Games with continuous action spaces

Maryam Kamgarpour

1 Outline

- Introduction: playing Bertrand and Cournot game
- Convex game formulation
- Kakutani's fixed point theorem and existence of equilibria
- Game pseudo-gradient and monotonicity
- Uniqueness of equilibria

2 Introduction

We will consider games in which each player has a continuous action space. We will study how to formulate them and to address existence and uniqueness of Nash equilibria in these games. We also look at connections of conditions for existence and uniqueness of Nash equilibria to convex optimization problems. This lecture is based mainly on the following references [4, 3, 1].

2.1 Motivation and objectives

We will consider two very common game theoretic models used in economy, and more recently in engineering. The Cournot and the Bertrand competitions.

Example 1 (Cournot competition)**.** Consider two producers competing in a market. Player (producer) i decides on the quantity to produce denoted by $x^i \in \mathbb{R}_+$ for $i = 1, 2$, and has a production marginal cost of $c \in \mathbb{R}_+$. The market price $p : \mathbb{R}^2 \to \mathbb{R}$, is a linearly decreasing function of the total production $x^1 + x^2$, that is, $p(x^1, x^2) = a - b(x^1 + x^2)$. The producers want to maximize profit or equivalently minimize their loss.

1. Write each producer's loss as a function of the quantities produced.
2. How would you formulate the conditions for $(x^1, x^2) \in \mathbb{R}^2$ to be a Nash equilibrium?
3. How would you find the Nash equilibrium based on the above condition?

Solution 1. Note that so far, except for linear quadratic games, all games we considered were over finite strategy spaces. Now, we are going to generalize our understanding to games over infinite (in this case continuous) action spaces.

1. The loss functions $J^i : \mathbb{R}^2 \to \mathbb{R}$, for $i = 1, 2$, are given as $J^i(x) = cx^i - \big(a - b(x^1 + x^2)\big)x^i$. Note, the profit will be negative of the loss. Maximizing the profit is equivalent to minimizing the loss.

2. A pair $(x^1_{NE}, x^2_{NE}) \in \mathbb{R}^2$ is a Nash equilibrium if and only if $J^1(x^1_{NE}, x^2_{NE}) \leq J^1(x^1, x^2_{NE})$, for all $x^1 \in \mathbb{R}$, and $J^2(x^1_{NE}, x^2_{NE}) \leq J^2(x^1_{NE}, x^2)$, for all $x^2 \in \mathbb{R}$. That is, no player has incentive to unilaterally deviate from the equilibrium strategy to any feasible strategy in her strategy set.

3. For player i, let $-i$ denote the other player. From above, it follows that

$$x^i_{NE} \in \arg\min_{x^i \in \mathbb{R}_+} J^i(x^i, x^{-i}_{NE}), \quad i = 1, 2.$$

Notice that J^i is strongly convex in x^i for a fixed x^{-i}. Hence, given x^{-i}, there exists a unique minimizer of J^i with respect to x^i. Furthermore, this minimizer is characterized by the point at which $\frac{\partial J^i}{x^i} = 0$. Setting $\nabla_{x^i} J^i(x) = 0$, for $i = 1, 2$, we obtain a pair of linear equations. This system of linear equations has a solution and the solution is given by $x^i = \frac{c-a}{3b}$.

Example 2 (Bertrand competition)**.** Consider two producers as before. This time, each producer decides on the price she will charge her consumers for the electricity provided. Hence, $x^i \in \mathbb{R}_+$ for $i = 1, 2$ denotes the price announced by each producer. As before, the production marginal costs are identically $c \in \mathbb{R}_+$. The total demand is one unit and the consumers choose to buy from the producer with the lowest price. Furthermore, if both firms declare the same price, then half of the demand chooses firm 1 and the other half chooses firm 2.

1. Write each producer's profit (note, not loss this time) as a function of the price they charge.

2. Derive the Nash equilibrium.

3. What is the Nash equilibrium if each producer has the capacity to serve maximum 2/3 of the unit demand?

Solution 2. We adopt the same notation as in the previous example.

1. The profits are given by

$$J^1(x) = \begin{cases} x^1 - c, & \text{if } x^1 < x^2 \\ \frac{x^1 - c}{2}, & \text{if } x^1 = x^2 \\ 0, & \text{if } x^1 > x^2. \end{cases}$$

The cost $J^2(x)$ is defined in exactly the same manner.

