Convex Analysis

Lecturer: Michiel Renger
Weierstraß Institut
renger@wias-berlin.de

Course Material: several books (will be more specific for each Lecture)

mostly:

> Juan Peypouquet - Convex Optimization in Normed Spaces: Theory, Methods & examples
> (Downloadable from his website!)

Language: English or Deutsch

Lecture 1

A Introduction to convexity

> **Def** X Banach space
> $F: X \to \mathbb{R} \cup \{\infty\}$ convex (cvx) $:\Leftrightarrow F((1-\sigma)x + \sigma y) \leq (1-\sigma)F(x) + \sigma F(y)$ $\forall x, y \in X, \sigma \in [0,1]$
> strictly convex $:\Leftrightarrow$ $\ldots < \ldots$ $\forall x \neq y, \sigma \in (0,1)$

$\to$ why Banach? - vector space
- connection with topology (later on)

$\to$ why $\mathbb{R} \cup \{\infty\}$? Useful convention: $\inf F = \inf_A F$ if $F|_{A^c} = \infty$

> **Def** • $\operatorname{Dom} F = \{x \in X : F(x) < \infty\}$ Domain
> • F is proper $:\Leftrightarrow \operatorname{Dom} F \neq \emptyset$

Typically X will be an L^p-space, Sobolev space or "Orlicz" space

Examples:

• $F: L^2(\Omega) \to \mathbb{R} \cup \{\infty\}$, $F(x) = \begin{cases} \frac{1}{2}\|\nabla x\|^2_{L^2(\Omega)}, & \text{if } x \in W^{1,2}(\Omega), \\ \infty, & \text{otherwise.} \end{cases}$

• $F(x) = \int_\Omega f(q, x(q))\,dq$ cvx? cvx if f cvx in x

• $F(x) = \int_\Omega f(q, \nabla(q))\,dq$ cvx if f cvx in ∇x

• $F(x) = \int_\Omega f(q, x(q), \nabla x(q))\,dq$ cvx if f cvx in x & f cvx in ∇x?

No! F only convex if f is cvx in $(x, \nabla x)$ jointly cvx

For example $f(q, x, y) = x^2 - 4xy + y^2$ is cvx in x and in y but on the diagonal $f(q, x, x) = -2x^2$.

LC Evans - PDEs, sec. 8.2
Peypouquet, sec. 2.2

B Motivating and important application:

The direct method in the calculus of variations

$$\inf_{x\in X} F(x) = \inf_{x \in X} \int_{\Omega} f(\underline{q}, x(\underline{q}), \nabla x(\underline{q}))\,d\underline{q}$$

If a minimiser exists, than it must solve the Euler-Lagrange eq.:

$$\frac{\partial f}{\partial x}(q, x(q), \nabla x(q)) = \mathrm{div}_q \frac{\partial f}{\partial \nabla x}(q, x(q), \nabla x(q)).$$

So this would prove existence of a solution to this (possibly nonlinear) equation!

How to prove that a minimiser exists?

Strategy

(i) Assume F is bounded from below.
Then there exists a minimising sequence $(x_n)_n \subset X$, i.e.
- $\lim_{n\to\infty} F(x_n) = \inf F$
- $F(x_{n+1}) \leq F(x)$ for all $n \geq 1$.

(ii) Show that the sequence converges (somehow)

(iii) Show that the limit must be a minimiser of F

(ii) $(x_n)_n \subset \{x \in X : F(x) \leq F(x_1) =: C\} = \{F \leq C\}$ (sub)level set of F

1-d: • e^x has no minimiser 

• but if $\lim_{|x|\to\infty} f(x) = \infty$  then $\{F \leq C\}$ is bounded

∞-d: Assume $F(x) \geq \varphi(\|x\|) - \alpha$ for some $\alpha \in \mathbb{R}$ and non-negative increasing $\varphi: \mathbb{R} \to \mathbb{R}$ coercivity/growth condition

Then $\varphi(\|x\|) \leq F(x) + \alpha \leq C + \alpha$ on $\{F \leq C\}$

$\|x\| \leq \varphi^{-1}(C+\alpha) \Rightarrow$ sublevel set $\{F \leq C\}$ is bounded in X.

Compactness? $\rightarrow$ Intermezzo weak topologies

Assume X has a predual, i.e. $X = Y^*$ and $\|\cdot\|_X = \|\cdot\|_{Y^*}$

$(x_n)_n \subset \{F \leq C\}$ bounded $\Rightarrow$ By Banach-Alaoglu:

$(x_n)_n$ is relatively weak-$*$ compact.

$\Rightarrow$ there exists a weakly-$*$ convergent subsequence

$x_{n_m} \xrightarrow[m\to\infty]{*} x_\infty$ candidate for minimiser!

(iii) $\lim_{m\to\infty} F(x_{n_m}) = \lim_{n\to\infty} F(x_n) = \inf F$

$x_{n_m} \xrightarrow{*} x_\infty$

If F would be weakly-* continuous, then

$\inf F = \lim_{m\to\infty} F(x_{n_m}) = F(x_\infty)$, and hence x_∞ would indeed be a minimiser!

Generally not true ☹. In fact,

> **Prop** F is weakly or weakly-* continuous
> $\Rightarrow$
> F is "strongly" (in the norm topology) continuous

Claim: lower semicontintinnity is enough:

> Assume F is $\overbrace{\text{lower}}^{\text{weakly-*}}$ semicontinuous (l.s.c.) i.e.
> $\liminf_n F(\hat{x}_n) \geq F(\hat{x}) \quad \forall\, \hat{x}_n \xrightarrow{*} \hat{x}$

Then:

$$\inf F \geq \liminf_{m\to\infty} F(x_{n_m}) \geq F(x_\infty) \geq \inf F$$

Hence x_∞ is a minimiser!

> **Th** ("Direct Method") Assume
> - F bdd from below ($\Rightarrow$ minimising sequence)
> - coercivity ($\Rightarrow$ bounded level sets)
> - X has a predual ($\Rightarrow$ weakly-* compact level sets)
> - F is weakly-* l.s.c. ($\Rightarrow$ cluster points are minimisers)
>
> Then there exists a minimiser $\inf F = F(x)$

- Remarks:
  - the existence of a minimiser has nothing to do with topology. We chose a topology that was helpful to prove this.
  - By choosing a weaker (= coarser) topology, we make it easier to prove compactness, but more difficult to prove (lower semi-) continuity
  - I haven't told you <u>how to prove l.s.c. & coercivity</u>. Here convexity will play an important role.
  - Next week more convexity.

H. Brezis - Analyse Fonctionelle sec. III
H. Brezis - Functional Analysis, Sobolev spaces & PDE's sec. III
Peypouquet 1.1 & 1.3

Recap: weak topologies

dual $\quad X^* := \{x^* : X \to \mathbb{R} \text{ linear \& bounded}\}$ "dual space"

$\downarrow$

$|x^*(x)| \leq C\|x\| \ \forall x$

notation $\langle x^*, x\rangle := x^*(x)$

$\|x^*\|_{X^*} := \sup_{x \neq 0} \frac{\langle x^*, x\rangle}{\|x\|} = \sup_{\|x\| \leq 1} \langle x^*, x\rangle$ (minimal constant)

Th $(X^*, \|\cdot\|_{X^*})$ is a Banach space

weak topology

Def. • A sequence $(x_n)_n$ converges weakly, $x_n \rightharpoonup x$ $:\Leftrightarrow$
$\langle x^*, x_n\rangle \to \langle x^*, x\rangle$ for all $x^* \in X^*$

• The weak topology $\sigma(X, X^*)$ is the topology on X generated by the subbase $\{\{x \in X : |\langle x^*, x - x_0\rangle| \leq \varepsilon\} : x \in X^*, x_0 \in X, \varepsilon > 0\}$

Remarks:
- Weak topologies are "Hausdorff", which implies that limits are unique.
- Weak topologies are generally not metrisable, which implies that convergent sequences are not enough to fully characterise the topology!
- ~~Weak topologies~~ $(X, \sigma(X, X^*))$ are so-called "locally convex vector spaces" (LCS). This is beyond the ~~focus~~ scope of this course.
- $x_n \to x \Rightarrow x_n \rightharpoonup x$

weak-* topology

Def. • A sequence $(x^*_n)_n$ converges weak-*, $x^*_n \xrightarrow{*} x^*$ $:\Leftrightarrow$
$\langle x^*_n, x\rangle \to \langle x^*, x\rangle$ for all $x \in X$

• The weak-* topology is the topology on X^* generated by the subbase $\{\{x^* \in X^* : |\langle x^* - x^*_0, x\rangle| < \varepsilon\} : x \in X, x^*_0 \in X^*, \varepsilon > 0\}$

Th (Banach-Alaoglu) The balls $\{\|x^*\|_{X^*} \leq 1\} \subset X^*$ are weakly-* compact.

Cor Any bounded set $A \subset X^*$, $\sup_{x^* \in A} \|x^*\| \leq C$ is relatively weakly-* compact

Lecture 2

Recall $x^* \in X^* :\Leftrightarrow x^*: X \to \mathbb{R}$ linear & bounded ($|x^*(x)| \le C\|x\|$)

$\Leftrightarrow x^*: X \to \mathbb{R}$ linear & $\|x^*\|_{X^*} := \sup_{\|x\| \le 1} \langle x^*, x \rangle < \infty$

$\overset{\text{in fact}}{\Leftrightarrow} x^*: X \to \mathbb{R}$ linear & continuous

→ In literature sometimes written as X'; however X' is also sometimes used for the space of all linear functionals.

We write $\langle x^*, x \rangle := x^*(x)$ to stress "bilinearity", i.e.:

$\langle x^*, ax_1 + bx_2 \rangle = a\langle x^*, x_1 \rangle + b\langle x^*, x_2 \rangle$

and $\langle ax_1^* + bx_2^*, x \rangle = a\langle x_1^*, x \rangle + b\langle x_2^*, x \rangle$

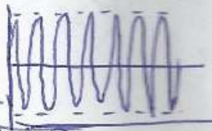
Three topologies:

$x_n \to x$	("strong / in norm")	$:\Leftrightarrow \|x_n - x\| \to 0$
$x_n \rightharpoonup x$	("weak")	$:\Leftrightarrow \langle x^*, x_n \rangle \to \langle x^*, x \rangle \ \forall x^* \in X^*$
$x_n^* \overset{*}{\rightharpoonup} x^*$	("weak-*")	$:\Leftrightarrow \langle x_n^*, x \rangle \to \langle x^*, x \rangle \ \forall x \in X$

From exercise 1: $\boxed{x_n^* \to x^* \Rightarrow x_n^* \rightharpoonup x^* \Rightarrow x_n^* \overset{*}{\rightharpoonup} x^*}$

A Examples and properties of weak / weak-* topologies

(I) $X := L^2(0, 2\pi)$ (Hilbert space!), $x_n(t) := \sin nt$



Th (Riesz representation Theorem for Hilbert spaces)

$L^2(0,2\pi)^* \simeq L^2(0,2\pi)$, i.e.:

any $x^* \in L^2(0,2\pi)^*$ is of the form $x^*(x) = \langle x^*, x \rangle = \int_0^{2\pi} \tilde{x}^*(t)\, x(t)\, dt$

for some $\tilde{x}^* \in L^2(0,2\pi)$, and $\|x^*\|_{X^*} = \|\tilde{x}^*\|_X$.

Peypouquet Th. 1.4.1
Brézis Th. 5.4 & 4.11

→ It is customary to identify x^* with $\tilde{x}^*$.

→ Since L^2 is Hilbert, it is its own dual and predual. Hence weak and weak-* convergence are the same!

→ $x_n(t) = \sin nt$ does not converge in norm.

However, $\|x_n\|_{L^2} = \int_0^{2\pi} \sin^2 nt\, dt \le 2\pi$ hence by Banach-Alaoglu there exists at least a weakly-* (= weakly) conv. subseq.

Prop $x_n \to 0$

Uses the following lemma:

Lem $C_c^\infty(0,2\pi) \overset{\text{dense}}{\subset} L^2(0,2\pi)$ in the norm topology | Brézis Th. 4.23 |

C_c^∞: smooth functions with compact support

(actually true for any L^p, $1 \le p < \infty$)

proof of prop:

Take an arbitray "test function" $\varphi \in L^2(0,2\pi)$, and approximate $C_c^\infty \ni \varphi_m \xrightarrow{L^2} \varphi$. For each such φ_m:

$$|\langle \varphi_m, x_n\rangle| = \left|\int_0^{2\pi} \varphi_m(t)\sin nt\, dt\right| = \frac{1}{n}\left|\int_0^{2\pi}\varphi_m'(t)\cos nt\, dt\right|$$
$$\le \frac{1}{n}\|\varphi_m'\|_{L^1} \xrightarrow[n\to\infty]{} 0.$$

Then

$$|\langle \varphi, x_n\rangle - \underbrace{\langle \varphi, 0\rangle}_{=0}| = |\langle \varphi_m, x_n\rangle - \langle \varphi_m - \varphi, x_n\rangle|$$
$$\text{(Cauchy-Schwarz)} \le |\langle \varphi_m, x_n\rangle| + \|\varphi_m - \varphi\|_{L^2}\|x_n\|_{L^2}$$
$$\le |\langle \varphi_m, x_n\rangle| + 2\pi\|\varphi_m - \varphi\|_{L^2}$$
$$\xrightarrow[n\to\infty]{} 2\pi\|\varphi_m-\varphi\|_{L^2}$$
$$\xrightarrow[m\to\infty]{} 0. \quad \blacksquare$$

II (Exercise 2)

Prop Assume X has a predual (and the norms coincide). Then $F(x) := \|x\|_X = \|x\|_{Y^*}$ is weakly-$*$ l.s.c.

proof: Take any sequence $x_n \xrightarrow{*} x$. Then: | Peypouquet prop. 1.22. |

$$\liminf_{n\to\infty}\|x_n\|_{Y^*} = \liminf_{n\to\infty}\sup_{\substack{y\in Y\\ \|y\|\le 1}}\langle x_n, y\rangle$$
$$\text{(for any } \|\hat{y}\|\le 1) \ge \liminf_{n\to\infty}\langle x_n, \hat{y}\rangle = \langle x, \hat{y}\rangle.$$

Now take the supremum over $\|\hat{y}\|\le 1$ on both sides:

$$\liminf_{n\to\infty}\|x_n\|_{Y^*} \ge \|x\| \quad \blacksquare$$

→ More general principle: a supremum over lsc functionals is always l.s.c!

→ We give a more precise definition of l.s.c. later...

→ Will see: any lsc cvx functional is $\sup_{i\in I}$ (affine, cont. functions)

(III) Exercise (3): Can you construct a convex discontinuous function $x:\mathbb{R}\to\mathbb{R}$?

→ Impossible in finite dimensions! (We come back to this)

B Topological properties of convex sets

Def A set $A\subset X$ is convex $:\Leftrightarrow \forall x_1, x_2\in A$ and $\sigma\in[0,1]$, $(1-\sigma)x_1+\sigma x_2\in A$.

Th (Hahn-Banach geometric / separation Theorem)

Brezis Th. 1.7 ≈ Peypouquet Th. 1.10

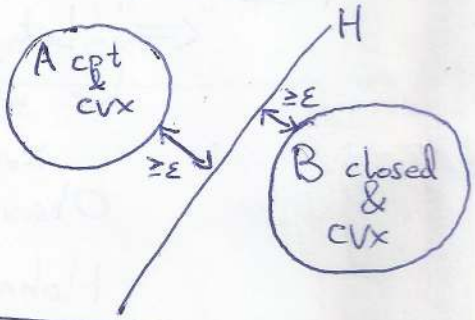
$A, B\subset X$ disjoint convex sets ($A\cap B=\emptyset$)

A compact, B closed (in norm topology)

Then there exists a "separating hyperplane"

$$H=\{x: \langle x^*, x\rangle=\gamma\} \text{ for some } x^*\in X^* \text{ and } \gamma\in\mathbb{R}, \text{ and } \varepsilon>0 \text{ s.t.}$$

$$\langle x^*, a\rangle\le\gamma-\varepsilon \quad \forall a\in A$$
$$\langle x^*, b\rangle\ge\gamma+\varepsilon \quad \forall b\in B$$



→ Many different equivalent "Hahn-Banach" Theorems exist in the Literature.

→ The proof is beyond the scope of this course, however … Keep in mind that the proof uses the axiom of choice!

Hence the proof is not constructive!

Recall the subbase of the weak topology:

$$\{\{|\langle x^*, x-x_0\rangle| \leq \varepsilon\} : x^*, x_0, \varepsilon\}.$$

In particular, sets of the form $\{\langle x^*, x\rangle < \gamma\}$ are weakly open.

Moreover $\sigma(X,X^*) \subset \sigma(\|\cdot\|_X)$, i.e. the weak topology is coarser/weaker than the norm topology.

Hence $C \subset X$ weakly closed

$\Rightarrow$

$C^c \in \sigma(X,X^*) \subset \sigma(\|\cdot\|_X)$

$\Rightarrow$

C also norm closed.

