2.3 Extremalprobleme unter Nebenbedingungen

Frage:

Welche Abmessungen sollte eine Metalldose haben, damit bei vorgegebenem Volumen der Materialverbrauch am geringsten ist?

Sei r der Radius und h die Höhe. Dann gilt

$$\begin{aligned} V &= \pi r^2 h \\ O &= 2\pi r^2 + 2\pi r h \end{aligned}$$

Setze bei vorgebenem $c \in \mathbb{R}_+$

$$\begin{aligned} f(x,y) &= 2\pi x^2 + 2\pi x y \\ g(x,y) &= \pi x^2 y - c = 0 \end{aligned}$$

Bestimme das Minimum der Funktion $f(x,y)$ auf der Menge

$$G := \{(x,y) \in \mathbb{R}^2_+ \; : \; g(x,y) = 0\}$$

Lösung:

Aus $g(x,y) = \pi x^2 y - c = 0$ folgt

$$y = \frac{c}{\pi x^2}$$

Einsetzen in $f(x,y)$ ergibt

$$h(x) := 2\pi x^2 + 2\pi x \frac{c}{\pi x^2} = 2\pi x^2 + \frac{2c}{x}$$

Bestimme das Minimum der Funktion $h(x)$:

$$h'(x) = 4\pi x - \frac{2c}{x^2} = 0 \quad \Rightarrow \quad 4\pi x = \frac{2c}{x^2} \quad \Rightarrow \quad x = \left(\frac{c}{2\pi}\right)^{1/3}$$

Hinreichende Bedingung

$$h''(x) = 4\pi + \frac{4c}{x^3} \quad \Rightarrow \quad h''\left(\left(\frac{c}{\pi}\right)^{1/3}\right) = 12\pi > 0$$

Allgemein:

Bestimme die Extremwerte der Funktion $f : \mathbb{R}^n \to \mathbb{R}$ unter den Nebenbedingungen

$$\mathbf{g}(\mathbf{x}) = 0$$

wobei $\mathbf{g} : \mathbb{R}^n \to \mathbb{R}^m$.
Die Nebenbedingungen lauten also

$$\begin{aligned} g_1(x_1, \ldots, x_n) &= 0 \\ &\vdots \\ g_m(x_1, \ldots, x_n) &= 0 \end{aligned}$$

Alternativ: Bestimme die Extremwerte der Funktion $f(\mathbf{x})$ auf der Menge

$$G := \{\mathbf{x} \in \mathbb{R}^n \ : \ \mathbf{g}(\mathbf{x}) = \mathbf{0}\}$$

Die Lagrange–Funktion:

Wir definieren folgende erweiterte Funktion $F(\mathbf{x})$:

$$F(\mathbf{x}) := f(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i g_i(\mathbf{x})$$

und suchen die Extremwerte von $F(\mathbf{x})$ für festes $\lambda = (\lambda_1, \ldots, \lambda_m)^T$.
Die Zahlen λ_i, $i = 1, \ldots, m$ nennt man **Lagrange–Multiplikatoren**.

Satz: (Lagrange–Lemma)

Minimiert (bzw. maximiert) $\mathbf{x}^0$ die Lagrange–Funktion $F(\mathbf{x})$ (für ein festes λ) über D und gilt $\mathbf{g}(\mathbf{x}^0) = 0$, so liefert $\mathbf{x}^0$ zugleich das Minimum (bzw. Maximum) von $f(\mathbf{x})$ über $G := \{\mathbf{x} \in D \ : \ \mathbf{g}(\mathbf{x}) = \mathbf{0}\}$.

Beweis: Für ein beliebiges $\mathbf{x} \in D$ gilt nach Vorausetzung

$$f(\mathbf{x}^0) + \lambda^T \mathbf{g}(\mathbf{x}^0) \leq f(\mathbf{x}) + \lambda^T \mathbf{g}(\mathbf{x})$$

Wählt man speziell $\mathbf{x} \in G$, so ist $\mathbf{g}(\mathbf{x}) = \mathbf{g}(\mathbf{x}^0) = \mathbf{0}$, also auch $f(\mathbf{x}^0) \leq f(\mathbf{x})$.

