0.1 Hauptsatz über implizite Funktionen

Ein lineares homogenes Gleichungssystem von q Gleichungen in $r+q$ Unbekannten kann bekanntlich verwendet werden um q Unbekannte durch die verbleibenden r Unbekannten auszudrücken falls die entsprechende $q \times q$ Untermatrix regulär ist, d.h. es gilt dann wenn das Gleichungssystem $Ax + By = 0$ durch $q \times r$ und $q \times q$ Matrizen A, B, B regulär gegeben ist eine eindutige (lineare) Lösungsfunktion nämlich $y = B^{-1}Ax$. Für nichtlineare Abbildungen ist dies im Allgemeinen klarerweise nicht möglich. Ist die Abbildung $F : \mathbb{R}^{p+q} \supseteq D \to \mathbb{R}^q$ aber in einer Umgebung von $(\mathbf{x}, \mathbf{u})$ stetig differenzierbar, so ist sie in $(\mathbf{x}, \mathbf{u})$ von höherer als erster Ordnung durch dF approximierbar und man kann hoffen, dass sich lokal um $(\mathbf{x}, \mathbf{u})$ $\mathbf{u}$ vermöge $\mathbf{F}$ darstellen lässt. Im Allgemenen ist eine explizite Darstellung nicht möglich, im Folgenden zeigen wir aber die Existenz solcher Lösungsfunktionen. Wir zeigen ihn zunächst für den Fall $q = 1$:

Satz 0.1.1 *Sei $F \in C^k(D)$, $k \geq 1$ für eine offene Teilmenge D von $\mathbb{R}^{q+1} = \mathbb{R}^q \times \mathbb{R}$. Für $(\mathbf{x}_0, u_0) \in D$ gelte $F(\mathbf{x}_0, u_0) = 0$, $\frac{\partial F}{\partial u}(\mathbf{x}_0, u_0) \neq 0$. Dann gibt es eine Umgebung V von $\mathbf{x}_0$ sowie ein $\alpha > 0$, sodass es für $\mathbf{x} \in V$ genau ein $u \in (u_0 - \alpha, u_0 + \alpha)$ mit $F(\mathbf{x}, u) = 0$ gibt.*

Fasst man dieses u als eine Funktion $g : V \to (u_0 - \alpha, u_0 + \alpha)$ auf, so ist g in V k-mal stetig differenzierbar mit

$$\frac{\partial g}{\partial x_i}(\mathbf{x}) = -\frac{\frac{\partial F}{\partial x_i}(\mathbf{x}, u)}{\frac{\partial F}{\partial u}(\mathbf{x}, u)}.$$

Beweis: Wir dürfen annehmen $\frac{\partial F}{\partial u}(\mathbf{x}_0, u_0) > 0$ (anderenfalls betrachten wir $-F$ statt F). Wegen der Stetigkeit der partiellen Ableitung $\partial/\partial u$ gibt es eine Umgebung $\tilde{V}$ von $\mathbf{x}_0$, eine Umgebung $(u_0 - \alpha, u_0 + \alpha)$ von u_0 und ein $\rho > 0$ sodass

$$\frac{\partial F}{\partial u}(\mathbf{x}, u) > \rho \tag{0.1}$$

für $(\mathbf{x}, u) \in \tilde{V} \times (u_0 - \alpha, u_0 + \alpha)$ gilt. Es folgt mit dem Mittelwertsatz

$$F(\mathbf{x}, u_0 \pm \alpha) - F(\mathbf{x}, u_0) = \pm\alpha \frac{\partial F}{\partial u}(\mathbf{x}, u_0 \pm \zeta\alpha),\ 0 < \zeta < 1.$$

Aus dem Zwischenwertsatz folgt mit (0.1), dass die Funktion $F(\mathbf{x}, \cdot)$ im Intervall $(u_0 - \alpha, u_0 + \alpha)$ alle Werte des Intervalls $(F(\mathbf{x}, u_0) - \alpha\rho, F(\mathbf{x}, u_0) + \alpha\rho)$ annimmt. Für eine geeignete offene Umgebung $V \subseteq \tilde{V}$ von $\mathbf{x}_0$ gilt wegen der Stetigkeit von F: $|F(\mathbf{x}, u_0)| < \alpha\rho$, dort gilt also

$$0 \in (F(\mathbf{x}, u_0) - \alpha\rho, F(\mathbf{x}, u_0) + \alpha\rho) \subset (F(\mathbf{x}, u_0 - \alpha), F(\mathbf{x}, u_0 + \alpha))$$

Es gibt also eine Funktion $g: V \to (u_0-\alpha, u_0+\alpha)$, die $F(\mathbf{x}, g(\mathbf{x})) = 0$ erfüllt. Wegen (0.1) ist $F(\mathbf{x}, \cdot)$ streng monoton steigend, daher ist diese Funktion eindeutig.

