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.

Lemma 0.4.

Every bounded linear operator over a Banach Space is weakly continuos.

Lemma 0.5 (Parallelogram law).

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2$$

1. Lecture 1

Fact 1.1.
Let $x \in H$, let U be a neighborhood of x, let G be a real Banach space, let $T : U \to G$, let V be a neighborhood of Tx, and let $R : V \to K$. Suppose that T is Frchet differentiable at x and that R is Gteaux differentiable at Tx. Then $R \circ T$ is Gateâux differentiable at x and $\mathsf{D}(R \circ T)(x) = (\mathsf{D}R(Tx)) \circ \mathsf{D}T(x)$. If R is Fréchet differentiable at x, then so is $R \circ T$.

Fact 1.2.
Let $x \in H$, let U be a neighborhood of x, let K be a real Banach space, and let $T : U \to K$. Suppose that T is twice Fréchet differentiable at x. Then $\forall (y, z) \in H \times H$, $(\mathsf{D}^2T(x)y)z = (\mathsf{D}^2T(x)z)y$.

Definition 1.1.
Let $x \in H$, let $C \in \mathcal{V}(x)$, and let $T : C \to K$. Then T is Fréchet differentiable at x if there exists an operator $\mathbf{D}T(x) \in B(H, K)$, called the Frchet derivative of T at x, such that

$$\lim_{0 \neq \|y\| \to 0} \frac{\|T(x+y) - Tx - \mathsf{D}T(x)y\|}{\|y\|} = 0$$

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 $J(u_t) \le tJ(u_t) + (1-t)J(u_t) = \alpha$. 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

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

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 $J(\overline{u}) \le J(u)$ 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,

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

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

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

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 $J'(\overline{u}; u - \overline{u}) \ge 0$, $\forall u \in C$.*
2. *If J is convex on C, then $J'(\overline{u}; u - \overline{u}) \ge 0$, $\forall u \in C$ 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 $J(\overline{u}) \le J(u)$, $\forall u \in B_\epsilon(\overline{u}) \cap C$, let $t \in [0,1]$, $u_t = \overline{u} + t(u - \overline{u})$, then $u_t \in C$, since C is convex.

   For small $t > 0$,

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

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

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

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.

$\square$

3. Lecture 3

Now consider Hilbert space $(H, \langle\cdot,\cdot\rangle)$ with the norm defined as $\|\cdot\| = \sqrt{\langle\cdot,\cdot\rangle}$.

Let be $J : H \to \mathbb{R}$ a functional over a Hilbert space H, we define the set,

$$\operatorname*{arg\,min}_{v \in C \subseteq H} J(x) := \{x \mid x \in C \wedge \forall v \in C:\ J(x) \leq J(v)\}.$$

By Riesz-Fréchet representation formula, exists a unique vector $\nabla J(x) \in H$ such that,

$$(\forall y \in H) \quad J'(x; y) = \langle y, \nabla J(x)\rangle$$

namely Gateâux gradient of J at x.

Lemma 3.1.
Let H Hilbert space and $C \subset H$ closed and convex. Define $P_C : H \to C$,

$$P_C(x) = \operatorname*{arg\,min}_{v \in C} [\|v - x\|].$$

Then,

1. *P_C is well defined, i.e. $\exists! u \in H$ such that $P_C(x) = \{u\}$.*
2. *$\forall u, v \in H$, we have $x = P_C(u) \iff x \in C$ and $\langle u - x, v - x\rangle \leq 0$.*
3. *$\|P_C(u) - P_C(\overline{u})\| \leq \|u - \overline{u}\| \quad \forall u, \overline{u} \in H$, i.e. The projection P_C is non expansive.*
4. *$\langle P_C(u) - P_C(\overline{u}), u - \overline{u}\rangle \leq 0, \quad \forall u, \overline{u} \in H$*
5. *Let be $t > 0$ a real number, then $\forall u \in C$, and $\forall v \in H$, $\phi(t) = \frac{1}{t}\,\|P_C\,(u + tv) - u\|$ is non-increasing.*

