VL Grundlagen der mathematischen Optimierung

Mitschrift zur Vorlesung von Prof. Dr. Jürgen Guddat im WS 2001/2002
HU Berlin, Institut für Mathematik

7. Juli 2003

Inhaltsverzeichnis

1 Einleitung

1.1 Formulierung des allgemeinen Optimierungsproblems

- X allgemeiner Raum, z.B. normierter Raum oder $X = \mathbb{R}^n$
- $f : X \to \mathbb{R}$
- $M \subseteq X$

$$(P) \quad \min \{ f(x) \,|x \in M \}$$

allgem. Ziel: bestimme $\bar{x} \in M$ mit $f(\bar{x}) \leq f(x), \forall x \in M$

Bezeichnungen

$\bar{x}$ heißt globaler Minimalpunkt

$f(x)$ heißt Zielfunktion

M heißt zulässiger Bereich oder Restriktionsmenge

$x \in M$ heißt zulässiger Punkt

Definition 1.1 (Lokaler Minimalpunkt)

$\exists \varepsilon > 0$, $\bar{x} \in M$ mit $f(\bar{x}) \leq f(x), \forall x \in M \cap U_\varepsilon(\bar{x})$ und $U_\varepsilon(\bar{x}) := \{x \in X \,| \|x - \bar{x}\| \leq \varepsilon \}$

1.2 Klassen von Optimierungsproblemen

1. (P) heißt Optimierungsproblem in unendlich-dimensionalen Räumen (infinite Optimierung).
2. $X = \mathbb{R}^n$ und $M := \{x \in \mathbb{R}^n \,|h_i(x) = 0, g_j(x) \geq 0, i \in \{1, \ldots, m\}, j \in \{1, \ldots, s\}\}$ heißt endlich-dimensionale oder finite Optimierung. Wenn mindestens eine der Funktionen nichtlinear ist $\to$ nichtlineare Optimierung.
3. Sind alle Funktionen linear $\to$ lineare Optimierung.
4. Ist f eine konvexe Funktion und M eine konvexe Menge $\to$ konvexe Optimierung.
5. $M := \{x \in \mathbb{R}^n \mid h_i(x) = 0, g_j(x) \geq 0, i \in I, j \in J\}$ und $|I \cup J| = \infty$ $\to$ semiinfinite Optimierung.
6. $X = \mathbb{R}^n$ und M ist eine diskrete Punktmenge $\to$ diskrete Optimierung.

2 Eigenschaften konvexer Funktionen und konvexer Optimierungsprobleme ([1])

2.1 Eigenschaften konvexer Funktionen

Definition 2.1 (Konvexität einer Menge)

M heißt konvexe Menge, wenn $\forall x, y \in M, x \neq y$ gilt: $\lambda x + (1-\lambda) y \in M$ für alle $\lambda \in [0,1]$.

Es sei $M \subseteq \mathbb{R}^n$ konvex, nicht leer und $f : M \to \mathbb{R}$.

Definition 2.2 (Konvexität einer Funktion)

1. f heißt (streng) konvex auf M, wenn gilt:
$$f(\lambda x + (1-\lambda) y) \leq (<) \lambda f(x) + (1-\lambda) f(y), \forall x, y \in M, \lambda \in [0,1]$$
2. f heißt (streng) konkav auf M, gdw. $-f$ konvex auf M ist.

Definition 2.3

1. Interior$M$ (int$M$) ist die Menge aller Punkte $x_0 \in M$, für die ein $\varepsilon > 0$ ex., so daß $U_\varepsilon(x_0) := \{x \in \mathbb{R}^n \, | \|x - x_0\| < \varepsilon\}$ ganz in M liegt.
2. RelativInterior$M$ (relint$M$) ist die Menge aller Punkte $x_0 \in M$ mit $U_\varepsilon(x_0) \cap \mathrm{aff}M \subseteq M$.
3. aff$M$ ist der kleinste affine Unterraum (bzgl. Dimension), der M enthält.

Satz 2.1

Seien $M \subseteq \mathbb{R}^n$ konvex und $f : M \to \mathbb{R}$ konvex. Dann ist f auf relint$M$ stetig.

Definition 2.4

1. $\mathrm{epi}(f) := \{(x,a) \in M \times \mathbb{R} | f(x) \leq a\}$ heißt Epigraph von f.
2. $f^c := \{x \in M \, | f(x) \leq c\}, c \in \mathbb{R}$ heißt untere Niveaumenge.

Satz 2.2

Sei M eine konvexe Teilmenge des $\mathbb{R}^n$

1. $f : M \to \mathbb{R}$ ist konvex auf M, gdw. epi(f) konvex ist.
2. Wenn f konvex auf M ist, so ist f^c konvex für alle $c \in \mathbb{R}$.
3. Sind $f, g : M \to \mathbb{R}$ konvex, so sind auch $h_1(x) := f(x) + g(x)$ und $h_2(x) := \sup\{f(x), g(x)\}$ konvex.
4. Jensens'sche Ungleichung: f konvex auf M
$$f\left(\sum_{i=1}^n \lambda_i x_i\right) \leq \sum_{i=1}^n \lambda_i f(x_i), \forall x_i \in M, \lambda_i \geq 0, \sum \lambda_i = 1$$

Satz 2.3

Sei $f \in C^1(\mathbb{R}^n, \mathbb{R})$. Dann ist f (streng) konvex auf $\mathbb{R}^n$ gdw. $\forall x, y \in \mathbb{R}^n$ gilt:
$$f(y) - f(x) \geq (>) Df(x)(y-x).$$

Bew.: $(\Rightarrow)$: Sei $\phi(\lambda) := f(x + \lambda(y-x)), \lambda \in [0,1] \Rightarrow \phi'(\lambda) = Df(x+\lambda(y-x)) \cdot (y-x)$

Taylor um 0 : $\phi(\lambda) = \phi(0) + \lambda\phi'(0) + o(\lambda)$ (Def.: $o: \mathbb{R} \to \mathbb{R}$ stetig in 0, $\frac{o(|\lambda|)}{\lambda} \overset{\lambda\to 0}{\to} 0$)

Konv.:

$$\begin{aligned}
(1-\lambda) f(x) + \lambda f(y) &\geq f((1-\lambda)x + \lambda y) \\
&= f(x + (y-x)\cdot\lambda) = \phi(\lambda) \\
&= f(x) + \lambda \cdot Df(x)(y-x) + o(\lambda) \\
\Rightarrow -\lambda f(x) + \lambda f(y) &\geq \lambda Df(x)(y-x) + o(\lambda) \\
\Rightarrow f(y) - f(x) &\geq Df(x)(y-x) + \underbrace{\frac{o(\lambda)}{\lambda}}_{\overset{\lambda\to 0^+}{\to} 0}
\end{aligned}$$

$(\Leftarrow)$ Mit $\lambda \in (0,1)$ und $x := \lambda y + (1-\lambda) z$ folgt aus $(\Rightarrow)$ $\forall y, z \in \mathbb{R}^n$:

$$\begin{aligned}
f(y) - f(x) &\geq Df(x)(y-x) \\
f(z) - f(x) &\geq Df(x)(z-x) \\
\overset{+}{\Rightarrow} \lambda f(y) + (1-\lambda) f(z) - f(x) &\geq Df(x)\underbrace{[\lambda(y-x) + (1-\lambda)(z-x)]}_{=0}
\end{aligned}$$

Satz 2.4

Sei $f \in C^2(\mathbb{R}^n, \mathbb{R})$. f ist konvex, gdw. $D^2 f(x)$ positiv semidefinit $\forall x \in \mathbb{R}^n$.

Bew.: $(\Leftarrow)$: Seien $x, y \in \mathbb{R}^n$ bel. Taylor:

$$\begin{aligned}
f(y) &= f(x) + Df(x)(y-x) + \frac{1}{2}(y-x)^T D^2 f(x)(y-x) \\
z &= x + \Theta(y-x), \Theta \in (0,1)
\end{aligned}$$

$D^2 f(z)$ ist pos. semidef. $\forall x, y \in \mathbb{R}^n$

$$\begin{aligned}
\Rightarrow (y-x)^T D^2 f(x)(y-x) &\geq 0 \\
\Rightarrow f(y) - f(x) &\geq Df(x)(y-x) \overset{S.2.3}{\Rightarrow} \text{Beh.}
\end{aligned}$$

$(\Rightarrow)$ siehe Ende vom 4. Kapitel

Corollar 2.1

Seien $f \in C^2(\mathbb{R}^n, \mathbb{R})$ und $D^2 f(x)$ positiv definit $\forall x \in \mathbb{R}^n$. Dann ist f streng konvex. Die Umkehrung ist im allgemeinen nicht gültig.

2.2 Eigenschaften konvexer Optimierungsprobleme

(P) $\min\{f(x) | x \in M\}$

$M := \{x \in \mathbb{R}^n | h_i(x) = 0, g_j(x) \leq 0, i \in I = \{1, \ldots, m\}, j \in J = \{1, \ldots, s\}\}$
$h_i(x) := a_i^T x = b_i$, $i = 1, \ldots, m$ affin-lineare Funktion
$g_j(x)$, $j = 1, \ldots, s$ konvexe Funktion und f konvex
$\Rightarrow M$ ist eine konvexe Menge

konvexe Optimierung: $\min\{f(x) | x \in M\}$, f konvex, M konvex
$\Rightarrow$ dann sind die folgenden Probleme ebenfalls konvexe Probleme:

1. $\min \{f(x) \,|h_i(x) = 0,\, g_j(x) \geq 0,\, i \in I,\, j \in J\}$
   mit f konvex, g_j konkav, h_i affin-linear

2. $\max \{f(x) \,|h_i(x) = 0,\, g_j(x) \leq 0,\, i \in I,\, j \in J\}$
   mit f konkav, g_j konvex, h_i affin-linear

3. $\max \{f(x) \,|h_i(x) = 0,\, g_j(x) \geq 0,\, i \in I,\, j \in J\}$
   mit f konkav, g_j konkav, h_i affin-linear

(P) $\min \{f(x) \,|x \in M\}$, $M \subseteq \mathbb{R}^n$ konvex, f konvex

Satz 2.5

Jeder lokale Minimalpunkt von (P) ist globaler Minimalpunkt.

Bew.: Sei $\bar{x}$ lok. Min.pkt. und kein glob. Min.pkt. $\rightarrow \exists y \in M : f(y) < f(\bar{x})$
Da M konvex, gilt $\lambda y + (1-\lambda)\bar{x} \in M \,\forall \lambda \in [0,1]$
Da f konvex, gilt

$$\begin{aligned} f(\lambda y + (1-\lambda)\bar{x}) &\leq \lambda f(y) + (1-\lambda) f(\bar{x}) \\ &< \lambda f(\bar{x}) + (1-\lambda) f(\bar{x}) \\ &= f(\bar{x}). \end{aligned}$$

Da λ beliebig klein gewählt werden kann, ist $\bar{x}$ kein lok. Min.pkt. $\rightarrow$ Wid.

Corollar 2.2

Jeder lokale Maximalpunkt (Minimalpunkt) eines linearen Optimierungsproblems ist globaler Maximalpunkt (Minimalpunkt).

$$\max(\min) \left\{ \sum_{\alpha=1}^{n} c_\alpha x_\alpha \,\middle|\, \sum_{\alpha=1}^{n} a_{r\alpha} x_\alpha \leq b_r,\, r = 1, \ldots, m,\, \sum_{\alpha=1}^{n} \tilde{a}_{r\alpha} x_\alpha \leq \tilde{b}_r,\, r = 1, \ldots, \tilde{m} \right\}$$

Satz 2.6

Die Optimalmenge $M_{opt} := \{x \in M \,| f(x) \leq f(y),\, \forall y \in M\}$ von (P) ist eine konvexe Menge.

Bew.: Seien $\hat{x}, \hat{y} \in M_{opt}$. Dann gilt $\hat{x}, \hat{y} \in M$ und $f(\hat{x}) \leq f(x), f(\hat{y}) \leq f(x) \;\forall x \in M$.
$\rightarrow z := \alpha\hat{x} + (1-\alpha)\hat{y} \in M \,\forall \alpha \in (0,1)$, da M konvex.
f konvex: $f(z) \leq \alpha f(\hat{x}) + (1-\alpha) f(\hat{y}) \leq \alpha f(x) + (1-\alpha) f(x) = f(x) \;\forall x \in M \Rightarrow z \in M_{opt}$.

Satz 2.7

Seien $M \subseteq \mathbb{R}^n$ konvex und $f : M \rightarrow \mathbb{R}$ streng konvex. Wenn $M_{opt} \neq \emptyset$, so ist M_{opt} einelementig.

Bew.: Ann.: $\{x^0, x^1\} \subseteq M_{opt}, x^0 \neq x^1$. Da f streng konvex auf M, gilt

$$\begin{aligned} f\left(\alpha x^0 + (1-\alpha) x^1\right) &< \alpha f\left(x^0\right) + (1-\alpha) f\left(x^1\right) \\ &\leq \alpha f\left(x^0\right) + (1-\alpha) f\left(x^0\right) \\ &= f\left(x^0\right) \;\forall \alpha \in (0,1). \end{aligned}$$

Wid. zu M_{opt} konvex bzw. $f\left(x^0\right)$ glob. Minpkt.

3 Optimalitätsbedingungen für freie Minimumprobleme und einfach restriktionierte Probleme

3.1 Einführende Bemerkungen

(P) $\min\{f(x)\,|x \in M\}$, $M := \{x \in \mathbb{R}^n\,|h_i(x) = 0,\, g_j(x) \leq 0,\, i \in I = \{1,\ldots,m\},\, j \in J = \{1,\ldots,s\}\}$

$\left(\tilde{P}\right)$ $\min\{f(x)\,|x_i \in \mathbb{R},\, i = 1,\ldots,p,\, x_i \geq 0,\, i = p+1,\ldots,n\}$ heißt einfach restrikt. Problem

Definition 3.1

$x \in M$ heißt strikter lokaler Minimalpunkt für (P), wenn eine Umgebung $U(x)$ existiert, so daß $f(x) < f(y)$, $\forall y \in M \cap U(x)$.

Bezeichnung

$\mathbb{H}^n := \{x \in \mathbb{R}^n\,|x \geq 0\}$ heißt nichtnegativer Orthant.

3.2 Optimalitätsbedingung 1. Ordnung

Satz 3.1

Sei $f \in C^1(\mathbb{R}^n, \mathbb{R})$ und $\bar{x} \in \mathbb{R}^n$ lokales Minimum für $\min\{f(x)\,|x \in \mathbb{R}^n\}$. Dann gilt:

$$Df(\bar{x}) = 0.$$

Definition 3.2

Es sei $f \in C^1(\mathbb{R}^n, \mathbb{R})$. Ein Punkt $\bar{x} \in \mathbb{R}^n$ mit $Df(\bar{x}) = 0$ heißt stationärer Punkt oder kritischer Punkt. Ein stationärer Punkt, der weder lokaler Minimalpunkt noch Maximalpunkt ist, heißt Sattelpunkt.

Satz 3.2

Es sei $f \in C^1(\mathbb{R}^n, \mathbb{R})$ und $0 \in \mathbb{H}^n$ ein lokaler Minimalpunkt für $\min\{f(x)\,|x \in \mathbb{H}^n\}$. Dann gilt:

$$\frac{\partial f}{\partial x_i}(0) \geq 0,\, i = 1,\ldots,n$$

Bew.: Ann.: $\exists i^* \in \{1,\ldots,n\}$ mit $\frac{\partial f}{\partial x_{i^*}}(0) < 0$.

Sei $\tilde{x}_{i^*} > 0$ fest. Bilden $\phi(t) := f(0,\ldots,t\cdot\tilde{x}_{i^*},0\ldots,0)$. Taylor um $t = 0$:

$$\begin{aligned}
\phi(t) &= \phi(0) + t\cdot\phi'(0) + o(t) = \phi(0) + t\left(\phi'(0) + \frac{o(t)}{t}\right),\, t > 0 \\
\phi'(0) &= \frac{\partial f}{\partial x_i^*}(0) < 0 \\
\rightarrow \phi'(0) &< \frac{1}{2}\phi'(0)
\end{aligned}$$

Da $\frac{o(t)}{t} \overset{t\to 0}{\to} 0$, ex. ein $\bar{t} > 0$, so daß

$$\begin{aligned}
\phi'(0) + \frac{o(t)}{t} &\leq \frac{1}{2}\phi'(0)\;\forall t \in (0,\bar{t}) \\
\rightarrow \phi(t) &\leq \phi(0) + \frac{1}{2}\phi'(0)\cdot t\,\forall t \in (0,\bar{t}) \\
\overset{\phi(0)<0}{\rightarrow} \phi(t) &< \phi(0)\;\forall t \in (0,\bar{t}) \\
\rightarrow f(0,\ldots,0,t\cdot\tilde{x}_{i^*},0,\ldots,0) &< f(0)\;\forall t \in (0,\bar{t})
\end{aligned}$$

und $(0,\ldots,0,t\cdot\tilde{x}_{i^*},0,\ldots,0) \in U \cap \mathbb{H}^n\;\forall t \in (0,\bar{t})$. Wid. zu 0 ist lok. Min.pkt.

