6.6 Extrema unter Nebenbedingungen

Ziel: Bestimme Extrema der Funktion $z = f(x_1, \ldots, x_n)$ unter den Nebenbedingungen

$$
\begin{array}{rcl}
g_1(x_1, \ldots, x_n) & = & 0 \\
g_2(x_1, \ldots, x_n) & = & 0 \\
\ldots\ldots\ldots & & \\
g_m(x_1, \ldots, x_n) & = & 0
\end{array}
$$

Lassen sich die Bedingungen

$$g_j(x_1, \ldots, x_n) = 0,\ j = 1, 2, \ldots, m\,,$$

nach m der Variablen auflösen, etwa nach $x_1, \ldots, x_m$, dann gilt:

$$
\begin{array}{rcl}
x_1 & = & \varphi_1(x_{m+1}, x_{m+2}, \ldots, x_n) \\
x_2 & = & \varphi_2(x_{m+1}, x_{m+2}, \ldots, x_n) \\
 & & \ldots\ldots\ldots \\
x_m & = & \varphi_m(x_{m+1}, x_{m+2}, \ldots, x_n)
\end{array}
$$

Einsetzen in $f(x_1, x_2, \ldots, x_n)$ liefert eine Funktion $\Phi(x_{m+1}, x_{m+2}, \ldots, x_n)$, deren Extrema dann gesucht werden.

Beispiel 14. (SCHWARZE Band 2, 13.3.1(b)) Ziel: Stelle quaderförmige Blechschachteln (Länge l, Breite b und Höhe h) vorgegebenen Gewichts G und größtem Volumen her. Ein cm^2 Blech wiege a Gramm. Wir erhalten so das Problem

$$\begin{aligned} \text{maximiere } V(l,b,h) &= l \cdot b \cdot h \\ \text{unter } G &= 2 \cdot a \cdot (l \cdot b + l \cdot h + b \cdot h) \end{aligned}$$

Wir lösen die Nebenbedingung nach l auf und erhalten

$$l = \frac{\frac{G}{2a} - bh}{b+h}$$

Für $\frac{G}{2a}$ setzen wir A. Einsetzen in V liefert dann

$$V(b,h) = \frac{A - bh}{b+h} bh = \frac{Abh - b^2h^2}{b+h}$$

Beachte dass V eine Funktion in den Variablen b und h ist. Die Variable l haben wir eliminiert.

Von dieser Funktion müssen wir nun ein Maximnum bestimmen. Dazu bestimmen wir die ersten partiellen

Ableitungen:

$$
\begin{aligned}
\frac{\partial V}{\partial b} &= \frac{h^2(A - b^2 - 2bh)}{(b+h)^2} \\
\frac{\partial V}{\partial h} &= \frac{b^2(A - h^2 - 2bh)}{(b+h)^2}
\end{aligned}
$$

Wir müssen nun b und h so bestimmen, dass beide partiellen Ableitungen 0 sind. Das wäre für $b = h = 0$ der FAll, was aber offensichtlich keine sinnvolle Lösung für unser Optimierungsproblem ist. Wir erhalten $b = \pm h$. Weil negative Lösungen ebenfalls nicht sinnvoll sind, gilt

$$b = h = \sqrt{\frac{A}{3}}.$$

Man rechnet leicht nach, dass dann auch $l = \sqrt{\frac{A}{3}}$. Unser Volumen hat also für den Würfel einen Extremwert. Weil dies das einzige Extremum ist, kann man sich leicht klarmachen, dass es ein Maximum sein muss: Denn sicherlich muss das Volumen irgendwo maximiert werden. Wenn dies nicht für den Würfel passiert, müsste es ja einen anderen stationären Punkt geben, was aber nicht der Fall ist. Alternativ können Sie auch die Hessematrix bestimmen:

$$\mathbf{H} = \begin{pmatrix} \frac{-2h^2(h^2+A)}{(b+h)^3} & \frac{-2bh(-A+b^2+3bh+h^2)}{(b+h)^3} \\ \frac{-2bh(-A+b^2+3bh+h^2)}{(b+h)^3} & \frac{-2b^2(b^2+A)}{(b+h)^3} \end{pmatrix}$$

Die Hessematrix $\mathbf{H} = (h_{i,j})$ ist negativ definit für $b = h = \sqrt{A/3}$, denn

$$\begin{aligned} h_{1,1} &= -\frac{\sqrt{3A}}{3} < 0 \\ \det(\mathbf{H}) &= \frac{A}{4} > 0 \end{aligned}$$

Der Würfel maximiert also in der Tat das Volumen.

