7 Optimalitätsbedingungen für Aufgaben mit Restriktionen

Im folgenden wollen wir notwendige und hinreichende Bedingungen für finite Optimierungsaufgaben mit Restriktionen betrachten. Gegeben sei also eine Optimierungsaufgabe des Typs, wie wir ihn schon im ersten Abschnitt angegeben haben:

Minimiere eine Funktion

$$f(\boldsymbol{x}), \ \ \boldsymbol{x} \in \mathbb{R}^n \tag{7.1}$$

unter den Nebenbedingungen

$$\begin{array}{ll} c_i(\boldsymbol{x}) = 0 & (i \in E := \{1, \ldots, n'\}) \\ c_i(\boldsymbol{x}) \geq 0 & (i \in I := \{n'+1, \ldots, n''\}). \end{array} \tag{7.2}$$

Wieder setzen wir voraus, daß die Funktionen $f,\ c_i : \mathbb{R}^n \longrightarrow \mathbb{R}$ hinreichend oft stetig differenzierbar sind.

A. Gleichungsnebenbedingungen

Zunächst betrachten wir den Spezialfall, daß nur Gleichungsnebenbedingungen auftreten, daß also $I = \emptyset$, und $n'' = n' =: m$ gilt. Die Menge der zulässigen Punkte ist dann gegeben durch $M := \boldsymbol{c}^{-1}(0)$ mit $\boldsymbol{c} := (c_1, \ldots, c_m)^{\mathrm{T}}$.

Wir wollen voraussetzen, daß die Gleichungsnebenbedingungen unabhängig voneinander sind. Mathematisch führt dies auf folgende Definition:

Definition (7.3)

Ein zulässiger Punkt $\boldsymbol{x} \in M$ heißt **regulär**, falls die Gradienten der Nebenbedingungen $\boldsymbol{\nabla} c_1(\boldsymbol{x}), \ldots, \boldsymbol{\nabla} c_m(\boldsymbol{x})$ linear unabhängig sind.

Die Menge aller regulären Punkte ist eine in M offene Menge. Sind alle Punkte von M regulär, so ist M eine $(n-m)$-dimensionale Untermannigfaltigkeit des $\mathbb{R}^n$. Der Tangentialraum $T_{\boldsymbol{x}_0}M$ an diese Mannigfaltigkeit im Punkte $\boldsymbol{x}_0 \in M$ ist dann gegeben durch

$$T_{\boldsymbol{x}_0}M \ := \ \{\boldsymbol{y} \in \mathbb{R}^n \,|\, \forall\, i \,:\, \boldsymbol{\nabla} c_i(\boldsymbol{x}_0)^{\mathrm{T}} \boldsymbol{y} \ = \ \boldsymbol{0}\,\}\,. \tag{7.4}$$

Der Tangentialraum ist damit ein $(n-m)$-dimensionaler linearer Teilraum des $\mathbb{R}^n$.

Man kann den Tangentialraum auch folgendermaßen geometrisch beschreiben: $T_{\boldsymbol{x}_0}M$ ist die Menge aller Geschwindigkeitsvektoren $\dot{\boldsymbol{x}}(0)$ von (differenzierbaren) Wegen $\boldsymbol{x} :]-\delta, \delta[\longrightarrow M$ in M, die zur Zeit $t = 0$ durch der Punkt $\boldsymbol{x}_0$ laufen, also $\boldsymbol{x}(0) = \boldsymbol{x}_0$. Zum Beweis der Äquivalenz läßt sich der Satz über implizite Funktionen verwenden; wir verweisen hier auf Lehrbücher über Analysis. Aus dem Satz über implizite Funktionen folgt zugleich, daß sich die Wege $\boldsymbol{x}$ auch als hinreichend oft stetig differenzierbar wählen lassen, sofern die Funktionen c_i ebenfalls entsprechende Differenzierbarkeitseigenschaften besitzen.

Mit dieser anschaulichen Vorstellung der Menge der zulässigen Punkte lassen sich nun folgende notwendige Bedingungen beweisen.

