6.2 Differentialrechnung in mehreren Variablen

Mathematik für Informatik, p.249

- Existenz der partiellen Ableitungen einer Funktion garantiert nicht einmal deren Stetigkeit
- deshalb ist es alleine kein brauchbares Werkzeug für Änderung von Funktionswert
- jetzt folgt wie man umfassende Ableitungen entwickelt.

1. Totale Ableitung

Mathematik für Informatik, p.249

Die Tangente ist eine lineare Approximation einer Funktion.
Für eine differenzierbare Funktion $f(x)$ verhält sich diese in der Nähe von x_0 ähnlich wie ihre Tangente:

$$t(x) = f(x_0) + f'(x_0)(x - x_0)$$

Das "ungefähr" bedeutet, dass der Fehler $(f(x) - t(x))$ für $x \to x_0$ die Größenordnung $o(|x - x_0|)$ hat.

- $o(|x - x_0|)$ bedeutet, dass der Fehler schneller gegen Null geht als $|x - x_0|$. Das Verhältnis des Fehlers zu $|x - x_0|$ strebt also für $x \to x_0$ gegen Null.

Übertragung auf Funktionen mit mehreren Variablen

Diese Idee kann auf Funktionen mit mehreren Variablen übertragen werden:

- Funktionen in zwei Variablen ($f : \mathbb{R}^2 \to \mathbb{R}$):
  - Verhalten sich lokal wie ihre Tangentialebenen.
  - Der Fehler ist hier ebenfalls verhältnismäßig klein.
  - Die lokale Änderung der Funktion ist eine lineare Abbildung $A : \mathbb{R}^2 \to \mathbb{R}$.
  - Diese lineare Abbildung lässt sich als Matrix darstellen.
- Allgemeiner Fall ($f : \mathbb{R}^n \to \mathbb{R}^m$):
  - Das Vorgehen lässt sich auf Funktionen anwenden, die von n Variablen abhängen und m Werte zurückgeben.

Definition

Definition 6.13

Mathematik für Informatik, p.249, Mathematik für Informatik, p.249

Sei $D \subseteq \mathbb{R}^n$ offen. Eine Funktion $f : D \to \mathbb{R}^m$ heißt im Punkt $x_0 \in D$ total differenzierbar, falls eine lineare Abbildung $\mathbf{f}' : \mathbb{R}^n \to \mathbb{R}^m$ existiert, so dass

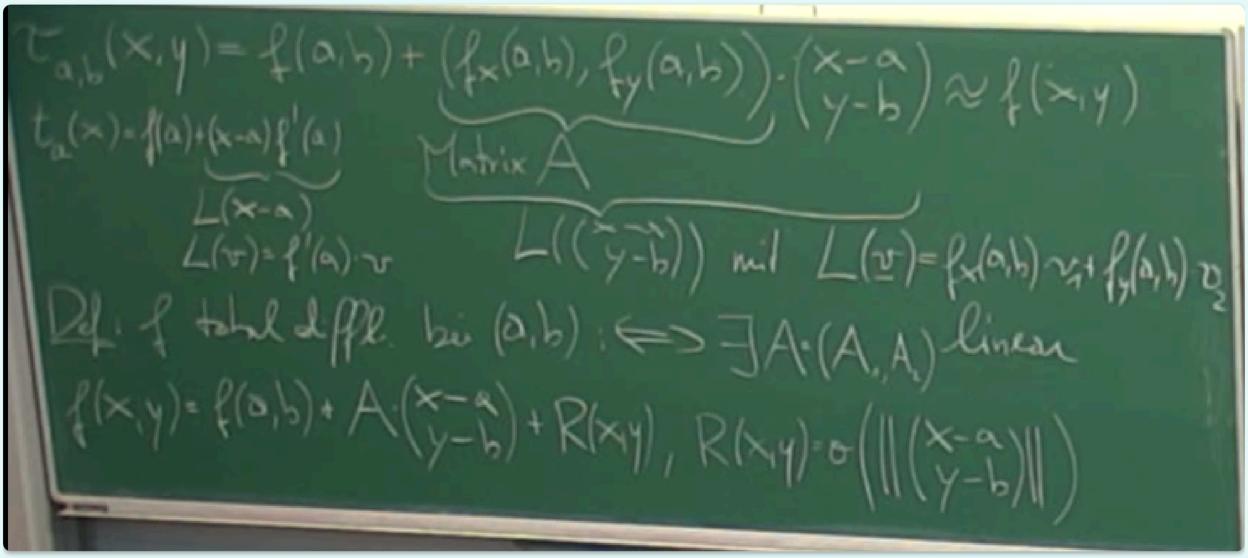
$$\mathbf{f}(\mathbf{x}) = \mathbf{f}\left(\mathbf{x}_0\right) + \mathbf{f}'\left(\mathbf{x} - \mathbf{x}_0\right) + \mathbf{R}(\mathbf{x})$$

gilt und der Rest $\mathbf{R}(\mathbf{x})$ die Bedingung

$$\lim_{x \to x_0} \frac{\|\mathbf{R}(x)\|}{\|x - x_0\|} = 0$$

erfüllt. Die lineare Abbildung f' heißt Ableitung von f im Punkt x_0, die dazu gehörige Matrix A heißt Jacobi-Matrix oder Funktionalmatrix. Setzen wir $\mathbf{x} = (x_1, \ldots, x_n)$ und $\mathbf{x}_0 = (x_{0,1}, \ldots, x_{0,n})$, so können wir die obige Gleichung ausführlicher schreiben als

$$\begin{pmatrix} f_1\left(x_1, \ldots, x_n\right) \\ \vdots \\ f_m\left(x_1, \ldots, x_n\right) \end{pmatrix} = \begin{pmatrix} f_1\left(x_{0,1}, \ldots, x_{0,n}\right) \\ \vdots \\ f_m\left(x_{0,1}, \ldots, x_{0,n}\right) \end{pmatrix} + A \begin{pmatrix} x_1 - x_{0,1} \\ \vdots \\ x_m - x_{0,n} \end{pmatrix} + \mathbf{R}\left(x_1, \ldots, x_n\right)$$



