Nichtlineare Optimierung

Roland Griesse

TECHNISCHE UNIVERSITÄT CHEMNITZ

Numerische Mathematik

Chemnitzer Skiseminar

Gerlosberg, 07.–14. März 2009

Gliederung

1. Konvexe Optimierung
   - Bedeutung
   - Beispiele
   - Charakterisierung konvexer Funktionen
2. Optimalitätsbedingungen
   - KKT-Bedingungen
   - Constraint Qualifications
   - KKT-Bedingungen bei konvexen Aufgaben
   - Hinreichende Bedingungen zweiter Ordnung
3. Unendlich-dimensionale Optimierung
   - KKT-Bedingungen
   - Constraint Qualifications
   - Hinreichende Bedingungen zweiter Ordnung

Konvexe Optimierungsaufgaben

Konvexe Optimierungsaufgabe

$$\begin{array}{rll} \min & f(x) & \text{konvex} \\ \text{sodass} & x \in C & \text{konvex} \end{array}$$

Theorem (Eigenschaften konvexer Optimierungsaufgaben)

1. *Jedes lokale Minimum ist bereits globales Minimum.*
2. *Die Menge globaler Minima ist konvex (evtl. leer).*
3. *Ist f strikt konvex, dann ist die Lösung (falls sie existiert) eindeutig.*

Beachte: Diese Eigenschaften gehen verloren bei nicht-konvexen Aufgaben.

Beispiele konvexer Optimierungsaufgaben

Lineare Programme (LP)

$$\begin{aligned} \text{Minimiere} \quad & c^\top x \\ \text{sodass} \quad & A\,x \le b \end{aligned} \qquad\qquad \begin{aligned} \text{Minimiere} \quad & c^\top x \\ \text{sodass} \quad & A\,x = b, \quad x \ge 0 \end{aligned}$$

Quadratische Programme (QP)

$$\begin{aligned} \text{Minimiere} \quad & \tfrac{1}{2} x^\top Q\,x + b^\top x, \quad Q = Q^\top \geq 0 \\ \text{sodass} \quad & A_1\,x \le b_1 \\ \text{und} \quad & A_2\,x = b_2 \end{aligned}$$

Beispiele konvexer Optimierungsaufgaben

Projektionsaufgabe

Minimiere	$f(x) := \|x - y\|$	(konvex)
sodass	$x \in C$	konvex

Allgemeine nichtlineare Optimierungsaufgaben (NLP)

Minimiere	$f(x)$	konvex
sodass	$g(x) \leq 0$	konvex
und	$h(x) = 0$	affin-linear

Beachte: Hierbei kann man sich geschickt und ungeschickt anstellen.

Konvexe und nichtkonvexe Beschreibungen

Beispiel

$$C = \{x \in \mathbb{R}^2 : x_1 \le 0, \quad x_1 + x_2 = 0\} \quad \text{(konvex)}$$

Nichtkonvexe Beschreibung

$$g(x) = \frac{x_1}{1 + x_2^2} \le 0$$

$$h(x) = (x_1 + x_2)^2 = 0$$

Konvexe Beschreibung

$$g(x) = x_1 \le 0$$

$$h(x) = x_1 + x_2 = 0$$

Charakterisierung konvexer Funktionen

Theorem

f (stetig diffbar) ist konvex auf C

$$\Leftrightarrow \quad f(x) \geq f(y) + \nabla f(y)^\top (x - y) \quad \text{für alle } x, y \in C$$

$$\Leftrightarrow \quad \langle \nabla f(x) - \nabla f(y), \quad x - y \rangle \geq 0 \quad \text{für alle } x, y \in C$$

Das heißt:

- *f verläuft oberhalb aller Tangenten*
- *der Gradient $\nabla f(\cdot)$ ist ein monotoner Operator*

Gliederung

1. Konvexe Optimierung
   - Bedeutung
   - Beispiele
   - Charakterisierung konvexer Funktionen
2. Optimalitätsbedingungen
   - KKT-Bedingungen
   - Constraint Qualifications
   - KKT-Bedingungen bei konvexen Aufgaben
   - Hinreichende Bedingungen zweiter Ordnung
3. Unendlich-dimensionale Optimierung
   - KKT-Bedingungen
   - Constraint Qualifications
   - Hinreichende Bedingungen zweiter Ordnung

