Statistical bounds for EOT

* $P,\ P_n = \frac{1}{n}\sum_{i=1}^{n} \delta_{x_i}$ (same for Q) ($P, Q \in \mathcal{P}(\mathbb{R}^d)$)

(A) $W_2^2(P_n, Q_n) \to W_2^2(P,Q)$ @ $O(n^{-1/\boxed{d}})$

$(\varepsilon=1)$

(P) * $S_{\cancel{\varepsilon}}(P,Q) = \inf_{\gamma \in \Pi(P,Q)} \int \frac{1}{2}\|x-y\|_2^2\, d\gamma(x,y) + \cancel{\varepsilon} H(\gamma \mid P \otimes Q)$

[Genevay, Cuturi...]

(☆) $\sup_{P,Q} \mathbb{E}_{P,Q} |S(P,Q) - S(P_n, Q_n)| \le K_{D,d} \cdot \frac{1}{\sqrt{n}} \cdot e^{D^2}$

$(\varepsilon \neq 1)$ $\le K_{D,d}\left(1 + \frac{1}{\varepsilon^{\lfloor d/2 \rfloor}}\right) e^{D^2/\varepsilon} \cdot \frac{1}{\sqrt{n}}$

$D :=$ diameter of the domain of P, Q

P, Q are σ^2-subg. ? $\boxed{(☆) \le K_{d,\sigma^2} \cdot \frac{1}{\sqrt{n}}}$ (Theorem 2)

(in $\mathbb{R}^d$) • P is σ^2-subg. if $\mathbb{E}_P\, e^{\|X\|^2/2\sigma^2 d} \le 2$ ①

• If P is σ^2-subg, $\mathbb{E}_P \|X\|^{2k} \le (2d\sigma^2)^k k!$ ②

$$\mathbb{E}_P[e^{v \cdot X}] \le \mathbb{E}_P[e^{\|v\| \cdot \|X\|}]$$
$$\le 2e^{\frac{d\sigma^2}{2}\|v\|^2}$$ ③

- P is σ^2-subg, $\exists\ \sigma_u < \infty$ (random) s.t. ④

$\{P_n\}_{n}, P$ are uniformly σ_u^2-sub. (P a.s)

(D) $$S(P,Q) = \sup_{\substack{f \in L_1(P) \\ g \in L_1(Q)}} \left\{ \int f\,dP + \int g\,dQ + 1 - \int e^{f(x)+g(y) - \frac{1}{2}\|x-y\|^2}\,dP\,dQ \right\}$$

opt: $$\int e^{f \oplus g - \frac{1}{2}\|x-y\|^2}\,dQ = 1 \quad (P.\ a.s)$$ ⑤

$$\int e^{f \oplus g - \frac{1}{2}\|x-y\|^2}\,dP = 1 \quad (Q.\ a.s)$$

step 1: Want (*) as an emp. process in terms of potential fun.

Step 1.5: but the regularity of the functions comes from P, Q being sub.

Step 2: do some bounding.

[Prop 6, Appendix A]. P, Q are σ^2-sub. $\exists\ (f, g)$ smooth opt. potentials s.t.

$$-d\sigma^2\left(1 + \tfrac{1}{2}(\|x\| + \sqrt{2d}\,\sigma)^2\right) - 1 \le f(x) \le \tfrac{1}{2}(\|x\| + \sqrt{2d}\,\sigma)^2$$

$$\text{———} \quad \text{''} \quad \text{———} \le g(y) \le \text{—} \text{''} \text{—}$$

Proof | • $\overset{\nearrow\,\text{opt.}}{(f_0,g_0)} \mapsto \overset{\longrightarrow\,\text{opt.}}{(f_0+k,\, g_0-k)}$

