Ex 4-1

- x = tf. placeholder(.... [None, pic_size, pic_size, 1])

Batch size, not limited.

- show Tensor board.

Ex 4-2

a) $\varepsilon(\vec{w}) = \sum_{i=1}^{n} \| x_i - \sum_{i \neq j} w_{ij} x_j \|^2 = 0$

we have $x_i = \sum_{i \neq j} w_{ij} x_j$.

which means. the manifold is exactly linear around x_i.

b) translation w.r.t. each input x, the error (cost) function:

$$\varepsilon(\vec{w}) = \sum_{i=1}^{n} \| (x_i - c) - \sum_{j \neq i} w_{ij} (x_j - c) \|^2$$

$$= \sum_{i=1}^{n} \| x_i - c - \sum_{j \neq i} w_{ij} x_j + \sum_{j \neq i} w_{ij} \cdot c \|^2$$

$$= \sum_{i=1}^{n} \| x_i - c - \sum_{j \neq i} w_{ij} x_j + c \|^2 = \varepsilon(\vec{w})$$

c) irregular (ill-posed problem)

d)

$$\|x_i - \sum_j w_{ij} x_j\|^2 + \alpha \sum_j w_{ij}^2$$

$$\boxed{Af = F \quad f \in \mathbb{F}}$$

d) regularization

Ex 4-3.

a). train a generator that is able to produce samples

b) generator v.s. discriminator

c) high quality artificial samples.

d) conditional GAN

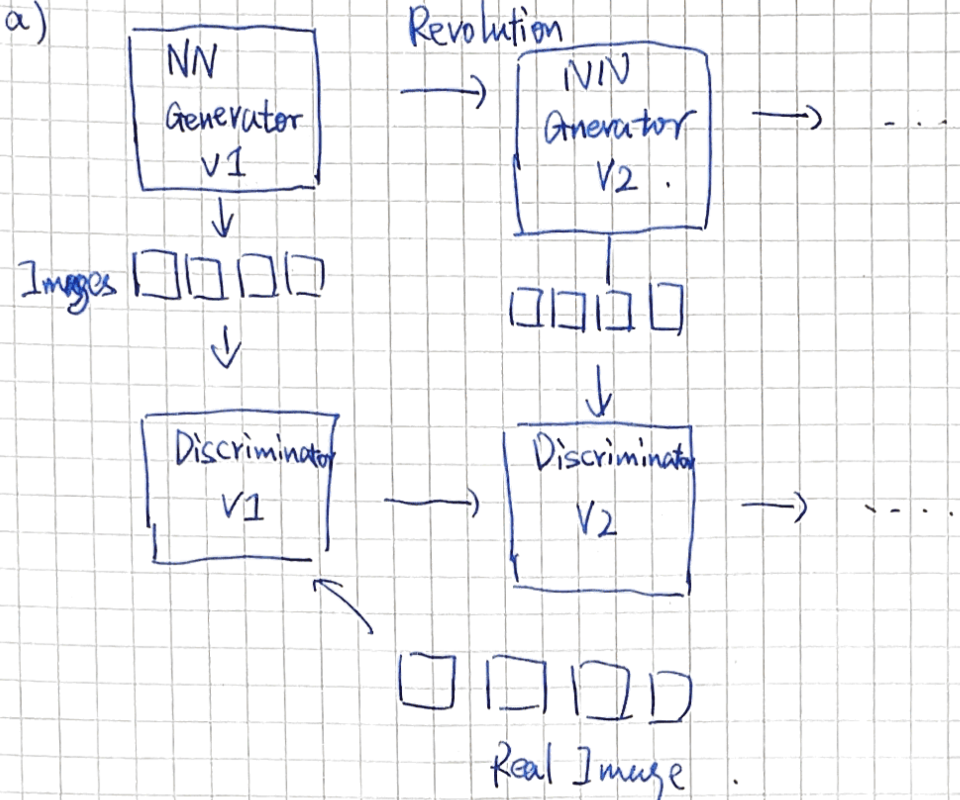
infoGAN.

cycle GAN.

BRUNNEN

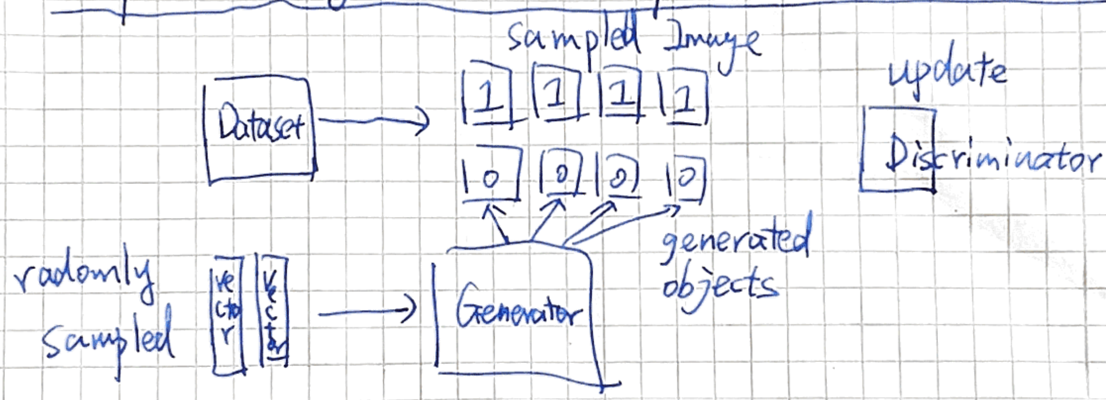
Ex 4-3

a)

