Kapitel 12

Lagrange-Funktion

Optimierung unter Nebenbedingungen

Aufgabe:
Berechne die Extrema der Funktion

$$f(x, y)$$

unter der Nebenbedingung

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

Beispiel:
Wir suchen die lokalen Extrema der Funktion

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

unter der Nebenbedingung

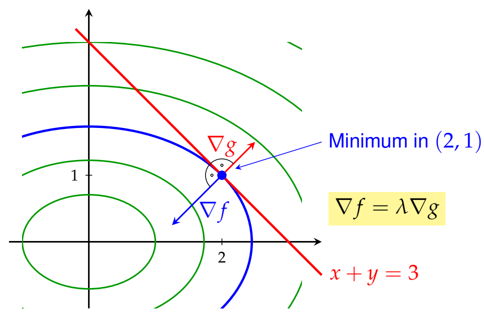
$$g(x, y) = x + y = 3$$

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.

Beispiel – Graphische Lösung



Extrema von $f(x, y) = x^2 + 2\,y^2$ gegeben $g(x, y) = x + y = 3$

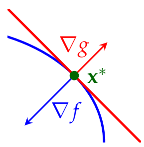
Lagrange-Ansatz

Sei $\mathbf{x}^*$ ein Extremum von $f(x, y)$ gegeben $g(x, y) = c$.
Dann müssen $\nabla f(\mathbf{x}^*)$ und $\nabla g(\mathbf{x}^*)$ proportional sein:

$$\nabla f(\mathbf{x}^*) = \lambda\,\nabla g(\mathbf{x}^*)$$

wobei λ eine geeignete Proportionalitätskonstante ist.

$$\begin{aligned} f_x(\mathbf{x}^*) &= \lambda\, g_x(\mathbf{x}^*) \\ f_y(\mathbf{x}^*) &= \lambda\, g_y(\mathbf{x}^*) \\ g(\mathbf{x}^*) &= c \end{aligned}$$



Umformen ergibt

$$\begin{aligned} f_x(\mathbf{x}^*) - \lambda\, g_x(\mathbf{x}^*) &= 0 \\ f_y(\mathbf{x}^*) - \lambda\, g_y(\mathbf{x}^*) &= 0 \\ c - g(\mathbf{x}^*) &= 0 \end{aligned}$$

Linke Seite ist Gradient von $\mathcal{L}(x, y; \lambda) = f(x, y) + \lambda\,(c - g(x, y))$

Lagrange-Funktion

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

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

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

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

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

Beispiel – Lagrangefunktion

Wir suchen die lokalen Extrema der

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

Lagrangefunktion:

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

Kritische 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 kritischer Punkt: $(\mathbf{x}_0; \lambda_0) = (2, 1; 4)$

Geänderte Hesse-Matrix

Die Matrix

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

heißt **geänderte Hesse-Matrix**.

Hinreichende Bedingung für lokales Extremum:

Sei $(\mathbf{x}_0; \lambda_0)$ ein kritischer Punkt von $\mathcal{L}$.

- $|\bar{\mathbf{H}}(\mathbf{x}_0; \lambda_0)| > 0 \quad \Rightarrow \quad \mathbf{x}_0$ ist *lokales Maximum*
- $|\bar{\mathbf{H}}(\mathbf{x}_0; \lambda_0)| < 0 \quad \Rightarrow \quad \mathbf{x}_0$ ist *lokales Minimum*
- $|\bar{\mathbf{H}}(\mathbf{x}_0; \lambda_0)| = 0 \quad \Rightarrow \quad$ keine Aussage möglich

Beispiel – Geränderte Hesse-Matrix

Wir suchen die lokalen Extrema der

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

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

Kritischer Punkt: $(\mathbf{x}_0; \lambda_0) = (2,1;4)$

Determinante der geränderten Hesse-Matrix:

$$|\bar{\mathbf{H}}(\mathbf{x}_0;\lambda_0)| = \begin{vmatrix} 0 & g_x & g_y \\ g_x & \mathcal{L}_{xx} & \mathcal{L}_{xy} \\ g_y & \mathcal{L}_{yx} & \mathcal{L}_{yy} \end{vmatrix} = \begin{vmatrix} 0 & 1 & 1 \\ 1 & 2 & 0 \\ 1 & 0 & 4 \end{vmatrix} = -6 < 0$$

