NASH EQUILIBRIUM VIA CONVEX ANALYSIS
Nguyen Mau Nam

Abstract. In this note we present a simple proof of Nash’s famous theorem on the existence of Nash Equilibrium. The argument (due to R. T. Rockafellar) uses very basic elements of convex analysis, and the Brower Fixed-Point Theorem.

Key words. Nash equilibrium, Brower Fixed-Point Theorem, Convex Analysis

1 Elements of Convex Analysis

In this section we review some elements of convex analysis to be used in the proof of Nash’s theorem. The detailed proofs are given for the convenience of the readers. The readers are also referred to the books [2, 6] for more complete study of convex analysis in finite dimensions. Throughout the note we consider the Euclidean space \( \mathbb{R}^n \) with the inner product denoted by \( \langle \cdot, \cdot \rangle \) and the Euclidean norm denoted by \( \| \cdot \| \).

For two points \( a \) and \( b \) in \( \mathbb{R}^n \), the line segment connecting them is
\[
[a, b] := \{ \lambda a + (1 - \lambda) b \mid \lambda \in [0, 1] \}.
\]

Note that if \( a = b \), then this interval reduces to a singleton \( [a, b] = \{a\} \).

A subset \( \Omega \) of \( \mathbb{R}^n \) is convex if \( [a, b] \subseteq \Omega \) whenever \( a, b \in \Omega \). Equivalently, \( \Omega \) is convex if and only if \( \lambda a + (1 - \lambda) b \in \Omega \) for all \( a, b \in \Omega \) and \( \lambda \in [0, 1] \).

Let \( f: \mathbb{R}^n \to \mathbb{R} \) be a real-valued function. The epigraph of \( f \) is a subset of \( \mathbb{R}^n \times \mathbb{R} \) defined by
\[
\text{epi } f := \{ (x, \alpha) \in \mathbb{R}^{n+1} \mid x \in \mathbb{R}^n \text{ and } \alpha \geq f(x) \}.
\]

The function \( f \) is called convex if
\[
f(\lambda x + (1 - \lambda) y) \leq \lambda f(x) + (1 - \lambda) f(y) \quad \text{for all } x, y \in \mathbb{R}^n \text{ and } \lambda \in (0, 1).
\]

From the definition, we can show that \( f \) is convex if and only if its epigraph is a convex set.

Given a nonempty set \( \Omega \subseteq \mathbb{R}^n \), the distance function associated with \( \Omega \) is defined by
\[
d(x; \Omega) := \inf \{ \| x - \omega \| \mid \omega \in \Omega \}, \quad x \in \mathbb{R}^n. \tag{1.1}
\]

For each \( x \in \mathbb{R}^n \), the Euclidean projection from \( x \) to \( \Omega \) is defined based on the distance function as follows
\[
P(x; \Omega) := \{ \omega \in \Omega \mid \| x - \omega \| = d(x; \Omega) \}. \tag{1.2}
\]
**Proposition 1.1** Let \( \Omega \) be a nonempty closed subset of \( \mathbb{R}^n \). Then for any \( x \in \mathbb{R}^n \), the Euclidean projection \( P(x; \Omega) \) is nonempty.

**Proof.** By definition (1.2), for each \( k \in \mathbb{N} \), there exists \( \omega_k \in \Omega \) such that

\[
d(x; \Omega) \leq \|x - \omega_k\| < d(x; \Omega) + \frac{1}{k}.
\]

Obviously, \( \{\omega_k\} \) is a bounded sequence. Thus it has a subsequence \( \{\omega_{k_m}\} \) that converges to \( \omega \). Since \( \Omega \) is closed, \( \omega \in \Omega \). Letting \( m \to \infty \) in the inequality

\[
d(x; \Omega) \leq \|x - \omega_{k_m}\| < d(x; \Omega) + \frac{1}{k_m},
\]

we have \( d(x; \Omega) = \|x - \omega\| \), which ensures that \( \omega \in P(x; \Omega) \). □

An interesting consequence of convexity is the following projection property.

