New Zealand Mathematical Olympiad Committee

Convex Functions

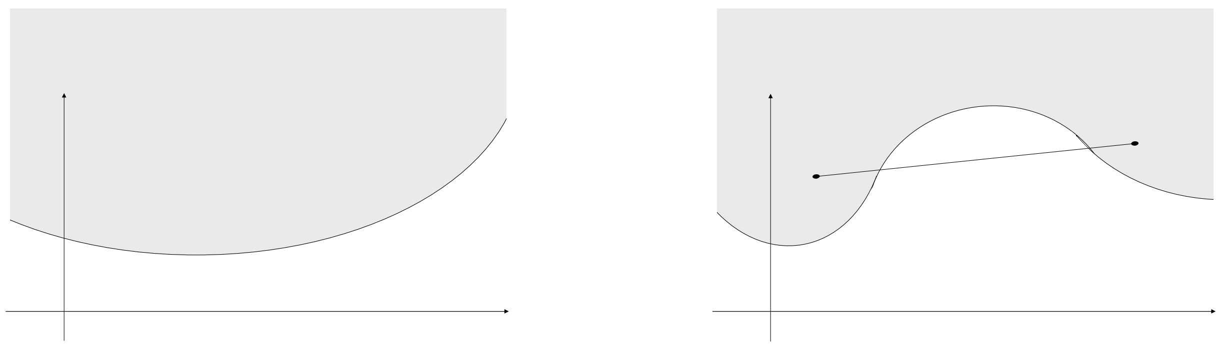
Heather Macbeth

1 Introduction

This lecture is about three tricks, two theorems, and one idea. The idea is that of *convex functions*. We formulate this idea more precisely for a function f on an interval I (this can be closed or open, bounded or unbounded) as follows:

Definition. A function $f : I \to \mathbb{R}$ is *convex*, if for all $x, y \in I$ and all $\mu \in [0, 1]$,

$$\mu f(x) + [1 - \mu]f(y) \geq f(\mu x + [1 - \mu]y).$$



(a) convex (b) not convex

Figure 1: Functions

For example, let's show that the function $f(x) = 1/x$ is convex. Indeed, for any $x, y \in (0, \infty)$ and $\mu \in [0, 1]$, $(x - y)^2 \geq 0$ and hence

$$(\mu y + [1 - \mu]x)(\mu x + [1 - \mu]y) = xy + \mu[1 - \mu](x - y)^2 \geq xy,$$

which means

$$\mu\frac{1}{x} + [1 - \mu]\frac{1}{y} = \frac{\mu y + [1 - \mu]x}{xy} \geq \frac{1}{\mu x + [1 - \mu]y}.$$

Exercise 1. Let $f : I \to \mathbb{R}$ be a convex function. Show that for all $x, y \in I$ and all $\mu \in \mathbb{R} \setminus [0, 1]$ such that $\mu x + [1 - \mu]y \in I$,

$$\mu f(x) + [1 - \mu]f(y) \leq f(\mu x + [1 - \mu]y).$$

Exercise 2 (Jensen's inequality)**.** Let $f : I \to \mathbb{R}$ be a convex function. Show that for any integer n, any real numbers $x_1, x_2, \ldots x_n \in I$, and any positive real numbers $\mu_1, \mu_2, \ldots \mu_n$ such that $\mu_1 + \cdots + \mu_n = 1$,

$$\mu_1 f(x_1) + \mu_2 f(x_2) + \cdots + \mu_n f(x_n) \geq f(\mu_1 x_1 + \mu_2 x_2 + \cdots + \mu_n x_n).$$

