Lecture 11 - Convex games II

Maryam Kamgarpour

1 Outline

To compute equilibria in games with continuous actions, our plan for the lecture is as follows:

- Banach fixed point theorem and its connection to algorithm design
- Characterizing Nash equilibria of convex games as fixed points of a projection operator
- Designing an algorithm, similar to gradient descent, using game pseudo-gradient.
- Prove convergence of the algorithm to a Nash equilibrium under suitable conditions on the game

2 Introduction

Last time, we discussed games with continuous action spaces. We developed sufficient conditions for the existence and uniqueness of their Nash equilibria. Furthermore, we characterized the Nash equilibria of convex games using the game pseudo-gradient. Our goal in this lecture is to develop an approach for computing these Nash equilibria.

Before presenting the approach, consider the following questions. It might be helpful to recall the specific games we looked at last time, namely, the Cournot and the Bertrand competitions.

1. Given a convex game $\Gamma(\mathcal{N}, K, \{J^i\}_{i\in\mathcal{N}})$ how can each player compute its Nash equilibrium strategy?
2. Suppose each player does not know the other players' constraint sets or cost functions. How does your answer above change?

3 Computing Nash equilibria in convex games

Our approach for computing the Nash equilibria will be to design an iterative algorithm, $x_{t+1} = f(x_t)$, so that its fixed point $x = f(x)$, is a Nash equilibrium. Here, x would be the stacked vector of decisions of the players. If we can ensure f is a contraction, then we are sure that the above iteration converges to the fixed point of f and hence, the Nash equilibrium. This is due to the *Banach fixed point Theorem*, also known as the contraction mapping theorem.

Theorem 1 (Banach fixed point theorem)**.** *Let $(C, \|.\|)$ be a Banach space (complete metric space). Let $f : C \to C$. Suppose that f is a contraction, that is, there exists $\tau \in [0, 1)$ such that*

$$\|f(x) - f(y)\| \leq \tau\|x - y\|, \quad \forall x, y \in C. \tag{1}$$

Then, f has a unique fixed point, that is, there exists a unique $x^\star \in C$ such that $x^\star = f(x^\star)$. Furthermore, starting from any $x_0 \in C$, the iteration

$$x_{t+1} = f(x_t),$$

converges to $x^\star$.

Recall that a Banach space is a complete normed vector space, for example, $\mathbb{R}^n$ and any compact subsets of $\mathbb{R}^n$ are Banach spaces. Note that in the definition of the contraction (1) the norm is not specified. It is sufficient to show the contraction property for any norm on C. Finally, the proof of the above is not difficult. It shows that the sequence generated by the iteration is Cauchy (the terms get close together as we look sufficiently far in the sequence) and since the space is complete, the Cauchy sequence will have a limit in the space. We can then argue this limit is the fixed point and must be unique.

Remark: We have already seen a few fixed-point theorems in game theory, for example, Brouwer's fixed-point theorem and Kakutani's fixed-point theorem. In past lectures, we used the above theorems to argue the existence of a Nash equilibrium since a Nash equilibrium is a fixed point of the best-response map. The Banach fixed point theorem is based on much stronger assumptions than the above theorems, but it also delivers more: namely, it provides a constructive way to compute the fixed point of the mapping. The first time I encountered this theorem was in the proof of the existence and uniqueness of solutions to order differential equations under Lipschitz continuity of the ODE. Recently, the attention in our community has been more on algorithms and learning. Hence, you might see this theorem first in the context of algorithm design as we do now. Also, note that fixed point theorems have other applications, such as computing equilibria of dynamical systems (closely related to control and dynamical systems). Given their importance, mathematicians have been attempting to derive them under reduced assumptions.

