

MEHRDIMENSIONALE FUNKTIONEN

1 Der Euklidische Raum

Wir bezeichnen die Menge aller n -Tupel von reellen Zahlen mit \mathbb{R}^n . Ein Element $x \in \mathbb{R}^n$ ist damit von der Form (x_1, \dots, x_n) .

Der \mathbb{R}^n ist ein \mathbb{R} -Vektorraum (siehe Lineare Algebra) über den reellen Zahlen mit Skalarmultiplikation und komponentenweiser Addition

$$a \cdot x = (a \cdot x_1, \dots, a \cdot x_n) \quad x + y = (x_1 + y_1, \dots, x_n + y_n).$$

Wir definieren die euklidische Struktur wie folgt:

Definition 1.1 (DIE EUKLIDISCHE STRUKTUR DES \mathbb{R}^n)

Seien $x, y \in \mathbb{R}^n$ gegeben, dann ist das euklidische **Skalarprodukt** gegeben als

$$\langle x, y \rangle := \sum_{i=1}^n x_i \cdot y_i,$$

die **Euklidische Norm** von x als

$$\|x\| := \sqrt{\langle x, x \rangle} = \sqrt{\sum_{i=1}^n x_i^2}$$

und die **Euklidische Distanz** als

$$\|x - y\| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}.$$

In dieser Vorlesung ist eigentlich nur der zweidimensionale Fall \mathbb{R}^2 relevant.

2 Grundlagen mehrdimensionaler Funktionen

In diesem Kapitel werden Vektoren immer als *Spaltenvektoren* $\begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2$ angesehen – auch wenn diese manchmal zur Übersicht als (x, y) oder auch $(x, y)^T$ notiert werden. Gemeint sind immer die Spaltenvektoren.

2.1 Definitionsbereich

Der **Definitionsbereich** D_f einer Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ ist die Menge aller Punkte in \mathbb{R}^2 , für die die Funktion definiert und wohldefiniert ist. Bei Funktionen müssen wir auf folgende typische Einschränkungen achten:

- **Divisionen:** Nenner darf nicht null sein
- **Wurzeln:** Argument muss nicht-negativ sein (für reelle Wurzeln)
- **Logarithmen:** Argument muss positiv sein

Beispiel 2.1

Betrachten wir die Funktion $f(x, y) = \frac{1}{x^2 + y^2 - 1}$. Die Funktion ist nicht definiert, wenn der Nenner verschwindet, also wenn $x^2 + y^2 - 1 = 0$ bzw. $x^2 + y^2 = 1$. Der Definitionsbereich ist daher

$$\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \neq 1\} = \mathbb{R}^2 \setminus \{(x, y) : x^2 + y^2 = 1\}.$$

Die Funktion ist also überall in der Ebene definiert, ausser auf dem Einheitskreis. Der Einheitskreis ist die Menge aller Punkte, die die Kreisgleichung mit Radius 1 erfüllen, also $\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$

Beispiel 2.2

Für die Funktion $g(x, y) = \ln(x) + \sqrt{y}$ benötigen wir $x > 0$ (wegen des Logarithmus) und $y \geq 0$ (wegen der Wurzel). Damit ist

$$\{(x, y) \in \mathbb{R}^2 : x > 0 \text{ und } y \geq 0\}.$$

Dies ist der erste Quadrant (inklusive der positiven y -Achse).

2.2 Darstellungen mehrdimensionaler Funktionen

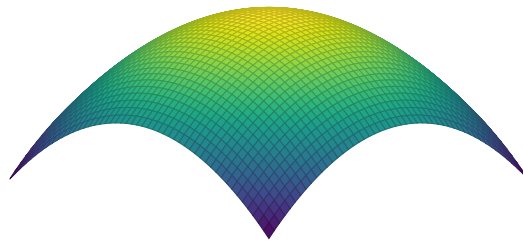
Mehrdimensionale Funktionen lassen sich auf verschiedene Weisen visualisieren und darstellen. Wir betrachten hier die wichtigsten Darstellungsformen für Funktionen $f : \mathbb{R}^2 \rightarrow \mathbb{R}$. Das sind sogenannte zweidimensionale *Skalarfelder*. Sie ordnen jedem Punkt in der Ebene \mathbb{R}^2 eine Zahl (Skalar) zu. Das einfachste Beispiel sind zum Beispiel topografische Landkarten, die jedem Punkt eine Höhe über dem Meeresspiegel zuordnen.

2.2.1 Graph einer zweidimensionalen Funktion

Der **Graph** einer Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ ist die Menge aller Punkte $(x, y, f(x, y))$ im dreidimensionalen Raum:

$$\text{Graph}(f) = \{(x, y, z) \in \mathbb{R}^3 : z = f(x, y), (x, y) \in D_f\}.$$

Der Graph ist somit eine Fläche im \mathbb{R}^3 , die über dem Definitionsbereich liegt. Jeder Punkt auf dieser Fläche hat die Höhe $z = f(x, y)$.



Der Graph in unserem Landkartenbeispiel wäre einfach die 3D-Darstellung der Region.

Beispiel 2.3

Die Funktion $f(x, y) = x^2 + y^2$ hat als Graphen ein Paraboloid (eine nach oben geöffnete Schale). Der tiefste Punkt liegt bei $(0, 0, 0)$ und die Fläche steigt in alle Richtungen quadratisch an.

Beispiel 2.4

Die Funktion $h(x, y) = \sin(x) \cos(y)$ hat als Graphen eine wellenförmige Fläche, die sich periodisch in beide Richtungen wiederholt und zwischen -1 und 1 oszilliert.

2.2.2 Niveaulinien

Niveaulinien (auch Höhenlinien oder Konturlinien genannt) sind eine zweidimensionale Darstellungsweise für Funktionen $f : \mathbb{R}^2 \rightarrow \mathbb{R}$. Sie sind analog zu Höhenlinien auf Landkarten und zeigen Kurven konstanter Funktionswerte respektive konstanter Höhe.

Definition 2.5 (NIVEAULINIE EINER ZWEIDIMENSIONALEN FUNKTION)

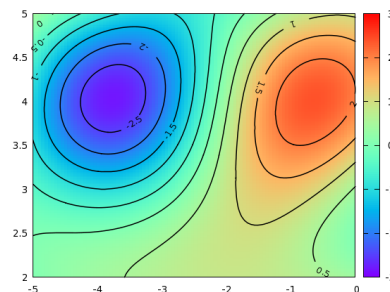
Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ eine Funktion und $c \in \mathbb{R}$ ein konstanter Wert. Die **Niveaulinie** (oder **Niveau-**

menge) zum Niveau c ist definiert als

$$\{(x, y) \in D_f : f(x, y) = c\}.$$

Bemerkung 2.6 (NIVEAULINIEN GRAPHISCH)

Niveaulinien ermöglichen es uns, den dreidimensionalen Graphen in die Ebene zu projizieren, indem wir zeigen, wo die Funktion welche Werte annimmt. Liegen Niveaulinien dicht beieinander, ändert sich die Funktion dort stark (steiler Anstieg); liegen sie weit auseinander, ändert sich die Funktion dort langsam (flacher Bereich). Man kann zusätzlich auch noch mit Farben arbeiten und so den Funktionswert visualisieren.



Beispiel 2.7

Für die Funktion $f(x, y) = x^2 + y^2$ sind die Niveaulinien konzentrische Kreise um den Ursprung:

$$\{(x, y) : x^2 + y^2 = c\}.$$

Für $c > 0$ ist dies ein Kreis mit Radius \sqrt{c} um $(0, 0)$. Für $c = 0$ ist es nur der Punkt $(0, 0)$. Für $c < 0$ ist die Niveaulinie leer, da $x^2 + y^2 \geq 0$ immer gilt.

Beispiel 2.8

Für die Funktion $g(x, y) = x - y$ sind die Niveaulinien Geraden:

$$\{(x, y) : x - y = c\} = \{(x, y) : y = x - c\}.$$

Dies sind parallele Geraden mit Steigung 1, die für verschiedene c -Werte vertikal verschoben sind.

Beispiel 2.9

Für die Funktion $h(x, y) = xy$ (eine Sattelfunktion) sind die Niveaulinien Hyperbeln:

$$\{(x, y) : xy = c\}.$$

Für $c > 0$ liegen die Hyperbeln im ersten und dritten Quadranten, für $c < 0$ im zweiten und vierten. Für $c = 0$ erhalten wir die beiden Koordinatenachsen $x = 0$ oder $y = 0$.

