy + 1}{(xy^2 + y)^2}$$

Gradient und Hesse-Matrix im Punkt $(0, 1)$:

$$\text{grad } f(0, 1) = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad Hf(0, 1) = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix} \Big|_{(0,1)} = \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}$$

↪ quadratische Taylor-Approximation im Punkt $(0, 1)$

$$\begin{aligned} p(x, y) &= (1, -1) \begin{pmatrix} x \\ y - 1 \end{pmatrix} \\ &\quad + \frac{1}{2}(x, y - 1) \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y - 1 \end{pmatrix} \\ &= x - y + 1 + \frac{1}{2}(-x^2 + 2x(y - 1) + (y - 1)^2) \end{aligned}$$

Fehler der Approximation für $(x, y) = (0.1, 0.9)$
exakter Wert:

$$f(0.1, 0.9) = \ln(0.1 + 1/0.9) = 0.1915 \dots$$

Näherung:

$$p(0.1, 0.9) = 0.1 - 0.9 + 1 + \frac{1}{2}(-0.1^2 + 2(0.1)(-0.1) + (-0.1)^2) = 0.19$$

Beispiel

Quadratische Taylor-Approximation von

$$f(x, y, z) = (xy)^z$$

im Punkt $(1, 1, 1)$

Regeln für die Differentiation von Potenzen,

$$\frac{d}{dt}(at)^b = ab(at)^{b-1}, \quad \frac{d}{dt}a^t = \ln a a^t$$

↪ partielle Ableitungen

$$\begin{aligned} f_x &= yz(xy)^{z-1}, & f_z &= \ln(xy)(xy)^z, \\ f_{xx} &= y^2z(z-1)(xy)^{z-2}, & f_{zz} &= (\ln(xy))^2(xy)^z, \\ f_{xy} &= z(xy)^{z-1} + xyz(z-1)(xy)^{z-2}, & f_{xz} &= y(xy)^{z-1} + yz \ln(xy)(xy)^{z-1} \end{aligned}$$

Vertauschen der Variablen ↪

$$f_y = xz(xy)^{z-1}, \quad f_{yy} = x^2z(z-1)(xy)^{z-2}, \quad f_{yz} = x(xy)^{z-1} + xz \ln(xy)(xy)^{z-1}$$

$$\text{grad } f(1, 1, 1) = (1, 1, 0)^t$$

$$Hf(1, 1, 1) = \begin{pmatrix} f_{xx} & f_{xy} & f_{xz} \\ f_{yx} & f_{yy} & f_{yz} \\ f_{zx} & f_{zy} & f_{zz} \end{pmatrix} \Big|_{(1,1,1)} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

↪ quadratische Taylor-Approximation im Punkt $(1, 1, 1)$

$$\begin{aligned} p(x, y, z) &= 1 + (1, 1, 0) \begin{pmatrix} x - 1 \\ y - 1 \\ z - 1 \end{pmatrix} \\ &\quad + \frac{1}{2} (x - 1, y - 1, z - 1) \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} x - 1 \\ y - 1 \\ z - 1 \end{pmatrix} \\ &= 1 + (x - 1) + (y - 1) \\ &\quad + (x - 1)(y - 1) + (x - 1)(z - 1) + (y - 1)(z - 1) \end{aligned}$$

Umkehrfunktion

Ist für eine stetig differenzierbare Funktion $f : \mathbb{R}^n \ni D \rightarrow \mathbb{R}^n$ die Jacobi-Matrix $f'(x_*)$ für einen Punkt x_* im Innern des Definitionsbereiches D nicht singulär, so ist f lokal invertierbar, d.h. f bildet eine Umgebung U von x_* bijektiv auf eine Umgebung V von $y_* = f(x_*)$ ab. Die Umkehrfunktion $g = f^{-1}$ ist auf V stetig differenzierbar, und es gilt

$$g'(y) = f'(x)^{-1}, \quad y = f(x) \iff x = g(y),$$

für alle $y \in V$.

Man beachte, dass im Gegensatz zu einer univariaten Funktion aus der Invertierbarkeit von $f'(x)$ für alle x aus einer zusammenhängenden Menge D nicht notwendig die Injektivität auf dem gesamten Definitionsbereich folgt; eine globale Umkehrfunktion muss nicht existieren.

Beweis

(i) Wählt man o.B.d.A. die Umgebung U von x_* als eine offene Teilmenge von D , auf der f' nicht singular ist (möglich aufgrund der Stetigkeit von $\det f'$), so genügt es, neben der lokalen Injektivität von f (Existenz von $g = f^{-1}$) die Differenzierbarkeit von g nur im Punkt y_* zu zeigen.

Begründung:

Existenz von $g'(y_*)$

\implies Stetigkeit von g im Punkt y_*

Anwendung des Resultates für beliebige Punkte $x \in U$ an Stelle von x_*

\implies Differenzierbarkeit von g an allen Punkten $y \in V = f(U)$

\implies Stetigkeit von g auf V

Stetigkeit und Invertierbarkeit von f' auf U , Stetigkeit von g auf V

\implies Stetigkeit von $g' : y \mapsto f'(g(y))^{-1}$

(ii) Invarianz des Resultates unter affinen Transformationen

↪ Betrachtung von

$$p(x) = f'(x_*)^{-1}(f(x_* + x) - y_*)$$

an Stelle von f

↪ vereinfachte Voraussetzungen:

$$p(O_n) = O_n, \text{ d.h. } x_*, y_* \rightarrow O_n = (0, \dots, 0), \quad p'(O_n) = E$$

mit $O_n = (0, \dots, 0)^t$ und E der $n \times n$ Einheitsmatrix

Stetigkeit und Differenzierbarkeit im Punkt $O_n \implies$

$$(1) \quad \|p'(x) - E\| \leq \varphi(\|x\|),$$

$$(2) \quad p(x) = x + R(x), \quad \|R(x)\| \leq \|x\|\varphi(\|x\|)$$

mit $\varphi(t) \searrow 0$ für $t \rightarrow 0$, d.h. $\varphi(t) = o(1)$

Verschiedene Funktionen φ in den Ungleichungen (1) und (2) können durch ihr Maximum, also eine gemeinsame Funktion, ersetzt werden.

(iii) Injektivität von p auf der Kugel $U : \|x\| < \delta$:

wähle δ so, dass $\varphi(\delta) < 1/2$

$x, \tilde{x} \in U, p(x) = p(\tilde{x}),$ multivariater Mittelwertsatz \implies

$$\begin{aligned} 0 &= u(\tilde{x}) - u(x) = \int_0^1 \underbrace{p'((1-t)x + t\tilde{x})}_{(E+p'(x_t)-E)} (\tilde{x} - x) dt \\ &= (\tilde{x} - x) + \int_0^1 (p'(x_t) - E)(\tilde{x} - x) dt \end{aligned}$$

und

$$\|\tilde{x} - x\| \leq \|p'(x_t) - E\| \|\tilde{x} - x\| \stackrel{(1)}{=} \varphi(\|x_t\|) \|\tilde{x} - x\|$$

$$x_t \in U \implies \varphi(\|x_t\|) < \varphi(\delta) < 1/2 \implies \tilde{x} = x$$

(iv) $V = p(U)$ enthält die Kugel $B : \|y\| < \delta/4$, ist also eine Umgebung von O_n :

zeige dazu mit Hilfe des Banachschen Fixpunktsatzes, dass die Abbildung

$$S : x \mapsto x - p(x) + y$$

einen eindeutigen Fixpunkt in der Kugel $\tilde{B} : \|x\| \leq \delta/2$ besitzt, d.h. $\exists x$ mit $y = p(x)$

verifiziere die notwendigen Voraussetzungen:

- $S(\tilde{B}) \subseteq \tilde{B}$:

$$\begin{aligned} \|S(x)\| &= \|x - \underbrace{(x + R(x))}_{p(x)} + y\| \stackrel{(2)}{\leq} \|x\| \varphi(\|x\|) + \|y\| \\ &\leq (\delta/2) \underbrace{\varphi(\delta/2)}_{\leq \varphi(\delta) < 1/2} + \delta/4 = \delta/2 \end{aligned}$$

- $\|S(\tilde{x}) - S(x)\| \leq c \|\tilde{x} - x\|$ mit $c < 1$:

Abschätzung der Kontraktionskonstante durch $c = \max_{x \in \tilde{B}} \|S'\|$

$$S'(x) = E - p'(x) \quad \stackrel{(1)}{\implies} \quad c = 1/2$$

(v) $q'(O_n) = E, q = p^{-1}$:

zu zeigen:

$$q(y) = \underbrace{q(O_n)}_{O_n} + y + \tilde{R}(y)$$

mit $\|\tilde{R}(y)\| = o(\|y\|)$ für $\|y\| \rightarrow 0$

$y = p(x) \rightsquigarrow$

$$\tilde{R}(y) = q(y) - y = x - p(x) = -R(x) \stackrel{(2)}{=} o(\|x\|)$$

\rightsquigarrow noch $\|x\|$ durch $\|y\|$ abzuschätzen:

$$\|y\| = \|p(x)\| = \|x + R(x)\| \stackrel{(2)}{\geq} \|x\| - \|x\|\varphi(\|x\|) \geq (1 - 1/2)\|x\|$$

für $y \in B, x \in \tilde{B}$

Beispiel

Lokale Invertierbarkeit der Funktion

$$f : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} xy \\ x/y \end{pmatrix}$$

im Punkt $(x_*, y_*) = (2, 1)$

Jacobi-Matrix

$$f'(x, y) = \begin{pmatrix} y & x \\ 1/y & -x/y^2 \end{pmatrix}$$

Einsetzen der konkreten Koordinaten

$$J = f'(2, 1) = \begin{pmatrix} 1 & 2 \\ 1 & -2 \end{pmatrix}$$

$\det J = -4 \neq 0 \implies \exists$ Umkehrfunktion g in Umgebung von

$$\begin{pmatrix} u_* \\ v_* \end{pmatrix} = f(x_*, y_*) = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$$

Jacobi-Matrix im Punkt (u_*, v_*)

$$g'(2, 2) = J^{-1} = \frac{1}{4} \begin{pmatrix} 2 & 2 \\ 1 & -1 \end{pmatrix}$$

explizite Bestimmung von g durch Auflösen der Ausdrücke für u und v nach x und y :

$$\begin{pmatrix} x \\ y \end{pmatrix} = g(u, v) = \begin{pmatrix} \sqrt{uv} \\ \sqrt{u/v} \end{pmatrix}$$

\rightsquigarrow Überprüfung der Formel für $g'(2, 2)$ auf direktem Weg

$$g'(u, v) = \frac{1}{2} \begin{pmatrix} \sqrt{v/u} & \sqrt{u/v} \\ 1/\sqrt{uv} & -\sqrt{u/v^3} \end{pmatrix}$$

Übereinstimmung mit J^{-1} für $(u_*, v_*) = (2, 2)$

Beispiel

Invertierbarkeit der komplexen Exponentialfunktion

$$z = x + iy \mapsto e^z = u + iv$$

Formel von Euler-Moivre

$$e^{it} = \cos t + i \sin t$$

↪ reelle Darstellung

$$f(x, y) = \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} e^x \cos y \\ e^x \sin y \end{pmatrix}$$

mit der Jacobi-Matrix

$$f'(x, y) = e^x \begin{pmatrix} \cos y & -\sin y \\ \sin y & \cos y \end{pmatrix}$$

$$\det f'(x, y) = (e^x \cos x)^2 + (e^x \sin x)^2 = e^{2x} > 0$$

\implies lokale Existenz einer Umkehrfunktion g (ein Zweig des komplexen Logarithmus) mit der Ableitung

$$g'(u, v) = (f'(x, y))^{-1} = e^{-x} \begin{pmatrix} \cos y & \sin y \\ -\sin y & \cos y \end{pmatrix}$$

keine globale Umkehrfunktion wegen

$$f(x, y + 2\pi k) = f(x, y), \quad k \in \mathbb{Z}$$

Periodizität \implies unendlich viele Urbilder für jeden Punkt $(u, v) \neq (0, 0)$

Beispiel

Paradox bei der Berechnung partieller Ableitungen von Polarkoordinaten

$$x = r \cos \varphi, \quad y = r \sin \varphi, \quad r = \sqrt{x^2 + y^2}$$

partielle Ableitung des Radius' nach der x-Koordinate

$$\frac{\partial r}{\partial x} = \frac{\partial}{\partial x} \sqrt{x^2 + y^2} = \frac{x}{\sqrt{x^2 + y^2}} = \frac{x}{r} = \cos \varphi$$

Paradox:

$$r = x / \cos(\varphi) \quad \stackrel{?}{\implies} \quad \frac{\partial r}{\partial x} = \frac{1}{\cos(\varphi)}$$

Grund: Bei den Berechnungen der partiellen Ableitungen werden nicht die gleichen Variablen konstant gehalten. In der ersten Gleichung ist y konstant und bei der zweiten φ .

