CORESPONDENȚE

Jensen's inequality for non-convex functions

Holger STEPHAN[1]

Jensen's inequality is well known: Let $f(x)$ be a convex function on $x \in I \subset \mathbb{R}$, $x_1, ..., x_n \in I$ and $q_i \geq 0$ are weights with $\sum_i q_i = 1$. Then, Jensen's inequality

$$\sum_{i=1}^{n} q_i f(x_i) \geq f\Big(\sum_{i=1}^{n} q_i x_i\Big)$$

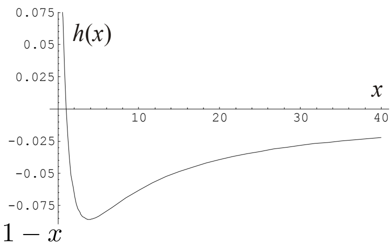
holds.

This inequality is in some sense equivalent to the definition of convexity. That's why, it is not well known that Jensen's inequality holds even for non-convex functions (moreover, this is believed to be false).

Let's start with an easy contest problem (from the journal $\sqrt[Die]{wurzel}$, Jena, 05/2005, problem $\mu 19$):

Let $x_1, \ldots, x_n$ be positive real numbers with $x_1 + \cdots + x_n = n$. Prove the following inequality.

$$\sum_{i=1}^{n} \frac{1}{1+x_i} \geq \sum_{i=1}^{n} \frac{2}{3+x_i} \qquad (1)$$



We use Jensen's inequality with the function

$$h(x) = \frac{1}{1+x} - \frac{2}{3+x} = \frac{1-x}{(1+x)(3+x)} \ .$$

and the weights $q_i = \frac{1}{n}$. We get

$$\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{1+x_i} - \frac{2}{3+x_i}\right) = \sum_{i=1}^{n}\frac{1}{n}h(x_i) \geq h\Big(\frac{1}{n}\sum_{i=1}^{n} x_i\Big) = h(1) = 0$$

and therefore (1). The result is true but the proof is false, because $h(x)$ is not convex for $x > 1 + \sqrt[3]{2} + \sqrt[3]{4} = 6.69464\ldots$ (see the picture). Why does the inequality holds, nevertheless? To investigate this question, we try to prove Jensen's inequality deriving an identity for it. We fix $x_0 = \sum\limits_{i=1}^{n} q_i x_i$ and define a function

$$g(x) = \frac{f(x) - f(x_0)}{x - x_0} \qquad (2)$$

(for $x = x_0$ we define $g(x_0) = f'(x_0)$). Now, we set

$$X = \sum_{i=1}^{n} q_i f(x_i) - f\Big(\sum_{i=1}^{n} q_i x_i\Big) = \sum_{i=1}^{n} q_i f(x_i) - f(x_0)$$

and check, when does $X \geq 0$ hold. Using easy calculations, we obtain

[1] Cercetător dr., Institutul Weierstrass, Berlin (e-mail: stephan@wias-berlin.de)

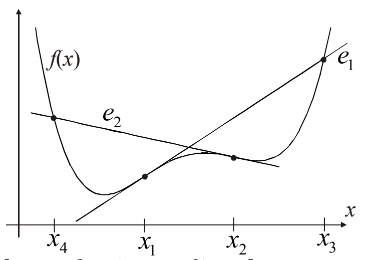
$$X=\sum_{i=1}^{n} q_i f(x_i)-f(x_0)=\sum_{i=1}^{n} q_i\big(f(x_i)-f(x_0)\big)=\sum_{i=1}^{n} q_i\frac{f(x_i)-f(x_0)}{x_i-x_0}(x_i-x_0)=$$
$$=\sum_{i=1}^{n} q_i g(x_i)\,(x_i-x_0)=\sum_{i=1}^{n} g(x_i)q_i\Big(x_i\sum_{j=1}^{n} q_j-\sum_{j=1}^{n} q_j x_j\Big)=$$
$$=\sum_{i,j=1}^{n} q_i q_j x_i g(x_i)-\sum_{i,j=1}^{n} q_i q_j x_j g(x_i)=\sum_{i,j=1}^{n} q_i q_j\big(x_i g(x_i)-x_j g(x_i)\big)=$$
$$=\sum_{1\le j<i\le n} q_i q_j (x_i-x_j)\big(g(x_i)-g(x_j)\big).$$

Thus, we get the identity

$$X=\sum_{i=1}^{n} q_i f(x_i)-f\Big(\sum_{i=1}^{n} q_i x_i\Big)=\sum_{1\le j<i\le n} q_i q_j (x_i-x_j)\big(g(x_i)-g(x_j)\big) \qquad (3)$$

We see that $X\ge 0$ holds, if the function $g(x)$ is monotone increasing. The function $g(x)$ is called slope function. $g(x)$ is the slope of the secant through the points in x_0 and x. If $f(x)$ in convex, $g(x)$ is increasing for any point x_0. But this is not the only case. Function $g(x)$ is increasing, if looking from the point $(x_0, f(x_0))$ on the graph of $f(x)$, no point of $f(x)$ "lies in the shadow" of the graph. This can happen for some points x_0 even if the function is not convex. The function $h(x)$ above is an example (here is $x_0 = 1$).

An other amazing example are polynomials of fourth order – typical non-convex functions. We consider such a function $f(x)$ and its inflection points (the x-coordinates are x_1 and x_2). The inflection point tangents e_1 and e_2 intersect the graph of $f(x)$ in points with x-coordinates x_3 and x_4. Now, Jensen's inequality holds for $x_0 \ge x_3$ or $x_0 \le x_4$.



The typical proof of Jensen's inequality starts from the inequality for convex functions

$$(x_i-x_0)f'(x_0)\le f(x_i)-f(x_0)\le (x_i-x_0)f'(x_i)$$

Multiplying by q_i and adding up over i we get

$$\sum_{i=1}^{n} q_i(x_i-x_0)f'(x_0)\le\sum_{i=1}^{n} q_i f(x_i)-f(x_0)\sum_{i=1}^{n} q_i\le\sum_{i=1}^{n} q_i(x_i-x_0)f'(x_i)$$

The left hand side in zero because of $\sum_{i=1}^{n} q_i = 1$ and $\sum_{i=1}^{n} q_i x_i = x_0$. The middle is X and transforming the right hand sind in a similar way as above (3) we obtain.

$$0\le\sum_{i=1}^{n} q_i f(x_i)-f\Big(\sum_{i=1}^{n} q_i x_i\Big)\le\sum_{1\le j<i\le n} q_i q_j (x_i-x_j)\big(f'(x_i)-f'(x_j)\big)$$

This is similar to identity (3), but holds only for convex functions.

Moreover, identity (3) is very useful, if we are not only interested in Jensen's inequality, but if we want to estimate the difference on the left hand side of (3).