Optimal Transport: Lecture #7

- Continuing our analysis of the OT problem for quadratic costs, we introduce the so-called double convexification trick:

$$x \cdot y \leq \varphi(x) + \psi(y) \qquad \forall \text{ a.e. } x, y$$

$$\Rightarrow x \cdot y - \varphi(x) \leq \psi(y) \qquad \forall \text{ a.e. } x, y$$

$$\Rightarrow \sup_x \{ x \cdot y - \varphi(x) \} \leq \psi(y) \qquad \forall y$$

But, we can identify the LHS as the Legendre-Fenchel transform of φ, namely

$$\varphi^*(y) = \sup_x \{ \langle x, y \rangle - \varphi(x) \},$$ noting that in $\mathbb{R}^d$, $\langle x, y \rangle = x \cdot y$ ($\mathbb{R}^d$ is self-dual in this sense).

- To summarize, $\psi(y) \geq \varphi^*(y)$. Applying J to the pairs (φ, ψ) and (φ, φ^*) and recalling J is linear, we have

$$J(\varphi, \psi) \geq J(\varphi, \varphi^*).$$

- By an analogous argument, we know $(\varphi, \varphi^*) \in \tilde{\Phi}$, so again

$$x \cdot y \leq \varphi(x) + \varphi^*(y)$$

$$\Rightarrow \sup_y \{ x \cdot y - \varphi^*(y) \} \leq \varphi(x)$$

and we can identify the LHS as the Legendre-Fenchel transform of ℓ^*, namely the double dual of ℓ!

· To summarize, $\ell^{**}(x) = \sup_y \{x\cdot y - \ell^*(y)\} \le \ell(x)$

$$\overset{\text{linearity of } J}{\Longrightarrow} \quad J(\ell^{**}, \ell^*) \le J(\ell, \ell^*)$$

· Having established the sequence of inequalities $J(\ell, \psi) \le J(\ell, \ell^*) \le J(\ell^{**}, \ell^*)$,

we can infimize as we want over (ℓ, ψ) and get

$$\inf_{(\ell,\psi)\in\hat{\Phi}} J(\ell, \psi) \ge \inf_{\ell\in L^1(d\mu)} J(\ell^{**}, \ell^*).$$

· Why does this help us analyze the RHS? Well, as long as $(\ell^{**}, \ell^*) \in L^1(d\mu)\times L^1(d\nu)$, then we clearly have the other inequality:

$$\inf_{(\ell,\psi)\in\tilde{\Phi}} J(\ell, \psi) \le \inf_{\ell\in L^1(d\mu)} J(\ell^{**}, \ell^*),$$

because the infimizing set on the LHS is larger than the infimizing set on the RHS.

· So, the punch line is

$$\inf_{(\ell,\psi)\in\tilde{\Phi}} J(\ell, \psi) = \inf_{\ell\in L^1(d\mu)} J(\ell^{**}, \ell^*),$$

which is an optimization over a nicer space of functions; as we shall see, ℓ^*, ℓ^{**} are both (i) convex

(ii) lower semicontinuous.

//

- To see this and some other nice properties we'll use later, we will make a quick tour of convex analysis. Recall $\ell : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ is convex if $\forall x, y \in \mathbb{R}^d$, $t \in [0,1]$,

$$\ell(tx + (1-t)y) \leq t\,\ell(x) + (1-t)\,\ell(y)$$

- Convex functions are continuous, locally Lipschitz, and $\nabla \ell$ exists a.e.
- Convex functions are above their tangents, in the sense that

$$\forall x \in \mathbb{R}^d \text{ s.t. } \nabla \ell(x) \text{ exists, and } \forall z \in \mathbb{R}^d, \quad \ell(z) \geq \ell(x) + \nabla \ell(x) \cdot [z - x]$$

A useful corollary is that, apply this also to z (assuming $\nabla \ell(z)$ exists) and adding the inequalities gives monotonicity:

$$\langle \nabla \ell(x) - \nabla \ell(y), x - z \rangle \geq 0.$$

- One can handle non-differentiable ℓ through a weaker notion of subdifferentiability.

<u>Defn</u>: Let $\ell : \mathbb{R}^d \to \mathbb{R}$ be convex. The subdifferential of ℓ is a set-valued

function $\partial\ell$, so that

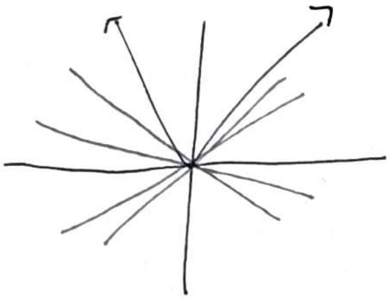
we'll use $\cdot$ and $\langle , \rangle$ when handling $\mathbb{R}^d$

$$y \in \partial\ell(x) \iff \left\{ \forall z \in \mathbb{R}^d,\ \ell(z) \geq \ell(x) + \langle y, z-x \rangle \right\}$$

We can consider the set $\text{Graph}(\partial\ell) = \{(x,y) \in \mathbb{R}^d \times \mathbb{R}^d \mid y \in \partial\ell(x)\}$.

- This covers $\nabla\ell$, in that $\nabla\ell(x)$ exists $\iff |\partial\ell(x)| = 1$ and $\partial\ell(x) = \{\nabla\ell(x)\}$. But, it exists and is non-empty for all $x \in \text{dom}(\ell)$, even the points of non-differentiability.

ex: $\ell(x) = |x|$:



- Subdifferentials have many useful properties (see pp. 55-57 in Villani), and can be understood as one way to capture some notion of differentiability for non-classically differentiable functions.
- An important connection between Legendre transforms and subdifferentials is:

Proposition (Connecting LT and ∂): Suppose $\ell : \mathbb{R}^d \to \mathbb{R}$ is lsc, convex, and proper (no $-\infty$, takes some non-$+\infty$ values). Then $\forall x, y \in \mathbb{R}^d$,

$$x \cdot y = \ell(x) + \ell^*(y) \iff y \in \partial\ell(x) \iff x \in \partial\ell^*(y).$$

Proof: Exercise/see Villani pp. 57.