↳ $\mathbb{E}_P[f_0(X)] = \mathbb{E}_Q[g_0(Y)] = \frac{1}{2} S(P,Q) \geq 0$

• $f(x) = -\log\left(\int e^{g_0(y) - \frac{1}{2}\|x-y\|^2} \, dQ(y)\right)$

• $g(y) = -\log\left(\int e^{f(x) - \frac{1}{2}\|x-y\|^2} \, dP(x)\right)$

=

$$g_0(y) = -\log\left(\int e^{f_0(x) - \frac{1}{2}\|x-y\|^2} \, dP(x)\right)$$

$$\leq -\mathbb{E}_P[\cancel{f_0(X)}] + \frac{1}{2}\mathbb{E}_P[\|X-y\|^2]$$

$$\to e^{g_0(y) - \frac{1}{2}\|x-y\|^2} \leq e^{\frac{1}{2}\mathbb{E}_P[\|X-y\|^2] - \frac{1}{2}\|x-y\|^2}$$

By ② ...... $\leq e^{d\sigma^2 + (\|x\| + \sqrt{2}\, d\sigma)\|y\|}$ .... ✓

=

We know $\int e^{f(x)+g(y) - \frac{1}{2}\|x-y\|^2} \, dP(x) = 1$ $\forall y$

↳ $\int e^{f \oplus g - \frac{1}{2}\|\cdot - \cdot\|^2} \, dP\, dQ = \cdots = \int e^{f_0 \oplus g_0 - \frac{1}{2}\|\cdot - \cdot\|^2} \, dP\, dQ$

$$\int (f - f_0)\, dP + \int (g - g_0)\, dQ \geq -\log\left(\int e^{f_0 - f} dP\right) - \log\left(\int e^{g_0 - g} dQ\right)$$

$$= -\log\left(\int e^{f_0 \oplus g_0 - \frac{1}{2}\|\cdot\cdot\|^2} dP\, dQ\right) = 0$$

$$-\log\left(\int e^{f + g_0 - \frac{1}{2}\|\cdot\cdot\|^2} dP\, dQ\right)$$

[Prop 1] P, Q .. σ^2-sub. $\exists (f, g)$ opt. s.t. for any multi-index α, $|\alpha| = k$

$(\|x\| \le \sqrt{d}\sigma)$ $|D^\alpha (f - \frac{1}{2}\|\cdot\|^2)(x)| \le C_{k,d} \begin{cases} 1 + \sigma^4 & k = 0 \\ \sigma^k (\sigma + \sigma^2)^k & \text{o.w.} \end{cases}$

$(\|x\| > \sqrt{d}\sigma)$

$|D^\alpha (f - \frac{1}{2}\|\cdot\|^2)(x)| \le C_{k,d} \begin{cases} 1 + (1+\sigma^2)\|x\|^2 & k = 0 \\ \cdots & \text{o.w.} \end{cases}$

$\mathcal{F}_\sigma \rightsquigarrow \mathcal{F}^s$: $|f(x)| \le C_{s,d}(1 + \|x\|^2)$

$|D^\alpha f(x)| \le C_{s,d}(1 + \|x\|^s)$ $\forall \alpha : |\alpha| \le s$

$\hookrightarrow$ If $C_{s,d}$ is large enough: $\boxed{\frac{1}{1+\sigma^{3s}} f \in \mathcal{F}^s \quad (f \in \mathcal{F}_\sigma)}$

[Prop 2] P, Q, P_n $\tilde{\sigma}^2$-sub.

$$|S(P_n, Q) - S(P, Q)| \le 2 \cdot \sup_{u \in \mathcal{F}_{\tilde{\sigma}}} |\mathbb{E}_P[u] - \mathbb{E}_{P_n}[u]|$$

α-Holder w/ M_j

2.7.4 Corollary. *Let $\mathbb{R}^d = \cup_{j=1}^{\infty} I_j$ be a partition of $\mathbb{R}^d$ into bounded, convex sets with nonempty interior, and let $\mathcal{F}$ be a class of functions $f: \mathbb{R}^d \mapsto \mathbb{R}$ such that the restrictions $\mathcal{F}_{|I_j}$ belong to $C^{\alpha}_{M_j}(I_j)$ for every j. Then there exists a constant K depending only on α, V, r, and d such that*

$$\log N_{[]}(\varepsilon, \mathcal{F}, L_r(Q)) \le K \left(\frac{1}{\varepsilon}\right)^V \left(\sum_{j=1}^{\infty} \lambda(I_j^1)^{\frac{r}{V+r}} M_j^{\frac{Vr}{V+r}} Q(I_j)^{\frac{V}{V+r}}\right)^{\frac{V+r}{r}},$$

for every $\varepsilon > 0$, $V \ge d/\alpha$, and probability measure Q.