Proof. 1. First we prove existence, let be $(v_k)_k$ a minimizing sequence in C, such that

$$\|x - v_k\| \to \alpha = \inf_{v \in C} \|x - v\|,$$

By the parallelogram law,

$$\begin{aligned}
2\|v_j - x\|^2 + 2\|v_i - x\|^2 &= \|v_j - v_i\|^2 + \|v_j + v_i - 2x\|^2 \\
2\|v_j - x\|^2 + 2\|v_i - x\|^2 &= \|v_j - v_i\|^2 + 4\left\|\frac{v_j + v_i}{2} - x\right\|^2 \\
\Longrightarrow 2\|v_j - x\|^2 + 2\|v_i - x\|^2 - 4\left\|\frac{v_j + v_i}{2} - x\right\|^2 &= \|v_j - v_i\|^2
\end{aligned}$$

Since C is convex $\frac{v_i+v_j}{2} \in C$, then by definition of α,

$$0 \leq \alpha \leq \left\|\frac{v_j + v_i}{2} - x\right\|$$

Therefore the above equations become in the following inequality,

$$2\|v_j - x\|^2 + 2\|v_i - x\|^2 - 4\alpha^2 \geq \|v_j - v_i\|^2$$

Since $\|v_i - x\| \to \alpha$ and $\|v_j - x\| \to \alpha$, we have that $\|v_j - v_i\| \to 0$, therefore the series is Cauchy and then converges. Since C is closed the series converges to a point $v \in C$.

Second we prove uniqueness, we proceed by contradiction, take $v, v' \in C$ such $v \neq v'$, and both of them minimizing the distant with respect the point x, i.e.

$$\|x - v\| = \|x - v'\| = \alpha = \min_{u \in C} \|u - x\|$$

By the parallelogram law,

$$2\|x - v\|^2 + 2\|x - v'\|^2 = \|2x - v - v'\|^2 + \|v - v'\|^2$$

Since C is convex, $\left\|\frac{v+v'}{2} - x\right\| \geq \alpha$

$$\begin{aligned}
\|v - v'\|^2 &= 2\|x - v\|^2 + 2\|x - v'\|^2 - \|2x - v - v'\|^2 \\
\|v - v'\|^2 &= 2\|x - v\|^2 + 2\|x - v'\|^2 - 4\left\|x - \frac{v - v'}{2}\right\|^2 \\
\|v - v'\|^2 &= 2\alpha^2 + 2\alpha^2 - 4\left\|x - \frac{v - v'}{2}\right\|^2 \leq 0
\end{aligned}$$

Therefore $\|v - v'\| = 0$, and $v = v'$.

By the uniqueness and existence $\arg\min_{u \in C} [\|u - x\|]$ is not empty set and has only one element for each $x \in H$.

Thus, P_C is well defined. □

Theorem 3.1.

Let H be Hilbert space, $C \subset H$ closed and convex, $J : C \to \mathbb{R}$, Gateâux differentiable at the local solution $\overline{u}$ of $\min_{u \in C} J(u)$. Thus, $J'(\overline{u}; u - \overline{u}) \geq 0$, $\forall u \in C$ and it is equivalent to $\overline{u} = P_C(\overline{u} - \delta\nabla J(\overline{u}))$

Proof. Since every Hilbert Space is a Banach space, and C is closed and Convex subset of H, and $\overline{u}$ is a solution of minimization problem; we can apply 2.1.
Thus $J'(\overline{u}; u-\overline{u}) \geq 0 \iff \langle u-\overline{u}, \nabla J(\overline{u})\rangle \geq 0 \; \forall u \in C$.
For all $\delta > 0$, we multiply the Gateâux gradient $(-\delta)$ and we have,

$$\langle u-\overline{u}, -\delta \nabla J(\overline{u})\rangle \leq 0 \; \forall u \in C,$$

adding zero to the gradient, $\langle u-\overline{u}, \overline{u} - \delta \nabla J(\overline{u}) - \overline{u}\rangle \leq 0$. Then we set $w \in H$ as $w := \overline{u} - \delta \nabla J(\overline{u})$, and applying lemma 3.1 we have,

$$\overline{u} = P_c(w) \iff \langle w-\overline{u}, u-\overline{u}\rangle$$

Thus,

$$\overline{u} = P_C(\overline{u} - \delta J(\overline{u}))$$

□

3.1. Application

Consider U, Y, Z Hilbert spaces. Let be $J : Y \times U \to \mathbb{R}$ a functional. Consider the minimization problem,

$$\begin{cases} \overline{u} = \min\limits_{y,u} J(y,u) \\ Ay = Bu \quad u \in U_{ad} \subset U \end{cases}$$