Bemerkung: Man beachte, dass die Bedingung (6.3) äquivalent zu

$$\lim_{x \to x_0} \frac{R(x)}{\|x - x_0\|} = 0$$

ist, wobei der Grenzwert koordinatenweise zu verstehen ist. Wir werden im Folgenden sowohl von (6.3) als auch von (6.5) Gebrauch machen.

ⓘ Satz aus vo (nicht im Buch gefunden)

Pasted image 20250513094737.png

Sei f eine total differenzierbare Funktion in (a, b). Dann existieren die partiellen Ableitungen $f_x(a, b)$ und $f_y(a, b)$, und der Gradient von f in (a, b) ist gegeben durch:

$$\nabla f(a, b) = (f_x(a, b), f_y(a, b))$$

Eine andere Schreibweise für den Gradienten, oft als Vektor dargestellt, ist:

$$\begin{pmatrix} A_x \\ A_y \end{pmatrix} = \begin{pmatrix} f_x(a, b) \\ f_y(a, b) \end{pmatrix} = \nabla f(a, b)$$

Beweis (Partielle Ableitung nach x):

Aus der Definition der partiellen Ableitung folgt:

$$\lim_{x \to a} \frac{f(x, b) - f(a, b)}{x - a}$$

Da f total differenzierbar in (a, b) ist, können wir die Zählerfunktion $f(x, b) - f(a, b)$ unter Verwendung des totalen Differentials ausdrücken. Die allgemeine Form des totalen Differentials für eine Funktion $f(x, y)$ ist:

$$f(x, y) - f(a, b) = A_x(x - a) + A_y(y - b) + R(x - a, y - b)$$

wobei $R(x - a, y - b)$ ein Restterm ist, für den gilt:

$$\lim_{(x,y) \to (a,b)} \frac{R(x - a, y - b)}{\sqrt{(x - a)^2 + (y - b)^2}} = 0$$

Im speziellen Fall der partiellen Ableitung nach x, betrachten wir nur die Änderung von x, während $y = b$ konstant bleibt. Daher wird der Term $A_y(y - b)$ zu $A_y(b - b) = 0$. Der Restterm vereinfacht sich ebenfalls zu $R(x - a, 0)$, den wir mit $R(x, b)$ bezeichnen.

Ersetzen wir dies in den Ausdruck für die partielle Ableitung:

$$\lim_{x \to a} \frac{A_x(x - a) + A_y(b - b) + R(x, b)}{x - a}$$

Da $b - b = 0$:

$$= \lim_{x \to a} \frac{A_x(x - a) + R(x, b)}{x - a}$$

Wir können den Bruch aufteilen:

$$= \lim_{x \to a} \left(A_x + \frac{R(x, b)}{x - a} \right)$$

Wir wissen, dass für den Restterm $R(x, b)$ gilt, wenn $y = b$:

$$R(x, b) = o(||(x - a, 0)||) = o(|x - a|)$$

Dies bedeutet, dass $\lim_{x \to a} \frac{R(x,b)}{|x-a|} = 0$.
Da der Nenner hier $x - a$ ist, gilt ebenso $\lim_{x \to a} \frac{R(x,b)}{x-a} = 0$.

Somit erhalten wir:

$$= A_x + 0$$

$$= A_x$$

Da die partielle Ableitung $f_x(a, b)$ per Definition dieser Grenzwert ist, folgt:

$$f_x(a, b) = A_x$$

Da A_x die Komponente der linearen Abbildung ist, die das totale Differential darstellt, und $f_x(a, b)$ die partielle Ableitung, bestätigt dies, dass $A_x = f_x(a, b)$. Ein analoger Beweis kann für $f_y(a, b)$ geführt werden.

ⓘ Satz aus der vo (nicht so direkt im Buch gefunden, aber leicht andere Version)

Pasted image 20250513095002.png, Mathematik für Informatik, p.249

Ist f total differenzierbar, so ist f stetig.

Beweis:
$f(x, y) \approx f(a, b) + A \cdot \begin{pmatrix} x - a \\ y - b \end{pmatrix}$

2. Ableitungsregeln

(Vergleiche 5.1 Ableitung > Ableitungsregeln)

ⓘ Summenregel

Mathematik für Informatik, p.252, Tafelbild

Summenregel überträgt sich direkt auf den mehrdimensionalen Raum.

Es gilt so wie im eindimensionalen:

$$(\mathbf{f} + \mathbf{g})' = \mathbf{f}' + \mathbf{g}'$$

Auch die Produktregel und die Kettenregel lassen sich übertragen.