Die skalare Ableitungsregel $dy/dx = (dx/dy)^{-1}$ ist im allgemeinen falsch für partielle Ableitungen:

$$x_u \neq (u_x)^{-1}$$

korrekte Berechnung für eine Bijektion $(x, y) \leftrightarrow (u, v)$ mit Hilfe der Jacobi-Matrix:

$$\frac{\partial(u, v)}{\partial(x, y)} = \begin{pmatrix} u_x & u_y \\ v_x & v_y \end{pmatrix}, \quad \frac{\partial(x, y)}{\partial(u, v)} = \begin{pmatrix} x_u & x_v \\ y_u & y_v \end{pmatrix} = \begin{pmatrix} u_x & u_y \\ v_x & v_y \end{pmatrix}^{-1}$$

im betrachteten Beispiel

$$J = \frac{\partial(x, y)}{\partial(r, \varphi)} = \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix}$$

Inverse \rightsquigarrow partielle Ableitungen nach x und y :

$$J^{-1} = \begin{pmatrix} \cos \varphi & \sin \varphi \\ -\frac{1}{r} \sin \varphi & \frac{1}{r} \cos \varphi \end{pmatrix} = \frac{\partial(r, \varphi)}{\partial(x, y)} = \begin{pmatrix} r_x & r_y \\ \varphi_x & \varphi_y \end{pmatrix}$$

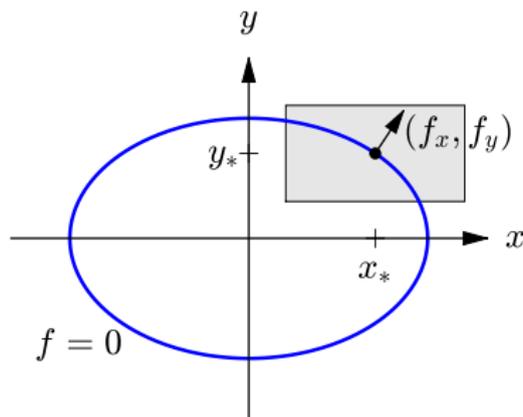
insbesondere: $r_x = \cos \varphi$

Implizite Funktionen

Für eine stetig differenzierbare bivariate Funktion f ist die Gleichung $f(x, y) = 0$ in der Umgebung einer Lösung (x_*, y_*) nach y auflösbar,

$$f(x, y) = 0 \iff y = g(x), \quad x \approx x_*,$$

falls $f_y(x_*, y_*) \neq 0$. Diese hinreichende Bedingung bedeutet, dass die Tangente an die durch die Gleichung definierte Kurve in der xy -Ebene im Punkt (x_*, y_*) nicht parallel zur y -Achse ist.



Die implizit definierte Funktion g lässt sich im allgemeinen nicht explizit angeben. Jedoch kann die Ableitung durch Differenzieren der Gleichung $f(x, g(x)) = 0$ bestimmt werden:

$$g'(x) = -f_y(x, g(x))^{-1} f_x(x, g(x)).$$

Die Berechnung höherer Ableitungen ist ebenfalls auf diese Weise möglich. Allgemeiner ist eine hinreichende Bedingung für die Auflösbarkeit einer Gleichung $f(z_1, \dots, z_{n+1}) = 0$ nach z_k in einer Umgebung einer Lösung z_* , dass $\partial_k f(z_*) \neq 0$, d.h. die k -te Komponente der Normale der durch die Gleichung definierten Fläche in \mathbb{R}^{n+1} muss im Punkt z_* ungleich null sein. Ein analoges Kriterium charakterisiert die lokale Auflösbarkeit eines Gleichungssystems

$$f_k(x_1, \dots, x_m, y_1, \dots, y_n) = 0, \quad k = 1, \dots, n,$$

mit stetig differenzierbaren $m + n$ -variaten Funktionen f_k . Ist

$$\det f_y(x_*, y_*) \neq 0, \quad f_y = \begin{pmatrix} \frac{\partial f_1}{\partial y_1} & \cdots & \frac{\partial f_1}{\partial y_n} \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial y_1} & \cdots & \frac{\partial f_n}{\partial y_n} \end{pmatrix},$$

so definiert das Gleichungssystem implizit eine Funktion

$$g : x \mapsto y = g(x), \quad x \approx x_*,$$

die die Lösungsmenge des Gleichungssystems in der Umgebung von (x_*, y_*) eindeutig parametrisiert: $f(x, y) = 0 \iff y = g(x_1, \dots, x_m)$, $x \approx x_*$. Die Funktion $g = (g_1, \dots, g_n)^t$ ist stetig differenzierbar und besitzt die $n \times m$ -Jacobi-Matrix

$$g' = \frac{\partial(g_1, \dots, g_n)}{\partial(x_1, \dots, x_m)} = -(f_y)^{-1} f_x, \quad f_x = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_m} \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_m} \end{pmatrix}.$$

Ein entsprechendes Resultat gilt nach Permutation der Variablen, d.h. man kann Gleichungen

$$f_k(z_1, \dots, z_{m+n}) = 0, \quad k = 1, \dots, n,$$

lokal nach z_{k_1}, \dots, z_{k_n} auflösen, wenn die Spalten k_1, \dots, k_n der Jacobi-Matrix $f'(z_*)$ linear unabhängig sind.

Das Kriterium für die Auflösbarkeit nichtlinearer Gleichungssysteme wird plausibel, wenn man linearisiert, d.h. die Funktion $f = (f_1, \dots, f_n)^t$ durch ihre lineare Taylor-Approximation ersetzt:

$$f(x, y) \approx \underbrace{f(x_*, y_*)}_{(0, \dots, 0)^t} + \underbrace{f_x(x_*, y_*)}_{A}(x - x_*) + \underbrace{f_y(x_*, y_*)}_{B}(y - y_*).$$

Offensichtlich ist das lineare Gleichungssystem

$$(0, \dots, 0)^t = A(x - x_*) + B(y - y_*)$$

nach y auflösbar, wenn die Matrix B invertierbar ist.

Beweis

betrachte die Abbildung

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto u(x, y) = \begin{pmatrix} x \\ f(x, y) \end{pmatrix}$$

von \mathbb{R}^{m+n} nach \mathbb{R}^{m+n} in einer Umgebung von (x_*, y_*)

invertierbare Jacobi Matrix

$$u'(x_*, y_*) = \left(\begin{array}{c|c} E & 0_{m \times n} \\ \hline f_x & f_y \end{array} \right) \Big|_{(x_*, y_*)}, \quad E : m \times m \text{ Einheitsmatrix}$$

\implies lokale Existenz einer Umkehrfunktion $u^{-1} = (v_1, \dots, v_{m+n})^t$

$$u(x, y) = \begin{pmatrix} x \\ f(x, y) \end{pmatrix} = \begin{pmatrix} x \\ 0_{n \times 1} \end{pmatrix} \iff \begin{pmatrix} x \\ y \end{pmatrix} = v(x, 0_{n \times 1}),$$

d.h. $y = g(x) = (v_{m+1}(x, 0, \dots, 0), \dots, v_{m+n}(x, 0, \dots, 0))^t$

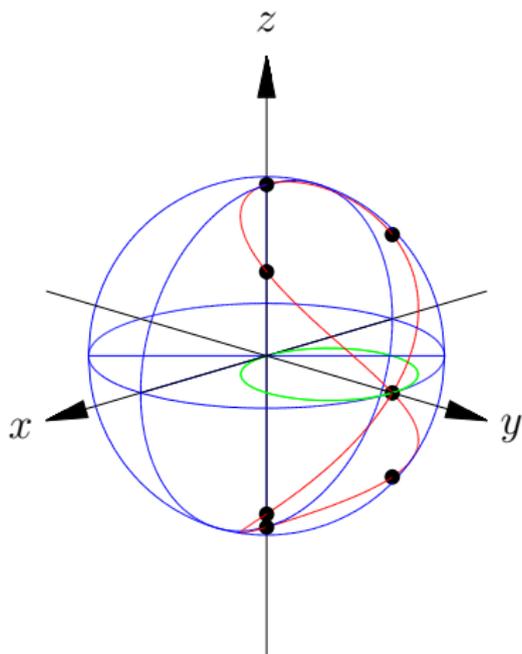
Differenzieren von $0_{n \times 1} = f(x, g(x))$ nach $x \implies$

$$0_{n \times m} = f_x + f_y g',$$

d.h. die explizite Formel für die Jacobi-Matrix g'

Vivianische Kurve (Schnitt der Einheitssphäre mit einem Zylinder)

$$C : t \mapsto (\sin t \cos t, \sin^2 t, \cos t)^t, \quad t \in [0, 2\pi]$$



implizite Form

$$f(x, y, z) = x^2 + y^2 + z^2 - 1 = 0$$

$$g(x, y, z) = x^2 + \left(y - \frac{1}{2}\right)^2 - \frac{1}{4} = 0$$

Jacobi-Matrix

$$\begin{pmatrix} f_x & f_y & f_z \\ g_x & g_y & g_z \end{pmatrix} = \begin{pmatrix} 2x & 2y & 2z \\ 2x & 2y - 1 & 0 \end{pmatrix}$$

Satz über implizite Funktion \implies

Bedingung für eine Parametrisierung durch eine der Koordinaten x, y, z

(i) Parametrisierung bezüglich x , d.h. auflösen nach y und z :
hinreichende Bedingung: Invertierbarkeit von

$$\frac{\partial(f, g)}{\partial(y, z)} = \begin{pmatrix} 2y & 2z \\ 2y - 1 & 0 \end{pmatrix}$$

erfüllt für $z \neq 0$ und $y \neq 1/2$

$$g = x^2 + (y - 1/2)^2 - 1/4 = 0, \quad f - g = z^2 - 1 + y = 0 \quad \rightsquigarrow$$

$$y(x) = \frac{1}{2} + \sigma \sqrt{\frac{1}{4} - x^2}, \quad z(x) = \sigma' \sqrt{\frac{1}{2} - \sigma \sqrt{\frac{1}{4} - x^2}}$$

wobei die Vorzeichen $\sigma, \sigma' \in \{-1, 1\}$ entsprechend den einzelnen Zweigen gewählt werden müssen

Die Parametrisierungen sind singulär in den Punkten mit $z = 0$ oder $y = 1/2$:

$$(0, 1, 0)^t, \quad (1/2, 1/2, \pm\sqrt{2}/2)^t, \quad (-1/2, 1/2, \pm\sqrt{2}/2)^t$$

geometrisch hinreichend für eine Parametrisierung bzgl. x :

Tangentenrichtung nicht orthogonal zu $(1, 0, 0)^t$, d.h. mit nichtrivialer Komponente in x -Richtung

(ii) Parametrisierung bezüglich y und z :
hinreichende Bedingungen

$$\det \begin{pmatrix} 2x & 2z \\ 2x & 0 \end{pmatrix} = -4xz \neq 0, \quad \det \begin{pmatrix} 2x & 2y \\ 2x & 2y - 1 \end{pmatrix} = -2x \neq 0$$

keine lokale Parametrisierung (weder nach x , y oder z) im Punkt $(0, 1, 0)^t$
Jacobi-Matrix

$$\left. \frac{\partial(f, g)}{\partial(x, y, z)} \right|_{(0,1,0)} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & 1 \end{pmatrix}$$

Rang der Jacobi-Matrix gleich 1 \rightsquigarrow Doppelpunkt der Kurve

Beispiel

Lokale Auflösbarkeit der Gleichung

$$f(x, y, z) = xe^y - yz = 0$$

nach einer Variablen auflösbar, falls die entsprechende partielle Ableitung nicht verschwindet

Gradient

$$(f_x, f_y, f_z)^t = (e^y, xe^y - z, -y)^t$$

(i) $f_x > 0 \implies f = 0$ für alle (x, y, z) nach x auflösbar:

$$x = yze^{-y}$$

(ii) Auflösen nach z für $y \neq 0$ möglich ($f_z \neq 0$):

$$z = xe^y / y$$

(iii) Keine elementare Auflösbarkeit nach y :

Satz über implizite Funktionen \implies Auflösbarkeit in einer Umgebung einer Lösung (x_*, y_*, z_*) , falls

$$f_y(x_*, y_*, z_*) = x_* e^{y_*} - z_* \neq 0$$

z.B. erfüllt für die Lösung $(0, 0, 1)$ der Gleichung: $f_y(0, 0, 1) = -1 \implies$

$$\exists g : f(x, y, z) = 0 \iff y = g(x, z), \quad (x, y, z) \approx (0, 0, 1)$$