Satz (7.5) (Notwendige Bedingungen)

Ist $\boldsymbol{x}^*$ *ein lokales Minimum von* f *auf* M *und ist* $\boldsymbol{x}^*$ *ein regulärer Punkt, so gibt es (eindeutig bestimmte)* **Lagrange-Multiplikatoren** $\lambda_1, \ldots, \lambda_{n'} \in \mathbb{R}$ *so daß für die* **erweiterte Funktion** (**Lagrange-Funktion**) $F(\boldsymbol{x}) := f(\boldsymbol{x}) - \sum_{i \in E} \lambda_i \, c_i(\boldsymbol{x})$ *gelten:*

a) $\boldsymbol{x}^*$ *ist stationärer Punkt der Lagrange-Funktion:* $\boldsymbol{\nabla} F(\boldsymbol{x}^*) = \boldsymbol{0}$,

b) die Hesse-Matrix der Lagrange-Funktion ist auf dem Tangentialraum positiv semidefinit, d.h. $\forall\, \boldsymbol{y} \in T_{\boldsymbol{x}^*}M \, : \, \boldsymbol{y}^{\mathrm{T}} \, \boldsymbol{\nabla}^2 F(\boldsymbol{x}^*) \, \boldsymbol{y} \geq 0$.

Beweis:

zu a) Zu einem Tangentialvektor $\boldsymbol{y} \in T_{\boldsymbol{x}_0}M$ sei $\boldsymbol{x} :]-\delta, \delta[\longrightarrow M$ ein Weg mit $\boldsymbol{x}(0) = \boldsymbol{x}_0$ und $\dot{\boldsymbol{x}}(0) = \boldsymbol{y}$. Da nun nach Voraussetzung $f(\boldsymbol{x}(t))$ in $t = 0$ minimal ist, folgt mit Satz (1.8) a)

$$\frac{d}{d\,t} \, f(\boldsymbol{x}(t))|_{t=0} \; = \; \boldsymbol{\nabla} f(\boldsymbol{x}^*)^{\mathrm{T}} \, \boldsymbol{y} \; = \; 0 \, ,$$

d.h. $\boldsymbol{\nabla} f(\boldsymbol{x}^*)$ steht senkrecht auf dem Tangentialraum an M in $\boldsymbol{x}_0$, ist also nach (7.4) als Linearkombination der Vektoren $\boldsymbol{\nabla} c_1(\boldsymbol{x}), \ldots, \boldsymbol{\nabla} c_{n'}(\boldsymbol{x})$ darstellbar.

zu b) Weiter folgt aus der Tatsache, daß $t = 0$ ein lokales Minimum der Funktion $f(\boldsymbol{x}(t))$ ist, nach Satz (1.8) b):

$$\frac{d^2}{d\,t^2} \, f(\boldsymbol{x}(t))|_{t=0} \; = \; \boldsymbol{y}^{\mathrm{T}} \, \boldsymbol{\nabla}^2 f(\boldsymbol{x}^*) \, \boldsymbol{y} \; + \; \boldsymbol{\nabla} f(\boldsymbol{x}^*)^{\mathrm{T}} \, \ddot{\boldsymbol{x}}(0) \; \geq \; 0 \, . \tag{7.6}$$

Aus $\sum_{i \in E} \lambda_i \, c_i(\boldsymbol{x}(t)) \equiv 0$ folgt weiterhin durch zweimaliges Differenzieren und Einsetzen von $t = 0$:

$$\boldsymbol{y}^{\mathrm{T}} \left(\sum_{i \in E} \lambda_i \, \boldsymbol{\nabla}^2 c_i(\boldsymbol{x}^*) \right) \boldsymbol{y} \; + \; \left(\sum_{i \in E} \lambda_i \, \boldsymbol{\nabla} c_i(\boldsymbol{x}^*) \right)^{\mathrm{T}} \ddot{\boldsymbol{x}}(0) \; = \; 0 \, . \tag{7.7}$$

Durch Subtraktion dieser Gleichung von (7.6) und Verwendung von a) erhält man hieraus die Behauptung. □

In den meisten Fällen bilden die notwendigen Bedingungen des Satzes (7.5) - auch als **Lagrange-Multiplikatorenregel** bezeichnet - den Schlüssel zur Bestimmung der gesuchten lokalen (oder globalen) Minima. Die Untersuchung hinreichender Bedingungen bzw. der notwendigen Bedingungen zweiter Ordnung (Bedingung (7.5) b)) lassen sich dabei mitunter aufgrund von Kompaktheitsargumenten vermeiden. Wir betrachten dazu ein einfaches Beispiel:

Beispiel (7.8)

Gesucht seien die Extrema der Funktion $f(x,y) := x\,y$ auf der Kreisscheibe

$$K := \left\{(x,y)^{\mathrm{T}} :\ x^2 + y^2 \ \leq\ 1\right\}.$$

Zunächst folgt aus der Stetigkeit von f und der Kompaktheit der Menge K, daß es globale Maxima und Minima geben muß.

