Implicit Competitive Regularization in GANs

Anonymous Author(s)
Affiliation
Address
email

Abstract
To improve GANs, we need to understand why they can produce realistic samples. Presently, GANs are understood as the generator minimizing a divergence given by the optimal discriminator. We point out a fundamental flaw of this interpretation that precludes it from explaining why GANs work in practice. Instead, we argue that the performance of GANs is due to the implicit competitive regularization (ICR) arising from the simultaneous optimization of generator and discriminator. We show that opponent-aware modelling of generator and discriminator, as present in competitive gradient descent (CGD), can significantly strengthen ICR and thus stabilize GAN training without explicit regularization. In our experiments, we use an existing implementation of WGAN-GP and show that by training it with CGD we can improve the inception score (IS) on CIFAR10 for a wide range of scenarios, without any hyperparameter tuning. The highest IS is obtained by combining CGD with the WGAN-loss, without any explicit regularization.

Generative adversarial networks (GANs): [Goodfellow et al., 2014] are a class of generative models based on a competitive game between a generator that tries to generate realistic new data, and a discriminator that tries to distinguish generated from real data. In practice, both players are parameterized by neural networks that are trained simultaneously by a variant of stochastic gradient descent.

The minimax interpretation: Presently, the success of GANs is mostly attributed to properties of the divergence or metric obtained under an optimal discriminator. For instance, an optimal discriminator in the original GAN leads to a generator loss equal to the Jensen-Shannon divergence between real and generated distribution. Optimization over the generator is then seen as approximately minimizing this divergence. We refer to this point of view as the minimax interpretation. The minimax interpretation has led to the development of numerous GAN variants that aim to use divergences or metrics with better theoretical properties.

The GAN dilemma: However, every attempt to explain GAN performance with the minimax interpretation faces one of the two following problems:

1. **Without regularity constraints, the discriminator can always be perfect.** This is because it can selectively assign a high score to the finite amount of real data points while assigning a low score on the remaining support of the generator distribution, as illustrated in Figure 1. Therefore, the Jensen-Shannon divergence between a continuous and a discrete distribution always achieves its maximal value, a property that is shared by all divergences that do not impose regularity constraints on the discriminator. Thus, these divergences can not compare the quality of different generators.

2. **Imposing regularity constraints needs a measure of similarity of images.** Imposing regularity on the discriminator amounts to forcing it to map similar images to similar results. To do so, we would require a notion of similarity between images that is congruent with human perception. This is a longstanding unsolved problem in computer vision. Commonly used gradient penalties use the Euclidean norm which is known to poorly capture visual similarity, as illustrated in Figure 2.
We believe that the different divergences underlying the various GAN formulations have little to do with their ability to produce realistic images. This is supported by the large scale studies of Lucic et al. [2017] that did not find systematic differences in the performance of GANs associated with different divergence measures. However, an understanding of GAN performance is crucial in order to improve training stability and reduce the amount of hyperparameter tuning required in their deployment.

**A way out?:** Due to the GAN-dilemma, every attempt at explaining the performance of GANs needs to go beyond the minimax interpretation and consider the dynamics of the training process. In this work, we argue that an implicit regularization due to the simultaneous training of generator and discriminator allows GANs to use the inductive biases of neural networks for the generation of realistic images.

**Implicit competitive regularization:** We define *implicit competitive regularization* (ICR) as the introduction of additional stable points or regions due to the simultaneous training of generator and discriminator that do not exist when only training the generator (or discriminator) with gradient descent while keeping the discriminator (or generator) fixed. It has been previously observed that performing simultaneous gradient descent (SimGD) on both players leads to stable points that are not present when performing gradient descent with respect to either player, while keeping the other player fixed [Mazumdar and Ratliff, 2018]. These stable points are not local Nash equilibria, meaning that they are not locally optimal for both players. This phenomenon is commonly seen as a shortcoming of SimGD and modifications that promote convergence only to local Nash equilibria which have been proposed by, for instance, [Balduzzi et al., 2018; Mazumdar et al., 2019]. In contrast to this view we believe that ICR is crucial to overcoming the GAN-dilemma and hence to explaining GAN performance in practice by allowing the inductive biases of the discriminator network to inform the generative model.

**Summary of Contributions**

In this work, we point out that a fundamental dilemma prevents the common minimax interpretation of GANs from explaining their successes. We then show that implicit competitive regularization (ICR), which so far was believed to be a flaw of SimGD, is key to overcoming this dilemma. Based on simple examples and numerical experiments on real GANs we illustrate how it allows to use the inductive biases of neural networks for generative modelling, resulting in the spectacular performance of GANs.

We then use this understanding to improve GAN performance in practice. Interpreting ICR from a game-theoretic perspective, we reason that strategic behavior and opponent-awareness of generator and discriminator during the training procedure can strengthen ICR. These elements are present in competitive gradient descent (CGD) [Schaefer and Anandkumar, 2019] which is based on the two players solving for a local Nash-equilibrium at each step of training. Accordingly, we observe that CGD greatly strengthens the effects of ICR. In comprehensive experiments on CIFAR 10, competitive gradient descent stabilizes previously unstable GAN formulations and achieves higher inception score compared to a wide range of explicit regularizers, using both WGAN loss and the original saturating GAN loss of Goodfellow et al. [2014]. In particular, taking an existing WGAN-GP implementation, dropping the gradient penalty, and training with CGD leads to the highest inception score in our experiments. We interpret this as additional evidence that ICR, as opposed to explicit regularization, is the key mechanism behind GAN performance.

---

2The pairs of images are ordered from left to right, in increasing order of distance. The first pair is identical, while the third pair differs by a tiny warping.
References