Satz 3.3

Es sei $f \in C^1(\mathbb{R}^n, \mathbb{R})$ und 0 lokaler Minimalpunkt für $\min\{f(x) | \mathbb{R}^p \times \mathbb{H}^q\}, n = p+q$. Dann gilt:

$$\frac{\partial f}{\partial x_i}(0) = 0,\ i = 1, \ldots, p \qquad \frac{\partial f}{\partial x_j} \geq 0,\ j = p+1, \ldots, n$$

Bezeichnung

$B(\bar{x}, r) := \{x \in \mathbb{R}^n | \|x - \bar{x}\| \leq r\}$ mit $r > 0$, $\bar{x} \in \mathbb{R}^n$ fest

Satz 3.4

Seien $f \in C^1(\mathbb{R}^n, \mathbb{R})$ und $\frac{\partial f}{\partial x_i}(0) > 0$, $i = 1, \ldots, n$. Dann gibt es eine Umgebung U von 0, und $\gamma > 0$, so daß

$$f(x) \geq f(0) + \gamma \sum_{i=1}^{n} x_i,\ \forall x \in U \cap \mathbb{H}^n$$

Insbesondere ist $0 \in \mathbb{H}^n$ ein strikter lokaler Minimalpunkt für $\min\{f(x) | x \in \mathbb{H}^n\}$.

Bew.: Letzte Aussage ist klar, da $\sum_{i=1}^n \bar{x}_i > 0$, falls $\bar{x} \in U \cap \mathbb{H}^n$ und $\bar{x} \neq 0$.

Da $f \in C^1(\mathbb{R}^n, \mathbb{R})$, sind $\frac{\partial f}{\partial x_i}, i = 1, \ldots, n$, stetige Funktionen.

Wegen $\frac{\partial f}{\partial x_i}(0) > 0, i = 1, \ldots, n$, gibt es ein $r > 0$ und $\gamma_i > 0$, so daß $\frac{\partial f}{\partial x_i}(0) \geq \gamma_i > 0\, \forall x \in B(0, r)$. Setze

$$\gamma := \min_{i=1,\ldots,n} \gamma_i \Rightarrow \gamma > 0.$$

Hauptsatz der Integral- und Diff.rechnung:

$$\begin{aligned}
f(x) - f(0) &= \int_0^1 \frac{d}{dt} f(tx)\, dt = \int_0^1 \frac{d}{dt} f(tx_1, \ldots, tx_n)\, dt \\
&= \int_0^1 \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(tx)\, x_i dt \\
&= \sum_{i=1}^{n} \left[\int_0^1 \frac{\partial f}{\partial x_i}(tx)\, dt\right] x_i \\
&\geq \sum_{i=1}^{n} \left[\int_0^1 \gamma dt\right] x_i \\
&= \gamma \cdot \sum_{i=1}^{n} x_i \Rightarrow \text{Beh.}
\end{aligned}$$

3.3 Optimalitätsbedingung 2. Ordnung

Satz 3.5

Es sei $f \in C^2(\mathbb{R}^n, \mathbb{R})$ und $\bar{x} \in \mathbb{R}^n$ ein lokaler Minimalpunkt für $\min\{f(x) | x \in \mathbb{R}^n\}$. Dann gilt $Df(\bar{x}) = 0$ und $D^2 f(\bar{x})$ ist positiv semidefinit.

Bew.: Aus Satz 3.1 folgt $Df(\bar{x}) = 0$.

Ann.: $D^2 f(\bar{x})$ ist nicht pos. semidefinit. Dann gibt es $\xi \in \mathbb{R}^n$ mit $\xi^T D^2 f(\bar{x})\, \xi =: \alpha < 0$.

$\phi(t) := f(\bar{x} + t\xi)$.

Taylor für $t = 0$: $\phi(t) = \phi(0) + t \cdot \phi'(0) + \frac{1}{2} t^2 \phi'(0) + o(t^2)$ mit $\frac{o(t^2)}{t^2} \overset{t \to 0}{\to} 0$.

Es gilt

$$
\begin{array}{rcl}
\phi'(t) & = & Df(\bar{x}+t\xi)\,\xi = \xi^T D^T f(\bar{x}+t\xi) \overset{Df(\bar{x})=0}{\to} \phi'(0) = 0 \\
\to \phi''(t) & = & \xi^T D^2 f(\bar{x}+t\xi)\,\xi = \alpha < 0 \\
\to \phi(t) & = & \phi(0) + t^2 \left[\dfrac{1}{2}\alpha + \dfrac{o(t^2)}{t^2}\right],\ t \neq 0 \\
\lim\limits_{t\to 0} \dfrac{o(t^2)}{t^2} = 0 & \Rightarrow & \exists \bar{t} > 0 : \dfrac{1}{2}\alpha + \dfrac{o(t^2)}{t^2} \leq \dfrac{1}{4}\alpha\, \forall t \in (0,\bar{t}) \\
\to \phi(t) & \leq & \phi(0) + \dfrac{1}{4}t^2\alpha \overset{\alpha<0}{<} \phi(0)\ \forall t \in (0,\bar{t}) \\
\to f(\bar{x}+t\xi) & < & f(\bar{x})\ \forall t \in (0,\bar{t})
\end{array}
$$

Wid. zu $\bar{x}$ lok. Min.pkt.

Satz 3.6

Es seien $f \in C^2(\mathbb{R}^n, \mathbb{R})$, $n = p + q$ und $0 \in \mathbb{R}^p \times \mathbb{H}^q$ ein lokaler Minimalpunkt für $\min\{f(x)\,|x \in \mathbb{R}^p \times \mathbb{H}^q\}$. Dann gilt:

$$
\frac{\partial f}{\partial x_i}(0) = 0,\ i = 1,\ldots,p \qquad \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(0)\right)_{i,j=1,\ldots,p} \text{ positiv semidefinit}
$$

$$
\frac{\partial f}{\partial x_j}(0) \geq 0,\ j = p+1,\ldots,n
$$

Satz 3.7

Es seien $f \in C^2(\mathbb{R}^n, \mathbb{R})$, $n = p + q$ und die folgenden Bedingungen erfüllt:

1. $\frac{\partial f}{\partial x_i}(0) = 0$, $i = 1,\ldots,p$, $\left(\frac{\partial^2 f}{\partial x_i \partial x_j}(0)\right)$ positiv definit
2. $\frac{\partial f}{\partial x_j}(0) > 0$, $j = p+1,\ldots,n$

Dann existiert eine Umgebung U von 0 und ein $\gamma > 0$, so daß

$$
f(x) \geq f(0) + \gamma\,\|x\|^2,\ \forall x \in U \cap (\mathbb{R}^p \times \mathbb{H}^q)
$$

Insbesondere ist $0 \in \mathbb{R}^p \times \mathbb{H}^q$ ein strikter lokaler Minimalpunkt für $\min\{f(x)\,|x \in \mathbb{R}^p \times \mathbb{H}^q\}$.

Lemma 3.1

Es sei $M \subseteq \mathbb{R}^n$ kompakt, nicht leer und $f \in C^0(\mathbb{R}^n \times \mathbb{R}^m, \mathbb{R})$, $z = (x,y)$, $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$ und

$$
\phi(x) := \min_{y\in M} f(x,y) \qquad \tilde{\phi}(x) := \max_{y\in M} f(x,y)
$$

Dann sind ϕ und $\tilde{\phi}$ stetig auf M.

Bew.: Wir zeigen die Stetigkeit von ϕ :

Sei $\{x^k\}$ bel. mit $x^k \to \bar{x}$. z.z.: $\phi(x^k) \to \phi(\bar{x})$

Nach dem Satz von Weierstraß ex. zu jedem x^k ein $y^k \in M$ mit $\phi(x^k) = f(x^k, y^k)$ (da M kompakt).

Ann.: $\phi(x^k) \not\to \phi(\bar{x})$

$\to (*)\ \exists \varepsilon > 0$ und Teilfolge $\{x^{k_i}\} \subseteq \{x^k\}$, so daß $\phi(x^{k_i}) \notin (\phi(\bar{x}) - \varepsilon, \phi(\bar{x}) + \varepsilon)$, $i \in \mathbb{N}$

$\{y^{k_i}\}$ hat einen Häufunktspunkt $\bar{y} \in M$, da M kompakt.

o.B.d.A. $y^{k_i} \to \bar{y}$ (ansonsten wählen wir eine geeignete Teilfolge)

$\Rightarrow$ $\left(x^{k_i}, y^{k_i}\right) \to (\bar{x}, \bar{y})$ $\Rightarrow$(Stetigkeit) $f\left(x^{k_i}, y^{k_i}\right) \to f(\bar{x}, \bar{y})$.
Andererseits folgt aus $(*)$: $f(\bar{x}, \bar{y}) \neq \phi(\bar{x})$
$\to$ $\exists \tilde{y} \in M$ und $\alpha > 0$ mit $f(\bar{x}, \tilde{y}) \leq f(\bar{x}, \bar{y}) - \alpha$ ($\tilde{y}$ realisiert $\min_{y \in M} f(\bar{x}, y)$)
$\to$ $\exists i_0$ mit $f\left(x^{k_i}, \tilde{y}\right) \leq f(\bar{x}, \bar{y}) - \frac{\alpha}{2}$ für $i \geq i_0$.
Wegen $f\left(x^{k_i}, y^{k_i}\right) \leq f\left(x^{k_i}, \tilde{y}\right)$ gilt: $f\left(x^{k_i}, y^{k_i}\right) \leq f(\bar{x}, \bar{y}) - \frac{\alpha}{2}$ für $i \geq i_0$. Wid. zu $f\left(x^{k_i}, y^{k_i}\right) \to f(\bar{x}, \bar{y})$.

Lemma 3.2

Es sei $x \in \mathbb{R}^p$ und $y \in \mathbb{R}^q$ und $A : (x, y) \to A(x, y)$ eine stetige Abbildung von $\mathbb{R}^p \times \mathbb{R}^q$ in den linearen Raum der symmetrischen $p \times p$-Matrizen (d.h. jedes Matrixelement $a_{ij}(x, y)$ ist eine stetige Funktion in (x, y)). Weiterhin sei $A(0,0)$ positiv definit. Dann gibt es eine Umgebung $U \subset \mathbb{R}^p \times \mathbb{R}^q$ von $(0,0)$ um $\gamma > 0$, so daß

$$x^T A(x, y)\, x \geq \gamma \|x\|^2, \; \forall (x, y) \in U$$

Bew.: Setze $\eta(x, y) := x^T A(x, y)\, x$.

1. $x = 0, \eta(0, y) = 0$ trivial für jedes γ
2. $x \neq 0$. Dann gilt

$$\begin{aligned} \eta(x, y) &= \left[\left(\frac{x}{\|x\|}\right)^T A(x, y) \left(\frac{x}{\|x\|}\right)\right] \|x\|^2 \\ &\geq \min_{\|\xi\|=1} \xi^T A(x, y)\, \xi \cdot \|x\|^2 \end{aligned}$$

$\to$(Lemma 3.1) $\phi(x, y) := \min_{\|\xi\|=1} \xi^T A(x, y)\, \xi$ stetig.
Es ist $A(0,0)$ pos. def. $\Rightarrow$ $\phi(0,0) = c > 0$
ϕ stetig $\to$ $\phi(x, y) \geq \frac{c}{2} > 0\, \forall (x, y)$ in einer Umgebung $U \subseteq \mathbb{R}^p \times \mathbb{R}^q$ von $(0,0)$.
Setze $\gamma := \frac{c}{2}$.

Lemma 3.3

Es sei $g : \mathbb{R}^n \to \mathbb{R}$ stetig und $g(0) > 0$. Dann existiert zu jedem $\gamma > 0$ eine Umgebung $U_\gamma \subset \mathbb{R}^n$ von 0, so daß $\forall i = 1, \ldots, n$ gilt:

$$x_i g(x) \geq \gamma x_i^2, \; \forall x \in U_\gamma \text{ mit } x_i \geq 0$$

Bew.: Sei $\delta > 0$ bel. Wir betrachten $\phi_i(x) := g(x) - \delta x_i, i = 1, \ldots, n$.
ϕ_i stetig auf $\mathbb{R}^n$. Außerdem gilt: $\phi_i(0) = g(0) > 0\, \forall i = 1, \ldots, n$.
$\to$ $\exists$ Umgebung U_δ von 0 mit $\phi_i(x) > 0\, \forall i = 1, \ldots, n\, \forall x \in U_\delta$
$\to$ $x_i \phi_i(x) > \delta^2 x_i^2\, \forall x \in U_\delta$ mit $x_i \geq 0$.

Bew. S.3.7: Es genügt, $\exists U(0), \gamma :\; f(x) \geq f(0) + \gamma \|x\|^2\; \forall x \in U \cap (\mathbb{R}^p \times \mathbb{H}^q)$ zu zeigen.

Hauptsatz der Diff.- und Int.rechnung:

$$\begin{aligned} f(x) - f(0) &= \int_0^1 \frac{d}{dt} f(tx)\, dt \\ &= \sum_{i=1}^n x_i g_i,\; g_i(x) = \int \frac{\partial g_i}{\partial x_i}(tx)\, dt,\; i = 1, \ldots, n \end{aligned}$$

$f \in C^2 \;\Rightarrow\; g_i \in C^1, i = 1, \ldots, n$

Vor. 1 $\Rightarrow$ $g_i(0) = 0$, $i = 1, \ldots, p$. Hauptsatz der Diff.- und Int.rechnung:

$$\begin{aligned}
g_i(x) &= \int_0^1 \frac{d}{dt} g_i(tx)\, dt \\
&= \sum_{j=1}^{n} x_j h_{i,j}(x),\ h_{i,j}(x) = \int_0^1 \frac{\partial g_i}{\partial x_j}(x)\, dt,\ h_{i,j} \in C^0, i = 1, \ldots, p \\
\Rightarrow f(x) &= f(0) + \sum_{i=1}^{p} \sum_{j=1}^{n} x_i x_j h_{i,j}(x) + \sum_{j=p+1}^{n} x_j g_j(x)
\end{aligned}$$

Sortierung $\tilde{h}_{i,j} = \frac{1}{2}(h_{i,j} + h_{j,i})$, $\tilde{g}_j(x) = g_j(x) + \sum_{i=1}^{p} h_{i,j}(x)\, x_i$

$$\Rightarrow f(x) = f(0) + \underbrace{\sum_{i,j=1}^{p} x_i x_j \tilde{h}_{i,j}(x)}_{\geq \gamma \|x\|^2 \text{ (L.3.2)}} + \underbrace{\sum_{j=p+1}^{n} x_j \tilde{g}_j(x)}_{> \gamma \|x\| \text{ (L.3.3.)}}$$

$\hat{g}_j(0) = \frac{\partial f}{\partial x_j}(0) > 0$ nach Vor. 2.

4 Optimalitätsbedingungen für allg. Optimierungsprobleme mit Nebenbedingungen

4.1 Optimalitätsbedingungen für nichtlineare Optimierungsprobleme

(P) $\min \{f(x) \,|x \in M\}$, $f : \mathbb{R}^n \to \mathbb{R}$, $M \subseteq \mathbb{R}^n$

$M = M[h, g] = \{x \in \mathbb{R}^n \,|h_i(x) = 0, i \in I, g_j(x) \geq 0, j \in J\}$

$h = \begin{pmatrix} h_1 & \cdots & h_m \end{pmatrix}^T$, $g = \begin{pmatrix} g_1 & \cdots & g_s \end{pmatrix}^T$, $I = \{1, \ldots, m\}$, $J = \{1, \ldots, s\}$

$J_0(x) := \{j \in J \,|g_j(x) = 0\}$ Indexmenge der aktiven Restriktionen

Satz 4.1

Es sei $h_i, g_j \in C^0(\mathbb{R}^n, \mathbb{R})$, $i \in I$, $j \in J$. Dann ist M eine abgeschlossene Menge, und es gilt: Zu jedem $\bar{x} \in M$ gibt es eine Umgebung U von $\bar{x}$, so daß:

$$J_0(x) \subseteq J_0(\bar{x}),\ \forall x \in U \cap M$$

Definition 4.1

Es sei $h_i, g_j \in C^1(\mathbb{R}^n, \mathbb{R})$, $i \in I$, $j \in J$ und $\bar{x} \in M$.

1. Die lineare Unabhängigkeitsbedingung (LICQ) ist in $\bar{x} \in M$ erfüllt, wenn die Vektoren $Dh_i(\bar{x})$, $i \in I$, $Dg_j(\bar{x})$, $j \in J_0(\bar{x})$ linear unabhängig sind. Falls die LICQ in jedem Punkt von M erfüllt ist, so sagt man die LICQ ist auf M erfüllt.
2. Die Mangasarian-Fromoritz-Bedingung (MFCQ) ist in $\bar{x} \in M$ erfüllt, falls:
   (a) $Dh_i(x)$, $i \in I$ linear unabhängig
   (b) $\exists \xi \in R^n$ mit $Dh_i(\bar{x})\xi = 0$, $i \in I$, $Dg_j(\bar{x})\xi > 0$, $j \in J_0(\bar{x})$

Bemerkung 4.1

LICQ $\to$ MFCQ

Lemma 4.1

Wenn LICQ in $\bar{x} \in M$ erfüllt ist, so gibt es eine Umgebung U von $\bar{x}$, so daß LICQ erfüllt ist $\forall x \in U \cap M$.