Nichtlineare Optimierung

Nichtlineares Programm (NLP)

$$
\begin{array}{rll}
\text{Minimiere} & f(x) & \\
\text{sodass} & g(x) \leq 0 & g : \mathbb{R}^n \to \mathbb{R}^m \\
\text{und} & h(x) = 0 & h : \mathbb{R}^n \to \mathbb{R}^p
\end{array}
$$

(alles stetig diffbar, nicht notwendig konvex)

Ziel: notwendige Optimalitätsbedingungen für lokale Minima

KKT-Bedingungen

Lagrangefunktion

$$\mathcal{L}(x,\mu,\lambda) = f(x) + \mu^\top g(x) + \lambda^\top h(x)$$

Nebenbedingungen „angekoppelt“

KKT-Bedingungen

$$\mathcal{L}_x(x,\mu,\lambda) = \nabla f(x) + g'(x)^\top \mu + h'(x)^\top \lambda = 0$$

$$h(x) = 0$$

$$\mu \ge 0, \quad g(x) \le 0, \quad \mu^\top g(x) = 0 \quad \text{(Komplementarität)}$$

Ein Tripel (x,μ,λ), das diese Bedingungen erfüllt, heißt **KKT-Punkt**.

Motivation für KKT-Bedingungen

Einfacher Fall: $h : \mathbb{R}^n \longrightarrow \mathbb{R}$, keine Ungleichungen

$$\nabla f(x) + \nabla h(x)^\top \lambda = 0 \quad \Leftrightarrow \quad \text{Parallelität}$$

Motivation für KKT-Bedingungen

Einfacher Fall: $h : \mathbb{R}^n \longrightarrow \mathbb{R}$, keine Ungleichungen

$$\nabla f(x) + \nabla h(x)^\top \lambda = 0 \quad \Leftrightarrow \quad \text{Parallelität}$$

Theorem (Notwendige Bedingungen)

Ist x^* ein lokales Minimum von (NLP) und sind gewisse Regularitätsbedingungen (Constraint Qualifications) erfüllt, dann existieren (nicht notwendig eindeutig bestimmte) Lagrange-Multiplikatoren μ^*, λ^*, sodass (x^*, μ^*, λ^*) ein KKT-Punkt ist.

Vorteil: System von Gleichungen und Ungleichungen

Brauchen wir CQs?

Beispiel

$$\begin{aligned} \text{Minimiere} \quad & -x_1 \\ \text{sodass} \quad & g_1(x) = x_2 + x_1^3 \le 0 \\ \text{und} \quad & g_2(x) = -x_2 \le 0. \end{aligned}$$

$x^* = (0,0)^\top$ ist eindeutiges (sogar globales) Minimum.

$$\nabla g_1(x) = (3\,x_1^2, 1)^\top \qquad \nabla g_1(x^*) = (0,1)^\top$$

$$\nabla g_2(x) = (0,-1)^\top \qquad \nabla g_1(x^*) = (0,-1)^\top$$

$$\nabla f(x) = (-1,0)^\top \qquad \nabla f(x^*) = (-1,0)^\top$$

Brauchen wir CQs?

Folgerungen

- Offenbar lässt sich $-\nabla f(x^*)$ nicht als $\mu_1 \nabla g_1(x^*) + \mu_2 \nabla g_2(x^*)$ schreiben.
- Lagrange-Multiplikatoren existieren nicht.
- Über KKT-Bedingungen lässt sich diese Aufgabe nicht lösen.

Übersicht Constraint Qualifications

Zusammenhang

allgemeine NLP	konvexe NLP
LICQ	
$\Downarrow$	
MFCQ	Slater-Bedingung
$\Downarrow$	$\Downarrow$
Abadie-CQ	Abadie-CQ

Bedeutung der Abadie-CQ

Tangentialkegel und Linearisierung

$$\mathcal{T}_X(x^*) := \{d \in \mathbb{R}^n : \exists \{x^k\} \subset X, \quad t_k \searrow 0, \quad \frac{x^k - x^*}{t_k} \to d\}$$

$$\mathcal{T}_{\text{lin}}(x^*) := \left\{ d \in \mathbb{R}^n : \begin{Bmatrix} \nabla g_i(x^*)^\top d \le 0 & \text{für alle } i \in \mathcal{A}(x^*) \\ \nabla h_j(x^*)^\top d = 0 & \text{für alle } j = 1, \ldots p \end{Bmatrix} \right\}$$

$d \in \mathcal{T}_{lin}(x^*)$ ist tangential zu den Gleichungen, nicht nach außen zeigend bzgl. der aktiven Ungleichungen.

