John Nachbar
Washington University
March 1, 2016

Existence of Optima.[1]

1 The Basic Existence Theorem for Optima.

Let X be a non-empty set and let $f : X \to \mathbb{R}$.

The MAX problem is

$$\max_{x \in X} \quad f(x).$$

Thus, x^* is a solution to MAX iff $x^* \in X$ and $f(x^*) \geq f(x)$ for all $x \in X$.

The MIN problem is

$$\min_{x \in X} \quad f(x).$$

Thus, x_* is a solution to MAX iff $x_* \in X$ and $f(x_*) \leq f(x)$ for all $x \in X$.

The basic existence theorem is the following.

Theorem 1. *Let (X, d_x) be a non-empty metric space. If X is compact and $f : X \to \mathbb{R}$ is continuous then MAX and MIN both have solutions.*

Proof. Since X is compact, $f(X)$ is compact and hence it is closed and bounded. Since $f(X) \subseteq \mathbb{R}$ and is bounded, $a = \inf f(X)$ and $b = \sup f(X)$ both exist. Since $f(X)$ is closed, $a, b \in f(X)$. Hence there are $x_*, x^* \in X$ such that $f(x_*) = a$, $f(x^*) = b$, from which the result follows. ■

Theorem 1 gives sufficient, not necessary, conditions for existence of extrema. For example, suppose $X = \mathbb{R}$ and $f : \mathbb{R} \to \mathbb{R}$ is the Dirichlet function:

$$f(x) = \begin{cases} 1 & \text{if } x \text{ is rational} \\ 0 & \text{otherwise.} \end{cases}$$

Then f is not continuous (or, anticipating the discussion in Section 5, either upper or lower semicontinuous), X is not compact, but f attains its maximum at any rational x and its minimum at any irrational x.

2 Approximate Optima.

The following result establishes that if $f : X \to \mathbb{R}$ and $C \subseteq X$ is dense in X (i.e., $\overline{C} = X$) then any solution to MAX or MIN has an approximate solution in C.

Theorem 2. *Let (X, d_x) be a non-empty metric space, let $f : X \to \mathbb{R}$ be continuous, and let $C \subseteq X$ be dense in X.*

1. *If x^* solves MAX, then for any $\varepsilon > 0$ there is a $c \in C$ such that $f(c) > f(x^*) - \varepsilon$.*
2. *If x_* solves MIN, then for any $\varepsilon > 0$ there is a $c \in C$ such that $f(c) < f(x_*) + \varepsilon$.*

Proof. I provide the proof for MAX; the proof for MIN is similar.

If $x^* \in C$ then set $c = x^*$. Otherwise, $x^* \notin C$ but x^* is a limit point of C. Fix $\varepsilon > 0$. By continuity, there is a $\delta > 0$ such that if $x \in N_\delta(x^*)$ then $f(x) \in N_\varepsilon(f(x^*))$. Since x^* is a limit point of C, there is a $c \in N_\delta(x^*) \cap C$. Take any such c. ■

In particular, if $X \subseteq \mathbb{R}$ is compact and C is the set of rationals in X then MAX and MIN have approximate rational solutions.

3 An Application: In Finite-Dimensional Vector Spaces, All Norms are Equivalent.

For a vector space X, recall that two norms, f and g are equivalent iff there are $0 < a \leq b$ such that for any $x \in X$,

$$ag(x) \leq f(x) \leq bg(x).$$

As discussed in the notes on Metric Spaces, equivalent norms induce metrics with the same convergence properties.

In the notes on Vector Spaces and Norms, I showed that the Euclidean and max norms on $\mathbb{R}^N$ are equivalent. I now show that in *any* finite-dimensional vector space, all norms are equivalent. (For a brief introduction to finite-dimensional vector spaces, see the notes on Vector Spaces and Norms.)

Theorem 3. *If X is a finite dimensional vector space, then all norms on X are equivalent.*

Proof. Suppose that X has dimension N and fix a basis $Z = \{z^1, \dots, z^N\}$. Thus any $x \in X$ can be represented (uniquely) as $x = \sum_{n=1}^{N} \lambda_n z^n$ for $\lambda_n \in \mathbb{R}$. The max norm with respect to this basis is $\|x\|_{\max} = \max_n |\lambda_n|$. It is almost immediate that norm equivalence is transitive. It suffices, therefore, to show that any norm on X is equivalent to the max norm with respect to Z. Consider, then, any norm f on X.

