Inequalities of Analysis

Andrejs Treibergs

University of Utah

Fall 2014

The URL for these Beamer Slides: *"Inequalities of Analysis"*

`http://www.math.utah.edu/~treiberg/InequalitiesSlides.pdf`

- G. Hardy, J. E. Littlewood & G. Pólya, *Inequalities,* 2nd ed., Cambridge university Press, Cambridge, 1991.
- G. Pólya & G. Szegö, *Problems and Theorems in Analysis I*, Springer Verlag, New York, 1972.

4. Outline.

- Triangle and Cauchy Schwarz Inequalities
- Arithmetic - Geometric - Harmonic Mean Inequality
  - Relations among the AGH means
  - Cauchy's proof
  - Applications: largest triangle of given perimeter and monotonicity of the compound interest sequence
- Jensen's Inequality
  - Convex functions and a proof for finitely many numbers
  - Probabilistic interpretation
  - Hölder's, Cauchy-Schwarz's and AG Inequalities follow from Jensen's
  - Application: largest polygons in circular arc
  - Another proof using support functions
  - Integral form of Jensen's, Hölder's and Minkowski's Inequalities
  - Application: least force exerted on magnetic pole at a point inside a loop wire of given length carrying a fixed current

For any two numbers $x, y \in \mathbf{R}$ we have the Triangle Inequality.
$|x + y| \leq |x| + |y|$.

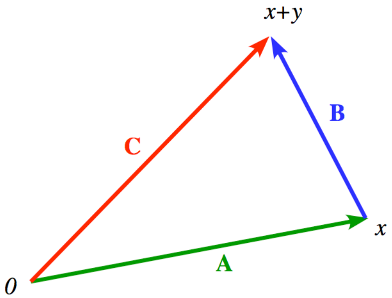


Figure 1: Euclidean Triangle.

The name comes from the fact that the sum of lengths of two sides of a triangle exceeds the length of the third side so the lengths satisfy

$$C \leq A + B.$$

If we have sides given as vectors x, y and $x + y$ then the lengths satisfy

$$|x + y| \leq |x| + |y|.$$

6. Triangle inequality for arbitrarily many terms and integrals.

For any finite set of numbers $x_1, x_2, \ldots, x_n \in \mathbf{R}$ we have

(Triangle Inequality) $\quad |x_1 + \cdots + x_n| \leq |x_1| + \cdots + |x_n|.$

(Equality Condition) $\quad$ "$=$" $\iff$ all x_i have the same sign.

This follows using induction on the inequality with two terms. By taking a limit the result also holds for infinite sums and integrals.
If $\{x_i\} \subset \mathbf{R}$ and the series is absolutely summable then

$$\left| \sum_{i=1}^{\infty} x_i \right| \leq \sum_{i=1}^{\infty} |\, x_i \,|$$

If f is integrable on the interval $[a, b]$ then $|f|$ is integrable and

$$\left| \int_a^b f(x)\, dx \right| \leq \int_a^b |\, f(x) \,|\, dx.$$

For any two vectors $A, B \in \mathbf{R}^n$, the Cauchy-Schwarz inequality amounts to the fact the the orthogonal projection of one vector A onto another B is shorter than the original vector: $|\operatorname{pr}_B(A)| \le |A|$.

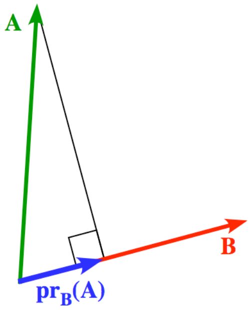


Figure 2: Euclidean Triangle.

Equality holds iff A and B are parallel. If we write vectors $A = (a_1, a_2, \ldots, a_n)$ and $B = (b_1, b_2, \ldots, b_n)$ then the Euclidean dot product is

$$A \bullet B = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n.$$

We can express the length of the projection using dot product

$$\begin{aligned} |\operatorname{pr}_B(A)| &= |A|\,|\cos\angle(AB)| \\ &= |A| \frac{|A \bullet B|}{|A|\,|B|} \\ &= \frac{|A \bullet B|}{|B|} \le |A|. \end{aligned}$$

8. Cauchy Schwarz Inequality for infinite sums and integrals.

Let $a_1, a_2, \ldots, a_n$ and $b_1, b_2, \ldots, b_n$ be arbitrary real numbers. Then the Cauchy Schwarz Inequality says

$$|A \bullet B| = \left| \sum_{i=1}^{n} a_i b_i \right| \leq \left(\sum_{i=1}^{n} a_i^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^{n} b_i^2 \right)^{\frac{1}{2}} = |A|\,|B|.$$

Equality holds iff there is $c \in \mathbf{R}$ such that $A = cB$ or $cA = B$.
By taking limits we also obtain for square-summable series

$$\left| \sum_{i=1}^{\infty} a_i b_i \right| \leq \left(\sum_{i=1}^{\infty} a_i^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^{\infty} b_i^2 \right)^{\frac{1}{2}}$$

and for square-integrable functions f, g on $[a, b]$,

$$\left| \int_a^b f(x) g(x)\, dx \right| \leq \left(\int_a^b f^2(x)\, dx \right)^{\frac{1}{2}} \left(\int_a^b g^2(x)\, dx \right)^{\frac{1}{2}}$$

We shall provide a proof of these later.