In einem lokalen Minimum x^* gilt:

$$\nabla f(x^*)^\top d \ge 0 \quad \text{für alle } d \in \mathcal{T}_X(x^*)$$

$$\Rightarrow \quad -\nabla f(x^*) \in \mathcal{T}_X(x^*)^\circ \quad \text{Normalenkegel}$$

Bedeutung der Abadie-CQ

Tangentialkegel und Linearisierung

$$\mathcal{T}_X(x^*) := \{d \in \mathbb{R}^n : \exists \{x^k\} \subset X, \quad t_k \searrow 0, \quad \frac{x^k - x^*}{t_k} \to d\}$$

$$\mathcal{T}_{\text{lin}}(x^*) := \left\{ d \in \mathbb{R}^n : \begin{Bmatrix} \nabla g_i(x^*)^\top d \le 0 & \text{für alle } i \in \mathcal{A}(x^*) \\ \nabla h_j(x^*)^\top d = 0 & \text{für alle } j = 1, \ldots p \end{Bmatrix} \right\}$$

$d \in \mathcal{T}_{lin}(x^*)$ ist tangential zu den Gleichungen, nicht nach außen zeigend bzgl. der aktiven Ungleichungen.

In einem lokalen Minimum x^* gilt:

$$\nabla f(x^*)^\top d \ge 0 \quad \text{für alle } d \in \mathcal{T}_X(x^*)$$

$$\Rightarrow \quad -\nabla f(x^*) \in \mathcal{T}_X(x^*)^\circ \quad \text{Normalenkegel}$$

Problem: Den Normalenkegel $\mathcal{T}_X(x^*)^\circ$ kann man nur schwer charakterisieren.

Bedeutung der Abadie-CQ

Abadie-CQ

Im Punkt x^* gelte $\mathcal{T}_X(x^*) = \mathcal{T}_{\text{lin}}(x^*)$.

$$\Rightarrow \quad -\nabla f(x^*) \in \mathcal{T}_X(x^*)^\circ = \mathcal{T}_{\text{lin}}(x^*)^\circ$$

Bedeutung der Abadie-CQ

Abadie-CQ

Im Punkt x^* gelte $\mathcal{T}_X(x^*) = \mathcal{T}_{\text{lin}}(x^*)$.

$$\Rightarrow \quad -\nabla f(x^*) \in \mathcal{T}_X(x^*)^\circ = \mathcal{T}_{\text{lin}}(x^*)^\circ$$

Vorteil

$$\mathcal{T}_{\text{lin}}(x^*)^\circ = \Big\{ \sum_{i=1}^m \mu_i \nabla g_i(x^*) + \sum_{j=1}^p \lambda_j \nabla h_j(x^*) : \ \mu_i \ge 0 \text{ für } i \in \mathcal{A}(x^*),$$
$$\mu_i = 0 \text{ für } i \in \mathcal{I}(x^*), \ \lambda_j \in \mathbb{R} \Big\}.$$

Beachte: $-\nabla f(x^*) \in \mathcal{T}_{\text{lin}}(x^*)^\circ \quad \Leftrightarrow \quad$ KKT-Bedingungen.

Bedeutung der Abadie-CQ

Einschätzung

Im Allgemeinen sind die Abadie-CQ schwer zu handhaben.

Jedoch:

Die Abadie-CQ gelten, falls g und h affin-linear sind.

Notorisches Gegenbeispiel für die Abadie-CQ

MPCCs

$$\begin{aligned} \text{Minimiere} \quad & f(x) \\ \text{sodass} \quad & 0 \le g_1(x), \quad 0 \le g_2(x), \quad g_1(x)\, g_2(x) = 0. \end{aligned}$$

- Die Abadie-CQ ist in bi-aktiven Punkten nicht erfüllt.
- Die MFCQ ist in keinem zulässigen Punkt erfüllt.

Anwendungen

Optimierungsaufgaben, in denen ein Optimalitätssystem einer anderen Aufgabe (lower-level) als Nebenbedingung vorkommt.