2. First, verify that $(x^1, x^2) = (c, c)$ is a Nash equilibrium as no player can unilaterally deviate and improve her payoff. Now, to see this is the only Nash equilibrium, assume without loss of generality, that player 1 announces a price $x^1 > x^2$. Then, all consumers will go to player 2 and thus, player 1 will make no profit. She has incentive to decrease her price until it reaches x^2. If $x^2 > c$, then she can still decrease x^1 and attract all consumers, hence making a profit. On the other hand, for any $x^1 < c$, she will make a loss rather than a profit. By symmetry, the same argument holds for player 2.

3. Let's write the loss functions again

$$J^1(x) = \begin{cases} \frac{2(x^1 - c)}{3}, & \text{if } x^1 < x^2 \\ \frac{x^1 - c}{2}, & \text{if } x^1 = x^2 \\ \frac{x^1 - c}{3}, & \text{if } x^1 > x^2. \end{cases}$$

Consider a strategy (x^1, x^2) and suppose $x^1 > x^2$. Then, player 1 will have 1/3 of the demand. But she can increase her cost and still keep 1/3 of the demand since the maximum capacity player 2 can serve is 2/3 of the demand. So, by symmetry one player charging a larger number than the other cannot be a Nash equilibrium at any finite x^1, x^2. Continuing with this reasoning, we see that for any strategy (x^1, x^2) the players have incentive to increase their cost. So, the Nash equilibrium is not attained at a finite value in $\mathbb{R}^2$.

The above are examples of games with continuous action spaces. For the rest of the lecture, we will see how to address existence and uniqueness of Nash equilibria for such games and how to compute them in a more systematic way.

2.2 Notation and background

Vectors and matrices. For a vector x in $\mathbb{R}^n$, we denote its transpose by x^T. Similarly, the transpose of the matrix $A \in \mathbb{R}^{m\times n}$ is denoted by A^T. We let $\mathcal{S}^n_+$ and $\mathcal{S}^n_{++}$ denote the set of (they are a cone) symmetric positive semi-definite and positive definite matrices in $\mathbb{R}^{n\times n}$, respectively. Recall that a symmetric positive definite matrix $M \in \mathcal{S}^n_{++}$ has n real positive eigenvalues $\lambda_i \in \mathbb{R}_{>0}$, $i = 1, \ldots, n$. In this case, we denote the smallest and largest eigenvalues of A by $\underline{\lambda}(M)$ and $\bar{\lambda}(M)$, respectively. Furthermore, we know that for any $x \in \mathbb{R}^n$, $\underline{\lambda}(M)\|x\|_2^2 \leq x^T M x \leq \bar{\lambda}(M)\|x\|_2^2$. Let $x^i \in \mathbb{R}^n$. We denote $x = (x^1, x^2, \ldots, x^N) \in \mathbb{R}^{nN}$ as the stacked vector; $x^{-i} \in \mathbb{R}^{n(N-1)}$ is constructed from x by removing the i-th vector from the stack. An n by m matrix of constant value $b \in \mathbb{R}$ on all entries is denoted by $b_{n\times m}$.

Norms and inner products. A *norm* on $\mathbb{R}^n$ (in general on a vector space over a field) is a function $\|.\| : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ that satisfies three axioms: (1) $\|ax\| = |a|\|x\|$, $\forall a \in \mathbb{R}$, (2) $\|x + y\| \leq \|x\| + \|y\|$, (3) $\|x\| = 0 \iff x = 0$. The Euclidean or 2−norm is $\|x\|_2 := (x^T x)^{1/2} = (x_1^2 + x_2^2 + \cdots + x_n^2)^{1/2}$. For any $A \in \mathcal{S}^n_{++}$, $\|x\|_A := (x^T A x)^{1/2}$ is a well-defined norm (verify this statement). We denote the inner product $\langle x, y\rangle_A = x^T A y$. The notation $\langle x, y\rangle$ refers to standard Euclidean inner product $\langle x, y\rangle = x^T y$. From the Cauchy-Schwarz inequality it follows that $|\langle x, y\rangle_A| = |x^T A y| \leq \|x\|_A \|y\|_A$. A celebrated result of linear spaces is that all norms on an *finite dimensional* vector space are equivalent. Namely, for any $\|.\|_a, \|.\|_b : \mathbb{R}^n \to \mathbb{R}$, there exists m_l, m_u such that for all $x \in \mathbb{R}^n$, $m_l\|x\|_a \leq \|x\|_b \leq \|m_u\|x\|_a$. This is useful because sometimes it is easier to work with one norm than another, and we can be rest assured that if we prove certain results such as the limit of a sequence based on a given norm, the result does not change if we use a different norm.

