FOLIA 102

Annales Universitatis Paedagogicae Cracoviensis

Studia Mathematica X (2011)

Sándor Kovács

Über Extrema mit Nebenbedingungen

Zusammenfassung. Zweck der vorliegenden Arbeit is es, eine gut handhabbare Methode zu zeigen, womit man die hinreichende Bedingung für die Existenz eines Extremums unter Nebenbedingungen behandeln kann. Das Resultat ist eigentlich nicht unbekannt, Einzelteile sind in mehreren Arbeiten wie etwa in [5], [10] oder in [16] enthalten. Es hat aber nicht Eingang in die neuere Lehrbuchliteratur gefunden (vgl. z. B. [1], [11] oder [13]) und ist nicht allgemein bekannt. Die Frage ist von einigem Interesse, da zum Beispiel zahlreiche Probleme in der angewandten Mathematik Extremwertaufgaben unter Nebenbedingungen sind.

1. Einleitung

In vielen Analysisbüchern und auch in vielen Analysisvorlesungen ist die Behandlung der hinreichenden Bedingung zweiter Ordnung für lokale Extrema unter Nebenbedingungen bezüglich Funktionen von mehreren Veränderlichen einfach vernachlässigt. An manchen Stellen dieser Werke gibt es vorsichtige Andeutungen darauf, daß zusätzliche Überlegungen notwendig sind, um zu entscheiden, ob ein und welche Art von Extremum vorliegt. In einigen Büchern wird die Überlegung über diese hinreichenden Bedingungen mit einer Aussage kurz abgeschlossen, nach der es hierfür kein einfaches allgemein andwendbares Verfahren gebe. In der Tat werden des öfteren andere Methoden als die hinreichende Bedingung zweiter Ordnung zur Rechtfertigung der Existenz eines Extremums herangezogen, wie etwa Kompaktheitsschluß oder die Tatsache, nach der es bei zahlreichen Abstandsaufgaben anschaulich klar ist, daß es eine Lösung gibt, und diese anschauliche Evidenz wird z. B. durch Ausnutzung der endlichdimensionaler Eigenschaft der gegebener Aufgabe untermauert. Es kommt aber oft vor, daß die Anzahl der Kandidaten für Extrema mehr als zwei ist und nicht nach globalen, sondern nach lokalen Extrema gesucht werden soll. In diesen Fällen ist das Kompaktheitsargument nicht immer ausreichend, da in diesen Stellen die zu optimierende Funktion recht verschiedene Werte annehmen kann. Unter anderem aus diesem Grunde heraus scheint die Kenntnis der hinreichenden Bedingung zweiter Ordnung vonnöten zu sein.

Diese Arbeit soll eine einfache und in vielen Aufgaben gut handhabbare Methode zur Anwendung der hinreichenden Bedingung zweiter Ordnung bekanntgeben,

AMS (2000) Subject Classification: 00A35 (97D40).

wobei vollständigkeitshalber und im Interesse der einheitlichen Behandlung auch die notwendige Bedingung (erster Ordnung) präsentiert wird. Um die Natur der Fragestellung zu verdeutlichen, werden mehrere nützliche Aufgaben gestellt und gelöst. Das Gemeinsame dieser Aufgaben ist das Auffinden einer Stelle, wo eine Funktion beschränkt auf eine Teilmenge ihres Definitionsbereiches ein (lokales) Minimum oder Maximum annimt. Dies wird präzisiert in der

Definition 1.1
Es sei Ω eine nichtleere offene Teilmenge des $\mathbb{R}^n$ und vorgelegt seien die Funktionen $f\colon \Omega \to \mathbb{R}$ (genannt auch Zielfunktion) und $\mathbf{g}\colon \Omega \to \mathbb{R}^m$, $1 \leq m, n \in \mathbb{N}$. Wir sagen, f besitze in $\mathbf{c} \in \Omega$ ein lokales Extremum unter der Nebenbedingung $\mathbf{g} = \mathbf{0}$, wenn $\mathbf{c}$ zur Nebenbedingungsmenge

$$\{\mathbf{g} = \mathbf{0}\} := \{\mathbf{r} \in \Omega : \ \mathbf{g}(\mathbf{r}) = \mathbf{0}\} \neq \emptyset$$

gehört und eine Umgebung U von $\mathbf{c}$ vorhanden ist, so daß

$$f(\mathbf{r}) \leq f(\mathbf{c}) \quad \text{bzw.} \quad f(\mathbf{r}) \geq f(\mathbf{c}), \qquad \mathbf{r} \in U \cap \{\mathbf{g} = \mathbf{0}\}$$

gilt.

2. Eine notwendige Bedingung

Die vektorielle Nebenbedingung $\mathbf{g} = \mathbf{0}$ zerfällt im Falle von $m < n$ nach Zerlegung in Komponenten in m skalare Nebenbedingungen

$$\begin{aligned} g_1(x_1, \ldots, x_{n-m}, x_{n-m+1}, \ldots, x_n) &= 0, \\ &\vdots \\ g_m(x_1, \ldots, x_{n-m}, x_{n-m+1}, \ldots, x_n) &= 0. \end{aligned}$$

Manchmal kann dieses System von m Gleichungen explizit nach m Veränderlichen, etwa nach $x_{n-m+1}, \ldots, x_n$ aufgelöst werden, so daß

$$x_{n-m+1} = \varphi_1(x_1, \ldots, x_{n-m}), \ldots, x_n = \varphi_m(x_1, \ldots, x_{n-m})$$

mit bekannten Funktionen $\varphi_1, \ldots, \varphi_m$ ist. In diesem Falle läuft die Aufgabe darauf hinaus, die „freien"lokalen Extrema der Funktion

$$\begin{aligned} &\Phi(x_1, \ldots, x_{n-m}) \\ &\qquad := f(x_1, \ldots, x_{n-m}, \varphi_1(x_1, \ldots, x_{n-m}), \ldots, \varphi_m(x_1, \ldots, x_{n-m})) \\ &\qquad\qquad ((x_1, \ldots, x_{n-m}) \in \mathbb{R}^{n-m} : \ (x_1, \ldots, x_{n-m}, x_{n-m+1}, \ldots, x_n) \in \Omega) \end{aligned}$$

zu bestimmen. Dies wird benutzt bei der Lösung der

Aufgabe 2.1
Für eine Bewässerungsanlage soll ein Kanal mit gleichschenklig trapezförmigem Querschnitt aus drei gleich großen Betonfertigplatten der Breite b gebaut werden. Man bestimme die Anordnung der Platten, so daß möglichst viel Wasser transportiert werden kann!

