Proof of the Minimax Theorem

Michael A. Goodrich

September 17, 2007

The proof of the minimax theorem follows the format given in Luce and Raiffa [2], but has been modified to use our class terminology.

Changes in the tutorial since it was first posted in Fall 2007 are highlighted in red.

Formalisms

We first present the properties of a two-person, zero-sum game. To help you through these properties, I will group the properties together in a way that makes sense to me. I hope it will help you.

Formalization of a 2 Person Zero-Sum Game

1. There are two players, P1 and P2.
2. P1 has a set $A = \{a_1, a_2, \ldots, a_m\}$ of m pure strategies (or actions).
3. P2 has a set $B = \{b_1, b_2, \ldots, b_n\}$ of n pure strategies (or actions).
4. Each player has a utility for each (a_i, b_j) pair of actions. The utility for P1 is denoted $U_1(a_i, b_j)$ and the utility for P2 is denoted $U_2(a_i, b_j)$. Since this is a zero-sum game, $U_1(a_i, b_j) = -U_2(a_i, b_j)$ for all i and j. To minimize the number of subscripts we will carry around, let $M(a_i, b_j) = U_1(a_i, b_j)$ denote the *mutual* utility for the game.

In essence, these properties state the number of players, the set of possible strategies available to each player, and the payoffs that occur when the game is played using pure strategies.

Payoffs for Mixed Strategies

The next properties identify what happens when players use mixed strategies.

5. Each player can use a mixed strategy by creating a probability mass function and playing each pure strategy with a fixed probability. Let p_i denote the probability that player 1 will play action a_i, and let q_j denote the probability that player 2 will play action b_j. Since p and q are probabilities, they must satisfy
   (a) $\forall i \;\; p_i \geq 0, \forall j \;\; q_j \geq 0$.

(b) $\sum_{i=1}^{m} p_i = 1$, $\sum_{j=1}^{n} q_j = 1$.

A mixed strategy that uses a particular probability mass function is denoted $\mathbf{p} = (p_1, p_2, \ldots, p_m)$ where $p_i = Pr(a_i)$ is the probability that action a_i will be played; similarly, for player 2 $\mathbf{q} = (q_1, q_2, \ldots, q_n)$.

6. For each randomized strategy pair $(\mathbf{p}, \mathbf{q})$, the payoff $M(\mathbf{p}, \mathbf{q})$ is defined to be

$$M(\mathbf{p}, \mathbf{q}) = \sum_{i=1}^{m} \sum_{j=1}^{n} p_i M(a_i, b_j) q_j.$$

We denote the payoff when player 1 uses pure strategy a_i and player 2 uses mixed strategy $\mathbf{q}$ as

$$M(a_i, \mathbf{q}) = \sum_{j=1}^{n} M(a_i, b_j) q_j$$

with similar notation for $M(\mathbf{p}, b_j)$.

7. In much the same way that A and B denote the set of *pure* strategies available to players 1 and 2, respectively, we use P and Q to denote the set of all *mixed* strategies available to players 1 and 2, respectively.

Maximum Security

Now that we have a language for expressing expected payoff for mixed strategies, we can select which of the possible mixed strategies are minimax or maximin strategies. We do this with the introduction of the notions of *security* and *regret*. In this section, we talk about maximizing security, and in the next section we discuss minimizing regret.

8. Player 1's objective is to select a randomized strategy $\mathbf{p}$ from P so as to maximize $M(\mathbf{p}, \mathbf{q})$. At the same time, player 2's objective is to select a randomized strategy $\mathbf{q}$ from Q so as to maximize its payoff, which is equivalent to minimizing $M(\mathbf{p}, \mathbf{q})$. The rules of the game require that each player choose its strategy in complete ignorance of the opponent's selection.

