Miscelánea Matemática **52** (2011) 63–73 SMM

Miscelánea de desigualdades: Bernoulli, Young y varias más

César L. García
clgarcia@itam.mx

Héctor E. Lomelí
lomeli@itam.mx

Departamento de Matemáticas
Instituto Tecnológico Autónomo de México
México, DF 01080

1. Introducción

Esta historia comienza en un primer curso de análisis matemático. De manera estándar, el curso inicia con los preliminares de rigor y sigue con la construcción de los números reales o su presentación axiomática. Las reglas del juego son, entre otras, no usar algo que no haya sido previamente demostrado y –con esta filosofía– alcanzar una buena descripción de las propiedades estructurales del conjunto de los números reales y de ahí hacer lo propio para el espacio $\mathbb{R}^n$. Ya después vendrá lo correspondiente para los espacios métricos y los espacios normados.

En este esquema, las desigualdades numéricas ofrecen todo un reto. Rápidamente uno encuentra desigualdades tan precisas (o justas) que decidir su validez y establecer su optimalidad es todo un *tour de force*. Desafortunadamente, uno es orillado –a veces sin remedio– a recurrir a una herramienta que permite lograr fineza y exactitud: el cálculo diferencial e integral. Sin embargo, las nociones de derivada e integral difícilmente llegan a discutirse en un primer curso de análisis matemático por lo que la herramienta aparece fuera de contexto y en conflicto con la filosofía del curso.

En este artículo queremos mostrar y resolver un ejemplo de la problemática anterior. Describiremos cómo manejar la desigualdad de Young dentro de un curso de análisis sin recurrir al cálculo que es la manera clásica en la que ésta se demuestra. La insistencia de querer comprobarla con "lo que tenemos a la mano" nos llevó a lo que es material de las siguientes secciones.

El reto que nos planteamos es obtener la desigualdad de Young y otras desigualdades clásicas a partir de una desigualdad aparentemente sencilla: la desigualdad de Bernoulli. Al final encontraremos también que las desigualdades que resultan de la desigualdad de Bernoulli son equivalentes entre sí.

2. ¿Por qué la desigualdad de Young?

La primera vez que uno puede encontrarse con la desigualdad de Young[1] es cuando se buscan otras maneras de normar el espacio vectorial $\mathbb{R}^n$. La norma usual con la que se equipa a $\mathbb{R}^n$ es la norma euclidiana:

$$\|(x_1, x_2, \ldots, x_n)\|_2 = \left(\sum_{i=1}^{n} |x_i|^2\right)^{1/2}.$$

Esta norma satisface la desigualdad del triángulo, $\|\bar{x}+\bar{y}\|_2 \leq \|\bar{x}\|_2 + \|\bar{y}\|_2$, donde $\bar{x} = (x_1, x_2, \ldots, x_n)$ y $\bar{y} = (y_1, y_2, \ldots, y_n)$, porque se tiene la desigualdad de Cauchy-Bunyakovskii-Schwarz[2] (CBS):

$$\sum_{i=1}^{n} |x_i y_i| \leq \left(\sum_{i=1}^{n} |x_i|^2\right)^{1/2} \left(\sum_{i=1}^{n} |y_i|^2\right)^{1/2}, \tag{1}$$

que a su vez es consecuencia del teorema egregio que rige a la geometría del espacio euclídeo, a saber, el teorema de Pitágoras[3]:

$$\|\bar{x}+\bar{y}\|_2^2 = \|\bar{x}\|_2^2 + \|\bar{y}\|_2^2 \quad \text{si y solo si} \quad \bar{x}\cdot\bar{y} = 0,$$

donde $\bar{x}\cdot\bar{y} := \displaystyle\sum_{i=1}^{n} x_i y_i$.

Existen otras normas "naturales" para $\mathbb{R}^n$ como la norma del taxista:

$$\|(x_1, x_2, \ldots, x_n)\|_1 = \sum_{i=1}^{n} |x_i|,$$

[1] William Henry Young, 1863-1942.

[2] Augustin-Louis Cauchy, 1789-1857; Victor Bunyakovskii, 1804-1889; Hermann Amandus Schwarz, 1843-1921.