Linear and quadratic functions. A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is linear if $f(\alpha x+\beta y) = \alpha f(x)+\beta f(y)$, for all $\alpha, \beta \in \mathbb{R}, x, y \in R^n$. A linear function is also referred to as a linear map. A function $f : \mathbb{R}^n \to \mathbb{R}$ is called affine if there exists $m \in \mathbb{R}$ such that $\tilde{f} = f - m$ is linear. Any matrix $A \in \mathbb{R}^{m\times n}$ gives rise to a linear map $f : \mathbb{R}^n \to \mathbb{R}^m$, with $f(x) = Ax$. [1] A function $f : \mathbb{R}^n \to \mathbb{R}$ is quadratic if $f(x) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j + \sum_i^n c_i x_i + b$. In this case, there exists (non-unique) matrix $C \in \mathbb{R}^{n\times n}$ such that $f(x) = x^T C x + c^T x + b$, with $c = (c_1, \ldots, c_n)$. When $c = 0, b = 0$, the function is also referred to as a quadratic form.

Gradient, Jacobian, and Hessian. Let $f : \mathbb{R}^n \to \mathbb{R}$ (referred to as a real-valued function since the codomain is $\mathbb{R}$) be differentiable. The gradient of f is the n-dimensional vector of its partial

[1] It can also be shown that any linear function whose domain and codomain are finite dimensional can be represented by a matrix. The representation depends on the choice of basis in the domain and codomain. Lastly, linear maps can be defined for more general domains, including function spaces. For example, integration and derivation are linear maps with suitable function spaces as domain and codomain. These are subject of more advanced courses in control.

derivatives, $\frac{\partial f}{\partial x_i} : \mathbb{R}^n \to \mathbb{R}$ for $i \in \{1, 2, \ldots, n\}$, and is denoted by $\nabla f : \mathbb{R}^n \to \mathbb{R}^n$:

$$\nabla f = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \ldots \\ \frac{\partial f}{\partial x_n} \end{pmatrix}.$$

The Hessian of f is the $n \times n$ dimensional matrix of the second order partial derivatives, $\frac{\partial^2 f}{\partial x_i \partial x_j}$: $\mathbb{R}^n \to \mathbb{R}$ for $i, j \in \{1, 2, \ldots, n\}$, and is denoted by $\nabla^2 f : \mathbb{R}^n \to \mathbb{R}^{n \times n}$.

$$\nabla^2 f = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}, \frac{\partial^2 f}{\partial x_1 \partial x_2}, \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}, \frac{\partial^2 f}{\partial x_2^2}, \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \ldots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}, \frac{\partial^2 f}{\partial x_n \partial x_2}, \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}.$$

A sufficient condition for the Hessian to be a symmetric matrix everywhere is that the function f has continuous second-order derivatives.

Exercise 1. Derive ∇f and $\nabla^2 f$ for $f : \mathbb{R}^n \to \mathbb{R}$ being (a) an affine function given by $f(x) = c'x + b$, $c \in \mathbb{R}^n$; and (b) a quadratic function $f(x) = x'Cx$, with $C \in \mathbb{R}^{n \times n}$.

Solution 3. In both cases, $\nabla f \in \mathbb{R}^n$ and $\nabla^2 f : \mathbb{R}^n \to \mathbb{R}^n$ since the domain of the function is $\mathbb{R}^n$ and are as follows.Verify the derivations here.

(a) $\nabla f(x) = c$, $\nabla^2 f(x) = 0_{n \times n}$.

(b) $\nabla f(x) = (C + C^T)x$, $\nabla^2 f(x) = C + C^T$ and if C is symmetric, this simplifies to $\nabla^2 f(x) = 2C$.

For a vector-valued function $f : \mathbb{R}^n \to \mathbb{R}^m$, the Jacobian is the (m, n) dimensional matrix of first-order partial derivatives of f, denoted by $Jf : \mathbb{R}^n \to \mathbb{R}^{n \times m}$.

$$Jf = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}, \frac{\partial f_1}{\partial x_2}, \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1}, \frac{\partial f_2}{\partial x_2}, \frac{\partial f_2}{\partial x_n} \\ \ldots \\ \frac{\partial f_n}{\partial x_1}, \frac{\partial f_n}{\partial x_2}, \frac{\partial f_n}{\partial x_n} \end{pmatrix}$$

Observe that for $f : \mathbb{R}^n \to \mathbb{R}$, the Jacobian is the transpose of the gradient of the function f.