Die Funktion g ist nicht explizit angebar; aber der Gradient g' kann bestimmt werden:

$$0 = f(x, g(x, z), z) \implies 0 = f_x + f_y g_x, \quad 0 = f_z + f_y g_z$$

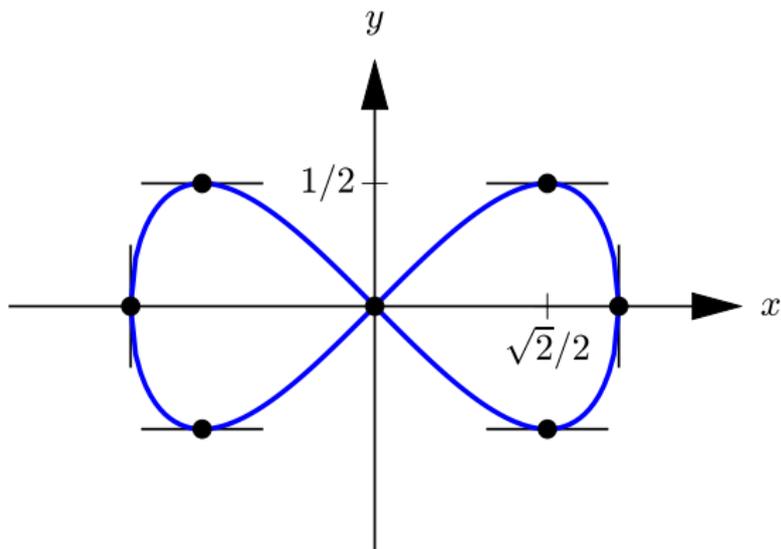
Auflösen nach den partiellen Ableitungen von g im Punkt (x_*, y_*, z_*)

$$g_x(0, 1) = \frac{-f_x(0, 0, 1)}{f_y(0, 0, 1)} = \frac{1}{-1} = -1$$

$$g_z(0, 1) = \frac{-f_z(0, 0, 1)}{f_y(0, 0, 1)} = \frac{0}{-1} = 0$$

Beispiel

Parametrisierung der Lemniskate $C : p(x, y) = x^4 - x^2 + y^2 = 0$



$$\text{grad } p = (-2x + 4x^3, 2y)^t$$

Doppelpunkt bei $(0, 0)$ ($\text{grad } p(0, 0) = (0, 0)^t$), keine eindeutige Auflösbarkeit nach x oder y

(i) Vertikale Tangente bei $(\pm 1, 0)$, $\text{grad } p = (\pm 2, 0)^t$:

lokale Auflösung von $p(x, y) = 0$ nach x (Parametrisierung bzgl. y) möglich:

$$x = \sigma \sqrt{\frac{1}{2} + \sqrt{\frac{1}{4} - y^2}}$$

mit $\sigma = 1$ für $x \approx 1$ und $\sigma = -1$ für $x \approx -1$

(ii) Waagrechte Tangente bei $(\sigma_1 \sqrt{2}/2, \sigma_2/2)$ mit $\sigma_k \in \{-1, 1\}$,
 $\text{grad } p \parallel (0, 1)^t$:

Auflösen von $p(x, y) = 0$ nach $y \rightsquigarrow$

$$y = \sigma_2 \sqrt{x^2 - x^4}$$

(iii) Punkte (x_0, y_0) mit weder vertikaler noch horizontaler Tangente:

Auflösbarkeit sowohl nach x oder y

Bei Auflösung nach y folgt aus

$$0 = \frac{d}{dx} p(x, y(x)) = p_x + p_y \frac{dy}{dx}$$

dass

$$\frac{dy}{dx} = -p_y^{-1} p_x = \frac{4x^3 - 2x}{2y}$$

↪ Gleichung der Tangente im Punkt (x_0, y_0)

$$y - y_0 = \frac{4x_0^3 - 2x_0}{2y_0} (x - x_0)$$

Fehlerfortpflanzung bei multivariaten Funktionen

Für eine stetig differenzierbare Funktion $(x_1, \dots, x_n)^t \mapsto y = f(x)$ lassen sich die Auswirkungen von inakuraten Argumenten $x + \Delta x \approx x$ mit Hilfe der partiellen Ableitungen von f beschreiben. Für den absoluten Fehler gilt bei Vernachlässigung von Termen der Ordnung $o(|\Delta x|)$

$$\Delta y = f(x + \Delta x) - f(x) \approx f_{x_1}(x)\Delta x_1 + \dots + f_{x_n}(x)\Delta x_n.$$

Sind $|x_k|, |y| \neq 0$, so folgt für den relativen Fehler

$$\frac{\Delta y}{|y|} \approx c_1 \frac{\Delta x_1}{|x_1|} + \dots + c_n \frac{\Delta x_n}{|x_n|}$$

mit den Konditionszahlen

$$c_k = \frac{\partial y}{\partial x_k} \frac{|x_k|}{|y|}.$$

Beispiel

Berechnung des Winkels φ von Polarkoordinaten aus kartesischen Koordinaten:

$$\varphi = f(x, y) = \arctan(y/x)$$

Bilden der partiellen Ableitungen: $\frac{d}{dt} \arctan t = \frac{1}{1+t^2}$, Kettenregel \implies

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

$$f_y(x, y) = \frac{1}{1 + (y/x)^2} \frac{1}{x} = \frac{x}{x^2 + y^2}$$

(i) Absoluter Fehler:

$$\Delta\varphi \approx f_x(x, y)\Delta x + f_y(x, y)\Delta y = -\frac{y\Delta x}{x^2 + y^2} + \frac{x\Delta y}{x^2 + y^2}$$

$|x|, |y| \leq r = \sqrt{x^2 + y^2} \implies$ mögliche Vergrößerung des absoluten Fehlers $\max(|\Delta x|, |\Delta y|)$ maximal um einen Faktor $2/r$

(ii) Relativer Fehler:

$$\frac{\Delta\varphi}{|\varphi|} \approx -\frac{y|x|}{(x^2+y^2)|\varphi|} \frac{\Delta x}{|x|} + \frac{x|y|}{(x^2+y^2)|\varphi|} \frac{\Delta y}{|y|}$$

$x = r \cos \varphi, y = r \sin \varphi$ und $|\sin \varphi / \varphi| \leq 1 \implies$ Betrag der rechten Seite

$$\leq \left| \frac{\cos \varphi \sin \varphi}{\varphi} \right| \left(\frac{|\Delta x|}{|x|} + \frac{|\Delta y|}{|y|} \right) \leq 2 \max \left(\frac{|\Delta x|}{|x|}, \frac{|\Delta y|}{|y|} \right)$$

Verstärkung des relativen Fehlers höchstens um einen Faktor 2

Steilster Abstieg

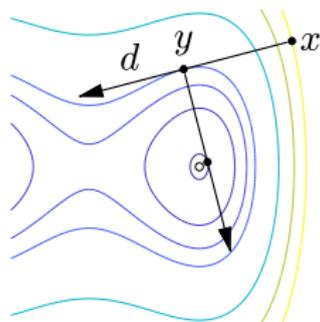
Die Methode des steilsten Abstiegs dient zur Minimierung multivariater Funktionen $f(x_1, \dots, x_n)$. Zur Durchführung eines Iterationsschritts $x \rightarrow y$ wird zunächst der negative Gradient

$$d = -\text{grad } f(x)$$

als lokal beste Abstiegsrichtung berechnet. Dann bestimmt man y als eine Minimalstelle von f in Richtung von d :

$$f(y) = \min_{t \geq 0} f(x + td).$$

Wie in der Abbildung illustriert, ist die Suchrichtung orthogonal zu der Niveaumenge durch x und berührt eine Niveaumenge zu einem kleineren Funktionswert in y .



Die Konvergenz der durch die Methode des steilsten Abstiegs erzeugten Folge x_0, x_1, \dots kann unter sehr allgemeinen Voraussetzungen gezeigt werden. Hinreichend ist, dass f nach unten beschränkt ist und $\text{grad } f$ in einer Umgebung U der Menge $\{x : f(x) \leq f(x_0)\}$ Lipschitz-stetig ist, d.h.

$$\|\text{grad } f(x) - \text{grad } f(\tilde{x})\| \leq L\|x - \tilde{x}\|, \quad x, \tilde{x} \in U.$$

Dann gilt

$$\sum_{\ell=0}^{\infty} \|\text{grad } f(x_\ell)\|^2 < \infty;$$

insbesondere ist $\|\text{grad } f(x_\ell)\|$ eine Nullfolge. Dies impliziert, dass jeder Häufungspunkt x_* der Folge x_0, x_1, \dots ein kritischer Punkt von f ist, d.h. $\text{grad } f(x_*) = (0, \dots, 0)^t$. Dass es sich um ein lokales Minimum handelt ist statistisch gesehen fast sicher, kann jedoch nicht zwingend gefolgert werden.

In dem Algorithmus braucht die eindimensionale Minimierung nur näherungsweise durchgeführt werden. Die Suchrichtung d muss nicht als der negative Gradient gewählt, und eine globale Minimalstelle y nicht

bestimmt werden. Entscheidend für die Konvergenz ist lediglich, dass in jedem Iterationsschritt eine Reduktion des Funktionswertes proportional zu $\|\text{grad } f(x)\|^2$ erreicht wird.

Beispiel

Steilster Abstieg für eine quadratische Funktion

$$f(x) = \frac{1}{2}x^tAx - b^tx$$

mit einer symmetrischen positiv definiten Matrix A

Iterationsschritt $x \rightarrow y = x + td$ mit

$$d = -\text{grad } f(x) = b - Ax$$

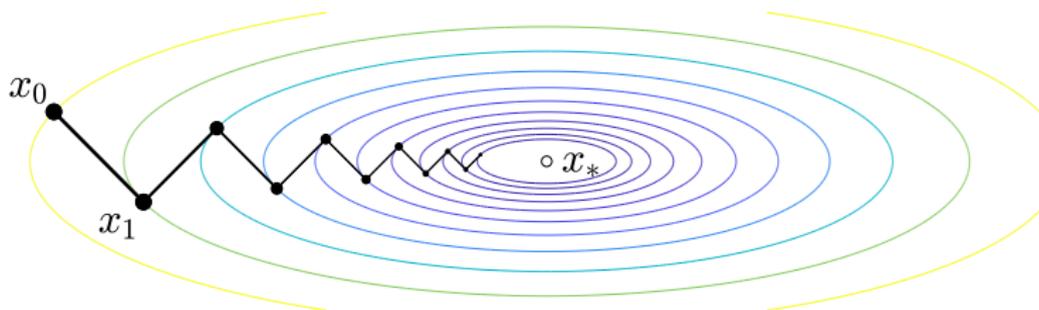
und t der Minimalstelle der univariaten Funktion

$$\begin{aligned} f(x + td) &= \frac{1}{2}(x + td)^tA(x + td) - b^t(x + td) \\ &= \frac{1}{2}d^tAd t^2 + (x^tAd - b^td) t + \frac{1}{2}(x^tAx - 2b^tx) \end{aligned}$$

Nullsetzen der Ableitung nach $t \rightsquigarrow$ Formel für den
Halbgeradenparameter t

$$0 = d^t A d t - (b - Ax)^t d = d^t A d t - d^t d \quad \Longrightarrow \quad t = \frac{d^t d}{d^t A d},$$

d.h. man erhält einen expliziten Ausdruck für y
unerwünschte Oszillationen bei Eigenwerten stark unterschiedlicher
Größenordnung von A



konkretes Beispiel

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 100 \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$x = (c, 1)^t \rightsquigarrow$$

$$d = -(Ax - b) = - \begin{pmatrix} c \\ 100 \end{pmatrix}, \quad Ad = - \begin{pmatrix} c \\ 10^4 \end{pmatrix}$$

und

$$d^t d = c^2 + 10^4, \quad d^t Ad = c^2 + 10^6, \quad t = \frac{d^t d}{d^t Ad} = \frac{c^2 + 10^4}{c^2 + 10^6}$$

sowie

$$y = x + td = \begin{pmatrix} c \\ 1 \end{pmatrix} - \frac{c^2 + 10^4}{c^2 + 10^6} \begin{pmatrix} c \\ 100 \end{pmatrix} = \frac{99c^2}{c^2 + 10^6} \begin{pmatrix} 10^4/c \\ -1 \end{pmatrix}$$

$c = 100 \rightsquigarrow$ Verbesserung um weniger als 1%:

$$y = \frac{99}{101} \begin{pmatrix} x_1 \\ -x_2 \end{pmatrix}$$

Newton-Verfahren

Eine Lösung $x_* \in \mathbb{R}^n$ eines nichtlinearen Gleichungssystems

$$f_1(x) = \cdots = f_n(x) = 0$$

kann mit der Newton-Iteration bestimmt werden, die auf einer linearen Approximation einer glatten Funktion f basiert:

$$(0, \dots, 0)^t \stackrel{!}{=} f(x + \Delta x) \approx f(x) + f'(x)\Delta x.$$

Das Inkrement Δx für einen Iterationsschritt $x \rightarrow y = x + \Delta x$ berechnet sich somit als Lösung des linearen Gleichungssystems

$$f'(x)\Delta x = -f(x).$$

Ist f zweimal stetig differenzierbar und die Jacobi-Matrix $f'(x_*)$ invertierbar, so konvergiert die durch das Verfahren erzeugte Folge x_0, x_1, \dots lokal quadratisch:

$$\|x_{k+1} - x_*\| \leq c \|x_k - x_*\|^2$$

für Startwerte x_0 in einer Umgebung U von x_* . Insbesondere ist $\det f'(x) \neq 0$ für $x \in U$, so dass das lineare Gleichungssystem für Δx eindeutig lösbar ist.

