8 Implizite Funktionen und Lagrange-Multiplikatoren

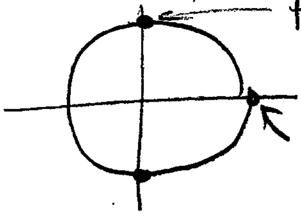
8.1 Beispiel: Betrachte $F:\mathbb{R}^2\to\mathbb{R}$, $F(x,y)=x^2+y^2-1$.

Dann ist $F(x,y)=0$ für (x,y) auf dem Einheitskreis.

Frage: Kann man y in Abhängigkeit von x schreiben auf der Linie $F(x,y)=0$, also finde $f:\mathbb{R}\to\mathbb{R}$ mit $F(x,f(x))=0$?

Antwort: Ja, lokal möglich, falls $\frac{\partial F}{\partial y}(x,y)=2y$ invertierbar ist, also für $y\neq 0$:

für $(x,y)=(0,1)$ ist lokal $f(x)=\sqrt{1-x^2}$ eine Fkt. mit $F(x,f(x))=0$ $\forall x\in(-1,1)$



für $(x,y)=(1,0)$ ist $y=0$ und keine eindeutige lokale Lösung möglich

für $(x,y)=(0,-1)$ ist $f(x)=-\sqrt{1-x^2}$ eine lokale Fkt. für $F(x,f(x))=0$

8.2 Satz (implizite Funktionen): Sei $U\subseteq\mathbb{R}^p\times\mathbb{R}^q$ offen und $(x_0,y_0)\in U$.
Sei $F:U\to\mathbb{R}^q$ stetig partiell differenzierbar.
Sei $F(x_0,y_0)=0$ und sei $\frac{\partial F}{\partial y}(x_0,y_0)=\left(\frac{\partial F_i}{\partial y_j}(x_0,y_0)\right)_{i,j=1,\dots,q}$ eine invertierbare $q\times q$-Matrix.

Dann gibt es $\varepsilon>0$ und $\delta>0$ mit $B(x_0,\delta)\times B(y_0,\varepsilon)\subseteq U$ und genau eine stetige Abbildung $f:\underbrace{B(x_0,\delta)}_{\subseteq\mathbb{R}^p}\to\underbrace{B(y_0,\varepsilon)}_{\subseteq\mathbb{R}^q}$, so dass $F(x,f(x))=0$ $\forall x\in B(x_0,\delta)$.

Für jedes $x\in B(x_0,\delta)$ ist $f(x)=y$ sogar die einzige Lösung der Gleichung $F(x,y)=0$.
Die Funktion f ist sogar differenzierbar in x_0 mit $Df(x_0)=-\left(\frac{\partial F}{\partial y}(x_0,y_0)\right)^{-1}\frac{\partial F}{\partial x}(x_0,y_0)$.

(Man kann also die Gleichung $F(x,y)=0$ implizit nach y auflösen, d.h. y kann als $f(x)$ geschrieben werden, selbst wenn man f nicht explizit angeben kann; f kann sogar implizit differenziert werden.
Außerdem: Ist F „in seiner y-Komponente nicht extremal" an der Stelle (x_0,y_0), so kann aus der Nullstelle (x_0,y_0) eine ganze „Nullstellen-Höhenlinie" gemacht werden.)

<u>Kurzversion von 8.2:</u> $\mathbb{R}^p \times \mathbb{R}^q \supseteq U \xrightarrow{F} \mathbb{R}^q$ stetig partiell diffbar.

1.) Ist F in einem Punkt Null, so findet man schon eine ganze Nullstellenkurve (falls $\frac{\partial F}{\partial y}$ invertierbar)

2.) Kann eine Funktion f implizit durch das Gleichungssystem $F(x, f(x)) = 0$ (q Gleichungen, $p+q$ Unbekannte) definieren.

Weiß: f ist stetig und diffbar, kenne sogar Ableitung.