Satz 4.2

Es sei $h_i, g_j \in C^k(\mathbb{R}^n, \mathbb{R}),\ i \in I, j \in J$ mit $k > 1$ und in $\bar{x} \in M$ die LICQ erfüllt $(J_0(\bar{x}) = \{1, \ldots, p\})$. Dann existiert eine Umgebung U von $\bar{x}$ und V von $0 \in \mathbb{R}^n$ und ein C^k-Diffeomorphismus $\phi : U \to V$, so daß

$$\phi(\bar{x}) = 0 \text{ und } \phi(U \cap M) = V \cap \left(\{0_m\} \times \mathbb{H}^p \times \mathbb{R}^{n-m-p}\right),\ 0_m \in \mathbb{R}^m$$

$$\phi(x) = (\phi_1(x), \ldots, \phi_n(x)),\ x \in \mathbb{R}^n,\ \phi_i(x) \in C^k(\mathbb{R}^n, \mathbb{R}),\ i = 1, \ldots, n$$

Bew.: LICQ $\to$ $D^T h_i(\bar{x}), i = 1, \ldots, m$, $D^T g_j(\bar{x}), j = 1, \ldots, p$ lin. unabh. und $m + p \leq n$.
Basisergänzungssatz $\to$ $\exists \xi^{m+p+1}, \ldots, \xi^n \in \mathbb{R}^n$, so daß wir zu einer Basis im $\mathbb{R}^n$ ergänzen.
$y = \phi(x), \phi \in C^k(\mathbb{R}^n, \mathbb{R}^m)$ wie folgt:

$$\begin{array}{ccc} y_1 = h_1(x), & y_{m+1} = g_1(x), & y_{m+p+1} = {\xi^{m+p+1}}^T (x - \bar{x}) \\ \vdots & \vdots & \vdots \\ y_m = h_m(x), & y_{m+p} = g_p(x), & y_n = {\xi^n}^T (x - \bar{x}) \end{array}$$

Es gilt $\phi(\bar{x}) = 0$ und die Jacobi-Matrix

$$D\phi(\bar{x}) = \begin{bmatrix} D^T h_1(\bar{x}) & \cdots & D^T h_m(\bar{x}) & D^T g_1(\bar{x}) & \cdots & D^T g_p(\bar{x}) & \xi^{m+p+1} & \cdots & \xi^n \end{bmatrix}^T$$

ist regulär.
Inversentheorem $\Rightarrow \exists U, V$ von $\bar{x}, 0 \in \mathbb{R}^n$, so daß $\phi : U \to V$ ein C^k-Diff. Durch eventuelle Verkleinerung von U kann man $J_0(x) \subseteq J_0(\bar{x}) \forall x \in U \cap M$ erreichen.
$U \cap M$ wird in y-Koo. wie folgt beschrieben:

$$\begin{array}{rcl} y_1 = \ldots = y_m & = & 0 \\ y_{m+1}, \ldots, y_m & > & 0 \\ y_{m+p+1}, \ldots, y_n & \in & \mathbb{R}. \end{array}$$

Bemerkung 4.2

Den C^k-Diffeomorphismus ϕ berechnet man als Standard-Diffeomorphismus. Mit Hilfe von ϕ werden die nichtlinearen Restriktionen in lineare Restriktionen überführt. Der nichtlineare Teil von M ist in ϕ versteckt. ϕ ist eine lokale Koordinatentransformation.

Lemma 4.2

Es sei $f \in C^1(\mathbb{R}^n, \mathbb{R})$ und die LICQ in $\bar{x} \in M$ erfüllt, ϕ der Standarddiffeomorphismus. Dann existieren eindeutig bestimmte Zahlen $\bar{\lambda}_i$, $i = 1, \ldots, m$, $\bar{\mu}_j$, $j = 1, \ldots, p$, $\bar{\nu}_r$, $r = p+1, \ldots, m$ mit:

1. $Df(\bar{x}) = \sum_{i=1}^m \bar{\lambda}_i Dh_i(\bar{x}) + \sum_{j=1}^p \bar{\mu}_j Dg_j(\bar{x}) + \sum_{r=m+p+1}^n \bar{\nu}_r {\xi^r}^T$
2. $\bar{\lambda}_i = \frac{\partial}{\partial y_i}\left(f \circ \phi^{-1}(0)\right),\ i = 1, \ldots, m$
   $\bar{\mu}_{j-m} = \frac{\partial}{\partial y_j}\left(f \circ \phi^{-1}(0)\right),\ j = m+1, \ldots, p$
   $\bar{\nu}_r = \frac{\partial}{\partial y_r}\left(f \circ \phi^{-1}(0)\right),\ r = p+m+1, \ldots, n$

Bew.:

1. $D^T h_i(\bar{x}), i = 1, \dots, m$, $D^T g_j(\bar{x}), j = 1, \dots, p$, $\xi^r, r = m+p+1, \dots, n$ bilden eine Basis. Daraus folgt insbesondere, daß die Zahlen $\bar{\lambda}_i, \bar{\mu}_{j-m}, \bar{\nu}_r$ eind. bestimmt sind.
2.

$$
\begin{aligned}
\frac{\partial}{\partial y_1}\left(f \circ \phi^{-1}(0)\right) &= \frac{\partial}{\partial y_1}\left(f\left(\phi_1^{-1}(y_1, \dots, y_n), \dots, \phi_n^{-1}(y_1, \dots, y_n)\right)\right)\Big|_{y=0} \\
&= Df(\bar{x}) \cdot D\phi^{-1}(0)\, e^1 \\
&= Df(\bar{x}) \cdot \begin{pmatrix} \frac{\partial \phi_1^{-1}(0)}{\partial y_1} & \dots & \frac{\partial \phi_1^{-1}(0)}{\partial y_n} \\ \vdots & & \vdots \\ \frac{\partial \phi_n^{-1}(0)}{\partial y_1} & \dots & \frac{\partial \phi_n^{-1}(0)}{\partial y_n} \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \\
&= Df(\bar{x}) \cdot \begin{pmatrix} \frac{\partial \phi_1^{-1}(0)}{\partial y_1} \\ \vdots \\ \frac{\partial \phi_n^{-1}(0)}{\partial y_1} \end{pmatrix} \\
&= \sum_{i=1}^{n} \frac{\partial f(\bar{x})}{\partial x_i} \cdot \frac{\partial \phi^{-1}(0)}{\partial y_1}
\end{aligned}
$$

Wir setzen $D\phi^{-1}(0)\, e^1 =: \eta$. Inversentheorem $\to$ $e^1 = D\phi(\bar{x}) \cdot \eta$, d.h. $Dh_1(\bar{x})\eta = 1$, $Dh_2(\bar{x})\eta = 0, \dots, \xi^{n^T}\eta = 0$

$$
\begin{aligned}
1. \to Df(\bar{x}) \cdot \eta &= \sum_{i=1}^{m} \bar{\lambda}_i Dh_i(\bar{x}) \cdot \eta + \sum_{j=1}^{p} \bar{\mu}_j Dg_j(\bar{x}) \cdot \eta + \sum_{r=m+p+1}^{n} \bar{\nu}_r \xi^{r^T} \cdot \eta. \\
\to Df(\bar{x}) \cdot \eta &= \bar{\lambda}_1
\end{aligned}
$$

Alle anderen Beziehungen analog. $\to$ Beh.

Satz 4.3 (Karush-Kuhn-Tucker-Satz)

Es seien $f, h_i, g_j \in C^1(\mathbb{R}^n, \mathbb{R})$, $i \in I$, $j \in J$, LICQ in $\bar{x} \in M$ erfüllt, und $\bar{x}$ ein lokaler Minimalpunkt für $\min\{f(x) \,|x \in M\}$. Dann gibt es $\bar{\lambda}_i$, $\bar{\mu}_j \in \mathbb{R}$, $i \in I$, $j \in J_0(\bar{x})$, so daß

1. $Df(\bar{x}) = \sum_{i \in I} \bar{\lambda}_i Dh_i(\bar{x}) + \sum_{j \in J_0} \bar{\mu}_j Dg_j(\bar{x})$
2. $\bar{\lambda}_i$, $\bar{\mu}_j$ sind eindeutig bestimmt
3. $\bar{\mu}_j \geq 0$, $j \in J_0(\bar{x})$

Bew.: Standarddiff. $\phi \to \bar{x}$ lok. Min.pkt. für $\min\{f(x)|\, x \in M\} \Leftrightarrow 0 = \phi(\bar{x})$ lok. Min.pkt. für $\min\left\{ f \circ \phi^{-1}(y) \middle|\, y \in \{0_m\} \times \mathbb{H}^p \times \mathbb{R}^{n-m-p} \right\}$.
Nach Satz 3.3 gilt:

$$
\begin{aligned}
\frac{\partial}{\partial y_r}\left(f \circ \phi^{-1}(0)\right) &= 0,\ r = m+p+1, \dots, n \ (= \bar{\nu}_r \text{ nach L.4.2}) \\
\frac{\partial}{\partial y_j}\left(f \circ \phi^{-1}(0)\right) &\geq 0,\ j = m+1, \dots, m+p \ (= \bar{\mu}_j \text{ nach L.4.2.})
\end{aligned}
$$

$\to \bar{\nu}_r = 0, \bar{\mu}_j \geq 0.$

KKT-System

Das System 4.3 a) und c) für beliebiges $x \in M$ heißt KKT-System und wird auch in folgender Weise geschrieben:

$$\begin{aligned}
&Df(x) = \textstyle\sum_{i \in I} \lambda_i Dh_i(x) + \sum_{j \in J} \mu_j Dg_j(x) \\
&h_i(x) = 0,\, g_j(x) \geq 0,\, \mu_j \geq 0,\, i \in I,\, j \in J \\
&\mu_j g_j(x) = 0,\, j \in J \text{ (Komplementaritätsbedingung)}
\end{aligned}$$

Definition 4.2

1. $\left(\bar{x}, \bar{\lambda}, \bar{\mu}\right)$ bzw. $\left(\bar{x}, \bar{\lambda}, \bar{\mu}^{J_0}\right)$, $J_0 = J_0(\bar{x})$, $\mu^{J_0} = \begin{pmatrix} \bar{\mu}_1 & \dots & \bar{\mu}_p \end{pmatrix}$ heißt KKT-Punkt.
2. $\bar{x} \in M$ heißt stationärer Punkt, wenn $\left(\bar{\lambda}, \bar{\mu}\right) \in \mathbb{R}^m \times \mathbb{R}^s$ bzw. $\left(\bar{\lambda}, \bar{\mu}^{J_0}\right) \in \mathbb{R}^m \times \mathbb{R}^q$ existiert, so daß das KKT-System (s.o.) und 4.3 a) und c) erfüllt sind.
3. $\bar{x}$ heißt kritischer Punkt, wenn $\bar{\lambda}, \bar{\mu} \in \mathbb{R}^m \times \mathbb{R}^s$ das KKT-System ohne $\mu_j \geq 0$ erfüllen.
4. μ_j heißen Lagrangemultiplikatoren.
5. $$\begin{aligned}
L(x, \lambda, \mu) &= f(x) - \sum \lambda_i h_i(x) - \sum \mu_j g_j(x) \\
\text{bzw. } L^{J_0}\left(x, \lambda, \mu^{J_0}\right) &= f(x) - \sum \lambda_i h_i(x) - \sum_{j \in J_0} \mu_j g_j(x)
\end{aligned}$$

   heißt Lagangefunktion.

Bemerkung 4.3

Für das freie Minimumproblem $\min\{f(x) \,|\, x \in \mathbb{R}^n\}$ ist kritischer Punkt gleich stationärer Punkt.

Bemerkung 4.4

Wenn LICQ nicht erfüllt, dann muß ein lokaler Minimalpunkt nicht notwendig stationärer Punkt sein.
Die KKT-Bedingungen sind auch notwendige Bedingungen unter der schwächeren Vorraussetzung der MFCQ.

Satz 4.4

Es seien $f, h_i, g_j \in C^1(\mathbb{R}^n, \mathbb{R})$, $i \in I$, $j \in J$ und in $\bar{x} \in M$ sei die LICQ erfüllt. Weiterhin seien die folgenden Bedingungen erfüllt:

1. $|I| + |J_0(\bar{x})| = n$
2. $Df(\bar{x}) = \sum_{i \in I} \bar{\lambda}_i Dh_i(\bar{x}) + \sum_{j \in J_0} \bar{\mu}_j Dg_j(\bar{x})$
3. $\bar{\mu}_j > 0$, $j \in J_0(\bar{x})$

Dann ist $\bar{x}$ ein strikter lokaler Minimalpunkt für $\min\{f(x) \mid x \in M\}$.

Lemma 4.3

Es sei $f \in C^2(\mathbb{R}^n, \mathbb{R})$, U, V Umgebungen von $\bar{x}, \bar{y} \in \mathbb{R}^n$ und $\phi : U \to V$ ein C^2-Diffeomorphismus mit $\phi(\bar{x}) = \bar{y}$. Wenn $Df(\bar{x}) = 0$, so gilt:

$$D^2\left(f \circ \phi^{-1}(\bar{y})\right) = \left(D\phi^{-1}(\bar{y})\right)^T D^2 f(\bar{x})\, D\phi^{-1}(\bar{y})$$

Bew.: Setze $g(y) = f \circ \phi^{-1}(y)$, $\phi^{-1} : \mathbb{R}^n \to \mathbb{R}^n$.
$D\phi^{-1} : \mathbb{R}^n \to L(\mathbb{R}^n, \mathbb{R}^n)$ lin. Raum der lin. beschränkten Operatoren.
$D\left(D\phi^{-1}\right) : \mathbb{R}^n \to L(\mathbb{R}^n, L(\mathbb{R}^n, \mathbb{R}^n))$.

$$\begin{aligned}
Dg(y) &= Df\left(\phi^{-1}(y)\right) \cdot D^T\phi^{-1}(y) \\
D^2g(y) &= D\left(D^Tg(y)\right) \\
&= D\left(D\phi^{-1}(y) \cdot D^Tf\left(\phi^{-1}(y)\right)\right) \\
&= D\left(D\phi^{-1}(y)\right) \cdot D^Tf\left(\phi^{-1}(y)\right) + D\phi^{-1}(y) \cdot D^2f\left(\phi^{-1}(y)\right) \cdot D\phi^{-1}(y)
\end{aligned}$$

Lemma 4.4

Es seien $f, h_i, g_j \in C^2(\mathbb{R}^n, \mathbb{R})$, $i \in I$, $j \in J$, LICQ in $\bar{x} \in M$ erfüllt, $\bar{x}$ kritischer Punkt, ϕ Standard-Diffeomorphismus. Dann gilt:

1. $\frac{\partial f \circ \phi^{-1}(0)}{\partial y_j} = 0$, $j = m+p+1, \ldots, n$
2. $\left(\frac{\partial^2 f \circ \phi^{-1}(0)}{\partial y_i \partial y_j}\right)_{i,j=m+p+1,\ldots,n} = \left(e^{iT}\left(D\phi^{-1}(0)\right)^T D^2L^{J_0}\left(\bar{x}, \bar{\lambda}, \bar{\mu}^{J_0}\right) D\phi^{-1}(0)\, e^j\right)$

Bew.: Es gilt

1. $DL^{J_0}\left(\bar{x}, \bar{\lambda}, \bar{\mu}\right) = 0$
2. $f(\bar{x}) = L^{J_0}\left(\bar{x}, \bar{\lambda}, \bar{\mu}^{J_0}\right)$, da $h_i(\bar{x}) = 0, i \in I$, $g_j(\bar{x}) = 0, j \in J_0(\bar{x})$

$$\begin{aligned}
\to\ f \circ \phi^{-1}(0, \ldots, 0, y_{m+p+1}, \ldots, y_n) &= L^{J_0} \circ \phi^{-1}(0, \ldots, 0, y_{m+p+1}, \ldots, y_n) \\
\frac{\partial}{\partial y_j}\left(f \circ \phi^{-1}(y)\right) &= \frac{\partial}{\partial y_j} L^{J_0} \circ \phi^{-1}(0) \\
&\overset{\bar{x}=\phi^{-1}(0)}{=} \frac{\partial}{\partial x_j} L^{J_0}\left(\bar{x}, \bar{\lambda}, \bar{\mu}^{J_0}\right) \Rightarrow 1.
\end{aligned}$$

Lemma 4.3 $\Rightarrow$ 2.

Definition 4.3

Es sei $T \subseteq \mathbb{R}^n$ ein linearer Unterraum des $\mathbb{R}^n$ und A eine symmetrische $(n \times n)$-Matrix. A heißt positiv definit auf T, wenn $x^TAx > 0$, $\forall x \in T$, $x \neq 0$. A heißt positiv semidefinit auf T, wenn $x^TAx \geq 0$, $\forall x \in T$. D.h. V^TAV ist positiv definit (semidefinit), wenn V eine Matrix ist, die aus Basisvariablen von T besteht.
Bezeichnung: $A|_T$, A eingeschränkt auf T.

Definition 4.4

Es seien $h_i, g_j \in C^1(\mathbb{R}^n, \mathbb{R})$, $i \in I$, $j \in J$, $\bar{x} \in M$.

$$T_{\bar{x}}M := \left\{\xi \in \mathbb{R}^n \,|h_i(\bar{x})\xi = 0,\, i \in I,\, Dg_j(\bar{x})\xi = 0,\, j \in J_0(\bar{x})\right\}$$

heißt Tangentialraum an M im Punkt $\bar{x} \in M$.

