Math Club Notes: 2005 November 3

1 Geometric mean vs sum

Problem A2 from the 2003 Putnam:

> Let $a_1, a_2, \dots, a_n$ and $b_1, b_2, \dots, b_n$ be nonnegative real numbers. Show that
>
> $$(a_1 a_2 \cdots a_n)^{1/n} + (b_1 b_2 \cdots b_n)^{1/n} \\ \le [(a_1 + b_1)(a_2 + b_2) \cdots (a_n + b_n)]^{1/n} .$$

In words: the sum of geometric means is at most the geometric mean of the sums. (We looked at geometric means once before: see the notes for May 9.)

The possibility that some of the numbers are zero is annoying — it prevents us from doing all kinds of useful things, such as dividing by them, and taking their logs. So let's deal with zeroes as a separate case:

First suppose that one of the a_k, or one of the b_k, is zero. Without loss of generality, $b_1 = 0$. Then

$$\begin{aligned}
&(a_1 a_2 \cdots a_n)^{1/n} + (b_1 b_2 \cdots b_n)^{1/n} \\
&\quad = (a_1 a_2 \cdots a_n)^{1/n} + 0 \\
&\quad = (a_1 a_2 \cdots a_n)^{1/n} \\
&\quad = [(a_1 + 0)(a_2 + 0) \cdots (a_n + 0)]^{1/n} \\
&\quad \le [(a_1 + b_1)(a_2 + b_2) \cdots (a_n + b_n)]^{1/n}
\end{aligned}$$

since each $b_k \ge 0$, and every function involved here is increasing.

So now we can assume that all of the a_k and all of the b_k are positive. Thus, for example, $(a_1 \cdots a_n)^{1/n}$ is positive, so we can divide by it and obtain an equivalent inequality:

$$1 + \left(\frac{b_1}{a_1} \cdot \frac{b_2}{a_2} \cdots \frac{b_n}{a_n} \right)^{1/n} \le \left(\frac{a_1 + b_1}{a_1} \cdot \frac{a_2 + b_2}{a_2} \cdots \frac{a_n + b_n}{a_n} \right)^{1/n} .$$

Simplifying a bit,

$$1 + \left(\frac{b_1}{a_1} \cdot \frac{b_2}{a_2} \cdots \frac{b_n}{a_n} \right)^{1/n} \le \left[\left(1 + \frac{b_1}{a_1} \right) \left(1 + \frac{b_2}{a_2} \right) \cdots \left(1 + \frac{b_n}{a_n} \right) \right]^{1/n} .$$

That looks a little nicer — now we're not really dealing with the numbers a_k and b_k (of which there are $2n$), but with the numbers b_k/a_k (of which there are n). Let's write it to express that: let $c_k = b_k/a_k$ (which is positive); we wish to show that

$$1 + (c_1 c_2 \cdots c_n)^{1/n} \le [(1 + c_1)(1 + c_2) \cdots (1 + c_n)]^{1/n} .$$

Products are a little annoying; sums are nicer. To convert products into sums, we use logarithms. To turn the right-hand side into a sum is pretty easy: just take the log of both sides of the inequality and apply the laws of logarithms. (Everything is positive, so this makes sense; it yields an equivalent inequality because ln is an increasing function.)

$$\ln(1 + (c_1 c_2 \cdots c_n)^{1/n}) \le \frac{1}{n}(\ln(1 + c_1) + \ln(1 + c_2) + \cdots + \ln(1 + c_n)) .$$

Let's also start using sigma notation now, to save ourselves some writing:

$$\ln(1 + (c_1 c_2 \cdots c_n)^{1/n}) \le \frac{1}{n}\sum_{k=1}^{n} \ln(1 + c_k) .$$

What about the left-hand side? It's now the log of a sum, and there's not much we can do with that; the product on the inside cannot be converted to a sum using that log.

But it's easy to introduce a new log: just balance it out with an exponentiation. We can replace any (positive) value x with $e^{\ln x}$. (For typographical reasons, we will use the "exp" function: $\exp(x) = e^x$.)

$$\ln(1 + \exp\ln[(c_1 c_2 \cdots c_n)^{1/n}]) \le \frac{1}{n}\sum_{k=1}^{n} \ln(1 + c_k) .$$

Now apply those good ol' laws of logarithms to get

$$\ln\left(1 + \exp\left(\frac{1}{n}\sum_{k=1}^{n} \ln c_k\right)\right) \le \frac{1}{n}\sum_{k=1}^{n} \ln(1 + c_k) .$$

Some new similarities between the left- and right-hand sides are beginning to emerge: both have this $\ln(1 + \cdots)$ structure in them. On the left, though, we have $\ln(1+\exp(\dots))$, and on the right we have just $\ln(1+c_k)$. Can we write c_k as $\exp(\dots)$? Sure; it's the same trick as before:

$$\ln\left(1 + \exp\left(\frac{1}{n}\sum_{k=1}^{n} \ln c_k\right)\right) \le \frac{1}{n}\sum_{k=1}^{n} \ln(1 + \exp\ln c_k) .$$

Now we're no longer talking about the c_k anywhere, but only about their logs. Adopt a suitable notation: let $x_k = \ln c_k$.

$$\ln\left(1 + \exp\left(\frac{1}{n}\sum_{k=1}^{n} x_k\right)\right) \le \frac{1}{n}\sum_{k=1}^{n} \ln(1 + \exp x_k) .$$

To make this a tad easier on the eyes, let's define an "average" notation:

$$\operatorname*{Avg}_{k=1}^{n} f(k) = \frac{1}{n}\sum_{k=1}^{n} f(k) .$$

Then our inequality is

$$\ln\left(1+\exp \operatorname*{Avg}_{k=1}^{n} x_k\right) \le \operatorname*{Avg}_{k=1}^{n} \ln(1+\exp x_k)\ .$$

Let's also define a function f by $f(x) = \ln(1+e^x)$; then our inequality is

$$f\left(\operatorname*{Avg}_{k=1}^{n} x_k\right) \le \operatorname*{Avg}_{k=1}^{n} f(x_k)\ .$$

In this form, the inequality has a nice geometric interpretation. We have some curve $y = f(x)$, and some points (x_k, y_k) along this curve. The centre of gravity $(\bar{x}, \bar{y})$ of these points is given by

$$\bar{x} = \operatorname*{Avg}_{k=1}^{n} x_k \qquad \bar{y} = \operatorname*{Avg}_{k=1}^{n} y_k = \operatorname*{Avg}_{k=1}^{n} f(x_k)$$

So our inequality is

$$f(\bar{x}) \le \bar{y}\ .$$

That is, we wish to show that the centre of gravity of points along the curve lies on or above the curve.

It turns out that this is (close to) one definition of convexity: a region is convex if, for any set of points in the region, their centre of gravity lies in the region. So we will have proven our inequality if we can show that the region on and above the curve $y = \ln(1+e^x)$ is convex.

It is intuitively clear (though it does need proof) that it's enough to show that the curve is concave up. We know how to do that: we compute that

$$f''(x) = \frac{e^x}{(1+e^x)^2} > 0 \quad \text{for all } x.$$

And we're done.

(Xi Chen, the university's resident Putnam guy, showed me this solution after one of our Putnam preparation sessions.)