Data Science Lab

Learning Latent Space Representations

Chloé Court
Rayane Dakhlaoui
Jules Roques

f-GAN: Generative Training via f-Divergence Minimization

- Generalize GANs by minimizing **f-divergences** between real and generated distributions.
  - More stable training and theoretical flexibility.

For a given f-divergence:

$$D_f(P \,||\, Q) = \int_{\chi} q(x) f\left(\frac{p(x)}{q(x)}\right) dx$$

With its convex conjugate, the f-GAN objective reduces to solving the following:

$$\min_G \max_D \mathbb{E}_{x \sim P_{\text{data}}} \left[f^*(D(x))\right] - \mathbb{E}_{x \sim P_G} \left[D(x)\right]$$

- Key f-Divergences in f-GANs: Jensen-Shannon, Kullback-Leibler, and Reverse Kullback-Leibler

f-GAN: Generative Training via f-Divergence Minimization

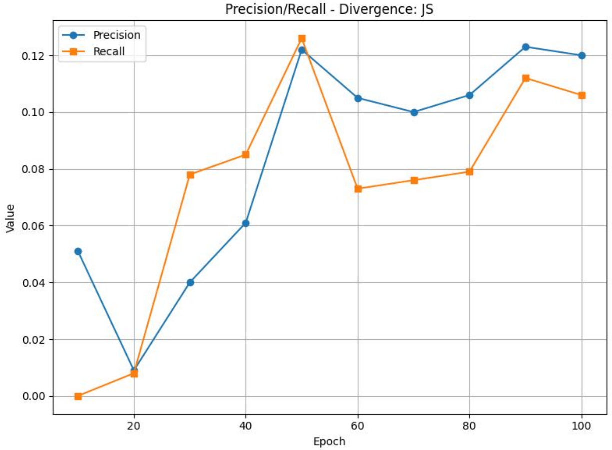
- ***Jensen-Shannon:*** Balanced, symmetric measure; commonly used in standard GANs

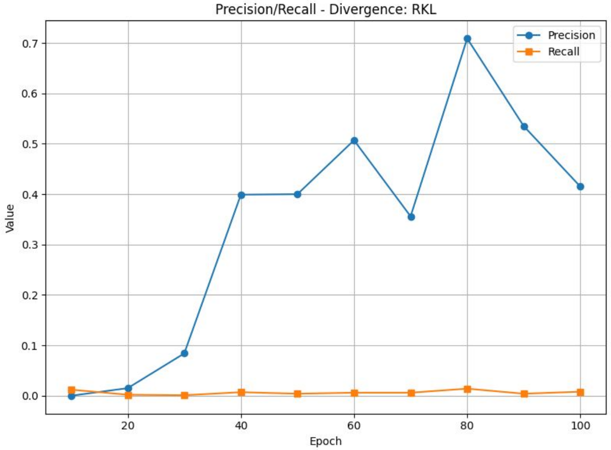
- ***Kullback-Leibler:*** Focuses on full coverage of real data distribution; prone to model collapse.

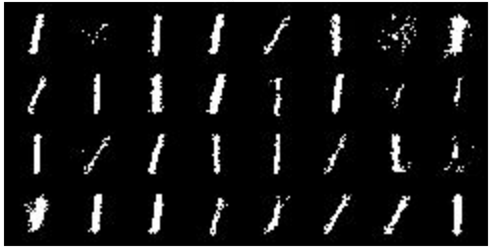
- ***Reverse KL:*** Emphasizes support matching; avoids model collapse but may produce blurry samples. Prone to mode collapse.