3 Grenzwerte und Stetigkeit

3.1 Grenzwerte

Definition 3.1 (GRENZWERT)

Der Grenzwert einer Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ an einer Stelle (x_0, y_0) ist definiert als

$$\lim_{(x, y) \rightarrow (x_0, y_0)} f(x, y) = L,$$

falls für jedes $\varepsilon > 0$ ein $\delta > 0$ existiert, sodass $\|f(x, y) - L\| < \varepsilon$ für alle (x, y) mit $0 < \|(x - x_0, y - y_0)\| < \delta$. Hier meint $\|\cdot\|$ die Euklidische Distanz im \mathbb{R}^2 .

Bemerkung 3.2 (VERGLEICH ZUM EINDIMENSIONALEN GRENZWERT)

Die „Neuheit“ hierbei ist jetzt, dass wir uns an (x_0, y_0) in jeder beliebigen Raumrichtung nähern können und nichtmehr nur von Links oder Rechts. Alle diese Werte müssen übereinstimmen und endlich sein, damit der **Grenzwert** L der Funktion bei (x_0, y_0) existiert.

3.2 Stetigkeit**Definition 3.3** (STETIGKEIT)

Eine Funktion $f : U \rightarrow \mathbb{R}$ heisst stetig an der Stelle $(x_0, y_0) \in U$, falls

$$\lim_{(x,y) \rightarrow (x_0,y_0)} f(x) = f(x_0, y_0).$$

Die Funktion heisst stetig auf U , falls sie an jedem Punkt von U stetig ist.

Bemerkung 3.4 (ZUSAMMENHANG ZU EINER DIMENSION)

Das ist nach wie vor dieselbe Definition wie die Folgenstetigkeit in einer Dimension. Hier ist jetzt aber der Grenzwertbegriff ein komplexerer, der wie oben beschrieben alle möglichen Raumrichtungen in Betracht zieht.

4 Differentialrechnung**4.1 Partielle Ableitungen**

Hat unser Definitionsbereich mehrere Dimensionen, so können wir in verschiedene Richtungen differenzieren. Der einfachste Fall sind die partiellen Ableitungen, also entlang der einzelnen Koordinatenachsen.

Definition 4.1 (PARTIELLE ABLEITUNG)

Sei $U \subseteq \mathbb{R}^2$ eine offene Teilmenge und $f : U \rightarrow \mathbb{R}$ eine Funktion.

Die **partielle Ableitung** in der j -ten Koordinate (oder der Variable x_j) bei $x_0 \in U$ ist definiert als

$$\partial_j f(x_0) = \lim_{h \rightarrow 0} \frac{f(x_0 + h e_j) - f(x_0)}{h},$$

wobei e_j der j -te Einheitsvektor ist. Man schreibt auch $\frac{\partial f}{\partial x_j}(x_0)$.

Existiert die partielle Ableitung in der j -ten Koordinate an jedem Punkt in U , so erhält man eine Funktion $\partial_j f : U \rightarrow \mathbb{R}$, die man j -te partielle Ableitung von f nennt.

Bemerkung 4.2

Explizit heisst das für eine Funktion $f(x, y)$

$$\partial_x f(x, y) = \lim_{h \rightarrow 0} \frac{f(x + h, y) - f(x, y)}{h} \quad \partial_y f(x, y) = \lim_{h \rightarrow 0} \frac{f(x, y + h) - f(x, y)}{h}$$

Bemerkung 4.3

Eine andere Notation für die partielle Ableitung ist

$$\partial_x = \frac{\partial}{\partial x} \quad \partial_y = \frac{\partial}{\partial y}$$

Beispiel 4.4

Haben wir eine Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ als

$$f(x, y) = y^2 + y \cos(x)$$

gegeben, so können wir die *partiellen Ableitungen* nach x und y einzeln berechnen, indem wir die andere Variable „*einfrisieren*“. Dadurch wird es einfach eine bekannte eindimensionale Ableitung.

$$\partial_x f(x, y, z) = \partial_x (y^2 + y \cos(x)) = -y \sin(x)$$

$$\partial_y f(x, y, z) = \partial_y (y^2 + y \cos(x)) = 2y + \cos(x)$$

Aufgabe Berechne die partiellen Ableitungen der Funktionen

$$f_1(x, y) = \sin(xy) + \cos(xy) \quad f_2(x, y) = xy + \ln(x^2 + 1) - \sin(y)$$

Bemerkung 4.5

Eine partielle Ableitung ist die tangentielle Steigung an den Graphen in der entsprechenden Koordinatenrichtung, also parallel zur entsprechenden Achse. Man kann jetzt die partiellen Ableitungen als „Basis“ (siehe Lineare Algebra) nutzen und aus ihnen jede beliebige *Richtungsableitung* konstruieren. Das sehen wir später im Kontext des Begriffs des *Gradienten*.

4.2 Das totale Differential

Wir erinnern uns an bereits bekannte Konzepte aus der Differentialrechnung in einer Dimension. Diese Interpretationen lassen sich problemlos verallgemeinern auf mehrere Dimensionen.

Bemerkung 4.6 (INTERPRETATIONEN DES DIFFERENTIALS)

(1) **Differential als Tangentensteigung**

Bereits in der Schule hat man gesehen, dass das Differential als die Steigung der Tangente an einem Punkt interpretiert werden kann.

(2) **Differential als lineare Änderungsrate**

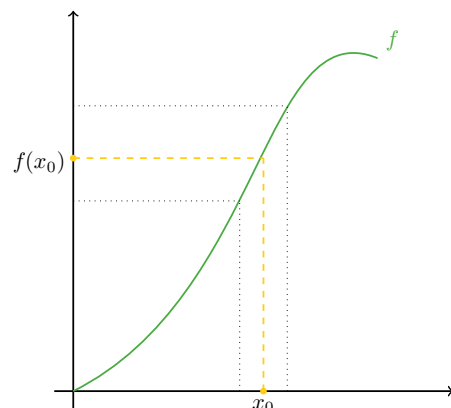
In direktem Zusammenhang steht die Interpretation als lokale Änderungsrate der Funktion. Das Differential gibt an, wie gross die Steigung (resp. Änderung) der Funktion an einer Stelle ist.

(3) **Differential als Zoom-In**

Stellt man sich vor, man könnte an den Graphen der Funktion hineinzoomen, so wird die Funktion in beliebiger Genauigkeit irgendwann linear (ähnlich wie die Erde von nahem flach aussieht, obwohl sie eigentlich eine Kugel ist). Die Steigung dieser linearen Funktion ist das Differential. Diese Interpretation ist essentiell im Mehrdimensionalen Fall!

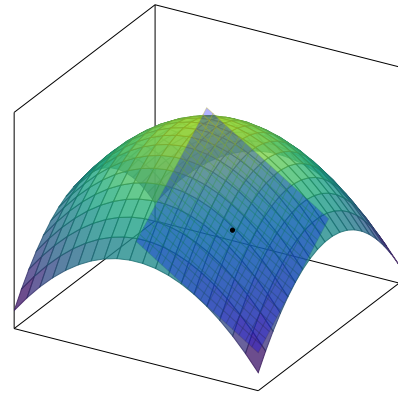
(4) **Differential als Skalierungsfaktor**

Betrachten wir ein Intervall im Definitionsbereich und bilden es durch die Funktion f ab, so sagt uns das Differential, wie sehr sich die Grösse des Intervalls nachher verändert hat. Ist die Funktion sehr steil, so skaliert sie das Intervall sehr gross (hohe Ableitung); ist die Funktion an der Stelle jedoch sehr flach, so skaliert sie das Intervall sehr klein (Ableitung nahe Null). Diese Interpretation wird besonders wichtig im Teil über mehrdimensionale Integration!



Bemerkung 4.7 (MEHRDIMENSIONALES DIFFERENTIAL)

Anders als in einer Dimension ist das Differential an einer Stelle im Definitionsbereich in mehreren Dimensionen nicht mehr nur eine einfache Zahl (vgl. Steigung an der Stelle, Steigung der Tangenten, Skalierungsfaktor, ...), **sondern eine lineare Abbildung**. Diese lineare Abbildung kann man lokal auch zeichnen. In zwei Dimensionen entspricht das keiner Tangenten mehr, sondern einer Ebene! Diese nennt man oft *Tangentialebene*.