9. For each mixed strategy $\mathbf{p}$ belonging to P, player 1's security level is defined to be

$$v_1(\mathbf{p}) = \min_{\mathbf{q}} M(\mathbf{p}, \mathbf{q}).$$

Since

$$M(\mathbf{p}, \mathbf{q}) = \sum_{j=1}^{n} M(\mathbf{p}, b_j) q_j$$

is a weighted sum of the n payoffs $M(\mathbf{p}, b_j)$, it is minimized when all of the weight is assigned to the least of these (do you see why? look at how we compute the maximin mixed strategy in the lecture notes.)

$$v_1(\mathbf{p}) = \min[M(\mathbf{p}, b_1), M(\mathbf{p}, b_2), \ldots, M(\mathbf{p}, b_n)].$$

You can think of $v_1(\mathbf{p})$ as the payoff that player 1 will receive if player 2 knows that P1 will do $\mathbf{p}$. (Why? Because if player 2 knows this, then it can choose its best response.) We can define $v_2(\mathbf{q})$ for player 2 in a similar way (but using maximums instead of minimums since high M means low payoff to player 2).

10. By assumption, player 1 wants to maximize its security level, so P1 must choose a strategy $\mathbf{p}^*$ such that
$$v_1(\mathbf{p}^*) \geq v_1(\mathbf{p}) \forall \mathbf{p} \in P.$$
Let v_1 denote this maximal security level (i.e., $v_1 = v_1(\mathbf{p}^*)$). Then
$$v_1 = \max_{\mathbf{p}} v_1(\mathbf{p}) \geq v_1(\mathbf{p}) \tag{1}$$
for all other mixed strategies. We also know that
$$v_1 = \min_{\mathbf{q}} M(\mathbf{p}^*, \mathbf{q}) \leq M(\mathbf{p}^*, \mathbf{q}) \forall \mathbf{q} \in Q. \tag{2}$$
Inequality (1) means that the strategy that produces v_1 is superior to all other strategies (in terms of maximizing security level). Inequality (2) means that v_1 is the worst (minimum payoff) that player 1 can expect given that player 1 plays p^*. If player 2 doesn't choose wisely then player 1 will get more than v_1. If player 1 doesn't play p^* then player 1 can receive a lower payoff. The strategy $\mathbf{p}^*$ is called the *maximin* strategy.

Minimum Regret

In this section, we repeat the analysis from the previous section but apply it to the minimizing player (player 2).

11. Because the game is zero-sum, we know that when player 2 maximizes its security level then it minimizes player 1's payoff. If player 2 uses strategy $\mathbf{q}$, player 1 cannot obtain a return greater than
$$v_2(\mathbf{q}) = \max_{\mathbf{p}} M(\mathbf{p}, \mathbf{q}).$$
The value v_2 is sometimes called *regret*, which kind of indicates that it is the negation of *security*. Just like player 1 tries to maximizes security, player 2 tries to minimize regret. Define
$$\mathbf{q}^* = \arg\min_{\mathbf{q}} v_2(\mathbf{q}),$$
and define
$$v_2 = v_2(\mathbf{q}^*) \leq v_2(\mathbf{q}) \quad \forall \mathbf{q} \in Q. \tag{3}$$
By repeating the analysis that we did for player 1 but with player 2 in mind, we learn that
$$v_2 \geq M(\mathbf{p}, \mathbf{q}^*) \ \ \forall \mathbf{p} \in P. \tag{4}$$
The strategy $\mathbf{q}^*$ is called the *minimax* strategy.

Relation Between Minimax and Maximin Values

As you may have guessed, there is a relationship between the minimax and maximin values for a zero-sum game. In this section, we develop this relationship.