[3] Pitágoras de Samos, -569 a -475.

o la norma del máximo:

$$\|(x_1, x_2, \ldots, x_n)\|_\infty = \text{máx}\{|x_i| : i = 1, 2, \ldots, n\}.$$

También se pueden considerar otras normas, como las llamadas de Minkowski[4]: si p es un número real mayor o igual a uno entonces

$$\|(x_1, x_2, \ldots, x_n)\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p}$$

es una norma en $\mathbb{R}^n$. La desigualdad del triángulo para las normas de Minkowski ya no es geométricamente evidente como en el caso euclidiano ($p = 2$). Ahora es una desigualdad analítica que depende de una desigualdad que generaliza la desigualdad de CBS y que se conoce como la desigualdad de Hölder[5]: si p y q son números reales, ambos mayores a uno y tales que $1/p + 1/q = 1$, entonces

$$\sum_{i=1}^{n} |x_i y_i| \leq \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p} \left(\sum_{i=1}^{n} |y_i|^q \right)^{1/q}. \tag{2}$$

La desigualdad de Hölder, como muchas otras desigualdades, puede comprobarse fácilmente si uno encuentra y comprueba la desigualdad real adecuada, en este caso es la desigualdad de Young:

Proposición 2.1 (Desigualdad de Young) *Si a y b son números reales no-negativos y p y q son números reales, ambos mayores a uno y tales que $1/p + 1/q = 1$ entonces*

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q} \tag{3}$$

con igualdad si y sólo si $a^p = b^q$.

La desigualdad de Young es una desigualdad muy justa o precisa y no es claro *a priori* como comprobarla sin recurrir al cálculo. De hecho, casi seguramente cualquier búsqueda de una demostración conduce a una de las siguientes dos:

2.1. Una prueba con cálculo integral

La prueba vía integración consiste en apelar a una figura escogida adecuadamente. Considerar la gráfica de $y = x^{p-1}$ sobre el intervalo $[0, a]$

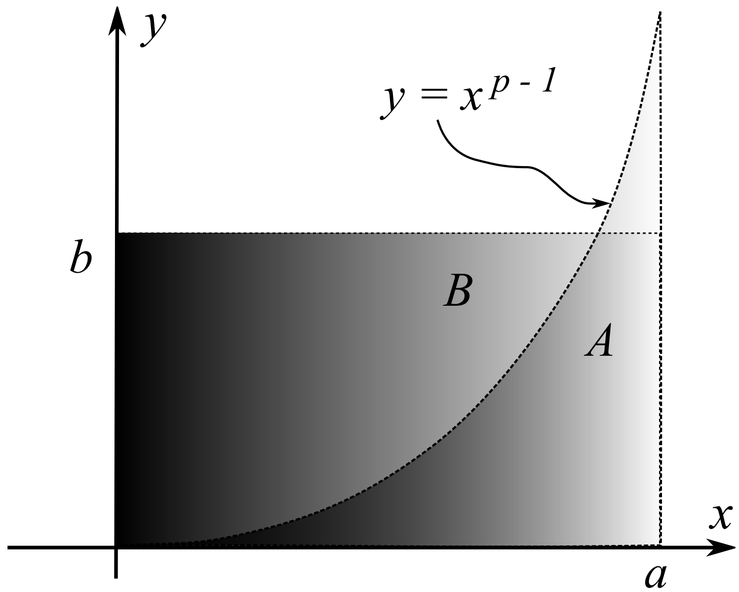

Figura 1: Prueba de la desigualdad de Young para el caso $b \leq a^{p-1}$.

(recordemos que $p > 1$) y marquemos el intervalo $[0, b]$ sobre el eje y. Supongamos además que $b \leq a^{p-1}$ como en la figura 1.

El área A en la figura 1 es

$$\int_0^a y\,dx = \int_0^a x^{p-1}\,dx = \frac{a^p}{p}.$$

mientras que el área B es

$$\int_0^b x\,dy = \int_0^b y^{1/(p-1)}\,dy = \frac{(p-1)b^{p/(p-1)}}{p} = \frac{b^q}{q}.$$

La suma de estas dos áreas ciertamente es mayor o igual al área, ab, del rectángulo $[0, a] \times [0, b]$ y de ahí la desigualdad de Young. La figura misma sugiere cuál es la condición para que se tenga igualdad en la desigualdad de Young. El caso $b \geq a^{p-1}$ corresponde a $b^{q-1} \geq a$ y se resuelve de manera similar.

