A Study of Convex Functions with Applications

Matthew Liedtke

May 14, 2012

1. Introduction

The study of concave up functions in calculus is often limited to describing one of the possible classifications of an increasing function. This is unfortunate because the simple generalization to a convex function greatly extends our scope for analysis. The theory developed in the study of convex functions, arising from intuitive geometrical observations, may be readily applied to topics in real analysis and economics.

This paper begins with a rigorous study of convex functions with the goal of developing a theoretical framework to extend to further topics. Our applications to real analysis include right-endpoint approximation for integrals and pointwise convergence. In our discussion of economics, we will use the theory of consumer behavior under uncertainty and Jensen's Inequality to show the relation between convex functions and the modeling of risk loving behavior. While a first course in real analysis is necessary to truly grasp these results, this paper has been written to be accessible to those who have completed a bridge to higher mathematics course.

2. Convex functions and theory

2.1. **Concave Functions.** The theoretical predecessor to convexity is a concave up function. As seen in Figure 1, the rate at which the concave up function f is increasing is itself increasing. This may be formally expressed in terms of f'.

Definition 2.1. Let f be a differentiable function defined on an interval I.

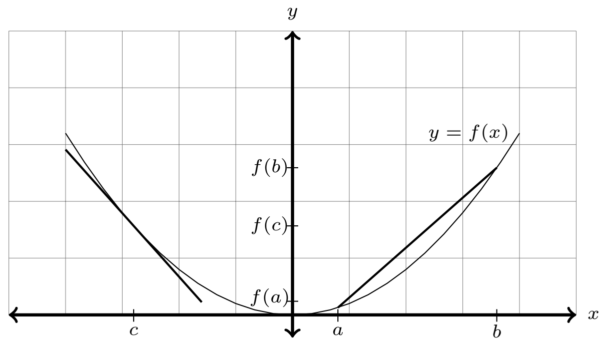
(a) The function f is **concave up** on I if f' is increasing on I.
(b) The function f is **concave down** on I if f' is decreasing on I.

Two important theoretical properties of a concave up function begin with geometric observations. Consider a function f that is concave up on an interval I. Referring to Figure 1, we see that the line tangent to the graph at a point $(c, f(c))$ lies below all other points on the graph of f. Similarly, a secant line drawn between $(a, f(a))$ and $(b, f(b))$ lies above the graph.

Definition 2.2. Let f be a function defined on an interval I.

(a) Suppose that f is differentiable on I. The graph of f **lies above its tangent lines** on I if for each $c \in I$, the inequality $f(x) \geq T_c(x)$ is valid for all $x \in I$, where T_c represents the function whose graph is the line tangent to the graph of $y = f(x)$ when $x = c$.
(b) The graph of f **lies below its secant lines** on I if for each pair of distinct points $a, b \in I$ with $a < b$, the inequality $f(x) \leq S_{ab}(x)$ is valid for all $x \in [a, b]$, where S_{ab} represents the function whose graph is the line joining the points $(a, f(a))$ and $(b, f(b))$.

FIGURE 1. A concave up function



Our observation concerning secant lines is particularly useful as a tool in further proofs. Indeed, throughout this paper we will be returning to this result as a basis for proofs and applications. For this reason we will carefully prove this result below while the complementary tangent line proof may be found in Gordon [1].

Theorem 2.3. Let f be a differentiable function defined on an interval I. The function f is concave up on I if and only if the graph of f lies below its secant lines on I.

Proof. Suppose that the function f is concave up on I, and let $[a, b] \subseteq I$. We define the secant line between the points $(a, f(a))$ and $(b, f(b))$ by

$$S_{ab}(x) = \frac{f(b) - f(a)}{b - a}(x - a) + f(a).$$

Let $D : [a, b] \to \mathbb{R}$ be defined

$$D(x) = \begin{cases} \frac{f(x)-f(a)}{x-a} & \text{for } a < x \leq c; \\ f'(a) & \text{for } x = a. \end{cases}$$

Since f is continuous on I and differentiable at a, it follows that D is continuous on $[a, b]$. Furthermore, using the fact that f is concave up, and thus $f'(x) - f'(a) \geq 0$ for $x \in (a, b)$, we have that $D'(x) \geq 0$ for $x \in (a, b)$. The Monotonicity Theorem states that $D(x) \leq D(b)$ for $x \in (a, b)$, and using the definition of D, we find that

$$\begin{aligned} \frac{f(x) - f(a)}{x - a} &\leq \frac{f(b) - f(a)}{b - a} \\ f(x) &\leq \frac{f(b) - f(a)}{b - a}(x - a) + f(a). \end{aligned}$$

Thus, the graph of f lies below its secant lines on I.

Now instead suppose that the graph of f lies below its secant lines on I. Let $a, b \in I$ such that $a < b$. By the Mean Value Theorem, there exists a

$c \in (a, b)$ such that $f'(c) = \frac{f(b)-f(a)}{b-a}$. Since the graph of f lies below its secant lines on I, we have for $x \in [a, b]$ that

$$f(x) \leq S_{ab} = f'(c)(x-a) + f(a),$$

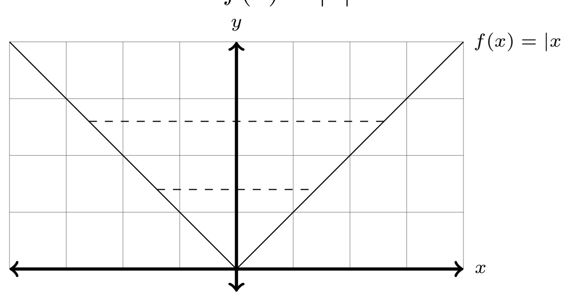
and therefore,

$$f'(a) = \lim_{x \to a^+} \frac{f(x) - f(a)}{x - a} \leq f'(c).$$

It can be shown that $f'(c) \leq f'(b)$; the proof is analogous. It follows $f'(a) \leq f'(b)$, and therefore f is concave up on $[a, b]$. [1] $\square$

2.2. **Convex functions.** While it is certainly convenient to work within the set of differentiable functions, this restriction will limit the scope of our analysis. For instance, there are functions that are not differentiable on the same interval for which their graph lies below its secant lines. The absolute value function, which is not differentiable at zero, is an example of a function with this property. See Figure 2.

FIGURE 2. The function $f(x) = |x|$ lies below its secant lines.



With the following definition of a convex function we will drop the requirement that f is differentiable while maintaining the geometrical properties of a concave up function.

Definition 2.4. Let f be a function defined on an interval I. The function f is **convex** on I if for each interval $[c, d] \subseteq I$, the inequality

$$f\Big((1-t)c + td\Big) \leq (1-t)f(c) + tf(d)$$

is valid for $0 \leq t \leq 1$.

As might be expected, this generalization comes with a trade-off. Proving convexity from the definition has two possible difficulties. First, we will show in Example 2.5 that even working with a simple function is in itself nontrivial. Second, recall that the motivating factor when introducing convexity was to further the reach of our analysis. As is often the case, it may be beneficial to have an equivalent definition. In Theorem 2.8 we will return to our original geometric observation and show that secant lines lie above the graph of a convex function.

[1]The details of this proof are referenced from Gordon [1].

Example 2.5. We will show that $f(x) = x^2$ is convex from the definition. Let $[c, d] \subseteq I$ and $0 \leq t \leq 1$. We find that

$$\begin{aligned} f\Big((1-t)c + td\Big) &= (1-t)^2c^2 + 2t(1-t)cd + t^2d^2 \\ &= (1-t)c^2 - t(1-t)c^2 + 2t(1-t)cd + t^2d^2 \\ &= (1-t)c^2 + t(1-t)c(2d-c) + t^2d^2 - td^2 + td^2 \\ &= (1-t)c^2 - t(1-t)(c^2 - 2cd + d^2) + td^2 \\ &= (1-t)c^2 - t(1-t)(c-d)^2 + td^2 \\ &\leq (1-t)c^2 + td^2 \end{aligned}$$

Thus, the function $f(x) = x^2$ is convex.

We begin the proof of Theorem 2.8 with two lemmas.

