Math 141: Lecture 10

The mean value theorem and extrema, Jensen's inequality

Bob Hough

October 10, 2016

Definition of relative maximum, minimum

Definition

Let f be a function defined on an interval I and let $c \in I$. The point c is a *relative maximum* of f if there is $\delta > 0$ such that, for all $x \in I$, $|x - c| < \delta$,

$$f(x) \leq f(c).$$

The notion of *relative minimum* is obtained by replacing $f(x) \leq f(c)$ with $f(x) \geq f(c)$.

Definition of extremum

Definition

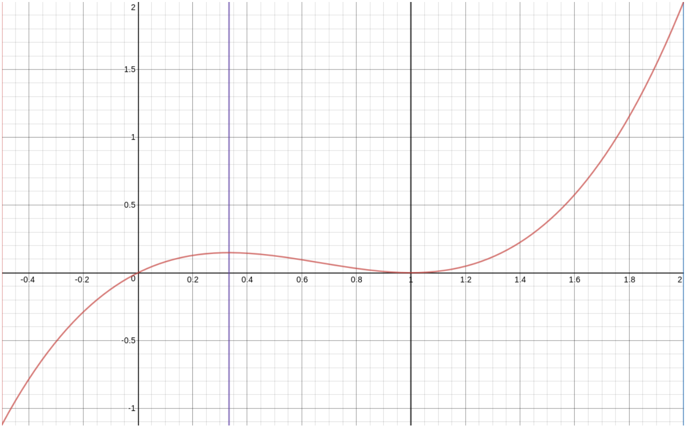
The point c is an *extremum* of f in I if it is either a local maximum or local minimum.

Definition of global maximum

Definition

Let f be a function defined on an interval I and let $c \in I$. The point c is a *global maximum* of f if for all $x \in I$, $f(x) \leq f(c)$. The point c is a *global minimum* of f if for all $x \in I$, $f(x) \geq f(c)$.

Examples of extrema



The function $f(x) = x(1-x)^2$ on $-\frac{1}{2} \le x \le 2$ has an global minimum at $-\frac{1}{2}$, a global maximum at 2, a local maximum at $\frac{1}{3}$, and a local minimum at 1.

Vanishing of the derivative at an interior extremum

Theorem

Let f be a function on an interval I, and let c be a point in the interior of I (not an endpoint). Suppose that c is an extremum of f and that f is differentiable at c. Then $f'(c) = 0$.

Proof.

Suppose that c is a relative maximum (if a relative minimum, make this argument replacing f with $-f$). Let $\delta > 0$ be such that $|x - c| < \delta$ implies $f(x) \leq f(c)$. For $|h| < \delta$,

$$\frac{f(c+h) - f(c)}{h} \begin{array}{ll} \geq 0 & \quad h < 0 \\ \leq 0 & \quad h > 0 \end{array}.$$

Since the limit exists, it is 0. □

Non-examples

The function $f(x) = x^3$ satisfies $f'(0) = 0$ but has no extremum at 0. The function $f(x) = |x|$ has a global minimum at 0, but its derivative doesn't exist there.

Rolle's Theorem

Theorem (Rolle's Theorem)

Let f be continuous on $[a, b]$ and differentiable on (a, b). Suppose $f(a) = f(b)$. Then there exists c, $a < c < b$, such that $f'(c) = 0$.

Proof.

If f is constant on $[a, b]$ then any $a < c < b$ suffices. Otherwise, since f achieves both its maximum and its minimum on $[a, b]$, there is a point $a < c < b$ at which f has an extremum. Hence $f'(c) = 0$. □

Mean Value Theorem for derivatives

Theorem (Mean Value Theorem)

Let f be continuous on $[a, b]$ and differentiable on (a, b). There is a point c, $a < c < b$ such that

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

Proof.

