Mathematics Letters
2021; 7(4): 54-58
http://www.sciencepublishinggroup.com/j/ml
doi: 10.11648/j.ml.20210704.12
ISSN: 2575-503X (Print); ISSN: 2575-5056 (Online)

A New Proof of Young’s Inequality Using Multivariable Optimization

Tadesse Bekeshie

Department of Mathematics, Addis Ababa University, Addis Ababa, Ethiopia

Email address:
tadesse.bekeshie@aau.edu.et, taddesebekeshie@gmail.com

To cite this article:
Tadesse Bekeshie. A New Proof of Young’s Inequality Using Multivariable Optimization. *Mathematics Letters*. Vol. 7, No. 4, 2021, pp. 54-58. doi: 10.11648/j.ml.20210704.12

Received: August 18, 2021; **Accepted**: December 16, 2021; **Published**: January 24, 2022

Abstract: In its standard form, Young's inequality for products is a mathematical inequality about the product of two numbers and it allows us to estimate a product of two terms by a sum of the same terms raised to a power and scaled. This inequality, though very simple, has attracted researchers working in many fields of mathematics due to its applications. Apart from the above standard form, there are numerous refinements and variants of Young’s inequality in the literature. Some of these variants are Young’s inequality for arbitrary products, Young’s inequality for increasing functions, Young’s inequality for convolutions, Young’s inequality for integrals, Young’s inequality for matrices, trace version of Young’s inequality, determinant version of Young’s inequality, and so on. The present study examines three variants of Young’s inequality, namely the standard Young’s inequality, Young’s inequality for increasing functions and Young’s inequality for arbitrary products. There are various proofs for these three variants in the literature. For example, just like several other classical important inequalities, these inequalities can be deduced from Jensen’s inequality. The objective of this article is to provide a new alternative proof for each of them. The significance of the article lies in its attempts to open a new direction of poof so that the same approach could be applied to other useful inequalities. The proofs to be presented are based on the methods of multivariable optimization theory.

Keywords: Young’s Inequality, Multivariable Extrema, Convolution, Cauchy-Schwarz Inequality, Holder’s Inequality, Jensen Inequality

1. Introduction

Inequalities are basic tools in the development of modern mathematical theories. Mathematical analysis is largely a systematic study and exploitation of inequalities. A large number of notions and theorems in mathematics are expressed in the language of inequalities. The definitions of such useful mathematical notions as *limit, convexity*, *monotonicity*, *continuity (of a linear operator)*, *extrema*, soon involve inequalities. When a numerical value of a quantity or a formula is approximated, it is often vital to know the error introduced. Error estimates are expressed in terms of inequalities. Mathematical analysis itself is devoted to finding judicious approximations for integrals, infinite sums, solutions of differential equations, etc.

Today, the subject of inequalities became a discipline. G, H. Hardy‘s essay titled “Prolegomena" can be considered as the start of the creation of this particular discipline. Since then several books [4-7, 10-12, 15, 17] and several papers have been published on the subject. Very often an inequality is deduced from another known inequality [3, 8, 9, 17].

Many of the classical inequalities are associated with the names of famous mathematicians such as Cauchy, Schwarz, Bunyakovsky, Young, Holder, Minkowski, Hilbert, Hardy, Littlewood, Polya, Jensen, Chebyshev, Hadamard, etc. The studies related to these classical inequalities remain an active field of research to date. One of the classical inequalities that has attracted the attention of many researchers is the inequality named after the English mathematician William Henry Young who introduced it in 1912 [18, 6]. It is called Young’s inequality. Young’s inequality has several variants. The term Young’s inequality could refer to Young’s inequality for products, Young’s inequality for increasing functions, Young’s inequality for convolutions, Young’s inequality for integrals, Young’s inequality for matrices and so on.

In this article, the author considers the standard Young’s inequality, Young’s inequality for arbitrary products and

Young's inequality for increasing functions. There are various proofs of these variants of Young's inequality, in the literature. For instance, like several other classical important inequalities, they can be deduced from Jensen's inequality. In the present article the author presents new proofs for all of these variants. The proofs are based on some simple techniques from the theory of multivariable extrema.

