Ein Buch

Für Anwendungen des Stoffs in den Wirtschaftswissenschaften:

Knut Sydsæter, Peter Hemmond:

Mathematik für Wirtschaftswissenschaftler

(Aber bei der Mathematik ein bisschen aufpassen!)

§4 Extremstellen

Maxima und Minima

Sei f eine Funktion in Variablen $x_1, \ldots, x_d$. Wir machen im Moment keine Annahme über den Definitionsbereich D_f.

Konvention. Wenn wir von einem Punkt $\underline{x}$ im Definitionsbereich reden, sprechen wir von einer **Stelle**. Wir haben dann den **Wert** $f(\underline{x})$.

Maxima und Minima

Definitionen

Eine **(globale) Maximalstelle** ist eine Stelle $\underline{x}^{(0)}$ für welche gilt:

Für alle $\underline{x} \in D_f$ ist

$$f(\underline{x}^{(0)}) \geq f(\underline{x}) \, .$$

$$\forall x \in D_f : f(\underline{x}^{(0)}) \geq f(\underline{x}) \, .$$

Der Wert $f(\underline{x}^{(0)})$ heißt das **Maximum** von f.

Wir sagen:

An $\underline{x}^{(0)}$ nimmt die Funktion f ihr **Maximum** an.

Die Funktion f wird / ist an $\underline{x}^{(0)}$ **maximal**.

Maxima und Minima

Eine **(globale) Miminalstelle** ist eine Stelle $\underline{x}^{(0)}$ für welche gilt:

Für alle $\underline{x} \in D_f$ ist

$$f(\underline{x}^{(0)}) \leq f(\underline{x}) \; .$$

$$\forall x \in D_f : f(\underline{x}^{(0)}) \leq f(\underline{x}) \; .$$

An so einer Stelle nimmt die Funktion f ihr **Minimum** an.

Sowohl globale Maximal- als auch globale Minimalstellen werden **globale Extremstellen** genannt.

Maxima und Minima

Eine Stelle $\underline{x}^{(0)}$ heißt **lokale Maximalstelle**, falls es ein $\epsilon > 0$ gibt, so dass für alle $\underline{x} \in D_f \cap \overline{U}_\epsilon(\underline{x}^{(0)})$ gilt:

$$f(\underline{x}^{(0)}) \geq f(\underline{x})$$

Dies kann man auch so ausdrücken:

Es gibt ein $\epsilon > 0$ derart, dass $\underline{x}^{(0)}$ eine globale Maximalstelle von $f|_{D_f \cap \overline{U}_\epsilon(\underline{x}^{(0)})}$ ist.

Maxima und Minima

Wenn $\underline{x}^{(0)}$ eine **lokale Maximalstelle** ist:

Der Wert $f(\underline{x}^{(0)})$ heißt ein **lokales Maximum** von f.

Wir sagen:

An / in $\underline{x}^{(0)}$ nimmt die Funktion f ein **lokales Maximum** an.

An $\underline{x}^{(0)}$ liegt ein **lokales Maximum** vor.

Wir haben ein **lokales Maximum** an $\underline{x}^{(0)}$.

Maxima und Minima

Eine Stelle $\underline{x}^{(0)}$ heißt **lokale Minimalstelle**, falls es ein $\epsilon > 0$ gibt, so dass für alle $\underline{x} \in D_f \cap \overline{U}_\epsilon(\underline{x}^{(0)})$ gilt:

$$f(\underline{x}^{(0)}) \geq f(\underline{x})$$

Dies kann man auch so ausdrücken:

Es gibt ein $\epsilon > 0$ derart, dass $\underline{x}^{(0)}$ eine globale Minimalstelle von $f|_{\overline{U}_\epsilon(\underline{x}^{(0)})}$ ist.

Maxima und Minima

Frage. Wie kann man Maxima und Minima berechnen?

Eine Variable

Wiederholung. Es sei f eine differenzierbare Funktion in einer Variablen x auf einem offenen Intervall (a, b). (Die Voraussetzung ist wichtig!)

Sei $x^{(0)} \in D_f = (a, b)$.

Wenn nun $x^{(0)}$ eine lokale Extremstelle ist, **dann** ist

$$f'(x^{(0)}) = 0 \, .$$

Dies ist ein **notwendiges Kriterium** dafür, dass $x^{(0)}$ eine lokale Extremstelle ist.

Eine Variable

Eine Stelle $x^{(0)}$ mit $f'(x^{(0)}) = 0$ heißt eine **kritische Stelle** oder auch **stationäre Stelle**.

Also: Wenn $x^{(0)}$ eine lokale Extremstelle ist, ist (notwendigerweise) $x^{(0)}$ eine kritische Stelle.

Eine Variable

Es sei nun f zweimal differenzierbar.

Wenn nun

$$f'(x^{(0)}) = 0 \text{ und } f''(x^{(0)}) < 0$$

ist, **dann** ist $x^{(0)}$ ein lokale Maximalstelle.

Dies ist ein **hinreichendes (=ausreichendes) Kriterium** dafür, dass $x^{(0)}$ eine lokale Maximalstelle ist.

Eine Variable

Wenn nun

$$f'(x^{(0)}) = 0 \text{ und } f''(x^{(0)}) > 0$$

ist, **dann** ist $x^{(0)}$ ein lokale Minimalstelle.

Dies ist ein **hinreichendes (=ausreichendes) Kriterium** dafür, dass $x^{(0)}$ eine lokale Minimalstelle ist.

Eine Variable

Notation

Es sei f eine zweimal differenzierbare Funktion. Dann bezeichnen wir f'' mit

$$\frac{\mathrm{d}^2 f}{\mathrm{d}x^2}$$

Frage. Warum d^2 in Zähler und $\mathrm{d}x^2$ im Nenner?

Antwort. Wenn wir "ganz formal" vorgehen, erhalten wir für die zweite Ableitung

$$\frac{\mathrm{d}\frac{\mathrm{d}f}{\mathrm{d}x}}{\mathrm{d}x} = \frac{\mathrm{d}\mathrm{d}f}{\mathrm{d}x\,\mathrm{d}x} = \frac{\mathrm{d}^2 f}{(\mathrm{d}x)^2}\,.$$

Eine Variable

Mit dieser Notation:

- Wenn x eine lokale Extremstelle ist, dann ist
$$\frac{\mathrm{d}f}{\mathrm{d}x}(x^{(0)}) = 0 \,.$$
- Wenn
$$\frac{\mathrm{d}f}{\mathrm{d}x}(x^{(0)}) = 0 \text{ und } \frac{\mathrm{d}^2 f}{\mathrm{d}x^2}(x^{(0)}) < 0$$
ist, dann ist $x^{(0)}$ eine lokale Maximalstelle.
- Wenn
$$\frac{\mathrm{d}f}{\mathrm{d}x}(x^{(0)}) = 0 \text{ und } \frac{\mathrm{d}^2 f}{\mathrm{d}x^2}(x^{(0)}) > 0$$
ist, dann ist $x^{(0)}$ eine lokale Minimalstelle.

Mehrere Variablen

Es sei f eine Funktion in Variablen $x_1, \ldots, x_n$ mit der üblichen wichtigen Voraussetzung:

Der Definitionsbereich D_f hat keinen Rand.

Dies heißt:

Für jedes $x^{(0)} \in D_f$ gibt es ein $\varepsilon > 0$ mit

$$\overline{U}_\varepsilon(\underline{x}^{(0)}) \subseteq D_f \ .$$

"Wir können von jedem Punkt aus 'ein bisschen' in jede Richtung gehen."

