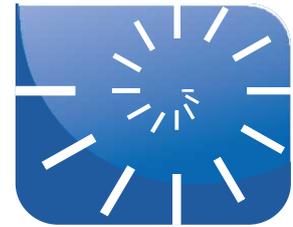


PH Schwäbisch Gmünd
University of Education



Prof. Dr. Helmut Albrecht:

Analysis II

Institut für Mathematik und Informatik

Inhalt

| | |
|---|-----------|
| Teil 1: Mehrdimensionale Analysis | 5 |
| 1 Zweidimensionale Funktionen | 6 |
| 1.1 Graphen zweidimensionaler Funktionen | 6 |
| 1.2 Höhenlinien | 8 |
| 2 Ableitungen von Funktionen zweier Veränderlicher | 12 |
| 2.1 Stetigkeit in $\mathbb{R} \times \mathbb{R}$ | 12 |
| 2.2 Die Ableitung im Raum | 13 |
| 2.3 Partielle Funktionen | 13 |
| 2.4 Partielle Ableitung | 14 |
| 2.5 Partielle Ableitung höherer Ordnung | 17 |
| 3 Methode der kleinsten Quadrate | 19 |
| 3.1 Idee | 19 |
| 3.2 Mathematisches Verfahren | 20 |
| 3.3 Beispielrechnung | 23 |
| 3.4 Bestimmung mit Maxima | 24 |
| 4 Elastizität | 26 |
| 4.1 Bogenelastizität | 26 |
| 4.2 Punktelastizität | 27 |
| 4.3 Funktionsgleichung der Elastizität | 28 |
| 4.4 Partielle Elastizität | 29 |
| 5 Gradient und Richtungsableitung | 31 |
| 5.1 Der Gradient | 31 |
| 5.2 Richtungsableitung | 34 |
| 6 Totales Differential | 38 |
| 6.1 Rückblick auf eindimensionale Funktionen | 38 |
| 6.2 Zweidimensionale Funktionen | 40 |
| 6.3 Anwendungen | 43 |
| 6.3.1 Lineare Näherung | 43 |
| 6.3.2 Fehlerrechnung | 45 |
| 7 Lokale Extremwerte | 47 |
| 7.1 Notwendige Bedingung | 47 |
| 7.2 Hinreichende Bedingung | 48 |
| 7.3 Anwendungsbeispiel | 50 |
| 7.4 Anwendung in der Betriebswirtschaft | 52 |
| 8 Extremwerte mit Nebenbedingungen | 53 |
| 8.1 Einsetzverfahren | 53 |
| 8.2 Lagrange-Optimierung | 54 |
| 8.2.1 Herleitung des Lagrange-Verfahrens | 54 |
| 8.2.2 Beispiel zum Lagrange-Verfahren | 57 |
| 8.2.3 Bestimmung der Extrema mit Hilfe der berandeten Hessematrix | 59 |
| 9 Doppelintegrale | 61 |
| 9.1 Die grundlegende Idee | 61 |
| 9.2 Integrieren eines Volumens | 63 |
| 9.2.1 Erster Ansatz | 63 |
| 9.2.2 Zweiter Ansatz | 64 |
| 9.2.3 Satz von Fubini | 65 |
| 9.2.4 Ein weiteres Beispiel | 66 |
| 9.2.5 Berechnung eines Doppelintegrals in Maxima | 66 |

| | |
|---|-----------|
| Teil 2: Differenzialgleichungen | 68 |
| 10 Hinführende Problemstellungen..... | 69 |
| 10.1 Radioaktiver Zerfall..... | 69 |
| 10.2 Abkühlungsvorgänge..... | 69 |
| 10.3 Das Federpendel..... | 70 |
| 10.4 Laplace- und Wellengleichung | 71 |
| 10.5 Zusammenfassung..... | 71 |
| 11 Differenzialgleichungen bei einfachen Bewegungen | 72 |
| 11.1 Die gleichförmige Bewegung..... | 72 |
| 11.2 Gleichmäßig beschleunigte Bewegung | 72 |
| 12 Anfangs- und Randwertprobleme..... | 73 |
| 13 Arten von Differenzialgleichungen..... | 75 |
| 13.1 Gewöhnliche vs. partielle Differenzialgleichungen | 75 |
| 13.2 Lineare vs. nichtlineare Differenzialgleichungen..... | 75 |
| 13.3 Homogenität..... | 76 |
| 13.4 Ordnung..... | 76 |
| 14 Lösungsmöglichkeiten von Differenzialgleichungen | 78 |
| 14.1 Lösen durch Raten | 78 |
| 14.2 Graphische Lösung | 78 |
| 14.3 Lösung von Differenzialgleichungen mit Hilfe des Computers | 85 |
| 15 Rechnerische Lösung von Differenzialgleichungen | 86 |
| 15.1 Lineare Differentialgleichungen erster Ordnung..... | 86 |
| 15.1.1 <i>Homogene Gleichungen</i> | 86 |
| 15.1.2 <i>Inhomogene Gleichungen</i> | 89 |
| 15.1.3 <i>Zusammenfassung</i> | 91 |
| 15.2 Nicht-lineare Differentialgleichungen erster Ordnung..... | 93 |
| 15.3 Abschließende Anmerkungen | 93 |
| 15.3.1 <i>Umgang mit der Integrationskonstanten</i> | 93 |
| 15.3.2 <i>Umformung der Gleichung</i> | 93 |
| 16 Anfangswertprobleme..... | 94 |
| 16.1 Hinführende Betrachtungen..... | 94 |
| 16.1.1 <i>Gleichförmige Bewegung</i> | 94 |
| 16.1.2 <i>Gleichmäßig beschleunigte Bewegung</i> | 94 |
| 16.2 Rechnerische Lösung von Anfangswertproblemen | 96 |
| 16.2.1 <i>Erstes Beispiel</i> | 96 |
| 16.2.2 <i>Zweites Beispiel</i> | 96 |
| 16.3 Das Eulerverfahren..... | 97 |
| 16.3.1 <i>Prinzipielles Vorgehen</i> | 97 |
| 16.3.2 <i>Das Eulerverfahren im Beispiel</i> | 98 |
| 16.3.3 <i>Das Eulerverfahren in EXCEL</i> | 99 |
| 16.3.4 <i>Das Eulerverfahren mit Maxima</i> | 100 |

Literatur:

- Günter Bräuning: Gewöhnliche Differentialgleichungen, Frankfurt: Harri Deutsch, 1965
- Büchter/Henn: Elementare Analysis, Heidelberg: Spektrum, 2010
- Sybille Handrock-Meyer: Differenzialgleichungen für Einsteiger, München: Carl Hanser, 2007

- Harald Scheid: Folgen und Funktionen, Münster: WTM, 2007
- Timm Sigg: Grundlagen der Differenzialgleichungen für Dummies, Weinheim: Wiley-VCH, 2012
- Smirnow, W. I.: Lehrgang der höheren Mathematik, Teil 1, Berlin: VEB, 1956
- Smirnow, W. I.: Lehrgang der höheren Mathematik, Teil 2, Berlin: VEB, 1960
- Thomas/Weir/Hass: Analysis 1, Hallbergmoos: Pearson, 2013
- Thomas/Weir/Hass: Analysis 2, Hallbergmoos: Pearson, 2014
- und viele mehr ...

Studienorganisatorische Hinweise:

Verortung und Leistungsnachweis

Die Veranstaltung Analysis II ist Bestandteil des Moduls 1 im Masterstudium der Sekundarstufe. Zusammen mit der Veranstaltung *Koordinatengeometrie* ist sie Inhalt der zugehörigen Modulklausur.

Inverted Classroom

Diese Veranstaltung wird in der (hochschul-)didaktischen Form des *inverted classrooms* durchgeführt. Dies bedeutet, dass Sie keine Vorlesung im üblichen Stil erwartet. Vielmehr stelle ich Ihnen mit diesem Skript, den Angaben zur Begleitliteratur und ggf. weiteren Materialien alles zur Verfügung, was Sie für eine Erarbeitung der Thematik benötigen. Von tragender Bedeutung sind die Aufgabenzusammenstellungen zu den einzelnen Themen. Anhand dieser Aufgaben sollen Sie die jeweiligen mathematischen Inhalte üben und Ihre Erkenntnisse in der Präsenzveranstaltung diskutieren. Diese Veranstaltungsform ist natürlich auch aufwändiger für Sie wesentlich effizienter – aber – als übliche Vorlesungen! Deshalb noch ein Wort zu dem Arbeitsaufwand, den Sie mit dieser Veranstaltung haben werden, bzw. der von Ihnen erwartet wird:

Der Veranstaltung *Analysis II* sind 4 ECTS-Punkte zugeordnet, dies ergibt einen *workload* von 120 Stunden. Wenn Sie davon Ihre Präsenz im Umfang von 21 Stunden in Abzug bringen, verbleiben noch knapp 100 Stunden für Ihre Vor- und Nachbereitung. Umgerechnet auf eine Vorlesungswoche sind dies immerhin noch rund 7 Stunden! Berücksichtigen Sie diese Zahl, wenn Ihnen der Arbeitsaufwand hoch erscheint.

Vorbemerkung: Die nachfolgende Darstellung von Funktionen zweier Veränderlicher, partieller Ableitungen und Extrema ist anwendungsorientiert und pragmatisch aufgebaut. Es werden durchaus wesentliche Gesichtspunkte unterschlagen, so wird beispielsweise der Begriff der Stetigkeit auf mehrdimensionalen Funktionen nicht in der nötigen Tiefe erörtert. Die Darstellung folgt daher nicht dem strukturmathematischen Aufbau einer höheren Analysis-Vorlesung, sondern dient vielmehr einem Ausblick auf eine Analysis über die in der Oberstufe vermittelten Inhalten hinaus.

Dies gilt ebenso für den zweiten Teil der Veranstaltung, der lediglich einen ersten orientierenden Einblick in die Welt der Differenzialgleichungen geben möchte.

Teil 1: Mehrdimensionale Analysis

Bisher wurden im Wesentlichen Funktionen betrachtet, bei denen *einzelnen* Dingen (Zahlen, Personen, sonstigen Objekten) ebenfalls *einzelne* Funktionswerte zugeordnet wurden. Mathematisch ausgedrückt wurden Zuordnungen aus einer eindimensionalen Menge (dem Definitionsbereich) in eine eindimensionale Menge (Wertebereich) vorgenommen. Diese Beschränkung ist allerdings nicht zwingend, wir können grundsätzlich aus einer beliebig-dimensionalen Menge in eine beliebig-dimensionale Menge abbilden. Jede Maxima-Funktion, die mehr als einen Aufrufparameter enthält, ist hierfür ein treffendes Beispiel. Außerdem kommt es nicht selten vor, dass das Ergebnis einer Maxima-Funktion nicht nur eine Zahl ist, sondern ein Paar von Zahlen oder auch eine ganze Liste mit Zahlen.

1 Zweidimensionale Funktionen

Funktionen können sowohl auf der „Eingabe“- wie auch auf der „Ausgabe“-seite mehrdimensional sein. Wir werden uns im Folgenden zumeist auf Funktionen beschränken, die einem Paar von Zahlen eindeutig einen (einzelnen) Wert zuordnen. Das ist nicht wirklich etwas erschreckend Neues, bereits Grundschüler gehen problemlos mit solchen Funktionen um, wenn sie beispielsweise dem Paar $[12, 4]$ den Wert 16 zuordnen (Addition) oder demselben Paar den Wert 48 (Multiplikation). Unsere bekannten Grundrechenarten sind also bereits solche „Funktionen zweier Veränderlicher“, mit denen wir uns nun etwas näher auseinandersetzen werden. Zunächst definieren wir allgemein:

Eine Funktion in n Variablen ist eine Abbildung

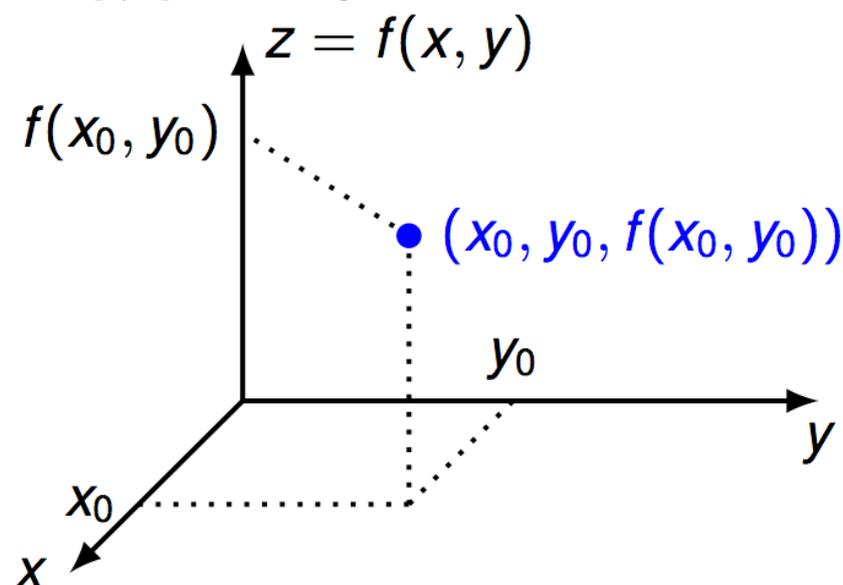
$$f: D \rightarrow \mathbb{R}, \quad D \subseteq \mathbb{R}^n,$$

welche jedem n -Tupel (x_1, x_2, \dots, x_n) eine Zahl $f(x_1, x_2, \dots, x_n) \in \mathbb{R}$ zuordnet.

Die Definitionsmenge ist im Allgemeinen \mathbb{R}^n oder auch nur eine Teilmenge davon, wir beschränken uns auf im Folgenden meist auf den \mathbb{R}^2 , also die xy -Ebene (oder auch nur einen Ausschnitt dieser Ebene) des bekannten Koordinatensystems. Jedem Punkt dieser „Definitions-Fläche“ wird dann eine Zahl z zugeordnet und zwar in Abhängigkeit von den Werten von x und y .

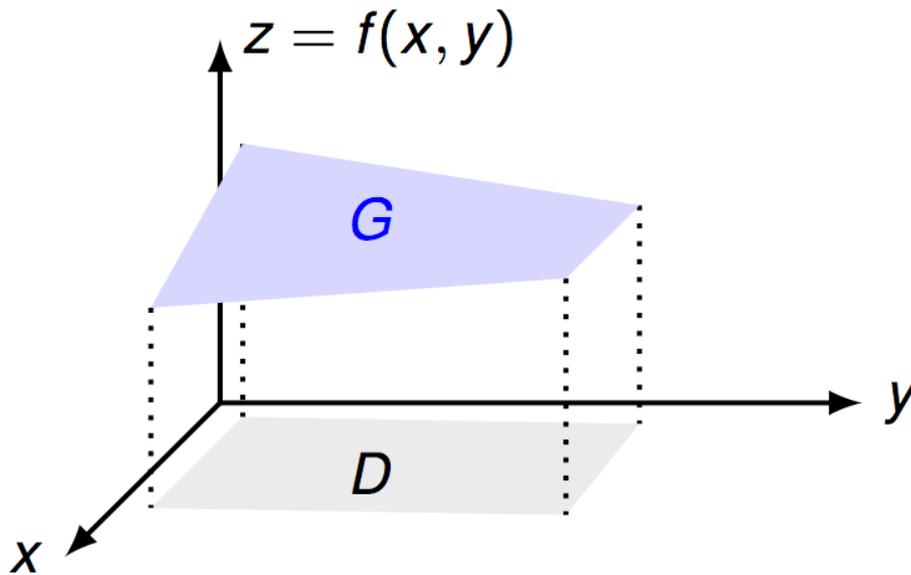
1.1 Graphen zweidimensionaler Funktionen

Damit ist bereits auch schon Wesentliches zur Darstellung solcher Funktionen in zwei Variablen gesagt. Seither war der Graph einer Funktion durch eine Kurve im Koordinatensystem gegeben – jedem x -Wert und damit jedem Punkt auf der x -Achse konnte eindeutig ein y -Wert und damit ein Punkt auf der y -Achse zugeordnet werden. Für diese Darstellung war ein zweidimensionales Koordinatensystem ausreichend. Wenn wir zukünftig zwei Variablen – x und y – einen dritten Wert z zuordnen wollen, dann benötigen wir für diese Darstellung ein dreidimensionales Koordinatensystem. Hierfür legen wir unser bekanntes xy -System in die Horizontale und lassen die (positive) z -Richtung nach oben laufen. Es wird dann jedem geordneten Paar $[x,y]$ und damit jedem Punkt der xy -Ebene ein genau darüber oder darunter liegender Punkt $[x,y,z]$ im Raum zugeordnet.



Ist die Definitionsmenge die ganze xy -Ebene – also der \mathbb{R}^2 – oder nur ein Ausschnitt davon, so ist der Graph einer Funktion mit zwei Variablen eine Fläche, die über diesem Definitions-

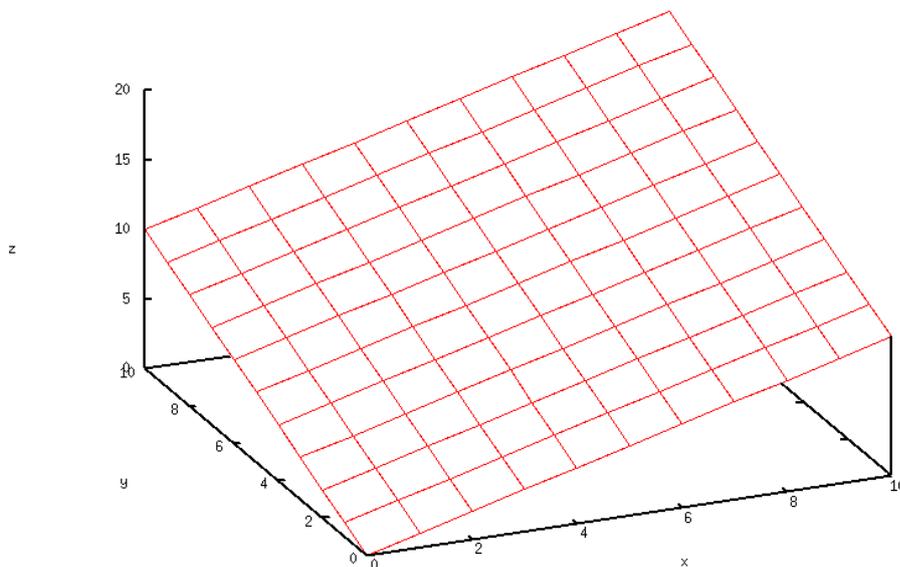
bereich liegt. Alle diese Bildpunkte bilden damit eine Fläche im Raum. Der Graph einer Funktion in zwei Variablen ist also eine (nicht notwendigerweise ebene) Fläche im Raum. Der Graph lässt sich vielmehr mit der Erdoberfläche vergleichen, die an manchen Stellen eben ist und an anderen Stellen Erhebungen und Senken aufweist. Man kann jedem Punkt der Erdoberfläche (gegeben durch seine geographischen Koordinaten: nördliche/südliche Breite und westliche/östliche Länge) seine Höhe über dem Bezugsniveau „Meeresspiegel“ zuordnen. Der Graph einer Funktion zweier Veränderlicher kann man also als eine „Landschaft“ über dem Gebiet der Definitionsmenge interpretieren.



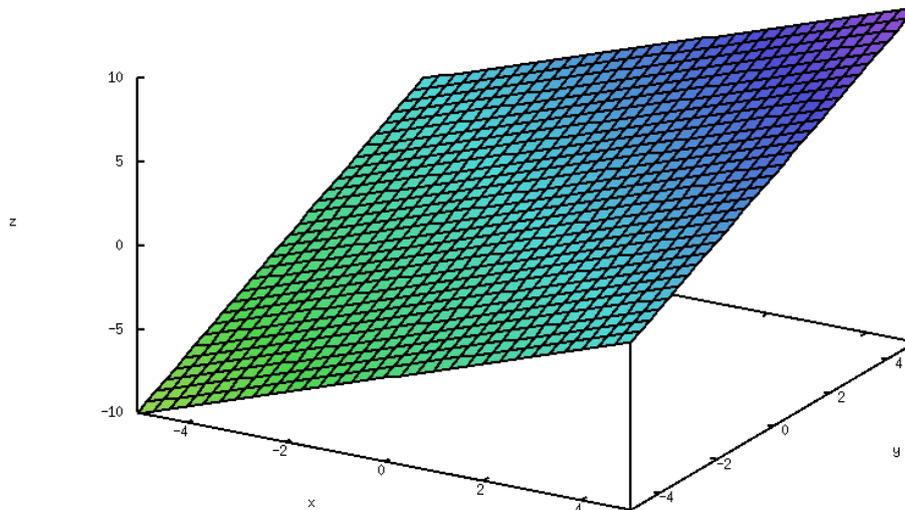
Machen wir uns diese Zusammenhänge anhand der oben beschriebenen Addition und damit mit der einfachen Funktion

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

deutlich: Jedem Gitterpunkt der xy -Ebene wird dabei ein z -Wert als Summe von $x+y$ zugeordnet. Beispielsweise ist $f(3,4) = 7$, es ist $f(6,8) = 14$ u.s.w. Von dem Gitterpunkt $[3,4]$ geht man also 7 Einheiten nach oben, vom Gitterpunkt $[6,8]$ geht man 14 Einheiten nach oben u.s.w. Bleibt man quasi in der Grundschule und lässt nur natürliche Zahlen für x und y zu, so erhält man als Graph diskrete Punkte im Raum (über der xy -Ebene). Im folgenden Plot sind dies die Kreuzungspunkte des roten Gitters:



Verlässt man die Grundschule und damit die Einschränkung auf den Bereich der natürlichen Zahlen, so erhalten wir als Graph der Additions-Funktion eine ebene Fläche, die schräg über und unter der xy -Ebene verläuft:



Beide Plots wurden mit Maxima erstellt. Der Befehl für den ersten Plot mit dem Gitter lautet:

```
plot3d(x+y, [x, 0, 10], [y, 0, 10], [grid, 10, 10], [palette, false]);
```

Der zweite Plot wurde mit

```
plot3d(x+y, [x, -5, 5], [y, -5, 5]);
```

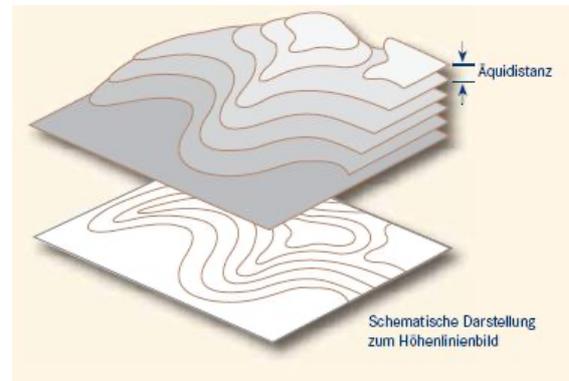
erzeugt. Experimentieren Sie weiter und lassen Sie sich andere Funktionen zweier Veränderlicher anzeigen.

1.2 Höhenlinien

Es wurde bereits die Analogie des Graphen einer Funktion zweier Variablen zur Erdoberfläche angesprochen. Für die (zweidimensionale) Darstellung der Erdoberfläche liefern Landkarten/Wanderkarten mit Höhenlinien eine sehr gute und aussagekräftige Möglichkeit:



Die Höhenlinien ergeben sich als Schnittlinien von parallel zur Meeresoberfläche liegenden Ebenen mit dem Gelände. Sehr gut kommt diese Vorstellungen bei terrassenförmig angelegten Reisfeldern zum Ausdruck:

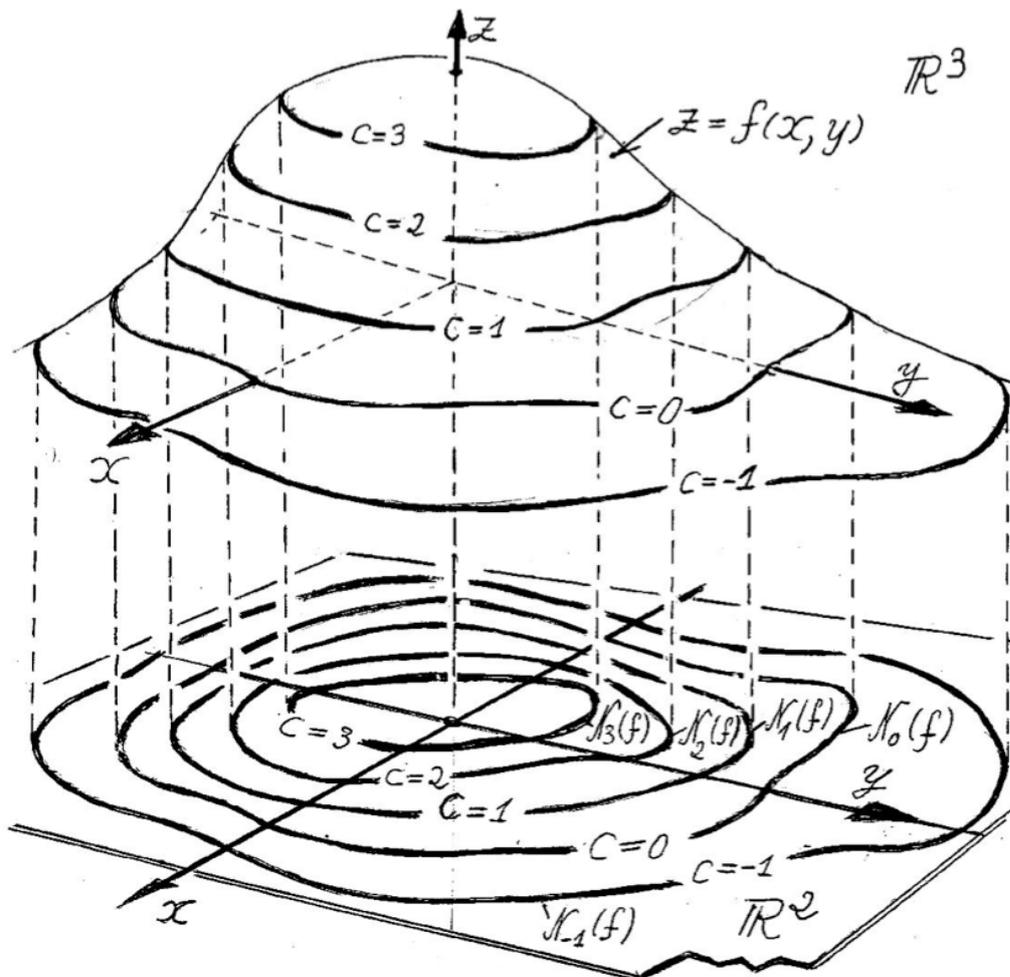


<http://www.geolinde.musin.de>

Auch für Funktionen in zwei Veränderlichen ist dies eine mögliche Darstellungsform: Man wählt eine Zahl c (die Höhe) und bestimmt alle Punkte $[x,y]$, deren Funktionswert $f(x,y) = c$ ist. Die Menge dieser Punkte bildet für ein festes c im Normalfall eine Kurve in der xy -Ebene. Formal gefasst ist eine Niveaulinie von f auf der Höhe c :

$$N_c(f) := \{[x,y] \in D \mid f(x,y) = c\}$$

Anschaulich sind diese Niveaulinien nichts anderes als Höhenlinien des Graphen einer Funktion in zwei Variablen. Diese Höhenlinien eines Funktionsgraphen werden auch als *Isoquanten* bezeichnet.