ⓘ Satz 6.19 (Produktregel)

Mathematik für Informatik, p.252, Tafelbild

Sei $D \subseteq \mathbb{R}^n$ eine offene Menge. Weiters seien f, g zwei total differenzierbare, skalarwertige Funktionen. Dann gilt für die Funktion $h(\mathbf{x}) = f(\mathbf{x})g(\mathbf{x})$ die Gleichung

$$\operatorname{grad} h\left(\mathbf{x}_0\right) = f\left(\mathbf{x}_0\right) \cdot \operatorname{grad} g\left(\mathbf{x}_0\right) + g\left(\mathbf{x}_0\right) \cdot \operatorname{grad} f\left(\mathbf{x}_0\right).$$

Beweis. Es gilt nach Satz 6.15

$$
\begin{aligned}
h(\mathbf{x})-h\left(\mathbf{x}_0\right) = & f(\mathbf{x}) g(\mathbf{x})-f\left(\mathbf{x}_0\right) g\left(\mathbf{x}_0\right) \\
= & \left(f\left(\mathbf{x}_0\right)+\operatorname{grad} f\left(\mathbf{x}_0\right) \cdot\left(\mathbf{x}-\mathbf{x}_0\right)+R_1(\mathbf{x})\right) \\
& \cdot\left(g\left(\mathbf{x}_0\right)+\operatorname{grad} g\left(\mathbf{x}_0\right) \cdot\left(\mathbf{x}-\mathbf{x}_0\right)+R_2(\mathbf{x})\right)-f\left(\mathbf{x}_0\right) g\left(\mathbf{x}_0\right) \\
= & \left(f\left(\mathbf{x}_0\right) \cdot \operatorname{grad} g\left(\mathbf{x}_0\right)+g\left(\mathbf{x}_0\right) \cdot \operatorname{grad} f\left(\mathbf{x}_0\right)\right) \cdot\left(\mathbf{x}-\mathbf{x}_0\right)+R(\mathbf{x})
\end{aligned}
$$

mit

$$
\begin{aligned}
R(\mathbf{x}) = & R_1(\mathbf{x})\left(g\left(\mathbf{x}_0\right)+\operatorname{grad} g\left(\mathbf{x}_0\right) \cdot\left(\mathbf{x}-\mathbf{x}_0\right)\right)+R_2(\mathbf{x})\left(f\left(\mathbf{x}_0\right)+\operatorname{grad} f\left(\mathbf{x}_0\right) \cdot\left(\mathbf{x}-\mathbf{x}_0\right)\right) \\
& +\left(\operatorname{grad} f\left(\mathbf{x}_0\right) \cdot\left(\mathbf{x}-\mathbf{x}_0\right)\right)\left(\operatorname{grad} g\left(\mathbf{x}_0\right) \cdot\left(\mathbf{x}-\mathbf{x}_0\right)\right)+R_1(\mathbf{x}) R_2(\mathbf{x}) .
\end{aligned}
$$

Unter Berücksichtigung von $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} R_1(\mathbf{x}) /\left\|\mathbf{x}-\mathbf{x}_0\right\|=0$ und $\lim_{\mathrm{x} \rightarrow \mathrm{x}_0} R_2(\mathbf{x}) /\left\|\mathrm{x}-\mathbf{x}_0\right\|=0$ ist leicht nachzurechnen, dass $\lim_{\mathrm{x} \rightarrow \mathrm{x}_0} R(\mathrm{x}) /\left\|\mathrm{x}-\mathrm{x}_0\right\|=0$. Daraus folgt die Behauptung.

ⓘ Satz 6.20 (Kettenregel)

Mathematik für Informatik, p.252, Tafelbild

Sei $D \subseteq \mathbb{R}^n$ eine offene Menge, $f: D \rightarrow \mathbb{R}$ und $\mathrm{g}: \mathbb{R} \rightarrow \mathbb{R}^n$ mit $\mathbf{g}(x)=\left(g_1(x), \ldots, g_n(x)\right)$ und $\mathbf{g}(\mathbb{R}) \subseteq D$. Weiters sei $F(x)=f(\mathbf{g}(x))$. Dann gilt

$$
F^{\prime}(x)=\sum_{i=1}^n f_{x_i}\left(g_1(x), \ldots, g_n(x)\right) g_i^{\prime}(x)
$$

In Leibniz'scher Notation:

$$
\frac{d F}{d x}=\sum_{i=1}^n \frac{\partial f}{\partial g_i} \cdot \frac{d g_i}{d x}
$$

Die Zusammensetzung zweier vektorwertiger Funktionen $\mathrm{f}: R^m \rightarrow \mathbb{R}^p$ und $\mathrm{g}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ ist durch $(\mathbf{f} \circ \mathbf{g})(\mathbf{x})=\mathbf{f}\left(g_1(\mathbf{x}), \ldots, g_m(\mathbf{x})\right)$ definiert, wobei $\mathbf{g}(\mathbf{x})=\left(g_1(\mathbf{x}), \ldots, g_m(\mathbf{x})\right)$. Für die entsprechenden Jacobi-Matrizen gilt

$$
\frac{\partial(\mathbf{f} \circ \mathbf{g})}{\partial \mathbf{x}}\left(\mathbf{x}_0\right)=\frac{\partial \mathbf{f}}{\partial \mathbf{g}}\left(\mathbf{g}\left(\mathbf{x}_0\right)\right) \cdot \frac{\partial \mathbf{g}}{\partial \mathbf{x}}\left(\mathbf{x}_0\right)
$$

Folgerung: Falls $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ total differenzierbar und bijektiv ist, dann ist die JacobiMatrix der Umkehrfunktion gleich der Inversen der Jacobi-Matrix von f, also

$$
\frac{\partial \mathbf{f}^{-1}}{\partial \mathbf{x}}\left(\mathbf{y}_0\right)=\left(\frac{\partial \mathbf{f}}{\partial \mathbf{x}}\left(\mathbf{x}_0\right)\right)^{-1}
$$

mit $\mathbf{y}_0=\mathbf{f}\left(\mathbf{x}_0\right)$.