Extremstellen und Richtungen

Grundlegende Idee

Es sei $\underline{x}^{(0)} \in D_f$.

Angenommen, $\underline{x}^{(0)}$ ist eine lokale Maximalstelle.

Wir wählen nun irgendeine Richtung $\vec{r}$. Nun darf f in diese Richtung nicht zunehmen.

Dies bedeutet:

Wir betrachten die Bewegung

$$\gamma : t \mapsto \underline{x}^{(0)} + t\vec{r}$$

und die Funktion

$$f \circ \gamma : t \mapsto f\left(\underline{x}^{(0)} + t\vec{r}\right) .$$

Nun muss diese Funktion ein lokales Maximum an $t = 0$ haben.

Extremstellen und Richtungen

Wir betrachten die Bewegung

$$\gamma : t \mapsto \underline{x}^{(0)} + t\vec{r}$$

und die Funktion

$$f \circ \gamma : t \mapsto f\left(\underline{x}^{(0)} + t\vec{r}\right) .$$

Nun muss diese Funktion ein lokales Maximum an $t = 0$ haben.

Somit insbesondere:

$$\frac{\mathrm{d}f(\gamma(t))}{\mathrm{d}t}\Big|_{t=0} = 0$$

Nun ist nach Definition

$$\frac{\mathrm{d}f(\gamma(t))}{\mathrm{d}t}\Big|_{t=0} = \frac{\partial f}{\partial \vec{r}}\Big|_{\underline{x}=\underline{x}^{(0)}} .$$

Also:

$$\frac{\partial f}{\partial \vec{r}}\Big|_{\underline{x}=\underline{x}^{(0)}} = 0$$

Extremstellen und Richtungen

Also:

Wenn $\underline{x}^{(0)}$ ein lokales Maximum ist, dann gilt für alle Richtungen $\vec{r}$:

$$\partial_{\vec{r}}(f)(\underline{x}^{(0)}) = 0$$

Dann muss

$$(\nabla f)(\underline{x}^{(0)}) = \vec{0}$$

gelten.

Denn: Wenn $(\nabla f)(\underline{x}^{(0)}) \neq \vec{0}$ ist, dann gilt für die Richtung $\vec{r} := \frac{1}{\|(\nabla f)(\underline{x}^{(0)})\|} \cdot \nabla(f)(\underline{x}^{(0)})$:

$$\partial_{\vec{r}}(f)(\underline{x}^{(0)}) \neq 0$$

Extremstellen und Richtungen

Dasselbe gilt auch für Minimalstellen.

Kritische Stellen

Erstes Resultat

Wenn $\underline{x}^{(0)}$ eine lokale Extremstelle ist, **dann** ist $(\nabla f)(\underline{x}^{(0)}) = \vec{0}$.

Somit ist

$$\nabla f(\underline{x}^{(0)}) = \vec{0}$$

eine **notwendige Bedingung** dafür, dass $\underline{x}^{(0)}$ eine lokale Extremstelle ist.

Kritische Stellen

Eine Stelle $\underline{x}^{(0)}$ mit $\nabla f(\underline{x}^{(0)}) = \vec{0}$ heißt **kritische Stelle** oder **stationäre Stelle**.

Also: Um lokale Extremstellen zu bestimmen, bestimmt man zuerst die kritischen Stellen. Hierfür löst man

$$\nabla f(\underline{x}) = \vec{0}$$

D.h. man bestimmt die Lösungen des Gleichungssystems

$$\partial_{x_1} f(\underline{x}) = 0, \ldots, \partial_{x_d} f(\underline{x}) = 0 \; .$$

Das kann schwierig / unmöglich sein ...

Kritische Stellen

Eine kritische Stelle, die keine lokale Extremstelle ist, heißt **Sattelstelle** (**Sattelpunkt**).

Für eine kritische Stelle $\underline{x}^{(0)}$ sind äquivalent:

- $\underline{x}^{(0)}$ ist eine Sattelstelle.
- In jeder ε-Umgebung von $\underline{x}^{(0)}$ liegt eine Stelle mit größerem Wert und eine Stelle mit kleinerem Wert als $f(\underline{x}^{(0)})$.
- Es gibt eine Folge $\underline{a}^{(n)}$ mit Grenzwert $\underline{x}^{(0)}$ und $f(\underline{a}) > f(\underline{x}^{(0)})$ und
  es gibt eine Folge $\underline{b}^{(n)}$ mit Grenzwert $\underline{x}^{(0)}$ und $f(\underline{a}) < f(\underline{x}^{(0)})$.

Beispiele

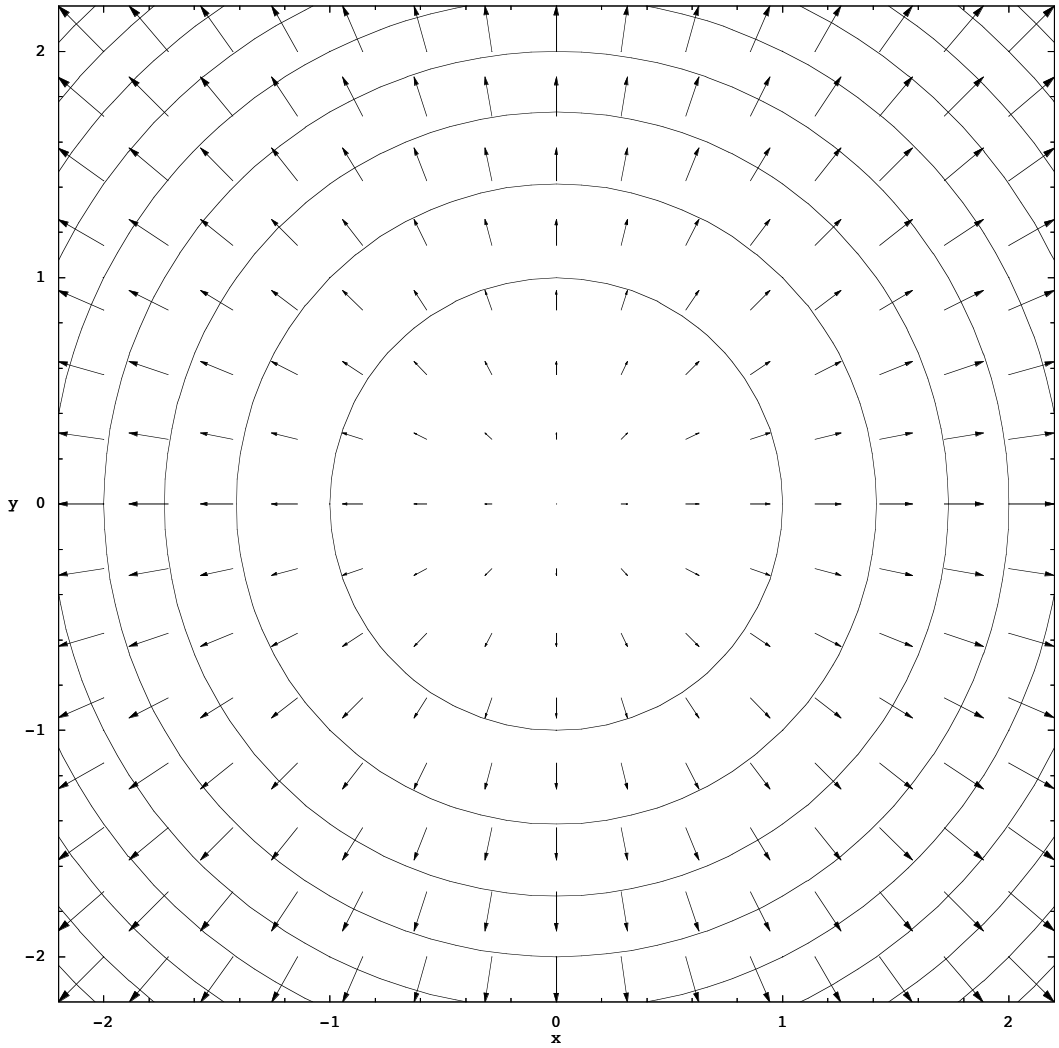
Wir betrachten f mit

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

Die Funktion hat an $(0, 0)$ ein globales Minimum und keine weiteren Extremstellen.

