

# Optimization

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  $f : H \rightarrow (\infty, \infty]$  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 continuous.

#### 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  $T : U \rightarrow G$ , let  $V$  be a neighborhood of  $Tx$ , and let  $R : V \rightarrow K$ . Suppose that  $T$  is Frchet differentiable at  $x$  and that  $R$  is Gateaux differentiable at  $Tx$ . Then  $R \circ T$  is Gateaux differentiable at  $x$  and  $D(R \circ T)(x) = (DR(Tx)) \circ DT(x)$ . 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  $T : U \rightarrow K$ . Suppose that  $T$  is twice Fréchet differentiable at  $x$ . Then  $\forall (y, z) \in H \times H$ ,  $(D^2T(x)y)z = (D^2T(x)z)y$ .

## Definition 1.1.

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

$$\lim_{0 \neq \|y\| \rightarrow 0} \frac{\|T(x+y) - Tx - DT(x)y\|}{\|y\|} = 0$$

# 2. Lecture 2

## 2.1. Convexity

### Definition 2.1.

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

$$J(tu_1 + (1-t)u_2) \leq tJ(u_1) + (1-t)J(u_2) \quad (1)$$

holds such that the right hand side 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  $J : C \rightarrow \mathbb{R}$ . Define  $\alpha = \inf_{u \in C} J(u)$ . Then the set  $\Psi = \{u \mid J(u) = \alpha\}$  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  $u_t = tu_1 + (1-t)u_2$ . Since  $J$  is convex, it holds that  $J(u_t) \leq tJ(u_1) + (1-t)J(u_2) = \alpha$ . Thus  $J(u_t) = \alpha, \forall t \in [0, 1]$ . Implying  $u_t \in \Psi$ . Hence  $\Psi$  is convex.  $\square$

**Lemma 2.4.**

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

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

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

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

For some  $t \in (0, 1)$ ,  $u_t \in B_\epsilon(\bar{u})$ .

Thus,

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

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

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

Therefore,  $\bar{u}$  is a local minimizer for  $C$ .  $\square$

**Theorem 2.1.**

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

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

1. Let  $\bar{u}$  be a local solution. Then  $J'(\bar{u}; u - \bar{u}) \geq 0, \forall u \in C$ .
2. If  $J$  is convex on  $C$ , then  $J'(\bar{u}; u - \bar{u}) \geq 0, \forall u \in C$  is necessary and sufficient for global optimality of  $\bar{u}$
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

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

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

*Proof.*

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

For small  $t > 0$ ,

$$0 \leq \frac{1}{t} [J(u_t) - J(\bar{u})] \leq \frac{1}{t} [J(\bar{u} + t(u - \bar{u})) - J(\bar{u})] \xrightarrow{t \downarrow 0} J'(\bar{u}; u - \bar{u})$$

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

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

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

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

$\implies (u_k)_k$  is bounded, because  $J \rightarrow \infty$  as  $\|u\| \rightarrow \infty$ .

$\implies (u_k)_k$  contains a weakly convergent subsequence  $u_{k_e} \xrightarrow{e \rightarrow \infty} \bar{u} \in C$ . Since  $C$  is closed and convex.

$\implies J$  is weakly-lower semicontinuous because it is convex and continuous.

□

### 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  $J : H \rightarrow \mathbb{R}$  a functional over a Hilbert space  $H$ , we define the set,

$$\arg \min_{v \in C \subseteq H} J(x) := \{x \mid x \in C \wedge \forall v \in C : J(x) \leq J(v)\}.$$

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 Gateaux gradient of  $J$  at  $x$ .

#### Lemma 3.1.

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

$$P_C(x) = \arg \min_{v \in C} [\|v - x\|].$$

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.  $\|P_C(u) - P_C(\bar{u})\| \leq \|u - \bar{u}\| \quad \forall u, \bar{u} \in H$ , i.e. The projection  $P_C$  is non expansive.
4.  $\langle P_C(u) - P_C(\bar{u}), u - \bar{u} \rangle \leq 0, \quad \forall u, \bar{u} \in H$
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

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

By the parallelogram law,

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

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

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

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  $\|v_i - x\| \rightarrow \alpha$  and  $\|v_j - x\| \rightarrow \alpha$ , we have that  $\|v_j - v_i\| \rightarrow 0$ , 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,  $J : C \rightarrow \mathbb{R}$ , Gateaux differentiable at the local solution  $\bar{u}$  of  $\min_{u \in C} J(u)$ . Thus,  $J'(\bar{u}; u - \bar{u}) \geq 0$ ,  $\forall u \in C$  and it is equivalent to  $\bar{u} = P_C(\bar{u} - \delta \nabla J(\bar{u}))$ ,  $\forall \delta > 0$ .