12. Putting these pieces together, we learn that if player 1 uses the maximin strategy, it is guaranteed at least v_1 units of (security) payoff

    $$v_1(\mathbf{p}) \leq v_1 \leq M(\mathbf{p}^*, \mathbf{q})\ \forall \mathbf{q} \in Q.$$

    (Note that $v_1(\mathbf{p})$ is a function of $\mathbf{p}$ which means that security is a function of what player 1 plans on doing.) Similarly, if player 2 uses the minimax strategy, it is guaranteed no more than v_2 units of (regret) loss, which is tantamount to guaranteeing that player 1 can receive no more than v_2 units of payoff

    $$M(\mathbf{p}, \mathbf{q}^*) \leq v_2 \leq v_2(\mathbf{q})\ \forall \mathbf{q} \in Q.$$

    (Again, note that $v_2(\mathbf{q})$ is a function of $\mathbf{q}$ which means that regret is a function of what player 2 plans on doing.) Thus,

    $$\begin{aligned} M(\mathbf{p}, \mathbf{q}^*) &\leq v_2\ \forall \mathbf{p} \in P \\ M(\mathbf{p}^*, \mathbf{q}^*) &\leq v_2 \\ v_1 &\leq M(\mathbf{p}^*, \mathbf{q})\ \forall \mathbf{q} \in Q \\ v_1 &\leq M(\mathbf{p}^*, \mathbf{q}^*) \\ v_1 &\leq v_2. \end{aligned} \tag{5}$$

Equilibrium Value

Now that we have established a relationship between the minimax and maximin value, we can formally define what an (Nash) equilibrium value for a game is.

13. A pair of strategies $(\mathbf{p}', \mathbf{q}')$ is said to be in equilibrium if $\mathbf{p}'$ is *good against* $\mathbf{q}'$ and vice versa, meaning

    $$M(\mathbf{p}, \mathbf{q}') \leq M(\mathbf{p}', \mathbf{q}') \leq M(\mathbf{p}', \mathbf{q}).$$

    In words, these two strategies are in equilibrium if neither player has an incentive to change (can increase its payoff by unilaterally changing its behavior). To help understand this, it is useful to recall that a Nash equilibrium is a solution pair such that no player has an incentive to unilaterally change his or her action.

Whew! That's quite a bit of information, but its nothing more than a formalism of the concepts we've been talking about informally. Since one of my goals for this course is to help you get confident about reading the literature, I want you to practice putting your thoughts into a concise, mathematical language.

A Big Picture Moment

It is sometimes useful when wrestling with concepts like these to take a moment and get the big picture. These notes aim at making clear three things: the formalization of a two person zero-sum game, the proof of the minimax theorem for two person games, and the relationship between minimax values and equilibrium values. You have just finished slogging your way through the formalization of a two person zero-sum game.

Before proceeding to the other two objectives, it is important that you understand the following:

- The minimax theorem was first proved by John von Neumann (I'll try to find the reference when I get a chance). This theorem states that for zero sum games there exists a unique value for this game. This value is the minimax value for the minimizing agent and the maximin value for the maximizing agent.
- It just so happens that the pair of solutions that generate the maximin value and the minimax value for this game form a Nash equilibrium. This holds because the Nash equilibrium corresponds to the minimax/maximin solution pair for zero-sum games. Thus, for two-person zero-sum games, the presence of a unique value for a game implies that a Nash equilibrium exists for that game.
- John Nash [1] showed that every game, not just two-person zero-sum games, has at least one equilibrium value. Since the minimax/maximin solution pair corresponds to this equilibrium, we will use a variant of Nash's more general proof as the basis for the proof of the minimax theorem.
- For non-zero sum games, the joint solution produced by players individually choosing a minimax strategy does not necessarily correspond to an equilibrium solution — consider the Battle of the Sexes game.

A Useful Theorem

We now turn to an interesting theorem that applies only to zero-sum games. The theorem states that (a) if an equilibrium exists then the maximin value v_1 equals the minimax value v_2, (b) if $v_1 = v_2$ then there exists a real number v and a pair of strategies $\mathbf{p}^*$ and $\mathbf{q}^*$ such that the payoffs for these strategies are bounded by v, and (c) if the conditions just mentioned hold then the game has an equilibrium. The theorem does not state that each zero-sum, two-person game satisfies any of these conditions; it only says that if it does then the maximin and minimax solutions are equilibrium solutions. Although we will state the theorem, we will omit the proof of this theorem (it follows almost immediately from our problem specification above) so that we can concentrate on the proof of the minimax theorem. In essence, this theorem states that we can use Nash's proof of the existence of an equilibrium in general-sum games to show that the minimax solution pair exists and is in equilibrium for zero-sum games.