- Define $g(x)$ on $[a, b]$ by $g(x) = \frac{x-a}{b-a}(f(b) - f(a))$.
- Then $h(x) = f(x) - g(x)$ satisfies $h(a) = h(b) = f(a)$.
- The conditions of Rolle's Theorem are met by $h(x)$, and so there is $a < c < b$ with $h'(c) = 0$. Calculate

$$f'(c) = h'(c) + g'(c) = 0 + \frac{f(b) - f(a)}{b - a}.$$

$\square$

Cauchy's mean-value formula

Theorem (Cauchy's Mean Value Theorem)

Let f and g be continuous on $[a, b]$ and differentiable on (a, b). Then there is a c, $a < c < b$, such that

$$f'(c)[g(b) - g(a)] = g'(c)[f(b) - f(a)].$$

Proof.

Define $h(x) = f(x)[g(b) - g(a)] - g(x)[f(b) - f(a)]$. Then

$$h(b) - h(a) = [f(b) - f(a)][g(b) - g(a)] - [g(b) - g(a)][f(b) - f(a)] = 0.$$

By Rolle's Theorem there is c, $a < c < b$ such that $h'(c) = 0$. At this point,

$$f'(c)[g(b) - g(a)] = g'(c)[f(b) - f(a)].$$

$\square$

Properties deducible from the derivative

Theorem

Let f be continuous on $[a, b]$ and differentiable on (a, b).

- *If $f'(x) > 0$ for all $x \in (a, b)$ then f is strictly increasing on $[a, b]$.*
- *If $f'(x) < 0$ for all $x \in (a, b)$ then f is strictly decreasing on $[a, b]$.*

Proof.

Let $x < y$. By the Mean Value Theorem there is a point z, $x < z < y$ where $f'(z)$ and $f(y) - f(x)$ are simultaneously positive/negative. The strictly increasing/strictly decreasing property follows. $\square$

First derivative test for extrema

Theorem

Let f be continuous on $[a, b]$ and differentiable on (a, b). Let $c \in [a, b]$.

- *If, for x in a neighborhood of c, $x < c$ implies $f'(x) > 0$ and $x > c$ implies $f'(x) < 0$ then c is a local maximum.*
- *If, for x in a neighborhood of c, $x < c$ implies $f'(x) < 0$ and $x > c$ implies $f'(x) > 0$ then c is a local minimum.*

Proof.

- Suppose the conditions of the theorem hold for $|x - c| < \delta$. Consider the intervals $I_1 = [c - \delta, c] \cap [a, b]$ and $I_2 = [c, c + \delta] \cap [a, b]$.
- Apply the previous theorem to conclude that f is strictly increasing on intervals where $f' > 0$ in the interior, and f is strictly decreasing on intervals where $f' < 0$ in the interior. This implies local max/min.

$\square$

Second derivative test for extrema

Let f be continuous on $[a, b]$ and differentiable on (a, b). A point $c \in (a, b)$ satisfying $f'(c) = 0$ is a *critical point*.

Theorem

Suppose f'' exists in (a, b) and $c \in (a, b)$ is a critical point.

- *If f'' is negative in (a, b), f has a relative maximum at c.*
- *If f'' is positive in (a, b), f has a relative minimum at c.*

Proof.

This follows from the first derivative test. $\square$

Derivative test for convexity

Theorem

Let f be continuous on $[a, b]$ and differentiable in (a, b). If f' is increasing in (a, b) then f is convex on $[a, b]$. If f' is decreasing on (a, b) then f is concave there. In particular, if f'' exists and $f'' > 0$ (resp. $f'' < 0$) then f is convex (resp. concave).

Proof.

