Characterizing Equilibria in Stackelberg Games

Tanner Fiez, Benjamin Chasnov, Lillian Ratliff
{fiezt, bchasnov, ratliffl}@uw.edu
Electrical & Computer Engineering
University of Washington

Abstract

This paper investigates the convergence of learning dynamics in Stackelberg games on continuous action spaces, a class of games distinguished by the hierarchical order of play between agents. We establish connections between the Nash and Stackelberg equilibrium concepts and characterize conditions under which attracting critical points of simultaneous gradient descent are Stackelberg equilibria in zero-sum games. Moreover, we show that the only stable critical points of the Stackelberg gradient dynamics are Stackelberg equilibria in zero-sum games. Using this insight, we develop two-timescale learning dynamics for which each stable critical point is guaranteed to be a Stackelberg equilibrium in zero-sum games and the dynamics converge to the set of stable attractors in general-sum games.

1 Introduction

In a Stackelberg game, there is a leader and a follower that interact in a hierarchical structure. The sequential order of play is such that the leader is endowed with the power to select an action using the knowledge that the follower will then play a best-response. We formulate and study a novel set of gradient-based learning rules in continuous, general-sum Stackelberg games. The dynamics analyzed in this work reflect the underlying game structure and characterize the expected outcomes of hierarchical games. The study of learning dynamics and equilibria in Stackelberg games we provide has implications for both multi-agent learning and adversarial learning applications.

We define and analyze the differential Stackelberg equilibrium solution concept, which is a local notion of the Stackelberg equilibrium. An analogous local minimax equilibrium concept was developed concurrently with this work, but strictly for zero-sum games [6]. Importantly, the equilibrium notion we present generalizes the local minimax equilibrium concept to general-sum games. The solution concept we study is a natural extension of the differential Nash equilibrium concept [10], which is a local notion of the Nash equilibrium in continuous games. The insights we present in this paper on the equilibria landscape of continuous games come as a direct consequence of focusing our attention on the Stackelberg equilibrium concept instead of the Nash equilibrium concept.

Contributions. We establish a number of connections between Nash and Stackelberg equilibria and characterize the conditions under which attracting critical points of simultaneous gradient descent are Stackelberg equilibria in zero-sum games. To summarize, we show stable differential Nash equilibria are differential Stackelberg equilibria in zero-sum games. Concurrently, Jin et al. [6] showed local Nash equilibria are local minimax equilibria. We also reveal that there exist non-Nash attractors of simultaneous gradient descent that are Stackelberg equilibria. Moreover, we give necessary and sufficient conditions under which the simultaneous gradient play dynamics can avoid Nash equilibria and converge to Stackelberg equilibria in zero-sum games. To demonstrate the relevancy to deep learning applications, we specialize the conditions to GANs satisfying the realizable assumption [8], which presumes the generator is able to create the underlying data distribution.

Abstract for: Smooth Games Optimization and Machine Learning Workshop (NeurIPS 2019), Vancouver, Canada.

Finally, we analyze the convergence behavior of gradient-based learning rules reflecting the underlying Stackelberg game structure. We show that the only stable critical points of the Stackelberg gradient dynamics are Stackelberg equilibria in zero-sum games. This is in contrast to the simultaneous gradient play dynamics, which can be attracted to non-Nash critical points in zero-sum games. We consider the follower to employ a gradient-play update rule and propose a two-timescale algorithm to learn Stackelberg equilibria. We show asymptotic convergence of the dynamics to Stackelberg equilibria in zero-sum games given an initialization in the region of attraction of a stable critical point and to the set of stable attractors in general-sum games.

2 Preliminaries

Consider a game between two agents where one agent is deemed the *leader* and the other the *follower*. The leader has cost $f_1 : X \to \mathbb{R}$ and the follower has cost $f_2 : X \to \mathbb{R}$, where $X = X_1 \times X_2$ with the action space of the leader being X_1 and the action space of the follower being X_2. We assume throughout that each f_i is sufficiently smooth, meaning $f_i \in C^q(X, \mathbb{R})$ for some $q \geq 2$ and for each agent $i \in \mathcal{I} = \{1, 2\}$. The designation of 'leader' and 'follower' indicates the order of play between the two agents, meaning the leader plays first and the follower second. The leader and the follower need not be cooperative. Such a game is known as a *Stackelberg game*.