Let $a_1, a_2, \ldots, a_n$ be arbitrary real numbers. Then the Arithmetic Mean is the the expression

$$\mathfrak{A}(a) = \frac{a_1 + a_2 + \cdots + a_n}{n}.$$

If all the numbers are positive, we define the Geometric and Harmonic Means as

$$\mathfrak{G}(a) = \sqrt[n]{a_1 \cdots a_n}, \qquad \mathfrak{H}(a) = \frac{n}{\dfrac{1}{a_1} + \cdots + \dfrac{1}{a_n}}.$$

After the Triangle and Schwartz inequalities, the next best known is Arithmetic-Geometric Mean Inequality: for arbitrary positive numbers which are not all equal,

$$\mathfrak{h}(a) < \mathfrak{G}(a) < \mathfrak{A}(a) \tag{1}$$

The three means are all averages. If $m = \min\{a_1, \ldots, a_n\}$ and $M = \max\{a_1, \ldots, a_n\}$ are the smallest and largest of the numbers, then

$$m \leq \mathfrak{A}(a) \leq M, \qquad m \leq \mathfrak{G}(a) \leq M, \qquad m \leq \mathfrak{H}(a) \leq M.$$

(For $\mathfrak{G}(a)$ and $\mathfrak{H}(a)$ we assume $m > 0$.) These inequalities and $\mathfrak{H}(a) \leq \mathfrak{G}(a) \leq \mathfrak{A}(a)$ become equalities if and only if the a_i's are equal.

The following properties are obvious from the definitions

$$\frac{1}{\mathfrak{G}(a)} = \mathfrak{G}\left(\frac{1}{a}\right), \qquad \frac{1}{\mathfrak{H}(a)} = \mathfrak{A}\left(\frac{1}{a}\right),$$

$$\mathfrak{A}(a+b) = \mathfrak{A}(a) + \mathfrak{A}(b), \quad \mathfrak{G}(ab) = \mathfrak{G}(a)\,\mathfrak{G}(b), \quad \log \mathfrak{G}(a) = \mathfrak{A}(\log a)$$

where $\frac{1}{a}$ means $\left(\frac{1}{a_1}, \frac{1}{a_2}, \cdots\right)$ *etc.*

We only need to prove the AG Inequality because the HG inequality follows from the AG inequality and properties of the means

$$\mathfrak{H}(a) = \frac{1}{\mathfrak{A}\left(\frac{1}{a}\right)} \le \frac{1}{\mathfrak{G}\left(\frac{1}{a}\right)} = \mathfrak{G}(a).$$

For two positive numbers, the AG inequality follows from the positivity of the square

$$\mathfrak{G}^2 = ab = \left(\frac{a+b}{2}\right)^2 - \left(\frac{a-b}{2}\right)^2 \le \left(\frac{a+b}{2}\right)^2 = \mathfrak{A}^2$$

with strict inequality if $a \neq b$. This inequality says more: if equality holds ($\mathfrak{A} = \mathfrak{G}$), then

$$0 = \mathfrak{A}^2 - \mathfrak{G}^2 = \left(\frac{a-b}{2}\right)^2$$

so $a = b$.

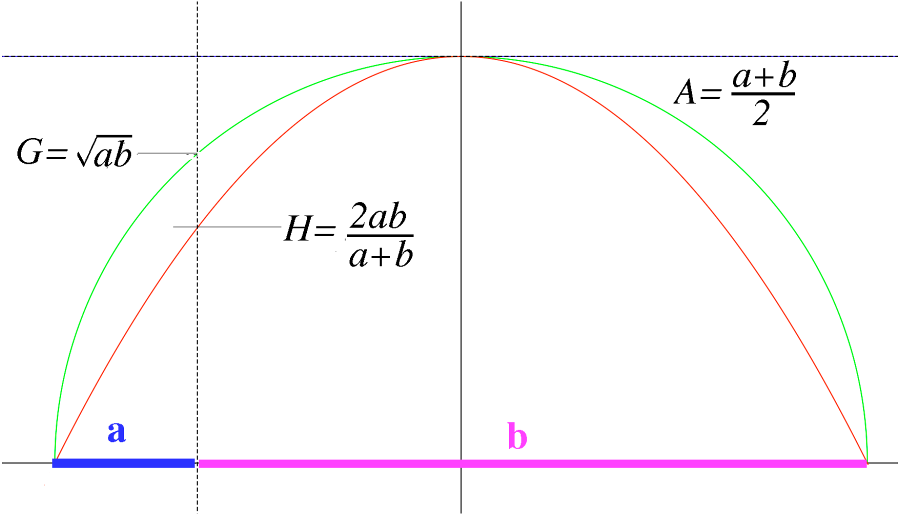


Figure 3: Semicircle / Parabola Construction for A / G / H Means

Cauchy's proof of 1897 first considers the case $n = 2^m$ is a power of two. By reordering if necessary, we may assume $a \neq b$. Then

$$abcd < \left(\frac{a+b}{2}\right)^2 \left(\frac{c+d}{2}\right)^2 \le \left(\frac{a+b+c+d}{4}\right)^4$$

$$abcdefgh < \left(\frac{a+b+c+d}{4}\right)^4 \left(\frac{e+f+g+h}{4}\right)^4$$

$$\le \left(\frac{a+b+c+d+e+f+g+h}{8}\right)^8$$

thus by induction,

$$a_1 a_2 \cdots a_{2^m} < \left(\frac{a_1 + a_2 + \cdots + a_{2^m}}{2^m}\right)^{2^m}$$

proving the inequality if $n = 2^m$. Note that "$\le$" holds without the assumption that $a \neq b$. Also, if $\mathfrak{A} = \mathfrak{G}$ then this argument also shows that all a_i must have been equal.