Weitere CQ

MFCQ / Linearisierte Slaterbedingung

- Die Gradienten $\{\nabla h_j(x^*)\}_{j=1}^p$ seien linear unabhängig.
- Es existiere ein Vektor $d \in \mathbb{R}^n$ mit

$$\begin{aligned} \nabla g_i(x^*)^\top d &< 0 \quad \text{für alle } i \in \mathcal{A}(x^*) \\ \nabla h_j(x^*)^\top d &= 0 \quad \text{für alle } j = 1, \ldots, p. \end{aligned}$$

LICQ

Die Gradienten $\{\nabla h_j(x^*)\}_{j=1}^p \cup \{\nabla g_i(x^*)\}_{i \in \mathcal{A}(x^*)}$ sind linear unabhängig.

Slater-Bedingung für konvexe Aufgaben

Sei also $g(\cdot)$ konvex und $h(x) = A\,x - b$.

Slater-Bedingung

Es existiere ein x_0 (**Slater-Punkt**) mit

$$g(x_0) < 0 \quad \text{und} \quad A\,x_0 = b.$$

Slater-Bedingung für konvexe Aufgaben

Sei also $g(\cdot)$ konvex und $h(x) = A\,x - b$.

Slater-Bedingung

Es existiere ein x_0 (**Slater-Punkt**) mit

$$g(x_0) < 0 \quad \text{und} \quad A\,x_0 = b.$$

Vorteil: Man muss x^* nicht kennen.

KKT-Bedingungen für konvexe Aufgaben

Theorem (Hinreichende Bedingungen)

Ist das (NLP) konvex, dann ist jeder KKT-Punkt (x^*, μ^*, λ^*) *ein lokales = globales Minimum.*

KKT-Bedingungen für konvexe Aufgaben

Theorem (Hinreichende Bedingungen)

Ist das (NLP) konvex, dann ist jeder KKT-Punkt (x^*, μ^*, λ^*) ein lokales = globales Minimum.

Beachte: CQs kommen nicht vor.

KKT-Bedingungen für konvexe Aufgaben

Theorem (Hinreichende Bedingungen)

Ist das (NLP) konvex, dann ist jeder KKT-Punkt (x^*, μ^*, λ^*) ein lokales = globales Minimum.

Beachte: CQs kommen nicht vor.

Beweis.

$$
\begin{aligned}
f(x) &\ge f(x^*) + \nabla f(x^*)^\top (x - x^*) \\
&= f(x^*) - (\mu^*)^\top g'(x^*)(x - x^*) - (\lambda^*)^\top \underbrace{A\,(x - x^*)}_{=b-b=0} \\
&\ge f(x^*) - (\mu^*)^\top \big(g_i(x) - g_i(x^*)\big) \\
&\ge f(x^*)
\end{aligned}
$$

$\square$

Hinreichende Bedingungen 2. Ordnung (SSC)

Theorem

Es sei (x^*, μ^*, λ^*) ein KKT-Punkt. Wenn

$$d^\top \mathcal{L}_{xx}(x^*, \mu^*, \lambda^*)\, d > 0 \quad \textit{für alle } d \in \mathcal{T}_{\text{lin}}(x^*) \setminus \{0\}$$

gilt, dann ist x^* ein striktes lokales Minimum von (NLP).

Hinreichende Bedingungen 2. Ordnung (SSC)

Theorem

Es sei (x^*, μ^*, λ^*) *ein KKT-Punkt. Wenn*

$$d^\top \mathcal{L}_{xx}(x^*, \mu^*, \lambda^*)\, d > 0 \quad \textit{für alle } d \in \mathcal{T}_{\text{lin}}(x^*) \setminus \{0\}$$

gilt, dann ist x^* *ein striktes lokales Minimum von (NLP).*

Beachte: (NLP) verhält sich dann in der Nähe von x^* wie ein strikt konvexes Problem.

Hinreichende Bedingungen 2. Ordnung (SSC)

Theorem

Es sei (x^*, μ^*, λ^*) *ein KKT-Punkt. Wenn*

$$d^\top \mathcal{L}_{xx}(x^*, \mu^*, \lambda^*)\, d > 0 \quad \textit{für alle } d \in \mathcal{T}_{\text{lin}}(x^*) \setminus \{0\}$$

gilt, dann ist x^* *ein striktes lokales Minimum von (NLP).*

Beachte: (NLP) verhält sich dann in der Nähe von x^* wie ein strikt konvexes Problem.

Verfeinerung

Diese Aussage gilt auch mit den kleineren Kegel $\mathcal{T}_2$, bei dem man stark aktive Ungleichungen wie Gleichungen behandelt.

