General Young's inequality for real numbers

27 April 2006

Abstract

This text contains the statement and proof of Young's inequality for real numbers and a generalization of it. Its generalization has a nice graphical interpretation which is also shown.

1 Statement

The most familiar form of Young's inequality, which is frequently used to prove the well-known Hölder inequality for L^p functions, is the following:

Theorem 1.1 (Young's inequality). *For $a, b \geq 0$ and $p, q \geq 1$ such that $\frac{1}{p} + \frac{1}{q} = 1$ one has*

$$ab \leq \frac{1}{p}a^p + \frac{1}{q}b^q.$$

The next theorem is a generalization:

Theorem 1.2 (General Young's inequality). *Let $c > 0$ and $f : [0, c] \to \mathbb{R}$ be a strictly increasing continuous function such that $f(0) = 0$. Let $a \in [0, c]$ and $b \in [0, f(c)]$. Then,*

$$ab \leq \int_0^a f(x)\,dx + \int_0^b f^{-1}(x)\,dx \tag{1}$$

Note that we obtain the previous inequality taking $f(x) := x^{p-1}$.

In these conditions, if we call

$$\Lambda(x) := \int_0^x f(y)\,dy$$
$$\Lambda^*(x) := \int_0^x f^{-1}(y)\,dy$$

then another way to state the same result is to say that:

$$ab \le \Lambda(a) + \Lambda^*(b).$$

As Λ^* can be obtained from Λ, we can rewrite theorem 1.2 as follows:

Theorem 1.3 (General Young's inequality, second form). *Take $c > 0$ and let $\Lambda : [0, c] \to \mathbb{R}$ be $\mathcal{C}^1$ and strictly convex with $\Lambda(0) = \Lambda'(0) = 0$. Then for any $a \in [0, c]$ and $b \in [0, \Lambda'(c)]$ it holds that*

$$ab \le \Lambda(a) + \Lambda^*(b) \tag{2}$$

where

$$\Lambda^*(x) := \int_0^x (\Lambda')^{-1}(y)\,dy \quad \text{for } x \in \Lambda([0, c]). \tag{3}$$

The following is a useful identity relating Λ and Λ^*:

Lemma 1.4. *Let $c > 0$ and $\Lambda : [0, c] \to \mathbb{R}$ be a $\mathcal{C}^1$ and strictly convex with $\Lambda(0) = \Lambda'(0) = 0$. Define Λ^* by (3). Then,*

$$x\Lambda'(x) = \Lambda(x) + \Lambda^*(\Lambda'(x)). \tag{4}$$

Remark 1.5. One can remove the requirement that Λ be strictly convex and still have the same results, but we will not do this here, as it involves some technical complications when defining the inverse function.

2 Proofs and explanations

For the proof of theorem 1.1 we just need the convexity of the exponential:

Proof of theorem 1.1. The function $x \mapsto e^x$ is convex. This means that for $x, y \in \mathbb{R}$ and $0 \le \theta \le 1$,

$$\exp(\theta x + (1-\theta)y) \le \theta \exp x + (1-\theta)\exp y.$$

When a or b are zero the inequality is trivial. Otherwise, the theorem is just this inequality with

$$x := \log a, \quad y := \log b, \quad \theta := \frac{1}{p}.$$

□

Theorem 1.2 can be understood with a picture (see figure 1). Recall that the inverse of a function can be drawn by reflecting its graph along the $x = y$ line, so *the area* below the inverse of a function is the area *to the left of the function* (the area between its graph and the vertical axis).

