339: Convergence Analysis of Gradient-Based Learning in Continuous Games

Introduction

We study learning and control in multi-agent systems in which agents have individual objectives and repeatedly interact in *simultaneous play*. We seek rigorous convergence guarantees on the limiting outcomes of such multi-agent interactions.

Definition (Continuous game)
A collection of costs $(f_1, \dots, f_n)$ on $X = X_1 \times \cdots \times X_n$ where $f_i : X \to R$ is agent i's cost and X_i is its action space.

Note: an individual's cost is coupled with other agents' actions. Agents repeated select actions through the process below.

Gradient-based learning (simultaneous)
Each agent updates their choice variable $x_i \in X_i$ by the
1) deterministic process

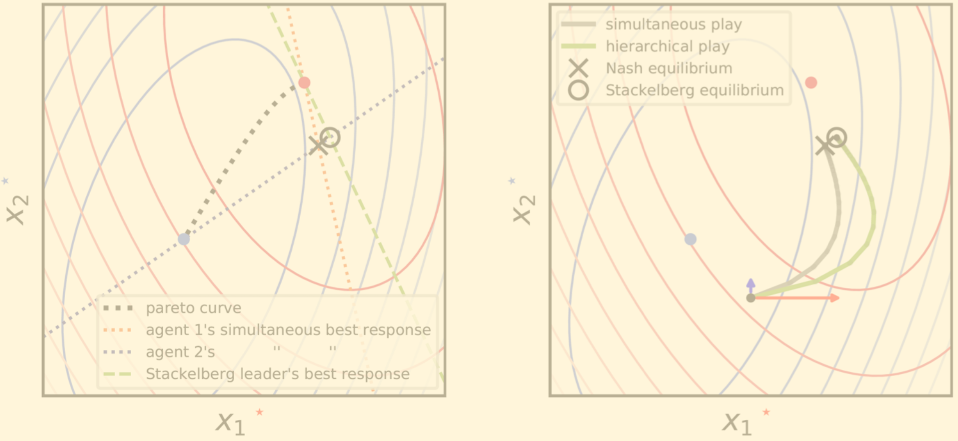
$$x_{i,k+1} = x_{i,k} - \gamma_i \, D_i f_i(x_{i,k}, x_{-i,k})$$

2) or stochastic process

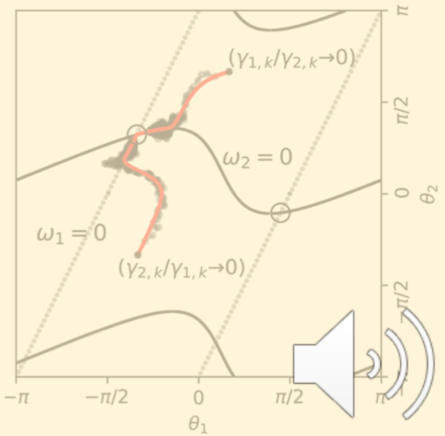
$$x_{i,k+1} = x_{i,k} - \gamma_i \, \widehat{D_i f_i}(x_{i,k}, x_{-i,k})$$

for all $i = 1, \dots, n$, where γ_i is the learning rate and $D_i f_i \equiv \frac{\partial f_i}{\partial x_i}$.

Deterministic example (two-player quadratic game)
Let $f_1(x_1, x_2) = x^T Q_1 x + q_1 x$ and $f_2(x_1, x_2) = x^T Q_2 x + q_2 x$, where joint action $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$. For games and learning rates that satisfy *Theorem 1,* we can guarantee convergence to a locally asymptotically stable Nash equilibrium.



Stochastic example (torus game)
We provide *concentration bounds* for agents that learn stochastically, with unbiased estimates of the gradient and *non-uniform* learning rates.



