Walter Wenzel
Universität Leipzig
Mathematik II für Wirtschaftswissenschaftler

Kommentare zum Kapitel 2.7 – Lokale Extremstellen und Taylorpolynom

In diesem Kapitel kehren wir zurück zu *reellwertigen* differenzierbaren Funktionen, die auf einer Teilmenge D von $\mathbb{R}^n$ definiert sind. Bei jeder Abbildung, die nur reelle Zahlen als Funktionswerte aufweist, kann man diese nach ihrer Größe miteinander vergleichen – und genau wie bei reellen Funktionen in nur einer reellen Variablen von Maximalstellen und Minimalstellen sprechen. Definition 2.26 ist daher die natürliche Verallgemeinerung von Definition 4.3 aus Teil I des Kurses.

Der wichtige Satz 2.14 besagt: Ist $D \subseteq \mathbb{R}^n$, ist $f : D \to \mathbb{R}$ in einem inneren Punkt x von D partiell differenzierbar, und befindet sich dort eine lokale Extremstelle, so ist $\bigtriangledown f(x) = 0$.

Die Begründung sei hier kurz angegeben: Wir schreiben $x = (x_1, ..., x_n)$. Für genügend kleines $t_0 > 0$ hat dann jede der n Funktionen $f_i : [-t_0, t_0] \to \mathbb{R}$ – für $1 \leq i \leq n$, definiert durch $f_i(t) := f(x_1, ..., x_{i-1}, x_i + t, x_{i+1}, ..., x_n)$ in $t = 0$ eine lokale Extremstelle; daher folgt:

$$\tfrac{\partial f}{\partial x_i}(x) = f_i'(0) = 0.$$

Ein innerer Punkt x von D mit $\bigtriangledown f(x) = 0$ heißt *kritischer Punkt* von f; jede lokale Extremstelle ist also ein kritischer Punkt.

Die Umkehrung gilt jedoch *nicht:*
Ist der innere Punkt x ein kritischer Punkt, aber keine lokale Extremstelle, so heißt x eine *Sattelstelle*; der Punkt $(x, f(x)) \in \mathbb{R}^{n+1}$ heißt dann *Sattelpunkt.* – Solche Sattelpunkte findet man auf üblichen Sattelflächen.
In der Natur sind diese wie folgt anzutreffen:

Wird ein Kammweg eines Berges von einer Passstraße gekreuzt, so befindet sich bei der Kreuzung normalerweise ein Sattelpunkt:
Der Wanderweg hat dort ein lokales Minimum, aber die Straße ein lokales Maximum.

Es stellt sich nun die Frage, wie man feststellen kann, ob ein ermittelter kritischer Punkt auch lokale Extremstelle ist. – Das geschieht, ähnlich wie in der Analysis in einer Variablen, mit Hilfe zweifacher Ableitungen. Genauer betrachten wir dazu für eine zweimal partiell differenzierbare Funktion $f : D \to \mathbb{R}$ die **Hesse-Matrix**

$$Hess\, f(x) = (\partial_{ij} f(x))_{i,j=1,...,n} \in \mathbb{R}^{n \times n}.$$

Diese Matrix ist übrigens symmetrisch, falls alle zweifachen partiellen Ableitungen in einer Umgebung von x stetig sind.
Setzen wir $H_f(x)(v, v) := v \bullet (Hess\, f(x) \bullet v)$ für $v \in \mathbb{R}^n$, so erhalten wir den folgenden wichtigen Satz 2.16, der in Aufgabe 41.) bewiesen werden soll:

Sei $D \subseteq \mathbb{R}^n$, und $f : D \to \mathbb{R}$ sei zweimal stetig partiell differenzierbar in einem inneren Punkt $x_0 \in D$. Ferner sei $v \in \mathbb{R}^n$ so, dass $x_0 + t \cdot v \in D$ ist für alle $t \in [0, 1]$. Dann existiert ein $t \in [0, 1]$ mit

$$f(x_0 + v) = f(x_0) + \bigtriangledown f(x_0) \bullet v + \tfrac{1}{2} \cdot H_f(x_0 + tv)(v, v).$$

Der Beweis soll auf eine Funktion in der einzigen Variablen t – und eine analoge Aussage aus Teil I des Kurses – zurückgeführt werden.