Convex sets. A set $D \subset \mathbb{R}^n$ is *convex* if $\forall x, y \in D$, $\forall t \in [0, 1]$, $tx + (1 - t)y \in D$. A function $f : D \to \mathbb{R}$ is convex if its domain D is a convex set and $\forall x, y \in D$, $\forall t \in [0, 1]$,

$$f(tx + (1 - t)y) \leq tf(x) + (1 - t)f(y).$$

If in the above, the inequality is strict $\forall t \in (0, 1)$, then the function is *strictly convex*. As checking the above conditions in general might be difficult, there are easier tests for verifying convexity of a function if more assumptions are made on f. Namely, a *differentiable* function $f : D \to \mathbb{R}$ is convex if and only if D is convex and

$$f(y) \geq f(x) + \nabla f(x)^T (y - x), \quad \forall x, y \in D. \tag{1}$$

The above says that the function f lies above its linear approximator. If the above inequality is strict, then the function is *strictly convex*. A *twice differentiable* function $f : D \to \mathbb{R}$ is convex if and only if D is convex and the Hessian $\nabla^2 f(x)$ is positive semi-definite ($\nabla^2 f(x) \in \mathcal{S}^n_+$) $\forall x \in D$, and is strictly convex if $\nabla^2 f(x)$ is positive definite ($\nabla^2 f(x) \in \mathcal{S}^n_{++}$) $\forall x \in D$ [2]. The twice differentiable function is strongly convex if there exists $\eta > 0$ such that $x^T \nabla^2 f(x) x \geq \eta \|x\|_2^2 > 0$, for all $x \in D$.

Exercise 2. Verify that a quadratic function is strongly convex if and only if it is strictly convex.

Solution 4. The Hessian of a quadratic function $x^T Cx + c^T x + b$ is given by $(C + C^T)$. Note that since $C + C^T$ is symmetric, it has real-valued eigenvalues and thus, the eigenvalues can be ordered $\lambda_{max} \geq \cdots \geq \lambda_{min}$. Now, $x^T(C + C^T)x \geq \lambda_{min}(C + C^T)\|x\|_2^2$. It follows that if $\lambda_{min} > 0$ the quadratic function is strongly (and thus, strictly) convex. However, if $\lambda_{min} \leq 0$, the quadratic function is not strictly (and thus, not strongly) convex.

Minimizer. A point $x^\star \in D$ is a minimizer of f if and only if $f(x^\star) \leq f(y)$, $\forall y \in D$. Note that sometimes x^* is referred to as a global minimizer in contrast to a *local minimizer*. A point $x^l \in D$ is a local minimizer of f, if there exists some neighborhood of x^l in D, denoted by $N(x^l)$ such that $f(x^\star) \leq f(y)$, $\forall y \in N(x^l) \cap D$. In general, searching for a minimizer is hard. However, if the function f is convex and differentiable, we can use necessary and sufficient optimality conditions based on the gradient of f.

Fact 1. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a continuously differentiable convex function, with derivate $\nabla f : \mathbb{R}^n \to \mathbb{R}^n$. Let $K \subset \mathbb{R}^n$ be a convex set. Consider the following constrained optimization problem

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & x \in K. \end{aligned} \tag{2}$$

A feasible point, $x^\star \in K$, is a solution of the above problem (in other words it achieves the minimum above or equivalently $x^\star \in \arg\min_{x \in K} f(x)$) if and only if $\nabla f(x^\star)^T(y - x^\star) \geq 0$, $\forall y \in K$.

Note that unless K is compact, a minimizer for the above may not exist (consider minimizing e^{-x^2}). Furthermore, if a minimizer exists, it may not be unique and hence, we write $x^\star \in \arg\min_{x \in K} f(x)$ rather than $x^\star = \arg\min_{x \in K} f(x)$.

Exercise 3. Prove Fact 1.