Es gilt für $\mathbf{x} \in V$ und hinreichend kleine $h \in \mathbb{R}$ nach dem Mittelwertsatz

$$\begin{aligned} 0 =& F(\mathbf{x}+h\mathbf{e}_i, g(\mathbf{x}+h\mathbf{e}_i)) - F(\mathbf{x}, g(\mathbf{x}+h\mathbf{e}_i)) \\ &+ F(\mathbf{x}, g(\mathbf{x}+h\mathbf{e}_i)) - F(\mathbf{x}, g(\mathbf{x})) \\ =& hF_{x_i}(\mathbf{x}+\zeta h\mathbf{e}_i, g(\mathbf{x}+h\mathbf{e}_i)) \\ &+ (g(\mathbf{x}+h\mathbf{e}_i) - g(\mathbf{x}))F_u(\mathbf{x}, g(\mathbf{x}) + \eta(g(\mathbf{x}+h\mathbf{e}_i) - g(\mathbf{x}))) \end{aligned}$$

ζ, $\eta \in (0,1)$, womit

$$\frac{g(\mathbf{x}+h\mathbf{e}_i) - g(\mathbf{x})}{h} = -\frac{F_{x_i}(\mathbf{x}+\zeta h\mathbf{e}_i, g(\mathbf{x}+h\mathbf{e}_i))}{F_u(\mathbf{x}, g(\mathbf{x}) + \eta(g(\mathbf{x}+h\mathbf{e}_i) - g(\mathbf{x})))} \tag{0.2}$$

folgt. Die Funktion $h \mapsto g(\mathbf{x}+h\mathbf{e}_i)$ ist aber stetig, denn aus dem Mittelwertsatz folgt

$$\begin{aligned} &\left|(g(\mathbf{x}+h\mathbf{e}_i) - g(\mathbf{x}))\frac{\partial F}{\partial u}(\mathbf{x}, tg(\mathbf{x}+h\mathbf{e}_i) + (1-t)g(\mathbf{x})\right| \\ =& |F(\mathbf{x}, g(\mathbf{x}+h\mathbf{e}_i)) - F(\mathbf{x}, g(\mathbf{x}))| = |F(\mathbf{x}, g(\mathbf{x}+h\mathbf{e}_i))| \\ =& |F(\mathbf{x}, g(\mathbf{x}+h\mathbf{e}_i)) - F(\mathbf{x}+h\mathbf{e}_i, g(\mathbf{x}+h\mathbf{e}_i))| \\ \le& h \mathrm{max}_{t\in[0,1]} \left|\frac{\partial F}{\partial x_i}(\mathbf{x}+th\mathbf{e}_i, g(\mathbf{x}+h\mathbf{e}_i))\right| \end{aligned}$$

und mit (0.1)

$$|g(\mathbf{x}+h\mathbf{e}_i) - g(\mathbf{x})| \le \frac{h}{\rho}\mathrm{max}_{t\in[0,1]} \left|\frac{\partial F}{\partial x_i}(\mathbf{x}+th\mathbf{e}_i, g(\mathbf{x}+h\mathbf{e}_i))\right|.$$

Für $h \to 0$ erhalten wir damit aus (0.2) die partielle Differenzierbarkeit von g mit

$$\frac{\partial g}{\partial x_i}(\mathbf{x}) = -\frac{F_{x_i}(\mathbf{x}, g(\mathbf{x}))}{F_u(\mathbf{x}, g(\mathbf{x}))}$$

und damit die stetige partielle Differenzierbarkeit also die stetige Differenzierbarkeit. Aus obiger Darstellung von $\frac{\partial g}{\partial x_i}$ folgt, dass g k-mal stetig differenzierbar ist, wenn F dieser Differenzierbarkeitsforderung genügt. $\square$

Wir verallgemeinern die Aussage dieses Satzes auf die Lösbarkeit eines Gleichungssystems

$$F_i(\mathbf{x}, \mathbf{u}) = 0, \quad \mathbf{x} \in \mathbb{R}^q, \mathbf{u} \in \mathbb{R}^r, \quad 1 \le i \le r$$

nach den Variablen u_i, d.h. wir fragen ob es lokal um einen Punkt $(\mathbf{x}_0, \mathbf{u}_0)$, der $F_i(\mathbf{x}_0, \mathbf{u}_0) = 0$ erfüllt eindeutige Funktionen $g_i: \mathbb{R}^q \to \mathbb{R}$ gibt für die $F_i(\mathbf{x}, g_1(\mathbf{x}), \dots, g_r(\mathbf{x})) = 0$ gilt.

Wir schreiben für die Ableitungsmatrizen der Funktionen $\mathbf{F} = (F_i)^t$, $\mathbf{G} = (g_i)^t$

$$\frac{\partial \mathbf{F}}{\partial \mathbf{x}} = \begin{pmatrix} \frac{\partial F_1}{\partial x_1}, & \dots, & \frac{\partial F_1}{\partial x_q} \\ \vdots & & \vdots \\ \frac{\partial F_r}{\partial x_1}, & \dots, & \frac{\partial F_r}{\partial x_q} \end{pmatrix} \in \mathbb{R}^{r \times q}, \quad \frac{\partial \mathbf{F}}{\partial \mathbf{u}} = \begin{pmatrix} \frac{\partial F_1}{\partial u_1}, & \dots, & \frac{\partial F_1}{\partial u_r} \\ \vdots & & \vdots \\ \frac{\partial F_r}{\partial u_1}, & \dots, & \frac{\partial F_r}{\partial u_r} \end{pmatrix} \in \mathbb{R}^{r \times r},$$

$$\frac{\partial \mathbf{G}}{\partial \mathbf{x}} = \begin{pmatrix} \frac{\partial g_1}{\partial x_1}, & \dots, & \frac{\partial g_1}{\partial x_q} \\ \vdots & & \vdots \\ \frac{\partial g_r}{\partial x_1}, & \dots, & \frac{\partial g_r}{\partial x_q} \end{pmatrix} \in \mathbb{R}^{r \times q}$$

