不一定要 is divergence.

f-divergence.

$$D_f(P\|Q) = \int_x q(x) f\left(\frac{p(x)}{q(x)}\right) dx$$

$$\begin{cases} f \text{ is convex} \\ f(1) = 0 \end{cases}$$

If $p(x) = q(x)$ for all x $\quad D_f(P\|Q) = 0$

$\because f$ is convex,

$$\therefore D_f(P\|Q) \geq f\left(\int_x q(x) \frac{p(x)}{q(x)} dx\right)$$

$$= f(1)$$

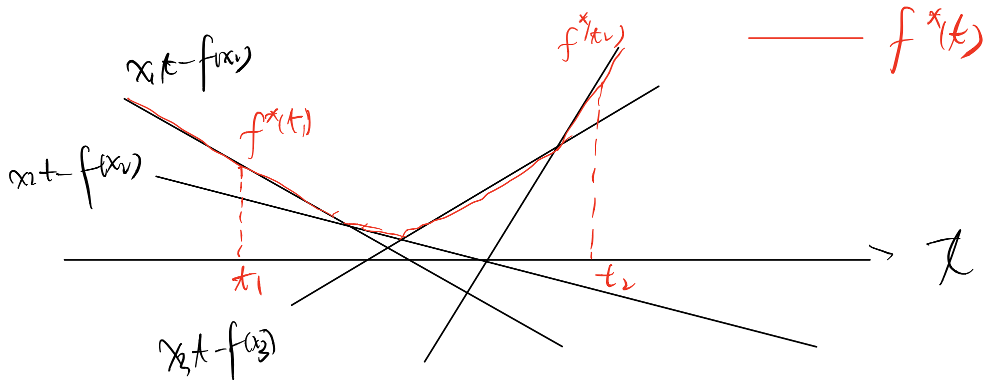
$$= 0 \qquad \text{0 为最小值}$$

$f(x) = x\log x$ KL.

$-\log x$ Reverse KL

$(x-1)^2$ Chi Square

f 共轭函数 $f^*(t) = \max\limits_{x \in dom(f)} \{xt - f(x)\}$



$$\begin{cases} f(x) = x\log x \\ f^*(t) = e^{t-1} \end{cases}$$

Connection with GAN

$$f^*(t) = \max_{x \in dom(f)} \{xt - f(x)\} \longleftrightarrow f(x) = \max_{t \in dom(f^*)} \{xt - f^*(t)\}$$

$$D_f(P\|Q) = \int_x q(x) f\left(\frac{p(x)}{q(x)}\right) dx$$

$$= \int_x q(x) \left(\max_{t \in \mathrm{dom}(f^*)} \left\{ \frac{p(x)}{q(x)} t - f^*(t) \right\} \right) dx$$

$$\simeq \max_D \int_x q(x) p(x) dx - \int_x q(x) f^*(D(x)) dx$$

D is a function, whose input is x and output is t

$$D_f(P\|Q) \geq \int_x q(x) \left(\frac{p(x)}{q(x)} D(x) - f^*(D(x)) \right) dx$$

$$= \int_x p(x) D(x)\, dx \int_x q(x)\, f^*(D(x)) dx$$

$$D_f(P\|Q) \simeq \max_D \left\{ E_{x \sim P}[D(x)] - E_{x \sim Q}[f^*(D(x))] \right\}$$

$$D_f(P_{data}\| P_G) = \max_D \left\{ E_{x \sim P_{data}}[D(x)] - E_{x \sim P_G}[f^*(D(x))] \right\}$$

$$G^* = \arg\min_G D_f(P_{data}\| P_G)$$

$$= \arg\min_G \max_D \left\{ E_{x \sim P_{data}}[D(x)] - E_{x \sim P_G}[f^*(D(x))] \right\}$$

$$= \arg\min_G \max_D V(G, D)$$

divergence