Oder nicht?

Examples of gradient vectors and level curves



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

Beispiele

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

$$\nabla f = \begin{pmatrix} 2x \\ 2y \end{pmatrix}$$

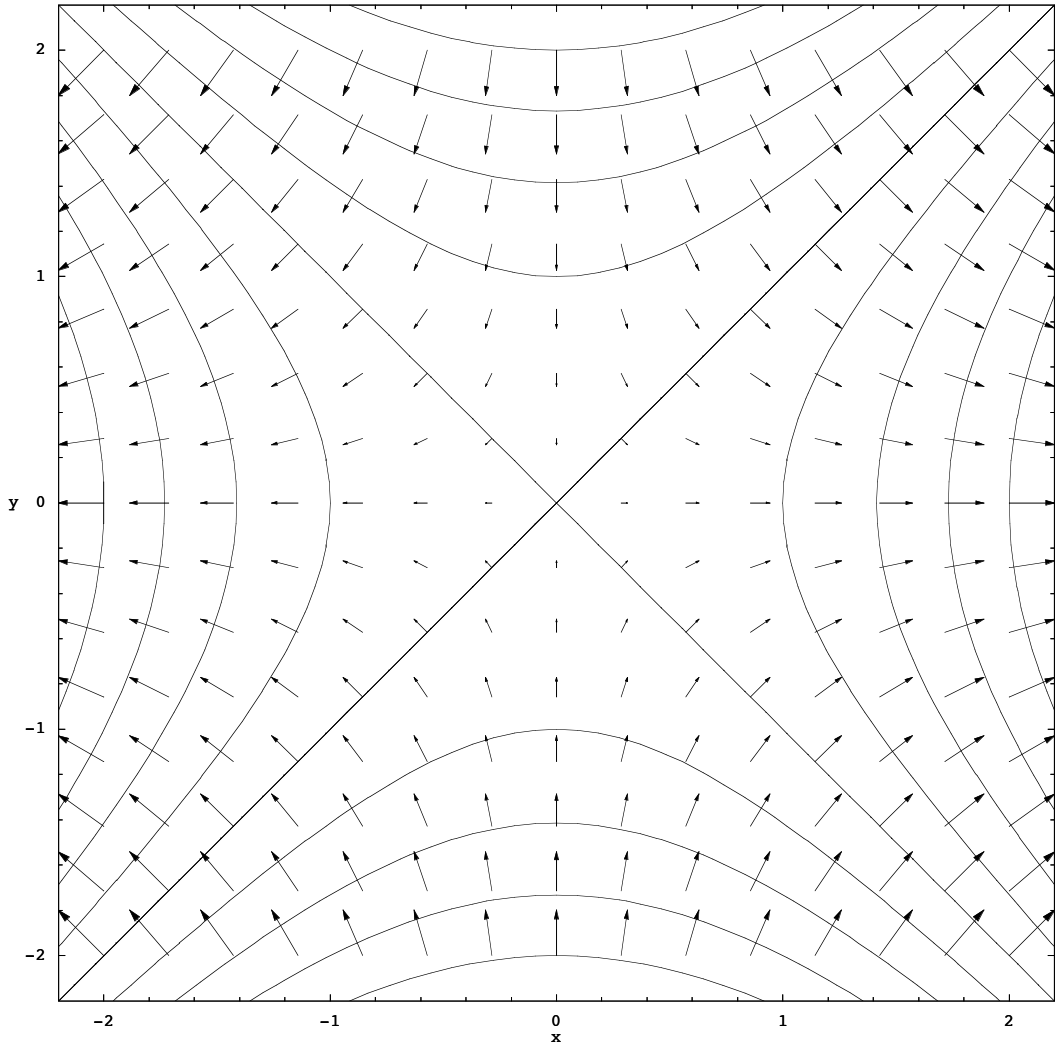
Es gibt also nur eine kritische Stelle: $(0, 0)$.

Hier nimmt die Funktion ihr globales Minimum an.

Beispiele

Wir betrachten f mit

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



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

Beispiele

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

$$\nabla f = \begin{pmatrix} 2x \\ -2y \end{pmatrix}$$

Es gibt wieder nur eine kritische Stelle: $(0,0)$. Es ist

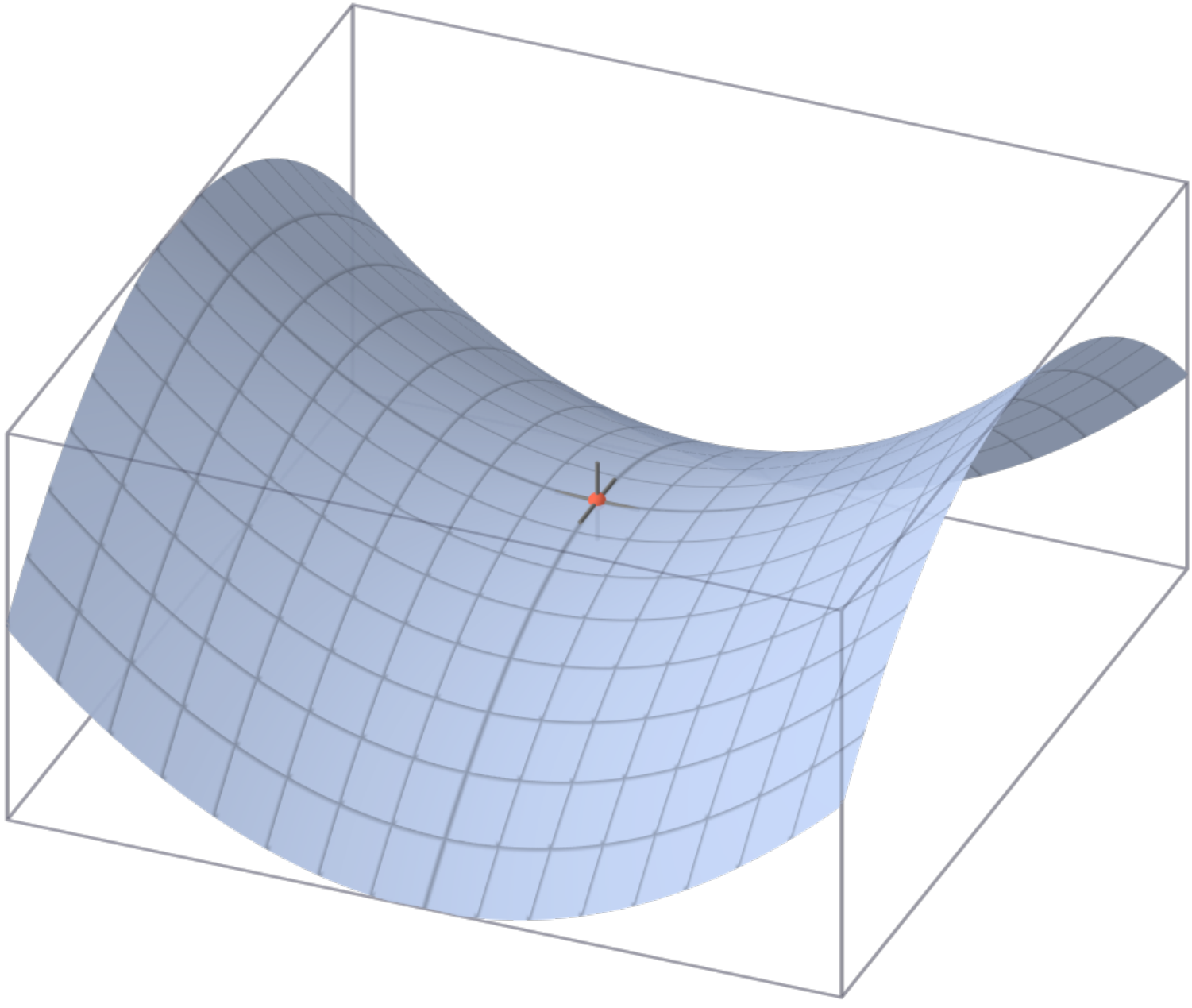
$f(x,0) = x^2$; dies ist > 0 für $x \neq 0$,

