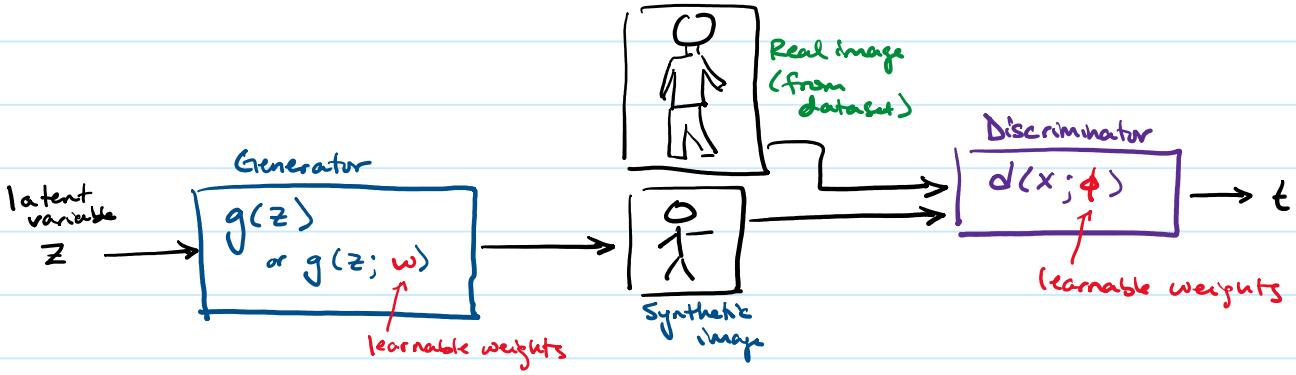
# GANS (ch. 17 Bishop)

Sunday, November 17, 2024 4:37 PM

**Generative Adversarial Networks** (ch. 17 Bishop) Goodfellow et al. '14




$z \sim p(z)$, something simple we can sample from
(almost always $N(0, I)$ Gaussian)

Train generator and discriminator jointly!
- discriminator → tries to distinguish synthetic from real
- generator → tries to fool discriminator

output: $t = 1$ if it thinks $x$ is real
$t = 0$ if it thinks $x$ is synthetic

or more precisely,

$$P[t = 1] = d(x; \phi)$$

(last layer has sigmoid so output is in $(0,1)$)

for discriminator, use usual cross-entropy loss:

$$E(w, \phi) = -\frac{1}{N} \sum_{n=1}^{N} t_n \log(d_n) + (1 - t_n) \log(1 - d_n)$$

($t_n$: true label, $d_n$: predicted)

$$d_n = d(x_n; \phi)$$

... but use a mix of real + synthetic

$$= \frac{-1}{N_{real}} \sum_{n \in real} \underbrace{\log(d_n)}_{t_n = 1} - \frac{1}{N_{synth}} \sum_{n \in synth} \underbrace{\log(1 - d(g(z_n, w), \phi))}_{t_n = 0}$$

Key idea: minimize $E$ with respect to $\phi$ (discriminator)
but maximize $E$ with respect to $w$ (generator)

i.e.
$$\phi \leftarrow \phi - \eta \cdot \nabla_\phi E(w, \phi)$$ gradient descent
$$w \leftarrow w + \eta \nabla_w E(w, \phi)$$ gradient ascent
} alternate

"Saddle-point problem"

$$\max_w \min_\phi E(w, \phi) \quad \left( \text{sometimes} = \min_\phi \max_w E(w, \phi) \right)$$

# GANS (p. 2)

Tuesday, December 10, 2024 6:33 PM

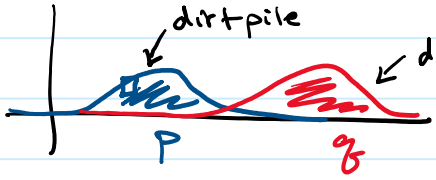
## Details, variants...

- conditional GANs
- not easy to train due to saddle-point nature
  - no clear metric of progress either
- mode collapse

  eg. on MNIST, generator learns how to generate realistic "3"s but no other digits

  - a few proposed fixes
- Change metric to... Wasserstein distance

  aka Earth mover's distance, part of optimal transport



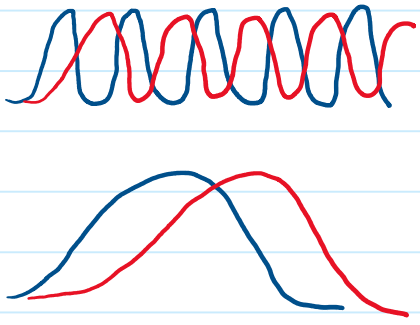

Assuming you are as efficient as possible, how much work moving dirt is required?

why?

- KL divergence requires shared support

  (so $\log\left(\frac{p(x)}{q(x)}\right)$ well-defined)

  If barely so, can have issues



← huge squared error loss

← similar Wasserstein loss

← small squared error loss

- CycleGANS

  used for things like style transfer
- Representation learning - features in latent space

  path from $z_1$ to $z_2$ leads to meaningful path $x_1$ to $x_2$