I first show that there is a $b > 0$ such that $f(x) \leq b\|x\|_{\max}$ for all $x \in X$. Let

$$\gamma = \max_n \{f(z^n)\}.$$

Note that γ is well-defined (in particular $\gamma < \infty$) since X is finite-dimensional. Then for any $x \in X$, by the norm properties, and by the definition of the max norm,

$$f(x) = f\left(\sum_{n=1}^{N} \lambda_n z^n\right) \leq \sum_{n=1}^{N} f(\lambda_n z^n) = \sum_{n=1}^{N} |\lambda_n| f(e^n) \leq \sum_{n=1}^{N} \gamma \|x\|_{\max} = \gamma N \|x\|_{\max}.$$

Let

$$b = \gamma N.$$

Thus $f(x) \leq b\|x\|_{\max}$ for all $x \in X$. Note that $b > 0$ since $\gamma > 0$ (since no $z^n = 0$).

It remains to show that there is an $a > 0$ such that $a\|x\|_{\max} \leq f(x)$ for all $x \in X$ (by transitivity, necessarily $a \leq b$).

To find a, I claim first that f is continuous with respect to the metric induced by the max norm. This follows from the fact that $f(x) \leq b\|x\|_{\max}$ and the fact that any norm is continuous with respect to the metric that it itself induces (proved in the notes on Continuity). To keep these notes more self-contained, however, here is an explicit proof.

Consider any sequence (x_t) and point x^* in X such that $x_t \to x^*$ under the max norm. Fix $\varepsilon > 0$. Since $x_t \to x^*$ under the max norm, there is a T such that for $t > T$, $\|x_t - x^*\|_{\max} < \varepsilon/b$. By the triangle inequality, $|f(x_t) - f(x^*)| \leq f(x_t - x^*)$. Hence, by the above inequalities and the definition of b, for all $t > T$,

$$|f(x_t) - f(x^*)| \leq f(x_t - x^*) \leq b\|x_t - x^*\|_{\max} < \varepsilon.$$

It follows that $f(x_t) \to f(x^*)$ and hence f is continuous under the metric induced by the max norm.

As discussed in the notes on Completeness in $\mathbb{R}^N$, the proof of Heine-Borel for $\mathbb{R}^N$ implies that Heine-Borel holds in any finite-dimensional vector space with the max norm. In particular, the unit sphere $S = \{x \in X : \|x\|_{\max} = 1\}$ is compact under the max norm, and hence, by Theorem 1, since f is continuous, there is an x_a that solves $\min_{x \in S} f(x)$. Let $a = f(x_a)$. Note that $a > 0$ since $x_a \neq 0$ (since $0 \notin S$).

For $x = 0$, $a\|x\|_{\max} = f(x)$ (since $\|0\|_{\max} = f(0) = 0$). Consider any $x \neq 0$, hence $\|x\|_{\max} > 0$. To make the notation a bit less cluttered, let $\gamma = \|x\|_{\max}$. Then, since f is a norm, $f(x) = f(\gamma(1/\gamma)x) = \gamma f((1/\gamma)x)$. Moreover, since $(1/\gamma)x = (1/\|x\|_{\max})x \in S$, it follows that $f((1/\gamma)x) \geq a$. Hence, $f(x) \geq a\gamma = a\|x\|_{\max}$ for all $x \in X$, as was to be shown. ■

Theorem 3 does not extend to general vector spaces; the notes on Vector Spaces and Norms provide a counterexample in ℓ^∞.

4 Optimization in ℓ^∞.

Using an example, I illustrate here a few of the issues that can arise when optimizing in infinite-dimensional spaces.

Let $X \subseteq \ell^\infty$ comprise the points, $x^0 = 0$, $x^1 = (1, 0, 0, \dots)$, $x^2 = (0, 1+r, 0, \dots)$, $x^3 = (0, 0, (1+r)^2, 0, \dots)$, and so on. The general term x^s has 0s everywhere except for coordinate $n = s$, where the value is $(1+r)^{s-1}$. Assume $r > 0$.

Define $f : X \to \mathbb{R}$ by

$$f(x) = \sum_n \delta^{n-1} x_n,$$

with $\delta > 0$. Thus $f(x^0) = 0$, $f(x^1) = 1$, $f(x^2) = \delta(1+r)$, and in general $f(x^s) = [\delta(1+r)]^{s-1}$.

