Mathematik für Ingenieure III

I Funktionen von mehreren reellen Variablen

Definition I.1: Funktion $u = f(x)$ mit $x \in D$ $(D = f(x)) \subset R^n$ und $u \in R$ heißen Funktionen von n Variablen (auch (reelle-) skalarwertige Funktionen). Dabei sei $X = (x_1 ... x_n)$ mit $x_i \in R(1 \le i \le n)$
Wir schreiben auch
$U = f(x_1 ... x_n), (x_1 ... x_n) \in D \subset R^n$ (Je nach Zusammenhang schreibt man auch
$u = f(x, y), u = f(r, \varphi), u = f(x, y, z)$

Beispiel I.2: Kegelfunktion gegeben durch
$u = f(x, y) = \sqrt{x^2 + y^2}, (x, y) \in D := \{(x, y) \in R^2 \,|\, 0 \le x^2 + y^2 \le R^2\}$
und Höhenlinien (bzw. Niveau o. Äquvipotentiallinien)
$\sqrt{h} = \{(x, y) \in D \,|\, f(x, y) = h\}$

Partielle Ableitungen der Gradienten

Definition I.3: $D \subset R^n offen$
$f : D \to R$ und $(a_1, ..., a_n) \in D$. Existieren die Ableitungen der partiellen Funktion
$$x_i \mapsto f(a_1, a_2, ..., a_{i-1}, x_i, a_{i+1}, ..., a_n)$$
an der Stelle $x_i = a_i$, so nennt man dies die partielle Ableitung von f nach x_i im Punkt a_i.

Man bezeichnet
$$\frac{\partial F}{\partial x_i}(a) oder \frac{\partial f(x)}{\partial x} | x = a$$
(man beachte hier ∂ anstatt d !
Für partielle Ableitungen verwendet man auch die Notation
$$f_{xx} = \frac{\partial^2 f}{\partial x \partial x} = \frac{\partial^2 f}{\partial x^2} = f_{|u} \text{ bzw. } f_{xy} = \frac{\partial^2 f}{\partial y \partial x} = (fx)_y = f_{|12} = D_y D_x f$$

Kann die Reihenfolge der Differenzierung vertauscht werden?

Beispiel I.2: Sei $f(x, y) = \left\{ xy \frac{x^2 - y^2}{x^2 + y^2} für (x, y) \ne (0 \,|\, 0) \right\}$
Außerhalb des Punktes (0|0) gelten
$$\frac{\partial f}{\partial y} = \frac{x^4 y - 4x^2 y^3 - y^5}{(x^2 + y^2)^2} \text{ bzw. } \frac{\partial f(0, y)}{\partial x} = -y$$
und

$$\frac{\partial f}{\partial y} = \frac{-x^4 y - 4x^3 y^2 + x^5}{(x^2+y^2)^2} \text{ bzw. } \frac{\partial f(x,0)}{\partial y} = x$$

Da $f(0,y) = f(x,0) = 0$ gilt folgt:

$$\frac{\partial f(0,y)}{\partial x} = -y \text{ und } \frac{\partial f(x,0)}{\partial y} = x \qquad \left((x,y) \in R^2\right)$$

Es gilt nun

$$\frac{\partial f(0,0)}{\partial y \partial x} = -1 + 1 = \frac{\partial f(0,0)}{\partial x \partial y}$$

Satz I.5: $\frac{\partial^2 f}{\partial x \partial y} und \frac{\partial^2 f}{\partial y \partial x}$ sein in einer gewissen ε - Umgebung U vorhanden und in (x_0, y_0) stetig. Dann gilt:

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$$

Definition I.6: Die partielle Ableitung einer Funktion zu einem Vektor zusammen gefaßt, bezeichnet man als Gradienten von f an Der Stelle x

$$grad\ f(x) := \begin{pmatrix} \frac{\partial f(x)}{\partial x_1} \\ . \\ . \\ . \\ \frac{\partial f(x)}{\partial x_n} \end{pmatrix} \left(= \nabla f(x)\right) \in R^4$$

(bzw. $\nabla = \left(\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n}\right)^T$)

f heißt k-mal partiell diffbar. ,wenn alle k-ten Ableitungen existieren, sind sämtliche Ableitungen stetig, dann heißt f k-mal stetig partiell Diffbar.

Definition I.7: Sei $D \subset R^4$ offen. Die Menge der k-mal stetig part. diffb. Funktionen $f : D \to R$ bezeichnen wir jetzt mit $C^k(D,R)$ (oder kurz C^k, wenn D und R aus dem Zusammenhang klar sind)und nennen $f \in C^k(D,R)$ eine C^k - Funktion.

$C^k(D,R) := \{f : D \to R | f :$ k-mal stetig part. diffbar. $\}$

$C^0(D,R)\ (= C(D,R)) := \{f : D \to R | f$ ist stetig $\}$

Definition I.8: (Satz von Schwarz)
Sei Ø$\neq D = R^n$ offen . Dann sind für jedes $f \in C^m(D)\ (m \in N)$ die part. Ableitung der Ordnung $\leq m$ unabhängig von der Reihenfolge der Differentiation.

Definition I.9: (Landau-Symbol, klein o)

Für $f, g : R^n \subset D \to R$ und $x_0 \in D$, $k \in N$ schreibt man

$$f(x) = g(x) + o(|x - x_0|)^k$$

falls

$$\lim_{x \to x_0} \frac{f(x) - g(x)}{|x - x_0|^k} = 0$$

<u>Bemerkung:</u> O-Symbol ist keine eindeutig festgelegte Funktion.

Die Approximation einer diffb. Funktion einer Variablen $f : I \to R$ ($I \subset R$ offenes Intervall) in der Umgebung von $x_0 \in I$ durch eine lineare Funktion

$$g(x) = f(x_0) + f'(x_0)(x - x_0)$$

wird mit dem O-Symbol geschrieben als

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + o(|x - x_0|)$$

Für Funktionen von mehreren Variablen gewährt die partielle Differenzierbarkeit allein noch nicht die analoge Approximation durch lineare Funktionen.