Lemma 2.6. *Let f be defined on an interval $[c, d]$. If $0 \leq t \leq 1$, then $c \leq (1-t)c + td \leq d$.*

Proof. Simple algebra reveals that

$$0 \leq t \leq 1$$

$$0 \leq t(d-c) \leq d - c$$

$$c \leq c + t(d-c) \leq d$$

$$c \leq (1-t)c + td \leq d \ ,$$

which is the desired result. □

Lemma 2.7. *Let S_{cd} be the function that represents the line through the points $(c, f(c))$ and $(d, f(d))$ and suppose that $0 \leq t \leq 1$. Then*

$$S_{cd}\Big((1-t)c + td\Big) = (1-t)f(c) + tf(d).$$

Proof. Note that the function S_{cd} is defined by

$$S_{cd}(x) = \frac{f(d) - f(c)}{d - c}(x - c) + f(c).$$

For $0 \leq$ t ≤ 1, we find that

$$\begin{aligned} S_{cd}\Big((1-t)c + td\Big) &= \frac{f(d) - f(c)}{d - c}\Big((1-t)c + td - c\Big) + f(c) \\ &= \frac{f(c) - f(d)}{c - d}(t)(d - c) + f(c) \\ &= t\Big(f(d) - f(c)\Big) + f(c) \\ &= (1-t)f(c) + tf(d). \quad \square \end{aligned}$$

Theorem 2.8. *Let f be a function defined on an interval I. Then f is convex on I if and only if the graph of f lies below its secant lines on I.*

Proof. Let f be defined on an interval I. Suppose that the graph of f lies below its secant lines on I, and let $c, d \in I$ with $c < d$. By hypothesis the inequality $f(x) \leq S_{cd}(x)$ is valid for all $x \in [c, d]$, and we note that the point $x = (1-t)c + td$ for $0 \leq t \leq 1$ is in the interval $[c, d]$ by Lemma 2.6. It follows from Lemma 2.7 that

$$f(x) = f\Big((1-t)c + td\Big) \leq S_{cd}\Big((1-t)c + td\Big) = (1-t)f(c) + tf(d).$$

Hence, the function f is convex by definition.

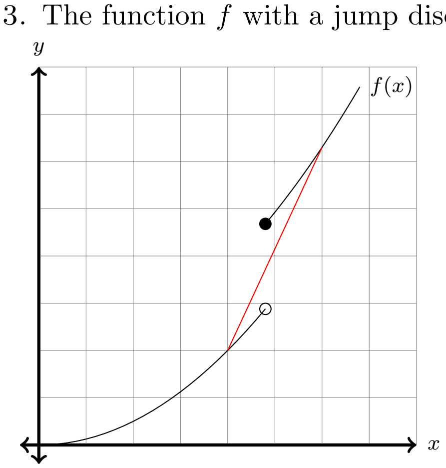
Now suppose f is convex on I, and let $c, d \in I$ with $c < d$. For each $x \in [c, d]$ there exists a $t \in [0, 1]$ such that $x = (1-t)c + td$. From the definition of convexity we have that

$$f(x) = f\Big((1-t)c + td\Big) \leq (1-t)f(c) + f(d) = S_{cd}\Big((1-t)c + td\Big).$$

This shows that the function f lies below its secant lines on I. $\square$

2.3. **Continuity.** With Theorem 2.8 we have sufficient theoretical tools to answer an important question: how well behaved are convex functions? That is, what can we say about continuity and differentiability? Recall that a function may not be continuous if it has a jump discontinuity, a removable discontinuity, or is wildly oscillating about a point. The fact that the secant lines must lie above the graph of a convex function should be convincing evidence that none of these cases are possible. See Figure 3 for an illustration of this argument with a jump discontinuity.

FIGURE 3. The function f with a jump discontinuity.



In proving this result, we will make use of Lemma 2.9, a proof of which may be found in Gordon [1].

Lemma 2.9. Let f be a convex function defined on an interval I. If a,b, and c are points in I such that $a < b < c$, then

$$\frac{f(b) - f(a)}{b - a} \leq \frac{f(c) - f(a)}{c - a} \leq \frac{f(c) - f(b)}{c - b}.$$

The inequality in Lemma 2.9 represents the slope of secant lines along the graph of f. This allows us to visualize lower and upper bounds for the slope of a line segment by choosing appropriate secant lines. This idea will be illustrated in the proof of the following theorem.

Theorem 2.10. If f is a convex function defined on an open interval (a, b), then f is continuous on (a, b).

Proof. Suppose f is convex on (a, b), and let $[c, d] \subseteq (a, b)$. Choose c_1 and d_1 such that

$$a < c_1 < c < d < d_1 < b.$$

If $x, y \in [c, d]$ with $x < y$, we have from Lemma 2.9 (see Figure 4) that

$$\frac{f(y) - f(x)}{y - x} \leq \frac{f(d) - f(y)}{d - y} \leq \frac{f(d_1) - f(d)}{d_1 - d}$$

and

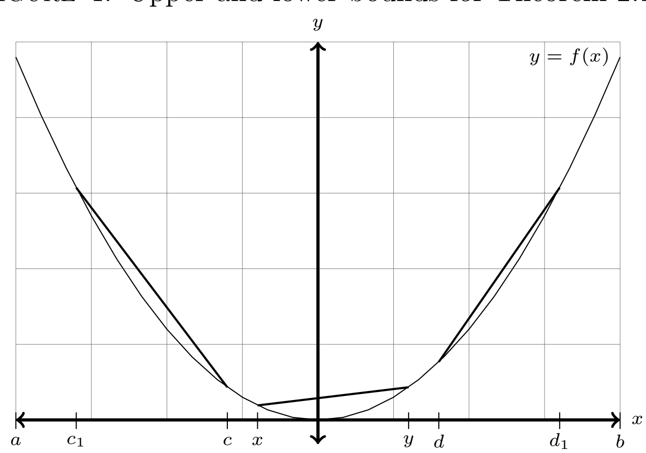
$$\frac{f(y) - f(x)}{y - x} \geq \frac{f(x) - f(c)}{x - c} \geq \frac{f(c) - f(c_1)}{c - c_1},$$

showing the set

$$\left\{ \left| \frac{f(y) - f(x)}{y - x} \right| : c \leq x < y \leq d \right\}$$

is bounded by $M > 0$. It follows $|f(y) - f(x)| \leq M|y - x|$, and therefore f is uniformly continuous on $[c, d]$. Recalling that uniform continuity implies continuity, we have shown that f is continuous on $[c, d]$. Since the interval $[c, d]$ was arbitrary, f is continuous on (a, b). [2] $\square$

FIGURE 4. Upper and lower bounds for Theorem 2.10.



The following corollary follows immediately from the fact that any closed interval may be contained in an open interval.

[2]The details of this proof are referenced from Gordon [1].

Corollary 2.11. If f is a convex function defined on $\mathbb{R}$, then f is continuous on $\mathbb{R}$.

2.4. **Differentiability.** We now consider the differentiability of a convex function. Remember that we introduced convexity in order to remove the differentiability constraint from a concave up function. Since a convex function is still geometrically related to a concave function, we may suspect they still share similar properties. The derivative of a concave up function is increasing. However, we can not readily draw an analogous conclusion for convex functions because we are not guaranteed the derivative of a convex function exists at a point. We need an additional definition to expand the concept.

Definition 2.12. Let f be defined on an interval I and let $c \in I$. The **left and right derivatives** of f at c are denoted and defined by

$$f^{-}(c) = \lim_{x \to c^{-}} \frac{f(x) - f(c)}{x - c}$$

and

$$f^{+}(c) = \lim_{x \to c^{+}} \frac{f(x) - f(c)}{x - c}$$

respectively, provided the appropriate limits exist. (If c is an endpoint of I, only one of these expressions is defined.) To say that f has **one-sided derivatives** at c means that both $f^{-}(c)$ and $f^{+}(c)$ exist.

The concept of left and right derivatives is similar to that of left and right limits; a function f is differentiable at c if and only if $f^{-}(c) = f^{+}(c)$. The following theorem provides the analogous result for convex functions. A proof may be found in Gordon [1].

Theorem 2.13. If f is a convex function defined on an open interval (a, b), then f has one-sided derivatives at each point of (a, b), $f^{-}(x) \leq f^{+}(x)$ for all $x \in (a, b)$, and the functions f^{-} and f^{+} are increasing on (a, b). Furthermore, if f is continuous on $[a, b]$, then these conclusions may be extended to the closed interval $[a, b]$.

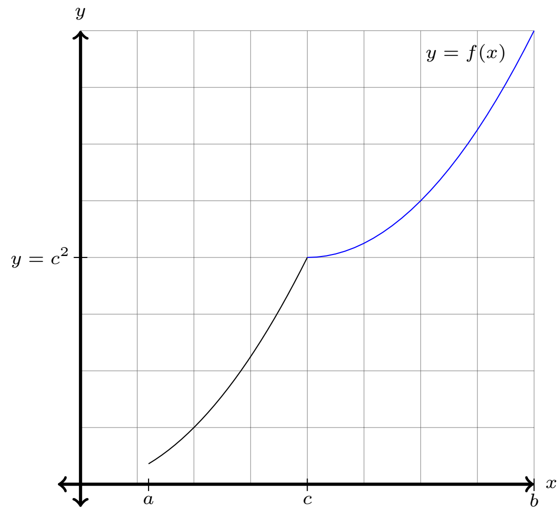
Thus, the difference between a convex and concave up function in terms of differentiablility is related to the inequality of the left and right derivatives. The following example continues our discussion of this result.