In vielen Fällen ist es nicht möglich, die Gleichungsrestriktionen aufzulösen. Um auch diesen Fall zu behandeln, definieren wir zunächst die **Lagrange-Funktion**:

$$\begin{aligned} L(x;\lambda) &= L(x_1,\ldots,x_n;\ \lambda_1,\ldots,\lambda_m) \\ &= f(x_1,\ldots,x_n) + \sum_{j=1}^{m} \lambda_j \cdot g_j(x_1,\ldots,x_n) \end{aligned}$$

Die λ_j, $j = 1,\ldots,m$, heißen **Lagrange-Multiplikatoren**.

Ferner definieren wir die **Jacobi-Matrix J** der g_i an der Stelle $\mathbf{a}$:

$$\mathbf{J}_g(\mathbf{a}) = \left(\frac{\partial g_j}{\partial x_k}(\mathbf{a})\right)_{j=1,\dots,m;\, k=1,\dots,n}$$

Die Matrix $\mathbf{J}$ ist eine $m \times n$ Matrix. Man kann Sie als *Ableitung* von $(g_1, \dots, g_m$ an der Stelle $\mathbf{a} \in \mathbb{R}^n$ auffassen.

Wir erhalten folgende notwendige Bedingung für Extrempunkte:

> Seien $f : D \to \mathbb{R}$ und $g_j : D \to \mathbb{R}$, $j = 1, 2, \dots, m$, $D \subseteq \mathbb{R}^n$, stetig partiell differenzierbare Funktionen. Die Funktion f habe an der Stelle $\mathbf{a} \in D^o$ (dem Innern von D) ein relatives Extremum unter den Nebenbedingungen $g_j(\mathbf{a}) = 0$, $j = 1, \dots, m$, und die Jacobi-Matrix $\mathbf{J}_g(\mathbf{a})$ habe den Rang m. Dann gibt es $\lambda^* = (\lambda_1^*, \dots, \lambda_m^*) \in \mathbb{R}^m$, so dass für $i = 1, \dots, n$ gilt:
>
> $$L_{x_i}(\mathbf{a}, \lambda^*) = \frac{\partial f}{\partial x_i}(\mathbf{a}) + \sum_{j=1}^{m} \lambda_j^* \frac{\partial g_j}{\partial x_i}(\mathbf{a}) = 0\ .$$

Beachte, dass die partiellen Ableitungen

$$\frac{\partial L}{\partial \lambda_j} = g_j$$

gerade die Gleichungsrestriktionen sind. Wenn wir also die partielle Ableitung der Lagrangefunktion L nach λ_i gleich 0 setzen, ist das gleichbedeutend damit, g_i gleich 0 zu setzen. Wir können also sagen, dass die Lagrange-Funktion $L(x, \lambda)$ an der Stelle (a, λ^0) einen stationären Punkt hat.

Um potentielle relative Extrema für f in der obigen Situation zu finden, werden also $n + m$ Gleichungen für die $n + m$ Unbekannten $x_1, \ldots, x_n, \lambda_1, \ldots, \lambda_m$ gelöst, um stationäre Punkte der Lagrange-Funktion zu finden.

Beispiel 15. Wir wollen

$$f(x_1, x_2, x_3, x_4) = x_1^2 + x_2^2 + x_3^2 + x_4^2$$

unter den Nebenbedingungen

$$\begin{aligned} x_1 + x_2 \quad &= 2 \\ x_2 + x_3 + x_4 \quad &= 4 \end{aligned}$$

minimieren. Die beiden Gleichungen g_1 und g_2 sind also

$$\begin{aligned} g_1 &= 2 - x_1 - x_2 \\ g_2 &= 4 - x_2 - x_3 - x_4 \end{aligned}$$

Die Nebenbedingungen sind

$$g_1 = 0, \qquad g_2 = 0.$$

Die Lagrangefunktion ist

$$\begin{aligned} L(x_1, x_2, x_3, x_4, \lambda_1, \lambda_2) &= x_1^2 + x_2^2 + x_3^2 + x_4^2 + \\ &\quad +\lambda_1(2 - x_1 - x_2) + \\ &\quad +\lambda_2(4 - x_2 - x_3 - x_4). \end{aligned}$$