La figurita es tan buena (no nos referimos a ningún mérito artístico) que fácilmente puede enunciarse la versión de la desigualdad de Young para una función estrictamente creciente y continua sobre el intervalo $[0, a]$: si $f : [0, a] \to \mathbb{R}$ es una tal función y $f(0) = 0$ entonces,

$$\int_0^a f(t)dt + \int_0^b f^{-1}(t)dt \geq ab.$$

[4]Hermann Minkowski, 1864-1909.

[5]Ludwig Otto Hölder, 1859-1937.

Con igualdad si y sólo si $b = f(a)$ (ver, por ejemplo [1] para una discusión completa de la desigualdad en esta forma).

2.2. Una prueba con cálculo diferencial

Una demostración, usando la noción de derivada, es vía la siguiente desigualdad

$$t^\alpha \leq \alpha t + (1-\alpha) \quad \text{si} \quad 0 < \alpha < 1 \text{ y } t > 0. \tag{4}$$

El lado derecho de la desigualdad (4) es la ecuación de la recta tangente a la gráfica de t^α en $(1,1)$. Notemos que la función $h(t) = \alpha t + (1-\alpha) - t^\alpha$ tiene un mínimo absoluto en $t = 1$ y es convexa (cóncava hacia arriba) luego $h(t) \geq h(1) = 0$ y de aquí la desigualdad. En particular, para $t = a^p/b^q$ y $\alpha = 1/p$, la desigualdad (4) implica que

$$t^{1/p} \leq \frac{1}{p}t + \frac{1}{q},$$

y por lo tanto,

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

Notemos que

$$\frac{b^q}{b^{q/p}} = b^{q-\frac{q}{p}} = b.$$

Si se multiplica la desigualdad anterior por b^q, se obtiene la desigualdad de Young (3).

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

Young comprobó esta desigualdad para la convolución de funciones en espacios L_p en el contexto de series de Fourier alrededor de 1912 (ver por ejemplo, [6, secciones 8.3, 8.4] y las referencias ahí citadas). Esta versión de la desigualdad, al igual que la desigualdad de Minkowski (del triángulo para normas p), se obtiene a partir de la desigualdad de Young (3).

3. Dándole la vuelta al cálculo

¿Cómo comprobar la desigualdad de Young sin apelar al cálculo? Partiendo de primeros principios, lo único que se tendría disponible es la definición de potencias reales para números reales: si $t > 1$ y $0 < \alpha < 1$ entonces

$$t^{\alpha} := \sup\{t^q : q \in \mathbb{Q}, q \leq \alpha\}. \tag{5}$$

Por supuesto, para llegar a la definición (5) se define o comprueba la existencia, según sea el caso, de las potencias: t^α, $\alpha \in \mathbb{N}$; t^α, $\alpha \in \mathbb{Z}$ y t^α, $\alpha = 1/r$, $r \in \mathbb{N}$ (la raíz r-ésima de t). Amén de las extensiones naturales de (5) a rangos más grandes de t y α (el rango aquí elegido, $t > 1$ y $0 < \alpha < 1$, será suficiente para nuestros fines). Veamos pues una demostración de la desigualdad de Young usando la definición de potencia de un número real. La idea es deducir (4) a partir de una desigualdad elemental, a saber, la desigualdad de Bernoulli:[6]

Teorema 3.1 (Desigualdad de Bernoulli) *Si $x > -1$ entonces*

$$(1+x)^n \geq 1 + nx \quad \textit{para todo} \quad n \in \mathbb{N}. \tag{6}$$

La desigualdad de Bernoulli es obvia si $x \geq 0$ pues el lado derecho de la desigualdad está constituido solamente por los primeros dos sumandos de la expansión del binomio $(1+x)^n$ y para $-1 < x < 0$ una prueba sencilla por inducción matemática resuelve el problema (de hecho el caso para cualquier $x > -1$). En efecto, sea $S = \{n \in \mathbb{N} : (1+x)^n \geq 1 + nx\}$. Primero, $(1+x)^1 = 1 + 1 \cdot x$, luego $1 \in S$. Por otro lado si $k \in S$, es decir, $(1+x)^k \geq 1 + kx$, entonces

$$\begin{aligned}(1+x)^{k+1} = (1+x)(1+x)^k &\geq (1+x)(1+kx)\\ &= 1 + (k+1)x + kx^2 \geq 1 + (k+1)x.\end{aligned}$$