Example 2.14. Let f be a continuous function defined on $[a, b]$, let $c \in (a, b)$, and suppose that f is convex on $[a, c]$ and $[c, b]$. We will construct a function that shows that f may not be convex on $[a, b]$. The clear point of interest is where the two intervals intersect at c. If we can piecewise define a function f such that $f^{-}(x) \geq f^{+}(x)$, then we may proceed using the contrapositive of Theorem 2.13. Suppose $a > 0$, and define $f : [a, b] \to \mathbb{R}$, shown in Figure 5 by

$$f(x) = \begin{cases} x^2 & \text{for } a \le x \le c; \\ (x-c)^2 + c^2 & \text{for } c \le x \le b. \end{cases}$$

We note f is continuous and that $f^-(x) = 2c \ge 0 = f^+(x)$. By Theorem 2.13, the function f is not convex on $[a, b]$.

FIGURE 5. The function f from Example 2.14.



In the previous example, the function f fails to be convex on the interval $[a, b]$ because the condition $f^-(x) \le f^+(x)$ is not valid at the point c. The following theorem explores a special case in which f is defined by piecewise linear segments. Since the slope of each segment is constant, we will satisfy this inequality if the slope of the first line is less than the slope of the second.

Theorem 2.15. Let f be a continuous function defined on $[a, b]$, let $c \in (a, b)$, and suppose that f is convex on $[a, c]$ and $[c, b]$. If f is linear on each of the intervals $[a, c]$ and $[c, b]$ and the slope of f on $[a, c]$ is less than the slope of f on $[c, b]$, then f is convex on $[a, b]$.

Proof. Let f_1 and f_2 denote the line segment joining the points $(a, f(a))$,$(b, f(b))$ and $(b, f(a)), (c, f(c))$ respectively (see Figure 6). We will first consider the case that $a < x \le c$ and show

$$S_{ab}(x) = \frac{f(b) - f(a)}{b - a}(x - a) + f(a) \ge \frac{f(c) - f(a)}{c - a}(x - a) + f(a) = f_1.$$

By hypothesis, the slope of f_2 is greater than the slope of f_1. This means

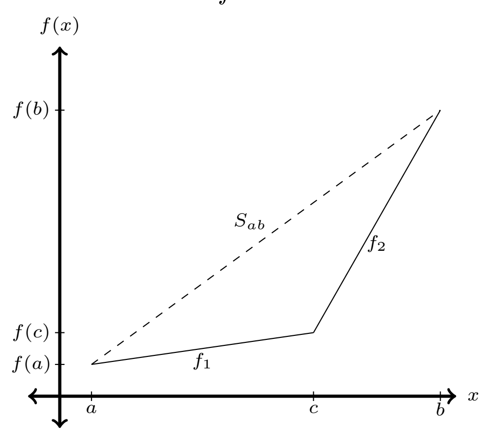
$$\frac{f(b) - f(a)}{b - c} - \frac{f(c) - f(a)}{c - a} \ge 0.$$

Noting that $\frac{b-c}{b-a} > 0$, we find

$$
\begin{aligned}
0 &\le \Big(\frac{b-c}{b-a}\Big)\Big(\frac{f(b)-f(c)}{b-c} - \frac{f(c)-f(a)}{c-a}\Big) \\
&= \Big(\frac{c-b}{b-a}\Big)\Big(\frac{f(b)-f(c)}{c-b} + \frac{f(c)-f(a)}{c-a}\Big) \\
&= \frac{f(b)-f(c)}{b-a} + \frac{c-b}{b-a}\Big(\frac{f(c)-f(a)}{c-a}\Big) \\
&= \frac{f(b)-f(c)}{b-a} + \Big(\frac{1}{b-a} - \frac{1}{c-a}\Big) f(c)-f(a) \\
&= \frac{f(b)-f(c)+f(c)-f(a)}{b-a} - \frac{f(c)-f(a)}{c-a} \\
&\le \frac{f(b)-f(a)}{b-a}(x-a) + f(b) - \frac{f(c)-f(a)}{b-a}(x-a) - f(b).
\end{aligned}
$$

It follows that $S_{ab}(x) \ge f_1$. The case in which we consider $c < x < b$ and show $S_{ab}(x) \ge f_2$ is analogous. The graph of f lies below the secant line S_{ab}, and therefore f is convex on $[a, b]$. $\square$

FIGURE 6. The function f as defined in Theorem 2.15.



Here we take an aside to prove an important result: the number of points at which a convex function defined on an open interval is not differentiable is countable. We will then return to our previous theorem in order to provide a worst case example of a convex function that is not differentiable at a countably infinite number of points.

Theorem 2.16. If f is a convex function defined on an open interval (a, b), then the set of points in (a, b) at which f is not differentiable is countable.

Proof. The function f^+ is increasing by Theorem 2.13. It can be shown that since f^+ is monotone, f^+ has a countable number of discontinuities. We will show that f is differentiable at every point when f^+ is continuous.

Let $\epsilon > 0$, and suppose f^+ is continuous at c. There exists a $\delta > 0$ such that $|f^+(x) - f^+(c)| < \epsilon$ for all $x \in (a, b)$ that satisfy $|x - c| < \delta$. Suppose $c - \delta < x < c$. By Theorem 2.13, we find $f^+(x) \leq f^-(c)$. Using this fact, we have

$$f^+(c) - \epsilon < f^+(x) \leq f^-(c) \leq f^+(c) < f^+(c) + \epsilon.$$

Since ϵ was arbitrary, it follows that $f^-(c) = f^+(c)$. Thus, f is differentiable at c. [3] □

Example 2.17. For each positive integer n, let $x_n = \frac{(2^n-1)}{2^n}$ and $y_n = \frac{n}{n+1}$. Let f be the piecewise linear function that joins the points (x_n, y_n). The function f is convex on the interval $[0.5, 1)$ and the set of points at which f is not differentiable is countably infinite. [1]

First note that the range of the sequence $\{x_n\}$ is the interval $[0.5, 1)$. Our goal will be to show that the slope of any linear segment of f is less than the slope of the subsequent segment. This amounts to proving that the following inequality is true for all positive integers $n \geq 2$:

$$\frac{y_{n+1} - y_n}{x_{n+1} - x_n} \geq \frac{y_n - y_{n-1}}{x_n - x_{n-1}}. \tag{1}$$

This will allow us to use Theorem 2.15 to prove that f is convex on $[0.5, 1)$. However, before we proceed it would be prudent to simplify inequality (1) to a more workable form. Substituting for the sequence definitions of $\{x_n\}$ and $\{y_n\}$, inequality (1) becomes

$$\frac{\frac{n+1}{n+2} - \frac{n}{n+1}}{\frac{2^{n+1}-1}{2^{n+1}} - \frac{2^n-1}{2^n}} \geq \frac{\frac{n}{n+1} - \frac{n-1}{n}}{\frac{2^n-1}{2^n} - \frac{2^{n-1}-1}{2^{n-1}}}. \tag{2}$$

Consider the left hand side of inequality (2). Simple algebra reveals that

$$\begin{aligned}
\frac{n+1}{n+2} - \frac{n}{n+1} &= \frac{(n+1)(n+1) - n(n+2)}{(n+2)(n+1)} \\
&= \frac{n^2 + 2n + 1 - n^2 - 2n}{(n+2)(n+1)} \\
&= \frac{1}{(n+2)(n+1)} \\
\frac{2^{n+1} - 1}{2^{n+1}} - \frac{2^n - 1}{2^n} &= \frac{2^n(2^{n+1} - 1) - (2^n - 1)(2^{n+1})}{(2^{n+1})(2^n)} \\
&= \frac{-2^n + 2^{n+1}}{(2^{n+1})(2^n)} \\
&= \frac{1}{2^{n+1}}.
\end{aligned}$$

[3] The details of this proof are referenced from Gordon [1].