Gliederung

1. Konvexe Optimierung
   - Bedeutung
   - Beispiele
   - Charakterisierung konvexer Funktionen
2. Optimalitätsbedingungen
   - KKT-Bedingungen
   - Constraint Qualifications
   - KKT-Bedingungen bei konvexen Aufgaben
   - Hinreichende Bedingungen zweiter Ordnung
3. Unendlich-dimensionale Optimierung
   - KKT-Bedingungen
   - Constraint Qualifications
   - Hinreichende Bedingungen zweiter Ordnung

Aufgabenstellung

Nichtlineares Programm (NLP)

$$\begin{array}{rlll} \text{Minimiere} & f(x) & & \\ \text{sodass} & g(x) \le 0 & g : X \to Z_g & \\ \text{und} & h(x) = 0 & h : X \to Z_h & \text{(PDE)} \end{array}$$

X, Z_g, Z_h Banachräume

Aufgabenstellung

Nichtlineares Programm (NLP)

$$\text{Minimiere} \quad f(x)$$

$$\text{und} \quad h(x) = 0 \qquad\qquad h : X \to Z_h \quad \text{(PDE)}$$

X, Z_g, Z_h Banachräume
Zunächst: nur Gleichungsbeschränkungen

Aufgabenstellung

Nichtlineares Programm (NLP)

Minimiere $\quad f(x)$

und $\quad h(x) = 0 \qquad\qquad h : X \to Z_h \quad$ (PDE)

X, Z_g, Z_h Banachräume

Zunächst: nur Gleichungsbeschränkungen

Lagrangefunktion

$$\mathcal{L}(x, \lambda) = f(x) + \langle \lambda, h(x) \rangle$$

Nebenbedingungen „angekoppelt“mit $\lambda \in Z_h^*$

Notwendige Bedingungen

KKT-Bedingungen ohne Ungleichungsbeschränkungen

$$\mathcal{L}_x(x,\lambda)\,\delta x = f'(x)\,\delta x + \langle \lambda, h'(x)\,\delta x\rangle = 0 \quad \text{für alle } \delta x \in X$$

$$h(x) = 0$$

Notwendige Bedingungen

KKT-Bedingungen ohne Ungleichungsbeschränkungen

$$\begin{aligned} & \mathcal{L}_x(x,\lambda)\,\delta x = f'(x)\,\delta x + \langle \lambda, h'(x)\,\delta x\rangle = 0 \quad \text{für alle } \delta x \in X \\ \Leftrightarrow \quad & f'(x) + h'(x)^\star \lambda = 0 \quad \text{in } X^* \\ & h(x) = 0 \end{aligned}$$

Notwendige Bedingungen

KKT-Bedingungen ohne Ungleichungsbeschränkungen

$$\begin{aligned} & \mathcal{L}_x(x,\lambda)\,\delta x = f'(x)\,\delta x + \langle \lambda, h'(x)\,\delta x\rangle = 0 \quad \text{für alle } \delta x \in X \\ \Leftrightarrow \quad & f'(x) + h'(x)^\star \lambda = 0 \quad \text{in } X^* \\ & h(x) = 0 \end{aligned}$$

Theorem (Notwendige Bedingungen)

Ist x^* ein lokales Minimum von (NLP) und sind gewisse Regularitätsbedingungen (Constraint Qualifications) erfüllt, dann existieren (nicht notwendig eindeutig bestimmte) Lagrange-Multiplikatoren $\lambda^* \in Z_h^*$, sodass (x^*, λ^*) ein KKT-Punkt ist.

Constraint Qualifications

LICQ

Die lineare Abbildung $h'(x^*)$ ist surjektiv auf Z_h.

Bedeutung für PDEs

Die linearisierte PDE

$$h'(x^*)\,\delta x = f$$

ist für jede rechte Seite $f \in Z_h$ lösbar in X.

Ungleichungsbeschränkungen

Frage: Was soll $g(x) \le 0$ für $x \in X$ bedeuten?

Eigentlicher (proper) Kegel

Ein konvexer Kegel $K \subset Z$ heißt **eigentlich**, wenn gilt:

- $K \cap (-K) = \{0\}$
- K ist abgeschlossen
- K hat nichtleeres Inneres

Bedeutung eigentlicher Kegel

Ein eigentlicher Kegel induziert mittels

$$\begin{aligned} x \le_K y \quad &\Leftrightarrow \quad y - x \in K \\ x <_K y \quad &\Leftrightarrow \quad y - x \in \operatorname{int}(K) \end{aligned}$$

eine [strikte] partielle Ordnung, die verträglich ist mit der Addition und mit Limiten in Z.

