Generalization of Von-Neumann’s Minimax Theorem for Compact Metric Spaces

Atul Shekhar
Chennai Mathematical Institute

September 14, 2009

Abstract

One of the famous results in Game theory is Von Neumann’s Minimax Theorem which says that a two person zero sum finite game has a value and each player has optimal strategies.A similar result holds for games on compact metric spaces also.Here we give a proof of this result which uses some results from Analysis and Measure Theory like Riesz representation theorem and Hahn-Banach separation Theorem.

Von Neumann’s first significant contribution to economics was the minimax theorem of 1928 saying about existence of value for two person zero sum finite game and optimal strategies for the two players.The same result holds for games on compact metric spaces also. This means that given a compact metric space, suppose player1 chooses a point x and player2 chooses a point y and the payoff is given to be A(x,y)(assume it to be amount of money player2 gives to player1) which is continuous.then there exist a real number v and strategies for both the players such that on an average player1 gets at least v and player2 doesn’t give more than v.This v is called value of the Game. I will state the result formally.

**Theorem***. Let X be a compact metric space and let A(x,y) be a continuous function A:X × X → R, then for probability measures P1 and P2 on the borel measure space (X,BX),where BX is the borel σ-algebra associated with X, both \( \min_{P_1} \max_{P_2} \int A(x,y) dP_1 dP_2 \) and \( \max_{P_2} \min_{P_1} \int A(x,y) dP_1 dP_2 \) exists (maximisation is over P1 which integrates with respect to first coordinate and minimisation is over P2 which integrates with respect to second coordinate),and both of them are equal.

The proof of this theorem uses a lemma.

**Lemma1.** Consider the space \( C(X) \) of all the continuous functions from X to R under the sup norm defined by \( ||f|| = \sup_{x \in X} |f(x)| \).

Now for each probability measure P on (X,BX) we get a continuous function \( f_P \) X to R defined by \( f_P(x) = \int A(x,y) dP \). (where P integrates with respect to second coordinate).(continuity of \( f_P \) can be proved easily).

Then set \( S = \{ f_P | P \) is a probability measure on (X,BX) \} is a compact subset of \( C(X) \).

**Proof.** I will first explain a theorem used to prove this lemma. Given a norm linear space N we have the norm linear space \( N^* \) of all the continuous linear maps from N to R and \( N^{**} \) of all the continuous linear maps from \( N^* \) to R. Given any x in N it induces a function \( F_x \) from \( N^* \) to R defined by \( F_x(f) = f(x) \).The weak* topology on set \( N^* \) is defined to be the weakest topology such that all these induced functions \( F_x \) becomes continuous \( \forall x \in N \).Also the subbasis for this topology is the collection of sets \( S(f_0,\epsilon) \) as \( x,f_0,\epsilon \) varies where \( x \in N,f_0 \in N^*,\epsilon > 0 \) and \( S(f_0,\epsilon) \) is defined to be the set \( \{ f \in N^* | |f(x) - f_0(x)| < \epsilon \} \).Now there is a famous result saying.

**Theorem1**. The closed unit ball B in \( N^* \) defined by \( B = \{ f | ||f|| \leq 1 \} \) is a compact hausdorff topological space in weak* topology on it.

But how does this say anything about probability measures which we want to deal with. This connection is provided by Riesz Representation Theorem.

**Theorem2.** Consider the set M of all the finite signed measures on (X,BX). Then \( \forall \mu \in M \) the map \( T_\mu \) defined by \( T_\mu(f) = \int f d\mu \) is a continuous linear map from \( C(X) \) to R.Also, given a continuous linear map T from \( C(X) \) to R there exist a finite signed measure \( \mu \in M \) such that \( T = T_\mu \). Also , \( ||\mu|| = ||T_\mu|| \), where \( ||\mu|| = |\mu|(X) \), where \( |\mu| \) the total variation measure of \( \mu \).

So in Theorem1, if we set \( N = C(X) \), then by using Theorem2 we conclude that

In \( M \) set \( B = \{ \mu \ | ||\mu|| \leq 1 \} \) is a compact hausdorff space in weak* topology on it.( induced from \( C(X)^* \) through the bijection \( \mu \leftrightarrow T_\mu \)).
Now for proving the required Lemma 1 we want to prove first that the set of probability measures in $M$ in weak* topology on it is compact. Since probability measures have norm 1, so set of probability measures is subset of closed unit ball $B$ which is compact itself. So for proving the compactness, we need to show that the set of probability measures is closed in $B$.

