Lecture 9

Recall:

<u>Def</u>: Given nvs $X$, $f: X \to \mathbb{R} \cup \{+\infty\}$ proper, $x \in D(f)$, the subdifferential of $f$ at $x$ is the set valued operator

$$\partial f(x) = \{ y \in X^* : f(x') \geq f(x) + \langle y, x' - x \rangle + o(\|x' - x\|), \text{ as } x' \to x \}$$

If $x \notin D(f)$, $\partial f(x) = \emptyset$.

<u>Thm</u>: If $f$ is differentiable at $x \in D(f)$, then $\partial f(x) = \{ \nabla f(x) \}$.

<u>Prop</u>: If $f: X \to \mathbb{R} \cup \{+\infty\}$ is convex and $x \in D(f)$, then

$$\partial f(x) = \{ y \in X^* : f(x') \geq f(x) + \langle y, x' - x \rangle \ \forall x' \in X \}$$

<u>Prop</u>: Given $f: X \to \mathbb{R} \cup \{+\infty\}$ proper and convex,

$$y \in \partial f(x_0) \iff f(x_0) + f^*(y_0) = \langle y_0, x_0 \rangle$$

**Prop:** Suppose $f$ is convex and it is lsc at $x_0 \in D(f)$. Then $x_0 \in \partial f^*(y_0) \Rightarrow y_0 \in \partial f(x_0)$

**Thm:** Suppose $f$ is proper, convex, and lsc. Then $y \in \partial f(x) \Leftrightarrow x \in \partial f^*(y)$.

**Def:** For $C \subseteq X$, its characteristic fn is

$$\chi_C(x) := \begin{cases} 0 & \text{if } x \in C \\ +\infty & \text{otherwise} \end{cases}$$

---

## Primal / Dual Optimization Problems

Goal of convex optimization: given $f$ convex, $C$ convex, solve

$$\inf_{x \in C} f(x) = \inf_{x \in X} f(x) + \chi_C(x)$$

Key trick: observe the behavior of this optimization problem under perturbations.

Def: Given normed vector spaces $\mathcal{X}$ and $\mathcal{U}$ and a convex function $F: \mathcal{X} \times \mathcal{U} \to \mathbb{R} \cup \{+\infty\}$.

primal problem: $\mathcal{P}_0 := \inf_{x \in \mathcal{X}} f(x)$ , $f(x) = F(x,0)$

dual problem: $D_0 := \sup_{v \in \mathcal{U}^*} g(v)$, $g(v) = -F^*(0,v)$.

Remark: $\mathcal{X} \times \mathcal{U}$ is a nvs with dual $\mathcal{X}^* \times \mathcal{U}^*$ and duality pairing

$$\langle (y,v), (x,u) \rangle = \langle y,x \rangle + \langle v,u \rangle$$

In this case, Young's inequality is

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

In particular,

$$f(x) - g(v) \geq 0 \iff f(x) \geq g(v) \quad \forall x \in \mathcal{X}, v \in \mathcal{U}^*$$

$$\iff \mathcal{P}_0 \geq \mathcal{D}_0$$

Thus, we will seek conditions on F that ensure $\mathcal{P}_0 = \mathcal{D}_0$, i.e. "there is no duality gap"

Thm (Equivalence of Primal and Dual Problems): Given $F : \mathcal{X} \times \mathcal{U} \to \mathbb{R} \cup \{+\infty\}$ convex, suppose $\mathcal{P}_0 < +\infty$. "the primal problem is feasible"

Define the inf-projection $\mathcal{P}(u) := \inf_{x \in \mathcal{X}} F(x,u)$. Then

(i) $\mathcal{P}_0 = \mathcal{D}_0 \iff \mathcal{P}$ is lsc at $u = 0$.

(ii) $\mathcal{P}_0 = \mathcal{D}_0$ and a maximizer of dual problem exists $\iff \partial \mathcal{P}(0) \neq \emptyset$.

Pf ① Show $\mathcal{P}(u)$ is proper and convex.

