Generative → learns joint distribution $p(x, y) = p(x|y)\,p(y)$

fake pair $G(x$, 0 (label))
real pair $(x, 1)$

Discriminator → (0, 1) → D loss → objective function
↑ fake ↑ real

noise ↓ G loss → objective function.

$z \to G(z) \to x_{fake}$

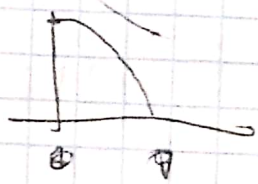
$x \to$ training x_{real}

$D(x) \to$ Discrim $\quad x_{real} \to P(y \mid x_{real}) = \{0, 1\}$

$D(G(z)) \to$ from $x_{fake} \quad P(y \mid x_{fake}) = \{0, 1\}$

$D(x) \uparrow$
$D(G(z)) \downarrow$ // Discri.

$D(G(z)) \uparrow$ // generator



$D_{loss} = D_{loss\,real} + D_{loss\,fake}$

$= \log(D(x)) + \log(1 - D(G(z)))$

Ds

total cost

$$\frac{1}{m}\sum_{i=1}^{m} \log\left(D(x^{(i)})\right) + \log\left(1 - D(G(z^{(i)}))\right)$$

Gen

$$\log(1 - D(G(z)))$$

$$\Rightarrow \frac{1}{m}\sum_{i=1}^{m} -\log\left(D(G(z^{(i)}))\right)$$

$$\min_{G}\max_{D} V(D, G) = \mathbb{E}_{x \sim p_{data}}\left[\log(D(x))\right] + \mathbb{E}_{z \sim p(z)}\left[1 - \log(1 - D(G(z)))\right]$$

KL : $D_{KL}(P||q) = \int_x P(x) \log \frac{P(x)}{q(x)} dx \geq 0$ unbounded upon

JS diver: $D_{JS}(P||q) = \frac{1}{2} D_{KL}\left(P \,||\, \frac{P+q}{2}\right) + \frac{1}{2} D_{KL}\left(q \,||\, \frac{P+q}{2}\right)$

Bounded $[0, 1]$

Optimal value of D:

$$L(G, D) = \int_x \left(P_r(x) \log(D(x)) + P_g(x) \log(1 - D(x)) \right) dx$$

$$\tilde{x} = D(x), \quad A = P_r(x), \quad B = P_g(x)$$

$$f(\tilde{x}) = A \log \tilde{x} + B \log(1 - \tilde{x})$$

$$\Rightarrow \frac{\partial f(\tilde{x})}{\partial \tilde{x}} = \frac{1}{\log 10} \frac{A - (A+B)\tilde{x}}{\tilde{x}(1-\tilde{x})} \Rightarrow 0 \text{ then}$$

$$D^*(x) = \tilde{x}^* = \frac{A}{A+B} = \frac{P_r(x)}{P_r(x) + P_g(x)} \in [0, 1]$$

$\boxed{\text{forward trained } g}$ if $P_g = P_r$ then $D^*(x) = \frac{1}{2}$

Both at optimal value:

$$L(G, D^*) = \int_x \left[P_r(x) \log(D^*(x)) + P_g(x) \log(1 - D^*(x)) \right] dx$$

$$= -2 \log 2 \quad // \; (P_r = P_g)$$

$$D_{JS}(P_r \| P_g) = \frac{1}{2} D_{KL}\left(P_r \| \frac{P_r + P_g}{2}\right) + \frac{1}{2} D_{KL}\left(P_g \| \frac{P_r + P_g}{2}\right)$$

$$= \frac{1}{2}\left[\log 4 + \cancel{\log} \; L(G, D^*)\right]$$

$$L(G, D^*) = 2 D_{JS}(P_r \| P_g) - 2\log 2$$

$$\left[L(G^*, D^*) = -2\log 2 \right]$$

Wasserstein loss:

$$W(P_r, P_g) = \inf_{\gamma \sim \Pi(P_r, P_g)} \mathbb{E}_{(x,y)\sim\gamma}\left[\|x - y\|\right]$$

$$W(P_r, P_g) = \frac{1}{K} \sup_{\|f\|_L \le K} \mathbb{E}_{x\sim P_r}\left[f(x)\right] - \mathbb{E}_{x\sim P_g}\left[f(x)\right]$$

$$L(P_r, P_g) = W(P_r, P_g) = \max_{w \in W} \mathbb{E}_{x\sim P_r}\left[f_w(x)\right] - \mathbb{E}_{z\sim P_r}\left[f_w(g_\theta(z))\right]$$

$\left\{ \{f_w\}_{w\in W} \rightarrow K \text{ Lipschitz continuous function} \right\}$

GAN family of Alternatives

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