2.5 Der Satz über implizite Funktionen

Bisher haben wir Funktionen fast ausschließlich explizit definiert durch Angabe von Definitions- und Zielbereich sowie einer Zuordnungsvorschrift, also $f: X \to Y,\ x \mapsto f(x) = \ldots$ (")

Eine Ausnahme bildete die Umkehrfunktion

$$f^{-1}: Y \to X,$$

die wir definieren können durch

$$f^{-1}(y) = x, \quad \text{falls } f(x) = y,$$

sofern zu jedem $y \in Y$ die Gleichung $f(x) = y$ genau eine Lösung $x \in X$ besitzt, d.h. wenn f bijektiv ist. $f^{-1}(y)$ wird also definiert als die eindeutig bestimmte Lösung $x \in X$ der Gleichung $f(x) = y$.

Diese Situation läßt sich verallgemeinern: Gegeben sei eine (explizit bekannte) Funktion

$$F: X \times Y \to Z, \quad (x, y) \mapsto F(x, y)$$

und ein $z_0 \in Z$. Zu jedem $x \in X$ gebe es genau ein $y \in Y$, so daß $F(x, y) = z_0$ gilt, also genau eine Lösung dieser Gleichung.

Dann wird durch

$$g: X \to Y, \quad x \mapsto g(x) := \text{Lösung } y \in Y \text{ der Gleichung}$$
$$F(x,y) = z_0$$

eine Abbildung definiert. Man sagt

- die Funktion g ist durch die Gleichung $F(x,y) = z_0$ implizit definiert und nennt g eine "implizite Funktion", bzw.
- die Gleichung $F(x,y) = z_0$ ist nach y auflösbar (in der Regel nicht durch elementare Umformungen).

Die Auflösbarkeit einer Gleichung $F(x,y) = z_0$ hängt wesentlich vom Definitionsbereich von F ab. Dazu

Bsp 1 Durch die Kreisgleichung $x^2 + y^2 = 1$ wird nur bei geeigneter Wahl des Definitionsbereich von $F_i : (x,y) \mapsto F_i(x,y) = x^2 + y^2$ eine Funktion $g : x \mapsto g(x)$ mit $x^2 + g(x)^2 = 1$ definiert.

(a) $F_1 : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$: $x^2 + y^2 = 1$ ist nicht nach y auflösbar, für $|x| > 1$ existiert keine Lösung

(b) $F_2 : [-1,1] \times \mathbb{R} \to \mathbb{R}$: Hier existiert zwar eine Lösung $y(x) = \pm\sqrt{1-x^2}$, für $|x| < 1$ ist sie aber nicht eindeutig bestimmt.

Also: nicht auflösbar.

(c) $F_3 : [-1,1] \times [0,\infty) \to \mathbb{R}$: $g(x) = \sqrt{1-x^2}$ ist eine "Auflösung der Gleichung nach y" D64

(d) $F_4 : [-1,1] \times (-\infty, 0] \to \mathbb{R}$: $g(x) = -\sqrt{1-x^2}$ ist eine Auflösung

(Konsequenz aus (c) und (d): "Die richtige Auflösung" wird festgelegt durch einen Punkt auf der Kurve.)

Im Folgenden betrachten wir allgemeiner Funktionen $F : \mathbb{R}^{m+n} \supset \Omega \to \mathbb{R}^n$, $(\underset{\mathbb{R}^m}{\underset{\cap}{x}}, \underset{\mathbb{R}^n}{\underset{\cap}{y}})^T \mapsto \underbrace{F(x,y)}_{\in \mathbb{R}^n}$,

in Komponenten

$$(x_1, \dots, x_m, y_1, \dots, y_n) \longmapsto \begin{pmatrix} F_1(x_1, \dots, x_m, y_1, \dots, y_n) \\ \vdots \\ F_n(x_1, \dots, x_m, y_1, \dots, y_n) \end{pmatrix}.$$

Die Gleichung $F(x,y) = 0$ (standardmäßig wird $z_0 = 0$ gewählt) ist dann ein – im allgemeinen nichtlineares Gleichungssystem

$$F_1(x_1, \dots, x_m, y_1, \dots, y_n) = 0$$
$$\vdots$$
$$F_n(x_1, \dots, x_m, y_1, \dots, y_n) = 0,$$

wobei die x_i gegeben und die y_i gesucht sind.

Vorbereitend betrachten wir den einfachen Spezialfall,

daß die Funktion F linear ist:

Bsp. 2: $F: \mathbb{R}^{m+n} \to \mathbb{R}^n$, $(x,y)^T \mapsto A\begin{pmatrix} x \\ y \end{pmatrix} =: F(x,y)$

mit einer $n \times (m+n)$-Matrix

$$A = \left(\begin{array}{cccccc} a_{11} & \cdots & a_{1m} & a_{1\,m+1} & \cdots & a_{1\,m+n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} & a_{n\,m+1} & \cdots & a_{n\,m+n} \end{array}\right)$$

$$\underbrace{\qquad\qquad}_{=: A_x} \qquad \underbrace{\qquad\qquad}_{=: A_y}$$

$n \times m$-Matrix $\qquad$ $n \times n$-Matrix ,

so daß $A\begin{pmatrix} x \\ y \end{pmatrix} = A_x x + A_y \cdot y$. Die Gleichung

$F(x,y) = 0$ lautet dann

$$A_x x + A_y y = 0 ,$$

was wir genau dann nach y auflösen können, wenn die Teilmatrix A_y invertierbar ist. In diesem Fall erhalten wir nämlich durch elementare Umformung

$$y = -(A_y)^{-1} \cdot A_x x .$$

Hier ist eine Auflösung sogar global möglich, was wir i. allg. nicht erwarten können (vgl. Bsp. 1 (a), (b)).