Por el principio de inducción se concluye que $S = \mathbb{N}$; es decir, la fórmula es válida para todo número natural. Para ser honestos, notemos que aunque estamos exhibiendo la desigualdad de Bernoulli como una propiedad meramente aritmética de números reales, la desigualdad esconde una propiedad geométrica que se decide fácilmente con cálculo: la recta tangente en $x = 0$ a la gráfica de la función $y = (1+x)^n$ es $y = 1 + nx$ y esta última va por abajo de la gráfica de $(1+x)^n$ en el rango $x \geq -1$.

La desigualdad de Bernoulli resulta, a pesar de ser elemental, una llave mágica: no sólo se logran corolarios como la desigualdad del triángulo para las normas de Minkowski como ya mencionamos, sino también, puede ser usada para dar estimaciones elegantes y sencillas de las cuales se obtienen teoremas fundamentales en el análisis matemático como el teorema de aproximación polinomial de Weierstrass (ver por ejemplo, [7]). Para una revisión completa de la desigualdad, así como

[6]Jacob Bernoulli, 1654-1705.

de sus extensiones ver [9] y [10]. También recomendamos [3] y las referencias ahí citadas pues comparte la filosofía de este artículo y, por supuesto, usa la desigualdad de Bernoulli entre otras.

Aunque la desigualdad de Bernoulli aparece en el artículo de Jacob Bernoulli: "Positiones arithmeticae de seriebus infinitis earumque summa finita" de 1689, no es la primera noticia sobre esta desigualdad: se sabe que la desigualdad ya era conocida por Isaac Barrow[7] (contemporáneo de Isaac Newton[8] y solamente por esto segundo gran científico de la época según sus colegas, [4]) desde al menos veinte años antes del artículo de Bernoulli. De hecho, la primera aparición por escrito de la desigualdad es en el *Mesolabum* de Sluze[9] según M. Feingold, [4].

4. La desigualdad de Bernoulli a buen uso

Demostremos algunas de las sorprendentes consecuencias de la desigualdad de Bernoulli. El primero de tres pasos hacia la desigualdad de Young y en el que se usa la desigualdad de Bernoulli es la siguiente proposición:

Proposición 4.1 *Sea $m \in \mathbb{N}$ y $z > -m$. Entonces, para todo $n \in \mathbb{N}$ con $n \geq m$ se cumple*

$$\left(1 + \frac{z}{n}\right)^n \geq \left(1 + \frac{z}{m}\right)^m. \tag{7}$$

El avezado lector habrá notado que en esta desigualdad aparecen términos de la conocida sucesión asociada a la función exponencial. No es coincidencia, pues recordemos que en la desigualdad de Young involucramos potencias reales de números reales. La desigualdad (7) resulta además pedagógicamente muy ventajosa pues lleva –casi de manera inmediata– a la definición de función exponencial y todas sus propiedades. Los términos de la forma $\left(1 + \frac{z}{n}\right)^n$ fueron asociados a la función exponencial por primera vez por Euler[10] quién además fue el primero en entender el papel fundamental del número e y la función exponencial $y = e^z$ en el análisis. Hasta antes de Euler la función exponencial era considerada como la función "al revés" de la función logarítmica. Euler es el primero en definir e^z como límite de la sucesión $\left(\left(1 + \frac{z}{n}\right)^n\right)$. Un hecho simpático, según Maor (ver, [8, pág. 156]), es que Euler en su

[7] Isaac Barrow, 1630-1677.
[8] Sir Isaac Newton, 1643-1727
[9] René François Walter de Sluze, 1622-1685.
[10] Leonhard Euler, 1707-1783.

Introductio escribió $\lim_{n\to\infty}\left(1+\frac{z}{n}\right)^n$ como $\left(1+\frac{z}{i}\right)^i$ donde i denota un "número infinito" cosa que, agrega Maor, ningún estudiante de primer año universitario se atrevería a hacer.