Lagrange-Multikplikatoren für Ungleichungen

Ungleichungsbeschränkung

$$\begin{aligned} & g(x) \le_K 0, \quad K \subset Z_g \text{ eigentlicher Kegel} \\ \Leftrightarrow \quad & -g(x) \in K \end{aligned}$$

Lagrange-Multiplikatoren leben im **Dualkegel**:

Dualkegel

$$K^* = \{\mu \in Z_g^* : \langle \mu, k \rangle \ge 0 \quad \text{für alle } k \in K\}$$

Wenn K eigentlicher Kegel ist, dann auch K^* ($\rightsquigarrow$ Ordnung auf Z_h^*).

Lagrange-Funktion und KKT-Punkt

Lagrangefunktion

$$\mathcal{L}(x,\mu,\lambda) = f(x) + \langle \mu, g(x)\rangle + \langle \lambda, h(x)\rangle$$

KKT-Bedingungen

$$\mathcal{L}_x(x,\mu,\lambda) = f'(x) + g'(x)^\star\mu + h'(x)^\star\lambda = 0$$

$$h(x) = 0$$

$$\mu \geq_{K^*} 0, \quad g(x) \leq_K 0, \quad \langle \mu, g(x)\rangle = 0 \quad \text{(Komplementarität)}$$

Beachte: Die Regularität der Multiplikatoren ist bei Optimalsteueraufgaben oft besser als erwartet ($\mu \in Z_g^*$, $\lambda \in Z_h^*$).

Slater-Bedingung für konvexe Aufgaben

Sei also $g(\cdot)$ konvex und $h(x) = A\,x - b$.

Slater-Bedingung

Es existiere ein x_0 (**Slater-Punkt**) mit

$$g(x_0) <_K 0 \quad \text{und} \quad A\,x_0 = b.$$

Slater-Bedingung für konvexe Aufgaben

Sei also $g(\cdot)$ konvex und $h(x) = A\,x - b$.

Slater-Bedingung

Es existiere ein x_0 (**Slater-Punkt**) mit

$$g(x_0) <_K 0 \quad \text{und} \quad A\,x_0 = b.$$

Beachte: $-g(x_0) \in \operatorname{int}(K)$

Slater-Bedingung für konvexe Aufgaben

Sei also $g(\cdot)$ konvex und $h(x) = A\,x - b$.

Slater-Bedingung

Es existiere ein x_0 (**Slater-Punkt**) mit

$$g(x_0) <_K 0 \quad \text{und} \quad A\,x_0 = b.$$

Beachte: $-g(x_0) \in \operatorname{int}(K)$

Beispiele konvexer Aufgaben

Optimalsteuerungsaufgaben mit

- quadratischem Zielfunktional
- linearer PDE (Gleichungsnebenbedingung)
- linearen punktweisen Ungleichungen (z.B. Box-Beschränkungen)

CQ für allgemeine Aufgaben

Allgemeine Theorie: [Zowe, Kurcyusz (1979)]

MFCQ / Linearisierte Slater-Bedingung

- $h'(x^*)$ sei surjektiv
- Es existiere $d \in X$ mit

$$\begin{cases} g(x^*) + g'(x^*)\,d <_K 0 \\ \underbrace{h(x^*)}_{=0} + h'(x^*)\,d = 0 \end{cases}$$

Vergleiche MFCQ im Endlich-Dimensionalen

- Die Gradienten $\{\nabla h_j(x^*)\}_{j=1}^p$ seien linear unabhängig.
- Es existiere ein Vektor $d \in \mathbb{R}^n$ mit

$$\begin{aligned} \nabla g_i(x^*)^\top d &< 0 \quad \text{für alle } i \in \mathcal{A}(x^*) \\ \nabla h_j(x^*)^\top d &= 0 \quad \text{für alle } j = 1, \ldots, p. \end{aligned}$$

Erweiterungen

Konvexe Grundmenge

Die Theorie kann erweitert werden auf

$$\begin{aligned} \text{Minimiere} \quad & f(x) \quad \text{über } x \in C, \quad C \text{ konvex} \\ \text{sodass} \quad & g(x) \le 0 \\ \text{und} \quad & h(x) = 0 \end{aligned}$$

- Die $\mathcal{L}_x(x^*, \mu^*, \lambda^*) = 0$ wird dann ersetzt durch

$$\mathcal{L}_x(x^*, \mu^*, \lambda^*)(x - x^*) \ge 0 \quad \text{für alle } x \in C.$$

- Man bekommt keine Lagrange-Multiplikatoren für $x \in C$.

