Lower semicontinuous and Convex Functions

James K. Peterson

Department of Biological Sciences and Department of Mathematical Sciences
Clemson University

October 6, 2017

Outline

- We are now going to expand our treatment of functions a bit and look at the class of **lower semicontinuous functions** and the class of **convex** functions.

- We are now going to expand our treatment of functions a bit and look at the class of **lower semicontinuous functions** and the class of **convex** functions.
- These functions gives us some new insights into how we can try to find extreme values of functions even when there is no compactness.

- We are now going to expand our treatment of functions a bit and look at the class of **lower semicontinuous functions** and the class of **convex** functions.
- These functions gives us some new insights into how we can try to find extreme values of functions even when there is no compactness.
- The function $|x|$ clearly has an absolute minimum over $\Re$ of value 0 and its domain is not compact. Note the function $f(x) = |x|$ does not have a derivative at 0 but the Left hand derivative at 0 is -1 and the right hand derivative is 1. It turns out $|x|$ is a convex function and we can define an extension of the idea of derivative, the **subdifferential** ∂f which here would be $\partial f(0) = [-1, 1]$, a compact set! Note also as $|x| \rightarrow \infty$, $f(x) \rightarrow \infty$ too. Also note $0 \in \partial f(0)$ which is like the condition that extreme values may occur when the derivative is zero.

- We are now going to expand our treatment of functions a bit and look at the class of **lower semicontinuous functions** and the class of **convex** functions.
- These functions gives us some new insights into how we can try to find extreme values of functions even when there is no compactness.
- The function $|x|$ clearly has an absolute minimum over $\Re$ of value 0 and its domain is not compact. Note the function $f(x) = |x|$ does not have a derivative at 0 but the Left hand derivative at 0 is -1 and the right hand derivative is 1. It turns out $|x|$ is a convex function and we can define an extension of the idea of derivative, the **subdifferential** ∂f which here would be $\partial f(0) = [-1, 1]$, a compact set! Note also as $|x| \to \infty$, $f(x) \to \infty$ too. Also note $0 \in \partial f(0)$ which is like the condition that extreme values may occur when the derivative is zero.
- The function $f(x) = x^2$ also has an absolute minimum of value 0 where $f'(0) = 0$. It also satisfies $|x| \to \infty$ implies $f(x) \to \infty$, it is a convex function and its domain is not compact.

Let's start with a new result with continuous functions in the spirit of the two examples we just used.

Theorem

Let Ω be a nonempty unbounded closed set of real numbers and let $f : \Omega \to \Re$ be continuous. Assume $f(x) \to \infty$ when $|x| \to \infty$. Then $\inf(f(\Omega))$ is finite and there is a sequence $(x_n) \subset \Omega$ so that $x_n \to x_0 \in \Omega$ and $f(x_n) \to f(x_0) = \inf(f(\Omega))$.

Proof

Let $\alpha = \inf(f(\Omega))$. First, there is a sequence $(y_n = f(x_n)) \subset \Omega$ that converges to α. If $(x_n) \subset \Omega$ satisfies $|x_n| \to \infty$, we would have $f(x_n) \to \infty$. Then, we would know $\inf(f(\Omega)) = \infty$ implying $f(x) \geq \infty$ for all x in Ω. But we assumed f is finite on Ω, so this is not possible. Hence, (x_n) must be bounded. Thus, by the Bolzano Weierstrass Theorem, there is a subsequence (x_n^1) of (x_n) converging to x_0. This means x_0 is either an accumulation point in Ω or a boundary point of Ω.

Proof

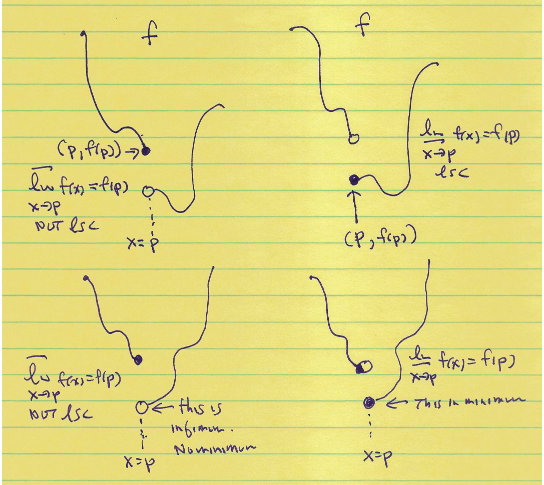
Since Ω is closed, Ω contains its boundary points. So we know $x_0 \in \Omega$. Since f is continuous at x_0, we must have $f(x_n^1) \to f(x_0)$ as $n^1 \to \infty$. Since $f(x_n) \to \alpha$ we also have $f(x_n^1) \to \alpha$. Thus, $f(x_0) = \alpha$ and we have shown f has an absolute minimum at the point x_0. $\square$

