f-div : $D_f(p\|q) = \int f(p/q)\, q \, dx$

$f: \mathbb{R}_+ \to \mathbb{R}$ convex $\downarrow$

$f(x) = x\log x \qquad \int \frac{p}{q}\log\frac{p}{q}\cdot q = \int p\log\frac{p}{q}$

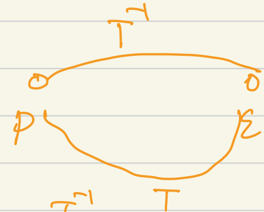
$f(x) = -\log x \quad \Rightarrow \quad \int q\log q/p$

$p_t = p * N(0,2t), \quad q_t = q * N(0,2t)$

$$\frac{d\, D_f(p_t\|q_t)}{dt} = -\boxed{\quad}$$

$p \sim N(0,1) \quad , \quad q \sim X_i$

$q = T^{-1}_{\#}p \Leftrightarrow p = T_{\#}q$



$p_t = (T_t^{-1})_{\#}\, p$

$q_t = (T_t)_{\#}\, q$

$$\min_{T_t} KL(q\,\|\,\underbrace{p_t}_{(T_t^{-1})_{\#}p}) = KL(\underbrace{q_t}_{(T_t)_{\#}q}\,\|\,p)$$

apply T_t $\Longrightarrow$ $\geq$

$\leq$ apply T_t^{-1}

$$KL(q \,\|\, p_t) = \int q(x) \log \frac{q(x)}{p_t(x)} \, dx = KL(q_t \,\|\, p)$$

$$= \int q_t(x) \log \frac{q_t(x)}{p(x)} \, dx$$

$$= \int \underline{q_t(T_t y)} \log \frac{q_t(T_t y)}{p(T_t y)} \, \underline{d T_t y}$$

$$\boxed{q_t(T_t y) \cdot |\det T_t' y| = q(y)} \Rightarrow$$

$$= \int q(y) \log \frac{q_t(T_t y)}{p(T_t y)} \, dy$$

$$\frac{\delta KL}{\delta T} = \left. \frac{d \int q(y) \log \frac{q_t((T+\varepsilon S)y)}{p((T+\varepsilon S)y)} \, dy}{d\varepsilon} \right|_{\varepsilon=0}$$

$$= \int q(y) \, \nabla \log q_t(Ty) \cdot \left. \frac{d(T+\varepsilon S)}{d\varepsilon} \right|_{\varepsilon=0}$$

$$\dots$$

$$= \int \underbrace{q(y) \left[S_{q_t}(Ty) - S_p(Ty) \right]}_{= \frac{\delta KL}{\delta T}} \cdot \underbrace{S(y)} \, dy$$

$$\langle f, g \rangle \triangleq \int f(y) g(y) \, dy$$

$\nabla f(x) =$

$= \nabla_i f(x)$

$x \in \mathbb{R}^d$

$x : \{1, 2, \cdots, d\} \to \mathbb{R}$

$x(1) = x_1$

$x(2) = x_2$

$\vdots$

$x(d) = x_d$

$$\left. \frac{d f(x + \varepsilon \cdot e_i)}{d\varepsilon} \right|_{\varepsilon=0} = \partial_i f(x)$$

$$= \langle \nabla f(x), e_i \rangle$$

$$\left. \frac{d f(x + \varepsilon \cdot d)}{d\varepsilon} \right|_{\varepsilon=0} = \langle \nabla f(x), d \rangle$$

$$\| $$

$$\nabla f(x)^T d$$

$\langle x, y \rangle_{(n)} = 2 x^T y$

under $\langle x, y \rangle_{(n)}$ we have $\nabla_{(n)} f = \frac{1}{2} \nabla f(x)$

$$\nabla_{L_2} KL = \left[S_{q_t}(T_t x) - S_p(T_t x)\right] q(x)$$

$$= \left[S_{q_t}(T_t x) - S_p(T_t x)\right] \cdot q_t(T_t x) \cdot \left[\cancel{\det T_t'(x)}\right]$$

$$\dot T_t = -b(T_t), \quad \text{where } b(z) = \left[S_{q_t}(z) - S_p(z)\right] q_t(z)$$

$$T_t(x; u)$$

Neural ODE (chen et al '2018)