Definition I.10: Sei $D \subset R^n$ offen. Eine Funktion $f : D \to R$ heißt in $x_0 \in D$ total differenzierbar (oder linear approximierbar, wenn es einen Vektor $a \in R^n$ mit

(i) $\quad f(x) = f(x_0) + < a, (x - x_0) > + o(|x - x_0|)$

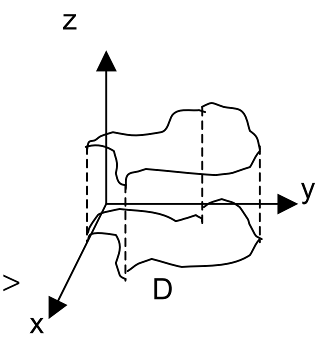
(für x nahe x_0) gilt

Anschauliche Deutung:

Der „über“ $D \subset R^2$ liegende Graph $z = f(x, y)$ wird in der Nähe des Flächenpunktes $(x_0, y_0, f(x_0, y_0))$ durch die Ebene

$$z = f(x_0, y_0) + < \begin{pmatrix} f_x(x_0, y_0) \\ f_y(x_0, y_0) \end{pmatrix}, \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix} >$$

mit einem Fehler $R(x, y) = o\sqrt{(x - x_0)^2 + (y - y_0)^2}$



Satz i.11: Ist f in $x_0 \in D$ total diffb., $f(x) = f(x_0) + < a, (x - x_0) > + o(|x - x_0|)$,

Dann gilt:

a) $\quad f$ ist stetig in x_0

b) $\quad \lim_{h \to 0} \frac{1}{h}(f(x_0 + hv) - f(x_0)) = < a, v > \quad (v \in R^n, v \neq 0)$

c) $\quad f$ ist part. diffb. und a eindeutig bestimmt als $a = grad\ f$.

Beweis:

a) $\quad \lim_{x \to x_0} f(x) = f(x_0) = \lim_{x \to x_0} < a, (x - x_0) > + o(|x - x_0|) = 0$

b) $\quad$ Für Punkte der Geraden $x_0 + hv$ ergibt (i)(in Def.I10) $f(x_0 + hv) - f(x_0) = h < a, v > + o(|hv|)$ Division durch h und anschließenden Grenzübergang liefert die Behauptung.

c) Man nimmt $v = e_i$ in b) und erhält

$$\frac{\partial f}{\partial x_i} = \lim_{h \to 0} \frac{1}{h} (f(x_0 + he_i) - f(x_0))$$ dann

$$\frac{\partial f}{\partial x_i}(x_0) = < a, e_i > = a_i$$

Im Sinne von Def.I.10 bedeutet

$$f(x) = f(x_0) + < grad\ f(x_0), x - x_0 > + o(|x - x_0|)$$

die lineare Approximation von $f(x)$ nahe bei x_0.

Satz I.12: Jede C^1- Funktion $f : D \to R$ $(D \subset R^n$ offen) ist auf D total diffbar.

Beispiel: In der Umgebung von $(1,1)$ lautet die lineare Approximation der Funktion

$$f(x, y) = x^4 + 2x^3 y^2 + y$$

$$f(x, y) = f(1,1) + < \nabla f(1,1), \begin{pmatrix} x-1 \\ y-1 \end{pmatrix} > + o(|\sqrt{(x-1)^2 + (y-1)^2}|)$$

$$= 4 + 10(x-1) + 5(y-1) + o$$

Dementsprechend

$$z = 4 + 10(x-1) + 5(y-1) \text{ bzw. } 10x + 5y - z = 11$$

die Gleichung der Tangentialebene der Fläche

$$z = x^4 + 2x^3 y^2 + y$$

im Flächenpunkt $(1,1,4)$

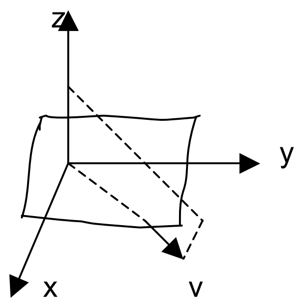
Richtungsableitung

Partielle Ableitungen $\frac{\partial f}{\partial x_i}(x)$ geben die Änderung der Funktion in Richtung der Koordinatenachsen an.
Wie sieht es mit anderen Richtungen aus?
Betrachten wir den Fall $n = 2$!

Man lege jetzt eine zur z-Achse parallele Ebene durch die Gerade

$$\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} + t \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}, \quad v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}, \quad |v| = 1$$

Die Schnittläufe des Graphen $z = f(x, y)$ besitzt die Darstellung:



$$x(t) = \begin{pmatrix} x_0 + tv_1 \\ y_0 + tv_2 \\ f(x_0 + v_1, y_0 + v_2) \end{pmatrix}$$

Mit dem Tangentialvektor:

$$\dot{x}(t) = \begin{pmatrix} v_1 \\ v_2 \\ \lim\limits_{t\to 0} \frac{1}{t}\left(f(f(x_0 + v1, y_0 + v_2) - f(x_0, y_0) \right) \end{pmatrix}$$

im Kurvenpunkt $(x_0, y_0, f(x_0, y_0))$.
Die Tangente hat den Anstieg

$$\lim_{t\to 0} \frac{1}{t}\left(f(f(x_0 + v1, y_0 + v_2) - f(x_0, y_0) \right)$$

Die motiviert den Begriff Richtungsableitung. Zu jedem Vektor $v \in R^n$, $|v| = 1$, nennen wir den Grenzwert

$$\partial_v f(x) = \lim_{t\to 0} \frac{1}{t}\left(f(x + tv_1) - f(x) \right)$$

(so fern er existiert) die Richtungsabweisung von f an der Stelle x in Richtung v_1.