Beispiel zur Kettenregel

☰ Beispiel 6.21

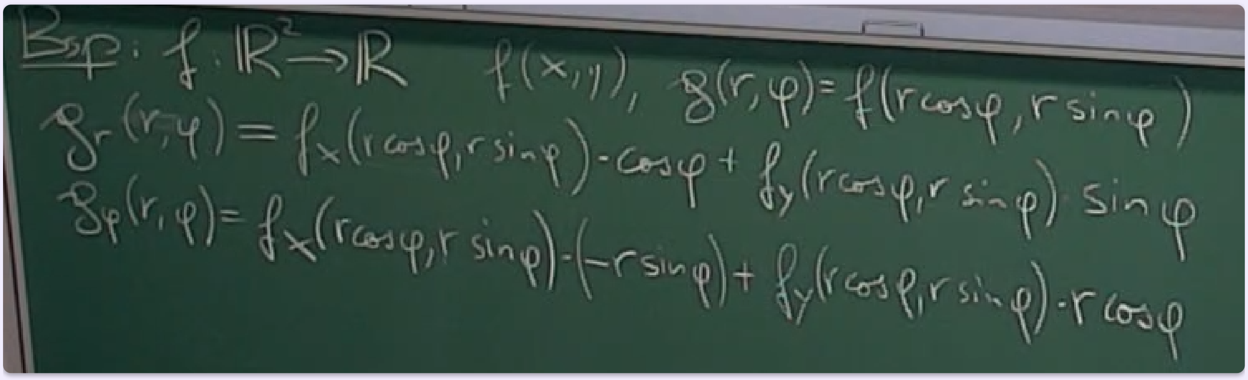
Mathematik für Informatik, p.253,

Wir betrachten eine Funktion $f : \mathbb{R}^2 \to \mathbb{R}$. Wie lässt sich die Änderung der Funktion beschreiben, wenn wir nicht in kartesischen, sondern in Polarkoordinaten rechnen. Die Transformation auf Polarkoordinaten geschieht mittels der Substitution $x = r\cos\varphi$ und $y = r\sin\varphi$. Aus der Funktion f entsteht dann die Funktion $F(r,\varphi) = f(r\cos\varphi, r\sin\varphi)$. Die partiellen Ableitungen von F ergeben sich nun aus der Kettenregel gemäß

$$\begin{aligned} F_r &= f_x \cos\varphi + f_y \sin\varphi \\ F_\varphi &= -f_x r \sin\varphi + f_y r \cos\varphi \end{aligned}$$

und nach Lösen dieses Gleichungssystems (in den Variablen f_x und f_y) folgt

$$\begin{aligned} f_x &= F_r \cos\varphi - \frac{1}{r} F_\varphi \sin\varphi, \\ f_y &= F_r \sin\varphi + \frac{1}{r} F_\varphi \cos\varphi. \end{aligned}$$



3. Richtungsableitung

Mathematik für Informatik, p.254

Die partiellen Ableitungen einer Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ geben den Anstieg der Funktion entlang der durch die Koordinatenachsen bestimmten Richtungen an. Sie sind also die Ableitungen von f in Richtung der Koordinatenachsen. Nun wollen wir entlang beliebiger Richtungen differenzieren.

