Kapitel 6

Multivariate Analysis

Lernziele

- Funktionen in mehreren Variablen
- Graph und Niveaulinien einer Funktion in zwei Variablen
- Partielle Ableitung und Gradient
- Lokale und globale Extrema
- Lagrange-Ansatz

Funktionen in mehreren Variablen

Eine **reelle Funktion in mehreren Variablen** ist eine Abbildung, die jedem Vektor $\mathbf{x}$ eine reelle Zahl zuordnet.

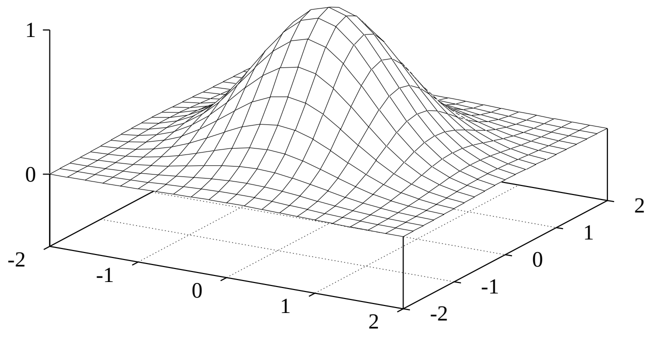
$$f\colon \mathbb{R}^n \to \mathbb{R}, \mathbf{x} \mapsto f(\mathbf{x}) = f(x_1, x_2, \dots, x_n)$$

Die Komponenten x_i des Vektors $\mathbf{x}$ heißen die **Variablen** der Funktion f.

Funktionen in *zwei* Variablen lassen sich durch den **Graphen** (Funktionengebirge) veranschaulichen.
Dabei wird für jeden Punkt in der (x, y)-Ebene der Funktionswert $f(x, y)$ in die z-Richtung eingezeichnet.

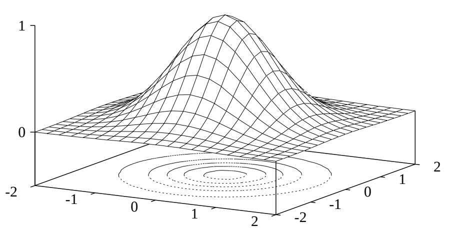
Graph einer Funktion in zwei Variablen

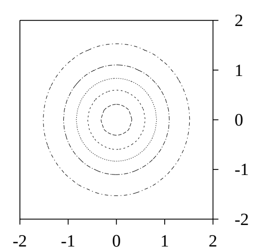
$$f\colon \mathbb{R}^2 \to \mathbb{R}, (x,y) \mapsto f(x,y) = \exp(-x^2 - y^2)$$



Niveaulinien

Eine **Isoquante** (**Höhenlinie**, **Niveaulinie**) ist die Menge aller Punkte (x,y) mit $f(x,y) = c$ ($c \in \mathbb{R}$ konstant).
Die Funktion f hat daher auf einer Höhenlinie den gleichen Funktionswert.





Partielle Ableitung

Wir untersuchen die Änderung des Funktionswertes wenn wir eine Variable x_i variieren und alle anderen konstant lassen. Wir erhalten dadurch die (erste) **partielle Anleitung** von f nach x_i:

$$f_{x_i}(\mathbf{x}) = \frac{\partial f}{\partial x_i} = \lim_{\Delta x_i \to 0} \frac{F(\ldots, x_i + \Delta x_i, \ldots) - F(\ldots, x_i, \ldots)}{\Delta x_i}$$

Wir erhalten die partielle Ableitung nach x_i wenn wir alle anderen Variablen als Konstante auffassen und f nach den bekannten Regeln für Funktionen in einer Variable nach x_i ableiten.