KL/RKL are numerically unstable due to exponentials in the loss.
=> **Apply stabilization techniques (gradient clipping, clamping discriminator outputs…)**

f-GAN: Generative Training via f-Divergence Minimization







Jensen-Shannon

Reverse Kullback-Leibler

Discriminator Rejection Sampling (Azadi et al., 2018)

Goal: Enhance GAN sample quality by filtering generated data (**post training)**

Key Principle: Accept samples based on discriminator confidence

$$D^*(x) = \frac{p_{data}(x)}{p_{data}(x) + p_g(x)} \quad \Rightarrow \quad \frac{D^*(x)}{1 - D^*(x)} = \frac{p_{data}(x)}{p_g(x)} \quad \Rightarrow \quad P_{accept}(x) = \frac{D(x)}{1 - D(x)} / M$$

Methods implemented: DRS and DRS with Soft-Truncation

Intuition:
- Keep samples with high discriminator confidence
- Remove unrealistic generations without retraining the model

Discriminator Rejection Sampling (Azadi et al., 2018)

Data: generator **G** and discriminator **D**
Result: Filtered samples from **G**
$D^* \leftarrow$ `KeepTraining(`D`);`
$\bar{M} \leftarrow$ `BurnIn(`$\boldsymbol{G}, D^*$`);`
$samples \leftarrow \emptyset$;
while $|samples| < N$ **do**
  $x \leftarrow$ `GetSample(`$\boldsymbol{G}$`);`
  $ratio \leftarrow e^{\tilde{D}^*(x)}$;
  $\bar{M} \leftarrow$ `Maximum(`$\bar{M}, ratio$`);`
  $p \leftarrow \sigma(\hat{F}(x, \bar{M}, \epsilon, \gamma))$;
  $\psi \leftarrow$ `RandomUniform(`$0,1$`);`
  if $\psi \leq p$ **then**
    `Append(`$X, samples$`);`
  end
end

$$\hat{F}(x) = \tilde{D}^*(x) - \tilde{D}^*_M - \log(1 - e^{\tilde{D}^*(x) - \tilde{D}^*_M - \varepsilon}) - \gamma$$

⇒ γ controls overall acceptance rate in

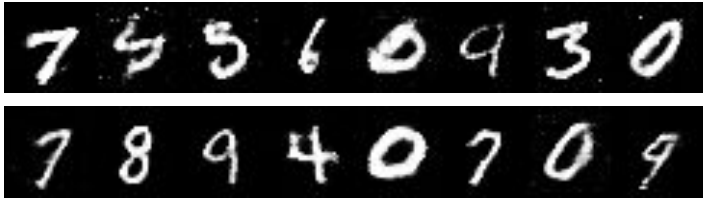
- **if too high:** almost no samples accepted **(model collapses)**
- **if too low:** almost all samples accepted **(no correction)**

⇒ Poor tuning of γ makes results **unstable** or **diverging**

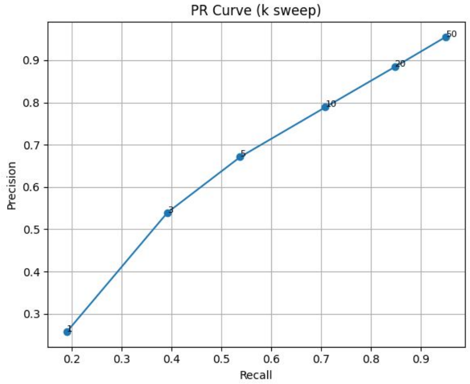
⇒ **Still calibrating the parameter to achieve stable convergence.**

Discriminator Rejection Sampling (Azadi et al., 2018)

Method	Div	FID	Precision	Recall
DRS only	JS	47.7	0.61	0.61
DRS /w Soft-Truncation	JS	59.7	0.67	0.53



(For 1k real and 10k fake)

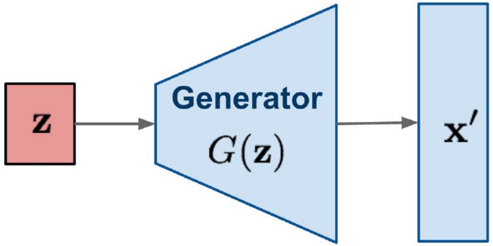


Observations:

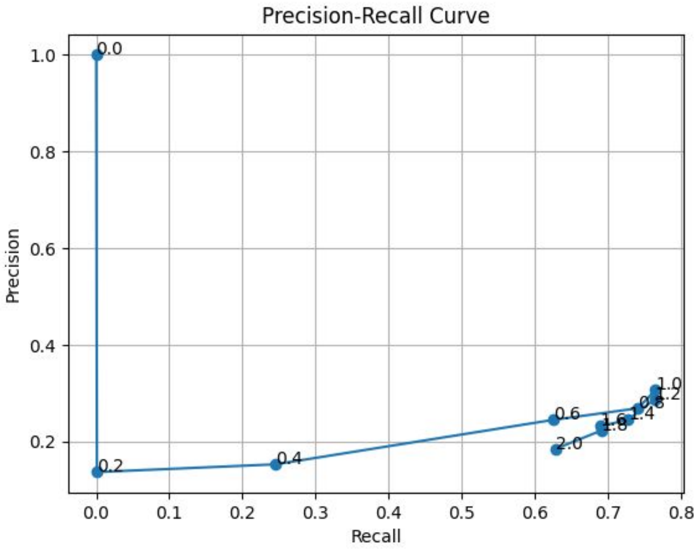
- **DRS /w Soft-Truncation** increases precision but reduces **recall**
- and **higher FID** suggests that this stronger filtering **discards too many valid sample**

Soft Truncation - Impact on Precision and Recall

$$z \sim \mathcal{N}(0, \sigma^2 I)$$

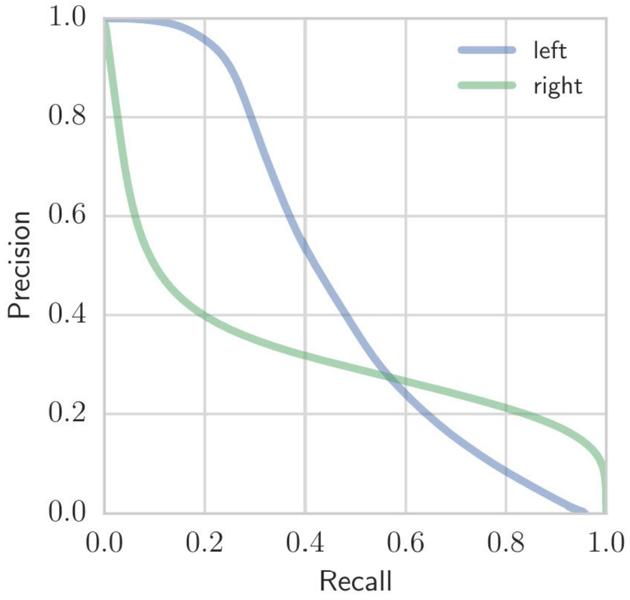


Ψ = 0 Ψ = 1 Ψ = 2



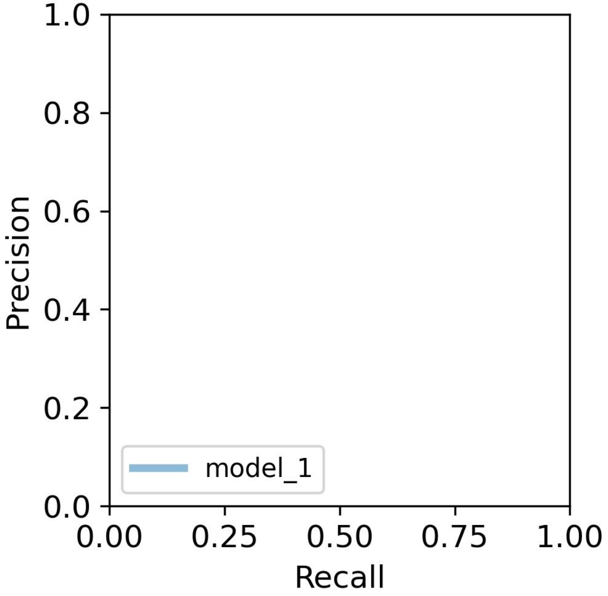
PR curve - JS GAN

We would like...



Assessing Generative Models via Precision and Recall, Sajjadi et al.

We have



Next steps

- Experiment with alternative divergences
- After DRS : Metropolis-Hasting ? Adaptive truncation ?
- Test soft truncation across all f-GANs (KL to boost recall, RKL to boost precision)
- Explore Gaussian Mixtures