So consider the space $B$ under the weak* topology on it induced from $M$ as a subspace of it, which is compact hausdorff. Now the subspace for this topology on $B$ is given by the collection $S(f, \mu_0, \epsilon)$ as $f$ varies over $C(X), \mu_0$ varies over $M$ and $\epsilon > 0$, where $S(f, \mu_0, \epsilon) = \{ \mu: \mu \in B \mid \| f \| \mu - \int f d\mu_0 \| < \epsilon \}$.

An another property of this $B$ under weak* topology is used in proving the closedness of probability measures in $B$.

**Lemma.** The space $B$ in weak* topology is first countable. (i.e. given any point $\mu \in B$, we can find a countable collection of open sets around $\mu$ such that any open set around $\mu$ contains one of the open sets from that countable collection).

**Proof.** As a consequence of Stone-weirstrass Theorem, we know that the $C(X)$ is separable i.e. it has a countable dense subset. Let $\{ f_i \}$ for $i = 1, 2, \ldots$ be the countable dense subset of $C(X)$.

Now given any $\mu_0 \in B$ we have to exhibit a countable collection of open sets around $\mu_0$ such that any open set around $\mu_0$ will contain one from that collection.

So consider $S(f_i, \mu_0, \frac{1}{n}) \forall i = 1, 2, n = 1, 2, \ldots$ where $S(f_i, \mu_0, \frac{1}{n})$ is as defined earlier. Now take all possible finite intersection of elements from this collections to make a new collections of open sets in $B$. Clearly each of set in this new collection contain $\mu_0$. Also, since countable union of countable sets is countable and number of finite subsets of a countable set is countable, the new collection is countable too.

Now, the claim is that this collection does the work. Since any open set is union of basic open sets, so given any basic open set around $\mu_0$ we need to give an element from this collection contained in that basic open set. Now since each basic open set is finite intersection of some $S(f, \mu, \epsilon)$’s, so if we prove that whenever $\mu_0 \in S(f, \mu, \epsilon)$ there exist $i$ and $n$ such that $S(f_i, \mu_0, \frac{1}{n}) \subset S(f, \mu, \epsilon)$, then we are done.

Now given $\mu_0 \in S(f, \mu, \epsilon)$, so $\| f \| \mu_0 - \| f \| \mu < \epsilon$, and then there exist large enough $k$ such that $\| f \| \mu_0 - \| f \| \mu < \frac{\epsilon}{k}$. Now choose $\delta = \frac{1}{k}$ and choose a $f_n$ such that $\| f_n - f \| < \delta$. (We can choose such $f_n$ because $\{ f_i \}$ for $i = 1, 2, \ldots$ is a countable dense subset of $C(X)$).

Now we show that $S(f_n, \mu_0, \frac{1}{kn}) \subset S(f, \mu, \epsilon)$.

Now consider any $\mu' \in S(f_n, \mu_0, \frac{1}{kn})$, then $\| f_n d\mu' - f_n d\mu_0 \| < \frac{\epsilon}{kn}$. So choose any $\mu \in S(f, \mu, \epsilon)$.

Now $\int f_n d\mu' - \int f_n d\mu_0 = | \int (f_n - f) d\mu' | \leq \int | f_n - f | d\mu' = \int \delta d\mu' = \delta | \mu' | (X) \leq \delta, (| \mu' |$ is total variation of $\mu'$), and $| \mu' | (X) \leq 1$ as $\mu \in B$.

And also $\| f_n d\mu_0 - f d\mu_0 \| = \| (f_n - f) d\mu_0 \| \leq \| f_n - f \| d\mu_0 \leq \delta | d\mu_0 | \leq \delta | \mu_0 | (X) \leq \delta, (| \mu_0 |$ is total variation of $\mu_0$, and $| \mu_0 | (X) \leq 1$ as $\mu_0 \in B$).

Now $\| f d\mu' - f d\mu_0 \| = | \int f d\mu' - \int (f_n - f) d\mu_0 | \leq \| f_n - f \| d\mu_0 = \delta | d\mu_0 | \leq \delta + \frac{\epsilon}{kn} = 2\delta + \frac{\epsilon}{kn}$.