Partielle Ableitung / Beispiel

Gesucht sind die ersten partiellen Ableitungen von

$$f(x_1, x_2) = \sin(2\,x_1) \cdot \cos(x_2)$$

$$\begin{aligned} f_{x_1} &= 2 \cdot \cos(2\,x_1) \cdot \underbrace{\cos(x_2)}_{\text{als Konstante betrachtet}} \\ f_{x_2} &= \underbrace{\sin(2\,x_1)}_{\text{als Konstante betrachtet}} \cdot (-\sin(x_2)) \end{aligned}$$

Partielle Ableitung / Beispiel

Gesucht sich die ersten partiellen Ableitungen von

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

$$\begin{aligned} f_x &= 2\,x + 3\,y \\ f_y &= 0 + 3\,x \end{aligned}$$

Der Gradient

Wir fassen die partiellen Ableitungen erster Ordnung zu einem Vektor, dem **Gradienten** an der Stelle $\mathbf{x}$, zusammen.

$$\boxed{\nabla f(\mathbf{x}) = (f_{x_1}(\mathbf{x}), \dots, f_{x_n}(\mathbf{x}))}$$

- ∇f heißt auch „nabla f“.
- Der Gradient wird oft als Zeilenvektor geschrieben.
- Der Gradient einer Funktion f zeigt in die Richtung des steilsten Anstieges von f. Seine Länge gibt diese Steigung an.
- Der Gradient steht immer normal auf die entsprechende Niveaulinie.
- Der Gradient „spielt“ die gleiche Rolle wie die erste Ableitung bei Funktionen in einer Variablen.

Der Gradient / Beispiel

Gesucht ist der Gradient von

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

an der Stelle $\mathbf{x} = (3,2)$.

$$\nabla f(\mathbf{x}) = (2x + 3y,\ 3x)$$
$$\nabla f(3,2) = (12,9)$$

Höhere partielle Ableitungen

Analog zu den Funktionen in einer Variablen können wir partielle Ableitungen nochmals ableiten und erhalten so **höhere partielle Ableitungen**.

$$f_{x_i x_k}(\mathbf{x}) = \frac{\partial^2 f}{\partial x_k \partial x_i}(\mathbf{x}) \quad \text{bzw.} \quad f_{x_i x_i}(\mathbf{x}) = \frac{\partial^2 f}{\partial x_i^2}(\mathbf{x})$$

Es kommt auf die Reihenfolge beim Differenzieren nicht an (falls alle zweiten partiellen Ableitungen existieren und stetig sind):

$$\frac{\partial^2 f}{\partial x_k \partial x_i}(\mathbf{x}) = \frac{\partial^2 f}{\partial x_i \partial x_k}(\mathbf{x})$$

Höhere partielle Ableitungen / Beispiel

Gesucht sind alle ersten und zweiten partiellen Ableitungen von

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

Erste partielle Ableitungen:

$$f_x = 2\,x + 3\,y \qquad f_y = 0 + 3\,x$$

Zweite partielle Ableitungen:

$$\begin{aligned} f_{xx} &= 2 & \qquad f_{xy} &= 3 \\ f_{yx} &= 3 & \qquad f_{yy} &= 0 \end{aligned}$$

Extrema (Optimierung)

x_0 heißt **globales Maximum** von f, falls für alle x gilt:

$$f(x_0) \geq f(x)$$

x_0 heißt **lokales Maximum** von f, falls für alle x in einer geeigneten Umgebung von x_0 (i.e. Menge aller Punkte in der Nähe von x_0) gilt:

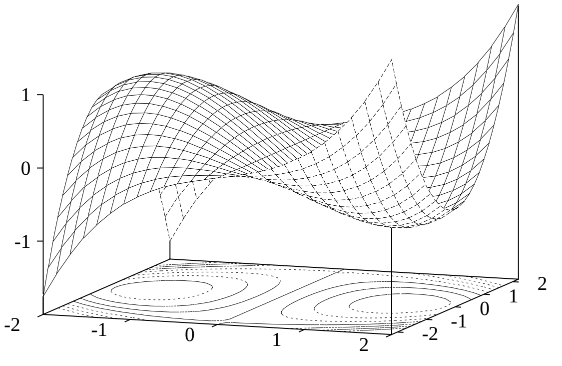
$$f(x_0) \geq f(x)$$

Analog werden **globales Minimum** und **lokales Minimum** definiert.

Minimum und Maximum werden als **Extremum** von f bezeichnet.

Aufgabe:
Suche Extrema einer gegebenen Funktionen.

