Intuition for Cauchy-Schwarz and Hölder

Daniel Rui

tags: exposition, exercises, inequalities, basic

Sadly it still takes me a while to see CBS (Cauchy-Bunyakovsky-Schwarz) or Hölder. It's kinda embarrassing.

1. Classic $|\langle u, v\rangle| \leqslant \|u\| \cdot \|v\|$.
   - $\sum fg \leqslant (\sum f^2)^{1/2}(\sum g^2)^{1/2}$
   - $\int fg \leqslant (\int f^2)^{1/2}(\int g^2)^{1/2}$
   - $\mathbb{E}fg \leqslant (\mathbb{E}f^2)^{1/2}(\mathbb{E}g^2)^{1/2}$
2. "Classic squared":
   - to upper bound, pull/distribute 2 inside, then distribute $\sum$ inside.
   - to lower bound, push squares outside
3. CBS also often applied with just one term (of the product, e.g. $g \equiv 1$ or something) especially when the measure space is finite. In that case, CBS will have extra factor of total-measure on the RHS (or BHS, bigger hand side).
   - to upper bound, pull/distribute 2 inside, multiply by total-measure.
     A different example of upper bound: assuming $|a_i|^2 \leqslant f_i g_i$, then $(\sum |a_i|)^2 = (\sum \sqrt{f_i}\sqrt{g_i}) \leqslant (\sum f_i)(\sum g_i)$ (e.g .246CNotes2 Thm.31)
   - to lower bound, push 2 outside, divide by total-measure (e.g. 246CNotes2 Thm.29)
4. The Thm.31 example has given me a better mental picture of when/how CBS works: anytime see $\big(\sum[\ldots\ldots]\big)^2$, first pull 2 inside to $\big(\sum[\ldots\ldots]^2\big)$ and then can split $[\ldots\ldots]^2$ however you want into 2 pieces (!!!), then get upper bound $\leqslant (\sum\ldots)(\sum\ldots)$.

   So what I said above "to upper bound, pull/distribute 2 inside, then distribute $\sum$ inside. " turns out to be exactly on the nose.

   I.e.
   $$(\sum\ldots\ldots)^2 \leqslant (\sum\ldots)(\sum\ldots) \text{ for any splitting } [\ldots\ldots]^2 = [\ldots]\cdot[\ldots]$$

   Or if you have 2 copies of the same sum $(\sum \ddots)(\sum \ddots)$, you can shuffle around factors of $[\ddots]$ between the 2 sums willy-nilly, but get upper bound.
   "Most symmetric sum state (with total factors fixed) = lowest energy"

   To be even clearer (since I tested myself with e.g. good MSE problem and failed), if we have $(a_1 + a_2 + \ldots)^2$, can upper bound by $(b_1 + \ldots)(c_1 + \ldots)$ where $a_i^2 = b_i c_i$, i.e. a summand by summand splitting, where each summand is (multiplicatively) weighted by the power/exponent on the sum it's in.
   The MSE example was $(x^3+y^2+z)(\frac{1}{x}+1+z) \geqslant (x+y+z)^2$ and $\sum_{\text{cyc}} xz \leqslant (x+y+z)^{1/2}(z+x+y)^{1/2}$.
5. Converse CBS "linearization", also "rid of abs. val." $(\sum |a_i|^2)^{1/2} = \sum \bar{c}_i a_i$ for some c_i with $\sum |c_i|^2 = 1$ (indeed, take $\vec{c} = \hat{a}$). Used in 254ANotes3

6. Weighted CBS $\sum fg \leqslant (\sum f^2/\nu)^{1/2}(\sum g^2\nu)^{1/2}$ (for $\nu \geqslant 0$; remove n for which $\nu(n) = 0$). Also used in 254ANotes3

7. Reverse CBS

$$\sum \frac{a_i^2}{b_i} \geqslant \frac{(\sum a_i)^2}{\sum b_i}$$

(e.g .246CNotes2 Thm.34 — this is the CBS application I remembered from how Terry looked at us lol. I search CBS in these notes, and by accident also some nice examples I didn't remember, which I wrote about above)

I.e. reverse CBS is: if I have any sum $(\sum \ldots)$, I can LOWER bound by any $\geqslant (\sum \ddots)^2$ I want, but just have to pay cost of **divide** RHS by $(\sum \cdots)$ where the only restriction is to keep total factors (i.e. ignore $\sum$ symbols) on both sides the same.

And classic CBS is: if I have any sum $(\sum \ddots)^2$, I can UPPER bound by any $\leqslant (\sum \ldots)$ I want, but just have to pay cost of **multiply** RHS by $(\sum \cdots)$ where the only restriction is to keep total factors (i.e. ignore $\sum$ symbols) on both sides the same.

0.0.1 Hölder

1. Classic Hölder: can split any norm L^p into pieces that reciprocal-sum to $\frac{1}{p}$. (Make numerators of subscript, e.g. $\| \bullet \|_{pD/D}$ all the same, say pD; denominators of subscripts on RHS need to add to D)

   I try for more conceptual picture (things inside sum are $\geqslant 0$): $(\sum \ldots)^k \leqslant (\sum \cdots) \cdots (\sum \cdots)$ as long as total factors on both sides are the same?

   e.g. $(\sum abc)^3 \leqslant (\sum f)(\sum g)(\sum h)$ as long as $(abc)^3 = fgh$? Yes, indeed $abc = f^{1/3}g^{1/3}h^{1/3}$, and then Hölder gives

$$\|abc\|_1 \leqslant \|f^{1/3}\|_3 \|g^{1/3}\|_3 \|h^{1/3}\|_3 \iff (\sum abc)^3 \leqslant (\sum f)(\sum g)(\sum h)$$

   and

$$\|abc\|_{3/3} \leqslant \|f^{1/3}\|_{3/1} \|g^{1/3}h^{1/3}\|_{3/2} \iff (\sum abc)^3 \leqslant (\sum f)(\sum g^{1/2}h^{1/2})^2$$

   So in fact Classic Hölder is: if I have any $(\sum \ldots)^p$ (for any power $p \geqslant 1$!), I can upper bound it by $\leqslant \prod_j (\sum \cdots_j)^{p_j}$ where the only restrictions are $p = \sum_j p_j$, and total factors on both sides (i.e. ignore $\sum$ symbols, $[\ldots]^p = \prod_j [\cdots_j]^{p_j}$)) are the same!

   As I explain in my MSE treatise on the subject, Hölder inequality follows in 3 steps: weighted AM-GM, integrating, and then arbitraging symmetry.

   Minimal energy idea I said above, reminds me very much of foundational entropy inequalities, e.g. uniform distribution maximizes entropy, which can be proven using weighted AM-GM. (Where weighted AM-GM is literally just concavity of log.) Hmm...