Extrema mit Nebenbedingungen

Kajetan Bentele

1 Einführung

Die Notwendigkeit, Extrema (Minima oder Maxima) einer Funktion zu bestimmen ergibt sich in einer Vielzahl von wissenschaftlichen Anwendungen. Oft sollen jedoch auch gleichzeitig bestimmte Relationen zwischen den Variablen erfüllt sein, sodass die Variablen nicht mehr voneinander unabhängig sind.

Es gilt also, die Extrema der Funktion

$$f(x_1, x_2, \ldots, x_m, x_{m+1}, \ldots, x_{m+n}) \tag{1}$$

der $m+n$ Veränderlichen x_i zu bestimmen, zwischen denen die n Beziehungen

$$\phi_j(x_1, x_2, \ldots, x_m, x_{m+1}, \ldots, x_{m+n}) = 0, \quad j = 1, 2, \ldots, n \tag{2}$$

bestehen. Diese Aufgabe kann im Prinzip auch direkt gelöst werden, indem man die n Relationen (2) beispielsweise nach den Variablen $x_{m+1}, x_{m+2}, \ldots, x_{m+n}$ auflöst[1] und in die Funktion f einsetzt. Wir erhalten dann eine Funktion, die nur noch von m Variablen $x_1, x_2, \ldots, x_m$ abhängt,

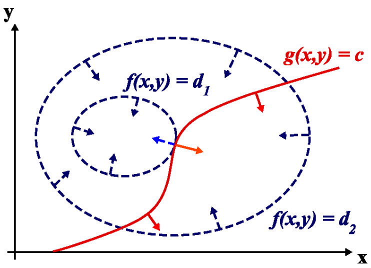
$$\begin{aligned}\tilde{f}(x_1, x_2, \ldots, x_m) = &\\ & f(x_1, x_2, \ldots, x_m, x_{m+1}(x_1, x_2, \ldots, x_m), \ldots, x_{m+n}(x_1, x_2, \ldots, x_m)),\end{aligned} \tag{3}$$

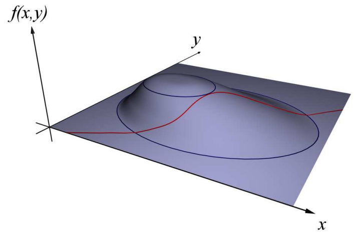
und deren Extrema dann zu bestimmen sind. Man hat damit die gestellte Aufgabe auf die Ermittlung von Maxima und Minima bei unabhängigen Variablen zurückgeführt. Allerdings ist dieses Verfahren mühsam und oft sogar unausführbar. Wir führen daher das Verfahren der Lagrangeschen Multiplikatoren ein, das dieses Problem erheblich einfacher und effizienter löst.

2 Veranschaulichung

Wir wollen uns das Verfahren zunächst im Fall einer Funktion zweier Veränderlicher, $f(x, y)$ und der Nebenbedingung $g(x, y) = c$ anschaulich klar machen. In der obenste-

[1] Im konkreten Fall hängt dies davon ab, nach welchen Variablen sich die Ausdrücke am einfachsten auflösen lassen.





Linke Abbildung zeigt die Höhenlinien der Funktion $f(x, y)$ in blau, die Punkte, welche die Nebenbedingung $g(x, y) = c$ erfüllen, als rote Linie. Die Gradienten der beiden Funktionen sind als Pfeile dargestellt. Die rechte Abbildung zeigt das gleiche Problem, die Funktionswerte $f(x, y)$ auf der Höhenachse abgetragen.

henden Abbildung sind die Höhenlinien der Funktion f blau dargestellt, die Menge aller Punkte, welche die gegebene Nebenbedingung erfüllen, als rote Linie.

Beim Verfolgen der Höhenlinie $g(x, y) = c$ berühren oder kreuzen wir Höhenlinien von f. Ein gemeinsamer Punkt (x, y) der Nebenbedingung $g(x, y) = c$ und einer Höhenlinie $f(x, y) = d$ kann nur dann Lösung des Optimierungsproblems sein, wenn unsere Bewegung auf der Höhenlinie $g(x, y) = c$ tangential zu $f(x, y) = d$ verläuft. Andernfalls könnten wir durch Vorwärts- oder Rückwärtsbewegung auf der vorgegebenen g-Höhenlinie den Funktionswert von f vergrößern oder verkleinern, ohne die Nebenbedingung zu verletzen. Da der Gradient einer Funktion senkrecht auf der Höhenlinie steht, zeigen die Gradienten zweier Funktionen, deren Höhenlinien tangential zueinander verlaufen entweder in die gleiche oder genau entgegengesetze Richtungen.