Satz 4.5

Es seien $f, h_i, g_j \in C^1(\mathbb{R}^n, \mathbb{R})$, $i \in I$, $j \in J$, $\bar{x}$ lokaler Minimalpunkt für $\min\{f(x)\,|x \in M\}$, LICQ in $\bar{x}$ erfüllt, $\left(\bar{\lambda}, \bar{\mu}^{J_0}\right)$ der zugehörige Lagrangemultiplikatorvektor. Dann gilt:

1. $D_xL^{J_0}\left(\bar{x}, \bar{\lambda}, \bar{\mu}^{J_0}\right) = 0$
2. $D_x^2L^{J_0}\left(\bar{x}, \bar{\lambda}, \bar{\mu}^{J_0}\right)|_{T_{\bar{x}}M}$ positiv semidefinit

Bew.:

1. $(P)\ \min\{f(x)\,|x\in M[h,g]\}$
   $(P^*)\ \min\left\{f\circ\phi^{-1}(y)\,|y\in\{0_m\}\times\mathbb{H}^p\times\mathbb{R}^{n-m-p}\right\}$
   $\bar{x}$ ist lok. Min.pkt. für (P) gdw. $0=\phi(\bar{x})$ lok. Min.pkt. für (P^*) ist. S.3.6 $\Rightarrow$

   $$\left(\frac{\partial^2\left(f\circ\phi^{-1}(0)\right)}{\partial y_i\partial y_j}\right)_{i,j=m+p+1,\dots,n}=\left(e^iD\phi^{-1}(0)^T D^2L^{J_0}\left(\bar{x},\bar{\lambda},\bar{\mu}^{J_0}\right)D\phi^{-1}(0)e^j\right)_{i,j=m+p+1,\dots,n}$$

   pos. semidef.
   $\rightarrow V^TD^2L^{J_0}\left(\bar{x},\bar{\lambda},\bar{\mu}^{J_0}\right)V$ pos. semidef. mit $\left(D^T\phi^{-1}(0)e^{m+p+1},\dots,D^T\phi^{-1}(0)e^n\right)$.
2. $D\phi^{-1}(0)e^i$, $i=m+p+1,\dots,n$ Basis von $T_{\bar{x}}M$. Es gilt

   $$D\phi(\bar{x})=\begin{bmatrix} D^Th_1(\bar{x}) & \cdots & D^Th_m(\bar{x}) & D^Tg_1(\bar{x}) & \cdots & D^Tg_p(\bar{x}) & \xi^{m+p+1} & \cdots & \xi^n\end{bmatrix}^T \text{ reg.}$$

   $$\begin{aligned}\dim T_{\bar{x}}M &= n-m-p\\ D\phi^{-1}(0)e^i &\in T_{\bar{x}}M,\ i=m+p+1,\dots,n\text{ lin. unabh.}\rightarrow\text{ Basis von }T_{\bar{x}}M\end{aligned}$$

Satz 4.6

Es seien $f,h_i,g_j\in C^2(\mathbb{R}^n,\mathbb{R})$, $i\in I$, $j\in J$, $\bar{x}$ kritischer Punkt für $\min\{f(x)\,|x\in M\}$, LICQ in $\bar{x}$ erfüllt, $\left(\bar{\lambda},\bar{\mu}^{J_0}\right)$ der zugehörige Lagrangemultiplikatorvektor mit folgenden Bedingungen:

1. $\bar{\mu}_j>0$, $j\in J_0(\bar{x})$ (strikte Komplementaritätsbedingung)
2. $\mathrm{D}^2L^{J_0}\left(\bar{x},\bar{\lambda},\bar{\mu}^{J_0}\right)|_{T_{\bar{x}}M}$ positiv definit

Dann ist $\bar{x}$ ein strikter lokaler Minimalpunkt.

4.2 Optimalitätsbedingungen für konvexe Optimierungsprobleme

$(P)\ \min\{f(x)\,|x\in\mathbb{R}^n\}$

Satz 4.7

Es sei $f\in C^1(\mathbb{R}^n,\mathbb{R})$ und f konvex auf $\mathbb{R}^n$. Dann sind die folgenden Aussagen äquivalent:

1. $\bar{x}$ ist globaler Minimalpunkt
2. $\bar{x}$ ist lokaler Minimalpunkt
3. $\bar{x}$ ist stationärer Punkt

Bew.: $1\Leftrightarrow 2$: Satz 2.6
$2\Rightarrow 3$: Satz 3.1
g.z.z.: $3\Rightarrow 2$.
Anwendung von Satz 2.3:
$f(x)$ ist konvex auf $\mathbb{R}^n$ $\Leftrightarrow \forall x,y\in\mathbb{R}^n\ f(y)-f(x)\geq Df(x)(y-x)$.
Sei $\bar{x}$ stat. Pkt., d.h. $Df(\bar{x})=0\rightarrow f(y)-f(\bar{x})\geq 0\,\forall y\in\mathbb{R}^n\Rightarrow$ glob. Min.pkt.

Bemerkung 4.5

$Df(x)=0$ ist notwendige und hinreichende Bedingung für konvexes $f(x)$.

$(P)\ \min\{f(x)\,|x\in M\}$, $M:=\{x\in\mathbb{R}^n\,|h_i(x)=0,\ g_j(x)\geq 0,\ i\in I,\ j\in J\}$

Satz 4.8

Es sei $f,\, h_i,\, g_j \in C^1(\mathbb{R}^n, \mathbb{R})$, $i \in I$, $j \in J$ und $f, -g_j$ konvex. $h_i(x) := a^{iT} x + b_i$, $i \in I$. Dann gilt: Wenn $\bar{x} \in M$ ein stationärer Punkt ist, so ist $\bar{x}$ globaler Minimalpunkt.

Bew.: Es seien $\bar{x}$ stat. Pkt. für (P). Dann ex. Zahlen $\lambda_i, i \in I$, $\mu_j, j \in J_0(\bar{x})$ mit

$$Df(\bar{x}) = \sum_{i\in I} \lambda_i Dh_i(\bar{x}) + \sum_{j\in J_0(\bar{x})} \mu_j Dg_j(\bar{x}),\ \mu_j \geq 0, j \in J_0(\bar{x}).$$

Es sei $\bar{x} \in M$ bel. Nach Satz 2.3. gilt $f(x) - f(\bar{x}) \geq Df(\bar{x})(x - \bar{x})$

$$\Rightarrow f(x) - f(\bar{x}) \geq \underbrace{\sum_{i\in I} \lambda_i Dh_i(\bar{x})(x-\bar{x})}_{=0\,(\text{Beh. }1)} + \underbrace{\sum_{j\in J_0(\bar{x})} Dg_j(\bar{x})(x-\bar{x})}_{\geq 0\,(\text{Beh. }2)}$$

Beh. 1:

$$\begin{aligned} i \in I, h_i(x) = h_i(\bar{x}) &= 0 \\ \Rightarrow a^{i^T} x + b_i &= a^{i^T} \bar{x} + b_i \\ \Rightarrow a^{i^T}(x - \bar{x}) &= 0 \\ \Rightarrow Dh_i(\bar{x})(x - \bar{x}) &= 0 \end{aligned}$$

Beh. 2: Aus Satz 2.3. folgt ($-g_j(x)$ ist konvex)

$$\begin{aligned} -\underbrace{g_j(x)}_{\geq 0} - \underbrace{(-g_j(\bar{x}))}_{=0} &\geq -Dg_j(\bar{x})(x - \bar{x}) \\ \Rightarrow Dg_j(\bar{x})(x - \bar{x}) &\geq 0 \end{aligned}$$

Bew. Satz 2.4: Es sei $f \in C^2(\mathbb{R}^n, \mathbb{R})$. f ist konvex auf $\mathbb{R}^n$, gdw. $D^2 f(x)$ pos. semidef. $\forall x \in \mathbb{R}^n$.

($\Rightarrow$) Es seien $x \in \mathbb{R}^n$ bel. und $F(x) := f(x) - Df(\bar{x})(x - \bar{x})$. Dann gilt:

- $F(\bar{x}) = f(\bar{x})$
- $F(\bar{x})$ konvex auf $\mathbb{R}^n$ (als Summe einer konvexen und einer affin-lin. Fkt.)
- $DF(x) = Df(x) - Df(\bar{x})$

$\Rightarrow DF(\bar{x}) = 0 \overset{S.4.7}{\Rightarrow} \bar{x}$ glob. Min.pkt. von F.

notw. Bed. 2. Ordnung (S.3.6) $\Rightarrow D^2F(\bar{x})$ pos. semidef. $\overset{D^2F(x)=D^2f(x)}{\Rightarrow} D^2 f(\bar{x})$ ist pos. semidef.

Bemerkung 4.6

1. Die KKT-Bedingungen (Optimalitätsbedingungen 1. Ordnung) sind notwendige und hinreichende Bedingung für einen globalen Minimalpunkt $\bar{x}$, wenn die LICQ bzw. die schwachen MFCQ in $\bar{x}$ erfüllt sind.
2. Anstelle von LICQ und MFCQ wird in der Literatur häufig eine andere Regularitätsbedingung benutzt, die sogenannte Slater-Bedingung:

$$\exists \hat{x} \in M :\ g_j(\hat{x}) > 0, j \in J,\ h_i(x), i \in I \text{ affinlinear}$$

3. Die Slater-Bedingung ist für konvexe Restriktionsmengen äquivalent zur MFCQ.

5 Trennungssätze für konvexe Mengen und der Satz von Farkas ([1])

5.1 Trennungssätze für konvexe Mengen

Satz 5.1

Es sei $C \subseteq \mathbb{R}^n$ nicht leer, konvex und abgeschlossen und $0 \notin C$. Dann existiert ein $a \in \mathbb{R}^n$, $a \neq 0, \gamma \in \mathbb{R}$ mit $\gamma > 0$ mit $a^T x \geq \gamma, \forall x \in C$, d.h. es existiert eine Hyperebene $\left\{x \in \mathbb{R}^n \left| a^T x = \gamma \right.\right\}$ und $C \subseteq \left\{x \in \mathbb{R}^n \left| a^T x \geq \gamma \right.\right\}$ und $0 \in \left\{x \in \mathbb{R}^n \left| a^T x < \gamma \right.\right\}$.

Bew.idee: $B(0,r) = \{x \in \mathbb{R}^n \, | \|x - 0\| \leq r\}$.

Sei $r > 0$ so gewählt, daß relint$(C \cap B(0,r)) \neq \emptyset$, d.h. ex. ein rel. innerer Pkt. $\bar{x}$.

$C \cap B(0,r)$ ist kompakt.

(P) min $\{\|x\| | \, x \in C \cap B(0,r)\}$ konv. Opt.problem.

Nach dem Satz von Weierstraß ex. ein glob. Min.pkt. x^0, d.h.

$$\left\|x^0\right\| \leq \|x\| \; \forall x \in C \cap B(0,r), x^0 \in C.$$

Da $0 \notin C$ und C abg., gilt $0 \notin C \cap B(0,r) \;\Rightarrow\; \left\|x^0\right\| > 0$.

Da relint$(C \cap B(0,r)) \neq \emptyset$, gilt $r > \left\|x^0\right\| \;\Rightarrow\; x^0$ liegt auf Rand von $C. \, a := x^0$

Beh.: $a^T x > 0 \, \forall x \in C$.

Da C konvex, gilt $x^0 + \lambda\left(x - x^0\right) \in C \, \forall x \in C, \lambda \in [0,1]$.

$$\Rightarrow \left\|x^0\right\| \quad \leq \quad \left\|x^0 + \lambda\left(x - x^0\right)\right\| \; \forall x \in C, \lambda \in [0,1].$$

Es gilt

$$\left(x^0 + \lambda\left(x - x^0\right)\right)^T \left(x^0 + \lambda\left(x - x^0\right)\right) = \left\|x^0 + \lambda\left(x - x^0\right)\right\|^2 \geq \left\|x^0\right\| = {x^0}^T x^0 \; \forall x \in C, \lambda \in [0,1].$$

Bin. Formel

$$\begin{aligned}
\Rightarrow {x^0}^T x^0 + \left[2\lambda\left(x - x^0\right) x^0 + \lambda^2 \left(x - x^0\right)^T \left(x - x^0\right)\right] - {x^0}^T x^0 \quad &\geq \quad 0 \, \forall x \in C, \lambda \in [0,1]. \\
\lim_{\lambda \to 0^+} \left[2\left(x - x^0\right)^T x^0 + \lambda\left(x - x^0\right)^T \left(x - x^0\right)\right] \quad &\geq \quad 0 \\
\to 2x^T x^0 - 2{x^0}^T x^0 \quad &\geq \quad 0 \, \forall x \in C \\
\Rightarrow x^T x^0 \quad &\geq \quad \left\|x^0\right\| =: \gamma > 0
\end{aligned}$$

Corollar 5.1

Sei $C \in \mathbb{R}^n$ nicht leer, konvex und abgeschlossen und $x^0 \notin C$. Dann existiert ein $a \in \mathbb{R}^n$, $a \neq 0$, $\gamma \in R$ mit $a^T x \geq \gamma > a^T x^0$, $\forall x \in C$.

Satz 5.2 (Allgemeiner Trennungssatz im $\mathbb{R}^n$, siehe [1])

Es seien $C_1, C_2 \subseteq \mathbb{R}^n$ nicht leer, konvex mit relint$C_1 \cap$ relint$C_2 = \emptyset$. Dann existiert ein $a \neq 0$, $a \in \mathbb{R}^n$, $\gamma \in \mathbb{R}$ und $a^T x \geq \gamma \geq a^T y$, $\forall x \in c\ell C_1$, $\forall y \in c\ell C_2$ und es gilt:

$$\left\{x \in \mathbb{R}^n \left| a^T x = \gamma \right.\right\} \not\supseteq C_1 \cup C_2$$

5.2 Der Satz von Farkas

Satz 5.3

Sei $A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$ und $A^T = \left(a^1, \dots, a^m\right)$, $a^i \in \mathbb{R}^n$. Dann gilt:

$$\{x \in \mathbb{R}^n \,|Ax \leq 0\} \subseteq \left\{x \in \mathbb{R}^n \,\middle|c^T x \leq 0\right\} \quad \Leftrightarrow \quad \exists u \in \mathbb{R}^m : u \geq 0,\, c = A^T u \; (*)$$

Bew.: $(\Leftarrow)$: Sei x mit $Ax < 0$ und $u \geq 0 \;\rightarrow\; 0 \geq u^T Ax \overset{(*)}{=} c^T x$.

$(\Rightarrow)$: Es sei $K := \left\{x \in \mathbb{R}^n \,\middle|x = A^T u, u \geq 0\right\}$ polyedrischer Kegel.
$(*) \Leftrightarrow c \in K$.
Ann. 1: $c \notin K$, $K \neq \emptyset$, konvex, abg.
Cor. 5.1 $\Rightarrow \exists a \in \mathbb{R}^n, a \neq 0, \gamma \in \mathbb{R}$ mit

$$a^T x \geq \gamma \,\forall x \in K \;\Rightarrow\; a^T c < \inf_{x \in K} a^T x =: \alpha \; (**)$$

Da $0 \in K \;\Rightarrow\; \alpha \leq 0$. Wir zeigen: $\alpha = 0$.
Ann. 2: $\alpha < 0 \;\Rightarrow\; \exists \tilde{x} \in K :\, a^T \tilde{x} < 0,\, \lambda \tilde{x} \in K \,\forall \lambda \geq 0$
$\Rightarrow\; \lambda a^T \tilde{x} < 0 \,\forall \lambda > 0 \;\Rightarrow\; \inf_{x \in K} a^T x = -\infty$. Wid. zu $(**) \;\Rightarrow\; \alpha = 0$
$\Rightarrow\; a^T c < 0 \leq a^T x \,\forall x \in K$. Da $a^i \in K, i = 1, \dots, m \;\Rightarrow\; a^T a^i \geq 0 \,\forall a^i$
$\Rightarrow\; Aa \geq 0$, $c^T a < 0$
$\Rightarrow\; A(-a) \leq 0$, $c^T(-a) > 0$ Wid. zu $\{x \in \mathbb{R}^n \,|Ax \leq 0\} \subseteq \left\{x \in \mathbb{R}^n \,\middle|c^T x \leq 0\right\}$.

6 Dualitätstheorie der nichtlin. Optimierung ([6])

6.1 Das Subdifferential

Definition 6.1

Es sei $f : \mathbb{R}^n \rightarrow \mathbb{R}$

1. Ein Vektor $a \in \mathbb{R}^n$ heißt Subgradient von f in $\bar{x}$, wenn gilt:
$$f(x) \geq f(\bar{x}) + a^T (x - \bar{x}),\; \forall x \in \mathbb{R}^n$$
2. Das Subdifferential $\partial f(\bar{x})$ ist die Menge aller Subgradienten von f in $\bar{x}$.
3. Wenn $\partial f(\bar{x}) \neq \emptyset$, so heißt f subdifferenzierbar.

Bemerkung 6.1

Das Subdifferential ist auch für nichtkonvexe Funktionen definiert.