By performing similar operations on the right hand side, it can be shown that inequality (2) reduces to

$$\frac{2^{n+1}}{(n+2)(n+1)} \geq \frac{2^n}{(n+1)(n)}.$$

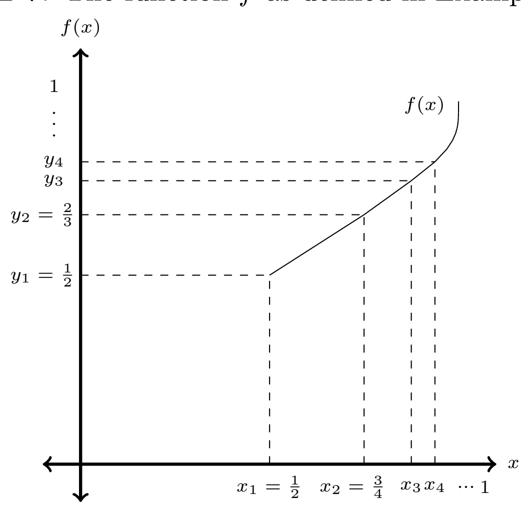
Let $n \geq 2$. We find that

$$\begin{aligned} n &\geq 2 \\ \frac{2}{n+2} &\geq \frac{1}{n} \\ \frac{2^{n+1}}{(n+2)(n+1)} &\geq \frac{2^n}{(n+1)(n)} \end{aligned}$$

Hence, inequality (2) is true for all $n \geq 2$. By Theorem 2.15 it follows that the function f is convex on [0.5,1).

Furthermore, since f is convex on $[0.5, 1)$. Theorem 2.16 states that the set of points in $(0.5, 1)$ at which f is not differentiable is countable. Extending this observation to the interval $[0.5, 1)$ is valid because we are adding at most one more point for which f is not differentiable. Thus, the set of points in $[0.5, 1)$ for which f is not differentiable is countable. The function f is illustrated in Figure 7. While f may appear smooth in the figure, it should be noted that the graph of f is defined as a countably infinite number of line segments.

FIGURE 7. The function f as defined in Example 2.17.



3. Alternative criteria for convexity

Before proceeding to applications of convexity in real analysis and economics we will first derive a useful result applicable to continuous functions. Our motivating factor is clear from Example 2.5; proving convexity from the definition may be very cumbersome. Any method to streamline this will make a welcome addition to our criterion for convexity.

Theorem 3.1. If f is continuous on (a,b) and

$$f\Big(\frac{x+y}{2}\Big) \leq \frac{1}{2}f(x) + \frac{1}{2}f(y)$$

for all $x, y \in (a,b)$, then f is convex on (a,b).

This is a rather remarkable result for two reasons. First, Theorem 3.1 simplifies the definition of convexity so that we need only consider the case in which $t = \frac{1}{2}$ for a large class of functions. This result also provides a means to approach unconventional functions that we know are continuous, an example of which will be shown following the proof.

This theorem is also notable for its long and possibly unfamiliar proof technique. We begin by leveraging our hypothesis to show that the number of terms in the numerator, and the corresponding integer divisor, can be extended to any power of two, and ultimately to any $n \in \mathbb{Z}^+$ (Lemma 3.3 and Lemma 3.5, respectively). Using these results, we can show the function satisfies the definition of convexity for rational value of t. Finally, sequences of rational numbers will be used to approximate irrational values of t.

Proof of Theorem 3.1. Each of the following lemmas will be preceded by an example in which we work through a specific case. While our algebra is not necessarily difficult, it is quite possible to be overwhelmed in the general proof.

Example 3.2. Suppose that f is continuous on (a,b) and

$$f\Big(\frac{x+y}{2}\Big) \leq \frac{1}{2}f(x) + \frac{1}{2}f(y)$$

for all $x, y \in (a,b)$. We will illustrate how the number of terms in Theorem 3.1 may be extended from two to four. By applying our hypothesis twice, we see that

$$\begin{aligned}
f\Big(\frac{x_1+x_2+x_3+x_4}{4}\Big) &= f\Big(\frac{\frac{x_1+x_2}{2}+\frac{x_3+x_4}{2}}{2}\Big) \\
&\leq \frac{1}{2}f\Big(\frac{x_1+x_2}{2}\Big) + \frac{1}{2}f\Big(\frac{x_3+x_4}{2}\Big) \\
&\leq \frac{1}{2}\Big(\frac{1}{2}f(x_1) + \frac{1}{2}f(x_2) + \frac{1}{2}f(x_3) + \frac{1}{2}f(x_4)\Big) \\
&= \frac{1}{4}f(x_1) + \frac{1}{4}f(x_2) + \frac{1}{4}f(x_3) + \frac{1}{4}f(x_4).
\end{aligned}$$

Lemma 3.3. If f is defined on (a, b) and

$$f\Big(\frac{x+y}{2}\Big) \leq \frac{1}{2}f(x) + \frac{1}{2}f(y)$$

for all $x, y \in (a, b)$, then for $n \in \mathbb{Z}^+$, f has the property that

$$f\left(\frac{1}{2^n}\sum_{i=1}^{2^n} x_i\right) \leq \frac{1}{2^n}\sum_{i=1}^{2^n} f(x_i).$$

Proof. We will use the Principle of Mathematical Induction. When $n = 1$, the inequality is valid by hypothesis, and hence $1 \in S$. Suppose that the inequality is true for some positive integer k. We then have that

$$\begin{aligned} f\Big(\frac{x_1 + x_2 + \ldots + x_{2^k} + \ldots + x_{2^{k+1}}}{2^{k+1}}\Big) &= f\left(\frac{\frac{x_1+x_2+\ldots+x_{2^k}}{2^k} + \frac{x_{2^k+1}+x_{2^k+2}+\ldots+x_{2^{k+1}}}{2^k}}{2}\right) \\ &\leq \frac{1}{2}f\Big(\frac{x_1 + x_2 + \ldots + x_{2^k}}{2^k}\Big) + \frac{1}{2}f\Big(\frac{x_{2^k+1} + x_{2^k+2} + \ldots + x_{2^{k+1}}}{2^k}\Big) \\ &\leq \frac{1}{2}\left(\frac{1}{2^k}\sum_{i=1}^{2^k} f(x_i)\right) + \frac{1}{2}\left(\frac{1}{2^k}\sum_{i=2^k+1}^{2^{k+1}} f(x_i)\right) \\ &= \frac{1}{2^{k+1}}\sum_{i=1}^{2^{k+1}} f(x_i). \end{aligned}$$

Thus, the inequality is true for $k + 1$. It follows from the Principle of Mathematical Induction that the inequality

$$f\left(\frac{1}{2^n}\sum_{i=1}^{2^n} x_i\right) \leq \frac{1}{2^n}\sum_{i=1}^{2^n} f(x_i)$$

holds for all positive integers. □

Our next goal is to show that the number of terms in Theorem 3.1 may be extended to any $m \in Z^+$. Note that $2^n \leq m < 2^{n+1}$ for some $n \in Z^+$. We will work through an example where $m = 6$.

Example 3.4. We wish to show that

$$f\left(\frac{1}{6}\sum_{i=1}^{6} x_i\right) \leq \frac{1}{6}\sum_{i=1}^{6} f(x_i).$$

Let $\bar{x}$ equal the average of $x_1, x_2, \ldots, x_6$; that is, let $\bar{x} = \frac{1}{6}\sum\limits_{i=1}^{6} x_i$. By adding $\bar{x}$ to the average of the x_i's we find that the resulting average remains unchanged. Therefore $\frac{\sum\limits_{i=1}^{6}(x_i)+\bar{x}+\bar{x}}{8} = \bar{x}$. The idea is to continue adding $\bar{x}$ until we are considering a number of terms that is a power of two. This allows us to use our previous result. We find that

$$\begin{aligned}
f(\bar{x}) &= f\left(\frac{\sum\limits_{i=1}^{6} x_i + \bar{x} + \bar{x}}{8}\right) \\
&\leq \frac{\sum\limits_{i=1}^{6} f(x_i) + f(\bar{x}) + f(\bar{x})}{8} \\
&= \frac{1}{8}\sum_{i=1}^{6} f(x_i) + \frac{2}{8} f(\bar{x}) \\
\frac{3}{4} f(\bar{x}) &\leq \frac{1}{8}\sum_{i=1}^{6} f(x_i) \\
f(\bar{x}) &\leq \frac{1}{6}\sum_{i=1}^{6} f(x_i) \\
f\left(\frac{1}{6}\sum_{i=1}^{6} x_i\right) &\leq \frac{1}{6}\sum_{i=1}^{6} f(x_i).
\end{aligned}$$

Having reached the desired result, we will now prove the general case.

