

Generative Adversarial Nets and Variational Autoencoders at ICML 2018

2018-08-13

1. どんなもの？

- 著者が面白いと感じたトピックをシェアする.
- 生成モデルは, 古典的にデータ x とラベル y を持つ同時分布 $p(x, y)$ のモデルを記述する.
- しかし, ここでは生成モデルを新しいサンプル $x \sim X$ を提供するために, 近似データ分布 X からサンプリングするメカニズムを要するものを意味すると解釈する.
- 生成モデルの一般的なものは, GANsとVAEである
 - Ian Goodfellow's [GAN tutorial from NIPS 2016](#)
 - Irhum Shafkat's [Intuitively Understanding Variational Autoencoders](#)
- GANsは, 学習の不安定さの問題があり, mode collapseに苦しむ. しかし, 理想的な状態では, とても現実的な画像を生成できる.



Samples generated by a progressively grown GAN trained on the CelebA-HQ dataset illustrating how realistic the generated images from a GAN can look. Figure 5 in [Progressive Growing of GANs for Improved Quality, Stability, and Variation](#) by Karras et al.

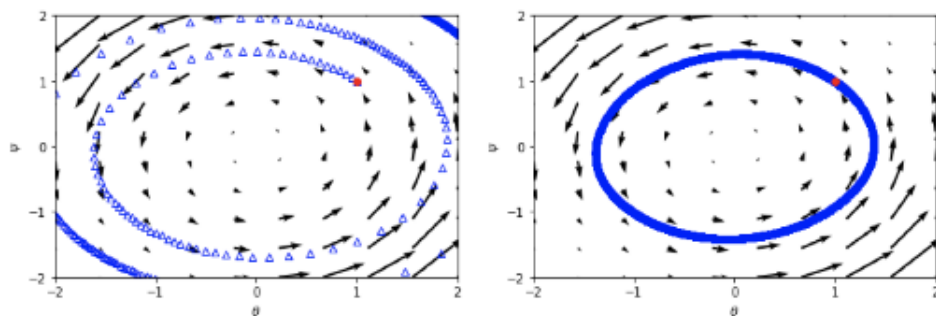
Figure 1

- 一方、VAEでは訓練することは簡単だが、尤度最大化によってよりぼけた画像を生成する傾向がある。
- 高い表現力を持つdecoderを使ったとき、VAEはときどき潜在ベクトルを無視する傾向があり、posterior collapseとして知られるものにつながる。

生成モデルの理解

GANs

[Which Training Methods for GANs do actually Converge?](#) - この論文は、2つだけのパラメータを持つ単純化されたGANsを組み立てることによって始まる。 - Generatorを構成するDirac分布の中心を制御するパラメータ - 線形分類器 Discriminatorの勾配を制御するパラメータ - 名づけてDirac-GAN - 同時更新（勾配を計算し、DiscriminatorとGeneratorの重みを同時に更新する）は、非安定な訓練になる傾向がある。



Trajectory of a standard Dirac-GAN trained with simultaneous updates (left) and alternating updates (right) with gradients illustrated with arrows in the θ -plane. Initial parameters are shown in red. Simultaneous updates lead to a gradient field with a positive divergence, forcing the trajectory away from the local optima in the middle. Figure 2 from [Which Training Methods for GANs do actually Converge?](#) by Mescheder, Geiger and Nowozin.

Figure 2

- GANの訓練軌跡，勾配のベクトル場を表している。
- 勾配ベクトル場は高い回転成分を持つようである。また，多くの手法は限りある円に突き当たる傾向がある。
- 定留点から解を強制的に移動させる可能性があるので，optimizerに momentumを使うべきでないという著者の推奨の理由である。
- よくまとまっている

[Toward Theoretical Understanding of Deep Learning tutorial](#)

Loading [MathJax]/jax/output/HTML-CSS/jax.js

VAEs

[Fixing a Broken ELBO](#)

- evidence lower bound(ELBO) loss項を調査する
- ELBOの導出の古典的な手法を取る代わりに，データ X と潜在コード Z 間の相互情報量 $I(X; Z)$ を研究することによって始まる．
- 以前から知られていた相互情報量の境界に基づいて，境界を述べる

