Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Hamburg University

Optimization

Notes

Department of Mathematics, Hamburg University, Bundesstrasse 55 , 20146, Hamburg, Germany

Abstract:

Keywords: Optimization• Convexity

Introduction

0.1. Definitions

0.2. Useful lemmas and Theorems.

Lemma 0.1.
Let $(x_n)_{n\in\mathbb{N}}$ be a bounded sequence in a Hilbert Space H. Then $(x_n)_{n\in\mathbb{N}}$ possesses a weakly convergent subsequence.

Lemma 0.2.
Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in a Hilbert Space H. Then $(x_n)_{n\in\mathbb{N}}$ converges if and only if it is bounded and possesses at most one weak sequential cluster point.

Theorem 0.1.
Let $f : H \to (\infty, \infty]$ be a convex functional on a Hilbert space. Then the following are equivalent:

(i) f is weakly sequentially lower semicontinuous.

(ii) f is sequentially lower semicontinuous.

(iii) f is lower semicontinuous.

(iv) f is weakly lower semicontinuous.

Lemma 0.3.
A convex set is closed if and only if it is weakly closed.

1. Lecture 1

2. Lecture 2

2.1. Convexity

Definition 2.1.
Let U be linear space. A functional $J : U \to \overline{\mathbb{R}}$ is called convex, if for $t \in [0, 1]$ and $u_1, u_2 \in U$.

$$J(tu_1 + (1-t)u_2) \leq tJ(u_1) + (1-t)J(u_2) \tag{1}$$

holds such that the right hand sid is well defined.

- J is strictly convex if (1) holds strictly for $\forall u_1, u_2 \in U$, $u_1 \neq u_2$ and $t \in (0, 1)$ with $J(u_1) < \infty$ and $J(u_2) < \infty$.
- An optimization problem,

$$\min_{u \in C} J(u)$$

  is called convex if both C and J are convex.

Lemma 2.1.
If C and V are convex in U, then

- $\alpha V = \{w = \alpha v, v \in C\}$ *is convex.*
- $C + V$ *is convex.*

Proof. □

Lemma 2.2.
Let V be a collection of convex sets in U, then $C = \bigcap_{K \in V} K$ is convex.

Proof. If $C = \emptyset$, then C the statement is vacuously true. Consider $C \neq \emptyset$ and $u_1, u_2 \in C$ then $u_1, u_2 \in K$ for all $K \in V$

$$\implies tu_1 + (1-t)u_2 \in K, \quad \forall K \in V \implies tu_1 + (1-t)u_2 \in \bigcap_{K \in V} K$$

□

Lemma 2.3.
Let $C \in U$ convex and $J : C \to \mathbb{R}$. Define $\alpha = \inf_{u \in C} J(u)$. Then the set $\Psi = u | J(u) = \alpha$ is convex, i.e. the solution of

$$\min_{u \in C} J(u)$$

is a convex set.

Proof. Let $u_1, u_2 \in \Psi$ and $u_t = tu_1 + (1-t)u_2$.Sinc J is convex, it holds that $J(u_t) \leq tJ(u_t) + (1-t)J(u_t) = \alpha$. Thus $J(u_t) = \alpha$, $\forall t \in [0, 1]$. Implying $u_t \in \Psi$ Hence Ψ is convex. □

Lemma 2.4.
Let U be linear normed space, and $C \subset U$ a convex set and $J : U \to \overline{\mathbb{R}}$ convex functional. Let $\overline{u} \in C$ such that

$$J(\overline{u}) \leq J(u) \quad \forall u \in B_\epsilon(\overline{u}) \cap C,$$

for some ball $B_\epsilon(\overline{u})$ in U with center in $\overline{u}$. Then $J(\overline{u}) = \inf_{u \in C} J(u)$. In other words, the local minimizer of a convex optimization problem is also a global minimizer.

Proof. Let $B_\epsilon(\overline{u})$ be an open neighborhood of $\overline{u}$ with $J(\overline{u}) \leq J(u)$ for all $u \in B_\epsilon(\overline{u}) \cap C$. Take an arbitrary $u^* \in C$ and consider $u_t = t\overline{u} + (1-t)u^*$. Since C is convex $u_t \in C$.
For some $t \in (0,1)$, $u_t \in B_\epsilon(\overline{u})$.
Thus,

$$J(\overline{u}) \leq J(u_t) \leq tJ(\overline{u}) + (1-t)J(u^*).$$

We have $\forall t \in [0,1]$ that $(1-t) \leq 0$, then

$$(1-t)J(\overline{u}) \leq (1-t)J(u^*) \qquad \forall u^* \in C$$

Therefore, $\overline{u}$ is a local minimizer for C. □

Theorem 2.1.