*Proof.* Since every Hilbert Space is a Banach space, and  $C$  is closed and Convex subset of  $H$ , and  $\bar{u}$  is a solution of minimization problem; we can apply 2.1.

Thus  $J'(\bar{u}; u - \bar{u}) \geq 0 \iff \langle u - \bar{u}, \nabla J(\bar{u}) \rangle \geq 0 \forall u \in C$ .

For all  $\delta > 0$ , we multiply the Gateaux gradient  $(-\delta)$  and we have,

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

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

$$\bar{u} = P_C(w) \iff \langle u - \bar{u}, w - \bar{u} \rangle$$

Thus,

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

□

### 3.1. Application

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

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

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

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

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

#### Proposition 3.1.

If  $J$  is continuous and weakly lower semicontinuous, then  $\bar{u} = \arg \min_{u \in U_{ad}} [F(u)]$ .

*Proof.* If  $J$  is weakly lower semicontinuous

$$J(y(\bar{u}), \bar{u}) \leq \liminf_{l \rightarrow \infty} J(y(u_k), u_k)$$

That is,

$$F(\bar{u}) \leq \liminf_{l \rightarrow \infty} F(u_k) = \alpha$$

Since  $u_{k_l} \xrightarrow{l \rightarrow \infty} \bar{u}$ ,  $\implies y(u_k) \rightharpoonup y(\bar{u})$  and  $A^{-1}Bu_k \rightharpoonup A^{-1}B\bar{u}$

□

$J$  is Gateaux differentiable, applying the chain rule to  $F$  and valuating in  $u$  we have

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

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

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

## 4. Lecture 4

### Lemma 4.1.

Let  $U$  be linear space and  $J : U \rightarrow \bar{\mathbb{R}}$ . Then

1. If  $J$  is convex, then the effective domain  $\text{dom}(J) = \{u \in U \mid J(u) < \infty\}$  is convex.
2.  $J$  is convex  $\iff \text{epi}(J) = \{(u, \alpha) \in U \times \mathbb{R} \mid J(u) \leq \alpha\}$  is convex.

*Proof.* Since  $U$  and  $\mathbb{R}$  are linear spaces, is easy to see that scalar multiplications and sums are well defined over  $U \times \mathbb{R}$  and so over  $\text{epi}(J)$ .

1. Assume  $J$  convex. If  $u_1 \in \text{dom}(J)$  and  $u_2$  are elements of  $\text{dom}(J)$ . Therefore,  $J(u_1) < \infty$ , and  $J(u_2) < \infty$ , therefore for  $t \in [0, 1]$ , we have  $tJ(u_1) < \infty$  and  $(1-t)J(u_2) < \infty$ . Since  $J$  is convex,

$$J(tu_1 + (1-t)u_2) \leq tJ(u_1) + (1-t)J(u_2) < \infty$$

,

Therefore,  $tu_1 + (1-t)u_2 \in \text{dom}(J)$ . Hence  $\text{dom } J$  is convex.

2. First consider  $J$  a convex functional, then we have for all  $u_1, u_2 \in U$ ,

$$J(tu_1 + (1-t)u_2) \leq tJ(u_1) + (1-t)J(u_2) \quad \forall t \in [0, 1]$$

Let  $(u_1, \alpha_1), (u_2, \alpha_2)$  elements of  $\text{epi}(J)$ , then  $J(u_1) < \alpha_1$  and  $J(u_2) < \alpha_2$ . Since  $J$  is convex.

$$J(tu_1 + (1-t)u_2) \leq t\alpha_1 + (1-t)\alpha_2$$

Then  $(tu_1 + (1-t)u_2, t\alpha_1 + (1-t)\alpha_2) \in \text{epi}(J)$ . Therefore, if  $J$  is convex, and  $(u_1, \alpha_1), (u_2, \alpha_2)$  are elements of  $\text{epi}(J)$  then,

$$t(u_1, \alpha_1) + (1-t)(u_2, \alpha_2) = (tu_1 + (1-t)u_2, t\alpha_1 + (1-t)\alpha_2) \in \text{epi}(J)$$

Hence  $\text{epi}(J)$  is convex.