Satz 0.1.2 *(Hauptsatz über implizite Funktionen) Es seien in einer offenen Teilmenge D von $\mathbb{R}^q \times \mathbb{R}^r$ C^k, $k \geq 1$, Funktionen $F_i : \mathbb{R}^{q+r} \to \mathbb{R}$ gegeben, die für einen Punkt $(\mathbf{x}_0, \mathbf{u}_0) \in D$*

$$\mathbf{F}(\mathbf{x}^0, \mathbf{u}^0) = 0, \quad \det\left(\frac{\partial \mathbf{F}}{\partial \mathbf{u}}(\mathbf{x}^0, \mathbf{u}^0)\right) \neq 0$$

erfüllen. Dann gibt es Umgebungen V von $\mathbf{x}^0$ und W von $\mathbf{u}^0$, sodass es für $\mathbf{x} \in V$ genau ein $\mathbf{u} \in W$ mit $F_i(\mathbf{x}, \mathbf{u}) = 0$ gibt. Die so definierte Lösungsfunktion $\mathbf{G}(\mathbf{x}) = (g_1(\mathbf{x}), \dots, g_q(\mathbf{x}))$ erfüllt in V also $\mathbf{F}(\mathbf{x}, \mathbf{G}(\mathbf{x})) = 0$, ist k-mal stetig differenzierbar und hat Ableitungsmatrix

$$\frac{\partial \mathbf{G}}{\partial \mathbf{x}}(\mathbf{x}) = -\left(\frac{\partial \mathbf{F}}{\partial \mathbf{u}}(\mathbf{x}, \mathbf{G}(\mathbf{x}))\right)^{-1} \frac{\partial \mathbf{F}}{\partial \mathbf{x}}(\mathbf{x}, \mathbf{G}(\mathbf{x})) \qquad (\mathbf{x}, \mathbf{G}(\mathbf{x})) \in V \times W.$$

Beweis: Wir führen den Beweis durch Induktion nach r. Für $r = 1$ ist die Aussagen Satz 0.1.1. Wir nehmen an der Satz gilt für $r - 1$ und zeigen ihn für r:

Wir dürfen o.B.d.A. (gegebenenfalls nach Permutation der Varialblen u_i) annehmen, dass

$$\begin{pmatrix} \frac{\partial F_2}{\partial u_2}, \dots, & \frac{\partial F_2}{\partial u_r} \\ \vdots & \vdots \\ \frac{\partial F_r}{\partial u_2}, \dots, & \frac{\partial F_r}{\partial u_r} \end{pmatrix} \in \mathbb{R}^{r-1 \times r-1},$$

regulär ist. Nach Induktionsvoraussetzung gibt es C^1-Funktionen $\phi_2, \dots, \phi_r$, die in einer Umgebung von $(\mathbf{x}^0, u_1^0)$

$$F_i(\mathbf{x}, u_1, \phi_2(\mathbf{x}, u_1), \dots, \phi_r(\mathbf{x}, u_1)) = 0, \qquad 2 \leq i \leq r \tag{0.3}$$

erfüllen. Wir betrachten die Funktion

$$H(\mathbf{x}, u_1) := F_1(\mathbf{x}, u_1, \phi_2(\mathbf{x}, u_1), \dots, \phi_r(\mathbf{x}, u_1)).$$

Um Satz 0.1.1 auf die Funktion H anwenden zu können müssen wir zeigen, dass $\frac{\partial H}{\partial u_1}(\mathbf{x}^0, u_1^0) \neq 0$ gilt.

Wegen

$$\frac{\partial H}{\partial u_1} = \frac{\partial F_1}{\partial u_1} + \sum_{l=2}^{r} \frac{\partial F_1}{\partial u_l}\frac{\partial \phi_l}{\partial u_1}$$

würde aus $\frac{\partial H}{\partial u_1}(\mathbf{x}^0, u_1^0) = 0$

$$\frac{\partial F_1}{\partial u_1} = -\sum_{l=2}^{r} \frac{\partial F_1}{\partial u_l}\frac{\partial \phi_l}{\partial u_1} \tag{0.4}$$

folgen. Aus (0.3) folgt aber mit der Kettenregel

$$\frac{\partial F_i}{\partial u_1} = -\sum_{l=2}^{r} \frac{\partial F_i}{\partial u_l}\frac{\partial \phi_l}{\partial u_1}, \qquad 2 \le i \le r. \tag{0.5}$$

Aus (0.4) und (0.5) würde

$$\frac{\partial \mathbf{F}}{\partial u_1} = -\sum_{l=2}^{r} \frac{\partial \mathbf{F}}{\partial u_l}\frac{\partial \phi_l}{\partial u_1}$$

folgen und der erste Spaltenvektor von $\frac{\partial \mathbf{F}}{\partial \mathbf{u}}$ wäre eine Linearkombination der anderen Spaltenvektoren, was der Regularität der Ableitungsmatrix in $\frac{\partial \mathbf{F}}{\partial \mathbf{u}}$ widerspricht.