Ähnlich wie beim Satz über die Umkehrfunktion setzen wir jetzt F als stetig diffbar voraus und schreiben die Jacobi-Matrix $DF(x,y)$ als

$$DF(x,y) = (D_x F(x,y), D_y F(x,y))$$

mit $\quad D_x F(x,y) := \left(\frac{\partial F_i}{\partial x_j}(x,y) \right)_{\substack{1 \le i \le n \\ 1 \le j \le m}}$

und $\quad D_y F(x,y) := \left(\frac{\partial F_i}{\partial y_j}(x,y) \right)_{1 \le i,j \le n}$,

also ganz analog zur Zerlegung $A = (A_x, A_y)$ in Bsp. 2.

Vermutung / Ziel: Unter der Voraussetzung der Invertierbarkeit von $D_y F(a,b)$ in einem Punkt $(a,b) \in \Omega$ mit $F(a,b) = 0$ können wir zumindest eine lokale Auflösung von $F(x,y) = 0$ finden. Genauer:

D67

Satz über implizite Funktionen: Es seien $\Omega \subset \mathbb{R}^{m+n}$ offen und $F: \Omega \to \mathbb{R}$ stetig diff'bar. In $(a,b) \in \Omega$ gelte (i) $F(a,b) = 0$ (ii) $\det D_y F(a,b) \neq 0$.

Dann existieren offene Umgebungen W_a von a und W_b von b und eine stetig diff'bare Funktion

$g: W_a \to W_b$

mit $g(a) = b$, $F(x, g(x)) = 0 \quad \forall x \in W_a$ und

$$Dg(x) = -D_y F(x, g(x))^{-1} D_x F(x, g(x)).$$

Bew.: ① Wir definieren die Hilfsfunktion

$$H: \Omega \to \mathbb{R}^{m+n}, \quad (x,y)^T \mapsto H(x,y) := \begin{pmatrix} x \\ F(x,y) \end{pmatrix}.$$

Die Jacobi-Matrix von H ist dann gegeben durch

$$DH(x,y) = \begin{pmatrix} I_m & 0 \\ D_x F(x,y) & D_y F(x,y) \end{pmatrix} \qquad \Big(I_m = \underbrace{\begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}}_{m \times m} \Big)$$

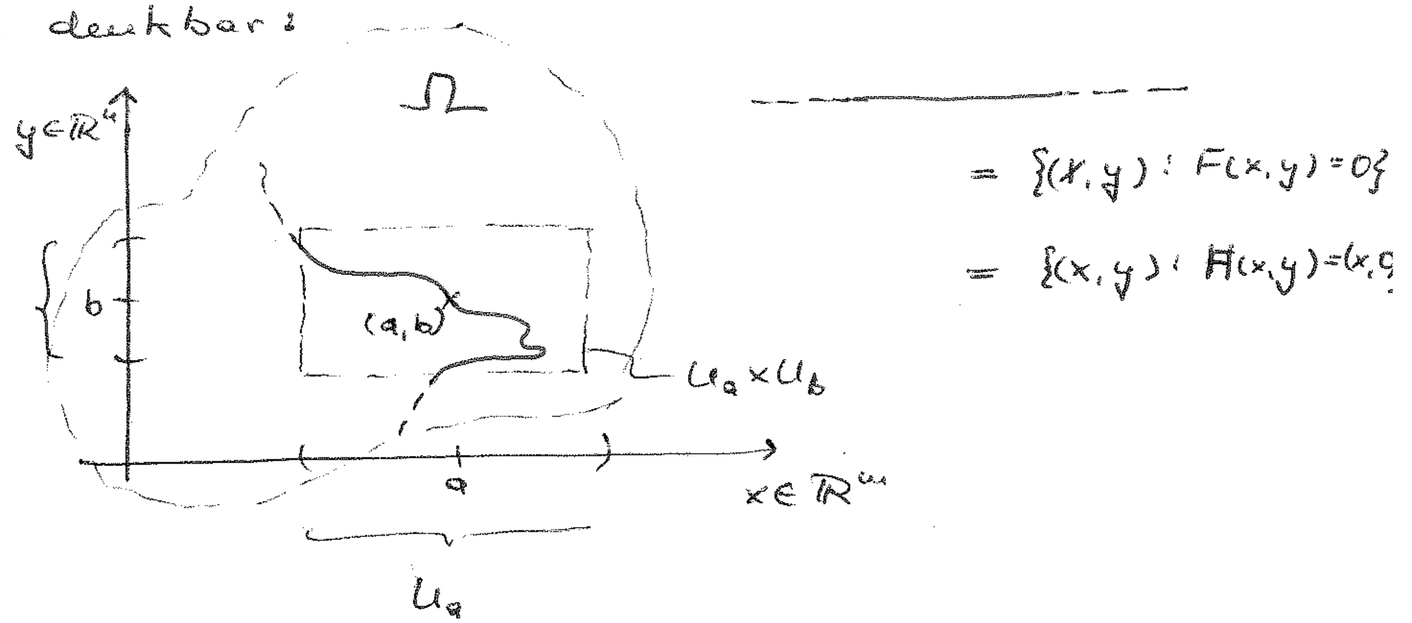
mit $\det(DH(x,y)) = \det(D_y F(x,y))$. Insbesondere ist $\det(DH(a,b)) = \det(D_y F(a,b)) \neq 0$, so dass die Funktion H die Voraussetzungen des Satzes über inverse Abbildungen erfüllt.