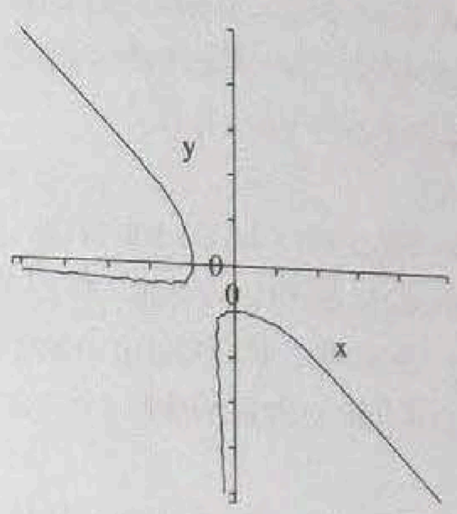


Abbildung 6.6 Die durch $e^{xy} + x + y = 0$ bestimmte ebene Kurve

Definition

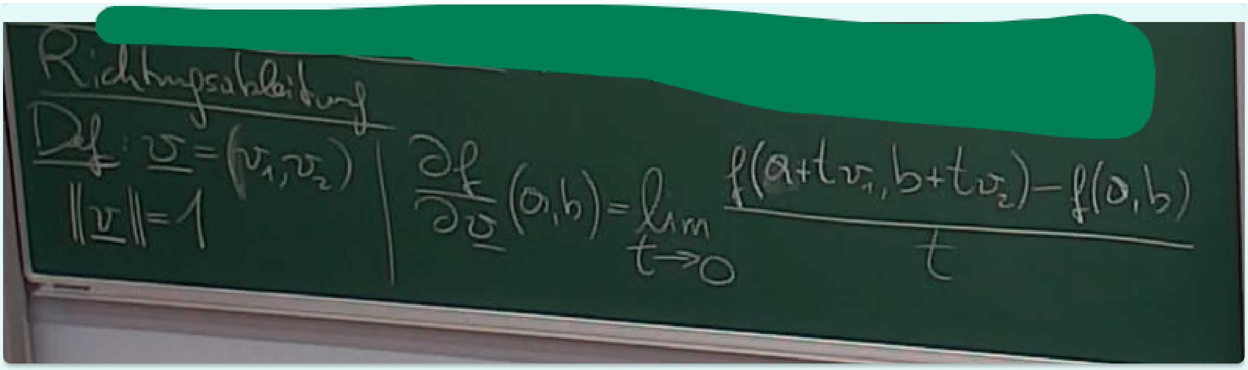
Definition 6.24

Mathematik für Informatik, p.255

Sei $D \subseteq \mathbb{R}^n$ eine offene Menge, $f : D \to \mathbb{R}$ eine skalarwertige Funktion und $v \in \mathbb{R}^n$ ein normierter Vektor, d.h. $\|v\| = 1$. Unter der Richtungsableitung von f an der Stelle $\mathrm{x} \in D$ nach v versteht man den Grenzwert

$$\frac{\partial f}{\partial \boldsymbol{v}}(\mathbf{x}) = \lim_{t \to 0} \frac{f(\mathbf{x} + t\boldsymbol{v}) - f(\mathbf{x})}{t}$$

In der vo haben wir es leicht anders definiert, weil wir's für 2d Funktionen gemacht haben:



ⓘ Satz 6.25

Mathematik für Informatik, p.255

Sei $D \subseteq \mathbb{R}^n$ eine offene Menge, $f : D \to \mathbb{R}$ eine an der Stelle $\mathbf{x} = (x_1, \ldots, x_n) \in D$ total differenzierbare Funktion und $\boldsymbol{v} = (v_1, \ldots, v_n) \in \mathbb{R}^n$ ein beliebiger normierter Vektor: Dann existiert die Richtungsableitung nach v, und es gilt

$$\frac{\partial f}{\partial \boldsymbol{v}}(\mathbf{x}) = f_{x_1}(\mathbf{x})v_1 + \cdots + f_{x_n}(\mathbf{x})v_n = \operatorname{grad} f(\mathbf{x}) \cdot \boldsymbol{v}$$

Beweis. Da f total differenzierbar ist, gilt

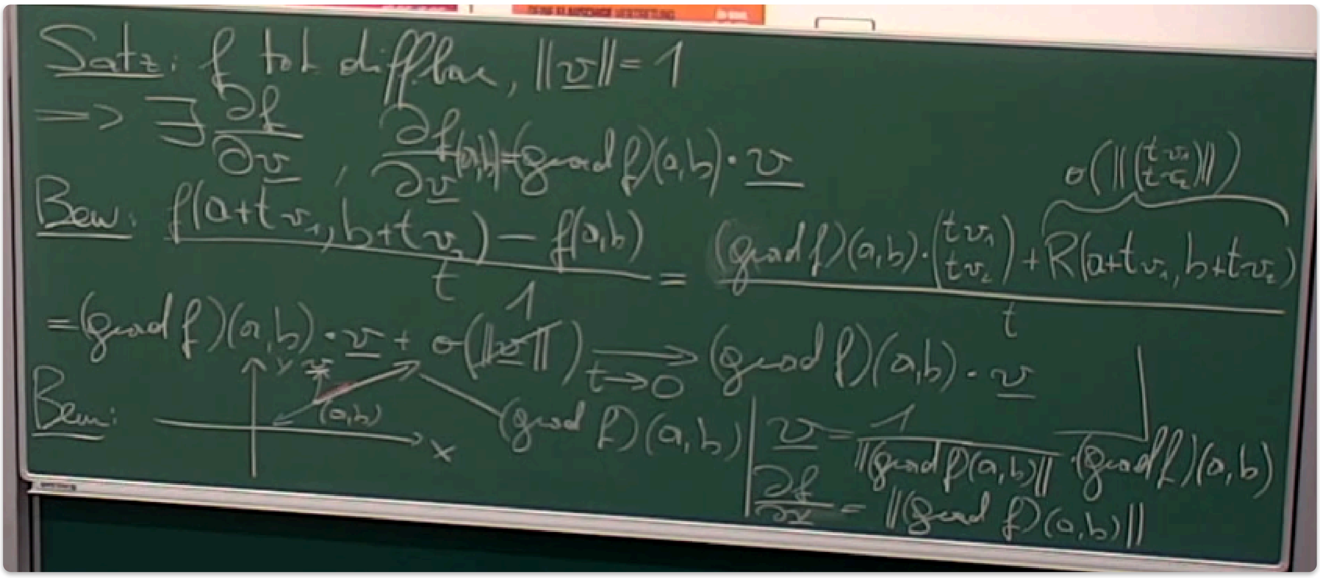
$$\begin{aligned}\frac{f(\mathbf{x} + t\boldsymbol{v}) - f(\mathbf{x})}{t} &= \frac{f_{x_1}(\mathbf{x})tv_1 + \cdots + f_{x_n}(\mathbf{x})tv_n + R(\mathbf{x} + t\boldsymbol{v})}{t} \\ &= f_{x_1}(\mathbf{x})v_1 + \cdots + f_{x_n}(\mathbf{x})v_n + \frac{R(\mathbf{x} + t\boldsymbol{v})}{t}\end{aligned}$$

Aus $t = \|\mathbf{x} + t\boldsymbol{v} - \mathbf{x}\|$ und (6.3) folgt $\lim_{t \to 0} \frac{R(\mathbf{x}+t\boldsymbol{v})}{t} = 0$. Nach Grenzübergang für $t \to 0$ erhält man nun die Behauptung.