And $\| f d\mu_0 - f d\mu_0 \| \leq \| f d\mu - f d\mu_0 \| + \| f d\mu_0 - f d\mu_0 \| \leq \| f d\mu' - f n d\mu_0 \| + \| f n d\mu_0 - f d\mu_0 \| \leq \delta + \frac{\epsilon}{kn} = 2\delta + \frac{\epsilon}{kn}$. (as $\delta = \frac{1}{k}$).

So this implies that $\mu' \in S(f, \mu, \epsilon)$. So $S(f_n, \mu_0, \frac{1}{kn}) \subset S(f, \mu, \epsilon)$.

So this completes the the proof of the fact that $B$ in weak* topology is first countable.

Now we go on proving the closedness of probability measures in $B$.

Since $B$ is first countable, so proving closedness of probability measures in $B$ is equivalent of proving that if $P_n$ is a sequence of probability measures converging to $P$ in $B$ then $P$ is also a probability measure.

So consider a sequence of probability measures $P_n$ converging to $P$ in $B$. So any open set around $P$ will contain all but finitely many $P_n$. So consider $S(1_X, P, \epsilon), (1_E$ is the indicator function of $E)$ is obviously continuous. So there exist a $P_n \in S(1_X, P, \epsilon). So, | \int 1_X dP_n - \int 1_X dP | < \epsilon \Rightarrow | P_n(X) - P(X) | < \epsilon \Rightarrow | P(X) - 1 | < \epsilon$

So choosing $\epsilon$ small enough, we get that $P(X) = 1$. 

2
We now again need a lemma.

**Lemma 1.** If $P$ is finite signed measure on $(X, B_X)$ such that $P(E) < 0$ for some $E \in B_X$. Then there exist a continous function from $X$ to $R$ such that $f(x) \geq 0 \forall x \in X$ and $\int f dP < 0$.

**Proof.** Assume the contrary that for all continuous $f$ $X$ to $R$ with $f \geq 0$, we have $\int f dP \geq 0$. Now since $P(E) < 0$, so we have $\int 1_E dP < 0$. Clearly we can choose an $\epsilon > 0$ such that $\int 1_E dP < -\epsilon$. So by using the assumption we get that $|\int f dP - \int 1_E dP| > \epsilon \Rightarrow |\int (f - 1_E) dP| > \epsilon \Rightarrow |\int |f - 1_E| dP| > \epsilon$ (Let $\mu = |P|$)

Now we will construct a continuous $f$ $\geq 0$ such that $\int |f - 1_E| d\mu < \epsilon$, and that will give a contradiction and hence prove the lemma.

For this, since $\mu$ is a finite positive measure and $X$ is compact, so by regularity of measure there exist open set $U$ and compact set $K$ with $K \subset E \subset U$ such that $\mu(U \setminus K) < \epsilon$. Now $K$ and $U^c$ are disjoint closed set in $X$, so by Urysohn’s lemma, we get continuous function $f$ from $X$ to $[0, 1]$ such that $f(x) = 1 \forall x \in K$ and $f(x) = 0 \forall x \in U^c$. We note that $f(x) = 1_E(x)(\forall x \in K \cup U^c$ and $0 \leq f, 1 - f \leq 1$.

So, this proves the lemma.

Now, we had that probability measures $P_n$ is converging to $P$. So, suppose that $P$ is not a probability measure. Since we have proved $P(X) = 1$, so there exist an $E \in B_X$ such that $P(E) < 0$. So, by above lemma we get a continuous function $f \geq 0$ and $\int f dP < 0$.

Now we can find a $\epsilon > 0$ such that $\int f dP < -\epsilon$. Also since $P_n$ are probability measure and $f \geq 0$, $\int f dP_n > 0$ $\Rightarrow |\int f dP_n - \int f dP| > \epsilon$. This gives that no $P_n$ can belong to $S(f, P, \epsilon)$. This gives a contradiction because $P_n$ is converging to $P$ (by def. of convergence).

So $P$ is a probability measure.

So, we have proved that set of probability measures is closed in $B$ and hence a compact space.

Now we go back to proving compactness of $S$ in $(C(X)$ defined in Lemma 1.

Since $(C(X)$ is a metric space, proving compactness is equivalent of proving sequential compactness. So, for proving compactness of $S$, we need to show that given any sequence $f_{P_n}$ in $S$, it has a convergent subsequence converging in $S$ itself.