Also existieren offene Umgebungen U von (a,b) und V von $H(a,b)$, so daß

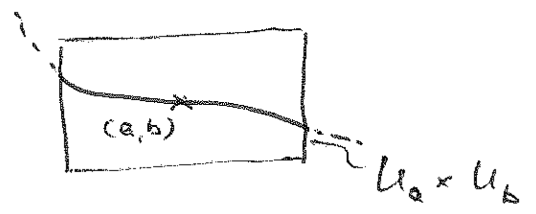
$$H\big|_U : U \to V$$

bijektiv ist. Nach eventl. Verkleinerung können wir $U = U_a \times U_b$ mit offenen Umgebungen U_a von a bzw. U_b von b annehmen. (2) Jetzt sind drei Situationen denkbar:

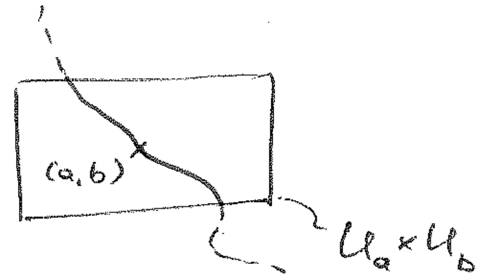
(i)



(ii)



(iii)



(i) Wäre katastrophal, ist aber ausgeschlossen. Denn in diesem Fall gäbe es $x \in U_a$ und $y_{1,2} \in U_b$ mit $y_1 \neq y_2$ und $F(x,y_1) = F(x,y_2) = 0$, also mit

$H(x,y_1) = H(x,y_2) = (x,0)$, im Widerspruch zur Injektivität von H auf $U = U_a \times U_b$

(ii) wäre perfekt, denn dann wäre bereits die Menge

$\{(x,y) \in U_a \times U_b : F(x,y) = 0\}$ der Graph einer

Funktion $g : U_a \to U_b$. Ist i. allg. aber nicht

gegeben.

(iii) Ist der tatsächliche Sachverhalt, der eine weitere

Verkleinerung erfordert. Wir setzen

$$
\begin{aligned}
W_a &:= \{x \in U_a : \exists\, y \in U_b, \text{ so daß } F(x,y) = 0\} \\
&= \{x \in U_a : \exists\, y \in U_b, \text{ " " } H(x,y) = (x,0)\} \\
&= \{x \in U_a : (x,0) \in V = H(U_a \times U_b)\} \\
&= U_a \cap f^{-1}(V) \quad \text{für } f(x) = (x,0).
\end{aligned}
$$

Da f stetig und V offen ist, ist auch W_a offen.

Ferner: $W_b := U_b$. Dann existiert also zu jedem

$x \in W_b$ genau ein $y \in W_b$ mit $F(x,y) = 0$. Wir

setzen $g : W_a \to W_b$, $x \mapsto g(x) :=$ eindeutige Lösung y von $F(x,y) = 0$

③ g ist stetig diff'bar, denn wir haben

$H(x, g(x)) = (x, F(x,g(x)) = (x,0)$, also

$(x, g(x)) = H^{-1}(x,0)$ und nach dem Satz über

die Umkehrabbildung ist H^{-1} stetig diff'bar.

Die Aussage über Dg folgt aus der Kettenregel:

D69

(a)

$0 = F(x, g(x))$

$\Rightarrow \quad 0 = D\, F(x, g(x)) = D_x F(x, g(x)) + D_y F(x, g(x))\, Dg(x).$

Nun ist aber $\det(D_y F(x, g(x))) = \det H(x, g(x)) \neq 0$ für alle $x \in U_a$ (Satz über inverse Abbn.) und damit $D_y F(x, g(x))$ invertierbar. Es folgt

$$Dg(x) = -\left(D_y F(x, g(x))\right)^{-1} D_x F(x, g(x)). \qquad \square$$

Der Satz über implizite Funktionen soll jetzt angewendet werden auf Extremwertaufgaben mit Nebenbedingungen.

<u>Problemstellung:</u> Gegeben sei eine offene Menge $\Omega \subset \mathbb{R}^n$ und eine Funktion $f : \Omega \to \mathbb{R}$. f soll optimiert werden, jedoch nicht auf ganz Ω, sondern nur auf einer Teilmenge

$$M = \{x \in \Omega : \varphi(x) = 0\},$$

die als Nullstellenmenge einer Funktion

$$\varphi : \Omega \to \mathbb{R}^\ell, \quad 1 \le \ell < n$$

gegeben ist.

