Mathematical
Inequalities
& Applications
Volume 1, Number 1 (1998), 131–136

A GENERALIZATION OF YOUNG'S ℓ^p INEQUALITY

JONATHAN M. BORWEIN

(*communicated by J. Pečarić*)

Abstract. We show that, for positive real numbers with $a \geqslant 1 + \sum_i^N \alpha_i$, the function

$$\frac{\frac{r^a}{a}}{\prod_{i=1}^N x_i^{\alpha_i}}$$

has a convex conjugate of the same form and so, in particular, obtain a clean proof that f is convex.

1. Introduction

The simplest version of *Young's inequality* ([6], [3], [5]) asserts that for complementary real numbers $1 < p, q < \infty$ with $(p-1)(q-1) = 1$

$$\frac{1}{p}|x|^p + \frac{1}{q}|y|^q \geqslant x\,y$$

for all real numbers x and y with equality exactly if

$$y = x^{p-1}\mathrm{sign}(x) \iff x = y^{q-1}\mathrm{sign}(y).$$

This result plays a fundamental role in the theory of $\mathscr{L}^p$ spaces ([6]) leading, for example, to a transparent proof of Hölder's inequality and to a simple proof of Minkowski's inequality. One attractive way of proving this is by introducing the *convex conjugate* (also called the Fenchel conjugate)

$$f^*(s) := \sup_r r\,s - f(r)$$

of an extended real–valued function $f : \mathbf{R} \to [-\infty, \infty]$ and its multidimensional analogue

$$f^*(y) := \sup_x \langle x, y\rangle - f(x)$$

of an extended real-valued function $f : \mathbf{R^N} \to [-\infty, \infty]$. The conjugate is necessarily convex as the supremum of affine functions ([5]). Moreover, a proper (i.e., somewhere

Mathematics subject classification (1991): Primary 52A41, Secondary 90C25, 26A51.
Key words and phrases: Fenchel conjugate, Young's ℓ^p -inequality, convexity.
Research supported by NSERC and the Shrum endowment of Simon Fraser University.

Paper MIA-01-11

finite) function f is lower semicontinuous and convex if and only if $f = f^{**}$. Our usage throughout is consistent with that of Rockafellar ([5]).

This also suggests that a powerful way of establishing convexity of a function is by realizing it as the conjugate of some other function. Indeed, with the advent of symbolic computation this also leads to a class of functions that can be proved convex in Maple. In this language, Young's inequalilty is recaptured from the observation that for conjugate real numbers $1 < p, q < \infty$ with $(p-1)(q-1) = 1$

$$f\colon x \to \frac{1}{p}|x|^p$$

has conjugate

$$g\colon y \to \frac{1}{q}|y|^q$$

with the equality being a statement about the *subgradients* (see below) of a function and its conjugate ([5]). Correspondingly, in N dimensions we are asserting that the conjugate of $\frac{1}{p}\|\cdot\|_p^p$ is the dual $\frac{1}{q}\|\cdot\|_q^q$.

Observe that for any function f we have from the definition of the conjugate that

$$f(x) + f^*(y) \geqslant \langle x, y \rangle \tag{YI}$$

for all vectors x and y. This is the general form of Young's inequality. It produces the best inequality of the given form, in the sense that if $f(x) + g(y) \geqslant \langle x, y \rangle$ holds universally, then necessarily $g(y) \geqslant f^*(y)$. Moreover, for a proper lower semicontinuous convex function, the *subgradient*, ∂f, defined by

$$\partial f(x) := \{y : \langle y, h \rangle \leqslant f(x+h) - f(x), \forall h \in \mathbf{R^N}\}$$

is easily seen to coincide with

$$\{y : \langle y, x \rangle = f(x) + f^*(y)\}$$

and so, in addition,

$$y \in \partial f(x) \iff x \in \partial f^*(y)$$

if and only if equality is obtained in (YI). Finally we note that when f is differentiable at x then $\{\nabla f(x)\} = \partial f(x)$ so that the conjugate allows one to invert derivatives.

The main business of this note is to compute the conjugate of the function

$$\frac{\frac{r^a}{a}}{\prod_{i=1}^N x_i^{\alpha_i}}$$

whose conjugate is of the same form and so, in particular, is convex as soon as $a \geqslant 1 + \sum_i^N \alpha_i$. This both illustrates conjugate analysis in action and provides a somewhat surprising explicit formula. It also supplies a large class of test examples for symbolic convex computation.

When $\alpha_i \equiv 0$ we see that we reduce to $\frac{r^a}{a}$.