Mit Satz 0.1.1 erhalten wir somit eine differenzierbare Funktion $g_1(\mathbf{x})$ für die, wenn wir $g_i(\mathbf{x}) := \phi_i(\mathbf{x}, g_1(\mathbf{x})),\ 2 \le i \le r$ definieren gilt

$$F_i(\mathbf{x}, g_1(\mathbf{x}), g_2(\mathbf{x}), \ldots, g_r(\mathbf{x})) = 0, \qquad 1 \le i \le r.$$

Aus der stetigen Differenzierbarkeit von g_1 und der von ϕ_i folgt dass die Funktionen $g_i,\ 1 \le i \le r$ stetig differenzierbar sind.

Durch Differentiation der Gleichungen $F_i(\mathbf{x}, G(\mathbf{x})) = 0$ erhalten wir

$$\frac{\partial \mathbf{F}}{\partial \mathbf{x}} + \frac{\partial \mathbf{F}}{\partial \mathbf{u}}\frac{\partial \mathbf{G}}{\partial \mathbf{x}} = 0 \quad \text{bzw.} \quad \frac{\partial \mathbf{G}}{\partial \mathbf{x}} = -\left(\frac{\partial \mathbf{F}}{\partial \mathbf{u}}\right)^{-1}\frac{\partial \mathbf{F}}{\partial \mathbf{x}}. \tag{0.6}$$

Es bleibt zu zeigen dass für $\mathbf{F} \in C^k$ auch $\mathbf{G} \in C^k$ gilt: Für $\mathbf{F} \in C^k$ sind $\frac{\partial \mathbf{F}}{\partial \mathbf{x}}$ und $\frac{\partial \mathbf{F}}{\partial \mathbf{u}}$ in C^{k-1}. Da die Inversenbildung eine C^∞-Funktion auf dem Raum der regulären $r \times r$ Matrizen ist, wie man etwa durch die Darstellung der Inversen mithilfe der algebraischen Komplemente sieht, folgt aus (0.6), dass $\frac{\partial \mathbf{G}}{\partial \mathbf{x}}$ aus C^{k-1} ist. $\square$

Ein wichtiger Spezialfall des vorangegangenen Satzes ist der Fall $q = r$. Eine bijektive Abbildung $\mathbf{F}$ von einer offenen Menge $D \subseteq \mathbb{R}^r$ auf eine offene Menge $\tilde{D} \subseteq \mathbb{R}^r$ heißt ein C^k-Diffeomorphismus, wenn $\mathbf{F}$ und $\mathbf{F}^{-1}$ C^k-Funktionen sind.

Satz 0.1.3 *(Umkehrsatz) Ist* $\mathbf{F}$ *auf der offenen Menge* $D \subseteq \mathbb{R}^r$ *eine* C^k*-Abbildung mit* $d\mathbf{F}(\mathbf{x}_0)$ *regulär, dann gibt es offene Umgebungen* V *von* $\mathbf{x}_0$ *und* W *von* $\mathbf{F}(\mathbf{x}_0)$*, sodass* $\mathbf{F}$ *eingeschränkt auf* V *ein* C^k*-Diffeomorphismus von* V *auf* W *ist. Ist* $d\mathbf{F}$ *auf ganz* D *regulär, dann ist* $\mathbf{F}$ *eine offene Abbildung, d.h. das Bild offener Teilmengen von* D *ist offen.*

Beweis: Dies folgt unmittelbar aus Satz 0.1.2, wenn wir die Funktion $\tilde{\mathbf{F}} : \mathbb{R}^r \times \mathbb{R}^r \to \mathbb{R}^r$, $\tilde{\mathbf{F}}(\mathbf{z}, \mathbf{w}) = \mathbf{F}(\mathbf{w}) - \mathbf{z}$ betrachten: Dann gibt es eine Funktion $\mathbf{G}$ von einer Umgebung von $\mathbf{z}_0$ in eine Umgebung von $\mathbf{w}_0$, die $\tilde{\mathbf{F}}(\mathbf{z}, \mathbf{G}(\mathbf{z})) = 0$, also $\mathbf{F}(\mathbf{G}(\mathbf{z})) - \mathbf{z} = 0$ erfüllt. $\mathbf{G}$ ist also lokal die Umkehrfunktion zu $\mathbf{F}$. Insbesondere liegt mit $\mathbf{F}(\mathbf{z})$ eine Umgebung dieses Punktes in der Bildmenge von $\mathbf{F}$, $\mathbf{F}$ ist also eine offene Abbildung. $\square$

Beispiel 0.1.4 *Wir untersuchen welche Variable implizit in einer Umgebung des Punktes* $\mathbf{x}_0 = (r_0, s_0, t_0, u_0) = (0, 0, 1, 0)$ *durch die Gleichungen*

$$F_1(r, s, t, u) = \tan r + rs + \sqrt{t} \sin s = 0 \tag{0.7}$$

$$F_2(r, s, t, u) = s^2 - u^2 + tu = 0 \tag{0.8}$$

durch die anderen beiden definiert sind und berechnen das Differential $d\mathbf{G}$ *für diese Funktionen.*

Die Ableitungsmatrix bei $\mathbf{x}_0 = (0, 0, 1, 0)$ *ist*

$$d\mathbf{F} = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Die einzige 2x2 Untermatrix mit Rang 2 ist die aus der 1. und der 4. Spalte bestehende Untermatrix

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \qquad \text{von } d\mathbf{F}.$$

Aufgrund des Hauptsatzes über implizite Funktionen kann das Gleichungssystem lokal um $\mathbf{x}_0$ *nach* r *und* u *aufgelöst werden, d.h. es gibt in einer Umgebung von* $(0, 1)$ *eine Funktion* $\mathbf{G} = (g_1, g_2)^t$ *mit* $F_i(g_1(s, t), s, t, g_2(s, t)) = 0$, $i = 1, 2$.

