Bregman divergence regularization of optimal transport problem on a finite set

D_U (U : convex fct) $\varepsilon > 0$

Asuka TAKATSU
j/ Koya Sakakibara
& Keiichi Morikuni

$$\inf_{\pi \in \Pi(x,y)} \left(\langle C, \pi \rangle + \varepsilon D_U(\pi, x \otimes y) \right)$$

$$\mathcal{C}_U^{\varepsilon}(x,y)$$

$U(r) = r \log r \Rightarrow D_U = D_{KL}$

fix $N \in \mathbb{N}$ & $C \in M_N(\mathbb{R})$

$\mathcal{P}_N := \{ x \in \mathbb{R}^N \mid x_n \geq 0 \ \& \ \sum_{n=1}^{N} x_n = 1 \}$

take $x, y \in \mathcal{P}_N$

$\Pi(x,y) := \{ \pi \in \mathcal{P}_{N \times N} \mid \sum_{n=1}^{N} \pi_{in} = x_i, \ \sum_{n=1}^{N} \pi_{nj} = y_j \}$

Thm.

Under $^{\exists}$ assumption

$$\mathcal{C}_U^{\varepsilon}(x,y) - \mathcal{C}(x,y) \leq \lambda_1 \cdot (U')^{-1}\left(-\frac{\lambda_1}{\varepsilon} + \lambda_2\right)$$

$\lambda_1 = \lambda_1(C, x, y), \ \lambda_2 = \lambda_2(U, x, y) > 0$

Ex. $(x \otimes y)_{ij} := x_i y_j$

$$\langle C, \pi \rangle := \sum_{i,j=1}^{N} C^{ij} \pi_{ij}$$

Assumption (for x & y)

① $x_i, y_j > 0$ — We can consider $x \in \mathcal{P}_I$, $y \in \mathcal{P}$

② $x \otimes y$ is NOT optimal — $x \otimes y$: optimal $\Rightarrow {}^\forall \pi \in \Pi(x, y)$: optimal

$\Rightarrow x_i, y_j < 1$ & $\pi^{opt} \notin \mathrm{int}\, \Pi(x, y)$

Rem. $0 \leq \pi_{ij} \leq x_i y_j < 1.$

Assumption (for $U \in C([0,1]) \cap C^1((0,1])$: strictly convex)

Def. (Bregman divergence)

$U(r) = r \log r$

$D_U(z, w) = \sum_{n=1}^{N} z_n \log \frac{z_n}{w_n} = D_{KL}(z, w)$

$$D_U : \mathcal{P}_N \times \mathcal{P}_N \longrightarrow [0, +\infty]$$

$$D_U(z, w) := \sum_{n=1}^{N} \left(U(z_n) - U(w_n) - (z_n - w_n) U'(w_n) \right) \in [0, \infty], \quad z, w \in [0, 1]$$

Assumption 2. ⓪ $U \in C([0,1]) \cap C^1((0,1]) \cap C^2((0,1))$ & $U'' > 0$ on $(0,1)$

① $U(0) = U(1) = 0$ $\quad \tilde{U}(r) = U(r) + r(-U(1) + U(0)) - U(0) \Rightarrow D_U = D_{\tilde{U}}$.

② $\lim_{\varepsilon \downarrow 0} U'(\varepsilon) = -\infty$ $\quad \Rightarrow \exists ! \, \pi^{\varepsilon} \in \operatorname*{argmin}_{\pi \in \Pi(x,y)} (\langle c, \pi \rangle + \varepsilon \underbrace{D_U(\pi, x \otimes y)}_{\text{str. cvx. w.r.t } \pi})$ c in $\Pi(x,y)$

$\downarrow \varepsilon \downarrow 0$

$\pi^0 \in \operatorname*{argmin}_{\pi \in \Pi(x,y)} \langle c, \pi \rangle$ w/ $D_U(\pi^0, x \otimes y) = \min_{\pi : \text{opt}} D_U(\pi, x \otimes y)$