Let us go back to our convex game $\Gamma(\mathcal{N}, K, \{J^i\}_{i\in\mathcal{N}})$, which satisfies the conditions for the existence and uniqueness of a Nash equilibrium as discussed in the previous lecture. Namely, $K = K^1 \times \dots K^N$ is compact convex, J^i are continuous in all variables, convex and continuously differentiable in each x^i for fixed x^{-i} and F_Γ, the game mapping, is strictly monotone. How shall we design the function f such that it is a contraction and its fixed point is the Nash equilibrium of the game $\Gamma(\mathcal{N}, K, \{J^i\}_{i\in\mathcal{N}})$?

A first idea might be to try to apply the Banach fixed point theorem to the best-response map. After all, a Nash equilibrium is a fixed point of the best-response map. The problem is the best-response map is generally non-contractive. Hence, our goal is to characterize Nash equilibria of convex games as fixed points of a different map. Here, we can leverage the connection we established between convex games to convex optimization problems. In particular, we consider a standard projected gradient descent approach commonly used in optimization. Given a starting point x_0 at time $t = 0$, the next iterate is obtained by $x_{t+1} = \Pi_K(x_t - \gamma F_\Gamma(x_t))$. The function Π_K is a projection operator, that is, it takes an element x and returns its projection to the compact convex constraint set K. The parameter $\gamma > 0$ is a step size that we choose.

We thus consider $x_{t+1} = f(x_t)$ with $f(x) = \Pi_K(x - \gamma F_\Gamma(x))$. Our goals for the rest of the lecture are (a) to show that the fixed point of f is the Nash equilibrium and (b) to design a step size γ such that the above iterative procedure becomes a contraction. Then, by application of Banach's fixed-point theorem, we are sure to converge to the Nash equilibrium. As you will show, in several problem instances such as the Cournot competition, such an algorithm can be implemented without players knowing others' cost functions or constraints.

Let us define more formally the projection operator.

Definition 1. Given a convex set $K \subset \mathbb{R}^n$, the *projection operator* $\Pi_K : \mathbb{R}^n \to K$ is defined as

$$\begin{aligned} \Pi_K(x) := \arg\min_y \quad & \frac{1}{2}(y-x)^T(y-x) \\ \text{s.t.} \quad & y \in K. \end{aligned} \tag{2}$$

The interpretation of the above optimization problem in the variable y is that the projection operator Π_K maps any $x \in \mathbb{R}^n$ to an $\bar{x} \in K$, which has the minimum Euclidean distance to x.

Exercise 1. Let $\bar{x} = \Pi_K(x)$. Show that

$$(y-\bar{x})^T(\bar{x}-x) \geq 0, \quad \forall y \in K.$$

Solution. Let $f(y) = \frac{1}{2}(y-x)^T(y-x)$ denote the objective function above. Notice that this is a convex function in the decision variable y with $\nabla_y f(y) = y - x$. If $\bar{x}$ is an optimum for the optimization problem, from convexity of $f(y)$ it follows that

$$\nabla f(\bar{x})^T(y-\bar{x}) = (\bar{x}-x)^T(y-\bar{x}) \geq 0, \quad \forall y \in K.$$

Hence, we have the desired result.

Exercise 2. Show that for any $\gamma > 0$

$$(y-x)^T F(x) \geq 0, \quad \forall y \in K \iff x = \Pi_K(x - \gamma F(x)). \tag{3}$$

Solution. From Exercise 1 we have $x = \Pi_K(x - \gamma F(x))$ if and only if

$$(y-x)^T(x-(x-\gamma F(x))) \geq 0, \quad \forall y \in K \iff (y-x)^T F(x) \geq 0, \quad \forall y \in K.$$

From the last lecture and the result above we conclude that the Nash equilibrium in the convex game $\Gamma(\mathcal{N}, K, \{J^i\}_{i\in\mathcal{N}})$ can be characterized as a fixed point of the operator $x \mapsto \Pi_K(x - \gamma F_\Gamma(x))$. The question of computing a Nash equilibrium is reduced to the computation of a fixed point of the operator $x \mapsto \Pi_K(x - \gamma F_\Gamma(x))$. We will design γ so that this map is a contraction. To do so, first we show that $\|\Pi_K(g(x)) - \Pi_K(g(y))\|_2 \leq \|g(x) - g(y)\|_2$. It then follows that if $g : x \mapsto x - \gamma F_\Gamma(x)$ is a contraction with respect to the 2−norm, then so is the composition map $\Pi_K \circ g : x \mapsto \Pi_K(x - \gamma F_\Gamma(x))$ (verify this). Hence, we obtain the desired result.