<u>Beweis:</u> Statt $\quad 0 = F(x, f(x))$

betrachte $\quad f(x) = \underbrace{f(x) - F(x, f(x))}_{=: Tf(x)}$ (äquivalent)

Fixpunktsatz. $\square$

Beweis: 1.) o.E. $\frac{\partial F}{\partial y}(x_0,y_0) = 1 = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$

Beweis zu 1.): $\tilde{F}(x,y) := \left(\frac{\partial F}{\partial y}(x_0,y_0)\right)^{-1} F(x,y) = 0 \Leftrightarrow F(x,y) = 0$

und $\frac{\partial \tilde{F}}{\partial y}(x_0,y_0) = \left(\frac{\partial F}{\partial y}(x_0,y_0)\right)^{-1} \frac{\partial F}{\partial y}(x_0,y_0) = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$. $\square$ (1.)

2.) Setze $G : U \to \mathbb{R}^q$, $G(x,y) := y - F(x,y)$.

Dann (i) $\frac{\partial G}{\partial y}$ stetig und $\frac{\partial G}{\partial y}(x_0,y_0) = 0$.

(ii) $\exists \varepsilon > 0 : \left\| \frac{\partial G}{\partial y}(x,y) \right\| < \frac{1}{2}$ falls $x \in B(x_0,\varepsilon), \overline{y \in B(y_0,\varepsilon)}$

(iii) $\exists\, 0 < \delta < \varepsilon : \| y_0 - G(x,y_0) \| < \frac{\varepsilon}{4}$ falls $x \in \overline{B(x_0,\delta)}$

(iv) $B(x_0,\delta) \times B(y_0,\varepsilon) \subseteq U$

Beweis zu 2.): (i) F stetig part. differenzierbar, $\mathrm{id} : \mathbb{R}^q \to \mathbb{R}^q$ ebenso $\Rightarrow \frac{\partial G}{\partial y}$ stetig.

und $\frac{\partial G}{\partial y}(x_0,y_0) = 1 - \frac{\partial F}{\partial y}(x_0,y_0) \overset{1.)}{=} 0$

(ii) Folgt direkt aus (i).

(iii) $y_0 - G(x_0,y_0) = F(x_0,y_0) = 0$ und Stetigkeit von G in x_0.

(iv) wähle ε und δ klein genug, U offen. $\square$ (2.)

3.) $X := \{ g : \underbrace{\overline{B(x_0,\delta)}}_{\subseteq \mathbb{R}^p} \to \mathbb{R}^q \text{ stetig} \mid g(x_0) = y_0,\ \|g(x) - y_0\| \leq \frac{\varepsilon}{2}\ \forall x \in \overline{B(x_0,\delta)} \}$ ist ein vollständiger, metrischer Raum mit $\|g\|_\infty := \sup_{x \in \overline{B(x_0,\delta)}} \|g(x)\|$

Beweis zu 3.): $\mathcal{C}(\overline{B(x_0,\delta)}, \mathbb{R}^q) := \{ f : \overline{B(x_0,\delta)} \to \mathbb{R}^q \text{ stetig} \}$ ist mit $\|f\|_\infty := \sup_{x \in \overline{B(x_0,\delta)}} \|f(x)\|$ ein vollständiger, metrischer Raum, da $(\mathbb{R}^q, \|\cdot\|)$ ein vollständiger normierter Raum ist und man den Beweis für $(\mathcal{C}[0,1], \|\cdot\|_\infty)$ wortwörtlich adaptieren kann. $X \subseteq \mathcal{C}(\overline{B(x_0,\delta)}, \mathbb{R}^q)$ abgeschlossen. $\square$ (3.)

4.) Die Abbildung $T : X \to \mathcal{C}(\overline{B(x_0,\delta)}, \mathbb{R}^q)$, $(Tg)(x) := G(x, g(x))$ hat ihr Bild in X, es gilt also $T : X \to X$.

Beweis zu 4.): Da g und G stetig sind, ist auch Tg stetig.

Und $(Tg)(x_0) = G(x_0, g(x_0)) = g(x_0) - F(x_0, g(x_0)) = y_0 - F(x_0,y_0) = y_0$

Sei $x \in \overline{B(x_0,\delta)}$. Sei $h \in X$ mit $h(x) = y_0$ (z.B. $h(x) \equiv y_0\ \forall x$).