The great thing about this theorem is that it establishes an equivalence between the minimax value, an equilibrium pair, and a concept called the *value* of the game. Can you think of a way to use the value of the game to create an algorithm that finds the equilibrium solution?

I should point out that many useful theorems have a form similar to this one; they show that several different conditions are equivalent. These theorems are very serviceable because they allow us to tie a handful of mixed ideas into an equivalence. The proof of these theorems usually follows a procedure wherein we show that condition 1 implies condition 2, condition 2 implies condition 3, and so on until we can show that condition ℓ implies condition 1.

Theorem 1 *For two-person, zero-sum games as we have presented them, each of the following three conditions implies the other two.*

1. *An equilibrium pair exists.*
2. $$v_1 = \max_{\mathbf{p}} \min_{\mathbf{q}} M(\mathbf{p}, \mathbf{q}) = \min_{\mathbf{q}} \max_{\mathbf{p}} M(\mathbf{p}, \mathbf{q}) = v_2.$$
3. *There exists a real number v, a mixed strategy $\mathbf{p}^*$, and a mixed strategy $\mathbf{q}^*$ such that*
    - *(a)* $\sum_i M(a_i, b_j)p_i^* \geq v$ *for* $j = 1, 2, \ldots, n$
    - *(b)* $\sum_j M(a_i, b_j)q_j^* \leq v$ *for* $i = 1, 2, \ldots, m$.

Note that condition 3(3a) says that the average loss for player 2 using any pure strategy is no less than v. Similarly, condition 3(3b) says that the average payoff for player 1 using any pure strategy is no greater than v. This value v is known as the value of a game in normal form.

The Minimax Theorem

We now present Nash's proof of the Minimax Theorem. Although there are several proofs of the theorem, Nash's is fairly easy to understand. Additionally, although we only give the proof for the specification of a zero-sum game given above, the proof extends to N-person general-sum games as well. This means that every general-sum, N-person game has at least one Nash equilibrium.

A simple interpretation of the steps in the proof goes something like this:

- We are going to make up a transformation that maps mixed strategies to new mixed strategies.
- We are going to show that this transformation is special: if the underlying game has an equilibrium then this transformation has a fixed point, and if the transformation has a fixed point only then the game has an equilibrium.
- Since the transformation is continuous and since probability spaces are "topologically close" to a multi-dimensional sphere (we'll discuss this in class), we know by Brouwer's fixed point theorem that the transformation has a fixed point.
- Since the fixed point is, in essence, indistinguishable from the equilibrium point, it follows that the game has an equilibrium point.

Theorem 2 *For every two-person, zero-sum game, there exists an equilibrium strategy.*

Proof: Consider a transformation T that maps mixed strategy pairs $(\mathbf{p}, \mathbf{q})$ into mixed strategy pairs $T(\mathbf{p}, \mathbf{q}) = (\mathbf{p}', \mathbf{q}')$. What we'll show is that this transformation T has the following two properties:

1. $\mathbf{p}^*$ and $\mathbf{q}^*$ are optimal (i.e., maximin and minimax) strategies if and only if $T(\mathbf{p}^*, \mathbf{q}^*) = (\mathbf{p}^*, \mathbf{q}^*)$. (Any point which is mapped to itself under a transformation is called a *fixed point* of this transformation. For example, consider the transformation $T : \Re^+ \mapsto \Re^+$ defined as $T(x) = x^2$. The value of $x = 1$ is a fixed point since $T(1) = 1^2 = 1$.)
2. T, defined below, has at least one fixed point.

In essence, $c_i(\mathbf{p},\mathbf{q})$ represents the improvement to player 1 for switching from strategy $\mathbf{p}$ to strategy a_i. Similarly, $d_j(\mathbf{p},\mathbf{q})$ represents the improvement to player 2 for switching from strategy $\mathbf{q}$ to strategy b_j.