Um stationäre Punkte zu finden, müssen wir die partiellen Ableitungen bilden und diese 0 setzen:

$$\begin{aligned} L_{x_1} &= 2x_1 - \lambda_1 \\ L_{x_2} &= 2x_2 - \lambda_1 - \lambda_2 \\ L_{x_3} &= 2x_3 - \lambda_2 \\ L_{x_4} &= 2x_4 - \lambda_2 \\ L_{\lambda_1} &= 2 - x_1 - x_2 \\ L_{\lambda_2} &= 4 - x_2 - x_3 - x_4 \end{aligned}$$

Beachte dass die letzten beiden Gleichungen nichts anderes als unsere Gleichungsrestriktionen $g_1 = 0$ und $g_2 = 0$ sind. Das liefert das inhomogene Gleichungssystem

$$\begin{pmatrix} 2 & 0 & 0 & 0 & -1 & 0 \\ 0 & 2 & 0 & 0 & -1 & -1 \\ 0 & 0 & 2 & 0 & 0 & -1 \\ 0 & 0 & 0 & 2 & 0 & -1 \\ -1 & -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & -1 & -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ -2 \\ -4 \end{pmatrix}$$

mit der eindeutigen Lösung

$$x_1 = 0,4; \quad x_2 = 1,6; \quad x_3 = x_4 = 1,2;$$
$$\lambda_1 = 0,8; \quad \lambda_2 = 2,4$$

Die Jacobimatrix ist (unabhängig von $\mathbf{a}$)

$$\mathbf{J} = \begin{pmatrix} -1 & -1 & 0 & 0 \\ 0 & -1 & -1 & -1 \end{pmatrix}$$

hat also den Rang 2.

Bevor wir nun zu hinreichenden Bedingungen kommen, ein Wort zu der Bedingung, dass die Jacobi-Matrix vollen Rang hat. Diese Bedingung wird beispielsweise in SCHWARZE unterschlagen:

Beispiel 16. Wir betrachten das Problem

$$\begin{array}{rrcl} & \text{minimiere } -x_1 & & \\ \text{unter} & x_2-(1-x_1)^3 & = & 0 \\ & -x_2-(1-x_1)^3 & = & 0 \end{array}$$

Beachte, dass die Nebenbedingungen hier nur eine vornehme Art sind, $(1-x_1)^3 = 0$ auszudrücken. Wir haben also ein Minimum an der Stelle $x_1 = 1$ (dem einzigen Punkt, der beide Gleichungsrestriktionen erfüllt). Die Lagrangefunktion lautet

$$L(x_1,x_2,\lambda_1,\lambda_2) = -x_1+\lambda_1(x_2-(1-x_1)^3)-\lambda_2(x_2+(1-x_1)^3)$$

An der Stelle $x_1 = 1, x_2 = 0$ haben wir aber keinen stationären Punkt, weil

$$\frac{\partial L}{\partial x_1}\begin{pmatrix}1\\0\end{pmatrix} = -1.$$

In diesem Fall hat die Jacobi-Matrix

$$\mathbf{J}_g\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}0 & 1\\0 & -1\end{pmatrix}$$

nicht den Rang 2.

Nun zu den hinreichenden Bedingungen: Ist (a, λ^*) ein stationärer Punkt der Lagrange-Funktion, dann betrachten wir die Hesse-Matrix von L nach den Variablen $x_1, \ldots, x_n$ für λ^*. Bezeichne diese Hesse-Matrix mit $\hat{\mathbf{H}}$:

$$\hat{\mathbf{H}}_L(x; \lambda^*) = \begin{pmatrix} L_{x_1x_1}(x; \lambda^*) & \cdots & L_{x_1x_n}(x; \lambda^*) \\ \vdots & \ddots & \vdots \\ L_{x_nx_1}(x; \lambda^*) & \cdots & L_{x_nx_n}(x; \lambda^*) \end{pmatrix}$$