Remarkably:

$$\boxed{\begin{array}{l}\textbf{Th} \text{ If } C \subset X \text{ is convex, then:} \\ \quad C \text{ weakly closed} \Leftrightarrow C \text{ norm closed}\end{array}}$$

Peypouquet Prop. 1.21

proof "$\Rightarrow$" trivial (just proven above)

"$\Leftarrow$" Let C be norm-closed, and take an arbitrary $x_0 \in C^c$.
We will show that there is a "weakly open ball" $x_0 \in V \subset C^c$, and hence C^c must be weakly open.
Observe that $\{x_0\}$ is convex and compact.
Hahn-Banach: $\exists\, x^* \in X^*, \gamma \in \mathbb{R}$ s.t.

$$\langle x^*, x_0\rangle < \gamma < \langle x^*, x\rangle \quad \forall x \in C.$$

Let $V := \{x \in X : \langle x^*, x\rangle < \gamma\}$. Then:

✓ $x_0 \in V$

✓ $V \cap C = \emptyset$

✓ V is weakly open □

$\sigma(X,X^*)$ is not metrisable: need to distinguish between topological and sequential properties

Peypouquet p. 12

Def $C\subset X$ closed (in some topology) $:\Leftrightarrow C^c \in \sigma$

$C\subset X$ sequentially closed $:\Leftrightarrow C$ is closed under σ-convergent sequences, i.e. $\forall\, x_n \to x,\ (x_n)_n \subset C \Rightarrow x \in C$.

Cor For $C\subset X$:

Peypouquet Prop 1.23

ⓐ C is weakly closed

$\Downarrow$ ("sequence lemma")

C is weakly sequentially closed

$\Downarrow$ (trivial; last week's exercise ①)

C is (norm) sequentially closed

$\Updownarrow$ ("sequence lemma for metrisable spaces")

C is (norm) closed

ⓑ All equivalent if C is convex

Lecture 3

Last week: distinction between top. & seq. notion

(a) C weakly closed
$\Downarrow$
C weakly sequentially closed
$\Downarrow$
C (norm) sequentially closed
$\Updownarrow$
C (norm) closed

(b) Equivalent if C is convex (by Hahn-Banach)

A Lower semicontinuity

Def. • $F: X \to \mathbb{R} \cup \{\infty\}$ is lsc (in some topology σ) $:\Longleftrightarrow$ All level sets $\{F \le C\}$ are closed

• $F: X \to \mathbb{R} \cup \{\infty\}$ is sequentially lsc $:\Longleftrightarrow$ For all convergent sequences $x_n \xrightarrow{\sigma} x$:
$$\liminf_{n\to\infty} F(x_n) \ge F(x)$$

Peypouquet p. 27 & 29

Lemma (in some topology)
F is seq. lsc. $\Leftrightarrow$ All level sets $\{F \le C\}$ are seq. closed

(should have been in Peypouquet) (5)

"$\Rightarrow$" Take a convergent sequence $(x_n)_n \subset \{F \le C\}$, $x_n \to x$.
Then $F(x) \le \liminf_n F(x_n) \le C$, hence $x \in \{F \le C\}$.

"$\Leftarrow$" Take a convergent sequence $(x_n)_n \subset X$, $x_n \to x$.
Pick an arbitrary $\varepsilon > 0$ and a (convergent) subsequence for which $F(x_{n_m}) \le (-\frac{1}{\varepsilon}) \vee \liminf_{n\to\infty} F(x_n) + \varepsilon =: C$.
Since $\{F \le C\}$ is seq.-closed: ($\vee$ denotes maximum)
$$F(x) \le C = (-\tfrac{1}{\varepsilon}) \vee \liminf_{n\to\infty} F(x_n) + \varepsilon.$$
As ε was chosen arbitrarily:
$$F(x) \le \liminf_{n\to\infty} F(x_n) \quad \square$$

xercise (5):

Lemma $F: X \to \mathbb{R} \cup \{\infty\}$ convex $\Rightarrow$ All level sets $\{F \le C\}$ convex

→ Beware: not necessarily the other way around!

Cor For $F: X \to \mathbb{R} \cup \{\infty\}$ (a)
F weakly lsc
$\Downarrow$
F weakly seq. lsc
$\Downarrow$
F (norm) seq. lsc
$\Updownarrow$
F (norm) lsc

(b) equivalent if F is convex

Peypouquet Prop 2.17

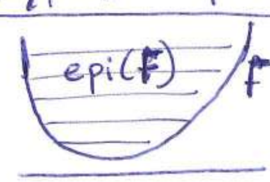
→ Recall: in the direct method one needs to prove
① compact level sets,
② lower semicontinuity (sequential)
in some topology. But if level sets are compact, then they are also closed, hence the l.s.c. is trivial!

B Epigraph (arbitrary topology)

Def (Epigraph). For an $F: X \to \mathbb{R} \cup \{\infty\}$

$$\mathrm{epi}(F) := \{(x, c) \in X \times \mathbb{R} : F(x) \le c\}$$

Peypouquet p. 26



sometimes useful in proofs, for example:

Prop For an $F: X \to \mathbb{R} \cup \{\infty\}$:

F convex
$\iff$
$\mathrm{epi}(F)$ convex
$\Longrightarrow$ (last week exercise ④ⓐ)
All level sets $\{F \le C\}$ are convex

(should have been in Peypouquet, but it's very easy!)

Prop $F: X \to \mathbb{R} \cup \{\infty\}$ is lsc
$\iff$
$\mathrm{epi}(F)$ is closed in $X \times \mathbb{R}$ (in product topology)

(Peypouquet Prop. 2.3)

→ Peypouquet uses a different, equivalent definition of lsc.

proof "$\Rightarrow$" Take an $(x_0, C_0) \in \mathrm{epi}(F)^c$, i.e. $F(x_0) > C_0$.
Let $\mu := \frac{F(x_0) + C_0}{2}$. By lsc, the level set
$\{F \le \mu\}$ is closed, hence
$\{F > \mu\}$ is open, hence
$(x_0, C_0) \in \underbrace{\{F > \mu\} \times (-\infty, \mu)}_{\{(x, C) \,:\, F(x) > \mu > C\}}$ is open, and disjoint with $\mathrm{epi}(F)$.
Hence $\mathrm{epi}(F)^c$ is open.

"$\Leftarrow$" If $\mathrm{epi}(F)$ is closed, then also (for any C)
$\mathrm{epi}(F) \cap (X \times \{C\}) = \{F \le C\} \times \{C\}$ closed.
It follows that $\{F \le C\}$ is closed $\square$

Cor $(F_i)_{i\in I}$ family of lsc functionals $\Rightarrow \sup_{i\in I} F_i$ is lsc (Peypouquet Example 2.4)

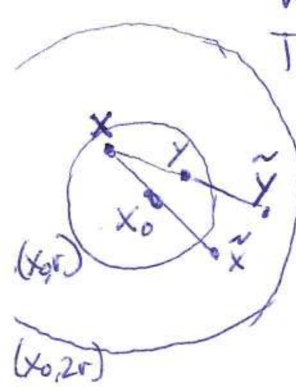
Proof All F_i lsc $\Leftrightarrow$ All $\mathrm{epi}(F_i)$ closed
$$\Rightarrow \mathrm{epi}(\sup_i F_i) = \bigcap_i \mathrm{epi}(F_i) \text{ closed}$$
$$\Leftrightarrow \sup_i F_i \text{ lsc.} \quad \square$$

[C] Continuity (in norm topology)

We first need a technical result; its implications will be super cool...

Lemma $F: X \to \mathbb{R}\cup\{\infty\}$ convex, $x_0 \in X$ (Peypouquet Prop. 3.2)
i) F bdd from above on a nbh of x_0
$\Leftrightarrow$ ii) F is Lipschitz cont. on a nbh of x_0
$\Leftrightarrow$ iii) F is cont. in x_0
$\Leftrightarrow$ iv) $F(x_0) < C \Rightarrow (x_0, C) \in \mathrm{int}(\mathrm{epi}(F))$

f i) $\Rightarrow$ ii) There exists a nbh $B(x_0, 2r) = \{\|x_0 - x\| < 2r\}$ on which $F(z) \le K \quad \forall z \in B(x_0, 2r)$.
Will prove Lipschitz cont. on $B(x_0, r)$.
Take any $x, y \in B(x_0, r)$ and construct:



$$\tilde{y} := y + r\frac{y-x}{\|y-x\|} \Leftrightarrow y = \lambda\tilde{y} + (1-\lambda)x, \ \lambda = \frac{\|y-x\|}{\|y-x\|+r} \le \frac{\|y-x\|}{r}$$
$$\tilde{x} := 2x_0 - x \quad \Leftrightarrow x_0 = \tfrac{1}{2}\tilde{x} + \tfrac{1}{2}x,$$
Then $\tilde{x}, \tilde{y} \in B(x_0, 2r)$ and so $F(\tilde{x}), F(\tilde{y}) \le K$. Convexity:
$$F(y) \le \lambda F(\tilde{y}) + (1-\lambda)F(x) \le \lambda K + (1-\lambda)F(x) \qquad (1)$$
$$F(x_0) \le \tfrac{1}{2}F(\tilde{x}) + \tfrac{1}{2}F(x) \le \tfrac{1}{2}K + \tfrac{1}{2}F(x) \qquad (2)$$
$$F(y) - F(x) \overset{(1)}{\le} \lambda(K - F(x)) \overset{(2)}{\le} 2\lambda\underbrace{(K - F(x_0))}_{\ge 0}$$
$$\le 2\frac{\|y-x\|}{r}(K - F(x_0)).$$

ii) $\Rightarrow$ iii) Trivial.

iii) $\Rightarrow$ iv) Take a $C > F(x_0)$ and an arbitrary $\eta \in (F(x_0), C)$. By continuity there exists a ball such that $F(z) < \eta \quad \forall z \in B(x_0, \delta)$.
Then $(x_0, C) \in \underbrace{B(x_0, \delta) \times (\eta, \infty)}_{\text{open nbh of } (x_0, C)} \subset \text{epi}(F)$

and so $(x_0, C) \in \text{int}(\text{epi}(F))$.

iv) $\Rightarrow$ i) Take any $C > F(x_0)$; since $(x_0, C) \in \text{int}(\text{epi}(F))$ there is again a ball and $\eta \in (F(x_0), C)$ s.t.

$$(x_0, C) \in B(x_0, \delta) \times (\eta, \infty) \subset \text{epi}(F).$$

Therefore $F(z) \leq \eta \quad \forall z \in B(x_0, \delta)$ $\square$

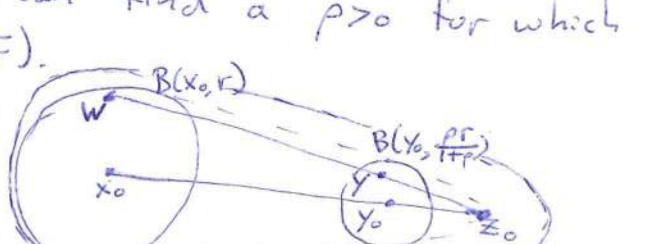
Proposition If $F: X \to \mathbb{R} \cup \{\infty\}$ is convex, and cont. at some $x_0 \in \text{dom}(F)$, then $x_0 \in \text{int}(\text{dom}(F))$ and F is cont. on whole $\text{int}(\text{dom}(F))$. (Peypouquet Prop 3.3)

$\to$ Note that $\text{dom}(F) = \bigcup_{C \in \mathbb{R}} \underbrace{\{F \leq C\}}_{\text{cvx}}$ is always convex!

proof. Assume F is cvx, and cont. at some $x_0 \in \text{dom}(F)$. By the previous lemma (iii) $\Rightarrow$ i)) $x_0 \in \text{int}(\text{dom}(F))$. By the same Lemma, $F(x) \leq K$ on some $B(x_0, r) \ni x$.
Take an arbitrary $y_0 \in \text{int}(\text{dom}(F))$; we shall prove that y_0 is also bdd on a small neighborhood, so that F is continuous in y_0.
Since y_0 lies in the interior, one can find a $\rho > 0$ for which

$$z_0 := y_0 + \rho(y_0 - x_0) \in \text{dom}(F).$$

$$\left(y_0 = \left(\tfrac{\rho}{1+\rho}\right) x_0 + \tfrac{1}{1+\rho} z_0 \right)$$



Take any $y \in B(y_0, \frac{\rho r}{1+\rho})$, and define w such that $y =: \frac{\rho}{1+\rho} w + \frac{1}{1+\rho} z_0$.
Then $w \in B(x_0, r)$ so $F(w) \leq K$. By convexity
$F(y) \leq \frac{\rho}{1+\rho} F(w) + \frac{1}{1+\rho} F(z_0) \leq K \vee F(z_0)$. $\square$

So far we haven't used completeness of the space X

In finite dimensions:

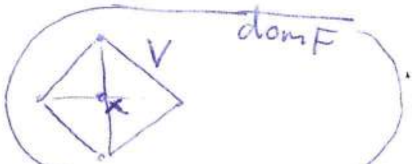
Prop Let X be finite-dimensional and $F: X \to \mathbb{R} \cup \{\infty\}$ convex. Then F is continuous on $\mathrm{int}(\mathrm{dom}(F))$. (Peypouquet prop 3.5)

→ First week exercise ③.

→ Recall: all norms on a finite-dim. space are equivalent ⇒ we may take $X = \mathbb{R}^d$.

proof For any $x \in \mathrm{int}(\mathrm{dom}(F))$, ~~take a ball $B(x_0, \delta)$ such that $B(x_0,\delta) \subset \mathrm{dom}(F)$. By convexity~~ take a $\rho > 0$ such that $x \pm \rho e_i \in \mathrm{dom}F$, $i = 1, \dots, d$, and let V be the convex hull of these points, i.e.:

$$V := \mathrm{co}(\{x \pm \rho e_i\}_i) = \{(1-\sigma)(x \pm \rho e_i) + \sigma(x \pm \rho e_j) : i, j = 1, \dots, d\}$$



By convexity, $\forall v \in V$ $F(v) \le \max(\{x \pm \rho e_i\}_i) < \infty$, hence x is bdd above on a neighborhood ⇒ cont. in x. □

In Infinite dimensions:

Prop Let X be a (complete!) Banach space and $F: X \to \mathbb{R} \cup \{\infty\}$ convex, and l.s.c.. Then F is continuous on $\mathrm{int}(\mathrm{dom}(F))$. (Peypouquet prop 3.6)

→ proof beyond the scope of this course (based on choice axiom, although it may be circumvented if the space is separable)

→ Beware that these continuity results are all in the norm topology

→ However, we saw before that norm (lower-semi-) continuity implies weak lower semicontinuity if the functional is convex.

Lecture 4: answers to exercises

L2, ex. 1 in $\mathbb{R}$ (or in any finite dimensions), the norm, weak and weak-* topologies coincide!

L2, ex 2 $C_c(\mathbb{R})^* \simeq$ {regular, signed measures on $\mathbb{R}$}
$C_b(\mathbb{R})^* \simeq$ {regular, finitely additive set functions on $\mathbb{R}$}

$(\delta_n)_{n\in\mathbb{N}} \subset C_c(\mathbb{R})^* \subset C_b(\mathbb{R})^*$

a) weak-* compact in $C_c(\mathbb{R})^*$? "vague topology"

$$\|\delta_n\|_{C_c^*} = \sup_{\substack{\varphi\in C_c \\ \|\varphi\|_{L^\infty}\le 1}} \varphi(n) = 1.$$

Hence $(\delta_n)_n \subset B(0,1)$; weak-* compactness follows by Banach-Alaoglu. (Rudin, Funct. Analysis Th. 3.17)

Theorem. If X is a separable Banach space, then any weakly-* compact $K\subset X$ is metrisable and hence weakly-* sequentially compact.

$C_c(\mathbb{R})$ is separable, so $(\delta_n)_n$ has a convergent subsequence (against $C_c(\mathbb{R})$).

b) Limit? For any $\varphi\in C_c(\mathbb{R})$:

$$\langle \delta_n, \varphi\rangle = \varphi(n) \xrightarrow[n\to\infty]{\text{ultimately}} 0 \quad \text{and so} \quad \delta_n \xrightarrow{*} 0.$$

But not in norm: $\|\delta_n - 0\|_{C_c^*} \equiv 1 \not\to 0$.
(Recall: strong convergence $\Rightarrow$ weak-* convergence & limits coincide)

c) weak-* compact in $C_b(\mathbb{R})^*$ "weak or narrow topology"

$$\|\delta_n\|_{C_b^*} = \sup_{\substack{\varphi\in C_b \\ \|\varphi\|_{L^\infty}\le 1}} \varphi(n) = 1 \Rightarrow \text{weakly-* compact by Banach-Alaoglu.}$$

However, $C_b(\mathbb{R})$ is not separable, and so we can not deduce weak-* sequential compactness!