Geometrische Interpretation

$\operatorname{epi}(f) = \left\{ \begin{pmatrix} x \\ \mu \end{pmatrix} \in \mathbb{R}^{n+1} \,\middle|\, f(x) \leq \mu \right\}$

Hyperebene im $\mathbb{R}^{n+1}$: $H(a, \beta, \alpha) := \left\{ \begin{pmatrix} x \\ \mu \end{pmatrix} \in \mathbb{R}^{n+1} \,\middle|\, a^T x + \beta \mu = \alpha \right\}$, $\begin{pmatrix} a \\ \beta \end{pmatrix} \neq 0$

sei $\beta = -1$: $H(a, \alpha) := \left\{ \begin{pmatrix} x \\ \mu \end{pmatrix} \in \mathbb{R}^{n+1} \,\middle|\, a^T x - \mu = \alpha \right\}$

$c\ell H^+(a, \alpha) := \left\{ \begin{pmatrix} x \\ \mu \end{pmatrix} \in \mathbb{R}^{n+1} \,\middle|\, a^T x - \mu \geq \alpha \right\}$

$c\ell H^-(a, \alpha) := \left\{ \begin{pmatrix} x \\ \mu \end{pmatrix} \in \mathbb{R}^{n+1} \,\middle|\, a^T x - \mu \leq \alpha \right\}$

Definition 6.2

$H(a,\alpha)$ heißt nichtvertikale Stützhyperebene an $\operatorname{epi}(f)$ durch $(\bar{x}, f(\bar{x}))$, wenn gilt:

1. $H(a,\alpha) = \left\{ \begin{pmatrix} x \\ \mu \end{pmatrix} \in \mathbb{R}^{n+1} \,\middle|\, a^T x - \mu = a^T \bar{x} - f(\bar{x}) \right\}$
2. $\operatorname{epi}(f) \subseteq c\ell H^-(a,\alpha)$

Satz 6.1

$\partial f(\bar{x})$ ist die Menge aller Vektoren a, für die $H(a,\alpha) = \left\{ \begin{pmatrix} x \\ \mu \end{pmatrix} \in \mathbb{R}^{n+1} \,\middle|\, a^T x - \mu = \alpha \right\}$ eine nichtvertikale Stützhyperebene an $\operatorname{epi}(f)$ im Punkt $\left(\bar{x}^T, f(\bar{x})\right)$ ist.

Bew.: $(\Leftarrow)$: Sei $H(a,\alpha)$ ein nichtvert. Stützhyperebene an epi(f) durch $(\bar{x}, f(\bar{x}))$.

$$\begin{aligned}
\to a^T x - \mu &\leq a^T \bar{x} - f(\bar{x})\ \forall \begin{pmatrix} x \\ \mu \end{pmatrix} \text{ mit } \mu \geq f(x)\ \forall x \in \mathbb{R}^n \\
\to a^T x - f(x) &\leq a^T \bar{x} - f(\bar{x})\ \forall x \in \mathbb{R}^n \\
\to f(x) - f(\bar{x}) &\geq a^T (x - \bar{x})\ \forall x \in \mathbb{R}^n \\
\to a &\in \partial f(\bar{x})
\end{aligned}$$

Corollar 6.1

Es sei $f \in C^1(\mathbb{R}^n, \mathbb{R})$ und konvex. Dann gilt: $\partial f(\bar{x}) = \{Df(\bar{x})\}$, $\forall \bar{x} \in \mathbb{R}^n$.

Bew.: Aus S. 6.1 folgt: $Df(\bar{x}) \in \partial f(x)$. Wir benutzen die Eind. des Tangentialraumes $\left\{ \begin{pmatrix} x \\ \mu \end{pmatrix} \in \mathbb{R}^{n+1} \,\middle|\, Df(\bar{x})\, x - \mu = 0 \right\}$ im Pkt. $(\bar{x}, f(\bar{x}))$.

Satz 6.2

Es sei $f : \mathbb{R}^n \to \mathbb{R}$ subdifferenzierbar in $\bar{x}$. Dann ist $\partial f(\bar{x})$ eine kompakte und konvexe Menge.

Satz 6.3

Es sei $f : \mathbb{R}^n \to \mathbb{R}$ konvex. Dann ist f subdifferenzierbar $\forall x \in \mathbb{R}^n$.

Bew.: Wir konstruieren einen Subgradienten mit Hilfe des Trennungssatzes, d.h. eine nichtvertikale Stützhyperebene an den Epigraphen epi$(f) = \left\{ \begin{pmatrix} x \\ \mu \end{pmatrix} \in \mathbb{R}^{n+1} \,\middle|\, f(x) \leq \mu \right\}$.

Es gilt: $(\bar{x}, f(\bar{x})) \in \operatorname{epi}(f) \setminus \operatorname{intepi}(f)$.

Nach dem Trennungssatz gibt es einen Vektor $\begin{pmatrix} \bar{a} \\ \bar{a}_{n+1} \end{pmatrix} \in \mathbb{R}^{n+1} \setminus \{0\}$ mit

$$\begin{aligned}
\begin{pmatrix} \bar{a} \\ \bar{a}_{n+1} \end{pmatrix}^T \left(\begin{pmatrix} x \\ \mu \end{pmatrix} - \begin{pmatrix} \bar{x} \\ f(\bar{x}) \end{pmatrix} \right) &\leq 0\, \forall \begin{pmatrix} x \\ \mu \end{pmatrix} \in \operatorname{epi}(f), \\
\text{d.h. } \bar{a}^T (x - \bar{x}) + \bar{a}_{n+1} (\mu - f(\bar{x})) &\leq 0\, \forall x \in \mathbb{R}^n, \mu \geq 0. \ (*)
\end{aligned}$$

Da μ bel. groß werden kann, muß $\bar{a}_{n+1} \leq 0$ sein.
Beh.: $\bar{a}_{n+1} < 0$ (nichtvertikale Stützhyperebene).
Ann.: $\bar{a}_{n+1} = 0 \Rightarrow \bar{a}^T (x - \bar{x}) \leq 0\, \forall x \in \mathbb{R}^n$.

Da $\bar{x} \in \mathbb{R}^n$ ist

$$\begin{array}{rcl} \bar{x} + \lambda\bar{a} & \in & \mathbb{R}^n \; \forall \lambda \geq 0 \\ \Rightarrow a^T(\bar{x} + \lambda\bar{a} - \bar{x}) = \lambda\bar{a}^T\bar{a} & \leq & 0 \\ \Rightarrow \bar{a} & = & 0 \\ \Rightarrow \begin{pmatrix} \bar{a} \\ \bar{a}_{n+1} \end{pmatrix} & = & 0 \, \text{Wid.} \\ \Rightarrow \bar{a}_{n+1} & < & 0 \rightarrow \bar{a}_{n+1} = -|\bar{a}_{n+1}|\,. \\ (*) \; \Rightarrow \; \bar{a}^T(x - \bar{x}) - |\bar{a}_{n+1}|(\mu - f(\bar{x})) & \leq & 0 \, \forall x \in \mathbb{R}^n, \mu \geq f(x)\,. \end{array}$$

Für $x \in \mathbb{R}^n$ bel. und $\mu = f(x)$ folgt

$$\begin{array}{rcl} \frac{1}{|\bar{a}_{n+1}|} \cdot a^T(x - \bar{x}) + f(\bar{x}) & \leq & f(x) \; \forall x \in \mathbb{R}^n \\ \Rightarrow \frac{a}{|\bar{a}_{n+1}|} & \in & \partial f(\bar{x})\,. \end{array}$$

Satz 6.4

Es sei $f : \mathbb{R}^n \rightarrow \mathbb{R}$ konvex. $\bar{x} \in \mathbb{R}^n$ ist optimal für $\{f(x) | \, x \in \mathbb{R}^n\}$ gdw. $0 \in \partial f(\bar{x})$.

Bew.: $(\Rightarrow)$: $\bar{x}$ ist opt., d.h. $f(\bar{x}) \leq f(x) \; \forall x \in \mathbb{R}^n$. $(*)$
S. 6.3., f konvex $\Rightarrow$ f ist subdiff.bar in $\bar{x}$, d.h. $f(\bar{x}) + a^T(x - \bar{x}) \leq f(x) \; \forall x \in \mathbb{R}^n$.
$\overset{(*)}{\Rightarrow} a = 0 \in \partial f(\bar{x})$.
$(\Leftarrow)$: $a^T\left(x - x^0\right) \leq f(x) - f(\bar{x})$; $a = 0 \Rightarrow$ Beh.

6.2 Dualitätssätze für allg. nichtlin. Optimierungsprobleme

Es seien f, g_j, $j \in J$ Funktionen definiert auf $\mathbb{R}^n$, $J := \{1, \ldots, s\}$, $I = \emptyset$

Primalproblem:

(P) $\inf\{f(x) | \, x \in M\}$, $M := \{x \in \mathbb{R}^n | \, g_j(x) \geq 0, \, j \in J\}$, $L(x, \mu) := f(x) - \sum_{j \in J} \mu_j g_j(x)$

Dualproblem:

(D) $\sup\left\{h(\mu) | \, \mu \in \mathbb{R}^s_+\right\}$, $\mathbb{R}^s_+ := \{\mu \in \mathbb{R}^s | \, \mu \geq 0\}$, $h(\mu) := \inf_{x \in \mathbb{R}^n} L(x, \mu)$ oder
(D) $\sup_{\mu \in \mathbb{R}^s_+} \inf_{x \in \mathbb{R}^n} L(x, \mu)$

Lemma 6.1

Wenn $M \neq \emptyset$, so gilt $f(x) \geq h(\mu)$, $\forall x \in M$, $\forall \mu \in \mathbb{R}^s_+$. Insbesondere gilt:

$$\varphi_P := \inf_{x \in M} f(x) \geq \varphi_D := \sup_{\mu \in \mathbb{R}^s_+} h(\mu)$$

Bew.: Es gilt

$$\psi(x) := \sup_{\mu \in \mathbb{R}^s_+} L(x, \mu) = \begin{cases} f(x) & , x \in M \\ \infty & , x \notin M \end{cases}$$

$\left(\bar{P}\right)$ $\inf\{\psi(x) | \, x \in \mathbb{R}^n\}$.
Da $M \neq \emptyset$, sind (P) und $\left(\bar{P}\right)$ äqu. im folgenden Sinne: die Lösungsmengen sind gleich und die Infima sind gleich. Es gilt für bel. $\bar{x} \in \mathbb{R}^n, \bar{\mu} \in \mathbb{R}^s_+$

$$\psi(\bar{x}) = \sup_{\mu \in \mathbb{R}^s_+} L(\bar{x}, \mu) \geq L(\bar{x}, \bar{\mu}) \geq \inf_{x \in \mathbb{R}^n} L(x, \bar{\mu}) = h(\bar{\mu})$$

$$\to \inf_{x\in M} f(x) = \inf_{x\in\mathbb{R}^n} \psi(x) = \inf_{x\in\mathbb{R}^n} \sup_{\mu\in\mathbb{R}^s_+} L(x,\mu) \geq \sup_{\mu\in\mathbb{R}^s_+} \inf_{x\in\mathbb{R}^n} L(x,\mu) = \sup_{\mu\in\mathbb{R}^s_+} h(\mu)$$

Bemerkung 6.2

(P) und (D) können auch in symmetrischer Form geschrieben werden:

$$(P) \quad \inf_{x\in\mathbb{R}^n} \sup_{\mu\in\mathbb{R}^s_+} L(x,\mu) \qquad (D) \quad \sup_{\mu\in\mathbb{R}^s_+} \inf_{x\in\mathbb{R}^n} L(x,\mu)$$

$M(y) := \{x \in \mathbb{R}^n \,|\, g_j(x) \geq y_j, j \in J\} \to$ param. Opt.problem

$P(y) : \inf\{f(x)|\, x \in M(y)\}$, $\varphi(y) = \inf\{f(x)|\, x \in M(y)\}$ Extremalwertfkt.

$\varphi : \mathbb{R}^s \to \bar{\mathbb{R}} = \mathbb{R} \cup \{+\infty, -\infty\}$, $\varphi(y) = +\infty$, falls $M(y) = 0$ (Vereinbarung)

Es gilt $P(0) = (P)$, $\varphi(0) = \varphi_P$, $M(0) = M$.

Satz 6.5 (starker Dualitätssatz)

Sei φ_P endlich ((P) nicht notwendig lösbar), dann gilt: $\varphi_P = \varphi_D$ und (D) hat einen globalen Maximalpunkt μ^0 gdw. φ subdifferenzierbar in 0 (und $\mu^0 \in \partial\varphi(0)$).

Bew.: Vorbetrachtungen: $h(\mu)$ für $\mu \geq 0$

$$\begin{aligned}
(*) : h(\mu) &= \inf_{x\in\mathbb{R}^n} \left\{ f(x) - \sum_{j\in J} \mu_j g_j(x) \right\} \\
&\overset{(da\, g_j \geq y_j)}{=} \inf_{x\in M(y)} \left\{ f(x) - \sum \mu_j y_j \right\} \\
&= \inf_{y\in\mathbb{R}^s} \inf_{x\in M(y)} \left\{ f(x) - \mu^T y \right\} \\
&= \inf_{y\in\mathbb{R}^s} \left\{ \varphi(y) - \mu^T y \right\}
\end{aligned}$$

($\Rightarrow$) Es gilt $\varphi_P = \varphi_D$ und (D) hat einen glob. Max.pkt. μ^0, d.h. $h(\mu^0) = \varphi_D$ und $\mu^0 \in \mathbb{R}^s_+$

$(*) \to \varphi(y) - {\mu^0}^T y \geq \varphi(0) = \varphi_P \,\forall y \in \mathbb{R}^s$

$\to \varphi(y) - \varphi(0) \geq {\mu^0}^T y \,\forall y \in \mathbb{R}^s \to \varphi$ ist subdiff.bar. in 0 und $\mu^0 \in \partial\varphi(0)$.

($\Leftarrow$) φ sei subdiff.bar in 0, d.h. $\exists \mu^0 \in \mathbb{R}^s$ mit $\varphi(y) - \varphi(0) \geq {\mu^0}^T y \,\forall y \in \mathbb{R}^s$.

$\Rightarrow \varphi_P = \varphi(0) \leq \varphi(y) - {\mu^0}^T y \,\forall y \in \mathbb{R}^+$ $(**)$

$(*) \Rightarrow h(\mu^0) \geq \varphi_P \overset{L.4.1, \varphi_P \geq \varphi_D}{\to} h(\mu^0) = \varphi_P$ $(***)$

z.z.: $\mu^0 \in \mathbb{R}^s_+$ (d.h. zulässig)

Sei $y \leq 0 \Rightarrow g_j(y) \geq 0 \geq y_j, j \in J \Rightarrow M(0) \subseteq M(y)$

$\Rightarrow \inf_{x\in M(0)} f(x) \geq \inf_{x\in M(y)} f(x)$, d.h. $\varphi(y) - \varphi(0) \leq 0 \overset{(**)}{\Rightarrow} {\mu^0}^T y \leq 0 \,\forall y \in \mathbb{R}^s_-$.

$\Rightarrow \mu^0 \geq 0$, zusammen mit $(***)$ $\Rightarrow \mu^0$ ist glob. Max.pkt. für (D).

Bemerkung 6.3

Wir haben $f, g_j, j \in J$ als Fkt.en auf $\mathbb{R}^n$ betrachtet. Analoges gilt auch für $X \subseteq \mathbb{R}^n, X \neq \emptyset$, $f : X \to \mathbb{R}, g_j : X \to \mathbb{R}, j \in J$.

Satz 6.6 (Dualitätssatz für konvexe Probleme)

Es sei $f : \mathbb{R}^n \to \mathbb{R}$ konvex, $g_j : \mathbb{R}^n \to \mathbb{R}$ konkav, $j \in J$ und die Slater-Bedingung sei erfüllt (d.h. $\exists \bar{x} \in \mathbb{R}^n : g_j(\bar{x}) > 0, j \in J$). Weiter sei φ_P endlich. Dann ist (D) lösbar und es gilt $\varphi_P = \varphi_D$.

Bew.: Erfolgt mit Hilfe von Satz 6.5. Man kann zeigen: φ subdiff.bar in 0. (siehe [6])

Bemerkung 6.5

Man kann mit Hilfe des Dualitätssatzes der konvexen Opt. die KKT-Bedingungen herleiten.

Satz 6.7 (Sattelpunkttheorem)

Es sei (P) ein konvexes Optimierungsproblem und es sei die Slater-Bedingung erfüllt. $\bar{x} \in \mathbb{R}^n$ ist optimal für (P) $\Leftrightarrow$ $\exists \bar{\mu} \geq 0$, so daß

$$L(\bar{x}, \mu) \leq L(\bar{x}, \bar{\mu}) \leq L(x, \bar{\mu}), \forall x \in \mathbb{R}^n, \forall \mu \geq 0$$

Bew.: siehe [6]: ($\Leftarrow$) gilt allg. ohne Konvexität und Slater-Bedingung
($\Rightarrow$) Wir benötigen den Trennungssatz.

Definition 6.3

$\varphi : \mathbb{R}^m \to \mathbb{R}$ heißt unterhalbstetig (uhs) in $\bar{y}$, wenn für jede Folge $\{y^t\}$, $y^t \to \bar{y}$ gilt:

$$\liminf_{t \to \infty} \varphi\left(y^t\right) \geq \varphi(\bar{y})$$

Satz 6.8

Es sei φ_P endlich. Dann gilt:

1. Falls $\varphi_P = \varphi_D$, so ist φ unterhalbstetig in 0.
2. Ist φ unterhalbstetig in 0 und (P) konvex, so gilt $\varphi_P = \varphi_D$.

