On the Mathematics of Utility Theory

Harshil Sahai
Swarthmore College

ON THE MATHEMATICS OF UTILITY THEORY

Harshil Sahai[a]

[a]Swarthmore College

On the Mathematics of Utility Theory

Harshil Sahai

Abstract. Utility theory is a field of economics which hopes to model the innate preferences humans have toward different objects. Though it is most obviously economic in spirit and application, the ever-growing discipline finds its theoretical roots in mathematics. This paper will explore the mathematical underpinnings of basic utility theory by following, divulging, and extending the work of Ok [*Real Analysis with Economics Applications*, 2007]. We will develop necessary analytic and algebraic concepts, and use this mathematical framework to support hypotheses in theoretical economics. Specifically, we establish classical existence theorems (Rader, Debreu, and von-Neumann Morgenstern) in both utility and expected utility contexts. The paper will require only a firm grasp of real analysis in $\mathbb{R}^n$, elementary group theory and linear algebra, and will proceed assuming no prior knowledge of economic theory.

1   Introduction

Suppose the set X consists of objects from your everyday grocery store. Apples, bread, steak, and potatoes are all elements of X. In economics, we call X a *commodity space.* As economists, we seek to ask questions about X, just as we would any other set. Can we compare any two elements of the set? How well-defined is this comparison? Is there a "supremum" or "infimum" in this set? And even if there is and we can compare any two elements, can we go further? Can we quantify the magnitude to which we prefer them? This question lies at the heart of utility theory.

For example, suppose we value steak more than we do apples, and apples more than we do bread. If there is any justice in the world, we would likely value steak more than we would bread. But by how much? That is, can we find a function u that assigns to each $x \in X$ a real number, so as to preserve our innate preferences for each object? Because this number would represent our "utility" of the object, we call such a function $u : X \to \mathbb{R}$ a *utility* function.

In Section 2 we explore what conditions on X allow the existence of a preference-preserving utility function u. In Section 3, we show under what further conditions is our utility function continuous. In Section 4, we extend by examining how we can establish a utility function on a set of lotteries, each of which yields prizes in X with certain probabilities. In Section 5, we conclude.

2   Existence of a Utility Function

We first define the comparison relation that presumably exists over our commodity space X, as well as formalize what we mean by a utility function u that "preserves" our preferences in X. Then, we define and expand on a class of spaces, called *separable* spaces, for which existence of such a utility function is feasible. Lastly, we show existence of such a utility function over this class of sets.

2.1   The Preference Relation

We hope to define a relation on our commodity space that is both intuitive to common sense, but also mathematically feasible. Further, we want our relation to be as loose as possible, yet still powerful in its interpretation.

Definition. *Let X be a metric space. A **preference relation** $\succsim$ is a relation on X that is transitive and reflexive. If $\boldsymbol{x} \succsim \boldsymbol{y}$, we say that x is **preferred** to y. We denote $\boldsymbol{x} \succ \boldsymbol{y}$ if $x \succsim y$ but $y \not\succsim x$. In this case, we say that x is **strictly preferred** to y. We denote $\boldsymbol{x} \sim \boldsymbol{y}$ if both $x \succsim y$ and $y \succsim x$, and say that x is **indifferent** to y [7, p.16].*

For example, $\geq$ on $\mathbb{R}$ is a preference relation, where $>$ is our strict preference relation and $=$ is our indifference relation. This follows easily by transitivity and reflexivity of $\geq$. Note that it can readily be shown that $\sim$ is an equivalence relation on X.

In a more abstract setting, recall our toy example, where our commodity space X is the set of grocery items. Our definition of $\succsim$ is natural if we read "steak $\succsim$ apple" as "we prefer a steak as much as we prefer an apple." Further, it is also easily interpreted that by "steak $\succ$ apple" we mean "we prefer steak strictly more than we prefer an apple." One would agree that our assumptions of transitivity and reflexivity are not entirely ridiculous in the context of grocery good preferences.

Now, we must keep in mind that this mere definition says nothing about how well behaved $\succsim$ is within X. This leads to our next definition.

Definition. *Let X be a metric space. A preference relation $\succsim$ on X is **complete** if for all $x, y \in X$,*

$$x \succsim y \ or \ y \succsim x$$

[7, p.101].

Note that this distinction is entirely non-trivial. One thinks of the age-old idiom "you cannot compare apples and oranges." By imposing completeness on our preference relation over our commodity space, we are making a bold claim that all objects in it can be compared nicely.

Further, we would like to have a measure of "smoothness" in our commodity space X. That is, if some good y is preferred to good x, then if another good z is in the neighborhood of x, it should also be the case that y is preferred to z. Else, our preference relation $\succsim$ would not be doing such a good job. In mathematics, ideas of "open-ness" and "continuity" come to mind.

Definition. *Let X be a metric space, equipped with preference relation $\succsim$. Let $x \in X$. Define the following sets:*

$$U_{\succ}(x) = \{y \in X : y \succ x\},$$
$$L_{\succ}(x) = \{y \in Y : x \succ y\}.$$

*We say that $\succsim$ is **upper semi-continuous** if for all $x \in X$, $L_{\succ}(x)$ is open. Similarly, we say that $\succsim$ is **lower semi-continuous** if for all $x \in X$, $U_{\succ}(x)$ is open. Finally, we say $\succsim$ is **continuous** if it is both upper and lower semi-continuous* [7, p.145].

Intuitively, upper semi-continuity of $\succsim$ would imply that if $y \succ x$ and we perturb x by a bit to get z, it still remains that $y \succ z$. This resonates very well with ideas of open-ness in analysis, where any perturbations are contained in the set.

Lastly, we define what we mean by our utility function u that preserves or *represents* the inherent preference structure of $\succsim$ on X.

Definition. *Let X be a metric space equipped with preference relation $\succsim$. We say that $\succsim$ is **representable** if there exists some $u : X \to \mathbb{R}$ such that for all $x, y \in X$,*

$$x \succsim y \ if \ and \ only \ if \ u(x) \geq u(y).$$

*The mapping u is called the **utility function** that represents $\succsim$* [7, p.102].

Suppose we have that $y \succsim x$. We want u to preserve this preference:
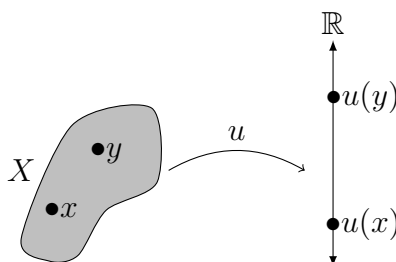
This is fairly intuitive. If we want a real valued function to represent preferences on our commodity space, why not have more preferential objects have higher function values? The key idea here is that once we have established our utility function u as above, we know all we need to about $\succsim$ on X. We can now do all the (mathematical) analysis we want on u, and learn more interesting things about how preferences behave on X.

2.2 Separable Spaces

Now that we have established some properties and potential characteristics of our preference relation, we hope to elaborate on types of commodity spaces for which representation is possible. Our first class of commodity spaces will be *separable* spaces.

Definition. *Let X be a metric space and $Y \subseteq X$. We say that Y is **dense** in X, if $\overline{Y} = X$. We say that X is **separable** if there exists a countable, dense subset of X [7, p.140].*

That is, we want X to be large and nice, but not too complicated. For example, $\mathbb{R}$ is separable. The set of rationals $\mathbb{Q}$ is a countable, dense subset of $\mathbb{R}$. Although $\mathbb{R}$ contains many many more numbers than $\mathbb{Q}$, we essentially can grab all of $\mathbb{R}$ by taking a neighborhood of a countable set $\mathbb{Q}$. Similarly, we want to be able to "grab" X by a countable subset.

Separable spaces offer some interesting properties with regards to open sets. Clearly, if our notion of continuity of $\succsim$ relies on an entirely topological concept of open sets in X, we want open sets in separable spaces to behave especially well if we have any chance of relating continuity of $\succsim$ with separability of X.