Extrema / Beispiel



$$f(x,y) = \frac{1}{6}x^3 - x + \frac{1}{4}xy^2$$

Stationärer Punkt

Ein Punkt x_0, in dem der Gradient verschwindet, i.e.,

$$\nabla f(x_0) = 0$$

heißt **stationärer Punkt** von f.

Es gilt (*notwendige Bedingung*):

Jedes Extremum von f ist ein stationärer Punkt von f.

Stationärer Punkt / Beispiel

Suche alle stationären Punkte von

$$f(x,y) = \frac{1}{6}x^3 - x + \frac{1}{4}xy^2$$

Bilde alle ersten partiellen Ableitungen und setze diese Null:

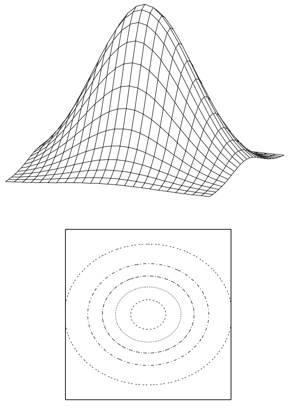
$$\begin{array}{llcll} (I) & f_x & = & \frac{1}{2}x^2 - 1 + \frac{1}{4}y^2 & = 0 \\ (II) & f_y & = & \frac{1}{2}xy & = 0 \end{array}$$

Die stationären Punkte von f sind nun die Lösungen dieses Gleichungssystems:

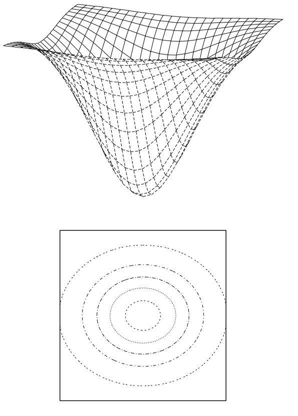
$$\mathbf{x}_1 = (0,2) \qquad \mathbf{x}_2 = (0,-2)$$

$$\mathbf{x}_3 = (\sqrt{2},0) \qquad \mathbf{x}_4 = (-\sqrt{2},0)$$

Stationäre Punkte: Lokale Extrema

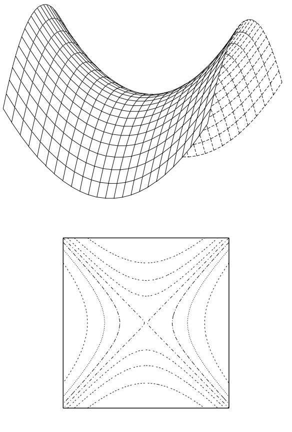


Lokales Maximum

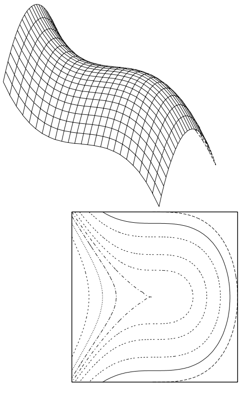


Lokales Minimum

Stationäre Punkte: Sattelpunkte und andere



Sattelpunkt



Beispiel für höhere Ordnung

Die Hesse-Matrix

Die Matrix aller zweiten partiellen Ableitungen einer Funktion f an der Stelle $\mathbf{x}$ wird als **Hesse-Matrix** $\mathbf{H}_f(\mathbf{x})$ bezeichnet.

$$\mathbf{H}_f(\mathbf{x}) = \begin{pmatrix} f_{x_1x_1}(\mathbf{x}) & f_{x_1x_2}(\mathbf{x}) & \dots & f_{x_1x_n}(\mathbf{x}) \\ f_{x_2x_1}(\mathbf{x}) & f_{x_2x_2}(\mathbf{x}) & \dots & f_{x_2x_n}(\mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ f_{x_nx_1}(\mathbf{x}) & f_{x_nx_2}(\mathbf{x}) & \dots & f_{x_nx_n}(\mathbf{x}) \end{pmatrix}$$

Die Hesse-Matrix bestimmt ob die Funktion f in der Nähe von $\mathbf{x}$ konvex oder konkav ist (oder nicht).