Bemerkung 6.6

Man kann mit Hilfe von Satz 6.8 die Optimalitätssätze der linearen Optimierung herleiten.
(P) $\max\left\{ c^T x \,\middle|\, Ax \leq b, x \geq 0\right\} = -\min\left\{ -c^T x \,\middle|\, Ax \leq b, x \geq 0\right\}$, $X := \{x |\, x \geq 0\}$
$P(y):\ \min\left\{ -c^T x \,\middle|\, Ax - b \leq y, x \in X\right\}$.
z.z.: $\varphi(y)$ ist uhs. in 0, sogar stetig in 0.
(D) $\min\left\{ b^T u \,\middle|\, A^T u \geq c, u \geq 0\right\}$

Satz 6.9

Wenn $M_P := \{x |\, Ax \leq b, x \geq 0\} \neq \emptyset$ und $M_D := \left\{u \,\middle|\, A^T u \geq c, u \geq 0\right\} \neq \emptyset$, so haben (P) und (D) Optimalitätspunkte $\hat{x}$ und $\hat{u}$ und es gilt:

$$c^T \hat{x} = \max_{x \in M_P} c^T x = \min_{u \in M_D} b^T u = b^T \hat{u}$$

Satz 6.10

(P) ist lösbar, gdw. (D) lösbar ist und es gilt die Aussage von Satz 6.9.

7 Morsetheorie ([2])

Satz 7.1

Es seien $k \geq 1$, $f \in C^k(\mathbb{R}^n, \mathbb{R})$ und $Df(\bar{x}) \neq 0$. Dann existieren Umgebungen U und V von $\bar{x}$ und $0 \in \mathbb{R}^n$ und ein C^k-Diffeomorphismus $F : U \to V$ und $F(x) = y$, $F(\bar{x}) = 0$ und $f \circ F^{-1} = f(\bar{x}) + y_1$.

Bew.: Nach Vor. $Df(\bar{x}) \neq 0$ (o.B.d.A. $\frac{\partial F}{\partial x_1}(\bar{x}) \neq 0$).

Def. von F wie folgt:

$$\begin{aligned} y &= F(x) \\ y_1 &= f(x) - f(\bar{x}) \\ y_j &= x_j - \bar{x}_j,\ j = 2, \ldots, n \end{aligned}$$

F hat folgende Eigenschaften:

1. $F \in C^k(\mathbb{R}^n, \mathbb{R}^n)$
2. $DF(\bar{x}) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(\bar{x}) & \frac{\partial f}{\partial x_2}(\bar{x}) & \cdots & \frac{\partial f}{\partial x_n}(\bar{x}) \\ 0 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 \end{pmatrix}$ hat Rang n (wegen $\frac{\partial f}{\partial x_1}(\bar{x}) \neq 0$)

Inversentheorem $\Rightarrow$ $\exists$ Umgebungen U, V von $\bar{x}, 0 \in \mathbb{R}^n$, so daß $F : U \to V$ ein C^k-Diff., und es gilt $F(\bar{x}) = 0$ und $f \circ F^{-1}(y_1, \ldots, y_n) = y_1 + f(\bar{x})$.

Bemerkung 7.1

$\bar{x}$ mit $Df(\bar{x}) \neq 0$ ist kein Kandidat für einen lokalen Minimalpunkt.

Bemerkung 7.2

Spezialfall: $f(0) = 0$, $\mathrm{D}f(0) \neq 0 \Rightarrow f \circ F^{-1}(y) = y_1$. Sei $f \in C^2(\mathbb{R}^n, \mathbb{R}) \Rightarrow D^2 f(x)$ ist symmetrisch $\Rightarrow$ relle Eigenwerte.

Definition 7.1

1. Ein stationärer (kritischer) Punkt $\bar{x} \in \mathbb{R}^n$ von $f(x)$ ($\min\{f(x) \,|x \in \mathbb{R}^n\}$) heißt nichtentartet, wenn $D^2 f(\bar{x})$ regulär.
2. $QI := QI\left(D^2 f(\bar{x})\right)$ ist die Anzahl der negativen Eigenwerte von $D^2 f(\bar{x})$ und heißt quadratischer Index (Morse-Index).
3. $QCI := QCI\left(D^2 f(\bar{x})\right)$ ist die Anzahl der positiven Eigenwerte und heißt quadratischer Co-Index.

Bemerkung 7.3

Sei $\bar{x}$ ein nichtentarteter stationärer Punkt von $f(x)$. Wenn $QI = 0$, so ist $\bar{x}$ lokaler Minimalpunkt. Wenn $QI = n$, so ist $\bar{x}$ ein lokaler Maximalpunkt. Wenn $0 < QI < n$, so ist $\bar{x}$ ein Sattelpunkt.

Satz 7.2 (Morse-Lemma für das freie Minimumproblem)

Sei $f \in C^r(\mathbb{R}^n, \mathbb{R})$, $r \geq 4$, $f(0) = 0$, $Df(0) = 0$ und $D^2 f(0)$ regulär (d.h. 0 ist ein nichtentarteter kritischer Punkt) mit $k = QI\left(D^2 f(0)\right)$. Dann existieren Umgebungen U, V von $0 \in \mathbb{R}^n$ und ein $F \in C^{r-2}(U, V)$ (d.h. F ist ein C^{r-2}-Diffeomorphismus) mit $F(x) = y$, $F(0) = 0$ und es gilt:

$$f \circ F^{-1}(y) = -\sum_{i=1}^{k} y_i^2 + \sum_{j=k+1}^{n} y_j^2$$

Bew.:

Teil I: Umformung von $f(x)$ in eine quadr. Form $x^T A(x) x$.

Schritt I.1: Sei $x \in \mathbb{R}^n$ bel.

$$
\begin{aligned}
\Rightarrow\; f(x) = f(x) - f(0) = \int_0^1 \frac{d}{dt} f(tx)\,dt &= \int_0^1 \sum_{i=1}^n x_i \frac{\partial f}{\partial x_i}(tx)\,dt = \sum_{i=1}^n x_i g_i(x) \\
g_i(x) &= \int_0^1 \frac{\partial f}{\partial x_i}(tx)\,dt \in C^{r-1}(\mathbb{R}^n, \mathbb{R})
\end{aligned}
$$

Schritt I.2: Nach Vor. gilt $f(0) = 0$ und $g_i(0) = \frac{\partial f}{\partial x_i}(0) = 0$

$$
\begin{aligned}
i \in \{1,\ldots,n\} :\; g_i(x) = g_i(x) - g_i(0) &= \int_0^1 \frac{d}{dt} g_i(tx)\,dt = \sum_{i=1}^n x_j h_{i,j}(x) \\
h_{i,j}(x) &= \int_0^1 \frac{\partial g_i(tx)}{\partial x_j} dt \in C^{r-2}(\mathbb{R}^n, \mathbb{R}) \\
\Rightarrow\; f(x) &= \sum_{i,j=1} x_i x_j h_{i,j}(x) \\
A(x) &:= \frac{1}{2}\left(h_{i,j}(x) + h_{j,i}(x)\right)_{i,j=1,\ldots,n}
\end{aligned}
$$

Teil II: Die EW von $A(x)$ wechseln auf einer Umgebung von 0 das Vorzeichen nicht.
Die Elemente der Matrix $A(x)$ sind $C^{r-2}-$Funktionen von x. Sei $GL(\mathbb{R}^n) = GL(\mathbb{R}^n, \mathbb{R}^n)$ der Raum der invertierbaren stetigen lin. Abb. $\mathbb{R}^n \to \mathbb{R}^n$. Da $A(0)$ reg., folgt nach dem Inversentheorem für C^{r-2}-Abbildungen: $\exists$ Umg. $U_1(0) \subseteq \mathbb{R}^n$ und $A \in C^{r-2}(U_1(0), GL(\mathbb{R}^n))$ (d.h. $A(x) \in GL(\mathbb{R}^n)$ für $x \in U_1(0)$).
Wir betrachten die EW $L_1(x), \ldots, L_n(x)$ der Matrizen $A(x)$ auf $U_1(0)$. Da alle $A(x)$ auf $U_1(0)$ symm. sind, sind $L_i(x)$ auf $U_1(0)$ reell. Es gilt $L_2 \in C^{r-2}(U_2(0), \mathbb{R})$ für eine Umgebung $U_2(0) \subseteq U_1(0)$. (denn die NS des charakteristischen Polynoms $\det(A(x) - L_2(x) E)$ hängen k-mal stetig diff.bar von den Matrixel. ab)
$r - 2 \geq 1 \Rightarrow L_i \in C^0(U_2(0), \mathbb{R})$.
Wegen der Stetigkeit wechseln die L_i auf $U_i(0)$ ihr Vorzeichen nicht.

Teil III: Konstruktion einer quadr. Form mit $+1$ und -1 in der Hauptdiagonalen in der quadr. Form.

Schritt III.1: Beh.: Die Matrix $A \in C^{r-2}(U_2(0), GL(\mathbb{R}^n))$ läßt sich in einer Umgebung $U_2(0)$ mit Hilfe einer orthog. Matrixfkt. $\hat{C} \in C^{r-2}(U_2(0), GL(\mathbb{R}^n))$ auf Diag.gestalt der folg. Form bringen:

$$\hat{C}^T A \hat{C} =: \hat{D} = \operatorname{diag}(L_1, \ldots, L_n) \in C^{r-2}(U_2(0), GL(\mathbb{R}^n))$$

Bew.: Induktion über die Anzahl der Spalten:
IA: $n = 1 :\; A \in C^{r-2}(U_2(0), GL(\mathbb{R}^n))$ und $A(x) \neq 0$ auf $U_2(0) \Rightarrow$ Diag.gestalt
IV für $n-1$: Die symm. Teilmatrixform $A^* \in C^{r-2}\left(U^*(0), GL\left(\mathbb{R}^{n-1}\right)\right)$ läßt sich auf $U^*(0) \subseteq \mathbb{R}^{n-1}$ mit Hilfe orthog. Matrixfkt. $C^* \in C^{r-2}\left(U^*(0), GL\left(\mathbb{R}^{n-1}\right)\right)$ auf Diag.gestalt bringen:

$${C^*}^T A C^* =: D^* = \operatorname{diag}(L_2, \ldots, L_n) \in C^{r-2}\left(U^*(0), GL\left(\mathbb{R}^{n-1}\right)\right)$$

Ind.bew.: Aus Teil II wissen wir: $A(x)$ hat einen Eigenwert $L_1(x)$ mit $L_1 \in C^{r-2}(U_2(0), \mathbb{R})$. Wir wählen einen Eigenvektor $\hat{v}_1 \in C^{r-2}(U_2(0), \mathbb{R}^n)$ mit $(A(x) - L_1(x) E)\hat{v}_1(x) = 0$ auf $U_2(0)$:
Beh.: Wir können $\hat{v}_1(x)$ durch $\hat{v}_2(x), \ldots, \hat{v}_n(x)$ zu einer Orthonormalmatrix $B(x) = (\hat{v}_1(x), \ldots, \hat{v}_n(x)) \in C^{r-2}(U_2(0), B(\mathbb{R}^n))$ ergänzen.

Bew.: $\hat v_1(x) \in C^{r-2}(U_2(0), \mathbb{R}^n)$. Sei $x^* \in U_2(0)$ mit $\hat v_1(x^*) \neq 0$. Dann ex. Vektoren $a_2, \ldots, a_n$, so daß $\hat v_1(x^*), a_2, \ldots, a_n$ lin. unabh. auf $V(x^*) \subseteq U_2(x)$. Da $\hat v_1(x) \in C^{r-2}(U_2(0), \mathbb{R}^n)$ sind $\hat v_1(x), a_2, \ldots, a_n$ lin. unabh.
Wir konstruieren Vektoren $\hat v_2(x), \ldots, \hat v_n(x)$ mit Hilfe des Orthogonalisierungsverfahrens (E. Schmidt), so daß $\hat v_1(x), \ldots, \hat v_n(x)$ eine ON-Basis bilden: $\hat v_2(x) = a_2 - \hat v_1^T(x)\, a_1 \cdot \frac{v_1(x)}{\|\hat v_1(x)\|}$.
z.z.: $\hat v_i(x) \in C^{r-2}(U_2(0), \mathbb{R}^n), i = 1, \ldots, n$
Sei $\bar x \in U_2(0)$ fest. $\hat v_1(\bar x), \ldots, \hat v_n(x)$ bilden eine ON-Basis, d.h. $\hat v_i^T(\bar x)\, \hat v_j(\bar x) = 0, i, j = 1, \ldots, n, i \neq j$; $\hat v_i^T(x)\, \hat v_i(x) - 1 = 0$.
Es sei $\hat v := \left(\hat v_1^1, \ldots, \hat v_1^n, \hat v_2^T, \ldots, \hat v_n^T\right)$.

$$G(\hat v, x) = \left\{ \begin{array}{ll} \sum_{\alpha=1}^n \hat v_i^\alpha \hat v_j^\alpha & i, j = 1, \ldots, n, i \neq j \\ \sum_{\alpha=1}^n \hat v_i^\alpha \hat v_i^\alpha - 1 & \end{array} \right\} = G_n(\hat v_1, \ldots, \hat v_n; \hat v_1, \ldots, \hat v_n) - E_n = 0$$

Wegen 7.1 gilt $G(\hat x, \bar x) = 0$. Die Jacobimatrix $D_{\hat v} G(\hat v, \bar x)$ hat wegen $\hat v_i(\bar x) \neq 0, i = 1, \ldots, n$, vollen Rang, z.B. für $n = 2$ erhalten wir $\begin{pmatrix} 1 & 1 & 1 & 1 \\ 2\hat v_1^1 & 0 & 0 & 2\hat v_2^1 \end{pmatrix}$.
Satz über implizite Fkt.en $\Rightarrow \exists \hat v : U \to V$ mit $\hat v \in C^{r-2}(U, V) \Rightarrow$ Beh.
Aus der Beh. folgt $B^T A B = \begin{pmatrix} L_1 & \\ & A^* \end{pmatrix}$. Sei C^* aus IV. Wir bilden $\hat B = \begin{pmatrix} 1 & 0 \\ 0 & C^* \end{pmatrix} \in C^{r-2}(U_3(0), B(\mathbb{R}^n))$ mit $U_3(0) \subseteq U_2(0)$.
Man überprüft leicht

$$\begin{aligned} \hat B^T B^T A B \hat B &= \begin{pmatrix} 1 & 0 \\ 0 & C^* \end{pmatrix} \begin{pmatrix} L_1 & 0 \\ 0 & A^* \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & C^* \end{pmatrix} \\ &= \begin{pmatrix} L_1 & 0 \\ 0 & {C^*}^T A^* \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & C^* \end{pmatrix} \\ &= \begin{pmatrix} L_1 & 0 \\ 0 & {C^*}^T A^* C^* \end{pmatrix} \\ &= \begin{pmatrix} L_1 & & \\ & \ddots & \\ & & L_n \end{pmatrix} \end{aligned}$$

Mit $\hat C = B \cdot \hat B$ folgt $\hat D = \hat C^T \cdot A \cdot \hat C$.

Schritt III.2: Mit

$$N := \operatorname{diag}\left(\frac{1}{\sqrt{|\alpha_1|}}, \ldots, \frac{1}{\sqrt{|\alpha_n|}}\right) \in C^{r-2}(U_3(0), GL(\mathbb{R}^n))$$

und geeigneter Umformung der Koo. folgt mit $C := N \cdot \hat C$:

$$C^T(x)\, A(x)\, C(x) = \operatorname{diag}(-1, \ldots, -1, 1 \ldots, 1) =: \tilde D.$$

(hier verwenden wir Teil 2)

Teil IV: Def. des Diffeo. aus der Beh. des Satzes.

$$\begin{aligned} y = F(x) &:= C(x) \cdot x \\ \Rightarrow DF(0) &= DC(0) \cdot 0 + C(0) = C(0) \in GL(\mathbb{R}^n) \end{aligned}$$

Inv.theorem $\to \exists U(0)$ und $V(0)$ mit $F(U(0)) = V(0)$, $F \in C^{r-2}(U(0), V(0))$.
Es gilt $C^{-1} = C^T \in C^{r-2}(U(0), V(0)) \to F^{-1}(y) = x = C^T(x) \cdot y$.
Mit der Bez. $f(x) = x^T A x$ folgt T

$$f \circ F^{-1}(y) = \left(C^T(x)\, y\right)^T A(x) \left(C^T(x)\, y\right) = y^T C(x)\, A(x)\, C^T(x)\, y = y^T \tilde D y.$$

Bemerkung 7.4

Sei $\bar{x} \in \mathbb{R}^n$ ein nichtentarteter kritischer Punkt, dann gilt:

$$f(x) \circ F^{-1}(y) = f(\bar{x}) - \sum_{i=1}^{k} y_i^2 + \sum_{j=k+1}^{n} y_j^2$$

Bemerkung 7.5

Das Morse-Lemma gilt auch für $1 \leq r \leq 3$ (siehe [2]).