Für uns eher noch wichtiger sind die Konsequenzen – für *lokale Extremstellen:*
Ist die Hesse-Matrix $Hess\,f(x_0)$ positiv definit, so sind die Matrizen $Hess\,f(x)$ auch noch für alle Punkte x aus einer geeigneten Umgebung von x_0 ebenfalls positiv definit. Das bedeutet insbesondere: $H_f(x_0 + tv)(v, v) > 0$ für *genügend kleines* $t \in [0, 1]$ und *alle* Vektoren $v \in \mathbb{R}^n$, die vom Nullvektor verschieden sind. Aus der Gleichung in Satz 2.16 folgt dann sofort:
Bei der kritischen Stelle x_0 befindet sich eine lokale Minimalstelle.

Analog folgt, dass sich bei der kritischen Stelle x_0 eine lokale Maximalstelle befindet, wenn die Hesse-Matrix $Hess\,f(x_0)$ negativ definit ist.

Aufgrund der gerade hervorgehobenen Bedeutung positiv bzw. negativ definiter Matrizen heben wir auch noch hervor:

(I) Eine Diagonalmatrix ist genau dann positiv definit bzw. negativ definit, wenn alle Koeffizienten auf der Hauptdiagonalen positiv bzw. wenn alle negativ sind.
Sind gewisse Einträge in einer Diagonalmatrix positiv und andere negativ, so befindet sich an einer kritischen Stelle mit solch einer Hesse-Matrix eine Sattelstelle.

(II) Eine symmetrische Matrix $A = (a_{ij})_{i,j=1,\dots,n} \in \mathbb{R}^{n\times n}$ ist genau dann positiv definit, wenn für alle Matrizen

$$A_k = (a_{ij})_{i,j=1,\dots,k} \in \mathbb{R}^{k\times k} \text{ mit } 1 \le k \le n$$

gilt: $det(A_k) > 0$.

(III) Eine Matrix A ist genau dann positiv definit, wenn $-A$ negativ definit ist.

Abschließend aber auch noch folgende *Warnung:*
Der Einsatz der Hesse-Matrizen ist im allgemeinen "nur" sinnvoll, um **lokale** Extremstellen zu bestimmen, *nicht* aber, um **globale** Extremstellen zu berechnen. Ein ähnliches Phänomen ist uns ja schon aus der Analysis in einer Variablen bekannt, wo die zweifachen Ableitungen noch gar nichts über *globale* Extremstellen aussagen.

Auf einem *kompakten* Definitionsbereich D gibt es aber sehr oft folgenden Ausweg:
Es *sei* bekannt, dass $f : D \to \mathbb{R}$ im Inneren von D überall partiell differenzierbar ist – und dass kein Randpunkt von D Extremstelle sein kann (was aber im Einzelfall zu überprüfen ist!). Wenn es dann genau zwei verschiedene kritische Punkte x_1, x_2 im Innern von D gibt und etwa $f(x_1) < f(x_2)$ ist, so folgt automatisch:
Bei x_1 ist die eindeutig bestimmte globale Minimalstelle; bei x_2 ist die eindeutig bestimmte globale Maximalstelle.
Auf der kompakten Menge D nimmt die stetige Funktion f nämlich einen minimalen Funktionswert und einen maximalen Funktionswert an. Das kann laut unserer obigen Annahme nicht auf dem Rand von D passieren. Folglich werden die globalen Extremwerte im Innern von D angenommen – und dort muss nach Satz 2.14 der Gradient jeweils der Nullvektor sein. Das kann aber wiederum nach Annahme nur zweimal passieren. Folglich ist eine dieser beiden Stellen die globale Minimalstelle und die andere die globale Maximalstelle.

Es sei auch noch einmal erwähnt, dass wir in der letzten Überlegung Voraussetzungen gefordert haben, die zwar recht häufig, aber längst nicht immer gegeben sind. Falls sie doch erfüllt sind, wird die Hesse- Matrix gar nicht mehr gebraucht!

Beispiele werden in späteren Ergänzungen folgen.