For $2^{m-1} < n < 2^m$ not a power of two he uses a padding trick: Put

$$K = \frac{a_1 + a_2 + \cdots + a_n}{n}.$$

Then using the inequality with 2^m terms (assuming $a_1 \neq a_2$),

$$a_1 \cdot a_2 \cdots a_n K^{2^m - n} < \left(\frac{a_1 + a_2 + \cdots + a_n + (2^m - n)K}{2^m} \right)^{2^m} = K^{2^m}$$

Thus

$$a_1 \cdot a_2 \cdots a_n < K^n = \left(\frac{a_1 + a_2 + \cdots + a_n}{n} \right)^n.$$

"$\leq$" holds without the assumption that $a_1 \neq a_2$. Also, if $\mathfrak{A} = \mathfrak{G}$ then again, all a_i must have been equal. $\square$

15. Geometric Application of Arithmetic Geometric Mean Inequality.

Theorem

The area of a triangle with given perimeter $2p = a + b + c$ is maximum if the sides a, b, c are equal.

Proof.

For a nondegenerate triangle, the sum of the lengths of any two sides is strictly greater than the third, thus $2p = a + b + c > 2c$ and so on. So

$$p - a, \qquad p - b, \qquad p - c$$

are all positive. By Heron's formula for area and the AG Inequality

$$A^2 = p(p-a)(p-b)(p-c) \le p\left(\frac{p-a+p-b+p-c}{3}\right)^3 = \frac{p^4}{27}$$

with equality iff $a = b = c$. $\square$

Theorem

If $0 < m < n$ are integers and the nonzero real number $\xi > -m$, then

$$\left(1+\frac{\xi}{m}\right)^m < \left(1+\frac{\xi}{n}\right)^n .$$

Proof.

Apply AG Inequality to m copies of $1+\dfrac{\xi}{m}$ and $n-m$ copies of 1.

$$\left(1+\frac{\xi}{m}\right)^{\frac{m}{n}} 1^{\frac{n-m}{n}} \le \frac{m}{n}\left(1+\frac{\xi}{m}\right) + \frac{n-m}{n}1 = 1+\frac{\xi}{n}$$

which is a strict inequality if $\xi \neq 0$. □

17. Arithmetic - Geometric Mean Inequality with Weights.

Suppose a_i occurs p_i times. Let $q = p_1 + \cdots + q_n$. Then the AG inequality is

$$\left(a_1^{p_1}\, a_2^{p_2} \cdots a_n^{p_n}\right)^{\frac{1}{q}} \leq \frac{p_1 a_1 + p_2 a_2 + \cdots p_n a_n}{q}$$

In other words for positive rational numbers $w_i = \dfrac{p_i}{q}$ such that $w_1 + \cdots + w_n = 1$ we have

$$a_1^{w_1} \cdots a_n^{w_n} \leq w_1 a_1 + \cdots + w_n a_n$$

For any real $0 < \theta_i$ such that $\theta_1 + \cdots + \theta_n = 1$, by approximating with rationals this gives the Arithmetic Geometric Inequliaty with Weights

$$a_1^{\theta_1} \cdots a_n^{\theta_n} \leq \theta_1 a_1 + \cdots + \theta_n a_n.$$

By a separate argument we can show that equality occurs iff $a_1 = a_2 = \cdots = a_n$.

Theorem (Hölder's Inequality for Sums)

Let $\alpha, \beta, , \ldots, \lambda$ be positive numbers such that $\alpha + \beta + \cdots + \lambda = 1$. Let a_j, b_j, ..., l_j be positive numbers for all $j = 1, \ldots, n$. Then

$$\sum_{j=1}^n {a_j}^\alpha {b_j}^\beta \cdots {l_j}^\lambda \leq \left(\sum_{j=1}^n a_j\right)^\alpha \left(\sum_{j=1}^n b_j\right)^\beta \cdots \left(\sum_{j=1}^n l_j\right)^\lambda$$

"$<$" holds unless all a, b,...,l are proportional.

The proof follows from the AG Inequality with weights. Indeed,

$$\begin{aligned}
\frac{\sum_{j=1}^n {a_j}^\alpha {b_j}^\beta \cdots {l_j}^\lambda}{\left(\sum_{j=1}^n a_j\right)^\alpha \left(\sum_{j=1}^n b_j\right)^\beta \cdots \left(\sum_{j=1}^n l_j\right)^\lambda} &= \sum_{j=1}^n \left(\frac{a_j}{\sum_{k=1}^n a_k}\right)^\alpha \left(\frac{b_j}{\sum_{k=1}^n b_k}\right)^\beta \cdots \left(\frac{l_j}{\sum_{k=1}^n l_k}\right)^\lambda \\
&\leq \sum_{j=1}^n \left(\frac{\alpha a_j}{\sum_{k=1}^n a_k} + \frac{\beta b_j}{\sum_{k=1}^n b_k} + \cdots + \frac{\lambda l_j}{\sum_{k=1}^n l_k}\right) \\
&= \alpha + \beta + \cdots + \lambda = 1. \quad \square
\end{aligned}$$

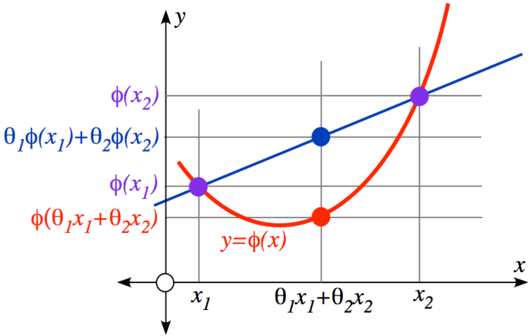


Figure 4: Convex Function $y = \varphi(x)$

Let I be an interval and $\varphi : I \to \mathbf{R}$ be a function.