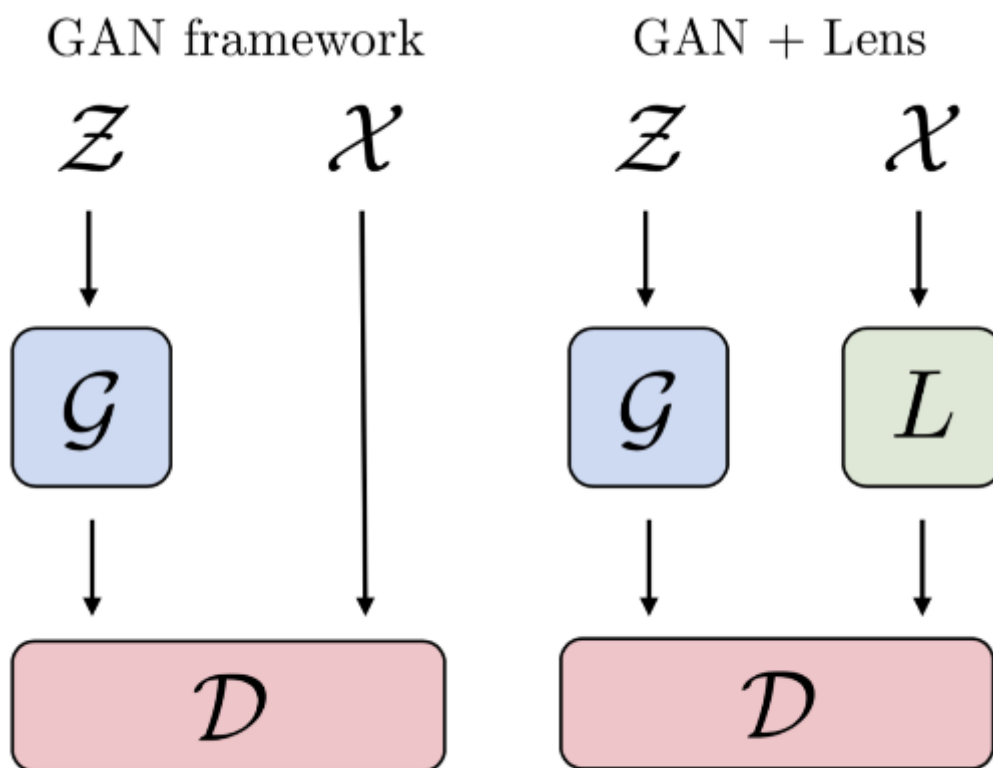
$$H - D \leq I(X; Z) \leq R$$

- H はデータのエントロピー， D はエンコーダを通した歪みの尺度， R は比率
- 潜在コードをエンコードするのに必要な追加情報の尺度は， Z 上の最適なコードと比較されるエンコーダからサンプルする．
- ゼロ分布 D は，潜在コードにおいて情報が失われないことを意味しており，サンプルは正確に再現できる．しかし，潜在コードが実際に役に立つものを獲得しているかは意味していない．一方，ゼロ比率 R は Z に X の独立を課す

Novel methods

1

[Tempered Adversarial Networks](#) - 重複していないサポートの問題に取り組んだ



Illustrates the added lens component in the GAN. Figure 1 in [Tempered Adversarial Networks](#) by Sajjadi et al.

[non-overlapping support](#)

2

[Mixed batches and symmetric discriminators for GAN training - Minibatch discrimination](#)の亜種を提案

Measuring and comparing quality of generated data

- 今までは、Inception ScoreやFIDを使って評価されてきた．一般的な Adversarial attacksによってもてあそばれる．また、いつも知覚的評価とよく相関があるわけではない

1

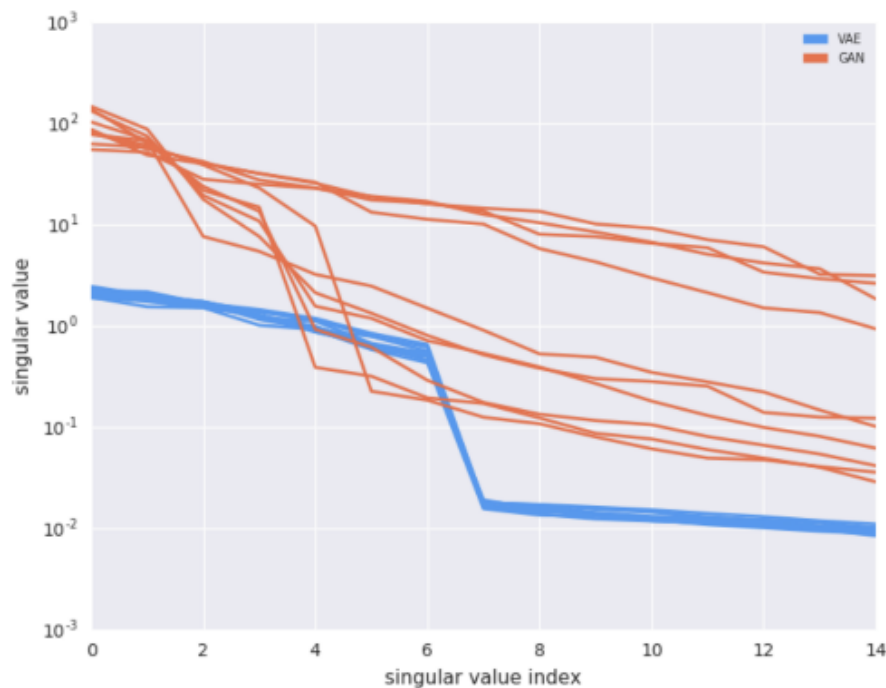
[Geometry Score: A Method For Comparing Generative Adversarial Networks](#)

- topological data analysis(TDA)を使って、生成データと実データのトポロジを比較する．

2

Is Generator Conditioning Causally Related to GAN Performance?

- Jacobian singular valuesを制御することの重要性を示した以前の研究を拡張した.
 - condition number(最大特異値と最小特異値間の比率)に注目している
 - 低いほどより安定
 - 理論が不安定さにリンクすることによって部分的に動機付けられている.
1. 実行のおよそ半分は、低いCNを獲得し、のこりの半分は高いCNを獲得している. 同じクラスタ内の値はかなり近づいている. したがって、良い、悪いの実行を特定することができる.
 2. GANsとVAE decodersの平均Jacobianの両方の特異値スペクトラムを研究した. VAEは特異値の小さな分散を持つ. VAEは低いCNを持つ傾向がある. これは、訓練の安定性を反映する定量的であるためおもしろい.



Here the singular value spectrum of the average Jacobian is shown for multiple training runs of GANs and VAEs. Figure 3 in Is Generator Conditioning Causally Related to GAN Performance? by Odena et al.

Figure 4

- CNを減らすために、特異値のレンジを制御するJacobian clampingという手法を提案している.

結論

2. 先行研究と比べてどこがすごいの？

3. 技術や手法の“キモ”はどこにある？

4. どうやって有効だと検証した？

5. 議論はあるか？

6. 次に読むべき論文はあるか？

論文情報・リンク

- [Generative Adversarial Nets and Variational Autoencoders at ICML 2018](#)