Семинар 2

2024 курс: Генеративные модели

Семинарист: Кондратенко Владимир

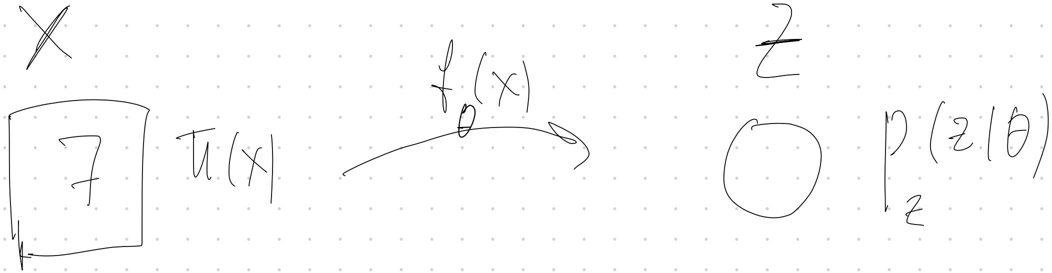
$\forall z \quad 1 + h'(w^T z + b) w^T u \neq 0 \quad > 0$

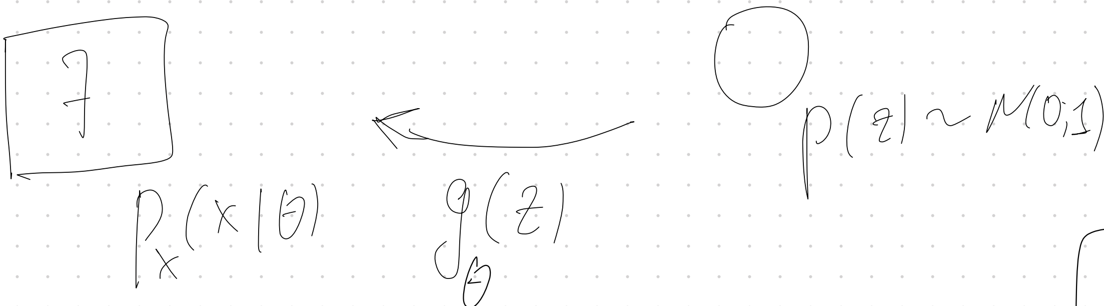
$\forall z \quad 0 < h'(\ldots) < 1$

$h'(\ldots) w^T u > -1$

$w^T u > \dfrac{-1}{h'(\ldots)} \quad (0;1) \quad \Rightarrow \boxed{w^T u \geq -1} \quad (-\infty; -1)$

$$\boxed{g(z) = z + u\, h(w^T z + b)}$$





$\pi(x)$ — страшное, не понятное

$$p(x|\theta) = p_z(f(x)) \cdot |\det J_f|$$

$$p(z|\theta) = \underline{\underline{p_x}}(g(z)) \cdot |\det J_g|$$

$\pi(x)$

$$p(z) \sim N(0;1)$$

$$\boxed{\begin{gathered} p_x(x)\,dx = \\ = p_z(z)\,dz \\ \Rightarrow \\ p(x) = p(z)\frac{dz}{dx} \end{gathered}}$$

Хотим оценивать $\pi(x)$? Forward KL

1. $X \to z$, в z умеем оценивать плотн.
2. $z \to X$, знаем $p(z)$ + знаем преобр

$$KL(\pi, p(x|\theta)) = \int \pi(x) \log \frac{\pi(x)}{p(x|\theta)} dx =$$

$$= -\int \pi(x) \log p(x|\theta) dx + \underbrace{\int \pi(x) \log \pi(x) dx}_{const(\theta)}$$

$$= -\mathbb{E}_\pi \log p(x|\theta) = -\mathbb{E}_\pi \log \left[p_z(f(x)) \cdot |\det J_f| \right]$$

$$= -\mathbb{E}_\pi \left[\log p_z(f(x)) + \log |\det J_f| \right]$$

$$KL[\pi ; p(x|\theta)] \simeq -\mathbb{E}_{\pi}\left[\log p_z(f(x)) + \log|\det J_f|\right]$$

1. $f(x) = z$ $\sim N(0;1)$

2. $p_z(f(x)) = p_z(z)$

3. $\det J_f$

4. sample $x \sim \pi$

1. $\boxed{gHH}$ — не можем семплировать

2. MLE

$$\boxed{KL[\pi ; p(x|\theta)]} = -\mathbb{E}_{\pi}\left[\log p_z(f(x)) + \log|\det J_f|\right] +$$

$$+\ \mathbb{E}_{\pi} \log \pi(x) =$$

$$|\det J_f| = \frac{1}{|\det J_g|}$$

$$= \mathbb{E}_{\pi}\left[\log \pi(x) - \log p_z(f(x)) - \log|\det J_f|\right] =$$

$$= \boxed{\text{LOTUS} \quad \mathbb{E}_{p(y)}\, g(y) = \mathbb{E}_{p(x)}\, g(f(x)) \qquad f(x) = z \quad g(x) = \cancel{z}} =$$

$$= \mathbb{E}_{p(z|\theta)}\left[\log \pi(g(z)) + \log|J_g| - \log p_z\right] =$$

$$= \mathbb{E}_{p(z|\theta)}\left[\log p_z(z|\theta) - \log p(z)\right] =$$

$$= \boxed{KL[p_z(z|\theta) ; p(z)]}$$

X

$\pi(x)$

$f(x) = z$

$p(z|\theta)$

$p(z)$

$g(z) = x$

$p(x|\theta)$

$$KL[p(x|\theta);\pi(x)] = KL[p(z);p(z|\theta)]$$

($\boxed{\text{Reverse KL}}$

$$\int p(x|\theta)\left[\log p(x|\theta) - \log \pi(x)\right]dx =$$

$$= \mathbb{E}_{p(x|\theta)}\left[\log p_z(f(x)) + \log|\det J_f| - \log \pi(x)\right]$$

$$= \mathbb{E}_{p(z)}\left[\log p_z(z) - \log|\det J_g| - \log \pi(g(z))\right]$$

RKL

$$\mathbb{E}_{p(z)}\left[\log p_z(z) - \log|\det J_g| - \log \pi(g(z))\right]$$

1. $g(z) = x$ $\qquad z \xrightarrow{g} x$
2. оценивать $\pi(x)$
3. $\det J_g$
4. sampling $p(z)$

$x \sin x$

$|x| > 1$