We will now construct a transformation that satisfies our two properties. To do this, define T as follows. Let

$$\begin{aligned} c_i(\mathbf{p},\mathbf{q}) &= \begin{cases} M(a_i,\mathbf{q}) - M(\mathbf{p},\mathbf{q}) & \text{if } M(a_i,\mathbf{q}) - M(\mathbf{p},\mathbf{q}) > 0 \\ 0 & \text{otherwise;} \end{cases} \\ d_j(\mathbf{p},\mathbf{q}) &= \begin{cases} M(\mathbf{p},\mathbf{q}) - M(\mathbf{p},b_j) & \text{if } M(\mathbf{p},\mathbf{q}) - M(\mathbf{p},b_j) > 0 \\ 0 & \text{otherwise;} \end{cases} \end{aligned}$$

Using the notation $T(\mathbf{p},\mathbf{q}) = (\mathbf{p}',\mathbf{q}')$, we define

$$\begin{aligned} p_i' &= \frac{p_i + c_i(\mathbf{p},\mathbf{q})}{1+\sum_{k=1}^m c_k(\mathbf{p},\mathbf{q})} \\ q_i' &= \frac{q_i + d_i(\mathbf{p},\mathbf{q})}{1+\sum_{k=1}^n d_k(\mathbf{p},\mathbf{q})}. \end{aligned}$$

Before proceeding, we need to check and see if this is indeed a transformation from a probability to another probability. Thus, we want to show that $T : P \times Q \mapsto P \times Q$ (i.e. that the things produced by the transformations are probabilities), so we need to show that $\mathbf{p}'$ and $\mathbf{q}'$ are probability mass functions. Clearly, since c_i and d_i are both positive then p_i' and q_i' are also both nonnegative. (Recall that probabilities cannot be negative.) The next requirement is that these probabilities sum to one. Since

$$\begin{aligned} \sum_{i=1}^m p_i' &= \sum_{i=1}^m \frac{p_i + c_i(\mathbf{p},\mathbf{q})}{1+\sum_{k=1}^m c_k(\mathbf{p},\mathbf{q})} \\ &= \frac{\sum_{i=1}^m p_i + \sum_{i=1}^m c_i(\mathbf{p},\mathbf{q})}{1+\sum_{k=1}^m c_k(\mathbf{p},\mathbf{q})} \\ &= \frac{1 + \sum_{i=1}^m c_i(\mathbf{p},\mathbf{q})}{1+\sum_{k=1}^m c_k(\mathbf{p},\mathbf{q})} \\ &= 1. \end{aligned}$$

Now that we know that the transformation is valid, we want to start the proof. The first step is the ”If” step. When you look carefully at the statement of the first property of the transformation, you’ll see an *if and only if A then B* statement. Most of you remember this, but just in case you don’t, the way you establish such a statement is first showing the *if A then B* part and then showing the *if B then A* part. We’ll start by showing that if $\mathbf{p}^*$ and $\mathbf{q}^*$ are optimal then $T(\mathbf{p}^*,\mathbf{q}^*) = (\mathbf{p}^*,\mathbf{q}^*)$. Observe that $c_i(\mathbf{p},\mathbf{q})$ is a measurement of the amount that a_i is better than $\mathbf{p}$ (if at all) as a response against $\mathbf{q}$. Similarly, $d_i(\mathbf{p},\mathbf{q})$ is a measurement of the amount that b_i is better than $\mathbf{q}$ (if at all) as a response against $\mathbf{p}$. When $\mathbf{p}^*$ and $\mathbf{q}^*$ are optimal, it follows that $c_i(\mathbf{p}^*,\mathbf{q}^*) = 0$ for all i (can you see why?) so $p_i^* = p_i$ for all i. Similary, $q_i^* = q_i$. Thus, $T(\mathbf{p}^*,\mathbf{q}^*) = (\mathbf{p}^*,\mathbf{q}^*)$.