Um die Iteration robuster zu gestalten und sicherzustellen, dass auch in größerer Entfernung von einer Lösung in jedem Schritt eine Verbesserung erzielt wird, kann man eine Dämpfungsstrategie verwenden:

$$x \rightarrow y = x + \lambda \Delta x$$

mit einem Dämpfungsparameter $\lambda \in (0, 1]$. Es liegt nahe, λ so zu wählen, dass $\|f(y)\| < \|f(x)\|$. In dieser Form ist die Iteration jedoch nicht affin invariant. Bereits eine unterschiedliche Skalierung der Funktionskomponenten f_k kann zu verschiedenen Ergebnissen führen.

Dieser Nachteil lässt sich beheben, indem man die Funktionswerte mit der Inversen der Jacobi-Matrix multipliziert. Fordert man zusätzlich, dass die Reduktion des Fehlers mindestens proportional zu $\lambda/2$ ist, so führt dies auf den Test

$$\|f'(x)^{-1}f(y)\| \stackrel{!}{\leq} (1 - \lambda/2) \underbrace{\|f'(x)^{-1}f(x)\|}_{\Delta}.$$

Beginnend mit $\lambda = 1$ (Normalfall) wird der Dämpfungsparameter sukzessive halbiert, bis diese Ungleichung erfüllt ist. Im allgemeinen wird durch diese Strategie das Konvergenzgebiet wesentlich vergrößert.

Beweis

(i) Beschreibung der Iteration:

lineare Approximation von $f \implies$

$$(0, \dots, 0)^t = f(x_*) = f(x) + f'(x)(x_* - x) + R,$$
$$R_k = \frac{1}{2} (x_* - x)^t H f_k(\tilde{x}_k)(x_* - x), \quad k = 1, \dots, n,$$

mit $H f_k$ der Hesse-Matrix der k -ten Komponente von f und \tilde{x}_k einem Punkt auf dem Segment zwischen x_* und x

Auflösen nach x_* und Vernachlässigen des Restglieds $R \rightsquigarrow$ verbesserte Näherung $y \approx x_*$,

$$y = x - f'(x)^{-1} f(x)$$

Auflösen nach $f(x) \implies$

$$f(x) = -f'(x)(x_* - x) - R$$

und nach Einsetzen in die Iterationsvorschrift

$$y = x_* + f'(x)^{-1} R$$

(ii) Abschätzung des Fehlers:

o.B.d.A. Verwendung der Maximumsnorm für Vektoren und der zugeordneten Zeilensummennorm für Matrizen

setze $C = \max(\|f'(x_*)^{-1}\|, \|H f_1(x_*)\|, \dots, \|H f_n(x_*)\|)$ und wähle ein positives $\delta < 1/(4C^2)$, so dass

$$\|x - x_*\| < \delta \quad \implies \quad \|f'(x)^{-1}\|, \|H f_k(x)\| \leq 2C$$

(möglich aufgrund der Stetigkeit der beteiligten Funktionen)

$$\|x - x_*\| < \delta \quad \implies$$

$$\|y - x_*\| \leq \|f'(x)^{-1}\| \|R\| \leq \underbrace{(2C) \frac{1}{2} (2C)}_c \|x - x_*\|^2$$

und nach Wahl von $\delta < 1/(4C^2)$ ebenfalls

$$\|y - x_*\| < \frac{1}{2} \|x - x_*\|$$

\implies δ -Umgebung von x_* enthält auch y

\implies Gültigkeit der Abschätzungen für alle generierten Approximationen

(iii) Dämpfung:

zu zeigen: $\|f'(x)^{-1}f(y)\| \leq (1 - \lambda/2)\|\Delta\|$ für hinreichend kleines λ

$$f(y) = f(x - \lambda f'(x)^{-1}f(x)) = f(x) - \lambda f(x) + R$$

mit $\|R\| = O((\lambda\|f(x)\|)^2) \implies$

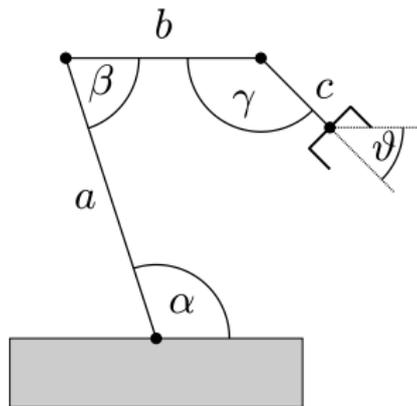
$$\|f(x)^{-1}f(y)\| \leq (1 - \lambda)\underbrace{\|f'(x)^{-1}f(x)\|}_{\Delta} + C\lambda^2\|f(x)\|^2$$

\implies behauptete Ungleichung, da $C\lambda^2\|f(x)\|^2 \leq (\lambda/2)\|\Delta\|$ für $\lambda \rightarrow 0$

Beispiel

Vereinfachte Robotersteuerung

Roboterarm mit drei Drehgelenken:
Position P und Orientierung ϑ des
Endeffektors als nichtlineare Funktion
 $q = (p_1, p_2, \vartheta)^t$ der Gelenkwinkel α ,
 β , γ



Winkel der Robotersegmente mit der x -Achse: α , $\alpha - (\pi - \beta)$,

$\alpha - (\pi - \beta) - (\pi - \gamma)$

\implies

$$q_1(\alpha, \beta, \gamma) = a \cos(\alpha) + b \cos(\alpha + \beta - \pi) + c \cos(\alpha + \beta + \gamma - 2\pi)$$

$$q_2(\alpha, \beta, \gamma) = a \sin(\alpha) + b \sin(\alpha + \beta - \pi) + c \sin(\alpha + \beta + \gamma - 2\pi)$$

$$q_3(\alpha, \beta, \gamma) = \alpha + \beta + \gamma - 2\pi$$

$a = 3$, $b = 2$, $c = 1$ und Substitution von $\beta' = \alpha + \beta$,

$\vartheta = q_3 = \alpha + \beta + \gamma - 2\pi \rightsquigarrow$ nichtlineares System

$$0 = f_1(\alpha, \beta') = p_1 - 3 \cos \alpha + 2 \cos \beta' - \cos \vartheta$$

$$0 = f_2(\alpha, \beta') = p_2 - 3 \sin \alpha + 2 \sin \beta' - \sin \vartheta$$

Ruhelage des Roboterarms: $\tilde{p} = (2, 2)^t$, $\tilde{\vartheta} = -\pi/2$ ($\alpha = \beta = \gamma = \pi/2$)

\rightsquigarrow Lösung $\alpha_0 = \pi/2$, $\beta'_0 = \pi$ mit der Jacobi-Matrix

$$f'(\alpha_0, \beta'_0) = \begin{bmatrix} 3 \sin \alpha_0 & -2 \sin \beta'_0 \\ -3 \cos \alpha_0 & 2 \cos \beta'_0 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & -2 \end{bmatrix}$$

erster Schritt des Newton-Verfahrens, $(\alpha_0, \beta'_0) \rightarrow (\alpha_1, \beta'_1)$, zur Erreichung der benachbarten Position $p = (2, 3)^t$, $\vartheta = -\pi/4$:

$$\begin{bmatrix} 3 & 0 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} \Delta\alpha \\ \Delta\beta' \end{bmatrix} = \begin{bmatrix} -\sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix}$$

$\implies \Delta\alpha = -\sqrt{2}/6$, $\Delta\beta' = -\sqrt{2}/4$ und

$$\begin{pmatrix} \alpha_1 \\ \beta'_1 \end{pmatrix} = \begin{pmatrix} \alpha_0 \\ \beta'_0 \end{pmatrix} - \begin{pmatrix} \Delta\alpha \\ \Delta\beta' \end{pmatrix} = \begin{pmatrix} \pi/2 + \sqrt{2}/6 \\ \pi + \sqrt{2}/4 \end{pmatrix}$$

Fehler zur Position p : $f(\alpha_1, \beta'_1) \approx (0.1172, 0.0976)$

Kritischer Punkt

Für eine Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ bezeichnet man $x = (x_1, \dots, x_n)^t$ als kritischen Punkt, wenn $\text{grad } f(x) = (0, \dots, 0)^t$. Ist f zweimal stetig differenzierbar, so kann f in einer Umgebung von x durch eine quadratische Taylor-Approximation angenähert werden:

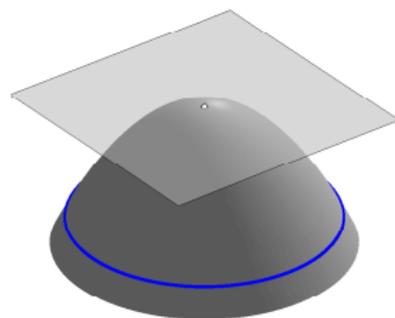
$$f(x + \Delta x) \approx f(x) + \frac{1}{2} \Delta x^t H f(x) \Delta x.$$

Der Typ des kritischen Punktes, d.h. die Form des Funktionsgraphen in einer Umgebung von x , wird somit durch die Eigenwerte λ_k der Hesse-Matrix $H f(x)$ bestimmt:

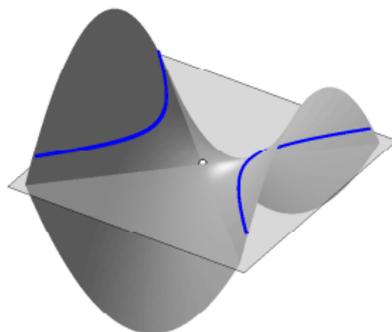
- Elliptischer Punkt:
Alle Eigenwerte λ_k sind ungleich Null und haben das gleiche Vorzeichen. Die Funktion f hat in diesem Fall ein lokales Minimum ($\lambda_k > 0$) oder lokales Maximum ($\lambda_k < 0$) bei x .

- **Hyperbolischer Punkt:**
Es gibt Eigenwerte λ_k mit verschiedenem Vorzeichen. In jeder Umgebung von x existieren dann sowohl kleinere als auch größere Funktionswerte als $f(x)$. Demzufolge bezeichnet man x als Sattelpunkt.
- **Parabolischer Punkt:**
Mindestens ein Eigenwert λ_k ist Null, und alle anderen Eigenwerte haben das gleiche Vorzeichen. In diesem Fall können Terme höherer Ordnung das lokale Verhalten beeinflussen. Der kritische Punkt kann ein Sattelpunkt oder ein lokales Extremum sein.
- **Flachpunkt:**
Alle Eigenwerte λ_k sind Null. Für eine glatte Funktion gilt dann $\|f(x + \Delta x) - f(x)\| \leq c\|\Delta x\|^3$, d.h. f wird in einer Umgebung von x mit hoher Ordnung durch eine waagrechte Ebene approximiert.

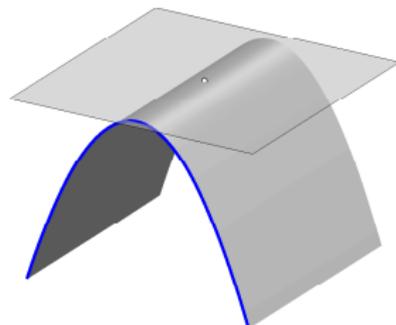
Die Motivation für die Bezeichnungen ist die Form der Höhenlinien im bivariaten Fall.



elliptischer Punkt



hyperbolischer Punkt



parabolischer Punkt

Bei Funktionen von zwei Veränderlichen kann der Typ anhand der Determinante und Spur der Hesse-Matrix klassifiziert werden. Ist $\det H f(x) > 0$ (< 0), so ist x ein lokales Extremum (ein Sattelpunkt). Für ein Minimum bzw. ein Maximum ist $\text{Spur } H f(x) > 0$ bzw. < 0 . Verschwindet die Determinante und ist die Hesse-Matrix nicht Null, so ist der Punkt parabolisch.