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

Viele Variablen und Gleichungen

Berechne die Extrema der Funktion

$f(x_1, \dots, x_n)$

unter den Nebenbedingungen

$$\begin{aligned} g_1(x_1, \dots, x_n) &= c_1 \\ &\vdots \qquad\qquad (k < n) \\ g_k(x_1, \dots, x_k) &= c_k \end{aligned}$$

Optimierungsproblem: $\min / \max \; f(\mathbf{x})$ gegeben $\mathbf{g}(\mathbf{x}) = \mathbf{c}$.

Lagrange-Funktion

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

Vorgangsweise – Kritische Punkte

1. Stelle Lagrange-Funktion $\mathcal{L}$ auf:

$$\begin{aligned} &\mathcal{L}(x_1, \dots, x_n; \lambda_1, \dots, \lambda_k) \\ &\quad = f(x_1, \dots, x_n) + \sum_{i=1}^{k} \lambda_i \,(c_i - g_i(x_1, \dots, x_n)) \end{aligned}$$

2. Berechne die ersten partiellen Ableitungen von $\mathcal{L}$.
3. Setze alle ersten partiellen Ableitungen gleich Null und lösen das so entstandene Gleichungssystem mit $n + k$ Unbekannten in $n + k$ Gleichungen.
4. Die ersten n Komponenten $(x_1, \dots, x_n)$ sind die Koordinaten der gesuchten kritischen Punkte.

Beispiel – Kritische Punkte

Wir suchen die kritischen Punkte von

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

unter den Nebenbedingungen

$$x_1 + 2\,x_2 = 2 \quad \text{und} \quad x_2 - x_3 = 3$$

Lagrange-Funktion:

$$\begin{aligned} \mathcal{L}(x_1, x_2, x_3; \lambda_1, \lambda_2) = {} & ((x_1 - 1)^2 + (x_2 - 2)^2 + 2\,x_3^2) \\ & + \lambda_1\,(2 - x_1 - 2\,x_2) + \lambda_2\,(3 - x_2 + x_3) \end{aligned}$$

Beispiel – Kritische Punkte

Partielle Ableitungen (Gradient):

$$\begin{aligned} \mathcal{L}_{x_1} &= 2\,(x_1 - 1) - \lambda_1 && = 0 \\ \mathcal{L}_{x_2} &= 2\,(x_2 - 2) - 2\,\lambda_1 - \lambda_2 && = 0 \\ \mathcal{L}_{x_3} &= 4\,x_3 + \lambda_2 && = 0 \\ \mathcal{L}_{\lambda_1} &= 2 - x_1 - 2\,x_2 && = 0 \\ \mathcal{L}_{\lambda_2} &= 3 - x_2 + x_3 && = 0 \end{aligned}$$

Die kritischen Punkte von $\mathcal{L}$ erhalten wir durch Nullsetzen der ersten partiellen Ableitungen: (lineares Gleichungssystem)

$$x_1 = -\tfrac{6}{7},\ x_2 = \tfrac{10}{7},\ x_3 = -\tfrac{11}{7};\ \lambda_1 = -\tfrac{26}{7},\ \lambda_2 = \tfrac{44}{7}\ .$$

Der einzige kritische Punkt von f unter dieser Nebenbedingungen ist somit

$$\mathbf{x}_0 = (-\tfrac{6}{7}, \tfrac{10}{7}, -\tfrac{11}{7})\ .$$

Geränderte Hesse-Matrix

$$\bar{\mathbf{H}}(\mathbf{x};\boldsymbol{\lambda}) = \begin{pmatrix} 0 & \dots & 0 & \frac{\partial g_1}{\partial x_1} & \dots & \frac{\partial g_1}{\partial x_n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \frac{\partial g_k}{\partial x_1} & \dots & \frac{\partial g_k}{\partial x_n} \\ \frac{\partial g_1}{\partial x_1} & \dots & \frac{\partial g_k}{\partial x_1} & \mathcal{L}_{x_1 x_1} & \dots & \mathcal{L}_{x_1 x_n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_1}{\partial x_n} & \dots & \frac{\partial g_k}{\partial x_n} & \mathcal{L}_{x_n x_1} & \dots & \mathcal{L}_{x_n x_n} \end{pmatrix}$$

Sei $B_r(\mathbf{x};\boldsymbol{\lambda})$ der $(k + r)$-te führende Hauptminor von $\bar{\mathbf{H}}(\mathbf{x};\boldsymbol{\lambda})$.