Im Innern $K^0 = \left\{(x,y,)^{\mathrm{T}} : x^2 + y^2 < 1\right\}$ hat man die notwendige Bedingung $\boldsymbol{\nabla} f = (y,x)^{\mathrm{T}} = \mathbf{0}$, die nur von $\boldsymbol{x}^0 = \mathbf{0}$ erfüllt wird. Da jedoch die Hesse-Matrix

$$\boldsymbol{\nabla}^2 f(\mathbf{0}) \ = \ \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

indefinit ist, ist $\boldsymbol{x}^0$ lediglich ein Sattelpunkt von f.

Für den Rand $c(x,y) := 1 - x^2 - y^2 = 0$ bilden wir die Lagrange-Funktion

$$F(x,y) \ := \ x\,y \ - \ \lambda\,(1 - x^2 - y^2).$$

Die notwendige Bedingung (7.5) a) liefert das (nichtlineare) Gleichungssystem

$$\begin{aligned} y \ + \ 2\,\lambda\,x \ &= \ 0 \\ x \ + \ 2\,\lambda\,y \ &= \ 0 \\ x^2 \ + \ y^2 \ &= \ 1 \end{aligned}$$

mit den vier Lösungen:

$$\begin{aligned} \lambda_{1/2} \ &= \quad 0.5 \ : \ \boldsymbol{x}^{(1)} \ = \ \left(\sqrt{0.5}, -\sqrt{0.5}\right), \quad \boldsymbol{x}^{(2)} \ = \ \left(-\sqrt{0.5}, \quad \sqrt{0.5}\right) \\ \lambda_{3/4} \ &= \ -0.5 \ : \ \boldsymbol{x}^{(3)} \ = \ \left(\sqrt{0.5}, \quad \sqrt{0.5}\right), \quad \boldsymbol{x}^{(4)} \ = \ \left(-\sqrt{0.5}, -\sqrt{0.5}\right). \end{aligned}$$

Aufgrund der Funktionswerte $f(\boldsymbol{x}^{(1)}) = f(\boldsymbol{x}^{(2)}) = -0.5$ und $f(\boldsymbol{x}^{(3)}) = f(\boldsymbol{x}^{(4)}) = 0.5$ sind folglich $\boldsymbol{x}^{(1)}$ und $\boldsymbol{x}^{(2)}$ globale Minima und $\boldsymbol{x}^{(3)}$ und $\boldsymbol{x}^{(4)}$ globale Maxima von $f(x,y)$ auf K .

Die notwendige Bedingung (7.5)b) braucht zur Lösung des Beispiels also garnicht herangezogen zu werden. Untersucht man denoch die Definitheit der Hesse-Matrix $\boldsymbol{\nabla}^2 F(\boldsymbol{x})$ auf den jeweiligen Tangentialräumen, so findet man, daß sogar die folgenden hinreichenden Bedingungen erfüllt sind.

Satz (7.9) (Hinreichende Bedingungen)

Existieren zu einem Punkt $\boldsymbol{x}^* \in M$ *Lagrange-Multiplikatoren* $\lambda_1, \ldots, \lambda_m \in \mathbb{R}$ *so daß die Lagrange-Funktion* $F(\boldsymbol{x}) := f(\boldsymbol{x}) - \sum_{i \in E} \lambda_i c_i(\boldsymbol{x})$ *die folgenden Bedingungen erfüllt:*

a) $\boldsymbol{\nabla} F(\boldsymbol{x}^*) = \boldsymbol{0}$,

b) die Hesse-Matrix der Lagrange-Funktion ist auf dem Tangentialraum positiv definit, d.h. $\forall \boldsymbol{y} \in T_{\boldsymbol{x}^*} M$, $\boldsymbol{y} \neq \boldsymbol{0}$: $\boldsymbol{y}^{\mathrm{T}} \boldsymbol{\nabla}^2 F(\boldsymbol{x}^*) \boldsymbol{y} > 0$,

so ist $\boldsymbol{x}^*$ *ein striktes lokales Minimum von* f *auf* M.

