(*October 20, 2019*)

02. Basic inequalities

Paul Garrett garrett@math.umn.edu http://www.math.umn.edu/˜garrett/

[This document is
http://www.math.umn.edu/˜garrett/m/real/notes_2019-20/02_basic_inequalities.pdf]

Although many of the inequalities here can be stated in much more general terms after the basics about *measure* and *integration* are developed, the mechanisms for these inequalities do not depend on any integration theory beyond Riemann's.

Therefore, for clarity, we state both *integral* forms of the inequalities, as well as *discrete* forms, although these seemingly disparate cases will be unified under the umbrella of abstract integration.

---

1. Cauchy-Schwarz-Bunyakowsky inequality

One more time, we recall:

[1.1] Claim: *(Cauchy-Schwarz-Bunyakowsky inequality)* For x, y an inner product space V,

$$|\langle x,y\rangle| \;\le\; |x|\cdot|y|$$

Assuming that neither x nor y is 0, *strict* inequality holds *unless* x and y are scalar multiples of each other.

Proof: For clarity, we first prove this for a *real* vector space V, with *real-valued* inner product. If $|y| = 0$, the assertions are trivially true. Thus, take $y \neq 0$. With real t, consider the quadratic polynomial function

$$f(t) \;=\; |x-ty|^2 \;=\; |x|^2 - 2t\langle x,y\rangle + t^2|y|^2$$

Certainly $f(t) \ge 0$ for all $t \in \mathbb{R}$, since $|x - ty| \ge 0$. Its minimum occurs where $f'(t) = 0$, namely, where $-2\langle x,y\rangle + 2t|y|^2 = 0$. This is where $t = \langle x,y\rangle/|y|^2$. Thus,

$$0 \;\le\; (\text{minimum}) \;\le\; f(\langle x,y\rangle/|y|^2) \;=\; |x|^2 - 2\frac{\langle x,y\rangle}{|y|^2}\langle x,y\rangle + \Big(\frac{\langle x,y\rangle}{|y|^2}\Big)^2\cdot|y|^2 \;=\; |x|^2 - \Big(\frac{\langle x,y\rangle}{|y|^2}\Big)^2\cdot|y|^2$$

Multiplying out by $|y|^2$,

$$0 \;\le\; |x|^2\cdot|y|^2 - \langle x,y\rangle^2$$

which gives the inequality. Further, for the inequality to be an *equality*, it must be that $|x - ty| = 0$, so x is a multiple of y.

Using complex scalars and hermitian complex-valued inner product introduces a minor technical complication, which is not so interesting. ///

2. *Young's inequality*

[2.1] Claim: *(Numerical Young's inequality)*

$$ab \;\le\; \frac{a^p}{p} + \frac{b^q}{q} \qquad\qquad (\text{for } a,b>0 \text{ and } \frac{1}{p}+\frac{1}{q}=1,\ p,q>0)$$

[2.2] Remark: The $p = q = 2$ case has an even simpler proof:

$$0 \;\le\; (a-b)^2 \;=\; a^2 - 2ab + b^2$$

rearranges to

$$ab \;\le\; \frac{a^2}{2} + \frac{b^2}{2} \qquad\qquad (\text{for } a,b \ge 0)$$

Proof: The convexity/concavity property of logarithm is that its graph lies *above* the line segment connecting two points on the graph:

$$t\cdot\log x + (1-t)\cdot\log y \;\le\; \log(t\cdot x + (1-t)\cdot y) \qquad\qquad (\text{for } x,y>0 \text{ and } 0\le t\le 1)$$

Thus,

$$\frac{1}{p}\log a^p + \frac{1}{q}\log b^q \;\;\le\;\; \log\big(\frac{a^p}{p} + \frac{b^q}{q}\big)$$

That is,

$$\log a + \log b \;\;\ge\;\; \log\big(\frac{a^p}{p} + \frac{b^q}{q}\big)$$

Exponentiating gives the inequality. ///

We mention the original papers [Young 1912a] and [Young 1912b], since the result has been so well assimilated that already the extensive bibliography of [Riesz-Nagy 1952] did not list these papers, although others of Young's did appear there.