<u>Bemerkung:</u> *Notation:* Anstatt $\partial_v f(x)$ verwendet man häufig auch Schreibweisen wie
$\frac{\partial f}{\partial v}(x)$ oder $D_v f(x)$

Satz i.14: Sei $D \subset R^n$ offen. Für alle total diffbar. Funktionen $f: D \to R$ und für jeden Vektor $v \in R^n$, $|v| = 1$, gilt

$$\partial v f(x) =< grad\ f(x), v >= \left(\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x) \cdot v_i \right)$$

Wie sieht es mit dem Änderungsverhalten einer Funktion längs eines im Def.bereich verlaufenden Kurve (z.B. Bahn eines bewegten Punktes) aus?

Diese Untersuchung führt zur Kettenregel.

Satz i.16: Für jede C^1- Funktion $f: D \to R$, $D \subset R^n$ offen, und für jedes Kurvenstück $x: R \supset [a,b] \to D$ gilt:

$$\frac{d}{dt} f(x(t)) = \frac{d}{dt} f(x_1(t), x_2(t), ..., x_n(t))$$
$$= \frac{\partial f}{\partial x_1}\left(x_1(t) \dot{x}_1(t) \right) + \frac{\partial f}{\partial x_2}\left(x_2(t) \dot{x}_2(t) \right) + ... + \frac{\partial f}{\partial x_n}\left(x_n(t) \dot{x}_n(t) \right)$$
$$=< grad\ f(x(t)), \dot{x}(t) >$$

<u>Bedeutung des Gradienten</u>

Sei $f: R^n \supset D \to R$ ein C^1- Funktion. Dann ist der Anstieg im Punkt $x \in D$ in Richtung eines Vektors $v \in R^n, |v| = 1$ gegeben durch

$$\partial v f(x) = < \nabla f(x), v > = |\nabla f(x)||v| \cos\alpha = |\nabla f(x)| \cos\alpha$$

wobei α der Winkel zwischen den Vektoren $\nabla f(x)$ und v ist.
$\partial v f(x)$ ist am größten, falls $\cos\alpha = 1$ d.h. $\alpha = 0$ ist

Also gilt: für $\nabla f(x) \neq 0$
Richtung von $\nabla f(x) =$ Richtung des stärksten Anstiegs von x f in x.

<u>Vektorwertige Funktionen</u>

Betrachten wir Funktionen, die jedem Ortsvektor $x \in D \subset R^n$ ein Vektor $f(\lambda) \in R^n$ zuordnen.

$$f: R^n \supset D \to R^n, f(x) = \begin{pmatrix} f_1(x) \\ . \\ . \\ . \\ f_n(x) \end{pmatrix} = \begin{pmatrix} f_1(x_1,..,x_n) \\ . \\ . \\ . \\ f_{n1}(x_1,..,x_n) \end{pmatrix}$$

und übertragen alle Begriffe skalarer Funktionen auf vektorwertige Funktionen.

Definition I.19: Sei $f: R^n \supset D \to R^n$ und $x_1, x_2 \in D$

i) Grenzwert, partielle Differentiation und Landau-Symbol „o“ sind komponentenweise definiert, d.h.

$$\lim_{x\to x_0} f(x) = \begin{pmatrix} \lim_{x\to x_0} f_1(x) \\ . \\ . \\ . \\ \lim_{x\to x_0} f_n(x) \end{pmatrix} ; \frac{\partial f}{\partial x_i}(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_i}(x) \\ . \\ . \\ . \\ \frac{\partial f_n}{\partial x_i}(x) \end{pmatrix}$$

$$f(x) = o(|x - x_0|) \Leftrightarrow f_k(x) = o|(x - x_0)| \quad (k = 1,...,m)$$

ii) f stetig, part. diffb. oder C^1- Funktion genau dann wenn sämtliche Komponentenfunktionen f_k von f $(k = 1,...,m)$ stetig part. diffb. bzw. C^1- Funktion sind.

iii) f heißt in $x_0 \in D$ total diffb. oder linear approximierbar, wenn es ein $(n \times n)$- Matrix A und eine ε- Umgebung $U_\varepsilon(x_0)$ in D gibt, so daß für alle $x \in U_\varepsilon(x_0)$ gilt:
$\otimes$ $f(x) = f(x_0) + A(x - x_0) + o(|x - x_0|)$
lineare Approximierbarkeit aller Komponentenfunktionen folgt aus $\otimes$, d.h.:

$$f_k(x_0) = f_k\left(x_0 + < \nabla f_k(x_0), (x - x_0) > + o(|x - x_0|)\right)$$

Man definiert daher die Funktionsmatrix oder Jacobi- Matrix von f in x:

$$J_f(x_0) = \begin{pmatrix} \nabla f_1(x_0)^T \\ . \\ . \\ . \\ \nabla f_{n1}(x_0)^T \end{pmatrix} = \left(\frac{\partial f}{\partial x_1}(x_0), ..., \frac{\partial f}{\partial x_n}(x_0)\right) = \left(\frac{\partial f_i}{\partial x_i}(x)\right)_{m \times n}$$

Mit Satz I.11 ist A in $\otimes$ eindeutig bestimmt als

$$A = J_f(x_0)$$

Die lin. App.eigenschaften schreibt man $\left(f \in C^1\right)$

$$f(x) \approx f(x_0) + J_f(x_0)(x - x_0)$$

mit einem Fehler

$$o(|x - x_0|)$$

Für $x \in R^n$, $v(x), w(x) \in R^m$, $f(x) \in R$ und $\alpha, \beta \in R$. Es gelten die Ableitungsregeln:

a) $J_{\alpha v + \beta w}(x) = \alpha J_\alpha(x) + \beta J_\beta(x)$

b) $J_{fv} = f(x)J_v(h) + v(x)J_f(x)$

$\quad = f(x)J_v(x) + v(x)(\nabla f(x))^T$

c) $J_{v^T w} = v(x)^T J_W(x) + w(x)^T J_v(x)$

d) $J_{v \times w} = v(x) \times J_w(x) - w(x) \times J_v(x)$

Kettenregel

Seien $f : R^n \supset D \to R^m, g : R^m \supset G \to R^2$ mit $f(D) \subset G$, f und g nacheinander ausgeführt ergeben die Komposition $g \circ f : D \to R^2, (g \circ f)(x) = g(f(x))$.

