$$W_2^2(P,Q) = \sup_{\phi \in \Psi_c(X)} \int_X \phi \, dP + \int_Y \phi^c \, dQ. = \int_X \varphi^* dP + \int_Y \varphi^{*c} dQ.$$

$$\frac{\delta W_2^2(P,Q)}{\delta P} = \frac{\delta}{\delta P}\left\{ \int_X \varphi \, dP + \int_Y \varphi^c dQ \right\}.$$

$$= \int_X \frac{\delta}{\delta P} \varphi \, dP$$

$$= \varphi^*.$$

$P = F_\theta$. $\quad \frac{\partial F_\theta}{\partial \theta}$

$$\frac{\partial W_2^2(F_\theta, Q)}{\partial \theta} = \int \varphi^*_{(x)} \cdot \nabla_\theta F_\theta(x) \, dx$$

$T_\theta(x) = \nabla_x \varphi_\theta(x) \quad x \sim \mu. \quad P =$

φ^* optimal potential.

$$\nabla_\theta W_2^2((T_\theta)_\# \mu, Q) =$$

$$\frac{\delta}{\delta \rho} W_2^2(\rho, Q) = \varphi_\rho$$

First variation 定义: $\quad F(\rho) - F(\rho_0) \approx \int \frac{\delta}{\delta \rho} F \cdot (\rho - \rho_0) + \underbrace{O(W_2^2(\rho - \rho_0))}_{\text{higher order}}.$

$\theta' \approx \theta,$

$$W_2^2((T_{\theta'})_\# \mu, Q) - W_2^2((T_\theta)_\# \mu, Q).$$

$$\approx \int \varphi_\rho \cdot (d(T_{\theta'})_\# \mu - d(T_\theta)_\# \mu).$$

$$= \int \varphi_\rho \circ T_{\theta'} \, d\mu - \int \varphi_\rho \circ T_\theta \, d\mu. \quad \text{(sample avg).}$$

$u_1, \dots, u_m \sim \mu$ (source).

$$= \frac{1}{M} \sum_{j=1}^{M} \left[\phi^*(T_{\theta'}(u_j)) - \phi^*(T_\theta(u_j)) \right]$$

$$= \frac{1}{M} \sum_{j=1}^{M} \nabla_\theta (\underbrace{\phi^*}_{\text{Sinkhorn.}} \circ T_\theta)(u_j)$$

$W_2^2(P, Q)$ $\quad \phi^*$, $\nabla \phi^* = T_P^Q$

$$\nabla_\theta (\phi^* \circ T_\theta) = T_P^Q \circ \nabla T_\theta.$$

- First variation.

F, gradient, $F(\theta) - F(\theta^*)$

$$= \langle \underset{\substack{\uparrow \\ \text{gradient.}}}{f}, (\theta - \theta^*) \rangle + o(\|\theta - \theta^*\|^2).$$

CS 540 Deep Learing Theorom.