Lemma 3.5. If f is defined on (a, b) and

$$f\Big(\frac{x+y}{2}\Big) \leq \frac{1}{2} f(x) + \frac{1}{2} f(y)$$

for all $x, y \in (a, b)$, then for $m \in Z^+$ the function f has the property that

$$f\left(\frac{1}{m}\sum_{i=1}^{m} x_i\right) \leq \frac{1}{m}\sum_{i=1}^{m} f(x_i).$$

Proof. Note that we have already covered the case where n is a power of two. Let $\bar{x} = \frac{1}{m}\sum\limits_{i=1}^{m} x_i$ and choose $n \in Z^+$ such that $2^n < m < 2^{n+1}$. Let

$p = 2^{n+1} - m$. Since $m + p$ is a power of two, we have by Lemma 3.3 that

$$\begin{aligned}
f(\bar{x}) &= f\left(\frac{\sum_{i=1}^{m} x_i + \sum_{i=1}^{p} \bar{x}_i}{m+p}\right) \\
&\leq \frac{\sum_{i=1}^{m} f(x_i) + \sum_{i=1}^{p} f(\bar{x}_i)}{m+p} \\
&= \frac{1}{m+p}\sum_{i=1}^{m} f(x_i) + \frac{p}{m+p} f(\bar{x}) \\
\frac{m}{m+p} f(\bar{x}) &\leq \frac{1}{m+p}\sum_{i=1}^{m} f(x_i) \\
f(\bar{x}) &\leq \frac{1}{m}\sum_{i=1}^{m} f(x_i) \\
f\left(\frac{1}{m}\sum_{i=1}^{m} x_i\right) &\leq \frac{1}{m}\sum_{i=1}^{m} f(x_i).
\end{aligned}$$

□

We now have the computational tools to prove Theorem 3.1.

Theorem 3.1. If f is continuous on (a, b) and

$$f\Big(\frac{x+y}{2}\Big) \leq \frac{1}{2} f(x) + \frac{1}{2} f(y)$$

for all $x, y \in (a, b)$, then f is convex on (a, b).

Proof. We will first prove the case where $t = \frac{m}{n}$ for $m, n \in \mathbb{Z}^+$. By Lemma 3.5 we have that

$$\begin{aligned}
f\Big((1-t)x + ty\Big) &= f\Big((1 - \frac{m}{n})x + \frac{m}{n} y\Big) \\
&= f\Big(\frac{(n-m)x + my}{n}\Big) \\
&\leq \frac{1}{n}\Big(\sum_{i=1}^{n-m} f(x) + \sum_{i=1}^{m} f(y)\Big) \\
&= \frac{n-m}{n} f(x) + \frac{m}{n} f(y) \\
&= (1-t) f(x) + t f(y).
\end{aligned}$$

Now consider the general case where $t \in [0, 1]$. Let $t \in [0, 1]$, and let $\{t_n\}$ be a sequence of rational numbers that converges to t. By our earlier result we find that

$$f\Big((1-t_n)x + t_n y\Big) \leq (1-t_n) f(x) + t_n f(y)$$

for every positive integer n. Since f is continuous, we have

$$\lim_{n\to\infty} f\Big((1-t_n)x + t_n y\Big) = f\Big((1-t)x + ty\Big)$$

and

$$\lim_{n\to\infty} \Big((1-t_n)f(x) + t_n f(y)\Big) = (1-t)f(x) + tf(y).$$

Therefore $f\Big((1-t)x+ty)\Big) \le (1-t)f(x)+tf(y)$. It follows that f is convex by definition. □

The benefit of Theorem 3.1, as we alluded to earlier, is that it makes determining convexity for continuous functions much easier. By considering only the $t = \frac{1}{2}$ case in the definition of a convex function, we may gain insight into algebraic techniques that would otherwise be obfuscated. The following theorem is an excellent example of how our result simplifies an otherwise difficult proof.

Theorem 3.6. Suppose that f is increasing on an interval $[a, b)$ and define a function ϕ by

$$\phi(x) = \int_a^x f.$$

The function ϕ is convex on (a, b).

Proof. By the Fundamental Theorem of Calculus, the function ϕ is continuous on the interval (a, b). Let $[c, d] \subseteq (a, b)$, and let $m = \frac{c+d}{2}$. We must show

$$\phi(m) \quad \le \quad \frac{1}{2}\phi(c) + \frac{1}{2}\phi(d)$$

or equivalently,

$$2\int_a^m f \quad \le \quad \int_a^c f + \int_a^d f.$$

Since f is increasing, the number $f(m)$ is an upper and lower bound for f on the intervals $[c, m]$ and $[m, d]$ respectively. Noting that m is the midpoint of c and d, we have

$$\int_c^m f \le f(m)(m-c) = f(m)(d-m) \le \int_m^d f.$$

From this statement we find that

$$\begin{aligned}\int_a^c f + \int_a^d f &= \int_a^c f + \int_a^m f + \int_m^d f \\ &\geq \int_a^c f + \int_a^m f + \int_c^m f \\ &= \int_a^m f + \int_a^m f \\ &= 2\int_a^m f.\end{aligned}$$

By Theorem 3.1, the function ϕ is convex on $[a, b)$.

We note in passing that if the increasing function f were continuous, we may have used the Fundamental Theorem of Calculus to show that $\phi'' > 0$, and hence that ϕ is concave up on. Using our previous observations about secant lines, we know that a concave up function is also convex. Since a convex function may not be differentiable, however, the converse of this statement is not true in general. □

4. Applications to Real Analysis

As we explored the behavior of convex functions we found that several of the results appealed to the fact that a convex function lies below its secant lines. A few of the specific applications to real analysis, however, instead refer to the definition of convexity. The result is that the proofs become notably more algebraic in nature. The extra work involved should not be a deterrent from the notable results. In this section, we will first discuss the relationship between a decreasing convex function and the sequence of right endpoint approximations to the area under the curve. We will then explore a curious exception involving pointwise convergence.

4.1. **Right Endpoint Approximation.** The Riemann integral is inextricably tied to the concept of determining the area under a function f on a closed interval I. Generally, the interval I is first divided into a number of subintervals from which a point, referred to as a tag, is chosen from each. The function evaluated at each tag, multiplied by the length of the respective interval, is a rectangle estimate of the area under f on that subinterval. The sum of these rectangles provides an estimate to the area under f on the entire interval.

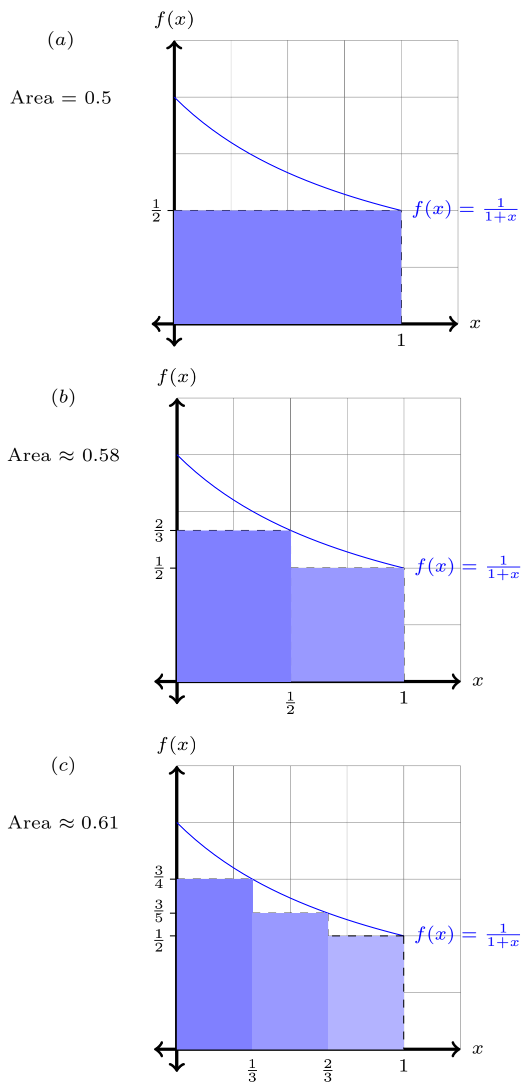
Right endpoint approximation is the specific case in which the right endpoint of each subinterval is chosen as the tag. The sequence of right endpoint approximations to a positive, continuous, decreasing function defined on $[0, 1]$ is given by

$$R_n = \sum_{i=1}^{n} f\left(\frac{i}{n}\right)\frac{1}{n}.$$

While each term of the sequence will underestimate the area under the curve, it can be shown that R_n converges to the Riemann integral. However, it is not necessarily true that each iteration provides a better estimate; that is, we need additional criteria to guarantee the sequence of right endpoint approximations $\{R_n\}$ is increasing. Convexity satisfies this requirement.

The first three values of R_n are illustrated in Figure 8. Note that since R_2 is simply a bisection of R_1, the total area approximated must be greater. With this argument we see that the subsequence R_{2^n} is increasing. However, it is not immediately obvious that $R_2 \leq R_3$.