The rest of this paper is organized as follows. In Section 2, preliminary results pertaining to Young's inequality are presented. In Section 3 the main results of the author are given. In this section new proofs are given for the standard Young's inequality, Young's inequality for three products and Young's inequality for increasing functions. The main body of the paper is culminated with a short section, section 4, which gives brief concluding remarks.

2. Preliminaries

In mathematics, Young's inequality for products is a mathematical inequality about the product of two or more numbers and it allows us to estimate a product of two terms by a sum of the same terms raised to a power and scaled. It is also called the standard Young's inequality or Young's inequality for conjugate Holder exponents. Formally it can be stated as follows.

Theorem 1 [The Standard Young's Inequality]: If a and b are nonnegative real numbers and p and q are real numbers greater than 1 such that $\frac{1}{p}+\frac{1}{q}=1$, then $ab \le \frac{a^p}{p}+\frac{b^q}{q}$. Equality holds if and only if $a^p = b^q$.

The inequality in Theorem1 can be written equivalently as $\alpha^\lambda \beta^{1-\lambda} \le \lambda\alpha + (1-\lambda)\beta$ where $\alpha, \beta \ge 0$ and $0 \le \lambda \le 1$. Various proofs of this inequality exist in the literature. In particular, it can be proved (when $\beta \ne 0$) by letting t: $=\frac{\alpha}{\beta}$ and then finding the minimum value of the single variable function $f(t) := -t^\lambda + \lambda t + (1-\lambda)$.

This inequality, though very simple, has attracted researchers working in many fields of mathematics due to its applications. For instance, the most famous classical inequalities such as Holder's inequality can be deduced easily from the standard Young's inequality [17, 9]. In [1], the authors deduced Holder's inequality from Cauchy-Schwarz inequality.

Generalizations and Extensions

Young's inequality has been generalized and extended along many directions [2, 14, 16, 19]. Several generalizations of Young inequality exist. The following inequality, called Young's inequality for increasing function, is one such generalization.

Theorem 2 [Young's Inequality for Increasing functions]:

Let f be a real-valued, continuous and strictly increasing function on $[0, c]$ with $c > 0$ and $f(0) = 0$. Let f^{-1} denote the inverse function of f. Then, for all $a \in [0, c]$ and $b \in [0, f(c)]$,

$$ab \le \int_0^a f(x) + \int_0^b f^{-1}(x)$$

with equality if and only if $b = f(a)$.

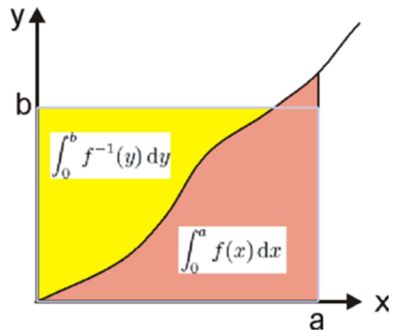

Figure 1. *Young's Inequality for increasing Functions.*

The standard Young's inequality can be deduced from Young's inequality for increasing functions by letting $f(x) = x^{p-1}$.

On the other hand, the classical Young's inequality for two scalars can be generalized to the product of n numbers as follows. First, we note that the standard Young's inequality could be equivalently written as $\alpha^\lambda\beta^{1-\lambda} \le \lambda\alpha + (1-\lambda)\beta$ where $\alpha, \beta \ge 0$ and $0 \le \lambda \le 1$.

Theorem 3 [Young's Inequality for product of n numbers]:

If $0 \le \lambda_1, \dots \lambda_n, \lambda_1 + \dots + \lambda_n = 1$ and $\alpha_1, \dots \alpha_n \ge 0$, then

$$\prod_{k=1}^n \alpha_k^{\lambda_k} \le \sum_{k=1}^n \lambda_k \, \alpha_k$$

In another development T. Ando [2] extended Young's inequality to matrices in 1995.

3. Main Result

In this section the author gives his main result, new proofs of the standard Young's inequality and its generalization to increasing functions.