Durch den Kanal kann möglichst viel Wasser fließen, wenn die Querschnittsfläche Q maximal wird. Je nachdem, in welchem Winkel α die Wände zur Horizontalen geneigt sind, ändern sich die Trapezhöhe h, die obere Breite β und mit ihnen die Querschnittsfläche Q. Für Q gilt nach der Flächenformel für ein Trapez

$$Q = \frac{\beta + b}{2} \cdot h \qquad \text{mit } h = b \cdot \sin(\alpha),\ \frac{\beta - b}{2} = b \cdot \cos(\alpha).$$

Die Zielfunktion und die Nebenbedingungsmenge sind also

$$f(\mathbf{r}) := \frac{\beta + b}{2} \cdot b \cdot \sin(\alpha),\ \{g = 0\} \qquad \text{mit } g(\mathbf{r}) := \frac{\beta - b}{2} - b \cdot \cos(\alpha)$$
$$(\mathbf{r} = (\alpha, \beta) \in \mathbb{R}^2 :\ 0 < \alpha < \pi,\ 0 < \beta < 3b).$$

Aus der Nebenbedingung errechnet man dann $\beta = \varphi(\alpha) := b \cdot (1 + 2\cos(\alpha))$. So sind die lokalen Extrema der Funktion

$$\Phi(\alpha) := f(\alpha, \varphi(\alpha)) = b^2 \cdot (1 + \cos(\alpha)) \cdot \sin(\alpha), \qquad \alpha \in (0, \pi)$$

aufzufinden. Wegen

$$\Phi'(\alpha) = b^2 \cdot (\cos^2(\alpha) - \sin^2(\alpha) + \cos(\alpha)) = b^2 \cdot (2\cos^2(\alpha) + \cos(\alpha) - 1),$$

und

$$\Phi''(\alpha) = -b^2 \cdot \sin(\alpha) \cdot (1 + 4\cos(\alpha)),$$

sowie

$$\Phi'\left(\frac{\pi}{3}\right) = 0 \quad \text{bzw.} \quad \Phi''\left(\frac{\pi}{3}\right) = -\frac{3\sqrt{3} \cdot b^2}{2} < 0$$

haben wir es mit einem Maximum zu tun. Der Kanal hat also den größten Durchfluß, wenn die Wände um 60° zur Horizontalen geneigt sind.

Im Falle $n = 2$, $m = 1$ könnte so eine Aufgabe zum Beispiel auch durch die Erfüllung der Nebenbedingung $g(x_1, x_2) = 0$ durch eine Parameterdarstellung

$$x_1 = \varphi_1(t),\ x_2 = \varphi_2(t), \quad t \in I \qquad \text{mit } \varphi_1, \varphi_2 \in \mathfrak{D}(I)$$

auf einem Intervall $I \subset \mathbb{R}$ gelöst und die Extremalstellen der Funktion F mit

$$F(t) := f(\varphi_1(t), \varphi_2(t)), \qquad t \in I$$

mit den üblichen Methoden bestimmt werden, vorausgesetzt f ist differenzierbar. Dieser Trick wird ausgenutzt bei der Lösung der

Aufgabe 2.2
Es sollen diejenigen Punkte einer durch die Gleichung

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \qquad (0 < a < b)$$

gegebenen Ellipse bestimmt werden, die von ihrem Mittelpunkt $(0, 0)$ maximalen bzw. minimalen Abstand haben.

Die Parametrisierung

$$x = \varphi_1(t) := a\cos(t), \quad y = \varphi_2(t) := b\sin(t), \qquad t \in [0, 2\pi]$$

der Ellipse führt zu der Funktion

$$F(t) := a^2\cos^2(t) + b^2\sin^2(t), \qquad t \in [0, 2\pi]$$

(statt des Abstandes wird – zur formalen Vereinfachung – das Quadrat des Abstandes verwendet: $f(x, y) := x^2 + y^2$, $(x, y) \in \mathbb{R}^2$; dabei werden dieselben Extremalstellen erhalten). Wegen

$$F'(t) = (b^2 - a^2)\sin(2t), \qquad t \in [0, 2\pi]$$

und $F'(s) = 0$, $s \in \{0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}, 2\pi\}$ sowie

$$F(0) = F(\pi) = a^2 < b^2 = F\left(\frac{\pi}{2}\right) = F\left(\frac{3\pi}{2}\right)$$

sind die gefundenen Punkte die Scheitelpunkte der Ellipse.

In der Praxis sind diese Methoden nur selten anwendbar: die explizite Auflösung ist in den meisten Fällen nicht möglich und auch die Bestimmung einer Parameterdarstellung ist oft ziemlich mühsam. Man denke nur an die Variante der Aufgabe 2.2, wo die gegebene Ellipse die Gleichung

$$ax^2 + 2bxy + cy^2 = 1 \qquad (a > 0,\ ac - b^2 > 0)$$

hat. Wie die Auflösung nach Variablen oder die Parameterbestimmung umgegangen werden kann, wird z. B. aus der folgenden Überlegung ersichtlich. Ist nämlich bei der Parameterbestimmung $\tau \in I$ derjenige Parameterwert, wofür $(\varphi_1(\tau), \varphi_2(\tau)) = \mathbf{c}$ gilt, so genügt er der Gleichung

$$0 = F'(t) = \partial_1 f(\varphi_1(t), \varphi_2(t)) \cdot \varphi_1'(t) + \partial_2 f(\varphi_1(t), \varphi_2(t)) \cdot \varphi_2'(t), \qquad t \in I.$$

Wegen $g(\varphi_1(t), \varphi_2(t)) = 0$, $t \in I$ gilt weiter

$$\partial_1 g(\varphi_1(t), \varphi_2(t)) \cdot \varphi_1'(t) + \partial_2 g(\varphi_1(t), \varphi_2(t)) \cdot \varphi_2'(t) = 0, \qquad t \in I.$$

Für $t = \tau$ sind also die Vektoren $f'(\mathbf{c})$ und $g'(\mathbf{c})$ linear abhängig, d. h. es gibt $\lambda \in \mathbb{R}$ derart, daß

$$\partial_1 f(\mathbf{c}) + \lambda\partial_1 g(\mathbf{c}) = 0, \quad \partial_2 f(\mathbf{c}) + \lambda\partial_2 g(\mathbf{c}) = 0 \quad \text{und} \quad g(\mathbf{c}) = 0$$

gelten.

Das hier skizzierte Verfahren wird für höhere Dimensionen im nächsten Satz formuliert.

Satz 2.3
Die folgenden drei Bedingungen seien erfüllt.

1. *f sei differenzierbar, $\mathbf{g}$ sei stetig differenzierbar.*

2. *f habe in $\mathbf{c} \in \Omega$ ein lokales Extremum unter der Nebenbedingung $\mathbf{g} = \mathbf{0}$.*

3. *Es gebe in $\mathbf{g}'(\mathbf{c})$ eine m-reihige Unterdeterminante, die nicht verschwindet, d. h. der Rang von $\mathbf{g}'(\mathbf{c})$ sei m (Rang- bzw. Regularitätsbedingung).*

Dann gibt es einen Vektor $\boldsymbol{\lambda} \in \mathbb{R}^m$ (Lagrange-Multiplikator), so daß mit $L := f + \langle \boldsymbol{\lambda}, \mathbf{g} \rangle$ (Lagrange-Funktion) gilt: $L'(\mathbf{c}) = \mathbf{0}$, d. h.

$$\partial_k f(\mathbf{c}) + \sum_{l=1}^{m} \lambda_l \partial_k g_l(\mathbf{c}) = 0, \qquad 1 \leq k \leq n.$$