**Bemerkung 4.8** (TANGENTIALEBENE)

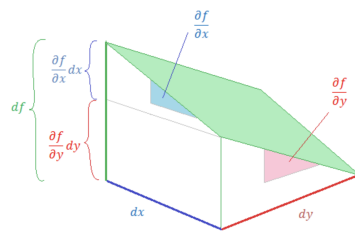
Die Funktionsgleichung der Tangentialebene an einer Stelle (x_0, y_0) ist gegeben durch

$$z = f(x_0, y_0) + \partial_x f(x_0, y_0)(x - x_0) + \partial_y f(x_0, y_0)(y - y_0).$$

Das ist wie wir später sehen werden einfach die lineare Approximation erster Ordnung mithilfe der totalen Ableitung oder auch die *Taylorapproximation erster Ordnung*.

Bemerkung 4.9 (HERLEITUNG/ INTUITION TOTALES DIFFERENTIAL)

Visualisieren wir ein zweidimensionales Skalarfeld durch seinen Graphen, so erhalten wir eine gekrümmte Fläche im dreidimensionalen Raum (zwei Input Dimensionen x, y und eine Output Dimension $f(x, y)$). Betrachten wir einen Punkt und zoomen ganz nah ran, so sieht diese gekrümmte Fläche genügend nah aus wie eine Ebene. Dasselbe Phänomen kennen wir aus dem Alltag mit unserer Erde. Sie ist global gesehen gekrümmt, sieht hinreichend nah aus unserer Perspektive aber flach aus wie eine Ebene.¹ Diese Ebene aus der nahen Perspektive nennen wir Tangentialebene. Fixieren wir einen Punkt (x, y) und schauen uns an, welche Steigung diese Ebene an unseren Graphen hat, so erhalten wir das (totale) Differential unserer zweidimensionalen Funktion. Hierbei ist jetzt zu beachten, dass diese Ebene in x und y Richtung verschiedene Steigungen haben kann und wir beide betrachten müssen!



Die Gesamtänderung unseres Funktionswertes zwischen den Punkten (x, y) und $(x + \Delta x, y + \Delta y)$ nennen wir Δf . Diese können wir erhalten, indem wir uns zuerst die Änderung von (x, y) zu $(x + \Delta x, y)$ mit konstantem y und dann von $(x + \Delta x, y)$ zu $(x + \Delta x, y + \Delta y)$ mit konstantem x betrachten. Denn jetzt gilt

$$\begin{aligned} \Delta f &= f(x + \Delta x, y + \Delta y) - f(x, y) \\ &= f(x + \Delta x, y + \Delta y) - \underbrace{f(x + \Delta x, y) + f(x + \Delta x, y) - f(x, y)}_{=0} \\ &= \frac{\partial f(x, y)}{\partial y} \cdot \Delta y + \frac{\partial f(x, y)}{\partial x} \cdot \Delta x. \end{aligned}$$

Denn die Steigung der Tangentialebene entlang x und y entspricht einfach den partiellen Ableitungen in dieser Richtung. Die Form, die wir erhalten haben entspricht genau dem sogenannten totalen

¹Objekte dieser Art nennen Mathematiker *differenzierbare Mannigfaltigkeiten*.

Differential eines zweidimensionalen Skalarfeldes

$$\Delta f = \frac{\partial f(x, y)}{\partial x} \cdot \Delta x + \frac{\partial f(x, y)}{\partial y} \cdot \Delta y.$$

Denn macht man nun einen Grenzübergang, indem man alle $\Delta x, \Delta y \rightarrow 0$ schickt, so werden diese *infinitesimal* klein und man schreibt dx und dy . Die korrespondierende Änderung in f notiert man mit df .

Definition 4.10 (TOTALES DIFFERENTIAL)

Sei $U \subseteq \mathbb{R}^2$ offen und $f : U \rightarrow \mathbb{R}$ eine Funktion. Das totale Differential der Funktion f an der Stelle $(x_0, y_0) \in U$ ist

$$df = \partial_x f(x_0, y_0) dx + \partial_y f(x_0, y_0) dy$$

Haben wir mehr als nur zwei Koordinaten, setzt sich das additive Prinzip hier einfach fort.

Beispiel 4.11 (ÄNDERUNG DURCH TOTALES DIFFERENTIAL BERECHNEN)

Wir suchen das totale Differential von der Funktion $f(x, y) = x^2 y + y^2$. Dafür berechnen wir einfach die partiellen Ableitungen

$$\partial_x f = 2xy \quad \partial_y f = x^2 + 2y.$$

Damit folgt für das totale Differential:

$$df = 2xy dx + (x^2 + 2y) dy.$$

Damit können wir jetzt abschätzen, wie sich der Funktionswert ungefähr ändern wird, wenn wir einen kleinen Schritt (dx, dy) von einer Stelle (x_0, y_0) bewegen. Wir fragen uns jetzt zum Beispiel, wie sehr sich der Funktionswert verändert, wenn wir von der Stelle $(1, 2)$ um $(0.1, 0.05)$ uns bewegen. Wir berechnen somit einfach:

$$df = 2 \cdot 1 \cdot 2 \cdot 0.1 + (1^2 + 2 \cdot 2) \cdot 0.05 = 0.65.$$

Also ändert sich der Funktionswert um ungefähr 0.65, wenn wir von $(1, 2)$ uns um $(0.1, 0.05)$ bewegen, also zu der Stelle $(1.1, 2.05)$ gehen. Das ist eine Näherung erster Ordnung!

Wir können auch die exakte Änderung ausrechnen. Diese wäre einfach

$$f(1.1, 2.05) - f(1, 2) = 0.683.$$

Wir sehen also, dass für kleine Schritte/Änderungen (dx, dy) ist die lokale lineare Näherung durch das totale Differential sehr genau! Das ist im Grunde die Taylorapproximation erster Ordnung, wie wir später sehen werden.

Aufgabe [2024 S]

Gegeben ist die folgende Funktion $f(x, y)$ in zwei Variablen $f(x, y) = \sin(xy)$. Welche ungefähre Veränderung der Funktionswerte erwarten Sie, wenn Sie die Argumente von $(x_0, y_0) = (1, 0)$ zu $(x, y) = (1.1, -0.1)$ verändern?

4.3 Tangentialebene und Linearisierung

Für eine differenzierbare Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ können wir an einem Punkt (x_0, y_0) die Tangentialebene definieren. Diese ist die beste lineare Approximation des Graphen von f an dieser Stelle.

Die Gleichung der Tangentialebene lautet:

$$z = f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0).$$

Sie ist die Visualisierung der zweidimensionalen Ableitung!

Beispiel 4.12 (TANGENTIALEBENE AUFSTELLEN)

Sei $f(x, y) = x^2 + y^2$. Wir suchen die Tangentialebene an der Stelle $(1, 1)$.

Die partiellen Ableitungen sind:

$$\frac{\partial f}{\partial x} = 2x, \quad \frac{\partial f}{\partial y} = 2y.$$

An der Stelle $(1, 1)$ gilt: $f(1, 1) = 2$, $\frac{\partial f}{\partial x}(1, 1) = 2$, $\frac{\partial f}{\partial y}(1, 1) = 2$.

Die Tangentialebene ist:

$$z = 2 + 2(x - 1) + 2(y - 1) = 2x + 2y - 2.$$

4.4 Kettenregel

Die Kettenregel für mehrdimensionale Funktionen verallgemeinert die bekannte eindimensionale Kettenregel. Hier formulieren wir sie für Funktionen der Form $f(x(t), y(t), t)$.

Satz 4.13 (KETTENREGEL FÜR SKALARFELDER)

Haben wir eine Funktion $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ der Form $f(x(t), y(t), t)$ und wollen diese (total) nach t ableiten, so besagt die Kettenregel:

$$\frac{df}{dt} = \frac{\partial f}{\partial x} \frac{dx}{dt} + \frac{\partial f}{\partial y} \frac{dy}{dt} + \frac{\partial f}{\partial t}.$$

Fragt man sich, wie sich f mit t ändert, so betrachtet man die direkte Änderung nach t (partielle Ableitung) als auch noch die indirekten Beiträge, wenn sich f mit x und y ändert, weil sich x und y mit t ändern.

Beispiel 4.14 (TEMPERATUR AUF EINER LANDKARTE)

Nehmen wir an f beschreibt die Temperatur auf einer Landkarte. Die Temperatur ist an jedem Punkt unterschiedlich, im Norden ist es am kältesten und im Süden am wärmsten. Ausserdem ändert sich die Temperatur auch mit der Tageszeit. Am Morgen und am Abend ist es am kühnsten und am Mittag am heissesten. Wenn wir uns jetzt fragen, wie sich die Temperatur verändert, dann kann das über zwei verschiedene Weisen passieren.

