SG-1

Mar 21 2016

In preparation for ADMM lecture:

Subgradients + Subdifferential of Convex Functions

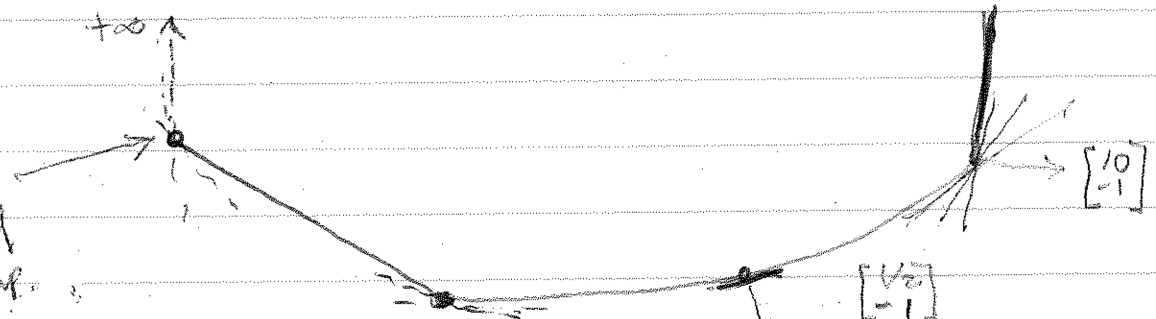
Oddly, not in BV.

Assume f is convex + proper: $\exists x$ s.t. $f(x) < +\infty$ + closed: all sublevel sets are closed.

$f: \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$. $\forall x$, $f(x) > -\infty$.

Def $y \in \mathbb{R}^n$ is a subgradient of f at x if

$$f(x+z) \geq f(x) + y^T z \quad \forall z \in \mathbb{R}^n$$



must include end pt of jump to ∞ or beyond

But $f = \begin{cases} +\infty & x \leq 0 \\ -\log x & x > 0 \end{cases}$ IS CLOSED although drawing of f is NOT

$n = 1$: y is the slope of a ~~two~~ line passing through $(x, f(x))$ and lying underneath the graph of x

$n \geq 1$ $\boxed{\text{NOT } 2}$ $\begin{bmatrix} y \\ -1 \end{bmatrix}$ is normal to a hyperplane in $\mathbb{R}^{n+1}$ passing through $\begin{bmatrix} x \\ f(x) \end{bmatrix}$ and lying below the graph of f.

a proper convex fun. is closed iff it is LSC.

The set of all subgradients of f at x is denoted $\partial f(x)$, the SUBDIFFERENTIAL of f at x.

e.g. $f(x) = |x|$, $\partial f(0) = [-1, 1]$

If f is differentiable at x then
$$\partial f(x) = \{\nabla f(x)\}.$$
In fact this is IFF.

Note for $x \in \mathrm{dom} f$, $\partial f(x)$ is always a
CLOSED, CONVEX, NON-EMPTY, COMPACT set.

e.g. $f(x) = \max_{1 \le i \le n} (x_i)$ $\quad (= x_{[1]}$ in BV notation$)$

What is $\partial f(x)$ for $x = \begin{bmatrix} 1 \\ 3 \\ 2 \\ 2 \\ 3 \end{bmatrix}$? Need

$$\max\left(\begin{bmatrix} 1+z_1 \\ 3+z_2 \\ 2+z_3 \\ 2+z_4 \\ 3+z_5 \end{bmatrix}\right) \ge 3 + y^T z \quad \forall z \in \mathbb{R}^n$$

Clearly $e_1 \notin \partial f(x)$ as RHS is $3+z_1$ (take $z=e_1$)
$e_2 \in \partial f(x)$ as RHS is $3+z_2$.

In fact $\partial f(x) = \mathrm{conv}(e_2, e_5) = \left\{\begin{bmatrix} 0 \\ \tau \\ 0 \\ 0 \\ 1-\tau \end{bmatrix} : \tau \in [0,1]\right\}$

Does this remind you of something?
Answer: (Fenchel) conjugate.

THM (Fenchel-Young)
$$f(x) + f^*(y) \ge x^T y$$
with equality IFF $y \in \partial f(x)$.
Pf: exercise in Borwein + Lewis.

S63

Relationship to Directional Derivative

$$f'(x;d) = \lim_{t \downarrow 0} \frac{f(x+td) - f(x)}{t}$$

Thm $\quad y \in \partial f(x)$ iff $y^T d \le f'(x;d) \ \forall d \in \mathbb{R}^n$.

Pf Ex in Borwein + Lewis.

Chain Rule - simplest version.

More general versions: Borwein + Lewis p. 52
Rockafellar p. 225.

Let $f: \mathbb{R}^n \to \mathbb{R}$ convex, $\operatorname{dom} f = \mathbb{R}^n$.
Let $A \in \mathbb{R}^{n \times m}$, $b \in \mathbb{R}^m$.
Let h be the convex function on $\mathbb{R}^m$ defined by

$$h(\xi) = f(A\xi + b) \quad \xi \in \mathbb{R}^m.$$

Then $\partial h(\xi) = \underbrace{A^T \partial f(A\xi + b)}_{\text{means}}$

$$\{A^T y;\ y \in \partial f(A\xi + b)\}$$

Works even if A does not have full rank
e.g. $A = 0$.

Optimality Condition

$$0 \in \partial f(x) \iff x \text{ is global minimizer of } f.$$

Pf: immediate from def'n.