L3, ex ①. $F: X \to \mathbb{R} \cup \{\infty\}$ convex & (norm) lsc $\Rightarrow$ weakly lsc.

short proof using the epigraph:

$$
\begin{aligned}
F \text{ convex \& lsc } &\Leftrightarrow \mathrm{epi}(F) \text{ convex \& closed} \\
&\Rightarrow \mathrm{epi}(F) \text{ convex \& weakly closed} \\
&\Leftrightarrow F \text{ convex \& weakly lsc}.
\end{aligned}
$$

Lecture 4

A Differentiation in Banach spaces

First consider finite dimensions: $F: \mathbb{R}^2 \to \mathbb{R}$.

$$\nabla F(x) := \begin{bmatrix} \frac{\partial}{\partial x_1} F(x) \\ \frac{\partial}{\partial x_2} F(x) \end{bmatrix}.$$

Why is it meaningful to know only the derivatives in the directions x_1 and x_2...? Well, if F differentiable, then

$$dF(x; \underbrace{h}_{\text{direction}}) := \lim_{\varepsilon \to 0} \frac{F(x+\varepsilon h) - F(x)}{\varepsilon} = \underbrace{\nabla F(x) \cdot}_{\text{linear operator acting on directions } h} h \quad \text{"directional derivative"}$$

Similarly in an (infinite-dim) Banach space: Peypouquet P.14

Def $F: X \to \mathbb{R}$ (or $\operatorname{dom} F \to \mathbb{R}$) is Gâteaux differentiable if

$$dF(x;h) := \lim_{\varepsilon \to 0} \frac{F(x+\varepsilon h) - F(x)}{\varepsilon}$$ is a linear & bounded functional on $h \in X$

We write $\ldots = {}_{X^*}\langle DF(x), h \rangle_X$, and $DF(x) \in X^*$ is called the Gâteaux derivative of F in x.

→ Peypouquet writes Gâteaux derivatives as ∇. There are good reasons not to do this...!

Def $F: X \to \mathbb{R}$ is twice Gâteaux differentiable if

$$d^2F(x; h_1, h_2) := \lim_{\varepsilon \to 0} \langle DF(x+\varepsilon h_2), h_1 \rangle - \langle DF(x), h_1 \rangle$$ is a linear & bounded functional on $h_1, h_2 \in X \times X$.

We write $\ldots = D^2F(x)[h_1, h_2]$ or $\cancel{{}_X\langle h_1, h_2 \rangle_{X, D^2F(x)}}$

Example (Euler-Lagrange, formal) (Peypouquet Example 1.28)

$$F(x) := \int_0^T \mathcal{L}(x(t), \dot{x}(t))\, dt \quad , X = L^2(0,T)$$

$$\lim_{\varepsilon \to 0} \frac{F(x+\varepsilon h) - F(x)}{\varepsilon} = \frac{d}{d\varepsilon} \int_0^T \mathcal{L}(x(t) + \varepsilon h(t), \dot{x}(t) + \varepsilon \dot{h}(t))\, dt \Big|_{\varepsilon \to 0}$$

$$= \int_0^T h(t)\, \partial_x \mathcal{L}(x(t) + \varepsilon h(t), \dot{x}(t) + \varepsilon \dot{h}(t))\, dt \Big|_{\varepsilon \to 0}$$

$$+ \int_0^T \dot{h}(t)\, \partial_{\dot{x}} \mathcal{L}(x(t) + \varepsilon h(t), \dot{x}(t) + \varepsilon \dot{h}(t))\, dt \Big|_{\varepsilon \to 0}$$

if h has compact support →

$$= \int_0^T h(t) \left[\partial_x \mathcal{L}(x(t), \dot{x}(t)) - \frac{d}{dt} \partial_{\dot{x}} \mathcal{L}(x(t), \dot{x}(t)) \right] dt$$

$$= \langle \underbrace{\partial_x \mathcal{L}(x, \dot{x}) - \frac{d}{dt} \partial_{\dot{x}} \mathcal{L}(x, \dot{x})}_{=: DF(x) \in L^2(0,T)}, h \rangle$$ (differentiable if this ...

Lecture 5

A Convexity & derivatives

Proposition. Let $A\,(=\operatorname{dom} F) \subset X$ be open and convex, (Peypouquet Prop 3.10) and $F: A \to \mathbb{R}$ be twice Gâteaux differentiable.

i) F is convex

$\iff$

ii) $F(y) \geq F(x) + \langle DF(x), y-x \rangle \quad \forall\, x,y \in A$ (gradient inequality)

$\iff$

iii) $\langle DF(x) - DF(y), x-y \rangle \geq 0 \quad \forall\, x,y \in A$ (monotonicity of DF)

$\iff$

iv) $D^2F(x) \underset{\substack{\uparrow \\ \text{pos. semidefinite}}}{\text{"}\geq\text{"}} 0$, i.e. $\forall x \in A,\ h \in X: D^2F(x)[h,h] \geq 0$.

proof

i)⇒ii)
$$F((1-\lambda)x + \lambda y) \leq (1-\lambda)F(x) + \lambda F(y)$$
$$\frac{F(x + \lambda(y-x)) - F(x)}{\lambda} \leq F(y) - F(x)$$
$$\downarrow \lambda \to 0$$
$$\langle DF(x), y-x \rangle \leq F(y) - F(x)$$

ii)⇒iii)
$$\langle DF(y), x-y \rangle \leq F(x) - F(y) \quad +$$
$$\langle DF(x) - DF(y), y-x \rangle \leq 0$$

iii)⇒i)
iv)⇒i) (not using twice differentiability! We don't really need this)

For any $x, y \in A$, define
$$\phi(\lambda) = F((1-\lambda)x + \lambda y) - (1-\lambda)F(x) - \lambda F(y), \quad \lambda \in [0,1]$$
$$\phi'(\lambda) = \langle DF((1-\lambda)x + \lambda y), y-x \rangle + F(x) - F(y)$$

If $0 < \lambda_1 < \lambda_2 < 1$ and $x_1 := (1-\lambda_1)x + \lambda_1 y$, $x_2 := (1-\lambda_2)x + \lambda_2 y$, then
$$\phi'(\lambda_2) - \phi'(\lambda_1) = \langle DF(x_2) - DF(x_1), y-x \rangle + F(x) - F(y) - F(x) + F(y)$$
$$= \langle DF(x_2) - DF(x_1), \frac{x_2 - x_1}{\underbrace{\lambda_2 - \lambda_1}_{\geq 0}} \rangle \geq 0 \text{ by assumption}$$

so ϕ' is non-decreasing.

iv)⇒i) If $\phi'' \geq 0$ then clearly ϕ' is non-increasing.

Since $\phi(0) = 0 = \phi(1)$, $\exists\, \bar{\lambda} \in (0,1)$ s.t. $\phi'(\bar{\lambda}) = 0$.

But ϕ' is non-decreasing, so $\phi'|_{[0,\bar{\lambda}]} \leq 0$ and $\phi'|_{[\bar{\lambda},1]} \geq 0$.

Hence $\phi(\lambda) \leq 0$.

iii)⇒iv)
$$\langle DF(x + \varepsilon h) - DF(x), \varepsilon h \rangle \geq 0$$
$$\downarrow \tfrac{\cdot}{\varepsilon^2},\ \varepsilon \to 0$$
$$D^2F(x)[h,h] \geq 0 \quad \square$$

$\boxed{\text{Proposition} \quad \text{strictly convex} \Leftrightarrow \text{strict inequalities}}$

Example (From lecture ①)

$X = \mathbb{R}^2, \quad F(x,y) = x^2 - \alpha xy + y^2$

$DF(x,y) = \nabla F(x,y) = \begin{bmatrix} 2x - \alpha y \\ -\alpha x + 2y \end{bmatrix}$

$D^2 F(x,y) = \nabla^2 F(x,y) = \begin{bmatrix} 2 & -\alpha \\ -\alpha & 2 \end{bmatrix}$

eigenvalues:

$$0 = \begin{vmatrix} 2-\lambda & -\alpha \\ -\alpha & 2-\lambda \end{vmatrix} = (2-\lambda)^2 - \alpha^2 \Rightarrow \lambda = 2 \pm \alpha$$

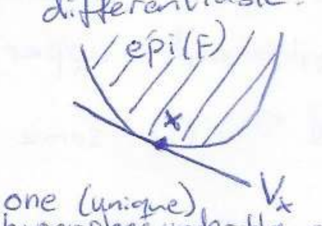
pos. semidef $\Leftrightarrow \alpha \in [-2,2]$.

$\rightarrow$ F is always convex in x and in y

$\rightarrow$ F only convex in (x,y) ("jointly convex") iff $\alpha \in [-2,2]$.

Cor The hyperplane $V_x := \{ (y,z) \in X \times \mathbb{R} : F(x) + \langle \nabla F(x), y-x \rangle = z \}$ ($x \in \operatorname{dom} F$ & F differentiable in x) lies below $\operatorname{epi}(F)$

differentiable:



one (unique) hyperplane under the epigraph

non-differentiable:

possibly many hyperplanes under the epigraph!

B Subdifferentials

Def For $F: X \to \mathbb{R} \cup \{\infty\}$ (proper), the subdifferential at $x \in X$ is the set:

(Peypouquet P. 41)

$$\partial F(x) := \{ \underset{\text{"slopes"}}{x^* \in X^*} : F(y) \geq \underbrace{F(x) + \langle x^*, y-x \rangle}_{\text{hyperplane below the (epi)graph}} \quad \forall y \in X \}$$

Example:

$F: \mathbb{R} \to \mathbb{R}$, $F(x) = x$

$$\partial F(x) = \begin{cases} \{-1\}, & x < 0, \\ [-1,1], & x = 0, \\ \{1\}, & x > 0. \end{cases}$$

(Peypouquet prop 3.20 & prop 3.52)

Prop. F is Gâteaux differentiable in $x \in X$, $\nabla F(x) = x^*$

$\iff$

$\partial F(x) = \{x^*\}$ (i.e. the subdifferential is a singleton)

proof "$\Rightarrow$" Clearly $\nabla F(x) \in \partial F(x)$; need to prove that this is the only one.

Take any $x^* \in \partial F(x)$, so that $\forall t > 0$, $h \in X$:

$$F(x+th) \geq F(x) + \langle x^*, th \rangle$$

$$\langle \nabla F(x), h \rangle \xleftarrow[t \to 0]{} \frac{F(x+th) - F(x)}{t} \geq \langle x^*, h \rangle \quad \forall h \in X.$$

Hence $\nabla F(x) = x^*$.

"$\Leftarrow$" Will be proven later... $\square$

Note that for the example $F(x) = |x|$, the subdiff. is always a closed interval. In higher (possibly infinite) dimensions, convex sets can be seen as generalisations of intervals...

(Peypouquet Prop 3.21)

Prop $F: X \to \mathbb{R} \cup \{\infty\}$ convex. For any $x \in X$, the subdiff. $\partial F(x)$ is (norm) closed and convex.

Proof

convex: Take $x_1^*, x_2^* \in \partial F(x)$ and $\lambda \in (0,1)$. For all $y \in X$:

$$F(y) \geq F(x) + \langle x_1^*, y-x \rangle \qquad \times \lambda$$
$$F(y) \geq F(x) + \langle x_2^*, y-x \rangle \qquad \times (1-\lambda) \quad +$$
$$F(y) \geq F(x) + \langle \lambda x_1^* + (1-\lambda) x_2^*, y-x \rangle, \text{ hence}$$
$$\lambda x_1^* + (1-\lambda) x_2^* \in \partial F(x).$$

closed: The norm topology is metric, so we only need to proof sequential closedness. Take $\partial F(x) \ni x_n^* \to x^* \overset{?}{\in} \partial F(x)$

$$F(y) \geq F(x) + \langle x_n^*, y-x \rangle$$
$$\downarrow n \to \infty$$
$$F(y) \geq F(x) + \langle x^*, y-x \rangle \qquad \blacksquare$$

(Peypouquet Prop. 3.22)

Prop (monotonicity) $F: X \to \mathbb{R} \cup \{\infty\}$. If $x^* \in \partial F(x)$, $y^* \in \partial F(y)$ then $\langle x^* - y^*, x-y \rangle \geq 0$

proof: exercise ☺

Th (Fermat's Rule) $F: X \to \mathbb{R} \cup \{\infty\}$ (proper and) convex. $\hat{x}$ is a global minimiser of F iff $0 \ni \partial F(\hat{x})$

proof: exercise ☺

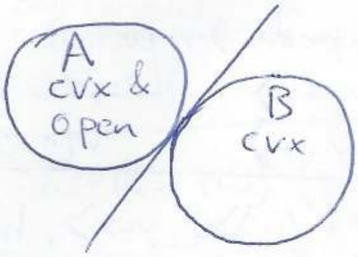
(Peypouquet 3.25)

Prop $F: X \to \mathbb{R} \cup \{\infty\}$ (proper and) convex and F cont. (lsc suffices) in $x \in \operatorname{dom}(F)$. Then $\partial F(x)$ is non-empty and bounded. (and closed and convex)

We need an alternative Hahn-Banach Theorem:

Theorem (Hahn-Banach geometric / separation Theorem)
$A, B \subset X$ disjoint, empty, convex & A is open.
Then $\exists\, x^* \in X^* \setminus \{0\}$ s.t.

$$\langle x^*, a\rangle < \langle x^*, b\rangle \quad \forall a \in A,\ b \in B$$



(Peypouquet Th. 1.10)
(Brézis Th. 1.6)

Proof of the prop. Take $A, B \subset X \times \mathbb{R}$:

ⓐ non-empty (open) $A := \mathrm{int}(\mathrm{epi}\, F) \neq \emptyset$ since F is cont. at at least one point

$B := \{(x, F(x))\}$ (and $(x, F(x)) \notin \mathrm{int}(\mathrm{epi}\, F)$ so $A \cap B = \emptyset$)

Hahn-Banach: $\exists\ (x^*, s) \in X \times \mathbb{R} \setminus \{0, 0\}$ s.t.

$$\underbrace{\langle (x^*, s), (y, C)\rangle}_{= \langle x^*, y\rangle + sC} < \underbrace{\langle (x^*, s), (x, F(x))\rangle}_{= \langle x^*, x\rangle + sF(x)} \quad \forall (y, C) \in \mathrm{int}(\mathrm{epi}(F))$$

For $y = x$,

$$\langle x^*, x\rangle + s\underbrace{C}_{\geq F(x)} < \langle x^*, x\rangle + sF(x) \qquad \text{so} \quad s \leq 0.$$

$$C > F(x) + \langle -\tfrac{x^*}{s}, y - x\rangle \quad \cancel{\tfrac{x^*}{s}}$$

$$\downarrow\ C \searrow F(y)$$

$$F(y) \geq F(x) + \langle -\tfrac{x^*}{s}, y - x\rangle \Rightarrow -\tfrac{x^*}{s} \in \partial F(x) \neq \emptyset.$$

ⓑ bounded. Take any $x^* \in \cancel{B}\, \partial F(x)$. Since F cont in x, F Lipschitz on nbh(x):

$$F(x) + \langle x^*, y - x\rangle \underset{\substack{\uparrow \\ x^* \in \partial F(x)}}{\leq} F(y) \underset{\substack{\uparrow \\ \text{Lipschitz}}}{\leq} F(x) + M\|y - x\| \quad \forall y \in \mathrm{nbh}(x)$$

Hence $\|x^*\| = \sup\limits_{y \neq x} \dfrac{\langle x^*, y - x\rangle}{\|y - x\|} = \sup\limits_{\substack{y \neq x \\ y \in \mathrm{nbh}(x)}} \dfrac{\langle x^*, y - x\rangle}{\|y - x\|} \leq M.$ □

Lecture 6 | A envelopes

$F : X \to \mathbb{R} \cup \{\infty\}$

If F is not convex, can it be "convexified"?
If F is not lsc, can it be "lsc-ified"?

For a set $A \subset X$:

$\bar{A}$:= closure of A := smallest closed set containing A

$\mathrm{co}\, A$:= convex hull of A := smallest convex set containing A

$= \{ \sum_{i=1}^{n} \sigma_i a_i : \sigma \in P(n) \text{ and } (a_i)_{i=1}^{n} \subset A,\ n \geq 1 \}$

For a functional $F : X \to \mathbb{R} \cup \{\infty\}$

$\bar{F}$ = lsc envelope of F := Largest lsc functional below F

well-defined since ($\bar{F}(x) =$) $\sup_{\substack{G \leq F \\ G \text{ lsc}}} G(x)$ is lsc (see below)

$\mathrm{co}F$ = convex envelope of F := Largest cvx functional below F

well-defined since ($\mathrm{co}F(x) =$) $\sup_{\substack{G \leq F \\ G \text{ cvx}}} G(x)$ is convex (see below)

Recall $\mathrm{epi}\, F := \{ (x, c) : F(x) \leq c \}$. Then (we already proved this!)