La demostración de (7) es por inducción sobre n. En efecto, notemos que para cualquier $n \geq m$ se cumple que $1+z/n > 0$. El paso inductivo se sigue de la desigualdad de Bernoulli (6), $(1+x)^{n+1} \geq 1+(n+1)x$, usando como x a

$$x = \frac{-z}{(n+1)n}\left(1+\frac{z}{n}\right)^{-1}.$$

En efecto, un poco de álgebra muestra que

$$1+x = \left(1+\frac{z}{n+1}\right)\left(1+\frac{z}{n}\right)^{-1} > 0 \quad \text{(luego se puede aplicar (6))}$$

y que

$$1+(n+1)x = \left(1+\frac{z}{n}\right)^{-1}.$$

Así, se tiene la desigualdad

$$(1+x)^{n+1} \geq \left(1+\frac{z}{n}\right)^{-1}.$$

Al multiplicar ambos lados de esta desigualdad por $(1+z/n)^{n+1}$ el lado derecho queda $(1+z/n)^n$, que por hipótesis de inducción es mayor o igual a $(1+z/m)^m$, mientras que el lado izquierdo se simplifica en $(1+z/(n+1))^{n+1}$.

El segundo paso en la demostración de la desigualdad de Young es usar la desigualdad (7) para comprobar la desigualdad (4). Sean entonces $t \geq 1$ y $0 < \alpha < 1$. Consideremos un número racional $m/n \in \mathbb{Q}$, con $m/n \leq \alpha$ y $m < n$. Definimos $z = m(t-1)$. Claramente $z > -m$ y por la desigualdad (7) se cumple que

$$\left(1+\frac{z}{m}\right)^m \leq \left(1+\frac{z}{n}\right)^n.$$

Pero entonces,

$$t^m \leq \left(1+\frac{m}{n}(t-1)\right)^n,$$

y por lo tanto

$$t^{m/n} \leq 1+\frac{m}{n}(t-1) \leq 1+\alpha(t-1).$$

Así, al tomar el supremo –tomando en cuenta la definición 5– se obtiene la desigualdad (4) enunciada anteriormente:

$$t^\alpha \leq \alpha t + (1-\alpha).$$

Finalmente, el tercer y último paso consiste en observar que si $a^p \geq b^q$ entonces al sustituir $t = a^p/b^q \geq 1$ y $\alpha = 1/p$ en (4) se obtiene la desigualdad de Young. La conclusión en el caso $a^p < b^q$ se obtiene de manera similar, usando ahora $t = b^q/a^p$ y $\alpha = 1/q$. La igualdad en la desigualdad de Young es consecuencia de que la igualdad en la desigualdad (4) ocurre si y sólo si $t = 1$ (dejamos al amable lector incursionar en los detalles).

5. Una desigualdad más

Sin duda una de las desigualdades más importantes de la geometría y del análisis convexo es la que compara la media geométrica con la media aritmética (*MG-MA*) de números reales positivos. Como corolario de la desigualdad de Young se obtiene fácilmente la siguiente versión de la desigualdad entre las medias mencionadas.

Corolario 5.1 (MG-MA) *Sean $x_1, \ldots, x_n$ números reales positivos. Sean $\alpha_1, \ldots, \alpha_n$ números no negativos tales que*

$$\sum_{k=1}^{n} \alpha_k = 1.$$

Entonces se cumple

$$x_1^{\alpha_1} \cdots x_n^{\alpha_n} \leq \alpha_1 x_1 + \cdots + \alpha_n x_n.$$

La versión "usual" de la desigualdad *MG-MA* se obtiene cuando cada α_i es igual a $1/n$ con n un número natural.

Haremos la demostración del corolario por inducción sobre n. Claramente la afirmación es cierta para $n = 1$ pues en este caso $\alpha_1 = 1$. Supongamos que la afirmación es cierta para el número natural n. Sean $x_1, \ldots, x_{n+1}$ números positivos y sean $\alpha_1, \ldots, \alpha_{n+1}$ números no negativos tales que

$$\sum_{k=1}^{n+1} \alpha_k = 1.$$

Los casos $\alpha_{n+1} = 0$ y $\alpha_{n+1} = 1$, son triviales. Supongamos, sin pérdida de generalidad, que $0 < \alpha_{n+1} < 1$. Definimos $q = 1/\alpha_{n+1}$. Claramente $q > 1$. Sea p el exponente conjugado a q; es decir, p satisface $1/p + 1/q = 1$.