Note, the condition that $f(x) \to \infty$ when $|x| \to \infty$ allows us to bound the (x_n) sequence. This is how we get around the lack of compactness in the domain. We can relax the continuity assumption too. We can look at functions which are **lower semicontinuous**.

Definition

Let $f : dom(f) \to \Re$ be finite. We say f is **lower semicontinuous** at p if $\underline{\lim}\, f(p) = f(p)$. By definition, this means for all sequences (x_n) with $x_n \to p$ and $f(x_n) \to a$, we have $\lim_{n \to \infty} f(x_n) \geq f(p)$.

Here are two pairs of functions which show what the condition of lower semicontinuity means graphically.



Let's relax our continuity condition into lower semicontinuity for the theorem we just proved.

Theorem

Let Ω be a nonempty unbounded closed set of real numbers and let $f : \Omega \rightarrow \Re$ be lower semicontinuous. Assume $f(x) \rightarrow \infty$ when $|x| \rightarrow \infty$. Then $\inf(f(\Omega))$ is finite and there is a sequence $(x_n) \subset \Omega$ so that $x_n \rightarrow x_0 \in \Omega$ and $f(x_n) \rightarrow f(x_0) = \inf(f(\Omega))$.

Proof

Again, let $\alpha = \inf(f(\Omega))$. There is a sequence $(y_n = f(x_n)) \subset \Omega$ that converges to α. The arguments we just used still show (x_n) must be bounded. Thus, by the Bolzano Weierstrass Theorem, there is a subsequence (x_n^1) of (x_n) converging to x_0 and since Ω is closed, x_0 is in Ω. Since f is lower semicontinuous at x_0, we must have $\lim_{n^1 \rightarrow \infty} f(x_n^1) \geq f(x_0)$. Since $f(x_n) \rightarrow \alpha$ we also have $f(x_n^1) \rightarrow \alpha$. Thus, $\alpha \geq f(x_0)$. But $f(x_0) \geq \alpha$ by the definition of an infimum. Thus, $f(x_0) = \alpha = \inf(f(\Omega))$ and f has an absolute minimum at x_0.

The **graph** of a function $f : dom(f) \rightarrow \Re$ is a subset of $dom(f) \times \Re$.

$$gr(f) = \{(x, a) : x \in dom(f) \text{ and } f(x) = a\}$$

and the **epigraph** of f is everything lying above the graph.

$$epi(f) = \{(x, b) : x \in dom(f) \text{ and } f(x) \leq b\}$$

We can now define what we mean by a convex function.

Definition

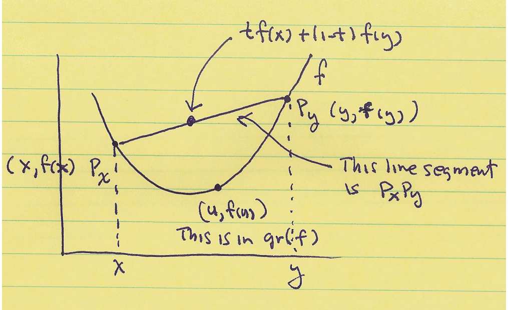
Let I be a nonempty interval of $\Re$. A function $f : I \rightarrow \Re$ is **convex** on I if

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y), \ \forall x, y \in I \text{ and } \forall t \in (0, 1)$$

f is **strictly convex** if

$$f(tx + (1-t)y) < tf(x) + (1-t)f(y), \ \forall x, y \in I \text{ and } \forall t \in (0, 1)$$

Let P_x be the point $(x, f(x))$ and P_y be the point $(y, f(y))$. The f is convex **means** the point $(u, f(u))$ on the graph of f, $gr(f)$, lies below the line segment $P_x P_y$ joining P_x and P_y. This is shown in the sketch below.