$t \in [0,1]$ $\quad \boxed{\dot X_t = \vec f(t, X_t, u)}$ $\quad X_0 \sim P_{data}$

$$\partial_t P_t = -\nabla \cdot (P_t \vec f)$$

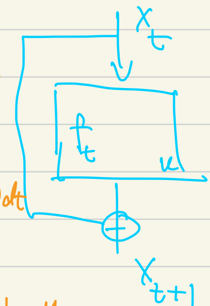
discretize

$$\min_u \; g(X_1)$$

$$X_{t+1} = X_t + \vec f_t(X_t, u)$$

$$L(t, x, u, \rho) = g(X_1) + \int_0^1 (\dot X_t - \vec f) \cdot \vec\rho(t)\, dt$$

$$= \int_0^1 \langle \dot X_t, \rho_t \rangle dt - \int_0^1 H(t, x, u, \rho)\, dt$$



$$\tilde u = u + \varepsilon \cdot v$$

$$\tilde X = X + \varepsilon \cdot y + o(\varepsilon)$$

$$\frac{dL(t, \tilde X, \tilde u, \vec\rho)}{d\varepsilon}\Big|_{\varepsilon=0} = \langle \nabla g(X_1), y_1 \rangle + \langle \rho_1, y_1 \rangle - \int_0^1 \langle \cancelto{y}{X_t}, \dot\rho_t \rangle dt$$

$$\frac{dL}{d\varepsilon}\Big|_{\varepsilon=0} = \underbrace{\langle \nabla g(x_1) + \rho_1,\ y_1 \rangle}_{=0} - \int_0^1 \langle \overbrace{\dot{\rho} + H_x}^{0}, y \rangle + \langle H_u, v \rangle \, dt$$

$$\begin{cases} ② \ \dot{\rho} = -H_x, \quad \rho_1 = -\nabla g(x_1) \\ ① \ \dot{x} = +H_p \\ \nabla_u L = -\int_0^1 H_u \, dt = -\int_0^1 \nabla_u f \cdot \rho \, dt \end{cases}$$

$$\frac{d \log P_t(x_t)}{dt} = \frac{1}{\cancel{P_t(x_t)}} \cdot \Big[\underbrace{\partial_t P_t}_{-\nabla\cdot(P_t \vec{f}_t) \,=\, \cancel{-\nabla_x P_t \cdot \vec{f}_t} - \cancel{P_t} \nabla\cdot\vec{f}_t} + \cancel{\partial_x P_t \cdot \underbrace{\dot{X}_t}_{f_t}} \Big]$$

$$= -\nabla \cdot \vec{f}_t \qquad \vec{f}_t = \nabla \varphi$$

$$\begin{cases} \dot{X}_t = \vec{f} \\ \dot{\overline{\log P_t(X_t)}} = -\nabla \cdot f_t \end{cases}$$

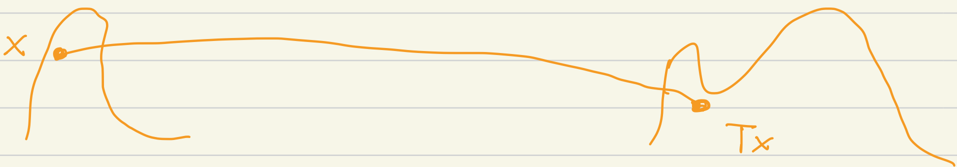
$$\nabla \cdot \vec{f} = \operatorname{tr}(\nabla^2 \varphi) = E[\xi^T \nabla^2 \varphi \, \xi]$$

$$\begin{pmatrix} \partial_1 \\ \partial_2 \\ \vdots \\ \partial_d \end{pmatrix}^T \begin{pmatrix} f_1 \\ f_2 \\ f_d \end{pmatrix} = \sum_i \partial_i \cancel{f_i} \, \partial_i \varphi$$

ρ_0, ρ_1 two densities

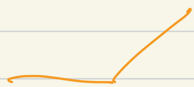
$$W_2^2(\rho_0, \rho_1) = \min_{T} \mathbb{E}_{x \sim \rho_0} \| Tx - x \|_2^2 , \quad T_{\#} \rho_0 = \rho_1$$

Brenier $\quad T = \nabla \varphi$, φ convex



$$f(x) = \sum_i w_i \, \sigma\Big(\sum_j u_j^i x_j \Big)$$

$\downarrow$

Relu  $\quad w \geq 0$

$$\dot{X}_t = \vec{b}_t(X_t) \qquad \partial_t P_t = -\nabla \cdot (P_t \vec{b}_t)$$

$$P_0 = \rho_0 , \quad P_1 = \rho_1$$

$$\int \int_0^1 P_t(x) \| \vec{b}_t(x) \|_2^2 \, dt \, dx = \int \int_0^1 \| b_t(X_t) \|_2^2 \, dt \cdot P_0(x) dx$$

$$= \int \int_0^1 \| \dot{X}_t \|_2^2 \, dt \cdot P_0(x) dx$$

$$\geq \int \Big\| \int_0^1 \dot{X}_t \Big\|_2^2 P_0(x) dx$$

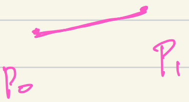
$$= \int \| X_1 - X_0 \|_2^2 P_0(x) dx \geq W_2^2$$

$$W_2^2(\rho_0, \rho_1) = \min_{P_t, \vec{b}_t} \int \int_0^1 P_t(x) \, \| b_t(x_t) \|_2^2 \, dt \, dx$$

$$\partial_t P_t = - \nabla \cdot (P_t \vec{b}_t) \quad ②$$

$$P_0 = \rho_0, \quad P_1 = \rho_1 \quad ①$$

$$= \min_{P_t \, ①} \int_0^1 \left[\min_{\vec{b}_t \, ②} \underbrace{\int P_t \cdot \| b_t \|_2^2 \, dx}_{f(b)} \right] dt$$