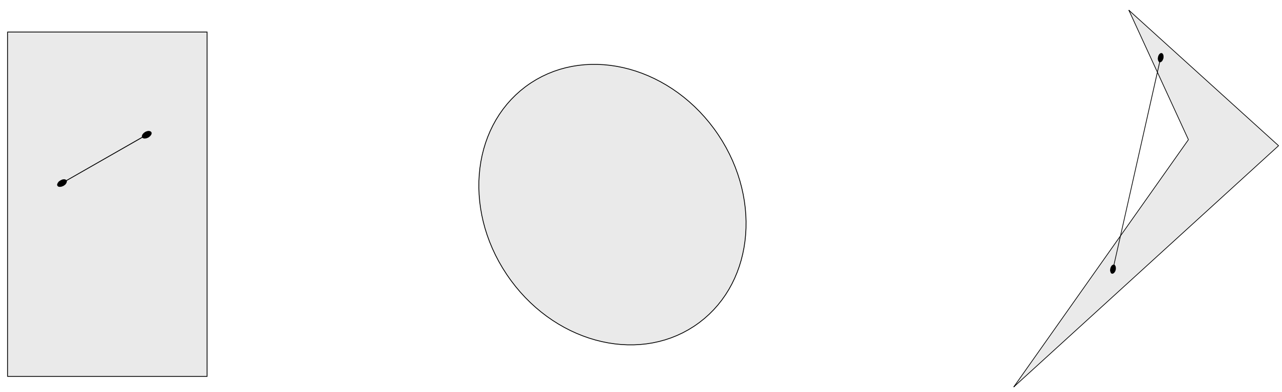
Among the nice consequences of Jensen's inequality are the power mean inequalities. For instance, the AM-HM inequality is obtained by setting $f(x) = 1/x$ and $\mu_1 = \cdots = \mu_n = \frac{1}{n}$: for all $x_1, \ldots x_n \in (0, \infty)$,

$$\frac{x_1 + x_2 + \cdots + x_n}{n} \geq \frac{n}{\frac{1}{x_1} + \cdots + \frac{1}{x_n}}.$$

You may have encountered the following related geometrical concept:

Definition. A set $A \subseteq \mathbb{R}^2$ of points in the plane is *convex*, if for all $\mathbf{x}, \mathbf{y} \in A$ and all $\mu \in [0, 1]$,

$$\mu\mathbf{x} + [1 - \mu]\mathbf{y} \in A.$$



(a) convex (b) convex (c) not convex

Figure 2: Shapes

How are these concepts related? A function $f : \mathbb{R} \to \mathbb{R}$ is convex precisely if the region above its graph is convex.

2 Not everyone is below average

Theorem 1. *A convex function on a closed bounded interval attains its maximum at one of its endpoints.*

Proof. Suppose $f : [a, b] \to \mathbb{R}$ is strongly convex. Then for any $t \in [a, b]$, taking x, y, μ to be a, b, $(b - t)/(b - a)$ respectively, we find

$$f(t) = f\left(\left[\frac{b-t}{b-a}\right] a + \left[1 - \frac{b-t}{b-a}\right] b\right) \leq \left[\frac{b-t}{b-a}\right] f(a) + \left[1 - \frac{b-t}{b-a}\right] f(b).$$

That is, $f(t)$ is at most some weighted average of $f(a)$ and $f(b)$. This means that at least one of $f(a)$ and $f(b)$ is greater than or equal to $f(t)$:

$$\begin{aligned} f(t) &\leq \left[\frac{b-t}{b-a}\right] f(a) + \left[1 - \frac{b-t}{b-a}\right] f(b) \\ &\leq \left(\left[\frac{b-t}{b-a}\right] + \left[1 - \frac{b-t}{b-a}\right]\right) \max\left[f(a), f(b)\right] \\ &= \max\left[f(a), f(b)\right]. \end{aligned}$$

It follows that

$$\max_{t \in [a,b]} f(t) = \max[f(a), f(b)].$$

$\square$

The idea here is:

> If the average of a set of numbers is m, then at least one of the numbers is at least m.

The Pigeonhole Principle is another version of this same trick.

Exercise 3. Show that there is some pair of Londoners, who have the same number of hairs on their heads.

Exercise 4 (APMO 2002)**.** Let $a_1, a_2, \dots a_n$ be natural numbers, and set

$$A = \frac{a_1 + \cdots + a_n}{n}.$$

Show that

$$a_1!a_2! \cdots a_n! \geq (\lfloor A \rfloor!)^n.$$

3 Iterate and pad

The rest of these notes are a diversion into the land of midpoint-convex functions.