Suppose f' is increasing. Let $x < z < y$. Apply the Mean Value Theorem to find x_1, x_2 such that

$$x < x_1 < z < x_2 < y$$

and

$$\frac{f(z) - f(x)}{z - x} = f'(x_1) \leq f'(x_2) = \frac{f(y) - f(z)}{y - z}.$$

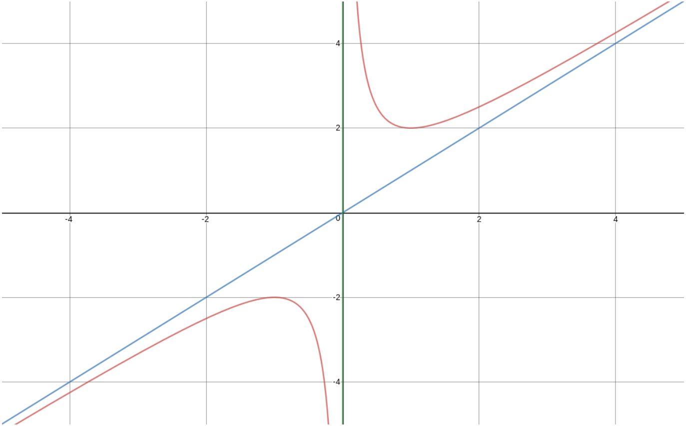
This proves that f is convex. When f' is decreasing, reverse the inequality to conclude f is concave. $\square$

Curve sketching

The *graph* of a function f is the set of points $(x, f(x))$ in $\mathbb{R}^2$.

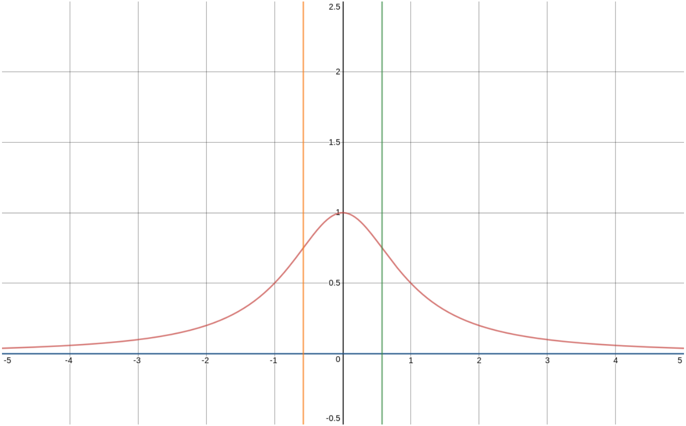
- An *intercept* (of the x axis) is a point $(x, 0)$ where $f(x) = 0$.
- A non-vertical line $y = mx + b$ is an *asymptote* of the graph of f if $f(x) - (mx + b)$ tends to 0 as $x \to \infty$ or $x \to -\infty$.
- A vertical line $x = a$ is a *vertical asymptote* of the graph of f if f takes arbitrarily large positive or negative values as x tends to a.
- A point $(a, f(a))$ such that f'' is defined in a neighborhood of a and changes sign at a is an *inflection point*.

Illustrations



The function $f(x) = x + \frac{1}{x}$ has a vertical asymptote at $x = 0$, and an asymptote $y = x$. There is a local minimum at $x = 1$, a local maximum at $x = -1$. The function is odd, convex for $x > 0$, concave for $x < 0$ and does not cross its asymptotes.

Illustrations



The function $f(x) = \frac{1}{1+x^2}$ has a horizontal asymptote at $y = 0$. The function is even, and concave between its inflection points at $\pm\sqrt{\frac{1}{3}}$, convex elsewhere.

Extrema problems

Problem

Find the rectangle of largest area that can be inscribed in a semicircle, the lower base being on the diameter.

Solution

- *Suppose the semicircle to have radius 1, given by* $\{(x, y) : y \geq 0, x^2 + y^2 \leq 1\}$.
- *Let the upper right corner of the rectangle have coordinates* $(\cos\theta, \sin\theta)$ *with* $0 \leq \theta \leq \frac{\pi}{2}$. *The area is* $A(\theta) = 2\sin\theta\cos\theta = \sin 2\theta$.
- *The maximum of* $\sin$ *on* $[0, \pi]$ *is 1, which occurs at* $2\theta = \frac{\pi}{2}$, *so* $\theta = \frac{\pi}{4}$.