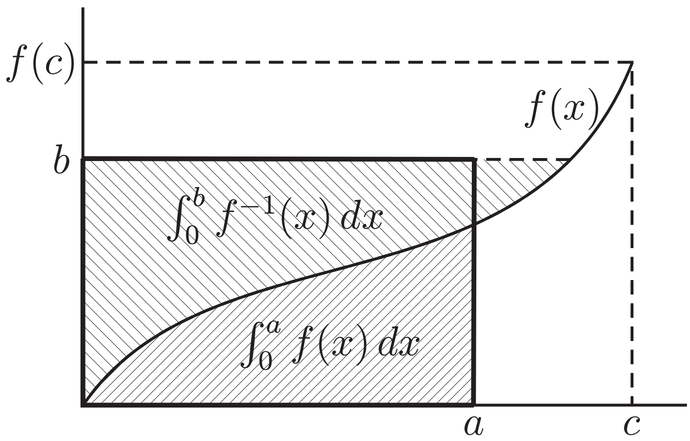

Figure 1: if $f(a) \leq b$

The following lemma, which we will use in the proof, can also be seen in the picture:

Lemma 2.1. *Let $c > 0$. For a strictly increasing continuous function $f : [0, c] \to \mathbb{R}$ and $b \in [0, f(c)]$ we have that*

$$bf^{-1}(b) = \int_0^b f^{-1}(x)\,dx + \int_0^{f^{-1}(b)} f(x)\,dx \tag{5}$$

(Note that this lemma is the identity in equation (4) when Λ is defined as $\Lambda(x) := \int_0^x f^{-1}(y)\,dy$.)

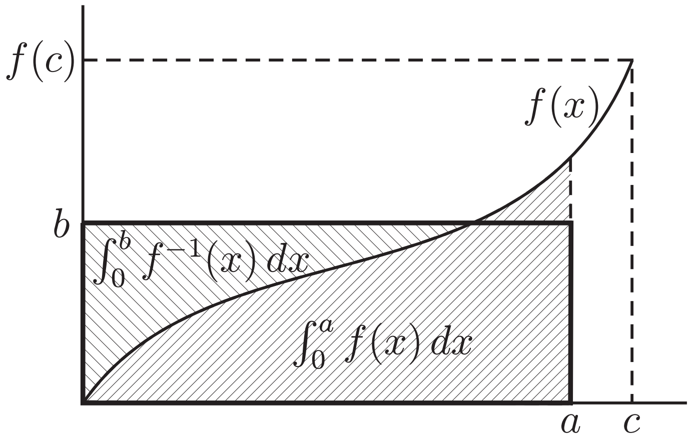

Figure 2: if $f(a) \geq b$

Proof. It is enough to prove it when f is $\mathcal{C}^1$, as then a continuous f can be uniformly approximated by strictly increasing $\mathcal{C}^1$ functions while f^{-1} is also uniformly approximated. It is also enough to prove it when $f(a) \geq b$, as otherwise we can just interchange both f,f^{-1} and a, b.

To prove it in this case we can change variables in the integral $\int_0^b f^{-1}(x)\,dx$ by putting $x = f(y)$, and then integrate by parts:

$$\int_0^b f^{-1}(x)\,dx = \int_0^{f^{-1}(b)} y f'(y)\,dy = bf^{-1}(b) - \int_0^{f^{-1}(b)} f(x)\,dx.$$

□

Proof of theorem 1.2. Use (5) and the fact that f is increasing:

$$\begin{aligned} ab &= bf^{-1}(b) + b(a - f^{-1}(b)) \\ &= \int_0^b f^{-1}(x)\,dx + \int_0^{f^{-1}(b)} f(x)\,dx + b(a - f^{-1}(b)) \\ &\leq \int_0^b f^{-1}(x)\,dx + \int_0^{f^{-1}(b)} f(x)\,dx + \int_{f^{-1}(b)}^a f(x)\,dx \\ &= \int_0^b f^{-1}(x)\,dx + \int_0^a f(x)\,dx. \end{aligned}$$

□

The rest is an easy consequence of these results. Theorem 1.3 is a rewrite of 1.2 and as it was said before, the identity in equation (4) is just lemma 2.1 when Λ is defined as $\Lambda(x) := \int_0^x f^{-1}(y)\, dy$.

3 About this text

This document has been written by José Alfredo Cañizo. The figures are by an anonymous vector artist who prefers to remain in the shadows. For comments or suggestions write to `ozarfreo@yahoo.com`. The latest version should be at `http://www.ugr.es/~ozarfreo/tex` .

3.1 History

11 August 2005: First version.
27 April 2006: Correction of some mistakes.