Sie „spielt" die gleiche Rolle, wie die zweite Ableitung von Funktionen in einer Variablen.

Die Hesse-Matrix / Beispiel

Berechne die Hesse-Matrix an den Stellen $\mathbf{x}_1 = (0, 2)$ und $\mathbf{x}_4 = (-\sqrt{2}, 0)$ von

$$f(x, y) = \frac{1}{6}x^3 - x + \frac{1}{4}xy^2$$

$$\begin{array}{lll} f_x = \frac{1}{2}x^2 - 1 + \frac{1}{4}y^2 & f_{xx} = x & f_{xy} = \frac{1}{2}y \\ f_y = \frac{1}{2}xy & f_{yx} = \frac{1}{2}y & f_{yy} = \frac{1}{2}x \end{array}$$

$$\mathbf{H}_f(\mathbf{x}) = \begin{pmatrix} f_{xx}(\mathbf{x}) & f_{xy}(\mathbf{x}) \\ f_{yx}(\mathbf{x}) & f_{yy}(\mathbf{x}) \end{pmatrix} = \begin{pmatrix} x & \frac{1}{2}y \\ \frac{1}{2}y & \frac{1}{2}x \end{pmatrix}$$

$$\mathbf{H}_f(\mathbf{x}_1) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \mathbf{H}_f(\mathbf{x}_4) = \begin{pmatrix} -\sqrt{2} & 0 \\ 0 & -\frac{\sqrt{2}}{2} \end{pmatrix}$$

Hauptminoren

Wir benötigen das „Vorzeichen" (die **Definitheit**) der Hesse-Matrix.

Die Definitheit der Hesse-Matrix kann mit Hilfe der sogenannten **Hauptminoren** festgestellt werden.

Die Determinante der linken oberen $k \times k$-Untermatrix heißt der **k-te (führende) Hauptminor**:

$$M_k = \begin{vmatrix} f_{x_1x_1}(\mathbf{x}) & f_{x_1x_2}(\mathbf{x}) & \dots & f_{x_1x_k}(\mathbf{x}) \\ f_{x_2x_1}(\mathbf{x}) & f_{x_2x_2}(\mathbf{x}) & \dots & f_{x_2x_k}(\mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ f_{x_kx_1}(\mathbf{x}) & f_{x_kx_2}(\mathbf{x}) & \dots & f_{x_kx_k}(\mathbf{x}) \end{vmatrix}$$

Hauptminoren / Beispiel

Berechne alle Hauptminoren der Hesse-Matrix an den Stellen $\mathbf{x}_1 = (0, 2)$ und $\mathbf{x}_4 = (-\sqrt{2}, 0)$ von

$$f(x, y) = \frac{1}{6}x^3 - x + \frac{1}{4}xy^2$$

$$\mathbf{H}_f(\mathbf{x}_1) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad M_1 = 0$$

$$M_2 = \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} = -1$$

$$\mathbf{H}_f(\mathbf{x}_4) = \begin{pmatrix} -\sqrt{2} & 0 \\ 0 & -\frac{\sqrt{2}}{2} \end{pmatrix} \qquad M_1 = -\sqrt{2}$$

$$M_2 = \begin{vmatrix} -\sqrt{2} & 0 \\ 0 & -\frac{\sqrt{2}}{2} \end{vmatrix} = 1$$

Hinreichende Bedingung für Lokale Extrema

Für Funktionen in zwei Variablen erhalten wir folgende Bedingung:

Sei $\mathbf{x_0}$ ein stationärer Punkt von f, und
M_1 und M_2 die Hauptminoren von $\mathbf{H}_f(\mathbf{x}_0)$.