<u>Bsp.</u> 1: $A \in M_n(\mathbb{R})$ sei symmetrisch. Gefragt ist nach den lokalen und globalen Extrema, die die Funktion

$$f : \mathbb{R}^n \to \mathbb{R}, \quad x \mapsto f(x) := \langle x, Ax \rangle$$

auf der Einheitssphäre $S^{n-1} \{x \in \mathbb{R}^n : |x| = 1\}$ annimmt. Die Menge M ist hier also die S^n, die wir als Nullstellengebilde von $\varphi(x) = |x|^2 - 1$ auffassen können.

(weitere Bspe. unten!)

Eine notwendige Bedingung für Extrema unter (D71)
Nebenbedingungen stellt der folgende Satz zur Verfügung:

<u>Satz</u> (über die Lagrange-Multiplikatoren): Es sei $\Omega \subset \mathbb{R}^n$ offen und

$$f : \Omega \to \mathbb{R}$$

sowie

$$\varphi : \Omega \to \mathbb{R}^\ell \qquad (1 \le \ell < n)$$

seien stetig d'bar. In $z_0 \in M := \{x \in \Omega : \varphi(x) = 0\}$ gelte

$$\operatorname{Rang} D\varphi(z_0) = \ell$$

und die Einschränkung von f auf M, also

$$f|_M : M \to \mathbb{R}$$

besitze in z_0 ein lokales Extremum. Dann existieren $\lambda_1, \dots, \lambda_\ell \in \mathbb{R}$, so daß

$$\nabla f(z_0) - \sum_{i=1}^{\ell} \lambda_i \nabla \varphi_i(z_0) = 0.$$

<u>Bez.</u>: Die Zahlen $\lambda_1, \dots, \lambda_\ell$ werden als Lagrange-Multiplikatoren bezeichnet.

Bew.: Wir wenden den Satz über implizite Funktionen an auf die Funktion $\varphi : \Omega \to \mathbb{R}^\ell$ und nehmen dazu o.E. an, daß

$$\det\left(\frac{\partial \varphi_i}{\partial z_j}(z_0)\right)_{\substack{1\le i\le \ell \\ n-\ell< j\le n}} \neq 0 \quad (*).$$

Die Variablen $z \in \Omega \subset \mathbb{R}^n$ schreiben wir als $z = (x_1, \dots, x_{n-\ell}, y_1, \dots, y_\ell)$ $=: (x, y)$ und $z_0 = (a, b)$ mit $a \in \mathbb{R}^{n-\ell}$ und $b \in \mathbb{R}^\ell$. Dann lautet $(*)$: $\det D_y \varphi(a,b) \neq 0$, $D_y \varphi(a,b)$ ist also invertierbar. Ferner gilt $\varphi(a,b) = 0$.

Damit sind die Voraussetzungen des Satzes erfüllt, es gibt also offene Umgebungen $W_a \subset \mathbb{R}^{n-\ell}$ von a und $W_b \subset \mathbb{R}^\ell$ von b, und eine Funktion $g \in C^1(W_a, W_b)$, so daß $g(a) = b$, $\varphi(x, g(x)) = 0 \ \forall x \in W_a$ und

$$Dg(x) = -D_y \varphi(x, g(x))^{-1} D_x \varphi(x, g(x)) \quad \forall x \in W_a.$$

Jetzt setzen wir $\tilde f(x) = f(x, g(x))$. Da f in $z_0 = (a,b)$ ein lokales Extremum besitzt, hat auch $\tilde f$ in a ein solches, also ist $0 = \nabla_x \tilde f(a) = \nabla_x f(z_0) + \nabla_y f(z_0) \cdot Dg(a)$, wobei $\nabla_x = \left(\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_{n-\ell}}\right)$ und $\nabla_y = \left(\frac{\partial}{\partial y_1}, \dots, \frac{\partial}{\partial y_\ell}\right)$.

Setzen wir die Formel für Dg ein, erhalten wir

$$\nabla_x f(z_0) = \underbrace{\nabla_y f(z_0)}_{\substack{\text{Spaltenvektor} \\ \text{mit } \ell \text{ Komponenten}}} \cdot \underbrace{D_y \varphi(z_0)^{-1}}_{\ell\times\ell\text{-Matrix}} \cdot \underbrace{D_x \varphi(z_0)}_{= \begin{pmatrix} - \nabla_x \varphi_1(z_0) - \\ \vdots \\ - \nabla_x \varphi_\ell(z_0) - \end{pmatrix}}$$

Wir setzen $(\lambda_1, \dots, \lambda_\ell) = \nabla_y f(z_0) (D_y \varphi(z_0))^{-1}$, so daß

$$\nabla_x f(z_0) = (\lambda_1,\dots,\lambda_e)\cdot D_x \varphi(z_0) = \sum_{i=1}^{e} \lambda_i \nabla_x \varphi_i(z_0)$$

und

$$\nabla_y f(z_0) = (\lambda_1,\dots,\lambda_e)\, D_y \varphi(z_0) = \sum_{i=1}^{e} \lambda_i \nabla_y \varphi_i(z_0).$$

Wg. $\nabla f = (\nabla_x f, \nabla_y f)$ können wir dies zusammenfassen zu

$$\nabla f(z_0) - \sum_{i=1}^{e} \lambda_i \nabla \varphi_i(z_0) = 0,$$ wie behauptet. □