Hinreichende Bedingung

Es seien $f : D \to \mathbb{R}$ und $g_j : D \to \mathbb{R}$, $D \subseteq \mathbb{R}^n$, zweimal stetig partiell differenzierbare Funktionen. Die zugehörige Lagrange-Funktion $L(x; \lambda)$ habe an der Stelle $(\mathbf{a}; \lambda^*)$, $\mathbf{a} \in D^o$, einen stationären Punkt. Ist $\hat{\mathbf{H}}_L(\mathbf{a}; \lambda^*)$ positiv definit (negativ definit), so ist $\mathbf{a}$ ein relatives Minimum (Maximum) unter den Nebenbedingungen $g_j(\mathbf{a}) = 0$ für $j = 1, 2, \ldots, m$.
Ist $\hat{H}_L(x; \lambda^*)$ sogar positiv definit (negativ definit) für alle $x \in D$, so ist $\mathbf{a}$ ein globales Minimum (Maximum) unter den Nebenbedingungen $g_j(\mathbf{a}) = 0$ für $j = 1, 2, \ldots, m$.

Beispiel 17. In Beispiel 15. ist die Matrix $\hat{\mathbf{H}}$ unabhängig von $\mathbf{a}$ und λ:

$$\hat{\mathbf{H}}_L = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

ist positiv definit, wir haben also ein Minimum.

Beispiel 18. Wir wollen Extrema der Funktion

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

unter den Restriktionen

$$\begin{aligned} x + 2y + z &= 1 \\ 2x - y - 3z &= 4 \end{aligned}$$

bestimmen. Die Lagrangefunktion L ist

$$\begin{aligned} L(x, y, z, \lambda_1, \lambda_2) &= x^2 + y^2 + z^2 + \\ &\quad +\lambda_1(-x - 2y - z + 1) + \\ &\quad +\lambda_2(-2x + y + 3z + 4) \end{aligned}$$

Die partiellen Ableitungen sind

$$\begin{aligned}
L_x &= 2x - \lambda_1 - 2\lambda_2 \\
L_y &= 2y - 2\lambda_1 + \lambda_2 \\
L_z &= 2z - \lambda_1 + 3\lambda_2 \\
L_{\lambda_1} &= -x - 2y - z + 1 \\
L_{\lambda_2} &= -2x + y + 3z + 4
\end{aligned}$$

Alle diese Ableitungen sollen gleich Null sein. Die ersten beiden Gleichungen zeigen dann

$$\lambda_1 = \frac{2}{5}x + \frac{4}{5}y, \quad \lambda_2 = \frac{4}{5}x - \frac{2}{5}y$$

Wenn wir dies in die dritte Gleichung einsetzen bekommen wir

$$x - y + z = 0.$$

Diese Gleichung zusammen mit den letzten beiden Gleichungen liefert das inhomogene Gleichungssystem

$$\begin{pmatrix} 1 & -1 & 1 \\ 1 & 2 & 1 \\ 2 & -1 & -3 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 4 \end{pmatrix}$$

mit der Lösung

$$x_0 = \frac{16}{15}, \quad y_0 = \frac{1}{3}, \quad z_0 = -\frac{11}{15}$$

Die Multiplikatoren sind

$$\lambda_1 = \frac{52}{75}, \quad \lambda_2 = \frac{54}{75}$$

Die zugehörige Hessematrix $\hat{\mathbf{H}}$ ist

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

also positiv definit. Deshalb haben wir an der Stelle (x_0, y_0, z_0) ein Minimum!

Lagrangemultiplikatoren haben eine wichtige ökonomische Interpretation. Wir erläutern dies an einem Beispiel mit nur einer Gleichungsrestriktion. Für den allgemeinen Fall verweisen wir auf die Literatur.

Unser Ziel ist

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

unter der Restriktion

$$g(\mathbf{x}) = c.$$

Angenommen, $\mathbf{x}^*$ löst dieses Problem. Dann ist $\mathbf{x}^*$ in der Regel eine Funktion abhängig von c, z.B. $\mathbf{x}^*(c)$. Auch der zugehörige Lagrange-Multiplikator λ ist eine Funktion von c, also $\lambda(c)$. Das Optimum von f ist

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

also eine Funktion abhängig von c. Unter geeigneten Annahmen (auf die wir hier nicht eingehen) kann man zeigen

$$\frac{\mathrm{d}f^*(c)}{\mathrm{d}c} = \lambda(c).$$

Das heisst, der Lagrangemultiplikator ist ein Maß, wie sich das Maximum $f^*(c)$ relativ zu c ändert. Bezeichnet f etwa den Profit, die Restriktion $g(\mathbf{x}) = c$ die Verfügbarkeit einer knappen Ressource, dann ist $\lambda(c)$ ein Maß dafür, wie sich der Profit relativ zu einer Änderung der knappen Ressource ändert.

