1　GAN

1.1　GAN の Loss について

Discriminator の Loss は以下の通り.

$$
\begin{aligned}
&\mathbb{E}_{\boldsymbol{x}\sim p_{data}(\boldsymbol{x})}\left[\log D(\boldsymbol{x})\right]+\mathbb{E}_{\boldsymbol{z}\sim p_z(\boldsymbol{z})}\left[\log\left(1-D\left(G\left(\boldsymbol{z}\right)\right)\right)\right]\\
&=\mathbb{E}_{\boldsymbol{x}\sim p_{data}(\boldsymbol{x})}\left[\log D(\boldsymbol{x})\right]+\mathbb{E}_{\boldsymbol{x}\sim p_g(\boldsymbol{x})}\left[\log\left(1-D(\boldsymbol{x})\right)\right] \qquad (1)
\end{aligned}
$$

ただし，$\boldsymbol{x}\sim p_g(\boldsymbol{x})$ を $\boldsymbol{x}=D\left(G(\boldsymbol{z})\right), \boldsymbol{z}\sim p_g(\boldsymbol{x})$ とした．また，Generator の Loss は以下の通り.

$$
\mathbb{E}_{\boldsymbol{x}\sim p_g(\boldsymbol{x})}\left[\log\left(1-D(\boldsymbol{x})\right)\right] \qquad (2)
$$

ここで，$p_{data}(\boldsymbol{x}), p_g(\boldsymbol{x}), p_z(\boldsymbol{z})$ はそれぞれデータの生成分布と生成モデルからの出力分布，ノイズの分布を表す.

ただし，式 (1), 式 (2) はデータの生成分布，生成モデルからの出力分布で期待値を取っているが，実際は以下の経験分布 (瀧 [1] p94 式 (5.2)　参照) を用いて近似的に汎化誤差を見積もった Loss を目的関数として最適化を行う.

$$
\frac{1}{m}\sum_{i=1}^{m}\left[\log D(\boldsymbol{x}_i)+\log\left(1-D(\boldsymbol{x}_i)\right)\right] \qquad (3)
$$

$$
\frac{1}{m}\sum_{i=1}^{m}\log\left(1-D\left(G(\boldsymbol{z}_i)\right)\right) \qquad (4)
$$

m はデータ数を表す.

2　Loss の数学的解析

GAN は以下の価値関数で表される，Discriminator と Generato の two-player minmax game を行う.

$$
\min_G\max_D V(D,G)=\mathbb{E}_{\boldsymbol{x}\sim p_{data}(\boldsymbol{x})}\left[\log D(\boldsymbol{x})\right]+\mathbb{E}_{\boldsymbol{z}\sim p_z(\boldsymbol{z})}\left[\log\left(1-D\left(G\left(\boldsymbol{z}\right)\right)\right)\right] \qquad (5)
$$

この時，Generator が固定のもと，$\max_D V(D,G)$ の最適値は

$$
D^*_G(\boldsymbol{x})=\frac{p_{data}(\boldsymbol{x})}{p_{data}(\boldsymbol{x})+p_g(\boldsymbol{x})}
$$

である.

Proof. GAN の評価関数より，

$$
\begin{aligned}
V(D,G)=&\mathbb{E}_{\boldsymbol{x}\sim p_{data}(\boldsymbol{x})}\left[\log D(\boldsymbol{x})\right]+\mathbb{E}_{\boldsymbol{z}\sim p_z(\boldsymbol{z})}\left[\log\left(1-D\left(G\left(\boldsymbol{z}\right)\right)\right)\right] \qquad (6)\\
=&\int_{\boldsymbol{x}}p_{data}(\boldsymbol{x})\log\left(D(\boldsymbol{x})\right)d\boldsymbol{x}+\int_{\boldsymbol{z}}p_{\boldsymbol{z}}(\boldsymbol{z})\log\left(1-D\left(G(\boldsymbol{z})\right)\right)d\boldsymbol{z} \qquad (7)
\end{aligned}
$$