339: Convergence Analysis of Gradient-Based Learning in Continuous Games

Stability of critical points and Nash equilibria

Definition (Differential Nash equilibria)
For continuous game $(f_1, \dots, f_n)$, x^* is a *differential Nash equilibrium* (DNE) if $D_i f_i(x^*) = 0$ and $D_i^2 f_i(x^*) > 0, \forall i$.

Whether the learning dynamics

$$x_{i,k+1} = x_{i,k} - \gamma_i \, g_i(x_{i,k}, x_{-i,k}), \quad \forall i \in [n],$$

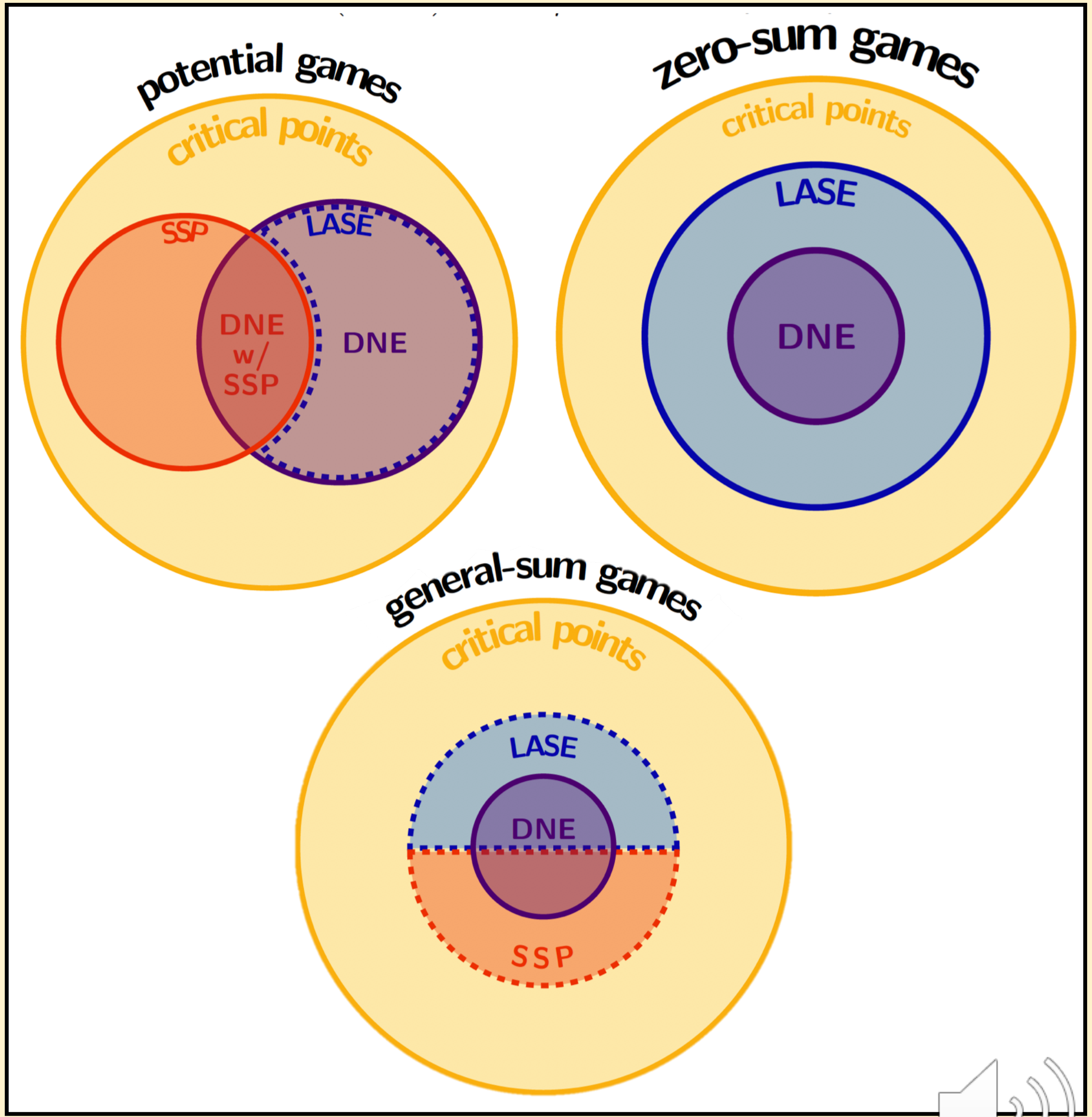
converges to a Nash equilibrium depends on the structure of the cost functions and the learning rate.