$f(0,y) = -y^2$; dies ist < 0 für $y \neq 0$.

Somit ist $(0,0)$ keine lokale Extremstelle und folglich eine Sattelstelle.

Die Funktion hat also keine lokalen Extrenstellen

(und erst recht keine globalen).



Kritische Stellen

Maximal- und Minimalstellen sind notwendigerweise kritische Stellen.

Wir wollen nun zuerst ein weiteres notwendiges und dann ein hinreichendes Kriterium für Maximal-/ Minimalstellen finden.

Kritische Stellen

Es sei nun $\underline{x}^{(0)}$ eine Maximalstelle.

Es sei wieder $\vec{r}$ eine **beliebige** Richtung und

$$\gamma : t \mapsto f(\underline{x}^{(0)} + t\vec{r}) \, .$$

Nun ist 0 eine Maximalstelle von

$$f \circ \gamma : t \mapsto f(\gamma(t)) \, .$$

Es gilt

$$\frac{\mathrm{d}f(\gamma(t))}{\mathrm{d}t}|_{t=0} = 0$$

und wenn $f \circ \gamma$ an 0 zweimal differenzierbar ist:

$$\frac{\mathrm{d}^2 f(\gamma(t))}{\mathrm{d}t^2}|_{t=0} \leq 0 \, .$$

Höhere Ableitungen

Es gilt

$$\frac{\mathrm{d}f(\gamma(t))}{\mathrm{d}t}|_{t=0} = 0$$

$$\frac{\mathrm{d}^2 f(\gamma(t))}{\mathrm{d}t^2}|_{t=0} \leq 0 \,.$$

Es ist

$$\frac{\mathrm{d}f(\gamma(t))}{\mathrm{d}t}|_{t=0} = \partial_{\vec{r}} f|_{\underline{X}=\underline{X}^{(0)}} \,.$$

Was ist nun

$$\frac{\mathrm{d}^2 f(\gamma(t))}{\mathrm{d}t^2}|_{t=0}$$

in Bezug auf f ?

Höhere Ableitungen

Es ist

$$\frac{\mathrm{d}f(\gamma(t))}{\mathrm{d}t}|_{t=0} = \partial_{\vec{r}}(f)|_{\underline{x}=\underline{x}^{(0)}}$$

$$\frac{\mathrm{d}f(\gamma(t))}{\mathrm{d}t} = \partial_{\vec{r}}(f)|_{\underline{x}=\gamma(t)} = (\partial_{\vec{r}}(f))(\gamma(t)) \; .$$

Sei $g := \partial_{\vec{r}} f : D_f \longrightarrow \mathbb{R}$. Diese Funktion sei differenzierbar.

Nun ist

$$\frac{\mathrm{d}f(\gamma(t))}{\mathrm{d}t} = g(\gamma(t)) \; .$$

und somit

$$\frac{\mathrm{d}^2 f(\gamma(t))}{\mathrm{d}t^2}|_{t=0} = \frac{\mathrm{d}g(\gamma(t))}{\mathrm{d}t}|_{t=0} = \partial_{\vec{r}} g|_{\underline{x}=\underline{x}^{(0)}} = \partial_{\vec{r}}\partial_{\vec{r}} f|_{\underline{x}=\underline{x}^{(0)}} \; .$$

Höhere Ableitungen

Ergebnis

Wir setzen voraus, dass alle Richtungsableitungen differenzierbar sind.

Wenn $x^{(0)}$ eine lokale Maximalstelle ist, dann gilt für alle Richtungen $\vec{r}$

$$\partial_{\vec{r}} f(\underline{x}^{(0)}) = 0$$

und

$$\partial_{\vec{r}} \partial_{\vec{r}} f(\underline{x}^{(0)}) \leq 0 \, .$$

Ebenso:

Wenn $x^{(0)}$ eine lokale Minimalstelle ist, dann gilt für alle Richtungen $\vec{r}$

$$\partial_{\vec{r}} f(\underline{x}^{(0)}) = 0$$

und

$$\partial_{\vec{r}} \partial_{\vec{r}} f(\underline{x}^{(0)}) \geq 0 \, .$$

Höhere Ableitungen

Ergebnis (Fortsetzung)

