Lagrange-Multiplikatoren

Ist x_* eine lokale Extremstelle einer Funktion $f : \mathbb{R}^n \ni D \to \mathbb{R}$ auf der durch die Nebenbedingungen $g_k(x_1, \ldots, x_n) = 0$ definierten Menge D, dann existieren Lagrange-Multiplikatoren $\lambda_k \in \mathbb{R}$, so dass

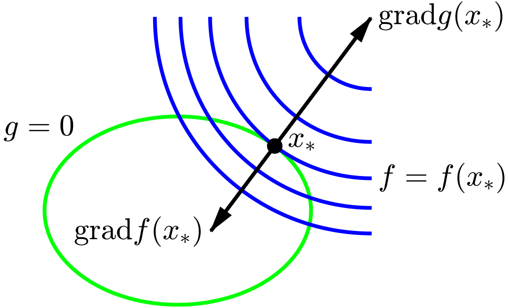
$$\operatorname{grad} f(x_*) = \sum_k \lambda_k \operatorname{grad} g_k(x_*)\,.$$

Dabei wird vorausgesetzt, dass f und die Funktionen g_k in einer Umgebung von x_* stetig differenzierbar sind und dass die Gradienten $\operatorname{grad} g_k(x_*)$ linear unabhängig sind.

Bei nur einer Nebenbedingung $g(x_1, \ldots, x_n) = 0$ hat die Lagrange-Bedingung die einfache Form

$$\operatorname{grad} f(x_*) \parallel \operatorname{grad} g(x_*)\,,$$

falls $\operatorname{grad} g(x_*) \neq (0, \ldots, 0)^{\mathrm{t}}$, d.h. die Niveauflächen von f und g berühren sich an einer Extremstelle x_*. Dies ist in der Abbildung für bivariate Funktionen veranschaulicht.



Die Lagrange-Bedingung ist nicht hinreichend, um zu entscheiden, ob ein lokales Extremum vorliegt und ob es sich um ein Minimum oder ein Maximum handelt. Dies lässt sich nur mit Hilfe weiterer Informationen feststellen.

Die globalen Extrema erhält man durch den Vergleich der Funktionswerte an den Punkten, welche die Lagrange-Bedingung erfüllen, sowie gegebenenfalls den Randpunkten des Definitionsbereichs D oder Punkten x, an denen die Gradienten grad $g_k(x)$ linear abhängig sind.

Beweis

n: Anzahl der Variablen, m: Anzahl der Nebenbedingungen g_k

(i) $m \geq n$:

Der n-Vektor $\operatorname{grad} f(x_*)$ ist immer als Linearkombination der nach Voraussetzung linear unabhängigen Gradienten $\operatorname{grad} g_k(x_*)$ darstellbar. ✓

Grund: Für $m \geq n$, besteht die zulässige Menge im Allgemeinen bereits aus diskreten Punkten, die durch die Nebenbedingungen festgelegt sind.

(ii) $m < n$:

fasse die Nebenbedingungen g_k zu einer Funktion $g = (g_1, \ldots, g_m)^{\mathrm{t}}$ zusammen

partitioniere die Variablen als $x = (u, v) \in \mathbb{R}^m \times \mathbb{R}^{n-m}$, wobei nach eventueller Permutation die Invertierbarkeit der Jacobi-Matrix $(\partial g(u, v)/\partial u)_{|(u_*, v_*)} = g_u(u_*, v_*)$ vorausgesetzt wird

Satz über implizite Funktionen $\implies$ lokale Auflösbarkeit der Nebenbedingungen

$$g(u, v) = (0, \ldots, 0)^{\mathrm{t}} \quad \Longleftrightarrow \quad u = \varphi(v), \quad (u, v) \approx (u_*, v_*)$$

Gradient der Funktion $v \mapsto h(v) = f(\varphi(v), v)$ Null an einem Extremum, d.h.

$$\operatorname{grad} h(v_*)^{\mathrm{t}} = f_u(u_*, v_*)\varphi'(v_*) + f_v(u_*, v_*) = (0, \ldots, 0)^{\mathrm{t}}$$

aufgrund der Kettenregel und mit φ' der Jacobi-Matrix von φ
Differenzieren der Nebenbedingungen $g(\varphi(v), v) = (0, \ldots, 0)^{\mathrm{t}} \quad \Longrightarrow$ $g_u(\varphi(v), v)\varphi'(v) + g_v(\varphi(v), v) = (0, \ldots, 0)^{\mathrm{t}}$, d.h.