We say that φ is convex in I if it satisfies

$$\varphi(\theta_1 x_1 + \theta_2 x_2) \leq \theta_1 \varphi(x_1) + \theta_2 \varphi(x_2)$$

for every $x_1 \neq x_2 \in I$ and every $0 < \theta_i$ such that $\theta_1 + \theta_2 = 1$. This condition says that all of the interior points of any chord of the curve $y = \varphi(x)$ lie above the curve.

Theorem

Let $\varphi : I \to \mathbf{R}$ be a convex function. Then φ is continuous at any $a \in I^\circ$.

Proof.

Let $b > 0$ so that $[a-b, a+b] \subset I$. Convexity tells us that for $0 < h < 1$, since $a - b < a - hb < a$, $a < a + hb < a + b$ and $a - hb < a < a + hb$,

$$\varphi(a - hb) \leq \varphi(a) + h[\varphi(a-b) - \varphi(a)] \tag{2}$$
$$\varphi(a + hb) \leq \varphi(a) + h[\varphi(a+b) - \varphi(a)] \tag{3}$$
$$\varphi(a) \leq \tfrac{1}{2}\varphi(a - hb) + \tfrac{1}{2}\varphi(a + hb) \tag{4}$$

Combining (2) and (3) with (4) yields

$$\varphi(a - hb) \geq \varphi(a) - h[\varphi(a+b) - \varphi(a)] \tag{5}$$
$$\varphi(a + hb) \geq \varphi(a) - h[\varphi(a-b) - \varphi(a)] \tag{6}$$

(2) and (5) say $\varphi(a - hb) \to \varphi(a)$ and (3) and (6) say $\varphi(a + hb) \to \varphi(a)$ as $h \to 0+$. Hence φ is continuous at a. $\square$

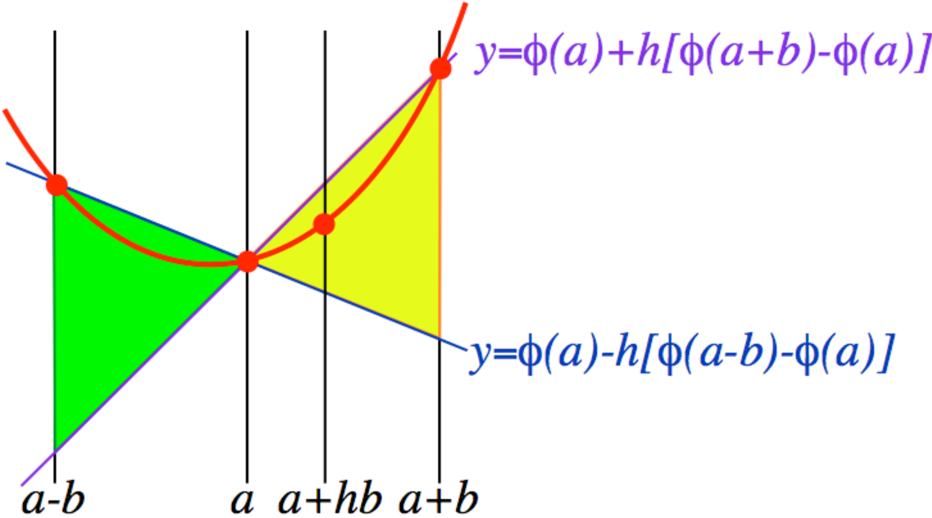


Figure 5: Convex Function Lies Between Secants so is Continuous at a.

Theorem (Jensen's Inequality)

Suppose $\varphi : I \to \mathbf{R}$ *is convex,* $x_1, \ldots, x_n \in I$ *arbitrary points and* $\theta_1, \ldots, \theta_n \in (0,1)$ *arbitrary weights. Then*

$$\varphi(\theta_1 x_1 + \cdots + \theta_n x_n) \le \theta_1 \varphi(x_1) + \cdots + \theta_n \varphi(x_n). \tag{7}$$

Proof.