Beweis:

Wäre $\boldsymbol{x}^*$ kein striktes lokales Minimum von f auf M, so gäbe es eine Folge von Punkten $\boldsymbol{x}^k \in M$, $\boldsymbol{x}^k \neq \boldsymbol{x}^*$, mit $\boldsymbol{x}^k \longrightarrow \boldsymbol{x}^*$ $(k \longrightarrow \infty)$ und $f(\boldsymbol{x}^k) \leq f(\boldsymbol{x}^*)$ für alle k. Durch Übergang auf eine Teilfolge läßt sich nun erreichen, daß folgender Grenzwert existiert:

$$\lim_{k \to \infty} \frac{\boldsymbol{x}^k - \boldsymbol{x}^*}{\|\boldsymbol{x}^k - \boldsymbol{x}^*\|} =: \boldsymbol{y}\,.$$

Offenbar ist $\boldsymbol{y}$ ein Tangenteneinheitsvektor: $\boldsymbol{y} \in T_{\boldsymbol{x}^*} M$, $\|\boldsymbol{y}\| = 1$. Dies sieht man etwa durch Taylor-Entwicklung der Nebenbedingungen c_i:

$$c_i(\boldsymbol{x}^k) = c_i(\boldsymbol{x}^*) + \boldsymbol{\nabla} c_i(\boldsymbol{x}^*)^{\mathrm{T}} (\boldsymbol{x}^k - \boldsymbol{x}^*) + o(\|\boldsymbol{x}^k - \boldsymbol{x}^*\|)\,.$$

Wegen $c_i(\boldsymbol{x}^k) = c_i(\boldsymbol{x}^*) = 0$ folgt hieraus nämlich $\boldsymbol{\nabla} c_i(\boldsymbol{x}^*)^{\mathrm{T}} \boldsymbol{y} = 0$, für alle $i = 1, \ldots, n'$.

Weiter liefert nun der Taylorsche Satz für die Lagrange-Funktion F:

$$\begin{aligned}
f(\boldsymbol{x}^*) &\geq f(\boldsymbol{x}^k) \\
&= F(\boldsymbol{x}^k) \\
&= F(\boldsymbol{x}^*) + \boldsymbol{\nabla} F(\boldsymbol{x}^*)^{\mathrm{T}} (\boldsymbol{x}^k - \boldsymbol{x}^*) \\
&+ \frac{1}{2} (\boldsymbol{x}^k - \boldsymbol{x}^*)^{\mathrm{T}} \boldsymbol{\nabla}^2 F(\boldsymbol{x}^k + \theta_k(\boldsymbol{x}^k - \boldsymbol{x}^*))\, (\boldsymbol{x}^k - \boldsymbol{x}^*) \\
&= f(\boldsymbol{x}^*) + \frac{1}{2} (\boldsymbol{x}^k - \boldsymbol{x}^*)^{\mathrm{T}} \boldsymbol{\nabla}^2 F(\boldsymbol{x}^k + \theta_k(\boldsymbol{x}^k - \boldsymbol{x}^*))\, (\boldsymbol{x}^k - \boldsymbol{x}^*)\,,
\end{aligned}$$

wobei $0 < \theta_k < 1$. Hieraus folgt

$$0 \geq (\boldsymbol{x}^k - \boldsymbol{x}^*)^{\mathrm{T}} \boldsymbol{\nabla}^2 F(\boldsymbol{x}^k + \theta_k(\boldsymbol{x}^k - \boldsymbol{x}^*))\, (\boldsymbol{x}^k - \boldsymbol{x}^*).$$

Dividiert man diese Ungleichung durch $\|\boldsymbol{x}^k - \boldsymbol{x}^*\|^2$ und bildet den Grenzübergang $k \to \infty$, so erhält man $\boldsymbol{y}^{\mathrm{T}} \boldsymbol{\nabla}^2 F(\boldsymbol{x}^*) \boldsymbol{y} \leq 0$, im Widerspruch zur Voraussetzung b). □

Bemerkungen

1) In Satz (7.9) wird die Regularität des strikten lokalen Minimums $\boldsymbol{x}^*$ nicht benötigt. Ist $\boldsymbol{x}^*$ jedoch ein regulärer Punkt, so sind die Lagrange-Parameter $\lambda_1, \ldots, \lambda_m$ gemäß Satz (7.5) eindeutig bestimmt.