FIGURE 8. The sequence $\{R_n\}$ respresented for a convex function: (a) $n = 1$, (b) $n = 2$, (c) $n = 3$.



Lemma 4.1. If f is convex on $[a, b]$ and i and n are positive integers, then the inequality

$$f\left(\frac{i}{n}\right) \leq \left(\frac{n-i}{n}\right) f\left(\frac{i}{n+1}\right) + \frac{i}{n} f\left(\frac{i+1}{n+1}\right)$$

is valid.

Proof. We find that

$$\frac{i}{n} = \frac{in+i}{n(n+1)} = \frac{in-i^2+i^2+i}{n(n+1)} = \frac{(n-i)i+i(i+1)}{n(n+1)} = \left(\frac{n-i}{n}\right)\left(\frac{i}{n+1}\right)+\left(\frac{i}{n}\right)\left(\frac{i+1}{n+1}\right),$$

and noting that f is convex, we have that

$$f\left(\frac{i}{n}\right) \leq f\left(\left(\frac{n-i}{n}\right)\left(\frac{i}{n+1}\right)\right) + f\left(\left(\frac{i}{n}\right)\left(\frac{i+1}{n+1}\right)\right).$$

$\square$

Theorem 4.2. If $f : [0, 1] \to \mathbb{R}$ is convex, positive, and decreasing on $[0, 1]$, then the sequence $\{R_n\}$ of right endpoint approximations to $\int_0^1$ is increasing.

Proof. We will show that $R_n \leq R_{n+1}$. We have from Lemma 4.1 that

$$\begin{aligned}
R_n &= \frac{1}{n}\sum_{i=1}^{n} f\Big(\frac{i}{n}\Big) \\
&\leq \frac{1}{n}\sum_{i=1}^{n}\Bigg(\Big(\frac{n-i}{n}\Big)f\Big(\frac{i}{n+1}\Big)+\frac{i}{n}f\Big(\frac{i+1}{n+1}\Big)\Bigg) \\
&= \frac{1}{n}\sum_{i=1}^{n}\Big(\frac{n-i}{n}\Big)f\Big(\frac{i}{n+1}\Big)+\frac{1}{n}\sum_{i=1}^{n+1}\frac{i-1}{n}f\Big(\frac{i}{n+1}\Big) \\
&= \frac{1}{n}\Bigg(\sum_{i=1}^{n}\frac{n-1}{n}f\Big(\frac{i}{n+1}\Big)\Bigg)+\frac{1}{n}f(1) \\
&= \sum_{i=1}^{n}\frac{1}{n+1}f\Big(\frac{i}{n+1}\Big)+\sum_{i=1}^{n}\Big(\frac{n-1}{n^2}-\frac{1}{n+1}\Big)f\Big(\frac{i}{n+1}\Big)+\frac{1}{n}f(1) \\
&= \sum_{i=1}^{n}\frac{1}{n+1}f\Big(\frac{i}{n+1}\Big)+\sum_{i=1}^{n}\Big(\frac{(n-1)(n+1)-n^2}{(n^2)(n+1)}\Big)f\Big(\frac{i}{n+1}\Big)+\frac{1}{n}f(1) \\
&= \sum_{i=1}^{n}\frac{1}{n+1}f\Big(\frac{i}{n+1}\Big)-\frac{1}{(n^2)(n+1)}\sum_{i=1}^{n}f\Big(\frac{i}{n+1}\Big)+\frac{1}{n}f(1)
\end{aligned}$$

and using the fact that f is a decreasing function,

$$\begin{aligned}
&\leq \sum_{i=1}^{n}\frac{1}{n+1}f\Big(\frac{i}{n+1}\Big)-\frac{1}{(n^2)(n+1)}\sum_{i=1}^{n}f(1)+\frac{1}{n}f(1) \\
&= \sum_{i=1}^{n}\frac{1}{n+1}f\Big(\frac{i}{n+1}\Big)+f(1)\Big(\frac{1}{n}-\frac{1}{n(n+1)}\Big) \\
&= \sum_{i=1}^{n}\frac{1}{n+1}f\Big(\frac{i}{n+1}\Big)+f(1)\frac{1}{n+1} \\
&= \frac{1}{n+1}\Big(\sum_{i=1}^{n}f\Big(\frac{i}{n+1}\Big)+f\Big(\frac{n+1}{n+1}\Big)\Big) \\
&= \frac{1}{n+1}\sum_{i=1}^{n+1}f\Big(\frac{i}{n+1}\Big) \\
&= R_{n+1}.
\end{aligned}$$

□

4.2. **Pointwise Convergence.** A sequence of functions $\{f_n\}$ converges pointwise to f on an interval I if the sequence $\{f_n(x)\}$ converges to $f(x)$ for each $x \in I$. One of the striking weaknesses of pointwise convergence is that certain properties of $\{f_n\}$ may fail to carry over to the limit function. That is to say, even if each of the functions f_n is bounded, continuous, or Riemann Integrable, it does not necessarily follow that f shares this property. For counterexamples demonstrating this fact, see Gordon [1].

One may wonder if a sequence of convex functions shares a similar fate. The answer may be surprising. Pointwise convergence is indeed sufficient to guarantee that convexity transfers to the limit function.

Theorem 4.3. Let $\{f_n\}$ be a sequence of convex functions defined on an interval I. If $\{f_n\}$ converges pointwise on I to a function f then f is convex on I.

Proof. Suppose that f_n is a sequence of convex functions defined on I that converges pointwise on I to a function f. This means that for $[c,d] \subseteq I$

$$\begin{aligned} f\Big((1-t)c+td\Big) &= \lim_{n\to\infty} f_n\Big((1-t)c+td\Big) \\ &\leq \lim_{n\to\infty} \Big((1-t)f_n(c)+tf_n(d)\Big) \\ &= \lim_{n\to\infty}(1-t)f_n(c)+\lim_{n\to\infty} tf_n(d) \\ &= (1-t)f(c)+tf(d). \end{aligned}$$

It follows that the function f is convex on I by definition.

□

5. Economic Modeling of Uncertainty

The purpose of this section will be to explore an application of convexity to economic consumer thoery. Economics is the study of how agents choose to allocate scarce resources. Given a set of rationality assumptions, economists hope to model and predict the choices of individuals. Since the 1960's, economists have been increasingly concerned with developing more rigorous mathematical models. The particular model we will be discussing is that of consumer behavior under situations of uncertainty. Before addressing our application to convexity, we will briefly discuss its theoretical foundation. [4]

5.1. Consumer Preference.

Let $x_i \in \mathbb{R}_+$ represent the number of units of good i. We assume that there is a finite number of goods n. The vector $\mathbf{x} = (x_1, \ldots, x_n)$ describes a particular distribution of goods, and is referred to as a **consumption bundle**. Note that a consumption bundle is a point $\mathbf{x} \in \mathbb{R}^n_+$. In situations of certainty, consumer theory looks at the preference of individuals over the set of consumption bundles X. An axiomatic assumption is that an individual is able to apply a preference relation to a collection of bundles. That is, given two bundles $\mathbf{x}_1$ and $\mathbf{x}_2$, an individual

[4]The axioms, definitions, and theorems in the section are referenced from Jehle and Reny's Advanced Microeconomic Theory.

has a preference or is indifferent between the two. The symbol $\succsim$ represents a binary relation defined on the set of consumption bundles X. For example, the statement $\mathbf{x}_1 \succsim \mathbf{x}_2$ means that $\mathbf{x}_1$ is at least as preferable as $\mathbf{x}_2$ to a particular consumer. We extend this concept to consumer behavior in situations of uncertainty by considering a vector of probability weighted outcomes called a **gamble**.

Let $A = \{a_1, \ldots, a_n\}$ be a finite set of certain **outcomes**. Depending on the situation, an outcome may be an amount of money, an event, or anything else. A **simple gamble** assigns a probability p_i, using the symbol $\circ$ as notation, to each of the outcomes a_i in A such that

$$g = (p_1 \circ a_1, \ldots, p_n \circ a_n) \text{ where } p_i \geq 0, \sum_{i=1}^{n} p_i = 1.$$

For example, the gamble in which a coin is flipped and an individual recieves \$10 if heads, and \$0 if tails, may be represented as $(\frac{1}{2} \circ 10, \frac{1}{2} \circ 0)$. The set of simple gambles $\mathcal{G}_S$ on A is given by

$$\mathcal{G}_S \equiv \{(p_1 \circ a_1, \ldots, p_n \circ a_n) |\ p_i \geq 0, \sum_{i=1}^{n} p_i = 1\}.$$

There are six axioms of choice under uncertainty; however, for the scope of this paper, we will only be concerned with two. The interested reader should refer to Jehle and Reny [2]. We assume that a consumer is able to apply a preference relation $\succsim$ over $\mathcal{G}_S$, just as they are able to over consumption bundles. The symbols $\sim$ and $\succ$ denote indifference and strict preferences, respectively.