Prop
- $\sup_{\substack{G \leq F \\ G \text{ lsc}}} G(\cdot)$ is lsc
- $\sup_{\substack{G \leq F \\ G \text{ cvx}}} G(\cdot)$ is cvx

Proof: $\mathrm{epi}\left(\sup_{\substack{G \leq F \\ G \text{ lsc/cvx}}} G(\cdot) \right) = \bigcap_{\substack{G \leq F \\ G \text{ lsc/cvx}}} \mathrm{epi}\, G$ is $\overbrace{\text{closed / cvx}}$ □

General principle: sup over lsc & cvx functions is lsc & cvx.
Special role played by cont. affine functions:

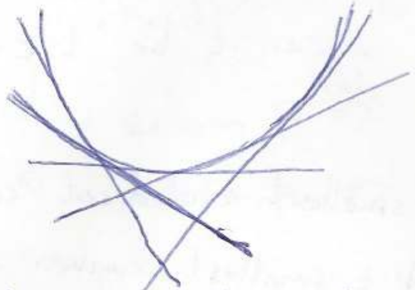
B affine functionals

Def A continuous affine functional $F : X \to \mathbb{R}\, (\cup \{\infty\})$ is a functional of the form

$$F(x) = \langle x^*, x \rangle + \alpha \quad \text{for some } x^* \in X^*,\ \alpha \in \mathbb{R}$$

Proposition. $F: X \to \mathbb{R} \cup \{\infty\}$ (proper). (Peypouquet prop 3.1)

F is convex & lsc $\iff$ $\exists$ family $(F_i)_{i \in I}$ of continuous affine functions such that $F(x) = \sup_{i \in I} F_i(x)$



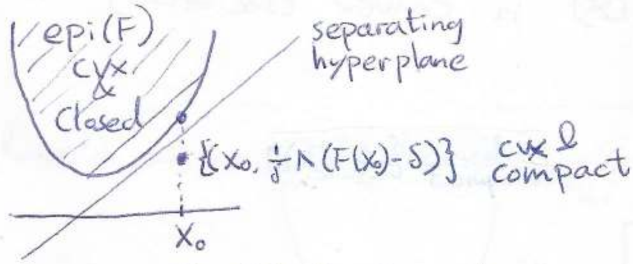
proof "$\Leftarrow$" we already proved this (epigraph)

"$\Rightarrow$" Let F be convex & lsc; Need to construct a family of continuous affine functions. We shall take $I = X \times \mathbb{R}_+$. Pick $x_0 \in X$, $\delta > 0$. We shall prove that ~~for any $\varepsilon > 0$~~ there exist a cont. affine $F_{x_0,\delta}: X \to \mathbb{R}$ such that (i) $F_{x_0,\delta}(x) \le F(x)$ $\forall x \in X$

(ii) $\frac{1}{\delta} \wedge (F(x_0) - \delta) \le F_{x_0,\delta}(x_0) \le F(x_0)$,

hence $F(x) = \sup_{x_0,\delta} F_{x_0,\delta}(x)$. How to construct $F_{x_0,\delta}$ ...?

Hahn-Banach:



$\exists$ $(x^*, s) \in X^* \times \mathbb{R} \setminus \{0,0\}$ and $\varepsilon > 0$ such that $\forall (x, C) \in \operatorname{epi} F$:

$$\langle x^*, x_0 \rangle + s\left(\tfrac{1}{\delta} \wedge (F(x_0) - \delta)\right) + \varepsilon = \langle (x^*, s), (x_0, \tfrac{1}{\delta} \wedge (F(x_0) - \delta)) \rangle + \varepsilon \le \langle (x^*, s), (x, C) \rangle = \langle x^*, x \rangle + sC \quad (*)$$

since C is arbitrarily large, $s \ge 0$. Two cases:

$\boxed{s > 0 \text{ (wlog assume } s = 1)}$

$$F_{x_0,\delta}(x) := \langle -x^*, x \rangle + \langle x^*, x_0 \rangle + \tfrac{1}{\delta} \wedge (F(x_0) - \delta) + \varepsilon.$$

From $(*)$ with $x \in \operatorname{dom} F$ and $C = F(x)$: (i) $F_{x_0,\delta}(x) \le F(x)$

(ii) $F_{x_0,\delta}(x_0) = F(x_0) - \delta + \varepsilon \ge (F(x_0) - \delta) \wedge \frac{1}{\delta}$

$\boxed{s = 0 \text{ (then } x_0 \notin \operatorname{dom} F)}$

$$G_{x_0,\delta}(x) := \langle -x^*, x \rangle + \langle x^*, x_0 \rangle + \varepsilon$$ and take $F_{\hat{x},\delta}(x)$ for some $\hat{x} \in \operatorname{dom} F$

$$F_{x_0,\delta}(x) := F_{\hat{x}_0,\delta}(x) + n_\delta G_{x_0,\delta}(x)$$ for some $n_\delta \xrightarrow[\delta \to 0]{} \infty$

From (*): $\forall x \in \operatorname{dom} F \quad G_{x_0,\delta}(x) = \langle -x^*, x\rangle + \langle x^*, \hat{x}\rangle + \varepsilon \le 0$

Hence (i) $\forall x \in \operatorname{dom} F \quad F_{x_0,\delta}(x) \le F_{\hat{x},\delta}(x) \le F(x)$

$\forall x \notin \operatorname{dom} F \quad F_{x_0,\delta}(x) \le \infty = F(x)$.

Moreover ii) $F_{x_0,\delta}(x_0) = F_{\hat{x},\delta}(x_0) + n_\delta \varepsilon \ge \frac{1}{\delta}$ for n_δ sufficiently large.

□

Of particular interest will be $F(x) = \sup_{x^*} \underbrace{\langle x^*, x\rangle - G(x^*)}_{\text{affine function in } x.}$

$\boxed{\text{C convex duals}}$

Recall the definition of the subdifferential:

$$\partial F(x) := \{ x^* \in X^* : \underbrace{F(y) \ge F(x) + \langle x^*, y-x\rangle}_{\text{this means that}} \ \forall y \in X \}$$

~~F(x) ... ⟨x*, x⟩ ≤~~

$$\langle x^*, x\rangle - F(x) \ge \sup_y \langle x^*, y\rangle - F(y)$$

Def Convex/Moreau/Fenchel dual/conjugate a.k.a Legendre transform (Peypouquet eq(3.17))

$F^* : X^* \to \mathbb{R} \cup \{\infty\}$

$$F^*(x^*) := \sup_{x \in X} \langle x^*, x\rangle - F(x) = \sup_{x \in \operatorname{dom} F} \langle x^*, x\rangle - F(x)$$

Examples

a) $F(x) = \frac{1}{p}\|x\|_{L^p}^p, \quad 1 < p < \infty. \qquad (X^* = L^{p^*}, \ \frac{1}{p} + \frac{1}{p^*} = 1)$

$F^*(x^*) = \sup_{x \in L^p} \langle x^*, x\rangle - \frac{1}{p}\|x\|_{L^p}^p$

Does a maximiser exist? Superlevel set: $\{ \langle x^*, x\rangle - \frac{1}{p}\|x\|_{L^p}^p \ge C \}$

$C \le \langle x^*, x\rangle - \frac{1}{p}\|x\|^p \le \|x^*\|\,\|x\| - \frac{1}{p}\|x\|^p = \|x\|\left(\|x^*\| - \frac{1}{p}\|x\|^{p-1}\right)$

if $\|x\|_{L^p} \le 1$ then $\|x\|_{L^p} \le 1$ (duh)

if $\|x\|_{L^p} \ge 1$ then

$$\frac{1}{p}\|x\|^{p-1} - \|x^*\|_{L^{p^*}} \le \frac{-C}{\|x\|} \le |C|$$

$$\|x\| \le \sqrt[p-1]{p(|C| + \|x^*\|)} \quad \text{hence bounded superlevel sets}$$

Banach-Alaoglu: weak-* cpt superlevel sets.

Hence also weak-* closed superlevel sets, i.e. upper semicontinuity.

Direct method: there exists a maximiser. (even unique by strict concavity)

Maximiser is a critical point. Gâteaux derivative:

$$\forall h \quad 0 = \lim_{\varepsilon \to 0} \frac{\langle x^*, x+\varepsilon h\rangle - \frac{1}{p}\|x+\varepsilon h\|^p - \langle x^*, x\rangle - \frac{1}{p}\|x\|^p}{\varepsilon}$$

$$= \frac{d}{d\varepsilon} \langle x^*, x+\varepsilon h\rangle - \frac{1}{p}\int (x+\varepsilon h)^p \Big|_{\varepsilon \to 0}$$

$$= \langle x^*, h\rangle - \int (x+\varepsilon h)^{p-1} h \Big|_{\varepsilon\to 0}$$

$$= \int (x^* - x^{p-1}) h$$

$$\frac{1}{p} + \frac{1}{p^*} = 1$$

$$\frac{1}{p^*} = 1 - \frac{1}{p} = \frac{p-1}{p}$$

$$p^* = \frac{p}{p-1} = 1 + \frac{1}{p-1}$$

$$\Rightarrow \quad x = \sqrt[p-1]{x^*} \quad \text{(assuming } x^* \geq 0 \text{ a.e.)}$$

$$F^*(x^*) = \sup_x \langle x^*, x\rangle - \frac{1}{p}\|x\|^p = \int |x^*|^{1+\frac{1}{p-1}} - \frac{1}{p}\int |x^*|^{\frac{p}{p-1}}$$

$$= \frac{1}{p^*}\int |x^*|^{p^*} = \frac{1}{p^*}\|x^*\|^{p^*}_{L^{p^*}}.$$

b) $F(x) = \langle y^*, x\rangle + \alpha$ (continuous affine) (Peypouquet example 3.46)

$$F^*(x^*) = \sup_x \langle x^*, x\rangle - \langle y^*, x\rangle - \alpha$$

$$= \sup_x \langle x^* - y^*, x\rangle - \alpha = \begin{cases} \infty, & x^* \neq y^* \\ -\alpha, & x^* = y^*. \end{cases}$$

c) $F(x) = \chi_C(x) := \begin{cases} \infty, & x \notin C \\ 0, & x \in C \end{cases}$ "characteristic function"

$$F^*(x^*) = \sup_x \langle x^*, x\rangle - \chi_C(x)$$ (Peypouquet example 3.47)

$$= \sup_{x\in C} \langle x^*, x\rangle \quad \text{"support function"}$$

Lecture 7

Answers to exercises

(3) $\boxed{\begin{array}{l}\underline{\underline{\text{Prop}}}\quad F: X \to \mathbb{R} \cup \{\infty\} \text{ (proper)} \\ F \text{ convex \& lsc} \iff \exists \text{ family } (F_i)_{i\in I} \text{ of cont. affine functions} \\ \qquad\qquad \text{such that } F(x) = \sup_{i \in I} F_i(x)\end{array}}$

We proved "$\Rightarrow$" by Hahn-Banach.
Can also be proven more directly!
We already know that convex & lsc implies nonempty subdifferential.
(also based on a Hahn-Banach theorem)

Hence $\quad F(x) = \sup_{\substack{x_0 \in X \\ x^* \in \partial F(x_0)}} F(x_0) + \langle x^*, x - x_0 \rangle .$

(4) $F: \mathbb{R} \to \mathbb{R}$ $\qquad F^*: \mathbb{R} \to \mathbb{R} \cup \{\infty\}$ "relative entropy"

$F(x) = b(e^x - 1) \Rightarrow F^*(x^*) = s(x^* | b) =$

$(b > 0)$

$$= b\lambda_B\left(\frac{x^*}{b}\right) = \begin{cases} b, & x = 0, \\ x \log \frac{x}{b} - x + b, & x > 0, \\ \infty, & x < 0, \end{cases}$$

"Boltzmann function"

$\Rightarrow F^{**}: \mathbb{R} \to \mathbb{R}.$

$F^{**}(x^{**}) = b(e^{x^{**}} - 1).$

Lecture 7 | A Properties of convex duals

Prop (properties of convex duals)

i) F^* is convex & lsc (in norm and even in weak-* topology)

ii) $F \le G \Rightarrow F^* \ge G^*$ (pointwise)

iii) $F^*(x^*) = \sup_{x \in X} \langle x^*, x\rangle - F(x) = \sup_{x \in \operatorname{dom} F} \langle x^*, x\rangle - F(x)$

iv) If $A: Z \to X$ is a linear bounded operator

$$\|A\| := \sup_{\substack{z \in Z \\ z \neq 0}} \frac{\|Az\|}{\|z\|} = \sup_{\substack{z \in Z \\ \|z\| \le 1}} \|Az\| < \infty$$

$${}_{X^*}\langle x^*, Az\rangle_X =: {}_{Z^*}\langle A^T x^*, z\rangle_Z$$

$A^T : X^* \to Z^*$ adjoint operator
(corresponds to transpose for matrices)
(More commonly denoted by A^*, but this gets confusing in convex analysis)

and $F(x) = \inf_{\substack{z \in Z: \\ Az = x}} H(z)$, $H : Z \to \mathbb{R} \cup \{\infty\}$

$\Rightarrow F^*(x^*) = H^*(A^T x^*)$.

v) $F(x) = G(\alpha x) \Rightarrow F^*(x^*) = G(\frac{x^*}{\alpha})$, $\alpha \in \mathbb{R} \setminus \{0\}$

vi) $F(x) = \alpha G(x) \Rightarrow F^*(x^*) = \alpha G(\frac{x^*}{\alpha})$, $\alpha \in \mathbb{R} \setminus \{0\}$

vii) $F(x) = G(x + x_0) \Rightarrow F^*(x^*) = G^*(x^*) - \langle x^*, x_0\rangle$

viii) $F(x) = G(x) - \langle y^*, x\rangle \Rightarrow F^*(x^*) = G^*(x^* + y^*)$

ix) $F(x) = G(x) + H(x) \Rightarrow F^*(x^*) = (G^* \square H^*)(x^*)$

$$:= \inf_{\substack{a^*, b^* \in X \\ a^* + b^* = x^*}} G^*(a^*) + H^*(b^*)$$

"inf-convolution" (sometimes denoted by $*_{\inf}$)

x) $F(x) + F^*(x^*) \ge \langle x^*, x\rangle$ Fenchel-Moreau-Young inequality

proofs: exercise (all but viii))

Remark: ix) is trivial but very important none-theless.
It generalises a "completing-the-squares" argument for quadratic functions (Young inequality)

$$\tfrac{1}{2}x^2 + \tfrac{1}{2}y^2 = \tfrac{1}{2}(x-y)^2 + xy \ge xy.$$

Corollary

$$F(x) + F^*(x^*) = \langle x^*, x\rangle \iff x \in \partial F^*(x^*)$$

proof $\forall y^*$ $F(x) + \underbrace{F^*(y^*) - \langle y^*, x\rangle}_{\text{convex in } y^*} \geq 0 = F(x) + F^*(x^*) - \langle x^*, x\rangle$ $\square$

(Alternatively, $F(x) + F^*(\cdot) - \langle \cdot, x\rangle$ is minimised by $x^* \Rightarrow$ Fermat's rule)

B The convex bidual

(Peypouquet p. 57)

Def For $F: X \to \mathbb{R} \cup \{\infty\}$ the bidual is
$$F^{**}: X \to \mathbb{R} \cup \{\infty\}$$
$$F^{**}(x) := \sup_{x^* \in X^*} \langle x^*, x\rangle - F^*(x^*).$$

(again clearly convex & lsc!)

$\to$ Why not defined on X^{**}? (Peypouquet remark 3.54)

$(F^*)^*(x^{**}) = \sup_{x^* \in X^*} \langle x^{**}, x^*\rangle - F^*(x^*)$

restricted to $X \subset X^{**}$:

$(F^*)^*|_X = F^{**}$.

Def (Canonical embedding)

For $x \in X$, let $\delta_x: X^* \to \mathbb{R}$ be the linear mapping
$$\langle x^*, \delta_x\rangle = \delta_x(x^*) := \langle x^*, x\rangle$$
Then $|\delta_x(x^*)| \leq \|x\| \|x^*\| \Rightarrow \delta_x$ bdd/cont.
$\Rightarrow \delta_x \in X^{**}$.

Hence, identifying x with δ_x:
$$X \subset X^{**}.$$

Prop $F: X \to \mathbb{R} \cup \{\infty\}$ proper.
$$F \text{ is convex \& lsc} \iff F = F^{**}$$

(Peypouquet prop. 3.56)

proof "$\Leftarrow$" F^{**} is a supremum over cont. affine functions.

"$\Rightarrow$" $F^{**}(x) = \sup_{x^*} \underbrace{\langle x^*, x\rangle - F^*(x^*)}_{\leq F(x) \text{ by Fenchel-Moreau-Young}} \leq F(x)$.

On the other hand, $F(x) = \sup_{i \in I} F_i(x)$ for family of cont. affine functions.

$F^*(x^*) = \sup_x \inf_i \langle x^*, x\rangle - F_i(x) \leq \inf_i \sup_x \langle x^*, x\rangle - F_i(x) = \inf_i F_i^*(x^*)$

$F^{**}(x) \geq \sup_{x^* \in X^*} \langle x^*, x\rangle - \inf_i F_i^*(x^*) = \sup_i \sup_{x^* \in X^*} \langle x^*, x\rangle - F_i^*(x^*)$

$= \sup_i F_i^{**}(x) = \sup_i F_i(x) = F(x)$. $\square$

$\boxed{\begin{array}{l} \text{Cor. } F: X \to \mathbb{R}\cup\{\infty\} \text{ proper.} \\ F^{**} = \mathrm{co}\,\overline{F} \end{array}}$ (Peypouquet cor. 3.57)

proof If $G \le F$ (pointwise) and G is convex & lsc, then

$$G = G^{**} \le F^{**}.$$

On the other hand, $F^{**} \le F$ and F^* convex & lsc.

$$F^{**} \le \sup_{\substack{(\text{pointwise}) \\ G\le F \\ G \text{ convex \& lsc}}} G = \mathrm{co}\,\overline{F} \le F^{**} \quad \square$$

$\boxed{\begin{array}{l} \text{Prop } F: X \to \mathbb{R}\cup\{\infty\} \text{ convex.} \\ F \text{ is Gâteaux differentiable in } x\in X,\ \nabla F(x) = x^* \\ \qquad\qquad \Longleftrightarrow \\ \partial F(x) = \{x^*\} \ \& \ F \text{ cont. in } x\ (\in \mathrm{int}(\mathrm{dom} F)) \end{array}}$ Prop 3.20 (Peypouquet prop 3.50)

"$\Rightarrow$" proven in lecture 5!!! Still needed to prove the other direction.