Erweiterungen

Gleichungen und Ungleichungen

Die Gleichungen $h(x) = 0$ können als verallgemeinerte Ungleichungen

$$g(x) \leq_K 0$$

geschrieben werden mit $K = \{0\}$. Dann ist

$$\mu \in K^* = Z_g^* \quad \text{der ganze Raum (keine Vorzeichenbedingung).}$$

CHEMNITZ UNIVERSITY OF TECHNOLOGY

Hinreichende Bedingungen 2. Ordnung (SSC)

Wir betrachten zunächst die Aufgabe

$$\text{Minimiere} \quad f(x), \quad x \in C \subset X, \quad C \text{ konvex.}$$

Notwendige Bedingung 1. Ordnung

$$f'(x^*)(x - x^*) \geq 0 \quad \text{für alle } x \in C.$$

Theorem (Hinreichende Bedingung)

x^* erfülle die notwendige Bedingung 1. Ordnung. Es existiere $\delta > 0$ mit

$$f''(x^*)(x, x) \geq \delta \, \|x\|^2 \quad \text{für alle } x \in X.$$

Dann ist x^* ein strikt lokales Minimum, und es gilt

$$f(x^* + h) \geq f(x^*) + \frac{\delta}{4} \|h\|^2 \quad \text{für } \|h\| \text{ hinreichend klein.}$$

(Gegen-)Beispiel (Tröltzsch, Kapitel 4.9)

Aufgabenstellung

$$\text{Minimiere} \quad f(u) := -\int_0^1 \cos(u(x))\,dx$$
$$\text{sodass} \quad u \in C = \{u \in L^2(0,1) : 0 \le u(x) \le 2\pi \text{ f.ü.}\}$$

Offenbar ist $u^* \equiv 0$ eine globale Lösung der Aufgabe.

Die notwendige Bedingung 1. Ordnung

$$f'(u^*)(u - u^*) = \int_0^1 \sin(u^*(x))(u(x) - u^*(x))\,dx = \int_0^1 \sin(0)\,u(x)\,dx = 0$$

für alle $u \in C$ ist erfüllt.

(Gegen-)Beispiel (Tröltzsch, Kapitel 4.9)

$$f''(u^*)(u,u) = \int_0^1 \cos(0)\, u^2(x)\, dx = 1 \cdot \|u\|^2_{L^2(0,1)}$$

für alle $u \in L^2(0,1)$. Also ist die hinreichende Bedingung 2. Ordnung erfüllt, und $u^* \equiv 0$ ist ein strikt lokales Minimum.

Aber:

Auch die Funktion

$$u_\varepsilon(x) = \begin{cases} 2\pi, & 0 \le x \le \varepsilon \\ 0, & \varepsilon < x \le 1 \end{cases}$$

ist für jedes $\varepsilon > 0$ ein globales Minimum, und es gilt

$$\|u^* - u_\varepsilon\|^2_{L^2(0,1)} = 2\pi\sqrt{\varepsilon}.$$

(Gegen-)Beispiel (Tröltzsch, Kapitel 4.9)

Fehler

Die Funktion

$$f(u) := -\int_0^1 \cos(u(x))\,dx$$

ist im Raum $L^2(0,1)$ nicht zweimal stetig diffbar, jedoch in $L^\infty(0,1)$.

(Gegen-)Beispiel (Tröltzsch, Kapitel 4.9)

Fehler

Die Funktion

$$f(u) := -\int_0^1 \cos(u(x))\,dx$$

ist im Raum $L^2(0,1)$ nicht zweimal stetig diffbar, jedoch in $L^\infty(0,1)$.