Fortsetzung von Bsp. 1: Hier haben wir

$$f(x) = \langle x, Ax\rangle = \sum_{i,j=1}^{n} a_{ij} x_i x_j \quad \text{und} \quad \frac{\partial f}{\partial x_k}(x) = \sum_{i,j=1}^{n} a_{ij}(\delta_{ik}x_j + \delta_{jk}x_i)$$

$$= \sum_{j=1}^{n} a_{kj} x_j + \sum_{i=1}^{n} a_{ik} x_i \underset{A^T=A}{=} 2\sum_{j=1}^{n} a_{kj} x_j = 2(Ax)_k$$

Also $\nabla f(x) = 2Ax$ und, mit $A = I_n$,

$$\nabla \varphi(x) = \nabla(|x|^2 - 1) = 2x.$$

Die notwendige Bedingung lautet also

$$0 = \nabla(f - \lambda\varphi)(x) = Ax - \lambda x, \quad \text{d.h.} \quad Ax = \lambda x.$$

Die Funktion f nimmt ihre lokalen Extrema(*) also in den Eigenvektoren der Matrix A an. Der Lagrange-Multiplikator ist (in diesem Bsp.!) gerade ein Eigenwert von A. Für den Funktionswert in den kritischen Stellen erhalten wir

$$f(x) = \langle x, Ax\rangle = \langle x, \lambda x\rangle = \lambda|x|^2 = \lambda \quad (\text{da } |x| = 1).$$

(*) falls solche existieren, was nicht durch den Satz geliefert wird.

Die Frage der Existenz zumindest der globalen Extrema ist aber leicht zu beantworten. Die Funktion

$$f : S^n \to \mathbb{R}, \quad x \mapsto \langle x, Ax \rangle$$

ist stetig auf einem Kompaktum, nimmt also ihr (globales) Maximum und Minimum an. Es folgt

$$\max \{ \langle x, Ax \rangle : |x| = 1 \} = \max_{j=1}^{n} \{ \lambda_j : \lambda_j \text{ ist EW von } A \}$$

und entsprechend für das Minimum.

Folgerung: Es sei $B : \mathbb{R}^n \to \mathbb{R}^m$ eine reelle $m \times n$-Matrix (bzw. lineare Abbildung) und $\| B \| = \sup \{ |Bx| : |x| \leq 1 \}$.

Dann gilt $\| B \| = \max_{j=1}^{n} \{ \sqrt{\lambda_j} : \lambda_j \text{ ist EW von } B^T B \}$.

Bew.: $\| B \|^2 = \sup \{ |Bx|^2 : x \in \mathbb{R}^n, |x| \leq 1 \}$

$$= \sup \{ \langle Bx, Bx \rangle : \text{—} \,"\, \text{—} \}$$

$$= \sup \{ \langle x, B^T B x \rangle : x \in \mathbb{R}^n, |x| = 1 \}$$

stetig auf kompakt $\longrightarrow = \max \{ \langle x, B^T B x \rangle : x \in S^n \}$

Bsp. 1 $\longrightarrow = \max_{j=1}^{n} \{ \sqrt{\lambda_j} : \lambda_j \text{ ist Eigenwert von } B^T B \}$. □

Dieses Beispiel enthält bereits die wesentliche Idee für den Beweis des Spektralsatzes für reelle symmetrische Matrizen (mit analytischen Argumenten!), den wir benutzt haben, um das Eigenwertkriterium für die Definitheit symmetrischer Matrizen herzuleiten.

<u>Spektralsatz</u> für reelle symmetrische Matrizen:

Sei $A \in M_n(\mathbb{R})$ symmetrisch. Dann existieren reelle Zahlen $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ und eine ONB $v_1, \ldots, v_n$ von $\mathbb{R}^n$, so daß $\quad A v_i = \lambda_i v_i \quad \forall\, i \in \{1, \ldots, n\}$.

Bew.: Per Induktion über $k \in \{1, \ldots, n\}$ zeigen wir die folgende Aussage: Es existieren $v_1, \ldots, v_k \in \mathbb{R}^n$ und $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ mit

1. $\langle v_i, v_j \rangle = \delta_{ij}$
2. $A v_i = \lambda_i v_i$
3. $\lambda_1 \geq \ldots \geq \lambda_k$

Ind. Anfang: $k=1$. Vorhergehendes Bsp.

Ind.-Schritt: $k-1 \to k$ $(k \in \{2, \ldots, n\})$. Seien also $v_1, \ldots, v_{k-1}$; $\lambda_1, \ldots, \lambda_{k-1}$ bereits konstruiert. Wir setzen

$$f_k : S^{n-1} \cap H_k := \langle v_1, \ldots, v_{k-1} \rangle^\perp \longrightarrow \mathbb{R},\ x \mapsto f_k(x) := \langle x, Ax \rangle.$$

Dann ist f_k eine stetige Funktion auf einem Kompaktum, besitzt also eine Maximalstelle v_k. Für diese gilt $|v_k| = 1$ und $v_k \perp \langle v_1, \ldots, v_{k-1} \rangle$. Also ist Eigenschaft 1. von $\{v_1, \ldots, v_k\}$ erfüllt.