1. Man bleibt konstant an einem Ort stehen aber die Tageszeit verändert sich und damit die Temperatur. Sie steigt von Sekunde zu Sekunde bis Mittags und fällt dann wieder.
2. Man bewegt sich zum Beispiel von Norden (kalt) nach Süden (warm) wodurch die Temperatur steigt (abhängig von der Bewegungsgeschwindigkeit. Fahre ich sehr sehr schnell dann ändert sich die Temperatur natürlich schneller.)

Beide diese Beiträge müssen wir aufsummieren! Aus dieser Überlegung folgt direkt die Kettenregel:

$$\frac{df}{dt} = \frac{\partial f}{\partial x} \frac{dx}{dt} + \frac{\partial f}{\partial y} \frac{dy}{dt} + \frac{\partial f}{\partial t}.$$

Dabei besagt der Term $\frac{\partial f}{\partial x}$ wie schnell sich die Temperatur mit Ost West auf der Karte verändert. $\frac{dx}{dt}$ ist meine aktuelle Ost-West Geschwindigkeit. $\frac{\partial f}{\partial y}$ und $\frac{dy}{dt}$ analog für Nord-Süd. Der Term $\frac{\partial f}{\partial t}$ gibt die Temperaturänderung mit der Zeit an.

Bemerkung 4.15 (ZUSAMMENHANG MIT DEM TOTALEN DIFFERENTIAL)

Das entspricht im Grunde einfach dem totalen Differential von $f(x, y, t)$. Denn schreiben wir

auf:

$$df = \partial_x f dx + \partial_y f dy + \partial_t f dt$$

und teilen diese Gleichung durch dt , so erhalten wir direkt die Kettenregel

$$\frac{df}{dt} = \partial_x f \frac{dx}{dt} + \partial_y f \frac{dy}{dt} + \underbrace{\partial_t f}_{=1} \frac{dt}{dt} = \frac{\partial f}{\partial x} \frac{dx}{dt} + \frac{\partial f}{\partial y} \frac{dy}{dt} + \frac{\partial f}{\partial t}.$$

Beispiel 4.16 (ZEITLICHE ÄNDERUNG DES LINEAREN IMPULSES AUF EINER KREISBAHN)

Wir betrachten ein Teilchen der Masse m , das sich auf einer Kreisbahn mit konstantem Radius r bewegt. Die Position des Teilchens zum Zeitpunkt t wird beschrieben durch:

$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = r \begin{bmatrix} \cos(\varphi(t)) \\ \sin(\varphi(t)) \end{bmatrix}.$$

Dabei ist $\varphi(t)$ der momentane Drehwinkel, der im Allgemeinen zeitabhängig sein kann.

Der lineare Impuls in x -Richtung ist definiert als $p_x = mv_x$. Die x -Komponente der Geschwindigkeit erhalten wir durch Ableiten der x -Koordinate nach der Zeit:

$$v_x = \frac{dx}{dt} = \frac{d}{dt} [r \cos(\varphi(t))].$$

Mit der eindimensionalen Kettenregel folgt:

$$v_x = -r\dot{\varphi}(t) \sin(\varphi(t)),$$

wobei $\dot{\varphi}(t) = \frac{d\varphi}{dt}$ die Winkelgeschwindigkeit ist. Damit ergibt sich:

$$p_x(t) = mv_x = -mr\dot{\varphi}(t) \sin(\varphi(t)).$$

Wir interessieren uns für die Beschleunigung in x -Richtung, also für die *totale* zeitliche Änderung:

$$\frac{dp_x}{dt}.$$

In dem Ausdruck für p_x kommen zwei zeitabhängige Größen vor:

- der Winkel $\varphi(t)$,
- die Winkelgeschwindigkeit $\dot{\varphi}(t)$.

Da p_x sowohl von φ als auch von $\dot{\varphi}$ abhängt, müssen wir die (zweidimensionale) Kettenregel anwenden:

$$\frac{dp_x}{dt} = \frac{\partial p_x}{\partial \varphi} \cdot \frac{d\varphi}{dt} + \frac{\partial p_x}{\partial \dot{\varphi}} \cdot \frac{d\dot{\varphi}}{dt} + \underbrace{\frac{\partial p_x}{\partial t}}_{=0}.$$

Ein expliziter Zeitanteil $\partial p_x / \partial t$ tritt nicht auf, da p_x nur über φ und $\dot{\varphi}$ von t abhängt.

Aus $p_x = -mr\dot{\varphi} \sin \varphi$ erhalten wir:

$$\frac{\partial p_x}{\partial \varphi} = -mr\dot{\varphi} \cos \varphi, \quad \frac{\partial p_x}{\partial \dot{\varphi}} = -mr \sin \varphi.$$

Zudem gilt:

$$\frac{d\varphi}{dt} = \dot{\varphi}, \quad \frac{d\dot{\varphi}}{dt} = \ddot{\varphi},$$

wobei $\dot{\varphi}$ die Winkelgeschwindigkeit und $\ddot{\varphi}$ die Winkelbeschleunigung ist.

Setzen wir alle Terme ein, folgt:

$$\frac{dp_x}{dt} = (-mr\dot{\varphi} \cos \varphi) \cdot \dot{\varphi} + (-mr \sin \varphi) \cdot \ddot{\varphi}.$$

Damit erhalten wir für die totale zeitliche Änderung des x -Impulses:

$$\frac{dp_x}{dt} = -mr(\dot{\varphi}^2 \cos \varphi + \ddot{\varphi} \sin \varphi).$$

Das ist die totale Änderung des Impulses mit t . Dieser Ausdruck beschreibt physikalisch sowohl den Beitrag der Zentripetalbeschleunigung (Term mit $\dot{\varphi}^2$) als auch den Beitrag der Tangentialbeschleunigung (Term mit $\ddot{\varphi}$).

Bemerkung 4.17

Wir können die Kettenregel auch in der folgenden Form notieren. Seien u und v jeweils von den Variablen x und y abhängig, also $u(x, y)$ und $v(x, y)$. Dann gilt für die partiellen Ableitungen von $f(u(x, y), v(x, y))$ nach x und y

$$\boxed{\frac{\partial f}{\partial x} = \frac{\partial f}{\partial u} \frac{\partial u}{\partial x} + \frac{\partial f}{\partial v} \frac{\partial v}{\partial x} \quad \frac{\partial f}{\partial y} = \frac{\partial f}{\partial u} \frac{\partial u}{\partial y} + \frac{\partial f}{\partial v} \frac{\partial v}{\partial y}}$$

Beachte, dass dies hier jeweils partielle Ableitungen sind, da die Funktion f im Endeffekt von u und v abhängt, also von zwei Variablen.

Beispiel 4.18 (PRÜFUNG S2024)

Wir betrachten die Funktion $f(u, v) = uv^2 - e^u$ mit $u = \sin(x)$ und $v = xy$. Was ist dann die partielle Ableitung nach x ?

Nach der Kettenregel gilt:

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial u} \frac{\partial u}{\partial x} + \frac{\partial f}{\partial v} \frac{\partial v}{\partial x}$$

Also berechnen wir

$$\frac{\partial f}{\partial u} = v^2 - e^u \quad \frac{\partial u}{\partial x} = \cos(x)$$

$$\frac{\partial f}{\partial v} = 2uv \quad \frac{\partial v}{\partial x} = y.$$

Daraus folgt mit der Kettenregel:

$$\frac{\partial f}{\partial x} = (v^2 - e^u) \cos(x) + 2uvy = (x^2y^2 - e^{\sin(x)}) \cos(x) + 2xy^2 \sin(x).$$

Am Ende haben wir alles in Abhängigkeit von x und y geschrieben, da das unsere letztendlichen Variablen sind. Denn durch $f(u(x, y), v(x, y))$ ist f letztendlich nur durch x und y bestimmt.

4.5 Gradient und Richtungsableitungen

Für Skalarfelder (also Funktionen $f : \mathbb{R}^2 \rightarrow \mathbb{R}$) hat das Differential eine besondere Darstellung durch den Gradienten.