Die Bestimmung von Höhenlinien ist ein eher algebraisch orientiertes Problem. Möchte man Höhenlinien der Funktion $f(z) = x + y$ der Höhe c bestimmen, so muss man die Gleichung

$$x + y = c$$

in der Menge der reellen Zahlen „lösen“. Da diese Gleichung für jedes beliebige c unendlich viele Lösungen hat, kann man sich für eine orientierende Skizze bei einem festen c einige x -Werte vorgeben und die zugehörigen y -Werte berechnen. Zum Beispiel wollen wir die Höhenlinie 3 für die angegebene Funktion bestimmen. Es müssen also die Lösungspaare der Gleichung

$$x + y = 3$$

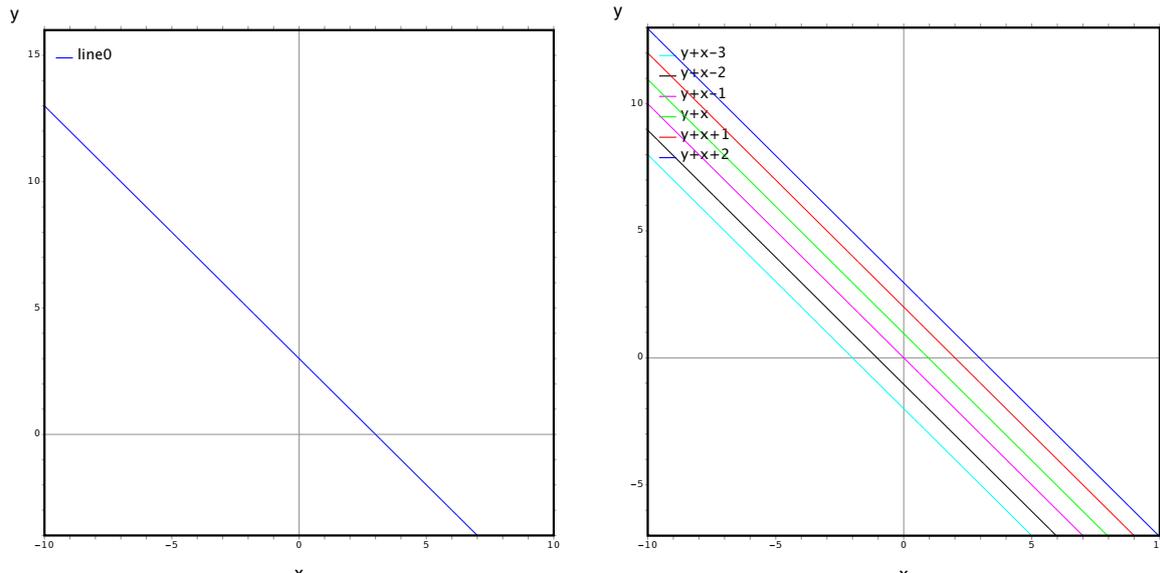
gefunden werden. Man kann dazu in einer Art Wertetabelle für einige x -Werte die zugehörigen y -Partner eintragen:

| | | | | | | | | | |
|----|----|----|----|---|---|---|---|----|----|
| x: | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 |
| y: | 6 | 5 | 4 | 3 | 2 | 1 | 0 | -1 | -2 |

Oder aber, was sich bei diesem Beispiel natürlich anbietet, die Gleichung nach y auflösen

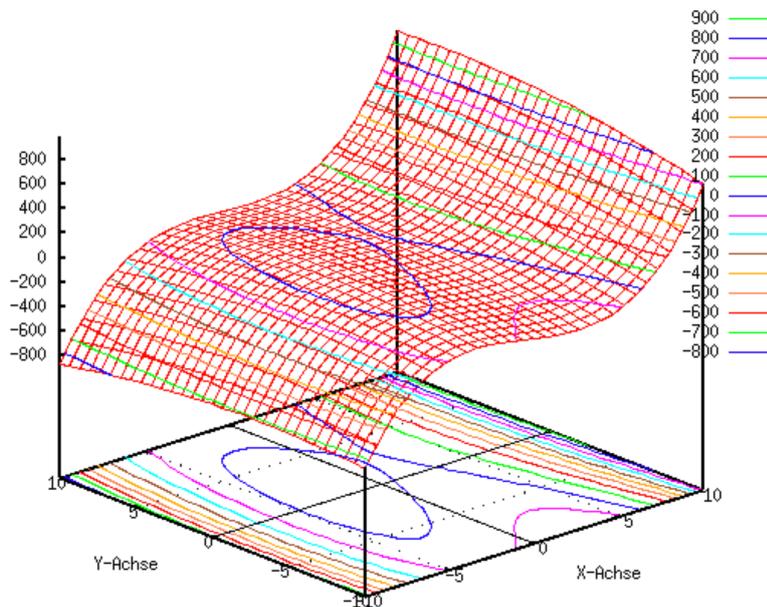
$$y = 3 - x \quad \Leftrightarrow \quad y = -x + 3$$

und erkennen, dass alle Paare $[x, y]$ auf einer Geraden mit der Steigung -1 und durch den Achsenabschnitt 3 in der xy -Ebene liegen.



Es wird schnell klar, dass andere Höhenlinien dieser Funktion ebenfalls Geraden mit der Steigung -1 sind, deren Achsenabschnitt jeweils der gewählten Höhe c entspricht. Demzufolge liegen alle Höhenlinien der untersuchten Funktion parallel zueinander.

In Maxima steht mit der Funktion `draw()` eine recht universelle Möglichkeit zum Zeichnen auch von dreidimensionalen Funktionsgraphen zur Verfügung, mit der diese Höhenlinien gezeichnet werden können. Informieren Sie sich in der *Maxima-Anleitung* über den korrekten Gebrauch dieser Funktion.



Damit ist es möglich, die Höhenlinien nicht nur in die xy -Ebene, sondern auch auf den Graphen zeichnen zu lassen.

2 Ableitungen von Funktionen zweier Veränderlicher

2.1 Stetigkeit in $\mathbb{R} \times \mathbb{R}$

Vor einer Frage nach der Differenzierbarkeit einer Funktion muss selbstverständlich die Frage nach deren Stetigkeit geklärt sein. Seither hatten wir wohl mathematisch unscharf aber doch sehr plastisch formuliert: „Eine Funktion ist dann stetig, wenn sich der Funktionsgraph von Hand ohne abzusetzen zeichnen lässt.“ In ihrem Wesen ist diese Erklärung immer noch passend, allerdings können wir den Funktionsgraphen, der nun ja eine mehr oder minder gewellte Fläche darstellt, ohnehin nicht mehr in einem Zug zeichnen. Gemeint ist nun vielmehr, dass wir auf dem Funktionsgraphen in jede beliebige Richtung fortschreiten können müssen, ohne dass irgendwelche „Sprünge“ oder „Absätze“ auftreten.

Eine klare Definition der Stetigkeit hatten wir in der Formulierung erreicht, dass an jeder Stelle x_0 der Funktionswert $f(x_0)$ gleich dem Grenzwert $\lim_{\substack{x \rightarrow x_0 \\ x \neq x_0}} f(x)$ sein muss.

Damit kommt der Begriff des Grenzwerts ins Spiel. Den Grenzwert für eindimensionale Folgen hatten wir folgendermaßen definiert:

Die Funktion f hat für $x \rightarrow a$ den Grenzwert b , wenn für jede x -Folge $(x_n)_{n \in \mathbb{N}}$ mit $x_n \neq a$ und mit Grenzwert a die zugehörige y -Folge $(y_n)_{n \in \mathbb{N}}$ mit $y_n = f(x_n)$ ebenfalls konvergiert und ebendiesen Grenzwert b hat. Dabei müssen alle Elemente im Definitionsbereich der Funktion f liegen. Existiert ein derartiger Grenzwert, so schreibt man auch

$$\lim_{x \rightarrow a} f(x) = b$$

Maßgeblich für den Nachweis eines Grenzwerts waren hierbei die x -Folgen, die beliebig zur interessierenden Stelle a konvergieren müssen. Dafür gab es seither nur drei Möglichkeiten: Entweder war diese x -Folge eine monoton fallende Folge, die sich von rechts der Stelle a annähert, oder es war eine monoton steigende Folge, die sich von links annähert oder aber es war eine alternierende Folge, deren Glieder von rechts und von links immer näher an die Stelle a heranrücken. In jedem Fall aber konnten sich die Folgenglieder nur auf der eindimensionalen x -Achse befinden und waren dadurch auf ihrem Weg hin zu a schon stark kanalisiert.

Nun aber, bei Funktionen mit zwei Veränderlichen, befinden sich die Folgenglieder als Punkte in einer Ebene. Sie können also nicht nur von der positiven oder negativen x -Richtung auf die interessierende Stelle zulaufen, sondern natürlich auch von der positiven und negativen y -Richtung. Und natürlich gibt es darüber hinaus noch unendlich viele weitere Richtungen, aus denen sich die Folgenglieder der gefragten Stelle nähern können. Sie können zudem „in Kurven“ auf a zulaufen oder können sich spiralförmig auf a zusammenziehen. Damit gibt es weit mehr Möglichkeiten und die Forderung, dass für jede dieser x -Folgen sich die zugehörigen Funktionswerte dem Grenzwert nähern müssen, wird im konkreten Fall deutlich schwieriger nachzuweisen sein.

Eine weitere Definition der Stetigkeit beruhte auf der anschaulichen Vorstellung, dass, wenn man ein wenig am x -Wert „wackelt“, der zugehörige y -Wert auch nur ein wenig „wackeln“ darf:

Die Funktion $f: D_f \rightarrow \mathbb{R}$ heißt genau dann stetig in $x_0 \in D_f$, wenn es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt, sodass gilt: Aus

$$|x - x_0| < \delta, x \in D_f, \quad \text{folgt} \quad |f(x) - f(x_0)| < \varepsilon$$

Allerdings „wackeln“ wir jetzt nicht mehr nur in x -Richtung – wir können vielmehr in jede beliebige Richtung der xy -Ebene „wackeln“. Dabei darf dann jetzt der z -Wert auch nur ein

ganz klein wenig „wackeln“. Ohne die exakte Definition zu formulieren, soll hier nur deren Umschreibung für die Stetigkeit im \mathbb{R}^2 angegeben werden:

Eine Funktion f ist in einem Punkt (x_0, y_0) stetig, wenn es zu jeder ε -Umgebung seines Bildpunktes $f(x_0, y_0)$ eine δ -Umgebung von (x_0, y_0) gibt, die durch f ganz in die ε -Umgebung des Bildpunktes abgebildet wird.

2.2 Die Ableitung im Raum

Vergegenwärtigt man sich, dass wir unter der Ableitung einer Funktion die Steilheit ihres Graphen in einem Punkt definiert hatten, so stellt sich bei Funktionen zweier Veränderlicher das Problem, dass diese Steigung in einem speziellen Punkt in aller Regel nicht eindeutig ist. Stellen Sie sich dazu vor, dass Sie auf einen Berg steigen und an einem Punkt kurz innehalten. Schauen Sie von diesem Punkt aus weiter bergan, so geht es (mehr oder weniger) steil bergauf. Blicken Sie aber um 90° gedreht entlang der Bergflanke und damit entlang einer Höhenlinie, so ist auf diesem Weg die Steigung gleich Null und blicken Sie gar bergab, dann ist die Steigung negativ!

Als Ableitung hatten wir bei eindimensionalen Funktionen die Tangente an den Graphen definiert und diese ist in einem bestimmten Punkt – eine stetige und differenzierbare Funktion vorausgesetzt – immer eindeutig. Für deren Steigung hatten wir immer gefragt, wie sich denn der Höhenwert y ändert, wenn wir in der Horizontalen (x -Richtung) um einen Schritt weiter gehen.

Für zweidimensionale Funktionen, die als Graph eine „gewellte Landschaft“ haben, können wir an jeden Punkt dieser Landschaft eindeutig eine Tangentialfläche heften, wobei die Lage dieser Tangentialfläche im Raum ebenfalls eindeutig ist. Bei der Bestimmung ihrer „Steilheit“ allerdings wird es „schwierig“. Wir können beispielsweise fragen, wie der Höhenwert z zu- oder abnimmt, wenn wir ...

... einen Schritt in x -Richtung laufen

... einen Schritt in y -Richtung laufen

... gleichzeitig $\sqrt{\frac{1}{2}}$ in x - und $\sqrt{\frac{1}{2}}$ in y -Richtung gehen

... gleichzeitig $\sqrt{\frac{2}{5}}$ in x und $\sqrt{\frac{1}{5}}$ in y -Richtung gehen u.s.w.

Die Steigung hängt also davon ab, in welche Richtung wir weiter fortschreiten, sie ist damit nicht mehr „eindeutig“!

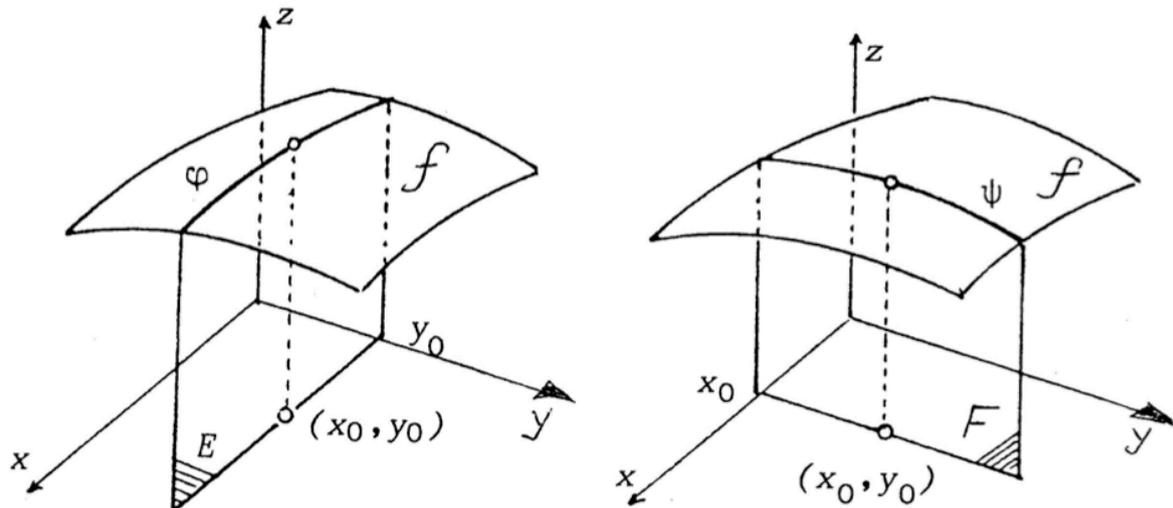
2.3 Partielle Funktionen

Wir haben Funktionen als linkstotale und rechtseindeutige Zuordnungen aus der Definitionsmenge in die Zielmenge kennen gelernt. Die Eigenschaft „linkstotal“ bedeutet, dass **jedem** Element der Definitionsmenge eindeutig ein Element zugeordnet werden muss. Im Zusammenhang mit unseren zweidimensionalen Funktionen bedeutet dies, dass man (bei üblicher Wahl des Definitionsbereichs) jedem Element aus $\mathbb{R} \times \mathbb{R}$ und damit jedem Punkt der xy -Ebene einen z -Wert zuordnen können muss. Genau so haben wir es seither auch gehandhabt und eine mehr oder weniger gewellte Fläche als Graph solcher Funktionen kennen gelernt.

Nun kann es aber sinnvoll sein, nicht die „ganze“ Funktion $\mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ zu untersuchen, sondern nur einen „Ausschnitt“ davon. Konkret heißt dies, dass wir nicht allen Elementen des Definitionsbereichs einen Funktionswert zuordnen, sondern eben nur einer Teilmenge des Definitionsbereichs. Damit erhalten wir nicht die definierte Funktion „als Ganzes“, sondern nur einen Ausschnitt davon – eben eine „partielle Funktion“.

Wir haben bereits solche partiellen Funktionen betrachtet, als wir gefragt haben, was als Schnitt des (ganzen) Funktionsgraphen mit einer zur xy -Ebene senkrechten Ebene entsteht. Diese zur xy -Ebene senkrechte Ebene greift aus der xy -Ebene ja nur ganz spezielle Punkte heraus – nämlich diejenigen Punkte welche auf der Schnittgeraden beider Ebenen liegen. Solche Schnittgeraden gibt es natürlich wieder unendlich viele, von besonderem Interesse sind aber diejenigen, welche parallel zur x - bzw. y -Achse liegen.

In jedem Fall erhalten wir als Graph einer solchen partiellen Funktion eine Linie als Graph und damit bewegen wir uns wieder auf bekanntem Terrain.



Der Graph φ ist die Schnittkurve des (kompletten) Graphen f mit der Schnittebene E parallel zur x -Achse und der Graph ψ ¹ ist die entsprechende Schnittkurve mit der Ebene F parallel zur y -Achse.

Etwas formaler ausgedrückt erhalten wir partielle Funktionen einer Funktion $f(x,y)$, indem wir y_0 festhalten und x als (einzige) Funktionsvariable verwenden oder aber, indem wir x_0 festhalten und y als Funktionsvariable verwenden. Wir erhalten so eine ...

... partielle Funktion in x -Richtung:

$$x \rightarrow f(x, y_0) =: \varphi(x)$$

bzw. eine

... partielle Funktion in y -Richtung:

$$y \rightarrow f(x_0, y) =: \psi(y)$$

2.4 Partielle Ableitung

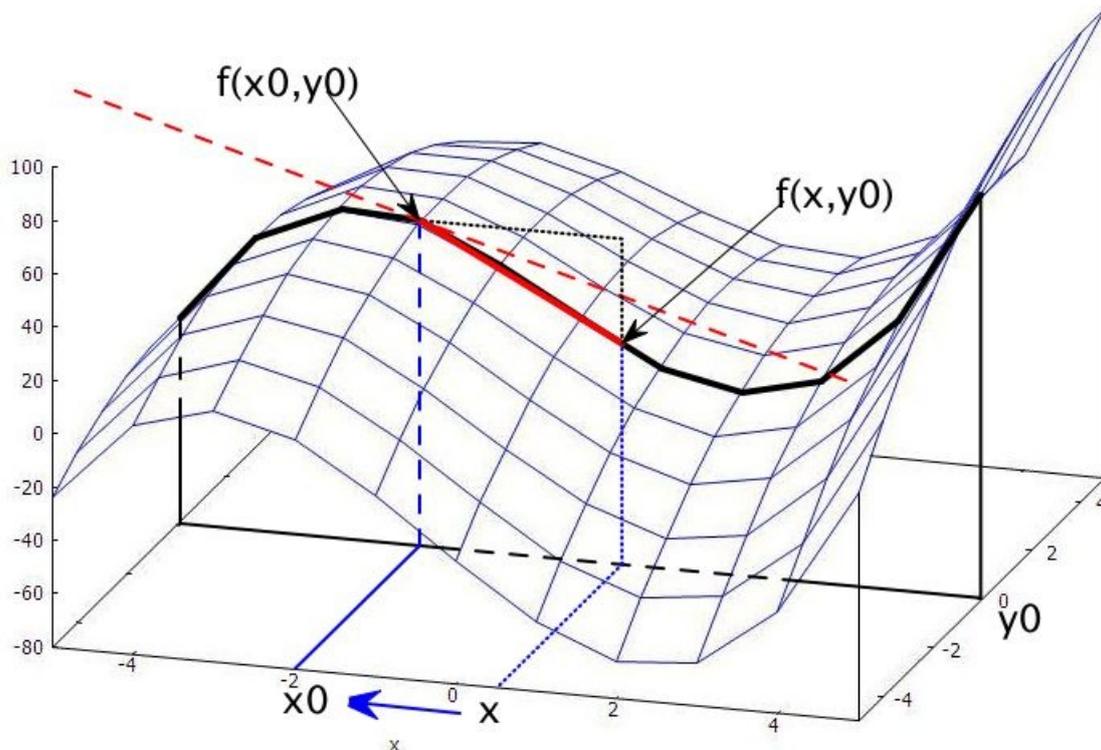
Mit Hilfe partieller Funktionen können wir recht einfach die Ableitung einer Funktion zweier Veränderlicher und damit die Steigung(en) einer an einen Punkt angelegten Tangentialfläche ermitteln. Dazu müssen wir die partiellen Ableitungen in x - und in y -Richtung ermitteln und erhalten daraus schließlich die Steigung in x -Richtung und die Steigung in y -Richtung an einem beliebigen Punkt des Graphen. Weil die Lage einer Ebene im Raum durch einen Punkt und zwei Richtungen festgelegt ist, ist damit auch die Lage der Tangentialfläche im interessierenden Punkt eindeutig und zweifelsfrei bestimmt.

Was es nun mit den partiellen Ableitungen auf sich hat, werden wir im Folgenden unter Rückgriff auf den seitherigen Ableitungsbegriff erörtern. Die Ableitung an einer Stelle x_0 hatten wir als Grenzwert des Differenzenquotienten

$$\lim_{x \rightarrow x_0} \frac{f(x_0) - f(x)}{x_0 - x}$$

¹ gr. Buchstabe psi

definiert. Diese Idee behalten wir grundsätzlich bei, indem wir durch „Festhalten“ einer Funktionsvariablen eine partielle Funktion erzeugen und diese dann wie gewohnt ableiten. Das „Festhalten“ von y meint lediglich, dass wir y als konstant annehmen. Wir müssen dafür keinen expliziten Wert (und schon gar nicht den Wert 0 wie in der untenstehenden Zeichnung) verwenden.



In dieser Zeichnung haben wir willkürlich y auf 0 gesetzt. Dadurch erhalten wir den in der Zeichnung dick schwarz hervorgehobenen Graphen der partiellen Funktion in x -Richtung. Deren Steigung wollen wir an der Stelle $[x_0, y_0]$ in x -Richtung ermitteln. Dazu ermitteln wir zum einen den zugehörigen Funktionswert $f(x_0, y_0)$ und zum anderen den Funktionswert $f(x, y_0)$ an der Stelle $[x, y_0]$, die – um es anschaulich auszudrücken – ein kleines Stück in x -Richtung entfernt ist. Mit dem Unterschied der beiden Funktionswerte $f(x_0, y_0) - f(x, y_0)$ und dem Abstand $x_0 - x$ können wir ein Steigungsdreieck zeichnen, das die Sekantensteigung im Punkt (x_0, y_0) in x -Richtung angibt. In der obigen Zeichnung ist dieses Steigungsdreieck schwarz gepunktet eingezeichnet. Das Steigungsdreieck ist also

$$\frac{f(x_0, y_0) - f(x, y_0)}{x_0 - x}$$

Nun lassen wir – wie bereits bekannt – den Punkt $[x, y_0]$ in x -Richtung immer näher an $[x_0, y_0]$ heranrücken und machen den Übergang vom Differenzen zum Differenzialquotienten für $x \rightarrow x_0$ und erhalten

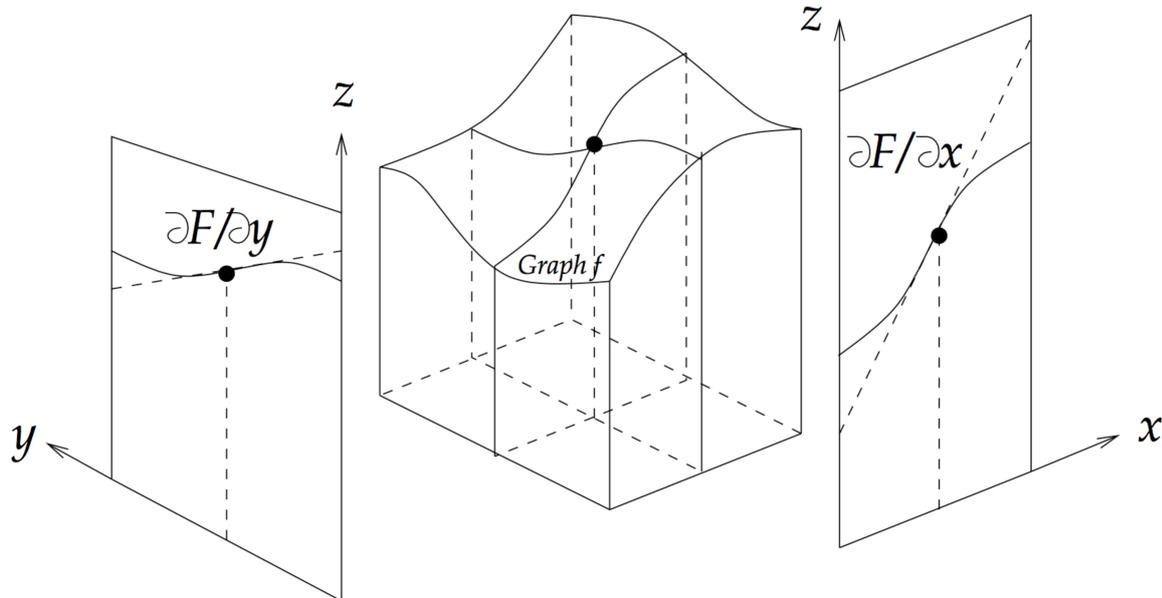
$$f_x(x_0, y_0) = \frac{\partial f}{\partial x}(x_0, y_0) = \lim_{x \rightarrow x_0} \frac{f(x_0, y_0) - f(x, y_0)}{x_0 - x}$$

Im Unterschied der partiellen Ableitung zur allgemeinen Ableitung schreibt man den Differenzialquotienten in der Form $\frac{\partial f}{\partial x}$ und macht dadurch die partielle Ableitung in x -Richtung deutlich.

Entsprechend geht man für die partielle Ableitung in y -Richtung vor und schreibt analog:

$$f_y(x_0, y_0) = \frac{\partial f}{\partial y}(x_0, y_0) = \lim_{y \rightarrow y_0} \frac{f(x_0, y_0) - f(x_0, y)}{y_0 - y}$$

Diese Definitionen bezeichnen die partiellen Ableitungen in x - und y -Richtung in einem bestimmten Punkt $[x_0, y_0]$. Nun kann man – bei fixiertem y_0 – jedem Punkt $[x, y_0]$ eine solche Ableitung zuordnen und erhält damit die partielle Ableitungsfunktion in x -Richtung. Ebenso kann man natürlich bei fixiertem x_0 jedem Punkt $[x_0, y]$ seine Ableitung in y -Richtung zuordnen und erhält so die partielle Ableitungsfunktion in y -Richtung.



Aus beiden obigen Definitionen wird bereits die schon kurz angesprochene Vorgehensweise beim partiellen Ableiten deutlich: Beim partiellen Ableiten nach x bleibt y_0 konstant, man lässt also den Punkt $[x, y_0]$ auf einer Parallelen zur x -Achse an den Punkt $[x_0, y_0]$ näher rücken. Beim partiellen Ableiten nach y bleibt x_0 konstant und entsprechend wandert der Punkt $[x_0, y]$ auf einer Parallelen zur x -Achse immer näher an den Punkt $[x_0, y_0]$ heran. Für die konkrete Durchführung der Ableitung heißt dies, dass man einmal x und das andere Mal y als Konstante betrachtet. Dies sei im Folgenden an einem ersten einfachen Beispiel der Funktion

$$f(x, y) = x^2 + 3x + 5 \cdot y - 5$$

gezeigt. Für die Ableitung nach x betrachtet man y als Konstante und erhält die partielle Ableitungsfunktion nach x :

$$f_x(x, y) = \frac{\partial f}{\partial x}(x, y) = 2x + 3$$

Für die Ableitung nach y betrachtet man andererseits x als Konstante und erhält die partielle Ableitungsfunktion nach y :

$$f_y(x, y) = \frac{\partial f}{\partial y}(x, y) = 5$$

Aus dem letzten Ergebnis wird deutlich, dass die Steigung in y -Richtung an jeder Stelle des Graphen immer 5 beträgt. Die Steigung in x -Richtung ist abhängig von der x -Koordinate des untersuchten Punktes.

Möchte man die Steigungen im Punkt $P:[3, 6]$ wissen, so setzt man diese Koordinaten in die Ableitungsgleichung nach x ein und erhält:

$$\frac{\partial f}{\partial x}(3, 6) = 2 \cdot 3 + 3 = 9$$

Damit beträgt die Steigung im Punkt P in x -Richtung 9, die Steigung in y -Richtung ist konstant und beträgt 5.

Führen wir hierzu noch ein zweites Beispiel durch anhand der Funktion

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

Wir leiten diese Funktion partiell nach x ab und erhalten

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

Leiten wir partiell nach y ab, so erhalten wir

$$f_y(x, y) = \frac{\partial f}{\partial y} = 2y + x.$$

Ermitteln Sie daraus die Steigungen in x - und y -Richtung im Punkt P : [3, 6].

2.5 Partielle Ableitung höherer Ordnung

Die erhaltenen partiellen Ableitungen sind ihrerseits Funktionen – und manchmal sind diese Funktionen nach wie vor Funktionen zweier Veränderlicher. Dies ist am oben durchgeführten zweiten Beispiel $f(x, y) = x^3 + xy + y^2$ so, da in beiden Ableitungen auch beide Funktionsvariablen x und y wieder auftauchen. Beim ersten dargestellten Beispiel war das nicht so, da war die partielle Ableitung nach x nur noch von x abhängig und die partielle Ableitung nach y war gar eine konstante Funktion.

