Universität Konstanz
FB Mathematik & Statistik

Prof. Dr. M. Junk
H. Kungl
J. Simons

Mathematik einüben in Gruppenarbeit
Lösungsvorschlag zu Aufgabenblatt 20

Aufgabe 1: Explizite Sorgfalt beim impliziten Ableiten

Gesucht ist das Maximum einer Funktion $f : \mathbb{R}^n \to \mathbb{R}$ unter gewissen Nebenbedingungen die durch $m < n$ Funktionen $g_i : \mathbb{R}^n \to \mathbb{R}$ beschrieben sind. Die zulässigen Punkte sind definiert durch die Bedingungen

$$g(x) = (g_1(x), \dots, g_m(x)) = (\bar{c}_1, \dots, \bar{c}_m) = \bar{c}$$

wobei $\bar{c} \in \mathbb{R}^m$ ein vorgegebener Vektor ist. Nehmen wir an, dass für die Vektoren c in einer Umgebung von $\bar{c}$ das Maximierungsproblem unter den Nebenbedingungen $g(x) = c$ jeweils genau eine Lösung hat, die wir mit $X(c)$ bezeichnen und dass X in $\bar{c}$ differenzierbar ist. Berechnen Sie die Ableitungen des maximalen Funktionswertes $M(c) = f(X(c))$ an der Stelle $\bar{c}$. Welche Bedeutung hat $M'(\bar{c})h$ für $h \in \mathbb{R}^m$? Beachten Sie bei Ihrer sorgfältigen Rechnung, dass nach dem Satz über die Optimierung unter Nebenbedingungen ein Vektor $\Lambda(\bar{c})$ existiert, so dass $F(x, \lambda) = f(x) - \lambda \cdot (g(x) - c)$ einen kritischen Punkt an der Stelle $(X(\bar{c}), \Lambda(\bar{c}))$ hat.

Zunächst fügen wir die Voraussetzung hinzu, dass f eine stetig differenzierbare Funktion ist. Sonst können wir nicht ohne weiteres die Ableitung von M bestimmen. Dann folgt mit der Kettenregel $M'(\bar{c}) = f'(X(\bar{c}))X'(\bar{c})$. Außerdem wissen wir, dass $F'(X(\bar{c}), \Lambda(\bar{c})) = 0$ ist, also insbesondere $F'(X(\bar{c}), \Lambda(\bar{c}))(k, 0) = 0$ für alle $k \in \mathbb{R}^n$. Mit anderen Worten

$$0 = k_1 \partial_1 F(X(\bar{c}), \Lambda(\bar{c})) + \dots + k_n \partial_n F(X(\bar{c}), \Lambda(\bar{c})) = f'(X(\bar{c}))k - \Lambda(\bar{c}) \cdot g'(X(\bar{c}))k$$

Wählen wir speziell $k = X'(\bar{c})$, so folgt

$$M'(\bar{c}) = f'(X(\bar{c}))X'(\bar{c}) = \Lambda(\bar{c}) \cdot g'(X(\bar{c}))X'(\bar{c}).$$

Nun ist aber $X(c)$ jeweils ein zulässiger Punkt zu den Nebenbedingungen $g(x) = c$, d.h. $g(X(c)) = c$ für alle c in der angegebenen Umgebung von $\bar{c}$. Leiten wir implizit ab (was die Glattheit der Nebenbedingungsfunktionen g voraussetzt und deshalb zu den Voraussetzungen dazugefügt werden muss), so erhalten wir $g'(X(\bar{c}))X'(\bar{c}) = I$, wobei I die Identitätsabbildung auf $\mathbb{R}^m$ ist (sorgfältig sein auf der rechten Seite: die Funktion $G(c) = c$, die in der Umgebung von $\bar{c}$ differenzierbar ist hat die Ableitung $G'(\bar{c})h = h_1 \partial_1 G(\bar{c}) + \dots + h_m \partial_m G(\bar{c}) = h_1 e_1 + \dots + h_m e_m = Ih$). Also gilt

$$M'(\bar{c})h = \Lambda(\bar{c}) \cdot g'(X(\bar{c}))X'(\bar{c})h = \Lambda(\bar{c}) \cdot (Ih) = \Lambda(\bar{c}) \cdot h = h_1 \Lambda_1(\bar{c}) + \dots + h_m \Lambda_m(\bar{c}).$$

Folglich ist $\partial_i M(\bar{c}) = \Lambda_i(\bar{c})$, d.h. der i-te Lagrange-Multiplikator $\Lambda_i(\bar{c})$ beschreibt gerade, wie stark sich der Maximalwert von f ändert, relativ zur Änderung von c_i in der i-ten Nebenbedingung $g_i(x) = c_i$. Genauer gesagt, ist $M(\bar{c})$ der Maximalwert unter den Nebenbedingungen $g(x) = c$, so ist

$$M(\bar{c} + \Delta c) \approx M(\bar{c}) + M'(\bar{c})\Delta c = M(\bar{c}) + \Lambda(\bar{c}) \cdot \Delta c.$$

Denken Sie in einem zweiten Schritt darüber nach, ob die Forderung der Differenzierbarkeit nötig ist. Muss $c \mapsto X(c)$ nicht automatisch glatt sein? Hinweis: Finden Sie, eine geeignete Bedingung $\Phi(x, \lambda, c) = 0$, die (unter anzugebenden Voraussetzungen) lokal eindeutig in der Form $\Phi(X(c), \Lambda(c), c) = 0$ aufgelöst werden kann. Der Satz über implizite Funktionen liefert dann die Glattheit.