Proposition 2.1. *Let X be a metric space. If X is separable, then there exists a countable collection $\mathcal{O}$ of open subsets of X such that for any open set $U \subseteq X$,*

$$U = \bigcup \{O \in \mathcal{O} : O \subseteq U\}$$

[7, p.143].

Proof. Suppose X is separable. Let Y be such a countable, dense subset of X. Define the collection $\mathcal{O}$ as

$$\mathcal{O} = \{B_\epsilon(y) \subseteq X : y \in Y \text{ and } \epsilon \in \mathbb{Q}^+\}.$$

That is, the set of open balls centered at points in Y, contained in X, with postive, rational radius. Note that $\mathcal{O}$ is countable, by the injection $f : \mathcal{O} \to Y \times \mathbb{Q}^+$, defined by

$$f(B_\epsilon(y)) = (y, \epsilon),$$

where $Y \times \mathbb{Q}^+$ is countable as a cross product of countable sets. Note further that $\mathcal{O}$ is a collection of open sets in X, by virtue of being a collection of open balls in X.

It remains to show the equality stated in Proposition 2.1, by showing containment in both directions. Now, consider any open set $U \subseteq X$, and let $x \in U$. Thus, there exists $B_\epsilon(x) \subseteq U$. We can readily shrink ϵ to $\bar{\epsilon} \in \mathbb{Q}^+$, so that $B_{\bar{\epsilon}}(x) \subseteq U$. However, $\overline{Y} = X$ implies that for all $x \in X$, there exists a $y \in Y$ such that $d(x, y) < \bar{\epsilon}/2$. That is, $x \in B_{\bar{\epsilon}/2}(y) \subseteq \mathcal{O}$, because $\bar{\epsilon}/2$ is rational. Also, $B_{\bar{\epsilon}/2}(y) \subseteq B_{\bar{\epsilon}}(x) \subseteq U$. So, $x \in O$, for some $O \in \mathcal{O}$ open and $O \subseteq U$. As x was selected arbitrarily, it is clear that $U \subseteq \bigcup \{O \in \mathcal{O} : O \subseteq U\}$.

Containment in the opposite direction is obvious since for all $x \in \bigcup \{O \in \mathcal{O} : O \subseteq U\}$, $x \in O \subseteq U$ for some $O \in \mathcal{O}$, by construction. $\square$

That is, we may write each open set in a separable space as a countable union of open subsets, where these subsets are drawn from a larger collection inside the space.

2.3 The Rader Utility Representation Theorem

Our goal is to show under what conditions on X and $\succsim$ does there exist a proper utility representation. John Rader found in 1963, that if we assume X is separable and $\succsim$ is upper semi-continuous and complete, we achieve such a representation [8].

Theorem 2.2 (Rader). *Let X be a separable metric space, and $\succsim$ a complete preference relation on X. If $\succsim$ is upper semi-continuous, then it can be represented by a utility function $u : X \to [0, 1]$ [7, p.146].*

Proof. Because X is separable, by Proposition 2.1, there exists a countable collection $\mathcal{O}$ of open sets of X such that for any open set $U \subseteq X$,

$$U = \bigcup \{O \in \mathcal{O} : O \subseteq U\}.$$

Because such a collection is countable, we enumerate $\mathcal{O} = \{O_1, O_2, \dots\}$. Let $x \in X$. Consider

$$M(x) = \{i \in \mathbb{N} : O_i \subseteq L_\succ(x)\}.$$

That is, the index set of all sets of $\mathcal{O}$ whose elements x is strictly preferred to.

Define $u : X \to [0, 1]$ as

$$u(x) = \sum_{i \in M(x)} 1/2^i$$

Let $x, y \in X$. We note that

$$x \succsim y \text{ if and only if } L_\succ(x) \supseteq L_\succ(y). \tag{2.1}$$

That is, x is preferred to y if and only if x is strictly preferred to everything that y is strictly preferred to.

Now, if $L_\succ(x) \supseteq L_\succ(y)$ then it follows that for any $O_i \in L_\succ(y)$, such an O_i must also be contained in $L_\succ(x)$. So, $M(x) \supseteq M(y)$. That is,

$$L_\succ(x) \supseteq L_\succ(y) \text{ only if } M(x) \supseteq M(y). \tag{2.2}$$

We now show the "if" direction of (2.2), by using a contrapositive argument. Suppose there exist $x, y \in X$ such that $L_\succ(x) \not\supseteq L_\succ(y)$. In other words, there exists $z \in L_\succ(y) \setminus L_\succ(x)$. Note that by hypothesis $\succsim$ is upper semi-continuous, so that $L_\succ(y)$ is an open set in X. By our construction of $\mathcal{O}$ we can write

$$L_\succ(y) = O_{n_1} \cup O_{n_2} \cup \ldots$$

for $O_{n_i} \in \mathcal{O}$. So, because $z \in L_\succ(y) \setminus L_\succ(x) \subset L_\succ(y)$, z must belong to some O_{n_k}. Further $O_{n_k} \not\subseteq L_\succ(x)$. Thus, there exists $O_{n_k} \subseteq L_\succ(y)$ such that $O_{n_k} \not\subseteq L_\succ(x)$, which implies directly that $M(x) \not\supseteq M(y)$. We may now conclude a stronger case of (2.2):

$$L_\succ(x) \supseteq L_\succ(y) \text{ if and only if } M(x) \supseteq M(y). \tag{2.3}$$

We finish our proof by showing containment in M corresponds to inequalities in u. Suppose $M(x) \supseteq M(y)$. Then, it is fairly clear that

$$u(x) = \sum_{i \in M(x)} 1/2^i \geq \sum_{i \in M(y)} 1/2^i = u(y),$$

as the latter summation is over a smaller set. Now, suppose $u(x) \geq u(y)$. Then, from the contrapositive of above,

$$u(y) \not> u(x) \Rightarrow M(y) \not\supseteq M(x)$$
$$\Rightarrow L_\succ(y) \not\supseteq L_\succ(x), \text{ by (2.3)}$$
$$\Rightarrow y \not\succ x, \text{ by definition of } L_\succ$$
$$\Rightarrow x \succsim y, \text{ by completeness of } \succsim$$
$$\Rightarrow M(x) \supseteq M(y), \text{ by (2.1), (2.3)}.$$

We have shown that
$$M(x) \supseteq M(y) \text{ if and only if } u(x) \geq u(y). \tag{2.4}$$

To conclude, we combine (2.1), (2.3), and (2.4) to achieve our desired result about the representability of $\succsim$:

$$x \succsim y \text{ if and only if } u(x) \geq u(y).$$

$\square$

3 Continuity of a Utility Function

Now that we have established our first representation theorem, we go further to explore under what conditions is such a utility function u nice. In the spirit of analysis, by "nice" we really mean continuous. Intuitively, if we want u to be continuous, it makes sense for $\succsim$ to be continuous. However, is it sufficient? To show this we first state a preliminary result.

Lemma 3.1 (The Open Gap Lemma)**.** *For any nonempty subset $S \subset \mathbb{R}$, there exists a strictly increasing function $f : S \to \mathbb{R}$ such that every $\subseteq$-maximal connected set in $\mathbb{R} \setminus f(S)$ is either a singleton or an open interval* [7, p.240].

Proof. This proof is left as an exercise for the reader. Please see [1] and [4] for reference. □

3.1 Function Continuity

If we hope to use Rader's result (Theorem 2.2) from before to say something about the continuity of u, we must first divide continuity of a function into upper and lower, just as we did for preference relations.