Jede Extremstelle $\underline{x}$ ist eine kritische Stelle, d.h. es gilt

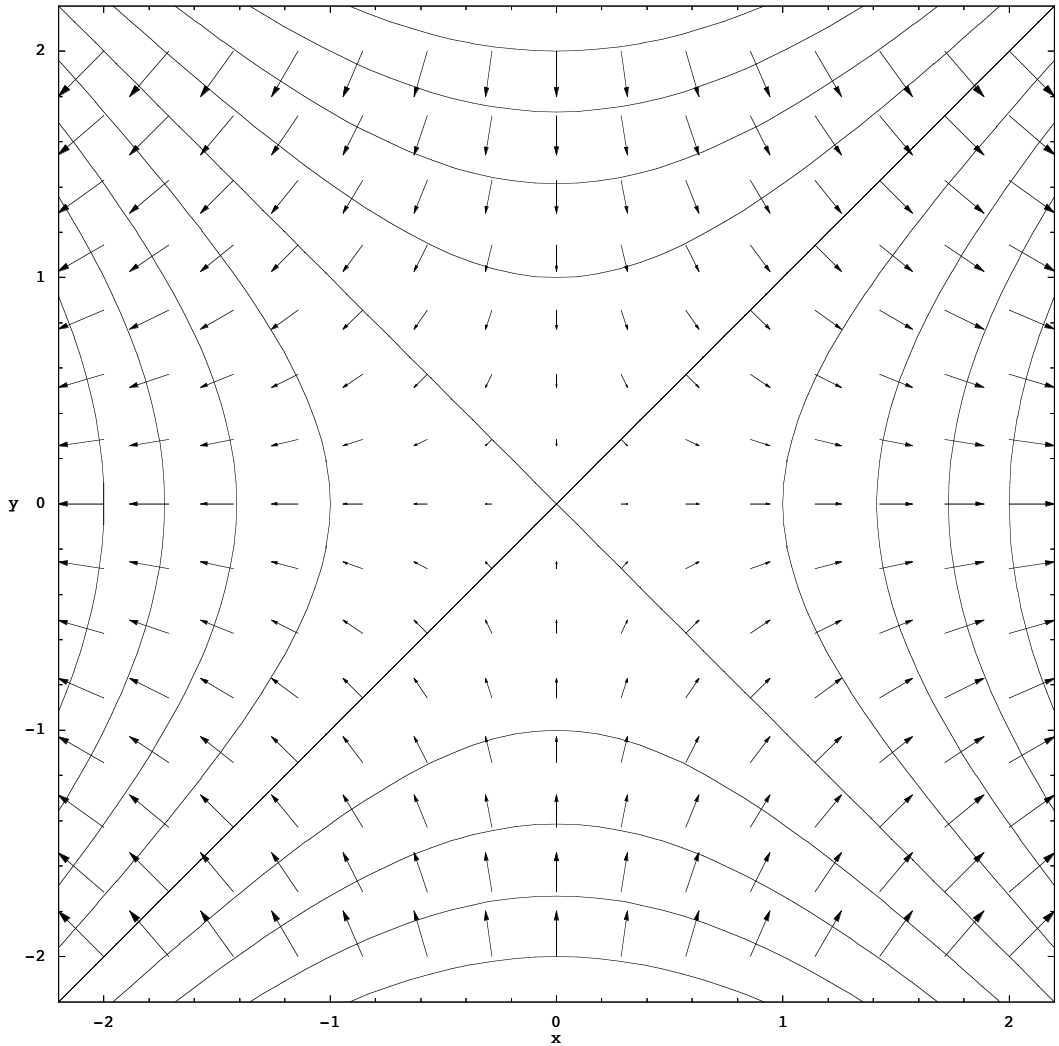
$$(\nabla f)(\underline{x}^{(0)}) = \vec{0}\,.$$

Es sei umgekehrt $\underline{x}^{(0)}$ ist eine kritische Stelle. Wenn es nun zwei Richtungen $\vec{r}$ und $\vec{s}$ mit

$$\partial_{\vec{r}}\partial_{\vec{r}} f(\underline{x}^{(0)}) > 0 \quad , \quad \partial_{\vec{s}}\partial_{\vec{s}} f(\underline{x}^{(0)}) < 0$$

gibt, ist $\underline{x}^{(0)}$ keine Extremstelle und folglich eine **Sattelstelle**.

Beispiel



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

Beispiel

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

$$\nabla f = \begin{pmatrix} 2x \\ -2y \end{pmatrix}$$

Kritische Stelle: $(0,0)$.

$$\partial_x \partial_x f = 2 > 0 \quad , \quad \partial_y \partial_y f = -2 < 0$$

[Und mit $\vec{r} = \frac{1}{\sqrt{2}} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}$:

$$\partial_{\vec{r}} \partial_{\vec{r}} f = 0$$

Ebenso mit $\vec{r} = \frac{1}{\sqrt{2}} \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix}$.]

Höhere Ableitungen

Für eine Richtung $\vec{r}$ haben wir

$$\partial_{\vec{r}} f = \vec{r} \bullet \nabla f = r_x \partial_x f + r_y \partial_y f \ .$$

Frage. Kann man $\partial_{\vec{r}} \partial_{\vec{r}} f$ auf ähnliche Weise ausrechnen?

Antwort. Ja, wenn wir entsprechende Voraussetzungen machen.

Höhere Ableitungen

Satz. Es sei vorausgesetzt, dass alle partiellen Ableitungen $\partial_{x_j}\partial_{x_i} f$ existieren und stetig sind. Dann gilt:

a) Die Richtungsableitung $\partial_{\vec{r}} f$ ist selbst in jede Richtung $\vec{s}$ stetig differenzierbar.

b) Für je zwei Richtungen $\vec{r}$ und $\vec{s}$ ist

$$\partial_{\vec{s}}\partial_{\vec{r}} f = \partial_{\vec{r}}\partial_{\vec{s}} f \ .$$

Insbesondere ist also

$$\partial_{x_j}\partial_{x_i} f = \partial_{x_i}\partial_{x_j} f \ .$$

(Ohne Beweis)

Definition. Eine Funktion, deren zweiten partiellen Ableitungen existieren und stetig sind, heißt f **zweimal stetig differenzierbar**.

Höhere Ableitungen

Für eine Funktion in zwei Variablen x, y gilt dann

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

Beispiel

Es sei $f(x, y) := x^a \cdot y^b$ mit nicht-negativen Konstanten a, b.

Nun ist

$$\partial_x f = a \cdot x^{a-1} y^b \quad , \quad \partial_y f = b \cdot x^a y^{b-1},$$

$$\partial_y \partial_x f = \partial_y a \cdot x^{a-1} y^b = ab \cdot x^{a-1} y^{b-1},$$

$$\partial_x \partial_y f = \partial_x b \cdot x^a y^{b-1} = ab \cdot x^{a-1} y^{b-1} \ .$$

Höhere Ableitungen

Es sei f eine zweimal stetig differenzierbare Funktion.

Dann ist

$$\partial_{\vec{r}} f = r_x \cdot \partial_x f + r_y \partial_y f$$

und

$$\begin{aligned}
& \partial_{\vec{r}} \partial_{\vec{r}} f \\
= \;& \partial_{\vec{r}} (r_x \cdot \partial_x f + r_y \cdot \partial_y f) \\
= \;& r_x \partial_{\vec{r}} \partial_x f + r_y \partial_{\vec{r}} \partial_y f \\
= \;& r_x \cdot (r_x \partial_x \partial_x f + r_y \partial_y \partial_x f) + r_y \cdot (r_x \partial_x \partial_y f + r_y \partial_y \partial_y f) \\
= \;& r_x^2 \, \partial_x \partial_x f + 2 r_x r_y \, \partial_x \partial_y f + r_y^2 \, \partial_y \partial_y f
\end{aligned}$$

.

Kritische Stellen

Es sei $(x^{(0)}, y^{(0)})$ eine kritische Stelle.

Wir wissen:

- Wenn $(x^{(0)}, y^{(0)})$ eine lokale Maximalstelle ist, ist stets $\partial_{\vec{r}}\partial_{\vec{r}} f(x^{(0)}, y^{(0)}) \leq 0$.
- Wenn $(x^{(0)}, y^{(0)})$ eine lokale Minimalstelle ist, ist stets $\partial_{\vec{r}}\partial_{\vec{r}} f(x^{(0)}, y^{(0)}) \geq 0$.

Satz

- Wenn stets $\partial_{\vec{r}}\partial_{\vec{r}} f(x^{(0)}, y^{(0)}) < 0$ ist, ist $(x^{(0)}, y^{(0)})$ eine lokale Maximalstelle.
- Wenn stets $\partial_{\vec{r}}\partial_{\vec{r}} f(x^{(0)}, y^{(0)}) > 0$ ist, ist $(x^{(0)}, y^{(0)})$ eine lokale Minimalstelle.

(Ohne Beweis)

Wie kann man das nun in ein gutes Kriterium umformen?

Kritische Stellen

Es sei $\underline{x}^{(0)}$ eine kritische Stelle.