Beispiel

kritische Punkte der Funktion

$$f(x, y) = y(1 - x^2 - y^2)$$

Gradient und Hesse-Matrix

$$\text{grad } f = \begin{pmatrix} -2xy \\ 1 - x^2 - 3y^2 \end{pmatrix}, \quad Hf = \begin{pmatrix} -2y & -2x \\ -2x & -6y \end{pmatrix}$$

$$\text{grad } f = (0, 0)^t \iff$$

$$xy = 0 \quad \wedge \quad x^2 = 1 - 3y^2$$

\rightsquigarrow kritische Punkte

$$(0, \pm 1/\sqrt{3}), \quad (\pm 1, 0)$$

entsprechende Hesse-Matrizen

$$\begin{pmatrix} \mp 2/\sqrt{3} & 0 \\ 0 & \mp 6/\sqrt{3} \end{pmatrix}, \quad \begin{pmatrix} 0 & \mp 2 \\ \mp 2 & 0 \end{pmatrix}$$

Punkte $(x, y) = (0, \pm 1/\sqrt{3})$:

$$\det(H f) = (\mp 2/\sqrt{3})(-\mp 6/\sqrt{3}) = 4 > 0$$

\implies lokale Extrema

$$\text{Spur}(H f) = (-\mp 2/\sqrt{3}) + (-\mp 6/\sqrt{3}) = \mp 8/\sqrt{3}$$

\implies lokales Minimum bei $(0, -1/\sqrt{3})$ (Spur positiv) und lokales Maximum bei $(0, 1/\sqrt{3})$ (Spur negativ)

Funktionswerte: $f(0, \pm 1/\sqrt{3}) = \pm \frac{1}{\sqrt{3}}(1 - 1/3) = \frac{2}{9}\sqrt{3}$

Punkte $(x, y) = (\pm 1, 0)$

$$\det(H f) = -(\mp 2)(\mp 2) = -4 < 0$$

\rightsquigarrow Sattelpunkte mit Funktionswert $f(\pm 1, 0) = 0$

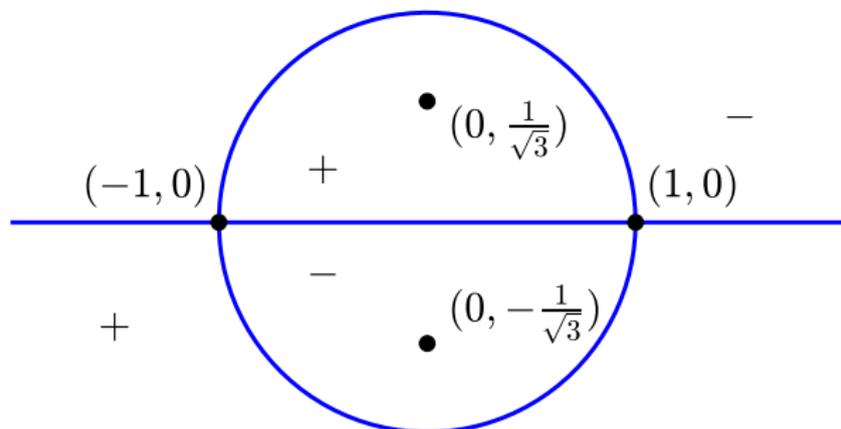
keine globalen Extrema, da

$$f(x, \pm 1) = \pm(1 - x^2 - 1) = \mp x^2 \rightarrow \mp \infty \text{ f\"ur } x \rightarrow \infty$$

Alternative Methode

Typbestimmung anhand der Nullstellenmenge und der sich daraus ergebenden Vorzeichenverteilung von f

$$f(x, y) = 0 \quad \Longleftrightarrow \quad y = 0 \quad \vee \quad x^2 + y^2 = 1$$



Sattelpunkte an Schnittpunkten mit Vorzeichenwechsel
lokale Extrema in den von der Nullstellenmenge eingeschlossenen
beschränkten Bereichen

Beispiel

kritische Punkte der Funktion

$$f(x, y) = (y - x + x^2)y$$

$$\text{grad } f = (-1 + 2x)y, 2y - x + x^2)^t = (0, 0)^t$$

↪ kritische Punkte $(0, 0)$, $(1, 0)$, $(1/2, 1/8)$

(i) Typbestimmung anhand der Vorzeichenverteilung von f :

positive und negative Bereiche begrenzt

durch die Gerade $G : y = 0$ und die Pa-

rabell $P : y = x - x^2$

Sattelpunkte an den Schnittpunkten von

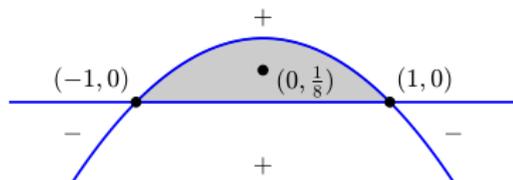
G und P , denn in jeder Umgebung existieren

sowohl positive als auch negative

Werte

lokales Minimum im grauen Bereich; f ist

im Innern negativ und Null auf dem Rand



(ii) Typbestimmung mit Hilfe der Hesse-Matrix:

$$Hf = \begin{pmatrix} 2y & 2x-1 \\ 2x-1 & 2 \end{pmatrix}$$

Einsetzen der kritischen Punkte \rightsquigarrow

$$Hf(0,0) = \begin{pmatrix} 0 & -1 \\ -1 & 2 \end{pmatrix}$$

$$Hf(1,0) = \begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix}$$

$$Hf(1/2, 1/8) = \begin{pmatrix} 1/4 & 0 \\ 0 & 2 \end{pmatrix}$$

Sattelpunkte bei $(0,0)$ und $(1,0)$, da $\det(Hf) < 0$

lokales Minimum bei $(1/2, 1/8)$, da $\det(Hf) > 0$ und $\text{Spur}(Hf) > 0$

Extrema multivariater Funktionen

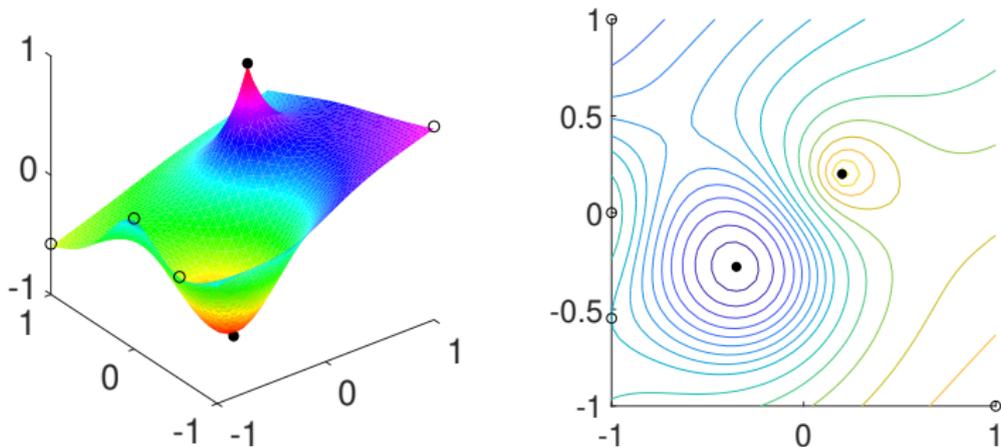
Lokale Extremstellen x_* (Minima oder Maxima) einer Funktion $f : \mathbb{R}^n \ni D \rightarrow \mathbb{R}$ können an folgenden Punkten auftreten:

- Unstetigkeitsstellen der partiellen Ableitungen,
- kritischen Punkten, d.h. Punkten mit $\text{grad } f(x_*) = (0, \dots, 0)^t$,
- Randpunkten des Definitionsbereichs D .

Eine hinreichende Bedingung für ein lokales Minimum (Maximum) in einem kritischen Punkt x_* im Innern des Definitionsbereichs D ist, dass ebenfalls die zweiten partiellen Ableitungen stetig sind und alle Eigenwerte der Hesse-Matrix $Hf(x_*)$ positiv (negativ) sind. Gibt es Eigenwerte mit verschiedenen Vorzeichen, so handelt es sich um einen Sattelpunkt, also kein lokales Extremum. Ist mindestens ein Eigenwert Null bei gleichen Vorzeichen der von Null verschiedenen Eigenwerte, so kann der Typ des kritischen Punktes x_* anhand der zweiten Ableitungen nicht klassifiziert werden.

Für lokale Minima (Maxima) an Randpunkten des Definitionsbereichs D muss die Richtungsableitung $\partial_v f(x_*)$ für jede ins Innere von D zeigende Richtung ≥ 0 (≤ 0) sein.

Eine globale Extremstelle von f kann durch Vergleich der Funktionswerte an allen Frage kommenden Punkten ermittelt werden.



Die Abbildung illustriert die verschiedenen Möglichkeiten für eine bivariate Funktion. Dabei sind lokale Extrema durch Kreise und globale Extrema

durch Punkte gekennzeichnet. Extrema im Innern werden von Niveaulinien umschlossen. In glatten Randbereichen berührt die durch eine Extremstelle verlaufene Niveaulinie den Rand.

Die Randbetrachtung ist für multivariate Funktionen komplizierter als für Funktionen einer Variablen. Bei der Bestimmung globaler Extrema wird man deshalb versuchen, Randpunkte als Kandidaten auszuschließen. Existiert beispielsweise für einen beschränkten Definitionsbereich D ein Punkt $y \in D$ mit einem kleineren Funktionswert als das Minimum der Funktionswerte auf dem Rand von D , dann kommen nur Punkte im Innern von D als Minimalstellen in Betracht. Gibt es nur eine Unstetigkeitsstelle einer partiellen Ableitung bzw. einen kritischen Punkt im Innern von D , so muss es sich um die Minimalstelle von f auf D handeln.

Für einen unbeschränkten Definitionsbereich muss man zusätzlich zeigen, dass $f(x) > f(y)$ für $|x|$ hinreichend groß, um die Existenz einer Minimalstelle im Innern von D zu folgern.

Bei der Bestimmung globaler Maxima verfährt man analog.

Beweis

betrachte die n -variate Funktion f entlang von Geraden, d.h. für einen beliebigen Vektor v mit $|v| = 1$ die univariate Funktion

$$g(t) = f(x_* + tv), \quad t \in \mathbb{R}$$

Kettenregel \implies

$$g'(t) = \sum_{k=1}^n \partial_k f(x_* + tv) v_k = \text{grad } f(x_* + tv)^t v$$

$$g''(t) = \sum_{j=1}^n \sum_{k=1}^n \partial_j \partial_k f(x_* + tv) v_j v_k = v^t H f(x_* + tv) v$$

(i) Notwendige Bedingung für Extrema:

x_* lokale Extremstelle von $f \implies t = 0$ lokale Extremstelle von g
und folglich

$$0 = g'(0) = \text{grad } f(x_*)^t v = \partial_v f(x_*)$$

Richtung v beliebig $\implies \text{grad } f(x_*) = (0, \dots, 0)^t$

Für ein lokales Minimum x_* am Rand folgt aus $g'(0) \geq 0$ für jede ins Innere von D zeigende Richtung v , dass $\partial_v f(x_*) \geq 0$
analoges Argument für ein lokales Maximum am Rand

(ii) Hinreichende Bedingung:

Eigenwerte von $Hf(x_*)$ positiv $\iff Hf(x_*)$ positiv definit \iff

$$v^t Hf(x_*)v \geq c > 0$$

Stetigkeit \implies

Die Ungleichung bleibt für $Hf(x)$ und $x = x_* + tv$ in einer Umgebung U : $t < \delta$ erhalten mit der kleineren Konstanten $c/2$ an Stelle von c .