In den Fällen, in denen die ersten partiellen Ableitungen wiederum Funktionen zweier Veränderlicher sind, kann man diese natürlich ein zweites Mal ableiten – das ist nun nichts wirklich Neues. Allerdings bekommen wir jetzt – ganz im Gegensatz zum Ableiten bei eindimensionalen Funktionen – deutlich mehr „Möglichkeiten“:

Die partiell nach x abgeleitete erste Ableitung können wir, wenn wir ein zweites Mal ableiten, partiell nochmals nach x und eben auch nach y ableiten. Ebenso ist es mit der ersten partiellen Ableitung nach y . Wir greifen zur Darstellung unser eben durchgeführtes Beispiel nochmals auf. Dort hatten wir für die partielle Ableitung nach x gefunden:

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

Diese Ableitung können wir sowohl nach x

$$f_{xx}(x, y) = \frac{\partial^2 f}{\partial x^2} = 6x$$

als auch nach y ableiten

$$f_{xy}(x, y) = \frac{\partial^2 f}{\partial x \partial y} = 1.$$

Ebenso können wir die oben gefundene partielle Ableitung nach y

$$f_y(x, y) = \frac{\partial f}{\partial y} = 2y + x$$

ein zweites Mal nach x

$$f_{yx}(x, y) = \frac{\partial^2 f}{\partial y \partial x} = 1$$

als auch nach y ableiten

$$f_{yy}(x, y) = \frac{\partial f^2}{\partial y^2} = 2.$$

In diesem Beispiel könnten wir nun nur noch $f_{xx}(x, y)$ einer dritten Ableitung unterziehen, aber prinzipiell ist es sicher vorstellbar, dass man – mit einer entsprechend hochgradigen Funktion beginnend – auch noch eine dritte und weitere Ableitungen bilden kann.² Auch dies ist nichts wesentlich Neues im Vergleich zu eindimensionalen Funktionen. Allerdings potenziert sich nun bei jedem Ableiten die Anzahl der zu bildenden partiellen Ableitungen. Im Rahmen dieser Veranstaltung werden wir in der Regel höchstens zweimal partiell ableiten.

Wenn man die eben erzeugten partiellen Ableitungen zweiter Ordnung betrachtet, so fällt auf, dass

$$f_{xy}(x, y) = \frac{\partial f^2}{\partial x \partial y} = 1 \quad \text{und} \quad f_{yx}(x, y) = \frac{\partial f^2}{\partial y \partial x} = 1$$

dasselbe Ergebnis haben. Dies könnte nun natürlich Zufall sein, allerdings garantiert uns der „Satz von Schwarz“³:

Ist eine Funktion $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ 2-mal stetig differenzierbar (d.h.: alle partiellen Ableitungen bis zur Ordnung 2 existieren und sind stetig) so ist die Differentiationsreihenfolge in den 2. Ableitungen unerheblich.

Diesen Satz werden wir nicht beweisen, Sie können aber im Weiteren einfach von dessen Gültigkeit ausgehen und sich damit die Arbeit etwas erleichtern.

Noch ein **Hinweis zur Schreibweise** der zweiten partiellen Ableitungen:

Die Notation $\frac{\partial^2 f}{\partial x \partial y}$ wird in der Literatur leider unterschiedlich interpretiert. An manchen

Stellen wird die im Nenner angegebene Ableitungsreihenfolge in umgekehrter Leserichtung interpretiert, also die erste Ableitung nach y und die zweite Ableitung nach x . Wir verwenden aber die Ableitungsreihenfolge in Leserichtung von links nach rechts, also die erste Ableitung nach x und die zweite Ableitung nach y . Letztendlich ist es ja ohnehin (meist) egal, da der Satz von Schwarz garantiert, dass beide Reihenfolgen im Ergebnis identisch sind.

Die Notation $f_{xy}(x, y)$ wird hingegen eindeutig aufgefasst.

² Auch eine e-Funktion oder trigonometrische Funktionen lassen sich beliebig oft ableiten!

³ hier in einer Formulierung für $\mathbb{R}^2 \rightarrow \mathbb{R}$

3 Methode der kleinsten Quadrate

Eine wichtige Anwendung partieller Differentiation ist die *Methode der kleinsten Quadrate*. Mit ihr ist es möglich, viele durch Messung gefundene Punkte $[x, y]$ durch einen Funktionsgraph anzunähern.

3.1 Idee

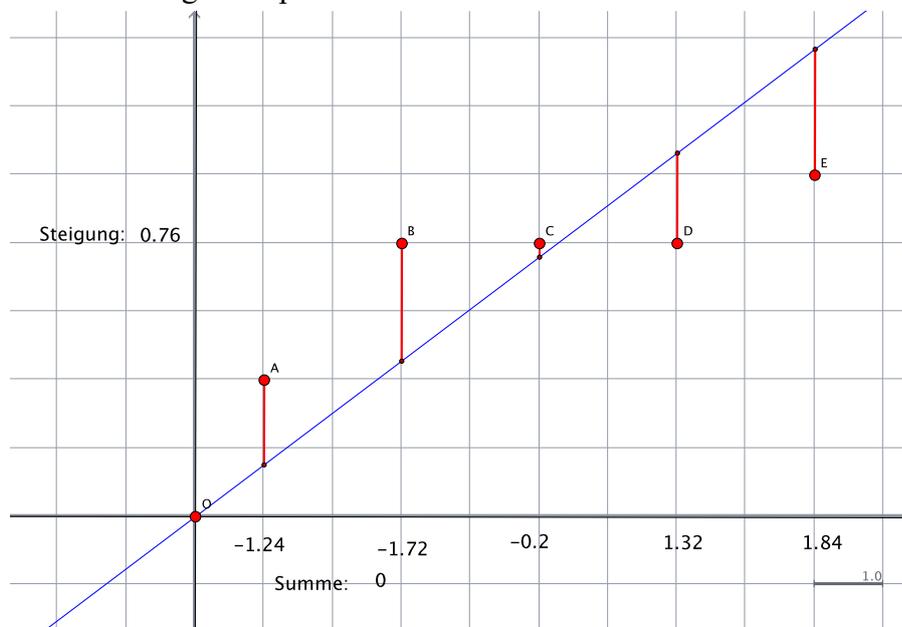
Anfangs des Jahres 1801 wurde der Zwergplanet Ceres von Guiseppe Piazzi entdeckt. Bevor dieser hinter der Sonne verschwand, konnten einige Bahndaten notiert werden. Carl Friedrich Gauß war damals bereits in der Lage, aus drei Beobachtungen die elliptische Bahn des neu entdeckten Himmelskörpers zu berechnen. Da deutlich mehr Bahnpunkte vorlagen – die ja alle einen gewissen Messfehler enthielten – wandte er die von ihm schon 1795 entwickelte Methode der kleinsten Quadrate an, um die Genauigkeit der Messungen zu erhöhen. Tatsächlich wurde Ceres im Dezember desselben Jahres an genau der Position wiederentdeckt, die Gauß vorhergesagt hatte.

Das Verfahren fußt auf einer Idee von Pierre Simon Laplace, aus vielen (fehlerbehafteten) Messwerten den „tatsächlichen“ Wert so festzulegen, dass sich die vorzeichenbehafteten Abstände der Messwerte zum tatsächlichen Wert zu Null aufaddieren.

Die grundsätzliche Idee des Verfahrens besteht darin, diskret gewonnene Messwerte (Wertepaare) durch einen Graphen und damit eine mathematische Funktion bestmöglich anzunähern. Für eine erste Annäherung gehen wir davon aus, dass in einem physikalischen Versuch die folgenden Messwertepaare gemessen wurden:

| | | | | | |
|---------|---|---|---|---|---|
| x_i : | 1 | 3 | 5 | 7 | 9 |
| y_i : | 2 | 4 | 4 | 4 | 5 |

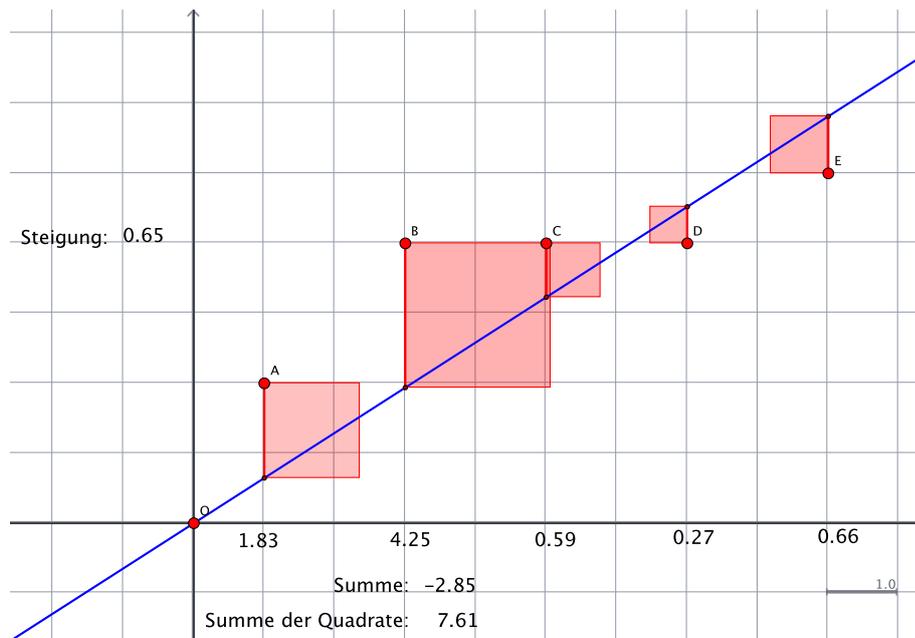
Wenn man zudem weiß, dass die zugrundeliegende Funktion eine Proportionalität ist, dann geht es darum, die Ursprungsgerade gerade so einzupassen, dass diese die gefundenen Messwerte bestmöglich repräsentiert:



Diese „bestmögliche Anpassung“ sollte nach der Idee von Laplace so geschehen, dass man von den gemessenen Punkten das Lot auf die x -Achse fällt und die Länge der Strecke vom Messpunkt bis zum Schnittpunkt des Lots mit dem Graphen als gerichtete Strecke misst. Die Entfernung des Punktes A vom Graphen beträgt in der obigen Darstellung $-1,24$, der Punkt D hingegen hat die Entfernung $1,3$. In einem dynamischen Geometriesystem kann man nun die

Ursprungsgerade so um den Ursprung drehen, dass die Summe der fünf Entfernungen genau Null ergibt. Laut dem obigen Bild ist dies für eine Geradensteigung von 0,76 der Fall.

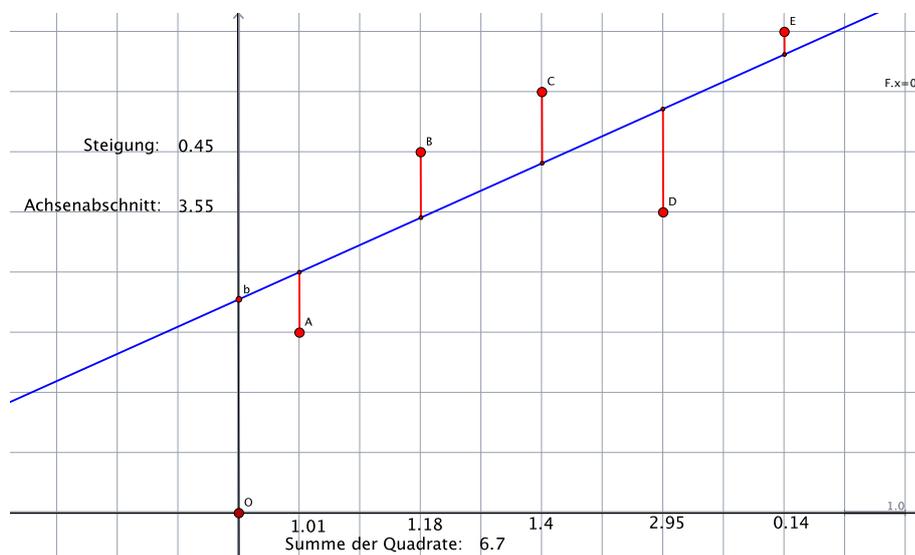
Gauß veränderte diese Laplace'sche Idee dahingehend, dass er nicht mit den vorzeichenbehafteten Entfernungen rechnete, sondern mit den (immer positiven) Quadraten dieser Entfernungen. Zum einen eliminiert man damit ein fehlerträchtiges Hantieren mit Vorzeichen ohne in die Bedrängnis des Rechnens mit Beträgen zu kommen, zum anderen wollte Gauß erreichen, dass größere Ablagen etwas stärker gewichtet werden. Die „bestmögliche“ Annäherung des Graphen an die Messpunkte ergibt sich nun genau dann, wenn die Summe der Abstandsquadrate minimal ist:



Für das obige Bild wurde diese optimale Lage wieder durch Probieren ermittelt, die hierbei erreichte Quadratsumme von 7,61 ist minimal. Es wird deutlich, dass dies bei einer etwas flacheren Geradensteigung von 0,65 der Fall ist. Rechnet man diesen Fall mit den vorzeichenbehafteten Entfernungen, so würde man nun ein Ergebnis von $-2,85$ erhalten.

3.2 Mathematisches Verfahren

Während man Ursprungsgeraden noch durch Probieren ermitteln kann, gilt dies für allgemein lineare oder quadratische Zusammenhänge nicht mehr, hierfür benötigen wir mathematische Verfahren.



Die im obigen Bild durch fünf Punkte A bis E dargestellten Messwerte sollen durch eine sogenannte Regressionsgerade in einen linearen Zusammenhang gebracht werden. Dabei wird eine solche Gerade durch die Funktionsgleichung $y = mx + b$ festgelegt. Um die tatsächliche Regressionsgerade festzulegen, müssen also deren Parameter m und b gefunden werden!

Dazu benötigen wir nach Gauß die Quadrate der Abstände der einzelnen Messwerte $A:[x_1, y_1]$ bis $E:[x_5, y_5]$ zur Geraden. Diese sind:

$$(y_1 - (mx_1 + b))^2, (y_2 - (mx_2 + b))^2, (y_3 - (mx_3 + b))^2, (y_4 - (mx_4 + b))^2, (y_5 - (mx_5 + b))^2$$

Die Summe dieser Abstandsquadrate hängt bei gegebenen Messwerten $[x_i, y_i]$ lediglich von der Wahl der Geradenparameter m und b ab. Anders formuliert: Die Summe dieser Abstandsquadrate ist eine Funktion von m und b :

$$A(m, b) = (y_1 - (mx_1 + b))^2 + (y_2 - (mx_2 + b))^2 + (y_3 - (mx_3 + b))^2 + (y_4 - (mx_4 + b))^2 + (y_5 - (mx_5 + b))^2$$

Kürzer und allgemein für n Messwerte formuliert:

$$A(m, b) = \sum_{i=1}^n (y_i - (mx_i + b))^2$$

Diese Abstandsquadrate summe soll minimal werden! Ganz in Anlehnung zur eindimensionalen Analysis und im Vorgriff auf ein folgendes Kapitel suchen wir das Minimum der Funktion $A(m, b)$, indem wir die beiden partiellen Ableitungen auf Null setzen. Dazu müssen wir zuerst die beiden partiellen Ableitungen von $A(m, b)$ nach m und nach b bilden. Das ist nicht so schwierig, wie es auf den ersten Blick aussehen mag, da die Ableitung einer Summe die Summe der Ableitungen ist. Leiten wir daher zunächst mal einen einzelnen Summanden $(y_i - (mx_i + b))^2$ nach m ab. Dazu müssen wir die Kettenregel anwenden:

$$2 \cdot (y_i - (mx_i + b)) \cdot -x_i = -2 \cdot (y_i - (mx_i + b)) \cdot x_i$$

Nun leiten wir nach b ab und erhalten:

$$-2 \cdot (y_i - (mx_i + b))$$

Damit gilt für die beiden partiellen Ableitungen:

$$\frac{\partial A}{\partial m} = -2 \cdot \sum_{i=1}^n ((y_i - (mx_i + b)) \cdot x_i)$$

und

$$\frac{\partial A}{\partial b} = -2 \cdot \sum_{i=1}^n (y_i - (mx_i + b))$$

Hinweis: Die Konstante -2 kann nach dem Distributivgesetz jeweils vor die Summe gezogen werden!

Diese beiden partiellen Ableitungen müssen gleich Null werden:

$$\frac{\partial A}{\partial m} = -2 \cdot \sum_{i=1}^n ((y_i - (mx_i + b)) \cdot x_i) = 0$$

und

$$\frac{\partial A}{\partial b} = -2 \cdot \sum_{i=1}^n (y_i - (mx_i + b)) = 0$$

Die letztgenannte partielle Ableitung nach b kann man noch etwas einfacher schreiben:

$$\frac{\partial A}{\partial b} = -2 \cdot \sum_{i=1}^n (y_i - mx_i - b) = 0$$

Wir schreiben die Summe des mittleren Terms für unser Beispiel mit den fünf Messwerten wieder ausführlich:

$$-2 \cdot [(y_1 - mx_1 - b) + (y_2 - mx_2 - b) + (y_3 - mx_3 - b) + (y_4 - mx_4 - b) + (y_5 - mx_5 - b)]$$

Wir lösen die Klammern auf und sortieren die Variablen:

$$-2 \cdot [y_1 + y_2 + y_3 + y_4 + y_5 - m \cdot (x_1 + x_2 + x_3 + x_4 + x_5) - (b + b + b + b + b)]$$

Bitte daran erinnern, dass dies für unser Beispiel mit fünf Messwerten zutrifft. Allgemein für n Messwerte gilt:

$$-2 \cdot [y_1 + y_2 + y_3 + \dots + y_n - m \cdot (x_1 + x_2 + x_3 + \dots + x_n) - (b + b + b + \dots + b_n)]$$

In Worten: Wir haben n verschiedene y_i und n verschiedene x_i sowie n (gleiche) b .

Die n gleichen b kann man einfacher als $n \cdot b$ schreiben.

Anstatt $y_1 + y_2 + y_3 + \dots + y_n$ können wir das arithmetische Mittel \bar{y} bilden und $n \cdot \bar{y}$ schreiben.

Genauso gehen wir bei den x -Werten vor und schreiben $n \cdot \bar{x}$:

Damit gilt:

$$-2 \cdot (n \cdot \bar{y} - n \cdot m \cdot \bar{x} - n \cdot b) = 0$$

Wir ziehen n vor die Klammer und dividieren durch $-2n$:

$$\bar{y} - m \cdot \bar{x} - b = 0$$

Das ergibt schließlich

$$b = \bar{y} - m \cdot \bar{x}$$

Die partielle Ableitung nach m wurde ebenfalls gleich Null gesetzt:

$$-2 \cdot \sum_{i=1}^n ((y_i - (mx_i + b)) \cdot x_i) = 0$$

Zunächst multiplizieren wir mit dem Faktor $-\frac{1}{2}$:

$$\sum_{i=1}^n ((y_i - (mx_i + b)) \cdot x_i) = 0$$

Dann multiplizieren wir aus:

$$\sum_{i=1}^n (y_i x_i - mx_i^2 - bx_i) = 0$$

Wir substituieren b mit dem eben gefundenen Term:

$$\sum_{i=1}^n (y_i x_i - mx_i^2 - (\bar{y} - m\bar{x}) x_i) = 0$$

Ausmultiplizieren:

$$\sum_{i=1}^n (y_i x_i - mx_i^2 - \bar{y} x_i + m\bar{x} x_i) = 0$$

Zusammenfassen:

$$\sum_{i=1}^n (y_i x_i - \bar{y} x_i - m(x_i^2 - \bar{x} x_i)) = 0$$

Trennen:

$$\sum_{i=1}^n (y_i x_i - \bar{y} x_i) = \sum_{i=1}^n (m(x_i^2 - \bar{x} x_i))$$

Faktor m vor die Summe ziehen:

$$\sum_{i=1}^n (y_i x_i - \bar{y} x_i) = m \cdot \sum_{i=1}^n (x_i^2 - \bar{x} x_i)$$

Dies ergibt für m :

$$m = \frac{\sum_{i=1}^n (y_i x_i - \bar{y} x_i)}{\sum_{i=1}^n (x_i^2 - \bar{x} x_i)} = \frac{\sum_{i=1}^n (y_i - \bar{y}) \cdot x_i}{\sum_{i=1}^n (x_i - \bar{x}) \cdot x_i}$$

Nun wird es etwas tricky: Da die Summen

$$\sum_{i=1}^n (y_i - \bar{y}) \quad \text{und} \quad \sum_{i=1}^n (x_i - \bar{x})$$

jeweils gleich Null sind, und damit auch die Summen

$$\sum_{i=1}^n (y_i - \bar{y}) \cdot \bar{x} \quad \text{sowie} \quad \sum_{i=1}^n (x_i - \bar{x}) \cdot \bar{x} \quad (\bar{x} \text{ ist ja de facto eine Konstante!})$$

subtrahieren wir diese im Zähler und im Nenner des Bruchs:

$$m = \frac{\sum_{i=1}^n (y_i - \bar{y}) \cdot x_i - \sum_{i=1}^n (y_i - \bar{y}) \cdot \bar{x}}{\sum_{i=1}^n (x_i - \bar{x}) \cdot x_i - \sum_{i=1}^n (x_i - \bar{x}) \cdot \bar{x}}$$

Das ist aber gleich:

$$m = \frac{\sum_{i=1}^n (y_i - \bar{y}) \cdot x_i + (y_i - \bar{y}) \cdot -\bar{x}}{\sum_{i=1}^n (x_i - \bar{x}) \cdot x_i + (x_i - \bar{x}) \cdot -\bar{x}}$$

Wir erhalten schließlich:

$$m = \frac{\sum_{i=1}^n (y_i - \bar{y}) \cdot (x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

Damit können wir die Geradensteigung m aus den Messwerten und deren Mittelwerten berechnen. Wissen wir m , dann können wir über die bereits eingangs gefundene Beziehung

$$b = \bar{y} - m \cdot \bar{x}$$

den Achsenabschnitt b der gesuchten Regressionsgerade berechnen.

3.3 Beispielrechnung

Auch hierfür rechnen wir ein einfaches Beispiel und verwenden dazu die in der obigen Abbildung dargestellten Messwerte. Um den Rechenaufwand erträglich zu halten, verwenden wir ein EXCEL-Sheet, in welchem die einzelnen Berechnungen (Mittelwertbildung, Differenzen, Produkte, ...) hinterlegt sind.

| | A | B | C | D | E | F | G |
|----|--------------------|-------|-------------------------|-------|-------------------------|---|-----------------------------|
| 1 | | | | | | | |
| 2 | | x_i | $x_i - x_{\text{quer}}$ | y_i | $y_i - y_{\text{quer}}$ | $(y_i - y_{\text{quer}}) * (x_i - x_{\text{quer}})$ | $(x_i - x_{\text{quer}})^2$ |
| 3 | | 1 | -4 | 3 | -2,8 | 11,2 | 16 |
| 4 | | 3 | -2 | 6 | 0,2 | -0,4 | 4 |
| 5 | | 5 | 0 | 7 | 1,2 | 0 | 0 |
| 6 | | 7 | 2 | 5 | -0,8 | -1,6 | 4 |
| 7 | | 9 | 4 | 8 | 2,2 | 8,8 | 16 |
| 8 | | | | | Summen: | 18 | 40 |
| 9 | $x_{\text{quer}}:$ | 5 | $y_{\text{quer}}:$ | 5,8 | | | |
| 10 | | | | | | m: | 0,45 |
| 11 | | | | | | b: | 3,55 |

3.4 Bestimmung mit Maxima

Die Methode der kleinsten Quadrate funktioniert nicht nur für lineare Funktionen. Es kann eine quadratische Funktion oder allgemein eine Potenzfunktion höheren Grades für eine experimentell überprüfte Situation modelliert werden oder eine Exponentialfunktion oder eine trigonometrische Funktion oder derlei mehr. Das Prinzip ist jedoch immer dasselbe: Der Graph der entsprechenden Funktion soll so an die gefundenen Messwerte angepasst werden, dass die Summe der Entfernungsquadrate minimal wird. Jede Funktionsklasse wird durch entsprechende Parameter eindeutig festgelegt, so beispielsweise eine allgemeine quadratische Funktion $ax^2 + bx + c$ durch die Parameter a , b und c , eine Funktion 5. Grades $ax^5 + bx^4 + cx^3 + dx^2 + ex + f$ benötigt schon 5 Parameter für eine eindeutige Festlegung. Während der prinzipielle Rechenweg derselbe bleibt, steigt damit der rechnerische Aufwand. Im letztgenannten Fall müssen fünf partielle Ableitungen gebildet und schließlich ein lineares Gleichungssystem mit 5 Variablen gelöst werden. Da es uns hier nicht um den Nachweis statistischer Verfahren geht, übergeben wir die konkrete Durchführung der Methode der kleinsten Quadrate getrost dem Computer und lassen Maxima den rechnerischen Aufwand für uns durchführen.

Maxima benötigt dafür eine Matrix mit den gefundenen Messwerten und die konkrete Funktion, welche wir hinter dem Zusammenhang vermuten. Zur Klarstellung muss Maxima auch noch wissen, welches die Funktionsvariablen sind (in der Regel x und y) und welche Bezeichnungen die Funktionsparameter haben.

Damit jedoch Maxima mit der Methode der kleinsten Quadrate überhaupt arbeiten kann, muss zunächst das entsprechende Paket mit

```
load(lsquares);
```

geladen werden. Die Funktion, die uns die gewünschten Parameter der zugrundeliegenden Funktion liefert, heißt `lsquares_estimates()`, sie hat die folgende Syntax:

```
lsquares_estimates(<matrix>, <fkt_var>, <fkt_gl>, <fkt_param>).
```

`matrix` ist eine Matrix mit den Messwerten, wobei eine Matrix in Maxima folgendermaßen erstellt wird:

```
M:matrix([x1, y1], [x2, y2], [x3, y3], ...)
```

Existiert bereits eine geschachtelte Liste mit den Messwerten, so geht man folgendermaßen vor:

```
messwerte: [[x1, y1], [x2, y2], [x3, y3], ...]
```

`M: apply(matrix, messwerte)`

`fkt_var` ist eine Liste mit den Funktionsvariablen, in aller Regel $[x, y]$,

`fkt_gl` ist die passende Funktionsgleichung, beispielsweise $y=m*x+b$ und

`fkt_param` ist eine Liste mit den Funktionsparametern, beispielsweise $[m, b]$.

4 Elastizität

4.1 Bogenelastizität

Um das Verhalten einer Funktion $y = f(x)$ bei Änderung der unabhängigen Variablen x zu analysieren, verwendet man die erste Ableitung. Sie beantwortet die Frage, um wie viele Einheiten sich y ändert, wenn x um eine (infinitesimale) Einheit steigt. Allerdings sind diese absoluten Änderungen von x und y häufig wenig aussagefähig, weil der Vergleichsmaßstab mit den ursprünglichen Größen fehlt.

Insbesondere die Wirtschaftswissenschaftler beschäftigen sich mit der Elastizität, deshalb sind die meisten Aufgaben auch in diesem Bereich angesiedelt und beschäftigen sich mit Preisen, Produktionsmengen und Absatzzahlen. Hierzu ein erstes allgemeines Beispiel:

Sie stellen ein Produkt her und verkaufen dies zum Stückpreis von 5 €. Nun möchten Sie mehr Geld mit diesem Produkt verdienen und planen einen höheren Verkaufspreis von 6 €. Dabei müssen Sie aber damit rechnen, dass Sie Kunden verlieren, weil denen Ihr Produkt nun zu teuer ist. Die Kunden kaufen nicht mehr Ihr Produkt, sondern eines der Konkurrenz oder sie kaufen Ihr Produkt nicht mehr so oft. In der Folge wird somit der Absatz Ihres Produkts zurück gehen. Die alles entscheidende Frage ist nun, mit welchem Absatzrückgang Sie rechnen müssen. Wenn Ihr Absatz nur geringfügig zurück geht, dann lohnt sich die Preissteigerung für Sie. Bricht er aber stark ein, so haben Sie unterm Strich mit der Preiserhöhung nichts gewonnen

Wir verdeutlichen dieses Problem an einem einfachen aber konkreten Beispiel:

In einem Supermarkt kosten 250 g Butter 1 €, pro Woche werden 2000 Stück verkauft. Nach einer Preiserhöhung um 1 € werden nur noch 500 Stück verkauft.

Derselbe Supermarkt verkauft Champagner, die Flasche für 20 € und davon 100 Flaschen pro Woche. Auch der Champagnerpreis wird um 1 € erhöht, wodurch der wöchentliche Verkauf auf 90 Flaschen sinkt.

Bei beiden Artikeln wurde der Preis um einen Euro erhöht. Der Butterabsatz sank um 1500 Stück, d.h., die Ableitung an dieser Stelle beträgt -1500 .

Der Champagnerabsatz sank um 10 Flaschen, also beträgt die Ableitung dort -10 .

Der Nachfragerückgang ist somit bei der Butter deutlich stärker. Allerdings hat sich der Butterpreis auch verdoppelt, während der Champagnerpreis nur um 5% gestiegen ist.