Definition. *Let X be a metric space. Let $\phi : X \to \mathbb{R}$. We say that ϕ is **upper semi-continuous** at $x \in X$ if given $\epsilon > 0$, there exists $\delta > 0$ such that $d(x, y) < \delta$ implies*

$$\phi(y) \leq \phi(x) + \epsilon.$$

*Similarly, we say that ϕ is **lower semi-continuous** at $x \in X$ if given $\epsilon > 0$, there exists $\delta > 0$ such that $d(x, y) < \delta$ implies*

$$\phi(y) \geq \phi(x) - \epsilon$$

[7, p.229].

That is, the notion of upper semi-continuity of ϕ implies that ϕ takes a neighborhood of x to no more than $\phi(x) + \epsilon$. In other words, we may locally bound ϕ from above. A similar intuition follows for lower semi-continuity. Further, we conclude this digression with the following result of combining upper and lower semi-continuity.

Corollary 3.2. *Let X be a metric space. Let $\phi : X \to \mathbb{R}$. The map ϕ is continuous if and only if ϕ is both upper and lower semi-continuous.*

Proof. The result follows directly from the definitions of upper and lower semi-continuity of ϕ [10]. □

As we know from analysis, if ϕ is a continuous function, it inversely maps open sets to open sets. This arises intuitively because continuity ensures that for any neighborhood in the image, there exists a neighborhood in the preimage that maps inside of it. That is, once again, continuity of a function is very closely tied to the properties of open sets in

the function preimage and image. However, to find an analogous result for upper (lower) semi-continuity, we focus on infinite intervals starting from a fixed right (left) end point, as the image of a neighborhood is now only bounded by one side, not two.

So, we achieve the following result about upper semi-continuity. Note that an analogous result holds for lower semi-continuity which can be shown identically.

Proposition 3.3. *Let X be a metric space. Let $\phi : X \to \mathbb{R}$. Then, ϕ is upper semi-continuous if and only if for all $\alpha \in \mathbb{R}$, $\phi^{-1}([\alpha, \infty))$ is closed* [7, p.233].

Proof. For the forward direction, suppose ϕ is upper semi-continuous. The desired result is equivalent to showing that $X \setminus \phi^{-1}([\alpha, \infty)) = \phi^{-1}(\mathbb{R} \setminus [\alpha, \infty)) = \phi^{-1}((-\infty, \alpha))$ is open, for all $\alpha \in \mathbb{R}$.

Fix $\alpha \in \mathbb{R}$. Let $x \in \phi^{-1}((-\infty, \alpha))$. Let $\epsilon \in (0, \alpha - \phi(x))$. Now, by the upper semi-continuity of ϕ, there exists $\delta > 0$ such that $d(x, y) < \delta$ implies

$$\alpha > \phi(x) + \epsilon \geq \phi(y).$$

That is, $\phi(B_\delta(x)) \subseteq (-\infty, \alpha)$, where $B_\delta(x)$ is the ball of radius δ centered at x. Taking ϕ^{-1} of both sides, we achieve

$$B_\delta(x) \subseteq \phi^{-1}((-\infty, \alpha)),$$

and so $\phi^{-1}((-\infty, \alpha))$ is open.

In the reverse direction, suppose for all $\alpha \in \mathbb{R}$, $\phi^{-1}([\alpha, \infty))$ is closed. Let $\phi(x) \in \mathbb{R}$. Given $\epsilon > 0$, consider the interval $(-\infty, \phi(x) + \epsilon) \subset \mathbb{R}$. So, because $\phi^{-1}([\phi(x) + \epsilon, \infty))$ is closed, we know that by above, $\phi^{-1}((-\infty, \phi(x) + \epsilon))$ is open.

Because, $\phi(x) \in (-\infty, \phi(x) + \epsilon)$, we know that $x \in \phi^{-1}((-\infty, \phi(x) + \epsilon))$. Thus, there exists a $\delta > 0$ such that

$$B_\delta(x) \subseteq \phi^{-1}((-\infty, \phi(x) + \epsilon)).$$

So, for all $y \in B_\delta(x)$, $\phi(y) \in (-\infty, \phi(x) + \epsilon)$. That is, $d(x, y) < \delta$ implies

$$\phi(y) < \phi(x) + \epsilon,$$

and thus ϕ is upper semi-continuous. $\square$

That is, if ϕ is upper semi-continuous, it inversely maps intervals of the form $(-\infty, \alpha)$ to open sets. Whereas, if ϕ were simply continuous, it would inversely map intervals of the form (α, β), and thus all open sets, to open sets.

3.2   The Debreu Utility Representation Theorem

We now show the classical result presented by French mathematician and economist Gerard Debreu in 1964 [2]. Debreu extended Rader's formulation by establishing that we can achieve a continuous utility representation by simply adding the additional feature of continuity to $\succsim$.

Theorem 3.4 (Debreu). *Let X be a separable metric space and $\succsim$ a complete preference relation on X. If $\succsim$ is continuous, then it can be represented by a continuous utility function* [7, p.241].

Proof. If X is separable and $\succsim$ is complete then by Theorem 2.2, there exists a map $u : X \to [0,1]$ that represents $\succsim$.

Suppose $\succsim$ is continuous. We use Lemma 3.1 on $u(X) \subseteq \mathbb{R}$. Thus, there exists a strictly increasing map $f : u(X) \to \mathbb{R}$ such that every $\subseteq$-maximal connected set in $\mathbb{R} \setminus f(u(X))$ is either a singleton or an open interval.

Let us define $v = f \circ u : X \to \mathbb{R}$. Note because f is strictly increasing,

$$x \succsim y \text{ iff. } u(x) \geq u(y) \text{ iff. } v(x) = f(u(x)) \geq f(u(y)) = v(y),$$

so that v represents $\succsim$, just as well. We show that v is continuous by showing that it is upper semi-continuous, and the proof of lower semi-continuity follows identically.

By Proposition 3.3, it is sufficient to show that for all $\alpha \in \mathbb{R}$, $v^{-1}([\alpha, \infty))$ is closed. Fix $\alpha \in \mathbb{R}$, and consider the following cases:

- If $\alpha \in v(X)$, then $\alpha = v(x)$, for some $x \in X$. Note that $L_\succ(x) = v^{-1}((-\infty, \alpha))$. That is, if α represents the object x, then all those values less than α must represent those objects in X that x is preferred to. We know that $L_\succ(x)$ is open by the supposed continuity of $\succsim$ (and thus, its upper semi-continuity). So, $v^{-1}((-\infty, \alpha))$ is open and its complement, the set in question, is closed.

- Suppose $\alpha \in \mathbb{R} \setminus v(X)$. Then, if $\alpha \leq \inf v(X)$, then $v^{-1}([\alpha, \infty)) = X$. Similarly, if $\alpha \geq \sup v(X)$, then $v^{-1}([\alpha, \infty)) = \emptyset$. In both cases, the set in question is closed in the ambient metric space X.

  Suppose $\inf v(X) < \alpha < \sup v(X)$. Let I be the $\subseteq$-maximal connected set in $\mathbb{R} \setminus v(X)$ that contains α. By our construction of v, we know that $I = \{\alpha\}$ or $I = (\alpha_1, \alpha_2)$, for some $\alpha_1, \alpha_2 \in v(X)$. This last condition on the boundary of the interval belonging to $v(X)$ is true, else we could fatten I by ϵ and still be inside $\mathbb{R} \setminus v(X)$, violating the $\subseteq$-maximality of I.

  If $I = (\alpha_1, \alpha_2)$, then

$$v^{-1}([\alpha, \infty)) = \{y \in X : v(y) \geq \alpha_2\},$$

because α_2 is the first element in $\mathbb{R}$ greater than α that belongs to $v(X)$. Consider the complement of the right hand side of the above set:

$$\begin{aligned}
\{y \in x : v(y) < \alpha_2\} &= \{y \in X : v(y) < v(v^{-1}(\alpha_2))\} \\
&= \{y \in X : v^{-1}(\alpha_2) \succ y\} \\
&= L_\succ(v^{-1}(\alpha_2)),
\end{aligned}$$

which is open because $v^{-1}(\alpha_2) \in X$ and by the upper semi-continuity of $\succsim$ as noted above. So, the complement of $v^{-1}([\alpha, \infty))$ is open, and so the set in question is closed.