Extrema problems

Problem

Find the trapezoid of largest area that can be inscribed in a semicircle, the lower base being on the diameter.

Solution

- *Parametrize the problem as before. The area of the trapezoid is*

$$f(\theta) = 2\sin\theta\cos\theta + \sin\theta(1-\cos\theta) = \sin\theta(1+\cos\theta).$$

- *One has* $f(0) = 0$ *and* $f\left(\frac{\pi}{2}\right) = 1$ *(endpoint check).*
- *Obtain* $f'(\theta) = \cos\theta + \cos^2\theta - \sin^2\theta = 2\cos^2\theta + \cos\theta - 1.$
- *Set* $x = \cos\theta$, $0 \le x \le 1$. *Solving* $2x^2 + x - 1 = 0$ *gives a single critical point at* $x = \frac{1}{2}$, $\theta = \frac{\pi}{3}$. *The value at this critical point is the global max,*

$$f\left(\frac{\pi}{3}\right) = \frac{\sqrt{3}}{2}\left(1 + \frac{1}{2}\right) = \frac{3\sqrt{3}}{4} > 1.$$

Jensen's inequality

Theorem (Jensen's inequality)

Let f be convex on $[a, b]$. Let $x_1, ..., x_n \in [a, b]$. Let $w_1, ..., w_n$ be positive with $w_1 + ... + w_n = 1$. Then

$$f(w_1 x_1 + ... + w_n x_n) \le w_1 f(x_1) + ... + w_n f(x_n).$$

In particular,

$$f\left(\frac{1}{n}(x_1 + ... + x_n)\right) \le \frac{1}{n}(f(x_1) + ... + f(x_n)).$$

Jensen's inequality

Proof.

The proof is by induction on n.

- Base case: If $n = 1$ there is nothing to prove.
- Suppose for some $n \geq 1$ that the claim has been established for all sets of weights $w_1, ..., w_n$.
- Given $n + 1$ points $x_1, ..., x_{n+1}$ and $n + 1$ weights $w_1, ..., w_{n+1}$, write

$$\begin{aligned} &w_1 x_1 + ... + w_{n+1} x_{n+1} \\ &= (1 - w_{n+1}) \left(\frac{1}{1 - w_{n+1}} (w_1 x_1 + ... + w_n x_n) \right) + w_{n+1} x_{n+1}. \end{aligned}$$

□

Jensen's inequality

Proof.

- Notice that $w_1' = \frac{w_1}{1-w_{n+1}}$, ..., $w_n' = \frac{w_n}{1-w_{n+1}}$ are a set of n non-negative weights that add to 1, so that the inductive assumption applies, showing

$$f(w_1'x_1 + ... + w_n'x_n) \leq w_1'f(x_1) + ... + w_n'f(x_n).$$

- By convexity of f, then the inductive assumption,

$$\begin{aligned}
&f(w_1x_1 + ... + w_{n+1}x_{n+1}) \\
&\leq (1 - w_{n+1})f\left(\frac{1}{1-w_{n+1}}(w_1x_1 + ... + w_nx_n)\right) + w_{n+1}f(x_{n+1}) \\
&\leq w_1f(x_1) + ... + w_nf(x_n) + w_{n+1}f(x_{n+1}).
\end{aligned}$$

$\square$

Positive and negative parts

Given f a function on $[a, b]$, define $f_+ = \max(f, 0)$ and $f_- = \min(f, 0)$ the positive and negative parts of f.

Theorem

Let f be integrable on $[a, b]$. Then f_+ and f_- are integrable.

Proof.