Satz I.20: Sind $f : R^n \supset D \to R^m$ in $x_0 \in D,\ g : R^m \supset G \to R^2$ in $f(x_0) \in G$ linear approximierbar, dann ist auch die Komposition $g \circ f$ linear approximierbar und es gilt:

$$J_{g \circ f}(x_0) = J_g\left(f(x_0) \cdot J_f(x_0)\right)$$

Räumliche Skalaren- und Vektorenfelder

Für den Fall R^3 betrachten wir spezielle Skalaren- und Vektorenfelder.

- ein Skalarfeld ordnet jedem Punkt $x \in D$ eine Zahl (Skalar) aus R zu.
- ein Vektorfeld ordnet jedem Punkt $x \in D$ einen Vektor $v(x) \in R^m, m > 1$.

Man spricht von einem C^k –Skalarfeld f oder einem C^k –Vektorfeld v auf D wenn f bzw. v C^k –Funktionen sind.

Jede C^1 – Funktion $f : R^3 \supset D \to R$ ordnen wir das Vektorfeld Gradient zu

$$\nabla f : D \to R^3 \quad , \quad \nabla f(x) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(x) \\ . \\ . \\ . \\ \frac{\partial f}{\partial x_3}(x) \end{pmatrix}$$

Jeder C^2 – Funktion f mit dem Laplace-Operator ∇ das Skalarfeld ∇f

$$\nabla f : D \to R \quad , \quad \nabla f(x) = \frac{\partial^2 f}{\partial x_1^2} + \frac{\partial^2 f}{\partial x_2^2} + \frac{\partial^2 f}{\partial x_3^2}$$ (kann man auch in bel. R^d def.)

Sei $v : R^3 \supset D \to R^3$ ein C^1 – Vektorfeld dann definieren wir ein Skalarfeld

„Divergenz" $div : D \to R \, , div \;\; v(x) = \frac{\partial v_1(x)}{\partial x_1} + \frac{\partial v_2(x)}{\partial x_2} + \frac{\partial v_3(x)}{\partial x_3}$

und ein Vektorfeld

„Rotation" $rot : D \to R^3 \, , rot \;\; v(x) = (\nabla \times v) = \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \\ \frac{\partial}{\partial x_3} \end{pmatrix} \times \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} \frac{\partial v_3}{\partial x_2} - \frac{\partial v_2}{\partial x_3} \\ \frac{\partial v_1}{\partial x_3} - \frac{\partial v_3}{\partial x_1} \\ \frac{\partial v_2}{\partial x_1} - \frac{\partial v_1}{\partial x_2} \end{pmatrix}$

Es beschreiben „$div \;\; v$"die Quellendichte und „$rot \;\; v$"die Wirbeldichte des Vektorfeldes v.

Rechenregeln

i) $rot(grad \;\; f) = 0$

ii) $div(rot \;\; v) = 0$

iii) $div(grad \;\; f) = \nabla f$

iv) $div(f \cdot v) = < grad \;\; f , v > + f \cdot div \;\; v$

v) $rot(f \cdot v) = grad \;\; f \times v + f \cdot rot \;\; v$

vi) $rot(rot \;\; v) = grad(div \;\; v) - \nabla v$

(Komponentenweise auswerten)

Man bezeichnet ein Vektorfeld v als ein Gradientenfeld, wenn es ein Skalarfeld f auf G gibt so daß

$$v(x) = grad \;\; f(x) \qquad (x \in G)$$

Die Funktion $U := -f(x)$ nennt man das Potential des Vektorfeldes v, und man sagt v besitzt ein Potential.

Satz I.22 Sei v ein C^1 – Vektorfeld auf der offenen Menge $D \subseteq R^n$. Dann kann v höchstens ein Potential besitzen, wenn seine Ableitungen symmetrisch sind, d.h. $\frac{\partial v_j}{\partial x_k} = \frac{\partial v_k}{\partial x_j} \; (j, k = 1, ..., n)$

Beweis: Im Falle $v = grad\ f$ ist nämlich $v_j = \frac{\partial f}{\partial x_j}$

$$\frac{\partial v_j}{\partial x_k} = \frac{\partial^2 f}{\partial x_j \partial x_k} = \frac{\partial^2 f}{\partial x_k \partial x_j} = \frac{\partial v_k}{\partial x_j}$$

Die Taylor-Formel:
Zur Erinnerung die Taylor-Formel im Eindimensionalen

Satz I.23 $I \subset R$ offen $f \in C^{k+1}(I)$

$$f(x) = f(x_0) + \frac{(x-x_0)}{1!} f'(x_0) + \frac{(x-x_0)^2}{2!} + \ldots + \frac{(x-x_0)^k}{k!} + R_{k+1}(x_k)$$

mit dem Lagrange-Restglied

$$R_{k+1}(x_k) = \frac{f^{k+1}(\xi)}{(k+1)!}(x-x_0) \quad \xi \in (x, x_0)$$

Verallgemeinern wir die Taylor-Formel aus Satz I.23 auf C^{k+1} – Funktionen $f: R^n \supset I \to R$ in $n > 1$ Veränderlichen.
Betrachten wir für $x \in D$ und $v \in R^n$ die Funktion $h(t) := f(x+tv) \quad (0 \le t \le 1)$
Dies setzt voraus ,daß die ganze x und $x+v$ verbindende Strecke in I liegt.
Der Einfachheit wegen setzen wir voraus, daß D konvex ist, d.h.

i) D offen
ii) $x, y \in D \Rightarrow x + t(y-x) \in D \quad (0 \le t \le 1)$

Die Taylor-Formel für $h(t)$ ausgewertet für $t = 0$ lautet:

$$h(1) = h(0) + h'(0) + \frac{1}{2} h''(0) + \ldots + \frac{1}{(k+1)!} h^{(k+1)}(\xi)$$