Dann: $\|Tg(x) - y_0\| \leq \underbrace{\|Tg(x) - Th(x)\|}_{\underset{5.)}{\leq} \frac{1}{2}\|g(x) - \underbrace{h(x)}_{= y_0}\| \leq \frac{\varepsilon}{2} \text{ da } g \in X} + \underbrace{\|Th(x) - y_0\|}_{= \|G(x,y_0) - y_0\| \underset{2.)}{\leq} \frac{\varepsilon}{4}} \leq \frac{\varepsilon}{2} \Rightarrow Tg \in X$.

$\square$ (4.)

5.) Es gilt $\|Tg(x) - Th(x)\| \le \frac{1}{2} \|g(x) - h(x)\| \quad \forall x \in \overline{B(x_0,\delta)}$

und $\|Tg - Th\|_\infty \le \frac{1}{2} \|g - h\|_\infty$.

Beweis von 5.): Sei $x \in \overline{B(x_0,\delta)}$, $g, h \in X$. Setze $y := h(x)$, $z := g(x)$.

Dann $y + t(z-y) \overset{(**)}{\in} B(y_0,\varepsilon) \quad \forall t \in [0,1]$

Setze $a(t) := G(x, y + t(z-y))$ für $t \in [0,1]$

$$\Rightarrow G(x,z) - G(x,y) = a(1) - a(0) \overset{(*)}{=} \int_0^1 a'(t)\,dt = \int_0^1 \frac{\partial G}{\partial y}(x, y+t(z-y))(z-y)\,dt$$

$$\Rightarrow \|G(x,z) - G(x,y)\| \le \int_0^1 \underbrace{\left\|\frac{\partial G}{\partial y}(x, y+t(z-y))\right\|}_{\overset{2.)}{\le} \frac{1}{2}} \|z-y\|\,dt \le \frac{1}{2}\|z-y\|$$

$$\Rightarrow \|Tg(x) - Th(x)\| = \|G(x,z) - G(x,y)\| \le \frac{1}{2}\|g(x) - h(x)\|$$

Nehme sup über $x \in \overline{B(x_0,\delta)}$: $\|Tg - Th\|_\infty \le \frac{1}{2}\|g - h\|_\infty$ □ (5.)
(zuerst rechts, dann links)

6.) Also ist $T : X \to X$ kontrahierend, nach 1.6 ex. also $f \in X$ mit $Tf = f$, d.h. $f(x) = Tf(x) = G(x, f(x)) = f(x) - F(x, f(x)) \Rightarrow F(x, f(x)) = 0$

Die Abbildung f ist eindeutig. b.z.z.: f ist differ.

[(*): Für $h : [a,b] \to \mathbb{R}^n$ mit h_i, $i = 1, \dots, n$ Regelfkt., definiere $\int_a^b h(t)\,dt := \left(\int_a^b h_i(t)\,dt\right)_{i=1,\dots,n} \in \mathbb{R}^n$. Dann gilt $\left\|\int_a^b h(t)\,dt\right\| \le \int_a^b \|h(t)\|\,dt$

7.) Ist f differ. in x_0, so ist $Df(x_0) = -\left(\frac{\partial F}{\partial y}(x_0,y_0)\right)^{-1} \frac{\partial F}{\partial x}(x_0,y_0)$.

Beweis von 7.): Setze $h(x) := F(x, f(x))$, also $h : B(x_0,\delta) \xrightarrow{g} \mathbb{R}^p \times \mathbb{R}^q \xrightarrow{F} \mathbb{R}^q$
$\subseteq \mathbb{R}^p \quad x \mapsto (x, f(x))$

$$\overset{\text{Kettenregel 6.14}}{\underset{h \equiv 0}{\Longrightarrow}} 0 = D(F \circ g)(x_0) = DF(g(x_0)) \cdot Dg(x_0) = \left(\frac{\partial F}{\partial x}(x_0,y_0), \frac{\partial F}{\partial y}(x_0,y_0)\right) \begin{pmatrix} E \\ \frac{\partial f}{\partial x}(x_0) \end{pmatrix}$$

$$= \frac{\partial F}{\partial x}(x_0,y_0) + \frac{\partial F}{\partial y}(x_0,y_0) \frac{\partial f}{\partial x}(x_0)$$

$$\Rightarrow \frac{\partial f}{\partial x}(x_0) = -\left(\frac{\partial F}{\partial y}(x_0,y_0)\right)^{-1} \frac{\partial F}{\partial x}(x_0,y_0).$$ □ (7.)