Let U is Banach Space, $C \subset U$ convex and $J : C \to \mathbb{R}$ Gateâux differentiable. Consider the minimization problem.

$$\min_{u \in C} J(u)$$

1. *Let $\overline{u}$ be a local solution. Then $J'(\overline{u}, u - \overline{u}) \geq 0$, $\forall u \in C$.*
2. *If J is convex on C, then $J'(\overline{u}, u - \overline{u}) \geq 0$, $\forall u \in C$ is necessary and sufficient for global optimality of $\overline{u}$*
3. *If J is strictly convex on C, then the minimization problem admits at most one solution.*
4. *If C is closed, and J is convex and continuous with*

$$\lim_{\substack{u \in C \\ \|u\| \to \infty}} J(u) = \infty.$$

   Then a global solution $\overline{u} \in C$ exists.

Proof.

1. Let $\overline{u}$ be a local solution $J(\overline{u}) \leq J(u)$, $\forall u \in B_\epsilon(\overline{u}) \cap C$, let $t \in [0,1]$, $u_t = \overline{u} + t(u - \overline{u})$, then $u_t \in C$, since C is convex.

   For small $t > 0$,

$$0 \leq \frac{1}{t}\left[J(u_t) - J(u)\right] \leq \frac{1}{t}\left[J(\overline{u} + t(u - \overline{u})) - J(u)\right] \xrightarrow{t \downarrow 0} J'(\overline{u}, u - \overline{u})$$

2. Since J is convex we have for $u \in C$, $J(\overline{u} + t(u - \overline{u})) \leq J(\overline{u}) + t\left[J(u) - J(\overline{u})\right]$, for $t > 0$

$$\implies J(u) - J(\overline{u}) \geq \frac{1}{t}\left[J(\overline{u} + t(u - \overline{u})) - J(\overline{u})\right] \xrightarrow{t \downarrow 0} J'(\overline{u}, u - \overline{u}) \geq 0.$$

   Therefore $\overline{u}$ is a global minimizer.

3. Assume, that there are two solution for the minimization problem, $\overline{u}, u^* \in C$, such that $\overline{u} \neq u^*$ and $J(\overline{u}) = J(u^*) = \inf_{u \in C} J(u)$. Since J is strictly convex $J(u_t) = J(t\overline{u} + (1-t)u^*) < tJ(\overline{u}) + (1-t)J(u^*) = \alpha$ for all $t \in [0,1]$. Contradicting our assumption that u^* and $\overline{u}$ are solutions.

4. $\alpha = \inf_{u \in C} J(u) \in \mathbb{R} \cup \{-\infty\}$, choose a minimizing sequence $(u_k)_k \subset C$ with $J(u_k) \xrightarrow{k \to \infty} \alpha$

   $\implies (u_k)_k$ is bounded, because $J \to \infty$ as $\|u\| \to \infty$.
   $\implies (u_k)_k$ contains a weakly convergent subsequence $u_{k_e} \xrightarrow[e \to \infty]{} \overline{u} \in C$. Since C is closed and convex.
   $\implies J$ is weakly-lower semicontinuos because it is convex and continuos.

□

3. Lecture 3

Now consider Hilbert space $(H, < . >)$ with $\|x\| = \sqrt{(x,x)}$.

4. Lecture 4

5. Lecture 5

6. Lecture 6

7. Lecture 7

8. Lecture 8

Definition 8.1.
A function $\Phi : U \times Y \to \overline{\mathbb{R}}$ is said to be a perturbation function of J for the minimization problem

$$\inf_{u \in U} J(u) \tag{P}$$

if $\Phi(u, 0) = J(u)$ for all $u \in U$. For each $p \in Y$, the minimization problem (Pp)

$$\inf_{u \in U} \Phi(u, p) \tag{Pp}$$

is called a perturbation problem. The variable p is called perturbation parameter. If we denote by Φ^* the convex conjugate function of Φ, the *dual problem*, with respect to Φ is defined by

$$\sup_{p^* \in Y^*} -\Phi^*(0, p^*) \tag{P*}$$

where $\Phi^* : (U \times Y)^* \cong U^* \times Y^* \to \overline{\mathbb{R}}$, a function defined as follows.

$$\Phi^*(u^*, p^*) = \sup_{\substack{u \in U \\ p \in Y}} \left((u^*, u)_{U^*U} + (p^*, p)_{Y^*Y} - \Phi(u, p)\right)$$

Remark 8.1.
For $p = 0$, (P*) $\equiv$ (Pp). We denote the infimum for problem (P) by inf (P) and the supremum for problem (P*) by sup (P*)