Wegen

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \text{ gilt } \qquad d\mathbf{G}(0, 0) = -\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

0.2 Extrema unter Nebenbedingungen, Lagrangemultiplikatoren

Häufig werden Extrema einer Funktion $f : \mathbb{R}^n \to \mathbb{R}$ gesucht, wobei die Variablen $\mathbf{x}$ gewissen Nebenbedingungen $\mathbf{G}(\mathbf{x}) = 0$ genügen sollen, d.h. wir suchen lokale Maxima (Minima) in der Nullstellenmenge von $\mathbf{G} = (g_1, \ldots, g_l)$.

Ein lokales Extremum der Funktion f unter der Nebenbedingung $\mathbf{G}(\mathbf{x}) = 0$ ist demnach ein Punkt $\mathbf{x}$ des Definitionsbereiches von f, der $\mathbf{G}(\mathbf{x}) = 0$ und $f(\mathbf{x}) \geq f(\mathbf{y})$ $(f(\mathbf{x}) \leq f(\mathbf{y}))\,\forall \mathbf{y} \in U, \mathbf{G}(\mathbf{y}) = 0$ für eine Umgebung U von $\mathbf{x}$ erfüllt. Hat $d\mathbf{G}(\mathbf{x})$ in einem Punkt $\mathbf{x}$ der Nullstellenmenge von $\mathbf{G}$ vollen Rang, so können wir nach dem Hauptsatz über implizite Funktionen l der Variablen $x_1, \ldots, x_n$ mithilfe der Gleichungen $\mathbf{G}(\mathbf{x}) = 0$ durch die restlichen $n - l$ Variablen ausdrücken. Sei zunächst angenommen, dass $\frac{\partial \mathbf{G}}{\partial \mathbf{u}}$ vollen Rang hat. Dann ist $(\mathbf{y}, \mathbf{u}) \in D$ ein lokales Extremum von f unter den Nebenbedingungen $\mathbf{G}(\mathbf{x}) = 0$ genau dann wenn $\tilde{f}(\mathbf{y}) := f(\mathbf{y}, \mathbf{H}(\mathbf{y}))$ in $\mathbf{y}$ ein lokales Extremum ohne Nebenbedingungen hat, wobei $\mathbf{H}$ die durch Satz 0.1.2 gegebene Lösungsfunktion der Gleichung $\mathbf{G}(\mathbf{y}, \mathbf{H}(\mathbf{y})) = 0$ ist. Eine notwendige Bedingung für ein lokales Extremum in $\mathbf{x}_0$ von f unter den Nebenbedingungen $\mathbf{G}(\mathbf{x}) = 0$ in Punkten in denen $d\mathbf{G}$ vollen Rang hat ist also $d\tilde{f}(\mathbf{x}_0) = 0$. Diese Bedingung ist wegen

$$\frac{d\tilde{f}}{d\mathbf{y}} = \frac{df}{d\mathbf{y}} + \frac{df}{d\mathbf{u}} \frac{d\mathbf{H}}{d\mathbf{y}}$$

äquivalent zu

$$\frac{df}{d\mathbf{y}}(\mathbf{x}_0, \mathbf{u}_0) + \frac{df}{d\mathbf{u}}(\mathbf{x}_0, \mathbf{u}_0) \frac{d\mathbf{H}}{d\mathbf{y}}(\mathbf{x}_0, \mathbf{u}_0) = 0. \tag{0.9}$$

Wegen $\mathbf{G}(\mathbf{y}, \mathbf{H}(\mathbf{y})) = 0$ folgt

$$\frac{d\mathbf{G}}{d\mathbf{y}} + \frac{d\mathbf{G}}{d\mathbf{u}} \frac{d\mathbf{H}}{d\mathbf{y}} = 0 \text{ bzw.} \quad \frac{d\mathbf{H}}{d\mathbf{y}} = -\left(\frac{d\mathbf{G}}{d\mathbf{u}}\right)^{-1} \frac{d\mathbf{G}}{d\mathbf{y}}$$

Mit

$$\boldsymbol{\lambda} := \frac{df}{d\mathbf{u}} \left(\frac{d\mathbf{G}}{d\mathbf{u}}\right)^{-1} \quad \text{bzw.} \quad \frac{df}{d\mathbf{u}} - \boldsymbol{\lambda} \frac{d\mathbf{G}}{d\mathbf{u}} = 0 \tag{0.10}$$

(die Faktoren $\boldsymbol{\lambda}$ werden *Lagrangemultiplikatoren* genannt) folgt aus (0.9)

$$\frac{df}{d\mathbf{y}}(\mathbf{x}_0, \mathbf{u}_0) - \boldsymbol{\lambda} \frac{d\mathbf{G}}{d\mathbf{y}}(\mathbf{x}_0, \mathbf{u}_0) = 0. \tag{0.11}$$

