IV Selected topics

1 (Introduction) to primal-dual methods

Tools enabling to reformulate the problem in an equivalent way. May be the resultant problem is easier to solve.

1.1 Convex conjugate functions

V real normed vector space, reflexive $V^{**} = V$

eg. $\mathbb{R}^n$ is reflexive.

$\text{epi}\, F = \{(u, r) \in V \times \mathbb{R} : F(u) \leq r\}$ the epigraph of F

Thm $\quad F$ convex $\Longleftrightarrow$ epi F convex.

F convex $\Longrightarrow$ $\{u : F(u) \leq r\}$ convex called the r-sublevel set

$\not\Longleftarrow$ false.

convex conjugacy – dual representation of epi F.

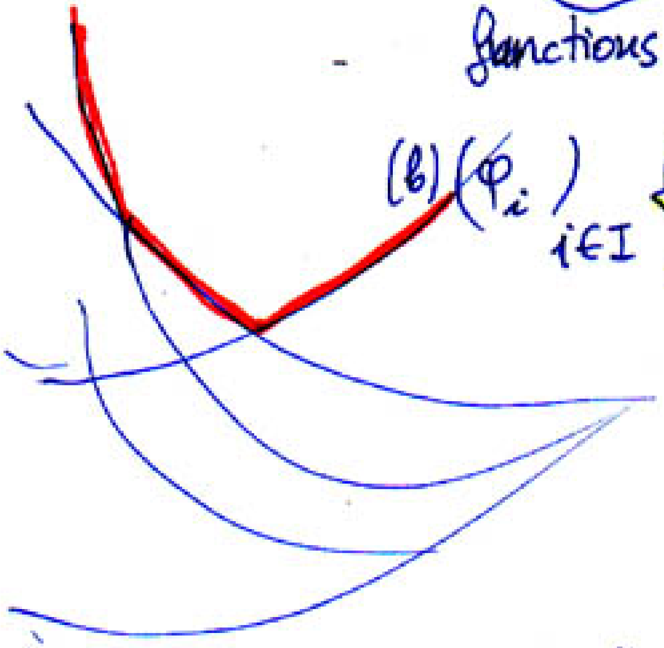
Definition $\quad F : V \to \mathbb{R}$ (may be non convex)

The function conjugated to F is defined by

$$F^*(u^*) = \sup_{u \in V} \left(\langle u, u^* \rangle - F(u) \right)$$

(if $u \in \mathbb{R}^n$ then $u^* \in \mathbb{R}^n$)

Basic results (a) $(\varphi_i)_{i\in I}$ family of lower semi-continuous (l.s.c.) functions $\Rightarrow$ $\varphi(u) = \sup_{i\in I} \varphi_i(u)$ is l.s.c.



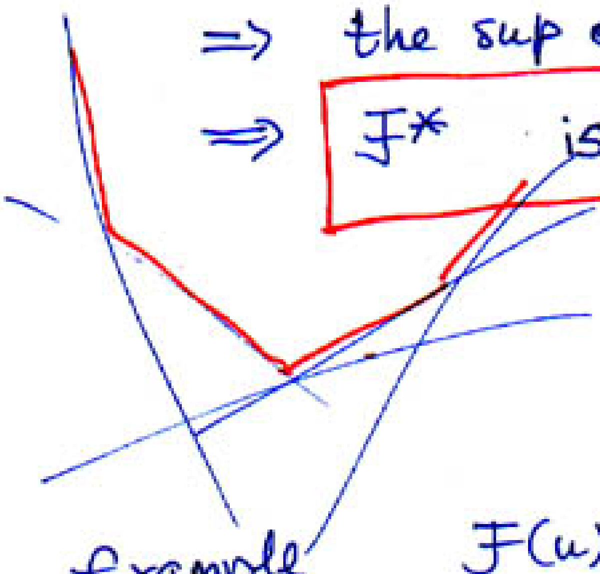
(b) $(\varphi_i)_{i\in I}$ family of convex functions then $\varphi(\overset{*}{u}) = \sup_{i\in I} \varphi_i(\overset{*}{u})$ is convex