If $I = \{\alpha\}$, let $A_\alpha = \{\beta \in v(X) : \alpha \geq \beta\}$. That is, the set of elements of $v(X)$ that bound α from below. Consider the set $\bigcap v^{-1}(\{[\beta, \infty) : \beta \in A_\alpha\})$. Suppose $x \in v^{-1}([\alpha, \infty))$. Then, $v(x) \in [\alpha, \infty) \subseteq [\beta, \infty)$, for all $\beta \in A_\alpha$. So, $x \in v^{-1}([\beta, \infty))$, for all $\beta \in A_\alpha$. So, $x \in \bigcap v^{-1}(\{[\beta, \infty) : \beta \in A_\alpha\})$, and finally $v^{-1}([\alpha, \infty)) \subseteq \bigcap v^{-1}(\{[\beta, \infty) : \beta \in A_\alpha\})$.

Now suppose $x \in \bigcap v^{-1}(\{[\beta, \infty) : \beta \in A_\alpha\})$. Then $x \in v^{-1}([\beta, \infty))$, for all $\beta \in A_\alpha$. So, $v(x) \in [\beta, \infty)$, for all $\beta \in A_\alpha$. Suppose $v(x) \in [\beta, \alpha)$, for all $\beta \in A_\alpha$ such that $\beta \neq \alpha$. Let $\beta_0 = \sup A_\alpha$. If $\beta_0 < \alpha$ then the interval $(\beta_0, \alpha] \in \mathbb{R} \setminus v(X)$, and $I = \{\alpha\}$ is no longer the $\subseteq$-maximal connected set containing α. If $\beta_0 = \alpha$, that implies that $\beta \to \alpha$, from the left, so that $[\beta, \alpha) \to \emptyset$ and such a $v(x)$ does not exist. So it must be that $v(x) \in [\alpha, \infty)$, and $x \in v^{-1}([\alpha, \infty))$. That is, $\bigcap v^{-1}(\{[\beta, \infty) : \beta \in A_\alpha\}) \subseteq v^{-1}([\alpha, \infty))$.

We see that:
$$v^{-1}([\alpha, \infty)) = \bigcap v^{-1}(\{[\beta, \infty) : \beta \in A_\alpha\}).$$

As before, we comment that $v^{-1}([\beta, \infty))$ is closed for all $\beta \in A_\alpha \subseteq v(X)$, by virtue of its complement $v^{-1}((-\infty, \beta)) = \{y \in X : v(y) < \beta\} = L_\succ(v^{-1}(\beta))$ being open by upper semi-continuity of $\succsim$. Thus, our set in question is closed by virtue of being an arbitrary intersection of closed sets.

So, in all cases, $v^{-1}([\alpha, \infty))$ is closed, and we are done. $\square$

Although separability is a sufficiently general assumption for our commodity space, it is not entirely obvious why that is. We show a special case of Debreu's result, one that makes more sense if our commodity space is somehow numeric in nature (this is where our example of items in a grocery store would be harder to use).

Corollary 3.5. *Let X be a non-empty subset of $\mathbb{R}^n$ and $\succsim$ a complete preference relation on X. If $\succsim$ is continuous, then it can be represented by a continuous utility function* [7, p.242].

Proof. Suppose $\succsim$ is continuous. So, for all $x \in X$, we know that $L_\succ(x)$ and $L_\prec(x)$ are open in X. Further, because $\succsim$ is complete, for all $x \in X$, there exists a $y \in X$ such that either $x \succ y$ or $y \succ x$. Else, for all $x, y \in X$, $x \sim y$, in which case $u(x) = 0$ would do nicely. In the latter case, $x \in L_\succ(y)$ and in the former, $x \in L_\prec(y)$. So, for all $x \in X$, x belongs to some open set in X. So, there is an open ball around x contained in such an open set (in X), so that X is open.

Now, consider $D = \mathbb{Q}^n \cap X$. We show that D is a countable, dense set contained in X. Because X is open in $\mathbb{R}^n$, for any ϵ-ball around $x \in X$, there exists some $q \in \mathbb{Q}^n$ inside of it, contained in X (and thus $X \cap \mathbb{Q}^n$). So, D is dense in X. Further, D is countable by virtue of being an intersection with $\mathbb{Q}^n$, a countable set (by a finite cross product of countable sets). Lastly, D is contained inside X, by virtue of being an intersection with X. So, X contains a countable, dense subset D, and is thus separable.

Finally, X is a metric space simply by restricting the $\mathbb{R}^n$-metric to X, and so X is a separable metric space. Thus the hypotheses of Theorem 3.4 are established, and the result holds. $\square$

4 Extension: Expected Utility Theory

Suppose now that after shopping at your grocer, you head to an antique art auction. At this odd auction you have the ability to purchase raffle tickets to enter into a lottery of your choosing. Each lottery allows you to win a different painting with different probabilities, suppose.

Call the set of all prizes your commodity space X, and set of lotteries on these prizes $\mathfrak{L}_X$. Clearly, if one has preference with regards to the prizes, they must exhibit some sort of preferences with regards to these lotteries, say $\succsim$. In this section we establish a framework to answer the next question: under what conditions on $\succsim$ and X, can we represent $\succsim$ with a real-valued function?

4.1 Expected Utility

We start by defining what exactly we mean by "lottery," and the space of all lotteries we will be working with.

Definition. *Let X be a non-empty finite set. A map $p : X \to [0,1]$ is a **lottery** (or probability distribution) on X if*

$$\sum_{x \in X} p(x) = 1.$$

We denote $\mathfrak{L}_X$ as the set of all lotteries on X:

$$\mathfrak{L}_X = \left\{ p : X \to [0,1] \text{ such that } \sum_{x \in X} p(x) = 1 \right\}.$$

*We denote a **degenerate lottery** by δ_y as one whose support is only the element $y \in X$:*

$$\delta_y(x) = \begin{cases} 1 & \text{if } x = y \\ 0 & \text{otherwise} \end{cases}$$

[7, p.395-96].

An excellent introduction to probability can be found in [9]. Note that our construction of δ_y is very intuitive. It is simply the lottery which gives the prize y with certainty.

We recall that our goal is to find a utility representation for $\succsim$, our preference relation on $\mathfrak{L}_X$. We want such a utility function to have an intuitive meaning. More specifically, we note that individuals have preferences for each lottery only insofar as they prefer the underlying prizes. So, our function must take into account both the probability of the prize as well as the value of the prize.

Definition. *Let X be a non-empty finite set. Let $p \in \mathfrak{L}_X$, and $u : X \to \mathbb{R}$. We define the map $E_p(u) : \mathfrak{L}_X \to \mathbb{R}$ (the expected value of u with respect to p) as*

$$E_p(u) = \sum_{x \in X} p(x) u(x)$$

[7, p.396].

Suppose we have that $p_2 \succsim p_1$. We want $E_{p_i}(u)$ to preserve $\succsim$ on $\mathfrak{L}_X$:
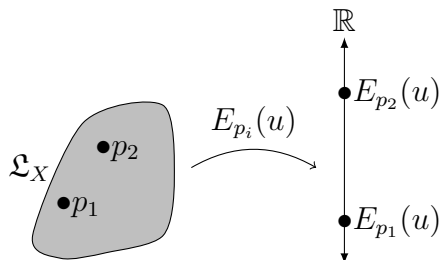
Keep in mind that the goal of any representation is to capture all essential information of the underlying preference relation. Thus, an expected value function over a set of probability distributions on X accomplishes exactly that.

4.2   Linear Spaces and Affine Functions

We hope to find under which conditions does our construction of $\succsim$ on $\mathfrak{L}_X$ give rise to an intuitive utility representation, namely $E_p(u)$. Because any such ambition must rely on the structure of $E_p(u)$, we set up a mathematical foundation of functions of a similar form.