$$\varphi'(v) = -g_u(u, v)^{-1} g_v(u, v)$$

Setzen von $(\lambda_1, \ldots, \lambda_m) = f_u(u_*, v_*) g_u(u_*, v_*)^{-1}$ und Einsetzen des Ausdrucks für φ' in den Gradienten von $h \quad \rightsquigarrow$

$$f_u = \lambda g_u, \quad f_v = -f_u(-g_u^{-1} g_v) = \lambda g_v$$

(u- und v-Komponenten der Lagrange-Bedingung $f' = \lambda g'$ im Punkt (u_*, v_*))

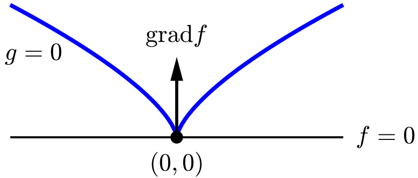
Beispiel

Minimierung von

$$f(x, y) = y$$

unter der Nebenbedingung

$$g(x, y) = y^3 - x^2 = 0$$



$\rightsquigarrow$ Minimum bei $(x_*, y_*) = (0, 0)$

Die Lagrange-Bedingung $(f_x(0,0), f_y(0,0)) = \lambda(g_x(0,0), g_y(0,0))$ ist nicht erfüllt:

$$(0, 1) \neq (0, 0) = \lambda(-2x_*, 3y_*^2)$$

Grund: $\operatorname{grad} g(x_*, y_*) = (0, 0)^{\mathrm{t}}$
Die Lagrange-Bedingung ist in singulären Punkten (kein maximaler Rang der Jacobi-Matrix g' der Nebenbedingungen) nicht anwendbar.

Beispiel

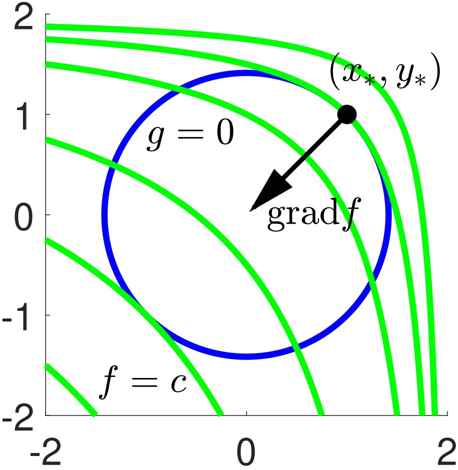
Minimierung von $f(x, y) = (x-2)(y-2)$ unter der Nebenbedingung $g(x, y) = x^2 + y^2 - 2 = 0$

Lagrange-Bedingung

$$\underbrace{(f_x, f_y)}_{\text{grad } f} = (y-2, x-2)$$

$$\stackrel{!}{=} \lambda(2x, 2y) = \lambda \underbrace{(g_x, g_y)}_{\text{grad } g}$$

Die Niveaulinien von f im Punkt (x_*, y_*) sind tangential zu der durch die Nebenbedingung $g = 0$ definierten Kurve.



Elimination von λ durch Bilden der Differenz $yf_x - xf_y$ in der Lagrange-Bedingung $\quad \rightsquigarrow$

$$y(y-2) - x(x-2) = 0 \quad \Longleftrightarrow \quad (y-x)(y+x-2) = 0$$

zwei Fälle: (i) $x = y$ und (ii) $y + x - 2 = 0$
Berücksichtigung der Nebenbedingung $x^2 + y^2 - 2 = 0 \quad \rightsquigarrow$
$(x, y) = (1, 1)$ oder $(-1, -1)$ im Fall (i)
und
$x^2 + (2-x)^2 - 2 = 2(x-1)^2 = 0$, d.h. ebenfalls $(x, y) = (1, 1)$ im Fall (ii)
Existenz von Minimum und Maximum für eine stetige Funktion auf einer kompakten Menge (Kreis mit Radius $\sqrt{2}$) und Vergleich der Funkionswerte

$$f(1,1) = 1, \quad f(-1,-1) = 9$$

$\Longrightarrow \quad f$ bei $(1, 1)$ minimal

Beispiel

Bestimmung der Extrema der Funktion

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

unter den Nebenbedingungen

$$g_1(x, y, z) = x^2 + y^2 - 8 = 0, \quad g_2(x, y, z) = x + z - 4 = 0$$

(Ellipse: Schnitt eines Zylinders mit einer Ebene)