The leader aims to solve the optimization problem given by

$$\min_{x_1 \in X_1} \{f_1(x_1, x_2) |\, x_2 \in \arg\min_{y \in X_2} f_2(x_1, y)\}$$

and the follower aims to solve the optimization problem $\min_{x_2 \in X_2} f_2(x_1, x_2)$. The learning algorithms we study are such that the agents follow myopic update rules which take steps in the direction of steepest descent with respect to these optimization problems.

Before formalizing the learning rules, we discuss the equilibrium concept studied for simultaneous and hierarchical play games. The typical equilibrium notion in continuous games is the pure strategy Nash equilibrium in simultaneous play games and the Stackelberg equilibrium in hierarchical play games. Each notion of equilibria can be defined as the intersection points of the reaction curves of the players [1]. We characterize local Nash and Stackelberg equilibrium using sufficient conditions.

To define the equilibrium concepts, we need some notation. We denote $D_i f_i$ as the derivative of f_i with respect to x_i, $D_{ij} f_i$ as the partial derivative of $D_i f_i$ with respect to x_j, and $D(\cdot)$ as the total derivative[1]. Denote by $\omega(x) = (D_1 f_1(x), D_2 f_2(x))$ the vector of individual gradients for simultaneous play and $\omega_{\mathcal{S}}(x) = (D f_1(x), D_2 f_2(x))$ as the equivalent for hierarchical play where x_2 is implicitly a function of x_1, which captures the fact that the leader operates under the assumption that the follower will play a best response to its choice of x_1[2]. The following are local equilibrium concepts defined using sufficient conditions.

Definition 1 (Differential Nash Equilibrium [10])**.** *The joint strategy $x^* \in X$ is a differential Nash equilibrium if $\omega(x^*) = 0$ and $D_i^2 f_i(x^*) > 0$ for each $i \in \mathcal{I}$.*

Definition 2 (Differential Stackelberg Equilibrium)**.** *The pair $(x_1^*, x_2^*) \in X$ with $x_2^* = r(x_1^*)$, where r is implicitly defined by $D_2 f_2(x_1^*, x_2^*) = 0$, is a differential Stackelberg equilibrium for the game (f_1, f_2) with player 1 as the leader if $D f_1(x_1^*, r(x_1^*)) = 0$ and $D^2 f_1(x_1^*, r(x_1^*))$ is positive definite.*

Remark 1. *In zero-sum games, the differential Stackelberg equilibrium notion is equivalent to a local minimax equilibrium. This is a known concept in optimization (see, e.g., [1, 4, 3]), and it has recently been presented in the learning literature (see, e.g., [6]). For general-sum games, such a characterization derives from sufficient conditions for local optimality respecting the game structure.*

We utilize Definition 2 to formulate the Stackelberg gradient dynamics we study; indeed, the combined learning dynamics given an appropriate learning rate $\gamma > 0$ are defined by

$$x_{k+1} = x_k - \gamma \omega_{\mathcal{S}}(x_k). \tag{1}$$

The learning dynamics approximate the continuous time dynamical system $\dot{x} = -\omega_{\mathcal{S}}(x)$ as the learning rate $\gamma \to 0$. We study the limit points of this system to gain novel insights into the equilibrium landscape of continuous games.

[1] For example, given a function $f(x, r(x))$, $Df = D_1 f + D_2 f \partial r / \partial x$.

[2] Under sufficient regularity assumptions, the total derivative for the hierarchical gradient dynamics can be computed using the implicit function theorem to be $D f_1 = D_1 f_1 - D_2 f_1 (D_2^2 f_2)^{-1} D_{21} f_2$.

3 Linking Nash and Stackelberg Equilibria in Zero-Sum Games

In this section we show the distinct connections between the limit points of simultaneous gradient descent and the Stackelberg gradient dynamics in zero-sum games, which demonstrate Nash equilibria and Stackelberg equilibria are closely intertwined in this class of games.