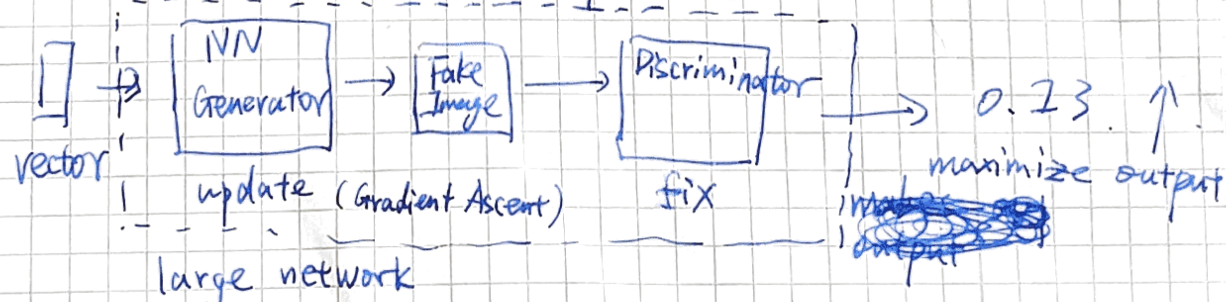


Algorithm: ① Initializ generator and discriminator $\boxed{G}$ $\boxed{D}$.
② In each training iteration:

Step 1: Fix generator G, update discriminator D



Step 2: Fix discriminator D, update generator G.



b) artificial data generation:

- Image Inpainting
- Super-resolution
- ....

c) - Conditional GAN: use class lable as an additional input.

- Info GAN
- Cycle GAN
- Introduce Zoo-GAN
- check Lecture slides.

Ex 4-4

a) This is the first step of Algorithm.

$$d_g^* = \arg\max_d V(g,d) = \arg\max_d \left\{ \int_x p_{data}(x) \log d(x) + p_g(x) \log(1-d(x))\, dx \right\}.$$

According to the Hint, $\begin{cases} a = p_{data}(x) \\ b = p_g(x) \end{cases}$

therefore: $d_g^* = \dfrac{p_{data}(x)}{p_{data}(x) + p_g(x)}$

b) ~~This is the second step of Algorithm~~: Given a fixed g^*, $p_{g^*} = p_{data}$.

we have $d_g^* = \dfrac{1}{2}$.

The $V(g, d_g^*) = \int_x p_{data}(x) \log \frac{1}{2} + p_g(x) \log \frac{1}{2}\, dx$

$= \log\frac{1}{2} \int_x [p_{data}(x) + p_g(x)]\, dx = 2\log\frac{1}{2} = \log\frac{1}{4} = -\log 4$

$= -2\log 2$

BRUNNEN

c) We still need to show the optimal solution (g^*, d^*) is the unique global optimum of V.

For all g, $V(g^*, d^*) \le V(g, d_g^*)$ if and only if $g = g^*$

Step 2. fix discriminator, update generator:

$$V(g, d^*) = \int_x \left[P_{data}(x) \log \frac{P_{data}(x)}{P_{data}(x) + P_g(x)} + P_g(x) \log \left(1 - \frac{P_{data}(x)}{P_{data}(x) + P_g(x)}\right) \right] dx$$

$$= \int_x \left[P_{data}(x) \log \frac{P_{data}(x)}{P_{data}(x) + P_g(x)} + P_g(x) \log \frac{P_g(x)}{P_{data}(x) + P_g(x)} \right] dx$$

$$= \int_x \left[P_{data}(x) \log \frac{\frac{1}{2} P_{data}(x)}{\frac{1}{2}(P_{data}(x) + P_g(x))} + P_g(x) \log \frac{\frac{1}{2} P_g(x)}{\frac{1}{2}(P_{data}(x) + P_g(x))} \right] dx$$