We check this for f_+. Given $\epsilon > 0$, choose upper and lower step functions $u(x), \ell(x)$ such that

$$\int_a^b u(x)dx - \epsilon < \int_a^b f(x)dx < \int_a^b \ell(x)dx + \epsilon.$$

Then u_+ and ℓ_+ are upper and lower step functions for f_+, and $|u_+(x) - \ell_+(x)| \leq |u(x) - \ell(x)|$ for all x. Thus $\int_a^b u_+(x) - \ell_+(x)dx \leq 2\epsilon$, which suffices to prove that the integral of f_+ exists. $\square$

Products of integrable functions

Theorem

Let f and g be integrable (and bounded) on $[a, b]$. Then fg is integrable.

Proof.

Assume $0 \leq f, g \leq M$ by splitting into positive and negative parts. Given $\epsilon > 0$, let $0 \leq \ell_1 \leq u_1, \leq M$, $0 \leq \ell_2 \leq u_2 \leq M$ be upper and lower step functions for f, g, with all integrals making error $< \frac{\epsilon}{2M}$. Then

$$\begin{aligned}
&\int_a^b u_1 u_2(x)dx - \int_a^b \ell_1\ell_2(x)dx \\
&= \int_a^b (u_1 u_2(x) - u_1\ell_2(x)) + (u_1\ell_2(x) - \ell_1\ell_2(x))dx \\
&\leq M \int_a^b u_2(x) - \ell_2(x)dx + M \int_a^b u_1(x) - \ell_1(x)dx < \epsilon.
\end{aligned}$$

Since $\ell_1\ell_2, u_1 u_2$ are step functions, $\ell_1\ell_2 \leq fg \leq u_1 u_2$, fg is integrable. $\square$

Subintervals

Theorem

Let f be integrable on $[a, b]$. Then f is integrable on all sub-intervals of $[a, b]$.

Proof.

Let $[c, d] \subset [a, b]$ and let $1_{[c,d]}(x)$ equal 1 if $x \in [c, d]$, 0 otherwise. Since this is a step function, it is integrable. The theorem now follows from the theorem regarding products. $\square$

Integral Jensen's inequality

Theorem (Integral Jensen's inequality)

Let $w \geq 0$ on $[a, b]$ be integrable with $\int_a^b w(x)dx = 1$, and let f be convex and continuous on $[a, b]$. Then

$$f\left(\int_a^b xw(x)dx\right) \leq \int_a^b f(x)w(x)dx.$$

Proof.

For convenience, assume $a = 0$, $b = 1$.

- Define step functions s_n and f_n by partitioning $[0, 1]$ into n equal sub-intervals and assign s_n, f_n the values of $x, f(x)$ at each right endpoint.
- Let $w_{1,n}, ..., w_{n,n}$ given by taking $w_{j,n}$ the integral of $w(x)$ on the jth subinterval. Thus $w_{1,n} + ... + w_{n,n} = 1$.

□

Integral Jensen's inequality

Proof.

- Recall for $1 \leq j \leq n$, $w_{j,n} = \int_{\frac{j-1}{n}}^{\frac{j}{n}} w(x)dx$ and for $\frac{j-1}{n} < x \leq \frac{j}{n}$, $s_n(x) = \frac{j}{n}$ and $f_n(x) = f(\frac{j}{n})$.
- $\int_0^1 s_n(x)w(x)dx = w_{1,n}\frac{1}{n} + ... + w_{n,n}\frac{n}{n}$, $\int_0^1 f_n(x)w(x)dx = w_{1,n}f(1/n) + ... + w_{n,n}f(n/n)$.
- By Jensen's inequality with points $1/n, ..., n/n$ and weights $w_{1,n}, ..., w_{n,n}$,

$$f\left(\int_0^1 s_n(x)w(x)dx\right) \leq \int_0^1 f_n(x)w(x)dx.$$

$\square$

Integral Jensen's inequality

Proof.

