Famous Inequalities

Herman Jaramillo

May 10, 2016

2 Introduction

Just a list of inequalities which are useful.

3 Convex Functions

Let $I \subset \mathbb{R}$ be an interval. A function $f : I \to \mathbb{R}$ is called convex if

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y). \tag{3.1}$$

for all $x, y \in I$ and all $\lambda \in [0,1]$.

It is interesting to observe that if the function f is linear then

$$f(\lambda x + (1-\lambda)y) = \lambda f(x) + (1-\lambda)f(y).$$

which is always convex. But a concave (or convex down) function satisfies the inequality $f(\lambda x + (1-\lambda)y) \geq \lambda f(x) + (1-\lambda)f(y)$. So a linear function is simultaneously convex and concave. A function f is convex if and only its negative $-f$ is concave.

Figure 1 shows an illustration of a convex function.

The interesting functions in this area are functions which satisfy strictly inequality. That is, they have curvature. and we prove first that for a convex

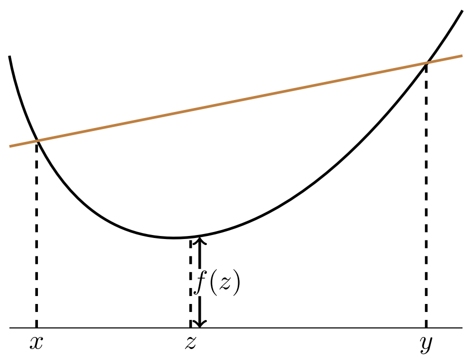

Figure 1: A convex function

function, the second derivative is always greater or equal to 0 (equal if it is linear), or in symbols $f''(z) \geq 0$, $\forall z \in I$. The proof of this is easy. Pick

$$x = z + h \quad y = z - h \quad \lambda = 1/2$$

in equation 3.1, and find

$$f((z+h)/2 + (z-h)/2) \leq f(z+h)/2 + f(z-h)/2$$

that is

$$f(z+h) - 2f(z) + f(z-h) \geq 0$$

and dividing by $2h$

$$\frac{f(z+h) - 2f(z) + f(z-h)}{2h} \geq 0$$

we take the limit as $h \to 0$ and find

$$f''(z) \geq 0.$$

Since the second derivative of $-\log x$ is $1/x^2$ then $-\log x$ is convex in the interval $(0, \infty)$. We now prove the reverse. That is if $f''(x) \geq 0$ for all $x \in I$ where I is an interval over the real numbers, then f is convex in I. Let us first prove that if $f'(x)$ is monotonically increasing. From the mean value theorem, pick $x < y$ in I we know that there is a $c \in [x, y]$ such that

$$f''(c) = \frac{f'(y) - f'(x)}{y - x}$$

and since $f''(c) \geq 0$, and $y > x$ then

$$f'(y) \geq f'(x).$$

So f' is monotinically increasing. Now, let us choose $s = \lambda x + (1 - \lambda)y$. Then $x < s < y$. By the mean value theorem exists c_1 , $x < c_1 < s$ and, c_2, $s < c_2 < y$ such that

$$f'(c_1) = \frac{f(s) - f(x)}{s - x} \qquad f'(c_2) = \frac{f(y) - f(s)}{y - s},$$

since f' is monotonically increasing and $c_1 \leq c_2$ then

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

$$\frac{f[\lambda x + (1-\lambda)y] - f(x)}{\lambda x + (1-\lambda)y - x} \leq \frac{f(y) - f(\lambda x + (1-\lambda)y)}{y - (\lambda x + (1 - \lambda y)}$$

$$\frac{f[\lambda x + (1-\lambda)y] - f(x)}{(1-\lambda)\cancel{(y-x)}} \leq \frac{f(y) - f(\lambda x + (1-\lambda)y)}{\lambda\cancel{(y-x)}}$$

$$\cancel{\lambda f[\lambda x + (1-\lambda)y]} - \lambda f(x) \leq (1-\lambda)f(y) - f[\lambda x + (1-\lambda)y] + \cancel{\lambda f[\lambda x + (1-\lambda)y]}$$

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

So f is convex.

