Analysis II

Apl. Prof. Dr. Axel Grünrock

Heinrich-Heine-Universität Düsseldorf

Sommersemester 2020

Extrema mit Nebenbedingungen

Der Satz über implizite Funktionen, der in der letzten Videoaufzeichnung formuliert und bewiesen wurde, soll jetzt angewendet werden auf Extremwertaufgaben mit Nebenbedingungen. Was ist hierbei die Problemstellung?

Gegeben sei eine offene Menge $\Omega \subset \mathbb{R}^n$ und eine Funktion $f : \Omega \to \mathbb{R}$. f soll optimiert werden, jedoch nicht auf der gesamten Menge Ω sondern nur auf einer Teilmenge

$$M = \{x \in \Omega : \varphi(x) = 0\},$$

die als Nullstellenmenge einer Funktion

$$\varphi : \Omega \to \mathbb{R}^\ell, \quad \text{wobei} \quad 1 \leq \ell < n,$$

gegeben ist.

Beispiel 1:

Es sei A eine symmetrische, reelle $n \times n$ - Matrix. Gefragt ist nach den lokalen und globalen Extrema, die die Funktion

$$f : \mathbb{R}^n \to \mathbb{R},\, x \mapsto f(x) := \langle x, Ax \rangle$$

auf der Einheitssphäre $S^{n-1} := \{x \in \mathbb{R}^n : |x| = 1\}$ annimmt. Die Menge M ist hier also die S^{n-1}, die wir als Nullstellengebilde von $\varphi(x) = |x|^2 - 1$ auffassen können.

(Weiter unten werden wir dieses Beispiel fortsetzen und weitere Beispiele im Tutorium und in den Übungen diskutieren.)

Der Satz über die Lagrange-Multiplikatoren

Eine notwendige Bedingung für Extrema unter Nebenbedingungen stellt der folgende Satz zur Verfügung:

Satz: Es sei $\Omega \subset \mathbb{R}^n$ offen und $f : \Omega \to \mathbb{R}$ sowie $\varphi : \Omega \to \mathbb{R}^\ell$ seien stetig differenzierbar. In $z_0 \in M = \{x \in \Omega : \varphi(x) = 0\}$ gelte

$$\mathrm{Rang} D\varphi(z_0) = \ell$$

und die Einschränkung $f\big|_M : M \to \mathbb{R}$ besitze in z_0 ein lokales Extremum. Dann existieren $\lambda_1, ..., \lambda_\ell \in \mathbb{R}$, so dass

$$\nabla f(z_0) - \sum_{i=1}^{\ell} \lambda_i \nabla \varphi_i(z_0) = 0.$$

Bemerkungen

(1) Der Rang einer Matrix ist die Anzahl ihrer linear unabhängigen Zeilen bzw. Spalten.

Bemerkungen

(1) Der Rang einer Matrix ist die Anzahl ihrer linear unabhängigen Zeilen bzw. Spalten.

(2) Die Zahlen $\lambda_1, ..., \lambda_\ell \in \mathbb{R}$ werden als *Lagrange-Multiplikatoren* bezeichnet.

Beweis

Die Variablen $z \in \Omega$ von f und φ schreibt man als

$$z = (x_1, ..x_{n-\ell}, y_1, ..y_\ell) =: (x, y),$$

wobei die Nummerierung so gewählt wird, dass die ℓ Vektoren $\frac{\partial \varphi}{\partial y_1}(z_0), ..., \frac{\partial \varphi}{\partial y_\ell}(z_0)$ linear unabhängig sind. Ebenso schreibt man $z_0 = (a, b)$ mit $a \in \mathbb{R}^{n-\ell}$ und $b \in \mathbb{R}^\ell$. Mit diesen Bezeichnungen gehen die Voraussetzungen des Satzes über in

$$\det D_y \varphi(a, b) \neq 0 \quad \text{und} \quad \varphi(a, b) = 0,$$

der Satz über implizite Funktionen kann also auf φ angewendet werden. Demnach gibt es offene Umgebungen $W_a \subset \mathbb{R}^{n-\ell}$ von a und $W_b \subset \mathbb{R}^{n-\ell}$ von b und eine Funktion $g \in C^1(W_a, W_b)$,

