VI Optimierung mit Restriktionen

Im letzten Kapitel haben wir uns damit beschäftigt, Extrema von Funktionen zu bestimmen, oder, anders ausgedrückt, Optimierungsprobleme ohne weitere Nebenbedingungen / Restriktionen zu lösen. In vielen ökonomischen Problemen müssen die Variablen aber gewisse Nebenbedingungen / Restriktionen erfüllen.

Beispiel: Ein Verbraucher möchte Mengen x und y für zwei Güter so bestimmen, dass die Nutzenfunktion $u(x,y)$ maximal wird. Dabei muss er seine Budgetbeschränkung $px+qy \leq m$ beachten.

Beispiel: Ein Unternehmen möchte seine Produktionskosten in einer bestimmten Zeitspanne möglichst gering halten. Auf Grund von Vorbestellungen muss aber eine Mindestmenge produziert werden.

Da die Suche nach dem Maximum einer Funktion f äquivalent zu der Suche nach dem Minimum von $-f$ ist, beschränken wir uns häufig auf die Betrachtung von Minimierungsproblemen.

Allgemein lassen sich Minimierungsprobleme mit Restriktionen mathematisch in folgender Form angeben.

$$\min f(x_1, x_2, \dots, x_n)$$

$$\text{so dass} \quad \left.\begin{array}{l} g_1(x_1,x_2,\dots,x_n)=0 \\ \quad \vdots \\ g_m(x_1,x_2,\dots,x_n)=0 \end{array}\right\} \text{Gleichungsrestriktionen}$$

$$\left.\begin{array}{l} h_1(x_1,x_2,\dots,x_n)\leq 0 \\ \quad \vdots \\ h_\ell(x_1,x_2,\dots,x_n)\leq 0 \end{array}\right\} \text{Ungleichungsrestriktionen}$$

Wir werden im Rahmen dieser Vorlesung nur Probleme mit Gleichungsrestriktionen behandeln, d.h.

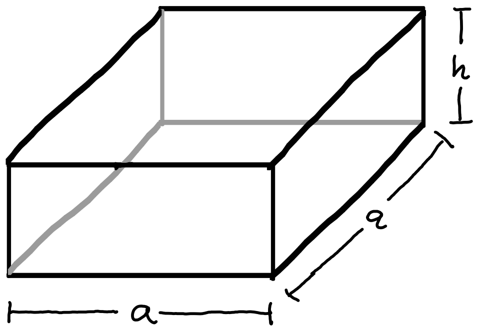
$$\min f(x_1, x_2, \ldots, x_n)$$
$$\text{so dass } g_1(x_1, x_2, \ldots, x_n) = 0$$
$$\vdots$$
$$g_m(x_1, x_2, \ldots, x_n) = 0 \quad (m < n)$$

und zwei grundlegende Methoden zur Lösung solcher Probleme kennen lernen: Verfahren der Variablensubstitution, Lagrangesches Multiplikatorverfahren.

Verfahren der Variablensubstitution

Wir erläutern die Methode zunächst an einem Beispiel.

Beispiel: Das Unternehmen Zack-und-Pack hat von der Bäckerei Knack-und-Back einen Auftrag zur Herstellung von Keksschachteln erhalten. Die oben offenen Schachteln sollen ein Fassungsvermögen von $V = 2048\,cm^3$ haben und eine quadratische Grundfläche besitzen. Ansonsten bleibt die Wahl der Abmessungen dem Hersteller überlassen.



$$V = a^2 \cdot h = 2048\,cm^3$$

Zack-und-Pack möchte die Materialkosten zur Herstellung der Schachteln minimieren. Zu lösen ist also das Problem

$$\min f(a,h) = a^2 + 4ah$$
$$\text{so dass } a^2 h = 2048$$

Da nach der Aufgabenstellung $a > 0$ sein muss, kann man die Restriktion nach h auflösen: $h = \frac{2048}{a^2}$

Gibt man den Wert der Variablen a vor, so ist h mit $h = \frac{2048}{a^2}$ eindeutig festgelegt.