Next, we show a variation of Jensen's inequality [1], for a discrete set.

[1] http://en.wikipedia.org/wiki/Jensen%27s_inequality

3.1 Jensen's Inequality

Suppose $I \in \mathbb{R}$ is an interval and $f : I \to \mathbb{R}$ is convex. Then for all $n \in \mathbb{N}$, all $\lambda_i \geq 0$, $i = 1, \cdots, n$ and $\sum_{j=1}^{n} \lambda_j = 1$, and all $z_i \in I$,

$$f\left(\sum_{j=1}^{n} \lambda_j z_j\right) \leq \sum_{j=1}^{n} \lambda_j f(z_j)$$

That is, the function is "under-linear" for more than just two points. It is like the definition of convexity extended to many points.

We use induction. For $n = 1$ it is trival that $f(z_1) \leq f(z_1)$. Let us assume that up to some $n \geq 2$ the inequality is valid. Then we want to evaluate

$$f\left(\sum_{j=1}^{n+1} \lambda_j z_j\right).$$

The idea is to split the sum in two sets of points each with less than $n+1$ elements. By assumption

$$\sum_{i=1}^{n+1} \lambda_i = 1$$

If any of the λ_i is equal to 1, all others have to be equal to 0 because $\lambda_i \geq 0$, $i = 1, \cdots, n+1$. So, necessarily $\lambda_i < 1$, for all i's. Let us separate λ_1 from the rest of λ_i's and write

$$\sum_{j=1}^{n+1} \lambda_j z_j = \lambda_1 z_1 + \sum_{j=2}^{n+1} \lambda_j z_j$$

Since z_1 has already a coefficient λ_1, we want to write the second sum as some number u with coefficient $1 - \lambda_1$. That is, we call

$$(1 - \lambda_1)u = \sum_{j=2}^{n+1} \lambda_j z_j,$$

or define

$$u = \frac{\sum_{j=2}^{n+1} \lambda_j z_j}{1 - \lambda_1} = \sum_{j=2}^{n+1} \frac{\lambda_j z_j}{1 - \lambda_1}.$$

We obtained what we wanted. That is,

$$\sum_{j=1}^{n+1} \lambda_j z_j = \lambda_1 z_1 + (1 - \lambda)u$$

and we can apply the convexity of f to these two terms to obtain

$$f\left(\sum_{j=1}^{n+1} \lambda_j z_j\right) = f(\lambda_1 z_1 + (1 - \lambda_1)u) \leq \lambda_1 f(z_1) + (1 - \lambda_1) f(u).$$

Now we should expand $f(u)$. We note that the coefficients of u add to 1, because

$$\sum_{i=1}^{n+1} \lambda_i = 1 \qquad \Rightarrow \qquad \lambda_1 = 1 - \sum_{i=2}^{n+1} \lambda_i \qquad \Rightarrow \qquad \sum_{i=2}^{n+1} \lambda_i = 1 - \lambda_1$$

and so

$$\sum_{i=2}^{n+1} \frac{\lambda_i}{1-\lambda_1} = \frac{1-\lambda_1}{1-\lambda_1} = 1,$$

and then by the induction hypothesis

$$f(u) \leq \sum_{i=2}^{n+1} \frac{\lambda_i}{1-\lambda_1} f(z_j),$$

and finally

$$f\left(\sum_{j=1}^{n+1} \lambda_j z_j\right) = f(\lambda_1 z_1) + (1-\lambda_1)u) \leq \lambda_1 f(z_1) + \frac{\cancel{1-\lambda_1}}{\cancel{1-\lambda_1}} \sum_{j=2}^{n+1} \lambda_i f(z_i),$$

that is

$$f\left(\sum_{j=1}^{n+1} \lambda_j z_j\right) \leq \sum_{j=1}^{n+1} \lambda_i f(z_i),$$

which proves the theorem.

The Jensen's inequality is useful to prove other inequalities. This is the case of the following inequality.