Beweis. (vgl. [14]) Aus der Rangbedingung folgt, daß zwischen m und n die Relation $m \leq n$ bestehen muß. Im Falle von $m = n$ folgt aus dem Satz über inverse Funktionen, daß der Punkt $\mathbf{c}$ eine Umgebung hat, die gemeinsam mit der Nebenbedingungsmenge nur den Punkt $\mathbf{c}$ innehat. In diesem – hinsichtlich des Extremums – uninteressanten Fall, ist die Existenz eines $\boldsymbol{\lambda} \in \mathbb{R}^m$ wegen der dritten Bedingung offensichtlich. Des weiteren sei o. B. d. A. angenommen, daß $m < n$ gilt und die Matrix, die aus den letzten m Spalten von $\mathbf{g}'(\mathbf{c})$ gebildet wird, eine von Null verschiedene Determinante hat. (Notfalls benennt man die Variablen entsprechend um.) $\mathbf{g}'(\mathbf{c})$ läßt sich also in der Form $\mathbf{g}'(\mathbf{c}) = [G_1, G_2]$ schreiben, wobei $G_1 \in \mathbb{R}^{m \times (n-m)}$, $G_2 \in \mathbb{R}^{m \times m}$ mit

$$G_1 := \partial_1 \mathbf{g}(\mathbf{c}) := [\partial_j g_i(\mathbf{c})]_{i,j=1,1}^{m,n-m}, \qquad G_2 := \partial_2 \mathbf{g}(\mathbf{c}) := [\partial_j g_i(\mathbf{c})]_{i,j=1,n-m+1}^{m,n}$$

und $\det(\partial_2 \mathbf{g}(\mathbf{c})) \neq 0$. Nach dem Satz von der impliziten Funktion (angewandt auf $\mathbf{g}$) gibt es dann eine Umgebung U von $\mathbf{a} := (c_1, \ldots, c_{n-m}) \in \mathbb{R}^{n-m}$ und eine Umgebung V von $\mathbf{b} := (c_{n-m+1}, \ldots, c_n) \in \mathbb{R}^m$ mit $U \times V \subset \Omega$ und eine stetig differenzierbare Funktion $\boldsymbol{\varphi}\colon U \to V$ mit

$$\{\mathbf{g} = \mathbf{0}\} \cap (U \times V) = \{(\mathbf{r}, \boldsymbol{\varphi}(\mathbf{r})) : \ \mathbf{r} \in U\},$$

d. h.

$$\mathbf{g}(\mathbf{r}, \boldsymbol{\varphi}(\mathbf{r})) = \mathbf{0}, \qquad \mathbf{r} \in U$$

bzw.

$$\boldsymbol{\varphi}'(\mathbf{r})) = -[\partial_2 \mathbf{g}(\mathbf{r}, \boldsymbol{\varphi}(\mathbf{r})]^{-1} \cdot \partial_1 \mathbf{g}(\mathbf{r}, \boldsymbol{\varphi}(\mathbf{r})), \qquad \mathbf{r} \in U.$$

Da die Beschränkung von f auf $\{\mathbf{g} = \mathbf{0}\} \cap (U \times V)$ im Punkte $\mathbf{c} = (\mathbf{a}, \mathbf{b})$ ($\mathbf{b} = \boldsymbol{\varphi}(\mathbf{a})$) ein lokales Extremum besitzt, hat auch die Funktion

$$\Phi(\mathbf{r}) := f(\mathbf{r}, \boldsymbol{\varphi}(\mathbf{r})), \qquad \mathbf{r} \in U$$

ein (freies) lokales Extremum in $\mathbf{c}$. Nach der notwendigen Bedingung für lokale Extrema gilt also

$$\mathbf{0} = \Phi'(\mathbf{c}) = \partial_1 f(\mathbf{a}, \mathbf{b}) + \partial_2 f(\mathbf{a}, \mathbf{b}) \cdot \boldsymbol{\varphi}'(\mathbf{a}),$$

woraus mit

$$\boldsymbol{\lambda} := -\partial_2 f(\mathbf{a}, \mathbf{b}) \cdot [\partial_2 \mathbf{g}(\mathbf{a}, \mathbf{b})]^{-1}$$

die Gleichheit

$$\mathbf{0} = \partial_1 f(\mathbf{a},\mathbf{b}) - \partial_2 f(\mathbf{a},\mathbf{b}) \cdot [\partial_2 \mathbf{g}(\mathbf{a},\mathbf{b})]^{-1} \cdot \partial_1 g(\mathbf{a},\mathbf{b}) = \partial_1 f(\mathbf{a},\mathbf{b}) + \boldsymbol{\lambda}\partial_1 \mathbf{g}(\mathbf{a},\mathbf{b}),$$

sowie (aus der Definition für $\boldsymbol{\lambda}$)

$$\partial_2 f(\mathbf{a},\mathbf{b}) + \boldsymbol{\lambda}\partial_2 \mathbf{g}(\mathbf{a},\mathbf{b}) = \mathbf{0}$$

folgt.

Um die kritische Stelle $\mathbf{c}$ zu bestimmen, muß also das System der $(m+n)$ Gleichungen

$$\begin{aligned} \partial_k f(\mathbf{r}) + \sum_{l=1}^{m} \lambda_l \partial_k g_l(\mathbf{r}) &= 0, \qquad 1 \le k \le n, \\ g_l(\mathbf{r}) &= 0, \qquad 1 \le l \le m \end{aligned} \tag{1}$$

für $\mathbf{r}$ und $\boldsymbol{\lambda}$ gelöst werden. Als Beispiel betrachten wir zwei Aufgaben.

Aufgabe 2.4
Es sei $A \in \mathbb{R}^{n\times n}$ eine symmetrische Matrix. Gesucht sind die Extrema der Funktion

$$f(\mathbf{r}) := \langle A\mathbf{r}, \mathbf{r}\rangle = \sum_{i,j=1}^{n} a_{ij} x_i x_j \qquad (\mathbf{r} = (x_1, \ldots, x_n) \in \mathbb{R}^n : \ \|\mathbf{r}\|_2 = 1).$$

Da g die Form

$$g(\mathbf{r}) := \sum_{i=1}^{n} x_i^2 - 1, \qquad \mathbf{r} = (x_1, \ldots, x_n) \in \mathbb{R}^n$$

hat, ist die Rangbedingung $2\mathbf{c} \neq \mathbf{0}$, d. h. $\mathbf{c} \neq \mathbf{0}$ erfüllt, denn $\mathbf{0} \notin \{\mathbf{g} = \mathbf{0}\}$. Anhand (1) erhalten wir also das folgende Gleichungssystem:

$$\sum_{j=1}^{n} a_{kj} x_j + \sum_{i=1}^{n} a_{ki} x_i + 2\lambda x_k = 0, \quad 1 \le k \le n, \qquad \sum_{i=1}^{n} x_i^2 - 1 = 0.$$

$\mathbf{c}$ genügt der Gleichung (1), wenn

$$2\sum_{i=1}^{n} a_{ki} c_i + 2\lambda c_k = 0, \quad 1 \le k \le n \qquad \text{und} \qquad \sum_{i=1}^{n} c_i^2 = 1$$

gelten. Es gibt also ein $\mu \in \mathbb{R}$, so daß $A\mathbf{c} = \mu\mathbf{c}$ gilt, d. h. $\mathbf{c}$ ist ein normierter Eigenvektor der Matrix A zum Eigenwert $-\lambda$.

Aufgabe 2.5
Die lokalen Extrema der Funktion

$$f(x,y,z) := x + 2y + 3z \qquad ((x,y,z) \in \mathbb{R}^3 : \ x^2 + y^2 = 2, \ y + z = 1)$$

sollen bestimmt werden.

Die Rangbedingung ist erfüllt wenn

$$\operatorname{Rang}(\mathbf{g}'(\mathbf{c})) = \operatorname{Rang}\begin{pmatrix} 2c_1 & 2c_2 & 0 \\ 0 & 1 & 1 \end{pmatrix} = 2$$

gilt. Anhand (1) hat das zu lösende Gleichungssystem die Form

$$1+2\lambda_1 x=0, \quad 2+2\lambda_1 y+\lambda_2=0, \quad 3+\lambda_2=0, \quad x^2+y^2-2=0, \quad y+z-1=0.$$

Als Lösung haben wir zwei kritische Stellen: $\mathbf{c} \in \{(1,-1,2);(-1,1,0)\}$ (mit $\lambda_1 = -\frac{1}{2}$, $\lambda_1 = \frac{1}{2}$ bzw. $\lambda_2 = -3$), mit denen die Rangbedingung erfüllt ist.