Hinreichende Bedingung

Sei $(\mathbf{x}_0;\boldsymbol{\lambda}_0)$ ein kritischer Punkt von $\mathcal{L}$.

- $(-1)^k\,B_r(\mathbf{x}_0;\boldsymbol{\lambda}_0) > 0$ für alle $r = k+1, \dots, n$
  $\Rightarrow$ $\mathbf{x}_0$ ist *lokales Minimum*
- $(-1)^r\,B_r(\mathbf{x}_0;\boldsymbol{\lambda}_0) > 0$ für alle $r = k+1, \dots, n$
  $\Rightarrow$ $\mathbf{x}_0$ ist *lokales Maximum*

(n ist die Anzahl der Variablen x_i und
k ist die Anzahl der Nebenbedingungen.)

Beispiel – Hinreichende Bedingung

Suchen Extrema von $\quad f(x_1, x_2, x_3) = (x_1 - 1)^2 + (x_2 - 2)^2 + 2\,x_3^2$

unter den Nebenbedingungen $\quad x_1 + 2\,x_2 = 2 \quad$ und $\quad x_2 - x_3 = 3$

Lagrange-Funktion:

$$\begin{aligned} \mathcal{L}(x_1, x_2, x_3; \lambda_1, \lambda_2) = {} & ((x_1 - 1)^2 + (x_2 - 2)^2 + 2\,x_3^2) \\ & + \lambda_1\,(2 - x_1 - 2\,x_2) + \lambda_2\,(3 - x_2 + x_3) \end{aligned}$$

Kritischer Punkt von $\mathcal{L}$:

$$x_1 = -\tfrac{6}{7},\ x_2 = \tfrac{10}{7},\ x_3 = -\tfrac{11}{7};\ \lambda_1 = -\tfrac{26}{7},\ \lambda_2 = \tfrac{44}{7}\ .$$

Beispiel – Hinreichende Bedingung

Geränderte Hesse-Matrix:

$$\bar{\mathbf{H}}(\mathbf{x};\boldsymbol{\lambda}) = \begin{pmatrix} 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 & -1 \\ 1 & 0 & 2 & 0 & 0 \\ 2 & 1 & 0 & 2 & 0 \\ 0 & -1 & 0 & 0 & 4 \end{pmatrix}$$

3 Variablen, 2 Nebenbedingungen: $n = 3,\ k = 2 \quad \Rightarrow \quad r = 3$

$B_3 = |\bar{\mathbf{H}}(\mathbf{x};\boldsymbol{\lambda})| = 14$

$(-1)^k B_r = (-1)^2 B_3 = 14 \quad > 0$ Bedingung erfüllt

$(-1)^r B_r = (-1)^3 B_3 = -14 \quad < 0$ nicht erfüllt

Der kritische Punkt $\mathbf{x}_0 = (-\frac{6}{7}, \frac{10}{7}, -\frac{11}{7})$ ist ein *lokales Minimum*.

Globale Extrema

Sei $(\mathbf{x}^*, \boldsymbol{\lambda}^*)$ ein kritischer Punkt der Lagrange-Funktion $\mathcal{L}$ des Optimierungsproblems

$$\min / \max \ f(\mathbf{x}) \ \text{ gegeben } \ \mathbf{g}(\mathbf{x}) = \mathbf{c}$$

Falls $\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}^*)$ *konkav* (konvex) in $\mathbf{x}$ ist, dann ist $\mathbf{x}^*$ ein **globales Maximum** (globales Minimum) von $f(\mathbf{x})$ gegeben $\mathbf{g}(\mathbf{x}) = \mathbf{c}$.