Aber:

In $L^\infty(0,1)$ bekommen wir nicht die benötigte Abschätzung

$$f''(u^*)(u,u) \geq \delta \, \|u\|^2_{L^\infty(0,1)} \quad \text{für alle } u \in L^\infty(0,1).$$

Zwei-Norm-Diskrepanz

Ausweg

Arbeiten mit zwei Normen

Im Beispiel:

$$f(u+h) - f(u) = \text{Taylorpolynom 2. Ordnung} + r_2(u,h)$$

mit der Restgliedabschätzung

$$\frac{|r_2(u,h)|}{\|h\|^2_{L^2(0,1)}} \to 0 \quad \text{für } \|h\|_{L^\infty(0,1)} \to 0.$$

Es folgt:

$$f(u^*+h) \ge f(u^*) + \delta' \, \|h\|^2_{L^2(0,1)}, \quad \text{falls } \|h\|_{L^\infty(0,1)} \le \varepsilon.$$

Beispiel: (Tröltzsch, Satz 4.27)

Aufgabe

$$\text{Minimiere} \quad \int_\Omega \varphi(x,y)\,dx + \int_\Omega \psi(x,u)\,dx$$

$$\text{unter} \quad \begin{cases} -\triangle y + y + d(x,y) = u & \text{in } \Omega \\ \dfrac{\partial y}{\partial n} = 0 & \text{auf } \Gamma \end{cases}$$

$$\text{und} \quad u_a(x) \le u(x) \le u_b(x) \quad \text{f.ü. in } \Omega$$

Voraussetzungen: siehe Tröltzsch, Vor. 4.12

Notwendige Bedingung 1. Ordnung

$$\begin{aligned} -\triangle p^* + p^* + d_y(x,y^*)\,p^* &= \varphi_y(x,y^*) && \text{in } \Omega \\ \frac{\partial p^*}{\partial n} &= 0 && \text{auf } \Gamma \end{aligned}$$

$$\int_\Omega \big[p^* + \psi_u(x,u^*)\big](u-u^*)\,dx \ge 0 \quad \text{für alle } u \in U_{\text{ad}}$$

Beispiel: (Tröltzsch, Satz 4.27)

Theorem (Hinreichende Bedingung 2. Ordnung)

Falls (y^*, u^*, p^*) die notwendige Bedingung 1. Ordnung und

$$\mathcal{L}_{xx}(y^*, u^*, p^*)[(y, u), (y, u)] \geq \delta \, \|u\|^2_{L^2(\Omega)}$$

für alle $u \in C(u^*)$ mit zugehörigem linearisertem Zustand y

$$\begin{aligned} -\triangle y + y + d_y(x, y^*) \, y &= u \quad \text{in } \Omega \\ \frac{\partial y}{\partial n} &= 0 \quad \text{auf } \Gamma \end{aligned}$$

erfüllt, dann gilt die quadratische Wachstumsbedingung

$$J(y, u) \geq J(y^*, u^*) + \delta' \, \|u - u^*\|^2_{L^2(\Omega)}$$

für alle $u \in U_{\text{ad}}$ mit $\|u - u^*\|_{L^\infty(\Omega)} \leq \varepsilon$.

Beispiel: (Tröltzsch, Satz 4.27)

Kegel zulässiger Richtungen

$$C(u^*) = \left\{ u \in L^\infty(\Omega) : \left\{ \begin{array}{ll} u(x) \geq 0 & \text{wo } u^* = u_a \\ u(x) \leq 0 & \text{wo } u^* = u_b \end{array} \right\} \right\}$$

Dieser kann noch verkleinert werden, indem man stark aktive Ungleichungen berücksichtigt, vgl. $\mathcal{T}_{\text{lin}}(x^*)$ und $\mathcal{T}_2(x^*)$.

Zusammenfassung

- Eigenschaften konvexer Optimierungsaufgaben
  - lokale = globale Minima
  - notwendige Bedingungen sind hinreichend
- allgemeines (NLP)
  - Constraint Qualifications und KKT-Bedingungen
  - hinreichende Bedingungen 2. Ordnung
- unendlich-dimensionale Optimierung
  - Constraint Qualifications und KKT-Bedingungen
  - hinreichende Bedingungen 2. Ordnung: Zwei-Norm-Diskrepanz

Referenzen

S. Boyd and L. Vandenberghe.
Convex Optimization.
Cambridge University Press, Cambridge, 2004.

J. Zowe and S. Kurcyusz.
Regularity and stability for the mathematical programming problem in Banach spaces.
Applied Mathematics and Optimization, 5(1):49–62, 1979.

H. Maurer and J. Zowe.
First and second order necessary and sufficient optimality conditions for infinite-dimensional programming problems.
Mathematical Programming, 16:98–110, 1979.

F. Tröltzsch.
Optimale Steuerung partieller Differentialgleichungen.
Vieweg, Wiesbaden, 2005.