

敵対的学習におけるハイパーパラメータの自動制御

神谷 俊樹[†] 坂上 文彦[†] 佐藤 淳[†]名古屋工業大学大学院[†]

1. はじめに

本稿では学習バランスを取ることが難しい敵対的学習[1]を安定化させるために、ハイパーパラメータを自動的に制御する手法を提案する。GANとも呼ばれる敵対的学習は一般に Generator と Discriminator が敵対的に学習する。そのため、これらのバランスを保つためにハイパーパラメータを適切に設定する必要がある。しかし、GAN のハイパーパラメータの調整は従来のネットワークに比べて遥かに難しいことが知られている。そこで本稿では GAN の学習において、Generator と Discriminator の現在までの学習状態から将来の学習状態を予測し、学習が最も成功するようハイパーパラメータを自動的に適切に制御する手法を提案する。提案手法を複数の GAN を用いて評価し提案手法の有効性を示す。

2. 学習成功度

学習率を制御して訓練を改善するためには、訓練がどれほど進展しているかを表す尺度が必要となる。そこで、GAN の訓練の進行度を表す指標を導入し、これを学習成功度 S と呼ぶこととする。

画像に対して条件ラベルが与えられる conditional GAN[2]では、生成された画像がラベルに符合しているほど訓練が進んでいると言える。そこで、画像評価ネットワークとして訓練済みの識別器を使用し、この訓練済み識別器による生成画像の識別精度により conditional GAN における学習成功度 S_c を定義する。

$$S_c = \frac{1}{N_u} \sum_{i=1}^{N_u} C_i^*(u_i) \in [0,1]$$

ここで、 N_u は生成された画像の数であり、 u_i と l_i はそれぞれ i 番目の画像とそのラベルを表す。また、 $C_k^*(u)$ は画像 u がクラス k に分類される確率を表す。学習成功度 S_c は値が大きいほど訓練が進んでいることを表す。

画像に対して条件ラベルが与えられない unconditional GAN ではデータセットと生成画像が類似しているほど訓練が進んでいると言える。したがって、画像評価ネットワークとして Inception Network を使用し、データセットの画像と Generator が生成した画像の間の Fréchet Inception Distance [3] を学習成功度と定義する。画像データセットを D 、生成された画像群を U としたとき、unconditional GAN における学習成功度 S_u は次式のように表せる。

$$S_u = FID(D, U) \in [0, \infty)$$

ここで、 FID は Fréchet Inception Distance を表す。また、 S_u は値が小さいほど訓練が進んでいることを表す。

3. 学習成功度の予測

次に、将来のある時点における学習成功度 S を予測することを考える。これまでの訓練の学習率や訓練誤差、学習成功度は GAN の訓練の様子を表しており、今後の訓練で使用する学習率はあらかじめ定めておくことができる。そこで以下のように挙動データセット $\{x_b, x_f, S\}$ を定める。

x_b は過去 b エポックの間に Generator および Discriminator が訓練で用いた学習率、Generator および Discriminator の訓練誤差、Generator が生成した画像から計算した学習成功度より成る 5 チャンネルの時系列データである。また、 x_f は今後 f エポックの訓練で用いる Generator および Discriminator の学習率より成る 2 チャンネルの時系列データであり、 S は f エポック先の真の学習成功度である。

学習成功度 S の予測は、予測器 P と呼ぶネットワークを用いて行う。これは次式に示すように、 x_b および x_f から f エポック先の学習成功度を予測するものである。

$$\hat{S} = P(x_b, x_f)$$

この予測器 P の訓練は次式に示すように、予測器による学習成功度の予測値 \hat{S} と真の学習成功度 S の間の二乗誤差を最小化することで行う。

$$P^* = \arg \min_P \sum_i \|S_i - P(x_{b_i}, x_{f_i})\|^2$$

Deep Automatic Control of Learning Rates for GANs

[†]Toshiki Kamiya, [†]Fumihiko Sakae, [†]Jun Sato[†]Nagoya Institute of Technology

4. 学習率の制御

GAN の訓練中のある時点において、過去 b エポックの学習率、誤差、学習成功率より成るベクトル \mathbf{x}_b が得られているとする。このとき、以降の訓練において GAN の学習成功率が最も高くなるような現時点での Generator と Discriminator の学習率の組 (g^*, d^*) を予測器 P により推定する。このために、学習率の組 (g, d) より成る 2 次元空間 T を考える。すると求めたい学習率はこの 2 次元空間中の一点と考えることができる。ある学習率を採用した場合の GAN の将来の学習成功率は、 \mathbf{x}_b と (g, d) を時間軸方向に並べた \mathbf{x}_f の組から予測器 P を用いて次式のように予測できる。

$$\hat{S} = P(\mathbf{x}_b, \mathbf{x}_f)$$

よって予測した学習成功率が最大となる学習率 x_f^* は次式に示す最大化問題を解くことで得られる。

$$\mathbf{x}_f^* = \arg \max_{\mathbf{x}_f \in T} P(\mathbf{x}_b, \mathbf{x}_f)$$

