TRY AGAN

Dhruv Madeka
Quantitative Researcher

GENERATIVE MODELLING

Goodfellow NIPS [2016]

Given a set of features x and labels y
Try to learn:

$$\mathbb{P}(x, y)$$

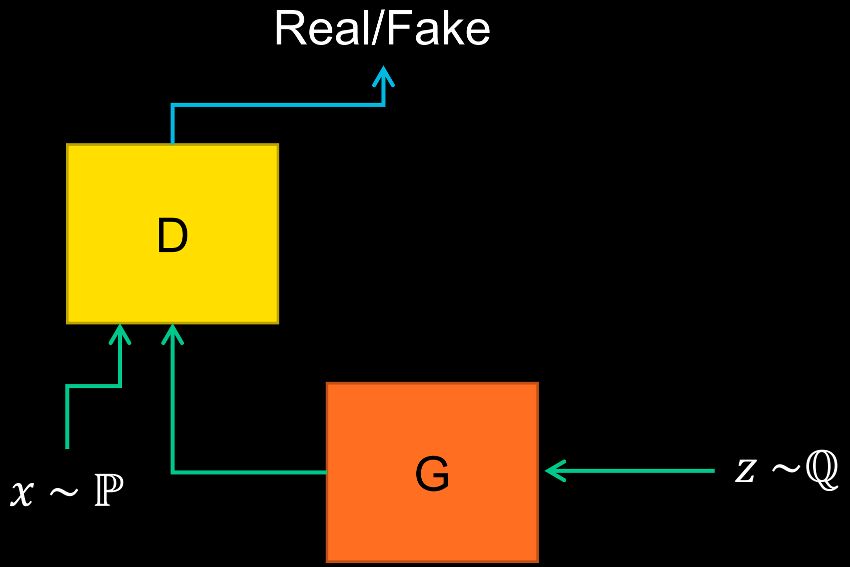
- Learn a distribution
- Generate samples

GENERATIVE ADVERSARIAL NETWORKS

Goodfellow et al [2014]

$$\min \max V(G, D)$$

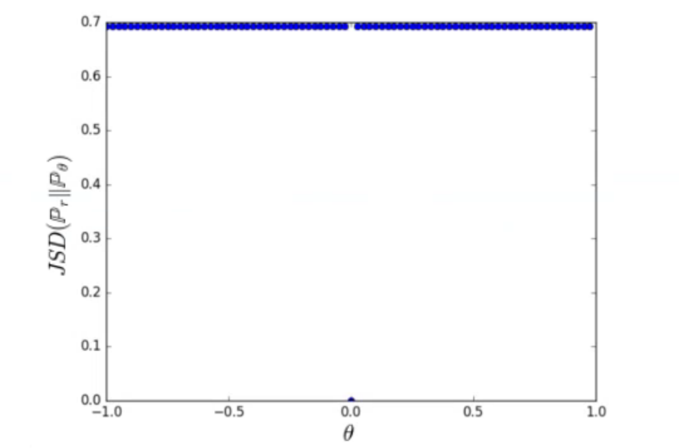
$$\min_{G} \max_{D} \mathbb{E}_{\mathbb{P}}\left[\log\left(D(x)\right)\right] + \mathbb{E}_{\mathbb{Q}}\left[\log\left(1 - D\left(G(z)\right)\right)\right]$$

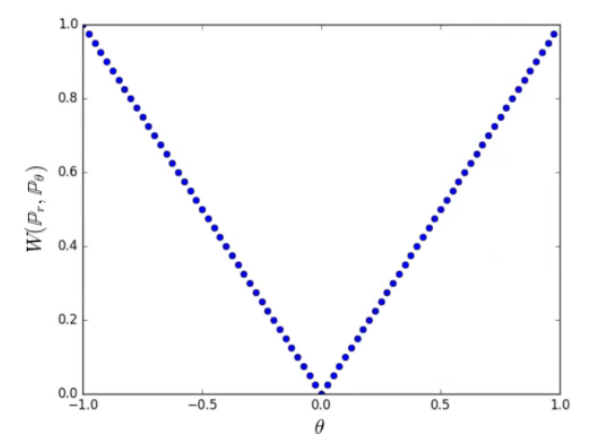


WHATS THE ISSUE?