**Proposition 1.2** If \( \Omega \) is a nonempty closed convex subset of \( \mathbb{R}^n \), then for each \( x \in \mathbb{R}^n \), the Euclidean projection \( P(x; \Omega) \) is a singleton.

**Proof.** The nonemptiness of the projection \( P(x; \Omega) \) follows from Proposition 1.1. To prove the uniqueness, suppose that \( \omega_1, \omega_2 \in P(x; \Omega) \) with \( \omega_1 \neq \omega_2 \). Then

\[
\|x - \omega_1\| = \|x - \omega_2\| = d(x; \Omega).
\]

By the classical parallelogram equality, we have that

\[
2\|x - \omega_1\|^2 = \|x - \omega_1\|^2 + \|x - \omega_2\|^2 = 2\left\|x - \frac{\omega_1 + \omega_2}{2}\right\|^2 + \left\|\frac{\omega_1 - \omega_2}{2}\right\|^2.
\]

This directly implies that

\[
\left\|x - \frac{\omega_1 + \omega_2}{2}\right\|^2 = \|x - \omega_1\|^2 - \frac{\|\omega_1 - \omega_2\|^2}{4} < \|x - \omega_1\|^2 = \left[d(x; \Omega)\right]^2,
\]

which is a contradiction due to the inclusion \( \frac{\omega_1 + \omega_2}{2} \in \Omega \). □

Next we characterized the Euclidean projection to convex sets in \( \mathbb{R}^n \).

**Proposition 1.3** Let \( \Omega \) be a nonempty closed convex subset of \( \mathbb{R}^n \). Then \( \bar{\omega} \in P(\bar{x}; \Omega) \) if and only if

\[
\langle \bar{x} - \bar{\omega}, \omega - \bar{\omega} \rangle \leq 0 \text{ for all } \omega \in \Omega.
\]

**Proof.** Suppose that \( \bar{\omega} \in P(\bar{x}; \Omega) \). For any \( \omega \in \Omega \) and \( \lambda \in (0, 1) \), we have \( \bar{\omega} + \lambda(\omega - \bar{\omega}) \in \Omega \). Thus

\[
\|\bar{x} - \bar{\omega}\|^2 = \left[d(\bar{x}; \Omega)\right]^2 \leq \|\bar{x} - \left[\bar{\omega} + \lambda(\omega - \bar{\omega})\right]\|^2
= \|\bar{x} - \bar{\omega}\|^2 - 2\lambda\|\bar{x} - \bar{\omega}, \omega - \bar{\omega}\| + \lambda^2\|\omega - \bar{\omega}\|^2.
\]
This readily implies that
\[ 2\langle \bar{x} - \bar{\omega}, \omega - \bar{\omega} \rangle \leq \lambda \|\omega - \bar{\omega}\|^2. \]

Letting \( \lambda \downarrow 0 \), we obtain (1.3)

Let us now prove the converse by assuming that (1.3) is satisfied. For any \( \omega \in \Omega \), the following estimates show that \( \bar{\omega} \in P(\bar{x}; \Omega) \):

\[
\|\bar{x} - \omega\|^2 = \|\bar{x} - \bar{\omega} + \bar{\omega} - \omega\|^2 \\
= \|\bar{x} - \bar{\omega}\|^2 + \|\bar{\omega} - \omega\|^2 + 2\langle \bar{x} - \bar{\omega}, \bar{\omega} - \omega \rangle \\
= \|\bar{x} - \bar{\omega}\|^2 + \|\bar{\omega} - \omega\|^2 - 2\langle \bar{x} - \bar{\omega}, \omega - \bar{\omega} \rangle \geq \|\bar{x} - \bar{\omega}\|^2.
\]

The proof is now complete. \( \square \)

We know from the above that for any nonempty closed set \( \Omega \) in \( \mathbb{R}^n \) and for any \( x \in \mathbb{R}^n \), the Euclidean projection \( P(x; \Omega) \) is a singleton. Now we show that the projection mapping is in fact nonexpansive, i.e., satisfies the Lipschitz property in (1.4), which also implies that it is continuous.