Turning to the *only if* portion of the proof, suppose that $(\mathbf{p},\mathbf{q})$ is a fixed point. We need to show that $(\mathbf{p},\mathbf{q})$ is optimal. We first show that there exists an i such that both $p_i > 0$ and $c_i(\mathbf{p},\mathbf{q}) = 0$. Since, by definition (in class, ask me about how convex combinations can be graphically depicted),

$$M(\mathbf{p},\mathbf{q}) = \sum_{i=1}^m p_i M(a_i,\mathbf{q})$$

we conclude that $M(\mathbf{p},\mathbf{q}) < M(a_i,\mathbf{q})$ cannot be true for all i such that $p_i > 0$. To see this, we'll do a mini-proof within this proof. Suppose that, $\forall i$ such that $p_i > 0$, $M(\mathbf{p},\mathbf{q}) < M(a_i,\mathbf{q})$. Then

$$\begin{aligned} M(\mathbf{p},\mathbf{q}) &= \sum_{i=1}^{m} p_i M(a_i,\mathbf{q}) \\ &> \sum_{i=1}^{m} p_i M(\mathbf{p},\mathbf{q}) \\ &= M(\mathbf{p},\mathbf{q}) \sum_{i=1}^{m} p_i \\ &= M(\mathbf{p},\mathbf{q}). \end{aligned}$$

This is a contradiction, which means that for at least one $p_i > 0$ it must follow that $M(\mathbf{p},\mathbf{q}) \geq M(a_i,\mathbf{q})$. This ends our mini-proof. We will now use this to show that fixed points are optimal points.

But this implies that, for this i, $0 \geq M(a_i,\mathbf{q}) - M(\mathbf{p},\mathbf{q})$ so, by definition of c_i, $c_i(\mathbf{p},\mathbf{q}) = 0$ for this i. For this i, the fact that $(\mathbf{p},\mathbf{q})$ is a fixed point implies that

$$p_i = \frac{p_i + 0}{1 + \sum_{k=1}^{m} c_k(\mathbf{p},\mathbf{q})}.$$

Since $p_i > 0$ for this i, it follows that $\sum_{k=1}^{m} c_k(\mathbf{p},\mathbf{q}) = 0$. But the terms c_k are all non-negative (by definition), so they must all equal 0. This, in turn, means that $M(\mathbf{p},\mathbf{q}) \geq M(a_i,\mathbf{q})$ for all a_i. Since this is true regardless of $\mathbf{q}$, it follows that no other pure or mixed strategy has higher payoff for all $\mathbf{q}$. Thus, $M(\mathbf{p},\mathbf{q}) \geq M(\mathbf{p}',\mathbf{q})$ for all $\mathbf{p}'$ and for all $\mathbf{q}$ so $\mathbf{p}$ is good against all $\mathbf{q}$.

Similarly, we can show that $\mathbf{q}$ is good against all $\mathbf{p}$, so when $(\mathbf{p},\mathbf{q})$ is a fixed point of the transformation T then it is also optimal.

The only thing we have to do now to complete the proof is show that the transformation T has a fixed point. This existence follows from the Brouwer fixed-point theorem. We won't show how the fixed point theorem applies, but you might want to think about it a little bit. I'll paraphrase Luce and Raiffa's statement of the theorem. The theorem says that *any continuous transformation that maps a point of a spheroid (or something topologically "close" to a spheroid) in a finite dimensional Euclidean space into another point of the sphereoid has at least one fixed point.* The space $P \times Q$ is topologically "close" to a spheroid and our transformation is continuous, so we know that a fixed point exists. Any optimal point is a fixed point, and any fixed point is an optimal point so we know that every zero-sum, two-player game has a mixed strategy equilibrium point.

Yippee! We did it.

References

[1] Jr. J. F. Nash. The bargaining problem. *Econometrica*, 18:155–162, 1950. Reprinted in *Classics in Game Theory*, H. W. Kuhn, ed.

[2] R. D. Luce and H. Raiffa. *Games and Decisions*. John Wiley, New York, 1957.