Proposition (asymptotic convergence to DNE)
Suppose agents use the gradient-based update with learning rates γ_i such that $\rho(I - \Gamma\, Dg) < 1$. Then, for x_0 initialized in the region of attraction of x^*, $x_k \to x^*$ exponentially.

We can choose learning rates that bound the convergence rates.

Theorem (non-asymptotic convergence guarantees)
Suppose g is Lipchitz and let $\gamma = \sqrt{\alpha}/\beta$ where α, β relate to the singular values of Dg, then if x_0 is initialized within radius r of x^*, then iteration x_k for $k \geq T$ achieves ε-Nash

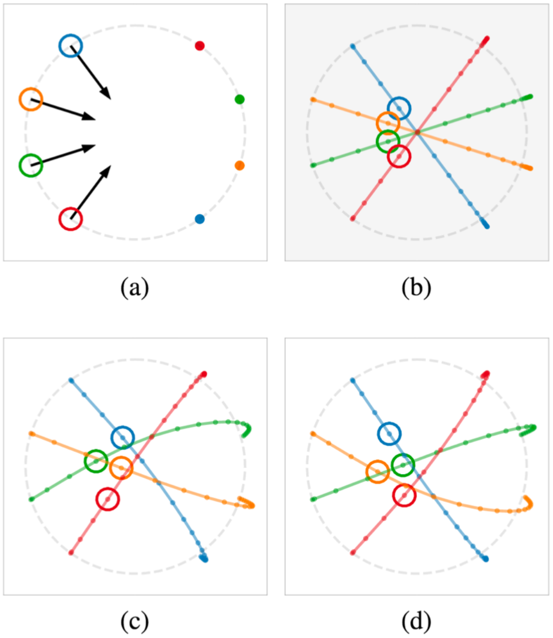
$$T = \lceil \beta/\alpha \log(r/\varepsilon) \rceil$$



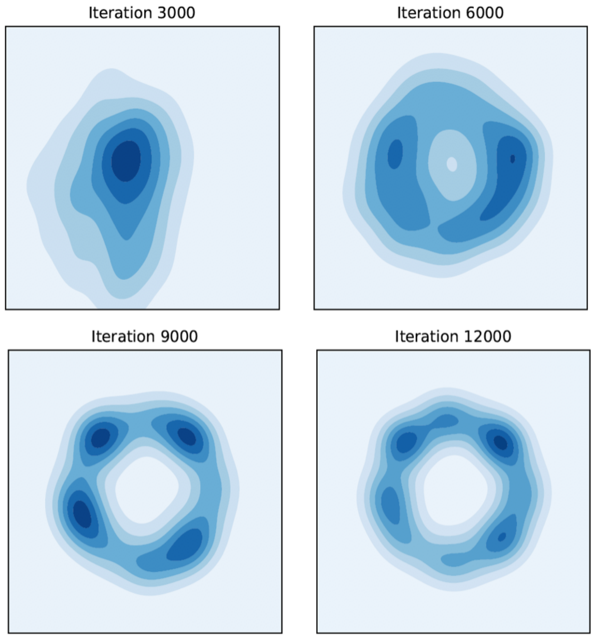
339: Convergence Analysis of Gradient-Based Learning in Continuous Games

Applications in controls and machine learning

Collision avoidance and control



Zero-sum game between generator and discriminator networks



Dynamic games with coupled interaction