Bemerkung:

Sind f und g_i, $i = 1, \ldots, m$, $\mathcal{C}^1$–Funktionen, so ist eine notwendige Bedingung für eine Extremstelle $\mathbf{x}^0$ von $F(\mathbf{x})$ gegeben durch

$$\text{grad}\, f(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i \text{grad}\, g_i(\mathbf{x}) = \mathbf{0}^T$$

Zusammen mit den Nebenbedinungen $\mathbf{g}(\mathbf{x}) = 0$ ergibt sich ein (nichtlineares) Gleichungssystem mit $(n+m)$ Gleichungen und $(n+m)$ Unbekannten $\mathbf{x}$ und λ.

Die Lösungen $(\mathbf{x}^0, \lambda^0)$ sind die Kandidaten für die gesuchten Extremstellen (**Notwendige Bedingung**).

Alternativ: Definiere eine Langrange–Funktion

$$G(\mathbf{x}, \lambda) := f(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i g_i(\mathbf{x})$$

und suche die Extremstellen von $G(\mathbf{x}, \lambda)$ bezüglich $\mathbf{x}$ **und** λ.

Bemerkung:

Man kann auch eine **hinreichende** Bedingung aufstellen:

Sind die Funktionen f und $\mathbf{g}$ sogar $\mathcal{C}^2$–Funktionen und ist die Hesse–Matrix $\mathbf{H}F(\mathbf{x}^0)$ der Lagrange–Funktion positiv (bzw. negativ) definit, so ist $\mathbf{x}^0$ tatsächlich ein strenges lokales Minimum (bzw. Maximum) von $f(\mathbf{x})$ auf G.

In den meisten Anwendungen ist die hinreichende Bedingung allerdings **nicht** erfüllt, obwohl $\mathbf{x}^0$ ein strenges lokales Extremum **ist**.

Insbesondere kann man aus der Indefinitheit der Hesse–Matrix $\mathbf{H}F(\mathbf{x}^0)$ **nicht** schließen, dass $\mathbf{x}^0$ kein Extremwert ist.

Ähnlich problematisch ist die hinreichende Bedingung, die man aus der Hesse–Matrix für die Lagrange–Funktion $G(\mathbf{x}, \lambda)$ bezüglich $\mathbf{x}$ **und** λ erhält.

Beispiel:

Gesucht seien die Extrema von $f(x,y) := xy$ auf der Kreisscheibe

$$K := \{(x,y)^T \, : \, x^2 + y^2 \leq 1\}$$

Da die betrachte Funktion stetig ist und $K \subset \mathbb{R}^2$ kompakt ist, folgt aus der Min–Max–Eigenschaft die Existenz von globalen Maxima und Minima auf K.

Wir betrachten zunächst das Innere K^0 von K, also

$$K^0 := \{(x,y)^T \, : \, x^2 + y^2 < 1\}$$

Dies ist eine offene Menge und die notwendige Bedingung für einen Extremwert lautet

$$\text{grad}\, f = (y,x) = \mathbf{0}^T$$

Also erhalten wir als Kandidaten den Ursprung $\mathbf{x}^0 = \mathbf{0}$.

Beispiel: (Fortsetzung)

Die Hesse–Matrix im Punkt $\mathbf{x}^0 = \mathbf{0}$ lautet

$$\mathbf{H}f(\mathbf{0}) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

und ist **indefinit**. Daher ist $\mathbf{x}^0$ ein **Sattelpunkt**.

Die Extrema der Funktion müssen also auf dem Rand liegen, der ein Gleichungsnebenbedingung darstellt:

$$g(x,y) = x^2 + y^2 - 1 = 0$$

Wir suchen also die Extremwerte von $f(x,y) = xy$ unter der Nebenbedingung $g(x,y) = 0$.

