

# TJ USAMO Practice/Lecture 2 - Inequalities

PDiao05\*

October 6, 2005

Inequalities often come up on the USAMO. To learn to do these it is important to have seen many inequalities. That way you gain a larger base of inequalities you know to be true. There are several techniques and theorems that come up over and over again. The most fundamental inequality of all is! drumroll...

- **Jensen!** - Let  $f : [a, b] \rightarrow \mathbb{R}$  be a convex function. Then for any  $x_1, \dots, x_n \in [a, b]$  and any nonnegative reals  $\lambda_1, \dots, \lambda_n$  summing to one, Jensen says:

$$\lambda_1 f(x_1) + \dots + \lambda_n f(x_n) \geq f(\lambda_1 x_1 + \dots + \lambda_n x_n)$$

With equality iff all the  $x_i$  are equal. You might be wondering what a convex function is. Here are some ways you can tell if something is convex:

1. You can check if  $f$  is a convex function by looking at the area above the graph of your function. If that region is convex in the geometric sense then your function is also convex. Geometrically, a region is convex iff (if and only if) all the points on the line segment joining any two points in the region is also in the region.
2. The formal definition is  $(1 - \lambda)f(x) + \lambda f(y) \geq f((1 - \lambda)x + \lambda y) \forall \lambda \in [0, 1]$  Most of the time we deal with continuous  $f$  and in those cases it suffices to check for just a single value of  $\lambda$ , generally  $1/2$ . The reason I say suffice and not iff is because conversely, if  $f$  is convex, then  $f$  is continuous except possibly at  $a$  and  $b$ .
3. Lastly, you may already know this from calculus, but if  $f$  is twice differentiable, you can just check for  $f''(x) \geq 0$ .

Some common examples of convex functions:

1.  $y = |x - a|$
2.  $y = cx^\alpha$  where  $\alpha \geq 1$  or  $\alpha \leq 0$  and  $x > 0$ .
3.  $y = a^x$  for  $a \in \mathbb{R}^+$ .

---

\*Although based on Mildorf and Reid Barton Lectures

Jensen follows from the definition of convexity and smoothing. Smoothing is essentially when we keep on changing the inequality to be less and less true and then show that even when it is least true, it is still true.

- **Weighted Power Mean**- If  $x_1, \dots, x_n$  are nonnegative reals and  $\lambda_1, \dots, \lambda_n$  are nonnegative reals summing to one again, then The Weighted Power Mean says:

$$(\lambda_1 x_1^r + \dots + \lambda_n x_n^r)^{1/r}$$

is a nondecreasing function of  $r$ . It is strictly increasing unless all of the  $x_i$  are equal. When  $r = 0$ , this expression represents the weighted geometric mean of the  $x_i$ . In fact our beloved AM-GM-HM (Arithmetic Mean, Geometric Mean, Harmonic Mean) inequality is just a specific application of this. That inequality states  $\forall a_1, \dots, a_n \in \mathbb{R}^+$ , we have

$$\frac{\sum_{i=1}^n a_i}{n} \geq \left( \prod_{i=1}^n a_i \right)^{\frac{1}{n}} \geq \frac{n}{\sum_{i=1}^n \frac{1}{a_i}}$$

with equality iff  $a_1 = \dots = a_n$ . This inequality is extremely important. Jensen on a special function yields this inequality.

- **Hölder** - Let  $(a_1, \dots, a_n), (b_1, \dots, b_n), \dots, (z_1, \dots, z_n)$  be sequences of nonnegative real numbers and  $\lambda_a, \lambda_b, \dots, \lambda_z$  be nonnegative reals summing to one. Then Hölder says:

$$(a_1 + \dots + a_n)^{\lambda_a} (b_1 + \dots + b_n)^{\lambda_b} \dots (z_1 + \dots + z_n)^{\lambda_z} \geq a_1^{\lambda_a} b_1^{\lambda_b} \dots z_1^{\lambda_z} + \dots a_n^{\lambda_a} b_n^{\lambda_b} \dots z_n^{\lambda_z}$$

In reality, the most commonly used special case of Hölders is the Cauchy Schwartz inequality. In vector land it is just a statement about the fact that the dot product of two vectors cannot exceed the product of the magnitudes of the two vectors and are only equal when the vectors are parallel. It states  $\forall a_1, \dots, a_n, b_1, \dots, b_n \in \mathbb{R}$ , we have

$$(a_1^2 + \dots + a_n^2) (b_1^2 + \dots + b_n^2) \geq (a_1 b_1 + \dots + a_n b_n)^2$$

with equality iff  $b_k = ca_k$  for  $k = 1, \dots, n$ . As a random note, Hölder's Inequality can be used to prove Minkowski's inequality an even more obscure thing. However a special case of it is the triangle inequality which is useful sometimes, especially in geometric inequalities:

$$\sqrt[r]{x_1^r + \dots + x_n^r} + \sqrt[r]{y_1^r + \dots + y_n^r} \geq \sqrt[r]{(x_1 + y_1)^r + \dots + (x_n + y_n)^r}$$

for  $x_1, \dots, x_n, y_1, \dots, y_n$ , nonnegative reals,  $r \geq 1$ . To prove Hölder, try reducing it to the case of two sequences and renormalization of the sequences to sum to 1.