Mathematik für Informatik, p.255

- Frage: In welcher Richtung wächst bzw. fällt eine Funktion f am stärksten?
- Es reicht, sich auf einen der beiden Fälle (Wachsen oder Fallen) zu beschränken.
- Die Existenz der Richtungsableitung nach v impliziert, dass die Funktion $g(t) = f(x + tv)$ an der Stelle $t = 0$ differenzierbar und daher linear approximierbar ist.
- Wichtige Eigenschaft: Die Richtungsableitung ändert ihr Vorzeichen, wenn man v durch $-v$ ersetzt.
- Erkenntnis: Wenn f in Richtung v am stärksten ansteigt, dann ist die Richtung des stärksten Abstiegs genau $-v$.

Also Länge vom Gradienten sagt wie steil es ist und die Länge sagt wo es am steilsten ist.



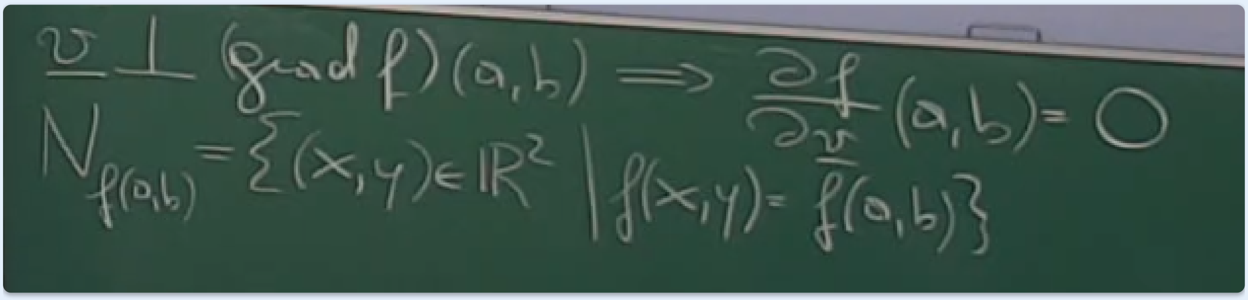
Satz 6.27

Mathematik für Informatik, p.256

Seien D, f und x wie in Satz 6.25. Dann ist die Richtung des größten Anstiegs genau die Richtung des Gradienten grad f. Der Wert des größten Anstiegs ist $\|\operatorname{grad} f\|$. Im Fall $\operatorname{grad} f = 0$ sind alle Richtungsableitungen gleich 0.

Beweis. Wir suchen jenen Vektor v, für den die zugehörige Richtungsableitung am größten ist. Nach dem vorigen Satz ist die Richtungsableitung nach v gleich $\operatorname{grad} f \cdot v$, und diese wird genau dann maximal, wenn v und $\operatorname{grad} f$ dieselbe Richtung haben. In diesem Fall gilt $\operatorname{grad} f \cdot \boldsymbol{v} = \|\operatorname{grad} f\|$. Falls $\operatorname{grad} f = 0$, dann gilt für jeden Vektor v natürlich $\operatorname{grad} f \cdot v = 0$.

Bemerkung: Es besteht ein einfacher Zusammenhang zwischen den Niveaulinien einer Funktion f, also jenen Kurven, entlang derer der Funktionswert konstant ist, und dem Gradienten von f. Es gilt: Falls $\operatorname{grad} f(\mathbf{x}) \neq 0$, dann steht $\operatorname{grad} f(\mathbf{x})$ normal auf die Niveaulinie, auf der $\mathbf{x}$ liegt.



Implizite Funktionen

Tafelbild

- **Definition:** Funktionen können implizit durch eine Gleichung der Form $F(x, y) = 0$ gegeben sein.
  - **Beispiel:** Der Einheitskreis wird durch die Gleichung $x^2 + y^2 - 1 = 0$ beschrieben.
- **Problemstellung:** Bei implizit gegebenen Funktionen ist die Lösbarkeit einer solchen Gleichung eine zentrale Frage.
- **Ziel:** Gesucht ist eine reellwertige Funktion $y(x)$ (d.h., y ist eine Funktion von x), sodass $F(x, y(x)) = 0$ gilt.
- **Lösung:** Diese Frage wird durch den folgenden Satz (Hauptsatz über implizite Funktionen) geklärt.

Satz 6.22 (Hauptsatz über implizite Funktionen)

Mathematik für Informatik, p.253, Mathematik für Informatik, p.254, Tafelbild

Seien $D \subseteq \mathbb{R}^2$ eine offene Menge und $F : D \to \mathbb{R}$ eine stetig differenzierbare Funktion. Weiters sei $F\left(x_0, y_0\right) = 0$ und $F_y\left(x_0, y_0\right) \neq 0$. Dann gibt es eine Umgebung U von (x_0, y_0), so dass die Gleichung $F(x, y) = 0$ in U eine eindeutig bestimmte stetige Lösung $y(x)$ hat. Die Funktion $y(x)$ ist darüber hinaus stetig differenzierbar und erfüllt

$$y'(x) = -\frac{F_x(x, y(x))}{F_y(x, y(x))}$$

Der Beweis dieses Satzes würde den Rahmen unseres Buches sprengen, aber die Gleichung für $y'(x)$ ist leicht zu zeigen. Man muss nur die definierende Gleichung nach der Kettenregel differenzieren. Aus $F(x, y(x)) = 0$ folgt

$$\frac{d}{dx} F(x, y(x)) = F_x(x, y(x)) + F_y(x, y(x)) y'(x) = 0$$

und damit

$$y'(x) = -\frac{F_x(x, y(x))}{F_y(x, y(x))}$$

Beispiel

Beispiel 6.23

Tafelbild, Ergänzung, Mathematik für Informatik, p.254

Die Lösung $y(x)$ der Gleichung $F(x, y) = e^{xy} + x + y = 0$ ist keine elementare Funktion. Die Lösungskurve ist in Abb. 6.6 dargestellt. Es gilt $F_x(x, y) = y e^{xy} + 1$ und

