Basic inequalities

N.A.

07

A function $\Phi : I \to \mathbb{R}$ defined on an interval $I \subseteq \mathbb{R}$ is *convex* if for all integer $n \geq 2$:

$$\sum_{j=1}^{n} t_j = 1,\ t_1, \ldots, t_n \geq 0,\ a_1, \ldots, a_n \in I \implies \Phi\left(\sum_{j=1}^{n} t_j a_j\right) \leq \sum_{j=1}^{n} t_j \Phi\left(a_j\right).$$

By induction, Φ is convex iff the inequality above holds for $n = 2$.
Jensen's inequality.[1] Let $\Phi : [0, +\infty) \to [0, \infty)$ be a convex function and let (X, μ) be a probability measure space. If $f \geq 0$ is a measurable function on X, then

$$\Phi\left(\int_X f d\mu\right) \leq \int_X \Phi(f) d\mu. \tag{1}$$

Proof. Let $f = \sum_j a_j \chi_{E_j}$ be a simple function: $\{E_j\}$ is a countable, measurable partition of X. Then, by convexity,

$$\begin{aligned}\Phi\left(\int_X f d\mu\right) &= \Phi\left(\sum_{j=1}^{n} a_j \mu(E_j)\right)\\ &\leq \sum_{j=1}^{n} \Phi\left(a_j\right) \mu(E_j) = \int_X \Phi(f) d\mu.\end{aligned}$$

For general $f \geq 0$, let $\{f_n\}$ be a sequence of simple functions such that $f_n \nearrow f$. The desired inequality follows by a simple limiting argument[2]. ■

We can dispense with the positivity assumption provided f is integrable.

Proposition 1 *Let $\Phi : (a, b) \to \mathbb{R}$ be convex, $-\infty \leq a < b \leq +\infty$, and let (X, μ) be a probability space. If $f : X \to \mathbb{R}$ is integrable and $f(X) \subseteq (a, b)$, then*

$$\Phi\left(\int_X f d\mu\right) \leq \int_X \Phi(f) d\mu.$$

Proof. Let $a < u < w < v < b$. By convexity,

$$\begin{aligned} w &= \frac{v-w}{v-u} u + \frac{w-u}{v-u} v \implies \\ \Phi(w) &\leq \frac{v-w}{v-u} \Phi(u) + \frac{w-u}{v-u} \Phi(v) \implies \end{aligned}$$

[1] Some words on extremals?

[2] It is useful to split $\Phi = \Phi_1 + \Phi_2$, with Φ_1 increasing and Φ_2 decreasing. Use Monotone Convergence with Φ_1 and Dominated Convergence with Φ_2

$$\begin{aligned}(\Phi(w)-\Phi(u))(v-w) &\leq (\Phi(v)-\Phi(w))(w-u) \implies \\ \frac{\Phi(w)-\Phi(u)}{w-u} &\leq \frac{\Phi(v)-\Phi(w)}{v-w}.\end{aligned}$$

[3] Then, there is $C(w) \in \mathbb{R}$ such that

$$\Phi(t) \geq \Phi(w) + C(w)(t-w)$$

whenever $t \in (a,b)$.

Let now $w = \int_X f d\mu \in (a,b)$, by the Mean Value Theorem, and let $t = f(x)$. Integrating w.r.t. μ,

$$\begin{aligned}\int_X \Phi(f(x))d\mu(x) &\geq \Phi\left(\int_X f d\mu\right) + C\int_X \left(f(x) - \int_X f d\mu\right) d\mu(x) \\ &= \Phi\left(\int_X f d\mu\right).\end{aligned}$$

■ [4]

Exercise 4 *Suppose that Φ is also increasing and that for all*[5] *$T > 0$ there is*

---

[3]Observe that the inequality in the second line also gives

$$\frac{\Phi(v)-\Phi(u)}{v-u} \leq \frac{\Phi(v)-\Phi(w)}{v-w}.$$

[4]**A different proof.**

Lemma 2 (An extension of the Monotone Convergence Theorem.) *Suppose that $\varphi_n \in L^(\mu)$ for $n \geq 1$ and that $\varphi_n \nearrow \varphi$. Then,*

$$\int \varphi_n d\mu \nearrow \int \varphi d\mu.$$

Proof. Let $\psi_n = \varphi_n \vee 0 \nearrow \psi = \varphi \vee 0$ and $\eta_n = \varphi_n \wedge 0 \nearrow \eta = \varphi \wedge 0$. Use MCT for ψ_n and DCT for η_n. ■

Lemma 3 *Let $\Phi : (a,b) \to \mathbb{R}$ be a convex function and let $a < \alpha < \beta < b$. Then, there exist*

$$-\infty < \Phi'(\alpha+) \leq \Phi'(\beta-) < +\infty.$$

Proof. Whenever $0 < h, k < \frac{\alpha+\beta}{2}$, we have

$$\frac{\Phi(\alpha+h)-\Phi(\alpha)}{h} \leq \frac{\Phi(\beta)-\Phi(\beta-k)}{k}.$$