Jacobi-Matrix der Nebenbedingungen

$$g'(x, y) = \begin{pmatrix} \operatorname{grad} g_1^{\mathrm{t}} \\ \operatorname{grad} g_2^{\mathrm{t}} \end{pmatrix} = \begin{pmatrix} 2x & 2y & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

voller Rang ($\Longleftrightarrow$ lineare Unabhängigkeit der Gradienten) für $(x, y) \neq (0, 0)$; auf zulässiger Menge erfüllt

Lagrange-Bedingung für Extremstellen (x, y, z)

$$\underbrace{(1, 2, -1)}_{\text{grad } f} = (\lambda_1, \lambda_2) \begin{pmatrix} 2x & 2y & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

bzw.

$$1 = 2\lambda_1 x + \lambda_2, \quad 2 = 2\lambda_1 y, \quad -1 = \lambda_2$$

Einsetzen von $\lambda_1 = 1/y$ und $\lambda_2 = -1$ in die erste Gleichung $\quad \rightsquigarrow \quad x = y$
Nebenbedingungen $\quad \rightsquigarrow \quad$ mögliche Extrema $(2, 2, 2)$ und $(-2, -2, 6)$

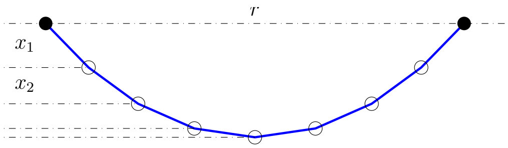
Existenz von Minimum und Maximum auf der Ellipse und Vergleich der Funktionswerte,

$$f(-2, -2, 6) = -12 < 4 = f(2, 2, 2)\,,$$

$\implies \quad f$ ist minimal bei $(-2, -2, 6)$ und maximal bei $(2, 2, 2)$.

Beispiel

Gleichgewichtslage einer an zwei Punkten aufgehängten Kette mit $2n$ Kettengliedern der Länge 1



potentielle Energie unter Berücksichtigung der Symmetrie

$$\begin{aligned} f(x) &= -2\left(\frac{x_1}{2}\right) - 2\left(x_1 + \frac{x_2}{2}\right) - \cdots - 2\left(x_1 + \cdots + x_{n-1} + \frac{x_n}{2}\right) \\ &= -a_1x_1 - \cdots - a_nx_n \end{aligned}$$

mit $a_k = 2(n-k)+1$

Länge der Kette $\quad\rightsquigarrow\quad$ Nebenbedingung

$$g(x) = r/2 - \sum_{k=1}^{n} \sqrt{1 - x_k^2} = 0$$

$\rightsquigarrow$ Optimierungsproblem

$$f \to \min, \quad g = 0$$

Lagrange-Bedingungen $\operatorname{grad} f = \lambda \operatorname{grad} g$

$$-a_k = \lambda \frac{x_k}{\sqrt{1 - x_k^2}}, \quad k = 1, \ldots, n$$

Quadrieren und Auflösen nach $x_k \quad \Longrightarrow$

$$a_k^2(1 - x_k^2) = \lambda^2 x_k^2, \quad x_k^2 = \frac{a_k^2}{a_k^2 + \lambda^2}$$

Einsetzen in die Nebenbedingung $r/2 = \sum_k \sqrt{1 - x_k^2} \quad \rightsquigarrow$

$$\frac{r}{2} = \sum_{k=1}^{n} \sqrt{\frac{\lambda^2}{a_k^2 + \lambda^2}}$$

$\sqrt{\ldots}$ monotone Funktion von λ

$\rightsquigarrow$ einfach zu berechnende numerische Lösung λ_*

$\rightsquigarrow$ Bestimmung von x_k aus den Lagrange-Bedingungen