For some set U_{ad} closed, convex and bounded. And $A \in \mathcal{L}(Y,Z)$ bounded and invertible with $A^{-1} \in \mathcal{L}(Z,Y)$ and $B \in \mathcal{L}(U,Z)$.

Then we can write $y \in Y$ as a function of $u \in U$,

$$y = y(u) = A^{-1}Bu$$

Consider the reduced cost functional $F(u) := J(y(u), u)$, then our problem is equivalent to

$$\overline{u}_{ad} = \min_{u \in U_{ad}} F(u)$$

Let $(u_k)_k \in U_{ad}$ denote a minimizing sequence, i.e. $F(u_k) \to \inf_{u \in U_{ad}} F(u)$, since $u_k \in U_{ad}$ the sequence is bounded. Therefore we can find a convergent subsequence $u_{k_l} \xrightharpoonup[l\to\infty]{} \overline{u}$, moreover since U_{ad} is closed and convex U_{ad} is weakly closed, implying $\overline{u} \in U_{ad}$

Proposition 3.1.
If J is continuous and weakly lower semicontinuous, then $\overline{u} = \arg\min\limits_{u \in U_{ad}} [F(u)]$.

Proof. If J is weakly lower semicontinuos

$$J(y(\overline{u}), \overline{u}) \leq \liminf_{l \to \infty} J(y(u_k), u_k)$$

That is,

$$F(\overline{u}) \leq \liminf_{l \to \infty} F(u_k) = \alpha$$

Since $u_{k_l} \xrightarrow[l\to\infty]{} \overline{u}, \implies y(u_k) \rightharpoonup y(\overline{u})$ and $A^{-1}Bu_k \rightharpoonup A^{-1}B\overline{u}$
J is Gateâux differentiable, applying the chain rule to F and valuating in u we have

$$F_u(u;h) = J_y(y;\alpha(u;h),u) + J_u(y,u;h)$$

$$\begin{aligned}
0 &\leq \langle u-\overline{u}, \nabla_u F(\overline{u})\rangle \quad \forall u \in U_{ad}\\
&= \langle A^{-1}B(u-\overline{u}), \nabla_y J(\overline{y},\overline{u})\rangle_{Y^*Y} + \langle u-\overline{u}, \nabla_u J(\overline{y},\overline{u})\rangle_{U^*U}\\
&= \langle A^{-1}B(u-\overline{u}), \nabla_y J(\overline{y},\overline{u})\rangle_{Y^*Y} + \langle u-\overline{u}, \nabla_u J(\overline{y},\overline{u})\rangle_{U^*U}\\
&= \langle u-\overline{u}, (A^{-1}B)^*\nabla_y J(y,\overline{u})\rangle_{U^*U} + \langle u-\overline{u}, \nabla_u J(y,\overline{u})\rangle_{U^*U}\\
&= \langle u-\overline{u}, (A^{-1}B)^*\nabla_y J(\overline{y},\overline{u}) + \nabla_u J(\overline{y},\overline{u})\rangle_{U^*U}
\end{aligned}$$

Setting $p^* = (A^{-1}B)^*\nabla_y J(\overline{y},\overline{u})$. We have that $\overline{u} = P_{U_{ad}}(\overline{u} - \delta(p^* + \nabla_u J(\overline{y},\overline{u})))$ $\square$

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

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

□

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): $\Phi(\overline{u},0)+\Phi^*(0,\overline{p^*})=0=(0,\overline{u})+(\overline{p^*},0)=\left((0,\overline{p^*}),(\overline{u},0)\right) \iff (0,\overline{p^*})\in\partial\Phi(\overline{u},0)\ \forall u\in U, \forall p^*\in\partial J(u) \iff J(u)+J^*(p^*)=(p^*,u)$

□

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,

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

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

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

References