Die Ableitung $h^{(l)}(t)$ bzw. $h^{(l)}(0)$ lassen sich mit der Kettenregel bestimmen. Dazu verwenden wir den Diverentialoperator

$$\partial_v =< \nabla, v > \quad \text{(d.h.} \ \partial v := v_1 \frac{\partial}{\partial x_1} + v_2 \frac{\partial}{\partial x_2} + \ldots + v_n \frac{\partial}{\partial x_n}$$

mit der Eigenschaft

$$\partial_v C^l(D,R) \to C^l(D,R) f(x) \to \partial_v f(x)$$

$$h(1) = f(x+v)$$

$$h(0) = f(x)$$

$$\dot{h}(t) =< \nabla f(x+tv), v >; \dot{h}(0) = \partial_v f(x)$$

$$\ddot{h}(t) = \sum_{i=1}^{n} v_i \cdot < (\nabla f)_{x_i}(x+tv), V >; \ddot{h}(0) = \partial_v^2 f(x)$$

Satz I.24: (Taylor-Formel für n-Variablen)
Sei $D \subset R^n$ ein konvexes Gebiet, $f \in C^{k+1}(D,R)$, $x \in D$. Dann gilt mit $x+v \in D$ die Approximationsformel

$$\otimes \ f(x+v) = f(x) + \partial_v f(x) + \frac{1}{2!}\partial_v^2 f(x) + \ldots + \frac{1}{k!}\partial_v^k f(x) + R_{k+1}(x,v)$$

mit dem Restglied $R_{k+1}(x,v) := \frac{1}{(k+1)!}\partial_v^{k+1} f(x+\xi v) \quad \xi \in (0,1)$

Bemerkung: Mit Hilfe von Satz I.24 läßt sich eine Funktion $f(x)$ $\left(x \in R^n\right)$ in der Umgebung eines festen Punkts $x_0 \in D$ durch ein Polynom (sog. Taylor-Plynom) approximieren.

$$p(x) := f(x_0) + \partial_v f(x_0) + \frac{1}{2!}\partial_v^2 f(x_0) + \ldots + \frac{1}{k!}\partial_v^k f(x_0)$$

Für den Fehler gilt $f(x) - p(x) = R_{k+1}(x_0, x-x_0) = o\left(\left|x-x_0\right|\right)^2$

Für eine C^2 – Funktion heißt die nach dem Satz von Schwarz symmetrische Matrix

$$H_f(x) := \frac{\partial^2 f}{\partial x_i \partial x_j}(x) = \begin{pmatrix} f_{x_1x_1}(x) & . & . & . & f_{x_1x_m}(x) \\ . & & & & . \\ . & & & & . \\ . & & & & . \\ f_{x_mx_1}(x) & . & . & . & f_{x_mx_m}(x) \end{pmatrix}$$

die Hesse-Matrix im Punkt x.

Da $\partial_v f(x) = <\nabla f(x), v>$ und $\partial_v^2 f(x) = v^T H_f(x) v = <v, H_f(x)v>$

Ergibt sich die Taylor-Formel für einige spezielle $k = 0,1,2$ in $\otimes$ einsetze $x = x_0$ dann $v = x - x_0$

Korollar I.25: Spezialfälle der Taylor-Formel

a) $f(x) = f(x_0) + <\nabla f(\tilde{x}), x - x_0>$ (Mittelwertsatz) mit

$\tilde{x} = x + \lambda(x_0 - x) \quad (\lambda \in [0,1])$

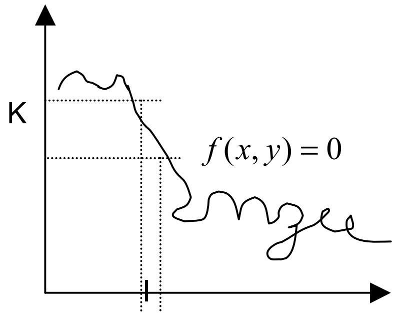
b) $f(x) = f(x_0) + <\nabla f(x_0), x - x_0> + \frac{1}{2}(x-x_0)^T \cdot H_f(\tilde{x}) \cdot (x - x_0)$ mit

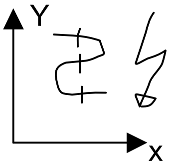
$\tilde{x} = x + \lambda(x_0 - x) \quad (\lambda \in [0,1])$

c) $f(x) = f(x_0) + <\nabla f(x_0), x - x_0> + \frac{1}{2}(x-x_0)^T \cdot H_f(\tilde{x}) \cdot (x - x_0) + o\left(\left|x-x_0\right|^2\right)$

Implizite Funktionen

Bisher hatten Funktionen $g: R \supset D \to R$ die wir Mit den Methoden der Differentialrechnung Untersucht haben, die explizite Form $y = g(x)$. Ist der Zusammenhang zwischen x und y jedoch eine Gleichung $f(x,y) = 0$ implizit gegeben, so ist y i.allg. nicht eindeutig festgelegt.





Definition I.25: Sei $f : R^2 \supset D \to R$. Man nennt, die durch $f(x,y)=0$ auf $I \subset R$ erklärte Funktion $g : I \to K \subset R$ eine implizite Funktion, wenn es zu jedem $x \in I$ genau ein $y \in I$ gibt mit $(x,y) \in D$ und $f(x,y)=0$. Dieses y wird mit $g(x)$ bezeichnet.

Beispiel: Durch $f(x,y) = e^y + y^3 + x^3 + x^2 - 1 = 0$ $(D = R^2)$ ist genau eine implizite Funktion $g(x) : R \to R$ erklärt, da zu jedem $x \in R$ $e^y + y^3 = 1 - x^2 + x^3$ genau ein $y = g(x)$ existiert. Durch formale Umrechnung kann die Gleichung

$$e^y + y^3 + x^3 + x^2 - 1 = 0$$

nicht nach y aufgelöst werden.