2) Mitunter wird die erweiterte Funktion (Lagrange-Funktion) auch als Funktion von $\boldsymbol{x}$ und $\boldsymbol{\lambda} := (\lambda_1, \ldots, \lambda_m)^{\mathrm{T}}$ betrachtet und dann mit $L(\boldsymbol{x}, \boldsymbol{\lambda})$ bezeichnet. Ist $\boldsymbol{x}^*$ ein reguläres lokales Minimum von f auf M und ist $\boldsymbol{\lambda}^*$ der zugehörige Lagrange-Multiplikator gemäß Satz (7.5), so ist $(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)$ ein stationärer Punkt der Lagrange-Funktion $L(\boldsymbol{x}, \boldsymbol{\lambda})$.

3) Bilden die Spaltenvektoren einer $(n, n-m)$–Matrix $\boldsymbol{W} = (\boldsymbol{z}_1, \ldots, \boldsymbol{z}_{n-m})$ eine Basis des Tangentialraumes $T_{\boldsymbol{x}^*}M$, so sind die notwendigen bzw. hinreichenden Bedingungen zweiter Ordnung äquivalent zur Bedingung

$$\boldsymbol{R} \;:=\; \boldsymbol{W}^{\mathrm{T}}\, \boldsymbol{\nabla}^2 F(\boldsymbol{x}^*)\, \boldsymbol{W} \qquad \text{positiv (semi-)definit.}$$

Die Matrix $\boldsymbol{R}$ wird auch als **reduzierte Hesse-Matrix** bezeichnet.

4) Für praktische Anwendungen ist die folgende **Stabilitätsaussage** wichtig. Sie liefert eine unmittelbare Interpretation der Lagrange-Parameter als Empfindlichkeit des „optimalen“ Zielfunktionswertes gegenüber Störungen in den Nebenbedingungen.

Sei wieder $\boldsymbol{x}^*$ ein reguläres striktes lokales Minimum von f auf M, welches die hinreichenden Bedingungen (7.9) erfüllt. Ferner seien $\lambda_1, \ldots, \lambda_m$ die zugehörigen Lagrange-Multiplikatoren. Dann besitzt die „gestörte“ Optimierungsaufgabe

$$\begin{aligned} &\text{Minimiere} \quad f(\boldsymbol{x}),\ \ \boldsymbol{x} \in \mathbb{R}^n,\ \ \text{unter den Nebenbedingungen} \\ &c_i(\boldsymbol{x}) = \varepsilon_i \qquad (i \in E := \{1, \ldots, m\}) \end{aligned} \tag{7.10}$$

für alle hinreichend kleinen $|\varepsilon_i|$ ein (eindeutig bestimmtes) striktes und reguläres lokales Minimum $\boldsymbol{x}(\boldsymbol{\varepsilon})$ in der Nähe von $\boldsymbol{x}^*$, $\boldsymbol{\varepsilon} := (\varepsilon_1, \ldots, \varepsilon_m)^{\mathrm{T}}$. $\boldsymbol{x}(\boldsymbol{\varepsilon})$ hängt stetig differenzierbar von den $\varepsilon_1, \ldots, \varepsilon_m$ ab und es gilt

$$\lambda_i \;=\; \frac{\partial}{\partial \varepsilon_i} f(\boldsymbol{x}(\boldsymbol{\varepsilon}))|_{\boldsymbol{\varepsilon}=\boldsymbol{0}}\,. \tag{7.11}$$

Der Lagrange-Parameter λ_i beschreibt also die Empfindlichkeit der Zielfunktion gegenüber Störungen in der Nebenbedingung c_i.

B. Gleichungs- und Ungleichungsnebenbedingungen

Wir betrachten nun die allgemeine Problemstellung (7.1), (7.2), die Gleichungs- und Ungleichungsnebenbedingungen enthält. Die beteiligten Funktionen seien wieder hinreichend oft stetig differenzierbar und die Menge zulässigen Punkte M sei nichtleer.

Am Beispiel (7.8) haben wir schon gesehen, daß sich Ungleichungsnebenbedingungen im wesentlichen wie Gleichungsnebenbedingungen behandeln lassen. Ist nämlich in einem lokalen Minimum $\boldsymbol{x}^*$ die Nebenbedingung c_i positiv, also $c_i(\boldsymbol{x}^*) > 0$, so gilt dies auch in einer Umgebung von $\boldsymbol{x}^*$, so daß c_i keinen Einfluß auf die Optimalität von $\boldsymbol{x}^*$ hat und in Bezug auf die notwendigen und hinreichenden Bedingungen einfach weggelassen werden kann. Ist dagegen $c_i(\boldsymbol{x}^*) = 0$, so ist $\boldsymbol{x}^*$ zugleich ein Minimum unter der Gleichungsnebenbedingung $c_i(\boldsymbol{x}) = 0$. Es ergibt sich jedoch die Schwierigkeiten, daß der letzte Schluß nicht umkehrbar ist (ein Minimum der gleichungsrestringierten Optimierungsaufgabe muß nicht notwendig auch das Problem mit der Ungleichungsnebenbedingung lösen). Ferner weiß man a priori nicht, welche Ungleichungsnebenbedingungen in der Lösung mit Gleichheit erfüllt werden. Die folgenden Definitionen sind daher nützlich.