Taylor-Approximation für $x \in U$ \implies

$$\begin{aligned} f(x) &= g(t) = g(0) + \underbrace{g'(0)}_0 t + \frac{1}{2} g''(\underbrace{s}_{\in [0, \delta)}) t^2 \\ &= f(x_*) + \frac{1}{2} v^t \underbrace{Hf(x + sv)}_{\in U} v t^2 \geq f(x_*) + \frac{1}{2} t^2 \frac{c}{2} > f(x_*) \end{aligned}$$

$\implies x_*$ lokale Minimalstelle von f in U

analoges Argument für ein lokales Maximum im Innern

Beispiel

Extrema einer quadratischen Funktion

$$f(x) = \frac{1}{2}x^tAx - x^tb + c, \quad A = A^t$$

Gradient und Hesse-Matrix

$$\text{grad } f = Ax - b, \quad Hf = A$$

kritische Punkte x_* : Lösungen von $Ax = b$

Eigenwerte von A ausschließlich positiv oder negativ $\implies \det A \neq 0$
und $Ax = b$ eindeutig lösbar

\rightsquigarrow genau ein lokales und damit ebenfalls globales Extremum von f

z.B.

$$f(x, y) = \frac{3}{2}x^2 + 2xy + \frac{3}{2}y^2 + x + 4y - 3, \quad A = \begin{pmatrix} 3 & 2 \\ 2 & 3 \end{pmatrix}, \quad b = \begin{pmatrix} -1 \\ -4 \end{pmatrix}$$

$$\text{grad } f = (0, 0)^t \quad \rightsquigarrow$$

$$\begin{pmatrix} 3 & 2 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -1 \\ -4 \end{pmatrix}$$

mit der Lösung $(x_*, y_*) = (1, -2)$

$\det A = 5$ und $\text{Spur } A = 6$ positiv

\implies Eigenwerte von A positiv

$\implies (1, -2)$ ist das globale Minimum von f

Beispiel

Bestimmung der globalen Extrema der Funktion

$$f(x, y) = \cos x + \cos y + \cos(x + y)$$

f ist 2π -periodisch bezüglich x und y und $f(y, x) = f(x, y) = f(-x, -y)$
 \implies Es genügt, den Bereich

$$D = (-\pi, \pi] \times [0, \pi]$$

zu untersuchen.

keine Randpunkte und Unstetigkeitsstellen von partiellen Ableitungen

\rightsquigarrow nur kritische Punkte relevant

$$\text{grad } f = \begin{pmatrix} -\sin x - \sin(x + y) \\ -\sin y - \sin(x + y) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

\iff

$$\sin x = -\sin(x + y) = \sin y$$

$\sin x = \sin y$ mit $0 \leq y \leq \pi$ und $-\pi < x \leq \pi$

\rightsquigarrow zwei mögliche Fälle

$$x = y \quad \vee \quad y = \pi - x$$

betrachte für beide Fälle die zweite Gleichung $\sin x = -\sin(x + y)$

(i) $x = y$:

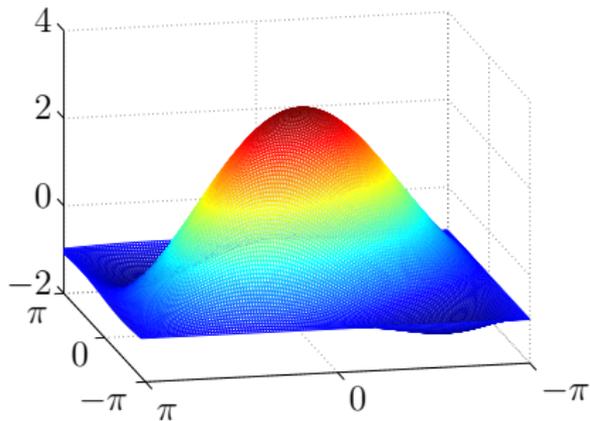
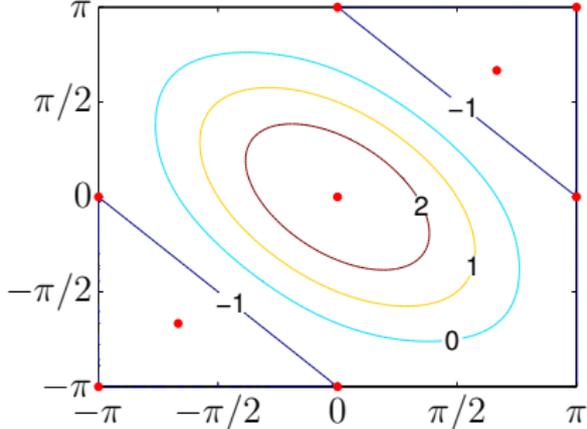
$$\sin x = -\sin(2x) = -2 \sin x \cos x \quad \iff \quad (1/2 + \cos x) \sin x = 0$$

\rightsquigarrow kritische Punkte $(0, 0)$, (π, π) und $(2\pi/3, 2\pi/3)$ im betrachteten Bereich

(ii) $y = \pi - x$:

$$\sin(x) = -\sin(\pi) = 0$$

\rightsquigarrow kritische Punkte $(0, \pi)$ und $(\pi, 0)$ im betrachteten Bereich



Vergleich der Funktionswerte \implies

globales Maximum mit Wert $f(0, 0) = 3$ bei $(0, 0)$

globales Minimum mit Wert $f(2\pi/3, 2\pi/3) = -3/2$ bei $(2\pi/3, 2\pi/3)$

Typbestimmung der anderen kritischen Punkte mit Hilfe der Hesse Matrix

$$Hf = \begin{pmatrix} -\cos x - \cos(x+y) & -\cos(x+y) \\ -\cos(x+y) & -\cos y - \cos(x+y) \end{pmatrix}$$

Sattelpunkt bei (π, π) , da

$$Hf(\pi, \pi) = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}, \quad \det Hf = -1 < 0$$

Sattelpunkte bei $(\pi, 0)$ und $(0, \pi)$ (symmetrische Lage), da

$$Hf(\pi, 0) = \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix}, \quad Hf(0, \pi) = \begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix}, \quad \det Hf = -1 < 0$$

Punktsymmetrie bzgl. $(0, 0)$ und Periodizität \rightsquigarrow

globale Maxima bei $(2k\pi, 2\ell\pi)$,

globale Minima bei $(2\pi/3 + 2k\pi, 2\pi/3 + 2\ell\pi)$,

$(-2\pi/3 + 2k\pi, -2\pi/3 + 2\ell\pi)$,

Sattelpunkte bei $(k\pi, \ell\pi)$,

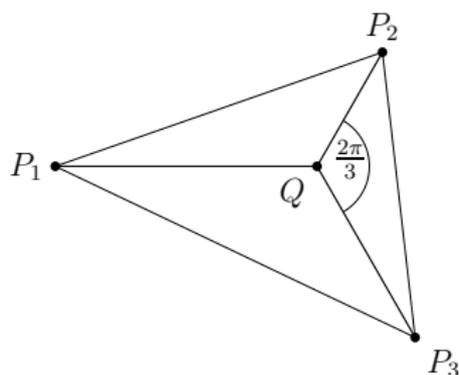
jeweils mit $k, \ell \in \mathbb{Z}$

Beispiel

Steiners Problem: Bestimme den Punkt $Q \in \mathbb{R}^2$, so dass die Summe der Abstände zu vorgegebenen Punkten P_k minimal wird.

Spezialfall dreier Punkte $P_1, P_2, P_3 \rightsquigarrow$

Q Randpunkt des Dreiecks $[P_1, P_2, P_3]$ oder $\sphericalangle(P_k, Q, P_\ell) = 2\pi/3 \forall k, \ell$



zeige die Charakterisierung für einen inneren Punkt:

d_k : Abstand zwischen $Q = (x, y)$ und $P_k = (x_k, y_k)$, d.h.

$$d_k^2 = (x - x_k)^2 + (y - y_k)^2$$

Kettenregel \implies

$$2d_k \frac{\partial d_k}{\partial x} = 2(x - x_k)$$

bzw.

$$\frac{\partial d_k}{\partial x} = \frac{x - x_k}{d_k}$$

und entsprechend $\partial d_k / \partial y = (y - y_k) / d_k$

\implies

$$\text{grad } d_k = \frac{1}{d_k} \begin{pmatrix} x - x_k \\ y - y_k \end{pmatrix}, \quad d_k = \left| \begin{pmatrix} x - x_k \\ y - y_k \end{pmatrix} \right|$$

Minimum von

$$f = d_1 + d_2 + d_3$$

bei innerem Punkt $Q \implies$

$$\text{grad } f(Q) = \sum_k \text{grad } d_k = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

da f in einer Umgebung von Q stetig differenzierbar ist

$$|\text{grad } d_k| = 1 \implies$$

Gradienten bilden gleichseitiges Dreieck (Vektorsumme null)

\rightsquigarrow Winkel $2\pi/3$ zwischen den Gradientenrichtungen

Ist x_* eine lokale Extremstelle einer Funktion $f : \mathbb{R}^n \ni D \rightarrow \mathbb{R}$ auf der durch die Nebenbedingungen $g_k(x_1, \dots, x_n) = 0$ definierten Menge D , dann existieren Lagrange-Multiplikatoren $\lambda_k \in \mathbb{R}$, so dass

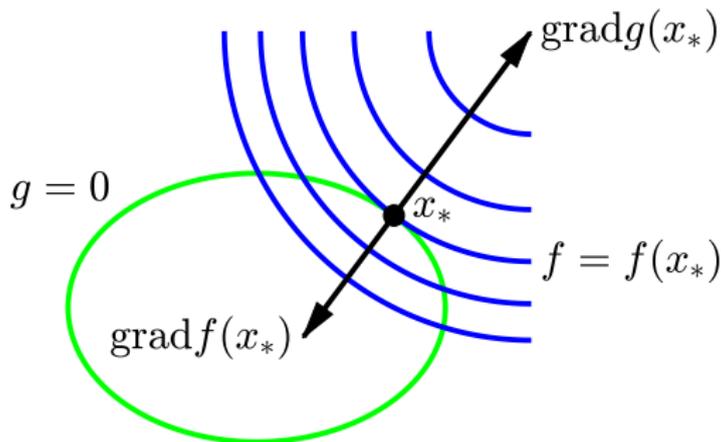
$$\text{grad } f(x_*) = \sum_k \lambda_k \text{grad } g_k(x_*).$$

Dabei wird vorausgesetzt, dass f und die Funktionen g_k in einer Umgebung von x_* stetig differenzierbar sind und dass die Gradienten $\text{grad } g_k(x_*)$ linear unabhängig sind.

Bei nur einer Nebenbedingung $g(x_1, \dots, x_n) = 0$ hat die Lagrange-Bedingung die einfache Form

$$\text{grad } f(x_*) \parallel \text{grad } g(x_*),$$

falls $\text{grad } g(x_*) \neq (0, \dots, 0)^t$, d.h. die Niveauflächen von f und g berühren sich an einer Extremstelle x_* . Dies ist in der Abbildung für bivariate Funktionen veranschaulicht.



Die Lagrange-Bedingung ist nicht hinreichend, um zu entscheiden, ob ein lokales Extremum vorliegt und ob es sich um ein Minimum oder ein Maximum handelt. Dies lässt sich nur mit Hilfe weiterer Informationen feststellen.

Die globalen Extrema erhält man durch den Vergleich der Funktionswerte an den Punkten, welche die Lagrange-Bedingung erfüllen, sowie gegebenenfalls den Randpunkten des Definitionsbereichs D oder Punkten x , an denen die Gradienten $\text{grad} g_k(x)$ linear abhängig sind.

Beweis

n : Anzahl der Variablen, m : Anzahl der Nebenbedingungen g_k

(i) $m \geq n$:

Der n -Vektor $\text{grad } f(x_*)$ ist immer als Linearkombination der nach Voraussetzung linear unabhängigen Gradienten $\text{grad } g_k(x_*)$ darstellbar.

✓

Grund: Für $m \geq n$, besteht die zulässige Menge im Allgemeinen bereits aus diskreten Punkten, die durch die Nebenbedingungen festgelegt sind.