Proposition 1. *For any continuous zero-sum game $(f, -f)$, stable critical points of $\dot{x} = -\omega_{\mathcal{S}}(x)$ are differential Stackelberg equilibria.*

The result follows from examination of the structure of the Jacobian of $\omega_{\mathcal{S}}$, which is block lower triangular with player 1 and 2 as the leader and follower, respectively. The preceding result implies the only stable critical points of the dynamics in (1) are Stackelberg equilibria and thus, unlike simultaneous gradient descent, will not converge to spurious stable critical points of the dynamics.

Proposition 2. *For any continuous zero-sum game $(f, -f)$, stable differential Nash equilibria are differential Stackelberg equilibria.*

This result is obtained by examining the first Schur complement of the Jacobian of the simultaneously play dynamics, and noting that it is exactly the second order sufficient conditions for the leader in the Stackelberg game. In zero-sum games, the fact that Nash equilibria are a subset of Stackelberg equilibria (or minimax equilibria) for finite games is well-known [1]. We show the result for the differential Stackelberg equilibria in continuous games. Concurrently, Jin et al. [6] showed local Nash equilibria are local minimax solutions in continuous zero-sum games. The result indicates algorithms seeking Nash are seeking Stackelberg simultaneously in zero-sum games.

The simultaneous gradient play dynamics can avoid Nash equilibria and converge to locally asymptotically stable critical points of the dynamics. Previously, such points have been viewed as lacking game-theoretic significance. However, the following results give conditions under which a non-Nash stable critical point of the dynamics $\dot{x} = -\omega(x)$ is an attractor of the Stackelberg dynamics $\dot{x} = -\omega_{\mathcal{S}}(x)$. From Proposition 1, it follows that such attractors are Stackelberg equilibria.

For a non-Nash attractor $x^* = (x_1, x_2) \in \mathbb{R}^{m+n}$ of a zero-sum game $(f, -f)$, let $\mathrm{spec}(D_1^2 f(x^*)) = \{\mu_j,\ j \in \{1, \dots, m\}\}$ where $\mu_1 \leq \cdots \leq \mu_r < 0 \leq \mu_{r+1} \leq \cdots \leq \mu_m$, and let $\mathrm{spec}(-D_2^2 f(x^*)) = \{\lambda_i,\ i \in \{1, \dots, n\}\}$ where $\lambda_1 \geq \cdots \geq \lambda_n > 0$, and define $p = \dim(\ker(D_1^2 f(x^*)))$.

Proposition 3 (Necessary conditions)**.** *Consider a non-Nash attractor x^* of the individual gradient dynamics $\dot{x} = -\omega(x)$ such that $-D_2^2 f(x^*) > 0$. Given $\kappa > 0$ such that $\|D_{21} f(x^*)\| \leq \kappa$, if x^* is an attractor of $\dot{x} = -\omega_{\mathcal{S}}(x)$, then $r \leq n$ and $\kappa^2 \lambda_i + \mu_i > 0$ for all $i \in \{1, \dots, r-p\}$.*

Proposition 4 (Sufficient conditions)**.** *Let x^* be a non-Nash attractor of the individual gradient dynamics $\dot{x} = -\omega(x)$ such that $D_1^2 f(x^*)$ and $-D_2^2 f(x^*)$ are Hermitian, and $-D_2^2 f(x^*) > 0$. Suppose that there exists a diagonal matrix (not necessarily positive) $\Sigma \in \mathbb{C}^{m \times n}$ with non-zero entries such that $D_{12} f(x^*) = W_1 \Sigma W_2^*$ where W_1 are the orthonormal eigenvectors of $D_1^2 f(x^*)$ and W_2 are orthonormal eigenvectors of $-D_2^2 f(x^*)$. Given $\kappa > 0$ such that $\|D_{21} f(x^*)\| \leq \kappa$, if $r \leq n$ and $\kappa^2 \lambda_i + \mu_i > 0$ for each $i \in \{1, \dots, r-p\}$, then x^* is an attractor of $\dot{x} = -\omega_{\mathcal{S}}(x)$.*