So consider $f_{P_n}$ in $S$. Now sequence of corresponding $P_n$ is a sequence of probability measures. Now the set of probability measures in $B$, we proved to be compact. Also, since $B$ was first countable, so the subspace of first countable is first countable, so the set of probability measures is compact and first countable in the subspace topology induced from $B$. Now we know that compactness implies limit point compactness (i.e., any infinite subset of a compact space has a limit point in that space) and by using first countability we get that any sequence has a convergent subsequence. So sequence $P_n$ has a convergent subsequence $P_{n_k}$ going to probability measure say $P$.

Now the claim is that this $f_{P_{n_k}}$ again has a subsequence which converges uniformly to $f_P$. Suppose we prove this, then actually we have got a subsequence of $f_{P_{n_k}}$, so actually a subsequence of $f_{P_{n_k}}$ which converges uniformly to $f_P$, that means in $S$ itself, which in turn implies sequential compactness and so compactness, because $S$ is a metric space. For simplicity, as we have to look at only $P_{n_k}$ from now, we drop the index $k$.

Now given that $P_n$ converges to $P$ in $B$ (where $P_n$ and $P$ are probability measures), we have to prove that $f_{P_n}$ has a subsequence which converges uniformly to $f_P$.

Define continuous functions $A_x$ for each fixed $x \in X$ defined by $A_x(y) = A(x, y)$. Now since $P$ is in the open set $S(A_x, P, \epsilon)$, so by convergence of $P_n$ to $P$ there exists a $N_x$ such that $P_n \in S(A_x, P, \epsilon)$ $\forall n \geq N_x \Rightarrow |\int A_x dP_n - \int A_x dP| < \epsilon$ $\forall n \geq N_x$. $\Rightarrow |f_{P_n}(x) - f_P(x)| < \epsilon$ $\forall n \geq N_x$. This says that sequence $f_{P_n}$ is pointwise convergent to $f_P$.

We need the theorem of Ascoli-Arzelà.

**Theorem.** Let $X$ be a compact metric space and $\{f_n\}$ is sequence of continuous functions which is equicontinuous on $X$ and pointwise bounded, then it has a uniformly convergent subsequence.
We prove that \( \{f_{P_n}\} \) is equicontinuous and pointwise bounded sequence, which by using Ascoli-Arzela will say that it will have a uniformly convergent subsequence (say it goes to a continuous \( g \)). But since \( \{f_{P_n}\} \) is pointwise convergent to \( f_P \), so any subsequence of \( \{f_{P_n}\} \) will pointwise converge to \( f_P \). Since uniform convergence implies pointwise convergence, so we get that \( g = f_P \), giving \( \{f_{P_n}\} \) has a subsequence converging uniformly to \( f_P \).

Given \( x \in X, \forall n, |f_{P_n}(x)| = |\int A(x,y) dP_n| \leq \int |A(x,y)| dP_n \leq \int M_x dP_n = M_x \), where \( M_x = \max_y |A(x,y)| \), which exists as \( X \) is compact. So \( \{f_{P_n}\} \) is pointwise bounded.

Now we prove the equicontinuity of \( \{f_{P_n}\} \).

Now, as \( X \) is compact metric space say in metric \( d \), then we know that \( X \times X \) is a compact metric space in metric \( d^*(x_1, y_1, x_2, y_2) = \max(d(x_1, x_2), d(y_1, y_2)) \). Also, since \( A \) is continuous and \( X \times X \) is compact, so \( A \) is uniformly continuous. So \( \forall \epsilon > 0 \) there exist a \( \delta > 0 \) such that if \( d^*((x_1, y_1), (x_2, y_2)) < \delta \) then \( |A(x_1, y_1) - A(x_2, y_2)| < \epsilon \). Taking \( y_1 = y_2 = y \), we get

\[
d(x_1, x_2) < \delta \Rightarrow |A(x_1, y) - A(x_2, y)| < \epsilon, \forall y \in X.
\]

Now integrating both sides with respect to \( P_n \) for any \( n \), we get

\[
\int |A(x_1, y) - A(x_2, y)| dP_n < \int \epsilon dP_n = \epsilon
\]

\[
\Rightarrow |\int (A(x_1, y) - A(x_2, y)) dP_n| < \epsilon \Rightarrow |f_{P_n}(x_1) - f_{P_n}(x_2)| < \epsilon, \forall n.
\]

So we have got that given any \( \epsilon > 0 \) there exist a \( \delta > 0 \) such that \( d(x_1, x_2) < \delta \) implies \( |f_{P_n}(x_1) - f_{P_n}(x_2)| < \epsilon, \forall n \geq 1 \). So this sequence of functions is equicontinuous.