③ $\mathfrak{g}(U) :\overset{\text{“ ”}}{=} -\inf_{r \in (0,1)} \frac{r U'''(r)}{U''(r)}$

$= 1$

② $\Rightarrow \mathfrak{g}(U) \geq 1$

$\mathfrak{g}(U) < \frac{3}{2} \Rightarrow \mathcal{D}(x,y)$: conti. on $\mathcal{P}_N \times \mathcal{P}_N$

Rem. · $\mathfrak{g}(U = r \log r) = 1$

· $U \in C^2((0,\infty))$ & $\mathfrak{g}(U) < 2 \Rightarrow U \in DC_N$ w/ $N = \frac{1}{\mathfrak{g}(U) - 1}$

Thm. Under Assumptions 1 & 2

$$\langle C,\pi^{\varepsilon}\rangle - \langle C,\pi^{0}\rangle \le \Delta_C(x,y)\, e_U\left(-\frac{\Delta_C(x,y)}{\varepsilon} + \mathcal{D}(x,y) + \delta(x,y)\right)$$

$$\Delta_C(x,y) := \min_{\pi\in\Gamma'}\langle C,\pi\rangle - \underbrace{\min_{\pi\in\Gamma^{opt}}\langle C,\pi\rangle}_{=\langle C,\pi^0\rangle}$$

$\Gamma' \cup \Gamma^{opt}$: vertex set of $\Pi(x,y)$

e_U : inverse fct of $U' : (0,1] \to (-\infty, U'(1)]$

$$\mathcal{D}(x,y) = D_U(\pi^0, x\otimes y) = \min_{\pi:\mathrm{opt}} D_U(\pi, x\otimes y)$$

$\exists!\, R \in (\tfrac{1}{2}, 1)$ s.t. $U'(R) - U'(1-R) = \mathcal{D}(x,y)$

$$\delta(x,y) := \sup_{r\in(0,R)} \Big(\underbrace{U'(1-r)}_{\le U'(1)} + \underbrace{r\,U''(r)}_{\le RU''(R)}\Big) < +\infty$$

$U'' > 0$

non decreasing $\Leftrightarrow$ $\xi(U) \le 1$

Rem.

idea from J. Weed (2018)

for $U(r) = r\log r$

$e_U(\tau) = e^{\tau-1}$

$\delta(x,y) = 2.$

Ex. ($g(U)=1$) ∵ <u>deformed log. fct</u> $\quad a \in (0,\infty]$

$$\varphi : (0,a) \longrightarrow (0,\infty) \qquad \left(\varphi = \frac{1}{U''}, \quad g(U) = \sup_{s\in(0,a)} \frac{s\varphi'(s)}{\varphi(s)} \right)$$

$g(\varphi)$

$$\ln_\varphi(t) = \int_1^t \frac{1}{\varphi(s)}\, ds$$

$U_\varphi(r) = \int_0^r \ln_\varphi(t)\, dt$ … well-defined if $g(\varphi) < 2$.

① $\varphi_g(s) = s^g \Rightarrow \ln_g(t) = \dfrac{t^{1-g}-1}{1-g}$ $\quad (g \neq 1)$ $\quad g(\varphi) = g$

② $\varphi_\delta(s) = s(-\log s)^{1-\delta}$, $\quad g(\varphi_\delta) = \begin{cases} 1 & \delta \le 1 \\ \infty & \delta > 1 \end{cases}$ faster than exp

$$\ln_{\varphi_\delta}(t) = -\frac{1}{\delta}\left((-\log t)^\delta - 1 \right) \rightsquigarrow e_\delta(\tau) = \exp\left(-(-\delta\tau+1)^{\frac{1}{\delta}} \right)$$

$\ln_{\varphi_\delta}(0) = -\infty \iff \delta \ge 0$, $\quad \ln_{\varphi_\delta}(1) < \infty \iff \delta > 0$

$\delta = 1$: KL

$0 < \delta \le 1$