$u^* \longrightarrow \langle u, u^*\rangle - \mathcal{F}(u)$ affine hence convex $\forall u \in V$

family of convex func. when u goes throug V

$\Rightarrow$ the sup over V is convex

$\Rightarrow$ $\boxed{\mathcal{F}^* \text{ is convex}}$



Example $\mathcal{F}(u) = \|u\|$ any norm on $\mathbb{R}^n$ $\Rightarrow$ convex

$$\mathcal{F}^*(u^*) = \sup_{u\in\mathbb{R}^n}\left(\langle u, u^*\rangle - \mathcal{F}(u)\right) = \sup_{u\in\mathbb{R}^n}\left(\langle u, u^*\rangle - \|u\|\right)$$

Schwarz: $\langle u, u^*\rangle \le \|u\| \cdot \|u^*\|$ equality reached iff $u = \lambda u^*$, $\lambda \in \mathbb{R}$

$$F^*(u^*) = \sup_{\substack{\|u\|\in\mathbb{R} \\ \text{in fact} \\ \geq 0}} \Big(\|u\|\,\|u^*\| - \underbrace{\|u\|}_{\lambda} \Big)$$

$$(*) \quad = \sup_{\lambda \geq 0} \lambda\big(\|u^*\| - 1\big) = \begin{cases} 0 & \text{if } \|u^*\| < 1 \\ +\infty & \text{if } \|u^*\| > 1 \end{cases}$$

Rmk

$$F^{**}(u) \leq F(u) \quad \forall u \in V$$

1.2. Dual problems

Fenchel-Moreau theorem

V reflexive real normed vector space (eg $\mathbb{R}^n$)
$V = V^{**}$
F - convex l.s.c. $\not\equiv \infty$ defined on $(-\infty, +\infty]$

$$F^{**} = F$$

the bi-conjugate $\quad F^{**}(u) = \sup_{u^* \in V^*} \big(\langle u, u^* \rangle - F(u^*) \big)$

Important theorem (Fenchel-Rockafellar)

Let F, G convex functions on V
Suppose $\exists\, u_0$ such that $F(u_0) < +\infty$, $G(u_0) < +\infty$
and F continuous at u_0. Then

$$\inf_{u \in V} \{F(u) + G(u)\} = \max_{u^* \in V^*} \{-F^*(-u^*) - G^*(u^*)\}$$

Eckeland Temam – Convex ...–
Luenberger – Optim. on vector spaces

Exemple $C \subset V$ has non-empty interior, closed convex

$$I_C(u) =: \begin{cases} 0 & \text{if } u \in C \\ +\infty & \text{if } u \notin C \end{cases}$$ the indicator function

$I_C(u)$ is convex and lsc $I_C \not\equiv +\infty \Rightarrow$

$$I_C^{**} = I_C$$

$$I_C^*(u^*) = \sup_{u \in V} \Big(\langle u, u^* \rangle \overbrace{- I_C(u)}^{-\infty \text{ if } u \notin C \text{ hence no supremum}} \Big)$$

$$= \sup_{u \in C} \langle u, u^* \rangle \stackrel{\text{def}}{=} \boxed{\sigma_C(u^*)}$$ the support function of the set C

$$\operatorname{Inf}_{u \in C} \mathcal{F}(u) = \operatorname{Inf}_{u \in V} \big(\mathcal{F}(u) + I_C(u) \big) \;=$$

Let $\mathcal{F}(u) = \|u - v\|$ (we look for the projection of v on C

$$\mathcal{F}^*(-u^*) = \sup_{\substack{u \in V \\ w \in V}} \Big(\langle -u^*, \underbrace{u - v}_{w} \rangle - \| \underbrace{u - v}_{w} \| \Big) - \langle u^*, v \rangle$$

use ⊛ page 91 TOP

$$= \begin{cases} -\langle u^*, v \rangle & \text{if } \|u^*\| \le 1 \\ +\infty & \text{if } \|u^*\| > 1 \end{cases}$$

Fenchel = Rockafel.

$$\max_{\|u^*\| \le 1} \{ \langle u^*, v \rangle - \sigma_C(u^*) \} = \operatorname{dist}(v, C)$$