Wir schreiben im Folgenden (nur zur Abkürzung) $\partial_{\vec{r}}\partial_{\vec{r}}f$ für $\partial_{\vec{r}}\partial_{\vec{r}}f(x^{(0)}, y^{(0)})$.

Es geht aber immer um die Ableitungen an $(x^{(0)}, y^{(0)})$.

Kritische Stellen

Es sind äquivalent:

- Für alle Richtungen $\vec{r}$ ist $\partial_{\vec{r}}\partial_{\vec{r}} f > 0$.
- Für alle Richtungen $\vec{r}$ ist

$$r_x^2 \cdot \partial_x\partial_x f + 2r_x r_y \cdot \partial_x\partial_y f + r_y^2 \cdot \partial_y\partial_y f > 0 \,.$$

- Für alle $(r_x, r_y) \in \mathbb{R}^2 \backslash \{(0,0)\}$ ist

$$r_x^2 \cdot \partial_x\partial_x f + 2r_x r_y \cdot \partial_x\partial_y f + r_y^2 \cdot \partial_y\partial_y f > 0 \,.$$

- $$\partial_x\partial_x f > 0$$

  und für alle $r \in \mathbb{R}$

$$r^2\, \partial_x\partial_x f + 2r\, \partial_x\partial_y f + \partial_y\partial_y f > 0 \,.$$

Kritische Stellen

- $$\partial_x \partial_x f > 0$$

  und für alle $r \in \mathbb{R}$

  $$r^2\, \partial_x \partial_x f + 2r\, \partial_x \partial_y f + \partial_y \partial_y f > 0\ .$$

- $$\partial_x \partial_x f > 0$$

  und das Polynom in r

  $$r^2 \partial_x \partial_x f + 2r \partial_x \partial_y f + \partial_y \partial_y f$$

  hat keine Nullstelle.

- $$\partial_x \partial_x f > 0$$

  und

  $$(\partial_x \partial_y f)^2 - \partial_x \partial_x f \cdot \partial_y \partial_y f < 0$$

Kritische Stellen

- $$\partial_x \partial_x f > 0$$

und

$$\partial_x \partial_x f \cdot \partial_y \partial_y f - (\partial_x \partial_y f)^2 > 0$$

Minima und Maxima

Satz (hinreichende Kriterien)

Es sei f eine zweimal stetig differenzierbare Funktion in zwei Variablen x, y.

Es sei $(x^{(0)}, y^{(0)})$ eine kritische Stelle, d.h. es gelte

$$\nabla f(x^{(0)}, y^{(0)}) = \begin{pmatrix} 0 \\ 0 \end{pmatrix} .$$

- Wenn
$$(\partial_x \partial_x f)(x^{(0)}, y^{(0)}) > 0 \,,$$
$$(\partial_x \partial_x f)(x^{(0)}, y^{(0)}) \cdot (\partial_y \partial_y f)(x^{(0)}, y^{(0)}) - ((\partial_x \partial_y f)(x^{(0)}, y^{(0)}))^2 > 0$$
ist, dann ist $(x^{(0)}, y^{(0)})$ eine lokale Minimalstelle.
- Wenn
$$(\partial_x \partial_x f)(x^{(0)}, y^{(0)}) < 0 \,,$$
$$(\partial_x \partial_x f)(x^{(0)}, y^{(0)}) \cdot (\partial_y \partial_y f)(x^{(0)}, y^{(0)}) - ((\partial_x \partial_y f)(x^{(0)}, y^{(0)}))^2 > 0$$
ist, dann ist $(x^{(0)}, y^{(0)})$ eine lokale Maximalstelle.
- Wenn
$$(\partial_x \partial_x f)(x^{(0)}, y^{(0)}) \cdot (\partial_y \partial_y f)(x^{(0)}, y^{(0)}) - ((\partial_x \partial_y f)(x^{(0)}, y^{(0)}))^2 < 0$$
ist, dann ist $(x^{(0)}, y^{(0)})$ eine Sattelstelle.

Extremstellen

Vorbemerkung. Im Folgenden untersuchen wir eine Funktion auf lokale Extremstellen, wobei gewisse so genannte **Nebenbedingungen** gelten sollen.

Der Einfachheit halber beschränken wir uns auf lokale Maximalstellen. Analoge Überlegungen gelten für lokale Minimalstellen.

Maximalstellen

Wiederholung. Es sei eine Funktion f in d Variablen gegeben. Der Definitionsbereich D_f sei ohne Rand und die Funktion sei differenzierbar.

Wenn nun $\underline{x}^{(0)}$ eine lokale Maximalstelle ist, ist $(\nabla f)(\underline{x}^{(0)}) = 0$.

Maximalstellen auf Höhenmengen

Wir betrachten die folgende Situation:

Gegeben seien zwei stetig differenzierbare Funktionen f und g in d Variablen und $c \in \mathbb{R}$.

Wir setzen voraus, dass die Höhenmenge

$$H_g(c) := \{\underline{x} \in \mathbb{R}^d \ , \ g(\underline{x}) = c\}$$

eine Teilmenge von D_f ist.

Wir wollen nun die Funktion f **eingeschränkt auf die Höhenmenge** auf lokale Maximalstellen untersuchen.

Mit anderen Worten:

Wir wollen ein Kriterium dafür angeben, dass ein $\underline{x}^{(0)}$ eine lokale Maximalstelle von

$$f|_{H_g(c)} : H_g(c) \longrightarrow \mathbb{R}$$

ist.

Maximalstellen auf Höhenmengen

"Dummerweise" besteht die Menge $H_g(c)$ nur aus ihrem Rand.

Man muss sich etwas neues einfallen lassen ...

Maximalstellen auf Höhenmengen

Sprachgebrauch. Die Bedingung

$$g(\underline{x}) = c$$

heißt **Nebenbedingung**.

Die Aufgabe, eine (globale) Maximalstelle von f unter der Bedingung $g(\underline{x}) = c$ zu finden, formuliert man so:

"Maximiere f unter der Nebenbedingung $g(\underline{x}) = c$."

In der Praxis werden oftmals auch mehrere Nebenbedingungen betrachtet, aber das machen wir nicht.

Maximalstellen unter Nebenbedingungen

Idee. Sowohl ohne Nebenbedingung als auch mit Nebenbedingung gilt:

Es sei $\underline{x}^{(0)}$ eine lokale Maximalstelle (eventuell unter einer Nebenbedingung $g(\underline{x}) = c$).

Dann gilt für jede differenzierbare Bewegung γ mit $\gamma(0) = \underline{x}^{(0)}$ mit $D_\gamma \subseteq D_f$ (beziehungsweise mit $D_\gamma \subseteq H_g(c)$):

0 ist eine lokale Maximalstelle von

$$f \circ \gamma : t \mapsto f(\gamma(t)) \ .$$

Wenn $[-\epsilon, \epsilon] \subseteq D_\gamma$ für ein $\epsilon > 0$, folgt:

$$0 = \frac{\mathrm{d} f(\gamma(t))}{\mathrm{d} t}|_{t=0} = (\nabla g)(\underline{x}^{(0)}) \bullet \dot{\gamma}\ (0) \ .$$

Wenn $\dot{\gamma}\ (0) \neq 0$: Mit $\vec{r} := \frac{\dot{\gamma}(0)}{\|\dot{\gamma}(0)\|}$:

$$(\partial_{\vec{r}} f)(\underline{x}^{(0)}) = 0$$

Maximalstellen unter Nebenbedingungen

Wenn $\dot{\gamma}(0) \neq 0$: Mit $\vec{r} := \frac{\dot{\gamma}(0)}{\|\dot{\gamma}(0)\|}$:

$$(\partial_{\vec{r}} f)(\underline{x}^{(0)}) = 0$$

Ohne Nebenbedingung

Wir haben zu jeder Richtung $\vec{r}$ eine Bewegung mit $\|\dot{\gamma}(0)\| = \vec{r}$, nämlich

$$t \mapsto \underline{x}^{(0)} + t\vec{r}\,.$$

Damit haben wir

$$(\nabla f)(\underline{x}^{(0)}) = \vec{0}\,.$$

Maximalstellen unter Nebenbedingung

Mit Nebenbedingung $g(\underline{x}) = c$

Wir machen die **Voraussetzung**: $\nabla g(\underline{x}^{(0)}) \neq \vec{0}$.