Arjovsky et al 2017

Standard GAN objective is equivalent to minimizing the ratio of the two densities: $\dfrac{d\mathbb{P}}{dG(\mathbb{Q})}$





WGAN

Arjovsky et al 2017

$$W(\mathbb{P}, G(\mathbb{Q})) = \sup_{||f||_L \leq 1} \left(\mathbb{E}_{\mathbb{P}}[f(x)] - \mathbb{E}_{G(\mathbb{Q})}[f(x)] \right)$$

- The suggestion of the paper is to clip the weights in some range [-c, c].
- Metz et al (2016) suggest unrolling GAN – but each step increases computational cost.

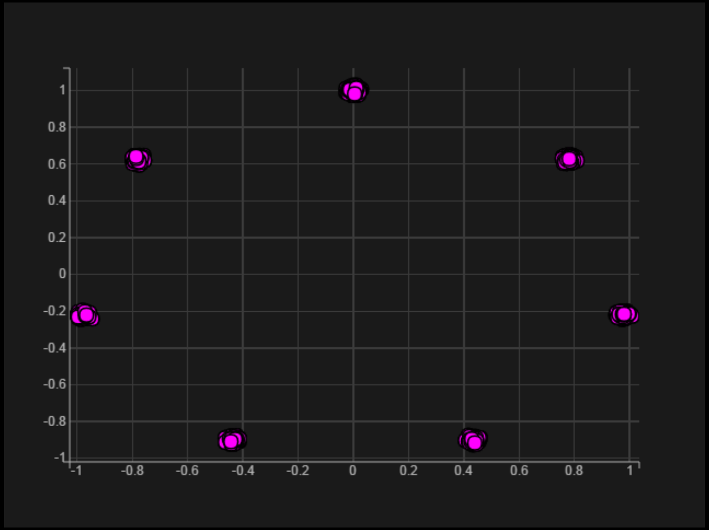
PENALIZE GRADIENTS

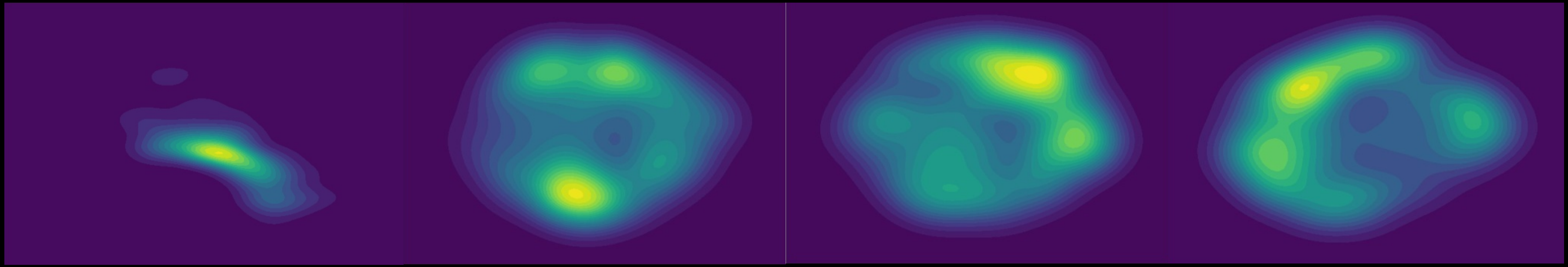
$$V = \left(\mathbb{E}_{\mathbb{P}}[\log(f)] - \mathbb{E}_{\mathbb{Q}}[\log(1-f)]\right) - \lambda[\mathbb{E}_{\mathbb{P}}\left[\frac{1}{\left(f(x)\right)^2}\left|\left|\nabla f(x)\right|\right|^2\right] + \mathbb{E}_{\mathbb{Q}}\left[\frac{1}{\left(1-f(x)\right)^2}\left|\left|\nabla(1-f(x))\right|\right|^2\right]$$

- Penalizing the derivative can be interpreted as a kind of weak Lipschitz constraint
- Better numerical properties than clipping, better computational cost to unrolling

GAUSSIAN MIXTURE

Metz et al 2016 propose a mixture of Gaussians with





GAUSSIAN MIXTURE