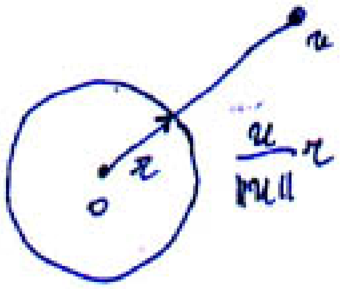
Suppose $C = H$ a vector subspace $\subset V$ (e.g. $\mathbb{R}^n$)

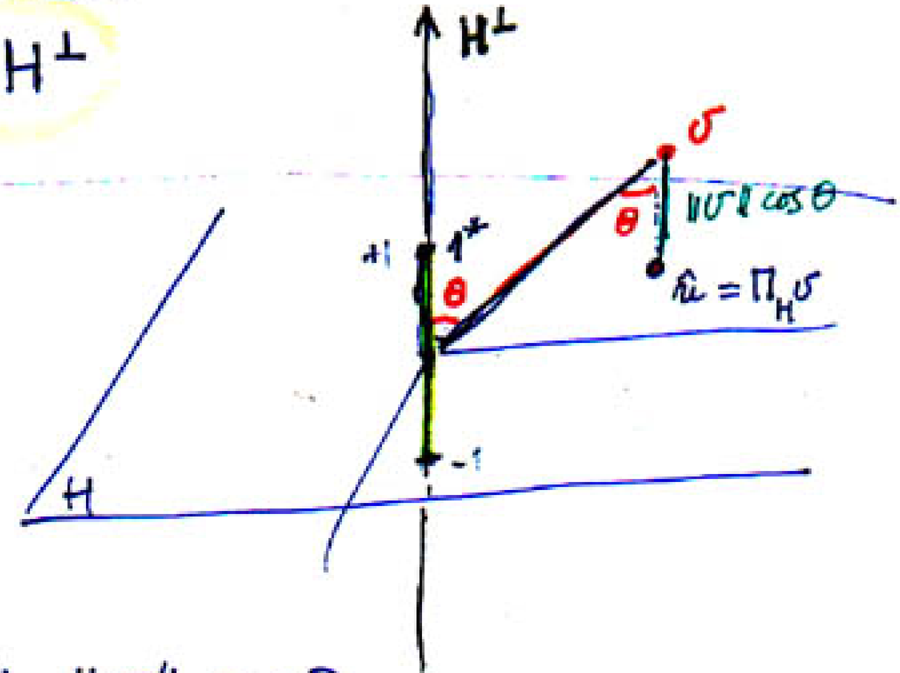
$$I_H^*(u) = \sup_{u \in H} \langle u, u^* \rangle = \begin{cases} 0 & \text{if } u^* \in H^\perp \\ +\infty & \text{if } u^* \notin H^\perp \end{cases}$$

Given $v \in V$

$$\operatorname*{Inf}_{u \in H} \|u - v\| = \max_{\substack{\|u^*\| \le 1 \\ u^* \in H^\perp}} \langle v, u^* \rangle$$

(the proj. of v on H)





$$\max \langle v, u^* \rangle = \underbrace{\|1^*\|}_{1} \cdot \|v\| \cos\theta$$

The Min-max theorem (see e.g. Leuenberger)

V reflexive normed vector space

A, B – compact convex nonempty : $A \subset V$, $B \subset V^*$

(if $V = \mathbb{R}^n$ + Euclid. norm, $A \subset \mathbb{R}^n$, $B \in \mathbb{R}^n$)

$$\boxed{\min_{u \in A}} \max_{u^* \in B} \langle u, u^* \rangle = \max_{u^* \in B} \boxed{\min_{u \in A}} \langle u, u^* \rangle$$