The LHS decreases as $h \to 0$, while the RHS increases as $k \to 0$. Observe that both RHS and LHS are bounded. Take limits. ■

Proof. of Proposition 1 For $a < \alpha < \beta < b$, let

$$\Phi_\alpha^\beta(t) = \begin{cases} \Phi(\beta) + \Phi'(\alpha+)(t-\alpha) \text{ if } t \in (a,\alpha] \\ \Phi(t) \text{ if } t \in [\alpha,\beta] \\ \Phi(\beta) + \Phi'(\beta-)(t-\beta) \text{ if } t \in [\beta,b). \end{cases}$$

Then, Φ_α^β is convex by the second lemma, $\Phi_\alpha^\beta \leq \Phi$ and, if $\alpha_n \searrow a$ and $\beta_n \nearrow b$, then $\Phi_{\alpha_n}^{\beta_n} \nearrow \Phi$.

If $f \in L^1(\mu)$, then $\varphi_n = \Phi_{\alpha_n}^{\beta_n} \circ f \in L^1(\mu)$ and $\varphi_n \nearrow \varphi = \Phi \circ f$. By the first lemma, the inequality is reduced to

$$\Phi_{\alpha_n}^{\beta_n}\left(\int f d\mu\right) \leq \int \Phi_{\alpha_n}^{\beta_n}(f)\, d\mu.$$

This last inequality can be proved similarly to (1): for simple f it reduces to the definition of convex function; for $f \in L^1$ use DCT. ■

[5]Or, which is the same, for just one such T.

$C > 0$ such that

$$\Phi(Tx) \le C\Phi(x). \tag{2}$$

($\Phi(t) = t^p,\ p \ge 1$ is a function with these properties).

Show that, if we replace the assumption that $\mu(X) = 1$ by $\mu(X) < \infty$, we obtain the inequality

$$\Phi\left(\int_X f d\mu\right) \le C(\mu(X)) \int_X \Phi(f) d\mu. \tag{3}$$

Find an example of a convex, increasing function Φ such that (2) and (3) both fail.

Exercise 5 *Let $\psi : [0, \infty) \to [0, \infty)$, $\psi(0) = 0$, $\psi(x) = x \log(1/x)$ if $x \neq 0$. Let $P = \{p_j\}_{1 \le j \le n}$ be a probability distribution: $\sum_{j=1}^n p_j = 1$, $p_j \ge 0$. The* entropy *of P is $\mathcal{E}(P) = \sum_{j=1}^n \psi(p_j)$. Prove that the estimates*

$$0 \le \mathcal{E}(P) \le \log n$$

hold and that they are sharp. What are the extremals?

Hölder's inequality. If $f, g \ge 0$ are nonnegative and measurable on the measure space (Y, dx), $1 \le p \le \infty$ and p' is the exponent conjugate to p, $\frac{1}{p} + \frac{1}{p'} = 1$, then

$$\int fg dx \le \|f\|_{L^p} \|g\|_{L^{p'}}.$$

Proof. The case $p = \infty$ or $p = 1$ is elementary, so we assume $1 < p < \infty$. We use the convexity of $t \to t^p$ and Jensen's inequality with the measure space (Z, μ), where Z is the support of g and $d\mu = \frac{g^{p'}}{\|g\|_{L^{p'}}^{p'}} dx$.

$$\begin{aligned}
\int fg dx &= \int f g^{1-p'} \frac{g^{p'}}{\|g\|_{L^{p'}}^{p'}} dx \cdot \|g\|_{L^{p'}}^{p'} \\
&\le \|g\|_{L^{p'}}^{p'} \left[\int \left(f g^{1-p'} \right)^p \frac{g^{p'}}{\|g\|_{L^{p'}}^{p'}} dx \right]^{1/p} \\
&\le \|f\|_{L^p} \|g\|_{L^{p'}}^{p' - \frac{p'}{p}} = \|f\|_{L^p} \|g\|_{L^{p'}}.
\end{aligned}$$

■

We have equality in Hölder's inequality if and only if $g^{p'} = f^p$ *a.e.*.

Iterated Hölder's inequality. If $p_j \in [1, \infty]$, $\sum_j \frac{1}{p_j} = 1$ and $f_j \ge 0$ is a family of measurable functions, then

$$\int \Pi_j f_j dx \le \Pi_j \|f_j\|_{L^{p_j}}.$$

The inequality follows from two-Hölder's by induction.