Satz I.26: (über implizite Funktionen)
Sei $D \subset R^2$ offen und $f : D \to R$ stetig diffbar. Ist nun $(x_0, y_0) \subset D$ ein Punkt der Niveaumenge $f(x_0, y_0) = 0$ mit $f_y(x_0, y_0) \neq 0$,dann gibt es ein Intervall $I \subset R$, $K \subset R$ mit Mittelpunkten x_0, y_0 so daß gilt:

i) $R := \{(x,y) | x \in I, y \in R\} \supset D$ und $f_y(x_0, y_0) \neq 0$ für alle $x, y \in R$

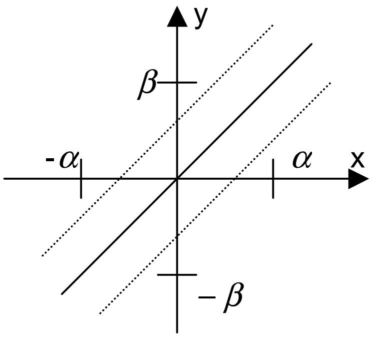
ii) Durch $f(x_0, y_0) = 0$ ist auf I eindeutig eine diffbare. implizite Funktion $g : I \to K$ erklärt mit der Ableitungen:

$$g'(x) = -\frac{f_x(x, g(x))}{f_y(x, g(x))} = -\frac{f_x(x,y)}{f_y(x,y)} \quad \text{für alle } x \in I$$

<u>Beweis:</u> OBdA ist $(x_0, y_0) = (0,0)$ und $f_y(0,0) > 0$

ad i) Da f_y stetig ist nach Voraussetzung, gilt $f_y(x,y) > 0$ in einer ganzen Umgebung $U_\varepsilon(0,0)$ $\varepsilon > 0$

ad ii) Betrachten wir nun f auf Rechtecken
$\{(x,y) | -\alpha \le x \le \alpha, -\beta \le y \le \beta\} \subset U_\varepsilon(0,0)$



Da die Funktion $y \to f(0,y)$ mit $f(0,0) = 0$ streng monoton wachsend ist, gilt $f(0,-\beta) < 0$ und $f(0,\beta) > 0$ für eine hinreichend kleines Intervall $-\alpha \le x \le \alpha$. Diese α und β definieren $I := (-\alpha, \alpha)$, $K := (-\beta, \beta)$. Für ein $x \in I$ ist $f(x,y)$ streng monoton wachsend und deshalb gibt es genau ein $y \in K$ mit $f(x,y) = 0$. Dieses eindeutig bestimmte y zu $x \in I$ mit $f(x,y) = 0$ bezeichnen wir mit $g(x)$. Für jede Folge (x_n) aus I mit $x_n \Rightarrow x$ gilt $g(x) \Rightarrow g$. Somit ist g stetig.

Sei $x \in I, y = g(x), \delta y = g(x + \delta x) - g(x)$. Für hinreichend kleines δx gilt nach dem Mittelwertsatz:

$f(x + \delta x, y + \delta y) - f(x,y) = f_x(\xi, \eta)\delta x + f_y(\xi, \eta)\delta y$ mit $\xi \in (x, x + \delta x)$, $\eta \in (y, y + \delta y)$

Nun folgt aus

$$f(x,g(x)) = f(x+\partial x, g(x+\delta x) = 0$$

und der Stetigkeit der partiellen Ableitung

$$\frac{\delta y}{\delta x} = -\frac{f_x(\xi,\eta)}{f_y(\xi,\eta)} \to -\frac{f_x(x,y)}{f_y(x,y)} \quad (\delta x \to o)$$

<u>Bemerkung:</u> Höhere Ableitungen von $g(x)$ erhält man mit der Kettenregel

$$f(x,g(x)) = 0 \overset{Abl.nach\,'x'}{\Rightarrow} f_x(x,g(x)) + f_y(x,g(x)) \cdot g'(x) = 0$$

$$\Rightarrow \quad g'(x) = -\frac{f_x(x,g(x))}{f_y(x,g(x))}$$

$$\overset{Abl.nach\,'x'}{\Rightarrow} f_{xx}(x,g(x)) + f_{xy}(x,g(x)) \cdot g'(x) + f_{yy}(x,g(x)) \cdot g'(x)$$

$$+ f_{yy}(x,g(x)) \cdot (g'(x))^2 + f_y(x,g(x)) \cdot g''(x) = 0$$

$$\Rightarrow g''(x) = -\frac{1}{f_y}\left(f_{yy} + 2 f_{xy} g' + f_{yy} (g')^2\right)$$

<u>Bemerkung:</u> Satz über implizite Funktionen läßt sich auf beliebige C^1 – Funktionen $f : R^n \supset D \to R$ verallgemeinern. Siehe Literatur!

<u>Beispiel:</u> $f(x,y) = e^y + y^3 + x^3 + x^2 - 1 \overset{!}{=} 0$

Es gilt: $g'(x) = -\dfrac{f_x(x,y)}{f_y(x,y)} = -\dfrac{3x^2 + 2x}{e^y + 3y^2}$

Für $x_1 = 0$ und $x_2 = -\dfrac{2}{3}$ gilt $g'(x) = 0$

$$g''(x) = -\frac{f_{xx}}{f_y} + \frac{g'}{f_y}\left(2 f_{xy} + f_{yy} g'\right) = \frac{6x+2}{e^y + 3y^2} + \ldots$$

$$g'(0) < 0, g'(-\frac{2}{3}) > 0$$

<u>Lokale Extrema</u>

Sei $f : R^n \supset D \to R$. Ein Punkt $x_0 \in D$ heißt lokale Maximalstelle (bzw. Minimalstelle) von f, wenn es eine Umgebung $U_\varepsilon(x_0)$ gibt, so daß gilt für alle $x \in U_\varepsilon(x_0)$ gilt

$$f(x) \leq f(x_0) \quad (\text{bzw. } f(x) \geq f(x_0))$$

Gelten diese Gleichungen für alle $x \in D$, so bezeichnet man diese Stellen als globale Maximal – bzw. Minimalstellen!