"$\Leftarrow$" $\phi_x : X \to \mathbb{R}\cup\{-\infty, \infty\}$

$$\phi_x(h) := \inf_{\varepsilon>0} \underbrace{\frac{F(x+\varepsilon h) - F(x)}{\varepsilon}}_{\substack{\text{non-decreasing in } \varepsilon \\ \text{by convexity}}} = \lim_{\varepsilon\to 0} \frac{F(x+\varepsilon h) - F(x)}{\varepsilon}$$

ϕ_x is convex in h by convexity of F.

F cont in $x \Rightarrow$ Lipschitz on $\mathrm{nbh}(x)$ (by the technical lemma):

$$-\mathrm{Lip}\,\|h\| \le \frac{F(x+\varepsilon h) - F(x)}{\varepsilon} \le \mathrm{Lip}\,\|h\|$$

$\Rightarrow |\phi_x(h)| < \infty \ \forall h$ ("$\mathrm{dom}\,\phi_x$" $= X$)

& ϕ_x bdd in a $\mathrm{nbh}(0)$

$\Rightarrow \phi_x$ cont. (& convex) on $\mathrm{int}(\mathrm{dom}\,\phi_x) = X$.

Now we know that

$$\phi_x(h) = \phi_x^{**}(h) = \sup_{y^*\in X^*} \langle y^*, h\rangle - \phi_x^*(y^*)$$

$$\boxed{\begin{aligned} \phi_x^*(y^*) &= \sup_h \sup_{\varepsilon>0} \langle y^*, h\rangle - \frac{F(x+\varepsilon h) - F(x)}{\varepsilon} \\ &= \sup_{\varepsilon>0} \underbrace{\frac{F(x) + F^*(y^*) - \langle y^*, x\rangle}{\varepsilon}}_{\substack{\text{Fenchel-Moreau-Young:} \\ =0 \text{ if } y^*\in\partial F(x) \\ >0 \text{ otherwise}}} \\ &= \chi_{\partial F(x)}(y^*) = \chi_{\{x^*\}}(y^*) \end{aligned}}$$

$$= \sup_{y^*\in Y^*} \langle y^*, h\rangle - \chi_{\{x^*\}}(y^*) = \langle x^*, h\rangle \quad \square$$

Lecture 8 | [A] Strict convexity & differentiability of the dual

Recall that a convex $F: X \to \mathbb{R} \cup \{\infty\}$:

- is Gâteaux differentiable at x, $DF(x) = x^*$ $\Leftrightarrow$ $\partial F(x) = \{x^*\}$ and F is cont. at x.
- is cont. in $x \Rightarrow \partial F(x) \neq \emptyset$ and $x \in \mathrm{int}(\mathrm{dom}\, F)$
- is lsc $\Rightarrow$ cont. on $\mathrm{int}(\mathrm{dom}\, F)$

Th Let $F: X \to \mathbb{R}$ lsc & convex (so $\mathrm{dom} F = X$ & ∂F nowhere empty)

F is strictly convex $\Leftrightarrow F^*$ is (everywhere) Gâteaux diff.

(Rockafellar, Convex Analysis 1970, Th. 26.3)

→ A bit more complicated if $\mathrm{dom} F \subsetneq X$. Has to do with the possibility that $\{\partial F \neq \emptyset\}$ may not be convex.

proof

"$\Leftarrow$" Suppose F is not strictly convex:

$$\exists\, x_1 \neq x_2,\ \lambda \in (0,1) \text{ s.t. } F(\underbrace{(1-\lambda)x_1 + \lambda x_2}_{=:x}) = (1-\lambda)F(x_1) + \lambda F(x_2). \quad (*)$$

Take $x^* \in \partial F(x)$, hence:

$$F(y) \geq F(x) + \langle x^*, y - x \rangle.$$

In particular:

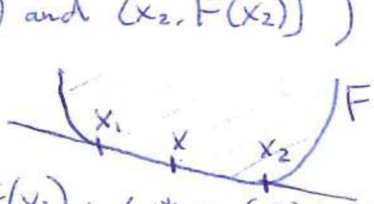
$$F(x_1) \overset{(*)}{\geq} (1-\lambda)F(x_1) + \lambda F(x_2) + \lambda \langle x^*, x_1 - x_2 \rangle$$
$$F(x_2) \overset{(*)}{\geq} (1-\lambda)F(x_1) + \lambda F(x_2) + (1-\lambda)\langle x^*, x_2 - x_1 \rangle$$

and so

$$F(x_1) \geq F(x_2) + \langle x^*, x_1 - x_2 \rangle \geq F(x_1) \qquad (**)$$
$$F(x_2) \geq F(x_1) + \langle x^*, x_2 - x_1 \rangle \geq F(x_2). \qquad (***)$$

(This means that the supporting hyperplane of epi F at $(x, F(x))$ passes through $(x_1, F(x_1))$ and $(x_2, F(x_2))$)

We can then rewrite, for any $y \in X$:



$$F(y) \geq F(x) + \langle x^*, y - x \rangle \overset{(*)}{=} (1-\lambda)F(x_1) + \lambda F(x_2) + \langle x^*, y - (1-\lambda)x_1 - \lambda x_2 \rangle$$

$$\begin{cases} \overset{(**)}{=} F(x_2) + \langle x^*, y - x_2 \rangle \\ \overset{(***)}{=} F(x_1) + \langle x^*, y - x_1 \rangle. \end{cases}$$

It follows that $x^* \in \partial F(x_1)$ and $x^* \in \partial F(x_2)$.
But then $x_1, x_2 \in \partial F^*(x^*)$, and so F^* can not be differentiable!

"$\Rightarrow$" (by a similar argument) Suppose F^* is not differentiable at some $x^* \in X^*$; then $x^* \in \partial F(x_1) \cap \partial F(x_2)$ for some $x_1 \neq x_2 \in X$.
For arbitrary $\lambda \in (0,1)$:

$$\begin{array}{ll} F(y) \geq F(x_1) + \langle x^*, y - x_1 \rangle & \times (1-\lambda) \\ F(y) \geq F(x_2) + \langle x^*, y - x_2 \rangle & \times \lambda \quad + \\ \hline F(x) \geq (1-\lambda) F(x_1) + \lambda F(x_2) & \end{array}$$

$(x = (1-\lambda)x_1 + \lambda x_2)$

Hence by convexity $F((1-\lambda)x_1 + \lambda x_2) = (1-\lambda)F(x_1) + \lambda F(x_2)$, ~~and~~ so F is not strictly convex. $\square$

B Inf-convolutions and Moreau-Yosida regularisation

"(Usual) convolution"	"Inf-convolution"
• $(F * G)(x) := \int F(x-z) G(z) dz = \int F(z) G(x-z) dz$	• $(F \square G)(x) = \inf_z F(x-z) + G(z) = \inf_z F(z) + G(x-z) = \inf_{a+b=x} F(a) + G(b)$
• $(F * \delta_0)(x) = F(x)$	• $(F \square \chi_0)(x) = F(x)$
• Typical smoothing kernel: $\theta_\varepsilon(x) = \frac{1}{\sqrt{(4\pi\varepsilon)^d}} e^{-\frac{\lvert x \rvert^2}{4\varepsilon}}$ $\theta_\varepsilon \xrightarrow{*} \delta_0$ (as measures, weakly-* against $C_b(\mathbb{R}^d)$)	• Typical smoothing kernel: $\theta_\varepsilon(x) = \frac{1}{2\varepsilon} \lVert x \rVert^2$ $\theta_\varepsilon \rightsquigarrow \chi_0$
• Regularising/smoothing effect: $(F * \theta_\varepsilon) \in C_b^\infty(\mathbb{R}^d)$	• Regularising/smoothing effect: $F \square \theta_\varepsilon$ convex & differentiable, $\frac{1}{\varepsilon}$-Lipschitz-cont. derivative
• Approximation: $F * \theta_\varepsilon \longrightarrow F$ (e.g. in L^1 if $F \in L^1$)	• Approximation: $F \square \theta_\varepsilon \longrightarrow F$ (pointwise)

$$F_\varepsilon(x) := \inf_{z\in X} F(z) + \frac{1}{2\varepsilon}\|x-z\|^2 \qquad (*)$$

Moreau-Yosida Regularization

In the following we shall take $X = H$ Hilbert space.
(Some results are generalizable to Banach, or even metric spaces)
(Peypouquet prop 3.35)

Prop. $F: H \to \mathbb{R}\cup\{\infty\}$ proper, cvx, lsc. For any $x\in H$, $\varepsilon>0$, $(*)$ has a unique minimiser $J_\varepsilon(x)$. Moreover,

$$-\frac{J_\varepsilon(x)-x}{\varepsilon} \in \partial F(J_\varepsilon(x))$$

Proof. F is proper, cvx, lsc $\Rightarrow$ $F(x) = \sup_{i\in I} F_i(x)$ for some family of cont. affine functions. Hence for the level sets of $F+\frac{1}{2\varepsilon}\|\cdot\|^2$:

$$-\|x_i^*\|\|x\| - |\alpha_i| + \frac{1}{2\varepsilon}\|x\|^2 \underset{(\text{any } i\in I)}{\leq} \sup_{i\in I} F_i(x) + \frac{1}{2\varepsilon}\|x\|^2 = F(x) + \frac{1}{2\varepsilon}\|x\|^2 \leq C \quad (**)$$

(if $F_i(x) = \langle x_i^*, x\rangle + \alpha_i$).

Bdd level sets $\xRightarrow{\text{Banach-Alaoglu}}$ weak-* compact level sets.
(Recall that H has a predual, namely H itself).

Norms are always weak-* lsc
F lsc & convex $\Rightarrow$ F weak seq. lsc $\Leftrightarrow$ weak-* seq. lsc (Hilbert space!)

$F+\frac{1}{2\varepsilon}\|\cdot\|^2$ has weak-* compact level sets, is weak-* seq. lsc, and bounded from below by $(**)$, and so the minimiser exists.
By strict convexity, the minimiser must be unique.
The optimality equation follows from Fermat's rule $\square$

Remarks

i) Since there exists a unique solution we can write
$$J_\varepsilon(x) = (I+\varepsilon\partial F)^{-1}(x).$$
This is similar to the "resolvent" of an operator $Q \approx \partial F$ which is studied in the Hille-Yosida Theorem to prove existence of a semigroup s.t. $\dot{P}_t = QP_t$.

ii) Clearly
$$\inf F \leq F_\varepsilon(x) \leq F(x),$$
And so
- $\inf F = \inf F_\varepsilon$
- $\inf F = F(\hat{x}) \Leftrightarrow \inf F_\varepsilon = F_\varepsilon(\hat{x}) \ \forall \varepsilon>0.$

(Peypouquet prop 3.38)

Lemma $F: H \to \mathbb{R} \cup \{\infty\}$ proper, lsc & cvx, then $J_\varepsilon: H \to H$ is a contraction, i.e. 1-Lipschitz continuous

proof For any $x, y \in H$

$$-\frac{J_\varepsilon(x)-x}{\varepsilon} \in \partial F(J_\varepsilon(x)) \text{ and } -\frac{J_\varepsilon(y)-y}{\varepsilon} \in \partial F(J_\varepsilon(y)).$$

By monotonicity of ∂F:

$$\left\langle -\frac{J_\varepsilon(x)-x}{\varepsilon} + \frac{J_\varepsilon(y)-y}{\varepsilon}, x-y \right\rangle \geq 0.$$

Hence

$$0 \leq \|J_\varepsilon(x) - J_\varepsilon(y)\|^2 \leq \langle J_\varepsilon(x) - J_\varepsilon(y), x-y \rangle \leq \|J_\varepsilon(x) - J_\varepsilon(y)\| \|x-y\|$$

and so $\|J_\varepsilon(x) - J_\varepsilon(y)\| \leq \|x-y\|$.

(Peypouquet prop. 3.39)

Prop $F: H \to \mathbb{R} \cup \{\infty\}$ proper, lsc & cvx, $\varepsilon > 0$, then F_ε is Gâteaux (even Fréchet) differentiable, and convex, and

$$DF_\varepsilon(x) = \frac{1}{\varepsilon}(x - J_\varepsilon(x)),$$

(and DF_ε is Lipschitz continuous with constant $\frac{2}{\varepsilon}$).

Proof For any $x, y \in H$

$$F_\varepsilon(y) - F_\varepsilon(x) = F_\varepsilon(J_\varepsilon(y)) - F_\varepsilon(J_\varepsilon(x)) + \frac{1}{2\varepsilon}\left(\|J_\varepsilon(y) - y\|^2 - \|J_\varepsilon(x) - x\|^2\right)$$

$$\geq \left\langle -\frac{J_\varepsilon(x)-x}{\varepsilon}, J_\varepsilon(y) - J_\varepsilon(x) \right\rangle + \frac{1}{2\varepsilon}\Big(\|J_\varepsilon(y) - y - J_\varepsilon(x) + x\|^2 + 2\langle J_\varepsilon(x) - x, J_\varepsilon(y) - y - J_\varepsilon(x) + x\rangle\Big)$$

(by definition of the subdifferential)

$$\geq \frac{1}{\varepsilon}\langle x - J_\varepsilon(x), y - x \rangle$$

Similarly

$$F_\varepsilon(x) - F_\varepsilon(y) \geq \frac{1}{\varepsilon}\langle y - J_\varepsilon(y), x-y \rangle.$$

Setting $y = x + \tau h$ for $\tau > 0$ and an arbitrary direction $h \in H$:

$$\frac{1}{\varepsilon\tau}\langle x - J_\varepsilon(x), \tau h \rangle \leq \frac{F_\varepsilon(x+\tau h) - F_\varepsilon(x)}{\tau} \leq \frac{1}{\varepsilon\tau}\langle x + \tau h - J_\varepsilon(x+\tau h), \tau h \rangle$$

Hence, by the Lipschitz continuity of J_ε:

$$\lim_{\tau \to 0} \frac{F_\varepsilon(x+\tau h) - F_\varepsilon(x)}{\tau} = \frac{1}{\varepsilon}\langle x - J_\varepsilon(x), h \rangle. \Rightarrow DF_\varepsilon(x) = \frac{1}{\varepsilon}(x - J_\varepsilon(x))$$

Note that $DF_\varepsilon(x) \in \partial F_\varepsilon(J_\varepsilon(x))$. Therefore

$$\langle DF_\varepsilon(y) - DF_\varepsilon(x), y-x\rangle = \left\langle \frac{y - J_\varepsilon(y) - x + J_\varepsilon(x)}{\varepsilon}, y - J_\varepsilon(y) - x + J_\varepsilon(x)\right\rangle$$

$$+ \langle y - J_\varepsilon(y) - x + J_\varepsilon(x), J_\varepsilon(y) - J_\varepsilon(x)\rangle \geq 0,$$

and so F_ε is convex. $\square$

(F proper, lsc & cvx)

$$\boxed{\text{Prop} \quad \lim_{\varepsilon \to 0} F_\varepsilon(x) = F(x) \quad \forall x \in H}$$ (Peypouquet prop. 3.41)

proof (For fixed $x \in \text{dom} F$)

$$F_\varepsilon(x) = F(J_\varepsilon(x)) + \frac{1}{2\varepsilon}\|J_\varepsilon(x) - x\|^2 \leq F(x)$$

$\leq$

$F_i(J_\varepsilon(x)) + \frac{1}{2\varepsilon}\|J_\varepsilon(x) - x\|^2$ for some cont. affine $F_i : H \to \mathbb{R}$.

Hence $J_\varepsilon(x)$ is ~~uniformly~~ bounded as $\varepsilon \to 0$.

Therefore $F_i(J_\varepsilon(x))$ $\cancel{\left(+\frac{1}{2\varepsilon}\|J_\varepsilon(x)-x\|^2\right)}$ is bounded from below,

and so is $F(J_\varepsilon(x))$. Then:

$$\underbrace{F(J_\varepsilon(x))}_{\text{bdd below}} + \underbrace{\frac{1}{2\varepsilon}\|J_\varepsilon(x) - x\|^2}_{\text{bdd}} \leq \underbrace{F(x)}_{\text{indep. of } \varepsilon}$$

and so $\|J_\varepsilon(x) - x\| \to 0$.