Proof [Theorem 1]:

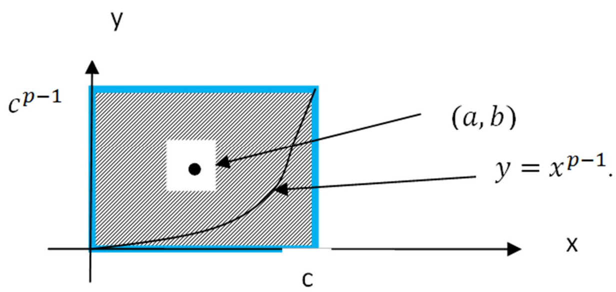

Figure 2. *Critical points of $f(x,y) := \frac{x^p}{p}+\frac{y^q}{q}-xy$ in the rectangle $R := [0,c]\times[0,c^{p-1}]$ lie on the curve $y = x^{p-1}$.*

Define $f(x,y) := \frac{x^p}{p}+\frac{y^q}{q}-xy$ where $(x,y)\in R$ and $R := [0,c]\times[0,c^{p-1}]$ is any rectangular region containing (a,b). It suffices to show that the minimum value of f on R is zero. To this end we can apply the theory of two-variable extrema as follows. Since f is continuous on the compact region R, so it has a minimum value on R (by Weirstrass Theorem for existence of optima). In view of a well-known fact in elementary calculus the extrema exist either at a critical point of f in R or at a boundary point of R. Since the partial derivatives f_x and f_y exist everywhere, so, at a local extreme point (x,y) of f we have $\nabla f(x,y) = \langle 0,0\rangle$. This implies that $y = x^{p-1}$.

Thus for every critical point (x,y) of f, we have $y = x^{p-1}$ and the value of f at such a point is 0. Moreover, the minimum value of f on the boundaries of R is greater than or equal to zero, as shown below.

Define

$$A := \{(x,y): x = 0 \,\&\, 0 \le y \le c^{p-1}\}$$

$$\mathrm{B} := \{(x,y): y = 0 \,\&\, 0 \le x \le c\}$$

$$D := \{(x,y): x = c \,\&\, 0 \le y \le c^{p-1}\}$$

$$E := \{(x,y): y = c^{p-1} \,\&\, 0 \le x \le c\}$$

These sets constitute the boundaries of rectangle$R$. Now, let us compute the extreme values of $f(x,y)$ on each of these sets.

(a) For $(x,y) \in A$, we have $x = 0$. Hence, on A, $f(x,y) := f(0,y) = \frac{y^q}{q}, 0 \le y \le c^{p-1}$.

Let $h(y) := \frac{y^q}{q}, 0 \le y \le c^{p-1}$. Using elementary calculus we get that the minimum value of h on $[0, c^{p-1}]$ is $h(0) = 0$. Hence, f assumes only non-negative values on set A.

(b) For $(x,y) \in B$, we have $y = 0$. Hence, on B, $f(x,y) := f(x,0) = \frac{x^p}{p}, 0 \le x \le c$.

Let $g(x) := \frac{x^p}{p}, 0 \le x \le c$. The minimum value of g on $[0,c]$ is $g(0) = 0$. Hence, f is nonnegative on set B.

(c) For $(x,y) \in D$, $f(x,y) := f(c,y) = \frac{c^p}{p} + \frac{y^q}{q} - cy, 0 \le y \le c^{p-1}$.

Let $k(y) := \frac{c^p}{p} + \frac{y^q}{q} - cy, 0 \le y \le c^{p-1}$. The derivative of k is 0 if and only if $y = c^{\frac{1}{q-1}} = c^{p-1}$, which is an end point of the domain of k. Hence, k has no critical point interior to its domain. Computing k at its endpoints gives us

$k(0) = \frac{c^p}{p}$ and $k(c^{p-1}) = \frac{c^p}{p} + \frac{c^p}{q} - c^p = 0$. Thus, the minimum value of k on $[0, c^{p-1}]$ is 0; hence, f assumes only non-negative values on set D.

(d) For $(x,y) \in E$, $f(x,y) := f(x, c^{p-1}) = \frac{x^p}{p} + \frac{c^p}{q} - xc^{p-1}, 0 \le x \le c$.