3. Hinreichende Bedingungen

3.1. Kompaktheitsargumente

Ist die Nebenbedingungsmenge $\{\mathbf{g}=\mathbf{0}\}$ kompakt und die Zielfunktion f stetig, so ist die Existenz der absoluten Extrema gesichert, wie das in den Aufgaben 2.4 und 2.5 der Fall ist. Dort sind nämlich die Nebenbedingungsmengen die Einheitskugel und eine Ellipse (die sich als Schnitt eines Zylinders und einer Ebene ergibt), die beschränkt und abgeschlossen, somit kompakt sind. Da im ersten Falle $f(\mathbf{c}) = \langle A\mathbf{c},\mathbf{c}\rangle = \mu$ gilt, wird das Maximum bzw. das Minimum bei einem Eigenvektor zum größten bzw. zum kleinsten Eigenwert angenommen. (Insbesondere folgt, daß eine reelle symmetrische Matrix mindestens einen reellen Eigenwert besitzt.) Im zweiten Falle entscheiden die Funktionswerte: $f(-1,1,0)=1<5=f(1,-1,2)$.

3.2. Abstandsaufgaben

Wenn es um Abstandsaufgaben geht, so erweist sich eine Behauptung, – die auf der Endlichdimensionalität des euklidischen Raumes $\mathbb{R}^d$ $(1 \le d \in \mathbb{N})$ beruht – als nützlich:

Satz 3.1
A sei eine kompakte und B eine abgeschlossene Teilmenge des $\mathbb{R}^d$; beide Mengen seien nicht leer. Dann gibt es in A einen Punkt $\mathbf{a}$ *und in B einen Punkt* $\mathbf{b}$ *mit*

$$\|\mathbf{a}-\mathbf{b}\| \le \|\mathbf{x}-\mathbf{y}\|, \qquad \mathbf{x}\in A,\ \mathbf{y}\in B,$$

wobei $\|\cdot\|$ irgendeine Norm auf $\mathbb{R}^d$ ist.

Statt des Beweises (siehe z. B. [11]), geben wir einen Fall an, wo dieser Satz vom Nutzen sein kann:

Aufgabe 3.2
Man bestimme denjenigen Punkt auf dem Schnitt der Ebenen $x+2y+z=1$ und $2x-y-3z=4$, der von der Koordinatenursprung den kleinsten (euklidischen) Abstand hat!

Die Zielfunktion und die Nebenbedingungsmenge sind also für $\mathbf{r}=(x,y,z)\in\mathbb{R}^3$

$$f(\mathbf{r}) := x^2+y^2+z^2,\ \{\mathbf{g}=\mathbf{0}\} \qquad \text{mit } \mathbf{g}(\mathbf{r}) := \begin{bmatrix} x+2y+z-1 \\ 2x-y-3z-4 \end{bmatrix}.$$

Die Rangbedingung ist überall erfüllt, denn

$$\operatorname{Rang}(\mathbf{g}'(\mathbf{r})) = \operatorname{Rang}\begin{bmatrix} 1 & 2 & 1 \\ 2 & -1 & -3 \end{bmatrix} = \operatorname{Rang}\begin{bmatrix} 1 & 2 & 1 \\ 0 & -5 & -5 \end{bmatrix} = 2, \qquad \mathbf{r}\in\mathbb{R}^3$$

gilt. Das Gleichungssystem (1) hat die Form

$$\begin{cases} 2x + \lambda_1 + 2\lambda_2 = 0, \\ 2y + 2\lambda_1 - \lambda_2 = 0, \\ 2z + \lambda_1 - 3\lambda_2 = 0, \\ x + 2y + z - 1 = 0, \\ 2x - y - 3z - 4 = 0. \end{cases}$$

Die einzige kritische Stelle $\mathbf{c} = (\frac{16}{15}, \frac{1}{3}), -\frac{11}{15}$ mit $\lambda_1 = -\frac{52}{75}$, $\lambda_2 = -\frac{54}{75}$ von L steht also im Verdacht, die gesuchte Stelle zu sein. Da wegen Satz 3.1 eine solche Stelle wirklich vorhanden ist, muß $\mathbf{c}$ tatsächlich die Lösung der Aufgabe 3.2 sein.

3.3. Hinreichende Bedingungen zweiter Ordnung

Es ist offensichtlich, daß im Falle von Funktionen $f, \mathbf{g} \in \mathfrak{D}^2[\mathbf{c}]$ die Existenz eines lokalen Extremums unter Nebenbedingungen durch die Definitheit der Hesse-Matrix $L''(\mathbf{c})$ gesichert wird. Aus der Definitheit folgt nämlich, daß die Lagrange-Funktion in $\mathbf{c}$ ein lokales Extremum besitzt, also eine Umgebung $U \subset \Omega$ von $\mathbf{c}$ gibt, wofür

$$L(\mathbf{r}) \leq L(\mathbf{c}) \quad \text{bzw.} \quad L(\mathbf{r}) \geq L(\mathbf{c}), \qquad \mathbf{r} \in U$$

gilt. Aus $\mathbf{c} \in \{\mathbf{g} = \mathbf{0}\}$ folgt sofort die Ungleichung

$$f(\mathbf{r}) = L(\mathbf{r}) \leq L(\mathbf{c}) = f(\mathbf{c}) \quad \text{bzw.} \quad f(\mathbf{r}) = L(\mathbf{r}) \geq L(\mathbf{c}) = f(\mathbf{c}) \qquad (\mathbf{r} \in U \cap \{\mathbf{g} = \mathbf{0}\}).$$

Dieses Resultat kann man z. B. bei der Lösung der folgenden Aufgabe ausnutzen.

Aufgabe 3.3
Man bestimme die Extrema der Funktion

$$f(\mathbf{r}) := x^2 - 2x + 2y^2 + z^2 + z \quad (\mathbf{r} = (x, y, z) \in \mathbb{R}^3 : \ x + y + z = 1,\ 2x - y - z = 5)!$$

Wie bei der Lösung der Aufgabe 3.2 kommt man hier mit der Kompaktheitsschluß ebenfalls nicht aus, denn die Nebenbedingunsmenge ist eine – als Schnitt von zwei Ebenen – ergebende Gerade, die unbeschränkt, somit nicht kompakt ist. Die Zielfunktion ist auch keine Abstandsfunktion, so scheint die Berechnung der zweiten Ableitung L'' in $\mathbf{c}$ sinnvoll zu sein. Die Lösung des Gleichungssystems (1) ist der einzige Vektor $\mathbf{c} = (2, 0, -1)$ mit $\lambda_1 = \lambda_2 = -\frac{3}{2}$. Die Rangbedingung ist überall erfüllt. Die Hesse-Matrix $L''(\mathbf{c}) = \mathrm{diag}\{2, 4, 2\}$ ist positiv definit, demzufolge hat f in $\mathbf{c}$ unter der Nebenbedingung $\{\mathbf{g} = \mathbf{0}\}$ ein lokales Minimum.

Oft kommt aber vor, daß $L''(\mathbf{c})$ nicht definit ist, wie das der Fall der folgenden Aufgabe zeigt.

Aufgabe 3.4
Man bestimme die lokalen Extrema der Funktion

$$f(\mathbf{r}) := xy + xz + xw + yz + yw + zw \quad (\mathbf{r} = (x, y, z, w) \in (\mathbb{R}^+)^4 : \ xyzw = 1)!$$

Die einzige kritische Stelle von L ist $\mathbf{c} := (1,1,1,1)$, mit $\lambda = 4$. Die Rangbedingung ist trivialerweise erfüllt, denn $g'(\mathbf{c}) = (1,1,1,1)$. Die Eigenwerte von $L''(\mathbf{c})$ sind -9 und 3 (dies letzte ist ein dreifacher Eigenwert), so ist $L''(\mathbf{c})$ nicht definit: indefinit.