(0.10) und (0.11) können zusammen als

$$\frac{df}{d\mathbf{x}}(\mathbf{x}_0, \mathbf{u}_0) - \boldsymbol{\lambda} \frac{d\mathbf{G}}{d\mathbf{x}}(\mathbf{x}_0, \mathbf{u}_0) = 0. \tag{0.12}$$

geschrieben werden. In dieser Darstellung spielt es keine Rolle nach welchen Variablen wir die Nebenbedingung $\mathbf{G}(\mathbf{x}) = 0$ lösen. Die getroffene Annahme, dass die $r \times r$. Matrix der letzen r Spaöten von $d\mathbf{G}$ vollen Rang hat kann also in dieser Darstellung durch die Bedingung $d\mathbf{G}$ hat vollen Rang ersetzt werden. Setzen wir $F(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) - \boldsymbol{\lambda}\mathbf{G}(\mathbf{x})$, so entspricht (0.12) der Gleichung

$$\frac{dF}{d\mathbf{x}} = 0. \tag{0.13}$$

Da $\frac{dF}{d\boldsymbol{\lambda}} = 0$ genau die Nebenbedingungen $\mathbf{G}(\mathbf{x}) = 0$ ergeben können wir diese mit (0.13) also $dF = 0$ schreiben.

Damit erhalten wir:

Satz 0.2.1 *Eine notwendige Bedingung für das Vorhandensein eines lokalen Extremums unter Nebenbedingungen $\mathbf{G} = 0$ an einer Stelle $\mathbf{x}_0$ für die $d\mathbf{G}(\mathbf{x}_0)$ vollen Rang hat ist, dass die Funktion*

$$F(\mathbf{x}, \boldsymbol{\lambda}) := f(\mathbf{x}) - \boldsymbol{\lambda}\mathbf{G}(\mathbf{x})$$

an der Stelle $(\mathbf{x}_0, \boldsymbol{\lambda})$ eine stationäre Stelle hat, also $dF = 0$ erfüllt.

Beispiel 0.2.2 *Gesucht sind Extrema der Funktion $f: \; f(x,y,z) = x+2y+3z$ unter der Nebenbedingung $x^2 + 2y^2 + z^2 = 1$.*

Wir erhalten

$$F(x,y,z,\lambda) = x + 2y + 3z - \lambda(x^2 + 2y^2 + z^2 - 1)$$

und aus $dF = 0$ das Gleichungssystem

$$\begin{aligned} 1 - 2x\lambda &= 0 \\ 2 - 4y\lambda &= 0 \\ 3 - 2z\lambda &= 0 \\ x^2 + 2y^2 + z^2 - 1 &= 0. \end{aligned}$$

Aus der ersten Gleichung folgt $\lambda = 1/2x$ und $x \neq 0$ und damit $y = x$ und $z = 3x$ aus der 2. und 3. Gleichung. Setzen wir in die 4. Gleichung ein erhalten wir $x^2 + 2x^2 + 9x^2 = 1$, also $12x^2 = 1$ und die beiden Lösungen für (x,y,z) $\pm\frac{1}{2\sqrt{3}}(1,1,3)$. Da das Differential dg der Funktion $g(x,y,z) = x^2 + 2y^2 + z^2 - 1$ außer in $(0,0,0)$ vollen Rang hat, aber $g(0,0,0) \neq 0$ gilt sind das die einzigen möglichen lokalen Extrema. Die Nullstellenmenge von g ist eine beschränkte und wegen der Stetigkeit von $\mathbf{G}$ abgeschlossene, also eine kompakte Teilmenge von $\mathbb{R}^3$, die stetige Funktion f nimmt darauf somit ihr Maximum und ihr Minimum an. Als globales Maximum resp. Minimum unter der NB sind diese Extrema auch lokale Extrema unter der NB. Wegen $f(\pm\frac{1}{2\sqrt{3}}(1,1,3)) = \pm 2\sqrt{3}$ ist das einzige lokale Minimum das globale Minimum in $-\frac{1}{2\sqrt{3}}(1,1,3)$ und das einzige lokale Maximum das globale Maximum in $\frac{1}{2\sqrt{3}}(1,1,3)$.

Man beachte, dass die Definitheit von d^2F keine Rückschlüsse auf das Vorhandensein lokaler Extrema von f unter der NB zulässt, da Satz 0.2.1 nur sagt dass lokale Extrema von f unter der NB stationäre Stellen von F sind, die aber keine lokalen Extrema von F sein müssen. Tatsächlich ist d^2F an $\frac{1}{2\sqrt{3}}(1,1,3)$ indefinit wie man aus den ersten beiden Hauptminoren von $d^2F(\frac{1}{2\sqrt{3}}(1,1,3))$ sieht, obwohl dort das Maximum unter der NB ist.