1.3. Exemple - $\boxed{\text{half-quadratic regularization}}$

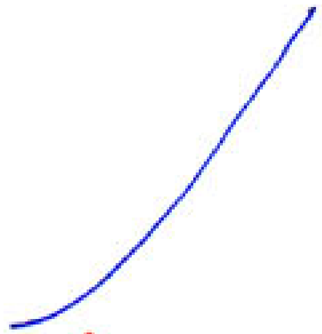
any linear

The pb: $\underset{u\in\mathbb{R}^n}{\text{minimize}} \quad \mathcal{F}(u) = \|Au - v\|^2 + \beta \sum_{i\in I} \varphi(\underbrace{\|G_i u\|}_{t})$

increasing on $\mathbb{R}_+$

e.g. $G_i u \approx \nabla_i u$

$$\begin{pmatrix} u_{ij} - u_{i-1j} \\ u_{ij} - u_{ij-1} \end{pmatrix} \qquad G_i u = u_{ij} - u_{ij-1}$$



examples for φ: $\begin{cases} \varphi(t) = \sqrt{t^2+\alpha} \\ \varphi(t) = t^p \quad p>1 \\ \varphi(t) = \log(\cosh(\alpha t)) \end{cases}$ $\boxed{\alpha>0 \text{ parameter}}$

for large t, φ is nearly affine, A often ill-conditioned or singular
minimization can be $\boxed{\text{VERY}}$ slow.

General idea: $\varphi(t) = \min\limits_{s} \{ Q(t,s) + \Psi(s) \}$

dual variable

where $\forall s$, $t \mapsto Q(t,s)$ is $\boxed{\text{quadratic}}$ $\qquad \mathbb{R}^{\#I \overset{\text{def}}{=} r}$

Alternate minimization of the augmented func. $F(u,b)$

$$F(u,b) = \|Au - v\|_2^2 + \beta \sum_{i\in I} \left(Q(\|G_i u\|, b_i) + \Psi(b_i) \right)$$

iterates

① $b^{(k)}$ such that $F(u^{(k-1)}, b^{(k)}) \le F(u^{(k-1)}, b) \quad \forall b \in \mathbb{R}^r$

② $u^{(k)}$ such that $F(u^k, b^{(k)}) \le F(u, b^{(k)}) \quad \forall u \in \mathbb{R}^n$

two main forms of interest .

multiplicative ("$*$") : $Q(t,s) = \frac{1}{2} t^2 s \quad (s \in \mathbb{R}_+)$

additive ("$+$") : $Q(t,s) = (t-s)^2$

(A) "$*$" form

Assumptions on φ :

H1 $t \to \varphi(t)$ convex

H2 $t \to \varphi(\sqrt{t})$ concave

H3 $\varphi(t)/t^2 \xrightarrow[t\to\infty]{} 0$

Equivalence :

$$\varphi(t) = \min_{s\in\mathbb{R}_+} \left\{ \frac{1}{2} s t^2 + \psi(s) \right\}$$

$$\psi(s) = \max_{t\in\mathbb{R}} \left\{ -\frac{1}{2} s t^2 + \varphi(t) \right\}$$

proof based on convex conjugacy.

$$F(u,b) = \|Au - v\|^2 + \frac{\beta}{2} (Gu)^T \operatorname{diag}(b) Gu + \beta \sum_i \psi(b_i)$$

① $$b^{(k)}(i) = \begin{cases} \dfrac{\varphi'(\|G_i u^{(k-1)}\|)}{\|G_i u^{k-1}\|} & \text{if } G_i u^{(k-1)} \neq 0 \\ \varphi''(0^+) & \text{if } G_i u^{(k-1)} = 0 \end{cases}$$

② $$u^{(k)} = \left(2A^TA + \beta G^T \operatorname{diag} b^{(k)} G\right)^{-1} 2A^T v.$$

$$\boxed{\ker G \cap \ker A = \{0\}}$$

This method amounts to a special form of Quasi-Newton.