Definition (7.12)

a) Für einen zulässigen Punkt $\boldsymbol{x} \in M$ heißt

$$I(\boldsymbol{x}) \;:=\; \{i \in I \,:\, c_i(\boldsymbol{x}) = 0\,\}$$

die Menge der Indizes der **aktiven Ungleichungsnebenbedingungen**.

b) Ein zulässiger Punkt $\boldsymbol{x} \in M$ heißt **regulär**, wenn die Gradienten der aktiven Nebenbedingungen $\boldsymbol{\nabla} c_i(\boldsymbol{x})$, $i \in E \cup I(\boldsymbol{x})$ linear unabhängig sind.

Mit diesen Begriffen lassen sich nun notwendige Bedingungen - analog zu Satz (7.5) - formulieren.

Satz (7.13) (Notwendige Bedingungen)

Ist $\boldsymbol{x}^*$ ein lokales Minimum von f auf M und ein regulärer Punkt von M, so gibt es (eindeutig bestimmte) Lagrange-Multiplikatoren $\lambda_1, \ldots, \lambda_{n''} \in \mathbb{R}$, so daß für die erweiterte Funktion (Lagrange-Funktion) $F(\boldsymbol{x}) \;:=\; f(\boldsymbol{x}) - \sum_{i\in E} \lambda_i\, c_i(\boldsymbol{x}) - \sum_{i\in I} \lambda_i\, c_i(\boldsymbol{x})$ folgende Bedingungen gelten:

a) **Notwendige Bedingungen erster Ordnung, Kuhn-Tucker Bedingungen:**

(i) $\boldsymbol{x}^*$ ist ein stationärer Punkt der Lagrange-Funktion: $\boldsymbol{\nabla} F(\boldsymbol{x}^*) \;=\; \boldsymbol{0}$.

(ii) Alle Lagrange-Parameter zu den Ungleichungsnebenbedingungen sind nichtnegativ, $\lambda_i \geq 0,\ i \in I$.

(iii) Es gelten die **Komplementaritätsbedingungen:** $\lambda_i \cdot c_i(\boldsymbol{x}^*) \;=\; 0,\ i \in I$.

b) Die Hesse-Matrix der Lagrange-Funktion ist positiv semidefinit auf dem Tangentialxraum der aktiven Nebenbedingungen

$$T_{\boldsymbol{x}^*}M_a \;:=\; \{\boldsymbol{y} \in \mathbb{R}^n \,|\, \forall\, i \in E \cup I(\boldsymbol{x}^*) \,:\, \boldsymbol{\nabla} c_i(\boldsymbol{x}^*)^{\mathrm{T}} \boldsymbol{y} \;=\; \boldsymbol{0}\,\}\,,$$

d.h. $\forall\, \boldsymbol{y} \in T_{\boldsymbol{x}^*}M_a \,:\, \boldsymbol{y}^{\mathrm{T}}\, \boldsymbol{\nabla}^2 F(\boldsymbol{x}^*)\, \boldsymbol{y} \;\geq\; 0\,.$

Beweis:

zu a) Setzt man, wie oben angedeutet

$$M_a \;:=\; \{\boldsymbol{x} \in \mathbb{R}^n \;:\; c_i(\boldsymbol{x}) \;=\; 0,\, i \in E \cup I(\boldsymbol{x}^*)\,\}\,,$$

so ist $\boldsymbol{x}^*$ zugleich ein lokales Minimum von f auf M_a und $\boldsymbol{x}^*$ ist auch ein regulärer Punkt von M_a im Sinne von (7.3). Nach Satz (7.5) gibt es daher (eindeutig bestimmte) Lagrange-Parameter $\lambda_i,\ i \in E \cup I(\boldsymbol{x}^*)$, mit $\boldsymbol{\nabla} F(\boldsymbol{x}^*) = \boldsymbol{0}$. Hierbei wird gesetzt $\lambda_i := 0,\ i \in I \setminus I(\boldsymbol{x}^*)$. Damit sind die notwendigen Bedingungen (i) und (iii) erfüllt.