Definition. *Let X be a non-empty set. The list $(X, +, \cdot)$, or simply X, is a **linear space** if $(X, +)$ is an abelian group; X is closed under scalar multiplication; the operation $\cdot$ is associative and distributive; and, $1 \cdot x = x$, for all $x \in X$ [7, p.361].*

Note that in an algebraic sense, "linear spaces" are no more than vector spaces [3]. Next, we define what we mean by a linear map, and note how the structure of linear maps resonates well with our construction of $E_p(u)$.

Definition. *Let S be a non-empty subset of linear space X. The map $\phi : S \to \mathbb{R}$ is **linear** if for any (1) $A \subseteq S$ and (2) $\lambda : A \to \mathbb{R}$ such that $\sum_{x \in A} \lambda(x)x \in S$,*

$$\phi\left(\sum_{x \in A} \lambda(x)x\right) = \sum_{x \in A} \lambda(x)\phi(x).$$

*If we add another condition (3) $\sum_{x \in A} \lambda(x) = 1$ to the equality above, then ϕ is **affine** [7, p.386-87].*

We see that linearity of a function ϕ, defined on a subset of linear space, is essentially a map that is preserved under the given summation. Note further that affinity of ϕ implies linearity over any probability distribution function (or lottery) λ on $A \subset S$.

Next, we show a result that establishes affinity of ϕ under an interesting condition of linearity in $\phi - \phi(\mathbf{0})$, and vice versa.

Lemma 4.1. *Let T be a subset of a linear space with $\mathbf{0} \in T$. Then any function $\phi : T \to \mathbb{R}$ is affine if and only if $\phi - \phi(\mathbf{0})$ is a linear function* [7, p.388].

Proof. For the forward direction, suppose ϕ is affine. Then, let any $A \subseteq T$ and $\lambda : A \to \mathbb{R}$, such that $\sum_{x \in A} \lambda(x)x \in T$. Then, consider

$$
\phi\left(\sum_{x \in A} \lambda(x)x\right) - \phi(\mathbf{0}) = \phi\left(\sum_{x \in A} \lambda(x)x + (1 - \sum_{x \in A} \lambda(x))\mathbf{0}\right) - \phi(\mathbf{0})
$$
$$
= \sum_{x \in A} \lambda(x)\phi(x) + (1 - \sum_{x \in A} \lambda(x))\phi(\mathbf{0}) - \phi(\mathbf{0}), \text{ by affinity of } \phi
$$
$$
= \sum_{x \in A} \lambda(x)(\phi(x) - \phi(\mathbf{0})),
$$

so that $\phi - \phi(\mathbf{0})$ is linear.

For the reverse direction, suppose $\phi - \phi(\mathbf{0})$ is linear. Then, let any $A \subseteq T$ and $\lambda : A \to \mathbb{R}$, such that $\sum_{x \in A} \lambda(x)x \in T$ and $\sum_{x \in A} \lambda(x) = 1$. Then, consider

$$
\phi\left(\sum_{x \in A} \lambda(x)x\right) = \left(\phi\left(\sum_{x \in A} \lambda(x)x\right) - \phi(\mathbf{0})\right) + \phi(\mathbf{0})
$$
$$
= \left(\sum_{x \in A} \lambda(x)(\phi(x) - \phi(\mathbf{0}))\right) + \phi(\mathbf{0}), \text{ by linearity of } \phi - \phi(\mathbf{0})
$$
$$
= \left(\sum_{x \in A} \lambda(x)\phi(x)\right) - \left(\sum_{x \in A} \lambda(x)\right)\phi(\mathbf{0}) + \phi(\mathbf{0})
$$
$$
= \sum_{x \in A} \lambda(x)\phi(x), \text{ by } \sum_{x \in A} \lambda(x) = 1,
$$

so that ϕ is affine. $\qquad \square$

This interesting relationship will be more useful in later sections. The next two results show some more interesting properties of linear and affine functions. Note that we use x_i to denote the ith entry of $x \in \mathbb{R}^n$.

Lemma 4.2. *Let $n \in \mathbb{N}$ and S a non-empty subset of $\mathbb{R}^n$. The map $\phi : S \to \mathbb{R}$ is linear if and only if there exist $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ such that for all $x \in S$,*

$$
\phi(x) = \sum_{i=1}^{n} \alpha_i x_i
$$

[7, p.386].

Proof. This proof is left as an exercise for the reader and follows without too much work by considering the basis of span(S). Please see [7] for reference. $\qquad \square$

That is, if a function ϕ is linear, inside $\mathbb{R}^n$, we can write the image of any vector as the linear combination of vector components. Next, we show a similar result for affine functions.

Proposition 4.3. *Let $n \in \mathbb{N}$ and S a non-empty subset of $\mathbb{R}^n$. Then any function $\phi : S \to \mathbb{R}$ is affine if and only if there exist $\alpha_1, \ldots, \alpha_n, \beta \in \mathbb{R}$ such that $\phi(x) = \sum_{i=1}^n \alpha_i x_i + \beta$ for all $x \in S$ [7, p.389].*

Proof. For the forward direction, let $\phi : S \to \mathbb{R}$ be affine. Let $x^* \in S$. Define $T = S \setminus \{x^*\}$, and $\psi : T \to \mathbb{R}$ as

$$\psi(y) = \phi(y + x^*).$$

Now, let any $A \subseteq T$ and $\lambda : A \to \mathbb{R}$, such that $\sum_{x \in A} \lambda(x)x \in T$ and $\sum_{x \in A} \lambda(x) = 1$. Then, consider

$$
\begin{aligned}
\psi\left(\sum_{x \in A} \lambda(x)x\right) &= \phi\left(\sum_{x \in A} \lambda(x)x + x^*\right) \\
&= \sum_{x \in A} \lambda(x)\phi(x) + \phi(x^*), \text{ by affinity of } \phi \\
&= \sum_{x \in A} \lambda(x)\phi(x) + \sum_{x \in A} \lambda(x)\phi(x^*), \text{ by } \sum_{x \in A} \lambda(x) = 1 \\
&= \sum_{x \in A} \lambda(x)(\phi(x) - \phi(x^*)) \\
&= \sum_{x \in A} \lambda(x)\psi(x),
\end{aligned}
$$

so that ψ is affine. So, by Lemma 4.1, we know that $\psi - \psi(\mathbf{0})$ is linear on T. Further, by Lemma 4.2, there exist $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ such that for all $y \in T$,

$$\psi(x) - \psi(\mathbf{0}) = \sum_{i=1}^n \alpha_i y_i.$$

So, for all $x \in S$, we can write $\phi(x) - \psi(\mathbf{0}) = \psi(x - x^*) - \psi(\mathbf{0}) = \sum_{i=1}^n \alpha_i(x_i - x_i^*)$. Lastly, we let $\beta = \psi(\mathbf{0}) - \sum_{i=1}^n \alpha_i x_i^*$, so that $\phi(x) = \sum_{i=1}^n \alpha_i x_i + \beta$, as desired.

In the reverse direction, let $\phi : S \to \mathbb{R}$ and suppose there exist $\alpha_1, \ldots, \alpha_n, \beta \in \mathbb{R}$ such that $\phi(x) = \sum_{i=1}^n \alpha_i x_i + \beta$ for all $x \in S$. Then, let any $A \subseteq S$ and $\lambda : S \to \mathbb{R}$, such that

$\sum_{x \in A} \lambda(x)x \in S$ and $\sum_{x \in A} \lambda(x) = 1$. Then, consider

$$
\begin{aligned}
\phi\left(\sum_{x \in A} \lambda(x)x\right) &= \sum_{i=1}^{n} \alpha_i \left(\sum_{x \in A} \lambda(x)x\right)_i + \beta \\
&= \sum_{i=1}^{n} \alpha_i \left(\sum_{x \in A} \lambda(x)\right) x_i + \beta \\
&= \sum_{x \in A} \lambda(x) \sum_{i=1}^{n} \alpha_i x_i + \beta \\
&= \sum_{x \in A} \lambda(x) \left(\sum_{i=1}^{n} \alpha_i x_i + \beta\right), \text{ by } \sum_{x \in A} \lambda(x) = 1 \\
&= \sum_{x \in A} \lambda(x)\phi(x),
\end{aligned}
$$

so that ϕ is affine. as desired. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ □

We see the key difference in this characterization between linear and affine functions is β. Although we can write affine images exactly as linear combinations, we can (almost) do the same for affine functions, at the cost of a translation by β.