In Fällen also, wo die obigen Hilfsmittel nicht ausreichen, könnte man eine hinreichende Bedingung zweiter Ordnung benutzen, die der eigentliche Gegenstand dieser Arbeit ist. Bevor wir aber die diesbezügliche Behauptung formulieren, wollen wir den folgenden Begriff bestimmen.

Definition 3.5
Es seien $0 < m, n \in \mathbb{N}$ mit $m < n$. Die symmetrische Matrix $A \in \mathbb{R}^{n\times n}$ bzw. die quadratische Form Q mit $Q(\mathbf{r}) := \langle A\mathbf{r}, \mathbf{r}\rangle$, $\mathbf{r} \in \mathbb{R}^n$ heißt positiv bzw. negativ definit bezüglich der Matrix $B \in \mathbb{R}^{m\times n}$ mit $\operatorname{Rang}(B) = m$, wenn Q positiv bzw. negativ definit auf dem Kernraum von B ist, d. h. aus $\mathbf{0} \neq \mathbf{r} \in \mathbb{R}^n$, $B\mathbf{r} = \mathbf{0}$ die Ungleichung $Q(\mathbf{r}) > 0$ bzw. $Q(\mathbf{r}) < 0$ folgt.

Beispiel 3.6
Im Fall $m = 1$, $n = 2$ habe die quadratische Form Q und die Matrix B die Gestalt

$$Q(\mathbf{r}) := ax^2 + 2bxy + cy^2, \quad \mathbf{r} = (x,y) \in \mathbb{R}^2, \qquad B := [d,\ e].$$

Der Vektor $\mathbf{r} = [x,y]^T$ gehört somit genau dann zum Kernraum von B, wenn $dx + ey = 0$ gilt. Die Forderung $\operatorname{Rang}(B) = 1$ heißt in diesem Falle, daß eine von den zwei Komponenten ungleich Null ist: $d^2 + e^2 > 0$. Ist z. B. $e \neq 0$, so bekommt man durch Einsetzen von $y = -\frac{dx}{e}$ in Q, daß

$$Q(\mathbf{r}) = ax^2 + 2bx\Big(-\frac{dx}{e}\Big) + c\Big(-\frac{dx}{e}\Big)^2 = \frac{(ae^2 - 2bde + cd^2)x^2}{e^2}, \quad \mathbf{r} = (x,y) \in \mathbb{R}^2.$$

Der Koeffizient von x^2 im Zähler lässt sich in der Form

$$ae^2 - 2bde + cd^2 = -\det \begin{bmatrix} a & b & d \\ b & c & e \\ d & e & 0 \end{bmatrix} \tag{2}$$

schreiben. So ist z. B. die positive Definitheit von Q bezüglich B damit gleichwertig, daß die obige Determinante negativ ist.

Somit kann schon das Resultat über eine hinreichende Bedingung zweiter Ordnung für die Existenz eines Extremums unter Nebenbedingungen formuliert werden, das übrigens schon gegen Ende des 19. Jahrhunderts bekannt war.

Satz 3.7
Die Funktionen f bzw. $\mathbf{g}$ aus Satz 2.3 seien zweimal differenzierbar, ferner sei $\mathbf{c} \in \{\mathbf{g} = \mathbf{0}\}$, wofür die notwendige Bedingung erster Ordnung erfüllt ist, d. h. $\operatorname{Rang}(\mathbf{g}'(\mathbf{c})) = m$ und für den zugehörigen Lagrange-Multiplikator $\boldsymbol{\lambda} \in \mathbb{R}^m$ und für die Lagrange-Funktion $L := f + \langle \boldsymbol{\lambda}, \mathbf{g}\rangle$ gilt $L'(\mathbf{c}) = \mathbf{0}$. Ist die quadratische Form

$$Q_{\mathbf{c}}^L(\mathbf{r})\colon \mathbb{R}^n \to \mathbb{R}, \qquad Q_{\mathbf{c}}^L(\mathbf{r}) := \langle L''(\mathbf{c})\mathbf{r}, \mathbf{r}\rangle$$

bezüglich $\mathbf{g}'(\mathbf{c})$ positiv bzw. negativ definit, so hat f ein lokales Minimum bzw. Maximum in $\mathbf{c}$ unter der Nebenbedingung $\mathbf{g} = \mathbf{0}$.

Beweis. (vgl. [10]) Wir werden nur die Behauptung bezüglich Minimum beweisen, da offensichtliche Modifikationen der folgenden Argumentation zum Fall des Maximums führen.

Angenommen, hätte f kein lokales Minimum in $\mathbf{c}$ unter der Nebenbedingung $\mathbf{g} = \mathbf{0}$, so gäbe es eine Folge

$$\mathbf{r}_k \in \{\mathbf{g} = \mathbf{0}\} \backslash \{\mathbf{c}\}, \quad k \in \mathbb{N} \qquad \text{mit } \lim(\|\mathbf{r}_k - \mathbf{c}\|) = 0 \text{ und } f(\mathbf{r}_k) < f(\mathbf{c}),$$

wobei $\|\cdot\|$ die euklidische Norm auf $\mathbb{R}^n$ bezeichnet: $\|\cdot\| := \|\cdot\|_2$. Da die Folge

$$\mathbf{h}_k := \frac{\mathbf{r}_k - \mathbf{c}}{\|\mathbf{r}_k - \mathbf{c}\|}, \qquad k \in \mathbb{N}$$

beschränkt ist, hat sie eine konvergente Teilfolge $(\mathbf{h}_{\nu_k})$, $k \in \mathbb{N}$ mit

$$\lim(\mathbf{h}_{\nu_k}) =: \mathbf{h} \quad \text{und} \quad \|\mathbf{h}\| = 1.$$

Wegen $\mathbf{g}(\mathbf{r}_{\nu_k}) = \mathbf{g}(\mathbf{c}) = \mathbf{0}$, $k \in \mathbb{N}$ folgt

$$\mathbf{0} = \lim\left(\frac{\mathbf{g}(\mathbf{r}_k) - \mathbf{g}(\mathbf{c})}{\|\mathbf{r}_k - \mathbf{c}\|}\right) = \mathbf{g}'(\mathbf{c})\mathbf{h} \quad \text{also} \quad Q^L_{\mathbf{c}}(\mathbf{h}) > 0, \tag{3}$$

denn laut Voraussetzung $Q^L_{\mathbf{c}}$ ist bezüglich $\mathbf{g}'(\mathbf{c})$ positiv definit. Da die Lagrange-Funktion L auf der Nebenbedingungsmenge $\{\mathbf{g} = \mathbf{0}\}$ mit f identisch ist, folgt

$$L(\mathbf{c}) = f(\mathbf{c}) \quad \text{bzw.} \quad L(\mathbf{r}_k) = f(\mathbf{r}_k).$$