Der Sektumsatz nur um 10 Stück, der Buttereinsatz ist hingegen um 1500 Stück zurückgegangen und damit um das 150-fache. Der prozentuale Rückgang des Butterumsatzes beträgt 75% und derjenige des Sektumsatzes 10%. Relativ betrachtet liegen die Umsatzrückgänge damit enger beieinander als es der absolute Vergleich ergibt.

Es kommt also für einen aussagekräftigen Vergleich auf die *relative* Preis- und Mengenänderung an!

Für die Ableitung gehen wir aus vom Differenzenquotienten

$$\frac{\Delta y}{\Delta x} = \frac{y_2 - y_1}{x_2 - x_1},$$

dieser berechnet jedoch die absoluten Änderungen des Preises und der Nachfragemenge. Diese absoluten Änderungen müssen wir ins Verhältnis setzen zu den Ausgangspreisen bzw.

Ausgangsmengen, dann erhalten wir die relativen Änderungen:

$$\epsilon_{y,x} = \frac{\frac{\Delta y}{y}}{\frac{\Delta x}{x}}$$

Die Elastizität ist somit der Quotient aus der relativen Änderung der abhängigen Variablen (y ist von x abhängig) und der relativen Änderung der unabhängigen Variablen (x kann im Definitionsbereich frei gewählt werden).

Diesen Quotienten aus den relativen Änderungen nennt man die *Elastizität*. Sie gibt an, um wieviel Prozent der y -Wert steigt, wenn der x -Wert um 1% erhöht wird. Rechnen wir mal mit den obigen Werten. Für Butter gilt:

$$\varepsilon_{\text{Butter}} = \frac{\frac{-1500}{2000}}{\frac{1}{1}} = -\frac{3}{4} = -0,75$$

Steigt der Butterpreis um 1 Prozent, so sinkt die Nachfrage um 0,75 Prozent.

$$\varepsilon_{\text{Champagner}} = \frac{\frac{-10}{100}}{\frac{1}{20}} = -\frac{10}{1} = -2$$

Beim Champagner führt eine einprozentige Preissteigerung hingegen zu einem Nachfragerückgang von 2 Prozent.

Die hier berechnete Elastizität mit den relativ großen Deltas nennt man die *Bogenelastizität*. Sie wird verwendet, wenn – wie in unserem Fall – die zugrundeliegende Funktion nicht bekannt oder kompliziert zu berechnen ist.

4.2 Punktelastizität

Wenn wir nun wieder unsere infinitesimale Betrachtungsweise einführen und Δx und Δy im Grenzprozess gegen 0 gehen lassen, dann erhalten wir den Differentialquotient $\frac{dy}{dx}$.

Für die Elastizität ergibt sich daraus dann

$$\varepsilon_{y,x} = \frac{\frac{dy}{dx}}{\frac{y}{x}} = \frac{dy}{dx} \cdot \frac{x}{y}$$

und weil $\frac{dy}{dx}$ die Ableitung $f'(x)$ ist, erhalten wir schließlich

$$\varepsilon_{y,x} = f'(x) \cdot \frac{x}{y}.$$

Diese Elastizität nennt man die Punktelastizität, weil sie – aufgrund der infinitesimalen Betrachtungsweise – die Elastizität genau für einen Punkt berechnet. Auch hierzu ein Beispiel: Gegeben ist die Preis-Absatz-Funktion:

$$y(p) = 600 - 3p$$

Deren Ableitung lautet

$$y'(p) = -3.$$

Damit erhalten wir die Elastizitätsfunktion

$$\varepsilon_{y,p} = -3 \cdot \frac{p}{600 - 3p} = -\frac{p}{200 - p}.$$

Welche Elastizität hat die Funktion bei einem Preis von 100?

$$\varepsilon_{y,100} = -\frac{100}{200-100} = -1$$

Steigt der Preis von 100 € um 1 Prozent, so sinkt die Nachfrage um 1 Prozent.

Wie sieht es bei einem Ausgangspreis von 200 € aus?

$$\varepsilon_{y,200} = -\frac{200}{200-200} = -\infty$$

Die Wirtschaftswissenschaftler nennen eine Elastizität, die gegen Unendlich geht, *vollkommen elastisch*. Dabei ist die relative Änderung von y – selbst bei kleinsten Änderungen von x unendlich hoch.

Wir rechnen noch einen Ausgangspreis von 300 €:

$$\varepsilon_{y,300} = -\frac{300}{200-300} = 3$$

Bei einem Ausgangspreis von 300 € steigt bei einer Preisänderung um 1 Prozent die Nachfrage um 3 Prozent. Ohne über die Realitätsnähe dieses einfachen Beispiels zu diskutieren, halten wir fest, dass die Elastizität ganz unterschiedliche Werte annehmen kann. Die Wirtschaftswissenschaftler haben daher verschiedene Kategorien definiert:

| Wert von $\varepsilon_{y,x}$ | Bezeichnung | Reaktion von y |
|------------------------------------|--------------------------------|---|
| $\varepsilon = 0$ | y ist vollkommen unelastisch | y reagiert nicht auf eine Änderung von x |
| $0 < \varepsilon < 1$ | y ist unelastisch | y ändert sich relativ weniger stark als x |
| $ \varepsilon = 1$ | y ist proportional elastisch | Die relative Änderung von y ist gleich der von x |
| $ \varepsilon > 1$ | y ist elastisch | y ändert sich relativ stärker als x |
| $ \varepsilon \rightarrow \infty$ | y ist vollkommen elastisch | Die relative Änderung von y ist unendlich hoch, selbst bei kleinsten Änderungen von x |

4.3 Funktionsgleichung der Elastizität

Für die Punktelastizität ε einer Funktion $f(x)$ haben wir oben den Zusammenhang

$\varepsilon_{y,x} = f'(x) \cdot \frac{x}{y}$ hergeleitet. Im vorhergehenden Beispiel wurde deutlich, dass die

Punktelastizität in aller Regel an jeder Stelle x der Funktion eine andere ist! Dies bedeutet, dass wir **jeder** Stelle x **eindeutig** eine Elastizität **zuordnen** können und dies bedeutet schließlich, dass es zu jeder Funktion eine Elastizitätsfunktion gibt. Diese Elastizitätsfunktion können wir einfach aus der Beziehung für die Punktelastizität herleiten, wenn wir dort y durch $f(x)$ ersetzen:

$$\varepsilon_{y,x} = \frac{f'(x) \cdot x}{f(x)}$$

Eigentlich haben wir inhaltlich überhaupt nichts verändert, daher ist auch die Notation gleich geblieben. Mit der Elastizität ist es genau wie mit der Ableitung: Unter dem Begriff *Ableitung* versteht man zum einen die Tangentensteigung an einer Stelle, zum anderen aber auch die Zuordnung dieser Steigung zu jeder Stelle x und damit die Ableitungsfunktion. Genau so gibt es die Elastizität an einer Stelle x und die Elastizitätsfunktion als Zuordnung der einzelnen Elastizitätswerte zu jeder Stelle x .

Machen wir hierzu ein einfaches Beispiel:

Gegeben sei die Funktion $f(x) = 2x^3 - 4x^2 + 3x - 5$

Ihre Ableitung lautet $f'(x) = 6x^2 - 8x + 3$. An der Stelle $x = 3$ beträgt die Ableitung

$f'(3) = 6 \cdot 3^2 - 8 \cdot 3 + 3 = 33$. Damit und dem Funktionswert $f(3) = 2 \cdot 3^3 - 4 \cdot 3^2 + 3 \cdot 3 - 5 = 22$ an der Stelle 3 können wir die Elastizität eben an dieser Stelle $x = 3$ bestimmen:

$$\varepsilon_{y,3} = 33 \cdot \frac{3}{22} = \frac{99}{22} = \frac{9}{2}$$

Die zugehörige Elastizitätsfunktion erhalten wir durch einfaches Einsetzen:

$$\varepsilon_{y,x} = \frac{f'(x) \cdot x}{f(x)} = \frac{(6x^2 - 8x + 3) \cdot x}{2x^3 - 4x^2 + 3x - 5} = \frac{6x^3 - 8x^2 + 3x}{2x^3 - 4x^2 + 3x - 5}$$

Man multipliziert im Zähler die Ableitung mit x und trägt im Nenner die Ausgangsfunktion $f(x)$ ein.

4.4 Partielle Elastizität

In höherdimensionalen Funktionen kann und muss die Elastizität für jede unabhängige Variable getrennt betrachtet werden. In dem uns selbst gesteckten Rahmen zweidimensionaler Funktionen, bei welchen sowohl x als auch y Auswirkungen auf den Wert der abhängigen Variablen z haben, gibt es somit die Elastizitäten $\varepsilon_{z,x}$ und $\varepsilon_{z,y}$. Da – wie oben bereits hergeleitet – die Punktelastizität über die Ableitung bestimmt werden kann, erhalten wir *partielle Elastizitäten*, die folgendermaßen bestimmt werden können:

$$\varepsilon_{z,x} = \frac{\partial f}{\partial x} \cdot \frac{x}{z} \quad \text{und} \quad \varepsilon_{z,y} = \frac{\partial f}{\partial y} \cdot \frac{y}{z}$$

Die partielle Elastizitäten $\varepsilon_{z,x}$ und $\varepsilon_{z,y}$ geben an, um wieviel Prozent sich der Funktionswert ändert, wenn sich x bzw. y um 1% ändert und die jeweils andere Variable unverändert bleibt.

Hierzu ein einfaches Beispiel: Gesucht sind die partiellen Elastizitäten der Funktion

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

Wir gehen gleich allgemein vor und bestimmen die Elastizitätsfunktionen:

$$\varepsilon_{z,x} = \frac{\partial f}{\partial x} \cdot \frac{x}{z} = \frac{\partial f}{\partial x} \cdot \frac{x}{f(x,y)} = x \cdot \frac{f_x(x,y)}{f(x,y)}$$

$$\varepsilon_{z,y} = \frac{\partial f}{\partial y} \cdot \frac{y}{z} = \frac{\partial f}{\partial y} \cdot \frac{y}{f(x,y)} = y \cdot \frac{f_y(x,y)}{f(x,y)}$$

Die partiellen Elastizitäten nach x bzw. y berechnen sich somit aus dem Produkt der jeweiligen Variablen x bzw. y mit dem Quotienten der jeweiligen partiellen Ableitung nach x bzw. y geteilt durch die Funktion.

Auch hierzu rechnen wir ein einfaches Beispiel: Die Produktionsmenge $f(x, y)$ ist von zwei Eingangsgrößen x und y abhängig und berechnet sich nach der Funktion

$$f(x, y) := x^2 + y^2 \cdot \ln\left(\frac{x}{y}\right).$$

Üblicher Weise werden als Einsatzmengen x und y je 10 Einheiten verwendet. Wie verändert sich die Produktion näherungsweise, wenn der Einsatz von x um 2% gesteigert wird?

Gesucht ist die partielle Elastizität in x -Richtung. Diese errechnet sich nach

$$\varepsilon_{z,x} = x \cdot \frac{f_x(x,y)}{f(x,y)}.$$

Wir benötigen somit den Wert der ersten partiellen Ableitung nach x und den Funktionswert an der Stelle $(10, 10)$:

Die erste partielle Ableitung nach x lautet

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

ihr Wert an der Stelle $(10, 10)$ ist 30.

Der Funktionswert an der Stelle $(10, 10)$ ist 100. Damit erhalten wir für die gesuchte Elastizität:

$$\varepsilon_{z,x} = 10 \cdot \frac{30}{100} = 3$$

Dieses Ergebnis besagt, dass bei einer einprozentigen Erhöhung von x die Produktionsmenge um drei Prozent steigt. Steigert man x gar um zwei Prozent und damit das Doppelte, so kann eine Produktionssteigerung um sechs Prozent erwartet werden.

Weiter in der Aufgabe:

Um wieviel verringert sich die Produktion näherungsweise, wenn die Einsatzmenge x konstant bei 10 Einheiten bleibt und die Einsatzmenge y sich von 10 Einheiten um 12% erhöht?

Nun ist die partielle Elastizität in y -Richtung gesucht. Wir bestimmen analog die partielle Ableitung nach y

$$f_y(x,y) = 2y \cdot \ln \frac{x}{y} - y,$$

ihr Wert an der Stelle $(10, 10)$ ist -10 .

Den Funktionswert an der fraglichen Stelle haben wir bereits mit 100 bestimmt. Damit können wir die partielle Elastizität in y -Richtung bestimmen:

$$\varepsilon_{z,y} = 10 \cdot \frac{-10}{100} = -1$$

Dies bedeutet, dass bei einer Erhöhung des Einsatzguts y und gleichbleibendem x um ein Prozent ein Produktionsrückgang um 1% erwartet werden muss. Mit einer Erhöhung von y um 12% geht somit ein Rückgang der Produktion um 12% einher.

5 Gradient und Richtungsableitung

5.1 Der Gradient

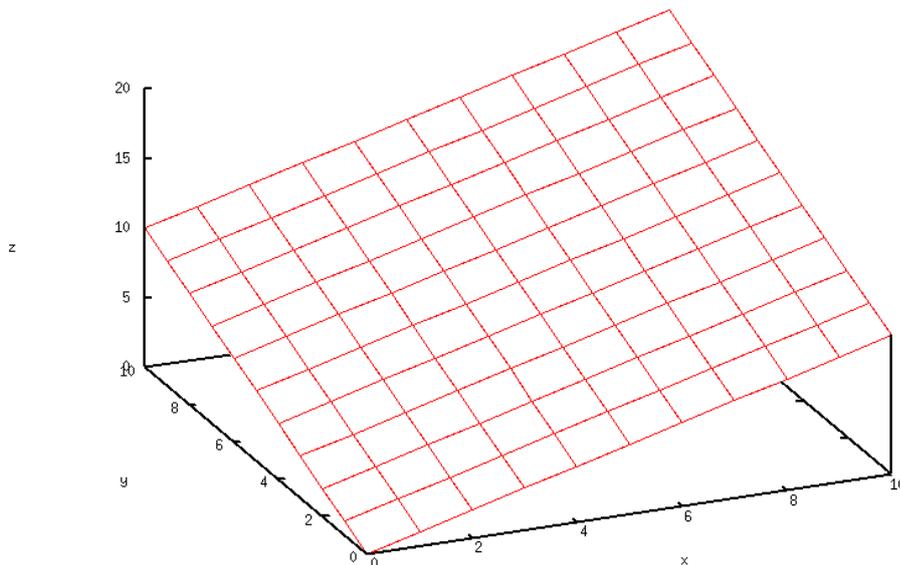
Wir greifen erneut auf die eingangs schon diskutierte einfache Funktion $f(x,y) = x + y$ zurück. Deren partielle Ableitung nach x lautet:

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

Die partielle Ableitung nach y lautet letztlich ebenso

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

Bei so viel Trivialität sollten wir uns genau überlegen, was wir denn hier bestimmt haben: Wenn die partielle Ableitung in x -Richtung gleich 1 ist, dann bedeutet dies doch, dass die Steigung unseres Flächengraphen in x -Richtung 1 beträgt. Ebenso steigt der Graph in y -Richtung um 1 an. Genau dies gibt ja auch der bereits weiter oben dargestellte 3D-Graph der Funktion wieder:

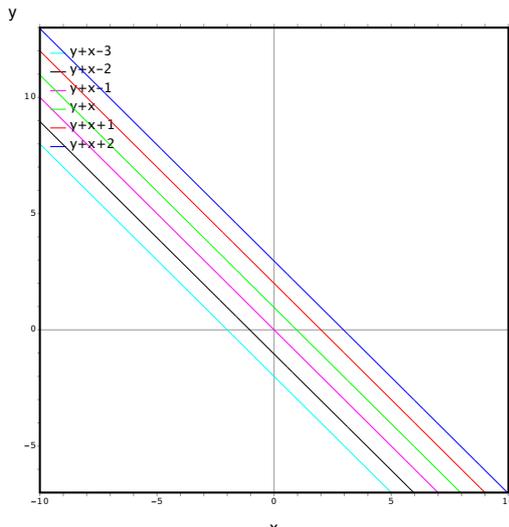


Laufen wir in x -Richtung, so haben wir die Steigung 1, dies gilt genauso, wenn wir in y -Richtung laufen.

Es wurde schon kurz angesprochen, dass die Steigung in einem Punkt auf einer (nicht unbedingt ebenen) Fläche richtungsabhängig ist: Schaut man beim Bergsteigen direkt zum Gipfel, so hat man eine relativ große positive Steigung vor sich, schaut man hinter sich ins Tal, so ist diese Steigung negativ. Bei einem Rundumblick wird die Steigung zwischen diesen beiden Extremen variieren. Anschaulich klar sollte auch sein, dass die Steigung gleich Null ist, wenn wir in Richtung der Höhenlinien schauen.

Wir fragen uns nun, in welche Richtung die durch die Funktion $f(x,y) = x + y$ festgelegte Ebene die größte Steigung hat. Gibt es eine größere Steigung als 1 (die Steigung in x - als auch in y -Richtung)?

Intuitiv wird man vermuten, dass die größte Steigung des Funktionsgraphen „genau dazwischen“ liegt – also derart, dass deren Richtungs-Projektion in die xy -Ebene dort die erste Winkelhalbierende ergibt. Dies kann man leicht überprüfen, wenn man einige Höhenlinien dieser Funktion bestimmt. Dies haben wir bereits weiter oben gemacht und dort das folgende Bild erhalten:



Die Höhenlinien verlaufen alle in Richtung der zweiten Winkelhalbierenden, sie haben damit jeweils die Steigung -1 in der xy -Ebene. Vom Wandern und Bergsteigen wissen wir, dass die größte Steigung immer senkrecht zu den Höhenlinien verläuft und damit – wie in diesem Beispiel bereits vermutet – in Richtung der ersten Winkelhalbierenden.

Die Richtung der größten Steigung haben wir aus den partiellen Ableitungen erhalten: Die Ableitung (Steigung) in x -Richtung ist 1 , die Ableitung (Steigung) in y -Richtung ist ebenfalls 1 und die größte Steigung haben wir, wenn wir **einen** Schritt in x - und **einen** Schritt in y -Richtung laufen.

Wir überprüfen unsere Vermutung sogleich an einem weiteren nur leicht veränderten Beispiel anhand der Funktion $f(x,y) = 2x + y$. Deren partiellen Ableitungen lauten:

$$f_x(x,y) = \frac{\partial f}{\partial x} = 2 + 0 = 2$$

und (nach wie vor)

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

Jetzt ist die Steigung in x -Richtung doppelt so groß wie die Steigung in y -Richtung. Die Richtung der größten Steigung liegt wieder „irgendwo dazwischen“, diesmal aber sicher eher zur x -Richtung hin orientiert.

Höhenlinien dieses Graphen sind ...

$$\text{Höhenlinie in der Höhe 1: } 2x + y = 1 \Rightarrow y = -2x + 1$$

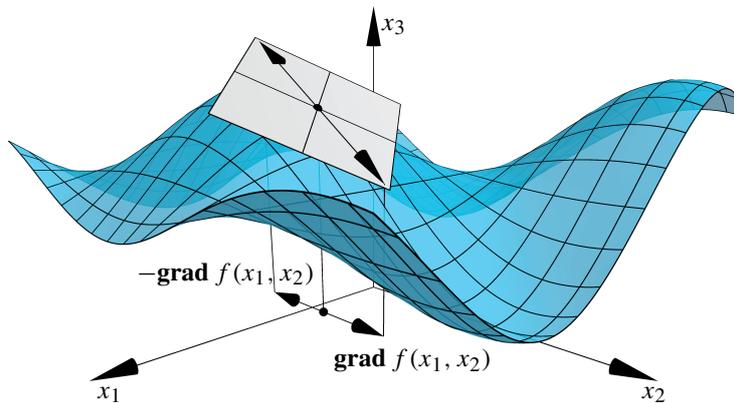
$$\text{Höhenlinie in der Höhe 3: } 2x + y = 3 \Rightarrow y = -2x + 3$$

...

Es wird schnell deutlich, dass alle Höhenlinien in der xy -Ebene mit der Steigung -2 verlaufen. Die Senkrechte dazu und damit die Richtung der größten Steigung des Funktionsgraphen muss dann der negative Kehrwert sein, also $\frac{1}{2}$.

Wenn wir unsere obige Überlegung der Ermittlung der größten Steigung aus den partiellen Ableitungen wieder aufgreifen, dann wird dies bedeuten, dass man die größte Steigung findet, indem man **zwei** Schritte in x - und **einen** Schritt in y -Richtung geht. In der xy -Ebene steht dies aber genau für die Steigung $\frac{1}{2}$. Damit ist unsere Überlegung an einem weiteren Beispiel gefestigt und wir halten (ohne Beweis) fest:

Die partiellen Ableitungen in x - und y -Richtung legen – als Vektor gedeutet – die Richtung der größten Steigung fest.



Diese Überlegungen haben wir anhand einfacher Funktionen und Graphen (Ebenen) angestellt. Sie bleiben aber auch gültig, wenn wir einen gebogenen, gewellten, gebirgigen, ... Graphen haben. Die Ableitung ist in einem solchen Fall ja wieder eine Ebene – nämlich die Tangentialebene an eben dem Punkt, für den wir die Ableitungen erstellt haben – so dass unsere Überlegungen auch in diesem Fall korrekt sind. Man definiert schließlich:

Der Vektor $\text{grad } f(x,y) := \text{grad } f(x,y) := \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}$ heißt der Gradient von f in (x,y) .

Hinweis: In vielen Veröffentlichungen wird dieser Spaltenvektor aus Platzgründen auch als Zeilenvektor geschrieben, außerdem ist für die Bezeichnung des Gradienten auch das Zeichen ∇^4 üblich:

$$\nabla f(x,y) := \left[\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right]$$

Übertragen auf eine Wanderkarte zeigt der Gradient immer die steilste Richtung bergauf an. Die Länge des Gradienten gibt an, wie steil der Berg ist!

Die Bestimmung des Gradienten kann in Maxima über die folgende Funktion erreicht werden. Das Besondere an dieser Realisation besteht darin, dass der Gradient einer global definierten Funktion $f(x,y)$ allgemein errechnet wird oder aber – so man als Parameter die x - und y -Koordinate eines interessierenden Punktes angibt – der Gradient konkret für den angegebenen Punkt berechnet wird.

```
gradient([P]) := block(
  [ab_x, ab_y],
  ab_x:diff(f(x,y), x, 1),
  ab_y:diff(f(x,y), y, 1),
  if length(P)>0 then
  (
    ab_x:subst(P[1][1], x, ab_x),
    ab_x:subst(P[1][2], y, ab_x),
    ab_y:subst(P[1][1], x, ab_y),
```

⁴ gesprochen „nabla“, geht zurück auf nàbla (gr.) bzw. nablium (lat.), womit eine antike Harfe bezeichnet wurde.

```

    ab_y:subst (P[1][2],y,ab_y)
),
[ab_x,ab_y]);

```

5.2 Richtungsableitung

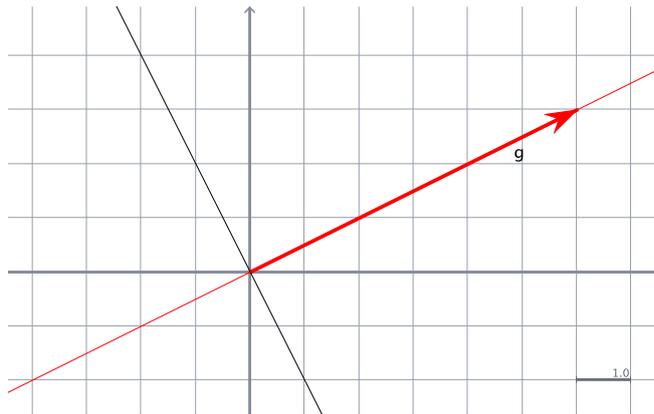


Mit den partiellen Ableitungen haben wir eine Möglichkeit gefunden, die Steigungen des Graphen in x - und y -Richtung an beliebigen Stellen zu ermitteln – aber eben nur die Steigungen in x - oder in y -Richtung. Wünschenswert wäre daher eine Methode, mit welcher die Steigung in jeder beliebigen Richtung ermittelt werden kann und genau hierfür ist der eben behandelte Gradient der Schlüssel zum Erfolg!

Vergegenwärtigen Sie sich zunächst noch einmal, dass die Ableitung einer Funktion zweier Veränderlicher in einem Punkt die Tangentialebene an den Funktionsgraphen im entsprechenden (Berühr-)Punkt ist. Mit den beiden partiellen Ableitungen haben wir die Steigungen dieser Tangentialebene in x - und in y -Richtung ermittelt und damit gleichzeitig auch die Lage dieser Tangentialebene eindeutig beschrieben. (Eine Ebene kann im Raum durch drei Punkte oder eben einen Punkt und zwei Richtungen festgelegt werden!)

Nun wollen wir die Steigung dieser Tangentialebene in jeder beliebigen Richtung angeben. Stellen Sie sich hierzu vor, dass Sie beim Wandern oder Bergsteigen die Steigung für eine beliebige Richtung (bergauf, bergab, schräg zum Berg, ...) wissen wollen – wobei Sie aber nur die Steigungen in zwei definierte Richtungen kennen, die zudem senkrecht aufeinander stehen. Bei der Lösung dieses Problems lassen wir uns von den Methoden der linearen Algebra

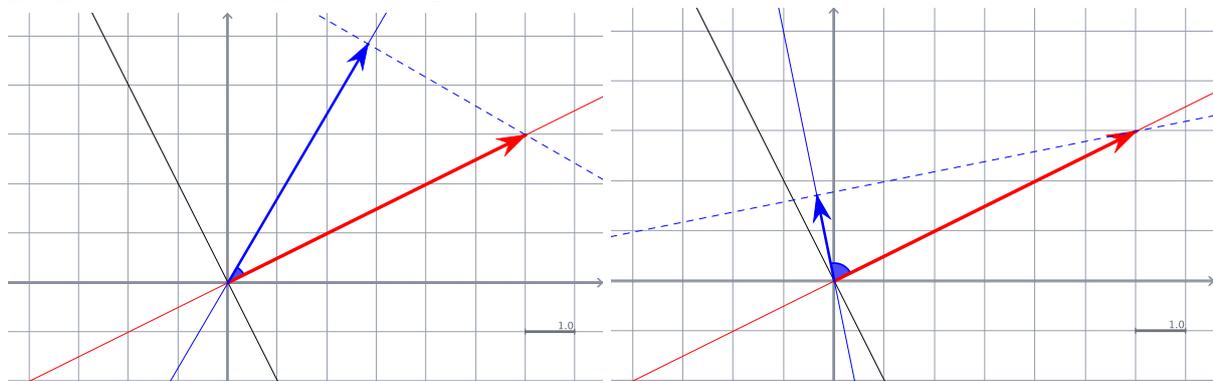
bzw. der analytischen Geometrie leiten und greifen dafür die Idee des eben behandelten Gradienten wieder auf.



Der Gradient einer Funktion ist ein Vektor, der in Richtung der größten Steigung der Tangentialebene weist und dessen Betrag (Länge) genau die Größe dieser Steigung angibt. Seine x -Komponente ist die partielle Ableitung in x -Richtung und seine y -Komponente die partielle Ableitung in y -Richtung. Für die Darstellung dieses Gradienten können wir uns somit in die xy -Ebene zurückziehen.

Im obenstehenden Bild ist der Gradient $[6, 3]$ dargestellt. Wenn wir auf dem Graphen in Richtung $[6, 3]$ blicken, dann schauen wir genau in Richtung der größten Steigung und die Größe dieser Steigung ist genau der Betrag des Gradienten $\sqrt{6^2 + 3^2} = \sqrt{45} = 3\sqrt{5}$.

Wenn wir in Richtung $[-2, -1]$ blicken, dann haben wir dort doch das größte „Gefälle“, also genau die negative Steigung $-3\sqrt{5}$. Und zwischen diesen beiden entgegengesetzten Richtungen variiert die Steigung vom größten (positiven) zum kleinsten (negativen) Wert. Anschaulich klar sollte auch noch sein, dass die Steigung in den Richtungen $[-1, 2]$ und $[1, -2]$ genau 0 ist. In diesem Fall blicken wir genau entlang der Höhenlinien. Wenn wir uns – ausgehend vom Blick in Richtung größte Steigung – einmal um 360° drehen, so wird der Betrag des Gradienten immer kleiner, er ist in senkrechter Blickrichtung gleich 0, um weiter auf seine Gegenzahl abzunehmen. Dann nimmt er wieder kontinuierlich zu. Genau dieses Verhalten zeigt auch die Cosinus-Funktion und in der Tat spielt der Cosinus bei der Ermittlung der Steigung in einer beliebigen Richtung eine wesentliche Rolle:



In den obigen Darstellungen ist der rote Pfeil der Gradient, die dünn blau gezeichnete Gerade gibt die Richtung an, für welche die Richtungsableitung ermittelt werden soll. Der Betrag der jeweiligen Richtungsableitung (blauer Pfeil) wird geometrisch durch die senkrechte Projektion der Pfeilspitze des Gradienten auf die geforderte Richtung ermittelt. In diesem rechtwinkligen Dreieck kann bei bekannter Länge des roten Pfeils die Länge des blauen Pfeils über den Cosinus des eingeschlossenen Winkels berechnet werden.