Wir haben nun den Tangentialraum von g an $\underline{x}^{(0)}$:

$$\{\underline{x} \in \mathbb{R}^d \mid (\underline{x} - \underline{x}^{(0)}) \bullet (\nabla f)(\underline{x}^{(0)}) = 0\}$$

Für eine Richtung $\vec{r}$ sind äquivalent:

- $\vec{r}$ zeigt von $\underline{x}^{(0)}$ ausgehend in den Tantentialraum, d.h. $\underline{x}^{(0)} + \vec{r}$ liegt im Tangentialraum.
- $\vec{r}$ ist orthogonal zu $(\nabla(g))(\underline{x}^{(0)})$, d.h. $\vec{r} \bullet (\nabla(g))(\underline{x}^{(0)}) = 0$.

Maximalstellen unter Nebenbedingung

Nun erscheint naheliegend:

Für jede Richtung $\vec{r}$,

die von $\underline{x}^{(0)}$ ausgehend in Richtung des Tangentialraums zeigt,

d.h. mit $\vec{r} \bullet (\nabla(g))(\underline{x}^{(0)}) = 0$,

gibt es eine Bewegung in der Höhenmenge $H_g(c)$ mit $\|\dot{\gamma}(0)\| = \vec{r}$.

Nehmen wir dies an.

Dann haben wir:

Für jede solche Richtung gilt:

$$0 = (\partial_{\vec{r}} f)(\underline{x}^{(0)}) = \vec{r} \bullet (\nabla f)(\underline{x}^{(0)}) \; .$$

Maximalstellen unter Nebenbedingungen

Wir haben also:

Für alle Richtungen $\vec{r}$ mit $\vec{r} \bullet (\nabla(g))(\underline{x}^{(0)}) = 0$ gilt $\vec{r} \bullet (\nabla(f))(\underline{x}^{(0)}) = 0$.

Hier bleibt nur noch: $(\nabla(g))(\underline{x}^{(0)})$ und $(\nabla(f))(\underline{x}^{(0)})$ zeigen in **dieselbe Richtung dieselbe oder die entgegengesetzte Richtung**.

Anders ausgedrückt:

Die Tangentialräume von f und g an $x^{(0)}$ sind identisch.

Maximalstellen unter Nebenbedingungen

Es bleibt nur noch: $(\nabla(g))(\underline{x}^{(0)})$ und $(\nabla(f))(\underline{x}^{(0)})$ zeigen in dieselbe oder in die entgegengesetzte Richtung.

Also: Es gibt ein $\lambda \in \mathbb{R}$ mit

$$(\nabla f)(\underline{x}^{(0)}) = -\lambda \cdot (\nabla g)(\underline{x}^{(0)}) \, .$$

Die Zahl λ heißt der **Lagrange-Multiplikator**.

Bemerkung. Das negative Vorzeichen ist klassisch. Im Buch von Sydæter und Hammond wird es allerdings weggelassen. Man muss also aufpassen. Für Rechnungen ist es allerdings egal, welches Vorzeichen man wählt.

Lagrange-Multiplikatoren

Satz. Es seien zwei stetig differenzierbare Funktionen f und g mit $H_g(c) \subseteq D_f$ gegeben.

Wenn nun $\underline{x}^{(0)}$ ein lokales Maximum von f unter der "Nebenbedingung" $g(\underline{x}) = c$ ist und $(\nabla g)(\underline{x}^{(0)}) \neq \vec{0}$ ist, dann gibt es ein $\lambda \in \mathbb{R}$ mit

$$(\nabla f)(\underline{x}^{(0)}) = -\lambda \cdot (\nabla g)(\underline{x}^{(0)}) \, .$$

Weitergehend. Der Wikipedia-Artikel zu "Lagrange-Multiplikator" ist gut. Das erste Bild dort auch sehr aufschlussreich. (Beachten Sie, dass sich die Höhenlinilien von f und g am Maximum berühren.) Das zweite Bild ist dieselbe Situation in 3D.

Lagrange-Multiplikatoren

Es seien zwei stetig differenzierbare Funktionen f und g mit $H_g(c) \subseteq D_f$ gegeben. Es sei $\underline{x}^{(0)}$ eine lokale Extremstelle von f unter der Nebenbedingung $g(\underline{x}) = c$.

Dann gilt:

$$(\nabla g)(\underline{x}^{(0)}) = \vec{0}$$

oder

$$(\nabla f)(\underline{x}^{(0)}) = -\lambda \cdot (\nabla g)(\underline{x}^{(0)})$$

Lagrange-Multiplikatoren

Die Gleichungen

$$(\nabla f)(\underline{x}^{(0)}) = -\lambda(\nabla g)(\underline{x}^{(0)}) \,,\; g(\underline{x}^{(0)}) = c$$

lauten ausgeschrieben und umgestellt:

$$\begin{aligned}
(\partial_{x_1} f)(\underline{x}^{(0)}) + \lambda(\partial_{x_1} g)(\underline{x}^{(0)}) &= 0 \\
\vdots & \quad \vdots \\
(\partial_{x_d} f)(\underline{x}^{(0)}) + \lambda(\partial_{x_d} g)(\underline{x}^{(0)}) &= 0 \\
g(\underline{x}^{(0)}) &= c
\end{aligned}$$

Diese Gleichungen erhält man, wenn man die partiellen Ableitungen der Funktion

$$(\underline{x}, \lambda) \mapsto f(\underline{x}) + \lambda \cdot (g(\underline{x}) - c)$$

gleich Null setzt.

Somit sind die Gleichungen äquivalent zu:

$(\underline{x}^{(0)}, \lambda)$ ist ein kritischer Punkt der obigen Funktion.

Ökonomische Anwendungen

Anwendung in Mikroökonomie

Wir wollen eine Nutzenfunktion f unter einer Budgetrestriktion $g(\underline{x}) \leq c$ maximieren.

Oft (aber nicht immer) führt dies zum Problem, f unter der Bedingung $g(\underline{x}) = c$ zu maximieren.

Ökonomische Anwendungen

Anwendung in der Betriebswirtschaftslehre

Wir haben eine Produktionsanlage für einen oder mehrere Stoffe. Die tägliche Produktion benötigt bestimmte Inputs $x_1, \ldots, x_d$, die varriert werden können. Aufgrund der Kapazität der Anlage haben wir eine Restriktion der Form $g(\underline{x}) \leq c$.

(Wahrscheinlich haben wir noch mehr Restriktionen, aber wir wollen es nicht zu kompliziert machen. Vielleicht ist auch a priori klar, dass die eine Bedingung realistischerweise die strengste ist.)