本研究ではこの最大化問題を、離散化された空間における最大値の選択問題として解く。まずは 2 次元空間 T を離散化する。次にこの離散化させた空間中の各点において学習成功率 \hat{S} を上記の手法を用いて計算する。 \hat{S} が最大となる学習率 (g, d) が最適な学習率 (g^*, d^*) である。

次に、予測器 P を用いて学習率を自動制御する方法について述べる。まずは学習率、計算された誤差、生成された画像による学習成功率を記録しながら GAN の訓練を b エポック行う。次に、記録した値を \mathbf{x}_b として最適な学習率 (g^*, d^*) を推定する。そして推定した学習率を用いて GAN の訓練を 1 エポック行い、 \mathbf{x}_b を更新する。以降同様に最適な学習率の推定と、推定した学習率を用いた訓練を交互に繰り返す。このように予測器 P を用いて学習率を適切に制御することによって、GAN を安定的に訓練することが可能となる。

5. 実験

提案した学習率制御法の有効性を確かめるため、学習率制御実験を行った。実験には MNIST と Fashion-MNIST の 2 種類の画像データセットと 4 種類の GAN を組み合わせた 8 種類の画像生成タスクを用いた。各試行では学習率の初期値をランダムに定め、GAN の訓練を 50 エポック行った。この試行を学習率の制御を行わない場合と提案法を用いて学習率を制御する場合のそれぞれにおいて 100 回ずつ行い、各試行の最終的な学習成功率の平均値を比較した。

実験の結果、学習率を制御せずに訓練した場合の学習成功率の平均が 0.50 であったのに対し、提案法を用いて学習率制御を行った場合の平均

は 0.82 であった。この結果より、提案法を用いることで、GAN をより安定的に訓練できることがわかる。また、訓練の一例として図 1 に示す画像が生成され、このときの学習率と誤差の挙動は図 2、図 3 の通りであった。提案法では、Generator の誤差が大きい場合には Generator の学習率を大きくするなど、訓練の状況に応じて学習率が適切に制御されている様子を確認することができる。

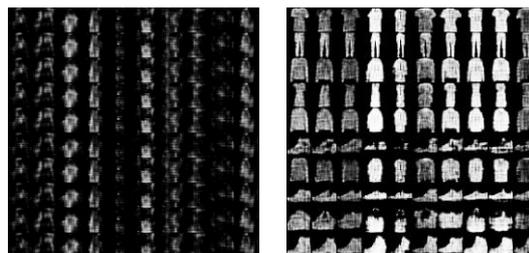


図 1. 同じ学習率から訓練した場合の生成画像 (左)制御なし (右)制御あり

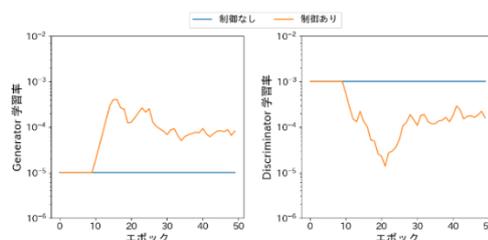


図 2. 学習率の挙動

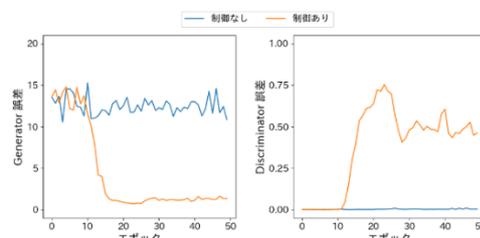


図 3. 誤差の挙動

6. まとめ

本研究では、将来の学習成功率を予測することで学習率を自動制御し GAN の訓練を安定化する手法を提案した。また実験により訓練の状況に応じて適応的に制御が行われることを確認した。

参考文献

- [1] Goodfellow ら, "Generative adversarial nets", Advances in neural information processing systems, pp. 2672-2680 (2014)
- [2] Mehdi Mirza ら, "Conditional generative adversarial nets" (2014)
- [3] Martin Heusel ら, "GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium" (2017)