Es gilt daher im Punkt (x, y), der die Funktion f bei Erfüllung der Nebenbedingung $g(x, y) = c$ maximiert, die Beziehung

$$\nabla f + \lambda \nabla g = 0, \tag{4}$$

für einen bestimmten Wert von λ. Wir haben somit im vorliegenden Fall drei Gleichungen (die zwei Gleichungen, die sich aus (4) sowie die Nebenbedingung $g(x, y) = c$) für drei Unbekannte (x, y und λ).

3 Allgemeine Herleitung

Wir betrachten jetzt wieder den allgemeine Fall $n + m$ Veränderlicher, die n Nebenbedingungen erfüllen müssen. Diese Nebenbedingungen führen dazu, dass n der $n + m$ Variablen nicht mehr unabhängig gewählt werden können. Aus diesem Grund können wir, wie in der Einführung, beispielsweise die Veränderlichen x_i mit $i > m$, als Funktion der verbleibenden m unabhängigen Variablen auffassen

$$x_{m+1}(x_1, x_2, \ldots, x_m), \ldots, x_{m+n}(x_1, x_2, \ldots, x_m). \tag{5}$$

Dabei spielt es für die folgende Betrachtung keine Rolle, ob diese Abhängigkeiten explizit oder nur implizit durch (2) gegeben sind. Mittels der Relationen (5) erhalten wir die

Funktion $\tilde{f}$,

$$\tilde{f}(x_1, x_2, \ldots, x_m) = \\ f(x_1, x_2, \ldots, x_m, x_{m+1}(x_1, x_2, \ldots, x_m), \ldots, x_{m+n}(x_1, x_2, \ldots, x_m)), \quad (6)$$

wie bereits in der Einführung dargestellt. Diese Funktion nimmt ein Extremalwert an, wenn

$$\sum_{i=1}^{m} \frac{\partial \tilde{f}}{\partial x_i} dx_i = 0 \quad (7)$$

für unabhängige Differentiale dx_i erfüllt ist. Das bedeutet aber, dass jeder der Summanden für sich alleine verschwinden muss. Diese m Gleichungen legen somit die verbleibenden x_i, $i \leq m$ fest. Diesen Ausdruck können wir mit Hilfe der Beziehungen (5) folgendermaßen umschreiben,

$$\begin{aligned}
\sum_{i=1}^{m} \frac{\partial \tilde{f}}{\partial x_i} dx_i &= \sum_{i=1}^{m} \left(\frac{\partial f}{\partial x_i} + \sum_{s=1}^{n} \frac{\partial f}{\partial x_{m+s}} \frac{\partial x_{m+s}}{\partial x_i} \right) dx_i \\
&= \sum_{i=1}^{m} \frac{\partial f}{\partial x_i} dx_i + \sum_{s=1}^{n} \frac{\partial f}{\partial x_{m+s}} \underbrace{\sum_{i=1}^{m} \frac{\partial x_{m+s}}{\partial x_i} dx_i}_{dx_{m+s}} \\
&= \sum_{i=1}^{m+n} \frac{\partial f}{\partial x_i} dx_i = 0, \quad (8)
\end{aligned}$$

wobei jetzt nicht mehr alle dx_i unabhängige Differentiale sind, d.h. nicht jeder der Summanden ist identisch Null.

Wenn wir die Beziehungen (5) in die Nebenbedingungen (2) einsetzen, so verschwinden die $\tilde{\phi}_j$ für alle x_i mit $i \leq m$ per definitionem:

$$\tilde{\phi}_j(x_1, x_2, \ldots, x_m) = \\ \phi_j(x_1, x_2, \ldots, x_m, x_{m+1}(x_1, x_2, \ldots, x_m), \ldots, x_{m+n}(x_1, x_2, \ldots, x_m)) \equiv 0 \\ j = 1, 2, \ldots, n. \quad (9)$$

Es gilt also insbesonders auch

$$\sum_{i=1}^{m} \frac{\partial \tilde{\phi}_j}{\partial x_i} dx_i = 0, \quad j = 1, 2, \ldots, n \quad (10)$$

für unabhängige dx_i. Genauso wie vorher können wir Gleichungen der Form

$$\sum_{i=1}^{m+n} \frac{\partial \phi_j}{\partial x_i} dx_i = 0, \quad j = 1, 2, \ldots, n \quad (11)$$

ableiten, wobei wiederum nicht alle Differentiale unabhängig sind, d.h. nicht jeder der Summanden verschwindet notwendigerweise für sich alleine.