**Remark.** Since we had proved that \( f_{P_n} \) is pointwise convergent to \( f_P \), one can try to prove that \( f_{P_n} \) is uniformly convergent to \( f_P \), not just its some subsequence. In fact, this can also be proved (though its not important here) by using the method that for each \( x \in X \) we find a ball around that \( x \) such that for each other point in that ball, the same \( N_n \) works i.e. \( f_{P_n} \) is uniformly convergent in that ball, and then by using compactness of \( X \), we take a finite cover made by these balls and then take \( N = \max(N_{x_1}, N_{x_2}, \ldots, N_{x_n}) \), which gives uniform \( N \) for each \( x \in X \) implying the uniform convergence. One can check this method works and \( f_{P_n} \) is actually uniformly convergent to \( f_P \).

So finally, as explained earlier, we get that \( S \) which was defined as before, is compact. So we have proved Lemma 1.

Now we go back to our main aim which was proving Theorem **3** stated in the beginning.

I first prove that \( \min_{P_1} \cdot \max_{P_2} \int A(x, y) dP_1 dP_2 \) exists.

Now for fixed \( P_2 \), the map \( f \mapsto \int f dP_2 \) is a continuous and that set of \( \int A(x, y) dP_1 \) as \( P_1 \) varies is compact by Lemma 1 (not exactly because here \( P_1 \) is integrating with respect to first co-ordinate, but this set will also be compact by the same proof as of Lemma 1). Since continuous image of compact set is compact, so we get that, for fixed \( P_2 \) and as \( P_1 \) varies, the set of \( \int A(x, y) dP_1 dP_2 \) is compact. So \( \max_{P_1} \int A(x, y) dP_1 dP_2 \) will exist. Now let this maxima is achieved at \( P_1^1 \). Now by Fubini theorem

\[
\int A(x, y) dP_2 dP_1 = \int \int A(x, y) dP_1 dP_2 dP_1.
\]

Again, \( f \mapsto \int f dP_1^* \) is continuous and by the same argument as above, we get that, as \( P_2 \) varies the set \( \int A(x, y) dP_2 dP_1^* \) is compact. So \( \min_{P_2} \int A(x, y) dP_2 dP_1^* \) will exist. So this proves that \( \min_{P_2} \max_{P_1} \int A(x, y) dP_1 dP_2 \) exists.

Similarly, \( \max_{P_2} \cdot \min_{P_1} \int A(x, y) dP_1 dP_2 \) also exists.

Now we prove that actually both of them are equal.

Assume the contrary that they are not equal. Now clearly

\[
\max_{P_1} \int A(x, y) dP_1 dP_2 \geq \int A(x, y) dP_1 dP_2 \geq \min_{P_2} \cdot \max_{P_1} \int A(x, y) dP_1 dP_2.
\]

\[
\Rightarrow \min_{P_2} \cdot \max_{P_1} \int A(x, y) dP_1 dP_2 \geq \max_{P_1} \cdot \min_{P_2} \int A(x, y) dP_1 dP_2. \quad \text{As since they are not equal, we get}
\]

\[
\min_{P_2} \cdot \max_{P_1} \int A(x, y) dP_1 dP_2 > v > \max_{P_1} \cdot \min_{P_2} \int A(x, y) dP_1 dP_2.
\]

We can assume this \( v \) to be \( v = 0 \), because otherwise, we can consider the function \( A(x, y) - v \) and then work with it. So we get that

\[
\min_{P_2} \cdot \max_{P_1} \int A(x, y) dP_1 dP_2 > 0 > \max_{P_1} \cdot \min_{P_2} \int A(x, y) dP_1 dP_2.
\]
Now consider the set $F = \{ f \in C(X) | f \geq 0 \}$. Now $F$ is a closed set because it is the inverse image of $[0, \infty)$ under the continuous map $f \mapsto \min(f)$, which is well defined as $X$ is compact. And $S$, as defined as before, is compact.