Ökonomische Anwendung

Weitere Annahmen:

- Beim Absatz und beim Einkauf haben wir keine Restriktion.
- Am Morgen eines Tages sind alle Einkaufs- und Verkaufspreise bekannt.
- Die Produktion kann ohne Kosten umgestellt werden. (Das ist wichtig, da wir am nächsten Tag wieder neu optimieren wollen und keine Abhängigkeiten erhalten wollen.)
- Insbesondere: In Abängigkeit von $x_1, \ldots, x_d$ erzielen wir einen bestimmten täglichen Gewinn $f(x_1, \ldots, x_d)$.

Wir wollen am Morgen eines Tages die optimale Produktion bestimmen.

Wir müssen wieder f unter der Nebenbedingung $g(\underline{x}) \leq c$ maximieren.

Maximalstellen

Wie bestimmt man die globalen Maximalstellen von f unter der Nebenbedingung $g(\underline{x}) \leq c$?

Wir setzen voraus, dass D_f keinen Rand hat.

Nun hat

$$\{\underline{x} \in D_f | g(\underline{x}) < c\}$$

auch keinen Rand.

Man kann nun so vorgehen: Zunächst bestimmt man "Kandidaten" für globale Maximalstellen.

- Man bestimmt die kritischen Stellen von f auf $\{\underline{x} \in D_f | g(\underline{x}) < c\}$.
- Man wendet die Lagrange-Methode auf f mit der Nebenbedingung $g(\underline{x}) = c$ an.
- Man bestimmt die Stellen, an denen $(\nabla g)(\underline{x}) = \vec{0}$ ist.

Man vergleicht. Jede globale Maximalstelle ist eine dieser Stellen.

Abhängigkeit von der Restriktion

Wir betrachten wieder Funktionen f und g wie oben vor.

Wir studieren nun das Verhalten einer lokalen Maximalstelle von f unter der Nebenbedingung $g(\underline{x}) = c$, wobei wir c kontinuierlich variieren.

Wir nehmen an, dass wir auf diese Weise eine differenzierbare Bewegung

$$c \mapsto \underline{x}^{\max}(c)$$

erhalten.

Also: $\underline{x}^{\max}(c)$ maximiert f lokal unter der Nebenbedingung $g(\underline{x}) = c$.

Frage. Was ist $\frac{\mathrm{d}f(\underline{x}^{\max}(c))}{\mathrm{d}c}$?

Abhängigkeit von der Restriktion

Frage. Was ist $\frac{\mathrm{d}f(\underline{x}^{\max}(c))}{\mathrm{d}c}$?

Wir rechnen:

$$
\begin{aligned}
& \frac{\mathrm{d}f(\underline{x}^{\max}(c))}{\mathrm{d}c} \\
= \quad & (\nabla f)(\underline{x}^{\max}(c)) \bullet \frac{\mathrm{d}\underline{x}^{\max}(c)}{\mathrm{d}c} \\
= \quad & -\lambda(c) \cdot (\nabla g)(\underline{x}^{\max}(c)) \bullet \frac{\mathrm{d}\underline{x}^{\max}(c)}{\mathrm{d}c} \\
= \quad & -\lambda(c) \cdot \frac{\mathrm{d}\, g(\underline{x}^{\max}(c))}{\mathrm{d}c} \\
= \quad & -\lambda(c) \cdot \frac{\mathrm{d}c}{\mathrm{d}c} \qquad\qquad\qquad\qquad , \text{ weil } g(\underline{x}^{\max}(c)) = c \\
= \quad & -\lambda(c)
\end{aligned}
$$

Abhängigkeit von der Restriktion

Ökonomische Interpretation

Nehmen wir an, wir haben eine Nutzenfunktion f und eine Restriktion $g(\underline{x}) = c$.

Nehmen wir auch an, wir können c kontinuierlich varriieren und die optimale Güterkombination hängt gutmütig von c ab.

Dann gilt:

Der Grenznutzen beim Verschieben der Restriktion von c auf $c + \mathrm{d}c$ beträgt $-\lambda \cdot \mathrm{d}c$.

Ökonomische Interpretation

Der Grenznutzen beim Verschieben der Restriktion von c auf $c + \mathrm{d}c$ beträgt $-\lambda \cdot \mathrm{d}c$.

Bemerkung. Wie kann man das negative Vorzeichen verstehen?

Im Optimum (Maximum) von f unter der Budgetbedingung $g(\underline{x}) = c$ sollte der Gradient von f in dieselbe Richtung zeigen wie der Gradient von g. Damit ist dann nach Definition λ negativ also $-\lambda$ wieder positiv.

Wie schon oben gesagt: Im Buch von Sydsæter und Hammond ist das Vorzeichen umgekehrt. Dies vermeidet dann einige Vorzeichenwechsel in der ökonomischen Anwendung.

Ökonomische Interpretation

Wenn "Nutzen = Gewinn":

Wenn die Restriktion um $\mathrm{d}c$ verschoben wird, verändert sich der Gewinn um $-\lambda\,\mathrm{d}c$.

Wir gehen davon aus, dass dies positiv ist, wenn $\mathrm{d}c$ positiv ist.

Ökonomische Interpretation

Verschiebung um $\mathrm{d}c > 0$.

Das Verschieben um $\mathrm{d}c$ könnte etwas kosten. Diese Kosten kann f "nicht sehen". (Weil f nur von $\underline{x}$ abhängt.)

Beispiel: Man benötigt zusätzliche Produktionskapazität.

Wir betrachten jetzt auch noch potentielle Kosten für das Verschieben von c.

Soeben hatten wir: Wenn es möglich wäre, die Restriktion um $\mathrm{d}c$ **ohne Kosten** zu verschieben, würde der Gewinn um $-\lambda\,\mathrm{d}c$ steigen.

Das heißt:

Wenn ein Verschieben der Restriktion um $\mathrm{d}c$ Kosten in Höhe von $-\lambda\,\mathrm{d}c$ verursacht, würde sich der Gewinn nicht ändern, wenn die Restriktion ein "marginales Bisschen" verschoben würde.

Ökonomische Interpretation

Verschiebung um $\mathbf{d}c$ mit $\mathbf{d}c < \mathbf{0}$

Wenn die Restriktion um $\mathrm{d}c$ verschoben wird, nimmt der Gewinn um $\lambda\,\mathrm{d}c$ ab.

Dies kann man so interpretieren: Ein Verschieben der Restriktion um $\mathrm{d}c$ führt zu "Kosten" (genauer: entgangenem Gewinn) in Höhe von $\lambda\,\mathrm{d}c$.

Nun könnte ein Verschieben der Restriktion um $\mathrm{d}c$ auch zu einem **zusätzlichen** Gewinn führen. (Einem Gewinn, der nicht durch f ausgedrückt ist.)

Ökonomische Interpretation

Beispiel: Man vermietet Produktionskapazität.

Frage: Wie hoch sollte dieser zusätzliche Gewinn sein, damit sich die Alternative lohnt, wenn sich c um $\mathrm{d}c$ verschiebt?

Antwort: Mindestens $\lambda\,\mathrm{d}c$.

Im Beispiel: Wir betrachten die Alternativen, die Produktionskapazität zu erhalten oder Produktionskapazität zu vermieten, wobei dabei die Restriktion um $\mathrm{d}c$ verändert würde.

Nun betragen die **Grenz-Opportunitätskosten** dafür, Produktionskapazität zu vermieten, $\lambda\,\mathrm{d}c$.