Definition 4.19 (GRADIENT)

Sei $U \subseteq \mathbb{R}^2$ eine offene Menge und $f : U \rightarrow \mathbb{R}$ eine C^1 -Funktion (also stetig differenzierbar). Wir definieren den **Gradienten** von f bei $(x, y) \in U$, bezeichnet als $\nabla f(x, y)$, als den Vektor

$$\nabla f = \begin{bmatrix} \partial_x f(x, y) \\ \partial_y f(x, y) \end{bmatrix}.$$

Somit ist $\nabla f : U \rightarrow \mathbb{R}^2$ eine stetige Abbildung, die jedem Punkt in U einen Vektor im \mathbb{R}^2 zuordnet (ein sogenanntes *Vektorfeld*). Die anschauliche Interpretation dieses Vektorfeldes sehen wir in den Bemerkungen (4.22 und 4.23).

Um den Gradienten als Vektor interpretieren zu können, führen wir die Definition der sogenannten Richtungsableitung ein. Vorher haben wir nur parallel zu den Koordinatenachsen x und y abgeleitet. Die sogenannten *partiellen Ableitungen*. Wir können uns aber an einer Stelle (x_0, y_0) beliebig Fragen, wie die Funktion tangential in eine Richtung v steigt. Diese Frage beantwortet die Richtungsableitung.

Definition 4.20 (RICHTUNGSABLEITUNG)

Sei $v = v_x e_x + v_y e_y = \begin{bmatrix} v_x \\ v_y \end{bmatrix}$ ein normierter Vektor, also $\|v\| = 1$ und $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ differenzierbar. Dann ist die Richtungsableitung von f an der Stelle (x_0, y_0) in Richtung v gegeben durch

$$\partial_v f(x_0, y_0) := \lim_{h \rightarrow 0} \frac{f(x_0 + hv_x, y_0 + hv_y) - f(x_0, y_0)}{h}$$

Man schreibt für die Richtungsableitung auch $D_v f(x_0, y_0)$ oder $\frac{\partial f}{\partial v}(x_0, y_0)$.

Wir können nun mithilfe der Kettenregel diese Ableitung auch umschreiben, denn wir können die Richtungsableitung auffassen, als die Ableitung nach h von der Funktion

$$f(x_0 + hv_x, y_0 + hv_y).$$

Nach der Kettenregel folgt dann

$$\frac{\partial f(x(h), y(h))}{\partial h} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial h} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial h},$$

mit $x(h) = x_0 + hv_x$ und $y(h) = y_0 + hv_y$. Damit gilt für die partiellen Ableitungen $\partial_h x(h) = v_x$ und $\partial_h y(h) = v_y$. Also folgt

$$\partial_v f(x_0, y_0) = \frac{\partial f(x(h), y(h))}{\partial h} = \frac{\partial f}{\partial x} v_x + \frac{\partial f}{\partial y} v_y = \left\langle \begin{bmatrix} \partial_x f(x_0, y_0) \\ \partial_y f(x_0, y_0) \end{bmatrix}, \begin{bmatrix} v_x \\ v_y \end{bmatrix} \right\rangle = \langle \nabla f(x_0, y_0), v \rangle.$$

Somit müssen wir, um die Richtungsableitung zu berechnen einfach das Skalarprodukt von Gradient ∇f und Richtungsvektor v berechnen. Wichtig ist, dass v normiert ist! Hat v nicht Länge Eins, so normieren wir ihn einfach, indem wir mit $\frac{1}{\|v\|}$ multiplizieren. Es folgt damit eine allgemeine Formel (auch für nicht normierte Vektoren)

$$\partial_v f(x_0, y_0) = \frac{1}{\|v\|} \langle \nabla f(x_0, y_0), v \rangle$$

Beispiel 4.21 (RICHTUNGSABLEITUNG BERECHNEN)

Wir suchen die Richtungsableitung der Funktion

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

im Punkt (π, π) in Richtung $\vec{v} = (-1, -3)$.

Zunächst berechnen wir den Gradienten von $f(x, y)$:

$$\nabla f(x, y) = \begin{bmatrix} -\sin(x) \sin(y) \\ \cos(x) \cos(y) \end{bmatrix}.$$

Einsetzen des Punktes (π, π) :

$$\nabla f(\pi, \pi) \begin{bmatrix} -\sin(\pi) \sin(\pi) \\ \cos(\pi) \cos(\pi) \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Der gegebene Richtungsvektor ist $\vec{v} = (-1, -3)$. Dessen Länge beträgt

$$\|\vec{v}\| = \sqrt{(-1)^2 + (-3)^2} = \sqrt{1 + 9} = \sqrt{10}.$$

Damit ist die Richtungsableitung:

$$\partial_{\vec{v}}f(\pi, \pi) = \frac{1}{\sqrt{10}} \left\langle \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ -3 \end{bmatrix} \right\rangle = \frac{0 \cdot (-1) + 1 \cdot (-3)}{\sqrt{10}} = -\frac{3}{\sqrt{10}}.$$

Aufgabe [2024 S] Gegeben ist die folgende Funktion in zwei Variablen

$$f(x, y) = \ln(x^2) + xy^2 - 3y \quad (x \neq 0),$$

welche wir insbesondere an der Stelle $(1, 1)$ betrachten.

Ausserdem ist die Richtung $\vec{v} = \begin{bmatrix} -3 \\ -4 \end{bmatrix}$ gegeben.

Wie gross ist die momentane Veränderung in Richtung dieses Vektors \vec{v} an der betrachteten Stelle?

Bemerkung 4.22 (GEOMETRISCHE INTUITION VOM GRADIENTEN)

Sei $U \subseteq \mathbb{R}^2$ offen und f eine Funktion $f : U \rightarrow \mathbb{R}$. Der Gradient ∇f ist ein Vektor, der mit der Ableitung der Funktion f zusammenhängt. Aber wie hängt er geometrisch mit der Steigung von f zusammen? In welche Richtung zeigt der Vektor? Was ist seine Länge?

Wir wissen, dass $\langle \nabla f, v \rangle$ den Wert des Anstiegs von f in Richtung v angibt. Wollen wir diesen Anstieg maximieren, so müssen wir einen Vektor parallel zu ∇f wählen, denn

$$\max_{\|v\|=1} \langle \nabla f, v \rangle = \max_{\|v\|=1} \|\nabla f\| \cdot \|v\| \cdot \cos(\theta) \iff \theta = 0.$$

Damit zeigt ∇f in die Richtung des steilsten Anstiegs, also die Richtung der grössten Richtungsableitung. Seine Länge gibt an, genau wie steil dieser Steilste Anstieg ist.

Bemerkung 4.23 (GRADIENT UND NIVEAULINIEN)

Was passiert, wenn wir uns orthogonal zum Gradienten bewegen?

Dann halten wir uns auf einer sogenannten Niveaumenge von f auf, wir ändern also unseren Funktionswert nicht. Denn es gilt, dass weil v orthogonal zu ∇f ist, die Richtungsableitung verschwindet:

$$\partial_v f = \langle \nabla f, v \rangle = 0.$$

Damit ändert sich entlang der Richtung v der Funktionswert nicht und wir bleiben damit auf einer Niveaumenge von f . Der Gradient von f steht also immer orthogonal auf der Niveaumenge von f .

4.6 Höhere Ableitungen und Hesse-Matrix

Der Satz von Schwarz besagt, dass für genügend reguläre Funktionen (d.h. C^2) die Reihenfolge der Differentiation keine Rolle spielt.

Satz 4.24 (SATZ VON SCHWARZ)

Sei $U \subseteq \mathbb{R}^2$ und $f : U \rightarrow \mathbb{R}$ eine C^2 -Funktion, dann gilt

$$\partial_x \partial_y f = \partial_y \partial_x f.$$

Das heisst für zweimal stetig differenzierbare Funktionen ist die Reihenfolge der partiellen Ableitungen egal!

Bemerkung 4.25

Zentral für diesen Satz ist die Stetigkeit der Ableitungen! Ist die Funktion zweimal differenzierbar, aber die zweite Ableitung nicht stetig, so gilt der Satz von Schwarz im Allgemeinen nicht. Dies tritt in der Praxis aber sehr selten auf.

Definition 4.26 (HESSE-MATRIX)

Sei $U \subseteq \mathbb{R}^2$ eine offene Menge und $f : U \rightarrow \mathbb{R}$ eine C^2 -Funktion. Wir definieren die Hesse Matrix von $f \in C^2(U)$ an der Stelle $x \in U$ als die 2×2 -Matrix

$$(Hf_x)_{ij} = \frac{\partial^2 f(x)}{\partial x_i \partial x_j}$$

für $i, j \in \{1, \dots, n\}$. Der Satz von Schwarz (siehe 4.24) impliziert, dass $Hf(x)$ eine symmetrische Matrix ist. Eine alternative Standardnotation für die Hesse Matrix ist $D^2 f(x)$.