Setzt man die Gleichung für h in die Funktion f ein, so muss man nur die Minimalstelle der Funktion

$$f\left(a, \frac{2048}{a^2}\right) = \tilde{f}(a) = a^2 + 4a \cdot \frac{2048}{a^2} = a^2 + \frac{8192}{a}$$

bestimmen.

Es gilt: $\tilde{f}'(a) = 2a - \dfrac{8192}{a^2}$

Somit: $\tilde{f}(a) = 0 \iff 2a = \dfrac{8192}{a^2}$

$\iff a^3 = 4096$

$\iff a = 16 \vee a = -16$

Nach Aufgabenstellung muss $a=-16$ nicht weiter betrachtet werden.

Weiter gilt: $\tilde{f}''(a) = 2 + \dfrac{16384}{a^3}$, d.h. $\tilde{f}''(a) > 0$ für $a > 0$.

Da $\mathbb{R}_+ = \{a \in \mathbb{R} : a > 0\}$ konvex ist, besitzt die konvexe Funktion $\tilde{f}$ somit an der Stelle $a = 16$ ein absolutes Maximum $\tilde{f}(16) = 768$.

Die optimale Keksschachtel hat also die Maße $a = 16\,cm$, $h = \frac{2048}{16^2}\,cm = 8\,cm$, bei einem Materialverbrauch von $\tilde{f}(16) = f(16, 8) = 768\,cm^3$.

Bei diesem Beispiel war es möglich, das Minimierungsproblem mit einer Nebenbedingung auf die Minimierung einer Funktion ohne Nebenbedingungen zurückzuführen. Verallgemeinert man diese Idee, so führt dies zu folgendem Verfahren.

Variablensubstitution

Wir gehen von folgendem Optimierungsproblem aus.

$$\min f(x_1, x_2, \ldots, x_n)$$

$$\text{so dass } \begin{aligned} g_1(x_1, x_2, \ldots, x_n) &= 0 \\ &\vdots \\ g_m(x_1, x_2, \ldots, x_n) &= 0 \end{aligned} \qquad m < n$$

lässt sich $g_1(x_1, x_2, \ldots, x_n) = 0$ nach einer der Variablen, z.B. nach x_1 auflösen zu $x_1 = z(x_2, \ldots, x_n)$, so kann man dies in die Funktion f und die restlichen Nebenbedingungen einsetzen und erhält das um eine Variable und eine Nebenbedingung reduzierte Problem

$$\min f(z(x_2,\dots,x_n),x_2,\dots,x_n)$$

so dass $g_2(z(x_2,\dots,x_n),x_2,\dots,x_n)=0$

$\vdots$

$g_m(z(x_2,\dots,x_n),x_2,\dots,x_n)=0$

Dieser Schritt wird m mal wiederholt, so dass schließlich die m Nebenbedingungen und m der n Variablen, z.B. $x_1,x_2,\dots,x_m$, eliminiert sind. Man erhält dann eine Funktion $\tilde{f}$, die nur noch von den restlichen Variablen, z.B. $x_{m+1},\dots,x_n$, abhängt und ohne Nebenbedingungen zu minimieren ist, d.h. die Aufgabe

$$\min \tilde{f}(x_{m+1},\dots,x_n)$$

Die Minimalstelle von $\tilde{f}$ liefert die Werte der Variablen $x^*_{m+1},\dots,x^*_n$ für die optimale Lösung des Problems. Die restlichen Werte für $x^*_1,\dots,x^*_m$ ergeben sich aus den Substitutionen.

Achtung: Dieses Verfahren kann nur verwendet werden, wenn man die Nebenbedingungen wie angegeben auflösen kann!

Beispiel: $\min f(x_1,x_2,x_3,x_4) = x_2^2+2x_3^2+x_2x_3-x_1+4x_2-x_3-2x_4+5$

so dass $g_1(x_1,x_2,x_3,x_4)=x_1-x_2+x_3-x_4+1=0$

$g_2(x_1,x_2,x_3,x_4)=2x_1-4x_2+2x_3+x_4-4=0$

Auflösen der 1. Restriktion nach x_1 liefert: $x_1=x_2-x_3+x_4-1$

Einsetzen in f und g_2 liefert:

$$\min f(x_2-x_3+x_4-1,x_2,x_3,x_4)=x_2^2+2x_3^2+x_2x_3-(x_2-x_3+x_4-1)+4x_2-x_3-2x_4+5$$
$$=x_2^2+2x_3^2+x_2x_3+3x_2-3x_4+6$$

so dass $g_2(x_2-x_3+x_4-1,x_2,x_3,x_4)=2(x_2-x_3+x_4-1)-4x_2+2x_3+x_4-4$

$$=-2x_2+3x_4-6=0$$

Das Problem hat nun eine Variable und eine Restriktion weniger als das Ausgangsproblem.