The previous results follow from examining the unique structure of the Jacobian of the simultaneous play dynamics, which resembles a J-frame on Krein spaces; with respect to the Krein space [11], the Jacobian is a self-adjoint operator which is afforded similar eigenstructure properties as self-adjoint operators on Hilbert spaces. The conditions imply some of the non-Nash attracting critical points of $\dot{x} = -\omega(x)$ are in fact Stackelberg equilibria. Some recent results show that several approaches to training GANs are not converging to stable Nash equilibria, but rather to stable non-Nash critical points of the dynamics [2]. In future work, we plan to explore whether or not such attractors satisfy the conditions we propose.

A common assumption in some of the GAN literature is that the discriminator network is zero in a neighborhood of an equilibrium configuration (see, e.g., [8, 9, 7]). This assumption limits the theory to the 'realizable' case, where the generator is capable of creating the underlying distribution. The work by [8] provides relaxed assumptions for the non-realizable case. In both cases, the Jacobian for the dynamics $\dot{x} = -\omega(x)$ is such that $D_1^2 f(x^*) = 0$. The following results specialize the conditions from Propositions 3 and 4 to zero-sum games satisfying the realizable assumption.

Proposition 5. *Consider a GAN satisfying the realizable assumption. Then, an attracting critical point x^* for the simultaneous gradient dynamics $\dot{x} = -\omega(x)$ at which $-D_2^2 f$ is positive semi-*

definite satisfies necessary conditions for a local Stackelberg equilibrium, and it will be a marginally stable point of the Stackelberg dynamics $\dot{x} = -\omega_{\mathcal{S}}(x)$.

This result follows from examining the Schur complement of the Jacobian of the simultaneous play dynamics under the realizable assumption, which guarantees the upper block component is zero.

Proposition 6. *Consider a GAN satisfying the realizable assumption and an attracting critical point x^* for the simultaneous gradient dynamics $\dot{x} = -\omega(x)$ at which $-D_2^2 f$ is positive definite. Suppose there exists a diagonal matrix Σ with non-zero entries such that $D_{12} f(x^*) = \Sigma W$ where W are the orthonormal eigenvectors of $-D_2^2 f(x^*)$. Then, x^* is an attractor of $\dot{x} = -\omega_{\mathcal{S}}(x)$.*

This result follows from the general sufficient conditions in Proposition 4.

4 Two-Timescale Learning Dynamics

Motivated by practical learning applications we focus on developing gradient-based algorithms with convergence guarantees under stochastic updates. To do so, we focus on the situation where the leader operates under the assumption that the follower is playing (locally) optimally at each round so that the belief is $D_2 f_2(x_{1,k}, x_{2,k}) = 0$, but the follower is actually performing the update $x_{2,k+1} = x_{2,k} + g_2(x_{1,k}, x_{2,k})$ where $g_2 \equiv -\gamma_{2,k}\mathbb{E}[D_2 f_2]$. The dynamics in this formulation are given by

$$\begin{aligned} x_{1,k+1} &= x_{1,k} - \gamma_{1,k}(Df_1(x_k) + w_{1,k+1}) \\ x_{2,k+1} &= x_{2,k} - \gamma_{2,k}(D_2 f_2(x_k) + w_{2,k+1}), \end{aligned} \tag{2}$$

where $Df_1(x) = D_1 f_1(x) + D_2 f_1(x)(D_2^2 f_2)^{-1}(x) D_{21} f_2(x)$. Now, suppose that $\gamma_{1,k} \to 0$ faster than $\gamma_{2,k}$ so that in the limit as $\tau \to 0$, the dynamics in (2) approximates the singularly perturbed system defined by

$$\begin{aligned} \dot{x}_1(t) &= -Df_1(x_1(t), x_2(t)) \\ \dot{x}_2(t) &= -\tau^{-1} D_2 f_2(x_1(t), x_2(t)). \end{aligned} \tag{3}$$