Beispiel 4.27 (HESSE-MATRIX BERECHNEN)

Für eine Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ ist die Hesse-Matrix eine 2×2 -Matrix:

$$Hf = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} = \begin{bmatrix} \partial_{xx} f & \partial_{xy} f \\ \partial_{yx} f & \partial_{yy} f \end{bmatrix}.$$

Nach dem Satz von Schwarz gilt $\partial_{xy} f = \partial_{yx} f$, also ist die Matrix symmetrisch.

Beispiel 4.28 (HESSE-MATRIX BERECHNEN)

Sei $f(x, y) = x^3 + xy^2 - y^3$. Wir berechnen die Hesse-Matrix:

Erste Ableitungen:

$$\partial_x f = 3x^2 + y^2, \quad \partial_y f = 2xy - 3y^2.$$

Zweite Ableitungen:

$$\partial_{xx} f = 6x, \quad \partial_{xy} f = 2y, \quad \partial_{yy} f = 2x - 6y.$$

Die Hesse-Matrix ist:

$$Hf = \begin{bmatrix} 6x & 2y \\ 2y & 2x - 6y \end{bmatrix}.$$

4.7 Taylorentwicklung

Wir wissen, dass das (totale) Differential die beste lineare Approximation der Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ darstellt. Also

$$f(x, y) \approx f(x_0, y_0) + df(x_0, y_0)(x - x_0, y - y_0).$$

Daraus folgt dann, dass

$$P_1(x, y) = f(x_0, y_0) + \partial_x f(x_0, y_0)(x - x_0) + \partial_y f(x_0, y_0)(y - y_0)$$

Bemerkung 4.29 (TAYLORPOLYNOM ZWEITER ORDNUNG EINES SKALARFELDES)

Das Taylor-Polynom zweiter Ordnung für eine Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ an der Stelle (x_0, y_0) , können wir mithilfe des Gradienten ∇f und der Hesse-Matrix Hf hinschreiben:

$$P_2(f(x, y)) = f(x_0, y_0) + \langle \nabla f(x_0, y_0), (x - x_0, y - y_0) \rangle + \frac{1}{2} \langle (x - x_0, y - y_0), Hf_{(x_0, y_0)}(x - x_0, y - y_0) \rangle.$$

Es gibt auch eine allgemeine Formel für die Approximation n -ter Ordnung, diese benötigt jedoch sogenannte Multiindexnotation und ist kompliziert aufzuschreiben.

Beispiel 4.30

Für eine Funktion $f(x, y)$ um den Punkt (x_0, y_0) lautet die Taylorentwicklung zweiter Ordnung explizit ausgeschrieben:

$$\begin{aligned} f(x, y) &\approx f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) \\ &\quad + \frac{1}{2} \frac{\partial^2 f}{\partial x^2}(x_0, y_0)(x - x_0)^2 + \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)(x - x_0)(y - y_0) \\ &\quad + \frac{1}{2} \frac{\partial^2 f}{\partial y^2}(x_0, y_0)(y - y_0)^2. \end{aligned}$$

Beispiel 4.31 (TAYLORPOLYNOM ZWEITER ORDNUNG EXPLIZIT)

Sei die Funktion $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ gegeben durch

$$f(x_1, x_2) = \sin(x_1) + \cos(x_2).$$

Wir möchten das Taylor-Polynom zweiter Ordnung an der Stelle $x_0 = (0, 0)$ berechnen.

Es gilt

$$f(0, 0) = \sin(0) + \cos(0) = 1.$$

Dann berechnen wir den Gradienten ∇f an der Stelle $x_0 = (0, 0)$:

$$\nabla f(0, 0) = \begin{bmatrix} \cos(0) \\ -\sin(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Die Hesse-Matrix an der Stelle $x_0 = (0, 0)$ ist:

$$Hf(0, 0) = \begin{bmatrix} -\sin(0) & 0 \\ 0 & -\cos(0) \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix}.$$

Nun setzen wir alles in das Taylor-Polynom zweiter Ordnung ein:

$$\begin{aligned} P_2(f(x_1, x_2)) &= f(0, 0) + \langle \nabla f(0, 0), (x - x_0) \rangle + \frac{1}{2} (x - x_0)^T Hf(0, 0) (x - x_0) \\ P_2(f(x_1, x_2)) &= 1 + [1 \ 0] \begin{bmatrix} x_1 - 0 \\ x_2 - 0 \end{bmatrix} + \frac{1}{2} [x_1 - 0 \ x_2 - 0] \begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 - 0 \\ x_2 - 0 \end{bmatrix} \end{aligned}$$

Damit ergibt sich das Taylor-Polynom zweiter Ordnung:

$$P_2(f(x_1, x_2)) = 1 + x_1 - \frac{1}{2} x_2^2.$$

Aufgabe [2023 S] Gegeben sei die Funktion

$$g(x, y) = \left(\frac{\cos(x)}{\sin(x)} + 5 \right) e^{y^2}.$$

Berechnen Sie das Taylorpolynom 2. Grades von der Funktion $g(x, y)$ an der Stelle $(\frac{\pi}{2}, 0)$.

Bemerkung 4.32 (ALTERNATIVE METHODE)

Wir hätten auch einfach die Taylorapproximationen von $\sin(x_1)$ und $\cos(x_2)$ um 0 einsetzen können und alle Terme bis zur Ordnung 2 behalten. So hätten wir

$$\sin(x_1) + \cos(x_2) \approx \left(x_1 - \frac{x_1^3}{6} + \dots \right) + \left(1 - \frac{x_2^2}{2} + \dots \right) = 1 + x_1 - \frac{1}{2} x_2^2.$$

Diese alternative Methode schauen wir uns jetzt genauer an:

Rechenmethode 4.33 (MEHRDIMENSIONALE TAYLOR-APPROXIMATION)

Um mehrdimensionale Taylor-Approximationen zu berechnen, haben wir die Möglichkeiten

1. für jede Variable einzeln die eindimensionale Taylor-Approximation einzusetzen
2. mehrere Argumente zu gruppieren und zu substituieren, z.B. $t = x^2 + y^2$, $s = xy, \dots$

Falls sich die Funktion nicht in bekannte eindimensionale Taylor-Approximationen zerlegen lässt, dann kann man auf die bekannten Formeln für eine und zwei Dimensionen zurückgreifen.

Beispiel 4.34 (TAYLORENTWICKLUNG)

Wir suchen das Taylor-Polynom von $f(x, y) = \sin(x) \cos(y) + 1$ bis zur 3. Ordnung. Dafür nutzen wir Methode 4.33 und setzen die Taylor-Polynome für $\sin(x)$ und $\cos(y)$ einzeln ein.

$$\left(x - \frac{x^3}{6}\right) \left(1 - \frac{y^2}{2}\right) + 1 = 1 + x - \frac{x^3}{6} - \frac{xy^2}{2} + \frac{x^3y^2}{12}.$$

Der letzte Term skaliert im Betrag jedoch wie ein Term 5. Ordnung. Das Taylor-Polynom bis zur 3. Ordnung ist damit

$$P_3(x, y) = 1 + x - \frac{x^3}{6} - \frac{xy^2}{2}.$$

Beispiel 4.35 (TAYLORENTWICKLUNG)

Wir suchen das Taylor-Polynom von $g(x, y) = \cos(x^2 + y^2) - 1$ bis zur 4. Ordnung. Dafür nutzen wir Methode 4.33 und substituieren mit $t = x^2 + y^2$. Es folgt für das Taylor-Polynom $\cos(t) = 1 - t^2/2 + \dots$. Damit können wir einsetzen

$$1 - \frac{t^2}{2} - 1 = -\frac{(x^2 + y^2)^2}{2} = -\frac{x^4 + 2x^2y^2 + y^4}{2} = -\frac{1}{2}x^4 - x^2y^2 - \frac{1}{2}y^4$$

Das Taylor-Polynom 4. Ordnung ist daher

$$P_4(x, y) = -\frac{1}{2}x^4 - x^2y^2 - \frac{1}{2}y^4.$$

Beispiel 4.36 (TAYLORENTWICKLUNG)

Wir möchten jetzt die Taylorapproximation 6. Ordnung für die Funktion $f(x, y) = \sin(xy)$ um die Stelle $(0, 0)$ finden. Wir kennen die Taylorapproximation von $\sin(t) \approx t - t^3/6 + \dots$ um 0. Damit folgt durch die Substitution $t = xy$

$$\sin(xy) \approx xy - \frac{x^3y^3}{6}.$$

5 Extremwertprobleme

5.1 Kritische Punkte und lokale Extrema

Sei $U \subseteq \mathbb{R}^2$ offen und nicht-leer. Wir untersuchen den Zusammenhang zwischen Ableitungen und Extrema von reellwertigen Funktionen $f : U \rightarrow \mathbb{R}$. Wie bei Funktionen einer Veränderlichen ist das Verschwinden der Ableitung eine notwendige, aber nicht hinreichende Bedingung für das Vorliegen eines Extremums.