**Proposition 1.4** Let \( \Omega \) be a nonempty closed convex subset of \( \mathbb{R}^n \). Then for any \( x_1, x_2 \in \mathbb{R}^n \), we have the estimate

\[
\|P(x_1; \Omega) - P(x_2; \Omega)\|^2 \leq \langle P(x_1; \Omega) - P(x_2; \Omega), x_1 - x_2 \rangle.
\]

In particular, it implies the Lipschitz continuity of the projection mapping with constant \( \ell = 1 \):

\[
\|P(x_1; \Omega) - P(x_2; \Omega)\| \leq \|x_2 - x_2\| \text{ for all } x_1, x_2 \in \mathbb{R}^n.
\]

**Proof.** It follows from the preceding proposition that

\[
\langle P(x_2; \Omega) - P(x_1; \Omega), x_1 - P(x_1; \Omega) \rangle \leq 0 \text{ for all } x_1, x_2 \in \mathbb{R}^n.
\]

Changing the role of \( x_1, x_2 \) in the above inequality and summing them up give us

\[
\langle P(x_1; \Omega) - P(x_2; \Omega), x_2 - x_1 + P(x_1; \Omega) - P(x_2; \Omega) \rangle \leq 0.
\]

This implies the first estimate in the proposition. Finally, the nonexpansive property of the Euclidean projection follows directly from

\[
\|P(x_1; \Omega) - P(x_2; \Omega)\|^2 \leq \langle P(x_1; \Omega) - P(x_2; \Omega), x_1 - x_2 \rangle \\
\leq \|P(x_1; \Omega) - P(x_2; \Omega)\| \cdot \|x_1 - x_2\|
\]

for all \( x_1, x_2 \in \mathbb{R}^n \), which completes the proof of the proposition. \( \square \)

Let \( \Omega \) be a nonempty, convex subset of \( \mathbb{R}^n \) and let \( \bar{x} \in \Omega \). The *normal cone* to \( \Omega \) at \( \bar{x} \) is

\[
N(\bar{x}; \Omega) := \{ v \in \mathbb{R}^n \mid \langle v, x - \bar{x} \rangle \leq 0 \text{ for all } x \in \Omega \}.
\]

The following proposition establish a useful relation between the normal cone and the projection to convex sets.
Proposition 1.5 Let $\Omega$ be a nonempty closed convex subset of $\mathbb{R}^n$ and let $\bar{x} \in \Omega$. Then $v \in N(\bar{x}; \Omega)$ if and only if $\bar{x} \in P(\bar{x} + v; \Omega)$.

Proof. By the definition, one has that $v \in N(\bar{x}; \Omega)$ if and only if
\[ \langle v, w - \bar{x} \rangle = \langle \bar{x} + v - \bar{x}, w - \bar{x} \rangle \leq 0 \quad \text{for all } w \in \Omega, \]
which is equivalent to the fact that $\bar{x} \in P(\bar{x} + v; \Omega)$ by Proposition 1.3. □

Consider the constrained optimization problem (P):
\[
\begin{align*}
\text{minimize} & \quad f(x) \\
\text{subject to} & \quad x \in \Omega,
\end{align*}
\]
where $f : \mathbb{R}^n \to \mathbb{R}$ is a convex function and $\Omega$ is a nonempty closed convex subset of $\mathbb{R}^n$.

Recall that an element $\bar{x} \in \Omega$ is called an optimal solution of problem (P) if and only if
\[ f(x) \geq f(\bar{x}) \quad \text{for all } x \in \Omega. \]

Proposition 1.6 Suppose that $f : \mathbb{R}^n \to \mathbb{R}$ is convex and $C^1$ and that $\Omega$ is a nonempty closed convex subset of $\mathbb{R}^n$. An element $\bar{x}$ is an optimal solution of problem (P) if and only if
\[ 0 \in \nabla f(\bar{x}) + N(\bar{x}; \Omega), \]
or, equivalently, $-\nabla f(\bar{x}) \in N(\bar{x}; \Omega)$.