Wie bestimmt man solche Extremstellen?
Man unterscheidet zwei Charakterisierungsmengen, je nachdem ob diese Extremstellen im

- inneren
- auf dem Rand von D liegen

Satz I.28: „lokale Extremstellen im inneren“

Sei $x_0 \in R^n$ und $f \in C^1\left(U_\varepsilon(x_0)\right) \quad (\varepsilon > 0)$

Dann gilt: x_0 lokale Extremstelle von f

$\Rightarrow \nabla f(x_0)$

<u>Beweis:</u> Nach der Vorlesung besitzt Funktion $h(t) := f(x_0 + t_{e_i})$ in $t = 0$ eine lokale Extremstelle.

(e_i kanonischer Einheitsvektor)

Somit gilt:

$$0 = \dot{h}(t) = \frac{\partial f}{\partial x_i}(x_0)$$

Man bezeichnet x_0 einen Stationären Punkt, wenn $\nabla f(x_0) = 0$ gilt.
Mit Satz I.28 wissen wir nun, daß wir die lokalen Extremstellen von f in D unter den Stationären Punkten zu suchen haben. Ist ein stationärer Punkt keine Extremstelle, so spricht man von einem Sattelpunkt.
Ähnlich wie im eindimensionalen Fall charakterisierten wir die stationären Punkte in dem wir die „zweite Ableitung“ verwendeten.

Satz I.29: Sei $D \supset R^n$ offen, $x_0 \in D$ ein stationärer Punkt von $f \in C^2(D, R)$ und $H_f(x_0) = \left(f_{x_i y_i}(x_0)\right) \in R^{n \times n}$ die Hesse-Matrix von f in x_0

Dann gilt:

i) $H_f(x_0)$ pos. def. $\Rightarrow x_0$ ist eine lokale Minimalstelle

ii) $H_f(x_0)$ neg. def. $\Rightarrow x_0$ ist eine lokale Maximalstelle

iii) $H_f(x_0)$ indefinit $\Rightarrow x_0$ ist ein Sattelpunkt

<u>Beweis:</u> Für jedes x_0 in einer ε – Umgebung von x_0 gilt , mit

$$f(x) = f(x_0) + < \nabla f(x_0), (x - x_0) > + \frac{1}{2}(x - x_0)^T H_f(x^*)(x - x_0)$$

und $\nabla f(x_0) = 0$

$$f(x) - f(x_0) = \frac{1}{2}(x - x_0)^T H_f(x^*)(x - x_0) \quad \otimes$$

mit einem x^* zwischen x und x_0.

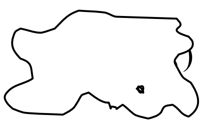
Da $H_f(x)$ stetig ist, ändern sich auch in einer ganzen Umgebung von x_0 die Vorzeichen der Eigenwerte von $H_f(x)$ nicht.

In dieser Umgebung ist die rechte Seite von $\otimes$ stets positiv (i), stets negativ (ii) oder sowohl positiv als auch negativ (iii). Somit

$f(x) > f(x_0)$ (i) oder

$f(x) < f(x_0)$ (ii)oder weder i) noch ii)

<u>Extremwertaufgaben mit Nebenbedingungen</u>

Beispiel: 

$$f(x, y) = \sqrt{(x - x_0)^2 + (y - y_0)^2} \overset{!}{=} \text{Minimiere}$$

NB: $h(x, y) - c = 0 \quad (C \in R)$

Wir untersuchen nun Extremwertaufgaben

$$f(x_1,...,x_n) \overset{!}{=} \text{Extr.}$$

Bei der die Menge der zulässigen Punkte $x=(x_1,...,x_n)$ durch eine oder mehrere Nebenbedingungen der Form

$$g_1(x_1,...,x_n)=0,...,g_k(x_1,...,x_n)=0$$

eingeschränkt ist.

Betrachten wir nun $k=1$. Sei $E:=\{x\in R^n \mid g(x)=0\}$. Nun bezeichnet $x_0\in E$ eine Maximalstelle (Minimalstelle) von f unter der NB $g(x)=0$, wenn es eine Umgebung $U_\varepsilon(x_0)$ gibt, so daß $f(x)\le f(x_0)\big(f(x)\ge f(x_0)\big)$ für alle $x\in U_\varepsilon(x_0)\cap E$ gilt. Die Aufgabenstellung schreiben wir kurz:

$$f(x)\overset{!}{=}Extr. \quad NB: g(x)=0$$ (maximiere $f(x)$, NB $g(x)=0$)

1 Methode: (expliziete Methode, funktioniert nur eingeschränkt)
Man löst $g(x)=0$ nach einer Variablen auf, z.B. $x_n=h(x_1,...,x_n)$
Und setzt diese in f ein, d.h. man eliminiert diese Variable aus f.
Man erhält somit eine Extremwertaufgabe <u>ohne</u> Nebenbedingung:

$$f(x_1,...,x_{n-1},h(x_1,...,x_{n-1}))\overset{!}{=}Extr.$$

2 Methode: (Parametrisierung der Nebenbedingung)
Sei $g:R^m\supset D\to R^n \ (n\le m)$ mit $g8x)=0$, d.h. es gibt n NB.
Man bestimme nun zu g eine Parametrisierung

$$f\big(x_1(Z),...,x_n(Z)\big) \qquad Z\in G\subset R^n$$

der Fläche $g(x)=0$ und löst für

$$F(z)=f(x_1(z),...,x_m(z))$$

das gewöhnliche EW-Problem

<u>Beispiel:</u> Man suche das Extremum von $f(x,y,z)=3x^2+z^3+4xy^4z$ auf der Fläche der Einheitskugel $s\in R^3$

$$f(x,y,z)\overset{!}{=}Extr. \qquad \text{NB } g(x,y,z):=x^2+y^2+z^2-1=0$$

Man wähle Kugelkoordinaten

$$x=r\cdot\sin\psi\cos\varphi$$
$$y=r\cdot\sin\psi\sin\varphi$$
$$z=r\cdot\cos\psi$$