Auflösen der 2. Restriktion nach x_4 liefert: $x_4 = 2 + \frac{2}{3}x_2$

Einsetzen in f liefert:

$$\min \tilde{f}(x_2, x_3) = x_2^2 + 2x_3^2 + x_2x_3 + 3x_2 - 3\left(2 + \frac{2}{3}x_2\right) + 6$$
$$= x_2^2 + 2x_3^2 + x_2x_3 + x_2$$

Die so entstandene Optimierungsaufgabe hat nun nur noch zwei Variablen und keine Nebenbedingungen mehr.

Es gilt: $\quad \operatorname{grad} \tilde{f}(x_2, x_3) = \begin{pmatrix} 2x_2 + x_3 + 1 \\ 4x_3 + x_2 \end{pmatrix}$

Notwendig für ein Extremum ist:

$$\operatorname{grad} \tilde{f}(x_2, x_3) = \vec{0} \iff \left\{ \begin{matrix} 2x_2 + x_3 + 1 \\ 4x_3 + x_2 \end{matrix} \right\} \iff \left(x_2 = -\frac{4}{7},\ x_3 = \frac{1}{7} \right)$$

Einziger stationärer Punkt von $\tilde{f}$ ist somit $P(-\frac{4}{7}, \frac{1}{7})$.

Weiter gilt für die Hesse-Matrix von $\tilde{f}$: $\quad H_{\tilde{f}}(x_2, x_3) = \begin{pmatrix} 2 & 1 \\ 1 & 4 \end{pmatrix}$

Da $2 > 0$ und $\Delta_{\tilde{f}}(x_2, x_3) = 7 > 0$, folgt nach dem Hurwitz-Kriterium, dass $H_{\tilde{f}}(x_2, x_3)$ positiv definit, d.h. $\tilde{f}$ strikt konvex ist.

$\tilde{f}$ nimmt also auf der konvexen Menge $\mathbb{R}^2$ an der Stelle $x_2^* = -\frac{4}{7}, x_3^* = \frac{1}{7}$ ihr eindeutig bestimmtes Minimum $\tilde{f}(-\frac{4}{7}, \frac{1}{7}) = -\frac{2}{7}$ an.

Aus den obigen Substitutionsgleichungen erhält man $x_4^* = 2 + \frac{2}{3}x_2^* = \frac{34}{21}$ und $x_1^* = x_2^* - x_3^* + x_4^* - 1 = -\frac{2}{21}$.

Das Minimum von f unter den gegebenen Restriktionen ist somit

$$f\left(-\frac{2}{21}, -\frac{4}{7}, \frac{1}{7}, \frac{34}{21}\right) = \tilde{f}\left(-\frac{4}{7}, \frac{1}{7}\right) = -\frac{2}{7}.$$

An dem Beispiel ist bereits erkennbar, dass die Methode der Variablensubstitution bei vielen Variablen und Nebenbedingungen zu langwierigen Rechnungen führt. Außerdem ist die Voraussetzung für die Durchführbarkeit (Nebenbedingungen jeweils nach einer Variablen auflösbar), häufig nicht erfüllt. Wir behandeln daher ein allgemeineres Verfahren.

Lagrangesches Multiplikatorverfahren

Bevor wir dieses Verfahren allgemein angeben, wollen wir an einem Beispiel anschaulich erläutern, worauf dieses Verfahren beruht.