Proof. Suppose that $\bar{x}$ is an optimal solution of the problem. Then
\[ f(x) \geq f(\bar{x}) \quad \text{for all } x \in \Omega. \]

Fix any $u \in \Omega$ and $t \in (0, 1)$. Since $\Omega$ is convex, $\bar{x} + t(u - \bar{x}) = tu + (1 - t)\bar{x} \in \Omega$. Thus,
\[ f(\bar{x} + t(u - \bar{x})) \geq f(\bar{x}). \]

This implies
\[ \frac{f(\bar{x} + t(u - \bar{x})) - f(\bar{x})}{t} \geq 0. \]

Taking the limit as $t \to 0^+$, one obtains
\[ \langle \nabla f(\bar{x}), u - \bar{x} \rangle \geq 0, \]
which can be rewritten as $\langle -\nabla f(\bar{x}), u - \bar{x} \rangle \leq 0$. Since this inequality holds true for any $u \in \Omega$, one has that $-\nabla f(\bar{x}) \in N(\bar{x}; \Omega)$.

Let us prove the converse. Suppose that $-\nabla f(\bar{x}) \in N(\bar{x}; \Omega)$. Then
\[ 0 \leq \langle \nabla f(\bar{x}), u - \bar{x} \rangle \quad \text{for all } u \in \Omega. \]

Since $f$ is a convex function, we always have
\[ \langle \nabla f(\bar{x}), u - \bar{x} \rangle \leq f(u) - f(\bar{x}) \quad \text{for all } u \in \mathbb{R}^n. \]

In particular, for every $u \in \Omega$,
\[ 0 \leq \langle \nabla f(\bar{x}), u - \bar{x} \rangle \leq f(u) - f(\bar{x}). \]

Thus, $f(\bar{x}) \leq f(u)$ for every $u \in \Omega$, which ensures that $\bar{x}$ is an optimal solution of (P). □
2 Nash Equilibrium

This section gives a brief introduction to noncooperative game theory and presents a simple proof of the existence of Nash equilibrium as a consequence of convexity and Brouwer’s fixed point theorem. For simplicity, we only consider two-person games while more general situations can be treated similarly.

**Definition 2.1** Let $\Omega_1$ and $\Omega_2$ be nonempty subsets of $\mathbb{R}^m$ and $\mathbb{R}^n$, respectively. A noncooperative game in the case of two players $I$ and $II$ consists of two strategy sets $\Omega_i$ and two real-valued functions $u_i : \Omega_1 \times \Omega_2 \to \mathbb{R}$ for $i = 1, 2$ called the payoff functions. Then we refer to this game as $\{\Omega_i, u_i\}$ for $i = 1, 2$.

The key notion of Nash equilibrium was introduced by John Forbes Nash, Jr. in 1950 who proved the existence of such an equilibrium and was awarded by the Nobel Memorial Prize in Economics in 1994; see, e.g., [8] and the references therein for more details.

**Definition 2.2** Given a noncooperative two-person game $\{\Omega_i, u_i\}, i = 1, 2$, a Nash equilibrium is an element $(\bar{x}_1, \bar{x}_2) \in \Omega_1 \times \Omega_2$ satisfying the conditions

$$u_1(x_1, \bar{x}_2) \leq u_1(\bar{x}_1, \bar{x}_2) \quad \text{for all } x_1 \in \Omega_1,$$

$$u_2(\bar{x}_1, x_2) \leq u_2(\bar{x}_1, \bar{x}_2) \quad \text{for all } x_2 \in \Omega_2.$$

The conditions above mean that $(\bar{x}_1, \bar{x}_2)$ is a pair of best strategies that both player agree with in the sense that $\bar{x}_1$ is the best response of Player I when Player II chooses the strategy $\bar{x}_2$, and $\bar{x}_2$ is the best response of Player II when Player I chooses the strategy $\bar{x}_1$. Let us now consider two examples to illustrate this concept.