Nach dem Satz über die Lagrange-Multiplikatoren ~~gibt~~ es Zahlen $\mu, \mu_1, \ldots, \mu_{k-1}$, so daß

$$\nabla f_k(v_k) - \mu\, \nabla(|x|^2 - 1)\Big|_{x=v_k} - \sum_{i=1}^{k-1} \mu_i\, \nabla \langle x, v_i \rangle \Big|_{x=v_k} = 0.$$

mit $\nabla f_k(x) = 2Ax$, $\nabla(|x|^2-1) = 2x$ (vorher berechnet) D76

und $\nabla\langle x, v_i\rangle = v_i$ ergibt sich

$$2Av_k - 2\mu v_k - \sum_{i=1}^{k-1} \mu_i v_i = 0.$$

Für $0 \le \ell \le k-1$ bilden wir das Skalarprodukt dieser Gleichung mit v_ℓ:

$$0 = 2\underbrace{\langle v_\ell, Av_k\rangle} - 2\mu \underbrace{\langle v_\ell, v_k\rangle}_{=0} - \sum_{i=1}^{k-1} \mu_i \underbrace{\langle v_\ell, v_i\rangle}_{=\delta_{i\ell}} = \mu_\ell$$

$$= \langle Av_\ell, v_k\rangle$$

$$= \langle \lambda_\ell v_\ell, v_k\rangle = 0 \qquad \Longrightarrow \mu_1, \dots, \mu_{k-1} = 0.$$

Also $Av_k = \mu v_k$. Wählen wir $\lambda_k := \mu$, ist auch die Eigenschaft 2. erfüllt. Ferner ist

$$f_k(v_k) = \langle v_k, Av_k\rangle = \lambda_k \ge \langle x, Ax\rangle \quad \forall x \in S^{n-1} \cap H_k$$

Gleiches gilt mit $k-1$ anstelle von k, also

$$f_{k-1}(v_{k-1}) = \lambda_{k-1} \ge \langle x, Ax\rangle \quad \forall x \in S^{n-1} \cap H_{k-1} = \langle v_1, \dots, v_{k-2}\rangle^\perp$$

insbes. also $\lambda_{k-1} \ge \langle v_k, Av_k\rangle = \lambda_k$, womit auch die dritte Eigenschaft gezeigt ist. □

Bsp. 2: Abstände von gekrümmten Flächen (stücken) zu einem gegebenen Punkt x_0.

Gegeben sei eine stetig diff'bare Funktion

$$\varphi : \mathbb{R}^4 \supset \Omega \longrightarrow \mathbb{R}$$

und ein "Hyperflächenstück" $A := \{x \in \Omega : \varphi(x) = 0\}$, dargestellt als Nullstellengebilde der Funktion φ. Zu berechnen ist der Abstand $\mathrm{dist}(x_0, A)$!

Wir wissen bereits: Da A abgeschlossen ist, und $\{x_0\}$ kompakt, existiert $x_{min} \in A$, für dass

$$\mathrm{dist}(x_0, A) = d(x_0, x_{min}) = |x_{min} - x_0| .$$

Um diesen Abstand zu bestimmen, ~~~~ minimieren wir die Funktion

$$f(x) = |x - x_0|^2 \longleftarrow \text{einfacher zu rechnen!}$$

unter der Nebenbedingung $x \in A$, d.h. unter der NB

$$\varphi(x) = 0 .$$

Nach dem Satz über die Lagrange-Multiplikatoren existiert eine reelle Zahl λ, so daß

$$\nabla(f - \lambda\varphi)(x) = 0 ,$$

für $f(x) = |x-x_0|^2$ also :

$$2(x-x_0) = \lambda \nabla \varphi(x) \qquad \text{(Notwendige Bedingung)}$$

Für den Punkt x_{min}, in dem der Abstand der Hyperfläche A zu x_0 minimal wird, gilt also

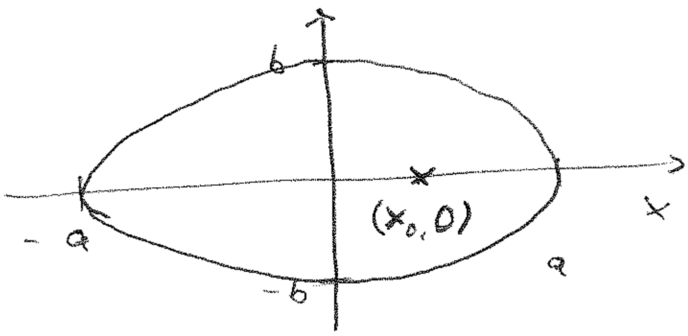
$$x_{min} - x_0 \parallel \nabla \varphi(x_{min})$$

Nun ist $A = \{x \in \Omega : \varphi(x) = 0\} = N_\varphi(0)$ eine Niveaumenge von φ, und wir haben festgestellt, daß der Gradient $\nabla\varphi(x_0)$ orthogonal ist zu jeder Tangente an A im Punkt x_0, somit zur Tangential(hyper-)ebene. Wir können also – etwas ungenau – feststellen:

"Der kürzeste Abstand eines Punktes x_0 zu einer (glatten) Fläche ist immer der senkrechte" (i.S.v.: senkrecht zur Tangentialebene).