- Note that $|x - s_n(x)| \leq \frac{1}{n}$. Thus

$$\begin{aligned}\left|\int_0^1 xw(x)dx - \int_0^1 s_n(x)w(x)dx\right| &= \left|\int_0^1 (x - s_n(x))w(x)dx\right| \\ &\leq \int_0^1 |x - s_n(x)|w(x)dx \\ &\leq \frac{1}{n}\int_0^1 w(x)dx = \frac{1}{n}.\end{aligned}$$

- Thus, as $n \to \infty$,

$$\int_0^1 s_n(x)w(x)dx \to \int_0^1 xw(x)dx.$$

$\square$

Integral Jensen's inequality

Proof.

- By a theorem regarding the continuous image of a limit (Lecture 8), since f is continuous, as $n \to \infty$,

$$f\left(\int_0^1 s_n(x)w(x)dx\right) \to f\left(\int_0^1 xw(x)dx\right).$$

- Since f is uniformly continuous on $[0,1]$, for each $\epsilon > 0$ there is N such that $n > N$ implies $|f_n(x) - f(x)| < \epsilon$. Using this as before, as $n \to \infty$,

$$\int_0^1 f_n(x)w(x)dx \to \int_0^1 f(x)w(x)dx.$$

□

Integral Jensen's inequality

Proof.

Since we've checked for each n that

$$f\left(\int_0^1 s_n(x)w(x)dx\right) \leq \int_0^1 f_n(x)w(x)dx$$

it follows that

$$f\left(\int_0^1 xw(x)dx\right) \leq \int_0^1 f(x)w(x)dx$$

by taking limits (see HW 7). □

The inequality between the harmonic and arithmetic means

- The *harmonic mean* of n positive numbers $x_1, x_2, ..., x_n$ is $H_n = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + ... + \frac{1}{x_n}}$.
- Harmonic means arise, for instance, when calculating average speed. For instance, if a driver drives for a mile at 30mph and a mile at 50 mph, the driver's average speed over two miles is the harmonic mean $H = \frac{2}{\frac{1}{30} + \frac{1}{50}} = 37.5$.
- Note that this is less than the arithmetic average of the two speeds, a fact which is true in general.

The inequality between the harmonic and arithmetic means

Theorem

Let $x_1, x_2, ..., x_n > 0$. *One has*

$$H_n = \frac{n}{\frac{1}{x_1} + ... + \frac{1}{x_n}} \le A_n = \frac{x_1 + ... + x_n}{n}.$$

Proof.

Let $f(x) = \frac{1}{x}$ on $(0, \infty)$. Then $f''(x) = \frac{2}{x^3}$ is positive on $x > 0$ so f is convex. By Jensen's inequality,

$$f(A_n) \le \frac{f(x_1) + ... + f(x_n)}{n} \qquad \Leftrightarrow \qquad A_n \ge H_n.$$

□

The arithmetic mean and the root mean square

- The quadratic mean or *Root Mean Square* of n real numbers $x_1, \ldots, x_n$ is
$$Q_n = \sqrt{\frac{x_1^2 + \ldots + x_n^2}{n}}.$$
- The RMS often arises in discussing mean error in measurements.
- In statistics, the *standard deviation* describes the 'range' around the average in which a measurement may be expected to occur. For instance, for data distributed according to the bell curve, 68% of data points lie within 1 standard deviation of the average, 95% lie within 2 s.d., 99.7 % within 3 s.d. and 99.9999998% within 6 s.d. (an industry standard).
- The standard deviation of the sum of n independent measurements is the RMS of the standard deviations of the individual measurements.

The arithmetic mean and the root mean square

Theorem

Let $x_1, \ldots, x_n$ be real numbers. Then

$$A_n = \frac{x_1 + \ldots + x_n}{n} \leq Q_n = \sqrt{\frac{x_1^2 + \ldots + x_n^2}{n}}.$$

Proof.

Consider $f(x) = x^2$ on $\mathbb{R}$, which is convex. By Jensen,

$$f(A_n) \leq \frac{f(x_1) + \ldots + f(x_n)}{n} \qquad \Leftrightarrow \qquad A_n \leq Q_n.$$

$\square$