Define $t(x) := \frac{x^p}{p} + \frac{c^p}{q} - xc^{p-1}, 0 \le x \le c$. The minimum value of t on $[0,c]$ is 0. Hence, f is nonnegative on set E.

It follows from (a) to (d) that f assumes only nonnegative values on the boundary of rectangle R

Thus $f(x,y) \ge 0$ on R and hence, $ab \le \frac{a^p}{p} + \frac{b^q}{q}$. Q. E. D.

The same approach used to prove Theorem 1 can also be used to prove Theorem 2 as shown next.

Proof [Theorem 2]:

Define $G(x,y) := \int_0^x f(t)dt + \int_0^y f^{-1}(t)dt - xy$, where $(x,y) \in R$ and $R \coloneqq [0,c] \times [0, f(c)]$ is any rectangular region containing (a,b). As in the previous proof, it suffices to show that the minimum value of f on R is zero. The two variable function G is continuous on the compact region R, so it has a minimum value on R. The extrema of G exist either at a critical point of G in R or at a boundary point of R. Since $G_x = f(x) - y$ and $G_y = f^{-1}(y) - x$, so $\nabla G(x,y) = \langle 0,0\rangle$ if and only if $y = f(x)$. Since the partial derivatives G_x and G_y exist everywhere in R, so the only critical points of G are exactly those points (x,y) in R satisfying $y = f(x)$.

At every critical point (x,y), the value of G is 0. It remains to compute the minimum value of f on the boundaries of R and make comparison with the value of G at critical points, which is 0.

Define

$$A := \{(x,y): x = 0 \,\&0 \le y \le f(c)\}$$

$$\mathrm{B} := \{(x,y): y = 0 \,\&\, 0 \le x \le c\}$$

$$D := \{(x,y): x = c \,\&0 \le y \le f(c)\}$$

$$E := \{(x,y): y = f(c) \,\&\, 0 \le x \le c\}$$

The extreme values of $G(x,y)$ on each of these sets are computed as follows.

(a) For $(x,y) \in A$, we have $x = 0$. Hence, on A, $G(x,y) := G(0,y) = \int_0^y f^{-1}(t)dt, 0 \le y \le f(c)$.

Let $h(y) := \int_0^y f^{-1}(t)dt, 0 \le y \le f(c)$.

Using elementary calculus we get that the minimum value of h on $[0, f(c)]$ is $h(0) = 0$. Hence, G assumes only non-negative values on set A.

(b) For $(x,y) \in B$, we have $y = 0$. Hence, on B, $f(x,y) := f(x,0) = \int_0^x f(t)dt, 0 \le x \le c$.

Let $g(x) := \int_0^x f(t)dt, 0 \le x \le c$. The minimum value of g on $[0,c]$ is $g(0) = 0$. Hence, G is nonnegative on set B.

(c) For $(x,y) \in D$, $G(x,y) := G(c,y) = \int_0^c f(t)dt + \int_0^y f^{-1}(t)dt - cy, 0 \le y \le f(c)$.

Let $k(y) := \int_0^c f(t)dt + \int_0^y f^{-1}(t)dt - cy, 0 \le y \le f(c)$. The critical point of k is $y = f(c)$. Now,

$k(0) = \int_0^c f(t)dt$ and $(f(c)) = \int_0^c f(t)dt + \int_0^{f(c)} f^{-1}(t)dt - cf(c) = 0$. Thus, the minimum value of k on $[0, f(c)]$ is 0; hence, G assumes only non-negative values on set D.

(d) For $(x,y) \in E$, $G(x,y) := G(x, f(c)) = \int_0^x f(t)dt + \int_0^{f(c)} f^{-1}(t)dt - xf(c), 0 \le x \le c$.

Define $t(x) := \int_0^x f(t)dt + \int_0^{f(c)} f^{-1}(t)dt - xf(c), 0 \le x \le c$. The minimum value of t on $[0,c]$ is 0. Hence, G is nonnegative on set E.

It follows from (a) to (d) that G assumes only nonnegative values on the boundary of rectangle R.