[(**): $\|(y + t(z-y)) - y_0\|$

$= \|(tz - ty_0) + ((1-t)y - (1-t)y_0)\|$

$\le t\underbrace{\|z - y_0\|}_{<\varepsilon} + (1-t)\underbrace{\|y - y_0\|}_{<\varepsilon} < \varepsilon$]

8.) f ist diffbar in x_0.

Beweisidee von 8.): $F(x,y) = \overset{=0}{F(x_0,y_0)} + \Big(\underbrace{\frac{\partial F}{\partial x}(x_0,y_0)}_{=:A}, \underbrace{\frac{\partial F}{\partial y}(x_0,y_0)}_{=:B}\Big)\begin{pmatrix} x-x_0 \\ y-y_0 \end{pmatrix} + \varphi(x-x_0, y-y_0)$

mit $\frac{\varphi(x-x_0,y-y_0)}{\|(x-x_0,y-y_0)\|} \to 0$ da F diffbar.

$\Rightarrow 0 = F(x,f(x)) = A(x-x_0) + B(f(x)-f(x_0)) + \varphi(\ldots)$, B invertierbar

$\Rightarrow f(x) = f(x_0) + B^{-1}A(x-x_0) + \underbrace{B^{-1}\varphi(x-x_0, f(x)-f(x_0))}_{=:\psi(x-x_0)}$

Zeige $\frac{\psi(x-x_0)}{\|x-x_0\|} \to 0 \Rightarrow f$ diffbar.

$\square$ (8.1)

$\square$ (8.2)

8.3 Beispiel: a) $F: \mathbb{R}^2 \to \mathbb{R}$, $F(x,y) := e^{x^2+y^2}xy - e^2$, $F(1,1) = 0$.

$\frac{\partial F}{\partial y}(x,y) = x e^{x^2+y^2} + 2xy^2 e^{x^2+y^2}$, $\frac{\partial F}{\partial y}(1,1) = e^2 + 2e^2 = 3e^2 \neq 0$ invertierbar.

$\overset{8.2}{\Longrightarrow}$ y kann um den Punkt $x_0 = 1$ als Funktion von x geschrieben werden

(Da $0 = e^{x^2+y^2}xy - e^2 \Rightarrow y = \frac{e^2}{xe^{x^2+y^2}}$ ist dies nicht explizit möglich)

Also ex. f mit $F(x,f(x)) = 0$, $x \in (1-\delta, 1+\delta)$ und $f'(x) = -\frac{\partial F}{\partial y}(1,1)^{-1}\frac{\partial F}{\partial x}(1,1)$

Da $\frac{\partial F}{\partial x}(x,y) = \frac{\partial F}{\partial y}(x,y)$ (F symmetrisch in x und y), ist $f'(x) = -1$.

b) $h: \mathbb{R}^2 \to \mathbb{R}$, $h(x,y) = x^2 y$. Wann ist h auf $\{(x,y) \mid x^2+y^2 = 1\}$ (~~Kreis~~) extremal? Suche also Extrema unter der Nebenbedingung $x^2+y^2=1$.

Betrachte $F: \mathbb{R}^2 \to \mathbb{R}$, $F(x,y) = x^2+y^2-1$ wie in 8.1, also $F(0,1) = 0$ und $\frac{\partial F}{\partial y}(0,1) = 2 \neq 0$, d.h. finde $F(x,f(x)) = 0$ mit $x \in (-\delta, \delta)$.

Und $f'(x) = -\left(\frac{\partial F}{\partial y}(x,y)\right)^{-1}\left(\frac{\partial F}{\partial x}(x,y)\right) = -(2y)^{-1}(2x) = -\frac{x}{y}$.