## 4.3   Convex Sets and Affine Preference Relations

We proceed to define a class of sets for which representation of $\succsim$ on $\mathfrak{L}_X$ is feasible, *convex sets*. After, we establish an analogous definition of affinity for preference relations just as we did above for functions.

Definition. *A subset T of $\mathbb{R}^n$ is **convex** if for all $x, y \in T$ and $\lambda \in [0,1]$,*

$$\lambda x + (1 - \lambda)y \in T.$$

*We call points of the form $\lambda(x) + (1 - \lambda)y$ a **convex combination** of elements x and y* [7, p.77].

One can think of a convex set as a set where the line segment between any two elements is contained within the set:
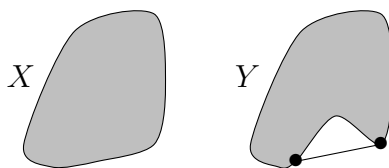
It is clear that X is convex, as it contains the line segment between any two points in X. In contrast, Y is not convex by the given line segment, for example, that is clearly not in Y.

Next, we recall that affinity of a function ϕ is one such that given a proper $\lambda : A \to \mathbb{R}$,

$$\phi\left(\sum_{x \in A} \lambda(x)x\right) = \sum_{x \in A} \lambda(x)\phi(x),$$

where $\sum_{x \in A} \lambda(x) = 1$, so that the expression on the right side is essentially a convex combination of the $\phi(x)$ terms. So, an affine ϕ behaves nicely on convex combinations. Similarly, we define a related concept for preference relations.

Definition. *Let S be a convex subset of linear space X. A preference relation $\succsim$ on S is **affine** if for all $p, q, r \in S$ and $\lambda \in (0, 1]$,*

$$p \succsim q \text{ if and only if } \lambda p + (1 - \lambda)r \succsim \lambda q + (1 - \lambda)r$$

[7, p.387].

That is, if $p \succsim q$, then any convex combination of p and r is preferred to the *same* combination of q and r. Suppose you prefer steak to apples. If you have (1) a combination of steak and potatoes and (2) as much apples as steak in (1) and as much potatoes as (1), you would likely prefer (1) over (2).

We now establish an interesting property of affine preferences defined over convex sets.

Lemma 4.4 (Shapley-Baucells). *Let X be a nonempty set. Let $\mathbb{R}^X = \{f : X \to \mathbb{R}\}$. Suppose S is a convex subset of $\mathbb{R}^X$, and $\succsim$ an affine preference relation on S. Then, for any $p, q \in S$, we have $p \succsim q$ if and only if there exist $\lambda > 0$ and $r, s \in S$ such that $r \succsim s$ and $p - q = \lambda(r - s)$* [7, p.401].

Proof. In the forward direction, let $p, q \in S$ and suppose $p \succsim q$. Pick any $\lambda > 0$, $k \in S$, and by convexity of S we may define $r, s \in S$ as

$$r = \frac{1}{\lambda}p + \frac{\lambda - 1}{\lambda}k,$$

$$s = \frac{1}{\lambda}q + \frac{\lambda - 1}{\lambda}k.$$

It is clear that $r - s = \frac{1}{\lambda}p - \frac{1}{\lambda}q$ so that $\lambda(r - s) = p - q$. Lastly, it is clear that $p \succsim q$ implies $r \succsim s$ by affinity of $\succsim$.

For the reverse direction, let $p, q \in S$. Take any $\lambda > 0$, and $r, s \in S$, such that $r \succsim s$ and $p - q = \lambda(r - s)$. Then, consider

$$p - q = \lambda(r - s) \Rightarrow p + \lambda s = q + \lambda r \Rightarrow \frac{1}{1 + \lambda}p + \frac{\lambda}{1 + \lambda}s = \frac{1}{1 + \lambda}q + \frac{\lambda}{1 + \lambda}r$$

But, because $\succsim$ is affine and $r \succsim s$, we know that

$$\frac{1}{1 + \lambda}q + \frac{\lambda}{1 + \lambda}r \succsim \frac{1}{1 + \lambda}q + \frac{\lambda}{1 + \lambda}s,$$

so that $p \succsim q$, also by affinity of $\succsim$, as desired. $\square$

This useful result states, essentially, that the difference between any two elements in a convex set can be pegged to a multiple of the difference of any two other elements in the set, under the condition that the pairs of elements prefer each other, affinely.

Lastly, we define an interesting property of $\succsim$ defined on convex sets.

Definition. *Let X be a non-empty finite set, S a nonempty convex subset of $\mathfrak{L}_X$, and $\succsim$ a preference relation on S. We say that $\succsim$ is **bounded** in S if for all $p \in S$, there exists a $p^*, p_* \in S$ such that*

$$p^* \succsim p \succsim p_*$$

*We denote $\boldsymbol{p^*}$ as the $\succsim$-**maximum** in S and $\boldsymbol{p_*}$ as the $\succsim$-**minimum** in S [7, p.398].*

The key distinction between usual set-boundedness and preference-boundedness is that the maximum and minimum are contained within the set. For example, $\geq$ is bounded on $S \subset \mathbb{R}$ if S is bounded *and closed* (compact). Put simply, $\succsim$ on X is bounded if X contains its $\succsim$-maximum and $\succsim$-minimum.

4.4 The von Neumann-Morgenstern Expected Utility Theorem

In 1947, mathematician John von Neumann and economist Oskar Morgenstern established a representation theorem for $\succsim$ on $\mathfrak{L}_X$ [6]. We first show a proposition and then prove the desired theorem as a special case.

Proposition 4.5. *Let X be a nonempty finite set, S a nonempty convex subset of $\mathfrak{L}_X$, and $\succsim$ a complete, bounded preference relation on S. Then, $\succsim$ is affine and continuous if and only if there exists a function $u : X \to \mathbb{R}$ such that for any $p, q \in S$,*

$$p \succsim q \text{ if and only if } E_p(u) \geq E_q(u)$$

[7, p.398-401].

Proof. The result is highly non-trivial. We proceed by proving the forward and reverse direction in two parts for clarity:

Part 1. *If $\succsim$ is affine and continuous, then there exists a function $u : X \to \mathbb{R}$ such that for any $p, q \in S$,*
$$p \succsim q \text{ if and only if } E_p(u) \geq E_q(u).$$

Suppose $\succsim$ is affine and continuous on S. By assumption, $\succsim$ is bounded on S, so there exist $p^*, p_* \in S$ as the $\succsim$-maximum and $\succsim$-minimum, respectively. We assume $p^* \succ p_*$, for the non-trivial case. Else, if $p^* \sim p_*$, then for all $p, q \in S$, $p \sim q$, in which case $u(x) = 0$ would do nicely. Now by Lemma 4.4, we see that for all $\alpha, \beta \in [0, 1]$,

$$\beta p^* + (1 - \beta)p_* \succ \alpha p^* + (1 - \alpha)p_*,$$

if and only if, there exists a $\lambda > 0$, and $r, s \in S$ with $r \succsim s$ such that

$$(\beta p^* + (1 - \beta)p_*) - (\alpha p^* + (1 - \alpha)p_*) = \lambda(r - s)$$

iff. $(\beta - \alpha)p^* + (\alpha - \beta)p_* = \lambda(r - s)$

iff. $(\beta - \alpha)(p^* - p_*) = \lambda(r - s)$

iff. $\beta - \alpha = \dfrac{\lambda(r - s)}{p^* - p_*} > 0$, by $p^* \succ p_*$, $r \succsim s$, and $\lambda > 0$

iff. $\beta > \alpha$.