Definition. The function $f : I \to \mathbb{R}$ is *midpoint-convex*, if for all $x, y \in I$,

$$\frac{f(x) + f(y)}{2} \geq f\left(\frac{x+y}{2}\right).$$

Clearly convexity implies midpoint-convexity. However, there exist midpoint-convex functions that are not convex. Such functions can be very strange and interesting. We will explore this distinction further in the next section.

In this section, we will prove a version of Jensen's inequality for midpoint-convex functions. We need a much cleverer argument here than we did for standard Jensen's in the previous section. Here's the idea, a variant of standard induction:

> Suppose that we are given a sequence $S_1, S_2, \dots$ of statements and an increasing sequence $a_1, a_2, \dots$ of natural numbers, and that
>
> 1. S_1 is true;
> 2. if S_{a_k} is true then $S_{a_{k+1}}$ is true; and,
> 3. if S_n is true, then for all $m < n$, S_m is also true.
>
> Then S_n is true for all n.

Theorem 2. *Let $f : I \to \mathbb{R}$ be a midpoint-convex function. Then for any integer n and any real numbers $x_1, x_2, \dots x_n \in I$,*

$$\frac{f(x_1) + f(x_2) + \cdots + f(x_n)}{n} \geq f\left(\frac{x_1 + x_2 + \cdots + x_n}{n}\right).$$

Proof. We are given a midpoint-convex function $f : I \to \mathbb{R}$. Let S_n be the statement,

For any real numbers $x_1, x_2, \dots x_n \in I$,

$$\frac{f(x_1) + f(x_2) + \cdots + f(x_n)}{n} \geq f\left(\frac{x_1 + x_2 + \cdots + x_n}{n}\right).$$

S_1 says that $f(x_1) \geq f(x_1)$ for all $x_1 \in I$, and this is certainly true.

For the second step, suppose that S_{2^k} is true. We'll deduce $S_{2^{k+1}}$ by applying the definition of midpoint-convexity to two copies of S_{2^k}. Indeed, take any real numbers $x_1, x_2, \dots x_{2^{k+1}} \in I$; then

$$\begin{aligned} & \frac{f(x_1) + f(x_2) + \cdots + f(x_{2^{k+1}})}{2^{k+1}} \\ \geq \; & \frac{f\left(\frac{x_1+x_2+\cdots+x_{2^k}}{2^k}\right) + f\left(\frac{x_{2^k+1}+x_{2^k+2}+\cdots+x_{2^{k+1}}}{2^k}\right)}{2} \\ \geq \; & f\left(\frac{x_1 + x_2 + \cdots + x_{2^{k+1}}}{2^{k+1}}\right). \end{aligned}$$

Finally, suppose that S_n is true. We'll show that S_m is true for all $m \leq n$, by applying S_n to some m variables $x_1, \dots x_m \in I$ that we care about, plus $n - m$ copies of the filler variable

$$X = \frac{x_1 + \cdots + x_m}{m}.$$

Indeed, S_n gives

$$\frac{f(x_1) + f(x_2) + \cdots + f(x_m) + (n-m)f(X)}{n} \geq f\left(\frac{x_1 + x_2 + \cdots + x_m + (n-m)X}{n}\right).$$

Simplifying the left-hand side,

$$\begin{aligned} & \frac{f(x_1) + f(x_2) + \cdots + f(x_m) + (n-m)f(X)}{n} \\ = \; & \frac{f(x_1) + f(x_2) + \cdots + f(x_m)}{n} + \left(\frac{n-m}{n}\right) f(X). \end{aligned}$$

In the right-hand side bracket,

$$\begin{aligned} & \frac{x_1 + x_2 + \cdots + x_m + (n-m)X}{n} \\ = \; & \frac{x_1 + x_2 + \cdots + x_m}{n} + \left(\frac{n-m}{n}\right) X \\ = \; & \left(\frac{m}{n} + \frac{n-m}{n}\right) X = X. \end{aligned}$$

