Lecture 6: Concentration via Optimal Transport

Tianpei Xie

Jan. 24th., 2023

Contents

1 Optimal Transport Basis

1.1 Optimal Transport Problem and its Dual Problem

1.2 Wasserstein Distance

1.3 Dual Formulation of Wasserstein Distance

2 The Transportation Method

2.1 Concentration via Transportation Cost Inequality

- **Remark** (***Equivalence of Transportation Cost Inequality and Sub-Gaussian***) [Boucheron et al., 2013]
  Let X be a real-valued integrable random variable. Let ϕ be a ***convex*** and ***continuously differentiable*** function on a (possibly unbounded) interval $[0, b)$ and assume that $\phi(0) = \phi'(0) = 0$. Define, for every $x \geq 0$, ***the Legendre transform*** $\phi^*(x) = \sup_{\lambda \in (0,b)}(\lambda x - \phi(\lambda))$, and let, for every $t \geq 0$, $\phi^{*-1}(t) = \inf\{x \geq 0 : \phi^*(x) > t\}$, i.e. the ***the generalized inverse*** of ϕ^*. Then the following two statements are equivalent:

  1. for every $\lambda \in (0, b)$,
  $$\psi_{X-\mathbb{E}[X]}(\lambda) \leq \phi(\lambda)$$
  where $\psi_X(\lambda) := \log \mathbb{E}_Q\left[e^{\lambda X}\right]$ is the logarithm of moment generating function;
  2. for any probability measure P absolutely continuous with respect to Q such that $\mathbb{KL}\left(P \,\|\, Q\right) < \infty$,
  $$\mathbb{E}_P\left[X\right] - \mathbb{E}_Q\left[X\right] \leq \phi^{*-1}\left(\mathbb{KL}\left(P \,\|\, Q\right)\right). \tag{1}$$

  In particular, given $\nu > 0$, X follows a ***sub-Gaussian distribution***, i.e.
  $$\psi_{X-\mathbb{E}[X]}(\lambda) \leq \frac{\nu\lambda^2}{2}$$
  for every $\lambda > 0$ ***if and only if*** for any probability measure P absolutely continuous with respect to Q and such that $\mathbb{KL}\left(P \,\|\, Q\right) < \infty$,
  $$\mathbb{E}_P\left[X\right] - \mathbb{E}_Q\left[X\right] \leq \sqrt{2\nu\mathbb{KL}\left(P \,\|\, Q\right)}. \tag{2}$$

- **Definition** (***d-Transportation Cost Inequality***) [Wainwright, 2019]
  Let $(\mathcal{X}, d)$ be a *metric space* with metric d, and $(\mathcal{X}, \mathscr{B})$ be a *measurable space*, where $\mathscr{B}$ is *the Borel σ-algebra* induced by metric d, ***the probability measure*** $\mathbb{P}$ is said to satisfy a ***d-transportation cost inequality*** with parameter $\nu > 0$ if
  $$\mathbb{E}_{\mathbb{Q}}\left[X\right] - \mathbb{E}_{\mathbb{P}}\left[X\right] \leq \sqrt{2\nu\mathbb{KL}\left(\mathbb{Q} \,\|\, \mathbb{P}\right)} \tag{3}$$
  for all probability measure $\mathbb{Q} \ll \mathbb{P}$ on $\mathscr{B}$.

- **Theorem 2.1** ***(Isoperimetric Inequality via Transportation Cost)****[Wainwright, 2019]* *Consider a metric measure space $(\mathcal{X}, \mathscr{B}, \mathbb{P})$ with metric d, and suppose that $\mathbb{P}$ satisfies the* ***d-transportation cost inequality***

$$\mathbb{E}_{\mathbb{Q}}\left[X\right] - \mathbb{E}_{\mathbb{P}}\left[X\right] \le \sqrt{2\nu \mathbb{KL}\left(\mathbb{Q} \,\|\, \mathbb{P}\right)}$$

for all probability measure $\mathbb{Q} \ll \mathbb{P}$ on $\mathscr{B}$. Then its ***concentration function*** *satisfies the bound*

$$\alpha_{\mathbb{P},(\mathcal{X},d)}(t) \le 2\exp\left(-\frac{t^2}{2\nu}\right) \tag{4}$$

Moreover, for any $Z \sim \mathbb{P}$ and any L-Lipschitz function $f : \mathcal{X} \to \mathbb{R}$, we have the ***concentration inequality***

$$\mathbb{P}\left\{\left|f(Z) - \mathbb{E}\left[f(Z)\right]\right| \ge t\right\} \le 2\exp\left(-\frac{t^2}{2\nu L^2}\right). \tag{5}$$

2.2 Tensorization for Transportation Cost

2.3 Bounded Difference Inequality via Transportation Methods

2.4 Conditional Transportation Inequality

2.5 Convex Distance Inequality via Conditional Transportation Cost

2.6 Talagrand's Gaussian Transportation Inequality

2.7 Transportation Cost Inequalities for Markov Chains

References

Stéphane Boucheron, Gábor Lugosi, and Pascal Massart. *Concentration inequalities: A nonasymptotic theory of independence*. Oxford university press, 2013.

Martin J Wainwright. *High-dimensional statistics: A non-asymptotic viewpoint*, volume 48. Cambridge University Press, 2019.