So, we summarize:

$$\beta p^* + (1 - \beta)p_* \succ \alpha p^* + (1 - \alpha)p_* \text{ if and only if } \beta > \alpha. \tag{4.1}$$

Intuitively if $\beta > \alpha$, then more weight is put towards p^* rather than p_*, and is thus preferential.

Note that $S \subseteq \mathfrak{L}_X \subseteq \mathbb{R}^X$. Consider $\mathbb{R}^X$. Because X is finite, let $|X| = k$, and enumerate $X = x_1, x_2, \ldots, x_k$. Note that R^k is the set of all unique k-dimensional vectors. But a k-dimensional vector is no more than a map from a k-tuple to a unique value, i.e. a function $f : X \to \mathbb{R}$. That is, $R^k \cong R^X$, and so R^X is no more than a Euclidean space.

So, S is a non-empty subset of a Euclidean space (namely R^k), and because $\succsim$ is complete and continuous, we may invoke Corollary 3.5. So there exists a continuous utility function $U : S \to \mathbb{R}$, that represents $\succsim$. Define $V : [0, 1] \to U(S)$ as

$$V(\lambda) = U(\lambda p^* + (1 - \lambda)p_*).$$

Note that V is increasing in λ, by (4.1). Note that because U is continuous, it clear that V is also continuous. We see that for all $p \in S$,

$$V(1) = U(p^*) \geq U(p) \geq U(p_*) = V(0).$$

So, because V is continuous on an interval $[0, 1]$, by the Intermediate Value Theorem there exists a $\lambda_p \in (0, 1)$ such that

$$V(\lambda_p) = U(p) = U(\lambda_p p^* + (1 - \lambda_p)p_*).$$

We further define the map $f : U(S) \to \mathbb{R}$ for all $U(p) \in U(S)$ as

$$f(U(p)) = \lambda_p.$$

Note, similarly, that f is increasing as a higher utility value equates more weight towards p^*. Thus, the utility function $L = f \circ U : S \to \mathbb{R}$ also represents $\succsim$.

We show that L is affine. Let $(p^{(i)}, \lambda_i) \in S \times \mathbb{R}$, such that for all $m \in \mathbb{N}$, $\sum_{i=1}^m \lambda_i = 1$ and $\sum_{i=1}^m \lambda_i p^{(i)} \in S$. Then, consider

$$L\left(\sum_{i=1}^m \lambda_i p^{(i)}\right) = L\left(\sum_{i=1}^m \lambda_i p^{(i)}\right)$$

$$= L\left(\sum_{i=1}^m \lambda_i(\lambda_{p^{(i)}} p^* + (1 - \lambda_{p^{(i)}})p_*)\right)$$

$$= L\left(p^* \sum_{i=1}^m \lambda_i \lambda_{p^{(i)}} + p_* \sum_{i=1}^m \lambda_i(1 - \lambda_{p^{(i)}})\right).$$

Note that $\sum_{i=1}^m \lambda_i \lambda_{p^{(i)}} + \sum_{i=1}^m \lambda_i(1 - \lambda_{p^{(i)}}) = \sum_{i=1}^m \lambda_i(\lambda_{p^{(i)}} + (1 - \lambda_{p^{(i)}})) = \sum_{i=1}^m \lambda_i = 1$, so that we can continue as

$$L\left(\sum_{i=1}^m \lambda_i p^{(i)}\right) = L\left(p^* \sum_{i=1}^m \lambda_i \lambda_{p^{(i)}} + p_*\left(1 - \sum_{i=1}^m \lambda_i \lambda_{p^{(i)}}\right)\right)$$

$$= L(k), \text{ such that } \lambda_k = \sum_{i=1}^m \lambda_i \lambda_{p^{(i)}}$$

$$= f(U(k)) = \lambda_k = \sum_{i=1}^m \lambda_i \lambda_{p^{(i)}} = \sum_{i=1}^m \lambda_i L(p^{(i)}),$$

so that L is affine.

We proceed by utilizing Proposition 4.3, because L is affine, there exists a k-tuple $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$ (or equivalently a $u : X \to \mathbb{R}$, as discussed) and a $\beta \in \mathbb{R}$ such that for all $x \in X$,

$$L(p) = \sum_{x \in X} u(x)p(x) + \beta.$$

So, because L represents $\succsim$ as mentioned before, for all $p, q \in S$,

$$p \succsim q \text{ iff. } L(p) \geq L(q)$$

$$\text{iff. } \sum_{x \in X} u(x)p(x) + \beta \geq \sum_{x \in X} u(x)q(x) + \beta$$

$$\text{iff. } \sum_{x \in X} u(x)p(x) \geq \sum_{x \in X} u(x)q(x)$$

$$\text{iff. } E_p(u) \geq E_q(u),$$

and we are done.

Part 2. *If there exists a function $u : X \to \mathbb{R}$ such that for any $p, q \in S$,*

$$p \succsim q \text{ if and only if } E_p(u) \geq E_q(u),$$

then $\succsim$ is affine and continuous

Suppose there exists a function $u : X \to \mathbb{R}$ such that for any $p, q \in S$,

$$p \succsim q \text{ if and only if } E_p(u) \geq E_q(u).$$

It is left to show that $\succsim$ is affine and continuous. For sake of clarity, we split the proof into two parts:

- We show $\succsim$ is affine. Let $\lambda \in (0, 1]$. Let $p, q, r \in S$. Note that by supposition,

$$
\begin{aligned}
p \succsim q \text{ iff. } & E_p(u) \geq E_q(u) \\
\text{iff. } & \sum_{x \in X} u(x)p(x) \geq \sum_{x \in X} u(x)q(x) \\
\text{iff. } & \sum_{x \in X} u(x)(\lambda p(x)) \geq \sum_{x \in X} u(x)(\lambda q(x)) \\
\text{iff. } & \sum_{x \in X} u(x)(\lambda p(x) + (1 - \lambda)r(x)) \geq \sum_{x \in X} u(x)(\lambda q(x) + (1 - \lambda)r(x)) \\
\text{iff. } & E_{\lambda p + (1-\lambda)r}(u) \geq E_{\lambda q + (1-\lambda)r}(u) \\
\text{iff. } & \lambda p + (1 - \lambda)r \succsim \lambda q + (1 - \lambda)r,
\end{aligned}
$$

  so that $\succsim$ is affine.

- We show $\succsim$ is continuous, by showing that it is upper semi-continuous, and an identical argument shows lower semi-continuity. Because $\succsim$ is bounded, give p^* and p_* their usual definitions. Define $f : [0, 1] \to \mathbb{R}$ as

$$f(\lambda) = E_{\lambda p^* + (1-\lambda)p_*}(u).$$

Given $\epsilon > 0$, $|\lambda - \lambda_0| < \delta = \epsilon/|E_{p^* - p_*}(u)|$ implies

$$
\begin{aligned}
|f(\lambda) - f(\lambda_0)| &= \left| \sum_{x \in X} u(x)(\lambda p^*(x) + (1 - \lambda)p_*(x)) - \sum_{x \in X} u(x)(\lambda_0 p^*(x) + (1 - \lambda_0)p_*(x)) \right| \\
&= \left| \sum_{x \in X} u(x)((\lambda - \lambda_0)p^*(x) + (\lambda_0 - \lambda)p_*(x)) \right| \\
&= \left| \sum_{x \in X} u(x)(\lambda - \lambda_0)(p^*(x) - p_*(x)) \right| \\
&= |\lambda - \lambda_0| \cdot \left| \sum_{x \in X} u(x)(p^*(x) - p_*(x)) \right| \\
&= |\lambda - \lambda_0| \cdot |E_{p^* - p_*}(u)| \\
&< \delta \cdot |E_{p^* - p_*}(u)| = \epsilon,
\end{aligned}
$$

  so that f is continuous.