(B) $\underline{(+)\ \text{form}}$

H1. - φ is convex, C^1 $\qquad$ $t>0$ bc. $\varphi(|t|)$

$G_i u = u_{i_x} - u_{i_y}$

H2 $\boxed{\frac{t^2}{2} - \varphi(t) \text{ is convex}} \to \varphi'(0)=0$

H3 $\lim_{|t|\to\infty} \varphi(t)/t^2 < \frac{1}{2}$

Equivalence :
$$\boxed{\begin{aligned} \psi(s) &= \max_{t\in\mathbb{R}} \left\{ -\frac{1}{2}(t-s)^2 + \varphi(t) \right\} \\ \varphi(t) &= \min_{s\in\mathbb{R}} \left\{ \frac{1}{2}(t-s)^2 + \psi(t) \right\} \end{aligned}}$$
(*)

$$F(u,b) = \|Au-v\|^2 + \frac{\beta}{2}\|Gu-b\|^2 + \beta \sum_{i=1}^{r} \psi(b_i)$$

$$\begin{cases} b_{(i)}^{(k)} = G_i u^{(k-1)} - \varphi'(G_i u^{(k-1)}) & 1 \le i \le r \\ u^{(k)} = (\underbrace{2A^TA + \beta G^TG}_{\text{fixed}})^{-1} (2A^Tv + \beta G^T b^{(k)}) \end{cases}$$

Rmk amounts to a form of Quasi-Newton.

$$\psi(s) = \sup_{t\in\mathbb{R}} \Big(\underbrace{-\tfrac{1}{2}(t-s)^2}_{-\frac{1}{2}t^2 + ts - s^2\frac{1}{2}} + \varphi(|t|) \Big) \iff$$

$$\psi(s) + \frac{s^2}{2} = \sup_{t\in\mathbb{R}} \Big(\langle t,s\rangle - \underbrace{\big(+\tfrac{1}{2}t^2 - \varphi(|t|)\big)}_{\text{convex by H2}} \Big)$$

Will use $F^{**} = F$

$$\Leftrightarrow +\frac{1}{2}t^2 - \varphi(|t|) = \sup_{s\in\mathbb{R}} \Big(\underbrace{\langle t,s\rangle}_{=ts} - \big(\tfrac{s^2}{2} + \psi(s)\big) \Big)$$

$$-\varphi(t) = \sup_{s\in\mathbb{R}} \Big(-\frac{1}{2}s^2 + ts - \frac{1}{2}t^2 - \psi(s) \Big)$$

$$\Longleftrightarrow \quad \varphi(t) = \inf_{s\in\mathbb{R}} \left(\tfrac{1}{2}(s-t)^2 + \psi(s) \right)$$

$$\forall s \qquad \psi(s) = \sup_{t\in\mathbb{R}} \Big\{ \underbrace{\varphi(t) - \tfrac{1}{2}t^2}_{\text{concave}} + ts - \tfrac{1}{2}s^2 \Big\}$$

goes to $-\infty$ with $|t| \to \infty$
hence the max is reached.

Equiv. ⊛ on p. 96:

$\psi(\hat{s})$ is max. for $\hat{t}$
$\Longleftrightarrow$
$\varphi(\hat{t})$ is min for $\hat{s}$

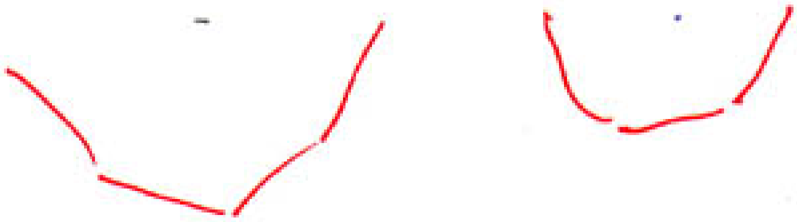
its calcalatition -
derivative $=0$

$$\varphi'(t) - t + s = 0$$

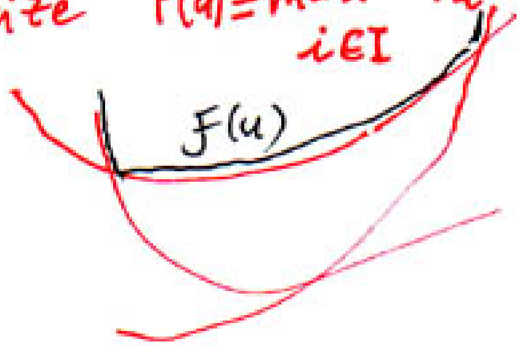
this is the formula for $b_i^{(k)}$ on p. 96.