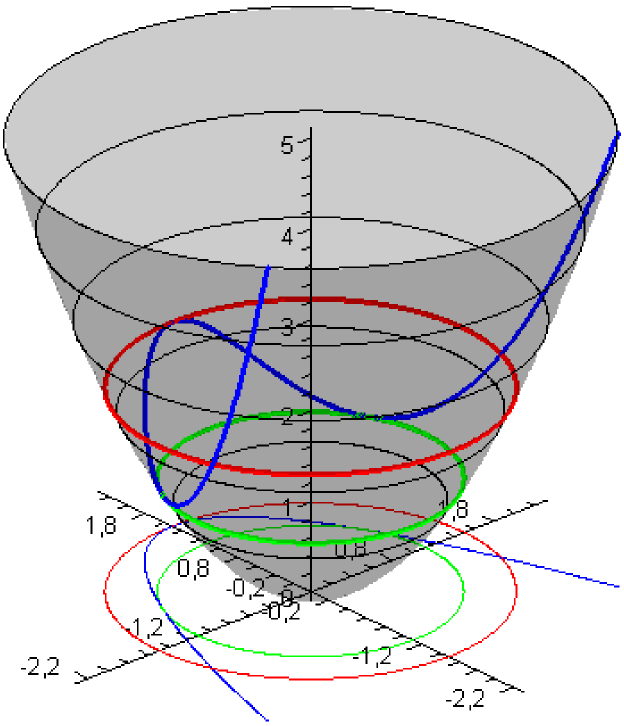
Beispiel: Wir betrachten das Problem

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

so dass $g(x,y) = y + x^2 - \frac{3}{2} = 0$

zunächst an Hand einer Graphik.

Die Funktion f beschreibt die grau schattierte Fläche im $\mathbb{R}^3$.

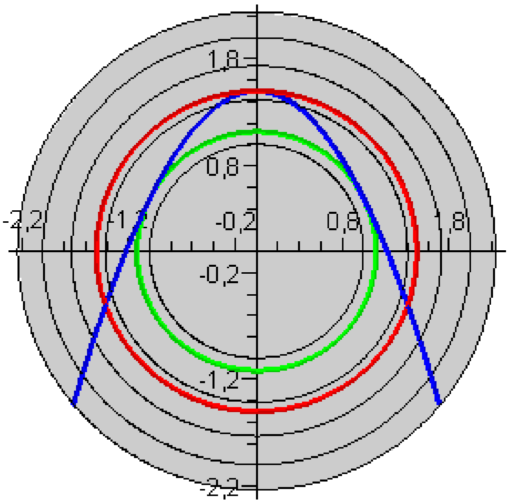


Extrema mit Restriktionenen

Zur Verdeutlichung sind einige Niveaulinien von f schwarz eingezeichnet. Durch $g(x,y)=0$ wird die in der xy-Ebene eingezeichnete blaue Parabel beschrieben. Schränkt man f auf diejenigen (x,y) ein, die die Restriktion erfüllen, so erhält man die auf der Fläche in blau hervorgehobene Kurve. Längs dieser Kurve sollen nun lokale Extrema bestimmt werden. Die lokalen Minima und das lokale Maximum sind in der Graphik deutlich zu erkennen. Die zugehörigen Niveaulinien sind auf der Fläche und in der xy-Ebene in grün bzw. rot dargestellt.

Man sieht, dass die blaue Kurve auf der Fläche die rote bzw. grüne Niveaulinie berührt.

Wir betrachten die Situation in der xy-Ebene noch einmal genauer.



Extrema mit Restriktionenen

Die Restriktionskurve und die zu den lokalen Extremalstellen gehörenden Niveaulinien berühren sich an den lokalen Extremalstellen. Anders ausgedrückt, haben diese Niveaulinien und die Restriktionskurve an den Extremalstellen jeweils dieselbe Steigung.

In Kapitel IV (vgl. Abschnitt Grenzrate der Substitution) haben wir gesehen, wie man die Steigungen der implizit gegebenen Kurven $f(x,y)=c$ (Niveaulinien) und $g(x,y)=0$ (Restriktion) berechnet.