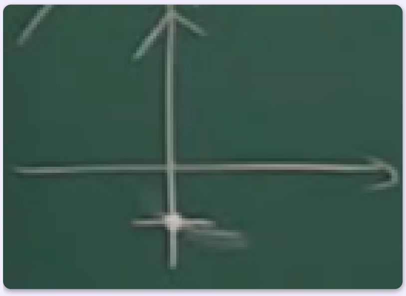
$F_y(x,y) = xe^{xy} + 1$. Obwohl die Lösungsfunktion nicht explizit durch einfache Funktionen ausgedrückt werden kann, ist es möglich, die Tangente an die Lösungsfunktion im Punkt (x_0, y_0) explizit anzugeben. Die Tangentengleichung ist nämlich durch $y = y_0 + y'(x_0)(x - x_0)$ gegeben, wobei man aus dem Hauptsatz die Darstellung

$$F_x(x_0, y_0)(x - x_0) + F_y(x_0, y_0)(y - y_0) = 0$$

konkret also

$$(ye^{xy} + 1)(x - x_0) + (xe^{xy} + 1)(y - y_0) = 0$$

erhält.



4. Taylorentwicklung

Mathematik für Informatik, p.256, p.257

- Einschränkung: Betrachtung des zweidimensionalen Falls.
- Gegeben: Offene Menge $D \subseteq \mathbb{R}^2$ und eine hinreichend oft stetig differenzierbare Funktion $f : D \to \mathbb{R}$.
- Jacobi-Matrix: Die Jacobi-Matrix von f an $(x_0, y_0) \in D$ ist die lineare Approximation von f.
  - Lokal wird f durch eine Ebene (Tangentialebene) angenähert.
- Verallgemeinerung der Taylorreihe:
  - Entwicklungspunkt: (x_0, y_0).
  - Weiterer Punkt: $(x, y) = (x_0 + h, y_0 + k)$.
  - Hilfsfunktion: $F(t) = f(x_0 + th, y_0 + tk)$ auf der Verbindungsstrecke von (x_0, y_0) nach (x, y).
  - Taylorreihe von $F(t)$ um $t_0 = 0$:

$$F(0) + F'(0)t + \frac{F''(0)}{2}t^2 + \frac{F'''(0)}{3!}t^3 + \cdots$$

  - Satz von Taylor für $t = 1$:

$$f(x, y) = F(1) = F(0) + F'(0) + \frac{F''(0)}{2} + \frac{F'''(0)}{3!} + \cdots + \frac{F^{(n)}(0)}{n!} + \frac{F^{(n+1)}(\xi)}{(n+1)!}$$

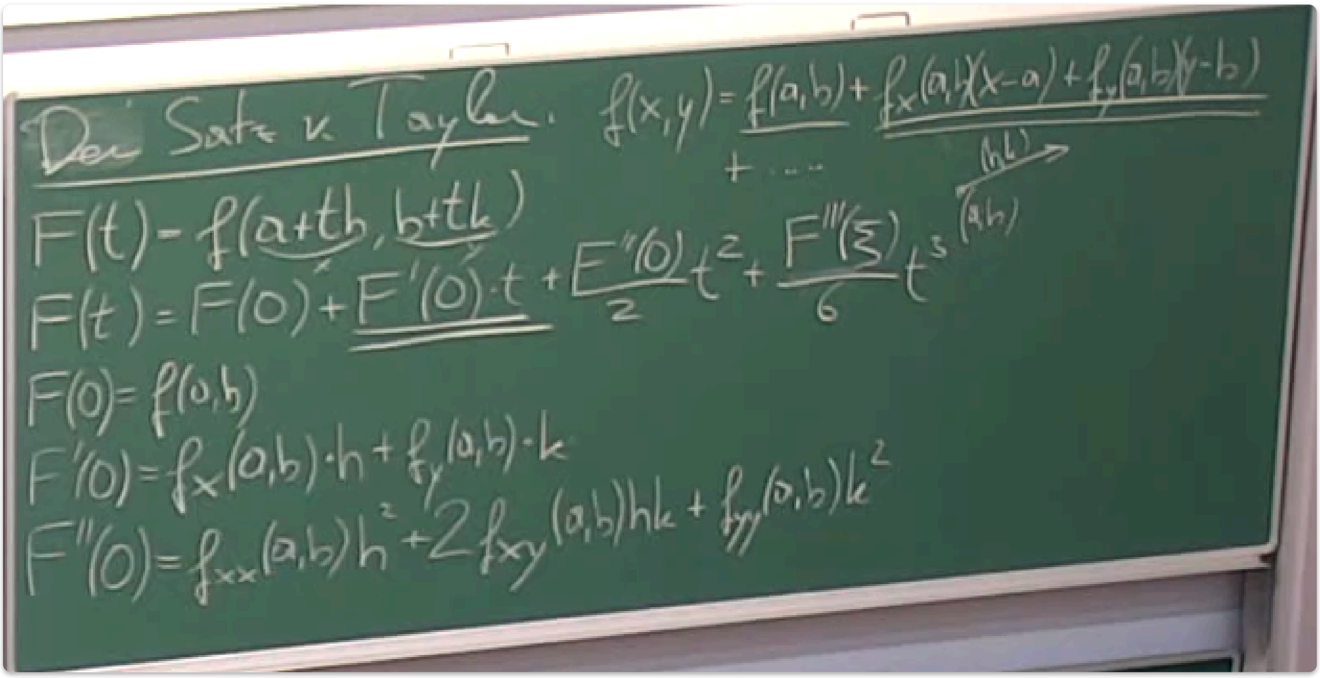
  mit $0 < \xi < 1$.