Axiom 1 (Completeness)**.** For any two gambles, g and g' in $\mathcal{G}_S$,

$$g \succsim g' \text{ or } g' \succsim g.$$

Axiom 2 (Transitivity)**.** For any three gambles g, g', g'' in $\mathcal{G}_S$,

$$\text{if } g \succsim g' \text{ and } g' \succsim g'', \text{ then } g \succsim g''.$$

5.2. **Von Neumann-Morgenstern Utility Functions.** It may be difficult to analyze the behavior of individuals when their preference information is expressed over a set of vectors. We therefore introduce the concept of a utility function. In general, a utility function is a transformation from a set of objects to the set of real numbers that preserves consumer preferences. In situations of certainty, the domain of a utility function is the set of consumption bundles. The corresponding domain of the utility function u in situations of uncertainity is the set of gambles $\mathcal{G}_S$. We also say that u has the expected utility property.

Definition 5.1. The utility function $u : \mathcal{G}_S \to \mathbb{R}$ has the **expected utility property** if, for every $g \in \mathcal{G}_S$,

$$u(g) = \sum_{i=1}^{n} p_i u(a_i)$$

where $(p_1 \circ a_1, \ldots, p_n \circ a_n)$ is the simple gamble induced by g.

Hence, the utility of a gamble is simply the probability weighted utilities of the outcomes a_i. Note that $u(a_i)$ is simply the utility function evaluated at a gamble in which the probability assigned to the outcome a_i is equal to one. A utility function in situations of uncertainty is refered to as a von Neumann-Morgenstern (VNM) utility function. It is named after economists John von Neumann and Oskar Morgenstern who proved its existence from four of the axioms of choice under uncertainty. A proof may be found in Jehle and Reny [2].

Theorem 5.2. Let preferences $\succsim$ over gambles in $\mathcal{G}_S$ satisfy the six axioms of choice under uncertainty. Then there exists a utility function $u : \mathcal{G}_S \to \mathbb{R}$ representing $\succsim$ on $\mathcal{G}_S$, such that u has the expected utility property.

5.3. **Risk.** One of the defining characteristics of the VNM utility function is the preservation of consumer preferences over $\mathcal{G}_S$. This result allows us to visualize consumer behavior. In particular, we will look at how an individual's attitude towards risk is shown by the VNM utility function.

For illustrative purposes, we will consider the case in which our outcomes are some units of wealth w_i. The expected value of a gamble g is given by

$$E(g) = \sum_{i=1}^{n} p_i w_i.$$

Intuitively, a risk averse individual is one who prefers to minimize their exposure to uncertainty. We would therefore expect a risk averse individual to prefer the expected value of a gamble instead of accepting the uncertainty inherently associated with a gamble. The following definitions reflect this idea.

Definition 5.3. Let the function u represent an individual's VNM utility function for gambles over wealth. For the gamble g, an individual is

(1) **risk averse at g** if $u\Big(E(g)\Big) > u(g)$.
(2) **risk neutral at g** if $u\Big(E(g)\Big) = u(g)$.
(3) **risk loving at g** if $u\Big(E(g)\Big) < u(g)$.

The following hypothetical gamble illustrates risk loving behavior.

Example 5.4. Consider the following gamble

$$g_1 \quad = \quad (0.75 \circ \$700, 0.25 \circ \$1900).$$

The expected value of the gamble is given by

$$E(g_1) = p_1 w_1 + p_2 w_2 = (0.75)(\$700) + (0.25)(\$1900) = \$1000.$$

By the definition of the expected utility property, the utility of the gamble is given by the individual's VNM utility function evaluated at each of the outcomes and weighted by the respective probabilities

$$u(g_1) = p_1 u(w_1) + p_2 u(w_2) = .75u(\$700) + .25u(\$1900).$$

Suppose that a risk loving individual over $\mathcal{G}_S$ is offered g_1. By our definition, we have that

$$u(\$1000) = u(E(g_1)) < u(g_1) = .75u(\$700) + .25u(\$1900).$$

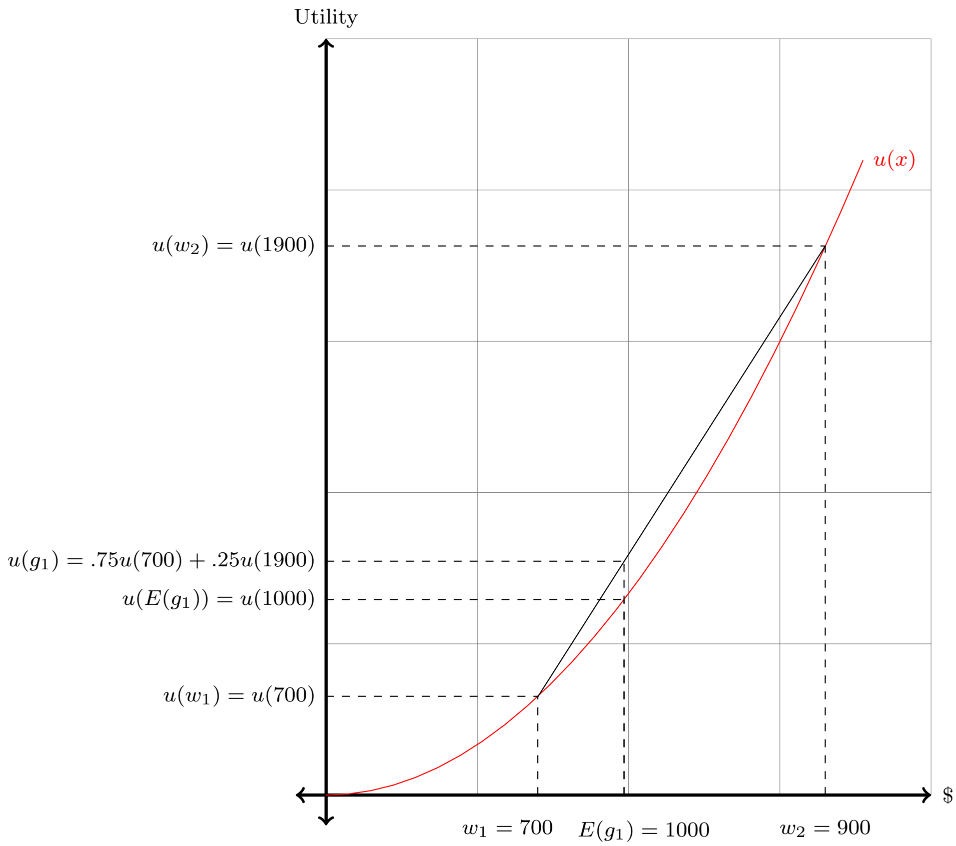
The individual's preference of the risk inherent in g_1 is illustrated in Figure 9. Note that the individual's VNM utility function u appears to be convex on $\mathbb{R}^+$.

We find intuition for this observation by recalling that the graph of a convex function lies below its secant lines. In the case of a gamble with two outcomes, it can be shown that $u(g_1)$ corresponds to a point along the secant line drawn from $(w_1, u(w_1))$ to $(w_2, u(w_2))$. In particular, we find that

$$S_{w_1 w_2}(E(g_1)) = u(g_1).$$

With reference to Figure 9, we note that a convex VNM utility function satisfies the requirement for risk loving behavior that $u(E(g_1)) < u(g_1)$.

FIGURE 9. Illustration of the gamble presented in Example 5.4.



5.4. **Jensen's Inequality.** We now return to the overlying topic of convexity. As we alluded to at the conclusion of our previous example, it can be shown that a VNM utility function that preserves risk loving behavior is convex. This follows quickly from Jensen's Inequality, a result that we prove below.