By (norm-sequential) lower semicontinuity:

$$F(x) \leq \liminf_{\varepsilon \to 0} F(J_\varepsilon(x)) \leq \limsup_{\varepsilon \to 0} F(J_\varepsilon(x)) \leq F(x). \quad \square$$

Lecture 9

A) Jensen's inequality

If $F: X \to \mathbb{R} \cup \{\infty\}$ convex then also (for any finite convex combination)

$$F\left(\sum_{i=1}^{n} \sigma_i x_i\right) \leq \sum_{i=1}^{n} \sigma_i F(x_i) \quad \forall (x_i)_{i=1}^{n} \subset X, \ \underbrace{\sigma \in \mathcal{P}(\{1, \dots, n\})}_{(\sigma_i)_{i=1}^{n} \subset [0,1], \ \sum_{i=1}^{n} \sigma_i = 1}.$$

How far can we push this? For example, if we use a "Schauder basis" and assume F is lsc, then

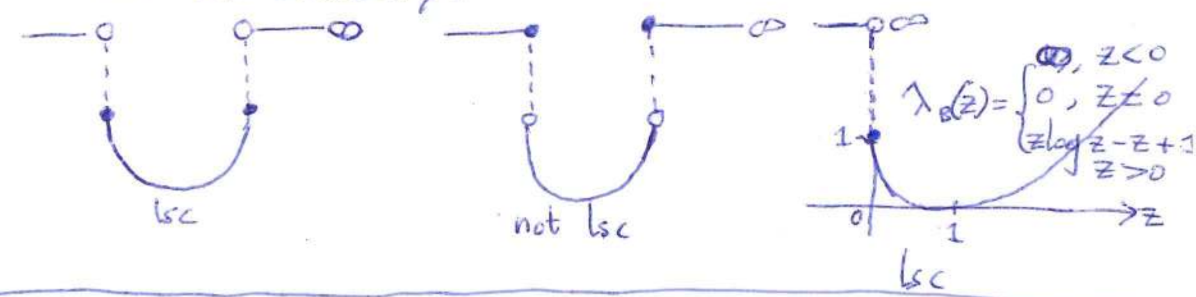
$$\sum_{i=1}^{n} \sigma_i x_i \xrightarrow[n\to\infty]{X} x$$

$$\Longrightarrow$$

$$F(x) \leq \liminf_n F\left(\sum_{i=1}^{n} \sigma_i x_i\right) \leq \liminf_n \sum_{i=1}^{n} \sigma_i F(x_i).$$

This again shows that lsc will be important!

How about integrals against a probability measure?
We will now restrict to real-valued (one-dimensional) convex functions. These are continuous on the interior of their (convex) domain. Lsc is needed only to control behaviour at the boundary:



Th (Jensen's inequality). Let $(\Omega, \mathcal{A}, \sigma)$ be a probability space, $F: \Omega \to \mathbb{R} \cup \{\infty\}$ be convex and lsc, and $x \in L^1_\sigma(\Omega)$. Then:

i) $F\left(\int x \, d\sigma\right) \leq \int F \circ x \, d\sigma$

ii) Moreover, if F is strictly convex, then

$$F\left(\int x \, d\sigma\right) = \int F \circ x \, d\sigma \iff \underbrace{\sigma(\{\omega \in \Omega : x(\omega) = q\}) = 1}_{\text{Prob}(x=q)=1} \quad (\text{for some } q \in \mathbb{R})$$

→ Can be found anywhere in the literature (not in Peypouquet), e.g.:

Bogachev - Measure Theory (2006) Th. I.2.12.19

Lieb & Loss - Analysis (2001) Th. 2.2

Evans - PDEs (1997) Th. B.1.2

However, not ~~ever~~ always very precise about lsc, or point (ii)...

→ A generalisation to Banach spaces would require "Bochner-integrals", i.e. Banach-valued integrals.

→ It is essential that $\sigma(\Omega)=1$. One often needs to rescale, for example $\int_{U\subset\mathbb{R}^d} \dots dx \to \int_U \dots \frac{dx}{|U|}$. Of course, this is impossible if $\sigma(\Omega)=\infty$, for example $\int_{\mathbb{R}^d} \dots dx$.

Proof (F∘x is measurable)

(i) $F(\int x\, d\sigma) = \sup_{i\in I} F_i(\int x\, d\sigma)$ for some family of cont. affine functions, since F is cvx & lsc.

$= \sup_{i\in I} \int F_i \circ x\, d\sigma$ since integrals are linear and $\sigma(\Omega)=1$

$\le \int \sup_{i\in I} F_i\circ x\, d\sigma$

$= \int F\circ x\, d\sigma.$

(ii) Assume σ is not "deterministic", i.e. $\sigma(\{\omega\in\Omega : x(\omega)=q\})<1\ \forall q\in\mathbb{R}$.

W.l.o.g. we may consider $\rho\in P(\mathbb{R})$, $\rho(A)=\sigma(x^{-1}(A))$
$=\sigma(\{\omega\in\Omega : x(\omega)\in A\})$.

The assumption means that $\rho\neq\delta_q\ \forall q\in\mathbb{R}$.

Hence there are two disjoint sets $A_1\cup A_2=\mathbb{R}$ such that $\rho(A_1)>0$ and $\rho(A_2)>0$. Then $\Omega_1=x^{-1}(A_1)$, $\Omega_2=x^{-1}(A_2)$ are also disjoint and $\Omega_1\cup\Omega_2=\Omega$. Now:

$$F\left(\int x\,d\sigma\right) = F\left(\sigma(\Omega_1)\int_{\Omega_1} x\frac{d\sigma}{\sigma(\Omega_1)} + \sigma(\Omega_2)\int_{\Omega_2} x\frac{d\sigma}{\sigma(\Omega_2)}\right)$$

$$\underset{\text{(strict convexity)}}{<} \sigma(\Omega_1)F\left(\int_{\Omega_1} x\frac{d\sigma}{\sigma(\Omega_1)}\right) + \sigma(\Omega_2)F\left(\int_{\Omega_2} x\frac{d\sigma}{\sigma(\Omega_2)}\right)$$

$$\underset{\text{(Jensen ineq.)}}{\le} \frac{\sigma(\Omega_1)}{\sigma(\Omega_1)}\int_{\Omega_1} F(x)\,d\sigma + \frac{\sigma(\Omega_2)}{\sigma(\Omega_2)}\int_{\Omega_2} F(x)\,d\sigma$$

$$= \int F(x)\,d\sigma \quad \square$$

→ Point (ii) is very important! It shows that Jensen's ineq. can be very powerfull when considering a sequence of probability measures that converge to a delta measure!

B Jensen and coercivity

Cor If $F(x) = \int f(|x(\omega)|)\mu(d\omega)$, $f: \mathbb{R} \to \mathbb{R} \cup \{\infty\}$ cvx & lsc and $(\Omega, \mathcal{A}, \mu)$ is a probability space, then F has $L^1_\mu(\Omega)$-uniformly bounded level sets.

Of course, a uniform L^1_μ-bound is usually not very helpfull since L^1_μ doesn't have a predual, and so we can not deduce weak-$*$ compactness. However, one can often do better:

Cor Let $(\Omega, \mathcal{A}, \mu)$ be a probability space and $p \geq 1$ and $x \in L^p_\mu(\Omega)$. Then for any $r \in (0, p]$:

$$\left(\int |x|^r d\mu\right)^{1/r} \leq \left(\int |x|^p d\mu\right)^{1/p}$$

(Bogachev - Measure Theory 2007. Cor I.2.12.21)

(Bogachev Th. I.2.12.24)

Th (Pinsker-Kullback-Csiszár). Let μ, ν be two probability measures on a measurable space $(\Omega, \mathcal{A})$, and assume $\frac{d\nu}{d\mu} > 0$.

$$\|\mu - \nu\|_{TV}^2 := \left(\int \left|\frac{d\nu}{d\mu} - 1\right| d\mu\right)^2 \leq 2 \underbrace{\int \left(\log \frac{d\nu}{d\mu}\right) d\nu}_{= \int \lambda_B\left(\frac{d\nu}{d\mu}\right) d\mu \text{ (relative entropy)}}$$

$\to$ In fact, we will see that $\int f(|x|) d\mu$ is related to the norm of so-called Orlicz-spaces. When estimating the L^1_μ-norm we throw away a lot of information that would otherwise be usefull for analysis in an Orlicz space!

C Jensen and convergence: an example

Let $(\Omega, \mathcal{A}, \sigma) = (\mathbb{R}^d, \text{Borel}, \mathcal{L}|_{[0,1]^d})$.

$$F(x) := \int_{[0,1]^d} \lambda_B\left(\frac{dx}{dq}\right) dx = \int_{(0,1)^d} (x(q)\log x(q) - x(q) + 1)\, dq$$

∞ if not $x \ll \mathcal{L}_{[0,1]^d}$, so we may assume $x \in L^1((0,1)^d)$.

Smoothing: $\theta_\varepsilon(q) = \frac{1}{\sqrt{(4\pi\varepsilon)^d}} e^{-\frac{|q|^2}{4\varepsilon}}$, $x_\varepsilon := x * \theta_\varepsilon \in C_b^\infty(\mathbb{R}^d)$.

Can we prove that $F(x_\varepsilon) \to F(x)$?
Very difficult to find a majorant for dominated convergence...
However we can exploit convexity!

$\boxed{\text{Lemma: If } x \in L^p \Rightarrow x * \theta_\varepsilon \xrightarrow{L^p} x, \quad 1 \le p < \infty}$ (Evans - PDEs 2010 Th. C.4.6)

$\boxed{\text{Prop: If } x \in L^1((0,1)^d) \text{ and } F(x) < \infty \text{ then } F(x_\varepsilon) \to F(x)}$

proof On the one hand,

$$F(x_\varepsilon) = \int_{(0,1)^d} \lambda_B\left(\int x(q)\theta_\varepsilon(q-z)\, dq\right) dz$$

$$\overset{\text{Jensen}}{\le} \int_{(0,1)^d} \int \lambda_B(x(q))\, \theta_\varepsilon(q-z)\, dq\, dz$$

$$\xrightarrow{\varepsilon \to 0} \int_{(0,1)^d} \lambda_B(x(q))\, dq = F(x), \text{ since } \lambda_B \circ x \in L^1$$

hence $(\lambda_B \circ x) * \theta_\varepsilon \xrightarrow{L^1} \lambda_B \circ x$.

On the other hand,

$$F(y) = \int_{(0,1)^d} \sup_{x^* \in \mathbb{R}} \Big(x^* y(q) - \underbrace{(e^{x^*} - 1)}_{\lambda_B^*(x^*)}\Big)\, dq$$

$$= \sup_{x^* \in \text{Measurable}((0,1)^d)} \int_{(0,1)^d} \left(x^*(q) y(q) - (e^{x^*(q)} - 1)\right) dq$$

$$\overset{\substack{\text{(monotone or}\\ \text{dominated convergence)}}}{=} \sup_{x^* \in L^\infty((0,1)^d)} \underbrace{\int_{(0,1)^d} \left(x^*(q) y(q) - (e^{x^*(q)} - 1)\right) dq}_{L^1\text{-continuous functional}}$$

Hence F is L^1-lsc (topologically & sequentially) and so

$$F(x) \le \liminf_{\varepsilon \to 0} F(x_\varepsilon) \le \limsup_{\varepsilon \to 0} F(x_\varepsilon) \le F(x) \qquad \square$$

10 A Smoothing as gradient flow

$$F(x) := S(x \mid \mathcal{L}|_{(0,1)^d}) = \begin{cases} \int \lambda_B\left(\frac{dx}{d\mathcal{L}|_{(0,1)^d}}(q)\right)\mathcal{L}|_{(0,1)^d}(dq), & x \ll \mathcal{L}|_{(0,1)^d}, \\ \infty, & \text{otherwise.} \end{cases}$$

"Relative entropy" of measure x w.r.t. measure $\mathcal{L}|_{(0,1)^d}$.

Last week we proved by Jensen that

$$F(x * \theta_t) \leq F(x), \quad \text{for } x \in L^1((0,1)^d).$$

There's another reason why this is true...

Instead of the heat kernel, let θ_ε be the solution of

$$\begin{cases} \dot{\theta}_t = \Delta \theta_t, & \text{on } (0,1)^d,\ t>0, \\ \frac{\partial \theta_t}{\partial n} = 0, & \text{on } \partial(0,1)^d,\ t>0, \quad \text{(Neumann BC)} \\ \theta_0 = \delta_0, & t=0. \end{cases}$$

Then for $x_t := x * \theta_t$ also:

$$\begin{cases} \dot{x}_t = \Delta x_t, & \text{on } (0,1)^d,\ t>0, \\ \frac{\partial x_t}{\partial n} = 0, & \text{on } \partial(0,1)^d,\ t>0, \\ x_0 = x, & t=0. \end{cases}$$

At least formally,

$$\langle DF(x), h\rangle = \lim_{\tau \to 0} \frac{F(x+\tau h) - F(x)}{h} = \int_{(0,1)^d} h \log x.$$

Therefore the evolution of x_t can also be written as

$$\dot{x}_t = \underbrace{\operatorname{div}(x_t \nabla DF(x_t))}_{=: -\operatorname{grad}_{x_t} F(x_t)} = \operatorname{div}(x_t \nabla \log x_t) = \operatorname{div}\left(x_t \frac{\nabla x_t}{x_t}\right) = \Delta x_t.$$

Hence x_t is the <u>gradient flow</u> of F on some strange manifold.

Therefore (this is always true for a gradient flow):

$$\begin{aligned} \frac{d}{dt} F(x_t) &= \langle DF(x_t), \dot{x}_t\rangle = \langle DF(x_t), \operatorname{div}(x_t \nabla DF(x_t))\rangle \\ &= -\langle \nabla DF(x_t), x_t \nabla DF(x_t)\rangle + \int_{\partial(0,1)^d} DF(x_t) \underbrace{x_t \nabla DF(x_t) \cdot n}_{= \nabla x_t \cdot n = 0} \\ &= -\underbrace{\|\nabla DF(x_t)\|^2_{L^2(x_t)}}_{= \int_{(0,1)^d} \frac{|\nabla x_t|^2}{x_t}} \leq 0 \quad \text{"Fisher information"} \end{aligned}$$

[B] Lower semicontinuity of Lagrangian actions

Previous lecture: strategy how to prove lsc of $F(x) = \int f(x(q))dq$, for some convex & lsc f.

Now: how to prove lsc of $F(x) = \int L(\nabla x(q), x(q), q)dq$?

(Evans - PDE's Th. 8.2.2.1)

Th. $\Omega \subset \mathbb{R}^n$ open bounded with smooth boundary, $1 < p < \infty$. Assume that $L : \mathbb{R}^n \times \mathbb{R} \times \bar{\Omega} \to \mathbb{R}$ is bdd from below and convex in the first argument. Then F is sequentially weakly lsc in $W^{1,p}(\Omega)$.

$\to$ $\|x\|_{W^{1,p}} = \|x\|_{L^p} + \|\nabla x\|_{L^p}$.

$W^{1,p}$ reflexive, so weak = weak-$*$.

We first need a number of results: (Brezis Th. II.1)

Th (Banach-Steinhaus / Uniform Boundedness Principle) Any weakly (or weakly-$*$) convergent sequence $x_k \to x$ in a Banach space X is bounded: $\sup_k \|x_k\| < \infty$.

(Evans, Th. 5.7.1)

Th (Rellich-Kondrachov) For all $1 \le p \le \infty$, $W^{1,p}(\Omega) \subset\subset L^p(\Omega)$, i.e. i) $\|x\|_{L^p} \le C\|x\|_{W^{1,p}}$ (duh!)

"compact embedding" ii) $\sup_k \|x_k\|_{W^{1,p}} < \infty \Rightarrow x_k$ rel. cpt in L^p.

(Brezis Th. IV.9)

Th ("converse dominated convergence"). If $x_k \xrightarrow{L^p} x$, $(1 \le p \le \infty)$ then $\exists$ subsequence x_{k_n} such that i) $x_{k_n}(q) \to x(q)$ a.e.

ii) $\exists\, h \in L^p$ with $|x_{k_n}(q)| \le h(q)$ a.e.

(Evans Th. E.2.2)

Th (Egorov) $\Omega \subset \mathbb{R}^n$ measurable and $f_k, f \in Meas(\Omega)$ with $f_k(x) \to f(x)$ for almost every $x \in \Omega$. Then for each $\varepsilon > 0$ there exist a measurable $E_\varepsilon \subseteq \Omega$ such that

i) $|\Omega \setminus E_\varepsilon| \le \varepsilon$ (in Lebesgue measure),

ii) $f_k \longrightarrow f$ uniformly on E_ε.

proof (that F is seq. weakly lsc)

① Take a sequence $x_k \xrightarrow{W^{1,p}} x$. By uniform boundedness principle,

$\sup_k \|x_k\|_{W^{1,p}} < \infty$.

Rellich-Kondrachov: $\exists$ L^p-convergent subsequence

Converse dominated convergence: $\exists$ a.e.-convergent subsubsequence

Taking a further subsubsubsequence, we may assume that

i) $x_{k_\ell}(q) \to x(q)$ for a.e. $q \in \Omega$,

ii) $\lim_{\ell\to\infty} F(x_{k_\ell}) = \liminf_k F(x_k)$

② Egorov: $\exists E_\varepsilon$ such that $x_{k_\ell} \to x$ uniformly on E_ε and $|\;\setminus E_\varepsilon| \leq \varepsilon$.