Wir erläutern dies an einem Beispiel:

Beispiel 19. Die Firma AUD benutzt als Input K (Kapital; damit sind insbesondere Maschinen gemeint) und W (Arbeit), um ein Auto zu produzieren. Um ein Auto zu produzieren müssen insgesamt

$$Q = F(K, W) = K^{1/2} W^{1/4}$$

Produktionsmittel eingesetzt werden; d.h. die Firma hat eine gewisse Freiheit, wieviel Kapital und wieviel Arbeit sie einsetzt. Kapital ist hier wertvoller als Arbeit: Wenn Sie den Kapitaleinsatz verdoppeln, können Sie den Arbeitseinsatz um den Faktor 4 verringern. Kapital und Arbeit konkurrieren: Sowohl das Kapital kostet Geld (Zinsen, Refinanzierung) als auch Arbeit (was jedem klar ist). Die Kosten fürs Kapital seien r, die für Arbeit w. Wir erhalten das Optimierungsproblem

$$\min \; rK + wW$$

unter der Restriktion

$$Q = K^{1/2} W^{1/4}.$$

Die Lagrangefunktion ist

$$L(K, W, \lambda) = rK + wW + \lambda(K^{1/2} W^{1/4} - Q).$$

Partielle Ableitung nach K und W sind

$$\frac{\partial L}{\partial K} = r - \frac{1}{2}\lambda K^{-1/2} W^{1/4} \quad \frac{\partial L}{\partial W} = w - \frac{1}{4}\lambda K^{-1/2} W^{-3/4}$$

Beide partiellen Ableitungen sollen 0 sein, also

$$r = \frac{1}{2}\lambda K^{-1/2} W^{1/4}, \quad w = \frac{1}{4}\lambda K^{1/2} W^{3/4}$$

Auflösen nach λ gibt

$$\lambda = 2rK^{1/2}W^{-1/4} = 4wK^{-1/2}W^{3/4}$$

also

$$2rK = 4wW, \quad \text{also} \quad W = \frac{rK}{2w}.$$

Wir setzen dies in die Restriktion $Q = K^{1/2}W^{1/4}$ ein und erhalten

$$K^{3/4} = Q\sqrt[4]{\frac{2w}{r}}$$

Einfache Umformungen liefern als Lösung des Lagrangeproblems

$$\begin{aligned}
K^* &= 2^{1/3} r^{-1/3} w^{1/3} Q^{4/3} \\
W^* &= 2^{-2/3} r^{2/3} w^{-2/3} Q^{4/3} \\
\lambda^* &= 2^{4/3} r^{2/3} w^{1/3} Q^{1/3} \\
C^* &= 3 \cdot 2^{-2/3} r^{2/3} w^{1/3} Q^{4/3}
\end{aligned}$$

Die Matrix $\hat{\mathbf{H}}$ ist

$$\hat{\mathbf{H}} = \begin{pmatrix} \frac{1}{4}\frac{\lambda^* W^{1/4}}{K^{3/2}} & -\frac{1}{8}\frac{\lambda^*}{K^{1/2}W^{3/4}} \\ -\frac{1}{8}\frac{\lambda^*}{K^{1/2}W^{3/4}} & \frac{3}{16}\frac{\lambda K^{1/2}}{W^{3/4}} \end{pmatrix}$$

Der (1,1)-Eintrag dieser Matrix ist offenbar > 0. Die Determinante erhält man nach einigem Rechnen als

$$\det(\hat{\mathbf{H}}) = \frac{1}{32}\frac{\lambda^2}{KL^{3/2}}$$

sie ist also auch > 0 und somit ist die Matrix positiv definit, wir haben also ein Minimum. Man rechnet leicht nach, das

$$\frac{\mathrm{d}C^*}{\mathrm{d}Q} = 2^{4/3}r^{2/3}w^{1/3}Q^{1/3} = \lambda^*.$$