2. Some Conjugate Formulae.

Let positive numbers $a, \alpha_1, \cdots, \alpha_N$ be given satisfying $a > 1+\alpha, \alpha := \sum_{i=1}^N \alpha_i$. Let $x := (x_1, \cdots, x_N)$, $y := (y_1, \cdots, y_N)$ when appropriate and define

$$f(x,r) := f(x_1, \cdots, x_N, r) := \frac{\frac{r^a}{a}}{\prod_{i=1}^N x_i^{\alpha_i}}$$

for $r \geqslant 0, x_1, \cdots, x_N > 0$.

It is convenient to start by computing the conjugate for the function obtained when $r \equiv 1$.

LEMMA 1. *Let $x_1 > 0, \cdots, x_N > 0$ and $\gamma_1 > 0, \cdots, \gamma_N > 0$ be given. Let*

$$g(x) := \frac{1}{\prod_{i=1}^N x_i^{\gamma_i}}.$$

(a) Then

$$g^*(y) := -(\gamma+1) \prod_{i=1}^N \left\{ \frac{-y_i}{\gamma_i} \right\}^{\frac{\gamma_i}{\gamma+1}},$$

where $\gamma := \sum_{i=1}^N \gamma_i$.

(b) Moreover,

$$g(x) = g^{**}(x)$$

for all x and so g is convex.

Proof. Part (a). Let $\Delta(x) := \ln(g(x))$. Then $g(x) := e^{\Delta(x)}$ where

$$\Delta(x) := -\sum_{i=1}^N \gamma_i \ \ln(x_i).$$

Now we may apply the formula for the conjugate of a composition with an increasing convex function. In general terms, this is

$$(mF)^*(y) := \inf_{t \geqslant 0} \{m^*(t) + t\, F^*(\frac{y}{t})\},$$

which is certainly valid when m is everywhere finite ([1], p. 69). Here we use the convention that $0\ F^*(\frac{y}{0})$ denotes the *recession function* of F^* at y which gives $\sup_{F(x)<\infty} \langle y, x \rangle$. Applied to e and Δ this yields

$$g^*(y) := \inf_{t \geqslant 0} \{t\ \ln(t) - t + t\, \Delta^*\left(\frac{y}{t}\right)\} = \inf_t \{t\ \ln(t) - t + t \sum_{i=1}^N \gamma_i \{-1 - \ln\left(\frac{-y_i}{t\gamma_i}\right)\}\}.$$

This may be directly minimized and obtains its minimum at $t_0 := e^{\left(\frac{-L}{\gamma+1}\right)}$, where

$$L := -\sum_{i=1}^N \gamma_i \ \ln\left(\frac{-y_i}{\gamma_i}\right).$$

Now substitution and simplification completes the proof of (a).

Part (b) follows from a similar computation applied to $\ln(g^*(y))$. $\square$

Now we can prove our main formula.

THEOREM 1. (a) *Under the hypotheses above, for* $s \geqslant 0, y_1, \cdots, y_N < 0$

$$f^*(y_1, ..., y_N, s) := \Gamma \frac{\frac{s^b}{b}}{\prod_{i=1}^N (-y_i)^{\beta_i}}$$

where

$$b := \frac{a}{a-(\alpha+1)}, \qquad \beta_i := \frac{\alpha_i}{a-(\alpha+1)}$$

and

$$\Gamma := \prod_{i=1}^N \left\{\frac{\beta_i}{b}\right\}^{\beta_i} = \prod_{i=1}^N \left\{\frac{\alpha_i}{a}\right\}^{\beta_i}.$$

(b) *In particular* $b \geqslant 1+\beta, \beta := \sum_i^N \beta_i$, *and so in essence* f *and* f^* *are symmetric whence* $f^{**} = f$ *and* f *is convex.*

Proof. To establish (a) we argue as follows.

$$f^*(y_1, \cdots, y_N, s) := \sup_r \left\{ rs + \sup_x \langle x, y \rangle - \frac{\frac{r^a}{a}}{\prod_{i=1}^N x_i^{\alpha_i}} \right\}.$$

The inner conjugate is $\frac{r^a}{a} g^* \left(\frac{ay}{r^a}\right)$ where

$$g(x) := f(x_1, \cdots, x_n, 1) = \frac{1}{\prod_{i=1}^N x_i^{\alpha_i}}.$$

In combination with Lemma 1, this yields

$$f^*(y_1, \cdots, y_N, s) = \sup_r \left\{ rs - (-g^*(y)) \left(\frac{r^a}{a}\right)^{\frac{1}{1+\alpha}} \right\}$$

and now some care with computing the remaining one–dimensional conjugate produces the desired result.

Part (b) now follows easily since the function defined by $h(y) := g(-y)$ has conjugate satisfying $h^*(x) = g^*(-x)$. $\square$

Iterated conjugation, as used above is a very useful tool in symbolic or numeric computation of conjugates. Convexity is established in ([4]) by other methods but the conjugate is not considered. In ([4]) and in ([2]) a central motivation is the use of the function as a *barrier or interior penalty* function.