Substituting back, multiplying by n/m and subtracting a term, we get

$$\frac{f(x_1) + f(x_2) + \cdots + f(x_m)}{m} \geq f(X) = f\left(\frac{x_1 + \cdots + x_m}{m}\right).$$

$\square$

Exercise 5. A sequence $a_0, a_1, a_2 \dots$ of whole numbers is given, such that for any whole number k, there is exactly one pair i, j of whole numbers for which $a_i + 2a_j = k$. Find all possible values for a_{2009}.

Exercise 6. For each natural number n, define $rad(n)$ to be the product of the primes which divide n. Pick some natural number a, and define a sequence $a_0 = a, a_1, a_2, \dots$ by the recurrence

$$a_{n+1} = a_n + rad(a_n).$$

Show that the sequence contains arbitrarily long arithmetic progressions.

4 Approximation by rationals

In this section, we'll explore some situations in which midpoint-convexity does imply convexity.

So assume f is midpoint-convex. First, for any $0 \le k \le n$, applying Theorem 2 to the n real numbers $x_1 = x_2 = \cdots = x_k = x$, $x_{k+1} = \cdots = x_n = y$ gives

$$\frac{kf(x) + [n-k]f(y)}{n} \ge f(kx + [n-k]f(y));$$

that is,

$$\frac{k}{n}f(x) + \left[1 - \frac{k}{n}\right] f(y) \ge f\left(\frac{k}{n}x + \left[1 - \frac{k}{n}\right] y\right).$$

So midpoint-convexity implies that for all $x, y \in I$ and all *rational* $\mu \in [0, 1]$,

$$\mu f(x) + [1-\mu]f(y) \ge f(\mu x + [1-\mu]y).$$

Going further requires a new tool – approximation of reals by rationals – which is based on the following fact:

$\boxed{\text{The rationals are } \textit{dense in the line.}}$

More precisely, for any real number $x \in \mathbb{R}$ and any real $\epsilon > 0$, there is a rational number $\xi \in \mathbb{Q}$ for which $|x - \xi| < \epsilon$. Equivalently, for any real number $x \in \mathbb{R}$, there is a sequence $(\xi_j)_{j \in \mathbb{N}} \subseteq \mathbb{Q}$ of rational numbers with limit x.

For a quick justification, note that we can approximate, say, π, by its sequence 3, 3.1, 3.14, 3.141, . . . of truncated decimal representations, in which the j-th term differs by at most 10^{j-1} from π. Of course, there are many other sequences of rationals with limit π.

Now we'll show that either of two quite natural conditions on a midpoint-convex function imply convexity. The first is boundedness.

Lemma 3. *If $f : I \to \mathbb{R}$ is midpoint-convex and bounded above, then it is convex.*

Proof. Ross Atkins showed me this. Let f be bounded above by M. Take any $x, y \in I$. Without loss of generality, $x < y$. For shorthand, for any real number μ, write t_μ for the expression $\mu x + [1-\mu]y$.

Suppose for the sake of contradiction that for some μ,

$$f(t_\mu) > \mu f(x) + [1-\mu]f(y).$$

Say that

$$f(t_\mu) - \mu f(x) - [1-\mu]f(y) = h,$$

where $h > 0$. Choose $n \in \mathbb{N}$ big enough that $M - \min[f(x), f(y)] < nh$. Let $a \in \mathbb{R}$ be any real number for which

1. $\mu - a$ is rational; and,
2. a is small enough (in absolute value) that $\mu - a$ and $\mu + na$ are in $[0, 1]$.

Since the rationals are dense in the line, such an a exists.