There is a continuous generalization of Hölder's inequality, which can be stated as follows. Let μ be a probability measure on some space X and $h = h(t, x) : X \times Y \to \mathbb{R}$ be mesurable and nonnegative. Then,

$$\log \left[\int_Y \exp\left(\int_X h(t, x) d\mu(t) \right) dx \right] \le \int_X \log \left[\int_Y \exp(h(t, x)) dx \right] d\mu(t).$$

This inequality follows easily from iterated Hölder's and an approximation argument.[6]

Let f, g be nonnegative, measurable functions on $\mathbb{R}$. The *convolution* of f and g is $f * g : \mathbb{R} \to \mathbb{R} \cup \{\infty\}$:

$$f * g(x) = \int f(x-y)g(y)dy.$$

Note that the convolution can be defined as well among sequences with indeces in $\mathbb{Z}$:

$$(a * b)_n = \sum_{m \in \mathbb{Z}} a_{n-m} b_m,$$

and functions defined on the circle:

$$f * g(e^{i\alpha}) = \int_{-\pi}^{\pi} f(e^{i(\alpha-\theta)})g(e^{i\theta})\frac{d\theta}{2\pi}.$$

In general, it makes sense to define convolutions whenever we have a group with a (left) translation invariant measure.

Young's inequality. Suppose that f, g are nonnegative and measurable on $\mathbb{R}$ and that $p, q, r \in [1, +\infty]$ are such that

$$\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r} \geq 0. \tag{4}$$

(The conditions imply that $r \geq p, q$.) Then,

$$\|f * g\|_{L^r} \leq \|f\|_{L^p} \|g\|_{L^q}. \tag{5}$$

Proof. The case $r = +\infty$ is contained in Hölder's inequality. Consider $r \in (1, \infty)$ first. Then $p, q \in [1, \infty)$. Let $P, Q \in [1, \infty)$, to be chosen below, be such that $P^{-1} + Q^{-1} + r^{-1} = 1$ and let $a, b \in [0, 1]$ be such that

$$p = aP = (1-a)r, \ q = bQ = (1-b)r, \ i.e. \ a = \frac{r}{P+r} = \frac{r}{\frac{p}{a}+r}, \ b = \frac{r}{Q+r} = \dots.$$

Hence,

$$a = \frac{r-p}{r}, \ b = \frac{r-q}{r}.$$

Let's verify the condition on P, Q:

$$\frac{1}{P} + \frac{1}{Q} = \frac{a}{p} + \frac{b}{q} = \frac{1}{p} - \frac{1}{r} + \frac{1}{q} - \frac{1}{r} = 1 - \frac{1}{r},$$

as wished. By Hölder's inequality,

$$\begin{aligned} f * g(x) &= \int f(x-y)g(y)dy = \int f(x-y)^a g(y)^b f(x-y)^{1-a} g(y)^{1-b} dy \\ &\leq \left[\int (f(x-y)^{aP} dy\right]^{1/P} \left[\int (g(y)^{bQ} dy\right]^{1/Q} \\ &\cdot \left[\int f(x-y)^{(1-a)r} (g(y)^{(1-b)r} dy\right]^{1/r}. \end{aligned}$$

[6] Is there a direct proof?

Taking into account the fact that the Lebesgue measure is translation invariant and the relations on a, b, P, Q, p, q, we have

$$\begin{aligned}\int [f(x-y)g(y)]^r dx &\leq \|f\|_{L^p}^{pr/P} \|g\|_{L^q}^{qr/Q} \int \left[\int f(x-y)^p g(y)^q dy\right] dx \\ &= \|f\|_{L^p}^{p(1+r/P)} \|g\|_{L^q}^{q(1+r/Q)} = \|f\|_{L^p}^r \|g\|_{L^q}^r.\end{aligned}$$

■

Exercise 6 *Prove that Young's inequality (5) in $\mathbb{R}$ holds in the cases (4) only.* **Suggestion:** *let $\delta_\lambda f(x) = f(x/\lambda)$, $\lambda > 0$; insert $\delta_\lambda f$, $\delta_\lambda g$ instead of f and g in (5) and let λ renge in $(0, \infty)$.*

We have not proved the important (easier) case $p = q = r = 1$. For these values of the exponents, Young's inequality says that $(L^1, *)$ is a *Banach algebra*.

Exercise 7 *Prove the case $p = q = r = 1$ of Young's inequality.*

Problem. Find an iterated version for Young's inequality and, if there is one, write down a continuos version of it.

The best constant in Young's inequality in $\mathbb{R}^n$ was found by W. Beckner in 1975 [Beck].

Schur' Lemma. Let X be a measure space and let $K : X \times X \to \mathbb{R}$ be a nonnegative, measurable functions. Define an operator T defined by the kernel K. If f is a nonnegative, measurable function, then

$$Tf(x) = \int K(x,y) f(y) dy.$$

Let $1 < p < \infty$. Suppose that there is a strictly positive function λ on X such that

$$\int K(x,y)\lambda^{p'}(y)dy \leq C\lambda^{p'}(x), \quad \int K(x,y)\lambda^p(x)dx \leq C\lambda^p(y).$$

Then, T is bounded from L^p to L^p.