(a) $M_2 < 0$ $\Rightarrow$ $\mathbf{x}_0$ ist Sattelpunkt

(b) $M_2 > 0$ und $M_1 > 0$ $\Rightarrow$ $\mathbf{x}_0$ ist lokales Minimum

(c) $M_2 > 0$ und $M_1 < 0$ $\Rightarrow$ $\mathbf{x}_0$ ist lokales Maximum

(d) $M_2 = 0$ $\Rightarrow$ Keine Aussage möglich

Lokale Extrema / Beispiel

Berechne alle lokalen Extrema der Funktion

$$f(x, y) = \frac{1}{6}x^3 - x + \frac{1}{4}xy^2$$

Stationäre Punkte:

$\mathbf{x}_1 = (0, 2)$: $M_2 = -1 < 0$ $\Rightarrow$ Sattelpunkt

$\mathbf{x}_2 = (0, -2)$: $M_2 = -1 < 0$ $\Rightarrow$ Sattelpunkt

$\mathbf{x}_3 = (\sqrt{2}, 0)$: $M_2 = 1 > 0$, $M_1 = \sqrt{2} > 0$ $\Rightarrow$ lokales Minimum

$\mathbf{x}_4 = (-\sqrt{2}, 0)$: $M_2 = 1 > 0$, $M_1 = -\sqrt{2} < 0$ $\Rightarrow$ lokales Maximum

Hinreichende Bedingung für Lokale Extrema / Allgemein

Sei $\boldsymbol{x_0}$ ein stationärer Punkt von f, und
M_k der k-te Hauptminoren von $\boldsymbol{H}_f(\boldsymbol{x}_0)$.

(a) Alle Hauptminoren $M_k > 0$
$\Rightarrow$ $\boldsymbol{x}_0$ ist ein **lokales Minimum** von f.

(b) Für alle Hauptminoren gilt $(-1)^k M_k > 0$
$\Rightarrow$ $\boldsymbol{x}_0$ ist ein **lokales Maximum** von f.

(c) $\det(\boldsymbol{H}_f(\boldsymbol{x}_0)) \neq 0$, aber weder (a) noch (b) sind erfüllt
$\Rightarrow$ $\boldsymbol{x}_0$ ist ein **Sattelpunkt** von f.

(d) Andernfalls ist keine Aussage möglich, d.h. $\boldsymbol{x}_0$ kann ein lokales Extremum sein, muß aber nicht.

Lokale Extrema / Beispiel

Wir suchen die lokalen Extrema der Funktion

$$f(x_1, x_2, x_3) = (x_1 - 1)^2 + (x_2 + 2)^2 + (x_3 + 1)^2$$

$$\nabla f = (2(x_1 - 1), 2(x_2 + 2), 2(x_3 + 1)) = (0, 0, 0)$$

Der einzige stationäre Punkt ist $\boldsymbol{x}_0 = (1, -2, -1)$.

$$\boldsymbol{H}_f(\boldsymbol{x}) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

Alle Hauptminoren sind positiv: $M_1 = 2$, $M_2 = 4$, $M_3 = 8$

$\Rightarrow$ $\boldsymbol{x}_0$ ist ein lokales Minimum.

Globale Extrema

Aufgabe:
Suche den höchsten Punkt der Welt.

Die Berechnung globaler Extrema von Funktionen in mehr als einer Variablen ist im allgemeinen sehr schwierig.

Wir wollen hier nur folgenden Spezialfall behandeln:

Sei $\boldsymbol{x_0}$ ein stationärer Punkt von f, und
M_k der k-te Hauptminoren von $\boldsymbol{H}_f(\boldsymbol{x})$.

- Falls alle Hauptminoren $M_k > 0$ für **alle** $\boldsymbol{x}$, dann ist $\boldsymbol{x}_0$ ist ein **globales Minimum** von f.
- Falls für alle Hauptminor $(-1)^k M_k > 0$ für **alle** $\boldsymbol{x}$, dann ist $\boldsymbol{x}_0$ ist ein **globales Maximum** von f.

Globale Extrema / Beispiel

Wir suchen die globalen Extrema der Funktion

$$f(x_1, x_2, x_3) = (x_1 - 1)^2 + (x_2 + 2)^2 + (x_3 + 1)^2$$

Der einzige stationäre Punkt ist $\mathbf{x}_0 = (1, -2, -1)$.

$$\mathbf{H}_f(\mathbf{x}) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

Die Hesse-Matrix ist unabhängig von $\mathbf{x}$ und alle Hauptminoren sind (für alle $\mathbf{x}$) positiv: $M_1 = 2$, $M_2 = 4$, $M_3 = 8$

$\Rightarrow$ $\mathbf{x}_0$ ist ein globales Minimum.