<u>Konkretisierung</u>: $A = \partial E = \{(x,y) \in \mathbb{R}^2 : \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1\}$

sei der Rand einer Ellipse mit großer Halbachse a und kleiner Halbachse b; der Punkt $\vec{x}_0 = (x_0, 0)$ befinde sich auf der x-Achse.



Hier ist $\varphi(x,y) = \frac{x^2}{a^2} + \frac{y^2}{b^2} - 1$ und

$$\nabla\varphi(x,y) = 2\left(\frac{x}{a^2}, \frac{y}{b^2}\right),$$

so dass die notwendige Bedingung lautet

$$(x - x_0, y) = \lambda\left(\frac{x}{a^2}, \frac{y}{b^2}\right).$$

Berücksichtigen wir auch die NB, haben wir das folgende GLS. zu lösen:

$$x\left(1 - \frac{\lambda}{a^2}\right) = x_0 \ ; \quad y\left(1 - \frac{\lambda}{b^2}\right) = 0 \ ; \quad \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 .$$

1. Fall: $y = 0 \Rightarrow x = \pm a$ liefert tatsächlich zwei kritische Stellen $P_{1,2} = (\pm a, 0)$

2. Fall: $y \neq 0 \Rightarrow 1 - \frac{\lambda}{b^2} = 0 \rightarrow \lambda = b^2 \Rightarrow x\left(1 - \frac{b^2}{a^2}\right) = x_0$

$$\Rightarrow x\left(\frac{a^2 - b^2}{a^2}\right) = x_0 \Rightarrow x = \frac{a^2}{a^2 - b^2} \cdot x_0$$

$$\Rightarrow y = \pm b\sqrt{1 - \frac{x^2}{a^2}} = \pm b\sqrt{1 - \frac{a^2 x_0^2}{(a^2 - b^2)^2}},$$

was tatsächlich zwei weitere kritische Stellen liefert,

sofern $x_0^2 a^2 \leq (a^2 - b^2)^2 \Leftrightarrow x_0^2 \leq \frac{(a^2 - b^2)^2}{a^2}$

$$\Leftrightarrow |x_0| \leq \frac{a^2 - b^2}{a} \left(= \frac{c^2}{a}, \ c \text{ der Brennpunkt}\right).$$

In diesem Fall sind also auch

$$P_{3,4} = \left(\frac{a^2}{a^2 - b^2} x_0, \pm b\sqrt{1 - \frac{a^2 x_0^2}{(a^2 - b^2)^2}}\right) \text{ kritisch.}$$

Welcher dieser Punkte liefert nun den Abstand von $(x_0, 0)$ zu ∂E? Dazu nehmen wir o.E. $x_0 \geq 0$ an

$$\Rightarrow |P_1 - (x_0, 0)|^2 = (a - x_0)^2$$

Andererseits:

$$|P_{3,4} - (x_0, 0)|^2 = x_0^2 \left(\frac{a^2}{a^2-b^2} - 1\right)^2 + b^2\left(1 - \frac{a^2 x_0^2}{(a^2-b^2)^2}\right)$$

$$= x_0^2 \left(\frac{b^2}{a^2-b^2}\right)^2 + b^2 \left(\frac{(a^2-b^2)^2 - a^2 x_0^2}{(a^2-b^2)^2}\right)$$

$$= \frac{b^2}{(a^2-b^2)^2}\left(x_0^2 b^2 + (a^2-b^2)^2 - a^2 x_0^2\right)$$

liefert für $x_0 = 0$ das Min

$$= \frac{b^2}{a^2-b^2}\left(a^2 - b^2 - x_0^2\right) = b^2 - \frac{b^2 x_0^2}{a^2-b^2} \underset{?}{\leq} (a - x_0)^2$$

$$? \Leftarrow \quad b^2 - \frac{b^2 x_0^2}{a^2-b^2} \leq a^2 - 2a x_0 + x_0^2$$

$$\Leftarrow \quad 2a x_0 \leq a^2 - b^2 + x_0^2 \cdot \frac{a^2}{a^2-b^2}$$

$$\Leftarrow \quad 2 x_0 \leq \frac{a^2-b^2}{a} + \frac{a}{a^2-b^2} \cdot x_0^2 \quad \text{(gilt Young)}$$

<u>Ergebnis:</u>

$$\operatorname{dist}((x_0,0), \partial E) = \begin{cases} \frac{b}{\sqrt{a^2-b^2}} \sqrt{a^2 - b^2 - x_0^2}, & \text{falls } |x_0| < \frac{a^2-b^2}{a} \\ |a - |x_0||, & \text{falls } |x_0| \geq \frac{a^2-b^2}{a}. \end{cases}$$

Bsp. 3 Abstände von Kurven im $\mathbb{R}^n$ zu einem Punkt (mehrere L-Multiplikatoren)

Eine Kurve $C \subset \mathbb{R}^n$ sei gegeben als Durchschnitt von $n-1$ Hyperflächen: $C = \bigcap_{i=1}^{n-1} A_i$, dabei

$$A_i = \{x \in \Omega : \varphi_i(x) = 0\}, \quad \varphi_i : \Omega \to \mathbb{R}$$

(Dann können wir C auch schreiben als $C = \{x \in \Omega : \varphi(x) = 0\}$, $\varphi = \begin{pmatrix} \varphi_1 \\ \vdots \\ \varphi_{n-1} \end{pmatrix}$, so daß wir exakt die Formulierung des Satzes über die Lagrange-Multiplikatoren treffen.)