Exercise 3. Show that $\Pi_K : \mathbb{R}^n \to K$ is a non-expansive operator, that is,

$$\|\Pi_K(y) - \Pi_K(x)\|_2 \leq \|y - x\|_2, \quad \forall x, y \in \mathbb{R}^n.$$

Solution. Let $\bar{x} = \Pi_K(x)$. Then, from Exercise (1) we have that

$$(y-\bar{x})^T(\bar{x}-x) \geq 0, \quad \forall y \in K.$$

Since the above holds for any $x \in R^n$ and $y \in K$, by letting $u, v \in \mathbb{R}^n$ arbitrary and choosing $x = v$ and $y = \Pi_K(u)$, we have that

$$(\Pi_K(u) - \Pi_K(v))^T(\Pi_K(v) - v) \geq 0.$$

Repeating the same steps but with u and v interchanged, we obtain

$$(\Pi_K(v) - \Pi_K(u))^T(\Pi_K(u) - u) \geq 0.$$

Now, if we add the above two inequalities, we obtain

$$\Pi_K(u)^T\Pi_K(v) - \Pi_K(v)^T\Pi_K(v) - \Pi_K(u)^Tv + \Pi_K(v)^Tv$$
$$+ \Pi_K(v)^T\Pi_K(u) - \Pi_K(v)^Tu - \Pi_K(u)^T\Pi_K(u) + \Pi_K(u)^Tu \geq 0.$$

Grouping terms, we obtain

$$\|\Pi_K(u) - \Pi_K(v)\|_2^2 \leq (\Pi_K(u) - \Pi_K(v))^T(u - v). \tag{4}$$

From the Cauchy-Schwarz inequality, we have

$$|(\Pi_K(u) - \Pi_K(v))^T(u - v)| \leq \|\Pi_K(u) - \Pi_K(v)\|\|u - v\|.$$

Finally, dividing the above by $\|\Pi_K(u) - \Pi_K(v)\|_2$ and combining with Inequality (4), we obtain the desired result.

Now, to ensure $x \mapsto x - \gamma F_\Gamma(x)$ can be a contraction we make the following assumptions on F_Γ.

Assumption 1. Let $K \subset \mathbb{R}^n$, $F : K \to \mathbb{R}^n$, and let $L, \mu \in \mathbb{R}_{>0}$ be such that $\forall x, y \in K$

$$(F(x) - F(y))^T(x - y) \geq \mu\|x - y\|_2^2 \tag{5}$$
$$\|F(x) - F(y)\|_2 \leq L\|x - y\|_2. \tag{6}$$

Note that the first assumption is satisfied when F_Γ is strongly monotone. The second assumption requires that the game mapping is Lipschitz continuous.

Proposition 1. *Under Assumption 1, if* $0 < \gamma < \frac{2\mu}{L^2}$*, then the iteration*

$$x_{t+1} = \Pi_K(x_t - \gamma F_\Gamma(x_t)), \tag{7}$$

converges to the Nash equilibrium of game $\Gamma(\mathcal{N}, K, \{J^i\}_{i\in\mathcal{N}})$.