Optimierung unter Nebenbedingungen

Aufgabe:
Berechne

$$\max / \min f(x, y)$$

unter der Nebenbedingung (*constraint*)

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

Beispiel:
Wir suchen die lokalen Extrema der Funktion

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

unter der Nebenbedingung

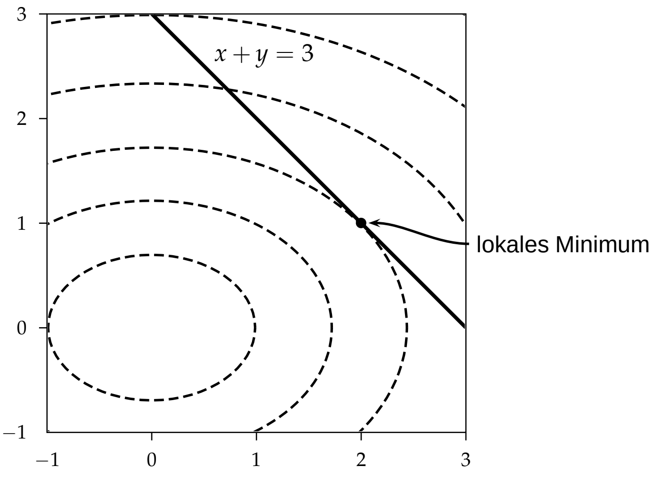
$$g(x, y) = x + y = 3$$

Graphische Lösung

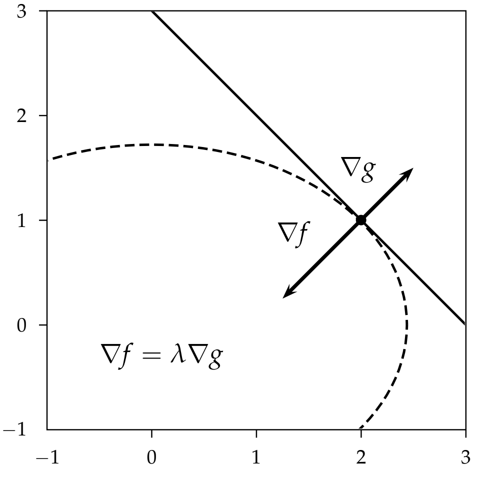
Im Falle von zwei Variablen können wir das Problem graphisch „lösen“.

1. Zeichne die Nebenbedingung $g(x, y) = c$ in die xy-Ebene ein. (Kurve in der Ebene)
2. Zeichne „geeignete“ Niveaulinien der zu optimierenden Funktion $f(x, y)$ ein.
3. Untersuche an hand der Zeichnung welche Niveaulinien den zulässigen Bereich schneiden und bestimme die ungefähre Lage der Extrema.

Graphische Lösung / Beispiel



Gradienten



Lagrangefunktion

Wir erzeugen uns aus f, g und einer Hilfsvariablen λ eine neue Funktion, die **Lagrange-Funktion**:

$$\mathcal{L}(x, y; \lambda) = f(x, y) - \lambda\,(g(x, y) - c)$$

Die Hilfsvariable λ heißt **Lagrange-Multiplikator**.

Wenn die Nebenbedingung erfüllt ist, so stimmen f und $\mathcal{L}$ überein.

Lokale Extrema von f gegeben $g(x, y) = c$ sind stationäre Punkte der Lagrangefunktion $\mathcal{L}$:

$$\begin{aligned}
\mathcal{L}_x &= f_x - \lambda g_x = 0 \\
\mathcal{L}_y &= f_y - \lambda g_y = 0 \\
\mathcal{L}_\lambda &= -(g(x, y) - c) = 0
\end{aligned}$$

Lagrangefunktion / Beispiel

Wir suchen die lokalen Extrema der

$$f(x,y) = x^2 + 2y^2 \qquad \text{gegeben} \qquad g(x,y) = x + y = 3$$

Lagrangefunktion:

$$\mathcal{L}(x,y,\lambda) = (x^2 + 2y^2) - \lambda(x + y - 3)$$

Stationäre Punkte:

$$\begin{aligned} \mathcal{L}_x &= 2x - \lambda & &= 0 \\ \mathcal{L}_y &= 4y - \lambda & &= 0 \\ \mathcal{L}_\lambda &= 3 - x - y & &= 0 \end{aligned}$$