Lemma 8.1 (Weak duality).
For the problem (P) and (P) it holds that*

$$-\infty \leq \sup\,(\text{P*}) \leq \inf\,(\text{P}) \leq \infty.$$

Proof. Let $p^* \in Y^*$. It follows

$$\begin{aligned}
-\Phi^*(0, p^*) &= -\sup_{\substack{u \in U \\ p \in Y}} \left((0, u) + (p^*, p) - \Phi(u, p)\right) && (2)\\
&= \inf_{\substack{u \in U \\ p \in Y}} \left(\Phi(u, p) - (p^*, p)\right) && (3)\\
&\leq \left(\Phi(u, 0) - (p^*, 0)\right) \quad \forall u \in U, p^* \in Y^* && (4)\\
\implies \sup_{p^* \in Y^*} \left(-\Phi\left(0, p^*\right)\right) &\leq \inf_{u \in U} \Phi(u, 0) = \inf(P) && (5)
\end{aligned}$$

$\square$

By iteration we can define, a bidual problem

$$-\sup_{u\in U}\left(-\Phi^*(u,0)\right)=\inf_{u\in U}\Phi^*(u,0) \tag{P**}$$

In case the space U is reflexive then $U^{**}=U$.

If the perturbation function $\Phi(u,p)$ is proper, convex and weakly lower semicontinuous. Then $\Phi^{**}=\Phi$. In this case $\Phi(u,0)=\Phi^{**}(u,0)$ i.e (P)≡ (P**)

Definition 8.2.

Consider the infimal value function

$$h(p^*)=\inf\,(\text{Pp})=\inf_{u\in U}\Phi(u,p)$$

The problem (P) is called stable if $h(0)$ is finite and its sub-differentiable in zero is not empty.

Theorem 8.1.

The primal problem (P) *is stable if and only if the following conditions are simultaneously satisfied:*

- *The dual problem* (P*) *has a solution.*
- *There is no duality gap, i.e.*

$$\inf\,(\text{P})=\sup\,(\text{P*})\leq\infty$$

Theorem 8.2 (Extremal relation).

Let $\Phi: U\times Y\to\overline{\mathbb{R}}$, be convex the the following statements are equivalent:

1. (P) *and* (Pp) *have solutions* $\overline{u}$ *and* $\overline{p^*}$ *and* $\inf(P)=\sup(P^*)$
2. $\Phi(\overline{u},0)+\Phi^*\left(0,\overline{p^*}\right)=0$
3. $\left(0,\overline{p^*}\right)\in\partial\Phi(u,0)$ *and* $(\overline{u},0)\in\partial\Phi^*(0,p^*)$

Proof. We proceed by parts:

1. (1) $\implies$ (2): $\overline{u}$ solution of $\inf$ (P) and $\overline{p^*}$ solution of $\sup$ (P*) and $\inf$ (P) $=$ $\sup$ (P*). This properties implies, $\Phi(\overline{u},0)=\inf\,(\text{P})=\sup\,(\text{P*})=-\Phi(0,\overline{p^*})\implies\Phi(\overline{u},0)+\Phi^*(0,\overline{p^*})=0$.
2. (2) $\implies$ (1): $-\Phi^*(0,\overline{p^*})=\sup\,(\text{P*})\leq\inf\,(\text{P})=\Phi(\overline{u},0)=-\Phi^*(0,\overline{p^*})\implies\sup\,(\text{P*})=\inf\,(\text{P})$
3. (2) $\iff$ (3): $\Phi(\overline{u},0)+\Phi^*(0,\overline{p^*})=0=(0,\overline{u})+(\overline{p^*},0)=\big((0,\overline{p^*}),(\overline{u},0)\big)\iff(0,\overline{p^*})\in\partial\Phi(\overline{u},0)\ \forall u\in U,\forall p^*\in\partial J(u)\iff J(u)+J^*(p^*)=(p^*,u)$

□

Fencel duality.

Consider the functional $J: U\to\overline{\mathbb{R}}$,

$$J(u)=F(u)+G(Au)$$

with $F: U\to\overline{\mathbb{R}}$, G convex function $G: V\to\overline{\mathbb{R}}$ and $A: U\to V$ bounded and linear.

We introduce the perturbation $\Phi(u,p)=F(u)+G(Au-p)$. The dual problem is obtained with,

$$\Phi^*(0,p^*)=\sup_{\substack{u\in U\\p\in V}}\left((p^*,p)-F(u)-G(Au-p)\right)$$

For fixed u we set $q : Au - p$.

$$
\begin{aligned}
\Phi^*(0, p^*) &= \sup_{u \in U} \sup_{p \in V} \left((p^*, Au - q) - F(u) - G(q) \right) \\
&= \sup_{u \in U} \sup_{p \in V} \left((A^* p^*, u) - (p^*, q) - F(u) - G(q) \right) \\
&= \sup_{u \in U} \left((p^*, Au) - F(u) \right) + \sup_{p \in V} \left((-p^*, q) - G(q) \right) \\
&= F^*(A^* p^*) + G^*(-p^*)
\end{aligned}
$$

References