<u>Konkret:</u> Im $\mathbb{R}^3$ betrachten wir die Zylinder

$A_1 = \{(x,y,z) \in \mathbb{R}^3 : x^2 + y^2 = 1\}$ und

$A_2 = \{(x,y,z) \in \mathbb{R}^3 : y^2 + z^2 = 4\}$

Deren Durchschnitt $C = A_1 \cap A_2$ ist dann eine Kurve im $\mathbb{R}^3$.

Wir fragen für einen festen Punkt $x_0 \in \mathbb{R}^n$ nach $\operatorname{dist}(x_0, C)$. Wieder ist C abgeschlossen und $\{x_0\}$ kompakt, also existiert $x_{min} \in C$, so daß

$$\operatorname{dist}(x_0, C) = d(x_0, x_{min})$$

Eine notwendige Bedingung zum Auffinden von x_{min} liefert wieder der Satz über die Lagrange-Multiplikatoren. Wir suchen nach einem Minimum der Fktn.

$f(x) = |x - x_0|^2$ unter den NBen

$\varphi_1(x) = \dots = \varphi_{n-1}(x) = 0$, also muß

$$\circledast \quad \nabla f(x) - \sum_{i=1}^{n-1} \lambda_i \nabla \varphi_i(x) = 2(x - x_0) - \sum_{i=1}^{n-1} \lambda_i \nabla \varphi_i(x) \overset{!}{=} 0$$

sein ($n + n-1$ Gleichungen für ebensoviele Unbekannte, nämlich $x_1, \dots, x_n, \lambda_1, \dots, \lambda_{n-1}$).

Kommen wir zu den beiden Zylindern zurück und bestimmen dist$(0, C)$!

$$\nabla \varphi_1(x,y,z) = 2(x,y,0) \quad , \quad \nabla \varphi_2(x,y,z) = 2(0,y,z)$$

und $\circledast$ lautet

$$(x,y,z) - \lambda_1(x,y,0) - \lambda_2(0,y,z) = (0,0,0), \text{ also}$$

$$(1-\lambda_1)x \overset{!}{=} (1-\lambda_1-\lambda_2)y \overset{!}{=} (1-\lambda_2)z \overset{!}{=} 0$$

ferner $\quad x^2 + y^2 \overset{!}{=} 1 \qquad y^2 + z^2 \overset{!}{=} 4$

(1) $x = 0 \Rightarrow y^2 = 1 \Rightarrow z^2 = 3$, ferner können noch $\lambda_2 = 1, \lambda_1 = 0$ gewählt werden

(2) $y = 0 \Rightarrow x^2 = 1$ und $z^2 = 4$, wähle $\lambda_1 = \lambda_2 = 1$

(3) $z = 0 \Rightarrow y^2 = 4$ ↯

Wählen wir $x \neq 0$ und $y \neq 0$ und $z \neq 0$ müssen $\lambda_1 = \lambda_2 = 1$ und auch $\lambda_1 + \lambda_2 = 1$ sein ↯

Insgesamt $2 \times 4 = 8$ kritische Stellen, die ersten vier liefern ein Minimum: $\underline{\text{dist}((0,0,0), C) = 2}$

Bsp. 4 : Abstand zweier Hyperflächen

$$F_1 = \{x \in \mathbb{R}^n : \varphi_1(x) = 0\}, \quad F_2 = \{y \in \mathbb{R}^n : \varphi_2(y) = 0\}$$

und fragen nach $\operatorname{dist}(F_1, F_2)$. ($\varphi_{1,2} \in C^1$!)

Wir wissen: Ist eine dieser Flächen beschränkt, gibt es $x_0 \in F_1$ und $y_0 \in F_2$, so daß

$$\operatorname{dist}(F_1, F_2) = d(x_0, y_0) = |x_0 - y_0|$$

Damit geht unsere Abstandsberechnung in die Lösung der folgenden Extremwertaufgabe über:

minimiere $F : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R} \quad (x,y) \longmapsto |x-y|^2$

unter den NBen $\varphi_1(x) = 0$ und $\varphi_2(y) = 0$.

Notwendige Bedingung nach dem Satz über Lagrange-Multiplikatoren

$$0 = \nabla F(x,y) - \lambda_1 \nabla \varphi_1(x) - \lambda_2 \nabla \varphi_2(y) \qquad (\nabla = (\nabla_x, \nabla_y))$$

$$= 2\big((x,0) - (0,y)\big) - \lambda_1(\nabla_x \varphi_1(x), 0) - \lambda_2(0, \nabla_y \varphi_2(y))$$

Also muß
$$\left.\begin{aligned} 2x &= \lambda_1 \nabla_x \varphi_1(x) \\ \text{und } -2y &= \lambda_2 \nabla_y \varphi_2(y) \end{aligned}\right\} \; 2 \times n \text{ Gleichungen}$$

und die beiden NBen erfüllt sein.

Konkrete Beispiele hierzu werden schnell sehr aufwändig, eines werden wir in den Übungen diskutieren.