3. Additional Comments

(a) For $N := 0$, this recovers Young's result that $\frac{r^a}{a}$ has conjugate $\frac{s^b}{b}$, whenever a and b are complementary. It is a simple matter to replace r and s by $|r|$ and $|s|$ respectively.

(b) Obviously it is now possible to explicitly characterize equality in the Young's inequality $f(x,r) + f^*(y,s) \geqslant rs + \langle x,y\rangle$. Moreover, the smoothness of the conjugate establishes the strict convexity of the function.

(c) The result remains true when $a = \alpha + 1$. One can argue directly (via homogenization), and one can also derive the result by taking limits in Theorem 1. In this case $f^*(y,s)$ is "roughly" the indicator function of $\{(y,s) : g^*(y) + s \leqslant 0\}$. Here as elsewhere we keep the implicit sign constraints on x and y.

(d) One may make the result totally symmetric (i.e., "distribute" Γ better). Let

$$\Lambda := \prod_{i=1}^{N} \left\{\frac{\alpha_i}{a}\right\}^{\frac{\alpha_i}{2}} \quad \text{whence} \quad \Lambda^* := \prod_{i=1}^{N} \left\{\frac{\beta_i}{b}\right\}^{\frac{\beta_i}{2}}$$

Let $\varepsilon := a - (\alpha+1)$ and $\eta := b - (\beta+1)$. Then $\eta\varepsilon = 1$ and

$$\Lambda \frac{r^a}{a\prod_{i=1}^N x_i^{\alpha_i}} \quad \text{and} \quad \Lambda^* \frac{s^b}{b\prod_{i=1}^N (-y_i)^{\beta_i}},$$

are conjugate. Theorem 1 can clearly be recast in this way. It is also then possible to write the conditions on α_i and β_i entirely symmetrically. Hence, we have

$$\Lambda \frac{r^a}{a\prod_{i=1}^N x_i^{\alpha_i}} + \Lambda^* \frac{s^b}{b\prod_{i=1}^N (-y_i)^{\beta_i}} \geqslant r\,s + \langle x,y\rangle .$$

(e) The formula is most symmetric when $\varepsilon = a - (\alpha+1) = 1$. For example, if $N := 1$ and $a := 3$ then $\alpha := 1$, and we see that $\frac{r^3}{3\sqrt{3x}}$ has conjugate $\frac{s^3}{3\sqrt{3(-y)}}$. More generally we see that for $a := 2 + \sum_{I=1}^N \alpha_i$, the function

$$f(x,r) := \prod_{i=1}^{N} \left\{\frac{\alpha_i}{a}\right\}^{\frac{\alpha_i}{2}} \frac{r^a}{a\prod_{i=1}^N x_i^{\alpha_i}}$$

satisfies $f^*(y,s) = f(-y,s)$ and so $f(x,r) + f(y,s) \geqslant r\,s - \langle x,y\rangle$, and f is self–conjugate up to a minus sign. For $\alpha_i \equiv \frac{\alpha}{N}$ we have

$$f(x,r) = \left\{\frac{\alpha}{Na}\right\}^{\frac{\alpha}{2}} \frac{r^a}{a\{\prod_{i=1}^N x_i\}^{\alpha/N}}.$$

(f) Finally, we observe that the argument in Imai ([2]) shows that for $a \geqslant N$ the function

$$\frac{\frac{r^a}{a}}{\prod_{i=1}^N x_i}$$

is convex for x in the simplex $\{x : \sum x_i = 1, x_i > 0\}$. Our result only shows this for $a \geqslant N+1$, but in Imai's setting the denominator parameters can not be generalized asymmetrically as above.

REFERENCES

[1] JEAN-BAPTISTE HIRIART-URRUTY AND CLAUDE LEMARÉCHAL, *Convex Analysis and Minimization Algorithms II*, Springer-Verlag, Berlin, 1993.
[2] *H. Imai*, *On the convexity of the multiplicative version of Karmarkar's potential function*, Mathematical Programming **40** (1988), 29–32.
[3] J. LINDENSTRAUSS AND L. TZAFRIRI, *Classical Banach spaces I: Sequence spaces*, Springer-Verlag, Berlin, 1979.
[4] M. AVRIEL, W. DIEWERT, S. SCHAIBLE AND I. ZANG, *Generalized Concavity*, Plenum Press, London and New York, 1988.
[5] R. T. ROCKAFELLAR, *Convex Analysis*, Princeton University Press, Princeton, 1970.
[6] H. L. ROYDEN, *Real Analysis*, second edition, MacMillan Publishing, New York, 1968.

(Received June 24, 1997)

Jonathan M. Borwein
Centre for Experimental and Constructive Mathematics
Department of Mathematics and Statistics
Simon Fraser University
Burnaby, BC. V5A 1S6
Canada
e-mail: jborwein@cecm.sfu.ca
URL: www.cecm.sfu.ca