Beispiel 0.2.3 *Wir untersuchen welcher Quader mit Volumen 1 minimale Oberfläche hat:*

Für die Lagrangefunktion der halben Oberfläche unter der Nebenbedingung $xyz - 1 = 0$ *erhalten wir*

$$F(x, yz, \lambda) = yz + xz + xy - \lambda(xyz - 1).$$

Wir setzen die partiellen Ableitungen 0:

$$\begin{aligned} y + z - \lambda yz &= 0 \\ x + z - \lambda xz &= 0 \\ x + y - \lambda xy &= 0 \\ xyz - 1 &= 0 \end{aligned}$$

Multiplikation der 1. Gleichung mit x*, der 2. mit* y *und der 3. mit* z *gibt als einzige Lösung* $x = y = z = 1$ *(Wegen der 4. Gleichung gilt* $x, y, z \neq 0$*). Da das Differential* dg *der Funktion* $g(x, y, z) = xyz - 1$ *auf ihrer Nullstellenmenge Rang 1 hat, ist dies das einzig mögliche lokale Extremum unter der Nebenbedingung.*

Auf der Nullstellenmenge von g *gilt*

$$O/2 = yz + xz + xy \geq xyz/\min\{x, y, z\} = 1/\min\{x, y, z\} \geq \sqrt{\max\{x, y, z\}}.$$

Allso strebt $O/2$ *für* $\max\{x, y, z\} \to \infty$ *gegen unendlich. Damit muss* $O/2$ *auf der Nullstellenmenge von* g *ein globales Minimum haben. Dieses ist auch ein lokales Extremum und kann also nur* $(1, 1, 1)$ *sein.*

Beispiel 0.2.4 *Wir bestimmen den Zylinder mit maximalem Volumen der in der Einheitskugel liegt durch Lagrangemultiplikatoren:*

Das Zylindervolumen ist $\pi r^2 h$ *unsere Nebenbedingung ist* $r^2 + h^2/4 = 1$*. Wir setzen* $F(r, h, \lambda) = \pi r^2 h - \lambda(r^2 + h^2/4 - 1)$ *und erhalten aus* $dF = 0$*:*

$$\begin{aligned} 2\pi r h - 2r\lambda &= 0 \\ \pi r^2 - \frac{1}{2} h\lambda &= 0 \\ r^2 + h^2/4 - 1 &= 0 \end{aligned}$$

Als Lösungen ergibt sich $(0, 2, 0)$ *mit* $V = 0$ *und* $(\frac{\sqrt{2}}{\sqrt{3}}, \frac{2}{\sqrt{3}}, \frac{2\pi}{\sqrt{3}})$ *mit Volumina 0 und* $\frac{4\pi}{3\sqrt{3}}$*. Da das Differential der Funktion* $g(r, h) = r^2 + h^2/4 - 1$ *auf ihrer Nullstellenmenge vollen Rang hat gibt es keine weitern lokalen Extrema. Da die Volumina für* $r = 0$ *und* $r = 1$ *0 sind muss die stationäre Stelle das Maximum sein.*

Beispiel 0.2.5 *Gesucht ist der minimale Abstand von Punkten auf der Ellipse* $x^2 + 2y^2 = 1$ *und der Geraden* $x + y = 4$.

Der Abstand zweier Punkten (x_1, y_1) *und* (x_2, y_2) *ist* $\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$. *Dieser ist offensichtlich genau dann minimal, wenn* $(x_1 - x_2)^2 + (y_1 - y_2)^2$ *minimal ist.*

Wir erhalten mit

$$F = (x_1 - x_2)^2 + (y_1 - y_2)^2 - \lambda(x_1^2 + 2y_1^2 - 1) - \mu(x_2 + y_2 - 4)$$

durch partielle Differentiation das Gleichungssystem

$$\begin{aligned} 2x_1 - 2x_2 - 2\lambda x_1 &= 0 && \text{(i)} \\ 2x_2 - 2x_1 - \mu &= 0 && \text{(ii)} \\ 2y_1 - 2y_2 - 4\lambda y_1 &= 0 && \text{(iii)} \\ 2y_2 - 2y_1 - \mu &= 0 && \text{(iv)} \\ x_1^2 + 2y_1^2 - 1 &= 0 && \text{(v)} \\ x_2 + y_2 - 4 &= 0 && \text{(vi)} \end{aligned}$$

Aus (i) und (ii) ergibt sich $-\mu = 2\lambda x_1$ *und aus (iii) und (iv)* $4\lambda y_1 = -\mu$. *Da wie man unmittelbar nachrechnet* $\lambda = 0$ *keine Lösung ergibt folgt* $x_1 = 2y_1$ *und mit (v)* $x_1 = \pm\frac{\sqrt{6}}{3}$, $y_1 = \pm\frac{\sqrt{6}}{6}$.

Aus (ii) und (iv) folgt $x_2 - y_2 = x_1 - y_1 = \pm\frac{\sqrt{6}}{6}$, *was mit (vi)* $x_2 = 2 \pm \frac{\sqrt{6}}{12}$ *und* $y_2 = 2 \mp \frac{\sqrt{6}}{12}$ *ergibt.*

Da $d\mathbf{G}$ *auf der Nullstellenmenge von* $\mathbf{G}$ *vollen Rang hat sind*

$$\left(\left(\pm\frac{\sqrt{6}}{3}, \pm\frac{\sqrt{6}}{6}\right), \left(2 \pm \frac{\sqrt{6}}{12}, 2 \mp \frac{\sqrt{6}}{12}\right)\right)$$

die einzigen möglichen Punkte mit minimalem Abstand mit Abständen

$$2\sqrt{2} \mp \frac{\sqrt{3}}{2}.$$

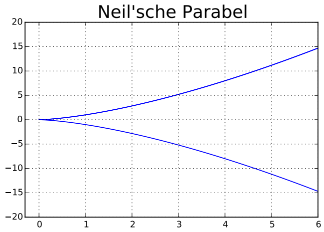
Da es offensichtlich Punkte mit minimalem Abstand gibt folgt dass

$$\left(\left(\frac{\sqrt{6}}{3}, \frac{\sqrt{6}}{6}\right), \left(2 + \frac{\sqrt{6}}{12}, 2 - \frac{\sqrt{6}}{12}\right)\right)$$

die gesuchte Lösung ist.