Da die Bestimmung des eingeschlossenen Winkels aufwändig ist und der Gradient ohnehin schon als Vektor vorliegt, bedienen wir uns der Möglichkeiten der analytischen Geometrie und ermitteln die Größe der Richtungsableitung über das *Skalarprodukt* zweier Vektoren. Geben wir neben dem Gradienten auch die gewünschte Richtung als Vektor an, so können wir die gesuchte Richtungsableitung einfach als Skalarprodukt beider Vektoren ermitteln. Wir müssen dabei allerdings darauf achten, den Richtungsvektor als *Einheitsvektor* anzugeben.

Berechnen wir beispielhaft die in den beiden obigen Bildern dargestellten Richtungsableitungen, so halten wir zunächst fest, dass der Gradient der Funktion (roter Pfeil) in beiden Fällen

gleich $\vec{g} = \begin{bmatrix} 6 \\ 3 \end{bmatrix}$ ist.

Die Richtung im linken Bild kann durch $\begin{bmatrix} 3 \\ 5 \end{bmatrix}$ angegeben werden. Diesen Richtungsvektor \vec{a} müssen wir aber zuerst auf den zugehörigen Einheitsvektor normieren:

$$|\vec{a}| = \sqrt{3^2 + 5^2} = \sqrt{34}$$

Der normierte Richtungsvektor ist dann $\vec{e}_a = \frac{\vec{a}}{|\vec{a}|} = \frac{1}{\sqrt{34}} \begin{bmatrix} 3 \\ 5 \end{bmatrix}$.

Das Skalarprodukt ergibt dann $\vec{e}_a \cdot \vec{g} = \frac{1}{\sqrt{34}} \begin{bmatrix} 3 \\ 5 \end{bmatrix} \cdot \begin{bmatrix} 6 \\ 3 \end{bmatrix} = \frac{1}{\sqrt{34}} \cdot (3 \cdot 6 + 5 \cdot 3) \approx 5,66$.

Im rechten Bild kann die Richtung mit $\begin{bmatrix} -1 \\ 5 \end{bmatrix}$ angegeben werden. Der zugehörige Einheits-

vektor ist dann $|\vec{a}| = \sqrt{-1^2 + 5^2} = \sqrt{26}$ bzw. $\vec{e}_a = \frac{\vec{a}}{|\vec{a}|} = \frac{1}{\sqrt{26}} \begin{pmatrix} -1 \\ 5 \end{pmatrix}$.

Für die Richtungsableitung ergibt sich

$$\vec{e}_a \cdot \vec{g} = \frac{1}{\sqrt{26}} \begin{bmatrix} -1 \\ 5 \end{bmatrix} \cdot \begin{bmatrix} 6 \\ 3 \end{bmatrix} = \frac{1}{\sqrt{26}} \cdot (-1 \cdot 6 + 5 \cdot 3) \approx 1,76.$$

Schließlich noch eine Aufgabe im vollständigen Zusammenhang mit beispielhafter Lösung: Gesucht sind die Richtungsableitungen der Funktion $f(x,y) = x + e^y$ im Punkt $P:[3, 4]$ in den Richtungen $\vec{a} = [2, 4]$, $\vec{b} = [-1, -3]$ und $\vec{c} = [55, -1]$.

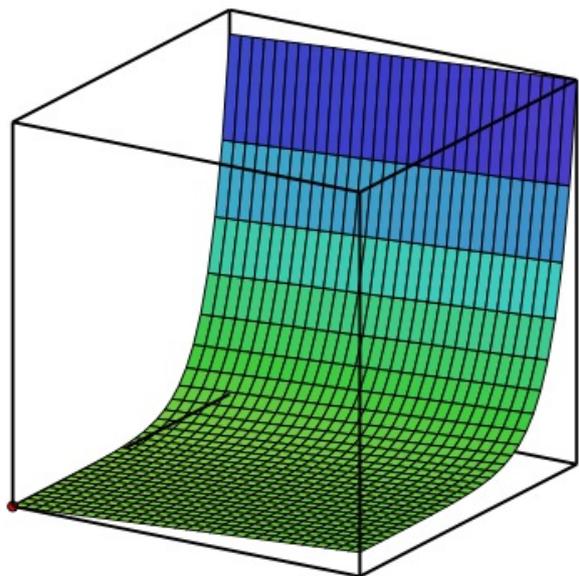
Zunächst ist es sicher keine schlechte Idee, sich mit einem Plot des Funktionsgraphen einen Überblick zu verschaffen. Dieser liefert nichts Aufregendes, in x -Richtung haben wir es mit einer linearen Funktion und in y -Richtung mit einer Exponentialfunktion zu tun. Um den Gradient zu ermitteln, benötigen wir zunächst die beiden partiellen Ableitungen in x - und in y -Richtung.

$$\frac{\partial f}{\partial x} = 1 \quad \text{und} \quad \frac{\partial f}{\partial y} = e^y.$$

Damit erhalten wir den Gradient allgemein als

$\nabla f(x,y) = \begin{bmatrix} 1 \\ e^y \end{bmatrix}$, im Punkt $P:[3, 4]$ beträgt er

konkret $\nabla f(3,4) = \begin{bmatrix} 1 \\ e^4 \end{bmatrix} \approx \begin{bmatrix} 1 \\ 54,6 \end{bmatrix}$.



Für die Richtungsableitung in Richtung \vec{a} benötigen wir zuerst den zugehörigen Einheitsvektor. Der Betrag des Richtungsvektors ist $2\sqrt{5}$ und damit der Einheitsvektor in Richtung \vec{a} :

$$\vec{e}_a = \frac{1}{2\sqrt{5}} \begin{bmatrix} 2 \\ 4 \end{bmatrix}.$$

Der gesuchte Richtungsableitung beträgt dann

$$\frac{\partial f}{\partial \vec{a}} = \frac{1}{2\sqrt{5}} \begin{bmatrix} 2 \\ 4 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ e^4 \end{bmatrix} = \frac{1}{2\sqrt{5}} \cdot (2 \cdot 1 + 4 \cdot e^4) \approx 49,28.$$

Zur Ermittlung der Richtungsableitung in Richtung $\vec{b} = [-1, -3]$ ermitteln wir wieder zuerst den zugehörigen Einheitsvektor

$$\vec{e}_b = \frac{1}{\sqrt{-1^2 + -3^2}} \begin{bmatrix} -1 \\ -3 \end{bmatrix} = \frac{1}{\sqrt{10}} \begin{bmatrix} -1 \\ -3 \end{bmatrix}$$

und multiplizieren diesen mit dem konkreten Gradienten:

$$\left[\frac{\partial f}{\partial \vec{b}} \right]_p = \frac{1}{\sqrt{10}} \begin{bmatrix} -1 \\ -3 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ e^4 \end{bmatrix} = \frac{1}{\sqrt{10}} \cdot (-1 \cdot 1 + -3 \cdot e^4) \approx -52,11$$

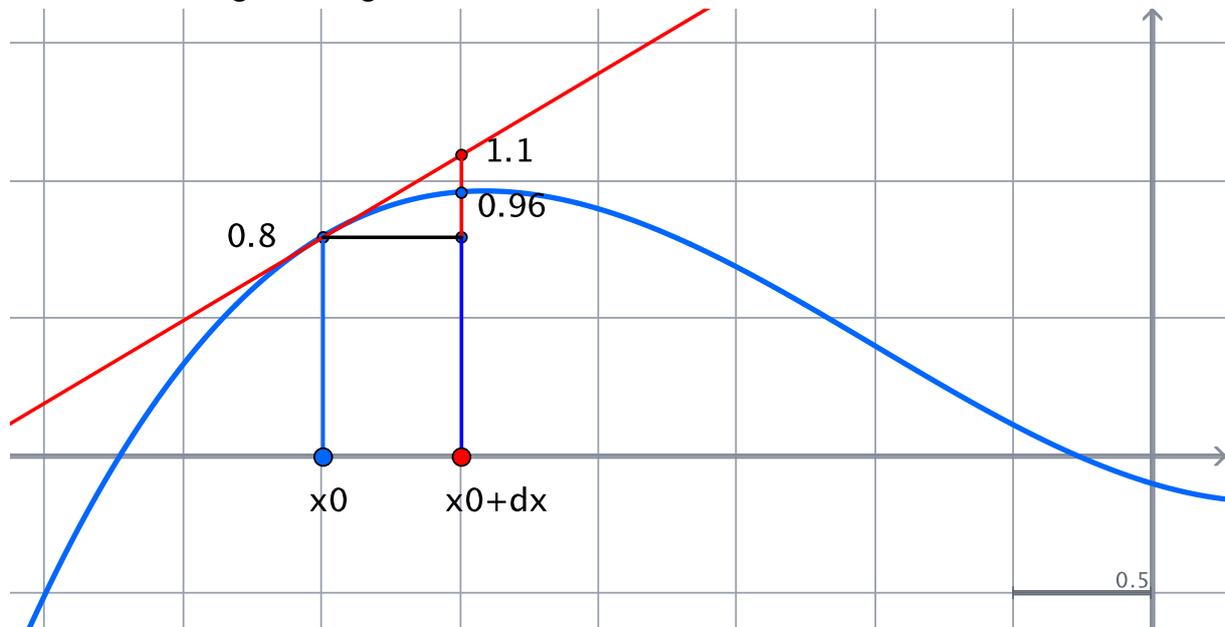
In Richtung $\vec{c} = [55, -1]$ beträgt die Ableitung

$$\left[\frac{\partial f}{\partial \vec{c}} \right]_p = \frac{1}{\sqrt{3026}} \begin{bmatrix} 55 \\ -1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ e^4 \end{bmatrix} = \frac{1}{\sqrt{3026}} \cdot (55 \cdot 1 + (-1) \cdot e^4) \approx 0,0073$$

6 Totales Differential

6.1 Rückblick auf eindimensionale Funktionen

Zur Erklärung des totalen oder vollständigen Differentials unternehmen wir zunächst nochmals ein Rückblick auf eindimensionale Funktionen und veranschaulichen den Zusammenhang mit Hilfe der Funktion $f(x) = 0,1x^3 + 0,3x^2 - 0,3x - 0,1$. Ihr Graph ist in der untenstehenden Abbildung blau dargestellt.



An der Stelle $x_0 = -3,0$ hat die Funktion den Funktionswert $0,8$. Um den Funktionswert an der Stelle $-2,5$ zu bestimmen, braucht man den Funktionswert an dieser Stelle nur über den Funktionsterm berechnen, man erhält $0,96$.

Da diese Stelle $-2,5$ nur ein kleines Stück $dx = 0,5$ von der erstgenannten Stelle x_0 entfernt liegt, kann man jedoch auch ein anderes Verfahren anwenden und den zweiten Funktionswert auf der Basis des erstgenannten annähern. Dies macht man, indem man an der Stelle x_0 die Tangente an den Graphen anlegt und entlang dieser Tangente um die kleine Distanz $0,5$ in positiver x -Richtung fortschreitet. An der Stelle $-2,5$ hat die Tangente den Wert $1,1$ und diesen Wert kann man zum einen als Näherung für den Funktionswert verwenden, zum anderen kann man über das Steigungsdreieck berechnen, dass der Funktionswert an der Stelle $-2,5$ um $0,3$ höher ist als an der Stelle $-3,0$. Den Einwand, dass diese Vorgehensweise aufwändiger ist als das bloße Berechnen des tatsächlichen Funktionswerts stellen wir für den Moment zurück und überlegen uns, wie diese Näherung geschehen kann:

Die Ableitung einer eindimensionalen Funktion gibt die Tangentensteigung an die Funktionskurve in einem Punkt x_0 an. Man gewinnt sie aus dem Differenzenquotienten

$$\frac{f(x) - f(x_0)}{x - x_0},$$

der Differenzen der y - und der x -Werte, indem man den Abstand $x - x_0$ infinitesimal klein macht und über den Limes den Differenzialquotienten erhält, der eben gleich der Ableitung ist:

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = f'(x_0)$$

Durch Multiplikation des Nenners erhalten wir – wobei wir den Limes aus Bequemlichkeitsgründen unter den Tisch fallen lassen –

$$\underbrace{f(x) - f(x_0)}_{df} = \underbrace{f'(x_0)}_{\frac{\partial f}{\partial x}} \cdot \underbrace{(x - x_0)}_{dx}.$$

Auf der linken Seite steht der Unterschied df der y -Werte (in der obigen Abbildung ist dies die rote, lotrechte Linie im Steigungsdreieck), rechts haben wir die Ableitung von f nach x multipliziert mit dem Unterschied der x -Werte. Damit haben wir bereits eine einprägsame Formel, um die Größe des Höhenunterschieds der Tangente zwischen der Stelle x_0 und einer nahegelegenen Stelle $x_0 + dx$ zu berechnen. Benötigt wird lediglich die Ableitung der Funktion an der Stelle x_0 . Wir halten fest:

Wir können ausgehend von einer Stelle x_0 und dem dortigen Funktionswert y_0 weitere in der nahen Umgebung liegenden Funktionswerte über den Höhenunterschied df mit Hilfe der Ableitung in x_0 berechnen:

$$df = \frac{\partial f}{\partial x} \cdot dx$$

Wenn wir in der obigen Gleichung $f(x) - f(x_0) = f'(x_0) \cdot (x - x_0)$ die Größe $f(x_0)$ auf die rechte Seite addieren

$$f(x) = f'(x_0) \cdot (x - x_0) + f(x_0),$$

erhalten wir übrigens die Geradengleichung im Punkt x_0 . Fallen x und x_0 zusammen, erhalten wir den Funktionswert an der Stelle x_0 . Bedenken Sie, dass $f(x)$ hier für einen y -Wert steht und die Überlegung ja nur für einen infinitesimal kleinen Wert von dx angestellt wird.

Die dargestellte Linearisierung einer Funktion in der Umgebung einer Stelle x_0 entspricht im Übrigen genau der ersten Taylornäherung. Der englische Mathematiker Brook Taylor (1685–1731) hat 1712 die Potenzreihenentwicklung einer differenzierbaren Funktion $f(x)$ gefunden. Dabei geht es darum, diese Funktion in der Umgebung einer Stelle x_0 durch eine noch zu findende Funktion $T(x)$ möglichst gut anzunähern. Diese „möglichst gute Annäherung“ soll so geschehen, dass in einer nullten Näherung natürlich die Funktionswerte an der Stelle x_0 übereinstimmen:

$$T_0(x_0) = f(x_0)$$

In einer ersten Näherung fordert man dann weiter, dass die erste Ableitung beider Funktionen an dieser Stelle x_0 ebenfalls übereinstimmen soll:

$$T_1'(x_0) = f'(x_0)$$

Mit der Ableitung als solcher ist es noch nicht getan, wir benötigen die Gleichung dieser Tangente an den Punkt $[x_0, f(x_0)]$. Mit der Ableitung und damit der Steigung im Punkt $[x_0, y_0]$ können wir allerdings über die Punkt-Steigungsform die Geradengleichung der Tangente ermitteln. Die Punkt-Steigungsform einer Geraden lautet:

$$m = \frac{y - y_0}{x - x_0} \Leftrightarrow y - y_0 = m \cdot (x - x_0) \Leftrightarrow y = y_0 + m \cdot (x - x_0)$$

Wenn wir berücksichtigen, dass y_0 der Funktionswert an der Stelle x_0 und damit gleich $f(x_0)$ ist, und dass die Steigung m die Ableitung an der Stelle x_0 und damit gleich $f'(x_0)$ ist, dann gilt für unsere verbesserte Näherungsfunktion $T_1(x)$:

$$T_1(x) = f(x_0) + f'(x_0) \cdot (x - x_0)$$

Diese Tangentengleichung ist eine verbesserte Näherung an unsere Funktion $f(x)$. Wir hatten gefordert, dass

1. diese Näherungsfunktion $T_1(x)$ an der Stelle x_0 denselben Funktionswert und
2. dieselbe Steigung haben muss wie die Funktion $f(x)$.

Dies lässt sich leicht nachweisen, wenn wir $T_1(x_0)$ bestimmen. Dann gilt nämlich:

$$1. \quad T_1(x_0) = f(x_0) + f'(x_0) \cdot (x_0 - x_0)$$

Der zweite Summand wird Null und es bleibt übrig:

$$T_1(x_0) = f(x_0) + 0,$$

d.h., die Funktionen $T_1(x)$ und $f(x)$ haben an der Stelle x_0 dieselben Funktionswerte.

2. Wir bilden die erste Ableitung von $T_1(x)$:

$$T_1'(x) = f'(x_0) + f'(x_0) \cdot x + f'(x_0) \cdot x_0$$

Ableiten führt auf

$$T_1'(x) = 0 + f'(x_0) + 0,$$

da der 1. und 3. Summand Konstanten sind, die beim Ableiten entfallen und die Ableitung von $f'(x_0) \cdot x$ ist $f'(x_0)$

Damit haben wir bestätigt, dass die erste Ableitung von $T'(x)$ genau der Ableitung von $f(x)$ an der Stelle x_0 entspricht.

Diese erste Näherung

$$T_1(x) = f(x_0) + f'(x_0) \cdot (x - x_0)$$

erfüllt somit die in sie gesetzten Anforderungen. Formal betrachtet haben wir unsere nullte Näherung (Übereinstimmung der Funktionswerte) durch den Summanden $f'(x_0) \cdot (x - x_0)$ erweitert und so die zweite Forderung nach Übereinstimmung der Steigungen hinzugewonnen.

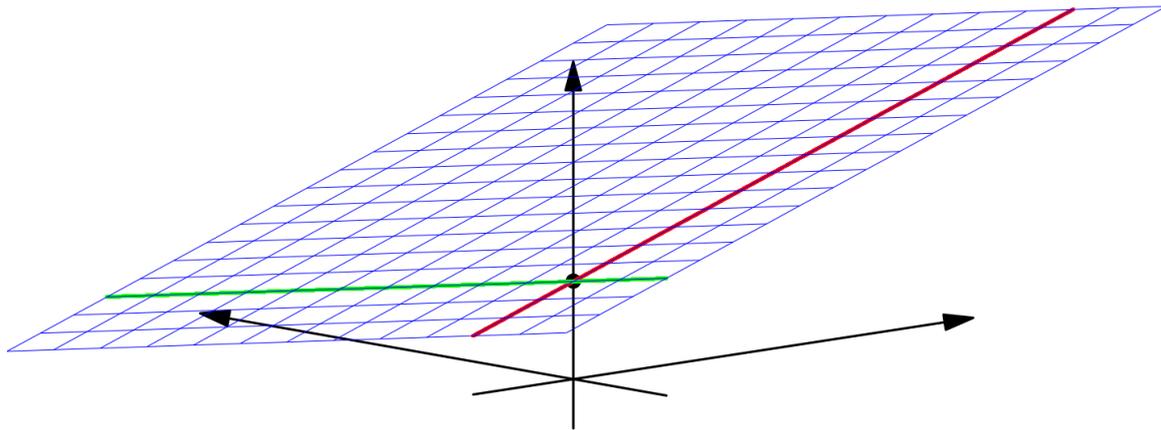
Weitere Verbesserungen im Sinne der Taylor'schen Potenzreihenentwicklung erhalten wir, wenn wir sukzessive fordern, dass auch die zweiten, dritten, ... Ableitungen übereinstimmen sollen. An dieser Stelle geben wir uns jedoch mit der ersten Näherung zufrieden, da diese einen gegebenen Funktionsgraph an einer Stelle x_0 linearisiert. Die Potenzreihenentwicklung wird in der Master-Veranstaltung *Geometrie und GPS* nochmals detaillierter aufgegriffen.

6.2 Zweidimensionale Funktionen

Übertragen wir diese Idee nun auf zweiwertige Funktionen: Der Funktionsgraph ist dort keine Linie, sondern eine mehr oder weniger gebogene oder gewellte Fläche im Raum – ein Funktions“gebirge“ – an welchem wir in jedem Punkt eine Tangentialfläche (statt einer Tangente) anlegen können

Mit den bisher kennengelernten partiellen Ableitungen ist diese Lage bisher nur sehr unständig beschrieben: Wir kennen aufgrund der Funktionsgleichung einen Punkt $[x_0, y_0, z_0]$ im Raum und müssen ausgehend von diesem Punkt die genaue Lage der Tangentialebene einmal in Richtung der x -Achse und zusätzlich in Richtung der y -Achse angeben.

Mit einer Gleichung der Form $z = a + b \cdot x + c \cdot y$ kann jedoch eine Ebene im Raum beschrieben werden:



Sind die Koordinaten für x und y jeweils Null, so hat die z -Koordinate den Wert a .

Setzen wir $y = 0$, so erhalten wir eine Geradengleichung $z = a + bx$ und damit die rote Gerade in der obigen Abbildung, welche durch den z -Achsenabschnitt a in Richtung der x -Achse verläuft und die Steigung b hat.

Setzen wir $x = 0$, so erhalten wir die grüne Gerade, welche ebenfalls den Achsenabschnitt a besitzt, in y -Richtung verläuft und die Steigung c hat. Auf diese beiden Geraden kann man eine Ebene auflegen, damit ist die Ebene im Raum eindeutig festgelegt! Wir benötigen dazu einen Punkt im Raum und – von diesem ausgehend – zwei in unterschiedliche Richtungen verlaufende Geraden mit jeweils eigener Steigung.

Übersetzen wir damit die allgemeine Ebenengleichung $z = a + b \cdot x + c \cdot y$ in die Bezeichnungen unserer zweiwertigen Funktionen:

Die linke Seite z ist der Funktionswert an einer beliebigen Stelle $f(x, y)$.

Die Größe a ist der Funktionswert z_0 an einer Stelle (x_0, y_0) , also $f(x_0, y_0)$.

Mit b und c haben wir die Steigungen in x - bzw. y -Richtungen bezeichnet, dies sind allgemein die partiellen Ableitungen $\frac{\partial f}{\partial x}$ und $\frac{\partial f}{\partial y}$ und konkret die jeweiligen Ableitungen an der

Stelle (x_0, y_0) . Damit können wir die allgemeine Form der Ebenengleichung $z = a + b \cdot x + c \cdot y$ übersetzen in die Gleichung

$$f(x, y) = f(x_0, y_0) + f_x(x_0, y_0) \cdot (x - x_0) + f_y(x_0, y_0) \cdot (y - y_0).$$

Bringen wir $f(x_0, y_0)$ auf die linke Seite, so erhalten wir

$$\underbrace{f(x, y) - f(x_0, y_0)}_{df} = \underbrace{f_x(x_0, y_0)}_{\frac{\partial f}{\partial x}} \cdot \underbrace{(x - x_0)}_{dx} + \underbrace{f_y(x_0, y_0)}_{\frac{\partial f}{\partial y}} \cdot \underbrace{(y - y_0)}_{dy}$$

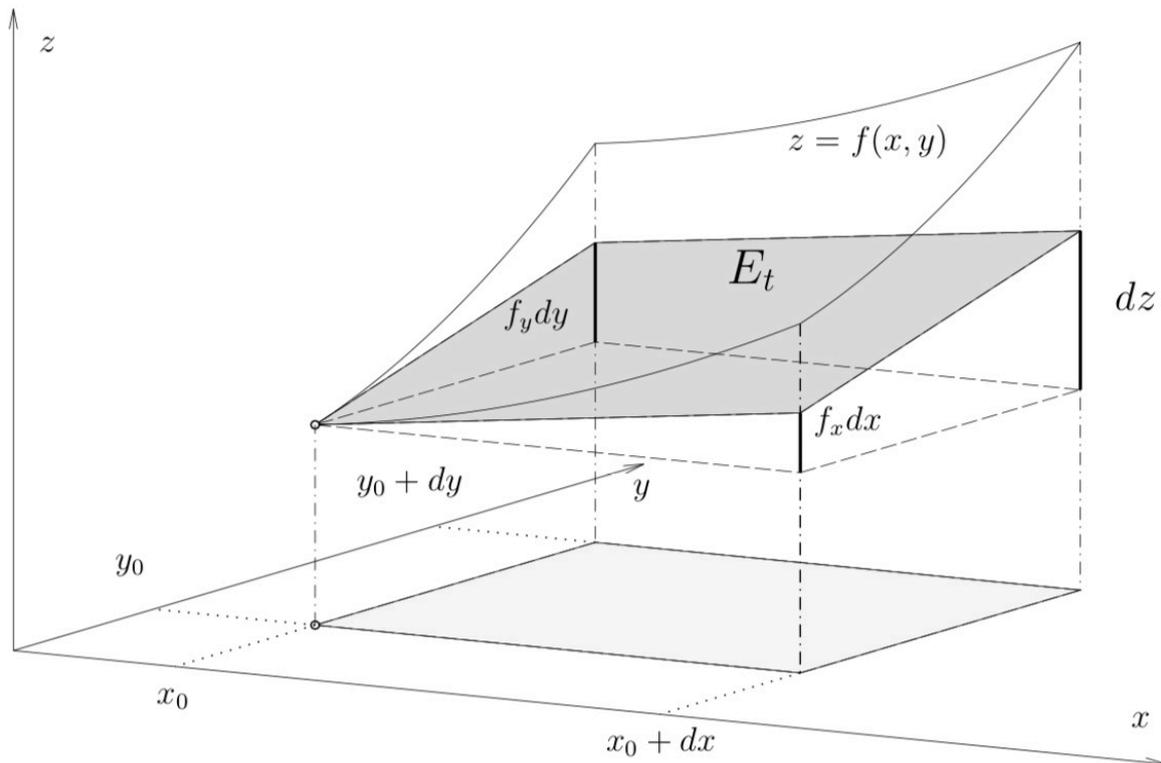
und damit

$$df = \frac{\partial f}{\partial x} \cdot dx + \frac{\partial f}{\partial y} \cdot dy$$

Auf der linken Seite steht die Differenz der z -Werte, auf der rechten Seite haben wir die Summe der partiellen Ableitungen nach x bzw. y , jeweils multipliziert mit den infinitesimal kleinen Differenzen in x - bzw. y -Richtung. Den Ausdruck

$$df = \frac{\partial f}{\partial x} \cdot dx + \frac{\partial f}{\partial y} \cdot dy$$

nennt man vollständiges oder *totales Differential*.



Darstellung des totalen Differentials und der Tangentialebene

Betrachten wir die rechte Seite unserer Gleichung

$$f(x, y) - f(x_0, y_0) = f_x(x_0, y_0) \cdot (x - x_0) + f_y(x_0, y_0) \cdot (y - y_0)$$

genauer, so erkennen wir darin ein Skalarprodukt und zwar das Skalarprodukt

$$\begin{bmatrix} f_x(x_0, y_0) \\ f_y(x_0, y_0) \end{bmatrix} \cdot \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix}$$

Den linken Faktor des Skalarprodukts kennen wir bereits: Dieser Vektor ist nichts anderes als der Gradient von f an der Stelle (x_0, y_0) und damit die Richtung der größten Steigung. Der rechte Faktor des Skalarprodukts ist zunächst ebenfalls ein Vektor. Im eindimensionalen Fall gibt $x - x_0$ an, wie weit wir uns in x -Richtung von der Stelle x_0 wegbewegen. Entsprechend gibt $y - y_0$ an, wie weit wir uns in y -Richtung von der Stelle y_0 wegbewegen. Als Vektor gibt somit

$$\begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix}$$

an, wie weit und in welche Richtung wir uns vom Punkt (x_0, y_0) wegbewegen.

Die grundsätzlich bei Ableitungen zugrundeliegende Fragestellung lautet doch, wie sich der Funktionswert ändert, wenn wir uns ein wenig von der Stelle x_0 und damit vom Funktionswert $f(x_0)$ wegbewegen. Diese Frage können wir bei eindimensionalen Funktionen beantworten, wenn wir den Funktionswert $f(x_0)$, die Steigung an der Stelle x_0 und die Größe der Abweichung kennen. Diese Funktionswertänderung kann dann mit der oben hergeleiteten Gleichung

$$df = \frac{\partial f}{\partial x} \cdot dx$$

näherungsweise berechnet werden. Dieselbe Frage lässt sich auch für zweiwertige Funktionen stellen:

Wie ändert sich der Funktionswert in linearer Näherung, wenn ich mit dem x -Wert ein Stückchen in x -Richtung und mit dem y -Wert ein Stückchen in y -Richtung zur Seite gehe? Die Antwort darauf liefert die bereits hergeleitete Gleichung

$$df = \frac{\partial f}{\partial x} \cdot dx + \frac{\partial f}{\partial y} \cdot dy,$$

die als totales oder vollständiges Differential bezeichnet wird. Diese Bezeichnung rührt daher, weil das totale Differential die gesamte Information über die Ableitung enthält, während die partiellen Ableitungen nur Informationen über die Ableitung in Richtung der Koordinatenachsen enthalten.