3. *Convexity and Jensen's inequality*

A function f on an interval $(a,b)\subset\mathbb{R}$ is *concave upward* or *convex downward* when its graph bends upward, in the sense that a line segment connecting two points on the graph lies *above* the graph. That is, when

$$tf(x) + (1-t)f(y) \;\ge\; f(tx + (1-t)y) \qquad\qquad (\text{for } 0\le t\le 1 \text{ and } a<x<y<b)$$

The prototype is the exponential function $x \to e^x$.

[3.1] Remark: It is sufficiently easy to muddle concave/convex up/down that in practice it is best to say explicitly whether the chord (line segment connecting two points on the graph) lies *above*, or *below* the graph.

[3.2] Claim: Convex (up or down) $\mathbb{R}$-valued functions on an open interval (a,b) (allowing $a=-\infty$ and/or $b=+\infty$) are *continuous.*

Proof: Let g be continuous on (a,b) and take $x\in(a,b)$. Fix any s,t such that $a<s<x<t<b$. For y in the range $x<y<t$, the point $(y,g(y))$ is on or above the line through $(s,g(s))$ and $(x,g(x))$, and is below

the line through $(x, g(x))$ and $(t, g(t))$, so $g(y) \to g(x)$ as $y \to x^+$. For $s < y < x$, the same argument gives *left*-continuity. ///

Before enlarging our notion of integrals, it is already very useful to see a preliminary version of *Jensen's inequality*:

[3.3] Theorem: *(Jensen)* Let g be an $\mathbb{R}$-valued function on $[0,1]$ with $a < g(x) < b$, where a, b can also be $-\infty$ and $+\infty$. For *convex* f on (a,b),

$$f\Big(\int_0^1 g\Big) \;\le\; \int_0^1 f \circ g$$

Proof: First, $a < g(x) < b$ gives $a < \int_0^1 g < b$. The convexity condition on f can be written as the condition that slopes of secants increase from left to right. Thus, for example,

$$\frac{f(y) - f(x)}{y - x} \;\le\; \frac{f(z) - f(y)}{z - y} \qquad\qquad (\text{for } x < y < z \text{ inside } (a,b))$$

Applying this with $y = \int_0^1 g$,

$$\frac{f(\int g) - f(x)}{\int g - x} \;\le\; \frac{f(z) - f(\int g)}{z - \int g} \qquad\qquad (\text{for all } a < x < \textstyle\int_0^1 g \text{ and for all } \int_0^1 g < z < b)$$

With

$$S \;=\; \sup_{x : a < x < \int_0^1 g} \frac{f(\int_0^1 g) - f(x)}{\int_0^1 g - x}$$

we have

$$\frac{f(\int_0^1 g) - f(x)}{\int_0^1 g - x} \;\le\; S \;\le\; \frac{f(z) - f(\int_0^1 g)}{z - \int_0^1 g} \qquad\qquad (\text{for all } a < x < \textstyle\int_0^1 g \text{ and for all } \int_0^1 g < z < b)$$

From the left half of the latter inequality,

$$f(x) \;\ge\; f(\int_0^1 g) + S \cdot (x - \int_0^1 g) \qquad\qquad (\text{for } a < x < \textstyle\int_0^1 g)$$

and from the right half

$$f(z) \;\ge\; f(\int_0^1 g) + S \cdot (z - \int_0^1 g) \qquad\qquad (\text{for } \textstyle\int_0^1 g < z < b)$$

Thus,

$$f(w) \;\ge\; f(\int_0^1 g) + S \cdot (w - \int_0^1 g) \qquad\qquad (\text{for } \textit{all } w \text{ in the range } a < w < b)$$

In particular, letting $w = g(x)$ now with $x \in [0,1]$,

$$f(g(x)) \;\ge\; f(\int_0^1 g) + S \cdot (g(x) - \int_0^1 g) \qquad\qquad (\text{for } \textit{all } w \text{ in the range } a < w < b)$$

