Lecture 20

Last time
- Preconditioning
- PDHG
- Primal-dual guarantee

Today
- P-D guarantee continued
- Fenchel subdifferential

Primal-dual guarantee continued

Theorem: Suppose that f, g are closed, convex and proper. Further, assume (P)-(D) exhibit strong duality with at least one primal dual solution. Let $z_k = (x_k, y_k)$ be a sequence defined via

positive definite

$$\underbrace{H}_{\text{positive definite}} (z_k - z_{k+1}) \in \underbrace{\begin{bmatrix} \partial_x L(x_{k+1}, y_{k+1}) \\ \partial_y (-L(x_{k+1}, y_{k+1})) \end{bmatrix}}_{\mathcal{F}(z_{k+1})}$$

Denote $\bar{z}_k = (\bar{x}_k, \bar{y}_k) = \frac{1}{k} \sum_{i=1}^{k} z_i$, then for any $k \geq 1$ and any primal-dual solution (x^*, y^*)

$$L(\bar{x}_k, y^*) - L(x^*, \bar{y}_k) \leq \frac{\|z^* - z_0\|_H^2}{2k}$$

$\|z\|_H^2 = \langle z, Hz \rangle$

Proof: Denote $\overbrace{\begin{bmatrix} u_k \\ v_k \end{bmatrix}}^{w_k} = H(z_k - z_{k+1}) \in \mathcal{F}(z_{k+1})$. Then by definition of the subdifferential

$$\begin{aligned}
&L(x_{k+1}, y^*) - L(x^*, y_{k+1}) \\
&= \left(L(x_{k+1}, y^*) - L(x_{k+1}, y_{k+1})\right) \\
&\quad + \left(L(x_{k+1}, y_{k+1}) - L(x^*, y_{k+1})\right) \\
&\leq \langle v_k, y_{k+1} - y^* \rangle + \langle u_k, x_{k+1} - x^* \rangle \\
&= \langle w_k, z_{k+1} - z^* \rangle \\
&= \langle z_k - z_{k+1}, z_{k+1} - z^* \rangle_H \\
&= \frac{1}{2}\|z_k - z^*\|_H^2 - \frac{1}{2}\|z_{k+1} - z^*\|_H^2 \\
&\qquad - \frac{1}{2}\|z_k - z_{k+1}\|_H^2.
\end{aligned}$$

$$\leq \frac{1}{2}\|z_k - z^*\|_H^2 - \frac{1}{2}\|z_{k+1} - z^*\|_H^2 .$$

Using convexity and concavity of $L(\cdot, y^*)$ and $L(x^*, \cdot)$ yield

$$L(\bar{x}_k, y^*) - L(x^*, \bar{y}_k)$$
$$\leq \frac{1}{k}\sum_{i=0}^{k-1} L(x_{i+1}, y^*) - L(x^*, y_{i+1})$$
$$\leq \frac{1}{2k}\left(\|z_0 - z^*\|_H^2 - \frac{1}{2}\|z_k - z^*\|_H^2\right)$$
$$\leq \frac{\|z_0 - z^*\|_H^2}{2k} .$$

$\square$

New topic: Variational Analysis

Next, we move away from algorithms and go back to understanding sets and functions (as we did at the start of class). Our goal is to generalize some of the ideas we consider for convex data.

Frechet subdifferential

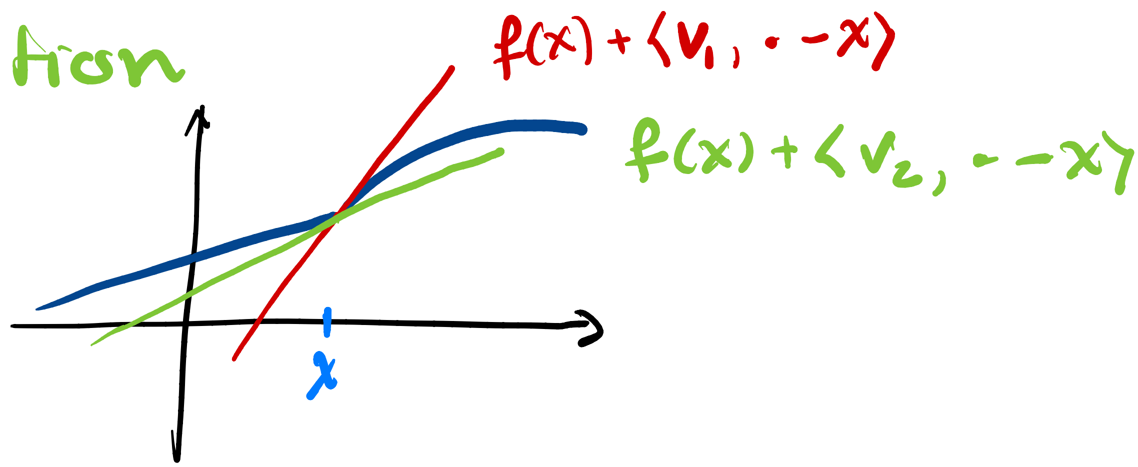
For smooth functions we defined gradients via linear approximations. For non-smooth convex functions we used minorizing functions. Here we combined these two for general functions.