Solution. We already showed that the fixed point of $x \mapsto \Pi_K(x - \gamma F_\Gamma(x))$ is the Nash equilibrium of the game. Hence, we now simply verify the conditions for the map above to be a contraction. In the derivation below, all norms are the 2−norm.

$$\begin{aligned}
&\|\Pi_K(x - \gamma F_\Gamma(x)) - \Pi_K(y - \gamma F_\Gamma(y))\|^2 \leq \|x - \gamma F_\Gamma(x) - (y - \gamma F_\Gamma(y))\|^2 \\
&= \|x - y\| + \gamma^2\|F_\Gamma(x) - F_\Gamma(y)\|^2 - 2\gamma(F(x) - F(y))^T(x - y) \\
&\leq (1 + \gamma^2L^2 - 2\mu\gamma)\|x - y\|^2,
\end{aligned}$$

where the first inequality above is due to the non-expansion of the projection operator, the second equality is simply writing out the 2−norm squared, the second inequality is due to the Assumption 1. Finally, we see that $0 < (1 + \gamma^2L^2 - 2\mu\gamma) < 1$ if and only if $0 < \gamma < \frac{2\mu}{L^2}$ as desired.

Notice that to implement the above algorithm, the i-th player needs to know only K^i (its own constraint set) as well as $\nabla_{x^i}J^i(x)$, which is the gradient of its own cost function with respect to its decision variables. To evaluate the gradient term $\nabla_{x^i}J^i(x_t)$, the player may need some information about the strategies x_t^{-i}. In several games, such information can be inferred from the game data. For example, in the Cournot competition, each producer only needs to know the total production of all firms rather than individual firms' production to compute this term. Another potential source for coordination in the above algorithm is that all players are using the same step size γ. The latter requirement is not so stringent and as we will see in the next section, we can relax the condition of the same step size.

Exercise 4. Consider the Cournot game introduced at the beginning of the previous lecture.

1. Propose a step size γ to ensure convergence of Equation (7) to the Nash equilibrium.
2. What does each player need to know about the game to implement this algorithm?

Solution. First, we derive the game mapping $F_\Gamma : \mathbb{R}^2 \to \mathbb{R}^2$ as the following quadratic function:

$$\begin{bmatrix} 2b+b & \\ b & 2b \end{bmatrix} \begin{bmatrix} x^1 \\ x^2 \end{bmatrix} - \begin{bmatrix} c-a \\ c-a \end{bmatrix}$$

Let the matrix above be $M \in \mathbb{R}^{2\times 2}$ and the vector above be $v \in \mathbb{R}^2$. From previous exercises, we know that the game mapping F_Γ is strongly monotone if the following holds

$$\langle F_\Gamma(x) - F_\Gamma(y), x-y\rangle = \langle Mx + v - (My+v), x-y\rangle = \langle Mx - My, x-y\rangle \geq \mu\|x-y\|^2$$

with monotonicity constant $c = \underline{\lambda}(M + M^\top)/2$, where $\underline{\lambda}$ is the smallest eigenvalue of $M + M^\top$. Thus, F_Γ is strongly monotone if μ is positive and this is the case if M is positive definite. Since matrix M is symmetric, we can compute the eigenvalues of M directly which are $3b$, and b. Given that $b > 0$, it follows that M is positive definite, and its strong monotonically constant is given by $b/2$. Thus, if $b > 0$ then Equation (5) in Assumption 1 is satisfied. Furthermore, Equation (6) in Assumption 1, that is, Lipschitz continuity of F_Γ, can be checked by observing that

$$\|F_\Gamma(x) - F_\Gamma(y)\|_2 = \|M(x-y)\|_2 \leq \|M\|_2\|x-y\|_2 \leq \sqrt{\bar{\lambda}(x)}\|x-y\|_2,$$

with $L = \bar{\lambda} = 3b$. For $0 < \gamma < \frac{2\mu}{L^2} = \frac{1}{3}$ the update $x_{t+1} = \Pi_K(x_t - \gamma F_\Gamma(x_t))$ converges to an NE.

The above example motivates us to also consider potential games in continuous action spaces.

Definition 2. Potential games in continuous action spaces Consider the game $\Gamma(\mathcal{N}, K, \{J^i\}_{i\in\mathcal{N}})$. Suppose there exists a single function $P : \mathbb{R}^{nN} \to \mathbb{R}$, such that

$$J^i(x^i, x^{-i}) - J^i(y^i, x^{-i}) = P(x^i, x^{-i}) - P(y^i, x^{-i}).$$

Then, the game is called an exact potential function.