Thus $G(x,y) \ge 0$ on R and from this the conclusion of Theorem 2 follows. Q. E. D.

Next, the proof of Young's inequality for the product of three numbers is presented. The proof uses optimization techniques, as in the proofs of Theorem 1 and Theorem 2. Whether or not such optimization techniques could be used to prove Young's inequality for the product of n numbers, for an arbitrary positive integer n, is left to the reader as an exercise. To prove Theorem 3 for $n = 3$, it suffices to prove the following restatement.

Restatement of Theorem 3 for the product of three numbers: If a, b and c are nonnegative real numbers and p, q and r are real numbers greater than 1 such that $\frac{1}{p}+\frac{1}{q}+\frac{1}{r}=1$, then $abc \leq \frac{a^p}{p}+\frac{b^q}{q}+\frac{c^r}{r}$.

Proof [Theorem 3 for n=3]:

Let

$$R:=\ [0,d]\times\left[0,d^{\frac{p}{q}}\right]\times[0,d^{\frac{p}{r}}]$$

be any rectangular box containing (a,b,c).

Define

$$f(x,y,z):=\frac{x^p}{p}+\frac{y^q}{q}+\frac{z^r}{r}-xyz,(x,y,z)\in R$$

It suffices to show that the minimum value of f on R is zero.

Step 1: We show that the value of f at its critical points is nonnegative. Let (x,y,z) be a critical point of f. Then$\nabla f(x,y,z)=\langle 0,0,0\rangle$. This implies $xyz = x^p = y^q = z^r$. At every critical point (x,y,z) of f we get that $f(x,y,z):=\frac{x^p}{p}+\frac{y^q}{q}+\frac{z^r}{r}-xyz = xyz\left(\frac{1}{p}+\frac{1}{q}+\frac{1}{r}\right)-xyz=0$, Thus f is zero at every critical pint (x,y,z).

Step 2: We show that the value of f at every boundary point of R is nonnegative. Let the six faces that form the boundary surface of R be labeled as follows:

$$A:=\left\{(x,y,z): x=0, 0\leq \mathrm{y}\leq \mathrm{d}^{\frac{p}{q}}\ \&0\leq z\leq d^{\frac{p}{r}}\right\}$$

$$\mathrm{B}:=\left\{(x,y,z): 0\leq x\leq d, \mathrm{y}=0\ \&0\leq z\leq d^{\frac{p}{r}}\right\}$$

$$D:=\left\{(x,y,z): 0\leq x\leq d, 0\leq \mathrm{y}\leq \mathrm{d}^{\frac{p}{q}}\ \&\mathrm{z}=0\right\}$$

$$E:=\left\{(x,y,z): x=\mathrm{d}, 0\leq \mathrm{y}\leq \mathrm{d}^{\frac{p}{q}}\ \&0\leq z\leq d^{\frac{p}{r}}\right\}$$

$$\mathrm{F}:=\left\{(x,y,z): 0\leq x\leq d, \mathrm{y}=\mathrm{d}^{\frac{p}{q}}\ \&0\leq z\leq d^{\frac{p}{r}}\right\}$$

$$G:=\left\{(x,y,z): 0\leq x\leq d, 0\leq \mathrm{y}\leq \mathrm{d}^{\frac{p}{q}}\ \&\mathrm{z}=d^{\frac{p}{r}}\right\}$$

Clearly, if (x,y,z) lies in A, B or D, then$f(x,y,z)=0$.

It remains to find the minima of f on the remaining three faces of R (i.e., E, F and G).

If *(x, y, z)*$\in E$, then $x = d$ and we get that

$$f(x,y,z):=f(d,y,z)=\frac{d^p}{p}+\frac{y^q}{q}+\frac{z^r}{r}-dyz$$

Define

$$h(y,z):=f(d,y,z)=\frac{d^p}{p}+\frac{y^q}{q}+\frac{z^r}{r}-dyz;$$

$$0\leq \mathrm{y}\leq \mathrm{d}^{\frac{p}{q}}\ \&0\leq z\leq d^{\frac{p}{r}}.$$

If (y,z) is a critical point of h, then $dyz = y^q = z^r$ and h(y, z) $=\frac{d^p}{p}+\frac{dyz}{q}+\frac{dyz}{r}-dyz=\frac{d^p}{p}-\frac{y^q}{p}\geq 0$. Hence, f is nonnegative at every critical point of h in E. On the other hand, it can be checked easily that the values of h (and hence f) on the boundaries of E are nonnegative, too. Thus f is nonnegative on E.