Verallgemeinerung des Morse-Lemma

(P) $\min \{ f(x) | \, h_i(x) = 0, \, i \in I = \{1, \ldots, m\}, \, g_j(x) \geq 0, \, j \in J = \{1, \ldots, s\}\}, \, f, h_i, g_j \in C^2(\mathbb{R}^n, \mathbb{R})$

Definition 7.2

Ein kritischer Punkt $\bar{x}$ für (P) heißt nichtentartet, wenn

1. LICQ ist in $\bar{x}$ erfüllt:
   - $h_i(\bar{x}) = 0, \, i \in I, \, g_j(\bar{x}) \geq 0, \, j \in J$
   - $\exists! \, \bar{\lambda}_i, \, \bar{\mu}_j, \, i \in I, \, j \in J_0(\bar{x})$ mit $Df(\bar{x}) = \sum_{i \in I} \bar{\lambda}_i Dh_i(\bar{x}) + \sum_{j \in J_0} \bar{\mu}_j Dg_j(\bar{x})$
2. $\bar{\mu}_j \neq 0, \, j \in J_0$
3. $D^2 L^{J_0}\left(\bar{x}, \bar{\lambda}, \bar{\mu}^{J_0}\right)\big|_{T_{\bar{x}}M}$ regulär, $T_{\bar{x}}M := \{\psi \in \mathbb{R}^n \,|\, Dh_i(\bar{x})\psi = 0, \, i \in I, \, Dg_j(\bar{x})\psi = 0, \, j \in J_0\}$

Definition 7.3

1. $LI := LI(\bar{x}) := |\{ j \in J_0(\bar{x}) | \, \bar{\mu}_j < 0\}|$ heißt linearer Index.
2. $LCI := LCI(\bar{x}) := |\{ j \in J_0(\bar{x}) | \, \bar{\mu}_j > 0\}|$ heißt linearer Co-Index.
3. $QI := QI\left(D^2 f(\bar{x})\right)$ Anzahl der neg. Eigenwerte von $D_x^2 L^{J_0}$, heißt quadratischer Index.
4. $QCI := QCI\left(D^2 f(\bar{x})\right)$ Anzahl der pos. Eigenwerte von $D_x^2 L^{J_0}$, heißt quadratischer Co-Index.

Satz 7.3

Seien $f, h_i, g_j \in C^3(\mathbb{R}^n, \mathbb{R}), \, i \in I, \, j \in J$ und $\bar{x}$ ein nichtentarteter kritischer Punkt mit $LI(\bar{x}) = k_1$, $QI(\bar{x}) = k_2$ und $J_0(\bar{x}) = \{1, \ldots, p\}$.

Dann existieren Umgebungen U, V von $\bar{x}, 0 \in \mathbb{R}^n$ und ein Diffeomorphismus $F : U \to V$ $(F \in C^1(U, V))$, $y = F(x)$, $F(\bar{x}) = 0$ und es gilt:

1. $y_{m+1} \geq 0, \, y_{m+p} \geq 0$
2. $f \circ F^{-1}(0, \ldots, 0, y_{m+1}, \ldots, y_n) =$

$$f(\bar{x}) = \sum_{i_1=m+1}^{m+k_1} y_{i_1} + \sum_{i_2=m+k_1+1}^{m+p} y_{i_2} - \sum_{i_3=m+p+1}^{m+p+k_2} y_{i_3}^2 + \sum_{i_4=m+p+k_2+1}^{n} y_{i_4}^2$$

Bemerkung 7.6

Aus den Optimalitätsbedingungen in Kapitel 3 folgt:

1. Wenn $LI(\bar{x}) = 0$, so ist $\bar{x}$ ein stationärer Punkt.
2. Wenn $LI(\bar{x}) = 0$ und $QI(\bar{x}) = 0$, so ist $\bar{x}$ ein lokaler Minimalpunkt.
3. Wenn $LI(\bar{x}) = 0$ und $QI(\bar{x}) = n$, so ist $\bar{x}$ ein lokaler Maximalpunkt in der Menge der stationären Punkte.

4. Wenn $0 < LI(\bar{x}) < p$ und $QI(\bar{x}) = n$, so ist $\bar{x}$ ein lokaler Maximalpunkt in der Menge der kritischen Punkte.
5. Wenn $LI(\bar{x}) = 0$ und $0 < QI(\bar{x}) < n$, so ist $\bar{x}$ ein lokaler Sattelpunkt in der Menge der stationären Punkte.
6. Wenn $0 < LI(\bar{x}) < p$ und $0 < QI(\bar{x}) < n$, so ist $\bar{x}$ ein lokaler Sattelpunkt in der Menge der kritischen Punkte.

8 Minimax- und Sattelpunktsätze ([1])

(P) $\min\{f(x) | x \in M\}$, $M := \{x \in \mathbb{R}^n | g_j(x) \geq 0,\ j \in J\}$, Slaterbedingung:
$\exists \bar{x} \in \mathbb{R}^n : g_j(\bar{x}) > 0,\ j \in J$, $L(x,\mu) = f(x) - \sum_{j \in J} \mu_j g_j(x)$

Satz 8.1

Es sei (P) ein konvexes Optimierungsproblem, Slaterbedingung erfüllt. $\bar{x} \in \mathbb{R}^n$ ist optimal für (P), gdw. $\exists \bar{\mu} \geq 0$, so daß $L(\bar{x},\mu) \leq L(\bar{x},\bar{\mu}) \leq L(x,\bar{\mu})$ $\forall x \in \mathbb{R}^n$, $\forall \mu \geq 0$

Dieses Minimax-Prinzip soll verallgemeinert werden:
Es seien $X \subseteq \mathbb{R}^n$, $Y \subseteq \mathbb{R}^n$, $X \neq \emptyset$, $Y \neq \emptyset$ und konvex. $\psi(x,y)$ sei eine Funktion mit $x \in X$, $y \in Y$, die für jeden Punkt $y \in Y$ konvex auf X und für jeden Punkt $x \in X$ konkav ist.

Ziel: Existenzsätze für Sattelpunkte bzw. Minimax-Punkte, d.h.

$$\psi(x,\bar{y}) \leq \psi(\bar{x},\bar{y}) \leq \psi(\bar{x},y), \ \forall x \in X, \forall y \in Y.$$

Geometrische Interpretation: $x \in X, y \in Y$, betrachten für jeden Punkt $y \in Y$ den Epigraph

$$\mathrm{E}_\psi(y) := \{(x,z) \in \mathbb{R}^n \times \mathbb{R} | z \geq \psi(x,y), \ \forall x \in X\}$$

und für jedes $x \in X$ den Hypograph

$$h_\psi(x) := \{(y,z) \in \mathbb{R}^n \times \mathbb{R} | z \leq \psi(x,y), \ \forall y \in Y\}$$

Epigraph über $X \times Y$:

$$\mathrm{E}_\psi := \{(x,y,z) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R} | z \geq \psi(x,y), \ (x,y) \in X \times Y\}$$

Hypograph über $X \times Y$:

$$h_\psi := \{(x,y,z) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R} | z \leq \psi(x,y), \ (x,y) \in X \times Y\}$$

Weiterhin betrachten man zwei lineare Unterräume von $\mathbb{R}^{n+m+1}$

$$\begin{aligned} L(\tilde{y}) &:= \left\{ (x,y,z) \in \mathbb{R}^{n+m+1} \,\middle|\, y = \tilde{y} \right\} \\ L(\tilde{x}) &:= \left\{ (x,y,z) \in \mathbb{R}^{n+m+1} \,\middle|\, x = \tilde{x} \right\} \end{aligned}$$

Offenbar gilt: $\mathrm{E}_\psi(\tilde{y}) = \mathrm{E} \cap L(\tilde{y})$ und $h_\psi(\tilde{x}) = h_\psi \cap L(\tilde{x})$.

Für einen festen Punkt $y \in Y$ nennt man deshalb die Menge $\mathrm{E}_\psi(y)$ einen y-Schnitt des Epigraphen E_ψ und für $x \in X$ $h_\psi(x)$ einen x-Schnitt des Hypographen h_ψ, der auf $X \times Y$ definierten Funktion $\psi(x,y)$.

Für jedes Punktepaar $y \in Y$, $x \in X$ existieren Punkte $x_y \in X$, $y_x \in Y$ derart, daß

$$\begin{aligned} \psi(x_y, y) &= \min_{x \in X} \psi(x, y), \\ \psi(x, y_x) &= \max_{y \in Y} \psi(x, y). \end{aligned}$$

Dann folgt mit obiger Definition des Epigraphen und Hypographen:

$$\begin{aligned} z_y := \psi(x_y, y) = \min_{x \in X} \psi(x, y) &\leq \psi(x, y) \leq z \, \forall (x, z) \in \mathrm{E}_\psi(y) \\ z_x := \psi(x, y_x) = \max_{y \in Y} \psi(x, y) &\geq \psi(x, y) \geq z \, \forall (y, z) \in h_\psi(x) \end{aligned}$$

Wenn es einen Punkt $y^0 \in Y$ derart gibt, daß $z_{y^0} \geq z_y$, $\forall y \in Y$ gilt, dann nennt man $\mathrm{E}_\psi(y^0)$ einen höchsten y-Schnitt von E_ψ. Wenn es einen Punkt $x^0 \in X$ mit $z_{x^0} \leq z_x$, $\forall x \in X$ gibt, dann bezeichnet man $h_\psi(x^0)$ einen tiefsten x-Schnitt des Hypographen.

Seien nun $x^0 \in X$, $y^0 \in Y$ Punkte mit diesen Eigenschaften, so gilt:

$$\begin{aligned} z_{y^0} = \psi(x_{y^0}, y^0) &= \max_{y \in Y} \min_{x \in X} \psi(x, y) \\ z_{x^0} = \psi(x^0, x_{x^0}) &= \min_{x \in X} \max_{y \in Y} \psi(x, y) \\ \Rightarrow z_{y^0} = \psi(x_{y^0}, y^0) &\leq \psi(x^0, y^0) \leq \psi(x^0, y_{x^0}) = z_{x^0} \\ \Rightarrow \max_{y \in Y} \min_{x \in X} \psi(x, y) &\leq \psi(x^0, y^0) \leq \min_{x \in X} \max_{y \in Y} \psi(x, y) \end{aligned}$$

Wenn $z_{y^0} = z_{x^0}$, d.h. wenn

$$\max_{y \in Y} \min_{x \in X} \psi(x, y) = \psi(x^0, y^0) = \min_{x \in X} \max_{y \in Y} \psi(x, y)$$

gilt, hat (x^0, y^0) die Eigenschaft, daß der entsprechende y^0-Schnitt ein höchster y-Schnitt des Epigraphen E_ψ $(y \in Y)$ ist und der entsprechende x^0-Schnitt ein tiefster x-Schnitt des Hypographen h_ψ $(x \in X)$ ist.

Lemma 8.1

Es seien $f_1(x), \ldots, f_k(x)$ konvexe Funktionen definiert auf einer nichtleeren konvexen Teilmenge $X \subseteq \mathbb{R}^n$. Dann ist die Menge

$$M' := \{x \in X \mid f_j(x) < 0, \, j = 1, \ldots, k\}$$

leer, gdw. es einen Punkt $\tilde{\mu} \in \mathbb{R}^k_+$ mit $\tilde{\mu} \neq 0$ gibt, so daß gilt:

$$\sum_{j=1}^{k} \tilde{\mu}_j f_j(x) \geq 0, \, \forall x \in X.$$

Bew.: Vorbetrachtungen: $\forall x \in \mathbb{R}^k$ ist die Menge $M'(x)$ konvex.

Beh.: Die Menge $A := \{r \in \mathbb{R}^k \mid M'(r) \neq \emptyset\}$ ist konvex und k-dimensional.

1. Konvexität: Es seien $x^0 \in X$ und $r^0_j > f_j(x^0)$, $j = 1, \ldots, k$.
Wegen $x^0 \in M'(r^0)$ ist $M'(r^0) \neq \emptyset$ und daher auch $A \neq \emptyset$.
Für bel. Punkte v^1, v^2 der Menge A gilt $M(r^i) \neq \emptyset, i = 1, 2$.
$\Rightarrow \exists x^i \in X$ mit $f_j(x^i) \leq r^i_j, j = 1, \ldots, k, i = 1, 2$
Konv. $\Rightarrow$ Sei $\lambda_1, \lambda_2 \geq 0$ mit $\lambda_1 + \lambda_2 = 1$:

$$f_j(\lambda_1 x^1 + \lambda_2 x^2) \leq \lambda_1 f(x^1) + \lambda_2 f(x^2) \leq \lambda_1 r^1_j + \lambda_2 r^2_j, \, j = 1, \ldots, k$$

$\lambda_1 x^1 + \lambda_2 x^2 \in X$ konvex $\Rightarrow$ $M'(\lambda_1 r^1 + \lambda_2 r^2) \neq \emptyset$ $\Rightarrow$ $\lambda_1 r^1 + \lambda_2 r^2 \in A$.

2. k-dimensional: Zu jedem $v^0 \in A$ gibt es ein Pkt. $x^0 \in X$ mit $f_j\left(x^0\right) < v_j^0, j = 1, \dots, k$.
   Dann gilt aber auch für jeden Pkt. v der $k-$dim. Menge

$$\mathbb{R}_+^k \quad := \quad \left\{ v \in \mathbb{R}^k \,\middle|\, r_j \geq r_j^0, j = 1, \dots, k \right\}$$

$$f_j\left(x^0\right) \leq r_j, j = 1, \dots, k \Rightarrow \mathbb{R}_+^k\left(r^0\right) \subseteq A \Rightarrow \dim A = k \; (*)$$

$(\Rightarrow)$: Sei $M' = M'(0) = \emptyset$ und daher $0 \notin A$. Nach dem Trennungssatz (S. 5.1) ex. in $\mathbb{R}^k$ ein Vektor $\tilde{u} \neq 0$ mit

$$A \subset \bar{H}^+ := \left\{ v \in \mathbb{R}^k \,\middle|\, \tilde{u}^T v \geq 0 \right\}$$

$\Rightarrow \tilde{u}^T v \geq 0, v \in A$, wobei ein Pkt. $\tilde{v} \in A$ mit $\tilde{u}^T \tilde{v} > 0$ ex.
Wegen $(*)$ liegen für $j_0 \in \{1, \dots, k\}$ die Halbgeraden

$$P_{j_0} := \{ v \in \mathbb{R}^n \,|\, v_j = \tilde{v}_j, j \in \{1, \dots, k\} \setminus \{j_0\}, v_{j_0} = \tilde{v}_{j_0} + t, t \geq 0 \}$$

in der Menge A.

$$\Rightarrow \tilde{u}^T \tilde{v} + \tilde{u}_{j_0} t \geq 0 \, \forall t \geq 0 \Rightarrow \tilde{u}_{j_0} \geq 0 \, \forall j_0 \in \{1, \dots, k\}. \; (**)$$

Es seien $\hat{x} \in X$ bel. fix und $t > 0$ bel. fix. Weiter sei $v(\hat{x}, t)$ ein Pkt. mit den Koo. $v_j(\hat{x}, t) := f_j(\hat{x}) + t, j = 1 \dots, k$ in $\mathbb{R}^k$.

$$\begin{aligned} \Rightarrow f_j(\hat{x}) &< v_j(\hat{x}, t), j = 1, \dots, k. \\ \Rightarrow v(\hat{x}, t) &\in A \\ \overset{(**)}{\Rightarrow} \tilde{u}^T v(\hat{x}, t) &\geq 0, \hat{x} \in X, t > 0 \\ \Rightarrow \sum_{j=1}^{k} \tilde{u}_j f_j(x) + t \cdot \sum_{j=1}^{k} \tilde{u}_j \geq 0, x \in X, t &> 0 \end{aligned}$$

Da $t > 0$ bel. $\Rightarrow$ Beh.

$(\Leftarrow)$ Ann.: Sei $M' \neq \emptyset$ und $x' \in M'$ bel., dann gilt für jedes $u \in \mathbb{R}_+^k$ mit $u \neq 0$ die Ungl. $\sum_{j=1}^k \mu_j f_j(x') < 0$. Wid. zur Vor.

Bezeichnung 8.1

Eine eindeutige Abbildung von $\mathbb{R}^n$ in $\mathbb{R}^* = \mathbb{R} \cup \{-\infty, \infty\}$ heißt verallgemeinerte Funktion.

Bezeichnung 8.2

Es sei f eine über einer nichtleeren Menge $M \subseteq \mathbb{R}^n$ definierte Funktion bzw. verallgemeinerte Funktion. Dann nennt man $\underline{f}(x) := \underline{\lim}_{\varepsilon \to 0} \{ f(z) | \, z \in U(x, \varepsilon) \cap M \}, x \in clM$ die der verallgemeinerten Funktion $f(x)$ zugeordnete verallgemeinerte Funktion.

Lemma 8.2

Es sei $f(x)$ eine auf $M \subseteq \mathbb{R}^n$, $M \neq \emptyset$ definierte Funktion.Dann besitzt $\underline{f}(x)$ die folgenden Eigenschaften:

1. $\underline{f}(x) \leq f(x), \forall x \in M$
2. für jedes $x \in M$ und für jede Punktfolge $\left\{x^i\right\}_{i=1}^{\infty}, x^i \in M \; (i = 1, 2, \dots)$ mit $\lim_{i \to \infty} x^i = x$ gilt: $\underline{f}(x) = \underline{\lim_{i \to \infty} f\left(x^i\right)}$

Definition 8.1

Es sei $f(x)$ eine auf $M \subseteq \mathbb{R}^n$, $M \neq \emptyset$ definierte Funktion und $\underline{f}(x)$ mit $\bar{x} \in c\ell M$ die ihr gemäß Bezeichnung 8.2 zugeordnete verallgemeinerte Funktion. $f(x)$ heißt in $x^0 \in M$ unterhalbstetig (uhs), wenn $\underline{f}(x^0) = f(x^0)$. $f(x)$ heißt uhs auf M, wenn $f(x)$ uhs $\forall x \in M$, d.h. $\underline{f}(x) = f(x)$, $\forall x \in M$.