(ii) $m < n$:

fasse die Nebenbedingungen g_k zu einer Funktion $g = (g_1, \dots, g_m)^t$ zusammen

partitioniere die Variablen als $x = (u, v) \in \mathbb{R}^m \times \mathbb{R}^{n-m}$, wobei nach eventueller Permutation die Invertierbarkeit der Jacobi-Matrix

$(\partial g(u, v) / \partial u)|_{(u_*, v_*)} = g_u(u_*, v_*)$ vorausgesetzt wird

Satz über implizite Funktionen \implies lokale Auflösbarkeit der Nebenbedingungen

$$g(u, v) = (0, \dots, 0)^t \iff u = \varphi(v), \quad (u, v) \approx (u_*, v_*)$$

Gradient der Funktion $v \mapsto h(v) = f(\varphi(v), v)$ Null an einem Extremum, d.h.

$$\text{grad } h(v_*)^t = f_u(u_*, v_*)\varphi'(v_*) + f_v(u_*, v_*) = (0, \dots, 0)^t$$

aufgrund der Kettenregel und mit φ' der Jacobi-Matrix von φ
Differenzieren der Nebenbedingungen $g(\varphi(v), v) = (0, \dots, 0)^t \implies$
 $g_u(\varphi(v), v)\varphi'(v) + g_v(\varphi(v), v) = (0, \dots, 0)^t$, d.h.

$$\varphi'(v) = -g_u(u, v)^{-1}g_v(u, v)$$

Setzen von $(\lambda_1, \dots, \lambda_m) = f_u(u_*, v_*)g_u(u_*, v_*)^{-1}$ und Einsetzen des
Ausdrucks für φ' in den Gradienten von $h \rightsquigarrow$

$$f_u = \lambda g_u, \quad f_v = -f_u(-g_u^{-1}g_v) = \lambda g_v$$

(u - und v -Komponenten der Lagrange-Bedingung $f' = \lambda g'$ im Punkt
 (u_*, v_*))

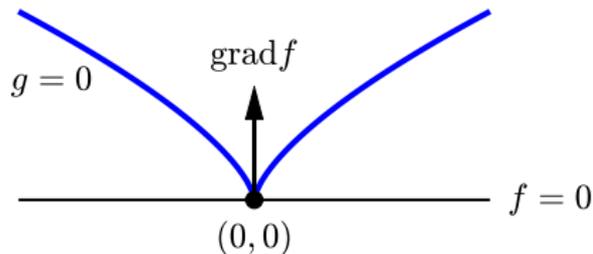
Beispiel

Minimierung von

$$f(x, y) = y$$

unter der Nebenbedingung

$$g(x, y) = y^3 - x^2 = 0$$



\rightsquigarrow Minimum bei $(x_*, y_*) = (0, 0)$

Die Lagrange-Bedingung $(f_x(0, 0), f_y(0, 0)) = \lambda(g_x(0, 0), g_y(0, 0))$ ist nicht erfüllt:

$$(0, 1) \neq (0, 0) = \lambda(-2x_*, 3y_*^2)$$

Grund: $\text{grad } g(x_*, y_*) = (0, 0)^t$

Die Lagrange-Bedingung ist in singulären Punkten (kein maximaler Rang der Jacobi-Matrix g' der Nebenbedingungen) nicht anwendbar.

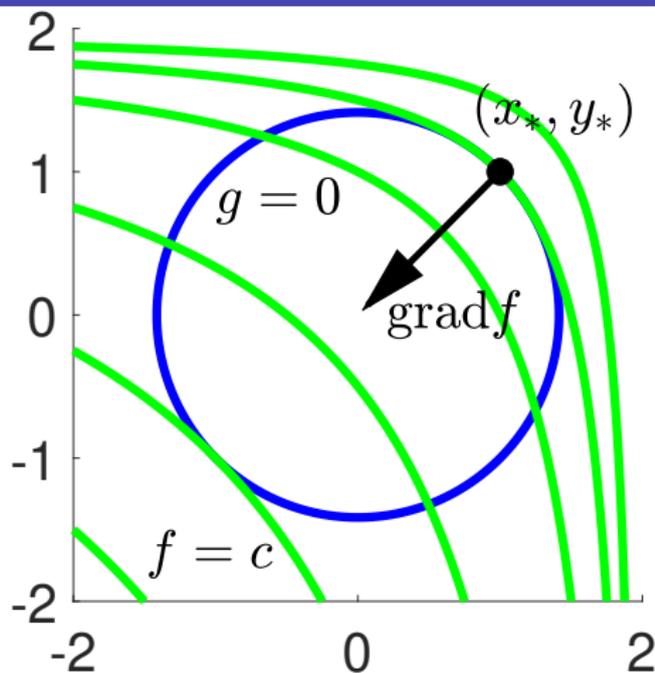
Beispiel

Minimierung von $f(x, y) = (x - 2)(y - 2)$ unter der Nebenbedingung
 $g(x, y) = x^2 + y^2 - 2 = 0$

Lagrange-Bedingung

$$\underbrace{(f_x, f_y)}_{\text{grad } f} = (y - 2, x - 2)$$
$$\stackrel{!}{=} \lambda(2x, 2y) = \lambda \underbrace{(g_x, g_y)}_{\text{grad } g}$$

Die Niveaulinien von f im Punkt (x_*, y_*) sind tangential zu der durch die Nebenbedingung $g = 0$ definierten Kurve.



Elimination von λ durch Bilden der Differenz $yf_x - xf_y$ in der Lagrange-Bedingung \rightsquigarrow

$$y(y - 2) - x(x - 2) = 0 \iff (y - x)(y + x - 2) = 0$$

zwei Fälle: (i) $x = y$ und (ii) $y + x - 2 = 0$

Berücksichtigung der Nebenbedingung $x^2 + y^2 - 2 = 0 \rightsquigarrow$

$(x, y) = (1, 1)$ oder $(-1, -1)$ im Fall (i)

und

$x^2 + (2 - x)^2 - 2 = 2(x - 1)^2 = 0$, d.h. ebenfalls $(x, y) = (1, 1)$ im Fall (ii)

Existenz von Minimum und Maximum für eine stetige Funktion auf einer kompakten Menge (Kreis mit Radius $\sqrt{2}$) und Vergleich der Funktionswerte

$$f(1, 1) = 1, \quad f(-1, -1) = 9$$

$\implies f$ bei $(1, 1)$ minimal

Beispiel

Bestimmung der Extrema der Funktion

$$f(x, y, z) = x + 2y - z$$

unter den Nebenbedingungen

$$g_1(x, y, z) = x^2 + y^2 - 8 = 0, \quad g_2(x, y, z) = x + z - 4 = 0$$

(Ellipse: Schnitt eines Zylinders mit einer Ebene)

Jacobi-Matrix der Nebenbedingungen

$$g'(x, y) = \begin{pmatrix} \text{grad } g_1^t \\ \text{grad } g_2^t \end{pmatrix} = \begin{pmatrix} 2x & 2y & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

voller Rang (\iff lineare Unabhängigkeit der Gradienten) für $(x, y) \neq (0, 0)$; auf zulässiger Menge erfüllt

Lagrange-Bedingung für Extremstellen (x, y, z)

$$\underbrace{(1, 2, -1)}_{\text{grad } f} = (\lambda_1, \lambda_2) \begin{pmatrix} 2x & 2y & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

bzw.

$$1 = 2\lambda_1 x + \lambda_2, \quad 2 = 2\lambda_1 y, \quad -1 = \lambda_2$$

Einsetzen von $\lambda_1 = 1/y$ und $\lambda_2 = -1$ in die erste Gleichung $\rightsquigarrow x = y$
Nebenbedingungen \rightsquigarrow mögliche Extrema $(2, 2, 2)$ und $(-2, -2, 6)$

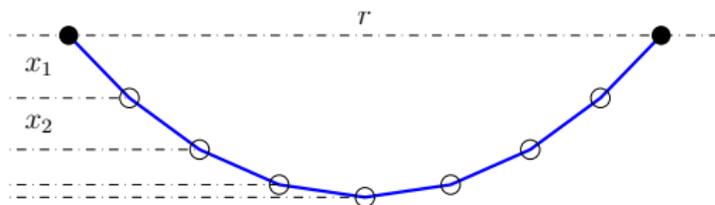
Existenz von Minimum und Maximum auf der Ellipse und Vergleich der Funktionswerte,

$$f(-2, -2, 6) = -12 < 4 = f(2, 2, 2),$$

$\implies f$ ist minimal bei $(-2, -2, 6)$ und maximal bei $(2, 2, 2)$.

Beispiel

Gleichgewichtslage einer an zwei Punkten aufgehängten Kette mit $2n$ Kettengliedern der Länge 1



potentielle Energie unter Berücksichtigung der Symmetrie

$$\begin{aligned} f(x) &= -2 \left(\frac{x_1}{2} \right) - 2 \left(x_1 + \frac{x_2}{2} \right) - \cdots - 2 \left(x_1 + \cdots + x_{n-1} + \frac{x_n}{2} \right) \\ &= -a_1 x_1 - \cdots - a_n x_n \end{aligned}$$

mit $a_k = 2(n - k) + 1$

Länge der Kette \rightsquigarrow Nebenbedingung

$$g(x) = r/2 - \sum_{k=1}^n \sqrt{1 - x_k^2} = 0$$

↪ Optimierungsproblem

$$f \rightarrow \min, \quad g = 0$$

Lagrange-Bedingungen $\text{grad } f = \lambda \text{ grad } g$

$$-a_k = \lambda \frac{x_k}{\sqrt{1-x_k^2}}, \quad k = 1, \dots, n$$

Quadrieren und Auflösen nach $x_k \implies$

$$a_k^2(1-x_k^2) = \lambda^2 x_k^2, \quad x_k^2 = \frac{a_k^2}{a_k^2 + \lambda^2}$$

Einsetzen in die Nebenbedingung $r/2 = \sum_k \sqrt{1-x_k^2} \rightsquigarrow$

$$\frac{r}{2} = \sum_{k=1}^n \sqrt{\frac{\lambda^2}{a_k^2 + \lambda^2}}$$

$\sqrt{\dots}$ monotone Funktion von λ

↪ einfach zu berechnende numerische Lösung λ_*

↪ Bestimmung von x_k aus den Lagrange-Bedingungen

Ist x_* ein lokales Minimum einer stetig differenzierbaren Funktion $f : \mathbb{R}^n \ni D \rightarrow \mathbb{R}$ auf der durch die Ungleichungen $g_k(x_1, \dots, x_n) \geq 0$ definierten Menge D und sind die Gradienten (ebenfalls als stetig vorausgesetzt) der aktiven Gleichungen $g_k(x_*) = 0$, $k \in I$, linear unabhängig, dann existieren Lagrange-Multiplikatoren $\lambda_k \geq 0$, so dass

$$\text{grad } f(x_*) = \sum_{k \in I} \lambda_k \text{grad } g_k(x_*).$$

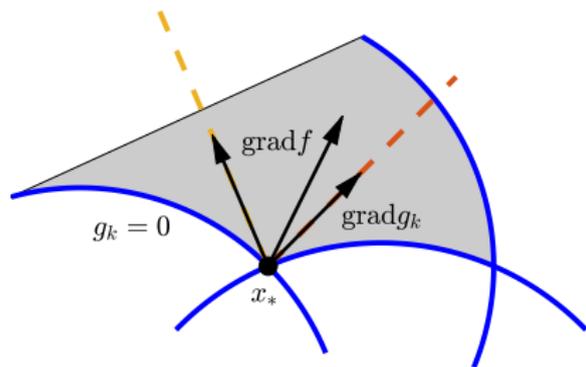
Für ein lokales Maximum ist entsprechend $\lambda_k \leq 0$.

Die Indexmenge I lässt sich auch implizit durch die Bedingungen

$$\lambda_k g_k(x_*) = 0$$

festlegen. Ist $g_k(x_*) > 0$, so folgt $\lambda_k = 0$, d.h. die nichttrivialen Multiplikatoren λ_k entsprechen den aktiven Nebenbedingungen ($g_k(x_*) = 0$).

Geometrisch bedeutet die Kuhn-Tucker-Bedingung für ein Minimum, dass der Gradient der Zielfunktion f in dem durch die Gradienten der aktiven Nebenbedingungen aufgespannten Kegel liegt. Dies ist in der Abbildung für bivariate Funktionen veranschaulicht. Bei einem Minimum x_* in einer Ecke liegt $\text{grad } f$ in dem durch die relevanten Gradientenrichtungen (gestrichelt) begrenzten Sektor.



Eine Gleichungsnebenbedingung $h(x_1, \dots, x_n) = 0$ kann durch die Ungleichungen $h \geq 0 \wedge -h \geq 0$ ersetzt werden. Damit ist das Kuhn-Tucker-Kriterium auch in diesem Fall anwendbar. Der entsprechende Term

$$\lambda_+ \text{grad } h + \lambda_- \text{grad}(-h) = \lambda \text{grad } h$$

besitzt einen Lagrange-Multiplikator $\lambda = \lambda_+ - \lambda_-$, dessen Vorzeichen nicht eingeschränkt ist.

Beweis

Die inaktiven Nebenbedingungen ($g_k(x_*) > 0$, $k \notin I$) sind irrelevant, da sie in der Umgebung von x_* keine Einschränkung bedeuten.