Solution 5. We need to show two statements: 1) if $x^\star$ is an optimizer then $\nabla f(x^\star)^T(y - x^\star) \geq 0$, $\forall y \in K$ and that 2) if $\nabla f(x^\star)^T(y - x^\star) \geq 0$, $\forall y \in K$ then $x^\star$ is an optimizer. Let us start with 2), namely, showing that $\nabla f(x^\star)^T(y - x^\star) \geq 0$, $\forall y \in K$ is a sufficient condition for optimality. If $\nabla f(x^\star)^T(y - x^\star) \geq 0$, $\forall y \in K$ then, due to convexity of f, from Inequality (1) we conclude that $f(y) \geq f(x^\star)$, $\forall y \in K$. Hence, $x^\star$ is a minimum. To show statement 1), namely, that $\nabla f(x^\star)^T(y - x^\star) \geq 0, \forall y \in K$ is a necessary condition for optimality, we can prove by contradiction. In particular, suppose there exists $z \in K$ such that $\nabla f(x^\star)^T(z - x^\star) < 0$. Since the function is differentiable, using first-order approximation of the function around the point $x^\star$, we know that $f(tz + (1-t)x^\star) = f(x^\star) + \nabla f(x^\star)'(t(z - x^\star)) + r(tz + (1-t)x^\star)(t(z - x^\star))$, where $r(t(z - x^\star)) \to 0$ as $t \to 0$. In other words, the third term in the first-order approximation is dominated by the second term. Hence, we can then show that for sufficiently small $t \in (0, 1]$, $\nabla f(x^\star)'(t(z - x^\star)) + r(tz + (1 - t)x^\star)(t(z - x^\star)) < 0$ and consequently, $f(tz + (1 - t)x^\star) < f(x^\star)$. This contradicts optimality of $x^\star$, since $tz + (1 - t)x^\star$ is a feasible solution (by convexity of set K) with a lower objective [2]

3 Convex Games

We consider an N player game with player set denoted by $\mathcal{N} = \{1, 2, \ldots, N\}$. For $i \in \mathcal{N}$, let $x^i \in K^i \subset \mathbb{R}^n$, where K^i is non-empty, closed and convex. Let $J^i : \mathbb{R}^{nN} \to \mathbb{R}$ denote player i's objective function. Let $K = K^1 \times K^2 \times \cdots \times K^N \subset \mathbb{R}^{nN}$. We equivalently represent this game by $\Gamma(\mathcal{N}, K, \{J^i\}_{i \in \mathcal{N}})$.

Definition 1. A point $x \in K$ is a Nash equilibrium for $\Gamma(\mathcal{N}, K, \{J^i\}_{i\in\mathcal{N}})$ if and only if

$$J^i(x^i, x^{-i}) \leq J^i(y^i, x^{-i}), \quad \forall y^i \in K^i, \quad \forall i \in \mathcal{N}. \tag{3}$$

Under which conditions on $\Gamma(\mathcal{N}, K, \{J^i\}_{i\in\mathcal{N}})$ does a Nash equilibrium exists? The following theorem provides sufficient conditions for the existence of Nash equilibria [3].

Theorem 1. *Suppose $K \subset \mathbb{R}^{nN}$ is compact convex and J^i are continuous in $x \in K$ and convex in x^i for fixed x^{-i}. Then, $\Gamma(\mathcal{N}, K, \{J^i\}_{i\in\mathcal{N}})$ has a Nash equilibrium.*

We follow the proof provided in [3]. It is similar to our argument for existence of mixed strategy Nash equilibrium in finite action games based on fixed point of best response maps. Note however that here we have an infinite action space ($\mathbb{R}^n$) but we are looking for pure strategies.

Proof. Define $\rho : K \times K \to \mathbb{R}$ as $\rho(x, y) = \sum_{i=1}^{N} J^i(x^1, \ldots, y^i, \ldots, x^N)$. By continuity of J^i, ρ is continuous on $K \times K$ and by convexity of J^i in x^i for fixed x^{-i}, ρ is convex in y for fixed x. Consider $\Omega(x) = \{y \mid \rho(x, y) = \min_{z\in K} \rho(x, z)\}$. In general Ω is a set-valued map, that is, it maps each point in K to a subset of K. Given that 1) K is convex, 2) ρ is convex in z and 3) ρ is continuous, it follows that the map $x \mapsto \Omega(x)$ is lower semi-continuous (see The Minimum Theorem below) and $\Omega(x) \subset K$ is compact for each $x \in K$. By Kakutani's fixed point theorem (see below), there exists $x_0 \in K$ such that $x_0 \in \Omega(x_0)$. We can verify that this x_0 is a Nash equilibrium as follows. By definition of x_0 being the fixed-point of the map Ω we have $\rho(x_0, x_0) = \min_{z\in K} \rho(x_0, z)$. This implies that x_0 satisfies the Nash equilibrium definition, namely, $J^i(x_0^i, x_0^{-i}) \leq J^i(\tilde{x}^i, x_0^{-i})$ for each i and for all $\tilde{x}^i \in K$. To see this, suppose that there exists i and $\tilde{x}_0 = (\tilde{x}^i, x_0^{-i}) \in K$ such that $J^i(\tilde{x}_0) = J^i(\tilde{x}^i, x_0^{-i}) < J^i(x^i, x_0^{-i})$. Then, by definition of ρ, we get $\rho(x_0, \tilde{x}_0) < \rho(x_0, x_0)$. This however contradicts $x_0 \in \Omega(x_0)$. $\square$

