

Partielle Ableitungen, Gradient, Lineare Näherung, Extrema, Fehlerfortpflanzung

Jörn Loviscach

Versionsstand: 29. Juni 2009, 18:41

1 Partielle Ableitungen, Gradient

Die Ableitung einer Funktion f an einer Stelle x gibt an, wie sich die Funktion ändert, wenn man einen „unendlich“ kleinen Schritt von x vorwärts oder rückwärts macht. Bei einer Funktion mehrerer Unabhängiger kann man diesen Schritt aber auch sozusagen quer machen. Das einfachste ist, entlang einer der Achsen zu schreiten. Das ergibt die jeweilige partielle Ableitung [partial derivative]: Man behandelt alle Unabhängige bis auf eine als Konstanten und leitet ganz normal nach der einen ab. Das Symbol dafür ist der Differentialquotient mit einem geschwungenen d, also ∂ .

Beispiele:

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

1

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

2

- $f(u, v, w) = uv^2 + ve^{3w}$

3

Wenn man die partiellen Ableitungen zu einem Vektor übereinanderstellt, hat man den Gradienten der Funktion, geschrieben $\text{grad}f$ oder ∇f , öfters auch mit Vektorpfeilen zu sehen: $\vec{\text{grad}}f$, $\vec{\nabla}f$. Das Symbol ∇ heißt „Nabla“ und kommt auch noch in anderen Zusammenhängen vor. Für die drei Beispiele ist der Gradient also:

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



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



- $f(u, v, w) = uv^2 + ve^{3w}$



Hier wird jedem Punkt des Definitionsbereichs nicht eine Zahl zugeordnet, sondern ein Vektor: Das sind Vektorfelder.

Wolfram Alpha scheint diese derzeit erst in zwei Schritten plotten zu können:

```
grad x^2+y^2
vectorfield (2x, 2y)
```

In Octave sieht das so aus:

```
x = [-3 : .5 : 3];
y = [-3 : .5 : 3];
[xx, yy] = meshgrid(x, y);
z = xx.^2.+yy.^2;
[gx, gy] = gradient(z, .5, .5);
quiver(x, y, gx, gy)
hold on, contour(x, y, z), hold off
```

Hier erkennt man schon die Bedeutung des Gradienten: Je länger er an einer Stelle ist, umso steiler ist die Funktion hier. Er zeigt in die Richtung des (örtlich!) steilsten Anstiegs der Funktion und steht deswegen senkrecht auf den Höhenlinien:



Achtung: Wie die Höhenlinien lebt der Gradient einer Funktion von n Veränderlichen im \mathbb{R}^n , also sozusagen in der Ebene der Landkarte. Er zeigt *nicht* steil den Berg hinauf!

Streng genommen müsste man diskutieren, was es heißt, dass eine Funktion von n Veränderlichen „total“ differenzierbar ist – also nicht nur partiell differenzierbar. Netterweise ist das in der Praxis kein Problem: Wenn auf einer Umgebung einer Stelle alle partiellen Ableitungen existieren und stetig sind, kann man in beliebige Richtungen ableiten.

2 Lineare Näherung, Tangentialebene

Erinnerung: Die lineare Näherung = Tangentengerade an der Stelle x_0 einer differenzierbaren Funktion f einer einzigen Veränderlichen war:

8

Eine differenzierbare Funktion f von zwei Veränderlichen hat an einer Stelle \mathbf{x}_0 als lineare Näherung eine Tangentenebene:

9

Deren Gleichung ist nicht allzu überraschend:

10

Wenn man das komplett ausbuchstabiert, ergibt sich:

11

Das schreibt man auch – eigentlich nur zum Angucken – als das „totale Differential“ df der Funktion f :

12

Entsprechend mit n Veränderlichen.

3 Fehlerfortpflanzung bei mehreren Veränderlichen

In viele Anwendungsproblemen bestimmt man eine Größe aus mehreren anderen, als $f(x_1, x_2, \dots, x_n)$. Mit Hilfe der linearen Näherung kann man nun schätzen, wie sich Fehler in den Ausgangsgrößen x_1, \dots, x_n auf den errechneten Wert von f auswirken.

Angenommen, die Ausgangsgrößen schwanken nur so stark, dass man f durch die lineare Näherung ersetzen kann. Ebenfalls angenommen, die Ausgangsgrößen schwanken voneinander unkorreliert; es soll also zum Beispiel nicht immer x_2 zu hoch sein, wenn auch x_1 zu hoch ist.

An zwei Veränderlichen untersucht: x schwanke mit der Breite (Standardabweichung, sozusagen der RMS-Fehler) σ_x um den Mittelwert \bar{x} , entsprechend für y . Dann ist $z = f(x, y)$ etwa:

13

und für seine Standardabweichung gilt:

14

Entsprechend mit n Veränderlichen.

4 Extrema von Funktionen mehrerer Veränderlicher

Wie bei Funktionen von einer Veränderlichen untersucht man Stellen^{c1} im Inneren des Definitionsbereichs, an denen der Funktionswert entweder größer oder aber kleiner ist als alle Funktionswerte in einer Umgebung^{c2}: lokale Extrema. Das globale Maximum oder globale Minimum wird – falls es existiert! – ein solches lokales Maximum sein oder aber am Rand des Definitionsbereichs liegen. Man hat also eine (hoffentlich) überschaubare Sammlung an Kandidaten, von denen man dann zu Fuß den größten bzw. den kleinsten sucht.

^{c1}jl: Extrema (Maxima und Minima)

^{c2} text added by jl

Damit an einer Stelle \mathbf{x} im Inneren des Definitionsbereichs ein lokales Extremum liegen kann, muss gelten (notwendige Bedingung):

15

Diese Bedingung ist allerdings nicht hinreichend, wie diese geometrischen Situationen zeigen:

16

Allerdings kann man einfach die Werte der Funktion f an allen solchen Stellen ausrechnen und darin nach dem größten bzw. den kleinsten suchen, denn die lokalen Extrema sind ja sicher dabei.

Wenn man die zweite Ableitung weiß, kann man typischerweise mit Bestimmtheit sagen, ob ein Punkt \mathbf{x} , der die notwendige Bedingung erfüllt, auch wirklich ein lokales Extremum ist (hinreichende Bedingung) – und ob es sich um

ein lokales Maximum oder ein lokales Minimum handelt. Bei einer Funktion einer Veränderlichen kommt es hier auf das Vorzeichen der zweiten Ableitung an:

17

Bei einer Funktion mehrerer Veränderlicher bildet man statt dessen diese Matrix (Hesse-Matrix) aller doppelten partiellen Ableitungen:

18

Diese Matrix ist immer symmetrisch. Wenn die notwendige Bedingung (s. o.) erfüllt ist und alle Eigenwerte dieser Matrix positiv sind, muss es sich um ein lokales Minimum (Minimum, nicht Maximum!) handeln, wenn alle Eigenwerte negativ sind, um ein lokales Maximum (Maximum!). Es gibt Hilfskriterien, um schnell festzustellen, wann das der Fall ist.