Argue by induction. The base case $n = 2$ is the definition of convexity. Assume that (7) is true for n. Then using $n = 2$ case and the induction hypothesis we have $\varphi(\theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_{n+1} x_{n+1}) =$

$$\begin{aligned}
&= \varphi\left(\theta_1 x_1 + (1-\theta_1)\left[\tfrac{\theta_2}{(1-\theta_1)} x_2 + \cdots + \tfrac{\theta_{n+1}}{(1-\theta_1)} x_{n+1}\right]\right) \\
&\le \theta_1 \varphi(x_1) + (1-\theta_1)\varphi\left(\tfrac{\theta_2}{(1-\theta_1)} x_2 + \cdots + \tfrac{\theta_{n+1}}{(1-\theta_1)} x_{n+1}\right) \\
&\le \theta_1 \varphi(x_1) + (1-\theta_1)\left[\tfrac{\theta_2}{(1-\theta_1)} \varphi(x_2) + \cdots + \tfrac{\theta_{n+1}}{(1-\theta_1)} \varphi(x_{n+1})\right] \\
&= \theta_1 \varphi(x_1) + \cdots + \theta_{n+1} \varphi(x_{n+1}).
\end{aligned}$$

Theorem (Equality in Jensen's Inequality)

Suppose $\varphi : I \to \mathbf{R}$ is convex, $x_1, \ldots, x_n \in I$ arbitrary points and $\theta_1, \ldots, \theta_n \in (0,1)$ arbitrary weights. If equality holds

$$\varphi(\theta_1 x_1 + \cdots + \theta_n x_n) = \theta_1 \varphi(x_1) + \cdots + \theta_n \varphi(x_n),$$

then either

- *all x_i are equal or*
- *$\varphi(x)$ is a linear function in an interval including all of the x_i's.*

$\varphi : I \to \mathbf{R}$ is strictly convex if $\varphi(\theta_1 x_1 + \theta_2 x_2) < \theta_1 \varphi(x_1) + \theta_2 \varphi(x_2)$ for every $x_1 \neq x_2 \in I$ and every $0 < \theta_i$ such that $\theta_1 + \theta_2 = 1$. If φ is strictly convex, then it cannot equal a linear function on any nontrivial interval.

If $\varphi \in \mathcal{C}^2((\alpha, \beta))$, then φ is convex iff $\varphi'' \geq 0$ in (α, β). If $\varphi'' > 0$ in (α, β) then φ is strictly convex.

24. Probabilistic Interpretation of Jensen's Inequality.

Suppose X is a discrete random variable that takes values in the set

$$X \in \mathcal{D} = \{x_1, x_2, \ldots, x_n\}$$

with probabilities $\mathcal{P}(X = x_i) = \theta_i$, where $\theta_i > 0$ with $\sum_i \theta_i = 1$. Then the expectation of X is

$$\mathcal{E}(X) = \sum_{i=1}^{n} x_i \theta_i$$

The expectation of a function $\varphi : \mathcal{D} \to \mathbf{R}$ is

$$\mathcal{E}(\varphi(X)) = \sum_{i=1}^{n} \varphi(x_i) \theta_i$$

Theorem (Jensen's Inequality)

Suppose $\varphi : I \to \mathbf{R}$ is convex and X a random variable taking values in $\mathcal{D} = \{x_1, \ldots, x_n\} \subset I$ with probabilities $\mathcal{P}(X = x_i) = \theta_i$ then

$$\varphi(\mathcal{E}(X)) \le \mathcal{E}(\varphi(X)).$$

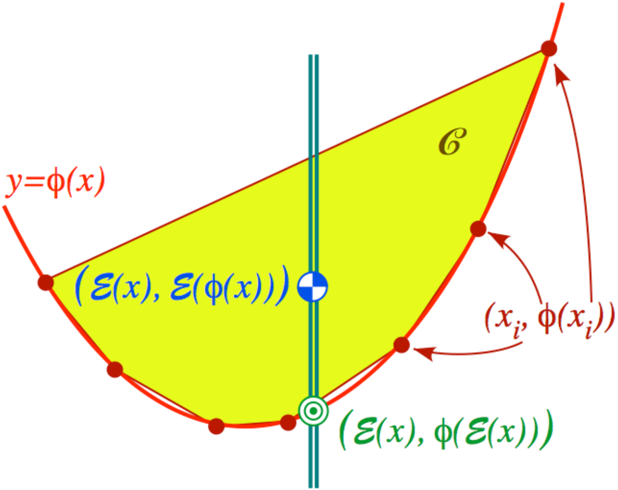


Figure 6: Jensen's Inequality $\varphi(\mathcal{E}(x)) \le \mathcal{E}(\varphi(x))$

The center of mass of all boundary points

$$\mathcal{E}(x, \varphi(x)) = (\mathcal{E}(x), \mathcal{E}(\varphi(x)))$$

is in the convex hull $\mathcal{C}$ of the points on the curve $(x_i, \varphi(x_i))$, which is above the mean point on the graph

$$(\mathcal{E}(x), \varphi(\mathcal{E}(x))).$$

Let $\varphi(t) = t^k$ for $k > 1$ on $I = (0, \infty)$. Then $\varphi''(t) = k(k-1)t^{k-2} > 0$ so φ is convex. Applying Jensen's Inequality yields for positive x_i and p_i so weights are $\theta_i = p_i/(\sum p_j)$,

$$\left(\frac{\sum x_i p_i}{\sum p_j} \right)^k \le \frac{\sum x_i^k p_i}{\sum p_j}.$$

with equality if and only if all x_i's are equal.