Proof. Using Hölder's from first to second line with measure $K(x,y)dy$ and the hypothesis

$$\begin{aligned}\int (Tf)^p(x)dx &= \int \left(\int K(x,y)\lambda(y)\lambda^{-1}(y)f(y)dy\right)^p dx \\ &\leq \int \left(\int K(x,y)\lambda^{p'}(y)dy\right)^{p/p'} \left(\int K(x,y)\lambda^{-p}(y)f(y)^p dy\right) dx \\ &\leq C\int \lambda^p(x) \left(\int K(x,y)\lambda^{-p}(y)f(y)^p dy\right) dx \\ &= C\int \lambda^{-p}(y)f(y)^p \left(\int K(x,y)\lambda^p(x)dx\right) dy \\ &\leq C' \int f(y)^p \lambda^{-p}(y)\lambda^p(y)dy = C' \int f(y)^p dy.\end{aligned}$$

■

Exercise 8 *Find a Schur's Lemma ensuring that $T : L^p \to L^q$, $1 < q \leq p < \infty$.*

Jensen's inequality deals with a probability space. A consequence of the inequality is that $L^q(\mu) \subset L^p(\mu)$ if μ is a probability measure and $p < q$. At the opposite end, we have the measure space $\mathbb{N}$ with the counting measure. Here, $\ell^1 \subset \ell^\infty$.

Proposition 9 *Let $a = \{a_k\}_{k\geq 0}$ be a sequence of nonnegative numbers. If $p < q$, then $\|a\|_q \leq \|a\|_p$.*

Proof. Since $t^{q/p} \leq t$ when $t \in [0,1]$,

$$\frac{\sum a_k^q}{(\sum a_h^p)^{q/p}} = \sum \left(\frac{a_k^p}{\sum a_h^p}\right)^{q/p} \leq \sum \left(\frac{a_k^p}{\sum a_h^p}\right) = 1.$$

■

Exercise 10 *Prove the following. Let $\Phi : [0,\infty) \to [0,\infty)$ be such that $\frac{\Phi(x)}{x}$ is increasing. Then,*

$$\sum \Phi(a_k) \leq \Phi(\sum a_k)$$

if $a_k \geq 0$.

For instance,

$$\sum (e^{a_k} - 1) \leq e^{\sum a_k} - 1.$$

Note that here, too, a differential inequality is the key to a class of integral inequalities.

An other application of convexity is the proof of *Minkowsky's inequality.* Let $\Phi : [0,\infty) \to [0,\infty)$ be a convex, increasing function such that $\Phi(0) = 0$. For a measurable function f (on some fixed measure space), let

$$\|f\|_\Phi = \inf\left\{C > 0 : \int \Phi\left(\frac{|f|}{C}\right) dx \leq 1\right\}.$$

Here, $\inf \emptyset = +\infty$, by definition.

Exercise 11 *If $\Phi(t) = t^p$, $p \geq 1$, then $\|f\|_\Phi = \|f\|_{L^p}$.*

Theorem 12

$$\|f+g\|_\Phi \leq \|f\|_\Phi + \|g\|_\Phi.$$

Proof. Let $a, b > 0$ be s.t. $\int \Phi\left(\frac{|f|}{a}\right) dx \leq 1$, $\int \Phi\left(\frac{|g|}{b}\right) dx \leq 1$. By convexity,

$$\begin{aligned}\int \Phi\left(\frac{|f+g|}{a+b}\right) dx &\leq \int \Phi\left(\frac{a}{a+b}\frac{|f|}{a} + \frac{b}{a+b}\frac{|g|}{b}\right) dx \\ &\leq \int \frac{a}{a+b}\Phi\left(\frac{|f|}{a}\right) + \frac{b}{a+b}\Phi\left(\frac{|g|}{b}\right) dx \\ &\leq 1. \end{aligned} \tag{6}$$

Hence, $a + b \geq \|f+g\|_\Phi$. The thesis follows by passing to infima. ■

Under suitable hypothesis, Minkowsky's inequality has the following integral generalization:

$$\left\|\int_X f(t,\cdot) d\lambda(t)\right\|_\Phi \leq \int_X \|f(t,\cdot)\|_\Phi d\lambda(t).$$

The classic of inequalities is [HLP].

References

[Beck] W. Beckner *Inequalities in Fourier analysis.* Ann. of Math. (2) **102** (1975), no. 1, 159–182.

[HLP] G. H. Hardy, J. E. Littlewood, G. Pólya, *Inequalities.*Reprint of the 1952 edition. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 1988. xii+324 pp.