Def: For $f: E \to \mathbb{R} \cup \{+\infty\}$ finite at x, we say that $y \in \partial f(x)$ is a Frechet subgradient if and only if

$$f(x+z) \geq f(x) + \langle y, z \rangle + o(\|z\|) \quad \text{as } z \to x$$

where $o(z)$ is a term such that $\frac{o(\|z\|)}{\|z\|} \to 0$ when $z \to 0$. $\dashv$

Intuition



Arguably, there where three things

we liked about convex subdifferentials (1) optimality conditions, (2) they exist in int dom f, and (3) very nice calculus rules. Next we aim to explore equivalent properties for the Frechet subdifferential.

Lemma (0) Let x be a local minimizer of $f: E \to \mathbb{R} \cup \{+\infty\}$ and suppose $x \in \operatorname{dom} f$. Then, $0 \in \partial f(x)$ holds. $\dashv$

Note that from now on we will only be able to distinguish local properties since ∂f is defined locally. We have been abusing notation since we used the symbol "∂f" for convex subdifferentials. But it is for a good reason:

Lemma (+): Suppose $h: E \to \mathbb{R} \cup \{+\infty\}$ and $g: E \to \mathbb{R}$ differentiable. Then,

$$\underset{\text{Frechet sub.}}{\partial (h+g)} = \underset{\text{convex sub.}}{\partial h(x)} + \nabla g(x) \quad \forall x \in \operatorname{dom} h.$$

Moreover the equality still holds when h is merely proper with the Frechet subdifferential on the right. $\dashv$

The next question is about existance. General nonsmooth functions can be pretty nasty (Weierstrass function). So we can only establish that the domain of the subdifferential is dense in dom f.

Lemma (Density of the subdifferential)
Consider a proper, closed function $f: E \to \mathbb{R} \cup \{+\infty\}$. Then, the set $\operatorname{dom} \partial f$ is dense in $\operatorname{dom} f$.

Proof: Fix a point in dom f. We show the existance of a sequence $x_i \in \operatorname{dom} f$ with $x_i \to x$. Since f is closed there is a closed ball $Q = \overline{B_\varepsilon(x)}$ s.t $f(y) \geq f(x) - 1 \quad \forall y \in Q$ (Why?)
Consider the sequence of potentials given by $f_n(y) = f(y) + \iota_Q(y) + \frac{n}{2}\|y - x\|^2$.

Fact from Nonlinear 1: If a closed function $h: E \to \mathbb{R} \cup \{\infty\}$ is such that $h(z) \to \infty$ when $\|z\| \to \infty$. (coercivity)
Then, h attains a minimizer. $\dashv$

Clearly f_n is coercive so $\exists\, x_n \in \operatorname{argmin} f_n$.
Note that

$$f(x_n) + \frac{n}{2}\|x_i - x\|^2 \le f(x)$$

$$\Rightarrow \quad \|x_i - x\|^2 \le \frac{2}{n}\big(f(x) - f(x_n)\big) \le \frac{2}{n}.$$

Thus, $x_n \to x$. Thus for large n, we have $x_n \in \operatorname{int} Q$, and so x_n minimizes $f(y) + \frac{n}{2}\|y - x\|^2$ without the indicator. By lemmas (0) and (+) we conclude that

$$0 \in \partial f(x_n) + n(x_n - x)$$

and so $\partial f(x_n) \ne \emptyset$. $\square$