Es bleibt also zu zeigen, daß $\lambda_k \geq 0$ für alle $k \in I(\boldsymbol{x}^*)$.

Sei dazu $k \in I(\boldsymbol{x}^*)$ fest gewählt und $M_k := \{\boldsymbol{x} \in \mathbb{R}^n \;:\; c_i(\boldsymbol{x}) \;=\; 0,\, i \in E \cup I(\boldsymbol{x}^*) \setminus \{k\}\,\}$. Aufgrund der Regularität von $\boldsymbol{x}^*$ ist

$$\boldsymbol{\nabla} c_k(\boldsymbol{x}^*) \;\notin\; \mathrm{Spann}\,\{\boldsymbol{\nabla} c_i(\boldsymbol{x}^*) \;:\; i \in E \cup (I(\boldsymbol{x}^*) \setminus \{k\})\,\} \;=:\; G_k.$$

Durch Orthogonalisierung von $\boldsymbol{\nabla} c_k(\boldsymbol{x}^*)$ bzgl. des linearen Raumes G_k findet man einen Vektor $\boldsymbol{y} \in G_k^{\perp}$ mit $\boldsymbol{\nabla} c_k(\boldsymbol{x}^*)^{\mathrm{T}}\,\boldsymbol{y} \;>\; 0$.

Nun ist der Orthogonalraum von G_k zugleich der Tangentialraum an M_k im Punkte $\boldsymbol{x}^*$. Daher gibt es einen stetig differenzierbaren Weg $\boldsymbol{x}(t) \in M_k,\ -\delta < t < \delta$, mit $\boldsymbol{x}(0) = \boldsymbol{x}^*$ und $\dot{\boldsymbol{x}}(0) = \boldsymbol{y}$.

Man sieht nun, daß $\boldsymbol{x}(t)$ alle Nebenbedingungen (7.2) auf einem hinreichend kleinen (einseitigen!) Intervall $0 \leq t \leq \delta'$ erfüllt, wobei $0 < \delta' \leq \delta$.

Für die Indizes $i \in E \cup I(\boldsymbol{x}^*) \setminus \{k\}$ folgt dies per Konstruktion ($\boldsymbol{x}(t) \in M_k$), für die Indizes $i \in I \setminus I(\boldsymbol{x}^*)$ folgt dies aus Stetigkeitsgründen und schließlich für den Index k mit Hilfe des Taylorschen Satzes:

$$c_k(\boldsymbol{x}(t)) \;=\; c_k(\boldsymbol{x}^*) \;+\; t\,\boldsymbol{\nabla} c_k(\boldsymbol{x}^*)^{\mathrm{T}}\,\boldsymbol{y} \;+\; o(t)\,.$$

Damit ist aber $t^* = 0$ ein Minimum von $f(\boldsymbol{x}(t))$ auf dem Intervall $0 \leq t < \delta'$ und es folgt

$$\begin{aligned}
0 \;&\geq\; \frac{d}{d\,t} f(\boldsymbol{x}(t))|_{t=0} \\
&=\; \boldsymbol{\nabla} f(\boldsymbol{x}^*)^{\mathrm{T}}\boldsymbol{y} \\
&=\; \sum_{i\in E} \lambda_i\, \boldsymbol{\nabla} c_i(\boldsymbol{x}^*)^{\mathrm{T}}\boldsymbol{y} \;+\; \sum_{i\in I(\boldsymbol{x}^*)} \lambda_i\, \boldsymbol{\nabla} c_i(\boldsymbol{x}^*)^{\mathrm{T}}\boldsymbol{y} \\
&=\; \lambda_k\, \boldsymbol{\nabla} c_k(\boldsymbol{x}^*)^{\mathrm{T}}\boldsymbol{y}\,,
\end{aligned}$$

und damit $\lambda_k \geq 0$.

zu b) Dies folgt nach gleicher Konstruktion aus Satz (7.5)b). $\square$

Leider lassen sich die hinreichenden Bedingungen des Satzes (7.9) nicht so ohne weiteres auf den Fall von Ungleichungsnebenbedingungen übertragen. Das liegt daran, daß der Tangentialraum der aktiven Ungleichungsnebenbedingungen nicht ausreicht, um das Verhalten von f beim Übergang ins Innere des zulässigen Bereiches zu beschreiben. Dies wird an dem folgenden einfachen Beispiel deutlich:

Beispiel (7.14)

Gesucht seien die Minima der Funktion $f(x_1, x_2, x_3) := -x_1^2 + x_2^2 + x_3^2$ unter den Nebenbedingungen $c_1(\boldsymbol{x}) := x_1 \geq 0$, $c_2(\boldsymbol{x}) := x_2 \geq 0$.