3.2 Weighted geometric/arithmetic mean inequality

Suppose $\sum_{j=1}^{n} \lambda_j a_j$ is a convex [2] combination of no-negative numbers $a_1, \cdots, a_n$. Then

$$a_1^{\lambda_1} a_2^{\lambda_2} \cdots a_n^{\lambda_n} \leq \sum_{j=1}^{n} \lambda_j a_j.$$

What this is telling us is that the weighted geometric mean is smaller than the weighted arithmetic mean. We assume that $0^0 = 0$ by definition, so if one of the $a_i = 0$, then the inequality is obvious and so we assume that all $a_i > 0$. We use Jensen's inequality applied to the convex function $f(x) = -\log x$, on the inteval $I = (0, \infty)$. That is

$$-\log\left(\sum_{j=1}^{n} \lambda_j a_j\right) \leq -\sum_{j=1}^{n} \lambda_j \log a_j = -\log(a_1^{\lambda_1} \cdots a_n^{\lambda_n})$$

and so, multiplying by -1 and taking the exponential function we observe the inequality. In particular when each $\lambda_i = 1/n$, we find

$$(a_1 a_2 \cdots a_n)^{\frac{1}{n}} \leq \frac{1}{n} \sum_{j=1}^{n} a_j$$

which is the classical geometric/arithmetic mean inequality.

[2]by convex combination we mean a weighted average where the weights λ_j add to 1.

3.3 The Hölder's inequality

For $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^n$ and $1 \leq p \leq \infty$

$$\sum_{j=1}^{n} |x_j y_j| \leq \|\boldsymbol{x}\|_p \|\boldsymbol{y}\|_q, \quad \text{where} \quad \frac{1}{p} + \frac{1}{q} = 1.$$

For any $1 \leq p \leq \infty$ the norm $\|.\|_p$ is defined as

$$\|\boldsymbol{x}\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p}$$

and

$$\|\boldsymbol{x}\|_\infty = \max_{i=1}^{n} |x_i|.$$

When $p = 2$, by default nothing is written as sub-index. That is $\|.\|_2 = \|.\|$.

Let us first prove this inequalities for the extrema values $p = 1$ and $p = \infty$. If $p = 1$, then $q = \infty$ we have

$$\sum_{j=1}^{n} |x_j y_j| \quad \leq \quad \max_{i=1}^{n} |y_i| \sum_{j=1}^{n} |x_j| = \|\boldsymbol{x}\|_1 \|\boldsymbol{y}\|_\infty.$$

The case $p = \infty$ and $q = 1$ is the equivalent to this (just change $\boldsymbol{x}$ by $\boldsymbol{y}$ and being those vectors arbitrary we do not have to repeat that proof here.

Let us then assume that $1 < p < \infty$. If one of or both vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ are zero the inequality turns into the equality $0 \leq 0$. So we assume that both $\boldsymbol{x}$ and $\boldsymbol{y}$ are non-zero vectors.

For any $a_j, b_j \geq 0$ the weighted arithmetic/geometric mean inequality indicates that

$$a_j^{\frac{1}{p}} b_j^{\frac{1}{q}} \leq \frac{a_j}{p} + \frac{b_j}{q}.$$

We set

$$a_j = \frac{|x_j|^p}{\|\boldsymbol{x}\|_p^p} \qquad b_j = \frac{|y_j|^q}{\|\boldsymbol{y}\|_q^q}$$

and so

$$\left(\frac{|x_j|^p}{\|\boldsymbol{x}\|_p^p} \right)^{\frac{1}{p}} \left(\frac{|y_j|^q}{\|\boldsymbol{y}\|_q^q} \right)^{\frac{1}{q}} \leq \frac{1}{p} \frac{|x_j|^p}{\|\boldsymbol{x}\|_p^p} + \frac{1}{q} \frac{|y_j|^q}{\|\boldsymbol{y}\|_q^q}$$

We add over $j = 1, \cdots n$, and find

$$\sum_{i=1}^{n} \left(\frac{|x_j|^p}{\|\boldsymbol{x}\|_p^p} \right)^{\frac{1}{p}} \left(\frac{|y_j|^q}{\|\boldsymbol{y}\|_q^q} \right)^{\frac{1}{q}} \leq \sum_{i=1}^{n} \frac{1}{p} \frac{|x_j|^p}{\|\boldsymbol{x}\|_p^p} + \frac{1}{q} \frac{|y_j^q|}{\|\boldsymbol{y}\|_q^q}.$$