Beispiel – Globale Extrema

$(x^*, y^*; \lambda^*) = (2, 1; 4)$ ist ein kritischer Punkt der Lagrange-Funktion des Optimierungsproblems

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

Lagrangefunktion:

$$\mathcal{L}(x, y, \lambda^*) = (x^2 + 2y^2) + 4 \cdot (3 - (x + y))$$

Hesse-Matrix:

$$\mathbf{H}_{\mathcal{L}}(x, y) = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix} \qquad \begin{array}{ll} H_1 = 2 & > 0 \\ H_2 = 8 & > 0 \end{array}$$

$\mathcal{L}$ ist konvex in (x, y).

$(x^*, y^*) = (2, 1)$ ist ein globales Minimum.

Beispiel – Globale Extrema

$$(\mathbf{x}^*; \boldsymbol{\lambda}^*) = (-\tfrac{6}{7}, \tfrac{10}{7}, -\tfrac{11}{7}; -\tfrac{26}{7}, \tfrac{44}{7})$$

ist ein kritischer Punkt der Lagrange-Funktion des Optimierungsproblems

$$\begin{aligned} \min / \max \ & f(x_1, x_2, x_3) = (x_1 - 1)^2 + (x_2 - 2)^2 + 2\,x_3^2 \\ \text{gegeben} \ & g_1(x_1, x_2, x_3) = x_1 + 2\,x_2 = 2 \\ & g_2(x_1, x_2, x_3) = x_2 - x_3 = 3 \end{aligned}$$

Lagrangefunktion:

$$\begin{aligned} \mathcal{L}(\mathbf{x}; \boldsymbol{\lambda}^*) = {} & ((x_1 - 1)^2 + (x_2 - 2)^2 + 2\,x_3^2) \\ & - \tfrac{26}{7}\,(2 - x_1 - 2\,x_2) + \tfrac{44}{7}\,(3 - x_2 + x_3) \end{aligned}$$

Beispiel – Globale Extrema

Hesse-Matrix:

$$\mathbf{H}_{\mathcal{L}}(x, y) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix} \qquad \begin{array}{ll} H_1 = 2 & > 0 \\ H_2 = 4 & > 0 \\ H_3 = 16 & > 0 \end{array}$$

$\mathcal{L}$ ist konvex in $\mathbf{x}$.

$\mathbf{x}^* = (-\frac{6}{7}, \frac{10}{7}, -\frac{11}{7})$ ist ein globales Minimum.

Interpretation des Lagrange-Multiplikators

Die Lage des Extremums $\mathbf{x}^*$ des Optimierungsproblems

$$\min / \max \ f(\mathbf{x}) \ \text{ gegeben } \ \mathbf{g}(\mathbf{x}) = \mathbf{c}$$

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

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

Wie ändert sich $f^*(\mathbf{c})$ mit $\mathbf{c}$?

$$\boxed{\frac{\partial f^*}{\partial c_j}(\mathbf{c}) = \lambda_j^*(\mathbf{c})}$$

Der Lagrange-Multiplikator λ_j gibt also an, wie sich der Extremalwert ändert, wenn die Konstante c_j der Nebenbedingung $g_j(\mathbf{x}) = c_j$ verändert wird.

Herleitung

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

$$\begin{aligned} \frac{\partial f^*(\mathbf{c})}{\partial c_j} &= \frac{\partial \mathcal{L}(\mathbf{x}^*(\mathbf{c}), \boldsymbol{\lambda}(\mathbf{c}))}{\partial c_j} \qquad \text{[Kettenregel]} \\ &= \sum_{i=1}^{n} \underbrace{\mathcal{L}_{x_i}(\mathbf{x}^*(\mathbf{c}), \boldsymbol{\lambda}(\mathbf{c}))}_{\substack{=0 \\ \text{da kritischer Punkt}}} \cdot \frac{\partial x_i^*(\mathbf{c})}{\partial c_j} + \left.\frac{\partial \mathcal{L}(\mathbf{x}, \mathbf{c})}{\partial c_j}\right|_{(\mathbf{x}^*(\mathbf{c}), \boldsymbol{\lambda}^*(\mathbf{c}))} \\ &= \left.\frac{\partial \mathcal{L}(\mathbf{x}, \mathbf{c})}{\partial c_j}\right|_{(\mathbf{x}^*(\mathbf{c}), \boldsymbol{\lambda}^*(\mathbf{c}))} \\ &= \left.\frac{\partial}{\partial c_j}\left(f(\mathbf{x}) + \sum_{i=1}^{k} \lambda_i (c_i - g_i(\mathbf{x}))\right)\right|_{(\mathbf{x}^*(\mathbf{c}), \boldsymbol{\lambda}^*(\mathbf{c}))} \\ &= \lambda_j^*(\mathbf{c}) \end{aligned}$$