Nach dem Satz über Extremwerte unter Nebenbedingungen gibt es ein $\bar{x} \in \mathbb{R}^n$ und ein $\bar{\lambda} \in \mathbb{R}^m$, so dass

$$\begin{aligned}0 &= \Phi((\bar{x}_1, \dots, \bar{x}_n, \bar{\lambda}_1, \dots, \bar{\lambda}_m), \bar{c}) \\ &= (\partial_1 f(\bar{x}) - \bar{\lambda} \cdot \partial_1 g(\bar{x}), \dots, \partial_n f(\bar{x}) - \bar{\lambda} \cdot \partial_n g(\bar{x}), \bar{c}_1 - g_1(\bar{x}), \bar{c}_m - \dots, g_m(\bar{x})).\end{aligned}$$

Außerdem ist $\Phi : \mathbb{R}^{n+m} \times \mathbb{R}^m \to \mathbb{R}^{n+m}$ stetig differenzierbar wenn wir zweifache stetige Differenzierbarkeit für f und g annehmen. Um den Satz über implizite Funktionen anwenden zu können, berechnen wir zunächst die Ableitung nach den ersten $n+m$ Variablen im Punkt $((\bar{x}, \bar{\lambda}), \bar{c})$ und stellen den entsprechenden Teil der Jordanmatrix von Φ auf. Dies ist eine Blockmatrix in $\mathbb{R}^{(n+m)\times(n+m)}$

$$\begin{pmatrix} H_{f-\bar{\lambda}\cdot g}(\bar{x}) & -J_g(\bar{x})^T \\ -J_g(\bar{x}) & 0 \end{pmatrix}$$

die im oberen linken $n \times n$-Block die Hessematrix der Funktion $x \mapsto f(x) - \bar{\lambda} g(x)$ enthält. Multipliziert man die Matrix mit einem Vektor $\binom{\alpha}{\beta}$, so ergeben sich die beiden Bedingungen

$$J_g(\bar{x})\alpha = 0, \qquad H_{f-\bar{\lambda}\cdot g}(\bar{x})\alpha = J_g(\bar{x})^T \beta.$$

Insbesondere folgt bei Multiplikation der zweiten Gleichung mit α^T

$$\alpha^T H_{f-\bar{\lambda}\cdot g}(\bar{x})\alpha = \alpha^T J_g(\bar{x})^T \beta = (J_g(\bar{x})\alpha)^T \beta = 0.$$

Unter der Voraussetzung, dass die Hesse Matrix definit ist folgt dann $\alpha = 0$. Damit ist aber auch $0 = H_{f-\bar{\lambda}\cdot g}(\bar{x})\alpha = J_g(\bar{x})^T \beta$, so dass $\beta = 0$ falls J_g vollen Rang hat. Letztere Bedingung wird sowieso für den Satz über Extrema unter Nebenbedingungen benötigt (wobei die Nebenbedingungen durch $\psi(x) = g(x) - c$ gegeben sind). Unter den beiden Bedingungen ist also $\binom{\alpha}{\beta} = \binom{0}{0}$ das einzige Element des Kerns und damit ist die Matrix invertierbar und der Satz über implizite Funktionen anwendbar. Der besagt nun, dass eine (lokal) eindeutig bestimmte und glatte Auflösung $\varphi(c) = (\tilde{X}(c), \tilde{\Lambda}(c))$ existiert, so dass $\Phi((\tilde{X}(c), \tilde{\Lambda}(c)), c) = 0$. Da zu jedem c aber nach Voraussetzung eine Extremstelle $X(c)$ existiert, gibt es auch mit dem Satz über Extrema unter Nebenbedingungen jeweils ein $\Lambda(c)$, so dass $F'(X(c), \Lambda(c)) = 0$ gilt und somit auch $\Phi((X(c), \Lambda(c)), c) = 0$, d.h. die Eindeutigkeit erzwingt $X(c) = \tilde{X}(c)$ womit die Glattheit der Funktion X gezeigt ist.

Zum Abschluss vielleicht noch ein Wort zur Definitheitsbedingung an die Hesse-Matrix $H_{f-\bar{\lambda}\cdot g}(\bar{x})$. Diese ist eine *hinreichende* Bedingung für das Vorliegen eines Extremums unter Nebenbedingungen (der Nachweis wird geführt wie im Beweis des Satzes von Extrema unter Nebenbedingungen). Fordert man also die Definitheit der Hesse-Matrix, so ist sogar die Existenz der Funktion $X(c)$ gesichert (!) und man benötigt gar keine weiteren Voraussetzungen (außer der Glattheit und dem vollen Rang von g'), um zu zeigen, dass die Lagrange-Multiplikatoren gerade die Ableitungen des Optimalwerts nach den Niveauwerten c_i der Nebenbedingungen $g_i(x) = c_i$ sind.