This can be given a more symmetric form. If a_i and b_i are any nonzero numbers, then by putting $x\theta = |a|\,|b|$ and $x^k\theta = |a|^k$, (or $x = |a|\,|b|^{1/(1-k)}$ and $\theta = |b|^{k/(k-1)}$) then

$$\sum |a_i|\,|b_i| \le \left(\sum |a_i|_i^k \right)^{\frac{1}{k}} \left(\sum |b_i|^{\frac{k}{k-1}} \right)^{\frac{k-1}{k}} .$$

It is convenient to call $k' = \frac{k}{k-1}$ the conjugate exponent. It satisfies

$$\frac{1}{k} + \frac{1}{k'} = 1.$$

Because $|\sum a_i b_i| \le \sum |a_i|\,|b_i|$ by the triangle inequality, the previous inequality, called Hölder's Inequality, becomes

$$\left|\sum a_i b_i\right| \le \left(\sum |a_i|^k\right)^{\frac{1}{k}} \left(\sum |b_i|^{k'}\right)^{\frac{1}{k'}}.$$

If equality holds then all $a_i b_i$ are positive or all negative for the equality to hold in the triangle inequality so all a_i and b_i have the same sign or opposite signs. Also there is a constant c so that $c = x_i$ all i for Jensen's Inequality to hold, hence $|a_i|^k = c^k |b_i|^{k'}$ for all i (*i.e.*, are proportional.)

The $k = k' = 2$ case is called the Cauchy-Schwarz Inequality.

$$\left(\sum a_i b_i\right)^2 \le \left(\sum a_i^2\right)\left(\sum b_i^2\right).$$

28. Jensen's Inequality Implies the AG Inequality.

Let $\varphi(t) = -\log t$ on $I = (0, \infty)$. Then $\varphi''(t) = t^{-2} > 0$ so φ is strictly convex. Applying Jensen's Inequality yields for positive x_i and weights θ_i such that $\sum \theta_i = 1$,

$$\sum \log(x_i)\,\theta_i \le \log\left(\sum x_i\theta_i\right)$$

with equality if and only if all x_i's are equal. Taking exponential

$$\mathfrak{G}(x) = \prod x_i^{\theta_i} = \exp\left[\sum \log\left(x_i^{\theta_i}\right)\right] \le \sum x_i\theta_i = \mathfrak{A}(x)$$

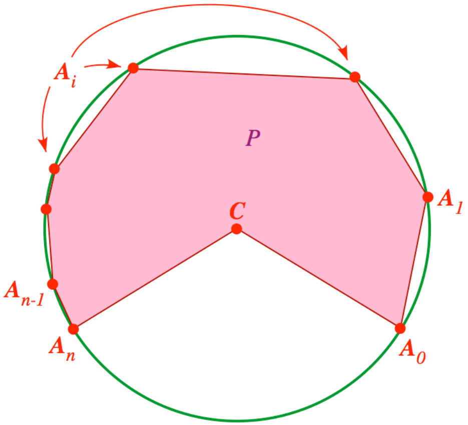


Figure 7: Polygon Whose Vertices are Center C and Points A_i on a Circle

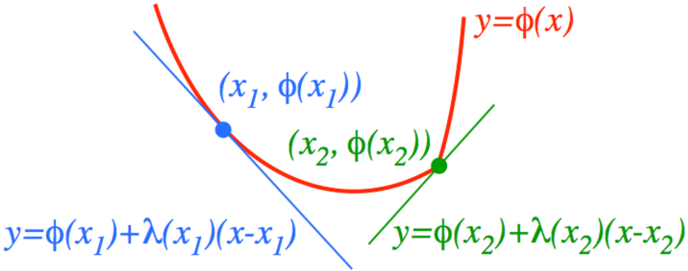
Suppose C is the center of a circle and $A_0, A_1, \ldots, A_n, C$ be a polygon, whose vertices except for C lie in order on a circle. C, A_0 and A_n are fixed and $A_1, \ldots, A_{n-1}$ vary. Show: the area and perimeter of the polygon are greatest when sides are equal $A_0A_1 = A_1A_2 = \cdots = A_{n-1}A_n$

Let α_i denote the angle $A_{i-1}CA_i$. Note $0 < \alpha_i < \pi$ since the polygon surrounds C. Let r be the radius and $\alpha = \sum \alpha_i$ the total angle. The the area of the sector is $\frac{1}{2}r^2 \sin \alpha_i$ and the distance $A_{i-1}A_i$ is $2r\sin(\alpha_i/2)$. $\varphi(t) = -\sin t$ and $\psi(t) = -2\sin(t/2)$ have $\varphi'' = \sin t > 0$ and $\psi'' = \frac{1}{2}\sin(t/2) > 0$, resp., so φ and ψ are convex on $(0,\pi)$.

By Jensen's Inequality and reversing signs, both inequalities hold

$$A = \frac{r^2}{2n}\sum \sin \alpha_i \le \frac{r^2}{2}\sin\left(\frac{\alpha}{n}\right)$$
$$L = \frac{2r}{n}\sum \sin\left(\frac{\alpha_i}{2}\right) \le 2r\sin\left(\frac{\alpha}{2n}\right)$$

with equalities if and only if all α_i are equal.



Let $\varphi : I \to \mathbf{R}$. A support function for φ at $x_1 \in I$ ia a linear function through the point with slope $\lambda(x_1)$ which is is below φ :

$$\varphi(x) \geq \varphi(x_1) + \lambda(x_1)(x - x_1) \qquad \text{for all } x \in I$$

We may take $\lambda(x_1) = \varphi'(x_1)$ if φ is differentiable at x_1.

Theorem

The continuous function $\varphi : I \to \mathbf{R}$ is convex if and only if there is a support function for φ at every point $x_1 \in I$.