Lemma 8.3

Sei M eine nichtleere und abgeschlossene Menge im $\mathbb{R}^n$ und f eine auf M definierte Funktion. Dann sind folgende Aussagen äquivalent:

1. $f(x)$ ist uhs auf M
2. der Epigraph $\mathrm{E}_f := \left\{ (x, x_{n+1}) \in \mathbb{R}^{n+1} \,\middle|\, x_{n+1} \geq f(x),\, x \in M \right\}$ ist eine abgeschlossene Menge
3. für jede Zahl $\alpha \in \mathbb{R}$ ist die obere Niveaumenge $M_\alpha := \{x \in \mathbb{R}^n \,|\, f(x) \leq \alpha,\, x \in M\}$ abgeschlossen im $\mathbb{R}^n$.

Bew.:

1⇒2: Sei $f(x)$ uhs. auf M, und es sei $\left(x^0, x^0_{n+1}\right)$ ein bel. HP in E_f.
Dann gibt es eine Pkt.folge $\left\{ \left(x^i, x^i_{n+1}\right) \right\} \in \mathrm{E}_f$ mit

$$\begin{aligned} \lim_{i\to\infty} x^i &= x^0, \\ \lim_{i\to\infty} x^i_{n+1} &= x^0_{n+1}. \end{aligned}$$

$$\begin{aligned} \overset{2}{\Rightarrow}\ x^i_{n+1} &\geq f\left(x^i\right) \\ \Rightarrow \underline{\lim}_{i\to\infty} x^i_{n+1} = \lim_{i\to\infty} x^i_{n+1} = x^0_{n+1} &\geq \lim_{i\to\infty} f\left(x^i\right) \end{aligned}$$

L.8.2,uhs. $\Rightarrow$ $x^0_{n+1} \geq \underline{f}\left(x^0\right) = f\left(x^0\right) \Rightarrow \left(x^0, x^0_{n+1}\right) \in \mathrm{E}_f$.

2⇒3: Sei E_f eine in $\mathbb{R}^{n+1}$ abg. Menge. Sei $\alpha \in \mathbb{R}$ bel.

$$\begin{aligned} R_\alpha &:= \left\{ (x, x_{n+1}) \in \mathbb{R}^{n+1} \,\middle|\, x \in \mathbb{R}^n, x_{n+1} = \alpha \right\} \\ \Rightarrow R_\alpha \cap \mathrm{E}_f &= \left\{ (x, x_{n+1}) \in \mathbb{R}^{n+1} \,\middle|\, \alpha \geq f(x), x \in M, x_{n+1} = \alpha \right\} \end{aligned}$$

ist als Durchschnitt abg. Mengen ebenfalls abgeschlossen.

3⇒1: Ann.: $\exists x^0 \in M$ mit $\underline{f}\left(x^0\right) < f\left(x^0\right)$.
Es sei α mit $\underline{f}\left(x^0\right) < \alpha < f\left(x^0\right)$. Nach Def. des Epigr. gilt $\left(x^0, \alpha\right) \notin \mathrm{E}_f$. $(*)$
Bez.8.2 $\Rightarrow$ $\underline{\lim}_{\varepsilon\to 0^+} \left\{ f(x) \,|\, x \in U\left(x^0, \varepsilon\right) \cap M \right\} < \alpha$
$\Rightarrow$ $\exists \varepsilon_0 > 0$ derart, daß $\inf \left\{ f(x) \,|\, x \in U\left(x^0, \varepsilon\right) \cap M \right\} < \alpha$, $\varepsilon \in (0, \varepsilon_0)$.
Für bel. Folge $\{\varepsilon_k\}$ mit $0 < \varepsilon_{k+1} < \varepsilon_k < \varepsilon_0$ gilt daher

$$\inf \left\{ f(x) \,|\, x \in U\left(x^0, \varepsilon_k\right) \cap M \right\} < \alpha.$$

$\Rightarrow$ $\exists \left\{x^k\right\}$ mit

- $x^k \in U\left(x^0, \varepsilon_k\right) \cap M$, $\lim_{k\to 0} x^k = x^0$
- $\inf \left\{ f(x) \,|\, x \in U\left(x^0, \varepsilon_k\right) \cap M \right\} \leq f\left(x^k\right) < \alpha$.

$\overset{2}{\Rightarrow}$ $\left(x^k, \alpha\right) \in \mathrm{E}_f$; $\lim_{k\to\infty} \left(x^k, \alpha\right) = \left(x^0, \alpha\right) \Rightarrow \left(x^k, \alpha\right) \in R_\alpha \cap \mathrm{E}_f$.
Nach Vor. ist M_α abg. $\Rightarrow$ $\lim_{k\to\infty} \left(x^k, \alpha\right) = \left(x^0, \alpha\right) \in R_\alpha \cap E_f \Rightarrow \left(x^0, \alpha\right) \in \mathrm{E}_f$.
Wid. zu $(*)$
$\Rightarrow$ $f\left(x^0\right) \geq f(x)\ \forall x \in M$.
L.8.2a $\Rightarrow$ $\underline{f}(x) = f(x)\ \forall x \in M \Rightarrow f(x)$ ist uhs. auf $M \Rightarrow 1$.

Lemma 8.4

Es sei X eine nichtleere, konvexe und kompakte Menge in $\mathbb{R}^n$, Y eine nichtleere Menge in $\mathbb{R}^m$ und $\psi(x,y)$ eine auf $X \times Y$ definierte Funktion, die in jedem Punkt $y \in Y$ uhs und konvex auf X ist. Es sei $\Phi = \{x \in X | \psi(x,y) \leq 0,\, y \in Y\}$. Dann gibt es Punkte $y_k^j \in Y_k$ $(j = 1, \ldots, k)$ und einen Punkt $\tilde{u} \in \mathbb{R}_+^k$ mit $\tilde{u} \neq 0$, $\sum_{j=1}^k \tilde{u}_j \psi\left(x, y^j\right) > 0$, $x \in X$.

Bew.: Es seien $y \in Y$ und $\varepsilon > 0$ bel. fix. Sei

$$C(y,\varepsilon) := \{x \in X | \psi(x,y) \leq \varepsilon\}.$$

Nach S. 2.2 ist $\psi(\cdot, y)$ uhs. auf X und X kompakt.
L.8.3c $\Rightarrow$ $C(y,\varepsilon)$ kompakt. Wir betrachten $C^*(y,\varepsilon) := \mathbb{R}^n \setminus C(y,\varepsilon)$ offen und zsh.
Beh.: $\bigcap_{y \in Y, \varepsilon > 0} C(y,\varepsilon) = \emptyset$.
Ann.: $\bigcap_{y \in Y, \varepsilon > 0} C(y,\varepsilon) \neq \emptyset \Rightarrow \exists \hat{x} \in X$ mit $\psi(\hat{x}, x) \leq \varepsilon \forall y \in Y \forall \varepsilon > 0 \Rightarrow \psi(\hat{x}, y) \leq 0 \, \forall y \in Y$ Wid.
$\Rightarrow \forall x \in X \, \exists y_x \in Y, \varepsilon_x > 0 : x \in C^x(y_x, \varepsilon_x)$
$\Rightarrow$ Das Sys. $\{C^*(y,\varepsilon) | y \in Y, \varepsilon > 0\}$ offener Mengen im $\mathbb{R}^n$ überdeckt die kompakte Menge X. Nach dem Borrelschen Überdeckungssatz gibt es endl. viele Mengen dieses Systems, die ebenfalls X überdecken. Es seien dies die Mengen $C^*(y_j, \varepsilon_j)$ $(j = 1, \ldots, k)$.

$$\Rightarrow \bigcap_{j=1}^{k} C(y_j, \varepsilon_j) = \emptyset \Rightarrow \{x \in X | \psi(x, y_j) < \varepsilon_j, j = 1, \ldots, k\} = \emptyset$$

L. 8.1 (hier benötigen wir Konvexivität) $\Rightarrow \exists \tilde{u} \in \mathbb{R}_+^k$ mit $\tilde{u} \neq 0$,

$$\sum_{j=1}^{k} \tilde{u}_j \left(\psi(x, y_j) - \varepsilon_j\right) \geq 0,\, x \in X.$$

Wegen $\varepsilon_j > 0$ folgt

$$\sum_{j=1}^{k} \tilde{u}_j \psi(x, y_j) \geq \sum_{j=1}^{k} \tilde{u}_j \varepsilon_j > 0,\, x \in X \Rightarrow \text{Beh.}$$

Lemma 8.5 (Verallg. des Satzes von Weierstraß)

Es sei $f(x)$ eine auf einer nichtleeren kompakten Menge $M \subseteq \mathbb{R}^n$ definierte Funktion, die uhs (bzw. ohs) auf dieser Menge ist. Dann gibt es einen Punkt $x^1 \in M$ (bzw. $x^2 \in M$), so daß:

$$f\left(x^1\right) = \min_{x \in M} f(x) \;\; (\text{bzw. } f\left(x^2\right) = \max_{x \in M} f(x))$$

Satz 8.1 (Minimax-Prinzip)

Es seinen $X \subseteq \mathbb{R}^n$, $Y \subseteq \mathbb{R}^m$ nichtleere kompakte und konvexe Mengen; weiter sei $\psi(x,y)$ eine auf $X \times Y$ definierte Funktion mit folgenden Eigenschaften:

1. Für jeden Punkt $y \in Y$ ist sie uhs und konvex auf X.
2. Für jeden Punkt $x \in X$ ist sie ohs und konkav auf Y.

Dann existiert ein Sattelpunkt $\left(x^0, y^0\right) \in X \times Y$ der Funktion $\psi(x,y)$ auf $X \times Y$, und es gilt:

$$\max_{y \in Y} \min_{x \in X} \psi(x,y) = \psi\left(x^0, y^0\right) = \min_{x \in X} \max_{y \in Y} \psi(x,y)$$

Bew.: Anw. von Lemma 8.5:

$$\begin{aligned} f(x) := \sup_{y\in Y}\psi(x,y) &= \max_{y\in Y}\psi(x,y)\ \forall x\in X \\ g(y) := \inf_{x\in X}\psi(x,y) &= \min_{x\in X}\psi(x,y)\ \forall y\in Y \end{aligned}$$

Offenbar gilt: $f(x)$ ist auf X konvex und $g(y)$ auf Y konkav.
Nach Lemma 8.3 folgt

$$\begin{aligned} \mathrm{E}_\psi(y) &= \{(x,z)\in\mathbb{R}^n\times\mathbb{R}|\ z\geq\psi(x,y), x\in X\}, y\in Y, \\ h(x) &= \{(x,z)\in\mathbb{R}^n\times\mathbb{R}|\ z\leq\psi(x,y), y\in Y\}, x\in X \end{aligned}$$

abg. und konvex.

$$\begin{aligned} \to\ \mathrm{E}_f := \bigcap_{y\in Y}\mathrm{E}_\psi(y) &= \left\{(x,z)\in\mathbb{R}^n\times\mathbb{R}|\ z\geq\max_{y\in Y}\psi(x,y), x\in X\right\} \\ h_g := \bigcap_{x\in X}h_\psi(x) &= \left\{(x,z)\in\mathbb{R}^n\times\mathbb{R}|\ z\leq\min_{x\in X}\psi(x,y), y\in Y\right\} \end{aligned}$$

abg. und konvex.
L.8.3. $\to$ $f(x)$ uhs. auf X und $g(y)$ ohs. auf Y.
$(*)$ Nach L. 8.5 ex. Zahlen μ_1, μ_2 mit

$$\begin{aligned} \mu_1 := \inf_{x\in X}f(x) = \min_{x\in X}f(x) &= \min_{x\in X}\max_{y\in Y}\psi(x,y) \\ \mu_2 := \sup_{y\in Y}g(y) = \max_{y\in Y}g(y) &= \max_{y\in Y}\min_{x\in X}\psi(x,y) \end{aligned}$$

Es seien x^0, y^0 Pkt. mit $f(x^0) = \mu_1 = \min_{x\in X}f(x)$, $g(y^0) = \mu_2 = \max_{y\in Y}g(y)$

$$\begin{aligned} \Rightarrow\ \mu_1 = f(x^0) = \max_{y\in Y}\psi(x^0,y) &\geq \psi(x^0,y^0) \\ \mu_2 = g(y^0) = \min_{x\in X}\psi(x,y^0) &\leq \psi(x^0,y^0) \\ \Rightarrow\ \mu_2 \leq \psi(x^0,y^0) &\leq \mu_1 \end{aligned}$$

z.z.: $\mu_2 = \mu_1$.
$(*)$ $\Rightarrow$ $\max_{y\in Y}\psi(x,y)\geq\mu_1$, $x\in X$
Für bel. $y\in Y$ und bel. $\varepsilon>0$ def. wir

$$\begin{aligned} C'(y,\varepsilon) &:= \{x\in X|\,\psi(x,y)-\mu_1+\varepsilon\leq 0\} \\ C'_\varepsilon &:= \bigcap_{y\in Y}C'(y,\varepsilon) \end{aligned}$$

Beh.: $C'_\varepsilon=\emptyset$. Ann.: $C'_\varepsilon\neq\emptyset$.
$\Rightarrow$ $\exists x_\varepsilon\in X:\psi(x_\varepsilon,y)-\mu_1+\varepsilon\leq 0\,\forall y\in Y$
$\Rightarrow$ $\exists y_{x_\varepsilon}\in Y:\max_{y\in Y}\psi(x_\varepsilon,y)=\psi(x_\varepsilon,y_{x_\varepsilon})<\mu_1$ Wid.
Nach Lemma 8.4 gibt es Pkt.e $y^j\in Y$ $(j=1,\ldots,k)$ und $\tilde{u}\in\mathbb{R}^k_+$ mit

$$\tilde{u}\neq 0, \sum_{j=1}^k\tilde{u}_j\left[\psi(x,y^j)-\mu_j+\varepsilon_j\right]>0, x\in X.$$

Da o.B.d.A. $\sum_{j=1}^{n} \tilde{u}_j = 1$ vorausgesetzt werden kann, und da für $x \in X\ \psi(x, y)$ konkav auf Y ist, gilt (mit Jensen's Ungl.)

$$\sum_{j=1}^{n} \tilde{u}_j \psi\left(x, y^j\right) \leq \psi\left(x, \sum_{j=1}^{k} \tilde{u}_j y^j\right) = \psi(x, \tilde{y}),\ x \in X$$

Dabei gilt $\tilde{y} := \sum_{j=1}^{k} \tilde{u}_j y^j \in Y \ \Rightarrow\ \psi(x, \tilde{y}) - \mu_1 + \varepsilon > 0\, \forall x \in X$

$$\overset{(*)}{\Rightarrow}\ \mu_2 = \max_{y \in Y} \min_{x \in X} \psi(x, y) \quad \geq \quad \min_{x \in X} \psi(x, \tilde{y}) \geq \mu_1 + \varepsilon$$

$\varepsilon > 0$ bel. $\Rightarrow\ \mu_2 \geq \mu_1\ \Rightarrow\ \mu_2 = \psi\left(x^0, y^0\right) = \mu_1\ \Rightarrow$ Beh.

Bemerkung 8.2

Die Aussage des Satzes bleibt erhalten, wenn wir die Konvex- bzw. Konkavität von $\psi(x, y)$ leicht abschwächen. (Satz von Soin-K...)

Definition 8.2

1. $F(x)$ heißt quasikonvex auf einer konvexen Menge M, wenn $\{x \in \mathbb{R}^n \,|\, F(x) \leq \lambda\}$ konvex $\forall \lambda \in \mathbb{R}$.
2. $F(x)$ heißt quasikonkav auf einer konvexen Menge M, wenn $\{x \in \mathbb{R}^n \,|\, F(x) \leq \lambda\}$ konkav $\forall \lambda \in \mathbb{R}$.

Bemerkung 8.3

Unter gewissen zusätzlichen Bed. läßt sich der Satz auch für nicht kompakte Mengen anwenden.

Satz 8.2 (Minimax-Satz von J. v. Neumann in der Spieltheorie)

Es seien $\psi(x, y) := \sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij} x_i y_j$,
$X := \{x \in \mathbb{R}^n \,|\, x_i \geq 0,\, i = 1, \ldots, n,\, \sum x_i = 1\}$
$Y := \{y \in \mathbb{R}^n \,|\, y_j \geq 0,\, j = 1, \ldots, m,\, \sum y_j = 1\}$
Dann existiert ein Punkt $\left(x^0, y^0\right) \in X \times Y$ mit:

$$\max_{y \in Y} \min_{x \in X} \left\{ \sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij} x_i y_j \right\} = \sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij} x_i^0 y_j^0 = \min_{x \in X} \max_{y \in Y} \left\{ \sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij} x_i y_j \right\}$$

Literatur

[1] Nozicka, Grygerva, Lommetzsch; Geometrie konvexer Mengen und konvexer Analysis, Akademie-Verlag, 1988

[2] Jougen, Jonker: ???

[3] Bank, Guddat, Klatte, Kummer, Taromer: Nonlinear parametric Optimization, Akademie Verlag, Berlin 1982

[4] Nozicka, Guddat,...: Linear parametric Optimization, Akademie-Verlag, 1974

[5] Guddat: Parametric optimization: singularities, pathfollowing and jumps, Teubner 1990

[6] Elster et.al.: Einführung in die nichtlineare Optimierung, Leipzig 1977