Now assume  $\text{epi}(J)$  convex. Let  $(u_1, \alpha_1), (u_2, \alpha_2)$  elements of  $\text{epi}(J)$  then  $(tu_1 + (1-t)u_2, t\alpha_1 + (1-t)\alpha_2)$ , then

$$J(tu_1 + (1-t)u_2) \leq t\alpha_1 + (1-t)\alpha_2 \quad \forall t \in [0, 1]$$

By definition of  $\text{epi}(J)$ , if  $u_1, u_2 \in \text{dom } J$ , then  $(u_1, J(u_1))$  and  $(u_2, J(u_2))$ , are elements of  $\text{epi}(J)$ , therefore

$$tJ(u_1) + (1-t)J(u_2) \leq t\alpha_1 + (1-t)\alpha_2$$

Implying that  $J$  is convex.

□

**Definition 4.1.**

## 5. Lecture 5

## 6. Lecture 6

## 7. Lecture 7

**Remark 7.1.**

Some elementary properties of conjugate functions

- **Young inequality**  $J(u) + J^*(p^*) \geq p^*(u) \quad \forall u \in U, \forall p^* \in U^*$
- $J^*(0) = \sup_{u \in U} ((0, u) - J(u)) = \sup_{u \in U} (-J(u)) = \inf_{u \in U} J(u)$

In many applications in optimization, is used the equivalent formulation,

$$\inf_{u \in U} J(u) = -J^*(0).$$

$$J \leq F \implies J^* \geq F^*$$

**Theorem 7.1.**

Let  $U$  a Banach space and  $J^* : U^* \rightarrow \overline{\mathbb{R}}$  be the conjugate of the  $J : U \rightarrow \overline{\mathbb{R}}$ . Then for all  $u \in U$ .

$$p^* \in \partial J(u) \iff J(u) + J^*(p^*) = p^*(u)$$

.

*Proof.* content...

□

**Corollary 7.1.**

It follows from previous theorem that  $\partial J(u) = \{p^* \in U^* | J(u) + J^*(p^*) = (p^*, u)\}$ .

**Theorem 7.2.**

Let  $U$  be a Banach space and  $J : U \rightarrow \mathbb{R}$  be proper function. If  $p^* \in \partial J(u)$  then  $u \in \partial J^*(p^*)$



*Proof.* Let  $p^* \in \partial J(u)$ . For any  $g^* \in U^*$ , it follows

$$J^*(g^*) = \sup_{v \in U} (g^*(v) - J(v)) \geq g^*(u) - J(u) \geq g^*(u) - J(u)$$

From theorem 7.1

$$J^*(g^*) \leq g^*(u) - p^*(u) + J^*(p^*) = (g^* - p^*)(u) + J^*(p^*) \implies u \in \partial J^*(p^*).$$

□

By iteration the definition, we obtain the bipolar function  $(J^*)^* = J^{**} : U^{**} \rightarrow \overline{\mathbb{R}}$ ,

$$J^{**}(u) = \sup_{p^* \in U^*} \{p^*(u) - J^*(p^*)\}$$

### Theorem 7.3.

Let  $U$  be a reflexive Banach space. The  $J^{**}$  is the maximum convex functional below  $J$  (also called convex envelope), i.e.  $J^{**}(u) \leq J(u)$ ,  $\forall u \in U$  and  $F(u) \leq J^{**}(u)$ ,  $\forall u \in U$  if  $F$  is also convex and  $F(u) \leq J(u)$ ,  $\forall u$ . In particular  $J^{**} = J$  if and only if  $J$  is convex.

*Proof.*

$$J^{**}(u) = \sup_{p^* \in U^*} \{p^*(u) - J^*(p^*)\} \tag{2}$$

$$= \sup_{p^* \in U^*} \left\{ p^*(u) - \sup_{v \in U} \{p^*(v) - J(v)\} \right\} \tag{3}$$

$$= \sup_{p^* \in U^*} \left\{ p^*(u) + \inf_{v \in U} \{p^*(v) - J(v)\} \right\} \tag{4}$$

$$\tag{5}$$

Since for any  $p^* \in U^*$ ,

$$\inf_{v \in U} \{p^*(u - v) + J(v)\} \leq p^*(u - u) + J(u)$$

We have that  $J^{**}(u) \leq J(u)$ .