6.3 Anwendungen

6.3.1 Lineare Näherung

Mit dem totalen Differential bzw. der Gleichung der Tangentialebene ist es möglich, gegebene Funktionen linear anzunähern: Genau wie differenzierbare eindimensionale Funktionen lokal linear sind und daher lokal begrenzt durch die jeweilige Tangente repräsentiert werden können, können differenzierbare zweidimensionale Funktionen lokal durch Tangentialebenen linear angenähert werden. Betrachten wir dazu die oben hergeleitete Gleichung für die Tangentialebene:

$$f(x, y) = f(x_0, y_0) + f_x(x_0, y_0) \cdot (x - x_0) + f_y(x_0, y_0) \cdot (y - y_0)$$

Die Tangentialebene berührt die Funktionsfläche an der Stelle (x_0, y_0) . In einer kleinen Umgebung um eben diesen Berührungspunkt kann – anstatt einen Funktionswert über die Funktionsgleichung zu berechnen, dieser auch über diese Gleichung der Tangentialebene näherungsweise bestimmt werden. Da hierin die Funktionsvariablen x und y immer nur in der ersten Potenz vorkommen, spricht man hierbei von einer linearen Näherung. Wir demonstrieren die Vorgehensweise anhand eines konkreten Beispiels:

Gegeben sei die Funktion

$$f(x, y) = \frac{\sqrt{x}}{y}.$$

Gesucht ist die lokale Näherung an der Stelle $(x_0, y_0) = (4, 2)$.

Dazu bestimmen wir die partiellen Ableitungen allgemein und konkret:

$$\frac{\partial f}{\partial x} = \frac{1}{2\sqrt{x}} \cdot \frac{1}{y}, \text{ konkret an der Stelle } (4, 2): \frac{\partial f}{\partial x} = \frac{1}{2\sqrt{4}} \cdot \frac{1}{2} = \frac{1}{8}$$

$$\frac{\partial f}{\partial y} = -\frac{\sqrt{x}}{y^2}, \text{ konkret an der Stelle } (4, 2): \frac{\partial f}{\partial y} = -\frac{\sqrt{4}}{2^2} = -\frac{1}{2}$$

Außerdem benötigen wir noch den Funktionswert an der Stelle $[4, 2]$:

$$f(4, 2) = \frac{\sqrt{4}}{2} = 1$$

Die lineare Näherung an der Stelle $[4, 2]$ lautet somit:

$$f(x, y) = 1 + \frac{1}{8} \cdot (x - 4) - \frac{1}{2} \cdot (y - 2)$$

Diese Näherung sagt aus, dass der Funktionswert an der Stelle $[4, 2]$ gleich 1 ist. Für jeden Schritt, den wir in x -Richtung gehen, vergrößert sich dieser um $1/8$. Für jeden Schritt, den wir

in y -Richtung gehen, wird er um $1/2$ kleiner. Wir nähern konkret den Funktionswert an der Stelle $[3.98, 2.03]$:

$$f(3.98, 2.03) = 1 + \frac{1}{8} \cdot (-0.02) - \frac{1}{2} \cdot (0.03) = 0.9825$$

Berechnen wir diesen Funktionswert anhand der Funktionsgleichung, so erhalten wir

$$f(3.98, 2.03) = \frac{\sqrt{3.98}}{2.03} = 0,98$$

Der genäherte und der korrekte Wert unterscheiden sich um $0,0025$. Dies ist der absolute Approximationsfehler. Bezieht man diesen absoluten Fehler auf den korrekten Wert, so erhält man den relativen Approximationsfehler in Prozent.

Nach so viel Theorie ist eine konkrete Anwendung in der Praxis hilfreich. Wir gehen dazu in den Bereich der idealen Gase und zitieren die allgemeine Gasgleichung für ein Mol eines Gases. Diese beschreibt das Volumen für ein Mol eines Gases in Abhängigkeit von dessen Druck und dessen Temperatur:

$$p \cdot V = R \cdot T$$

R ist die universelle Gaskonstante, sie beträgt gerundet $8,31 \frac{J}{K}$. Damit lautet die Funktion für das Gasvolumen in Abhängigkeit vom Druck und der Temperatur:

$$V(p, T) = 8,31 \frac{J}{K} \cdot \frac{T}{p}$$

Der Zusammenhang ist einfach: Sie wissen bestimmt, dass sich ein Gas ausdehnt, wenn dessen Temperatur erhöht wird und dass es komprimiert wird, wenn der Druck steigt.

Unsere Fragestellung lautet nun, wie sich das Volumen des Gases ändert, wenn sich dessen Umgebungsbedingungen Druck und Temperatur ändern. Als Umgebungsbedingungen legen wir eine Temperatur von $300K$ ($27^\circ C$) und einen Druck von $10^5 Pa$ ($1000 hPa$) fest.

Die Formel für das totale Differential lautet nun

$$dV = \frac{\partial V}{\partial T} \cdot dt + \frac{\partial V}{\partial p} \cdot dp.$$

Wir benötigen zunächst die partiellen Ableitungen nach T und p :

$$\frac{\partial V}{\partial T} = 8,31 \frac{J}{K} \cdot \frac{1}{p} \quad \text{und} \quad \frac{\partial V}{\partial p} = 8,31 \frac{J}{K} \cdot -\frac{T}{p^2} \quad (\text{Ableitung von } \frac{1}{p} \text{ ist } -\frac{1}{p^2})$$

Diese Ableitungen müssen für die gegebenen Umgebungsbedingungen berechnet werden:

$$\frac{\partial V}{\partial T} = 8,31 \frac{J}{K} \cdot \frac{1}{p} = 8,31 \frac{N \cdot m}{K} \cdot \frac{1}{10^5 \cdot \frac{N}{m^2}} \approx 10^{-4} \frac{m^3}{K}$$

$$\frac{\partial V}{\partial p} = 8,31 \frac{J}{K} \cdot -\frac{T}{p^2} = -8,31 \frac{N \cdot m}{K} \cdot \frac{300K}{(10^5 Pa)^2} \approx -2500 \cdot N \cdot m \cdot \frac{1}{10^{10} \cdot \frac{N}{m^2} \cdot Pa} \approx \frac{-2500 m^3}{10^{10} Pa} \approx 2,5 \cdot 10^{-7} \frac{m^3}{Pa}$$

Bei den gegebenen Bedingungen gilt somit:

$$dV = 10^{-4} \frac{m^3}{K} \cdot dt - 2,5 \cdot 10^{-7} \frac{m^3}{Pa} \cdot dp$$

Daraus können wir ersehen, wie sich das Volumen eines Mols eines idealen Gases bei Zimmertemperatur und Normaldruck verändert, wenn wir auf dessen Druck und dessen Temperatur einwirken. Für jedes Kelvin Temperaturänderung verkleinert oder vergrößert sich das Vo-

lumen um 10^{-4} m^3 und jede Druckänderung um 1 Pascal verändert das Volumen um $2,5 \cdot 10^{-7} \text{ m}^3$. Aus konkreten Änderungen der Umgebungsbedingungen können damit leicht die resultierenden Volumenänderungen berechnet werden – solange die Änderungen klein genug sind, da es sich ja nur um lineare Näherungen handelt.

6.3.2 Fehlerrechnung

Eine weitere Anwendung des totalen Differentials liegt in der Fehlerrechnung. Insbesondere die Physik gewinnt ihre Aussagen aus Messungen und diese sind immer mit einem kleinen Fehler behaftet. Aus verschiedenen Größen, die alle mit einem Messfehler behaftet sind, soll das Gesamtergebn aufgrund einer Formel o.Ä. berechnet werden. Da möchte man dann schon wissen, welche Auswirkungen diese kleiner Fehler der einzelnen Messwerte auf das Gesamtergebn haben. Übertragen in die Analysis bedeutet dies, dass wir eine mehrdimensionale Funktion f haben, die von mehreren Größen abhängt: $f(x, y, z, \dots)$. Alle Messgrößen x, y, z, \dots sind mit Messfehlern $\Delta x, \Delta y, \Delta z, \dots$ behaftet. Diese Messfehler führen zu einem Fehler Δf des daraus berechneten Messergebnisses f . Unter der Voraussetzung, dass diese Messfehler $\Delta x, \Delta y, \Delta z, \dots$ klein gegenüber den eigentlichen Messgrößen x, y, z, \dots sind, kann der Fehler Δf mit Hilfe des totalen Differentials abgeschätzt werden:

$$\Delta f = \frac{\partial f}{\partial x} \cdot \Delta x + \frac{\partial f}{\partial y} \cdot \Delta y + \frac{\partial f}{\partial z} \cdot \Delta z + \dots$$

Durch positive und negative partielle Ableitungen kann es vorkommen, dass sich Fehler aufheben. Diese „Fehlerkompensation“ ist in der Fehlerrechnung jedoch nicht erwünscht, so dass man immer die Beträge der partiellen Ableitungen verwendet:

$$\Delta f = \left| \frac{\partial f}{\partial x} \right| \cdot \Delta x + \left| \frac{\partial f}{\partial y} \right| \cdot \Delta y + \left| \frac{\partial f}{\partial z} \right| \cdot \Delta z + \dots$$

Auch hierzu ein einfaches Beispiel, in welchem das Volumen eines Zylinders bestimmt werden soll. Hierzu werden mehrere Messungen des Durchmessers d und der Höhe h im Millimetern durchgeführt. Die Messwerte des Durchmessers schwanken zwischen

$$98,5 \text{ mm} < d < 101,8 \text{ mm}$$

und die Messwerte der Höhe liegen zwischen

$$199,3 \text{ mm} < h < 201,1 \text{ mm}.$$

Daraus ergibt sich ein Durchschnittswert von 100,15 mm für den Durchmesser und von 200,2 mm für die Höhe. Der Messfehler Δd des Durchmessers beträgt

$$101,8 \text{ mm} - 100,15 \text{ mm} = 1,65 \text{ mm}$$

und der Messfehler Δh der Höhe beträgt

$$201,1 \text{ mm} - 200,2 \text{ mm} = 0,9 \text{ mm}.$$

Das Volumen V hängt von d und h ab:

$$V(d, h) = \pi \cdot \left(\frac{d}{2} \right)^2 \cdot h$$

Die partielle Ableitung nach d lautet:

$$\frac{\partial V}{\partial d} = \frac{1}{2} \cdot \pi \cdot h \cdot d,$$

konkret mit den errechneten Mittelwerten für d und h :

$$\frac{\partial V}{\partial d} = \frac{1}{2} \cdot \pi \cdot 200,2 \text{ mm} \cdot 100,15 \text{ mm} \approx 31494,513 \text{ mm}^2$$

Die partielle Ableitung nach h lautet:

$$\frac{\partial V}{\partial h} = \frac{1}{4} \cdot \pi \cdot d^2,$$

konkret mit dem errechneten Mittelwert für d :

$$\frac{\partial V}{\partial h} = \frac{1}{4} \cdot \pi \cdot (100,15 \text{ mm})^2 = 7877,5613 \text{ mm}^2$$

Damit können wir die Abweichung des Messergebnisses bestimmen:

$$\Delta V = 31.494,513 \text{ mm}^2 \cdot 1,65 \text{ mm} + 7.877,5613 \text{ mm}^2 \cdot 0,9 \text{ mm} = 59.055,75 \text{ mm}^3$$

Das Volumen des Zylinders – errechnet aus den Mittelwerten – ergibt:

$$V = \pi \cdot \left(\frac{100,15 \text{ mm}}{2} \right)^2 \cdot 200,20 \text{ mm} = 1.577.088 \text{ mm}^3$$

Als Messergebnis für den Zylinder geben wir somit den folgenden Wert an:

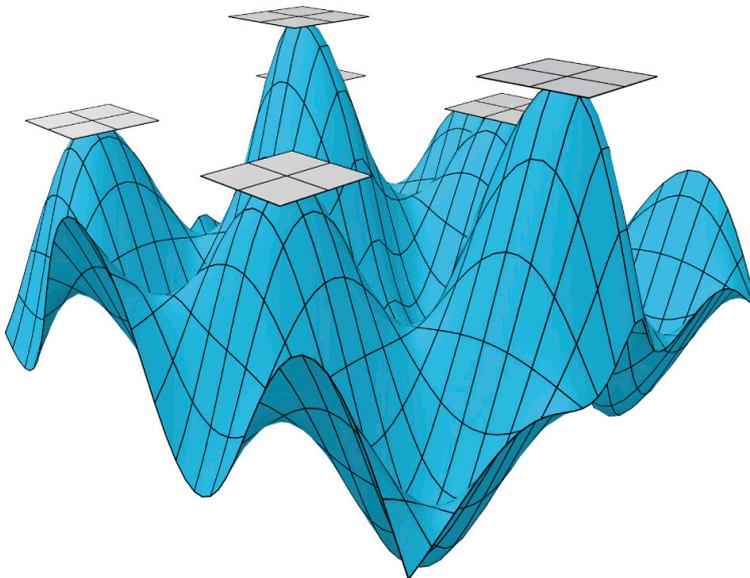
$$V_{\text{Zylinder}} = 1.577.088 \text{ mm}^3 \pm 59055,75 \text{ mm}^3$$

7 Lokale Extremwerte

Genau wie bei den seither behandelten eindimensionalen Funktionen stellt sich bei höher dimensionierten Funktionen ebenfalls die Frage nach den Extremstellen. Für die hier behandelten Funktionen zweier Veränderlicher hat diese Frage eine ganz anschauliche Bedeutung: Der Funktionsgraph einer solchen Funktion ist ja ein „Funktionsgebirge“ und kann damit ganz gut mit der Erdoberfläche konkretisiert werden. Wir befinden uns auf der Erdoberfläche auf einem Berggipfel, wenn es – egal in welche Richtung wir gehen –ringsum nur abwärts geht. Und wir befinden uns am tiefsten Punkt eines Tales, wenn es ringsum nur aufwärts geht.

7.1 Notwendige Bedingung

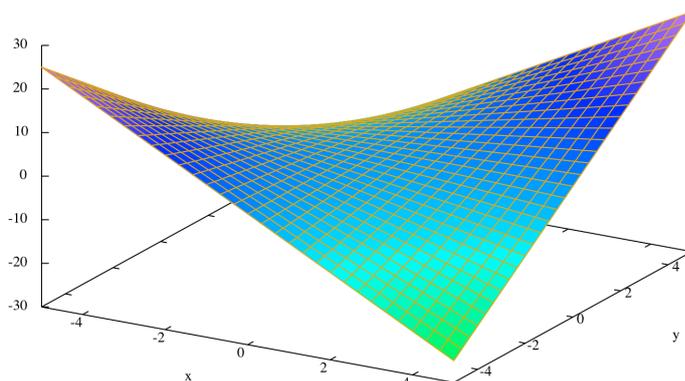
Wenn wir diese anschauliche Situation auf den Graphen einer Funktion übertragen, dann muss doch eine an einen Hochpunkt angelegte Tangentialfläche „absolut waagrecht“ sein und für eine an einen Tiefpunkt angelegte Tangentialfläche gilt genau dasselbe.



Jetzt müssen wir noch die umgangssprachliche Formulierung „absolut waagrecht“ übersetzen und finden dafür unschwer die Forderung, dass in solch einem Fall beide partiellen Ableitungen gleich 0 sein müssen. Insofern erhalten wir im Vergleich zu eindimensionalen Funktionen nichts wesentlich Neues, vielmehr können wir die dort gefundenen Konzepte leicht auf die höherwertigen Funktionen übertragen.

Die Forderung, dass beide partiellen Ableitungen gleich Null sein müssen, kann man auch in der Forderung kumulieren, dass der Gradient an der entsprechenden Stelle gleich Null sein muss.

Für eine erste Annäherung untersuchen wir die Funktion $f(x,y) = x \cdot y$. Ihr Plot sieht folgendermaßen aus ...



... und er offenbart uns sogleich eine Schwäche unserer seitherigen Überlegungen, die aber im vollen Einklang mit derselben Problematik bei eindimensionalen Funktionen steht: Die Forderung nach einer waagerechten Tangente bzw. Tangentialebene reicht nicht aus, es könnte sich dabei auch um einen Sattelpunkt handeln. In der Tat hat die Funktion $f(x, y) = x \cdot y$ die beiden partiellen Ableitungen

$$f_x(x, y) = \frac{\partial f}{\partial x} = y \quad \text{und} \quad f_y(x, y) = \frac{\partial f}{\partial y} = x$$

und an der Stelle $[0, 0]$ sind beide Ableitungen gleich Null, schließlich ist somit auch der Gradient im Nullpunkt gleich Null. Tatsächlich hat die Funktion an dieser Stelle aber kein Extremum, sondern einen Sattelpunkt. Solche Punkte, an denen die partiellen Ableitungen erster Ordnung allesamt gleich null sind, werden *stationäre Punkte* genannt. Diese Bezeichnung rührt daher, dass die Funktion an solchen Stellen ein Anwachsen oder verringern stoppt. Etwas salopp ausgedrückt: Es geht weder rauf noch runter! Solche stationären Punkte sind eine notwendige Bedingung für das Vorliegen eines Extremums: Ein Punkt, der nicht „stationär“ ist, kann keine Extremum sein! Ein stationärer Punkt ist jedoch für ein Extremum nicht hinreichend, es kann sich ja auch um einen Sattelpunkt handeln!

7.2 Hinreichende Bedingung

Die Lösung dieses Problems können wir (mindestens prinzipiell) ebenfalls aus der Analysis eindimensionaler Funktionen übernehmen – man muss sich eben die zweiten Ableitungen anschauen! Während es aber seither nur eine zweite Ableitung gab, haben wir im Fall von Funktionen mit zwei Veränderlichen deren vier! Es bleibt also zu klären, wie die Forderung *2. Ableitung untersuchen* hier konkret umgesetzt werden muss.

Hier kommt die nach dem Mathematiker Otto Hesse (1811–1874) benannte Hesse-Matrix zur Anwendung. Dabei werden die – in unseren hauptsächlich betrachteten Fällen zweiwertiger Funktionen – insgesamt vorliegenden vier partiellen Ableitungen 2. Ordnung

$$f_{xx}(x, y) = \frac{\partial^2 f}{\partial x^2}, \quad f_{xy}(x, y) = \frac{\partial^2 f}{\partial x \partial y}, \quad f_{yx}(x, y) = \frac{\partial^2 f}{\partial y \partial x}, \quad f_{yy}(x, y) = \frac{\partial^2 f}{\partial y^2}$$

in Form einer Matrix geschrieben:

$$H_f(x, y) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix}$$

Die Hesse-Matrix wird in Maxima mit Hilfe der folgenden Funktion erstellt. Auch hierbei kann durch die Angabe von Punktkoordinaten die allgemeine Form oder eine spezielle Form für den angegebenen Punkt ermittelt werden:

```
hessematrix([P]) := block(
[ab_xx, ab_xy, ab_yx, ab_yy],
ab_xx:diff(f(x, y), x, 2),
ab_xy:diff(diff(f(x, y), x, 1), y, 1),
ab_yx:diff(diff(f(x, y), y, 1), x, 1),
ab_yy:diff(f(x, y), y, 2),
if length(P)>0 then
(
ab_xx:subst(P[1][1], x, ab_xx),
```

```

ab_xx:subst (P[1][2], y, ab_xx) ,
ab_xy:subst (P[1][1], x, ab_xy) ,
ab_xy:subst (P[1][2], y, ab_xy) ,
ab_yx:subst (P[1][1], x, ab_yx) ,
ab_yx:subst (P[1][2], y, ab_yx) ,
ab_yy:subst (P[1][1], x, ab_yy) ,
ab_yy:subst (P[1][2], y, ab_yy)
) ,
matrix([ab_xx, ab_xy], [ab_yx, ab_yy]) );

```

Die Entscheidung, ob es sich bei einem Punkt mit dem Gradient 0 um einen Hoch-, Tief- oder Sattelpunkt handelt, ist nur zu treffen, wenn man die „Definitheit“ der Hesse-Matrix ermittelt. Der Begriff der „Definitheit“ führt in den Bereich der linearen Algebra, er beschreibt, „welche Vorzeichen reelle quadratische Formen annehmen können, die durch Matrizen erzeugt werden“, oder – um es trivial (und fachlich nicht ganz korrekt) auszudrücken – eben, welches „Vorzeichen“ eine Matrix hat. Dabei gibt es aber nicht nur „plus oder minus“, eine Matrix kann vielmehr positiv oder negativ definit sein, sie kann weiter positiv oder negativ semi-definit sein und sie kann schließlich indefinit sein.

Um die Definitheit einer Matrix zu bestimmen, gibt es mehrere Möglichkeiten, zum Beispiel über Eigenwerte oder Determinanten. Für die Lösung mit Hilfe von Determinanten müssen die Hauptminoren oder Unterdeterminanten berechnet werden. Dies sind die Determinanten aller Matrizen, die entstehen, wenn man – ausgehend von der linken oberen Ecke – quadratische Matrizen bildet. In einer 3×3-Matrix beispielsweise

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \text{ sind dies ...}$$

$$\det[a] = |a| = a$$

$$\det \begin{bmatrix} a & b \\ d & e \end{bmatrix} = \begin{vmatrix} a & b \\ d & e \end{vmatrix} = ae - bd \text{ und}$$

$$\det \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = aei + bfg + cdh - ceg - afh - bdi^5$$

Da in unseren hauptsächlich betrachteten Fällen nur 2×2-Matrizen auftreten und diese nach dem Satz von Schwarz symmetrisch sind, können wir uns bei der Bestimmung der Definitheit auf ein relativ einfaches Schema zurückziehen. Eine solche Matrix hat immer die Form

$$\begin{pmatrix} a & b \\ b & c \end{pmatrix} \text{ und es gilt: Die Matrix ist ...}$$

positiv definit, wenn: $a \cdot c > b^2$ und $a > 0$

negativ definit, wenn: $a \cdot c > b^2$ und $a < 0$

positiv semi-definit, wenn: $a \cdot c \geq b^2$ und $a \geq 0$ und $c \geq 0$

⁵ Regel von Sarrus („Jägerzaunregel“) für 3×3-Matrizen, für größere Matrizen verwendet man den Gauß-Jordan-Algorithmus oder den Laplace’schen Entwicklungssatz.

negativ semi-definit, wenn: $a \cdot c \geq b^2$ und $a \leq 0$ und $c \leq 0$

nicht definit, wenn: $a \cdot c < b^2$

Für größere Matrizen müssen wir eine andere Herangehensweise wählen, an dieser Stelle mag aber diese spezielle Definition genügen.

Auf das Problem der Feststellung von Extremwerten angewandt bedeutet dies, ganz im Gleichklang mit der Untersuchung bei eindimensionalen Funktionen:

Es liegt ein lokales Minimum vor, wenn die Hesse-Matrix positiv definit ist.

Es liegt ein lokales Maximum vor, wenn die Hesse-Matrix negativ definit ist.

Es liegt ein Sattelpunkt vor, wenn die Hesse-Matrix indefinit ist.

Keine Entscheidung ist möglich, wenn die Matrix semi-definit ist. Dies ist in unserem Kontext konkret der Fall, wenn $a \cdot c = b^2$ gilt.

Zusammengefasst bedeutet dies:

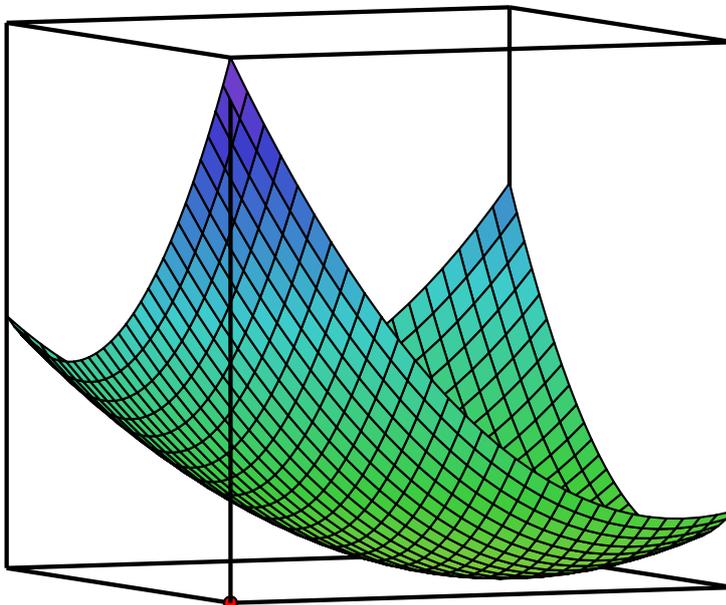
| Extremum in $f(x,y)$ | $a \cdot c > b^2$ | $a \cdot c = b^2$ | $a \cdot c < b^2$ |
|----------------------|-------------------|--------------------|-------------------|
| $a < 0$ | Maximum | nicht entscheidbar | Sattelpunkt |
| $a > 0$ | Minimum | | |

7.3 Anwendungsbeispiel

Am besten wenden wir dies alles gleich an einer einfachen Aufgabenstellung an. Gesucht sind Extremstellen der Funktion

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

Einen orientierenden Überblick bietet der Graph der Funktion:



Danach ist innerhalb des Plotbereichs ($-5 < x < 5$, $-5 < y < 5$) ein Minimum zu erwarten. Wir bestimmen zunächst die partiellen Ableitungen erster Ordnung

$$\frac{\partial f}{\partial x} = 2x - 3 + y \quad \text{und} \quad \frac{\partial f}{\partial y} = x + 2y.$$

An einer Extremstelle müssen beide partiellen Ableitungen erster Ordnung gleich Null werden, wir haben also ein Gleichungssystem zu lösen:

$$2x - 3 + y = 0 \quad (1)$$

$$x + 2y = 0 \quad (2)$$

Aus (2) entnehmen wir:

$$x = -2y$$

was wir in (1) einsetzen:

$$2(-2y) - 3 + y = 0$$

$$-4y - 3 + y = 0$$

$$y = -1$$

Damit in (2):

$$x + 2(-1) = 0$$

$$x = 2$$

Die notwendige Bedingung, dass die Tangentialebene horizontal verläuft, ist offenbar im stationären Punkt $P: [2, -1]$ erfüllt. Den Nachweis, dass es sich dort tatsächlich um ein Minimum handelt, müssen wir mit Hilfe der zweiten Ableitungen erbringen:

$$\frac{\partial f^2}{\partial x^2} = 2 \quad \frac{\partial f^2}{\partial x \partial y} = 1 \quad \frac{\partial f^2}{\partial y \partial x} = 1 \quad \frac{\partial f^2}{\partial y^2} = 2$$

Diese sortieren wir in die Hesse-Matrix ein:

$$H_f(x, y) = \begin{bmatrix} \frac{\partial f^2}{\partial x^2} & \frac{\partial f^2}{\partial x \partial y} \\ \frac{\partial f^2}{\partial y \partial x} & \frac{\partial f^2}{\partial y^2} \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

Um die Definitheit dieser Matrix zu bestimmen, untersuchen wir zuerst, ob $a \cdot c$ größer oder kleiner als b^2 ist. Bezogen auf unsere Ableitungen ist dies die Frage, ob

$$\frac{\partial f^2}{\partial x^2} \cdot \frac{\partial f^2}{\partial y^2} \text{ größer oder kleiner als } \frac{\partial f^2}{\partial x \partial y} \cdot \frac{\partial f^2}{\partial y \partial x} \text{ ist (wobei i.a. } \frac{\partial f^2}{\partial x \partial y} = \frac{\partial f^2}{\partial y \partial x} \text{ gilt}^6).$$

In unserem speziellen Fall ist $2 \cdot 2 > 1 \cdot 1$, die Matrix ist also auf jeden Fall definit und da au-

ßerdem $\frac{\partial f^2}{\partial x^2} > 0$ gilt, ist die Matrix positiv definit. Dies bedeutet schließlich und endlich, dass am Punkt $[2, -1]$ ein Minimum vorliegt.

