Generator G

Discriminator D

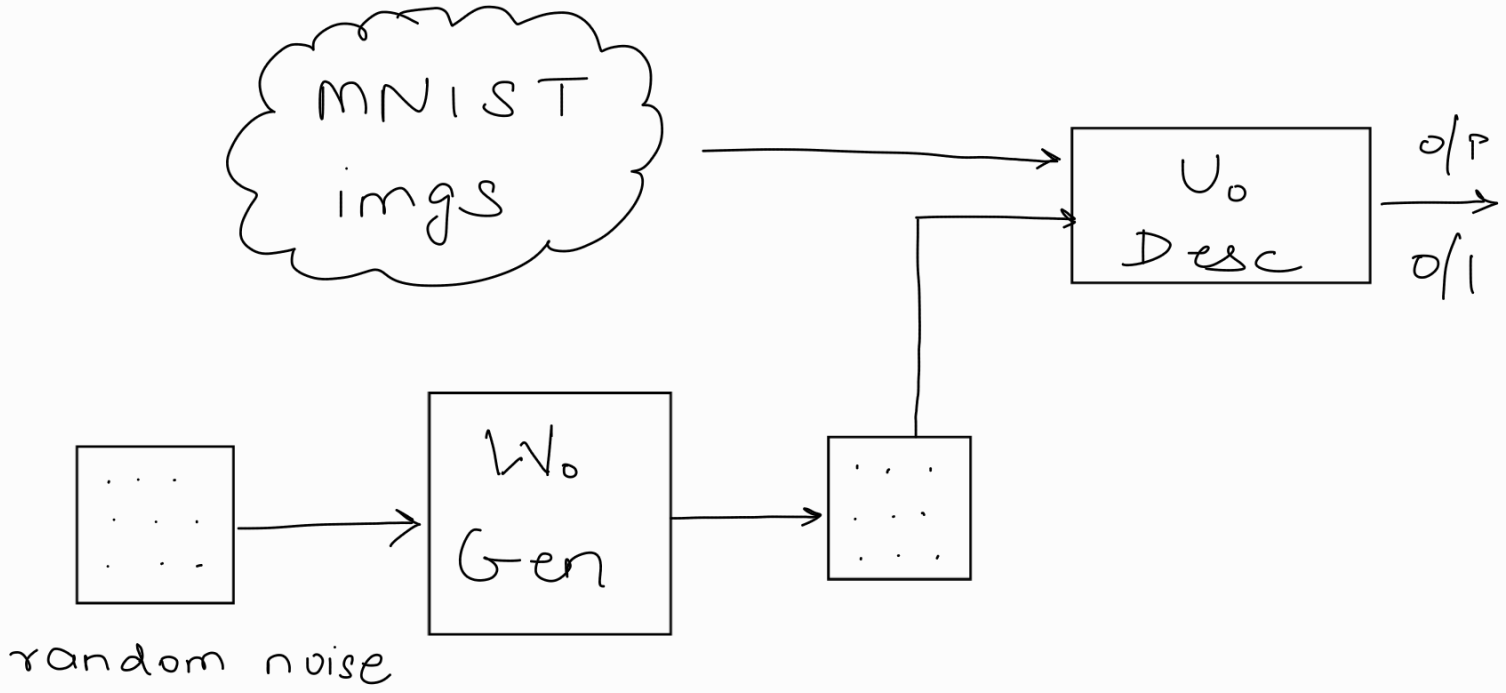
D tries to discriminate between:

1] A sample from data distribution

2] A sample from G.

G tries to trick D by generating samples that are hard for D to distinguish from data.

Called a zero sum game.



1] Generate 1000-2000 images using random initializations of W_0

2] Train the discriminator s.t. when the i/p is from MNIST, output 1 and when i/p is from the generator, output 0.

3] When discriminator starts performing well, freeze its weights. Now train the generator s.t. if disc. outputs 0 for gen images, backprop a higher loss. Basically, train the generator to fool the discriminator.

4] Repeat 1, 2, 3 till generator starts producing images similar to data distribution.

MLE

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

real i/p generated i/p

Generator Loss

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

Minimize this