Die Taylorsche Formel mit Restglied nach Peano liefert dann für L:

$$\begin{aligned} f(\mathbf{c}) > f(\mathbf{r}_k) &= L(\mathbf{r}_k) \\ &= L(\mathbf{c}) + \langle L'(\mathbf{c})(\mathbf{r}_k - \mathbf{c}), (\mathbf{r}_k - \mathbf{c})\rangle \\ &\quad + \frac{1}{2}\langle L''(\mathbf{c})(\mathbf{r}_k - \mathbf{c}), \mathbf{r}_k - \mathbf{c}\rangle + \eta(\mathbf{r}_k - \mathbf{c}) \cdot \|(\mathbf{r}_k - \mathbf{c})\|^2 \\ &= f(\mathbf{c}) + 0 + \frac{1}{2} Q^L_{\mathbf{c}}(\mathbf{r}_k - \mathbf{c}) + \eta(\mathbf{r}_k - \mathbf{c})\|\mathbf{r}_k - \mathbf{c}\|^2 \end{aligned}$$

mit $\lim_{\mathbf{r}\to\mathbf{0}} \eta(\mathbf{r}) = 0$ (vgl. [15]) also

$$0 > Q^L_{\mathbf{c}}(\mathbf{r}_k - \mathbf{c}) + 2\eta(\mathbf{r}_k - \mathbf{c})\|\mathbf{r}_k - \mathbf{c}\|^2.$$

Dividiert man diese Ungleichung durch $\|\mathbf{r}_k - \mathbf{c}\|^2$ und bildet den Grenzübergang $k \to \infty$, so erhält man $0 \geq Q^L_{\mathbf{c}}(\mathbf{h})$, was der Ungleichung in (3) widerspricht.

Als Beispiel wollen wir nochmal Aufgabe 2.5 behandeln. Dort ist die zweite Ableitung der Lagrange-Funktion semidefinit:

$$L''(\mathbf{c}) = \begin{bmatrix} 2\lambda_1 & 0 & 0 \\ 0 & 2\lambda_1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Die quadratische Form

$$Q_{\mathbf{c}}^{L}(\mathbf{r}) = \langle L''(\mathbf{c})\mathbf{r}, \mathbf{r}\rangle = 2\lambda_1(x^2+y^2), \qquad \mathbf{r} = (x,y,z) \in \mathbb{R}^3$$

ist aber bezüglich $\mathbf{g}'(\mathbf{c})$ definit, denn für alle $\mathbf{r} = (x,y,z) \in \mathbb{R}^3$ gilt

$$g'(\mathbf{c})\mathbf{r} = \begin{bmatrix} 2c_1 & 2c_2 & 0 \\ 0 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2c_1x + 2c_2y \\ y+z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$
$$\iff \left(x = -\frac{c_2}{c_1}y \;\&\; y = -z\right),$$

also

$$g'(\mathbf{c})\mathbf{r} = \mathbf{0} \iff \mathbf{r} = (\xi, \xi, -\xi), \qquad \xi \in \mathbb{R},$$

d. h. für $\xi \neq 0$ gilt

$$Q_{\mathbf{c}}^{L}(\mathbf{r}) = \begin{cases} 4\lambda_1\xi^2 > 0\,, & \lambda_1 = \dfrac{1}{2}, \\ 4\lambda_1\xi^2 < 0\,, & \lambda_1 = -\dfrac{1}{2}. \end{cases}$$

So hat man in $(-1,1,0)$ mit einem Minimum bzw. in $(1,-1,2)$ mit einem Maximum zu tun.

Durch Einführung der sog. erweiterten Lagrange-Funktion

$$\widetilde{L}(\mathbf{r}, \boldsymbol{\lambda}) := f(\mathbf{r}) + \langle \boldsymbol{\lambda}, \mathbf{g}(\mathbf{r})\rangle, \qquad (\mathbf{r}, \boldsymbol{\lambda}) \in \Omega \times \mathbb{R}^m$$

hat die notwendige Bedingung erster Ordnung die Form $\widetilde{L}'(\mathbf{c}, \boldsymbol{\lambda}_{\mathbf{c}}) = \mathbf{0}$. Das Vektorpaar $(\mathbf{c}, \boldsymbol{\lambda}_{\mathbf{c}})$ genügt also dem Gleichungssystem (1). Ferner gilt für die zweite Ableitung

$$\widetilde{L}''(\mathbf{c}, \boldsymbol{\lambda}_{\mathbf{c}}) = \begin{bmatrix} L''(\mathbf{c}) & \mathbf{g}'(\mathbf{c})^T \\ \mathbf{g}'(\mathbf{c}) & \mathbf{O}_m \end{bmatrix},$$

wobei $\mathbf{O}_m$ die Nullmatrix mit m Spalten und m Zeilen bezeichnet. Die Matrix $\widetilde{L}''(\mathbf{c}, \boldsymbol{\lambda}_{\mathbf{c}})$ ist nicht definit, da sie mindestens eine Null in der Hauptdiagonale hat. Wir zeigen aber, daß sie eine entscheidende Rolle bei der Bestimmung der Definitheit $L''(\mathbf{c})$ bezüglich $\mathbf{g}'(\mathbf{c})$ spielt.

Durch Einführung der Matrizen

$$\mathbf{R}_1 := [x_1, \ldots, x_{n-m}]^T \quad \text{und} \quad \mathbf{R}_2 := [x_{n-m+1}, \ldots, x_n]^T$$

läßt sich die Bedingung $\mathbf{g}'(\mathbf{c})\mathbf{r} = \mathbf{0}$ in der Form $G_1\mathbf{R}_1 + G_2\mathbf{R}_2 = \mathbf{0}$ aufschreiben, wobei G_1 und G_2 die im Beweis des Satzes 2.3 eingeführten Matrizen sind. Unter Berücksichtigung der Annahme $\det(G_2) \neq 0$ folgt

$$\mathbf{R}_2 = -G_2^{-1} \cdot G_1 \cdot \mathbf{R}_1, \quad \text{d. h.} \quad \mathbf{r} = \begin{bmatrix} E_{n-m} \\ -G_2^{-1} \cdot G_1 \end{bmatrix} \cdot \mathbf{R}_1 =: M \cdot \mathbf{R}_1,$$

wobei E_{n-m} die Einheitsmatrix mit $(n-m)$ Spalten und $(n-m)$ Zeilen ist. So hat die quadratische Form $Q_{\mathbf{c}}^{L}$ die Gestalt

$$\langle L''(\mathbf{c})\mathbf{r}, \mathbf{r}\rangle = \mathbf{R}_1^T \cdot M^T \cdot L''(\mathbf{c}) \cdot M \cdot \mathbf{R}_1, \qquad \mathbf{r} \in \mathbb{R}^n.$$

Die positive bzw. negative Definitheit von $Q_{\mathbf{c}}^{L}$ bezüglich $\mathbf{g}'(\mathbf{c})$ ist also mit der (bedingungslosen) positiven bzw. negativen Definitheit der Matrix

$$\widehat{Q}_{\mathbf{c}}^{L} := M^T \cdot L''(\mathbf{c}) \cdot M \in \mathbb{R}^{(n-m)\times(n-m)} \tag{4}$$

gleichwertig.