For completeness we provide the Kakutani's fixed point theorem. It generalizes the Brower's fixed point theorem to set-valued maps. Most Nash equilibrium existence results use fixed point theorems and there are generalizations of Kaktuani's fixed point theorem, to derive less existence conditions for more general classes of games.

Fact 2. [Kakutani's fixed point theorem] Let $K \subset \mathbb{R}^n$ be a non-empty convex compact set. Let $\Omega : K \to 2^K$ be a lower semicontinuous map such that for each $x \in K$, $\Omega(x)$ is non-empty, convex and closed subset of K. Then, there exists $x_0 \in K$ such that $x_0 \in \Omega(x_0)$.

The following theorem provides conditions on well-behavedness of set of minimizers of a function as a parameter changes. It is used in the proof of the theorem above on existence of Nash equilibria. It also has applications in optimal control theory. Note that the statement of the theorem can be generalized but the following version suffices for our problem.

Fact 3. [The Minimum theorem] Let $K \subset \mathbb{R}^n$ be compact and $f : X \times Y \to \mathbb{R}$ be continuous on $X \times Y$, and convex in Y for each fixed x. Then, for each point $x \in X$, $\Omega(x) = \arg\min_{y\in K}\{f(x, y) \mid y \in K\} \subset Y$ is lower semicontinuous and $\Omega(x) \subset Y$ is a compact convex set.

Exercise 4. Consider a finite action game, where each agent i has an action set $\mathcal{A}^i$ with cardinality d^i. Show that the mixed strategy extension of any finite action game is a convex game and satisfies the assumptions of the above theorem. In the mixed strategy extension, the agents can choose a mixed strategy from the probability simplex $\Delta(\mathcal{A}^i)$ which is defined as

$$\Delta(\mathcal{A}^i) = \left\{ x^i | x^i \geq 0, \; \sum_{d=1}^{d^i} x_d^i = 1 \right\}.$$

Solution 6. The probability simplex $\Delta(\mathcal{A}^i)$ is compact and convex for each $i = 1, \ldots, N$. Thus, the set $\Delta(\mathcal{A}) = \Delta(\mathcal{A}^1) \times \Delta(\mathcal{A}^2) \times \cdots \times \Delta(\mathcal{A}^N)$ is the cartesian product of N compact and convex sets and is therefore compact and convex as well. The cost function takes the form:

$$J^i(x^i, x^{-i}) = \sum_{a^1 \in \mathcal{A}^1} \cdots \sum_{a^N \in \mathcal{A}^N} x^1(a^1) \cdot \ldots \cdot x^N(a^N) J^i(a^1, \ldots, a^N),$$

where $x^i(a^i)$ is the probability of player i playing action $a^i \in \mathcal{A}^i$. The cost function is continuous on $\Delta(\mathcal{A})$. For a fixed x^{-i}, the cost J^i is also linear and thus convex on $\Delta(\mathcal{A}^i)$. All the assumptions of the Minimum theorem are thus satisfied.

Exercise 5. Go back to Cournot and Bertrand models at the beginning of the lecture and see if they meet the assumptions for existence of Nash equilibria as above.

Solution 7. In the Cournot competition, the losses are $J^i(x^1, x^2) = -px^i + c^i x^i$. It follows that the cost function of each player is continuous. However, with unbounded decision space the action spaces are not compact. Nevertheless, we were able to uniquely characterize the Nash equilibrium. This highlights that the above conditions are only sufficient and not necessary. When the action spaces are constrained, then the above theorem shows existence of Nash equilibrium. In the Bertrand competition, neither the cost functions are continuous, nor the action spaces are compact. This game is not a convex game. Hence, the above theorem does not apply. However, we saw that at least without capacity constraints, we could find the unique Nash equilibrium of the game.