Integrating in $x \in [0,1]$,

$$\int_0^1 f \circ g \;\ge\; f(\int_0^1 g) + S \cdot (\int_0^1 g - \int_0^1 g) \;=\; f(\int_0^1 g) + S \cdot 0$$

as claimed. ///

[3.4] **Remark:** The proof only needed the fact that integration preserves inequalities. Also, no use was made of the assumption that g was defined on $[0,1]$, only that $\int_0^1 1\,dx = 1$. Thus, the same argument has a *discrete* analogue:

[3.5] **Theorem:** *(Jensen)* Let $X = \{1,2,\ldots\}$, with weights $0 \le w_n$ such that $\sum_n w_n = 1$. Let g be an $\mathbb{R}$-valued function on X with $a < g(x) < b$, where a,b can also be $-\infty$ and $+\infty$. For *convex* f on (a,b),

$$f\Big(\sum_n w_n \cdot g(n)\Big) \;\le\; \sum_n w_n \cdot (f\circ g)(n)$$

For example, in the proof of Hölder's inequality below, we use g defined on a set with just two points, assigned weights *(measures)* $\frac{1}{p}$ and $\frac{1}{q}$ with $\frac{1}{p}+\frac{1}{q} = 1$. In that case the statement of Jensen's inequality becomes

[3.6] **Theorem:** *(Jensen)* Let g be an $\mathbb{R}$-valued function on the two-point set $\{0,1\}$ with $a < g(x) < b$, where a,b can also be $-\infty$ and $+\infty$. Let $1 < p,q < +\infty$ be such that $\frac{1}{p}+\frac{1}{q} = 1$. For *convex* f on (a,b),

$$f\Big(\frac{g(0)}{p} + \frac{g(1)}{q}\Big) \;\le\; \frac{f\circ g(0)}{p} + \frac{f\circ g(1)}{q}$$

---

4. *Arithmetic-geometric mean inequality*

[4.1] **Corollary:** *(Arithmetic-geometric mean inequality)* For positive real numbers $a_1,\ldots,a_n$,

$$(a_1 a_2 \ldots a_n)^{1/n} \;\le\; \frac{a_1 + a_2 + \ldots + a_n}{n}$$

Proof: In Jensen's inequality, take $f(x) = e^x$, take X a finite set with n (distinct) elements $\{x_1,\ldots,x_n\}$, with each point having measure $1/n$, and $g(x_i) = \log a_i$. Jensen's inequality gives

$$n \exp\Big(\frac{\log a_1 + \ldots + \log a_n}{n}\Big) \;\le\; \frac{e^{\log a_1} + \ldots + e^{\log a_n}}{n}$$

which gives the assertion. ///

---

5. *Hölder inequality*

Conjugate exponents are numbers $p,q > 1$ such that

$$\frac{1}{p} + \frac{1}{q} \;=\; 1$$

For example, p and $q = \frac{p}{p-1}$ are conjugate exponents.

[5.1] **Remark:** In the following, we need to assume that the integrals of the functions f,g exist and have the expected basic properties. But we do not need an explicit description of the class of functions in which they lie, nor an explicit description of the integration.

For the following theorem, let X be $[a,b]$ or $\mathbb{R}$ or $\mathbb{R}^n$, or any other space on which we know how to integrate functions. Generalizing the Cauchy-Schwarz-Bunyakowsky inequality, we have

[5.2] **Theorem:** *(Hölder)* For conjugate exponents p,q and $[0,+\infty]$-valued functions f,g on a set X

$$\int_X f\cdot g \;\le\; \Big(\int_X f^p\Big)^{\frac{1}{p}}\cdot\Big(\int_X g^q\Big)^{\frac{1}{q}}$$