ここで，$\boldsymbol{x}=G(\boldsymbol{z})$ では $\boldsymbol{x}\sim p_g(\boldsymbol{x})$ なので

$$
V(D,G)=\int_{\boldsymbol{x}}\left\{p_{data}(\boldsymbol{x})\log D(\boldsymbol{x})+p_g(\boldsymbol{x})\log\left(1-D\left(\boldsymbol{x}\right)\right)\right\}d\boldsymbol{x} \qquad (8)
$$

さらに式 (8) において Discriminator の最適値とその時の Discriminator の値を考える．

$$\max_D V(D,G) = \max_D \int_{\boldsymbol{x}} \left\{ p_{data}(\boldsymbol{x}) \log D(\boldsymbol{x}) + p_g(\boldsymbol{x}) \log \left(1 - D\left(\boldsymbol{x}\right)\right) \right\} d\boldsymbol{x} \tag{9}$$

今，積分範囲は $\boldsymbol{x}$ が取りうるすべての値についてなので，積分の中を最大にする D が最適値をとる時の Discriminator の値である．よって

$$\max_D \; p_{data}(\boldsymbol{x}) \log D(\boldsymbol{x}) + p_g(\boldsymbol{x}) \log \left(1 - D\left(\boldsymbol{x}\right)\right) \tag{10}$$

を考える．また，関数 $a \log x + b \log(1-x)$ は

$$x = \frac{a}{a+b} \tag{11}$$

で最大値をとる．よって，式 (9) において Discriminator の最適値を与える時のとる値 $D_G^*(\boldsymbol{x})$ は

$$D_G^*(\boldsymbol{x}) = \frac{p_{data}(\boldsymbol{x})}{p_{data}(\boldsymbol{x}) + p_g(\boldsymbol{x})} \tag{12}$$

である．またこの時の式 (7) の値 (最適値)$C(G)$ は

$$\begin{aligned}
C(G) &= \max_D V(D,G) && (13)\\
&= V(D_G^*, G) && (14)\\
&= \mathbb{E}_{\boldsymbol{x} \sim p_{data}(\boldsymbol{x})} \left[\log D_G^*(\boldsymbol{x})\right] + \mathbb{E}_{\boldsymbol{z} \sim p_z(\boldsymbol{z})} \left[\log \left(1 - D_G^*\left(G\left(\boldsymbol{z}\right)\right)\right)\right] && (15)\\
&= \mathbb{E}_{\boldsymbol{x} \sim p_{data}(\boldsymbol{x})} \left[\log D_G^*(\boldsymbol{x})\right] + \mathbb{E}_{\boldsymbol{x} \sim p_g(\boldsymbol{x})} \left[\log \left(1 - D_G^*\left(\boldsymbol{x}\right)\right)\right] && (16)\\
&= \mathbb{E}_{\boldsymbol{x} \sim p_{data}(\boldsymbol{x})} \left[\log \frac{p_{data}(\boldsymbol{x})}{p_{data}(\boldsymbol{x}) + p_g(\boldsymbol{x})}\right] + \mathbb{E}_{\boldsymbol{x} \sim p_g(\boldsymbol{x})} \left[\log \left(1 - \frac{p_{data}(\boldsymbol{x})}{p_{data}(\boldsymbol{x}) + p_g(\boldsymbol{x})}\right)\right] && (17)\\
&= \mathbb{E}_{\boldsymbol{x} \sim p_{data}(\boldsymbol{x})} \left[\log \frac{p_{data}(\boldsymbol{x})}{p_{data}(\boldsymbol{x}) + p_g(\boldsymbol{x})}\right] + \mathbb{E}_{\boldsymbol{x} \sim p_g(\boldsymbol{x})} \left[\log \frac{p_g(\boldsymbol{x})}{p_{data}(\boldsymbol{x}) + p_g(\boldsymbol{x})}\right] && (18)
\end{aligned}$$

□

参考文献

[1] 瀧 雅人 「深層学習」講談社

[2] Ian Goodfellow, Generative Adversarial Networks
https://arxiv.org/abs/1406.2661