A trivial example, yet relevant in application, is an identical interest game. Sometimes, we can work out the potential function of the game easily (see Communication game example in the problem set of this week). It follows that if the players' cost functions are differentiable, a game is an exact potential game if and only if $\nabla_{x^i} P(x) = \nabla_{x^i} J^i(x)$, for all $i \in \mathcal{N}$. Thus, the game map is precisely the gradient of the potential function. Now, considering the Jacobian of the game map, $JF_\Gamma \in \mathbb{R}^{nN\times nN}$ it follows that for a potential game, this matrix is the Hessian of P and is hence, symmetric. Thus, an equivalent condition for a game to be potential is that $JF_\Gamma(x)$ is symmetric for all $x \in K$.

Exercise 5. Show that in a potential game, any potential function optimizer is a Nash equilibrium. Note, however, that there may be other Nash equilibria that are not potential function optimizers. Next, show that a convex game which is potential, may not have a convex potential function. In particular, convex games only require player-wise convexity of the cost functions. For the potential function to be (strictly/strong) convex, the game pseudo-gradient needs to be (strictly/strongly) monotone.

Solution. First, we show that in a potential game, any potential function optimizer is a Nash equilibrium.

Let $\bar{x} \in \arg\min_{x \in K} P(x)$. Then, the following holds for all $i \in \mathcal{N}$ and $x^i \in K^i$:

$$J^i(\bar{x}^i, \bar{x}^{-i}) - J^i(x^i, \bar{x}^{-i}) = P(\bar{x}^i, \bar{x}^{-i}) - P(x^i, \bar{x}^{-i}) \leq 0.$$

Thus, $J^i(\bar{x}^i, \bar{x}^{-i}) \geq J^i(\bar{x}^i, \bar{x}^{-i})$ holds for all all $i \in \mathcal{N}$ and $x^i \in K^i$ and thus $\bar{x}$ is a Nash equilbrium.

Next, we show that a convex game which is potential, may not have a convex potential function. We show this by giving an example.

Consider a two-player game, where each player chooses an action $x^i \in \mathbb{R}$ (thus $x = (x^1, x^2)$). Define the cost function for $i = 1, 2$ as follows:

$$J^i(x^1, x^2) = \frac{1}{2}(x^1)^2 + \frac{1}{2}(x^2)^2 + 2x^1x^2,$$

The cost function J^i for $i = 1, 2$ is continuous in $x \in \mathbb{R}^2$. The cost is player-wise convex since for each player i, $J^i(x^i, x^{-i})$ is a convex quadratic function of x^i:

$$\frac{\partial^2 J^i}{\partial (x^i)^2} = 1 \geq 0.$$

The potential function P is given by:

$$P(x^1, x^2) = \frac{1}{2}(x^1)^2 + \frac{1}{2}(x^2)^2 + 2x^1x^2,$$

since the game is an identical interest. To check the convexity of P we compute its Hessian and check whether it is positive semidefinite:

$$H^P = \begin{pmatrix} \frac{\partial^2 P}{\partial (x^1)^2} & \frac{\partial^2 P}{\partial x^1 \partial x^2} \\ \frac{\partial^2 P}{\partial x^2 \partial x^1} & \frac{\partial^2 P}{\partial (x^2)^2} \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}.$$

The eigenvalues of H^P are given by -1 and 3, so there exist eigenvalues which are negative and thus the potential function P is not convex.

The lecture was mainly based on [2] and [1].

References

[1] G. Scutari, D.P. Palomar, F. Facchinei and J.S. Pang, *Convex optimization, game theory, and variational inequality theory*, IEEE Signal Processing Magazine. 2010 Apr 15;27(3):35-49.

[2] F. Facchinei and J. Pang, *Finite-dimensional variational inequalities and complementarity problems*, Springer Science & Business Media, 2007.