Suppose $\varphi : I \to \mathbf{R}$ is convex, $x_1, \ldots, x_n \in I$ arbitrary points and $\theta_1, \ldots, \theta_n \in (0,1)$ arbitrary weights. The arithmetic mean

$$\bar{x} = \sum x_i \theta_i$$

is a point of I, thus the support function satisfies for all $x \in I$,

$$\varphi(\bar{x}) + \lambda(\bar{x})(x - \bar{x}) \le \varphi(x).$$

Putting $x = x_i$, multiplying by θ_i and summing gives

$$\varphi(\mathcal{E}(X)) + 0 = \sum [\varphi(\bar{x})\theta_i + \lambda(\bar{x})(x_i - \bar{x})\theta_i)] \le \sum \varphi(x_i)\theta_i = \mathcal{E}(\varphi(X))$$

33. Integral Form of Jensen's Inequality.

Let $f(x)$ be a probability density function for a random variable X that takes values in I. So f is integrable, $f(x) \geq 0$ and $\int_I f(x)\,dx = 1$. The expectation of X is

$$\bar{x} = \mathcal{E}(X) = \int_I x\, f(x)\, dx.$$

If $\varphi : I \to \mathbf{R}$ is convex, then as before, for all $x \in I$,

$$\varphi(\bar{x}) + \lambda(\bar{x})(x - \bar{x}) \leq \varphi(x).$$

Integrating gives the Integral form of Jensen's Inequality

$$\varphi(\mathcal{E}(X)) = \int_I [\varphi(\bar{x}) + \lambda(\bar{x})(x - \bar{x})]\, f(x)\, dx \leq \int_I \varphi(x_i) f(x)\, dx = \mathcal{E}(\varphi(X))$$

Equality holds for strictly convex φ iff $f(x) = C$ is (essentially) constant.

This may also be achieved by approximating the Riemann Integral by finite sums, using the finite Jensen's inequality and passing to the limit.

In the Riemann Theory of integration, a bounded function $f : I \to \mathbf{R}$ on a bounded interval is called null if $f = 0$ at all points of continuity of f. If $f - g$ is null, we say f and g are equivalent and write $f \equiv g$. So $f \geq 0$ and $\int f\,dx = 0$ is a necessary and sufficient condition for $f \equiv 0$.

Put $\varphi(t) = -\log t$ in Jensen's Inequality and argue as before.

Theorem (Hölder's Theorem for Integrals)

Let $k > 1$ and $\frac{1}{k} + \frac{1}{k'} = 1$. Let $f^k, g^{k'} : I \to \mathbf{R}$ be integrable. Then fg is integrable and

$$\int fg\,dx \leq \left(\int f^k\,dx\right)^{\frac{1}{k}} \left(\int g^{k'}\,dx\right)^{\frac{1}{k'}}$$

Equality holds if and only if there are constants A, B, not both zero, such that $Af^k \equiv Bg^{k'}$.

If one of the functions is equivalent to zero, then equality condition holds.

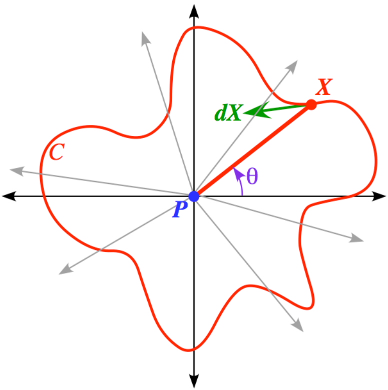


Figure 8: A Closed Star-shaped Plane Curve C and Pole P

C carries an electric current j_0 and induces a magnetic field B at P.

The Biot Savart law gives the force exerted on magnetic pole of strength μ_0 at P in the plane of and interior to C. Fix the area A enclosed by C. Show: B is minimum when C is a circle and P is at its center. Letting $X(\theta)$ be the position vector and r the distance from P to X, in polar coordinates $X - P = r(\cos\theta, \sin\theta)$. If $j_0 dX$ is the current density, then the infinitesimal force at P is

$$
\begin{aligned}
dB &= \frac{j_0\mu_0}{4\pi r^3}(X-P)\times dX \\
&= \frac{j_0\mu_0}{4\pi r^2}(\cos\theta,\sin\theta) \\
&\quad \times \left[r'(\cos\theta,\sin\theta) + r(-\sin\theta,\cos\theta)\right]\, d\theta \\
&= \frac{j_0\mu_0}{4\pi r}\, d\theta
\end{aligned}
$$

Integrating $0 \leq \theta < 2\pi$ in polar coordinates, using Hölder's Inequality,

$$\begin{aligned}
2\pi &= \int d\theta \\
&= \int \left(\frac{1}{r}\right)^{\frac{2}{3}} \left(r^2\right)^{\frac{1}{3}}\, d\theta \\
&\leq \left(\int \frac{d\theta}{r}\right)^{\frac{2}{3}} \left(\int r^2\, d\theta\right)^{\frac{1}{3}} \\
&= \left(\frac{4\pi B}{j_0 \mu_0}\right)^{\frac{2}{3}} (2A)^{\frac{1}{3}}
\end{aligned}$$

with equality if and only if r is constant. $\square$

Theorem (Minkowski's Inequality for Integrals)

Let $g^k, h^k : I \to \mathbf{R}$ be integrable, $f : I \to \mathbf{R}$ a probability density and $k > 1$. Then

$$\left[\int |g+h|^k f\,dx\right]^{\frac{1}{k}} \le \left[\int |g|^k f\,dx\right]^{\frac{1}{k}} + \left[\int |h|^k f\,dx\right]^{\frac{1}{k}}$$

Equality holds if and only if there are constants A, B, not both zero, such that $Ag \equiv Bh$.