Die Oberfläche der Einheitskugel hat dann die Parametrisierung:

$$r=1;\psi\in(0,\pi);\varphi=[0,2\pi)$$

$$f\big(x(r,\varphi,\psi),y(r,\varphi,\psi),z(r,\varphi,\psi)\big)=$$

$$3r^2\sin^2\psi\cos^2\varphi+r^3\cos^2\psi+4r^4\sin\psi\cos\varphi\sin^2\psi\sin^2\varphi\cos\psi$$

$$\tilde{f}(1,\varphi,\psi)=3\sin^2\psi\cos^2\varphi+...$$

3 Methode: (Lagrange-Methode)

Wir setzen voraus, daß f und g C^1-Funktionen sein und daß der Gradient ∇g nicht verschwindet in der gesuchten Extremstelle $x_0 \in E$. Da der $\nabla g \neq 0$ gilt, nehmen wir oBdA $\frac{\partial g}{\partial x_n}(x_0) \neq 0$. Nach dem Satz für implizite Funktionen ist die NB $g(x)=0$ lokal um x_0 nach x_n aufgelöst, d.h. es existiert die Funktion $x_n = h(x_1,...,x_{n-1})$. Da wir x_0 als Extremstelle angenommen haben gilt nun

$$\nabla f\underbrace{\left(\tilde{x}_1, \tilde{x}_2, ..., \tilde{x}_{n-1}, h\left(\tilde{x}_1, \tilde{x}_2, ..., \tilde{x}_{n-1}\right)\right)}_{x_0} = 0$$

Anwenden der Kettenregel und Ableitung einer implizierten Funktion liefert

$$0 = \frac{\partial f}{\partial x_i}(x_0) + \frac{\partial f}{\partial x_n}(x_0)\frac{\partial f}{\partial x_i}\left(\tilde{x}_1, ..., \tilde{x}_{n-1}\right)$$

$$= \frac{\partial f}{\partial x_i}(x_0) + \frac{\partial f}{\partial x_n}\left(-\frac{\frac{\partial g}{\partial x_i}(x_0)}{\frac{\partial g}{\partial x_n}(x_0)}\right)$$

$$0 = \frac{\partial f}{\partial x_i}(x_0) - \underbrace{\frac{\partial f}{\partial x_i}(x_0)\left(\frac{\partial g}{\partial x_n}(x_0)\right)^{-1}}_{=\lambda \in R}\frac{\partial g}{\partial x_i}(x_0)$$

$$\Rightarrow \exists \lambda \in R : \nabla f - \lambda \nabla g = 0$$

Mit obiger Schlußfolgerung haben wir den folgenden Satz gefunden

Satz I.30: Für jede Extremstelle x_0 des Problems

$f(x,y,z) \overset{!}{=} Extr.$, NB $g(x)=0$

mit C^1-Funktionen f und g und $\nabla g(x_0) \neq 0$ gilt, daß eine Zahl λ existiert, $\lambda \in R$, so daß

$$\nabla f(x_0) + \lambda_0 \nabla g(x_0) = 0$$

gilt.

Bemerkung: Unter der Voraussetzung $g, f \in C^1$ und $\nabla g(x_0) \neq 0$ lassen sich mit Satz I.30 mögliche Extremstellen finden.

Kochrezept I.31: (Anwenden von Lagrange-Multiplikatoren)

1.Schritt: Man definiert eine Hilfsfunktion
$L(x_1,...,x_n,\lambda) = f(x_1,...,x_n) + \lambda g(x_1,...,x_n)$

2.Schritt: Man löst das i.allg. nicht lineare Gleichungssystem
$\nabla L(x_1,...,x_n,\lambda) = 0$

3.Schritt: (schwierig)
Man stelle fest, welche gefundenen Punkte mit $\nabla L = 0$ wirklich Extremstellen sind. Das ist i.allg. recht schwierig und aufwendig. Häufig genügt ein direkter Vergleich der Funktionswerte.

Beispiel: Man maximiert das Volumen eines Quaders mit achsenparallelen Kanten innerhalb eines Ellipsoiden $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$. D.h. man untersucht das Problem maximiere $V(x,y,z) = 2x \cdot 2y \cdot 2z$ unter der Nebenbedingung $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$

Lagrange-Multiplikatoren:

Man setzt $L(x_1,...,x_n,\lambda) := 8xyz + \lambda\left(\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} - 1\right)$ und bekommt

$$\nabla L = \begin{pmatrix} 8yz + \frac{2\lambda x}{a^2} \\ 8xz + \frac{2\lambda x}{b^2} \\ 8xy + \frac{2\lambda x}{c^2} \\ \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} - 1 \end{pmatrix}$$

Man löst $\nabla L = 0$

$$\frac{\partial L}{\partial x}(x) = 8xyz + 2\lambda \frac{x^2}{a^2} \stackrel{!}{=} 0 \qquad (x \neq 0)$$

$$\Rightarrow \lambda \frac{x^2}{a} = -4xyz$$

ebenso

$$-4xyz = \lambda \frac{y^2}{b^2} = \lambda \frac{z^2}{c^2} \left(= \lambda \frac{x^2}{a^2} \right)$$

Da wir $\lambda = 0$ ausschließen können, gilt:

$\frac{y^2}{b^2} = \frac{z^2}{c^2} = \frac{x^2}{a^2}$ und es ergibt ich die Extremstelle aus

$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$ d.h. $x = \frac{a}{\sqrt{3}}$ und ebenso $y = \frac{b}{\sqrt{3}}, z = \frac{c}{\sqrt{3}}$

Maximales Volumen ist somit

$$V(x,y,z) = 8\frac{abc}{3\sqrt{3}} = \frac{8\sqrt{3}}{9}abc$$