As the reader can easily verify f is also nonnegative on the remaining two faces of R. Q. E. D.

4. Conclusion

This mini article never aims at providing easier proofs for the three forms of Young's inequality it considers; because simple proofs of these inequalities are abound in literature. Rather, it attempts to open a new direction of poof so that the same approach could be applied to other useful inequalities. Specially, the technique of proof used in the article could be applied to Young's inequality for arbitrary products.

Acknowledgements

The author wishes to thank the anonymous reviewer (s) for his/her/their helpful suggestions and comments.

References

[1] M. Akkouchi: Cauchy-Schwarz inequality implies Holder's inequality, RGMIA Research report collection, vol. 21, Art. ID. 48, 2018.

[2] T. Ando: Matrix Young inequality, Oper. Theory Adv. Appl. 75 (1995) 33–38.

[3] P. K. Bhandari and S. K. Bissu: Inequalities via Holder's inequality, Scholars journal of research in mathematics and computer science, vol. 2, no. 2, pp. 124-129, 2018.

[4] P. S. Bullen: A dictionary of inequalities, Pitman Monographs and Surveys in Pure and applied mathematics 97, Addison Wesley Longman Ltd. UK., 1998.

[5] P. S. Bullen: Handbook of Means and Their Inequalities, Kluwer Academic Publishers, Dordrecht, 2003. doi: 10.1007/978-94-017-0399-4.

[6] G. H. Hardy, J. E. Littlewood and G. Polya: Inequalities, Cambridge, 1934.

[7] Herman, R. Kucera, J. Simsa: Equations and inequalities, CMS Books in Mathematics, Springer 2000.

[8] C. A. Infantozzi: An introduction to relations among inequalities, The November Meeting in Cleveland, Ohio November 25, 1972, Notices of the American mathematical society, no. 141, pp. A819-A820, 1972.

[9] Y. C. Li and S. Y. Shaw: A proof of Hölder's inequality using the Cauchy-Schwarz inequality, Journal of inequalities in pure and applied mathematics, vol. 7, no. 2, Art. ID. 62, 2006.

[10] A. Lohwater: Introduction to Inequalities. e-book in PDF. http://www.mediafire.com/?1mw1tkgozzu, 1982.

[11] D. S. Mitrinovic: Analytic Inequalities, Springer-Verlag, Berlin Heiidelberg 1970.

[12] D. S. Mitrinovic, J. E. Pečarić and A. M. Fink: Classical and new inequalities in analysis, Dordrecht: Kluwer Academic Publishers, 1993.

[13] F. C. Mitroi & C. P. Niculescu: An extension of Young's inequality in Abstract and Applied Analysis (Vol. 2011). Hindawi.

[14] Z. Pales: On Young-type inequalities, Acta Scientiarum Mathematicarum, vol. 54, no. 3-4, pp. 327–338, 1990.

[15] J. M. Steel: The Cauchy-Schwarz Master Class: An Introduction to the Art of Mathematical Inequalities, Cambridge University Press, New York, 2004.

[16] W. T. Sulaiman: Notes on Youngs's inequality, International Mathematical Forum, vol. 4, no. 21–24, pp. 1173–1180, 2009.

[17] E. Tolsted: An elementary derivation of the Cauchy, Holder, and Minkowski inequality- ties from Young's inequality, Math. Mag., 37 (1964) 2-12.

[18] W. H. Young: On classes of summable functions and their Fourier series, Proc. Royal Soc., Series (A), 87 (1912) 225-229.

[19] J. Zhang, J. Wu: New progress on the operator inequalities involving improved Young's and its reverse inequalities relating to the Kantorovich constant. *J Inequal Appl* 2017, 69 (2017).