Ableitungen von $F(t)$ mit der Kettenregel

- Erste Ableitung $F'(0)$:
  - $$F'(0) = f_x(x_0, y_0)h + f_y(x_0, y_0)k$$
  - Dies entspricht der ersten Näherung durch die Jacobi-Matrix.
- Zweite Ableitung $F''(0)$ (mit Satz von Schwarz):
  - $$\begin{aligned} F''(0) &= \left[\frac{d}{dt} f_x(x_0 + th, y_0 + tk)h + \frac{d}{dt} f_y(x_0 + th, y_0 + tk)k\right]_{t=0} \\ &= f_{xx}(x_0, y_0)h^2 + 2f_{xy}(x_0, y_0)hk + f_{yy}(x_0, y_0)k^2 \end{aligned}$$
- Dritte Ableitung $F'''(0)$:
  - $$F'''(0) = f_{xxx}(x_0, y_0)h^3 + 3f_{xxy}(x_0, y_0)h^2k + 3f_{xyy}(x_0, y_0)hk^2 + f_{yyy}(x_0, y_0)k^3$$

Differentialoperatoren

- **Definition:** Partielle Ableitungen nach x und y können als lineare Funktionen im Vektorraum der (unendlich oft) differenzierbaren Funktionen verstanden werden.
- **Bezeichnung:** $D_x = \frac{\partial}{\partial x}$ und $D_y = \frac{\partial}{\partial y}$.
- **Konventionen beim Rechnen mit Operatoren:**
  - Hintereinanderausführung von Operatoren wird als Produkt oder Potenz geschrieben (wenn derselbe Operator mehrfach angewendet wird).
  - Konstante Faktoren in Produkten sind als Vielfache des identischen Operators zu verstehen.
- **Darstellung der Ableitungen mit Differentialoperatoren:**
  - $F''(0) = h^2 D_x^2 f(x_0, y_0) + 2hk D_x D_y f(x_0, y_0) + k^2 D_y^2 f(x_0, y_0) = (hD_x + kD_y)^2 f(x_0, y_0)$
  - Analog: $F'''(0) = (hD_x + kD_y)^3 f(x_0, y_0)$.
- **Muster:** Das Muster dieser Ableitungen ist nun leicht erkennbar. Der Satz von Taylor kann mittels vollständiger Induktion gezeigt werden.



Satz von Taylor für reellwertige Funktionen in 2 Variablen

--> Satz von Taylor

Hier genügt uns allerdings das hier, da wir nicht über Grad 2 hinausgehen werden:

(Satz) 6.30 a)

Mathematik für Informatik, p.258, Pasted image 20250520092608.png

Falls man beispielsweise quadratische Approximationen einer Funktion sucht, so muss man die Ableitungen bis zur Ordnung zwei bestimmen, um das Taylorpolynom zweiter Ordnung festzulegen. Dieses besitzt auch die Darstellung

$$f\left(x_0, y_0\right)+(h, k) \operatorname{grad} f\left(x_0, y_0\right)+\frac{1}{2!}(h, k)\begin{pmatrix} f_{xx}\left(x_0, y_0\right) & f_{xy}\left(x_0, y_0\right) \\ f_{yx}\left(x_0, y_0\right) & f_{yy}\left(x_0, y_0\right) \end{pmatrix}\begin{pmatrix} h \\ k \end{pmatrix}$$

Die hier auftretende Matrix der partiellen Ableitungen zweiter Ordnung heißt HesseMatrix. Allgemein gilt für zweimal stetig differenzierbare Funktionen in n Variablen (mit den Abkürzungen $\mathbf{x}=\left(x_1, \ldots, x_n\right)$ und $\mathbf{h}=\left(h_1, \ldots, h_n\right)$)

$$f(\mathbf{x}+\mathbf{h})=f(\mathbf{x})+\mathbf{h} \operatorname{grad} f(\mathbf{x})+\frac{1}{2!} \mathbf{h} \cdot H_f(\mathbf{x}) \cdot \mathbf{h}^T+R(\mathbf{x})$$

wobei H_f die durch

$$H_f=\begin{pmatrix} f_{x_1 x_1}\left(x_1, \ldots, x_n\right) & \cdots & f_{x_1 x_n}\left(x_1, \ldots, x_n\right) \\ \vdots & \vdots & \vdots \\ f_{x_n x_1}\left(x_1, \ldots, x_n\right) & \cdots & f_{x_n x_n}\left(x_1, \ldots, x_n\right) \end{pmatrix}$$

definierte Hesse-Matrix von f und $R\left(x_1, \ldots, x_n\right)$ das Restglied aus (6.9) bezeichnet.