Note the epigraph of f, $epi(f)$, is the **inside** of the curve in this sketch.

This leads to a new definition of the **convexity** of f.

Definition

Let I be a nonempty interval of $\Re$. A function $f : I \to \Re$ is **convex** on I if and only if $epi(f)$ is a convex subset of $\Re^2$, where we recall a **convex subset** of $\Re^2$ is a set $\mathbb{C}$ so that given any two points P and Q in $\mathbb{C}$, the line segment PQ is contained in $\mathbb{C}$.

Theorem

Let $P_x = (x, y)$, $P_{x'} = (x', y')$ and $P_u = (u, v)$ where $x < u < x'$ be three points in $\Re^2$. Then the following three properties are equivalent.

1. *P_u is below $P_x P_{x'}$.*
2. *the slope of $P_x P_u \leq$ the slope of $P_x P_{x'}$.*
3. *the slope of $P_x P_{x'} \leq$ the slope of $P_u P_{x'}$.*

Proof

The line segment $P_x P_{x'}$ *in point slope form is* $z = y + \frac{y'-y}{x'-x}(t-x)$ *for* $x \leq t \leq x'$. *So at* $t = u$, *we have* $z_u = y + \frac{y'-y}{x'-x}(u-x)$. *If Property (1) holds, we have* $v \leq y + \frac{y'-y}{x'-x}(u-x)$. *This implies*

$$\text{slope } P_x P_u = \frac{v-y}{u-x} \quad \leq \quad \frac{y'-y}{x'-x} = \text{ slope } P_x P_{x'}$$

which is Property (2). It is easy to see we can reverse this argument to show if Property (2) holds, then so does Property (1). So we have shown Property (1) holds if and only if Property (2) holds.

Next, if Property (1) holds, since P_u *is below the line segment, the slope of the line segment* $P_u P_{x'}$ *is steeper than or equal to the slope of the line segment* $P_x P_{x'}$. *Look at the earlier picture which clearly shows this although the role of* y *should be replaced by* x' *for our argument. Thus, the slope of* $P_x P_{x'} \leq$ *the slope of* $P_u P_{x'}$ *and so Property (3) holds.*

Proof

The same picture shows us we can reverse the argument to show if the slope of $P_x P_{x'} \leq$ the slope of $P_u P_{x'}$ then P_u must be below the line segment $P_x P_{x'}$. So Property (3) implies Property (1).

Hence, Property (2) implies Property (1) which implies Property (3). And it is easy to reverse the argument to show Property (3) implies Property (2). So all of these statements are equivalent.

Now if $x < u < x'$, then $u = tx + (1-t)x'$ for some $0 < t < 1$ and by the convexity of f, we have $f(u) \leq tf(x) + (1-t)f(x')$. Letting $P_x = (x, f(x))$ and $P_{x'} = (x', f(x'))$, convexity implies $P_u = (u, f(u))$ is below the line segment $P_x P_{x'}$ and by the previous Theorem,

$$\text{slope } P_x P_u \quad \leq \quad \text{slope } P_x P_{x'} \leq \text{ slope } P_u P_{x'}$$

or

$$\frac{f(u) - f(x)}{u - x} \quad \leq \quad \frac{f(x') - f(x)}{x' - x} \leq \frac{f(x') - f(u)}{x' - u}$$

Since $u = tx + (1-t)x' = t(x - x') + x'$, we see $t = \frac{u-x'}{x-x'} = \frac{x'-u}{x'-x}$. Thus, convexity can be written

$$f(u) \quad \leq \quad \left(\frac{x' - u}{x' - x}\right) f(x) + \left(\frac{u - x}{x' - x}\right) f(x')$$

For any $y \neq x_0$, et $S(y, x_0)$ denote the slope term $S(y, x_0) = \frac{f(y) - f(x_0)}{y - x_0}$.

Recall, we showed

$$\frac{f(u)-f(x)}{u-x} \leq \frac{f(x')-f(x)}{x'-x} \leq \frac{f(x')-f(u)}{x'-u}$$

for any $x < u < x'$. Switching to $u = x_0$, we have

$$S(x, x_0) = \frac{f(x)-f(x_0)}{x-x_0} \leq \frac{f(x')-f(x_0)}{x'-x_0} = S(x', x_0)$$

And this is true for all $x < x'$.