The learning rates can be seen as stepsizes in a discretization scheme for solving the dynamics. The condition that $\gamma_{1,k} = o(\gamma_{2,k})$ induces a *timescale separation* in which x_2 evolves on a faster timescale than x_1. That is, the fast transient player is the *follower* and the slow component is the *leader* since $\lim_{k\to\infty} \gamma_{1,k}/\gamma_{2,k} = 0$ implies that from the perspective of the follower, x_1 appears quasi-static and from the perspective of the leader, x_2 appears to have equilibriated, meaning $D_2 f_2(x_1, x_2) = 0$ given x_1. From this point of view, the learning dynamics in (2) approximate the follower playing an exact best response. Moreover, attracting critical points of the dynamics are such that the leader is at a local optima for f_1, not just along its coordinate axis but in both coordinates (x_1, x_2) constrained to the manifold $r(x_1)$; this is to make a distinction between differential Nash equilibria in that players are at local optima aligned with their individual coordinate axes. While the convergence analysis is outside the scope of this paper, we can show that the dynamics in (2) reach Stackelberg equilibria in zero-sum games if initialized in the region of attraction of a critical point and the set of stable attractors in general-sum games. Formal theoretical results and proofs for both asymptotic and finite time convergence are given in [5]. The analysis combines techniques from dynamical systems theory with the theory of stochastic approximation. We leverage the limiting continuous time dynamical system in (3) to characterize concentration bounds for iterates or samples generated by (2) along with results in Section 3 to obtain guarantees.

5 Conclusion

In this paper, we present a number of connections between the Nash and Stackelberg equilibrium concepts in continuous games. As a consequence of focusing on the Stackelberg equilibrium concept, we obtain conditions characterizing when non-Nash attracting critical points of simultaneous gradient descent are Stackelberg equilibria in zero-sum games. This result is of significant practical interest since previously such spurious stable points were not thought to be game-theoretically meaningful. Moreover, we show that the only stable critical points of the Stackelberg gradient dynamics are Stackelberg equilibria in zero-sum games. The result shows the highly desirable characteristics of attracting critical points in hierarchical play games in contrast to simultaneous play games. Finally, we develop learning dynamics that converge to Stackelberg equilibria if initialized in the region of attraction of a critical point in zero-sum games and the set of stable attractors in general-sum games. A longer version of this paper including numerical results is presented in [5].

References

[1] Tamer Basar and Geert Jan Olsder. *Dynamic Noncooperative Game Theory*. Society for Industrial and Applied Mathematics, 2nd edition, 1998.

[2] H Berard, G Gidel, A Almahairi, P Vincent, and S Lacoste-Julien. A closer look at the optimization landscapes of generative adversarial networks. *arXiv preprint arxiv:1906.04848*, 2019.

[3] John M. Danskin. The theory of max-min, with applications. *SIAM Journal on Applied Mathematics*, 14(4):641–664, 1966.

[4] John M. Danskin. *The Theory of Max-Min and its Application to Weapons Allocation Problems*. Springer, 1967.

[5] Tanner Fiez, Benjamin Chasnov, and Lillian J Ratliff. Convergence of learning dynamics in stackelberg games. *arXiv preprint arXiv:1906.01217*, 2019.

[6] Chi Jin, Praneeth Netrapalli, and Michael I Jordan. Minmax optimization: Stable limit points of gradient descent ascent are locally optimal. *arXiv preprint arXiv:1902.00618*, 2019.

[7] Lars Mescheder, Andreas Geiger, and Sebastian Nowozin. Which training methods for gans do actually converge? *arXiv preprint arXiv:1801.04406*, 2018.

[8] Vaishnavh Nagarajan and J Zico Kolter. Gradient descent gan optimization is locally stable. In *Advances in Neural Information Processing Systems*, pages 5585–5595, 2017.

[9] Weili Nie and Ankit B. Patel. Towards a better understanding and regularization of gan training dynamics. *arXiv preprint arxiv:1806.09235*, 2019.

[10] L. J. Ratliff, S. A. Burden, and S. S. Sastry. On the Characterization of Local Nash Equilibria in Continuous Games. *IEEE Transactions on Automatic Control*, 61(8):2301–2307, 2016.

[11] C. Tretter. *Spectral Theory of Block Operator Matrices and Applications*. Imperial College Press, 2008.