Note that if the action spaces are non-compact, a Nash equilibrium may not exist even if the cost function of each player is strongly convex in her decision variables. As an example, consider $J^1(x^1, x^2) = (x^1)^2 + 2x^1x^2 + 3x^1$, $J^2(x^1, x^2) = (x^2)^2 + 2x^1x^2 + 2x^2$. Verify that a Nash equilibrium must satisfy the following system of linear equations. However, this system does not have a solution:

$$\begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} x^1 \\ x^2 \end{bmatrix} = \begin{bmatrix} -3 \\ -2 \end{bmatrix}$$

The conditions of the theorem above may seem restrictive in light of the first two examples. Nevertheless, several classes of games arising in engineering problems do satisfy the above conditions. If a game does not satisfy these conditions, there may be alternative approaches to show existence of a Nash equilibrium, such as deriving the Nash equilibrium as we did in Examples 1, 2, or using fixed point theorems that are more general than Kakutani's (an active area in mathematics is to derive fixed point theorems and we now see how applicable they are to game theory, and as we will later see, to also algorithm design and dynamical system analysis).

From now on, we consider $\Gamma(\mathcal{N}, K, \{J^i\}_{i \in \mathcal{N}})$, with set K as Cartesian product of convex sets $K^i \subset \mathbb{R}^n$, that is, $K = K^1 \times K^2 \times \cdots \times K^N \subset \mathbb{R}^{nN}$ and whose J^i are continuous in $x \in K$ and convex in x^i for fixed x^{-i}.

4 Characterization of Nash equilibria of convex games

Definition 2. Given a convex game $\Gamma(\mathcal{N}, K, \{J^i\}_{i \in \mathcal{N}})$, the game map $F_\Gamma : \mathbb{R}^{nN} \to \mathbb{R}^{nN}$ is defined as

$$F_\Gamma = (\nabla_{x^1} J^1; \nabla_{x^2} J^2; \ldots; \nabla_{x^N} J^N). \tag{4}$$

The above is also referred to as the *game pseudo-gradient* due to its resemblance to gradient of a single function.

Exercise 6. Show that x is a Nash equilibrium for $\Gamma(\mathcal{N}, K, \{J^i\}_{i\in\mathcal{N}})$ if and only if $F_\Gamma(x)^T(y-x) \geq 0$ for all $y \in K$.

Solution 8. First, observe that

$$F_\Gamma(x)^T(y-x) = (\nabla_{x^1}J^1; \nabla_{x^2}J^2; \dots; \nabla_{x^N}J^N)^T(y-x) = \sum_{i=1}^{N} \nabla_{x^i}J^i(x^i, x^{-i})^T(y^i - x^i).$$

Next, x is a Nash equilibrium if and only if $x^i \in \arg\min_{y^i\in K^i} J^i(y^i, x^{-i})$, $\forall i \in \mathcal{N}$. From convexity of J^i in x^i it follows that x is a Nash equilibrium if and only if $\nabla_{x^i}J^i(x^i, x^{-i})^T(y^i - x^i) \geq 0$, $\forall y^i \in K^i$, $i \in \mathcal{N}$. This implies that $\sum_{i=1}^{N} \nabla_{x^i}J^i(x^i, x^{-i})^T(y^i - x^i) = F_\Gamma^T(x)(y-x) \geq 0$. On the other hand, if $\sum_{i=1}^{N} \nabla_{x^i}J^i(x^i, x^{-i})^T(y^i - x^i) \geq 0$ for all $y^i \in K^i$, then each of the terms in the summand has to be non-negative (verify). Hence, x is a Nash equilibrium if and only if $F_\Gamma(x)^T(y-x) \geq 0$ for all $y \in K$ as desired.

The above gives a first order characterization of Nash equilibria based on the game pseudo-gradient. We will use this characterization to determine conditions for uniqueness of equilibria and design of convergent algorithms. To dos, we need to look at properties of vector fields $F : D\mathbb{R}^n$, with $D \subset \mathbb{R}^n$.

Definition 3. The map $F : D \to \mathbb{R}^n$ is *monotone* if $\langle F(x) - F(y), x - y\rangle \geq 0$ for all $x, y \in D \subseteq \mathbb{R}^n$. If the above inequality is strict for all $x \neq y$, then F is *strictly monotone*.

Let us get more insight into monotonicity but looking at linear maps.

Exercise 7. Consider the linear map defined by the matrix multiplication $A : x \mapsto Ax$. Under which conditions on the matrix A is this operator monotone? strictly monotone? strongly monotone?