Note that for all $p \in S$, by supposition,

$$p^* \succ p \succ p_* \text{ iff. } E_{p^*}(u) > E_p(u) > E_{p_*}(u)$$
$$\text{iff. } f(1) > E_p(u) > f(0).$$

So, because f is continuous on the interval $[0, 1]$, by the Intermediate Value Theorem there exists a $\lambda_p \in (0, 1)$ such that for all $p \in S$,

$$f(\lambda_p) = E_p(u) = E_{\lambda_p p^* + (1-\lambda_p)p_*}(u).$$

However, this is only true if and only if

$$p \sim \lambda_p p^* + (1 - \lambda_p)p_*.$$

Note that $\succsim$ is affine and so we may use the result from (4.1). This implies that for all $p, q \in S$,

$$p \succ q \text{ iff. } p \sim \lambda_p p^* + (1 - \lambda_p)p_* \succ q \sim \lambda_q p^* + (1 - \lambda_q)p_*$$
$$\text{iff. } \lambda_p > \lambda_q. \tag{4.2}$$

Now, recall $L_\succ(p) = \{q \in S : p \succ q\}$. Define $\psi : S \to \mathbb{R}$ as

$$\psi(p) = \lambda_p.$$

It is clear that as $p \to q, \lambda_p \to \lambda_q$, so that ψ is continuous. By (4.2), we see that $\psi(L_\succ(p)) = \psi(\{q \in S : p \succ q\}) = \{\lambda_q \in \mathbb{R} : \lambda_p > \lambda_q\} = (-\infty, \lambda_p)$. Because ψ is continuous, because $\psi(L_\succ(p))$ is an open interval, and thus an open set, $L_\succ(p)$ is open in S. Therefore, $\succsim$ is upper semi-continuous.

So, we have shown both affinity and continuity of $\succsim$, and Part 2 is complete. Part 1 and Part 2 complete both directions of the stated Proposition, and so the proof is complete. $\square$

This classical result in utility theory establishes the existence of a utility representation of $\succsim$ on a subset S of $\mathfrak{L}_X$ provided (1) $\succsim$ is complete, bounded, affine, and continuous and (2) S is convex.

Recall that the conditions on the structure of $\succsim$ are not entirely unbelievable, and actually resonate well with common sense and how individuals prefer objects. It is very likely that you may have a most and least preferable lottery (boundedness), and that your preferences are preserved under a convex combination (affinity). Further, your preferences on lotteries are likely also the same under small perturbations of the lotteries (continuity). Perhaps the strictest assumption is that we have the ability to compare any two lotteries (completeness). It is often a dilemma whether to play Blackjack or Texas Hold'em at the neighborhood casino.

Perhaps also the convexity assumption on S, our commodity space subset, is a bit strict. Thus, to finish, we extend the Proposition into representability of $\succsim$ on all of $\mathfrak{L}_X$.

Theorem 4.6 (von Neumann-Morgenstern). *Let X be a non-empty finite set, $\succsim$ a complete preference relation on $\mathfrak{L}_X$. Then, $\succsim$ is affine and continuous if and only if there exists a function $u : X \to \mathbb{R}$ such that for any $p, q \in \mathfrak{L}_X$,*

$$p \succsim q \text{ if and only if } E_p(u) \geq E_q(u)$$

[7, p.399].

Proof. We prove the theorem by showing the hypotheses are consistent with those of Proposition 4.5, and so the identical result will hold. It is clear that $\mathfrak{L}_X$, the space of all lotteries (probability distributions) on X, is a convex subset of itself (any convex combination of two probability distributions is still a probability distribution!). It is left to show that $\succsim$ is bounded on $\mathfrak{L}_X$.

Because X is finite and $\succsim$-complete, for all $x \in X$, there exists $\delta_{x^*} \succsim \delta_x$. That is, there exists a prize in X that is the most preferred (and so the lottery with probability one of receiving such a prize is also the most preferred). Note that Part 2 of the proof of Proposition 4.5, showing affinity of $\succsim$, did not rely on its boundedness in $\mathfrak{L}_X$. So, we may use the result that $\succsim$ is affine. Consider any $p \in \mathfrak{L}_X$. Thus,

$$\delta_{x^*} = \sum_{x \in X} p(x)\delta_{x^*} \succsim \sum_{x \in X} p(x)\delta_x = p,$$

where we use affinity of $\succsim$ finitely many times for each $x \in X$. So δ_{x^*} is the $\succsim$-maximum on $\mathfrak{L}_X$. A similar argument shows that there exists a δ_{x_*}, which is the $\succsim$-minimum, and so $\succsim$ is bounded on $\mathfrak{L}_X$, and we are done. $\qquad\square$

Therefore, because boundedness comes naturally, we may represent a preference relation $\succsim$ on the entire $\mathfrak{L}_X$ provided only that $\succsim$ is complete, affine, and continuous.

5   Conclusion

How can we model preferences humans have towards different objects? This is a very important question in the field of economics and specifically, microeconomics, the study of behavior and decision making of firms and consumers [5].

Such a question has immediate mathematical implications. In particular, we start with two key concepts: *sets*, or the commodity space on which our preferences operate, and *relations*, or the structure on which we can "compare" any two objects in this commodity space.

By modeling or "representing" these preferences, we mathematically translate this question into whether or not there exists a real-valued *utility function* defined on our set that preserves our relation. So, at first glance, any set and any relation cannot induce such a function! We must think carefully about what properties of our set and our relation make existence of such a function feasible.

In this paper, we highlight a few of these special properties including separability of our set and completeness and upper semi-continuity of our relation that make representation possible (Rader [8, Section 2]). We then ask how we can make our target function continuous (in the analytic sense), which, we show, requires separability of our set and both completeness and continuity of our relation (Debreu [2, Section 3]). Lastly, we extend to show how we can represent preferences on the set of lotteries that give prizes in our commodity space, by assuming convexity of this set and completeness, affinity, and continuity of our relation (von Neumann-Morgenstern [6, Section 4]).

Utility theory is a field that dates back to 1776, where Adam Smith first wrote boldly trying to relate human psychology to mathematics, and continues its research today [11]. Ultimately, any new, interesting, and significant contribution to the field must find its roots in the analytic and algebraic properties of the sets and relations in question, making utility theory not a far cry away from a purely mathematical discipline.

References

[1] Beardon, A. 1992. "Debreu's Gap Theorem." *Economic Theory* 2: 150-52.

[2] Debreu, G. 1964. "Continuity Properties of Paretian Utility." *International Economic Review* 5: 285-293.

[3] Fraleigh, J., Katz, V. 2003. *A First Course in Abstract Algebra.* Boston: Addison-Wesley.

[4] Jaffray, J.-Y. 1975. "Existence of a Continuous Function: An Elementary Proof." *Econometrica* 43: 981-83.

[5] Kreps, D. 1990. *A Course in Microeconomic Theory.* Princeton: Princeton University Press.

[6] Neumann, J., Morgenstern, O. 1947. *Theory of Games and Economic Behavior* Princeton: Princeton University Press.

[7] Ok, E. 2007. *Real Analysis with Economic Applications.* Princeton: Princeton University Press.

[8] Rader, J. 1963. "The Existence of a Utility Function to Represent Preferences." *Review of Economic Studies* 30: 229-232.

[9] Rice, J. 2007. *Mathematical Statistics and Data Analysis.* Belmont: Thompson.

[10] Rosenlicht, M. 1985. *Introduction to Real Analysis.* New York: Dover Publications.

[11] Stigler, G. 1950. "The Development of Utility Theory." *Journal of Political Economy* 58: 307-327.