Beweis (Fortsetzung)

so dass $g(a) = b$ und - für alle $x \in W_a$ - $\varphi(x, g(x)) = 0$ und

$$Dg(x) = -D_y\varphi(x, g(x))^{-1} D_x\varphi(x, g(x)).$$

Nun setzt man $\tilde{f}(x) = f(x, g(x))$. Da f in $z_0 = (a, b)$ ein lokales Extremum besitzt, hat $\tilde{f}$ ein ebensolches Extremum in a und daher gilt mit $\nabla_x = (\frac{\partial}{\partial x_1}, ..., \frac{\partial}{\partial x_{n-\ell}})$ und $\nabla_y = (\frac{\partial}{\partial y_1}, ..., \frac{\partial}{\partial y_\ell})$, dass

$$0 = \nabla_x \tilde{f}(a) = \nabla_x f(z_0) + \nabla_y f(z_0) Dg(a).$$

Beim zweiten $=$ wurde die Kettenregel benutzt. Setzen wir nun die Formel für Dg ein, erhalten wir

$$\nabla_x f(z_0) = \underbrace{\nabla_y f(z_0) D_y\varphi(z_0)^{-1}}_{=:(\lambda_1,...,\lambda_\ell)} D_x\varphi(z_0) = \sum_{i=1}^{\ell} \lambda_i \nabla_x \varphi_i(z_0).$$

Beweis (Ende)

Ergänzen wir noch unsere obige Definition der Lambdas in der Form

$$\nabla_y f(z_0) = (\lambda_1, \ldots, \lambda_\ell) D_y \varphi(z_0) = \sum_{i=1}^{\ell} \lambda_i \nabla_y \varphi_i(z_0),$$

so erhalten wir wegen $\nabla = (\nabla_x, \nabla_y)$ die Behauptung.

$\square$

Fortsetzung von Beispiel 1

Hier haben wir

$$f(x) = \langle x, Ax \rangle = \sum_{i,j=1}^{n} a_{ij} x_i x_j$$

mit

$$\frac{\partial f}{\partial x_k}(x) = \sum_{i,j=1}^{n} a_{ij}(\delta_{ik} x_j + \delta_{jk} x_i).$$

Da A als symmetrisch vorausgesetzt sind, ergeben beiden Summen dasselbe, zusammen also

$$.. = 2 \sum_{j=1}^{n} a_{kj} x_j = 2(Ax)_k.$$

Also $\nabla f(x) = 2Ax$ und (mit $A = E_n$) $\nabla \varphi(x) = 2x$. Die notwendige Bedingung lautet daher (hier ist $\ell = 1$) ...

Fortsetzung von Beispiel 1

$$0 = \nabla(f - \lambda\varphi)(x) = 2Ax - 2\lambda x,$$

das heißt $Ax = \lambda x$. Die Funktion f nimmt ihre lokalen Extrema also in den Eigenwerten der Matrix A an. (Dass solche Extrema existieren, folgt aus dem Satz vom Maximum, der gerade bewiesene Satz liefert das *nicht*!) Der Langrange-Multiplikator ist in diesem Beispiel gerade ein Eigenwert von A. Für den Funktionswert in den kritischen Stellen erhalten wir wegen $|x| = 1$:

$$f(x) = \langle x, Ax \rangle = \langle x, \lambda x \rangle = \lambda.$$

Es folgt

$$\max\{\langle x, Ax \rangle : |x| = 1\} = \max_{j=1}^{n}\{\lambda_j : \lambda_j \text{ ist ein Eigenwert von } A\}.$$

Entsprechendes gilt für das Minimum.