- **Rearrangement** - Let  $a_1, \dots, a_n, b_1, \dots, b_n \in \mathbb{R}$  be non-decreasing sequences. Let  $\pi$  be a permutation of  $\{1, \dots, n\}$ . Then

$$\sum_{i=1}^n a_i b_i \geq \sum_{i=1}^n a_i b_{\pi(i)} \geq \sum_{i=1}^n a_i b_{n+1-i}$$

This inequality should be intuitively obvious, but how would you prove it?

- **Chebyshev** - Chebyshev's inequality applies for two sequences of real numbers  $a_1, \dots, a_n$  and  $b_1, \dots, b_n$ ,

$$\frac{\sum_{i=1}^n a_i b_i}{n} \geq \left( \frac{\sum_{i=1}^n a_i}{n} \right) \left( \frac{\sum_{i=1}^n b_i}{n} \right)$$

if the sequences  $A$  and  $B$  are sorted in the same order (ascending / descending). (If they are sorted in the opposite order, the inequality flips.) This follows from Rearrangement.

- **Schur** - Even though this inequality is easily proved, it is also one of the strongest. That is, it is closer to being not true than some inequalities such as AM-GM. It states that  $\forall a, b, c \in \mathbb{R}_0^+$  and  $r \in \mathbb{R}^+$ , we have

$$a^r(a-b)(a-c) + b^r(b-c)(b-a) + c^r(c-a)(c-b) \geq 0$$

with equality iff  $a = b = c$  or one of  $\{a, b, c\}$  is 0 and the other two are equal. For the mostly used case of  $r = 1$ , we have:

$$a^3 + b^3 + c^3 + 3abc \geq a^2b + b^2a + b^2c + c^2b + a^2c + c^2a$$

- **Trivial Inequality** - It states that  $\forall x \in \mathbb{R}$  we have  $x^2 \geq 0$ . Can't neglect this.
- **Majorization- Majorization and Muirhead** - First we must define what it means to majorize something. We say that the sequence  $(x_1, \dots, x_n)$  majorizes the sequence  $(y_1, \dots, y_k)$  iff  $x_1 + \dots + x_n \geq y_1 + \dots + y_k$  for all  $k = 1, 2, \dots, n$  and with equality when  $k = n$ . Note: the two sequences are first sorted into nonincreasing order. If  $f : [a, b] \rightarrow \mathbb{R}$  is a convex function and  $(x_1, \dots, x_n)$  majorizes the sequence  $(y_1, \dots, y_k)$  with  $x_i, y_i \in [a, b]$ . Then:

$$\sum_{i=1}^n f(x_i) \geq \sum_{i=1}^n f(y_i)$$

Although interesting, perhaps more interesting is Muirhead. Unfortunately, this is the only theorem in this packet you should never quote on the USAMO in your proof. Muirhead states that given the sequence  $(a_1, \dots, a_n)$  that majorizes the sequence  $(b_1, \dots, b_k)$  then for any positive reals  $x_1, \dots, x_n$ :

$$\sum_{\pi \in S_n} x_{\pi(1)}^{a_1} x_{\pi(2)}^{a_2} \cdots x_{\pi(n)}^{a_n} \geq \sum_{\pi \in S_n} x_{\pi(1)}^{b_1} x_{\pi(2)}^{b_2} \cdots x_{\pi(n)}^{b_n}$$

where the weird summation is a cyclic summation over all of the permutations of the  $x_i$ 's. I am told you can prove this with the majorization inequality from above. Sometimes it is known as "Bunching" because if a sequence of  $y_i$  is majorized by a sequence of  $x_i$  it is intuitively more bunched up. You can always use it to quickly check if you have gone too far. For example, if you have used AM-GM 20 times you want to make sure the expression you now have is still true, Muirhead is a very quick way to check the truth of your statement. In reality, all it does is do all the dirty work of AM-GM for you. When writing up your solution find the specific application of AM-GM that you want and do not just quote Muirhead.

- **Practice Problem** (USAMO '78)

$$a^2 + b^2 + c^2 + d^2 + e^2 = 16$$

$$a + b + c + d + e = 8$$

maximize  $e$ .