Bemerkung 5.1

Ein Punkt $(x_0, y_0) \in U$ heisst ein **lokales Maximum** von f , wenn es ein $r > 0$ gibt, sodass $f(x, y) \leq f(x_0, y_0)$ für alle $(x, y) \in B_r(x_0, y_0)$ gilt. Hierbei ist $B_r(x_0, y_0)$ eine Kreisscheibe im \mathbb{R}^2 um den Punkt

(x_0, y_0) mit Radius r . Das zweidimensionale Analogon zu einem offenen Intervall. Ein Punkt (x_0, y_0) heisst ein **isoliertes lokales Maximum** oder ein **striktes lokales Maximum**, wenn es ein $r > 0$ gibt, sodass $f(x, y) < f(x_0, y_0)$ für alle $(x, y) \in B_r(x_0, y_0)$ mit $(x, y) \neq (x_0, y_0)$. Die Definition eines **lokalen Minimums** ist analog, und zusammengefasst bezeichnen wir sie als **lokale Extrema**.

Satz 5.2 (KRITISCHE PUNKTE IN OFFENEN MENGEN)

Sei $U \subseteq \mathbb{R}^2$ offen, $f : U \rightarrow \mathbb{R}$ eine C^1 -Funktion, und sei $(x_0, y_0) \in U$ ein Punkt, an dem f differenzierbar ist und ein lokales Extremum annimmt. Dann gilt $\partial_i f(x_0, y_0) = 0$ für $i = x, y$. Mit anderen Worten, $\nabla f(x_0, y_0) = 0$ und (x_0, y_0) ist ein kritischer Punkt von f .

5.2 Definitheit von Matrizen

Definition 5.3 (DEFINITHEIT EINER SYMMETRISCHEN MATRIX)

Eine symmetrische Matrix $A \in \mathbb{R}^{n \times n}$ heisst:

- (1) **Positiv definit**, wenn alle ihre Eigenwerte positiv sind.
- (2) **Negativ definit**, wenn alle Eigenwerte negativ sind.
- (3) **Indefinit**, wenn mindestens ein Eigenwert positiv und mindestens ein Eigenwert negativ ist.
- (4) **Ausgeartet**, wenn null ein Eigenwert ist.

Bemerkung 5.4

Merke, dass wir im folgenden Teil nur die Hesse-Matrix betrachten und diese nach dem Satz von Schwarz immer symmetrisch ist.

5.3 Hesse-Test und Klassifikation kritischer Punkte

Satz 5.5 (HESSE-TEST)

Sei $U \subseteq \mathbb{R}^2$ offen und $f \in C^3(U)$, und sei $(x_0, y_0) \in U$ ein kritischer Punkt resp. $\nabla f(x_0, y_0) = 0$. Sei $Hf_{(x_0, y_0)}$ die Hesse-Matrix (siehe Definition 4.26) von f im Punkt (x_0, y_0) .

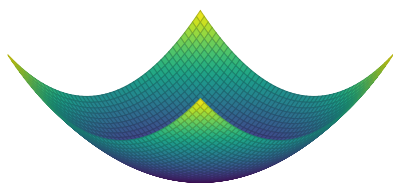
- (1) Wenn $Hf_{(x_0, y_0)}$ positiv definit ist, dann hat f ein striktes **lokales Minimum** bei (x_0, y_0) .
- (2) Wenn $Hf_{(x_0, y_0)}$ negativ definit ist, dann hat f ein striktes **lokales Maximum** bei (x_0, y_0) .
- (3) Wenn $Hf_{(x_0, y_0)}$ indefinit und nicht ausgeartet ist, dann hat f kein lokales Extremum bei (x_0, y_0) . In diesem Fall wird (x_0, y_0) als **Sattelpunkt** bezeichnet.

Bemerkung 5.6

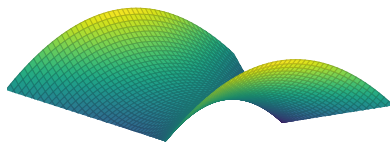
Wenn die Hesse-Matrix ausgeartet ist, das heisst wenn 0 ein Eigenwert von $Hf_{(x_0, y_0)}$ ist, dann ist der Hesse-Test (siehe Satz 5.5) nicht aussagekräftig und es kann nichts gesagt werden: Zum Beispiel hat die Funktion $f(x, y) = ax^4 + by^4$ ein lokales Maximum, ein lokales Minimum oder keines von beiden bei 0, abhängig von der Wahl von a und b . Aber die Hesse-Matrix bei 0 ist die Nullmatrix, unabhängig von der Wahl von a und b .

Beispiel 5.7

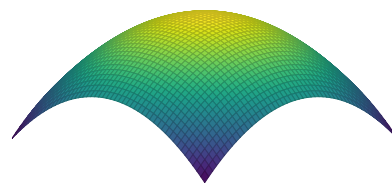
Wir betrachten die drei unten abgebildeten Funktionen.



$$f_1(x, y) = x^2 + y^2$$



$$f_2(x, y) = x^2 - y^2$$



$$f_3(x, y) = -x^2 - y^2$$

Wir merken, dass für alle drei Fälle $x_0 = (0, 0)$ ein kritischer Punkt ist. Wir möchten die Art dieses kritischen Punktes nun mit dem Hesse-Test analysieren. Die entsprechenden Hesse-Matrizen für die drei Funktionen sind

$$Hf_1 = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, \quad Hf_2 = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}, \quad \text{und} \quad Hf_3 = \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix}.$$

Wir können die Eigenwerte der Diagonalmatrizen direkt ablesen und sehen, dass Hf_1 positiv definit und Hf_3 negativ definit sind. Damit haben wir im ersten Fall ein Minimum und im zweiten Fall ein Maximum. Der erste Hauptminor ist positiv, aber der zweite negativ. Damit ist sie indefinit und nicht ausgeartet und f_2 hat der bei Null einen Sattelpunkt.

Aufgabe [2024 S] Gegeben ist die Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$,

$$f(x, y) = e^x + xy.$$

Welche der folgenden Aussagen über diese Funktion ist korrekt?

- (a) Die gegebene Funktion besitzt genau einen Sattelpunkt.
- (b) Die gegebene Funktion besitzt genau ein lokales Maximum.
- (c) Die gegebene Funktion besitzt genau ein lokales Minimum.
- (d) Die gegebene Funktion besitzt mehrere kritische Punkte.

5.4 Extrema unter Nebenbedingungen

Wir betrachten nun bedingte Optimierungsprobleme, also wir suchen Minima (oder Maxima) einer Funktion f nicht mehr auf einer offenen Menge, sondern auf einer Menge $M := \{(x, y) \in U : g(x, y) = 0\} \neq \emptyset$, die gegeben ist durch eine Nebenbedingung $g(x, y) = 0$. Dabei ist $g(x, y)$ eine Funktion $g : U \rightarrow \mathbb{R}$ und $g(x, y) = 0$ ist damit eine Niveaumenge dieser Funktion und parametrisiert eine Teilmenge in U – nämlich genau alle Punkte $(x, y) \in U$, die $g(x, y) = 0$ erfüllen resp. die Nullstellenmenge von g . Aus allen diesen Punkten $(x, y) \in M$, wollen wir nun die Extrema unserer Funktion f finden.