Proof: The assertion is trivial if either integral on the right-hand side is $+\infty$ or 0, so suppose the two quantities

$$I \;=\; \Big(\int_X f^p\Big)^{\frac{1}{p}} \qquad\qquad J \;=\; \Big(\int_X g^q\Big)^{\frac{1}{q}}$$

are finite and non-zero. Renormalize by taking $\varphi = f/I$ and $\psi = g/J$, so that $\int \varphi^p = 1 = \int \psi^q$. For $x\in X$ with $0<\varphi(x)<\infty$ and $0<\psi(x)<\infty$, there are real numbers u,v such that $e^{u/p}=\varphi(x)$ and $e^{v/q}=\psi(x)$. Invoking Jensen's inequality on a function defined on a set with just two points, with weights (*measures*) $\frac{1}{p}$ and $\frac{1}{q}$, using the convexity of the exponential function,

$$\varphi(x)\psi(x) \;=\; e^{\frac{u}{p}+\frac{v}{q}} \;\le\; \frac{e^u}{p}+\frac{e^v}{q} \;=\; \frac{\varphi(x)^p}{p}+\frac{\psi(x)^q}{q}$$

Integrating,

$$\int_X \varphi\cdot\psi \;\le\; \int_X \frac{\varphi(x)^p}{p}+\frac{\psi(x)^q}{q} \;=\; \frac{1}{p}+\frac{1}{q} \;=\; 1$$

From the renormalization, we are done. ///

Again, the argument uses no specifics about the integration process, and applies as well to discrete sums:

[5.3] **Theorem:** Let $X=\{1,2,\ldots\}$, with weights $0\le w_n$. For conjugate exponents p,q and $[0,+\infty)$-valued functions f,g on X,

$$\sum_n w_n\cdot f(n)\cdot g(n) \;\le\; \Big(\sum_n w_n\cdot f(n)^p\Big)^{\frac{1}{p}}\cdot\Big(\sum_n w_n\cdot g(n)^q\Big)^{\frac{1}{q}}$$

The case that all weights are $w_n=1$, namely,

[5.4] **Theorem:** Let $X=\{1,2,\ldots\}$. For conjugate exponents p,q and $[0,+\infty)$-valued functions f,g on X,

$$\sum_n f(n)\cdot g(n) \;\le\; \Big(\sum_n f(n)^p\Big)^{\frac{1}{p}}\cdot\Big(\sum_n g(n)^q\Big)^{\frac{1}{q}}$$

is already useful in showing that the alleged metric on spaces ℓ^p satisfies the triangle inequality, via a discrete form of *Minkowski's inequality* (in the next section).

---

6. *Minkowski's inequality*

Again, we assume that functions f,g below have integrals with basic properties. Take X to be $[a,b]$ or $\mathbb{R}$ or $\mathbb{R}^n$, for example.

For the triangle inequality in L^p spaces for general p, we need

[6.1] Corollary: *(Minkowski)* For $1 < p < +\infty$ and $[0, +\infty]$-valued functions f, g on X,

$$\Big(\int_X (f+g)^p\Big)^{\frac{1}{p}} \;\le\; \Big(\int_X f^p\Big)^{\frac{1}{p}} + \Big(\int_X g^p\Big)^{\frac{1}{p}}$$

Proof: We prove Minkowski's inequality from Hölder's, using the conjugate exponents p and $q = \frac{p}{p-1}$.

$$\int (f+g)^p \;=\; \int f\cdot (f+g)^{p-1} + \int g\cdot (f+g)^{p-1}$$

$$\le\; \Big(\int f^p\Big)^{\frac{1}{p}} \cdot \Big(\int (f+g)^{(p-1)q}\Big)^{\frac{1}{q}} \;+\; \Big(\int g^p\Big)^{\frac{1}{p}} \cdot \Big(\int (f+g)^{(p-1)q}\Big)^{\frac{1}{q}}$$

$$=\; \Big[\Big(\int f^p\Big)^{\frac{1}{p}} + \Big(\int g^p\Big)^{\frac{1}{p}}\Big] \cdot \Big(\int (f+g)^p\Big)^{\frac{p-1}{p}}$$

Dividing through by $\big(\int (f+g)^p\big)^{\frac{p-1}{p}}$ gives Minkowski's inequality. ///

Since no specifics about the integration were used, we also have discrete forms:

[6.2] Theorem: Let $X = \{1, 2, \ldots\}$, with weights $0 \le w_n$. For conjugate exponents p, q and $[0, +\infty)$-valued functions f, g on X,

$$\Big(\sum_n w_n \cdot (f(n) + g(n))^p\Big)^{\frac{1}{p}} \;\le\; \Big(\sum_n w_n \cdot f(n)^p\Big)^{\frac{1}{p}} + \Big(\sum_n w_n \cdot g(n)^p\Big)^{\frac{1}{p}}$$

[6.3] Theorem: Let $X = \{1, 2, \ldots\}$. For conjugate exponents p, q and $[0, +\infty)$-valued functions f, g on X,

$$\Big(\sum_n (f(n) + g(n))^p\Big)^{\frac{1}{p}} \;\le\; \Big(\sum_n f(n)^p\Big)^{\frac{1}{p}} + \Big(\sum_n g(n)^p\Big)^{\frac{1}{p}}$$

---

7. *Example: ℓ^p spaces*

For $1 \le p < \infty$ the usual ℓ^p norm on sequences $c = (c_1, c_2, \ldots)$ of complex numbers is

$$|c|_{\ell^p} \;=\; \left(\sum_n |c_n|^p\right)^{1/p}$$

The ℓ^p spaces are

$$\ell^p = \{c = (c_1, c_2, \ldots) : |c|_{L^p} < \infty\}$$

with associated metric

$$d(c, c') \;=\; |c - c'|_{\ell^p}$$

[7.1] Theorem: ℓ^p is a *complete* metric space.

Proof: For notational purposes, express elements f of ℓ^p as functions f from $X = \{1, 2, \ldots\}$ to $\mathbb{C}$. Write

$$\int_X f \;=\; \sum_{n=1}^{\infty} f(n)$$

The triangle inequality for the alleged metric is exactly *Minkowski's inequality.* To prove completeness, choose a subsequence f_{n_i} such that

$$|f_{n_i} - f_{n_{i+1}}|_p \ < \ 2^{-i}$$

and put

$$g_n(x) \ = \ \sum_{1 \le i \le n} |f_{n_{i+1}}(x) - f_{n_i}(x)| \qquad\qquad (\text{for } x \in X)$$

and

$$g(x) \ = \ \sum_{1 \le i < \infty} |f_{n_{i+1}}(x) - f_{n_i}(x)| \qquad\qquad (\text{for } x \in X)$$

The triangle inequality shows that $|g_n|_p \le 1$.

A discrete version of *Fatou's Lemma* asserts that for $[0, \infty]$-valued functions h_i on $X = \{1, 2, 3, \ldots\}$

$$\int_X \Big(\liminf_i h_i\Big) \ \le \ \liminf_i \int_X h_i$$

Thus, $|g|_p \le 1$, so is finite. Thus,

$$f_{n_1}(x) + \sum_{i \ge 1} (f_{n_{i+1}}(x) - f_{n_i}(x))$$

converges for all $x \in X$.

Now prove that this is the ℓ^p-limit of the original sequence. For $\varepsilon > 0$ take N such that $|f_m - f_n|_p < \varepsilon$ for $m, n \ge N$. Fatou's lemma gives

$$\int |f - f_n|^p \le \liminf_i \int |f_{n_i} - f_n|^p \le \varepsilon^p$$

Thus $f - f_n$ is in ℓ^p and hence f is in ℓ^p. And $|f - f_n|_p \to 0$. ///

---

8. *Appendix: discrete Fatou lemma and Lebesgue monotone convergence*

[8.1] Claim: (Fatou) For $[0, +\infty]$-valued functions f_j on $\{1, 2, 3, \ldots\}$,

$$\sum_{n=1}^{\infty} \liminf_j f_j(n) \ \le \ \liminf_j \sum_{n=1}^{\infty} f_j(n)$$