Die Lagrange–Funktion lautet

$$F(x,y) = xy + \lambda(x^2 + y^2 - 1)$$

Damit ergibt sich das (nichtlineare) Gleichungssystem

$$\begin{aligned} y + 2\lambda x &= 0 \\ x + 2\lambda y &= 0 \\ x^2 + y^2 &= 1 \end{aligned}$$

mit den vier Lösungen

$$\begin{array}{lcll} \lambda = \frac{1}{2} & : & \mathbf{x}^{(1)} = (\sqrt{0.5}, -\sqrt{0.5})^T & \mathbf{x}^{(2)} = (-\sqrt{0.5}, \sqrt{0.5})^T \\ \lambda = -\frac{1}{2} & : & \mathbf{x}^{(3)} = (\sqrt{0.5}, \sqrt{0.5})^T & \mathbf{x}^{(4)} = (-\sqrt{0.5}, -\sqrt{0.5})^T \end{array}$$

Minima und Maxima lassen sich nun einfach aus den Funktionswerten ablesen

$$f(\mathbf{x}^{(1)}) = f(\mathbf{x}^{(2)}) = -0.5 \qquad f(\mathbf{x}^{(3)}) = f(\mathbf{x}^{(4)}) = 0.5$$

d.h. Minima sind $\mathbf{x}^{(1)}$ und $\mathbf{x}^{(2)}$, Maxima $\mathbf{x}^{(3)}$ und $\mathbf{x}^{(4)}$.

Satz: (Lagrange–Multiplikatoren–Regel)

Seien $f, g_1, \ldots, g_m : D \to \mathbb{R}$ jeweils $\mathcal{C}^1$–Funktionen, und sei $\mathbf{x}^0 \in D$ ein lokales Extremum von $f(\mathbf{x})$ unter der Nebenbedingung $\mathbf{g}(\mathbf{x}) = \mathbf{0}$.

Ferner gelte die **Regularitätsbedingung**

$$\text{Rang}\,(\mathbf{J}\,\mathbf{g}(\mathbf{x}^0)) = m$$

Dann existieren Lagrange–Multiplikatoren $\lambda_1, \ldots, \lambda_m$, so dass für die Lagrange Funktion

$$F(\mathbf{x}) := f(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i g_i(\mathbf{x})$$

die folgende **notwendige Bedingung erster Ordnung** gilt:

$$\text{grad}\, F(\mathbf{x}^0) = \mathbf{0}$$

Bemerkung:

1) **Notwendige Bedingung zweiter Ordnung**

   Ist $\mathbf{x}^0 \in D$ ein lokales Minimum von $f(\mathbf{x})$ unter der Nebenbedingung $\mathbf{g}(\mathbf{x}) = 0$, ist die Regularitätsbedingung erfüllt und sind $\lambda_1, \ldots, \lambda_m$ zugehörige Lagrange–Multiplikatoren, so ist die Hesse–Matrix $\mathbf{H}F(\mathbf{x}^0)$ der Lagrange–Funktion positiv semidefinit auf dem Tangentialraum

   $$TG(\mathbf{x}^0) := \{\mathbf{y} \in \mathbb{R}^n \ : \ \text{grad}\, g_i(\mathbf{x}^0) \cdot \mathbf{y} = 0,\, i = 1, \ldots, m\}$$

   d.h., es gilt

   $$\mathbf{y}^T\, \mathbf{H}F(\mathbf{x}^0)\, \mathbf{y} \geq 0 \quad \forall\, \mathbf{y} \in TG(\mathbf{x}^0)$$

Bemerkung: (Fortsetzung)

2) **Hinreichende Bedingung**

   Ist für einen Punkt $\mathbf{x}^0 \in G$ die Regularitätsbedingung erfüllt, existieren Lagrange–Multiplikatoren $\lambda_1, \ldots, \lambda_m$, so dass $\mathbf{x}^0$ ein stationärer Punkt der zugehörigen Lagrange–Funktion ist, und ist die Hesse–Matrix $\mathbf{H}F(\mathbf{x}^0)$ positiv definit auf dem Tangentialraum $TG(\mathbf{x}^0)$, d.h., gilt

   $$\mathbf{y}^T\, \mathbf{H}F(\mathbf{x}^0)\, \mathbf{y} > 0 \quad \forall\, \mathbf{y} \in TG(\mathbf{x}^0)$$

   so ist $\mathbf{x}^0$ ein strenges lokales Minimum von $f(\mathbf{x})$ unter der Nebenbedingung $\mathbf{g}(\mathbf{x}) = \mathbf{0}$.