Dies führt auf die Gleichung

$$-\frac{f_x(x,y)}{f_y(x,y)} = -\frac{g_x(x,y)}{g_y(x,y)}$$

(Steigung Niveaulinie) (Steigung Restriktion)

bzw. anders aufgeschrieben

$$\frac{f_x(x,y)}{g_x(x,y)} = \frac{f_y(x,y)}{g_y(x,y)}$$

Diese Gleichheit kann man nun auch so beschreiben, dass für ein $\lambda \in \mathbb{R}$ die beiden Gleichungen

$$\frac{f_x(x,y)}{g_x(x,y)} = \lambda \quad \text{und} \quad \frac{f_y(x,y)}{g_y(x,y)} = \lambda$$

bzw.

$$f_x(x,y) - \lambda g_x(x,y) = 0 \quad \text{und} \quad f_y(x,y) - \lambda g_y(x,y) = 0$$

erfüllt sein müssen. Zusammen mit der Restriktion haben wir für unser Beispiel die drei Gleichungen:

1) $f_x(x,y) - \lambda g_x(x,y) = 2x - 2\lambda x = 0 \iff x(1-\lambda) = 0$

2) $f_y(x,y) - \lambda g_y(x,y) = 2y - \lambda = 0 \iff \lambda = 2y$

3) $g(x,y) = y + x^2 - \frac{3}{2} = 0$

Wegen 1) $x(1-\lambda) = 0 \iff x = 0 \vee \lambda = 1$, unterscheiden wir zwei Fälle:

1. Fall: $\underline{x = 0}$

Einsetzen in 3) liefert $\underline{y = \frac{3}{2}}$

Einsetzen in 2) liefert $\underline{\lambda = 3}$

2. Fall: $\underline{\lambda = 1}$

Einsetzen in 2) liefert $\underline{y = \frac{1}{2}}$

Einsetzen in 3) liefert $\underline{x = 1 \vee x = -1}$

Die Punkte $P_1(0, \frac{3}{2})$, $P_2(1, \frac{1}{2})$, $P_3(-1, \frac{1}{2})$ kommen also als Kandidaten für Extremalstellen in Frage.

Es gilt: $f(0, \frac{3}{2}) = \frac{9}{4}$ ist lokales Maximum, $f(1, \frac{1}{2}) = f(-1, \frac{1}{2}) = \frac{5}{4}$ sind lokale Minima von f unter der gegebenen Restriktion.

Bevor wir die Methode, die uns notwendige Bedingungen für lokale Extrema bei Optimierungsproblemen mit Nebenbedingungen liefert, in allgemeiner Form angeben, schauen wir die obigen Bedingungen noch einmal formal aus einem anderen Blickwinkel an.

Zu dem Problem aus unserem Beispiel definieren wir eine Hilfsfunktion (Lagrangefunktion) durch

$$L(x,y;\lambda) = f(x,y) - \lambda g(x,y).$$

Die Forderung, dass $\operatorname{grad} L(x,y;\lambda) = \vec{0}$ führt auf

$$L_x(x,y;\lambda) = f_x(x,y) - \lambda g_x(x,y) = 0$$

$$L_y(x,y;\lambda) = f_y(x,y) - \lambda g_y(x,y) = 0$$

$$L_\lambda(x,y;\lambda) = g(x,y) = 0$$

Dies sind genau dieselben Gleichungen, die wir aus der zunächst anschaulichen Argumentation hergeleitet haben.
Mit Hilfe der Lagrangefunktion lässt sich nun die Lagrangesche Multiplikatormethode sehr allgemein formulieren.

Langrangesches Multiplikatorverfahren

Es seien $f, g_1, g_2, \dots, g_m : \mathbb{R}^n \longrightarrow \mathbb{R}$, $m < n$, stetig partiell differenzierbare Funktionen in den Variablen $x_1, x_2, \dots, x_n$. $(x_1^*, x_2^*, \dots, x_n^*)$ sei ein innerer Punkt von $\mathbb{D}_f \cap \mathbb{D}_{g_1} \cap \dots \cap \mathbb{D}_{g_m}$. Der Rang der sogenannten Jacobi-Matrix (das ist die Matrix der ersten partiellen Ableitungen von $g_1, \dots, g_m$)