By assumption $\mathcal{P}(0) = \mathcal{P}_0 < +\infty$.

$$\mathcal{P}(\underbrace{(1-\alpha)u_0+\alpha u_1}_{u_\alpha}) = \inf_{x\in\mathcal{X}} F(x, u_\alpha)$$

$$\leq F(x_0, u_\alpha) \qquad \forall\, x_0 \in \mathcal{X}$$

$$= F((1-\alpha)x_0+\alpha x_0, (1-\alpha)u_0+\alpha u_1), \forall x_0$$

$$\leq (1-\alpha)F(x_0, u_0) + \alpha F(x_0, u_1), \forall x_0$$

Taking inf over $x_0$, $\mathcal{P}(u_\alpha) \leq (1-\alpha)\mathcal{P}(u_0) + \alpha\,\mathcal{P}(u_1)$.

② $\mathcal{P}_0 = D_0 \Leftrightarrow \mathcal{P}(0) = \mathcal{P}^{**}(0)$

By definition,

$$\underline{\underline{\mathcal{P}^*(v)}} = \sup_{u\in\mathcal{U}} \langle v, u\rangle - \mathcal{P}(u)$$
$$= \sup_{\substack{u\in\mathcal{U}\\ x\in\mathcal{X}}} \langle v, u\rangle - F(x, u)$$

$$= \sup_{\substack{u\in\mathcal{U}\\ x\in\mathcal{X}}} \langle 0, x\rangle + \langle v, u\rangle - F(x, u)$$

$$= F^*(0, v)$$
$$= \underline{\underline{-g(v)}}$$

$$\mathcal{P}^{**}(u) = \sup_{v\in\mathcal{U}^*} \langle v, u\rangle + g(v)$$

$$\mathcal{P}^{**}(0) = \sup_{v \in U^*} g(v) = D_0 .$$

Since $\mathcal{P}(0) = P_0$, this gives the result.

③ We now show (i).

Since $\mathcal{P}$ is convex, $0 \in D(\mathcal{P})$, Fenchel-Moreau ensures $\mathcal{P}(0) = \mathcal{P}^{**}(0) \iff \mathcal{P}$ is lsc at $0$.

④ We now show (ii).
Suppose $v_* \in \partial\mathcal{P}(0)$.

Since $\mathcal{P}$ is convex, $0 \in D(\mathcal{P})$, for any sequence $u_n \to 0$, we have

$$\begin{aligned}\liminf_{n\to\infty} \mathcal{P}(u_n) &\geq \liminf_{n\to\infty} \mathcal{P}(0) + \langle v_*, u_n - 0 \rangle \\ &\geq \mathcal{P}(0)\end{aligned}$$

Thus, $\mathcal{P}$ is lsc at zero, so by part (i), $P_0 = D_0$.

Furthermore, $v_* \in \partial\mathcal{P}(0)$ ensures
$$\mathcal{P}(0) + \mathcal{P}^*(v_*) = \langle v_*, 0 \rangle = 0$$

Thus,
$$\sup_{v \in U^*} g(v) = D_0 = P_0 = \overset{\text{step ①}}{-\mathcal{P}^*(v_*)} = g(v_*).$$

Converseley, suppose $\mathcal{P}_0 = \mathcal{D}_0$ and $v_*$ is a maximizer of the dual problem.

By (i), $\mathcal{P}_0 = \mathcal{D}_0 \Rightarrow \mathcal{P}$ lsc at $0$.
By Fenchel-Moreau,

$$\mathcal{P}(0) = \mathcal{P}^{**}(0) = \sup_{v \in \mathcal{U}^*} \overbrace{\langle v, 0 \rangle - \mathcal{P}^*(v)}^{g(v)}$$
$$= \sup_{v \in \mathcal{U}^*} g(v)$$
$$= g(v_*)$$
$$= -\mathcal{P}^*(v_*).$$

Thus, equality must hold in Young's inequality, $v_* \in \partial \mathcal{P}(0)$. $\square$

## Kantorovich Duality

$(\mathcal{X}, d)$ Polish space
$\mu, \nu \in \mathcal{P}(\mathcal{X})$
$c : \mathcal{X} \times \mathcal{X} \to [0, +\infty)$ lower semicontinuous

$$\min_{\gamma : \gamma \in \Gamma(\mu,\nu)} \overbrace{\int_{X \times X} c(x^1, x^2)\, d\gamma(x^1, x^2)}^{\mathbb{K}(\gamma)} \qquad (KP)$$

This is a convex optimization problem.
To find its dual...