\rightsquigarrow Minimalität von f ebenfalls auf der kleineren Menge, die durch die Gleichungsbedingungen $g_k(x) = 0$ beschrieben wird

Satz über Lagrange-Multiplikatoren \implies behauptete Identität mit $\lambda_k \in \mathbb{R}$

zu zeigen: $\lambda_k \geq 0$

Indirekter Beweis: Annahme $\lambda_\ell < 0$ für ein $\ell \in I$

lineare Unabhängigkeit der Gradienten \implies

$$\exists v : (\text{grad } g_\ell(x_*))^t v = 1, \quad (\text{grad } g_k(x_*))^t v = 0, \quad k \in I \setminus \ell,$$

d.h. v liegt in der Tangentialebene der durch g_k , $k \neq \ell$, definierten Fläche S und hat eine nichttriviale Komponente in Richtung der Normalen der Fläche $S_\ell : g_\ell(x) = 0$

wähle eine Kurve $t \mapsto x(t) \in S$ mit Anfangspunkt $x(0) = x_*$ und Anfangsrichtung $x'(0) = v \rightsquigarrow$

$$g_\ell(x(t)) = g_\ell(x_*) + ((\text{grad } g_\ell(x_*))^t v) t + O(t^2) = 0 + t + O(t^2), \quad t \rightarrow 0$$

$\implies x(t)$ für hinreichend kleines $t > 0$ zulässig:

$$g_\ell(x(t)) \geq 0, \quad g_k(x(t)) = 0, \quad k \in I \setminus \ell$$

Konstruktion von v und Identität für $\text{grad } f \implies$

$$\begin{aligned} \frac{d}{dt} f(x(t))|_{t=0} &= (\text{grad } f(x(t)))^t x'(t)|_{t=0} = (\text{grad } f(x_*))^t v \\ &= \sum_{k \in I} \lambda_k (\text{grad } g_k(x_*))^t v = \lambda_\ell (\text{grad } g_\ell(x_*))^t v = \lambda_\ell < 0 \end{aligned}$$

\implies Abnahme von f entlang der Kurve, im Widerspruch zur Minimalität von $f(x_*)$

Beispiel

Extrema der Funktion

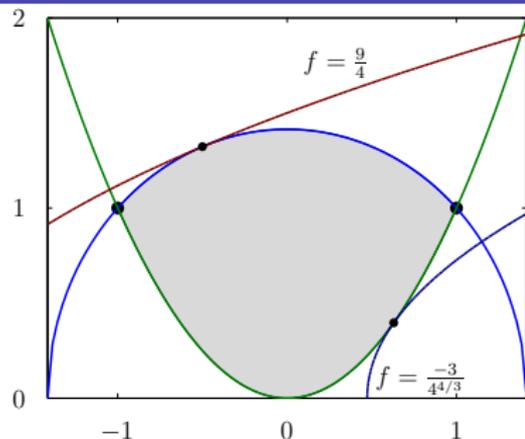
$$f(x, y) = y^2 - x$$

auf der durch

$$g_1(x, y) = y - x^2 \geq 0$$

$$g_2(x, y) = 2 - x^2 - y^2 \geq 0$$

definierten Menge D



Kuhn-Tucker-Bedingungen

$$\text{grad } f = (-1, 2y)^t = \underbrace{\lambda(-2x, 1)^t}_{\text{grad } g_1} + \underbrace{\varrho(-2x, -2y)^t}_{\text{grad } g_2}$$

$$\underbrace{\lambda(y - x^2)}_{g_1} = 0, \quad \underbrace{\varrho(2 - x^2 - y^2)}_{g_2} = 0$$

mit Lagrange-Multiplikatoren λ und ϱ gleichen Vorzeichens

lineare Unabhängigkeit der Gradienten der aktiven Nebenbedingungen für alle zulässigen Punkte (nicht Null und an den Schnittpunkten der Randkurven nicht parallel)

⇒ Notwendigkeit der Kuhn-Tucker-Bedingung für alle Extrema
verschiedene Fälle je nachdem welche Nebenbedingungen aktiv sind

(i) Keine Nebenbedingung aktiv:

$$\Rightarrow \lambda = \varrho = 0, \quad (-1, 2y) = (0, 0)$$

nicht erfüllbar, keine Extrema von f im Inneren von D

(ii) g_1 aktiv, d.h. $y - x^2 = 0$ und $g_2(x, y) > 0$:

$$\Rightarrow \varrho = 0, \quad (-1, 2y) = \lambda(-2x, 1)$$

$$\Rightarrow \lambda = 2y = 2x^2 \geq 0 \text{ und}$$

$$-1 = (2x^2)(-2x) \iff x = 4^{-1/3}, \quad y = 4^{-2/3}$$

⇒ Kuhn-Tucker-Bedingung für ein lokales Minimum

(iii) g_2 aktiv, d.h. $2 - x^2 - y^2 = 0$, $g_1(x, y) > 0$:

$$\implies \lambda = 0, \quad (-1, 2y) = \varrho(-2x, -2y)$$

$\implies \varrho = -1 < 0$ ($y = 0$ wegen $(\pm\sqrt{2}, 0) \notin D$ nicht möglich) und

$$x = -1/2, \quad y = \sqrt{7}/2$$

\implies Kuhn-Tucker-Bedingung für ein lokales Maximum

(iv) g_1 und g_2 aktiv, d.h. $y - x^2 = 0$ und $2 - x^2 - y^2 = 0$:

$\implies (x, y) = (1, 1)$ oder $(x, y) = (-1, 1)$

Einsetzen in die Gleichung für grad $f \rightsquigarrow$

$$(-1, 2) = \lambda(-2, 1) + \varrho(-2, -2) \implies \lambda = 1, \quad \varrho = -1/2$$

bzw.

$$(-1, 2) = \lambda(2, 1) + \varrho(2, -2) \implies \lambda = 1/3, \quad \varrho = -5/6$$

keine Extremstellen wegen der verschiedenen Vorzeichen der Lagrange-Multiplikatoren

Existenz von Extrema auf der kompakten zulässigen Menge \implies

Minimum im Fall (ii) bei $(4^{-1/3}, 4^{-2/3})$,

Maximum im Fall (iii) bei $(-1/2, \sqrt{7/2})$

alternative Typbestimmung durch Vergleich der Funktionswerte

$$f(4^{-1/3}, 4^{-2/3}) = -3\sqrt[3]{2}/8 < 9/4 = f(-1/2, \sqrt{7}/2)$$

Beispiel

Minimierung einer Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ auf einem Rechteck

$$D : a_k \leq x_k \leq b_k, \quad k = 1, \dots, n$$

Kuhn-Tucker-Bedingung für ein lokales Minimum

$$\text{grad } f(x_*) = \sum_{k=1}^n (\lambda_k - \varrho_k) e_k = (\lambda_1 - \varrho_1, \dots, \lambda_n - \varrho_n)^t$$

$$\lambda^t (x_* - a) = 0, \quad \varrho^t (b - x_*) = 0, \quad \lambda_k, \varrho_k \geq 0$$

mit e_k dem k -ten Einheitsvektor

$\pm\infty$ als Intervallgrenzen zugelassen

z.B. $a_k = -\infty \implies \lambda_k = 0$

\rightsquigarrow keine Einschränkung in den Kuhn-Tucker-Bedingungen

(i) $a_k < x_{*,k} < b_k$:

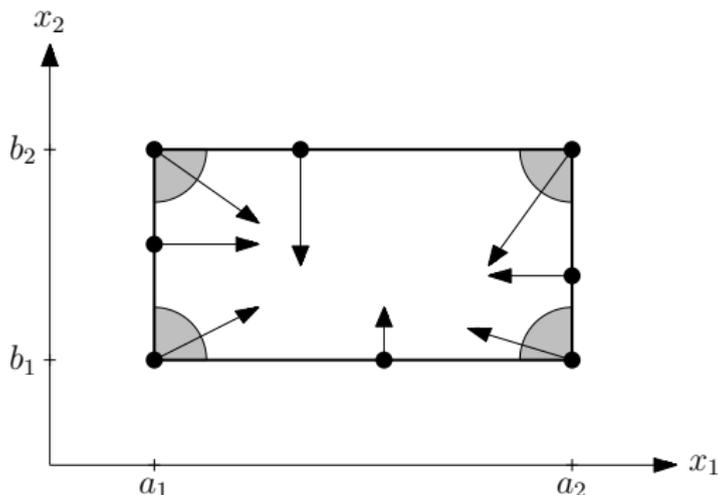
Lagrange-Multiplikatoren λ_k und ϱ_k Null und damit auch die k -te Komponente g_k von $\text{grad } f(x_*)$

(ii) Eine der Ungleichungen für x_k aktiv:

\rightsquigarrow entsprechender Lagrange-Multiplikator bestimmt Vorzeichen von g_k :

$$x_{*,k} = a_k \quad \Longrightarrow \quad \varrho_k = 0, \quad g_k = \lambda_k \geq 0$$

$$x_{*,k} = b_k \quad \Longrightarrow \quad \lambda_k = 0, \quad g_k = -\varrho_k \leq 0$$



mögliche Richtungen von $\text{grad } f(x_*)$ für lokale Minima von bivariaten Zielfunktionen auf einem Rechteck $[a_1, b_1] \times [a_2, b_2]$

- im Innern: keine Einschränkung für $\text{grad } f$
- auf einer Kante: $\text{grad } f$ orthogonal
- an einer Ecke: $\text{grad } f$ zeigt ins Innere

Beispiel

Lineares Programm:

Minimierung einer linearen Funktion

$$c^t x = \sum_{k=1}^n c_k x_k \rightarrow \min$$

unter den linearen Nebenbedingungen

$$Ax \geq b$$

mit einer $m \times n$ -Matrix A

Kuhn-Tucker-Bedingungen für ein lokales Minimum $x_* \in \mathbb{R}^n$

$$c^t = \lambda^t A, \quad \lambda^t (Ax_* - b) = 0$$

mit $\lambda_k \geq 0$ (entsprechend $\lambda_k \leq 0$ für ein Maximum)

z.B.

$$x + y \rightarrow \min, \quad x \geq 2, \quad y \geq 1, \quad x + 2y \geq 8,$$

d.h.

$$c = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 2 \end{pmatrix}, \quad b = \begin{pmatrix} 2 \\ 1 \\ 8 \end{pmatrix}$$

\implies Kuhn-Tucker-Bedingungen

$$1 = \lambda_1 + \lambda_3, \quad 1 = \lambda_2 + 2\lambda_3, \quad \lambda_1[x - 2] + \lambda_2[y - 1] + \lambda_3[x + 2y - 8] = 0$$

mit $\lambda_k, [\dots] \geq 0$

Bedingungen \implies genau einer der Lagrange-Multiplikatoren λ_k
gleich null

\rightsquigarrow drei Fälle (jeweils zwei Nebenbedingungen aktiv)

(i) $\lambda_1 = 0$:

$\implies \lambda_3 = 1, \lambda_2 = -1$

kein Extremum, da verschiedene Vorzeichen

(ii) $\lambda_2 = 0$:

$\implies \lambda_3 = 1/2, \lambda_1 = 1/2$

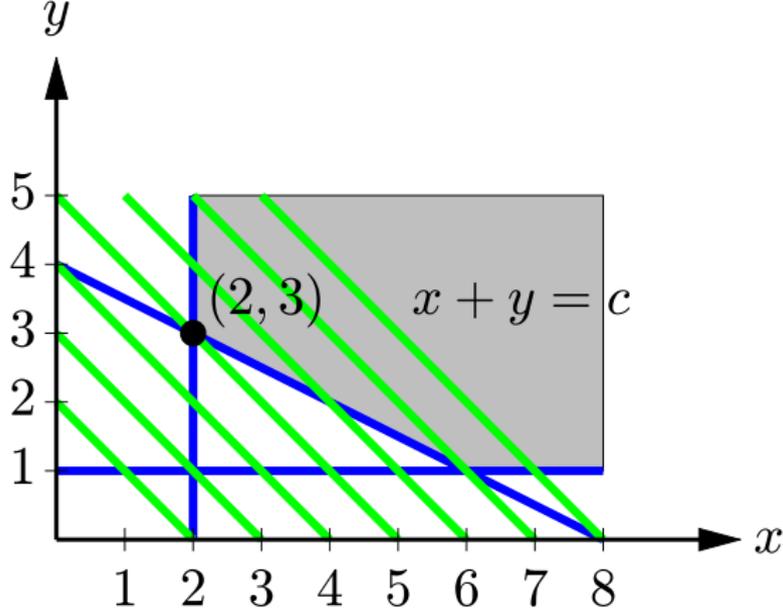
aktive Nebenbedingungen $x = 2, x + 2y = 8 \rightsquigarrow (x, y) = (2, 3)$

Kuhn-Tucker-Bedingungen für ein Minimum erfüllt

(iii) $\lambda_3 = 0$:

aktive Nebenbedingungen $x = 2, y = 1 \rightsquigarrow$ Punkt ausserhalb des zulässigen Bereichs

Beschränktheit von f nach unten auf dem zulässigen Bereich \implies
globales Minimum bei $(2, 3)$



geometrische Konstruktion der Lösung:

Niveaulinie der Zielfunktion berührt den zulässigen Bereich in $(2, 3)$

Zielfunktion steigt (fällt), wenn die Niveaugeraden den zulässigen Bereich schneiden (nicht schneiden)