Resilience under Wasserstein Norms

Last time:

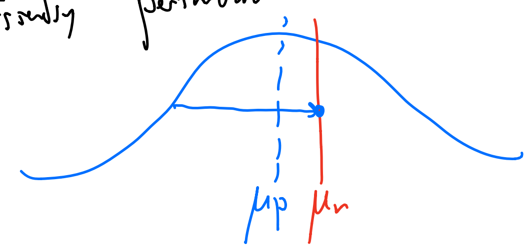
- Defined Wasserstein norm
- Kantorovich-Rubinstein duality
- Friendly perturbations:
  - "squeeze points towards mean"
  - generalization of deletions
- Started: midpoint lemma

any Lipschitz function lower-bound Wasserstein distance

This time,

- Finish midpoint lemma
- Derive resilience conditions for second moment estimation under W_1 ← Simplifying interesting problem
- Generalization of KR duality for non-Lipschitz

Friendly perturbation

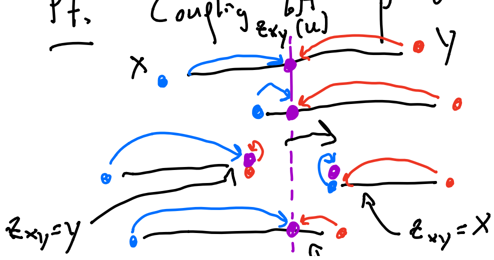


Midpoint lemma.
If $W_c(p,q) \le \varepsilon$, then $\exists\, r$ that is ε-friendly for both p and q.

Pf idea. Guess $\mu_r = u$

- Construct dist's $r(u)$ that can be obtained from p, q by squeezing towards u.
- Fixed-point lemma: There exists u s.t. $u = \mu_{r(u)}$.

Pf. Coupling b/t p and q.



$\Rightarrow$ For r, there is a coupling of cost ε where all points move towards u. (for both p, q)

$r(u)$

u

$z_{xy} = f^{-1}(u)$

$$s_{xy}(u) : \begin{cases} \max(f(x), f(y)) & : \text{if both} \le u \\ \min(f(x), f(y)) & ; \text{if both} \ge u \\ u & ; \text{else} \end{cases}$$

$$z_{xy}(u) = f^{-1}(s_{xy}(u))$$

Fixed-point lemma:

$\exists\ u$ s.t. $u = \mu_{r(u)}$.

$h(u) = \mu_{r(u)}$

$h(u)$: monotonic continuous function

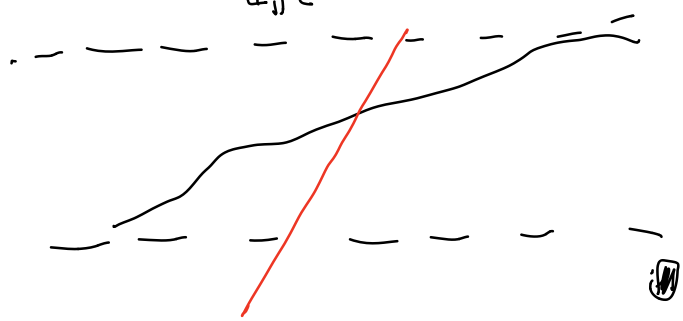
$h : \mathbb{R} \to \mathbb{R}$

$u - h(u)$: prove this crosses 0

↳ non-decreasing

$\mathbb{E}_{\pi}[\min(f(x), f(y))]$

$\mathbb{E}_{\pi}[\max(f(x), f(y))]$



Applying midpoint lemma.

2nd moment estimation under W_1

$$L(p, S) = \| \mathbb{E}_{x \sim p}[xx^T] - S \|_{op}$$

$$D(p, q) = W_1(p, q) = \min_{\pi} \mathbb{E}_{\pi}[\|x - y\|_2]$$

$f : \mathbb{R}^d \to \mathbb{R}$

↓

f?

what

p is (ρ, ε)-resilient for W_1 perturbations if stable for all ε-friendly perturbations

$$\mathcal{G}^{W_1}_{sec}(\rho, \varepsilon) = \{ p \mid |\underline{\mathbb{E}_r[\langle x, v\rangle^2]} - \mathbb{E}_p[\langle x, v\rangle^2]| \le \rho$$ whenever r is ε-friendly w/ $f_v(x) = \langle x, v\rangle^2$ and $\|v\|_2 = 1 \}$