Beispiel – Lagrange-Multiplikator

$(x^*, y^*) = (2, 1)$ ist ein Minimum des Optimierungsproblems

$$\min / \max \ f(x, y) = x^2 + 2y^2 \ \text{ gegeben } \ g(x, y) = x + y = c = 3$$

mit $\lambda^* = 4$.

Wie ändert sich der Minimalwert f^* von f, wenn sich c ändert?

$$\frac{df^*}{dc} = \lambda^* = 4$$

Umhüllungssatz (Envelope-Theorem)

Wie ändert sich das Extremum f^* des Optimierungsproblems

$$\min / \max \quad f(\mathbf{x}, \mathbf{p}) \quad \text{gegeben} \quad \mathbf{g}(\mathbf{x}, \mathbf{p}) = \mathbf{c}$$

wenn sich der Parameter (die exogene Variable) $\mathbf{p}$ ändert?

$$\boxed{\frac{\partial f^*(\mathbf{p})}{\partial p_j} = \left.\frac{\partial \mathcal{L}(\mathbf{x}, \mathbf{p})}{\partial p_j}\right|_{(\mathbf{x}^*(\mathbf{p}), \boldsymbol{\lambda}^*(\mathbf{p}))}}$$

Beispiel – Roys Identität

Maximiere Nutzenfunktion

$$\max \quad U(\mathbf{x}) \quad \text{gegeben} \quad \mathbf{p}^t \cdot \mathbf{x} = m$$

Maximaler Nutzen hängt von Preisen $\mathbf{p}$ und Einkommen m ab:

$$U^* = U^*(\mathbf{p}, m)$$ [indirekte Nutzenfunktion]

Lagrange-Funktion $\mathcal{L}(\mathbf{x}, \lambda) = U(\mathbf{x}) + \lambda\,(m - \mathbf{p}^t \cdot \mathbf{x})$

$$\frac{\partial U^*}{\partial p_j} = \frac{\partial \mathcal{L}}{\partial p_j} = -\lambda^* x_j^* \quad \text{und} \quad \frac{\partial U^*}{\partial m} = \frac{\partial \mathcal{L}}{\partial m} = \lambda^*$$

und somit

$$x_j^* = -\frac{\partial U^* / \partial p_j}{\partial U^* / \partial m}$$ [Marshallsche Nachfragefunktion]

Beispiel – Shephards Lemma

Minimiere Ausgaben

$$\min \quad \mathbf{p}^t \cdot \mathbf{x} = m \quad \text{gegeben} \quad U(\mathbf{x}) = \bar{u}$$

(Minimale) Ausgabenfunktion hängt von $\mathbf{p}$ und $\bar{u}$ ab: $e = e(\mathbf{p}, \bar{u})$

Lagrange-Funktion $\mathcal{L}(\mathbf{x}, \lambda) = \mathbf{p}^t \cdot \mathbf{x} + \lambda\,(\bar{u} - U(\mathbf{x}))$

$$\frac{\partial e}{\partial p_j} = \frac{\partial \mathcal{L}}{\partial p_j} = x_j^*$$ [Hicks'sche Nachfragefunktion]

Zusammenfassung

- Optimierung unter Nebenbedingungen
- Graphische Lösung
- Lagrange-Funktion und Lagrange-Multiplikator
- Extremum und kritischer Punkt
- Geränderte Hesse-Matrix
- Globale Extrema
- Interpretation des Lagrange-Multiplikators
- Umhüllungssatz