Lemma 5.5. For any points $x_1, x_2, \ldots, x_n$ in (a, b) and for any nonnegative numbers $p_1, p_2, \ldots, p_n$, we have

$$a < \frac{p_1x_1 + p_2x_2 + \cdots + p_nx_n}{p_1 + p_2 + \cdots + p_n} < b.$$

Proof. Let $x_m = \min\{x_1, x_2, \ldots, x_n\}$ and $x_M = \max\{x_1, x_2, \ldots, x_n\}$. We find that

$$\frac{p_1x_1 + p_2x_2 + \cdots + p_nx_n}{p_1 + p_2 + \cdots + p_n} \leq \frac{p_1x_M + p_2x_M + \cdots + p_nx_M}{p_1 + p_2 + \cdots + p_n} = x_M\left(\frac{p_1 + p_2 + \cdots + p_n}{p_1 + p_2 + \cdots + p_n}\right) = x_M < b,$$

and

$$\frac{p_1x_1 + p_2x_2 + \cdots + p_nx_n}{p_1 + p_2 + \cdots + p_n} \geq \frac{p_1x_m + p_2x_m + \cdots + p_nx_m}{p_1 + p_2 + \cdots + p_n} = x_m\left(\frac{p_1 + p_2 + \cdots + p_n}{p_1 + p_2 + \cdots + p_n}\right) = x_m > a.$$

$\square$

Theorem 5.6 (Jensen's Inequality)**.** A function f is convex on (a, b) if and only if for any points $x_1, x_2, \ldots, x_n$ in (a, b) and for any nonnegative numbers $p_1, p_2, \ldots, p_n$, the inequality

$$f\Big(\frac{p_1x_1 + p_2x_2 + \cdots + p_nx_n}{p_1 + p_2 + \cdots + p_n}\Big) \leq \frac{p_1f(x_1) + p_2f(x_2) + \cdots + p_nf(x_n)}{p_1 + p_2 + \cdots + p_n}$$

is valid.

Proof. Suppose that a function f is convex on (a, b). We will use the Principle of Mathematical Induction. For any points $x_1, x_2, \ldots, x_n$ in (a, b) and for any nonnegative numbers $p_1, p_2, \ldots, p_n$, consider the inequality

$$f\Big(\frac{p_1x_1 + p_2x_2 + \cdots + p_nx_n}{p_1 + p_2 + \cdots + p_n}\Big) \leq \frac{p_1f(x_1) + p_2f(x_2) + \cdots + p_nf(x_n)}{p_1 + p_2 + \cdots + p_n}.$$

When $n = 1$, we have that

$$f\big(\frac{p_1x_1}{p_1}\big) = f(x_1) \leq \frac{p_1f(x_1)}{p_1} = f(x_1).$$

Suppose that the inequality is true for some positive integer k. This may be expressed in summation notation as

$$f\left(\frac{\sum_{i=1}^{k+1} p_ix_i}{\sum_{i=1}^{k+1} p_i}\right) \quad = \quad f\left(\frac{\sum_{i=1}^{k} p_ix_i}{\sum_{i=1}^{k+1} p_i} + \frac{p_{k+1}x_{k+1}}{\sum_{i=1}^{k+1} p_i}\right)$$

Since $\frac{\sum_{i=1}^{k} p_i x_i}{\sum_{i=1}^{k} p_i}$ is in (a, b) by Lemma 5.5, we may use the definition of a convex function and the induction hypothesis. Referring to the definition of a convex function, it may be helpful to note

$$t = \frac{\sum_{i=1}^{k} p_i x_i}{\sum_{i=1}^{k+1} p_i} \text{ and } (1-t) = \frac{p_{k+1}}{\sum_{i=1}^{k+1} p_i}.$$

We find that

$$\begin{aligned}
f\left(\frac{\sum_{i=1}^{k+1} p_i x_i}{\sum_{i=1}^{k+1} p_i}\right) &= f\left(\frac{\sum_{i=1}^{k} p_i x_i}{\sum_{i=1}^{k} p_i} \cdot \frac{\sum_{i=1}^{k} p_i}{\sum_{i=1}^{k+1} p_i} + \frac{p_{k+1} x_{k+1}}{\sum_{i=1}^{k+1} p_i}\right) \\
&\leq f\left(\frac{\sum_{i=1}^{k} p_i x_i}{\sum_{i=1}^{k} p_i}\right) \frac{\sum_{i=1}^{k} p_i}{\sum_{i=1}^{k+1} p_i} + f(x_{k+1}) \frac{p_{k+1}}{\sum_{i=1}^{k+1} p_i} \\
&\leq \frac{\sum_{i=1}^{k} p_i f(x_i)}{\sum_{i=1}^{k} p_i} \cdot \frac{\sum_{i=1}^{k} p_i}{\sum_{i=1}^{k+1} p_i} + \frac{f(p_{k+1}) p_{k+1}}{\sum_{i=1}^{k+1} p_i} \\
&= \frac{\sum_{i=1}^{k} p_i f(x_i)}{\sum_{i=1}^{k+1} p_i} + \frac{f(p_{k+1}) p_{k+1}}{\sum_{i=1}^{k+1} p_i} \\
&= \frac{\sum_{i=1}^{k+1} p_i f(x_i)}{\sum_{i=1}^{k+1} p_i},
\end{aligned}$$

thus showing that the inequality is true for $k+1$. It follows from the Principle of Mathematical Induction that the inequality

$$f\Big(\frac{p_1 x_1 + p_2 x_2 + \cdots + p_n x_n}{p_1 + p_2 + \cdots + p_n}\Big) \leq \frac{p_1 f(x_1) + p_2 f(x_2) + \cdots + p_n f(x_n)}{p_1 + p_2 + \cdots + p_n}$$

is true for all positive integers, for any points $x_1, x_2, \ldots, x_n$ in (a, b), and for any nonnegative numbers $p_1, p_2, \ldots, p_n$.

Now suppose that for any points $x_1, x_2, \ldots, x_n$ in (a, b) and for any nonnegative numbers $p_1, p_2, \ldots, p_n$, the inequality

$$f\Big(\frac{p_1 x_1 + p_2 x_2 + \cdots + p_n x_n}{p_1 + p_2 + \cdots + p_n}\Big) \leq \frac{p_1 f(x_1) + p_2 f(x_2) + \cdots + p_n f(x_n)}{p_1 + p_2 + \cdots + p_n}$$

is valid. For $0 \leq t \leq 1$, let $p_1 = t$, $p_2 = (1-t)$, and $p_3 = \ldots = p_n = 0$. We find that

$$f\Big(t x_1 + (1-t) x_2\Big) \leq t f(x_1) + (1-t) f(x_2)$$

Hence, the function f is convex by definition. □

Theorem 5.7. An individual is risk loving over gambles involving nonnegative wealth levels if and only if their VNM utility function is strictly convex on $\mathbb{R}^+$.

Proof. Suppose that an individual is risk loving over gambles involving nonnegative wealth levels, let u represent their VNM utility function, and let

$g \in \mathcal{G}_S$. This means that $u\Big(E(g)\Big) < u(g)$, which can be equivalently expressed as

$$u\Big(p_1w_1 + p_2w_2 + \ldots + p_nw_n\Big) < p_1u(w_1) + p_2u(w_2) + \ldots + p_nu(w_n).$$

Noting that $\sum_{i=1}^{n} p_i = 1$, it follows from Jensen's Inequality that the function u is convex on $\mathbb{R}^+$.

Now suppose that an individual's VNM utility function u is convex on $\mathbb{R}^+$ and let $g \in \mathcal{G}_S$. Jensen's Inequality states that the following inequality

$$u\Big(p_1w_1 + p_2w_2 + \ldots + p_nw_n\Big) < p_1u(w_1) + p_2u(w_2) + \ldots + p_nu(w_n)$$

holds for each $w_i \in \mathbb{R}^+$. Hence, the individual is risk loving at g.

We have thus shown that an individual is risk loving over gambles involving nonnegative wealth levels if and only if their VNM utility function is strictly convex on $\mathbb{R}^+$. $\square$

Conclusion

Using convexity to discuss risk behavior has allowed for greater appreciation of the underlying economic theory. In the same way, our theoretical analysis of convex functions that began with geometrical observations led to rewarding topics in real analysis. We conclude our discussion of convex functions with a survey of areas of further research. With additional time, we may have explored the relationship between convex functions and metric preserving functions. For those interested in applications of convexity to economics, one may reference Jehle and Reny's discussion of indifference curve analysis in [2]. Furthermore, we note that convex functions are inextricably related to convex sets. The study of convex sets is essential in linear optimization, as well as graduate level economic proofs.

Acknowledgements

Sincere thanks to Price Hardman for his helpful comments on several drafts of the paper, to Professor Gordon and Professor Balof for their advisement and careful attention to details, to Professor Schueller, for without his support with TikZ this paper would be notably less interesting, and to Professor Crouter for her support in the economics section. Finally, I would like to thank all members of the Whitman Swim Team, past and present, for their daily support and encouragement.

References

[1] Russel A. Gordon, *Real Analysis: A First Course,* (New York: Addison-Wesley, 2002).
[2] Geoffrey A. Jehle and Philip J. Reny, *Advanced Microeconomic Theory,* (New York: Addison-Wesley, 2001).