Das folgende Beispiel zeigt, dass die Bedingung an $d\mathbf{G}$ vollen Rang in $\mathbf{x}$ zu haben tatsächlich notwendig ist, um durch Lagrange-Multiplikatoren zu prüfen ob in $\mathbf{x}$ ein lokales Extremem unter Nebenbedingungen vorliegen kann:

Die Neil'sche Parabel kann implizit durch die Gleichung $f(x, y) = 0$ dargestellt werden. Offensichtlich ist in $(0, 0)$ für alle Punkte auf der Neil'schen Parabel die x-Koordinate minimal.



Beispiel 0.2.6 *Gesucht sind die lokalen Extrema von $f(x, y) = x$ unter der Nebenbedingung $x^3 - y^2 = 0$.*

Für $F(x, y, \lambda) = x - \lambda(x^3 - y^2)$ erhalten wir

$$\frac{\partial F}{\partial x} = 1 - 3x^2\lambda = 0 \qquad \text{(i)}$$

$$\frac{\partial F}{\partial y} = 2y\lambda = 0 \qquad \text{(ii)}$$

$$\frac{\partial F}{\partial \lambda} = x^3 - y^2\lambda = 0 \qquad \text{(iii)}$$

Aus i) folgt $x\lambda \neq 0$, damit aus ii) $y = 0$ und mit iii) ein Widerspruch. Das Gleichungssystem hat also keine Lösung. Die Bedingung dG hat vollen Rang ergibt $(3x^2, 2y^2) \neq (0, 0)$. Da der Punkt (0,0) in der Nullstellenmenge von G liegt muss die Funktion f in $(0, 0)$ zusätzlich auf das Vorliegen eines lokalen Extremums unter der Nebenbedingung $x^3 - y^2 = 0$ untersucht werden. Offensichtlich gilt $x \geq 0$ auf der Nullstellenmenge von $x^3 - y^2$. Im Punkt $(0, 0)$ ist also ein Minimum von f unter der Nebenbedingung.

0.3 Übungsbeispiele

1, Finden Sie Punkte $(x, y) \in \mathbb{R}^2$ für die kokal durch

$$x^3 + 3y^2 + xz - 3z^4y = 1.$$

z als eine Funktion von x und y gegeben ist. Berechnen Sie in diesem Punkt die Ableitung dieser Funktion $z(x, y)$.

2. Zeigen Sie, dass die Funktion

$$\mathbf{F}(x, y) = (x^2 - y^2, xy)$$

außer um $(0, 0)$ in jedm Punkt lokal invertierbar ist.

3. Für Funktionen f, g von $\mathbb{R}^2 \supseteq D \to \mathbb{R}$ werden die Lösungen der Gleichungen $f(x, y) = \alpha$ resp. $g(x, y) = \beta$ die Höhenlinien von f resp. g genannt. Geben Sie Bedingungen an die partiellen Ableitugnen von f und g in einem

Punkt (x_0, y_0) die sicherstellen, dass sich die Höhenlinien in (x_0, y_0) senkrecht schneideb.

4. Diskutieren Sie wo durch die Funktion

$$x_1^4 + \ldots + x_r^4 = 1$$

je eine der Variablen x_i als Funktion der anderen $r - 1$ definiert wird.

Lösen Sie die folgenden Aufgaben mit Lagrange'schen Multiplikatoren

5. Für welchen Punkt des Kreises $x^2 + y^2 = 5$ ist $2x + y$ maximal?

6. Für welchen Punkt der Ellipse $x^2 + 2y^2 + 3z^2 = 1$ ist die Funktion $f(x, y, z) = 2x + 3y + z$ maximal?

7. Für welches Rechteck mit Fläche A ist der Umfang minimal?

8. Welcher Quader mit Oberfläche A hat das größte Volumen?

9. Finden sie das Minimum der Funktion

$$f(x, y, z, w) = x^2 + 2y^2 + z^2 + w^2.$$

unter den Nebenbedingungen

$$\begin{aligned} x + y + z + 3w &= 1 \\ x + y + 2z + w &= 2. \end{aligned}$$

10. Zeigen Sie für $x_i > 0$, $r_i > 0$, $1 \leq i \leq m$ die Ungleichung

$$\left(\prod_{i \leq m} x_i^{r_i} \right)^{1/r} \leq \frac{1}{r} \sum_{i \leq m} x_i r_i$$

für

$$r = \sum_{i \leq m} r_i$$

erfüllt, indem Sie das Maximum von $\prod_{i \leq m} x_i^{r_i}$ unter der Nebenbedingung $\sum_{i \leq m} x_i r_i = \alpha > 0$ bestimmen.