Die Frage, für welche Punkte der Gradient einer Funktion Null wird, kann mit der nachfolgenden Maxima-Funktion ermittelt werden. Dabei werden nur reelle Punkte angegeben:

```

extremum() := block(
  [ergebnis, loesung, i, x, y],
  ergebnis: [],
  loesung: solve
    ([diff(f(x, y), x, 1) = 0, diff(f(x, y), y, 1) = 0], [x, y]),
  for i:1 thru length(loesung) do
    (
      x: ev(x, loesung[i]),
      y: ev(y, loesung[i]),

```

⁶ Satz von Schwarz

```

    if (isreal_p(x) or isreal_p(y)) then
        ergebnis:endcons([x,y], ergebnis)
    ),
    ergebnis );

```

Um die hier verwendete Funktion `isreal_p()` aufrufen zu können, muss zuvor das zugehörige Paket mit `load(to_poly_solve)` eingebunden werden!

7.4 Anwendung in der Betriebswirtschaft

Häufig werden betriebswirtschaftliche Probleme mit Hilfe der Analysis gelöst: Der Gewinn, den ein Betrieb mit dem Verkauf zweier Produkte erzielt, hängt von den Mengen x und y ab, in denen er die beiden Produkte produziert. Dabei berechnet sich der Gewinn in Abhängigkeit von x und y nach dem folgenden Zusammenhang:

$$G(x,y) = 100x + 140y - x^2 - 2xy - 2y^2$$

Für welche Werte von x und y erzielt der Betrieb den höchsten Gewinn? Wie groß ist dieser?

Zur Beantwortung dieser Frage erstellen wir die beiden ersten partiellen Ableitungen:

$$G_x(x,y) = 100 - 2x - 2y$$

$$G_y(x,y) = 140 - 2x - 4y$$

Beide müssen gleich Null sein. Durch Gleichsetzen erhalten wir $y=20$ und daraus $x=30$.

Ob sich am stationären Punkt $[30, 20]$ tatsächlich ein Maximum befindet, müssen wir über die zweiten Ableitungen klären:

$$G_{xx}(x,y) = -2$$

$$G_{xy}(x,y) = -2$$

$$G_{yy}(x,y) = -4$$

$$G_{yx}(x,y) = -2$$

Daraus erhalten wir die Hesse-Matrix $\begin{pmatrix} -2 & -2 \\ -2 & -4 \end{pmatrix}$. Um die Definitheit dieser Matrix zu be-

stimmen, untersuchen wir, ob $a \cdot c$ größer oder kleiner als b^2 ist. In unserem Beispiel ist $-2 \cdot -4 > -2^2$, damit haben wir tatsächlich eine Extremstelle gefunden. Da außerdem $a < 0$ gilt, handelt es sich um ein Maximum.

Der Funktionswert $G(30,20)$ ist 2900.

8 Extremwerte mit Nebenbedingungen

Häufig sollen gegebene Funktionen, die dann *Zielfunktion* genannt werden, im Hinblick auf weitere Nebenbedingungen optimiert werden. Das Optimieren unter Nebenbedingungen hat eine große Relevanz für viele wissenschaftliche Gebiete. Die Wirtschaftswissenschaften wurden ja bereits beispielhaft erwähnt und schon einige Aufgaben gelöst, in denen es um die Optimierung des Gewinns oder der Produktionskosten geht. Bei der tatsächlichen Produktion von Gütern gibt es in der Realität feste Grenzen: Ein Unternehmen hat beispielsweise nur eine begrenzte Menge Geldes und Zeit für die Produktion zur Verfügung und die Optimierung kann nur innerhalb dieser Grenzen stattfinden.

Für eine solche Optimierung kommen grundsätzlich zwei Verfahren zur Anwendung. Wenn es möglich ist, die Nebenbedingung(en) nach einer Variablen aufzulösen, kann man diese Variable in der Zielfunktion durch den gefundenen Term substituieren und erhält so eine um eine Variable reduzierte Zielfunktion, deren Extrema nach dem herkömmlichen Verfahren bestimmt werden können.

Ist ein Auflösen der Nebenbedingung(en) nach einer Variablen nicht möglich, so kommt die Methode der Lagrange-Multiplikatoren zum Einsatz.

8.1 Einsetzverfahren

Eine Lieferfirma nimmt nur rechteckige Behälter an, bei denen die Summe aus Länge und Umfang (eines Querschnitts) maximal 270 cm ist. Bestimmen Sie die Abmessungen eines akzeptablen Behälters mit maximalem Volumen. Die Abmessungen des rechteckigen Behälters seien x (Länge), y (Breite) und z (Höhe).

Hier liegt eine Optimierungsaufgabe vor, das Volumen $V = x \cdot y \cdot z$ des Behälters soll maximal werden. Da dieses Volumen grundsätzlich beliebig an wachsen kann, kommt nun noch eine Nebenbedingung ins Spiel: Die Logistikfirma nimmt nur Behälter entgegen, deren Maße der folgenden Bedingung genügen: Die Länge des Quaders plus sein Gürtelmaß darf 270 cm nicht überschreiten:

$$x + 2y + 2z = 270$$

Damit haben wir eine *Zielfunktion* $\mathcal{Z}f(x, y, z) = x \cdot y \cdot z$, deren Maximum unter Beachtung der Bedingung $x + 2y + 2z = 270$ gefunden werden soll.

Die Vorgehensweise ist nun so, dass man die Nebenbedingung nach einer Variablen auflöst, z.B.

$$x = 270 - 2y - 2z$$

und damit die Variable x in der Zielfunktion substituiert:

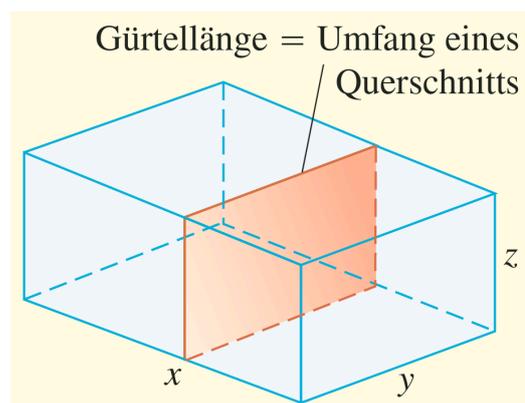
$$\mathcal{Z}f(y, z) = (270 - 2y - 2z) \cdot y \cdot z$$

$$\mathcal{Z}f(y, z) = 270yz - 2y^2z - 2yz^2$$

Damit können wir das Volumen des Behälters als eine Funktion von zwei Variablen schreiben. Von dieser Funktion ermitteln wir die beiden ersten Ableitungen:

$$\mathcal{Z}f_y(y, z) = 270z - 4yz - 2z^2 = z \cdot (270 - 4y - 2z)$$

$$\mathcal{Z}f_z(y, z) = 270y - 2y^2 - 4yz = y \cdot (270 - 2y - 4z)$$



Beide Ableitungen müssen zu Null werden, Maxima liefert aufgrund des Gleichungssystems die stationären Punkte $[[y=0, z=0], [y=0, z=135], [y=135, z=0], [y=45, z=45]]$

Die ersten drei Lösungen können wir im gegebenen Kontext ausschließen, da diese ein Volumen von Null liefern. Wir untersuchen daher die vierte Lösung $[y=45, z=45]$. Die zugehörige Hessematrix ist

ist $\begin{pmatrix} -180 & -90 \\ -90 & -180 \end{pmatrix}$, der Test auf Definitheit liefert wie erwartet ein

Maximum.

Die zulässige Länge des Behältnisses ergibt nach $x = 270 - 2y - 2z$ den Wert 90, das maximale Volumen beträgt 182,25.

8.2 Lagrange-Optimierung

Die oben dargestellte Methode des Einsetzungsverfahrens funktioniert nur, wenn es wie im obigen Beispiel gelingt, die Nebenbedingung eindeutig nach einer Variablen aufzulösen! In Aufgabe 7 der zugehörigen Übung soll das Maximalvolumen eines Quaders ermittelt werden, dessen Raumdiagonale gleich $2\sqrt{3}$ ist. Dabei ist es mathematisch gesehen nicht möglich, diese Nebenbedingung eindeutig nach einer Variablen aufzulösen, da hierbei eine Wurzel gezogen werden muss und damit sowohl der positive als auch der negative Wurzelwert als Lösungen erscheinen. Im konkreten Fall hilft der Bezug auf die zugrundeliegende geometrische Aufgabenstellung aus dem Dilemma: Da es sich um eine Länge handelt, kann diese selbstverständlich nicht negativ ist. Man kann somit in diesem Fall mit der positiven Wurzel weiterrechnen.

8.2.1 Herleitung des Lagrange-Verfahrens

Ist diese eindeutige Auflösung nicht möglich, so kommt das Lagrange-Verfahren zum Einsatz. Die grundsätzliche Funktionsweise dieses Verfahrens wird im Folgenden anhand eines einfachen Beispiels erläutert. In diesem ersten Beispiel ist die Nebenbedingung dergestalt, dass diese freilich eindeutig nach x oder y aufgelöst werden kann und damit auch das Einsetzungsverfahren zur Anwendung kommen könnte. Hier soll aber explizit an einem einfachen Beispiel das Lagrange-Verfahren erläutert werden. Wir gehen aus von der Zielfunktion

$$zf(x, y) = -2x^2 + 12x - y^2 + 8y - 4$$

und der Nebenbedingung

$$x + y = 2.$$

Diese Nebenbedingung können wir ebenfalls als zweiwertige Funktion $nb(x, y)$ formulieren

$$nb(x, y) = x + y$$

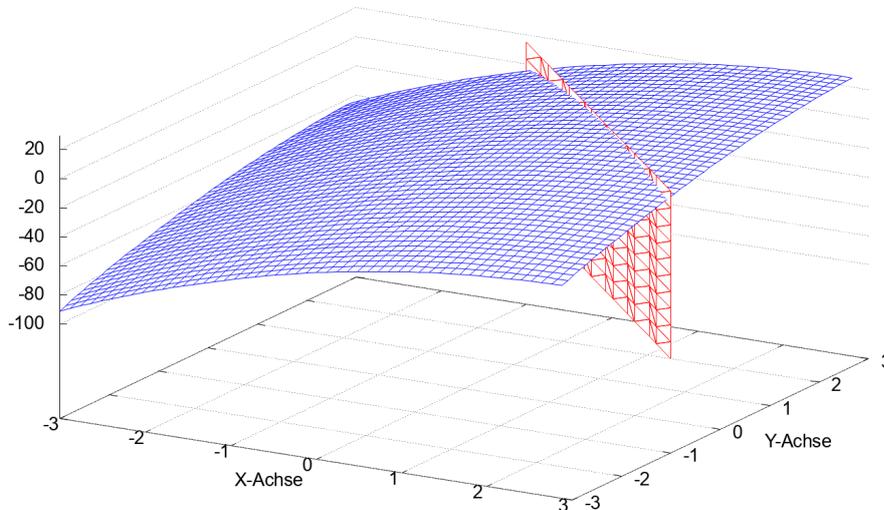
und fordern, dass der Funktionswert dieser Nebenbedingungs-Funktion gleich zwei sein muss!

Lassen wir uns diese Vorgaben graphisch darstellen, so erhalten wir aus unserer Zielfunktion ein nach unten geöffnetes Paraboloid und unsere Nebenbedingung stellt allgemein eine senkrecht auf xy -Ebene stehende Ebene dar, welche in der xy -Ebene die Funktionsgleichung

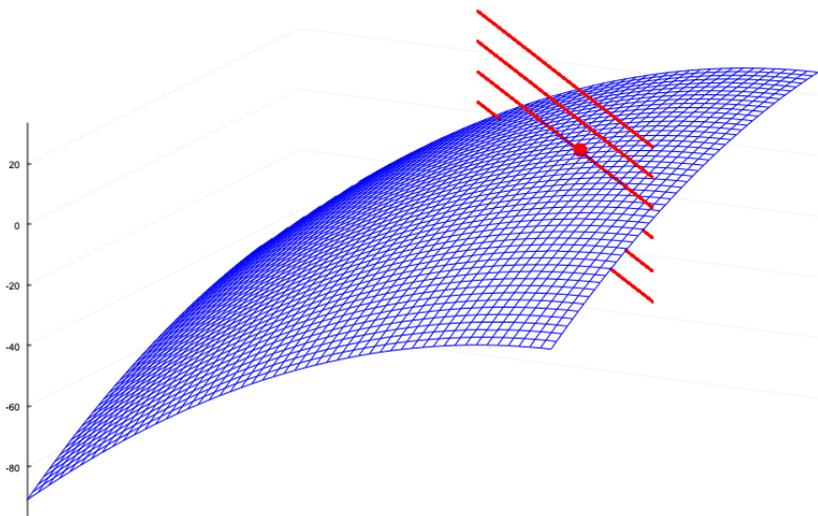
$$y = -x + 2$$

erfüllt. Diese Konstellation kommt im folgenden Plot zu Ausdruck. Sehr schön sichtbar ist dort auch bereits der Ansatz einer Lösungsidee: Die in rot dargestellte Nebenbedingung durchstößt den blauen Funktionsgraphen. Der höchste Punkt dieser Schnittlinie stellt genau das gesuchte Extremum – in diesem Fall ein Maximum – dar.

Diesen Punkt können wir finden, wenn wir die in die xy -Ebene gezeichnete Gerade sukzessive parallel nach oben verschieben und zwar so lange, bis diese Linie den Funktionsgraph exakt tangiert! Der Berührungspunkt dieser Tangente ist das gesuchte Maximum.



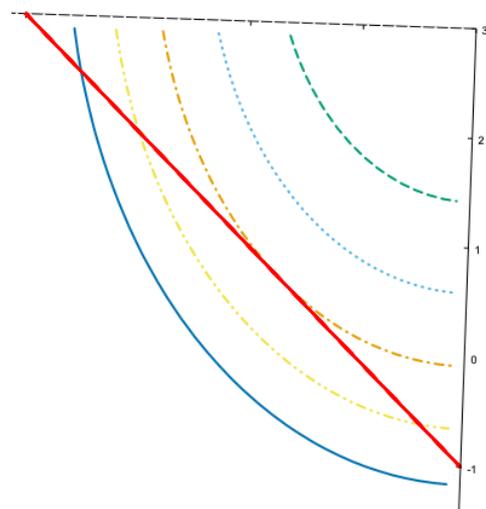
Dieses sukzessive Verschieben der Nebenbedingungs-Geraden nach oben ist in der nächsten Abbildung durch unterschiedliche Höhenlagen der Geraden angedeutet.



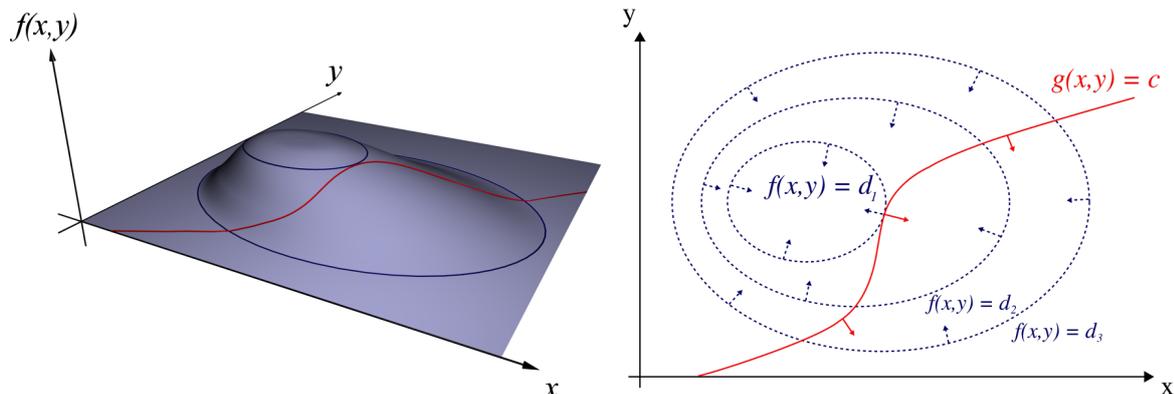
Die im Bild vierte Gerade von unten tangiert genau den Funktionsgraphen und liefert in ihrem Berührungspunkt das Maximum. Für eine Lösung unseres Optimierungsproblems muss es also darum gehen, genau diese Tangente und ihren Berührungspunkt zu finden!

Betrachten wir die Situation als „Wanderkarte“, also als Bild der Höhenlinien, so wird deutlich, dass die Gerade der Nebenbedingung im Maximum die dortige Höhenlinie der Funktion ebenfalls tangieren muss!

Wir hatten ja bereits herausgearbeitet, dass wir – an einem Berghang stehend – ganz unterschiedliche Steigungen wahrnehmen können, je nachdem, ob wir bergan, bergab, oder parallel zu den Höhenlinien schauen. Nur wenn wir parallel zu den Höhenlinien schauen, ist die Steigung gleich null.



Wenn wir genau entlang einer Höhenlinie laufen, wandern wir weder bergauf noch bergab! Solange wir nicht parallel zu den Höhenlinien laufen, sondern diese unter einem beliebigen Winkel schneiden, kommen wir auf unserer Wanderung immer entweder weiter bergan oder bergab. Auf unser Problem bezogen bedeutet dies, dass wir immer noch einen höher oder tiefer gelegenen Punkt erreichen können, wenn wir die Höhenlinien schneiden. Erst wenn wir parallel zu einer Höhenlinie laufen, geht es nicht mehr weiter bergauf oder bergab und wir haben auf unserem (vorgegebenen) Wanderweg die lokal höchste oder tiefste Stelle erreicht!



Die Voraussetzung, dass unser (durch die Nebenbedingung) vorgegebener Wanderweg genau parallel zu einer Höhenlinie verläuft und wir uns an einem Extremum befinden, bedeutet aber nichts anderes, als dass in diesem Extrempunkt die Gradienten sowohl der Zielfunktion als auch der Nebenbedingung in genau die gleiche Richtung weisen müssen! Formal ausgedrückt:

$$\nabla_z f(x,y) = -\lambda \cdot \nabla_{nb}(x,y)$$

bzw.

$$\nabla_z f(x,y) + \lambda \cdot \nabla_{nb}(x,y) = 0 \quad (1)$$

Der Faktor λ kommt ins Spiel, weil wir ja nur fordern, dass die Gradienten in die gleiche Richtung weisen. Dabei kann der Gradient der Nebenbedingung kürzer oder länger als der Gradient der Zielfunktion sein. Die formulierte Bedingung der gleichen Richtung der Gradienten ist aber nicht ausreichend, da es in aller Regel viele solcher Punkte gibt. Wir müssen natürlich auch noch fordern, dass unsere Nebenbedingungsfunktion einen ganz bestimmten Wert c einnimmt, also

$$nb(x,y) = c$$

bzw.

$$nb(x,y) - c = 0 \quad (2)$$

Die eben geschilderten Zusammenhänge soll Joseph-Louis Lagrange (1736–1813) als erster erkannt haben. Die beiden eben formulierten Voraussetzungen (1) und (2) werden von der nach ihm benannten *Lagrange-Funktion*

$$L(x,y,\lambda) = z f(x,y) + \lambda \cdot (nb(x,y) - c)$$

bzw. deren Gradient erfüllt. Unsere Voraussetzung (1) bedeutet als Vektorgleichung ausgeschrieben

$$\nabla_z f(x,y) + \lambda \cdot \nabla_{nb}(x,y) = \begin{bmatrix} z f_x(x,y) \\ z f_y(x,y) \end{bmatrix} + \lambda \cdot \begin{bmatrix} nb_x(x,y) \\ nb_y(x,y) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

und der Gradient der Lagrangefunktion ergibt

$$\Delta L(x, y, \lambda) = \begin{bmatrix} L_x(x, y, \lambda) \\ L_y(x, y, \lambda) \\ L_\lambda(x, y, \lambda) \end{bmatrix} = \begin{bmatrix} z f_x(x, y) + \lambda \cdot n b_x(x, y) \\ z f_y(x, y) + \lambda \cdot n b_y(x, y) \\ n b_\lambda(x, y) - c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Die jeweils blau markierten Terme in beiden Gleichungen stimmen in den Zeilen 1 und 2 überein, wobei der Gradient der Lagrange-Gleichung in der dritten rot markierten Zeile zusätzlich noch die Bedingung (2) enthält.

Dies bedeutet für uns konkret, dass wir aus der Zielfunktion und der Nullform der Nebenbedingung die Lagrange-Gleichung aufstellen und deren Ableitungen gleich Null setzen müssen. Konkret erhalten wir in unserem Beispiel die Lagrange-Funktion

$$L(x, y, \lambda) = -2x^2 + 12x - y^2 + 8y - 4 + \lambda \cdot (x + y - 2),$$

deren Ableitungen wir gleich Null setzen:

$$\begin{bmatrix} -4x + 12 + \lambda \\ -2y + 8 + \lambda \\ x + y - 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Wieder erhalten wir ein Gleichungssystem, welches wir lösen. Die Lösung hier lautet

$$\left[x = \frac{4}{3}, y = \frac{2}{3}, \lambda = -\frac{20}{3} \right].$$

Mit den Lösungen für x und y erhalten wir einen Punkt, der als Extrempunkt in Frage kommt. Dieser Punkt wird *stationärer* bzw. im Zusammenhang mit der Lagrange-Optimierung auch *kritischer Punkt* genannt. Ob es tatsächlich ein Extremum ist und falls ja, ob es sich um ein Maximum oder ein Minimum handelt, müssen wir natürlich noch überprüfen. Der erhaltene Wert für λ spielt bei betriebswirtschaftlichen Fragestellungen eine Rolle, da er – ähnlich wie die Elastizität – aussagen darüber macht, wie stark ausgeprägt das Extremum ist. Etwas umgangssprachlich formuliert gibt der λ -Wert an, wie steil oder flach die Flanken des Gipfels bzw. der Senke sind. Wir werden diesen Aspekt jedoch hier nicht weiter verfolgen.

8.2.2 Beispiel zum Lagrange-Verfahren

Hier noch ein weiteres Beispiel, in welchem es nicht möglich ist, die Nebenbedingung eindeutig in eine Variable umzuformen. Gesucht sind die Extremstellen der sogenannten Zielfunktion

$$z f(x, y) = 3x^2 - 5y^3$$

und zwar unter der Nebenbedingung

$$x^2 + y^2 = 4.$$

Diese Nebenbedingung lässt sich nun nicht eindeutig nach x oder y auflösen, es muss die Lagrange-Optimierung angewandt werden. Die Vorgehensweise wird anhand der oben dargestellten Herleitung des Verfahrens strukturiert dargestellt:

1. Schritt: Umstellen der Nebenbedingung in die Nullform

$$x^2 + y^2 - 4 = 0.$$

2. Schritt: Bilden der Lagrange-Funktion durch Anfügen des λ -fachen des Nebenbedingungs-terms an die Zielfunktion

$$L(x, y, \lambda) = 3x^2 - 5y^3 + (x^2 + y^2 - 4) \cdot \lambda$$

$$L(x, y, \lambda) = 3x^2 - 5y^3 + \lambda x^2 + \lambda y^2 - 4\lambda$$

Anmerkung: Gibt es mehr als eine Nebenbedingung, dann werden diese nach dem gleichen Schema angefügt, also beispielsweise:

$$L(x, y, \lambda) = 3x^2 - 5y^3 + (x^2 + y^2 - 4) \cdot \lambda_1 + (3 - 7y + 2) \cdot \lambda_2 + (y - 3) \cdot \lambda_3 + \dots$$

Die Vielfachen der Nebenbedingungen sind dabei verschieden, was oben in der Indizierung der Lambdas zum Ausdruck kommt.

3. Schritt: Bilden der partiellen Ableitungen der Lagrange-Funktion

$$\frac{\partial L}{\partial x} = 6x + 2\lambda x$$

$$\frac{\partial L}{\partial y} = -15y^2 + 2\lambda y$$

$$\frac{\partial L}{\partial \lambda} = x^2 + y^2 - 4$$

Die letzte Ableitung nach λ ergibt immer den Term der Nebenbedingung.

4. Schritt: Die ermittelten Ableitungen werden gleich null gesetzt.

$$6x + 2\lambda x = 0$$

$$-15y^2 + 2\lambda y = 0$$

$$x^2 + y^2 - 4 = 0$$

5. Schritt: Lösen des Gleichungssystems

Obwohl das vorliegende Gleichungssystem auch „von Hand“ lösbar ist, verwenden wir dazu Maxima. In jedem Fall erhalten wir sechs Lösungen:

$$E_1: [-2, 0, -3]$$

$$E_2: [2, 0, -3]$$

$$E_3: [0, 2, 15]$$

$$E_4: [0, -2, -15]$$

$$E_5: \left[\frac{4\sqrt{6}}{5}, -\frac{2}{5}, -3 \right]$$

$$E_6: \left[-\frac{4\sqrt{6}}{5}, -\frac{2}{5}, -3 \right]$$

Diese sechs stationären bzw. kritischen Punkte – nur die an der ersten und zweiten Stelle aufgeführten x - und y -Koordinaten sind interessant – sind Kandidaten für Extrempunkte. In einem letzten Schritt muss daher untersucht werden, welche dieser Punkte tatsächlich Extrema darstellen. Dies kann auf mehrere Arten geschehen.

6. Schritt: Bestimmung der konkreten Extrema

In vielen wirtschaftswissenschaftlichen Fragestellungen wird auf Auswertung der ermittelten stationären Punkte verzichtet oder aber nur die nachfolgend unter a) bzw. b) angesprochenen pragmatischen Methoden angewandt.

Möglichkeit a): Rechnerische Lösung über die Bestimmung der Funktionswerte

$$zf(-2, 0) = 12$$

$$zf(2, 0) = 12$$

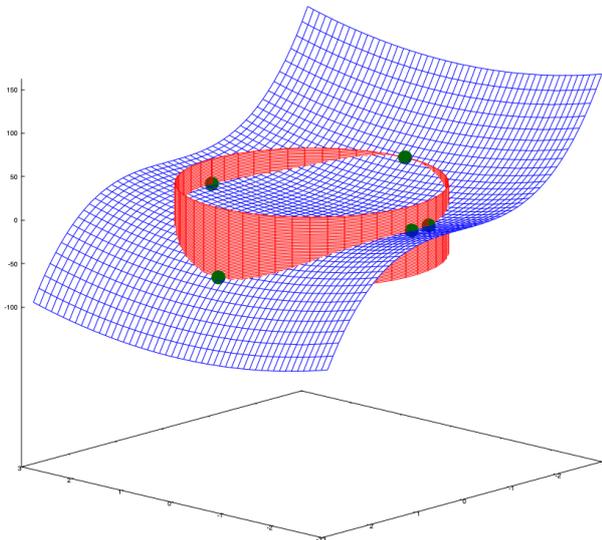
$$zf(0, 2) = -40 \qquad zf(0, -2) = 40$$

$$zf\left(\frac{4\sqrt{6}}{5}, -\frac{2}{5}\right) = 11,84 \qquad zf\left(-\frac{4\sqrt{6}}{5}, -\frac{2}{5}\right) = 11,84$$

Damit ist klar, dass der Punkt E_3 ein Minimum und der Punkt E_4 ein Maximum darstellt.

Möglichkeit b): Graphische Lösung

Zunächst einmal überlegt man sich, dass die Nebenbedingung $x^2 + y^2 = 4$ einen Kreis um den Koordinatenursprung der x - y -Ebene mit dem Radius 2 darstellt. Es kommen somit nur Extrema entlang der Kreislinie in Betracht. Plottet man die Zielfunktion und schneidet diese mit einem Zylinder um den Ursprung mit Radius 2, so wird deutlich, dass es ein Maximum und ein Minimum geben muss. Zeichnet man schließlich die Lage der Punkte am Kreis ein, so wird die Zuordnung zu den Extrema deutlich.



Möglichkeit c): Überprüfung mit Hilfe der „berandeten Hessematrix“. Diese werden wir jedoch im Rahmen dieser Veranstaltung zunächst nicht weiter thematisieren, da deren Anwendung tiefergehende Kenntnisse und die Anwendung spezieller Verfahren aus der linearen Algebra verlangt.

8.2.3 Bestimmung der Extrema mit Hilfe der berandeten Hessematrix

Eine formale Möglichkeit der Bestimmung der Extrempunkte geschieht wie im vorausgehenden Kapitel erläutert über die Hessematrix. Allerdings hatten wir uns dort bei der Auswertung der Hessematrix auf eine 2×2 -Matrix beschränkt deren Funktionsparameter immer x und y hießen. Diese Einschränkungen müssen wir aufgeben, da die Lagrangefunktion durch die Hinzunahme der Lambdas der Nebenbedingungen in aller Regel mehr als nur 2 unabhängige Variable besitzt, die zudem beliebige Namen haben können. Wir wollen das Verfahren der Bildung und Auswertung einer Hessematrix daher auf allgemeine Fälle erweitern.