Beispiel:

Man bestimme das globale Maximum der Funktion

$$f(x,y) = -x^2 + 8x - y^2 + 9$$

unter der Nebenbedingung

$$g(x,y) = x^2 + y^2 - 1 = 0$$

Die Lagrange–Funktion ist

$$F(x) = -x^2 + 8x - y^2 + 9 + \lambda(x^2 + y^2 - 1)$$

Die notwendige Bedingung ergibt das nichtlineare System

$$\begin{aligned} -2x + 8 &= -2\lambda x \\ -2y &= -2\lambda y \\ x^2 + y^2 &= 1 \end{aligned}$$

Beispiel: (Fortsetzung)

$$\begin{aligned} -2x + 8 &= -2\lambda x \\ -2y &= -2\lambda y \\ x^2 + y^2 &= 1 \end{aligned}$$

Aus der ersten Gleichung folgt $\lambda \neq 1$. Verwendet man dies in der zweiten Gleichung, so gilt $y = 0$. Aus der dritten Gleichung erkennt man sofort $x = \pm 1$.

Demnach sind die beiden Punkte $(x,y) = (1,0)$ und $(x,y) = (-1,0)$ Kandidaten für das globale Maximum. Wegen

$$f(1,0) = 16 \qquad f(-1,0) = 0$$

liegt das globale Maximum von $f(x,y)$ unter der Nebenbedingung $g(x,y) = 0$ im Punkt $(x,y) = (1,0)$.

Beispiel: Man bestimme die lokalen Extremwerte der Funktion

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

auf dem Durchschnitt des Zylinders

$$Z := \{(x, y, z)^T \in \mathbb{R}^3 \,:\, x^2 + y^2 = 2\}$$

mit der Ebene

$$E := \{(x, y, z)^T \in \mathbb{R}^3 \,:\, x + z = 1\}$$

Umformulierung: Bestimme die Extremwerte der Funktion $f(x, y, z)$ unter den Nebenbedingungen

$$\begin{aligned} g_1(x, y, z) &:= x^2 + y^2 - 2 = 0 \\ g_2(x, y, z) &:= x + z - 1 = 0 \end{aligned}$$

Beispiel: (Fortsetzung)

Die Jacobi–Matrix

$$\mathbf{Jg}(\mathbf{x}) = \begin{pmatrix} 2x & 2y & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

hat den Rang 2, d.h. wir können über die Lagrange–Funktion Extremwerte bestimmen:

$$F(x, y, z) = 2x + 3y + 2z + \lambda_1(x^2 + y^2 - 2) + \lambda_2(x + z - 1)$$

Die notwendige Bedingung ergibt das nichtlineare Gleichungssystem

$$\begin{aligned} 2 + 2\lambda_1 x + \lambda_2 &= 0 \\ 3 + 2\lambda_1 y &= 0 \\ 2 + \lambda_2 &= 0 \\ x^2 + y^2 &= 2 \\ x + z &= 1 \end{aligned}$$

Beispiel: (Fortsetzung)

$$\begin{aligned} 2+2\lambda_1 x+\lambda_2 &= 0 \\ 3+2\lambda_1 y &= 0 \\ 2+\lambda_2 &= 0 \\ x^2+y^2 &= 2 \\ x+z &= 1 \end{aligned}$$

Aus der ersten und dritten Gleichung folgt

$$2\lambda_1 x = 0$$

Aus der zweiten Gleichung folgt $\lambda_1 \neq 0$, also $x = 0$.
Damit ergeben sich die möglichen Extremwerte als

$$(x,y,z) = (0,\sqrt{2},1) \quad (x,y,z) = (0,-\sqrt{2},1)$$

Die möglichen Extremwerte sind also

$$(x,y,z) = (0,\sqrt{2},1) \quad (x,y,z) = (0,-\sqrt{2},1)$$

Man berechnet nun die zugehörigen Funktionswerte

$$\begin{aligned} f(0,\sqrt{2},1) &= 3\sqrt{2}+2 \\ f(0,-\sqrt{2},1) &= -3\sqrt{2}+2 \end{aligned}$$

Daher liegt im Punkt $(x,y,z) = (0,\sqrt{2},1)$ ein Maximum, im Punkt $(x,y,z) = (0,-\sqrt{2},1)$ ein Minimum.