$\log ab = \lg a + \lg b$

$$= \int_x \left[\log \frac{1}{2} P_{data}(x) + P_{data}(x) \log \frac{P_{data}(x)}{\frac{1}{2}(P_{data}(x) + P_g(x))} \right] dx$$

$$+ \int_x \left[\log \frac{1}{2} \cdot P_g(x)\, dx + P_g(x) \log \frac{P_g(x)}{\frac{1}{2}(P_{data}(x) + P_g(x))} dx \right]$$

$$= -2\log 2 + \int_x P_{data}(x) \log \frac{P_{data}(x)}{\frac{1}{2}(P_{data} + P_g)} dx + \int_x P_g(x) \log \frac{P_g(x)}{P_{data} + P_g} dx$$

$$= -2\log 2 + D_{KL}\left(P_{data} \,\|\, \frac{P_{data} + P_g}{2}\right) + D_{KL}\left(P_g \,\|\, \frac{P_{data} + P_g}{2}\right)$$

$$= -2\log 2 + 2 D_{JS}(P_{data} \,\|\, P_g) \ge -2\log 2 + 0$$

$$= -2\log 2$$

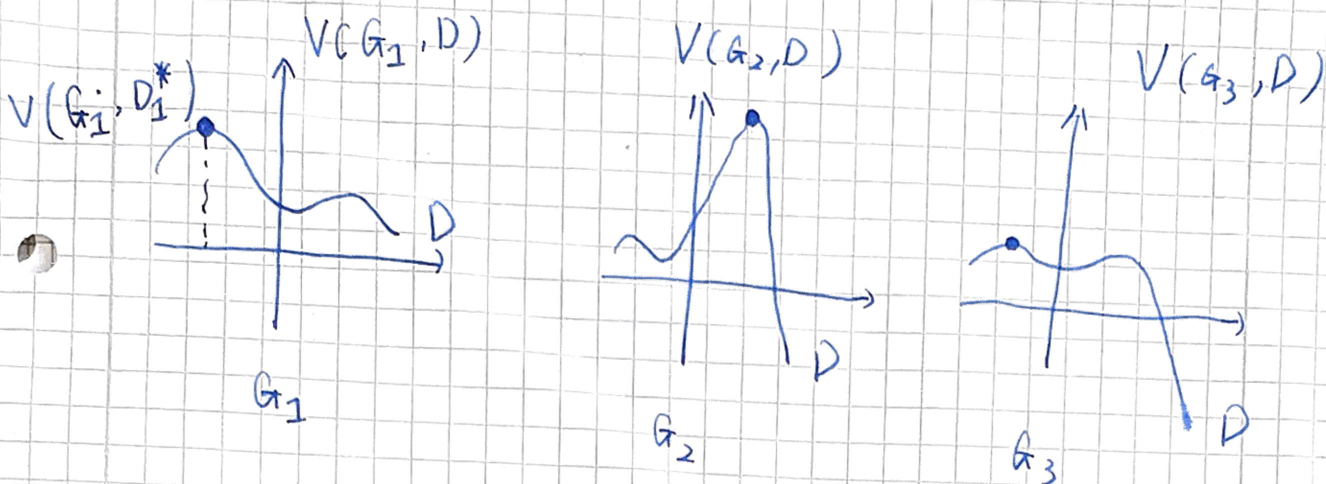
$\therefore \arg\min_g V(g, d_g^*) = -2\log 2$. if and only if $P_{data} = P_g$. i.e. $g = g^*$.

BRUNNEN

Example:

$$G^* = \arg\min_{g} \left[\max_{d_g} V(g, d_g) \right]$$

$$D^* = \arg\max_{d} V(d, g)$$

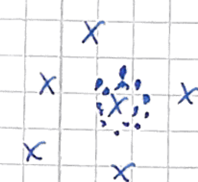


If G_1 is fixed.

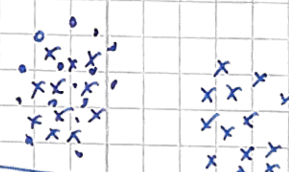
• is the best D^*.

✓

Mode collapse

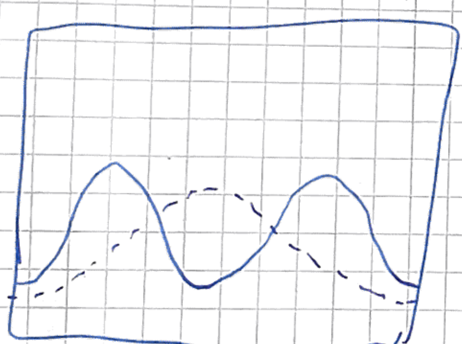


Mode dropping



Flaw in Optimization.

solution! maybe ensemble

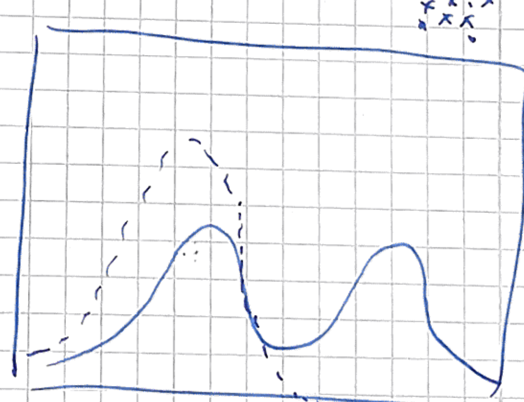


— Pdata

----- Pg.

minimize KL divergence.

$D_{KL}(P_{data} \| P_g)$



— Pdata.

---- Pg.

minimize reverse KL.

$D_{KL}(P_g \| P_{data})$