Die Hessematrix fasst bekanntlich die zweiten Ableitungen einer Funktion strukturiert zusammen. Da Maxima für das Erstellen der Hessematrix keine Funktion zur Verfügung stellt, müssen wir hierfür selbst tätig werden. Die Aufgabenstellung ist prinzipiell einfach, wir müssen lediglich eine $n \times n$ -Matrix für eine n -wertige Funktion erstellen. Ausgangspunkt für deren erste Zeile ist die erste Ableitung nach der ersten Variablen. Diese erste Ableitung nach der ersten Variablen leiten wir nacheinander je einmal nach der ersten, nach der zweiten bis zur n -ten Variablen ab und füllen damit fortlaufend die Spalten der ersten Zeile.

Ausgangspunkt für die zweite Zeile ist die erste Ableitung nach der zweiten Variablen. Diese wird nun ebenfalls nach der ersten, zweiten, ... n -ten Variablen abgeleitet und mit diesen Ableitungen werden die Spalten der zweiten Zeile gefüllt und so geht es weiter bis zur ersten Ableitung nach der n -ten Variablen als Grundlage für die Ableitungen in der n -ten Matrixzeile.

Für die Erstellung einer konkreten Hessematrix ist es daher wesentlich anzugeben, was denn die erste, die zweite, ..., die n -te Variable ist. Da dies Maxima nicht aus der Variablenliste der

Funktion ermitteln kann, muss als zweiter Aufrufparameter eine Liste mit den Funktionsvariablen in der gewünschten Reihenfolge angegeben werden.

```
hessematrix(funktion, fvar) := block(
  [numvar, zweite_ab, hesse],
  numvar:length(fvar),
  mx:[],
  zweite_ab:[],
  hesse:matrix(),
  for j:1 thru numvar do
    (
      ab1:diff(funktion, fvar[1], 1),
      ab2:diff(ab1, fvar[j], 1),
      zweite_ab:endcons(ab2, zweite_ab)
    ),
  hesse:matrix(zweite_ab),
  zweite_ab:[],
  for i:2 thru numvar do
    (
      ab1:diff(funktion, fvar[i], 1),
      for j:1 thru numvar do
        (
          ab2:diff(ab1, fvar[j], 1),
          zweite_ab:endcons(ab2, zweite_ab)
        ),
      hesse:addrow(hesse, zweite_ab),
      zweite_ab:[]
    ),
  hesse);
```

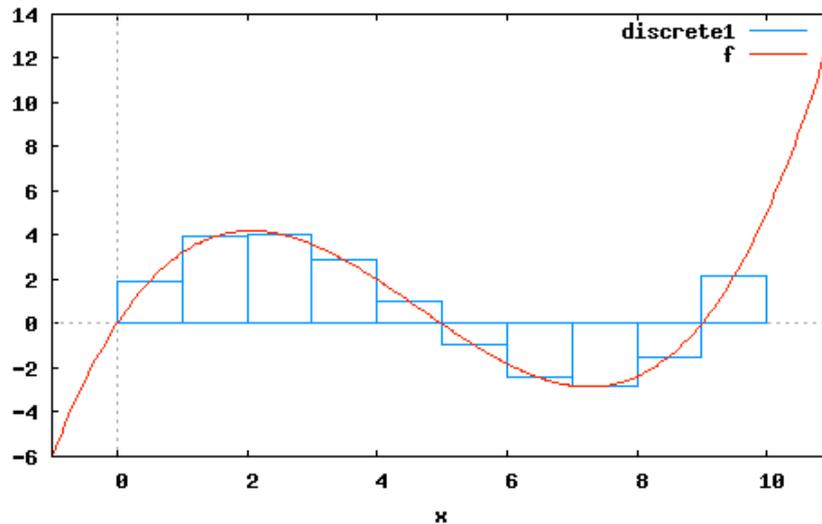
Aufgerufen wird die Funktion `hessematrix()` mit dem Funktionsnamen mitsamt den Parametern derjenigen Funktion, deren Hessematrix erstellt werden soll. Außerdem muss in einer Liste die gewünschte Reihenfolge der Funktionsvariablen angegeben werden.

(weiter siehe Röpcke/Wessler: Wirtschaftsmathematik (Hanser, 2019²) S. 258ff)

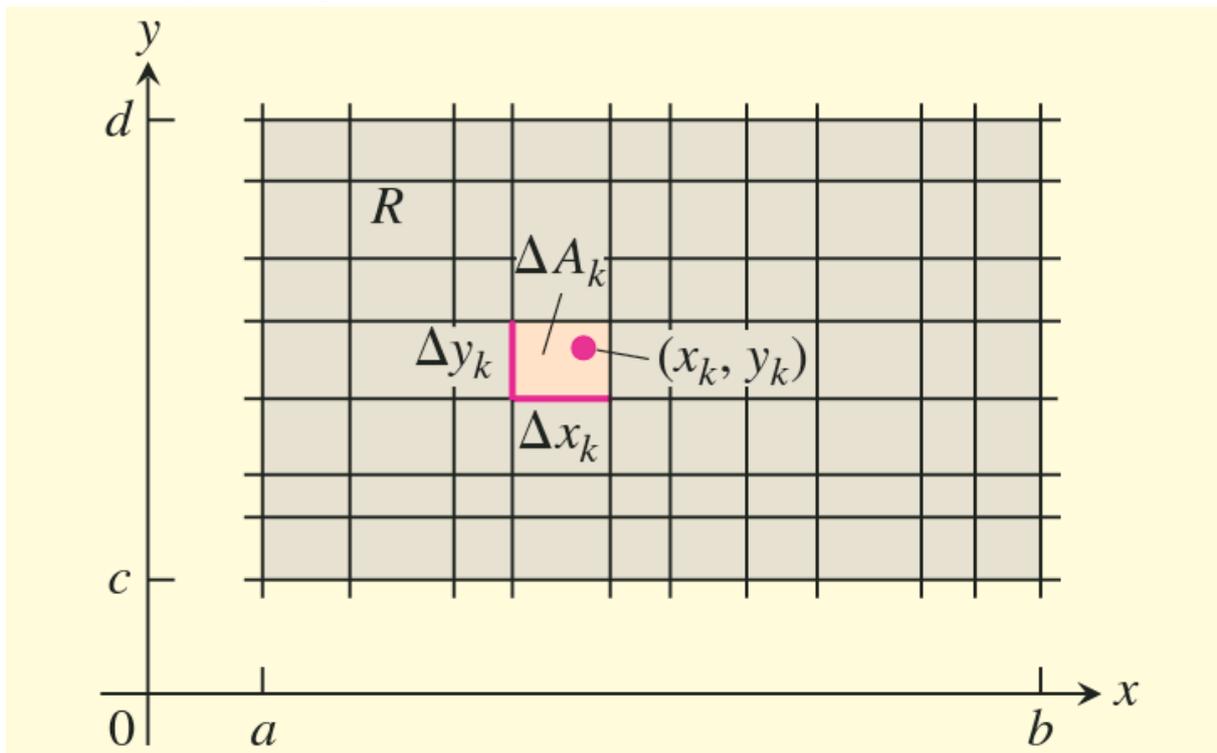
9 Doppelintegrale

9.1 Die grundlegende Idee

In der eindimensionalen Analysis haben wir ein Integral als Grenzwert der Summe von Rechteckstreifen unter dem Funktionsgraphen verstanden:

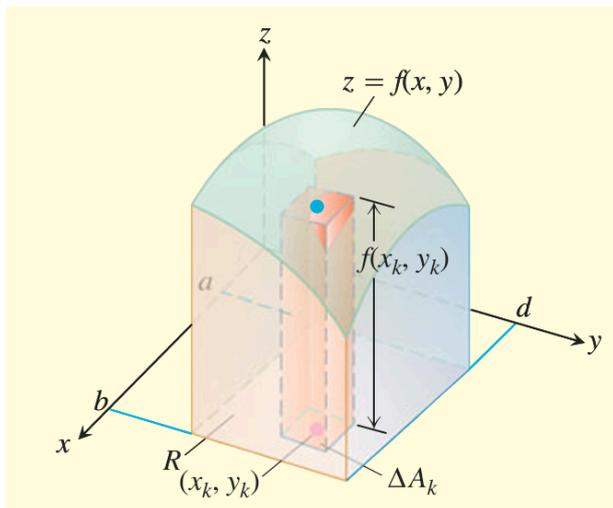


Das Aufsummieren der Rechteckstreifen führt zum Flächeninhalt unter der Kurve über der x -Achse. Um ein bestimmtes Integral zu berechnen, muss man dessen Grenzen angeben, also bestimmen, wo auf der x -Achse man beginnt und wo man endet. Diese Idee übertragen wir nun auf die zweidimensionale Analysis: Dort erhebt sich der Funktionsgraph als Fläche über der xy -Ebene. Um ein bestimmtes Integral berechnen zu können, muss man demnach ein Gebiet in der xy -Ebene bestimmen, über welchem man integrieren will. Für die Festlegung eines solchen Gebiets ist es zunächst am einfachsten, dieses als Rechteck zu wählen, dessen Seiten sich in x - und y -Richtung erstrecken.

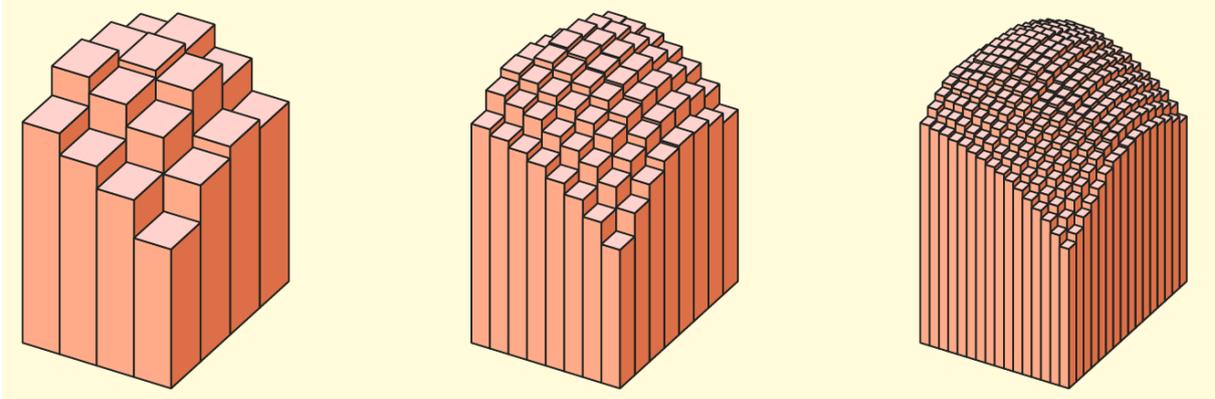


Man bestimmt somit einen Abschnitt ab der x -Achse und einen Abschnitt cd der y -Achse, welcher ein entsprechendes Rechteck definiert, über welchem integriert werden soll. Indem

nun die Strecken ab und cd unterteilt werden, wird das gesamte Rechteck in kleine Teilrechtecke unterteilt.



Misst man nun in jedem dieser Teilrechtecke nach oben (oder unten) bis zum Funktionsgraphen, so kann man durch Multiplikation der Grundfläche des Teilrechtecks mit der Höhe bis zum Graphen das Volumen des Rechteckprismas bestimmen und aus der Addition aller Prismen das Volumen zwischen der xy -Ebene und dem Funktionsgraphen über dem Rechteck $ab \times cd$ bestimmen. Dies geschieht natürlich umso genauer, je kleiner die Grundfläche der Rechteckprismen gewählt wird.



Dabei spielt es keine Rolle, ob innerhalb eines Teilrechtecks immer der größte oder der kleinste Funktionswert herangezogen wird, ob man immer den Funktionswert in der linken oberen oder rechten unteren Ecke wählt oder ob man die Höhe genau über dem Mittelpunkt des Teilrechtecks misst. Wir machen im Grenzprozess die Teilrechtecke beliebig klein, so dass schließlich alle Volumensummen gegen denselben Grenzwert streben. Dies ist jedenfalls bei den von uns betrachteten differenzierbaren Funktionen der Fall, die keine Löcher und senkrecht abfallenden Klüfte aufweisen. Existiert nun der Grenzwert der Volumensumme, dann nennen wir diese Summe das *Doppelintegral*.

Soweit die Idee in der Theorie. Um das Problem auch rechnerisch lösen zu können, verbinden wir gedanklich alle Prismen, die jeweils in y -Richtung nebeneinander liegen. Damit erhalten wir dünne Scheiben mit der Breite cd , die nach oben hin dem Funktionsgraphen treppenförmig angenähert sind. Wenn wir diese Scheiben direkt in x -Richtung betrachten, erhalten wir die bekannte Streifenform des Integrals aus der eindimensionalen Analysis und bewegen uns damit auf vertrautem Terrain.

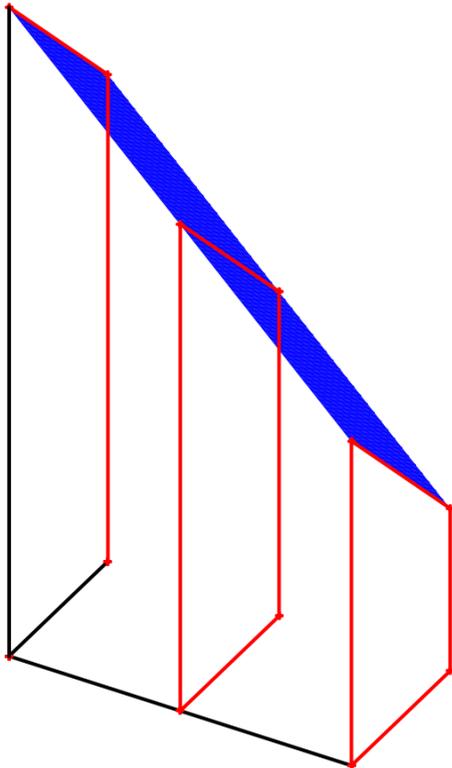
Wenn wir nun eine solche Scheibe infinitesimal dünn machen, dann können wir den Flächeninhalt dieser Scheibe mit Hilfe der Integralrechnung bestimmen. Das einzige, was wir dafür benötigen, ist die partielle Funktion an genau derjenigen Stelle x , an welcher sich diese Scheibe befindet.

Nun machen wir dies für all die unzählig vielen Scheiben die in x -Richtung von a bis b hintereinander stehen und summieren deren Flächen alle auf. Damit erhalten wir schließlich und endlich das Volumen im Rechteck $ab \times cd$ unter dem Funktionsgraphen.

9.2 Integrieren eines Volumens

9.2.1 Erster Ansatz

Wir werden in einem ersten Ansatz das Volumen des Körpers zwischen der Ebene $z=4-x-y$ über dem rechteckigen Gebiet $R: 0 \leq x \leq 2, 0 \leq y \leq 1$ in der xy -Ebene berechnen. Dabei handelt es sich um eine Ebene, deren Wert an der Stelle $[0, 0]$ gleich 4 ist und die sowohl in x - als auch in y -Richtung um -1 abfällt.



Ein Ausschnitt dieser Ebene ist in der nebenstehenden Grafik im angegebenen Ausschnitt blau dargestellt. Die x -Achse läuft nach rechts vorne, die y -Achse nach rechts hinten und die z -Achse nach oben. Rot eingezeichnet sind die Umrisse dreier Scheiben aus in y -Richtung nebeneinanderliegenden Prismen an den Stellen $x=0, x=1$ und $x=2$, deren Flächeninhalte wir berechnen wollen. Dazu bilden wir das (partielle) Integral der Funktion $f(x,y) = 4 - x - y$ nach der Variablen y :

$$\int f(x,y) dy = -\frac{1}{2}y^2 - xy + 4y$$

Zur Berechnung der Flächeninhalte benötigen wir die bestimmten Integrale jeweils im y -Intervall von 0 bis 1 an den Stellen $x=0, x=1$ und $x=2$:

$$\int_{y=0}^{y=1} f(x,y) dy = [F(x,y)]_0^1 = F(x,1) - F(x,0)$$

Wir berechnen ...

... die Fläche an der Stelle $x=0$:

$$F(0,1) - F(0,0) = \left(-\frac{1}{2} \cdot 1 - 0 \cdot 1 + 4 \cdot 1\right) - \left(-\frac{1}{2} \cdot 0 - 0 \cdot 0 + 4 \cdot 0\right) = 3,5 - 0$$

... die Fläche an der Stelle $x=1$:

$$F(1,1) - F(1,0) = \left(-\frac{1}{2} \cdot 1 - 1 \cdot 1 + 4 \cdot 1\right) - \left(-\frac{1}{2} \cdot 0 - 1 \cdot 0 + 4 \cdot 0\right) = 2,5 - 0$$

... die Fläche an der Stelle $x=2$:

$$F(2,1) - F(2,0) = \left(-\frac{1}{2} \cdot 1 - 2 \cdot 1 + 4 \cdot 1\right) - \left(-\frac{1}{2} \cdot 0 - 2 \cdot 0 + 4 \cdot 0\right) = 1,5 - 0$$

Die so errechneten Werte lassen sich leicht durch eine herkömmliche Flächeninhaltsberechnung von Trapezen bestätigen.

Ausgehend vom Trapez an der Stelle $x=0$ bis zum Trapez an der Stelle $x=2$ fallen somit die Flächeninhalte der Querschnittsflächen von 3,5 auf 1,5. Zur Bestimmung des Volumens gilt es nun, alle diese infinitesimal dünnen Querschnittsflächen $A(x)dx$ im Intervall $0 \leq x \leq 2$ aufzusummieren:

$$V = \int_{x=0}^{x=2} A(x) dx \quad (1)$$

Dabei ist $A(x)$ der Flächeninhalt der Querschnittsfläche bei x . Für jeden Wert von x lässt sich $A(x)$ – wie eben schon in drei Fällen gezeigt – mit dem Integral

$$A(x) = \int_{y=0}^{y=1} (4 - x - y) dy \quad (2)$$

berechnen. Berechnet man $A(x)$, so behandelt man x als festen Parameter und integriert über die Variable y . Kombiniert man die beiden Gleichungen durch Einsetzen von (2) in (1), so erhält man für das Gesamtvolumen des Körpers:

$$V = \int_{x=0}^{x=2} A(x) dx = \int_{x=0}^{x=2} \left(\int_{y=0}^{y=1} (4 - x - y) dy \right) dx$$

Wir bestimmen zunächst das innere, bestimmte Integral:

$$V = \int_{x=0}^{x=2} \left[4y - xy - \frac{y^2}{2} \right]_{y=0}^{y=1} dx = \int_{x=0}^{x=2} \left(\frac{7}{2} - x \right) dx$$

Nun kann das äußere Integral bestimmt werden:

$$V = \left[\frac{7}{2}x - \frac{x^2}{2} \right]_0^2 = 5$$

Damit haben wir das Volumen des Körpers konkret berechnet. Will man lediglich eine Gleichung für das Volumen aufstellen, aber keine der Integrationen wirklich ausführen, so schreibt man:

$$V = \int_0^2 \int_0^1 (4 - x - y) dy dx$$

Den Term auf der rechten Seite dieser Gleichung nennt man auch ein iteriertes Integral. Will man ihn berechnen, so integriert man zunächst $4 - x - y$ über y von $y=0$ bis $y=1$ und hält dabei x fest. Danach integriert man das Ergebnis über x von $x=0$ bis $x=2$. Man beachte, dass die Integrale geschachtelt geschrieben sind: Die Integrationsgrenzen 0 und 1 gehören in diesem Fall zum inneren Integral dy , die Grenzen 0 und 2 zum äußeren Integral dx .

9.2.2 Zweiter Ansatz

Wie haben in unserer obigen Rechnung Flächenstücke aufsummiert, welche in y -Richtung verlaufen. Diese Festlegung haben wir relativ willkürlich getroffen. Was ist, wenn wir Flächenstücke summieren, welche in x -Richtung verlaufen? In der untenstehenden Abbildung sind wieder drei solcher Flächenstücke eingezeichnet, sie befinden sich an den Stellen $y=0$, $y=0,5$ und $y=1$. Wieder handelt es sich um Trapeze, deren Flächeninhalt wir berechnen wollen.

Als Stammfunktion ermitteln wir

$$\int f(x, y) dx = -\frac{1}{2}x^2 - xy + 4x$$

und daraus die bestimmten Integrale an den Stellen y

$$\int_{x=0}^{x=2} f(x, y) dx = [F(x, y)]_0^2 = F(2, y) - F(0, y)$$

Wir berechnen konkret ...

... die Fläche an der Stelle $y=0$:

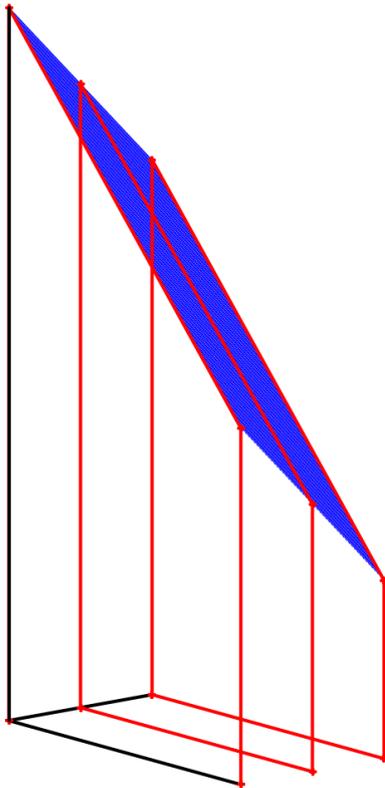
$$F(2, 0) - F(0, 0) = \left(-\frac{1}{2} \cdot 2^2 - 2 \cdot 0 + 4 \cdot 2 \right) - \left(-\frac{1}{2} \cdot 0^2 - 0 \cdot 0 + 4 \cdot 0 \right) = 6 - 0$$

... die Fläche an der Stelle $y=0,5$:

$$F\left(2, \frac{1}{2}\right) - F\left(0, \frac{1}{2}\right) = \left(-\frac{1}{2} \cdot 2^2 - 2 \cdot \frac{1}{2} + 4 \cdot 2\right) - \left(-\frac{1}{2} \cdot 0^2 - 0 \cdot \frac{1}{2} + 4 \cdot 0\right) = 5 - 0$$

... die Fläche an der Stelle $y=1$:

$$F(2,1) - F(0,1) = \left(-\frac{1}{2} \cdot 2^2 - 2 \cdot 1 + 4 \cdot 2\right) - \left(-\frac{1}{2} \cdot 0^2 - 0 \cdot 1 + 4 \cdot 0\right) = 4 - 0$$



Unsere Beispielrechnungen zeigen, dass der Flächeninhalt der Trapeze von 6 auf 4 fällt.

Um nun wieder das Volumen zu bestimmen, müssen wir alle infinitesimal dünnen Flächenstücke von $y = 0$ bis $y = 1$ aufsummieren.

Das Volumen berechnet sich nun folgendermaßen:

$$V = \int_{y=0}^{y=1} A(y) dy$$

Dabei ist $A(y)$ der Flächeninhalt der Querschnittsfläche bei y . Für jeden Wert von y lässt sich $A(y)$ – wie eben schon in drei Fällen gezeigt – mit dem Integral

$$A(y) = \int_{x=0}^{x=2} (4 - x - y) dx$$

berechnen.

Insgesamt ergibt sich damit für das Volumen:

$$V = \int_{y=0}^{y=1} A(y) dy = \int_{y=0}^{y=1} \left(\int_{x=0}^{x=2} (4 - x - y) dx \right) dy$$

Wieder errechnen wir zunächst das innere, bestimmte Integral:

$$V = \int_{y=0}^{y=1} \left[-\frac{1}{2} x^2 - xy + 4x \right]_0^2 dy = \int_{y=0}^{y=1} (6 - 2y) dy$$

Nun ist das äußere Integral an der Reihe:

$$V = \left[6y - y^2 \right]_0^1 = 6 - 1^2 = 5$$

Damit erhalten wir denselben Wert für das Volumen des Körpers. Es scheint also unerheblich zu sein, ob wie die Flächen in x - oder in y -Richtung anordnen und aufsummieren.

Die Gleichung für das iterierte Integral lautet in diesem Fall

$$V = \int_0^1 \int_0^2 (4 - x - y) dx dy.$$

In diesem Fall berechnen wir den Term auf der rechten Seite der Gleichung, indem wir $(4 - x - y)$ mit der Integrationsvariablen x und den Integrationsgrenzen $x = 0$ und $x = 2$ integrieren.

Das Ergebnis integrieren wir danach über y mit den Grenzen $y = 0$ und $y = 1$.

9.2.3 Satz von Fubini

Es ist unmittelbar einleuchtend, dass es für die Berechnung eines Volumens mit rechteckiger Grundfläche unter einem Funktionsgraph unerheblich ist, ob das Volumen in dünne Scheiben zerlegt wird, die in x -Richtung oder in y -Richtung verlaufen. Den mathematischen Beweis hierfür hat der italienische Mathematiker Guido Fubini 1907 geführt.

Es sei $f(x, y)$ stetig auf dem rechteckigen Gebiet $R: a \leq x \leq b, c \leq y \leq d$. Dann gilt:

$$\iint_R f(x,y) dA = \int_c^d \int_a^b f(x,y) dx dy = \int_a^b \int_c^d f(x,y) dy dx$$

Dieser Satz von Fubini besagt, dass sich Doppelintegrale über Rechtecken als iterierte Integrale bestimmen lassen. Wir können somit ein Doppelintegral berechnen, indem wir nacheinander über jeweils eine Variable integrieren. Der Satz besagt außerdem, dass die Reihenfolge dieser Integrationen beliebig ist, was viele Rechnungen deutlich erleichtert.

9.2.4 Ein weiteres Beispiel

Berechnen Sie das Volumen des Körpers zwischen dem Graph der Funktion $f(x,y) = 10 + x^2 + 3y^2$ und dem Rechteck $R: 0 \leq x \leq 1, 0 \leq y \leq 2$ in der xy -Ebene.

$$\begin{aligned} V &= \iint_R (10 + x^2 + 3y^2) = \int_0^1 \int_0^2 (10 + x^2 + 3y^2) dy dx \\ &= \int_0^1 [10y + x^2y + y^3]_{y=0}^{y=2} dx = \int_0^1 (20 + 2x^2 + 8) dx = \left[20x + \frac{2}{3}x^3 + 8x \right]_0^1 = \frac{86}{3} \end{aligned}$$

9.2.5 Berechnung eines Doppelintegrals in Maxima

Maxima enthält von Haus aus keine Funktion zur Berechnung eines Doppelintegrals. Da es sich aber um zwei nacheinander ausgeführte Integrationen handelt, kann die konkrete Berechnung eines Doppelintegrals genau wie in den obigen Beispielen erläutert auch in Maxima durchgeführt werden. Nachfolgend wird der Rechenweg in Maxima anhand des eben gelösten weiteren Beispiels demonstriert. Dazu definieren wir zunächst die zugrundeliegende Funktion:

$$f(x, y) := 10 + x^2 + 3 * y^2$$

Laut Aufgabenstellung soll zuerst das innere Integral nach dy in den Grenzen von 0 bis 2 ermittelt werden. Dazu definieren wir die zugehörige Stammfunktion $Fy(x,y)$:

$$\text{define}(Fy(x, y), \text{integrate}(f(x, y), y))$$

und erhalten:

$$Fy(x,y) := y^3 + x^2y + 10y$$

Dann definieren wir das bestimmte Intervall in den angegebenen Grenzen:

$$\text{define}(Fy_num(x, y), Fy(x, 2) - Fy(x, 0))$$

Das Ergebnis lautet:

$$Fy_num(x,y) := 2x^2 + 28$$

Dies ist nun das innere Integral, welches nun über dx in den Grenzen von 0 bis 1 integriert wird. Zunächst lassen wir uns die Stammfunktion ermitteln:

$$\text{define}(Fyx(x, y), \text{integrate}(Fy_num(x, y), x))$$

Diese lautet:

$$Fyx(x,y) := \frac{2x^3}{3} + 28x$$

Schließlich muss noch in den angegebenen Grenzen integriert werden. Dies machen wir hier direkt über die Differenz „obere Grenze minus untere Grenze“

$$Fyx(1, y) - Fyx(0, y)$$

und erhalten als Ergebnis $\frac{86}{3} \approx 28,67$.

Die aufgeführten Schritte können zu einer Funktion

`doppelintegral_yx(xrange, yrange)`

zusammengefasst werden. Darin müssen lediglich die oben konkret angegebenen Intervallgrenzen als Variable `xo`, `xu` sowie `yo` und `yu` verändert werden. Die andere Integrationsreihenfolge kann dann leicht angepasst werden.