Wenn wir die Gleichungen (8) und (11) mittels Lagrangescher Multiplikatoren λ_j kombinieren, erhalten wir

$$\sum_{i=1}^{m+n} \frac{\partial f}{\partial x_i} dx_i + \lambda_1 \sum_{i=1}^{m+n} \frac{\partial \phi_1}{\partial x_i} dx_i + \cdots + \lambda_n \sum_{i=1}^{m+n} \frac{\partial \phi_n}{\partial x_i} dx_i = \sum_{i=1}^{m+n} \left(\frac{\partial f}{\partial x_i} + \lambda_1 \frac{\partial \phi_1}{\partial x_i} + \cdots + \lambda_n \frac{\partial \phi_n}{\partial x_i} \right) dx_i = 0. \quad (12)$$

Durch die Einführung der Lagrangeschen Multiplikatoren haben wir n weitere Freiheitsgrade zu den m unabhängigen Variablen x_i gewonnen. Die Lagrangeschen Multiplikatoren werden jetzt so bestimmt, dass

$$\frac{\partial f}{\partial x_i} + \lambda_1 \frac{\partial \phi_1}{\partial x_i} + \cdots + \lambda_n \frac{\partial \phi_n}{\partial x_i} = 0, \quad i > m \quad (13)$$

gilt, d.h. dass auch für abhängige dx_i jeder Summand verschwindet. Die Gleichung (12) nimmt somit die Form

$$\sum_{i=1}^{m} \left(\frac{\partial f}{\partial x_i} + \lambda_1 \frac{\partial \phi_1}{\partial x_i} + \cdots + \lambda_n \frac{\partial \phi_n}{\partial x_i} \right) dx_i = 0 \quad (14)$$

an. Da die dx_i für $i \leq m$ unabhängige Differentiale sind, muss jeder der Summanden alleine verschwinden. Diese $n+m$ Gleichungen legen daher die m unabhängigen Variablen x_i, $i \leq m$ und die n Lagrangeschen Multiplikatoren λ_j fest.

Mit den Nebenbedingungen (2) erhalten wir also insgesamt das folgende Gleichungssystem

$$\frac{\partial f}{\partial x_i} + \lambda_1 \frac{\partial \phi_1}{\partial x_i} + \cdots + \lambda_n \frac{\partial \phi_n}{\partial x_i} = 0, \quad i = 1, 2, \ldots, m+n \quad (15)$$

$$\phi_j(x_1, x_2, \ldots, x_m, x_{m+1}, \ldots, x_{m+n}) = 0, \quad j = 1, 2, \ldots, n. \quad (16)$$

Wir können daher den Spieß auch umdrehen und die $m+n$ Gleichungen (15) als Bestimmungsgleichungen für die $m+n$ Variablen x_i, $i \leq n+m$ auffassen und die n Lagrange Multiplikatoren λ_j, falls wir diese benötigen, aus den n Nebenbedingungen (16) bestimmen. Man sollte jedoch beachten, dass die von uns entwickelten Gleichungen lediglich notwendige und keine hinreichenden Bedingungen für das Vorliegen eines Extremum darstellen.

4 Literatur

Der Abschnitt 2 ist weitgehend aus der deutschen Wikipedia übernommen. Ansonsten ist das „Lehrbuch der höheren Mathematik“ von Smirnow, sowie „Mathematical Methods for Physicists“ von Arfken und Weber empfehlenswert. Gerade im letzteren steht die Anwendbarkeit von Konzepten der höheren Mathematik für die Natur- und Ingenieurswissenschaften im Vordergrund. Wer an einem mathematisch „sauberen“ Beweis der Methode der Lagrangeschen Multiplikatoren interessiert ist, sollte das Buch „Lehrbuch der Analysis II“ von Harro Heuer zur Hand nehmen.