① Rewrite as unconstrained optimization problem.
② Identify "perturbation" function $F(x,u)$ so that $(KP) = D_0$.

We will do this via introducing a Lagrange multiplier.

Recall: Lagrange multipliers in Calculus...

Given $A \in M_{m \times n}(\mathbb{R})$, $b \in \mathbb{R}^m$

$$\inf_{Ax=b} f(x) = \inf_{x \in \mathbb{R}^n} f(x) + \chi_{\{x : Ax = b\}}(x)$$

$$= \inf_{x \in \mathbb{R}^n} \underbrace{\sup_{\lambda \in \mathbb{R}^m} f(x) + \langle \lambda, Ax - b \rangle}_{= \begin{cases} f(x) & \text{if } Ax = b \\ +\infty & \text{if } Ax \neq b \end{cases}}$$

## Relation to Primal/Dual Problem:

primal problem: $P_0 := \inf_{x \in \mathcal{X}} f(x)$, $f(x) = F(x, 0)$

dual problem: $D_0 := \sup_{v \in \mathcal{U}^*} g(v)$, $g(v) = -F^*(0, v)$.

Note that

$$g(v) = -F^*(0, v) = -\sup_{(x,u) \in \mathcal{X} \times \mathcal{U}} \langle 0, x \rangle + \langle v, u \rangle - F(x, u)$$

$$= \inf_{(x,u) \in \mathcal{X} \times \mathcal{U}} F(x, u) - \langle v, u \rangle$$

Thus,

$$D_0 = \sup_{v \in \mathcal{U}^*} \inf_{(x,u) \in \mathcal{X} \times \mathcal{U}} F(x, u) - \langle v, u \rangle$$

saddle point problem

Moral: Introducing a Lagrange multiplier to remove constraint can shed light on how to choose perturbation function $F(x, u)$.

How to do this for (KP)?

Recall: for $(X,d)$ locally compact,

| | | |
|---|---|---|
| Banach space | $(C_0(X), \|\cdot\|_\infty)$ | $(C_b(X), \|\cdot\|_\infty)$ |
| Dual space | $(\mathcal{M}^s(X), \|\cdot\|_{TV})$ | [big space, containing $(\mathcal{M}^s(X), \|\cdot\|_{TV})$] |
| Weak-* topology | wide topology | narrow topology |

If $(X,d)$ compact, all of above notions coincide.

Given $\varphi \in C_b(X)$, $\mu \in \mathcal{M}^s(X)$, let

$$\langle \mu, \varphi \rangle = \int_X \varphi(x)\, d\mu(x).$$

Fact: $\mu = \nu \iff \langle \mu, \varphi \rangle = \langle \nu, \varphi \rangle \quad \forall \varphi \in C_b(X)$.

<u>Lemma</u>: Given $\mu, \nu \in \mathcal{P}(X)$, $\gamma \in \mathcal{M}(X \times X)$

$$\sup_{\varphi, \psi \in C_b(X)} \langle \mu - \pi^1 \# \gamma, \varphi \rangle + \langle \nu - \pi^2 \# \gamma, \psi \rangle = \chi_{\Gamma(\mu,\nu)}(\gamma).$$

Pf: The equality is clearly true if $\gamma \in \Gamma(\mu,\nu)$.

If $\gamma \in \mathcal{M}(X \times X) \setminus \Gamma(\mu,\nu)$, then either $\pi^1 \# \gamma \neq \mu$ or $\pi^2 \# \gamma \neq \nu$. WLOG $\pi^1 \# \gamma \neq \mu$, so $\exists\ \varphi_0 \in C_b(X)$ so that $\langle \mu - \pi^1 \# \gamma, \varphi_0 \rangle = c_0 \neq 0$.

Define, for $n \in \mathbb{N}$,

$$\varphi_n(x) = \mathrm{sgn}(c_0)\, n\, \varphi_0(x) \in C_b(X)$$

Then

$$\lim_{n\to\infty} \langle \mu - \pi^1 \# \gamma, \varphi_n \rangle = \lim_{n\to\infty} \mathrm{sgn}(c_0)\, n \overbrace{\langle \mu - \pi^1 \# \gamma, \varphi_0 \rangle}^{c_0}$$

$$= \lim_{n\to\infty} |c_0|\, n$$

$$= +\infty \qquad \square$$