Lecture 24

Review: * Transportation Lemma

$$\psi_{Z-\mathbb{E}Z}(\lambda) \leq \frac{\lambda^2 v}{2} \quad \forall \lambda > 0 \text{ iff}$$

$$\mathbb{E}_Q[Z] - \mathbb{E}_P[Z] \leq \sqrt{2 v D(Q\|P)} \quad \forall\, Q \ll P.$$

* McDiarmid Revisited

Suffices to show: $\exists$ a coupling (X, Y) of P^n and Q s.t.

$$\underbrace{\sum_{i=1}^{n} \mathbb{P}(X_i \neq Y_i)^2 \leq \frac{1}{2} D(Q\|P^n).}_{\text{Marton's Transportation Cost Lemma shows}}$$

exactly this using:
- Pinsker's inequality
- Maximal coupling
- Chain rule for divergence:

$$D(Q_{Y^n}\|P_{X^n}) = \sum_{i=1}^{n} D\left(Q_{Y_i|Y^{i-1}} \,\middle\|\, P_{X_i|X^{i-1}} \,\middle|\, Q_{Y^{i-1}}\right)$$

$$\mathbb{E}_{y^{i-1}\sim Q_{Y^{i-1}}}\left[D\left(Q_{Y_i|Y^{i-1}=y^{i-1}} \,\middle\|\, P_{Y_i|Y^{i-1}=y^{i-1}}\right)\right]$$

$\boxed{F}$ Functions "Lipchitz" in d_T and beyond that

Talagrand's convex distance inequality told us the following:

Consider $f: \mathcal{X}^n \to \mathbb{R}$ s.t.

$$f(x) - f(y) \leq \sum_{i=1}^{n} c_i(x)\, \mathbb{1}_{\{x_i \neq y_i\}} \quad \forall x, y.$$

Then, $f / \left\| \sqrt{\sum_{i=1}^{n} c_i^2} \right\|_\infty$ is Lipchitz in d_T.

②

Therefore, $\boxed{\begin{aligned}&\mathbb{P}\Big(f(X) > \mathbb{M}f(X) + t\Big) \\ &\le \exp\Big(-t^2/4\,\|\sum_{i=1}^n c_i^2\|_\infty\Big)\end{aligned}}$.

Talagrand also showed the following conc. inequality:

$$\boxed{\mathbb{P}\big(f(X) > \mathbb{E}f(X) + t\big) \le \exp\left(-\frac{t^2}{2\|\sum_{i=1}^n c_i^2\|_\infty}\right)}$$

→ This can be derived using the entropy method:

Recall that using the modified log-Sobolev inequality we showed that for $Z = f(X)$, $Z_i = \inf_{x_i'} f(X^{i-1}, x_i', X_{i+1}^n)$

if $\sum_{i=1}^n (Z - Z_i)^2 \le v$,

then $\mathbb{P}(Z > \mathbb{E}Z + t) \le e^{-t^2/2v}$. (see lecture 14)

The condition we have yields $v = \sup_x \sum_{i=1}^n c_i^2(x)$.

→ Also, recall that the entropy method (see lec. 14) didn't allow us to get a lower tail bound. We had to additionally assume $Z - Z_i \le 1$, $1 \le i \le n$, which essentially is tantamount to getting a subGaussian bound with variance parameter $\sum_{i=1}^n \sup_x c_i(x)^2$.

* The transportation method will allow us to strengthen these bound in various ways:

(1) We will recover the upper tail bound above.

(2) We can also recover Talagrand's convex distance inequality. (HW: Get upper tail bound for $d_T(X, A)$).

(3) We will get a lower tail bound with even better variance parameter $\mathbb{E}\left[\sum_{i=1}^{n} c_i(X)^2\right]$.

Theorem $X = (X_1, \dots, X_n)$ indep

$$f(x) - f(y) \le \sum_{i=1}^{n} c_i(x) \mathbb{1}_{\{x_i \ne y_i\}}.$$

$$v = \mathbb{E}\left[\sum_{i=1}^{n} c_i(X)^2\right]$$

$$v_\infty = \sup_x \sum_{i=1}^{n} c_i(x)^2.$$

Then,

$$\mathbb{P}(Z > \mathbb{E}Z + t) \le e^{-t^2/2v_\infty}$$

$$\mathbb{P}(Z < \mathbb{E}Z - t) \le e^{-t^2/2v}.$$

Proof. Consider a $Q \ll P^n$ and let $(X, Y) \in \mathcal{P}(P, Q)$.

$$\begin{aligned}
\mathbb{E}_Q[Z] - \mathbb{E}_P[Z] &= \mathbb{E}\left[f(Y) - f(X)\right] \\
&\le \mathbb{E}\left[\sum_{i=1}^{n} c_i(Y) \mathbb{1}_{\{Y_i \ne X_i\}}\right] \\
&= \mathbb{E}\left[\sum_{i=1}^{n} c_i(Y)\, \mathbb{P}(Y_i \ne X_i \mid Y)\right] \\
&\le \mathbb{E}\left[\sqrt{\sum_{i=1}^{n} c_i(Y)^2} \cdot \sqrt{\sum_{i=1}^{n} \mathbb{P}(Y_i \ne X_i \mid Y)^2}\right] \\
&\le \sqrt{\mathbb{E}\left[\sum_{i=1}^{n} c_i(Y)^2\right]} \cdot \sqrt{\mathbb{E}\left[\sum_{i=1}^{n} \mathbb{P}(Y_i \ne X_i \mid Y)^2\right]}
\end{aligned}$$

Thus, $$\mathbb{E}_Q[Z]-\mathbb{E}_P[Z] \leq \sqrt{V_\infty}\cdot\sqrt{\mathbb{E}\left[\sum_{i=1}^n \mathbb{P}(Y_i \neq X_i|Y)^2\right]}$$

Similarly,
$$\mathbb{E}_Q[-Z]-\mathbb{E}_P[-Z] \leq \sqrt{\mathbb{E}\left[\sum_{i=1}^n c_i^2(X)\right]}\cdot$$
$$\sqrt{\mathbb{E}\left[\sum_{i=1}^n P(Y_i \neq X_i|X)^2\right]}.$$

Therefore, using the Transportation Lemma, the result will follow upon showing the following transportation cost inequality:

Marton's Conditional Transportation Cost Inequality

$$\min_{(X,Y)\in P(P^n,Q)} \mathbb{E}\left[\sum_{i=1}^n \mathbb{P}(X_i \neq Y_i|X)^2 + \mathbb{P}(X_i \neq Y_i|Y)^2\right]$$
$$\leq \quad 2D(Q\|P^n).$$

for all $Q \ll P^n = P_1 \times \ldots \times P_n$.