Solution 9. We need $\langle Ax - Ay, x - y\rangle \geq 0$ for all $x, y \in \mathbb{R}^n$. This is equivalent to $z^TAz \geq 0$. We have $z^TAz = (z^TAz)^T = z^TA^Tz = \frac{1}{2}z^T(A + A^T)z$. Notice that $A + A^T$ is a symmetric matrix that has real eigenvalues. We know that for a symmetric matrix M, $\underline{\lambda}(M)\|z\|_2^2 \leq z^TMz \leq \bar{\lambda}(M)\|x\|_2^2$ if and only if $M \in \mathcal{S}_+^n$. Hence, we conclude that $z^TAz \geq 0$ for all $z \in \mathbb{R}^n$ if and only if the symmetric part of A is positive semidefinite.

For strict monotonicity, we require $z^TAz > 0$ for all $z \in \mathbb{R}^n, z \neq 0$. Observe that $z^TAz \geq \underline{\lambda}(A + A^T)\|z\|^2/2$. Hence, if the smallest eigenvalue of $A + A^T$ is positive, then the map is both strictly and strongly monotone, with strong monotonicity constant of $\underline{\lambda}(A + A^T)/2$.

Note on quadratic functions: A common mistake is to consider a function $f(x) == x^TCx + c^Tx + b$ as convex and thus, easy to optimize. Unfortunately, this is not the case if Q is not positive semi-definite. Indeed, if Q is indefinite, the problem of optimizing the above function is NP-hard. The NP-hardness is shown by formulating boolean optimization problems as instances of above quadratic optimization.

Monotonicity generalizes the notion of convexity as can be verified below.

Exercise 8. Let $f : \mathbb{R}^n \to \mathbb{R}$ be continuously differentiable and convex. Show that $\nabla f : \mathbb{R}^n \to \mathbb{R}^n$ is monotone. If in addition, f is strictly convex, show that ∇f is strictly monotone.

Solution 10. We use the the equivalent characterization of convex differentiable functions in (1). Note that from convexity of f we have, for every $x, y \in \mathbb{R}^n$

$$f(y) \geq f(x) + \nabla f(x)^T(y - x),$$
$$f(x) \geq f(y) + \nabla f(y)^T(x - y).$$

If we add the above two inequalities, we obtain $(\nabla f(y) - \nabla f(x))^T(y - x) \geq 0$ as desired.

If the map F is not affine, we can verify monotonicity using the Jacobian of the map. For a map $F : \mathbb{R}^n \to \mathbb{R}^n$, let $JF := \left(\frac{\partial F_i}{\partial x_j}\right) \in \mathbb{R}^{n\times n}$ denote the Jacobian of F.

Proposition 1. *Let $F : D \to \mathbb{R}^n$ be continuously differentiable on the open convex set $D \subset \mathbb{R}^n$. Then, F is monotone on D if and only if $JF(x)$ is positive semi-definite for all $x \in D$. Furthermore, F is strictly monotone on D if and only if $JF(x)$ is positive definite for all $x \in D$.*

See [1] for proof. You can verify that for the map F arising from gradient of a function f, the above is a generalization of second-order condition for convexity. In particular, if F is a gradient map, then JF will be a symmetric matrix corresponding to the Hessian of f.

Finally, we can state a condition for uniqueness of a Nash equilibrium. This condition also helps in design of convergent algorithms.

Fact 4. From [1, Chapter 2, Corollary 2.2.5] Consider the convex game $\Gamma(\mathcal{N}, K, \{J^i\}_{i\in\mathcal{N}})$. The Nash equilibrium is unique if F_Γ is strictly monotone.

Exercise 9 (Cournot competition continued)**.** Let us return to Example 1. Consider the case in which each player's decision space is $[0, k^i] \subset \mathbb{R}$, $i = 1, 2$. Characterize the Nash equilibrium using the pseudo-gradient. Is there a unique Nash equilibrium in this setting?

Solution 11. Verify that $F_\Gamma(x) = [\nabla_{x^1} J^1(x^1, x^2); \nabla_{x^2} J^2(x^1, x^2)]$ is strictly monotone. Hence, the unique Nash equilibrium of this game is the solution of $\mathrm{VI}(K, F_\Gamma)$, where $K = [0, k^1] \times [0, k^2]$.

References

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

[2] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.

[3] J. B. Rosen, J. Ben, *Existence and uniqueness of equilibrium points for concave n-person games*, Econometrica: Journal of the Econometric Society, 1965.

[4] S. Kakutani, *A generalization of Brouwer's fixed point theorem*, Duke mathematical journal, 8(3), 457-459, 1941