We write the left sum as

$$\sum_{i=1}^{n} \left(\frac{|x_j|^p}{\|\boldsymbol{x}\|_p^p} \right)^{\frac{1}{p}} \left(\frac{|y_j|^q}{\|\boldsymbol{y}\|_q^q} \right)^{\frac{1}{q}} = \frac{1}{\|\boldsymbol{x}\|_p \|\boldsymbol{y}\|_q} \sum_{i=1}^{n} |x_j||y_j|$$

and the right side as

$$\sum_{i=1}^{n} \frac{1}{p} \frac{|x_j|^p}{\|\boldsymbol{x}\|_p^p} + \frac{1}{q} \frac{|y_j^q|}{\|\boldsymbol{y}\|_q^q} = \frac{1}{p} \frac{\cancel{\sum_{j=1}^n |x_j|^p}}{\cancel{\|\boldsymbol{x}\|_p^p}} + \frac{1}{q} \frac{\cancel{\sum_{j=1}^n |y_j|^p}}{\cancel{\|\boldsymbol{y}\|_q^q}} \frac{1}{p} + \frac{1}{q} = 1$$

and so, we showed that

$$\frac{1}{\|\boldsymbol{x}\|_p \|\boldsymbol{y}\|_q} \sum_{i=1}^{n} |x_j||y_j| \leq 1,$$

and from here

$$\sum_{i=1}^{n} |x_j||y_j| \leq \|\boldsymbol{x}\|_p \|\boldsymbol{y}\|_q$$

which is the Hölder's inequality. If $p = q = 2$ we find the famous

Cauchy-Schwartz inequality

$$\sum_{i=1}^{n} |x_j||y_j| \leq \|\boldsymbol{x}\|\|\boldsymbol{y}\|$$

3.4 Minkowski's inequality

Let $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^n$ and $1 \leq p \leq \infty$. Then

$$\|\boldsymbol{x} + \boldsymbol{y}\|_p \leq \|\boldsymbol{x}\|_p + \|\boldsymbol{y}\|_p.$$

Let us start with $p = 1$ here

$$\|\boldsymbol{x} + \boldsymbol{y}\|_1 = \sum_{i=1}^{n} |x_i + y_i| \leq \sum_{i=1}^{n} |x_i| + \sum_{i=1}^{n} y_i = \|\boldsymbol{x}\|_1 + \|\boldsymbol{y}\|_1.$$

and if $p = \infty$

$$\|\boldsymbol{x} + \boldsymbol{y}\|_\infty = \max_{i=1}^{n} |x_i + y_i| \leq \max_{i=1}^{n} |x_i| + \max_{i=1}^{n} |y_i| = \|\boldsymbol{x}\|_\infty + \|\boldsymbol{y}\|_\infty.$$

Let us now assume that $1 < p < \infty$. We write

$$\begin{aligned} \|\boldsymbol{x} + \boldsymbol{y}\|_p^p &= \sum_{j=1}^{n} |x_j + y_j|^p \\ &= \sum_{j=1}^{n} |x_j + y_j||x_j + y_j|^{p-1} \\ &\leq \sum_{j=1}^{n} |x_j||x_j + y_j|^{p-1} + \sum_{j=1}^{n} |y_j||x_j + y_j|^{p-1}. \end{aligned}$$

We apply the Hölder's inequality to each of the terms above. That is

$$\sum_{j=1}^{n} |x_j||x_j + y_j|^{p-1} \leq \|\boldsymbol{x}\|_p \|\boldsymbol{x} + \boldsymbol{y}\|_q^{p-1} \qquad \sum_{j=1}^{n} |y_j||x_j + y_j|^{p-1} \leq \|\boldsymbol{y}\|_p \|\boldsymbol{x} + \boldsymbol{y}\|_q^{p-1}.$$

Then

$$\|\boldsymbol{x} + \boldsymbol{y}\|_p^p \leq (\|\boldsymbol{x}\|_p + \|\boldsymbol{y}_p\|) \|\boldsymbol{x} + \boldsymbol{y}\|_q^{p-1}$$