Leb. measure

$$\|f\|_{L_2(P_n)} = \left(\frac{1}{n}\sum_{i=1}^{n} f(x_i)^2\right)^{1/2}$$

$\{x \mid d(x, I_j) \le 1\}$

- $r = 2$; $\mathcal{F} := \mathcal{F}^s$

Prop 3: • $\log N(\tau, \mathcal{F}^s, L_2(P_n)) \le C_d \cdot L^{d/2} \tau^{-d/s}(1+\sigma^{2d})$

$\hookrightarrow L = L(X_1, \dots, X_n)$

- $\max_{f \in \mathcal{F}^s} \|f\|^2_{L_2(P_n)} \le C_d(1 + L\sigma^4)$ $\rightsquigarrow$ subg ①

Theorem 2 (Proof): ④ $\to \tilde{\sigma}$

From Prop 2:

$$\mathbb{E}_{P,Q}|S(P,Q) - S(P_n, Q_n)| \le \mathbb{E} \sup_{u \in \mathcal{F}_{\tilde{\sigma}}} \left|\int u\,(dP - dP_n)\right| + \mathbb{E}_{v \in \mathcal{F}_{\tilde{\sigma}}} \left|\int v\,(dQ - dQ_n)\right|$$

$$\|P-P_n\|_{\mathcal{F}} := \sup_{u \in \mathcal{F}_{\tilde{\sigma}}} \left| \int u \, (dP - dP_n) \right|$$

$$\frac{1}{1+\tilde{\sigma}^{3s}} u \in \mathcal{F}^s$$

$$\mathbb{E}\|P-P_n\|_{\mathcal{F}} \le \mathbb{E}\left(1+\tilde{\sigma}^{3s}\right)\|P-P_n\|_{\mathcal{F}^s}$$

$$\le \left(\mathbb{E}(1+\tilde{\sigma}^{3s})^2\right)^{1/2} \underbrace{\left(\mathbb{E}\|P-P_n\|_{\mathcal{F}^s}^2\right)^{1/2}}$$

(Giné, Nickl ; Thm 3.5.1)

$$\mathbb{E}\|P-P_n\|_{\mathcal{F}^s}^2 \le \frac{1}{n}\mathbb{E}\left(\int_0^{\sqrt{\max_{f\in\mathcal{F}^s}\|f\|^2_{L^2(P_n)}}} \sqrt{2\log N(\tau, \mathcal{F}^s, L_2(P_n))}\, d\tau\right)^2$$

$$\le C_{d,\tilde{\sigma}} \cdot \frac{1}{n}$$

2.7.4 Corollary. *Let $\mathbb{R}^d = \cup_{j=1}^{\infty} I_j$ be a partition of $\mathbb{R}^d$ into bounded, convex sets with nonempty interior, and let $\mathcal{F}$ be a class of functions $f: \mathbb{R}^d \mapsto \mathbb{R}$ such that the restrictions $\mathcal{F}_{|I_j}$ belong to $C^{\alpha}_{M_j}(I_j)$ for every j. Then there exists a constant K depending only on α, V, r, and d such that*

$$\log N_{[\,]}\big(\varepsilon, \mathcal{F}, L_r(Q)\big) \le K \left(\frac{1}{\varepsilon}\right)^V \left(\sum_{j=1}^{\infty} \lambda(I_j^1)^{\frac{r}{V+r}} M_j^{\frac{Vr}{V+r}} Q(I_j)^{\frac{V}{V+r}}\right)^{\frac{V+r}{r}},$$

for every $\varepsilon > 0$, $V \ge d/\alpha$, and probability measure Q.