Mit der Lagrange-Funktion

$$F(x_1, x_2, x_3) \;=\; -x_1^2 + x_2^2 + x_3^2 - \lambda_1\, x_1 - \lambda_2\, x_2$$

erhält man $\boldsymbol{\nabla} F = (-2x_1 - \lambda_1, 2x_2 - \lambda_2, 2\, x_3)^{\mathrm{T}}$, so daß $\boldsymbol{x}^* := \mathbf{0}$ und $\boldsymbol{\lambda} := \mathbf{0}$ die notwendigen Bedingungen erster Ordnung erfüllen. $\boldsymbol{x}^*$ ist auch ein regulärer Punkt.

Für die aktiven Indizes hat man $I(\boldsymbol{x}^*) = \{1, 2\}$ und für den Tangentialraum an die aktiven Nebenbedingungen ergibt sich

$$T_{\boldsymbol{x}^*} M_a \;=\; T_{\boldsymbol{x}^*} M \;=\; \{(0, 0, \eta)^{\mathrm{T}} \,:\, \eta \in \mathbb{R}\}\,,$$

so daß die Hesse-Matrix

$$\boldsymbol{\nabla}^2 f(\mathbf{0}) \;=\; \begin{pmatrix} -2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

tatsächlich auf dem Tangentialraum positiv definit ist.

Anderseits ist es natürlich offensichtlich, daß $\boldsymbol{x}^*$ lediglich ein Sattelpunkt für das Optimierungsproblem darstellt.

Zur Formulierung hinreichender Bedingungen muß man den Tangentialraum der aktiven Nebenbedingungen in geeigneter Weise vergrößern. Der Vollständigkeit halber sei der folgende Satz über hinreichende Bedingungen angegeben. Auf den Beweis verzichten wir hier.

Satz (7.15) (Hinreichende Bedingungen)

Existieren zu einem Punkt $\boldsymbol{x}^* \in M$ *Lagrange-Multiplikatoren* $\lambda_i, i \in E \cup I$, *mit* $\lambda_i \geq 0$ *für* $i \in I$, *so daß die Lagrange-Funktion* $F(\boldsymbol{x}) := f(\boldsymbol{x}) - \sum_{i \in E} \lambda_i\, c_i(\boldsymbol{x}) - \sum_{i \in I} \lambda_i\, c_i(\boldsymbol{x})$ *die folgenden Bedingungen erfüllt:*

a) $\boldsymbol{\nabla} F(\boldsymbol{x}^*) \;=\; \mathbf{0}\,,$

b) die Hesse-Matrix der Lagrange-Funktion ist auf dem erweiterten Tangentialraum

$$T^0_{\boldsymbol{x}^*}M \;:=\; \{\boldsymbol{y} \in \mathbb{R}^n \;:\; \boldsymbol{\nabla} c_i(\boldsymbol{x}_0)^{\mathrm{T}}\boldsymbol{y} \;=\; \boldsymbol{0}\,, \forall\, i \in E \;\; \textit{und}\; \forall\, i \in I\, \textit{mit}\, \lambda_i > 0\}$$

positiv definit, d.h. $\forall\, \boldsymbol{y} \in T^0_{\boldsymbol{x}^*}M\,,\; \boldsymbol{y} \neq \boldsymbol{0} \;:\; \boldsymbol{y}^{\mathrm{T}}\, \boldsymbol{\nabla}^2 F(\boldsymbol{x}^*)\, \boldsymbol{y} \;>\; 0\,,$

so ist $\boldsymbol{x}^*$ *ein striktes lokales Minimum von* f *auf M.*

Versucht man, diese hinreichenden Bedingungen nun auf das Beispiel (7.14) anzuwenden, so erhält man für den erweiterten Tangentialraum $T^0_{\boldsymbol{x}^*}M \;=\; \mathbb{R}^3$. Auf $\mathbb{R}^3$ ist die Hesse-Matrix des beispiels aber offensichtlich nicht positiv definit, so daß die hinreichenden Bedingungen auch - wie erwartet - nicht erfüllt sind.