$\Rightarrow$ einziger stationärer Punkt: $(\boldsymbol{x}_0; \lambda_0) = (2, 1; 4)$

Geänderte Hesse-Matrix

Die Matrix

$$\bar{\boldsymbol{H}}(\boldsymbol{x}) = \begin{pmatrix} 0 & g_x & g_y \\ g_x & \mathcal{L}_{xx} & \mathcal{L}_{xy} \\ g_y & \mathcal{L}_{yx} & \mathcal{L}_{yy} \end{pmatrix}$$

heißt die **geänderte Hesse-Matrix**.

Hinreichende Bedingung für lokales Extremum:
Sei $(\boldsymbol{x}_0; \lambda_0)$ ein stationärer Punkt von $\mathcal{L}$.

- $|\bar{\boldsymbol{H}}(\boldsymbol{x}_0)| > 0 \quad \Rightarrow \quad \boldsymbol{x}_0$ ist lokales Maximum.
- $|\bar{\boldsymbol{H}}(\boldsymbol{x}_0)| < 0 \quad \Rightarrow \quad \boldsymbol{x}_0$ ist lokales Minimum.

Geänderte Hesse-Matrix / Beispiel

Wir suchen die lokalen Extrema der

$$f(x,y) = x^2 + 2y^2 \qquad \text{gegeben} \qquad g(x,y) = x + y = 3$$

Lagrangefunktion:$(\boldsymbol{x}_0; \lambda_0) = (2, 1; 4)$

$$\mathcal{L}(x,y,\lambda) = (x^2 + 2y^2) - \lambda(x + y - 3)$$

Stationärer Punkt: $(\boldsymbol{x}_0; \lambda_0) = (2, 1; 4)$
Determinante der geänderte Hesse-Matrix:

$$|\bar{\boldsymbol{H}}(\boldsymbol{x}_0)| = \begin{vmatrix} 0 & g_x & g_y \\ g_x & \mathcal{L}_{xx} & \mathcal{L}_{xy} \\ g_y & \mathcal{L}_{yx} & \mathcal{L}_{yy} \end{vmatrix} = \begin{vmatrix} 0 & 1 & 1 \\ 1 & 2 & 0 \\ 1 & 0 & 4 \end{vmatrix} = -6 < 0$$

$\Rightarrow \quad \boldsymbol{x}_0 = (2, 1)$ ist ein lokales Minimum.

Lagrangemultiplikator / Eine Interpretation

Die Lage des Extremums $\mathbf{x}^*$ hängt von c der Nebenbedingung ab, $\mathbf{x}^* = \mathbf{x}^*(c)$, und somit auch der Extremalwert von f^*:

$$f^*(c) = f(\mathbf{x}^*(c))$$

Wie ändert sich $f^*(c)$ mit c?

Im Optimimum stimmen $\mathcal{L}$ und f überein. Daher ist

$$\frac{df^*(c)}{dc} = \frac{\partial \mathcal{L}}{\partial c} = \frac{\partial}{\partial c}\big(f(\mathbf{x}) - \lambda g(\mathbf{x}) + \lambda c\big) = \lambda$$

$$\boxed{\frac{df^*(c)}{dc} = \lambda}$$

Optimierung unter Nebenbedingungen / Allgemeiner Fall

Aufgabe:
Berechne

$$\max / \min\ f(\mathbf{x})$$

unter den Nebenbedingungen

$$\begin{aligned} g_1(\mathbf{x}) &= c_1 \\ &\vdots \\ g_m(\mathbf{x}) &= c_m \end{aligned} \qquad (m < n)$$

Lagrangefunktion

$$\boxed{\mathcal{L}(\mathbf{x}; \lambda_1, \dots, \lambda_m) = f(\mathbf{x}) - \sum_{i=1}^{m} \lambda_i (g_i(\mathbf{x}) - c_i)}$$