Consider the problem,

$$\max_{x \in X} f(x).$$

This can be viewed as a "cut the tree" problem. The decision maker chooses a date t at which to consume. Waiting longer gives more to consume but the decision maker is impatient ($\delta > 0$). By inspection, the solution is as follows.

- If $\delta(1+r) < 1$ then the solution is $x^* = x^1$: consume immediately.
- If $\delta(1+r) = 1$ then any x^s, $s > 0$, is a solution.
- If $\delta(1+r) > 1$ there is no solution: for any s, x^{s+1} yields a higher value of f than x^s.

The question I want to address here is how does this analysis match up with the compact/continuous machinery just developed. To do this, we need to select a metric for X. Consider the two metrics introduced in the notes on $\mathbb{R}^\omega$, d_{sup} and d_p.

- Is X compact?
  - d_{sup}. NO. X has no limit points and hence the only convergent sequences are trivial: in any convergent sequence (x_t), there is a T such that for all $t > T$, all terms have the same value. The sequence (x^t) with $x_t = x^t$ is not convergent and has no convergent subsequence.
  - d_p. YES. In particular, X is a subset of the product set $A = \{0, 1\} \times \{0, 1+r\} \times \{0, (1+r)^2\} \times \dots$, and the later is compact under d_p by Tychonoff's Theorem. X is closed as a subset of A, since the non-trivial convergent sequences all converge to 0; I included 0 in X for precisely this reason. Hence X is compact since it is a closed subset of a compact set. Alternatively, every infinite subset of X has 0 as the unique limit point under d_p, and, again, $0 \in X$.
- Is f continuous?
  - d_{sup}. YES. Again, since X has no limit points, the only convergent sequences are trivial. Then f satisfies the sequence characterization of continuity almost vacuously.

– d_p. IT DEPENDS. Again, any non-trivial convergent sequence converges under d_p to 0. For concreteness, consider (x_t) where $x_t = x^t$. Recall that $f(x_t) = f(x^t) = [\delta(1+r)]^{t-1}$. On the other hand, $f(0) = 0$. Therefore, by the sequence characterization of continuity, f is continuous iff $[\delta(1+r)]^{t-1}$ converges to 0, which occurs iff $\delta(1 + r) < 1$.

In summary, if I use the metric $d_{\sup}$ then the answer to, "Does the maximization problem have a solution" is, "Possibly not, because X is never compact." I write "possibly" rather than "certainly" because Theorem 1 only gives a sufficient condition for existence of a solution. On the other hand, d_p gives the answer, "Yes if $\delta(1 + r) < 1$, possibly not otherwise." Thus, of these two metrics, d_p is the more helpful in this application, although even d_p misses the existence of a solution when $\delta(1 + r) = 1$. One should not conclude from this that $d_{\sup}$ is useless; it comes into its own in other applications.

Let me make two other remarks. First, we know that if $\delta(1 + r) > 1$ then no solution exists, and both $d_{\sup}$ and d_p flag this. But they flag it for different reasons. With $d_{\sup}$, f is continuous but X is not compact. With d_p, X is compact but f is not continuous.

Second, f is linear: for any $x, \hat{x} \in X$ and any $a \in \mathbb{R}$, $f(x+\hat{x}) = f(x)+f(\hat{x})$ and $f(ax) = af(x)$. In Euclidean spaces, any linear function is continuous; indeed, linear functions are the canonical examples of continuous functions in Euclidean spaces. But in infinite dimensional vector spaces, linear functions need not be continuous. In the above example, under d_p, f is not continuous, even though it is linear, when $\delta(1 + r) > 1$.

5 Semicontinuity.

Theorem 1 can be strengthened by weakening continuity to semicontinuity.

Theorem 4. *Let (X, d) be a compact metric space and let $f : X \to \mathbb{R}$.*

1. *If f is upper semicontinuous then MAX has a solution.*
2. *If f is lower semicontinuous then MIN has a solution.*

Proof. I provide a proof for upper semicontinuity. The proof for lower semicontinuity is similar.