Let $H_\varepsilon := \{q \in \Omega : |x(q)| + |\nabla x(q)| \leq \frac{1}{\varepsilon}\}$, so $|\Omega \setminus H_\varepsilon| \xrightarrow[\varepsilon\to 0]{} 0$.

Let $G_\varepsilon := E_\varepsilon \cap H_\varepsilon$, and so $|\Omega \setminus G_\varepsilon| \leq |\Omega \setminus E_\varepsilon| + |\Omega \setminus H_\varepsilon| \xrightarrow[\varepsilon\to 0]{} 0$.

③ Assume w.l.o.g. that $L \geq 0$ (recall L is bounded from below). Then

$$F(x_{k_\ell}) = \int_\Omega L(\nabla x_{k_\ell}(q), x_{k_\ell}(q), q)\,dq \geq \int_{G_\varepsilon} L(\nabla x_{k_\ell}(q), x_{k_\ell}(q), q)\,dq$$

$$\overset{\text{(convexity)}}{\geq} \underbrace{\int_{G_\varepsilon} L(\nabla x(q), x_{k_\ell}(q), q)\,dq}_{(\mathrm{I})} + \underbrace{\int_{G_\varepsilon} \nabla_{\nabla x} L(\nabla x(q), x_{k_\ell}(q), q) \cdot (\nabla x_{k_\ell}(q) - \nabla x(q))\,dq}_{(\mathrm{II})}$$

④ Since L is smooth, $L(\nabla x(q), \cdot, q)$ is uniformly continuous on $[-\frac{2}{\varepsilon}, \frac{2}{\varepsilon}]$.

Therefore $L(\nabla x(q), x_{k_\ell}(q), q) \xrightarrow[k]{} L(\nabla x(q), x(q), q)$ uniformly in $q \in G_\varepsilon$,

and thus $(\mathrm{I}) \to \int_{G_\varepsilon} L(\nabla x(q), x(q), q)\,dq$.

⑤ Similarly, $\nabla_{\nabla x} L(\nabla x(q), x_{k_\ell}(q), q) \to \nabla_{\nabla x} L(\nabla x(q), x(q), q)$ unif. in $q \in G_\varepsilon$, and so this convergence is in L^∞, and since $G_\varepsilon \subset \Omega$ is bounded, also (strongly) in L^p. Moreover, $\nabla x_{k_\ell} \xrightarrow{L^p} \nabla x$. This implies $(\mathrm{II}) \to 0$, i.e.

⑥
$$\liminf_{k\to\infty} F(x_k) = \lim_\ell F(x_{k_\ell}) \geq \int_{G_\varepsilon} L(\nabla x(q), x(q), q)\,dq.$$

monotone convergence as $\varepsilon \to 0$:

$$\liminf_{k\to\infty} F(x_k) \geq \int_\Omega L(\nabla x(q), x(q), q)\,dq = F(x). \quad \square$$

$\rightarrow$ Interestingly, this result also holds in the other direction...!

Lecture 11

A Generalised convexity notions

A1 Functionals on matrices

If $x: \Omega \to \mathbb{R}^m$, $\Omega \subset \mathbb{R}^n$ then $\nabla x \in \mathbb{R}^{m\times n}$.

$$F(x) = \int_\Omega L(\underbrace{\nabla x(q)}_{\in \mathbb{R}^{m\times n}}, x(q), q)\, dq$$

Def • $f: \mathbb{R}^{m\times n} \to \mathbb{R}\cup\{\infty\}$ is <u>polyconvex</u> iff

$$f(A) = \varphi(A, \text{minor}_1 A, \text{minor}_2 A, \dots) \text{ for some convex } \varphi.$$

(recall: a minor is the determinant of the matrix where row and columns have been removed)

• $f: \mathbb{R}^{m\times n} \to \mathbb{R}$ (locally bounded & measurable) is <u>quasiconvex</u> iff

$$f(A) \leq \frac{1}{|\Omega|} \int_\Omega f(A + \nabla\varphi(q))\, dq \quad \text{for every } \Omega \subset \mathbb{R}^n \text{ bdd \& open and } \varphi \in C_0^\infty(\Omega; \mathbb{R}^m).$$

Formal Theorem $F(x) = \int_\Omega L(\nabla x(q), x(q), q)\, dq$

L is quasiconvex in its first argument
$\Longleftrightarrow$
F is sequentially weakly lsc in $W^{1,p}(\Omega)$

(Dacorogna - Direct Methods in the calculus of variations, 2^{nd} ed, chapter 5 & 9)

→ ~~State~~ In practice, quasiconvexity is hard to check. Therefore one often replaces it by different notions (Dacorogna Th. 1.7)

Th $f: \mathbb{R}^{m\times n} \to \mathbb{R}$

f convex $\Rightarrow$ f polyconvex $\Rightarrow$ f quasiconvex $\Rightarrow$ f "rank one convex"

→ The exact terminology "quasiconvexity" changes in the literature...!

A2 λ - convexity (a.k.a. strong convexity or semi-convexity)

Recall: Let $A = \operatorname{dom} F \subset X$ be open and convex and $F: A \to \mathbb{R}$ be twice Gateaux differentiable.

i) F is convex
$\Longleftrightarrow$
ii) $F(y) \geq F(x) + \langle DF(x), y-x \rangle \quad \forall x, y \in A$ (gradient ineq.)
$\Longleftrightarrow$
iii) $\langle DF(x) - DF(y), x-y \rangle \geq 0 \quad \forall x, y \in A$ (monotonicity)
$\Longleftrightarrow$
iv) $D^2F(x) \geq 0$ (positive semidefinite)

What if we replace 0 by a different constant?

Def $F: X \to \mathbb{R} \cup \{\infty\}$ is λ-convex, $\lambda \in \mathbb{R}$ iff

$$F((1-\sigma)x + \sigma y) \leq (1-\sigma)F(x) + \sigma F(y) - \frac{\lambda}{2}(1-\sigma)\sigma \|x-y\|^2$$

(Peypouquet pp 25)

Remarks:

- $\lambda = 0 \Leftrightarrow$ convex
  - $\lambda > 0 \Rightarrow$ convex (stronger)
  - $\lambda < 0 \Leftarrow$ convex (weaker)
- F is λ-convex iff $F - \frac{\lambda}{2}\|\cdot\|^2$ is convex.
  (Peypouquet Prop. 3.12)
  Daneri & Savaré - Lecture notes on grad. flows and opt. transport 2010

Th Let $A = \operatorname{dom} F \subset X$ be open and convex and $F: A \to \mathbb{R}$ be twice Gateaux differentiable, $\lambda \in \mathbb{R}$.

i) F is λ-convex
$\Longleftrightarrow$
ii) $F(y) \geq F(x) + \langle DF(x), y-x \rangle + \frac{\lambda}{2}\|y-x\|^2$ (gradient ineq.)
$\Longleftrightarrow$
iii) $\langle DF(x) - DF(y), x-y \rangle \geq \lambda \|y-x\|^2$ (monotonicity)
$\Longleftrightarrow$
iv) $D^2F(x) \geq \lambda$ (positive definiteness)

- Convex (& lsc) functionals have supporting hyperplanes. Similarly, λ-convex functionals have supporting "hyperparabola"
- $\lambda < 0$: can still do a lot of convex analysis even though F not convex!
- $\lambda > 0$: more control & better estimates.

B Uniform integrability (Introduction to Orlicz spaces)

In what follows $(\Omega, \mathcal{A}, \mu)$ will always be a prob. meas. space.

Recall that
- L^p $p>1$ has a predual;
- A uniform L^p-bound yields weak-* compactness; (Banach-Alaoglu)
- L^1 doesn't have a predual, so a uniform L^1-bound (e.g. from Jensen) doesn't work.

However: (Brezis Th. 4.29/4.30)

Th (Dunford-Pettis)
A bounded sequence $(x_n) \subset L^1_\mu(\Omega)$ is relatively <u>weakly</u> compact $\iff$ it is "<u>uniformly integrable</u>"

Def A sequence $(x_n) \subset L^1_\mu(\Omega)$ is uniformly integrable iff

$$\forall \varepsilon>0 \; \exists \delta>0 \; \forall n \; \forall A \in \mathcal{A} \quad \mu(A)<\delta \Rightarrow \int_A |x_n(\omega)| \mu(d\omega) < \varepsilon$$

This is nice, but how to prove uniform integrability, e.g. of level sets $\{F \leq C\}$ of some functional $F(x) = \int f(|x(\omega)|) \mu(d\omega)$?
Can we exploit the information that we threw away by Jensen?

$$f\left(\int |x(\omega)| \mu(d\omega)\right) \overset{(\text{Jensen})}{\leq} \int f(|x(\omega)|) \mu(d\omega) \leq C, \quad (f \text{ convex})$$

(Bogachev, Th I.4.5.9)

Th (De la Vallée-Pousin) (Rao & Ren - Theory of Orlicz spaces (1991) Th. 1.2.2)
A sequence $(x_n) \subset L^1_\mu(\Omega)$ is uniformly integrable iff

there exists a non-negative increasing convex $\varphi : [0,\infty) \to [0,\infty)$ for which
- $\lim_{t \to \infty} \frac{\varphi(t)}{t} = \infty$ (superlinear growth)
- $\sup_n \int \varphi(|x_n(\omega)|) \mu(d\omega) < \infty$

Hence $\{F \leq C\}$ is (seq.) weakly compact in L^1

→ Another application of uniform integrability is Vitali's convergence theorem: $x_n(\omega) \to x(\omega)$ μ-a.e & x_n unif. int $\Rightarrow$ $x_n \xrightarrow{L^p} x$.

→ Instead of working with L^1 weak, maybe we can work directly with the "<u>Orlicz class</u>":

$$L^\varphi_\mu(\Omega) := \left\{ x : \Omega \to \mathbb{R} \text{ measurable} : \int_\Omega \varphi(|x(\omega)|) \mu(d\omega) < \infty \right\}.$$

C Relation with L^1_μ

(Rao&Ren, Cor. 1.2.3)

Th $L^1_\mu(\Omega) = \bigcup \{ \tilde{L}^{\varphi}_\mu(\Omega) : \varphi$ non-negative, increasing, convex, $\lim_{t\to\infty} \frac{\varphi(t)}{t} = \infty \}$

proof: "$\supseteq$" ~~φ is convex~~ $\tilde{L}^{\varphi}_\mu \subseteq L^1_{\varphi\mu}$ $\forall \varphi$ by Jensen.

hence $L^1_\mu \supseteq \bigcup_\varphi \tilde{L}^\varphi_\mu$.

"$\subseteq$" The one-element set $\{x\} \subset L^1_\mu$ is clearly "uniformly integrable", so by de la Vallée-Poussin $x \in \tilde{L}^\varphi_\mu$ for some φ satisfying the conditions. ∎

→ Compare this to:

$$L^1_\mu(\Omega) \supsetneq \bigcup_{p>1} L^p_\mu(\Omega).$$

→ Can we put more structure on $\tilde{L}^\varphi_\mu$? Challenges:

- $\int \varphi(|x|)d\mu$ can not be scaled to be a norm.
- $\tilde{L}^\varphi_\mu$ is not a vector space! (In general)

D Setting: (Young and) N-functions.

(we shall assume φ is defined on the whole $\mathbb{R}$, but even, i.e. $\varphi(-z) = \varphi(z)$)

Def $\varphi : \mathbb{R} \to \mathbb{R}^+$ is an <u>N-function</u> iff it is:

- continuous
- convex
- $\varphi(z) = 0 \iff z = 0$
- $\varphi(-z) = \varphi(z)$ (even)
- $\lim_{z\to\infty} \frac{\varphi(z)}{z} = \infty$ (superlinear growth)
- $\lim_{z\to 0} \frac{\varphi(z)}{z} = 0$ (differentiable in 0)

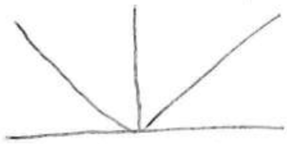
Prop If φ is an N-function then so is φ^*

proof (exercise).

→ Many results require less assumptions, e.g. $\varphi : \mathbb{R} \to \mathbb{R}^+ \cup \{\infty\}$, or μ may even be an infinite measure: $\mu(\Omega) = \infty$.

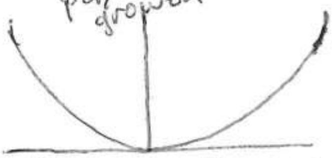
For consistency we focus on N-functions φ and prob. meas. μ.

examples:



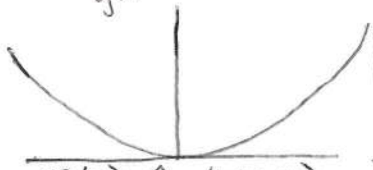
$\varphi(z) = |z|$
(strictly speaking not allowed)

polynomial growth



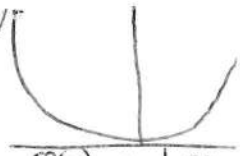
$\varphi(z) = |z|^p,\ p>1$

almost linear growth



$\varphi(z) = \lambda_B(|z|+1)$
$= (|z|+1)\operatorname{Log}(|z|+1) - z$

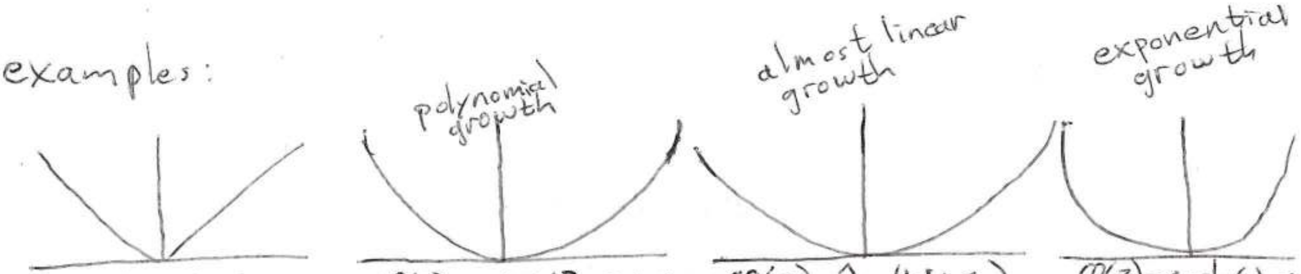
exponential growth



$\varphi(z) = \cosh(z) -$

$\int \varphi(|x|)\,d\mu = \|x\|_{L^1_\mu}$ $\quad \dots = \|x\|^p_{L^p_\mu}$ $\quad \dots = ?$ $\quad \dots = ?$

examples:



$\varphi(z) = |z|$ (strictly speaking not allowed) | $\varphi(z) = |z|^p,\ p>1$ | $\varphi(z) = \lambda_B(|z|+1) = (|z|+1)\operatorname{Log}(|z|+1) - z$ | $\varphi(z) = \cosh(z) - 1$

$\int \varphi(|x|)\,d\mu = \|x\|_{L^1_\mu}$ | $\ldots = \|x\|^p_{L^p_\mu}$ | $\ldots = ?$ | $\ldots = ?$

Lecture 12

A Linearising the Orlicz class

(Rao&Ren, Th. 3.1.2)

Th. Let $\varphi: \mathbb{R} \to \mathbb{R}^+$ be an N-function and $(\Omega, \mathcal{A}, \mu)$ a non-atomic probability space. Then:

i) $\tilde{L}^\varphi_\mu(\Omega)$ is a vector space iff

$$\exists K, z_0 \geq 0 \quad \forall z \geq 0 \quad \varphi(2z) \leq K\varphi(z)$$

"Δ_2-property" (or $\varphi \in \Delta_2$)

ii) In general,
- $x, y \in \tilde{L}^\varphi_\mu(\Omega)$, $|\alpha| + |\beta| \leq 1 \Rightarrow \alpha f + \beta g \in \tilde{L}^\varphi_\mu(\Omega)$,
- $y \in \tilde{L}^\varphi_\mu(\Omega)$ and x measurable with $|x| \leq |y| \Rightarrow x \in \tilde{L}^\varphi_\mu(\Omega)$.