Proof: Letting $g_j(n) = \inf_{i \ge j} f_j(n)$, certainly $g_j(n) \le f_j(n)$ for all n, and $\sum_n g_j(n) \le \sum_n f_j(n)$. Also, $g_1(n) \le g_2(n) \le \ldots$ for all n, and $\lim_j g_j(n) = \liminf_j f_j(n)$. A discrete form of the Monotone Convergence Theorem, proven just below, is

$$\sum_n \lim_j g_j(n) \ = \ \lim_j \sum_n g_j(n)$$

Thus,

$$\sum_n \liminf_j f_j(n) \ = \ \sum_n \lim_j g_j(n) \ = \ \lim_j \sum_n g_j(n) \ = \ \liminf_j \sum_n g_j(n) \ \le \ \liminf_j \sum_n f_j(n)$$

as claimed. ///

[8.2] Theorem: *(Discrete version of Lebesgue's Monotone Convergence Theorem)* For $[0,+\infty]$-valued functions f_j on $\{1,2,3,\ldots\}$, with $f_1(n) \le f_2(n) \le \ldots$ for all n,

$$\lim_j \sum_{n=1}^\infty f_j(n) \;=\; \sum_{n=1}^\infty \lim_j f_j(n) \qquad \text{(allowing value } +\infty)$$

Proof: Each non-decreasing sequence $f_1(n) \le f_2(n) \le \ldots$ has a limit $f(n) \in [0,+\infty]$. Similarly, since $\sum_n f_j(n) \le \sum_n f_{j+1}(n)$, the non-decreasing sequence of these sums has a limit $S = \lim_j \sum_n f_j(n)$. Since $f_j(n) \le f(n)$, certainly $\sum_n f_j(n) \le \sum_n f(n)$, and $S \le \sum_n f(n)$.

Fix N, and put $g(n) = f(n)$ for $n \le N$ and $g(n) = 0$ for $n > N$. For $\varepsilon > 0$, let

$$E_j \;=\; \{n : \sum_n f_j(n) \ge (1-\varepsilon)\cdot \sum_n g(n)\} \qquad \text{(for } j = 1,2,\ldots)$$

Certainly $E_1 \subset E_2 \subset \ldots$, since $f_{j+1}(n) \ge f_j(n)$ for all n. We claim that $\bigcup E_j = \{1,2,\ldots\}$: for $f(n) > 0$,

$$\lim_j f_j(n) \;=\; f(n) \;>\; (1-\varepsilon)\cdot f(n) \;\ge\; (1-\varepsilon)\cdot g(n) \qquad \text{(for all } n)$$

and for $f(n) = 0$, also $g(n) = 0$, and

$$f_1(n) \;\ge\; 0 \;\ge\; (1-\varepsilon)\cdot g(n)$$

Then

$$\sum_n f_j(n) \;\ge\; \sum_{n\in E_j} f_j(n) \;\ge\; (1-\varepsilon)\cdot \sum_{n\in E_j} g(n)$$

The set of n for which $g(n)$ is non-zero is finite, so there is j_o such that for $j \ge j_o$

$$\sum_{n\in E_j} g(n) \;=\; \sum_n g(n) \qquad \text{(for all } j \ge j_o)$$

That is, $\lim_j \sum_n f_j(n) \ge (1-\varepsilon)\sum_n g(n)$. Then

$$S \;=\; \lim_j \sum_n f_j(n) \;\ge\; (1-\varepsilon)\cdot \lim_j \sum_{n\in E_j} g(n) \;=\; (1-\varepsilon)\cdot \sum_n g(n)$$

This holds for every $\varepsilon > 0$, so $S \ge \sum_n g(n) = \sum_{n\le N} f(n)$. This holds for every N, so $S \ge \sum_n f(n)$. ///

[Riesz-Nagy 1952] F. Riesz, B. Szökefalvi-Nagy, *Functional Analysis*, English translation, 1955, L. Boron, from *Lecons d'analyse fonctionelle* 1952, F. Ungar, New York, 1955

[wiki/Young's_inequality 2012] http://en.wikipedia.org/wiki/Young's_inequality, retrieved Feb 29, 2012

[Young 1912a] W.C. Young, *On classes of summable functions and their Fourier series*, Proc. Royal Soc. Lond. **87** (1912), 225-229.

[Young 1912b] W.C. Young, *On the multiplication of successions of Fourier constants*, Proc. Royal Soc. Lond. **87** (1912), 331-339.