Beispiel 3.8
Im Fall $m = 1, n = 2$ haben $L''(\mathbf{c})$, $g'(\mathbf{c})$ und somit M die Form

$$\begin{aligned} L''(\mathbf{c}) &= \begin{bmatrix} \partial_{11} f(\mathbf{c}) + \lambda \partial_{11} g(\mathbf{c}) & \partial_{12} f(\mathbf{c}) + \lambda \partial_{12} g(\mathbf{c}) \\ \partial_{21} f(\mathbf{c}) + \lambda \partial_{21} g(\mathbf{c}) & \partial_{22} f(\mathbf{c}) + \lambda \partial_{22} g(\mathbf{c}) \end{bmatrix}, \\ g'(\mathbf{c}) &= [\partial_1 g(\mathbf{c}), \partial_2 g(\mathbf{c})] \end{aligned}$$

und $M = \left[1, -\frac{\partial_1 g(\mathbf{c})}{\partial_2 g(\mathbf{c})}\right]^T$. Demzufolge gilt für $M^T \cdot L''(\mathbf{c}) \cdot M$ in diesem Fall

$$M^T \cdot L''(\mathbf{c}) \cdot M = \frac{ae^2 - 2bde + cd^2}{e^2}$$

mit

$$\begin{aligned} a &:= \partial_{11} f(\mathbf{c}) + \lambda \partial_{11} g(\mathbf{c}), \quad b := \partial_{12} f(\mathbf{c}) + \lambda \partial_{12} g(\mathbf{c}) \\ c &:= \partial_{22} f(\mathbf{c}) + \lambda \partial_{22} g(\mathbf{c}), \quad d := \partial_1 g(\mathbf{c}), \quad e := \partial_2 g(\mathbf{c}). \end{aligned}$$

Somit hat $M^T \cdot L''(\mathbf{c}) \cdot M$ wegen der Identität (2) die Form

$$M^T \cdot L''(\mathbf{c}) \cdot M = -\frac{1}{[\partial_2 g(\mathbf{c})]^2} \cdot D(\mathbf{c}),$$

wobei

$$D(\mathbf{c}) := \det \begin{bmatrix} \partial_{11} f(\mathbf{c}) + \lambda \partial_{11} g(\mathbf{c}) & \partial_{12} f(\mathbf{c}) + \lambda \partial_{12} g(\mathbf{c}) & \partial_1 g(\mathbf{c}) \\ \partial_{21} f(\mathbf{c}) + \lambda \partial_{21} g(\mathbf{c}) & \partial_{22} f(\mathbf{c}) + \lambda \partial_{22} g(\mathbf{c}) & \partial_2 g(\mathbf{c}) \\ \partial_1 g(\mathbf{c}) & \partial_2 g(\mathbf{c}) & 0 \end{bmatrix}. \tag{5}$$

So hat f ein lokales Maximum bzw. Minimum in $\mathbf{c}$ unter der Nebenbedingung $\mathbf{g} = \mathbf{0}$, wenn $D(\mathbf{c}) > 0$ bzw. $D(\mathbf{c}) < 0$ gilt.

Dieses Ergenbis könnte z. B. bei der Lösung der Aufgabe 2.1 nützlich sein. Dort hat nämlich die Lagrange-Funktion die Form

$$L(\alpha, \beta) = \frac{\beta + b}{2} \cdot b \cdot \sin(\alpha) + \lambda\Big(\frac{\beta - b}{2} - b \cdot \cos(\alpha)\Big), \qquad 0 < \alpha < \pi,\ 0 < \beta < 3b.$$

Die Rangbedingung ist überall erfüllt, denn $2g'(\alpha, \beta) \equiv (2b\sin(\alpha), 1)$, und die einzige kritische Stelle ist $\mathbf{c} = (\frac{\pi}{3}, 2b)$ mit $\lambda = \frac{-b\sqrt{3}}{2}$. Die zweite Ableitung der Lagrange-Funktion ist indefinit:

$$L''(\mathbf{c}) = \begin{bmatrix} \frac{-b^2\sqrt{3}}{2} & \frac{b}{4} \\ \frac{b}{4} & 0 \end{bmatrix}.$$

Durch die Berechnung der Determinante in (5) bekommt man

$$\det \begin{bmatrix} \dfrac{-b^2\sqrt{3}}{2} & \dfrac{b}{4} & \dfrac{b\sqrt{3}}{2} \\ \dfrac{b}{4} & 0 & \dfrac{1}{2} \\ \dfrac{b\sqrt{3}}{2} & \dfrac{1}{2} & 0 \end{bmatrix} = \frac{b^2\sqrt{3}}{4} > 0,$$

so gibt es in der gefundenen Stelle ein Maximum.

Die Verallgemeinerung auf die Fälle $m, n \in \mathbb{N}$ mit $1 \leq m < n$ wird formuliert im

Satz 3.9
Für $k \in \{1, \ldots, m+n\}$ bezeichne H_k die Untermatrix der geränderten Hesse-Matrix

$$H := \begin{bmatrix} L''(\mathbf{c}) & \mathbf{g}'(\mathbf{c})^T \\ \mathbf{g}'(\mathbf{c}) & \mathbf{O}_m \end{bmatrix},$$

die aus den Einträgen der ersten k Zeilen und Spalten besteht, und die Voraussetzungen des Satzes 3.7 seien erfüllt. So hat f ein lokales Minimum bzw. Maximum in $\mathbf{c}$ unter der Nebenbedingung $\mathbf{g} = \mathbf{0}$, wenn die Ungleichungen

$$(-1)^m \det(H_k) > 0 \quad \textit{bzw.} \quad (-1)^{m+k} \det(H_k) > 0, \qquad k \in \{2m+1, \ldots, m+n\}$$

gelten.

Beweis. Da die Matrix in (4) über die Form

$$\widehat{Q}^L_{\mathbf{c}} = \begin{bmatrix} E_{n-m} & (-G_2^{-1}G_1)^T \end{bmatrix} \cdot L''(\mathbf{c}) \cdot \begin{bmatrix} E_{n-m} \\ -G_2^{-1}G_1 \end{bmatrix}$$

verfügt, ist es zweckmäßig die Hesse-Matrix $L''(\mathbf{c})$ als Übermatrix der Blöcke L_{ij}, $i, j \in \{1; 2\}$ aufzufassen:

$$L''(\mathbf{c}) =: \begin{array}{r} n-m\left\{\right. \\ m\left\{\right. \end{array} \overbrace{\underbrace{\begin{matrix} \boxed{L_{11}} \\ \boxed{L_{21}} \end{matrix}}_{n-m}}^{n-m} \; \overbrace{\underbrace{\begin{matrix} \boxed{L_{12}} \\ \boxed{L_{22}} \end{matrix}}_{m}}^{m} \begin{array}{l} \left.\right\} n-m \\ \left.\right\} m \end{array},$$

wobei $L_{21} = L_{12}^T$. So läßt sich die geränderte Hesse-Matrix in Blockgestalt

$$H = \begin{bmatrix} L_{11} & L_{12} & G_1^T \\ L_{21} & L_{22} & G_2^T \\ G_1 & G_2 & \mathbf{O}_m \end{bmatrix}$$

und $\widehat{Q}^L_{\mathbf{c}}$ in die Gestalt

$$\begin{aligned}\widehat{Q}^L_{\mathbf{c}} &= L_{11} - L_{12}(G_2^{-1}G_1) - (G_2^{-1}G_1)^T L_{21} + (G_2^{-1}G_1)^T L_{22}(G_2^{-1}G_1)\\ &= L_{11} - L_{12}(G_2^{-1}G_1) - G_1^T(G_2^{-1})^T L_{21} + G_1^T(G_2^T)^{-1} L_{22}(G_2^{-1}G_1)\\ &= L_{11} - L_{12}(G_2^{-1}G_1) - G_1^T(G_2^T)^{-1}\{L_{21} - L_{22}(G_2^{-1}G_1)\}\end{aligned}$$

schreiben. Laut Determinantensatz für Übermatrizen (vgl. [8]) gilt also

$$\begin{aligned}\det(\widehat{Q}^L_{\mathbf{c}}) &= \frac{1}{\det(G_2)} \cdot \det \begin{bmatrix} L_{11} - L_{12}(G_2^{-1}G_1) & G_1^T \\ L_{21} - L_{22}(G_2^{-1}G_1) & G_2^T \end{bmatrix}\\ &= \frac{1}{\det(G_2)} \cdot \frac{1}{\det(G_2)} \cdot \det \begin{bmatrix} L_{11} - L_{12}G_2^{-1}G_1 & G_1^T & L_{12} \\ L_{21} - L_{22}G_2^{-1}G_1 & G_2^T & L_{22} \\ \mathbf{O}_{n-m} & \mathbf{O}_m & G_2 \end{bmatrix}\\ &= \frac{\det(A)}{(\det(G_2))^2}\end{aligned}$$

mit

$$A := \begin{bmatrix} L_{11} & G_1^T & L_{12} \\ L_{21} & G_2^T & L_{22} \\ G_1 & \mathbf{O}_m & G_2 \end{bmatrix}.$$