**Example 2.3** In a two-person game, suppose that each player can only choose either strategy $A$ or $B$. If both players choose strategy $A$, they both get 4 points. If Player I chooses strategy $A$ and Player II chooses strategy $B$, then Player I gets 1 point and Player II gets 3 points. If Player I chooses strategy $B$ and Player II chooses strategy $A$, then Player I gets 3 points and Player II gets 1 point. If both choose strategy $B$, each player gets 2 points. The payoff function of each player is represented in the payoff matrix below.

$$
\begin{array}{cc}
\text{Player II} \\
\text{Player I} & A & B \\
A & 4 & 3 \\
B & 1 & 2 \\
\end{array}
$$

In this example, Nash equilibrium occurs when both players choose strategy $A$, and it also occurs when both players chooses strategy $B$. Let us consider, for instance, the case
where both players choose strategy $B$. In this case, given that Player II chooses strategy $B$, Player I also wants to keep strategy $B$ because a change of strategy would lead to a reduction of his/her payoff from 2 to 1. Similarly, given that Player I chooses strategy $B$, Player II wants to keep strategy $B$ because a change of strategy would also lead to a reduction of his/her payoff from 2 to 1.

Example 2.4 Let us consider another simple two-person game called *matching pennies*. Suppose that Player I and Player II each has a penny. Each player must secretly turn the penny to heads or tails and then reveal his/her choices simultaneously. Player I wins the game and gets Player II’s penny if both coins show the same face (heads-heads or tails-tails). In the other case where the coins show different faces (heads-tails or tails-heads), Player II wins the game and gets Player I’s penny. The payoff function of each player is represented in the payoff matrix below.

$$
\begin{array}{c|cc}
 & \text{head} & \text{tail} \\
\hline
\text{head} & 1 & -1 & 1 \\
\text{tail} & 1 & -1 & 1 \\
\end{array}
$$

In this game it is not hard to see that no Nash equilibrium exists. Indeed, denote by $A$ the matrix that represents the payoff function of Player I, and denote by $B$ the matrix that represents the payoff function of Player II:

$$
A = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}
$$

Now suppose that Player I is playing with a mind reader who knows Player I’s choice of faces. If Player I decides to turn heads, then Player II knows about it and chooses turn tails to win the game. In the case where Player I decides to turn tails, then Player II chooses to turn heads and again wins the game. Thus to have a fair game, he/she decides to randomize his/her strategy by, for instance, tossing the coin instead of putting the coin down. We describe the new game as follows.

$$
\begin{array}{c|cc}
 & \text{head, $q_1$} & \text{tail, $q_2$} \\
\hline
\text{head, $p_1$} & -1 & 1 \\
\text{tail, $p_2$} & 1 & -1 \\
\end{array}
$$

In this new game, Player I uses a coin randomly with probability of coming up heads $p_1$ and probability of coming tails $p_2$, where $p_1 + p_2 = 1$. Similarly, Player II uses another
coin randomly with probability of coming up heads $q_1$ and probability of coming tails $q_2$, where $q_1 + q_2 = 1$. The new strategies now are called mixed strategies while the original ones are called pure strategies. Now suppose that Player II uses mixed strategy $\{q_1, q_2\}$, then Player I’s expected payoff for playing heads is

$$u_H(q_1, q_2) = q_1 - q_2 = 2q_1 - 1.$$ 

Similarly, Player I’s expected payoff for playing tails is

$$u_T(q_1, q_2) = -q_1 + q_2 = -2q_1 + 1.$$ 

Thus Player I’s expected payoff for playing mixed strategy $\{p_1, p_2\}$ is

$$u_1(p, q) = p_1 u_H(q_1, q_2) + p_2 u_T(q_1, q_2) = p_1(q_1 - q_2) + p_2(-q_1 + q_2) = p^T Aq,$$

where $p = [p_1, p_2]^T$ and $q = [q_1, q_2]$. 