Now we assume that  $F$  is a convex functional and  $g^* \in \partial F(u)$  for  $u \in U$ .

$$\implies F(v) \geq F(v) + g^*(v - u) \tag{6}$$

$$F^{**}(u) = \sup_{p^* \in U^*} \inf_{v \in U} \{p^*(u - v) + F(v)\} \tag{7}$$

$$\geq \sup_{p^* \in U^*} \inf_{v \in U} \{(p^* - g^*)(u - v) + F(u)\} \tag{8}$$

$$\geq \inf_{v \in U} \{(g^* - q^*)(u - v) + F(u)\} \tag{9}$$

$$= F(u) \tag{10}$$

If  $F$  is convex,

$$\implies F(u) \leq F^{**}(u) \leq F(u) \implies F(u) = F^{**}(u), \tag{11}$$

$$F(u) = F^{**}(u) = \sup_{p^* \in U^*} \inf_{v \in U} \{p^*(u - v) + F(v)\} \leq J^{**}(u) \tag{12}$$

□

## 8. Lecture 8

### Definition 8.1.

Let  $U$  and  $Y$  Banach spaces and  $J : U \rightarrow \overline{\mathbb{R}}$  is a proper function. We consider an optimization problem (P) called primal problem

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

Then the problem is said to be nontrivial if there is  $\bar{u} \in U$  such that  $J(\bar{u}) < \infty$ . A function  $\Phi : U \times Y \rightarrow \overline{\mathbb{R}}$  is said to be a perturbation function of  $J$ ,

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

if  $\Phi(u, 0) = J(u)$  for all  $u \in U$ . For each  $p \in Y$ , the minimization problem (Pp) is called a perturbation problem. The variable  $p$  is called perturbation parameter. If we denote by  $\Phi^*$  the convex conjugate function of  $\Phi$ , the *dual problem*, with respect to  $\Phi$  is defined by

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

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

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

### Remark 8.1.

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

### Lemma 8.1 (Weak duality).

For the problem (P) and  $(\text{P}^*)$  it holds that

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

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

$$-\Phi^*(0, p^*) = - \sup_{\substack{u \in U \\ p \in Y}} ((0, u) + (p^*, p) - \Phi(u, p)) \quad (13)$$

$$= \inf_{\substack{u \in U \\ p \in Y}} (\Phi(u, p) - (p^*, p)) \quad (14)$$

$$\leq (\Phi(u, 0) - (p^*, 0)) \quad \forall u \in U, p^* \in Y^* \quad (15)$$

$$\implies \sup_{p^* \in Y^*} (-\Phi^*(0, p^*)) \leq \inf_{u \in U} \Phi(u, 0) = \inf(\text{P}) \quad (16)$$

□

By iteration we can define, a bidual problem

$$- \sup_{u \in U} (-\Phi^*(u, 0)) = \inf_{u \in U} \Phi^*(u, 0) \quad (\text{P}^{**})$$

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  $(\text{P}) \equiv (\text{P}^{**})$

### Definition 8.2.

Consider the infimal value function

$$h(p^*) = \inf(\mathbf{P}_p) = \inf_{u \in U} \Phi(u, p)$$

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

### Theorem 8.1.

The primal problem  $(\mathbf{P})$  is stable if and only if the following conditions are simultaneously satisfied:

- The dual problem  $(\mathbf{P}^*)$  has a solution.
- There is no duality gap, i.e.

$$\inf(\mathbf{P}) = \sup(\mathbf{P}^*) \leq \infty$$

### Theorem 8.2 (Extremal relation).

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

1.  $(\mathbf{P})$  and  $(\mathbf{P}_p)$  have solutions  $\bar{u}$  and  $\bar{p}^*$  and  $\inf(P) = \sup(P^*)$
2.  $\Phi(\bar{u}, 0) + \Phi^*(0, \bar{p}^*) = 0$
3.  $(0, \bar{p}^*) \in \partial\Phi(u, 0)$  and  $(\bar{u}, 0) \in \partial\Phi^*(0, p^*)$

*Proof.* We proceed by parts:

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

□

### Fenchel duality.

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

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

with  $F : U \rightarrow \overline{\mathbb{R}}$ ,  $G$  convex function  $G : V \rightarrow \overline{\mathbb{R}}$  and  $A : U \rightarrow V$  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

---