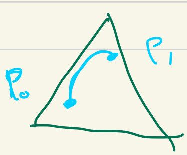
$$\frac{d f(b + \varepsilon d)}{d \varepsilon} \Big|_{\varepsilon = 0} = 2 \int \cancel{P_t} \cdot \langle b_t, P_t d \rangle \, dx = 0$$

$$b \perp p d$$

$$\cancel{\partial_t P_t} = - \nabla \cdot (P_t (b_t + \varepsilon d_t)) = - \nabla \cdot (\cancel{P_t b_t}) - \varepsilon \nabla \cdot (\underbrace{P_t d_t}_{=0})$$

Helmholtz decomposition $\Rightarrow$ $b = \nabla \varphi$

$$W_2^2 = \min_{P_t \, ①} \int_0^1 \underbrace{\min_{\nabla \cdot (p \nabla \varphi) = 0} \int P_t \cdot \| \nabla \varphi \|_2^2 \, dx}_{\| \partial_t P_t \|_{P_t}^2} \, dt$$



$h_1, h_2 : \mathbb{R}^d \to \mathbb{R}$ $\qquad \varphi_i : \mathbb{R}^d \to \mathbb{R}$

Weistein inner product

$$\langle h_1, h_2 \rangle_p = \int \langle \nabla \varphi_1(x), \nabla \varphi_2(x) \rangle \, p(x) \, dx$$

$$\boxed{\nabla \cdot (p \nabla \varphi_i) = -h_i}$$

$f(p)$: density $\to \mathbb{R}$

$$\partial_t p(t,x) \Big|_{t=0}$$

$$\frac{d f(p_t)}{dt}\Big|_{t=0} = \left\langle \nabla_{W_2} f(p(x)), \frac{d p_t(x)}{dt}\Big|_{t=0} \right\rangle_p$$

$$= \langle \nabla_{L_2} f(p), \underbrace{\frac{dp_t}{dt}\Big|_{t=0}}_{-\nabla \cdot (p \nabla \varphi_1)} \rangle_{L_2}$$

h_1

$$\boxed{\langle f, g \rangle_{L_2} = \int f(x) g(x) \, dx}$$

$$= \langle \nabla (\nabla_{L_2} f(p)), p \nabla \varphi \rangle_{L_2}$$

$$= \int \langle \nabla (\nabla_{L_2} f(p)), \nabla \varphi_1 \rangle \, p(x) \, dx$$

$$\underbrace{\nabla_{W_2} f(p)}_{h_2} = -\nabla \cdot \left(p \nabla \underbrace{(\nabla_{L_2} f(p))}_{\varphi_2} \right)$$

$$f(p) = \int u(p(x))\,dx, \qquad u: \mathbb{R}_+ \to \mathbb{R}$$

$$= \int p(x) \log p(x)\,dx, \qquad u(t) = t \log t$$

$$\frac{d f(p+\varepsilon q)}{d\varepsilon} = \int u'(p(x)) \cdot q(x)\,dx$$

$$= \langle u'(p), q \rangle_{L_2}$$

$$\nabla_{L_2} f = u'(p)$$

$$\nabla_{W_2} f(p) = -\nabla \cdot \left(p \nabla (\nabla_{L_2} f) \right)$$

$$= -\nabla \cdot \left(p\, u''(p) \cdot \nabla p \right)$$

$$\nabla_{W_2} f(p) = -\Delta p$$

$$q(x) \propto \exp(-\beta \cdot \varphi(x))$$

$$\boxed{dX_t = -\nabla\varphi(X_t)\,dt + \sqrt{2\beta^{-1}}\,dB_t}$$

$$\partial_t p = +\nabla\cdot\left(p\nabla\varphi(X_t)\right) + \frac{1}{2}\nabla^2\cdot\left(p\,2\beta^{-1}\cdot I\right)$$

$$= \nabla\cdot(p\nabla\varphi) + \beta^{-1}\Delta p$$

$$= \nabla\cdot\left(p\,(\underbrace{\nabla\varphi + \beta^{-1}s_p})\right) = 0$$

$$\nabla \log p = -\nabla(\beta^{-1}\varphi)$$

$$\Leftrightarrow \quad p \propto e^{-\beta^{-1}\varphi}$$

$$\frac{dp_t}{dt} = -\nabla_{W_2} KL(p_t \,\|\, q)$$

$$\frac{1}{2}\|p_t - q\|_1^2 \le KL(p_t \,\|\, q) \le e^{-\beta t} KL(p_0 \,\|\, q)$$

Pinsker