$$J(x_1^*, \dots, x_n^*) = \begin{pmatrix} \frac{\partial g_1}{\partial x_1}(x_1^*, \dots, x_n^*) & \cdots & \frac{\partial g_1}{\partial x_n}(x_1^*, \dots, x_n^*) \\ & & \\ \frac{\partial g_m}{\partial x_1}(x_1^*, \dots, x_n^*) & \cdots & \frac{\partial g_m}{\partial x_n}(x_1^*, \dots, x_n^*) \end{pmatrix}$$

sei gleich m. Eine notwendige Bedingung dafür, dass f an der Stelle $(x_1^*, \dots, x_n^*)$ eine lokale Extremalstelle unter Berücksichtigung der Restriktionen $g_1(x_1^*, \dots, x_n^*) = 0, \dots, g_m(x_1^*, \dots, x_n^*) = 0$, besitzt, ist, dass der Gradient der Lagrangefunktion

$$L(x_1, \dots, x_n; \lambda_1, \dots, \lambda_m) = f(x_1, \dots, x_n) - \sum_{i=1}^{m} \lambda_i g_i(x_1, \dots, x_n)$$

verschwindet. D.h. es müssen reelle Zahlen $\lambda_1, \dots, \lambda_m$ (die sogenannten Lagrange-Multiplikatoren) existieren, so dass

$$\begin{array}{|ll} L_{x_j}(x_1, \dots, x_n; \lambda_1, \dots, \lambda_m) = 0 & \text{für alle } j = 1, \dots, n \\ L_{\lambda_i}(x_1, \dots, x_n; \lambda_1, \dots, \lambda_m) = 0 & \text{für alle } i = 1, \dots, m. \end{array}$$

Bemerkung:

1) Die Bedingungen $L_{\lambda_i}(x_1, \dots, x_n; \lambda_1, \dots, \lambda_m) = 0$, $i = 1, \dots, m$, sind genau die Nebenbedingungen des Optimierungsproblems.
2) Die Bedingung, dass die Jacobi-Matrix maximalen Rang besitzt, garantiert, dass die Nebenbedingungen $g_i(x_1, \dots, x_m) = 0$, $i = 1, \dots, m$, lokal

nach m der n Variablen auflösbar sind. Dabei bedeutet lokal, dass die Auflösbarkeit gewährleistet ist, wenn $(x_1, ..., x_n)$ in einer Umgebung von $(x_1^*, ..., x_n^*)$ liegt.

Beispiel: Wir lösen die Keksschachtelaufgabe noch einmal mit Hilfe der Lagrangeschen Multiplikatormethode.

$$\min f(a,h) = a^2 + 4ah$$

so dass $g(a,h) = a^2h - 2048 = 0$

Die Lagrangefunktion ist

$$L(a,h;\lambda) = f(a,h) - \lambda g(a,h) = a^2 + 4ah - \lambda(a^2h - 2048).$$

Die partiellen Ableitungen sind:

$$L_a(a,h;\lambda) = 2a + 4h - 2\lambda ah = 2(a + 2h - \lambda ah)$$

$$L_h(a,h;\lambda) = 4a - \lambda a^2$$

$$L_\lambda(a,h;\lambda) = -(a^2h - 2048)$$

Die notwendigen Bedingungen lauten also:

1) $a + 2h - \lambda ah = 0$

2) $4a - \lambda a^2 = 0$

3) $a^2h = 2048$

Aus 3) folgt zunächst: $a \neq 0 \wedge h \neq 0$.

Division von 2) durch $a \neq 0$ liefert: 2)' $\lambda a = 4$

Einsetzen von $\lambda a = 4$ in 1) liefert: 1)' $a + 2h - 4h = 0 \iff a = 2h$

Einsetzen von $a = 2h$ in 3) liefert: 3)' $4h^3 = 2048 \iff \underline{h^* = 8}$

Mit 1)' und 2)' ergibt sich daraus: $\underline{a^* = 16}$, $\underline{\lambda^* = \frac{1}{4}}$.

Als Minimalstelle kommt also nur $(a^*, h^*) = (16, 8)$ mit $\lambda^* = \frac{1}{4}$ in Frage.

Für die Jacobi-Matrix gilt: $J(a,h) = (2ah, a^2)$, d.h. $J(a^*,h^*) = 16^2(1, 1)$.

Die Rangbedingung $Rg(J(a^*,h^*)) = 1$ ist also erfüllt.