Folgerung

Es sei B eine reelle $m \times n$ - Matrix bzw. $B : \mathbb{R}^n \to \mathbb{R}^m$ eine dadurch definierte lineare Abbildung und $\|B\| = \sup\{|Bx| : |x| \leq 1\}$ die Operatornorm von B. Dann gilt:

$$\|B\| = \max_{j=1}^{n}\{\sqrt{\lambda_j} : \lambda_j \text{ ist ein Eigenwert von } B^\top B\}.$$

Beweis: $\|B\|^2 = \sup\{|Bx|^2 : x \in \mathbb{R}^n, |x| \leq 1\}$ mit

$$|Bx|^2 = \langle Bx, Bx\rangle = \langle x, B^\top Bx\rangle.$$

Nun ist die abgeschlossene Einheitskugel im $\mathbb{R}^n$ kompakt, daher das Supremum tatsächlich ein Maximum, welches wegen der Linearität von B auf dem Rand der Kugel angenommen wird. Also ist nach Beispiel 1

$$\begin{aligned}\|B\|^2 &= \max\{\langle x, B^\top Bx\rangle : |x| = 1\}\\ &= \max_{j=1}^{n}\{\lambda_j : \lambda_j \text{ ist ein Eigenwert von } B^\top B\}.\end{aligned}$$

Der Spektralsatz für reelle symmetrische Matrizen

Das obige Beispiel und die Folgerung daraus enthält bereits das wesentliche Argument für den Beweis des folgenden Satzes, den wir benutzt haben, um das Eigenwertkriterium für die Definitheit symmetrischer Matrizen herzuleiten.

Spektralsatz für reelle symmetrische Matrizen: Es sei $A \in M_n(\mathbb{R})$ symmetrisch. Dann existieren reelle Zahlen $\lambda_1, ..., \lambda_n$ und eine Orthonormalbasis $v_1, ..., v_n$ des $\mathbb{R}^n$, so dass für $1 \leq i \leq n$ gilt $Av_i = \lambda_i v_i$.

Anmerkung zum Beweis dieses Satzes

Per Induktion über $k \in \{1, ..., n\}$ wird die folgende Aussage gezeigt: Unter den Voraussetzungen des Satzes existieren $v_1, ..., v_k \in \mathbb{R}^n$ und $\lambda_1, ..., \lambda_k \in \mathbb{R}$, so dass für $1 \leq i, j \leq k$ gilt:

Anmerkung zum Beweis dieses Satzes

Per Induktion über $k \in \{1, ..., n\}$ wird die folgende Aussage gezeigt: Unter den Voraussetzungen des Satzes existieren $v_1, ..., v_k \in \mathbb{R}^n$ und $\lambda_1, ..., \lambda_k \in \mathbb{R}$, so dass für $1 \leq i, j \leq k$ gilt:

(1) $\langle v_i, v_j \rangle = \delta_{ij}$,

Anmerkung zum Beweis dieses Satzes

Per Induktion über $k \in \{1, ..., n\}$ wird die folgende Aussage gezeigt: Unter den Voraussetzungen des Satzes existieren $v_1, ..., v_k \in \mathbb{R}^n$ und $\lambda_1, ..., \lambda_k \in \mathbb{R}$, so dass für $1 \leq i, j \leq k$ gilt:

(1) $\langle v_i, v_j \rangle = \delta_{ij}$,

(2) $Av_i = \lambda_i v_i$,

Anmerkung zum Beweis dieses Satzes

Per Induktion über $k \in \{1, ..., n\}$ wird die folgende Aussage gezeigt: Unter den Voraussetzungen des Satzes existieren $v_1, ..., v_k \in \mathbb{R}^n$ und $\lambda_1, ..., \lambda_k \in \mathbb{R}$, so dass für $1 \leq i, j \leq k$ gilt:

(1) $\langle v_i, v_j \rangle = \delta_{ij}$,

(2) $Av_i = \lambda_i v_i$,

(3) $\lambda_1 \geq ... \geq \lambda_k$.

Hierbei ist der Induktionsanfang durch das vorhergehende Beispiel gegeben. Eine ähnliche Anwendung des Satzes über die Lagrange-Multiplikatoren verwendet man im Induktionsschluss. Die Einzelheiten finden Sie im Skript ausgeführt.