By the same arguments, if Player I chooses mixed strategy $\{p_1, p_2\}$, then Player II’s expected payoff for playing mixed strategy $\{q_1, q_2\}$ is

$$u_2(p, q) = q^T Bp.$$ 

For this new game, an element $(\bar{p}, \bar{q})$ is a Nash equilibrium if

$$u_1(p, \bar{q}) \leq u_1(\bar{p}, \bar{q}) \quad \text{for all} \quad p \in \Delta,$$

$$u_2(\bar{p}, q) \leq u_2(\bar{p}, \bar{q}) \quad \text{for all} \quad q \in \Delta,$$

where $\Delta := \{(p_1, p_2) \mid p_1 \geq 0, p_2 \geq 0, p_1 + p_2 = 1\}$ is a nonempty, compact subset of $\mathbb{R}^2$.

Nash [3, 4] proved the existence of his equilibrium in the class of mixed strategies in the setting of Example 2.4 and in more general setting. His proof was based on Brouwer’s fixed point theorem as well as other, rather involved arguments. Now we present a much simpler proof of the Nash equilibrium theorem that also uses Brouwer’s fixed point theorem while applying in addition just elementary tools of convex analysis and optimization; see, e.g., [7].

**Theorem 2.5** Consider a two-person game $\{\Omega_i, u_i\}$, where $\Omega_1 \subset \mathbb{R}^m$ and $\Omega_2 \subset \mathbb{R}^n$ are nonempty compact sets. Let the payoff functions $u_i : \Omega_1 \times \Omega_2 \to \mathbb{R}$ be given by

$$u_1(p, q) := p^T Aq, \quad u_2(p, q) := q^T Bp,$$

where $A$ and $B$ are an $m \times n$ and $n \times m$ matrix, respectively. Then this game admits a Nash equilibrium in the sense of Definition 2.2.

**Proof.** It follows from the definition that an element $(\bar{p}, \bar{q}) \in \Omega_1 \times \Omega_2$ is a Nash equilibrium of the game under consideration if and only if

$$-u_1(p, \bar{q}) \geq -u_1(\bar{p}, \bar{q}) \quad \text{for all} \quad p \in \Omega_1,$$

$$-u_2(\bar{p}, q) \geq -u_2(\bar{p}, \bar{q}) \quad \text{for all} \quad q \in \Omega_2.$$
It follows from Proposition 1.5 that this holds if and only if we have the normal cone inclusions
\[ \nabla_p u_1(\bar{p}, \bar{q}) \in N(\bar{p}; \Omega_1), \quad \nabla_q u_2(\bar{p}, \bar{q}) \in N(\bar{q}; \Omega_2). \] (2.6)

By using the structures of \( u_i \), conditions (2.6) can be equivalently expressed as
\[ A\bar{q} \in N(\bar{p}; \Omega_1) \quad \text{and} \quad B\bar{p} \in N(\bar{q}; \Omega_2), \]
or \( (A\bar{q}, B\bar{p}) \in N(\bar{p}; \Omega_1) \times N(\bar{q}; \Omega_2) = N((\bar{p}, \bar{q}); \Omega_1 \times \Omega_2) = N((\bar{p}, \bar{q}); \Omega) \). By Proposition 1.3

The Euclidean projection, this is equivalent to
\[ (\bar{p}, \bar{q}) = \mathcal{P}((\bar{p}, \bar{q}) + (A\bar{q}, B\bar{p}); \Omega). \] (2.7)

Defining now the mapping \( F: \Omega \to \Omega \) by
\[
F(p, q) := \mathcal{P}((p, q) + (Aq, Bp); \Omega) \quad \text{with} \quad \Omega = \Omega_1 \times \Omega_2
\]
and employing another elementary projection property from Proposition 1.4 allow us to conclude that the mapping \( F \) is continuous, while the set \( \Omega \) is compact. By the classical Brouwer fixed point theorem (see, e.g., [9]), the mapping \( F \) has a \textit{fixed point} \( (\bar{p}, \bar{q}) \in \Omega \), which satisfies (2.7), and thus it is a Nash equilibrium of the game. \( \square \)

References