Prop. $m(\mathcal{G}_{sec}^{W_1}(\rho,\epsilon),\epsilon) \le 2\rho.$

Pf. $p, q \in \mathcal{G}_{sec}^{W_1}(\rho,\epsilon).$

$$\|\mathbb{E}_p[xx^T] - \mathbb{E}_q[xx^T]\|_{op}$$

$$= |\mathbb{E}_p[\langle x,v\rangle^2] - \mathbb{E}_q[\langle x,v\rangle^2]| \quad \text{for some } \|v\|_2 = 1$$

$\rightarrow$ take r from midpoint lemma

$$\le |\mathbb{E}_p[\langle x,v\rangle^2] - \mathbb{E}_r[\langle x,v\rangle^2]| + |\mathbb{E}_q[\langle x,v\rangle^2] - \mathbb{E}_r[\langle x,v\rangle^2]|$$

$$\le \rho + \rho \le 2\rho. \quad \blacksquare$$

TV case: resilience: stability under deletions

bounded 2^{nd} moment: $\Rightarrow (\mathcal{O}(\sigma\sqrt{\epsilon}), \epsilon)$-resilient

sub-Gaussian $\Rightarrow (\mathcal{O}(\sigma\epsilon\sqrt{\log(1/\epsilon)}), \epsilon)$-resilient

Orlicz norm:

$$\mathbb{E}\left[\psi\left(\frac{|v^T x|}{\sigma}\right)\right] \le 1 \quad \forall \|v\|_2 \le 1$$

$$\Rightarrow \boxed{\rho = \sigma\epsilon\,\psi^{-1}(1/\epsilon)}$$

Lemma. $\tilde{\psi}(x) = x\psi(2x)$, ψ orlicz: convex, non-decreasing, $\psi(0) = 0$, $\psi(x) \to \infty$ as $x \to \infty$

Suppose $\mathbb{E}_p\left[\tilde{\psi}\left(\frac{|v^T x|}{\sigma}\right)\right] \le 1$ and $\mathbb{E}[(v^T X)^2] \le \sigma^2$

for all $\|v\|_2 = 1$

Then $\rho = \max\left(\boxed{\sigma \epsilon\, \psi^{-1}\left(\frac{2\sigma}{\epsilon}\right)}, \boxed{4\epsilon^2 + 2\epsilon\sigma}\right)$.

$\psi(x) = x^2$

$\tilde{\psi}(x) = x \cdot (2x)^2 = 4x^3$

$\Rightarrow$ bounded 3^{rd} moment

$$\max\left(\sigma\epsilon \cdot \left(\frac{2\sigma}{\epsilon}\right)^{1/2}, \; 4\epsilon^2 + 2\epsilon\sigma\right)$$

$$= O\left(\sigma^{3/2}\epsilon^{1/2} + 4\epsilon^2 + \cancel{2\epsilon\sigma}\right)$$

$$= O\left(\sigma^{3/2}\boxed{\epsilon^{1/2}} + \epsilon^2\right)$$

$\psi(x) = x^k$: bounded $(k+1)^{st}$ moments

$$O\left(\sigma^{1+\frac{1}{k}}\boxed{\epsilon^{1-\frac{1}{k}}} + \epsilon^2\right)$$

Proposition (non-Lipschitz KR duality)

ψ Orlicz function

g any function

c cost metric

measures how "non-Lipschitz" g is

$\psi(1)$ if Lipschitz

$$\underbrace{\left| \mathbb{E}_{x\sim p}[g(x)] - \mathbb{E}_{y\sim q}[g(y)] \right|}_{(1)}$$

$$\leq \sigma \cdot \mathbb{E}_{\pi_{p,q}}[c(x,y)] \cdot \psi^{-1}\left(\frac{\mathbb{E}_{\pi_{p,q}}\left[c(x,y)\, \psi\left(\frac{|g(x)-g(y)|}{\sigma \cdot c(x,y)} \right)\right]}{\mathbb{E}_{\pi_{p,q}}[c(x,y)]} \right)$$

for any coupling $\pi_{p,q}$.

$$(1) = \left| \mathbb{E}_{\pi_{p,q}}[g(x) - g(y)] \right|$$

Want.
$$\frac{\left| \mathbb{E}_{\pi_{p,q}}[g(x) - g(y)] \right|}{\sigma \cdot \mathbb{E}_{\pi_{p,q}}[c(x,y)]} \leq \psi^{-1}\left(\sim\sim \right)$$