I claim first that $f(X)$ is bounded above. Suppose not. Then for any $t \in \mathbb{N}$ there is an x_t such that $f(x_t) > t$. By compactness, there is an $x^* \in X$ and a subsequence of (x_t) that converges to x^*. Restrict attention to this subsequence. Take any $T \geq f(x^*) + 1$. Then for every $t > T$, $f(x^*) < f(x_t) - 1$, hence f is not upper semicontinuous. By contraposition, $f(X)$ is bounded above.

Since $f(X)$ is bounded above, it has a least upper bound, $\overline{y}$. I am done if I can show that there is an $x^* \in X$ such that $f(x^*) = \overline{y}$.

Since $\overline{y}$ is a least upper bound of $f(X)$, there is a sequence (y_t) in $f(X)$ converging to $\overline{y}$. For each t, take $x_t \in X$ such that $f(x_t) = y_t$. By compactness of X, there is an $x^* \in X$ and a subsequence of (x_t) that converges to $\overline{y}$. Restrict attention to this subsequence.

Since f is upper semicontinuous, for any $\varepsilon > 0$ there is a T_1 such that for all $t > T_1$,

$$f(x^*) > f(x_t) - \varepsilon/2.$$

On the other hand, since $f(x_t)$ converges to $\overline{y}$, there is a T_2 such that for all $t > T_2$,

$$f(x_t) > \overline{y} - \varepsilon/2.$$

Combining,

$$f(x^*) > \overline{y} - \varepsilon.$$

Since this must hold for all $\varepsilon > 0$, $f(x^*) \geq \overline{y}$. On the other hand, since $\overline{y}$ is an upper bound of $f(X)$, $\overline{y} \geq f(x^*)$. Combine these inequalities and the result follows. ■

6 Further Remarks.

Questions of compactness and continuity are in an important sense not intrinsic to optimization problems. They are auxiliary concepts that we introduce to help with our analysis. A restatement of Theorem 1 may make this clearer (I don't offer a proof because this *is* just a restatement).

Theorem 5. *Let X be non-empty and let $f : X \to \mathbb{R}$. If there exists a metric such that X is compact and f is continuous then MAX and MIN have solutions.*

This restatement shifts the emphasis from working with a particular metric that happens to be given to us to instead *finding* a metric that makes X compact and f continuous. To show that a maximum exists it suffices to find *one* such metric, any metric, no matter how bizarre. If there is no maximum then there is no such metric. This facet of the maximization problem is obscured in $\mathbb{R}^N$, because in $\mathbb{R}^N$ there is a default metric; moreover, all metrics based on norms have the same convergence properties (Theorem 3).

As I mentioned in the Metric Space notes, there is a generalization of metric spaces called topological spaces that take open sets as primitive. The set of open sets is called a *topology*. It is easy to generalize the definition of compactness and continuity to topological spaces and for such spaces, Theorem 5 becomes the following.

Theorem 6. *Let X be non-empty and let $f : X \to \mathbb{R}$. If there exists a topology such that X is compact and f is continuous then MAX and MIN have solutions.*

In particular, to prove that $\max_{x \in X} f(x)$ has a solution, find a topology for which X is compact and f is (upper semi-) continuous. If no solution exists, then there will be no topology with this property.

There is, in particular, a tension between compactness and continuity. Say that one topology is stronger than another if the second is a subset of the first (the first has strictly more open sets than the second). In general, the stronger the topology, the harder it is for X to be compact but the easier it is for $f : X \to \mathbb{R}$ to be continuous. The example in Section 4 illustrated this. As a second example, in $\mathbb{R}$, two possible topologies are the trivial topology, where the only open sets are $\emptyset$ and $\mathbb{R}$ itself, and the discrete topology, where every set is open. The trivial topology cannot be generated by any metric. As discussed in the notes on Metric Spaces, the discrete topology can be generated by the metric $d_d(x, y) = 1$ if $x \neq y$, $d_d(x, x) = 0$, but it cannot be generated by any norm. In the trivial topology, every set is compact (since every sequence is convergent) but the only real-valued functions that are continuous are the constant functions (again, since every sequence is convergent). In the discrete topology, only finite sets are compact but every real-valued function is continuous. In general, there is an art to finding a topology of intermediate strength so that X is compact and f is continuous.

Last, in economic applications we typically do care about the metric/topology. In particular, we often assume that f is continuous with respect to some topology that has an interpretation that we find sensible. It remains true that we are free to consider other topologies in order to establish existence of optima, but in practice it often works out that some sensible topology yields existence, assuming that there is a maximum in the first place.