Betrachte $g(x) := h(x, f(x))$. Dann $\frac{\partial}{\partial x} g(x) = \left(\frac{\partial}{\partial x}h(x,y), \frac{\partial}{\partial y}h(x,y)\right)\begin{pmatrix} E \\ \frac{\partial f}{\partial x}(x) \end{pmatrix}$

$= 2xy + x^2 \cdot \left(-\frac{x}{y}\right) = 2xy - \frac{x^3}{y}$

g extremal $\Rightarrow \frac{\partial g}{\partial x}(x) = 0$, d.h. $2xy = \frac{x^3}{y} \Rightarrow y^2 = \frac{x^2}{2}$ oder $x = 0$.

Aber $x^2+y^2 = 1$ mit $y^2 = \frac{x^2}{2}$ liefert $x^2 = \frac{2}{3}$, d.h. $x = \pm\sqrt{\frac{2}{3}}$

$\Rightarrow$ Kandidaten für Extrema sind $(0,1)$, $(0,-1)$ $(x=0)$

oder $(x,y) \in \left\{\left(\pm\sqrt{\frac{2}{3}}, \pm\sqrt{\frac{1}{3}}\right)\right\}$.

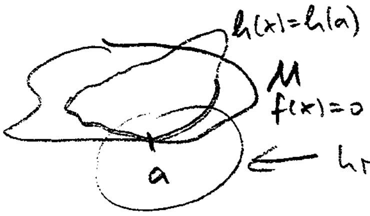
8.4 Satz (Lagrange-Multiplikator): Sei $\mathbb{R}^n \supseteq U \xrightarrow{f} \mathbb{R}$ stetig partiell differenzierbar und $M=\{x\in U \mid f(x)=0\}$. Sei $a\in M$ mit $\operatorname{grad} f(a)\neq 0$.

Sei $\mathbb{R}^n\supseteq U \xrightarrow{h}\mathbb{R}$ stetig diffbar, so dass h auf M in a extremal ist

(d.h. $\forall x\in \underline{M}$ mit $d(x,a)<\varepsilon$ gilt $h(a)\geq h(x)$, falls Extremum = Maximum).

Dann ex. $\lambda\in\mathbb{R}$ („Lagrange-Multiplikator") mit $\operatorname{grad} h(a)=\lambda \operatorname{grad} f(a)$.

[Idee:



$\leftarrow$ hier verlaufen h und f tangential, d.h. Gradienten zeigen in die gleiche Richtung]

__Beweis:__ $0\neq \operatorname{grad} f(a) = \left(\frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_n}(a)\right)$. o.E. $\frac{\partial f}{\partial x_n}(a)\neq 0$.

$a=(a',a_n)\in\mathbb{R}^n$ mit $a'\in\mathbb{R}^{n-1}$. Also $\mathbb{R}^{n-1}\times\mathbb{R}\supseteq U\xrightarrow{f}\mathbb{R}$ stetig part. diffbar, $f(a',a_n)=0$,
$\in\mathbb{R}^{n-1}$ $\in\mathbb{R}$

$\frac{\partial f}{\partial x_n}(a_n)$ invertierbar $\overset{8.2}{\Longrightarrow}$ $\exists g: B(a',\delta)\to B(a_n,\varepsilon)$ mit $f(x_1,\dots,x_{n-1},g(x_1,\dots,x_{n-1}))=0$
(lokale Parametrisierung von M)

und $\frac{\partial g}{\partial x_i} = -\left(\frac{\partial f}{\partial x_n}\right)^{-1}\left(\frac{\partial f}{\partial(x_1,\dots,x_{n-1})}\right)$, d.h. $\frac{\partial g}{\partial x_i} = -\left(\frac{\partial f}{\partial x_n}\right)^{-1}\frac{\partial f}{\partial x_i}$.

Betrachte $u(x_1,\dots,x_{n-1}) := h(x_1,\dots,x_{n-1},g(x_1,\dots,x_{n-1}))$.