ψ inside

Want.
$$\psi\left(\frac{\left| \mathbb{E}_{\pi_{p,q}}[g(x) - g(y)] \right|}{\sigma \cdot \mathbb{E}_{\pi_{p,q}}[c(x,y)]} \right) \leq \frac{\mathbb{E}_{\pi_{p,q}}\left[c(x,y)\, \psi\left(\frac{|g(x)-g(y)|}{\sigma \cdot c(x,y)} \right)\right]}{\mathbb{E}_{\pi_{p,q}}[c(x,y)]}.$$

ψ outside

$$\pi'(x,y) = \frac{c(x,y)\,\pi_{p,q}(x,y)}{\mathbb{E}_{\pi_{p,q}}[c(x,y)]}$$

$$\psi\left(\left|\mathbb{E}_{\pi'}\left[\frac{g(x)-g(y)}{\sigma\cdot c(x,y)}\right]\right|\right)$$

$$\leq \mathbb{E}_{\pi'}\left[\psi\left(\left|\frac{g(x)-g(y)}{\sigma\cdot c(x,y)}\right|\right)\right] \qquad \blacksquare$$

---

measures how "non-Lipschitz" g is

$$\left|\mathbb{E}_{x\sim p}[g(x)] - \mathbb{E}_{y\sim q}[g(y)]\right|$$

$$\leq \sigma\cdot\mathbb{E}_{\pi_{p,q}}[c(x,y)]\cdot\psi^{-1}\left(\frac{\mathbb{E}_{\pi_{p,q}}\left[c(x,y)\,\psi\left(\frac{|g(x)-g(y)|}{\sigma\cdot c(x,y)}\right)\right]}{\mathbb{E}_{\pi_{p,q}}[c(x,y)]}\right)$$

$c:\ \|x-y\|_2 \longrightarrow |\langle x-y, v\rangle| \quad (\leq \|x-y\|_2)$

$g:\ \langle x, v\rangle^2$

$$\left|\mathbb{E}_{x\sim p}[\langle x,v\rangle^2] - \mathbb{E}_{y\sim r}[\langle y,v\rangle^2]\right|$$

small if r ε-friendly perturb under $W1$

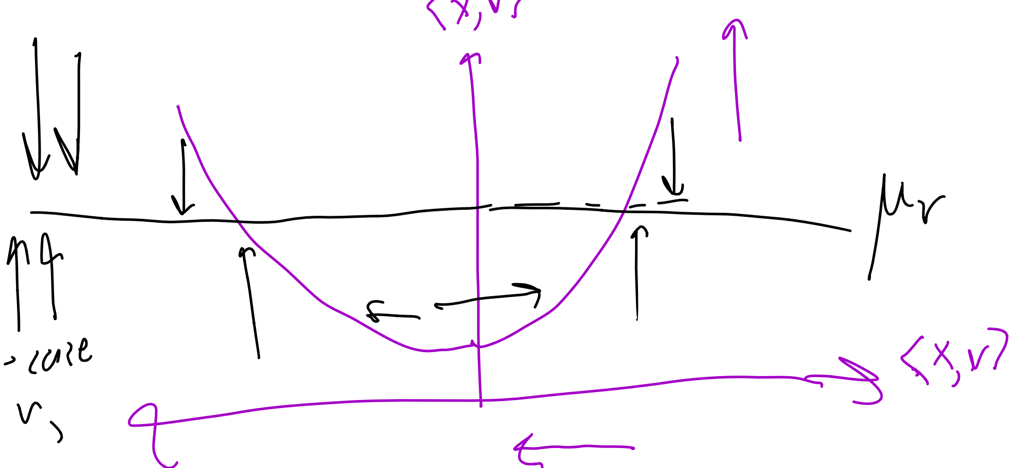
$$\left|\mathbb{E}_{x\sim p}[\langle x,v\rangle^2] - \mathbb{E}_{y\sim r}[\langle y,v\rangle^2]\right|$$

$$\lesssim \sigma \cdot \underbrace{\mathbb{E}_{\pi_{p,r}}[|\langle x-y,v\rangle|]}_{\le \varepsilon} \cdot \psi^{-1}\left(\frac{\mathbb{E}_{\pi_{p,r}}\left[|\langle x-y,v\rangle|\,\psi\left(\frac{|\langle x,v\rangle^2-\langle y,v\rangle^2|}{\sigma\cdot|\langle x-y,v\rangle|}\right)\right]}{\underbrace{\mathbb{E}_{\pi_{p,q}}[|\langle x-y,v\rangle|]}_{\le \varepsilon}}\right)$$

$$\lesssim \sigma\varepsilon\,\psi^{-1}\left(\frac{\mathbb{E}_{\pi_{p,r}}\left[|\langle x-y,v\rangle|\,\psi\left(\frac{|\langle x+y,v\rangle|}{\sigma}\right)\right]}{\varepsilon}\right)$$

$$\frac{a^2-b^2}{a-b}=a+b$$

(*)



observation:

Under worst-case perturbation v, all points move up or all move down.

Split into case:

(1) $\langle x,v\rangle^2 \ge \langle y,v\rangle^2 \quad \forall\ (x,y)\in \mathrm{supp}(\pi_{p,r})$

(2) $\langle x, v\rangle^2 \le \langle y, v\rangle^2 \quad \forall \quad (x,y) \in \mathrm{supp}(\pi_{p,r})$

$\le \mathbb{E}_r[\langle y, v\rangle^2]$ (by friendliness)

---

(1) $|\langle x+y, v\rangle| \le 2|\langle x, v\rangle|$

$|\langle x-y, v\rangle| \le 2|\langle x, v\rangle|$

$$(*) \le \sigma \varepsilon \psi^{-1}\left(\frac{\mathbb{E}_{\pi_{p,r}}\left[2|\langle x, v\rangle| \psi\left(\frac{2|\langle x, v\rangle|}{\sigma}\right)\right]}{\varepsilon}\right)$$

$$\le \sigma \varepsilon \psi^{-1}\left(\frac{\mathbb{E}_p\left[2|\langle x, v\rangle| \psi\left(\frac{2|\langle x, v\rangle|}{\sigma}\right)\right]}{\varepsilon}\right)$$

$$= \sigma \varepsilon \psi^{-1}\left(\frac{2\sigma}{\varepsilon} \cdot \underbrace{\mathbb{E}_p\left[\tilde{\psi}\left(\frac{|\langle x, v\rangle|}{\sigma}\right)\right]}_{\le 1}\right)$$

$$= \sigma \varepsilon \psi^{-1}\left(\frac{2\sigma}{\varepsilon}\right). \quad \blacksquare$$

(3) $|\langle x \pm y, v\rangle| \le 2|\langle y, v\rangle|$

$$\sigma \varepsilon \psi^{-1}\left(\frac{\mathbb{E}_{\pi_{p,r}}\left[|\langle x-y, v\rangle| \psi\left(\frac{|\langle x+y, v\rangle|}{\sigma}\right)\right]}{\varepsilon}\right)$$

(**)

$$\leq \sigma \varepsilon \psi^{-1}\left(\frac{\left|\mathbb{E}_{\pi_{p,r}}\left[\,|\underline{\langle x-y, v\rangle}|\cdot\psi\left(\frac{2|\langle y,v\rangle|}{\sigma}\right)\right]\right|}{\varepsilon}\right)$$

<u>Claim</u> Either $x=y$ or $\langle y,v\rangle^2 \leq \mathbb{E}_r[\langle y,v\rangle^2]$ by friendliness

$$\leq \sigma \varepsilon \cancel{\psi}^{-1}\left(\frac{\mathbb{E}_{\pi_{p,r}}\overset{p \leq \cancel{\varepsilon}}{[\langle x-y, v\rangle]}|\cdot\cancel{\psi}\left(\frac{2\,\mathbb{E}_r[\langle y', v\rangle^2]^{1/2}}{\sigma}\right)}{\cancel{\varepsilon}}\right)$$

constant

$$\leq 2\cancel{\sigma}\varepsilon \cdot \frac{\mathbb{E}_r[\langle y', v\rangle^2]^{1/2}}{\cancel{\sigma}}$$

$$\Big|\underbrace{\mathbb{E}_p[\langle x,v\rangle^2]}_{\sigma^2} - \underbrace{\mathbb{E}_r[\langle y,v\rangle^2]}_{a}\Big| \leq 2\varepsilon\cdot \mathbb{E}_r[\langle y,v\rangle^2]^{1/2}$$

$$|\sigma^2 - a| \leq 2\varepsilon\cdot\sqrt{a}$$

$$a \leq 4\varepsilon^2 + 2\varepsilon\sigma.$$