$$\Rightarrow \quad \boxed{\hat{s} = \hat{t} - \varphi'(\hat{t})}$$

2. Introduction to non-smooth optimization



e.g. minimize $F(u) = \max\limits_{i \in I} \varphi_i(u)$ where φ_i - convex & smooth



$\Downarrow$

$F(u)$ convex
usually nonsmooth

2.1. Functions on $\mathbb{R}$

Thm ①

$\mathcal{F} : \mathbb{R} \to \mathbb{R}$ convex $\quad \operatorname{dom} \mathcal{F} = \{u : \mathcal{F}(u) < +\infty\}$

$\forall\, u_0 \in \operatorname{Int} \operatorname{dom} \mathcal{F}$, $\mathcal{F}$ admits a finite left derivative and a finite right derivative

$$\mathcal{F}'(u_0^-) = \lim_{u \uparrow u_0} \frac{\mathcal{F}(u) - \mathcal{F}(u_0)}{u - u_0} = \sup_{u < u_0} \frac{\mathcal{F}(u) - \mathcal{F}(u_0)}{u - u_0}$$

$$\mathcal{F}'(u_0^+) = \lim_{u \downarrow u_0} \frac{\mathcal{F}(u) - \mathcal{F}(u_0)}{u - u_0} = \inf_{u > u_0} \frac{\mathcal{F}(u) - \mathcal{F}(u_0)}{u - u_0}$$

Definition. $\mathcal{F} : \mathbb{R} \to \mathbb{R}$ convex

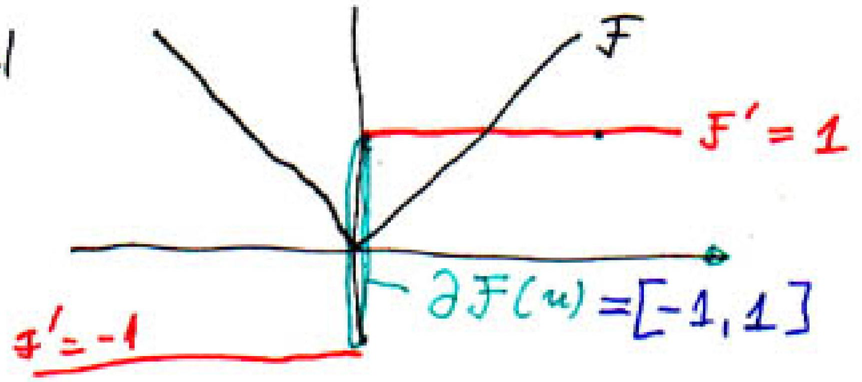
$g \in \mathbb{R}$ is a subderivative of $\mathcal{F}$ at $u \in \operatorname{dom} \mathcal{F}$ if

$$\mathcal{F}'(u^-) \le g \le \mathcal{F}'(u^+)$$

subdifferential $\partial F(u) =$ the set of all subderivatives of F at u

$$\partial F(u) = [F'(u^-), F'(u^+)]$$

Ex $F(u) = |u|$



if F is differentiable at u then $\partial F(u) = \{F'(u)\}$

- $\partial F(u_0)$ is a nonempty compact interval

Equivalent def. of subderivative

g is a subderivative of F at u_0 $\Longleftrightarrow$

$F(u) \geq F(u_0) + g(u - u_0) \quad \forall u \in \mathbb{R}$

$\Longleftrightarrow \partial F(u_0) = \{g : F(u) \geq F(u_0) + s(u - u_0), \forall u\}$