WLOG $f, g \ge 0$ and let $s = g + h$. Using $k' = \frac{k}{k-1}$, and Hölder's Ineq.

$$\begin{aligned}
\int s^k f\,dx &= \int g s^{k-1} f\,dx + \int h s^{k-1} f\,dx \\
&= \int \left[g f^{\frac{1}{k}}\right]\left[s f^{\frac{1}{k}}\right]^{k-1} dx + \int \left[h f^{\frac{1}{k}}\right]\left[s f^{\frac{1}{k}}\right]^{k-1} ds \\
&\le \left[\int g^k f\right]^{\frac{1}{k}}\left[\int s^k f\,dx\right]^{\frac{1}{k'}} + \left[\int h^k f\,dx\right]^{\frac{1}{k}}\left[\int s^k f\,dx\right]^{\frac{1}{k'}}. \quad \square
\end{aligned}$$

A metric space (X, d) is a set X and a distance function $d : X \times X \to [0, \infty)$ such that

- d is symmetric: $d(x, y) = d(y, x)$ for all $x, y \in X$.
- d is positive definite: $d(x, y) \geq 0$ for all $x, y \in X$ and $d(x, y) = 0$ if and only if $x = y$.
- d satisfies the Triangle Inequality: $d(x, z) \leq d(x, y) + d(y, z)$ for all $x, y, z \in X$.

If $X = \mathbf{R}^n$, the standard metric space struture is given by the Euclidean distance function

$$d_E(x, y) = \left[\sum_{k=1}^{n} (x_k - y_k)^2\right]^{\frac{1}{2}}.$$

Minkowski's Inequality says that there are infinitely many distance functions possible for $X = \mathbf{R}^n$, namely for $p > 1$ consider

$$d_p(x, y) = \left[\sum_{k=1}^{n} (x_k - y_k)^p \right]^{\frac{1}{p}}.$$

Thus $d_E(x, y) = d_2(x, y)$. Evidently $d_p(x, y)$ is symmetric and positive definite. Minkowski's Inequality for sums says that $d_p(x, y)$ satisfies the triangle inequality:

$$d_p(x, z) \leq d_p(x, y) + d_p(y, z) \qquad \text{for all } x, y, z \in \mathbf{R}^n.$$

Theorem

Let $\theta_1, \ldots, \theta_n$ be positive weights such that $\sum_{k=1}^n \theta_k = 1$. Let $a_1, \ldots, a_n$ be positive. Define the weighted and unweighted means for $r > 0$

$$\mathfrak{M}_{\mathrm{r}}(a) = \left(\sum_{k=1}^n \theta_k a_k^r\right)^{\frac{1}{r}}, \qquad \mathfrak{S}_{\mathrm{r}}(a) = \left(\sum_{k=1}^n a_k^r\right)^{\frac{1}{r}}$$

Then if $0 < r < s$,
(1) $\mathfrak{M}_{\mathrm{r}}(a) < \mathfrak{M}_{\mathrm{s}}(a)$ unless all a_k are equal and
(2) $\mathfrak{S}_{\mathrm{s}}(a) > \mathfrak{S}_{\mathrm{r}}(a)$ unless $n = 1$.

To see (1), let $r = s\alpha$ where $0 < \alpha < 1$. Let $u_k = \theta_k a_k^s$ and $v_k = \theta_k$. Thus $\theta_k a_k^{s\alpha} = (\theta_k a_k^s)^\alpha (\theta_k)^{1-\alpha} = u_k^\alpha v_k^{1-\alpha}$. By Hölder's Inequality,

$$\mathfrak{M}_{\mathrm{r}}(a) = \sum_{k=1}^n u_k^\alpha v_k^{1-\alpha} \le \left(\sum_{k=1}^n u_k\right)^\alpha \left(\sum_{k=1}^n v_k\right)^{1-\alpha} = (\mathfrak{M}_{\mathrm{s}}(a))^\alpha \cdot 1.$$

To see (2) we'll exploit the homogeneity of the inequality to simplify the proof.

Note that $\mathfrak{S}_{\mathrm{s}}(a)$ is homogeneous in a, namely for $c \geq 0$, $\mathfrak{S}_{\mathrm{s}}(ca) = c\,\mathfrak{S}_{\mathrm{s}}(a)$. Let $c = \mathfrak{S}_{\mathrm{r}}(a)$, and define $\tilde{a}_k = a_k/c$. Then

$$\mathfrak{S}_{\mathrm{r}}(\tilde{a}) = \left(\sum_{k=1}^{n} \tilde{a}_k^r\right)^{\frac{1}{r}} = 1.$$

It follows if $n > 1$ that $\tilde{a}_k < 1$ for every k, which means that $\tilde{a}_k^s < \tilde{a}_k^r$. Hence

$$\sum_{k=1}^{n} \tilde{a}_k^s < \sum_{k=1}^{n} \tilde{a}_k^r = 1.$$

Finally

$$\mathfrak{S}_{\mathrm{s}}(a) = c\left(\sum_{k=1}^{n} \tilde{a}_k^s\right)^{\frac{1}{s}} < c \cdot 1 = \mathfrak{S}_{\mathrm{r}}(a). \qquad \square$$

Thanks!