So f is convex on I implies the slope function $S(x, x_0)$ is increasing on the set $I \setminus \{x_0\}$ which is all of I except the point x_0. This is called the **criterion of increasing slopes**. It is a straightforward argument to see we can show the reverse: if the criterion of increasing slopes hold, then f is convex. We have a Theorem!

Theorem

The Criterion of Increasing Slopes

Let I be a nonempty interval in $\Re$. Then $f : I \rightarrow \Re$ is **convex** $\Leftrightarrow$ *the slope function $S(x, x_0)$ is increasing on $I \setminus \{x_0\}$ for all $x_0 \in I$.*

Proof

We have just gone over this argument. □

Now let's look at what we can do with this information. We will define what might be considerd a lower and upper form of a derivative next.

Theorem

Let f be convex on the interval I and let x_0 be an interior point. Then $\sup_{x<x_0} \frac{f(x)-f(x_0)}{x-x_0}$ *and* $\inf_{x>x_0} \frac{f(x)-f(x_0)}{x-x_0}$ *are both finite and*

$$\sup_{x<x_0} \frac{f(x)-f(x_0)}{x-x_0} \leq \inf_{x>x_0} \frac{f(x)-f(x_0)}{x-x_0}.$$

Proof

By the criterion of increasing slopes, $\frac{f(x)-f(x_0)}{x-x_0}$ *is increasing as* x *approaches* x_0 *from below. If* $\sup_{x<x_0} \frac{f(x)-f(x_0)}{x-x_0} = \infty$, *then there would be a sequence* (x_n) *in the interior of* I, *with* $x_n \neq x_0 < x$ *and* $\frac{f(x_n)-f(x_0)}{x_n-x_0} > n$. *Since* x_0 *is an interior point, choose a* $y > x_0$ *in* I *and apply the criterion of increasing slopes to see* $n < \frac{f(x_n)-f(x_0)}{x_n-x_0} \leq \frac{f(y)-f(x_0)}{y-x_0}$.

Proof

But this tells us $f(y) \geq n(y - x_0) + f(x_0)$ *implying* $f(y)$ *must be infinite in value. But* f *is finite on* I *so this is not possible. Thus, there is a constant* $L > 0$ *so that* $\sup_{x<x_0} \frac{f(x)-f(x_0)}{x-x_0} = L$.

We also know that by the criterion of increasing slopes that $\frac{f(x)-f(x_0)}{x-x_0} \leq \frac{f(y)-f(x_0)}{y-x_0}$ *for all* $x < x_0 < y$ *so we also have* $\sup_{x<x_0} \frac{f(x)-f(x_0)}{x-x_0} \leq \frac{f(y)-f(x_0)}{y-x_0}$ *for all* $y > x_0$ *too.*

But this then tells us $\sup_{x<x_0} \frac{f(x)-f(x_0)}{x-x_0} \leq \inf_{y>x_0} \frac{f(y)-f(x_0)}{y-x_0}$

We can do an argument similar to the one above to also show $\inf_{y>x_0} \frac{f(y)-f(x_0)}{y-x_0}$ *is finite and so there is a positive constant* K *so that* $\inf_{y>x_0} \frac{f(y)-f(x_0)}{y-x_0} = K$. *So we have shown*

$$L = \sup_{x<x_0} \frac{f(x) - f(x_0)}{x - x_0} \leq \inf_{y>x_0} \frac{f(y) - f(x_0)}{y - x_0} = K.$$

□

From this last theorem, these finite limits allows us to replace the idea of a derivative at an interior point x_0 for a convex function f. We set the **lower derivative**, $D_-f(x_0)$ and **upper derivative** of f, $D_+f(x_0)$ at an interior point to be the finite numbers

$$D_-f(x_0) = \sup_{x<x_0} \frac{f(x)-f(x_0)}{x-x_0}, \quad D_+f(x_0) = \inf_{x>x_0} \frac{f(x)-f(x_0)}{x-x_0}$$