Then

$$\left[\frac{1}{n+1}\right] t_{\mu+na} + \left[\frac{n}{n+1}\right] t_{\mu-a} = t_\mu,$$

and so by Jensen's inequality

$$\left[\frac{1}{n+1}\right] f(t_{\mu+na}) + \left[\frac{n}{n+1}\right] f(t_{\mu-a}) \ge f(t_\mu);$$

that is,

$$\begin{aligned} f(t_{\mu+na}) &\geq (n+1)f(t_\mu) - nf(t_{\mu-a}) \\ &= (n+1)\left(\mu f(x) + [1-\mu]f(y) + h\right) - nf(t_{\mu-a}). \end{aligned}$$

Also since $\mu - a$ is rational,

$$f(t_{\mu-a}) \leq (\mu - a)f(x) + [1-(\mu-a)]f(y).$$

Combining the last two equations gives

$$\begin{aligned} f(t_{\mu+na}) &\geq (\mu+na)f(x) + [1-(\mu+na)]f(y) + (n+1)h \\ &\geq \min[f(x), f(y)] + (n+1)h \\ &> M, \end{aligned}$$

a contradiction. Hence the convexity condition must hold for all μ. □

The other condition in question is continuity; that is:

Definition. A function $f : I \to \mathbb{R}$ is *continuous*, if for all convergent sequences $(x_j)_{j\in\mathbb{N}} \subseteq I$,

$$\lim_{j\to\infty} f(x_j) = f(\lim_{j\to\infty} x_j).$$

Lemma 4. *If $f : I \to \mathbb{R}$ is midpoint-convex and continuous, then it is convex.*

Proof. Let $f : I \to \mathbb{R}$ be midpoint-convex and continuous. Since the rationals are dense in the line, we can choose a sequence $(\mu_j)_{j\in\mathbb{N}} \subseteq [0,1]$ of rational numbers with limit μ. Let us do so. Then

$$\mu x + [1-\mu]y = \lim_{j\to\infty} [\mu_j x + [1-\mu_j]y]$$

and

$$\mu f(x) + [1-\mu]f(y) = \lim_{j\to\infty} [\mu_j f(x) + [1-\mu_j]f(y)].$$

So if f is continuous, then we can conclude that

$$\begin{aligned} \mu f(x) + [1-\mu]f(y) &= \lim_{j\to\infty} [\mu_j f(x) + [1-\mu_j]f(y)] \\ &\geq \lim_{j\to\infty} f(\mu_j x + [1-\mu_j]y) \\ &= f(\lim_{j\to\infty} [\mu_j x + [1-\mu_j]y]) \\ &= f(\mu x + [1-\mu]y). \end{aligned}$$

That is, f is convex. □

Exercise 7. Find all continuous functions $f : \mathbb{R} \to \mathbb{R}$, such that for all $x, y \in \mathbb{R}$,

$$f(x) + f(y) = f(x+y).$$

Exercise 8. Find all functions $f : \mathbb{R} \to \mathbb{R}$, which are bounded above on some interval, and which for all $x, y \in \mathbb{R}$ satisfy

$$f(x) + f(y) = f(x+y).$$

Exercise 9. Construct an increasing function $f : \mathbb{Q}^+ \to \mathbb{Q}^+$, so that for all $x \in \mathbb{Q}^+$,

$$f(f(x)) = 3x.$$

5 Problems

1. (IMO 2004) Let $n \geq 3$ be an integer. Let $x_1, x_2, \ldots x_n$ be positive real numbers such that

$$n^2 + 1 > (x_1 + x_2 + \cdots + x_n)\left(\frac{1}{x_1} + \frac{1}{x_2} + \cdots + \frac{1}{x_n}\right).$$

   Show that x_i, x_j, x_k are side lengths of a triangle for all i, j, k with $1 \leq i < j < k \leq n$.

2. (All-Union Olympiad 1978) Real numbers $x_1, x_2, \ldots x_n$ lie on the segment $[a, b]$, where $0 < a < b$. Prove that

$$(x_1 + x_2 + \cdots + x_n)\left(\frac{1}{x_1} + \frac{1}{x_2} + \cdots + \frac{1}{x_n}\right) \leq \frac{(a+b)^2}{4ab} n^2.$$

April 9, 2010

http://www.mathsolympiad.org.nz