Si $a = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ y $b = x_{n+1}^{\alpha_{n+1}}$ entonces la desigualdad de Young implica que

$$x_1^{\alpha_1} \cdots x_n^{\alpha_n} x_{n+1}^{\alpha_{n+1}} = ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

Ahora,

$$\frac{b^q}{q} = \alpha_{n+1} x_{n+1}$$

y además,

$$a^p = x_1^{p\alpha_1} \cdots x_n^{p\alpha_n}.$$

Pero como

$$\sum_{k=1}^{n} \alpha_k = 1 - \alpha_{n+1} = \frac{1}{p}$$

concluimos que

$$\sum_{k=1}^{n} p\alpha_k = 1.$$

Usando la hipótesis de inducción, tenemos que

$$\frac{a^p}{p} \leq \alpha_1 x_1 + \cdots + \alpha_n x_n,$$

y esto implica el resultado.

6. Epílogo

La versión usual de la desigualdad *MG-MA*,

$$(x_1 \cdot x_2 \cdots x_n)^{1/n} \leq \frac{x_1 + x_2 + \cdots + x_n}{n},$$

con n un número natural y $x_1, \ldots, x_n$ números reales positivos puede comprobarse, también a partir de primeros principios por inducción. La prueba clásica es la de Cauchy, quien resuelve la desigualdad para números naturales de la forma $n = 2^k$ y después, extiende la desigualdad para todo número natural sin usar más que mera aritmética. La demostración de Cauchy puede consultarse en el muy recomendable libro de desigualdades de Garling [5]. Otra demostración, también inductiva, puede verse en esta *Miscelánea* [2].

Sorprendentemente, la desigualdad (7) también es consecuencia de la desigualdad *MG-MA*. En efecto, si en la desigualdad *MG-MA* hacemos $x_1 = x_2 = \cdots = x_m = 1 + \frac{z}{m}$ y $x_{m+1} = \cdots = x_n = 1$ entonces

$$\left(1 + \frac{z}{m}\right)^{m/n} = (x_1 \cdots x_n)^{1/n} \leq \frac{x_1 + \cdots + x_n}{n} = 1 + \frac{z}{n}.$$

Se cierra así un círculo de desigualdades que pende de la desigualdad de Bernoulli:

$$\begin{array}{ccccc} & & \text{Bernoulli} & & \\ & & \downarrow & & \\ & & \left(1+\frac{z}{n}\right)^n \geq \left(1+\frac{z}{m}\right)^m & & \\ & \nearrow & & \searrow & \\ \text{MG-MA} & & & & t^\alpha \leq \alpha t + (1-\alpha) \\ & \nwarrow & & \swarrow & \\ & & \text{Young} & & \end{array}$$

Referencias

[1] J. B. Díaz, F. T. Metcalf, *An Analytic Proof of Young's Inequality*, Amer. Math. Monthly, **77**, no. 6 (1970), 603–609.

[2] J. Cruz Sampedro, M. Tetlalmatzi Montiel, *Una demostración inductiva directa de la desigualdad media geométrica media aritmética*, Miscelánea Mat. **28**, (1999), 11-15.

[3] R. M Dimitric, *Using Less Calculus in Teaching Calculus: an Historical Approach*, Math. Magazine, MAA **74**, no. 3 (2001), 201–211.

[4] M. Feingold, *Newton, Leibniz, and Barrow too: An Attempt at a Reinterpretation*, Isis **84**, no. 2 (1993), 310–338 (Disponible en JSTOR).

[5] D. J. H. Garling, *Inequalities, a Journey into Linear Analysis*, Cambridge University Press, Cambridge, 1999.

[6] G. H. Hardy, J. E. Littlewood, G. Pólya, *Inequalities*, Cambridge University Press, Cambridge, 1999.

[7] H. Kuhn, *Ein elementary Beweis des Weierstrasschen Approximationssatzes*, Arch. Math. **15** (1964), 316–317.

[8] E. Maor, *e: The Story of a Number*, Princeton University Press, Princeton, 1994.

[9] D. S. Mitrinović, J. E. Pečarić, *Bernoulli's inequality*, Rend. Circ. Mat. Palermo, serie II, tomo XLII (1993), 317–337.

[10] H. N. Shi, *Generalizations of Bernoulli's inequality with applications*, J. Math. Inequalities, **2**, no. 1 (2008), 101–107.