Now if, in a norm linear space, $C_1$ is a compact set and $C_2$ is a closed set, then $C_1 + C_2$ is a closed set. This is because that consider any sequence $a_n + b_n$ with $a_n \in C_1$ and $b_n \in C_2$ such that $a_n + b_n$ is convergent. Now since $C_1$ is compact, so $a_n$ will have a convergent subsequence going to some $a \in C_1$. This says, as $a_n + b_n$ is convergent, the same subsequence of $b_n$ will converge to some $b \in C_2$. So the limit of $a_n + b_n$ will be in $C_1 + C_2$ only. So $C_1 + C_2$ is closed set.

So, by using this, we get that $S + F$ is a closed set.

Now suppose $0 \in S + F$, then there exist probability measure $P_2$ and $f \geq 0$ such that $\int A(x, y) dP_2 + f(x) = 0$ for all $x \in X$. Then by integrating both sides with respect to $P_1$, we get that $\int A(x, y) dP_2 dP_1 \leq 0$. Again by using Fubini theorem, we get that $\int A(x, y) dP_1 dP_2 \leq 0$. Now taking maxima over $P_1$, then also $\max_{P_1} \int A(x, y) dP_1 dP_2 \leq 0$. So $\min_{P_2} \max_{P_1} \int A(x, y) dP_1 dP_2 \leq 0$. This gives a contradiction to (1).

So, $0 \not\in S + F$. Now, also, as $F$ and $S$ are clearly a convex set, $S + F$ is a convex set.

Now we will use Hahn-Banach Separation Theorem.

**Theorem.** Let $N$ be a norm linear space and $K$ and $C$ are two convex disjoint sets in $N$. Suppose $K$ is compact and $C$ is closed. Then there exist a non-zero continuous linear map $T : N \rightarrow R$ and a real number $\alpha$ such that $T(C) < \alpha < T(K)$.

So, applying this theorem on compact set $\{0\}$ and closed set $S + F$, we get a non-zero continuous linear map $T : C(X) \rightarrow R$ and $\alpha$ such that $T(S + F) < \alpha < T(0) = 0$. Again by applying Riesz Representation Theorem, we get a non-zero finite signed measure $\mu$ such that $T(f) = \int fd\mu \forall f \in C(X)$ i.e. $T = T_\mu$.

Now we prove that $\mu$ is a positive measure. Suppose not, then there exists $a \in B \subseteq X$ such that $\mu(E) < 0$. Then, using Lemma 1.1, we get a continuous function $f \geq 0$ such that $\int f d\mu < 0$. Also we can assume that this $\int f d\mu$ is arbitrarily large negative because we see that for any $b > 0$ the function $bf$ also give $\int bf d\mu < 0$ with $bf \geq 0$.

Now, as $T$ is continuous and $S$ is compact, so $T(S)$ bounded subset of $R$. Also, as proved above, since $T(F)$ contains arbitrarily large negative values, $T(S + F)$ will contain a negative value also. This gives a contradiction to $T(S + F) > \alpha > 0$.

So this proves that $\mu$ is a positive measure. Also since $\mu$ is non-zero, so $\mu(X)$ is non-zero, as otherwise this would say that $\mu = 0$. Then we can as well assume that $\mu$ is a probability measure, as we can divide $\mu$ by $\mu(X)$ to get probability measure $P_1$ and replace $\mu$ by $P_1$.

Now since $T(S) \subseteq T(S + F)$, so we get that $T(S) > \alpha > 0$. As explained above, $\mu$ can be assumed to be probability $P_1$, we get that $T_1(S) > \alpha > 0 \Rightarrow \forall$ probability $P_2$, (just by definition of $S$)

$\int A(x, y) dP_2 dP_1 > \alpha > 0$ for some probability measure $P_1$ and for all probability measure $P_2$.

Now we can take minimisation over $P_2$ giving that $\min_{P_2} \int A(x, y) dP_2 dP_1 > \alpha > 0$ and then maximisation over $P_1$ giving that $\max_{P_1} \min_{P_2} \int A(x, y) dP_2 dP_1 > \alpha > 0$. Again by Fubini theorem $\int A(x, y) dP_1 dP_2 = \int A(x, y) dP_2 dP_1$.

So we get that $\max_{P_1} \min_{P_2} \int A(x, y) dP_2 dP_2 > \alpha > 0$ which is again a contradiction to (1). So our assumption that these two values were different is false. So

$\min_{P_2} \max_{P_1} \int A(x, y) dP_2 dP_2 = \max_{P_1} \min_{P_2} \int A(x, y) dP_1 dP_2$.

So, thanks to probability and chance, we have proved that Von Neumann’s Minimax Theorem holds for any compact metric space.