Necessary and sufficient condition for a min. at u_0

$0 \in \partial F(u_0) \Longleftrightarrow F'(u_0^-) \leq 0 \leq F'(u_0^+)$

by Thm 1

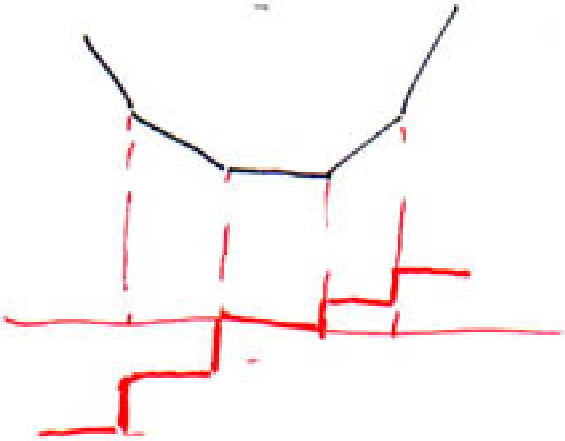
$F'(u_0^+) = \sup \{g : g \in \partial F(u_0)\}$

$F'(u_0^-) = \inf \{g : g \in \partial F(u_0)\}$

Theorem 2 $\quad F: \mathbb{R} \to \mathbb{R}$ convex

$\Longrightarrow \partial F$ is increasing on its domain.

$g_1 \le g_2$ for $\begin{cases} g_1 \in \partial \mathcal{F}(u_1) \\ g_2 \in \partial \mathcal{F}(u_2) \end{cases}$ $u_1 \le u_2$



$\partial \mathcal{F}$ is a multifunction (at some point u it is an interval)

Thm ③ the set of points where $\mathcal{F}$ fails to be differentiable is at most countable

2.2. Convex functions on $\mathbb{R}^n$

Fix $u, v \in \mathbb{R}^n$

difference quotient $q(t) \overset{def}{=} \dfrac{\mathcal{F}(u+tv)-\mathcal{F}(u)}{t}$ $t>0$

$t \to q(t)$ is increasing, bounded near 0
locally Lipschitz.

The directional derivative of $\mathcal{F}$ at u in the direction v

$$\delta \mathcal{F}(u)(v) = \lim_{t \searrow 0} q(t) = \inf_{t>0} q(t)$$

(in fact this is the right-side derivative)

The left-side $\displaystyle\lim_{t \nearrow 0} q(t) \ldots = -\delta \mathcal{F}(u)(-v)$

$$= \sup_{t<0} q(t)$$

Since g increasing $\quad -\delta F(u)(-v) \leq \delta F(u)(v)$

if F differentiable at u ~~(along v)~~

$$\delta F(u)(v) = \langle \nabla F(u), v \rangle = -\delta F(u)(-v)$$

linear with v

Subdifferential

$$\partial F(u) = \{ g \in \mathbb{R}^n : \langle g, v \rangle \leq \delta F(u)(v), \forall v \in \mathbb{R} \}$$

subgradient of F at u

$$\Rightarrow \delta F(u)(v) = \sup \{ \langle g, v \rangle : g \in \partial F(u) \}$$

called a support function for $\{\partial F(u)\}$

equivalent definition

$$\partial F(u) \stackrel{def}{=} \{ g \in \mathbb{R}^n : \langle g, v-u \rangle \leq F(v) - F(u) \ \forall v \in \mathbb{R}^n \}$$

- $v \to \delta F(u)(v)$ finite, convex, positively homogeneous $(\alpha > 0 \quad f(\alpha x) = \alpha f(x))$

Rademacher's theorem

$F : \mathbb{R}^n \to \mathbb{R}$ convex. The subset of Int dom F where F fails to be differentiable is of zero Lebesgue measure

These points are called kinks.

Often the minimum of F is at a kink.

Necessary and sufficient conditions for a minimum

Thm. $\mathcal{F} : \mathbb{R}^n \to \mathbb{R}$ convex. $\mathcal{F}$ has a minimum at $\hat{u}$

$$\iff \delta \mathcal{F}(\hat{u})(v) \geq 0 \quad \forall v \in \mathbb{R}^n$$

$$\iff 0 \in \partial \mathcal{F}(\hat{u})$$

KKT conditions for a minimum under constraint can also be formulated. (Lagrangian)

* see Book Hiriart Urruty - Lemarechal (in the bibliolist).