Next, if $[c,d] \subset int(I)$, choose $c < x < x' < d$. We can show $D_+f(c) \leq D_-f(d)$ using a very long chain of inequalities. Be patient!

$$\begin{aligned} D_+f(c) &= \inf_{\hat{x}>c} \frac{f(\hat{x})-f(c)}{\hat{x}-c} \leq \frac{f(x)-f(c)}{x-c} = \frac{f(c)-f(x)}{c-x}, \text{ increasing slopes} \\ &\leq \sup_{w<x} \frac{f(w)-f(x)}{w-x} = D_-f(x) \leq D_+f(x) \\ &= \inf_{\hat{x}>x} \frac{f(\hat{x})-f(x)}{\hat{x}-x} \leq \frac{f(x')-f(x)}{x'-x} = \frac{f(x)-f(x')}{x-x'}. \end{aligned}$$

So continuing we have

$$\begin{aligned}
&\leq \sup_{w<x'} \frac{f(w)-f(x')}{w-x'} = D_-f(x') \leq D_+f(x') \\
&= \inf_{w>x'} \frac{f(w)-f(x')}{w-x'} \leq \frac{f(d)-f(x')}{d-x'} = \frac{f(x')-f(d)}{x'-d} \\
&\leq \sup_{w<d} \frac{f(w)-f(d)}{w-d} = D_-f(d).
\end{aligned}$$

We conclude for $c < x < x' < d$, $D_+f(c) \leq \frac{f(x')-f(x)}{x'-x} \leq D_-f(d)$.

If $\frac{f(x')-f(x)}{x'-x} > 0$, we have $|\frac{f(x')-f(x)}{x'-x}| \leq D_-f(d)$. If $\frac{f(x')-f(x)}{x'-x} < 0$, we have $|\frac{f(x')-f(x)}{x'-x}| \leq -D_+f(c)$. Combining, we see

$$\left|\frac{f(x')-f(x)}{x'-x}\right| \leq \max\{-D_+f(c), D_-f(d)\}.$$

If $x = c$ or $x' = d$, we can do a similar analysis to get the same result. So we have for $c \leq x < x' \leq d$

$$\left| \frac{f(x') - f(x)}{x' - x} \right| \leq \max\{-D_+ f(c), D_- f(d)\}.$$

We have proven the following result.

Theorem

If f is convex on the interval I, if $[c, d] \subset int(I)$, then there is a positive constant $L^{[c,d]}$ so that if $c \leq x < x' \leq d$, we have

$$\left| \frac{f(x') - f(x)}{x' - x} \right| \leq L^{[c,d]}.$$

Proof

We have just argued this. The constant is
$L^{[c,d]} = \max\{-D_+ f(c), D_- f(d)\}$. □

Comment

From the above, we have $|f(x') - f(x)| \leq L^{[c,d]}\,|x' - x|$ *which immediately tells us* f *is continuous at each* x *in the interior of* I. *What happens at the boundary of* I *is not known yet. Here is the argument.*

Choose $\epsilon > 0$ *arbitrarily. Let* $\delta = \epsilon / L^{[c,d]}$. *Then* $|y - x| < \delta \Rightarrow |f(y) - f(x)| \leq L^{[c,d]}\,\epsilon / L^{[c,d]} = \epsilon$.

Comment

So this sort of a function can't be convex.

$$f(x) = \begin{cases} x^2, & x \leq 0 \\ x^2 - 2, & x > 0 \end{cases}$$

because f *is not continuous at* 0 *and a convex function is continuous at each point in the interior of its domain. You should look at the epigraph here and make sure you understand why it is not a convex subset of* $\Re^2$.

Homework 18

18.1 Draw $f(x) = x^2 + 4$ on $\Re$. Is f convex?

18.2 Draw

$$f(x) = \begin{cases} x^4, & x < 0 \\ a, & x = 0 \\ x^6 - 1, & x > 0 \end{cases}$$

- Find a so that f is lsc.
- Is it possible to choose a so that f is convex?

18.3 Let $f(x) = |x|$. Find $D_- f(0)$ and $D_+ f(0)$. This is just a matter of looking at slopes on the left and the right. You should get $D_- f(0) = -1$ and $D_+ f(0) = +1$.

18.4 Let $f(x) = |x - 2|$. Find $D_- f(2)$ and $D_+ f(2)$.