"circled solid subset"

(implicitly identifying a.e. equal functions)

Def $L^\varphi_\mu(\Omega) := \{x : \Omega \to \mathbb{R} \text{ measurable} : \exists \alpha > 0 \text{ s.t. } \int_\Omega \varphi(\alpha x)\,d\mu < \infty\}$

"Orlicz space"

Prop $L^\varphi_\mu(\Omega)$ is a vector space. (Rao&Ren prop. 3.1.6)

Proof Take $x, y \in L^\varphi_\mu(\Omega)$, hence $\exists \alpha_x, \alpha_y > 0$ s.t. $\alpha_x x, \alpha_y y \in \tilde{L}^\varphi_\mu(\Omega)$.
Take arbitrary $a, b \in \mathbb{R}$, and set $C := \frac{\alpha_x}{|a|} \wedge \frac{\alpha_y}{|b|}$.
Then
$$C(ax + by) = \underbrace{\frac{Ca}{\alpha_x}}_{|\cdot| \leq 1} \underbrace{\alpha_x x}_{\in \tilde{L}^\varphi_\mu} + \underbrace{\frac{Cb}{\alpha_y}}_{|\cdot| \leq 1} \underbrace{\alpha_y y}_{\in \tilde{L}^\varphi_\mu} \in \tilde{L}^\varphi_\mu(\Omega) \Rightarrow ax + by \in L^\varphi_\mu(\Omega)$$
$\square$

B Norming the Orlicz space I

(Rao & Ren prop 3.1.6)

prop $\forall x \in L^{\varphi}_{\mu}(\Omega)\ \exists \beta$ s.t. $\int \varphi(|\beta x|)d\mu \leq 1$

proof: Take an (arbitrary) sequence $a_n \to 0$ and set $\alpha_n := \alpha \wedge a_n$, where α is such that $\alpha x \in \tilde{L}^{\varphi}_{\mu}(\Omega)$.
Then $0 \leq \varphi(\alpha_n x) \leq \varphi(\alpha x) \in L^1_{\mu}(\Omega)$ and $\varphi(\underbrace{\alpha_n}_{\to 0} x(q)) \to 0$ for (ae) $q \in \Omega$
Dominated convergence: $\int_{\Omega} \varphi(\alpha_n x(q)) \mu(dq) \to 0$. Hence there exists an n_0 (set $\beta := \alpha_{n_0}$) such that $\int \varphi(|\beta x|) \leq 1$. □

$N_\varphi(x) := \inf\{k>0 : \int \varphi(|\frac{x}{k}|)d\mu \leq 1\}$ Luxemburg norm

Prop N_φ is a norm (Rao & Ren Th. 3.2.3)

sketch of proof:

- $N_\varphi(\alpha x) = \inf\{k>0 : \int\varphi(|\frac{\alpha x}{k}|)d\mu \leq 1\}$
$= \inf\{\alpha k > 0 : \int \varphi(\frac{|x|}{k})d\mu \leq 1\}$
$= \alpha N_\varphi(x)$.
- (triangle inequality follows from convexity of φ) □

Remark: more generally, norms of the type
$x \mapsto \inf\{k>0 : \frac{x}{k} \in B\}$ for some circled solid set B
are called gauge or Minkowski norms.

Th $(L^{\varphi}_{\mu}(\Omega), N_\varphi)$ is a Banach space (i.e. complete) (Rao & Ren Th. 3.3.10)
(when μ-a.e. equivalent functions are identified)

Very usefull unit ball property:

(Rao & Ren Th. 3.2.3)

Th $N_\varphi(x) <,=,> 1 \iff \int\varphi(|x|) <,=,> 1$

Let $0 \neq x \in L^{\varphi}_{\mu}$ and $0 \neq y \in L^{\varphi^*}_{\mu}$. By Young's inequality

$$\frac{|x\,y|}{N_\varphi(x)\,N_{\varphi^*}(y)} \leq \varphi\left(\frac{|x|}{N_\varphi(x)}\right) + \varphi^*\left(\frac{|y|}{N_\varphi(y)}\right).$$

Integrating:

$$\frac{1}{N_\varphi(x)N_{\varphi^*}(y)} \int |xy|\,d\mu \leq \int \varphi\left(\frac{|x|}{N_\varphi(x)}\right)d\mu + \int \varphi^*\left(\frac{|y|}{N_{\varphi^*}(y)}\right)d\mu$$

$$\leq \quad 1 \quad + \quad 1 \quad = 2,$$

and so we obtain a Hölder-type estimate:

$$\boxed{\int |xy|\,d\mu \leq 2\,N_\varphi(x)\,N_{\varphi^*}(y)}$$ (Rao & Ren prop. 3.3.1)

The factor 2 is not very nice here; we will see that one obtains a better estimate when choosing a different norm. However, this estimate does show that L^{φ}_{μ} and $L^{\varphi^*}_{\mu}$ may act as dual spaces (in some sense; we will be more precise later).

C Norming the Orlicz space II

Recall that, for a general Banach space,

$$\|x^*\|_{X^*} := \sup\{\langle x^*, x\rangle : \|x\|_X \leq 1\} \quad \text{and}$$

$$\|x\|_X = \sup\{\langle x^*, x\rangle : \|x^*\|_{X^*} \leq 1\}.$$

Motivated by this we define

Orlicz norm

$$\|x\|_\varphi := \sup\left\{\int |xy|\,d\mu : y \in L^{\varphi^*}_{\mu}(\Omega),\ \int \varphi^*(|y|)\,d\mu \leq 1\right\}$$

$$\underset{\substack{\text{(unit ball}\\\text{property)}}}{=} \sup\left\{\int |xy|\,d\mu : y \in L^{\varphi^*}_{\mu}(\Omega),\ N_{\varphi^*}(y) \leq 1\right\}$$

Not very difficult to see that this is a norm!

(Rao & Ren Th. 3.3.13)

Th (Krasnoselskii & Rutickii)

$$\|x\|_\varphi = \min_{k>0}\left\{\frac{1}{k}\left(1 + \int \varphi(|kx|)\,d\mu\right)\right\} \text{ (Amemiya norm)}$$

This is a very practical expression, since:

Cor $\|x\|_\varphi \leq 1 + \int \varphi(|x|)\,d\mu$

Hence a uniform bound on $\int \varphi(|x|)\,d\mu$ yields a uniform bound on the Orlicz norm!

Similarly to the unit ball property of the Luxemburg norm, we now have:

$$\boxed{\underline{\underline{\text{Prop}}} \quad \int \varphi\left(\frac{x}{\|x\|_\varphi}\right) \leq 1 \quad \forall\, 0 \neq x \in L^\varphi_\mu(\Omega)}$$ (Rao & Ren Prop. 3.3.3)

From this, for any $x \in L^\varphi_\mu$, $y \in L^{\varphi^*}_\mu$:

$$\|x\|_\varphi := \sup\left\{ \int |x\bar{y}|\,d\mu : \bar{y} \in L^{\varphi^*}_\mu \text{ with } \int \varphi^*(\bar{y}) \leq 1 \right\}$$

$$\geq \int \left| x \frac{y}{\|y\|_{\varphi^*}} \right| d\mu \qquad \text{since} \quad \int \varphi^*\left(\frac{y}{\|y\|_{\varphi^*}}\right) \leq 1.$$

Hence we find again a Hölder-type estimate:

$$\boxed{\int |xy|\,d\mu \leq \|x\|_\varphi \|y\|_{\varphi^*}}$$

Another Hölder-type estimate follows from the Luxemburg ball property:

$$\boxed{\int |xy|\,d\mu \leq \|x\|_\varphi N_{\varphi^*}(y)}$$

D) Relation between the Orlicz and Luxemburg norms

The two norms are certainly not equal. In fact

$$\|x\|_\varphi = N_\varphi(x) \iff x = 0 \text{ a.e.}$$

However,

(Rao & Ren Prop. 3.3.4)

$$\boxed{\underline{\underline{\text{Prop}}} \quad N_\varphi(x) \leq \|x\|_\varphi \leq 2 N_\varphi(x) \qquad \forall x \in L^\varphi_\mu(\Omega)}$$
(equivalence of norms!)

Proof. Note that $k = \|x\|_\varphi$ satisfies $\int \varphi\left(\frac{x}{k}\right) d\mu \leq 1$, hence

$$N_\varphi(x) := \inf\left\{ k > 0 : \int \varphi\left(\frac{x}{k}\right) \leq 1 \right\} \leq \|x\|_\varphi.$$

On the other hand,

$$\|x\|_\varphi = \sup\left\{ \int |xy|\,d\mu : y \in L^{\varphi^*}_\mu \text{ with } N_{\varphi^*}(y) \leq 1 \right\}$$

$$\overset{\text{Hölder}}{\leq} \sup\left\{ 2 N_\varphi(x) N_{\varphi^*}(y) : y \in L^{\varphi^*}_\mu \text{ with } N_{\varphi^*}(y) \leq 1 \right\}$$

$$= 2 N_\varphi(x). \quad \blacksquare$$

In other words, the topologies generated by N_φ and $\|\cdot\|_\varphi$ are the same! This also implies that

$\boxed{\underline{\text{Th}} \quad (L^\varphi_\mu(\Omega), \|\cdot\|_\varphi) \text{ is a (complete) Banach space}}$ (Rao & Ren Prop. 3.3.11)

Lecture 13

Last week we introduced the Orlicz class and space:

$\tilde{L}^{\varphi}_{\mu}(\Omega) := \{x : \Omega \to \mathbb{R}$ meas. with $\int \varphi(|x|)d\mu < \infty\}$,
$L^{\varphi}_{\mu}(\Omega) := \{x : \Omega \to \mathbb{R}$ meas. s.t. $\exists \alpha > 0, \int \varphi(|\alpha x|)d\mu < \infty\}$.

Infact, L^{φ}_{μ} is the linear hull of $\tilde{L}^{\varphi}_{\mu}$, i.e. the smallest vector space containing $\tilde{L}^{\varphi}_{\mu}$. This can be seen as follows:

- Closedness under addition is no problem if the space is closed under rescaling, since by convexity
$$\int \varphi(|x+y|)d\mu \leq \tfrac{1}{2}\int \varphi(|2x|)d\mu + \tfrac{1}{2}\int \varphi(|2y|)d\mu.$$
- In order to make $\tilde{L}^{\varphi}_{\mu}$ closed under any rescaling $\alpha > 0$, we need to take L^{φ}_{μ}!

Recall that $\tilde{L}^{\varphi}_{\mu} = L^{\varphi}_{\mu}$ iff φ has the Δ_2-property:
$$:\Leftrightarrow \exists K\ \forall z \text{ (suff. large) } \varphi(2z) \leq K\varphi(z).$$

Note that:
- $\varphi(z) = \frac{1}{p}|z|^p$ is special in that $\varphi(2z) \underset{\text{equality}}{=} \underbrace{2^p}_{K}\varphi(z)$.
- The factor 2 is arbitrary, since
$$\varphi(2^n z) \leq K^n \varphi(z).$$

Two equivalent norms:

$N_{\varphi}(x) := \inf\{k > 0 : \int \varphi(|\tfrac{x}{k}|)d\mu \leq 1\}$ Luxemburg (gauge/Minkowski)

$\|x\|_{\varphi} := \sup\{\int |xy| d\mu : y \in L^{\varphi^*}_{\mu}, \int \varphi^*(|y|)d\mu \leq 1\}$ Orlicz

A The M^{φ}_{μ}-space and its dual

$M^{\varphi}_{\mu}(\Omega) := \{x : \Omega \to \mathbb{R}$ meas. s.t. $\underline{\underline{\forall}}\, \alpha > 0, \int \varphi(|\alpha x|)d\mu < \infty\}$

Clearly, M^{φ}_{μ} is a vector space, and
$$M^{\varphi}_{\mu} \subseteq \tilde{L}^{\varphi}_{\mu} \subseteq L^{\varphi}_{\mu}.$$

Infact, M^{φ}_{μ} is also closed under either norm, and hence also a (complete) Banach space! (Rao & Ren prop. 3.4.3.)

Th If φ has the Δ_2-property then $M^{\varphi}_{\mu} = \tilde{L}^{\varphi}_{\mu} = L^{\varphi}_{\mu}$ (Rao & Ren cor. 3.4.5)

<u>proof</u> Take $x \in L^\varphi_\mu$ with $\alpha > 0$ s.t. $\int \varphi(|\alpha x|) d\mu < \infty$.
Need to prove that $\int \varphi(|\beta x|) < \infty$ for <u>any</u> $\beta > 0$, so that $x \in M^\varphi_\mu$.
Let $n \in \mathbb{N}$ be such that $\beta \leq 2^n \alpha$. Then

$$\int \varphi(|\beta x|) d\mu \leq \int \varphi(|2^n \alpha x|) d\mu \leq K^n \int \varphi(|\alpha x|) d\mu < \infty . \square$$

<u>Th</u> $(M^\varphi, N_\varphi)^* = (L^{\varphi^*}, \|\cdot\|_{\varphi^*})$ (Rao & Ren Th. 4.1.7)
and
$(M^\varphi, \|\cdot\|_\varphi)^* = (L^{\varphi^*}, N_{\varphi^*})$,
i.e. $(M^\varphi)^* = L^{\varphi^*}$ and

$$\|x^*\|_{(M^\varphi, N_\varphi)^*} := \sup \left\{ \int |x^* x| d\mu : x \in L^\varphi_\mu , N_\varphi(x) \leq 1 \right\} = \|x^*\|_{\varphi^*};$$

$$\|x^*\|_{(M^\varphi, \|\cdot\|)^*} := \sup \left\{ \int |x^* x| d\mu : x \in L^\varphi_\mu , \|x\|_\varphi \leq 1 \right\} = N_{\varphi^*}(x^*).$$

(Either norm corresponds to the other norm in the dual space!)

<u>Cor</u> 1) if φ has the Δ_2-property then $(M^\varphi)^* = (L^\varphi)^* = L^{\varphi^*}$.
2) if both φ and φ^* have the Δ_2-property then L^φ is reflexive.

<u>Important</u>: L^φ_μ always has a predual (whether Δ_2 holds or not)!!
Hence weak-$*$ compact level sets by Banach-Alaoglu!
If in addition, L^φ_μ is separable then weak-$*$ <u>sequentially</u> compact level sets!

<u>Prop</u> Let $\Omega \subseteq \mathbb{R}^n$ with Borel σ-algebra $\mathcal{A}$ and non-atomic measure $\mu \in \mathcal{P}(\Omega)$. L^φ_μ is separable iff φ has the Δ_2-property.

(Rao & Ren Th. 3.5.1)

$\boxed{B}$ An example (last week's exercise)

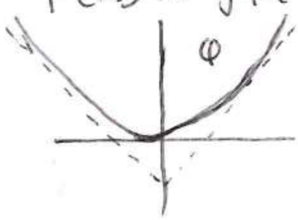
$\psi_1(z) := z\log z - z$

$\psi_2(z) := -z\log(-3z) + z$

$\psi(z) := (\psi_1 \,\square\, \psi_2)(z)$, $F(x) := \int \psi(x)\,d\mu$ Minimiser?

N-function:

$\varphi(z) := \cosh^* z + 1$



grows slightly faster than linearly, but slower than polynomial, for any $p>1$.

hence Δ_2-property satisfied:

$\varphi(2z) \leq 2^{1+\varepsilon}\varphi(z)$ for z sufficiently large

Horrible expressions if you'd want to calculate ψ and φ explicitly! Instead, work with their duals:

$$\psi_1^*(z^*) = \tfrac{1}{2}e^{z^*},\quad \psi_2^*(z^*) = \tfrac{1}{3}e^{-z^*}$$

$$\psi^*(z^*) = \psi_1^*(z^*) + \psi_2^*(z^*) = \tfrac{1}{2}e^{z^*} + \tfrac{1}{3}e^{-z^*}.$$

$$\varphi^*(z^*) = \cosh(z^*) - 1.$$

Clearly $\psi^*(z^*) \leq \varphi^*(z^*) + 1$, and so

$$\boxed{\psi(z) \geq \varphi(z) - 1}$$

Hence on level sets $\{F \leq C\}$:

(Krasnoselskii-Rutícki)

$$\|x\|_\varphi \leq 1 + \int \varphi(|x|)\,d\mu \leq 2 + \int (\varphi(|x|) - 1)\,d\mu$$

$$\leq 2 + \int \psi(x)\,d\mu \leq 2 + C \quad \text{uniformly bounded!}$$

compactness:

- Banach-Alaoglu: $\{F \leq C\}$ is weakly-$*$ compact in L^φ_μ.
  (This is meaningfull since $M^{\varphi^*}_\mu$ is the predual of L^φ_μ.
  Be aware however that φ^* grows exponentially and can not satisfy the Δ_2-property, hence $M^{\varphi^*}_\mu \subsetneq L^{\varphi^*}_\mu$.)

  If we take $\Omega \subseteq \mathbb{R}^n$ and μ a (rescaled) Lebesgue measure, then L^φ_μ is separable and $\{F \leq C\}$ is weakly-$*$ seq. cpt.

lsc:

- Write:

$$F(x) = \int \Big[\sup_{z^*} z^*x - \psi_1^*(z^*) - \psi_2^*(z^*)\Big]\,d\mu$$

$$\overset{\text{(to prove)}}{=} \sup_{x^* \text{ meas.}} \int [x^*x - \psi_1^*(x^*) - \psi_2^*(x^*)]\,d\mu$$

$$= \sup_{x^* \in M^{\varphi^*}_\mu} \int [\underbrace{x^*x}_{\text{weak-}*\text{ cont. in } x} - \underbrace{\psi_1^*(x^*) - \psi_2^*(x^*)}_{\text{constant in } x}]\,d\mu \Rightarrow F \text{ weakly-}* \text{ lsc.}$$

Direct method: $\exists!$ minimiser $x \in L^\varphi_\mu$ of $F(x)$ (unique by strict convexity)

Remark: we could also have worked with L^1_μ-weak and unif. integrability..