Sub-gradient minimization method.

for $k = 1, 2 \dots$

1) if $0 \in \partial \mathcal{F}(u_k)$ stop, else (unlikely)
2) find $g_k \in \partial \mathcal{F}(u_k)$
3) $u_{k+1} = u_k - \rho_k \dfrac{g_k}{\| g_k \|}$ $\quad \rho_k > 0$

not necessarily a descent direction

Thm assum $\mathcal{F}$ convex, min. at $\hat{u}$

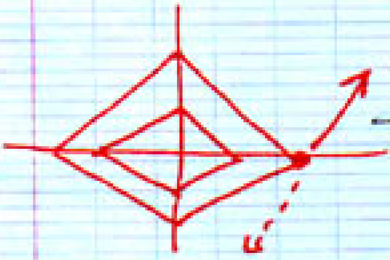
If $(\rho_k)_{k \geq 0}$ satisfies $\boxed{\sum_k \rho_k = +\infty \text{ and } \sum_k g_k^2 < +\infty}$

Then $u_k \to \hat{u}$

Une direction opposée à un sous-grad n'est pas forcément une direction de descente

d - dir de descente ssi $\boxed{\langle d.g\rangle < 0\ ,\ \forall g \in \partial F(u)}$

$F(u_1, u_2) = |u_1| + 2|u_2|$



$s = (1,2) \in \partial F(\cdot, 0)$

$-s$ n'est pas de descente

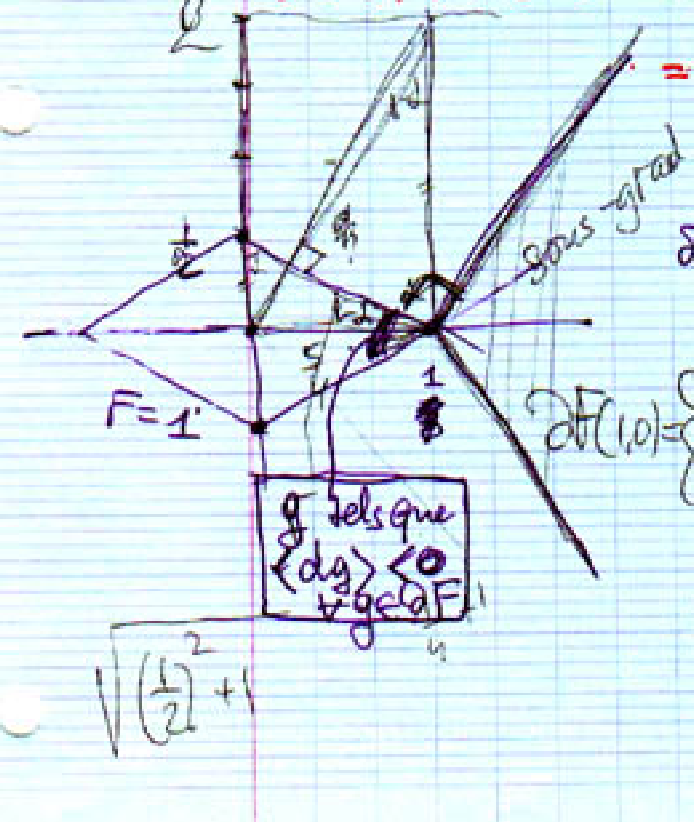
CNS de minimum

$F(u) \geq F(\hat{u})\ \forall u \iff 0 \in \partial F(\hat{u})$

évident de: $g \in \partial F(\hat{u}) \iff F(u) \geq F(\hat{u}) + \langle g, u - \hat{u}\rangle$

$0 \in \partial F(\hat{u}) \iff \delta F(\hat{u})(v) \geq 0\ \ \forall v \in \mathbb{R}^n$

$= \{ g : \langle g, v\rangle \leq \delta F(\hat{u})(v)\ \ \forall v \}$



$$\delta F(1,0) = \lim_{t \searrow 0} \frac{|1 + tv_1| - 1 + 2|tv_2|}{t}$$

$$= v_1 + 2|v_2|$$

$$g :$$

$$\partial F(1,0) = \{ v_1 + 2|v_2| \geq g_1 v_1 + g_2 v_2\ \ \forall v \in \mathbb{R}^2$$

$g_1 = 1$

$|g_2| \leq 2$

$= \{1\} \times [-2, 2]$