Also $u: \mathbb{R}^{n-1}\xrightarrow{b}\mathbb{R}^n\xrightarrow{h}\mathbb{R}$, $x'\mapsto(x',g(x'))\mapsto h(x',g(x'))$, also $b_i(x_1,\dots,x_{n-1}) = \begin{cases} x_i & i\leq n-1\\ g(x_1,\dots,x_{n-1}) & i=n\end{cases}$

Dann ist u in a' maximal ($u(a')=h(a',g(a'))=h(a',a_n)=h(a)$)

$$\Rightarrow 0 = Du(a') \overset{6.14}{=} Dh(\underbrace{a',g(a')}_{=a})\, Db(a') = \left(\frac{\partial h}{\partial x_1}(a) \dots \frac{\partial h}{\partial x_n}(a)\right)\begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \\ \frac{\partial g}{\partial x_1}(a) & \dots & \frac{\partial g}{\partial x_{n-1}}(a)\end{pmatrix}$$

$$\left(\frac{\partial u}{\partial x_1}(a'), \dots, \frac{\partial u}{\partial x_{n-1}}(a')\right)$$

$$\Rightarrow 0=\frac{\partial u}{\partial x_i}(a') = \frac{\partial h}{\partial x_i}(a)+\frac{\partial h}{\partial x_n}(a)\frac{\partial g}{\partial x_i}(a) = \frac{\partial h}{\partial x_i}(a) - \underbrace{\frac{\partial h}{\partial x_n}(a)\left(\frac{\partial f}{\partial x_n}(a)\right)^{-1}}_{=:\lambda}\frac{\partial f}{\partial x_i} \quad \forall i$$

$$\Rightarrow \operatorname{grad} h(a) = \left(\frac{\partial h}{\partial x_1}(a), \dots, \frac{\partial h}{\partial x_n}(a)\right) = \lambda\left(\frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_n}(a)\right) = \lambda \operatorname{grad} f(a).$$

$\square$

8.5 Beispiel (8.3 revisited): $h: \mathbb{R}^2 \to \mathbb{R}$, $(x,y) \mapsto x^2 y$ extremal auf $\{(x,y) \mid x^2+y^2=1\}$?
$f(x,y) := x^2+y^2-1$

Nach 8.4 ex. dann $\lambda \in \mathbb{R}$ mit grad $h(x,y) = \lambda$ grad $f(x,y)$.

grad $h(x,y) = (2xy, x^2)$, grad $f(x,y) = (2x, 2y)$

Also $(2xy, x^2) = \lambda(2x, 2y) \Rightarrow 2xy = \lambda 2x$
$x^2 = \lambda 2y$

$\Rightarrow x=0$ oder $\lambda = y$.

Im letzteren Fall: $x^2 = 2y^2$, d.h. $y = \pm\sqrt{\frac{x^2}{2}}$.

Mit $x^2+y^2=1$ folgt $x^2 + \frac{x^2}{2} = 1$, d.h. $x = \pm\sqrt{\frac{2}{3}}$ und $y = \pm\sqrt{\frac{1}{3}}$.

8.6 Korollar (von 8.2): Sei $\mathbb{R}^n \supseteq U \xrightarrow{f} \mathbb{R}^n$ stetig partiell differenzierbar, $x_0 \in U$ und $Df(x_0)$ invertierbar.

Dann ex. $\delta > 0$ und $g: B(f(x_0), \delta) \to U$ differ. mit $f \circ g(y) = y$, $g \circ f(x) = x$ für alle $y \in B(f(x_0), \delta)$, $x \in f^{-1}(B(f(x_0), \delta))$. Die Funktion g heißt dann „lokale Umkehrfunktion".

Beweis: O.E. $x_0 = 0$, $f(x_0) = 0$ (betrachte sonst $\tilde{f}(x) := f(x+x_0) - f(x_0)$)

Setze $F: U \times U \to \mathbb{R}^n$, $F(x,y) = x - f(y)$. Also $\frac{\partial F}{\partial y}(x_0, x_0) = Df(x_0)$ invertierbar.

$\overset{8.2}{\Longrightarrow}$ $\exists g: B(0,\delta) \to g(B(0,\delta))$ differ. mit $0 = F(x, g(x)) = x - f(g(x))$.

D.h. $f \circ g(x) = x$ und da g surj. (nach Def.) und injektiv $(f \circ g(x) = x)$ auch $g \circ f(x) = x$. □