Bemerkung 5.8 (INTUITION HINTER LAGRANGE-MULTIPLIKATOREN)

Wir wissen, dass wir uns, um auf der Niveaulinie $g(x, y) = 0$ zu bleiben uns orthogonal zu ∇g bewegen müssen (siehe 4.22). Damit wir einen kritischen Punkt von f finden müssen wir eine Stelle auf $g(x, y) = 0$ finden, sodass wenn wir uns infinitesimal auf $g(x, y) = 0$ bewegen, sich f nicht ändert – wir halten uns also **lokal** auf einer Niveaulinie von f auf. Dafür muss an dieser Stelle ∇f also orthogonal zur Bewegungsrichtung zeigen, da wir unseren Funktionswert lokal nicht ändern wollen. Damit sind sowohl ∇f als auch ∇g orthogonal zur Bewegungsrichtung und damit parallel. Es gilt notwendigerweise für einen kritischen Punkt damit

$$\nabla f = \lambda \nabla g \iff \nabla f - \lambda \nabla g = 0.$$

Satz 5.9 (LAGRANGE MULTIPLIKATOREN)

Sei $U \subseteq \mathbb{R}^2$ offen und seien f, g Funktionen in $C^1(U, \mathbb{R})$. Sei

$$M := \{(x, y) \in U : g(x, y) = 0\} \neq \emptyset,$$

und es sei angenommen, dass $f|_M$ ein lokales Extremum bei $(x_0, y_0) \in M$ hat. Dann existiert eine reelle Zahl $\lambda \neq 0$, sodass

$$\nabla f(x_0, y_0) + \lambda \nabla g(x_0, y_0) = 0.$$

Mit anderen Worten, die Vektoren $\nabla f(x_0, y_0)$ und $\nabla g(x_0, y_0)$ sind linear abhängig.

Bemerkung 5.10 (LAGRANGE-FUNKTION)

Man kann in der Praxis die sogenannte **Lagrange-Funktion** definieren als:

$$\mathcal{L} : U \times \mathbb{R} \rightarrow \mathbb{R}, \quad \mathcal{L}(x, y, \lambda) = f(x, y) - \lambda g(x, y).$$

Die Zahl $\lambda \in \mathbb{R}$ wird als **Lagrange-Multiplikator** bezeichnet. Dann besagt Satz 5.9 dass, wenn (x_0, y_0) ein lokales Minimum für das Nebenbedingungsproblem ist, ein $\lambda \in \mathbb{R}$ existiert, sodass die Gleichungen

$$\partial_x \mathcal{L}(x_0, y_0, \lambda) = 0 \quad \text{und} \quad \partial_\lambda \mathcal{L}(x_0, y_0, \lambda) = 0$$

erfüllt sind.

Rechenmethode 5.11 (LAGRANGE MULTIPLIKATOREN)

Ein Minimum einer Funktion mit Nebenbedingungen können wir mithilfe von Lagrange-Multiplikatoren finden. Dafür können wir wie folgt vorgehen:

1. Lagrange-Funktion definieren
2. Ableitungen berechnen
3. Gleichungssystem lösen
4. Einsetzen und Lösung finden

Beispiel 5.12 (LAGRANGE AUF EINEM EBENENSCHNITT)

Betrachten wir die Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ mit $f(x, y) = x^2 + y^2$ die minimiert werden soll unter der Nebenbedingung $x + y = 1$.

SCHRITT 1: LAGRANGE-FUNKTION DEFINIEREN

Die Lagrange-Funktion ist gegeben durch:

$$\mathcal{L}(x, y, \lambda) = x^2 + y^2 - \lambda(x + y - 1).$$

SCHRITT 2: ABLEITUNGEN BERECHNEN

Die notwendigen Bedingungen für ein Extremum sind:

$$\begin{cases} \frac{\partial \mathcal{L}}{\partial x} = 2x - \lambda = 0, \\ \frac{\partial \mathcal{L}}{\partial y} = 2y - \lambda = 0, \\ \frac{\partial \mathcal{L}}{\partial \lambda} = -x - y + 1 = 0. \end{cases}$$

SCHRITT 3: GLEICHUNGSSYSTEM LÖSEN

Aus den ersten beiden Gleichungen folgt:

$$2x = \lambda \quad \text{und} \quad 2y = \lambda \Rightarrow 2x = 2y \Rightarrow x = y.$$

Setzen wir dies in die Nebenbedingung ein:

$$x + x = 1 \Rightarrow 2x = 1 \Rightarrow x = \frac{1}{2}, \quad y = \frac{1}{2}.$$

SCHRITT 4: EINSETZEN UND LÖSUNG FINDEN

Die Grenzwerte im unendlichen divergieren gegen unendlich. Daher ist das Minimum der Funktion $f(x, y) = x^2 + y^2$ unter der Bedingung $x + y = 1$ liegt bei

$$\left(\frac{1}{2}, \frac{1}{2}\right) \quad \text{mit Funktionswert} \quad f\left(\frac{1}{2}, \frac{1}{2}\right) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}.$$

Beispiel 5.13 (LAGRANGE AUF EINEM KREISSCHNITT)

Betrachten wir die Funktion $f(x, y) = e^x \cdot y$ auf dem Einheitskreis ($x^2 + y^2 = 1$) minimiert werden soll. Man schneidet den Graphen der Funktion f also mit einem Zylinder $x^2 + y^2 = 1$ und findet Minima und Maxima auf dem resultierenden Pfad (siehe Bild).

SCHRITT 1: LAGRANGE-FUNKTION DEFINIEREN

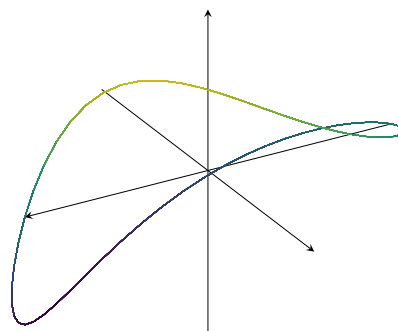
Die Lagrange-Funktion ist gegeben durch:

$$\mathcal{L}(x, y, \lambda) = e^x \cdot y - \lambda(x^2 + y^2 - 1).$$

SCHRITT 2: ABLEITUNGEN BERECHNEN

Die notwendigen Bedingungen für ein Extremum sind:

$$\begin{cases} \frac{\partial \mathcal{L}}{\partial x} = e^x \cdot y - 2\lambda x = 0, \\ \frac{\partial \mathcal{L}}{\partial y} = e^x - 2\lambda y = 0, \\ \frac{\partial \mathcal{L}}{\partial \lambda} = -x^2 - y^2 + 1 = 0. \end{cases}$$



SCHRITT 3: GLEICHUNGSSYSTEM LÖSEN

Aus der zweiten Gleichung folgt:

$$e^x = 2\lambda y \quad \Rightarrow \quad \lambda = \frac{e^x}{2y}.$$

Setzen wir dies in die erste Gleichung ein:

$$e^x \cdot y - 2 \cdot \frac{e^x}{2y} \cdot x = 0 \quad \Rightarrow \quad e^x \cdot y - \frac{e^x \cdot x}{y} = 0.$$

Multiplizieren wir beide Seiten mit y :

$$e^x \cdot y^2 - e^x \cdot x = 0 \quad \Rightarrow \quad e^x(y^2 - x) = 0.$$

Da $e^x \neq 0$ für alle x , folgt:

$$y^2 - x = 0 \quad \Rightarrow \quad x = y^2.$$

Setzen wir dies in die Nebenbedingung ein:

$$x^2 + x = 1.$$

$$x = \frac{-1 \pm \sqrt{5}}{2} \quad y = \pm \sqrt{\frac{-1 + \sqrt{5}}{2}}.$$

SCHRITT 4: EINSETZEN UND LÖSUNG FINDEN

Die Funktion $f(x, y) = e^x \cdot y$ hat unter der Bedingung $x^2 + y^2 = 1$ ein Maximum bei $\left(\frac{-1+\sqrt{5}}{2}, \sqrt{\frac{-1+\sqrt{5}}{2}}\right)$ und ein Minimum bei $\left(\frac{-1+\sqrt{5}}{2}, -\sqrt{\frac{-1+\sqrt{5}}{2}}\right)$.

Die entsprechenden Funktionswerte sind ungefähr -1.456 und 1.456 .

Aufgabe [2023 S] Gegeben ist die Funktion

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

Sei zudem $g(x, y) = x^2 + y^2$. Wir schränken die Argumente der Funktion $f(x, y)$ auf den Kreis

$$\{(x, y) \in \mathbb{R}^2 \mid g(x, y) = 1\}$$

ein. In diesem Fall nimmt die Funktion $f(x, y)$ unter der Nebenbedingung $g(x, y) = 1$

- (A) in $(0, 1)$ ihr bedingtes Maximum an.
- (B) in $\left(\frac{\sqrt{3}}{2}, \frac{1}{2}\right)$ und $\left(-\frac{\sqrt{3}}{2}, \frac{1}{2}\right)$ ihr bedingtes Maximum an.
- (C) in $(0, -1)$ ihr bedingtes Maximum an.
- (D) in $\left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right)$ und $\left(-\frac{1}{2}, \frac{\sqrt{3}}{2}\right)$ ihr bedingtes Maximum an.