So sind das $(\det(G_2))^2$-fache der Hauptminoren von $\widehat{Q}^L_{\mathbf{c}}$ der Ordnung k gleich den Hauptminoren von A der Ordnung $2m + k$, $k \in \{1, \ldots, n - m\}$. $\widehat{Q}^L_{\mathbf{c}}$ ist also genau dann z. B. positiv definit, wenn diese Hauptminoren von A positiv sind. Diese Hauptminoren sind aber gleich dem $(-1)^m$-fachen der entsprechenden Hauptminoren von H, denn A bekommt man von H durch die Vertauschung der $(n-m+l)$-ten und der $(n + 1 + l)$-ten Spalten, $l \in \{1, \ldots, m\}$.

Beispiel 3.10
Man rechnet leicht nach, daß die geränderte Hesse-Matrix im Falle der Aufgabe 3.4 die Form

$$H := \begin{bmatrix} 0 & -3 & -3 & -3 & 1 \\ -3 & 0 & -3 & -3 & 1 \\ -3 & -3 & 0 & -3 & 1 \\ -3 & -3 & -3 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{bmatrix} \tag{6}$$

hat. Für das Vorzeichen der entsprechenden Hauptminoren von H gilt wie folgt

$$(-1)^1 \det(H_3) > 0, \quad (-1)^1 \det(H_4) > 0 \quad \text{bzw.} \quad (-1)^1 \det(H_5) > 0,$$

so haben wir mit einem lokalen Minimum unter Nebenbedingungen zu tun.

Das im Satz 3.9 formulierte Kriterium war zuerst in [12] zu lesen. Etwas einfachere Beweismethode findet sich in [16], woraus die Idee der obigen Herleitung stammt. Das Besondere an diesem Kriterium war, daß man durch die Berechnung der Hauptminoren viel einfacher vorgehen konnte als bei dem früher bekannten Kriterium von H. Hancock (vgl. [9]), nach der die positive bzw. negative Definitheit von $Q^L_{\mathbf{c}}$ begüglich $\mathbf{g}'(\mathbf{c})$ damit gleichwertig ist, daß die Koeffizienten $a_0, \ldots, a_{n-m}$ des Polynoms

$$\det \begin{bmatrix} L''(\mathbf{c}) - zE_n & \mathbf{g}'(\mathbf{c})^T \\ \mathbf{g}'(\mathbf{c}) & \mathbf{O}_m \end{bmatrix} = \sum_{k=0}^{n-m} a_k(-z)^{n-m-k}, \qquad z \in \mathbb{R} \tag{7}$$

das gleiche Vorzeichen haben bzw. alternieren.

Ein anderes Kriterium haben Chabriallac und Crouzeix bewiesen (vgl. [2]), nach der die positive bzw. negative Definitheit der Hesse-Matrix $L''(\mathbf{c})$ bezüglich $\mathbf{g}'(\mathbf{c})$ damit gleichwertig ist, daß die Trägheit der geränderten Hessematrix H (also das Tripel, bestehend aus der Anzahl der negativen, nullgleich bzw. positiven Eigenwerten) gleich $(m, 0, n)$ bzw. $(n, 0, m)$ ist. So sieht man wiederum, daß im Falle der Aufgabe 3.4 ein lokales Minimum vorliegt, indem man das Polynom (7) für die geränderte Hesse-Matrix H in (6) aufschreibt:

$$-4(-z)^3 - 36(-z)^2 - 108(-z) - 108, \qquad z \in \mathbb{R}.$$

Dies beweist auch die Trägheit der Matrix H: $(1, 0, 4)$, denn die Eigenwerte von H sind: $-4 - 2\sqrt{5}$, $-4 + 2\sqrt{5}$, 3, 3, 3.

Abgesehen davon, daß es immer die Frage der Beurteilung im Einzelfall ist, welche dieser Kriterien einfacher sind, zeichnet sich der Chabriallac-Crouzeix-Test dadurch aus, daß er gut auf einem Rechner implementierbar ist. Die Berechnung der Trägheit einer Matrix kann nämlich z. B. mit Hilfe des Rangreduktionsverfahrens von J. Egerváry (s. [3], [4]) durchgeführt werden. Eine gute Übersicht über die verschiedenen Rangreduktionsverfahren findet man z. B. in den Werken [6], [7].

References

[1] M. Barner, F. Flohr, *Analysis II*, Walter de Gruyter&Co, Berlin, 1989.

[2] Y. Chabrillac, J.-P. Crouzeix, *Definiteness and semidefiniteness of quadratic forms revisited*, Linear Algebra Appl. **63** (1984), 283–292.

[3] E. Egerváry, *Über die Faktorisation von Matrizen und ihre Anwendung auf die Lösung von linearen Gleichungssystemen*, Z. Angew. Math. Mech. **35** (1955), 111–118.

[4] E. Egerváry, *On rank-diminishing operations and their applications to the solution of linear equations*, Z. Angew. Math. Phys. **11** (1960), 376–386.

[5] M. Farkas, *On the conditional extremum*, Mat. Lapok **24** (1973), 113–129.

[6] A. Galántai, *Rank reduction: theory and applications*, Int. J. Math. Game Theory Algebra **13** (2003), 173–189.

[7] A. Galántai, *The rank reduction procedure of Egerváry*, CEJOR Cent. Eur. J. Oper. Res. **18** (2010), 5–24.

[8] F.R. Gantmacher, *Matrizentheorie*, VEB Deutscher Verlag der Wissenschaften, Berlin, 1986.

[9] H. Hancock, *Theory of maxima and minima*, Boston, New York, etc: Ginn and Company, 1917.

[10] M.R. Hestenes, *Calculus of variations and optimal control theory*. Robert E. Krieger Publishing Co., Inc., Huntington, N.Y., 1980.

[11] H. Heuser, *Lehrbuch der Analysis* (Teil 2) B. G. Teubner, Stuttgart, 1991.

[12] H.B. Mann, *Quadratic Forms with Linear Constraints*, Amer. Math. Monthly **50** (1943), 430–433.

[13] K. Meyberg, P. Vachenauer, *Höhere Mathematik 1*, Berlin, New York, etc: Springer Verlag, 1990.

[14] J. Pál, F. Schipp, P. Simon, *Analysis II*, Budapest: Lehrbuchverlag, 1990.

[15] P. Simon, *Kapiteln aus der Analysis*, Budapest: Eötvös-Kiadó, 2006.

[16] H. Väliaho, *On the definity of quadratic forms subject to linear constrains*, J. Optim. Theory Appl. **38** (1982), 143–145.

Lehrstuhl für Numerische Analysis der
Eötvös-Loránd-Universität,
H-1117 Budapest, Pázmány Péter sétány 1/C, Ungarn
e-Post: alex@ludens.elte.hu

Received: 12 April 2011; final version: 22 October 2011;
available online: 2 January 2012.